INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
34100 TRIESTE (ITALY) - P.O.B. 586 - MIRAMARE - STRADA COSTIERA 11 - TELEPHONE: 2240-1
CABLE: CENTRATOM - TELEX 460392 - I

SMR 281/9

COLLEGE ON VARIATIONAL PROBLEMS IN ANALYSIS
(11 January - 5 February 1988)

ON THE SINGULARITIES OF HARMONIC MAPS

J.M. CORON
Mathématique
Universite de Paris-Sud
Batiment 425
Orsay 91405
FRANCE

Commun. Math. Phys. 107, 649–705 (1986)

Communications in Mathematical Physics

Harmonic Maps with Defects

Haïm Brezis[1], Jean-Michel Coron[2], and Elliott H. Lieb[3,*]

[1] Département de Mathématiques, Université P. et M. Curie, 4 pl. Jussieu, F-75252 Paris Cedex 05, France

[2] Département de Mathématiques, Ecole Polytechnique, F-91128 Palaiseau Cedex, France

[3] Institut des Hautes Etudes Scientifiques, F-91440 Bures-sur-Yvette, France

Abstract. Two problems concerning maps φ with point singularities from a domain $\Omega \subset \mathbb{R}^3$ to S^2 are solved. The first is to determine the minimum energy of φ when the location and topological degree of the singularities are prescribed. In the second problem Ω is the unit ball and $\varphi = g$ is given on $\partial\Omega$; we show that the only cases in which $g(x/|x|)$ minimizes the energy is $g = \text{const}$ or $g(x) = \pm Rx$ with R a rotation. Extensions of these problems are also solved, e.g. points are replaced by "holes," $\mathbb{R}^3, S^2$ is replaced by $\mathbb{R}^N, S^{N-1}$ or by $\mathbb{R}^N, \mathbb{R}P^{N-1}$, the latter being appropriate for the theory of liquid crystals.

I. Introduction

Suppose $U \subset \mathbb{R}^3$ is open and $a \in U$. Consider maps $\varphi : U \to S^2$ which are continuous except (possibly) at a. If S is a sphere in U centered at a, φ restricted to S defines a map from S^2 to S^2 and so has a topological degree in $\mathbb{Z}$ (also known as winding or covering number). By continuity this number is independent of S and we shall denote it by d. If φ is also continuous at a, then $d = 0$.

Suppose now that $\varphi \in C^1(U \backslash \{a\}; S^2)$ and consider its energy

$$E(\varphi) = \int_U |\nabla \varphi|^2 \tag{1.1}$$

possibly finite or infinite. The fact that $E(\varphi) < \infty$ does not imply that φ is continuous at a or even that $d = 0$. An example with $d = 1$, U bounded and $a = 0$ is $\varphi(x) = x/|x|$. However if $U = \mathbb{R}^3$ and $E(\varphi) < \infty$, then d must be zero (since φ goes to a constant at infinity).

A natural problem is to minimize $E(\varphi)$ given the degree, d, of φ at a (assuming $U \neq \mathbb{R}^3$). We shall prove that the minimum energy is

$$E = 8\pi L, \tag{1.2}$$

where L is $|d|$ times the distance of a to ∂U.

* Permanent address: Departments of Mathematics and Physics, Princeton University, Jadwin Hall, P.O. Box 708, Princeton NJ 08544, USA. Work partially supported by U.S. National Science Foundation grant PHY 85-15288-A02

Another simple case is to consider two points $a_1, a_2 \in \mathbb{R}^3$ and maps

$$\varphi \in C(\mathbb{R}^3 \setminus \{a_1, a_2\}; S^2).$$

As above, one can define $\deg(\varphi, a_i)$, $i=1,2$, by restricting φ to small spheres around a_i, $i=1,2$. Assuming that $\varphi \in C^1(\mathbb{R}^3 \setminus \{a_1, a_2\}; S^2)$ and $E(\varphi) < \infty$, then we must have $d \equiv \deg(\varphi, a_1) = -\deg(\varphi, a_2)$. A natural problem is to minimize $E(\varphi)$ given d. We shall prove that the minimum energy is given by (1.2) with $L = |a_1 - a_2||d|$. The infimum is not achieved; however if φ^n is a minimizing sequence, we shall prove that φ^n tends to a constant a.e. and $|\nabla \varphi^n|^2$ tends to a uniform measure on the segment $[a_1, a_2]$ (after passing to a subsequence if necessary).

There are various generalizations of the two-point problem just mentioned, and they all give rise to the same formula (1.2) provided L is interpreted appropriately. We shall discuss four examples of increasing generality. Let U be an open set in $\mathbb{R}^3$. Let $H_1, \ldots, H_k$ be k disjoint compact subsets of U, which will be called the *holes*. Let $\Omega = U \setminus \left(\bigcup_{i=1}^{k} H_i \right)$. If $\varphi \in C(\Omega; S^2)$, then it is possible to define $\deg(\varphi, H_i)$, the degree of φ around H_i. If H_i is a point, $\deg(\varphi, H_i)$ is the usual topological degree, as above. For general H_i the degree can also be defined, but a bit of analysis is required; this is carefully discussed in Appendix B. Essentially, $\deg(\varphi, H_i)$ is the degree of φ restricted to a surface surrounding H_i.

Given integers $d_1, \ldots, d_k \in \mathbb{Z}$ (possibly including zero), consider the class

$$\mathscr{E} = \left\{ \varphi \in C(\Omega; S^2) \,|\, \deg(\varphi, H_i) = d_i \text{ and } \int_\Omega |\nabla \varphi|^2 < \infty \right\}. \tag{1.3}$$

Set

$$E = \inf_{\varphi \in \mathscr{E}} \int_\Omega |\nabla \varphi|^2. \tag{1.4}$$

[Note that E is unchanged if $C(\Omega; S^2)$ is replaced by $C^1(\Omega; S^2)$; this is explained in Appendix A.]

Example 1. $U = \mathbb{R}^3$ and the H_i are points a_i in $\mathbb{R}^3$.

Example 2. $U = \mathbb{R}^3$ and the H_i are not necessarily points.

Example 3. $U \neq \mathbb{R}^3$ and the H_i are not necessarily points.

Example 4. This is the same as Example 3, except that we consider the smaller class

$$\mathscr{E}' = \{ \varphi \in C(\bar{U} \setminus (\cup H_i); S^2) \,|\, \varphi \in \mathscr{E} \text{ and } \varphi = \text{const on } \partial U \}.$$

and let

$$E = \inf_{\varphi \in \mathscr{E}'} \int_\Omega |\nabla \varphi|^2. \tag{1.5}$$

In Examples 1, 2 (respectively 4), $\mathscr{E}$ (respectively $\mathscr{E}'$) is empty unless $\sum d_i = 0$.

Our main result concerning this problem is

Theorem 1.1. *In all four examples,*

$$E = 8\pi L, \tag{1.6}$$

where L is defined in Sect. II.

L is a quantity which has the dimension of a length and depends on U, on the relative distances between the holes and on the d_i's. It is easiest to visualize L in Example 1 and when $d_i = \pm 1$ for all i. We shall say that a_i is a positive (respectively negative) point if $d_i = +1$ (respectively -1). Since $\sum d_i = 0$ we can pair the positive points with the negative points. This pairing, or *connection* as we call it in Sect. II, has a length which is the sum of the distances between the paired points. L is defined to be the minimum possible length. If the d_i's are not ± 1, then simply repeat the point a_i $|d_i|$ times.

In Example 2 the rule is the same as for Example 1, except that one has to use the following reduced distance between holes. Given two holes H_i, H_j we let dist (H_i, H_j) be the usual Euclidean distance between the holes. Then we define the reduced distance to be

$$D(H_i, H_j) = \min \sum_{m=1}^{p} \operatorname{dist}(H_{i_{m-1}}, H_{i_m}),$$

where $i_0, \ldots, i_p$ is a finite sequence with $i_0 = i$, $i_p = j$ and the above minimum is over all such sequences.

In Example 3 just pretend that $H_0 \equiv \mathbb{R}^3 \setminus U$ is a hole of degree $d_0 = -\sum d_i$ and use the above rule to compute L.

The rule in Example 4 is the same as in Example 2 except that $\operatorname{dist}(H_i, H_j)$ is replaced by the geodesic distance in U.

The proof of Theorem 1.1 has two steps. In Sect. III we show that $E \leq 8\pi L$ by an explicit construction of an almost minimizer, which is obtained by gluing together "*dipoles*," i.e. almost minimizers for the two-point problem which are concentrated near the lines joining paired points. The lower bound $E \geq 8\pi L$ is more delicate. For this purpose, we introduce in Sect. IV a useful vector field D associated to $\varphi \in \mathscr{E}$, with components

$$D = (\varphi \cdot \varphi_y \wedge \varphi_z, \varphi \cdot \varphi_z \wedge \varphi_x, \varphi \cdot \varphi_x \wedge \varphi_y). \tag{1.7}$$

In all examples $\operatorname{div} D = 0$ in Ω and $2|D| \leq |\nabla \varphi|^2$. We sketch the essence of the argument for Example 1. In that case,

$$\operatorname{div} D = 4\pi \sum_{i=1}^{k} d_i \delta_{a_i} \equiv 4\pi \varrho \text{ in } \mathscr{D}'(\mathbb{R}^3), \tag{1.8}$$

so that

$$E \geq 8\pi \inf \left\{ \int_{\mathbb{R}^3} |D| \,|\, \operatorname{div} D = \varrho \right\}. \tag{1.9}$$

By duality, as explained in Appendix C,

$$\inf \left\{ \int_{\mathbb{R}^3} |D| \,|\, \operatorname{div} D = \varrho \right\} = \max \left\{ \int_{\mathbb{R}^3} \zeta d\varrho \,|\, \zeta \in K \right\},$$

where

$$K = \{ \zeta : \mathbb{R}^3 \to \mathbb{R} \,|\, \|\zeta\|_{\mathrm{Lip}} \leq 1 \} \quad \text{and} \quad \|\zeta\|_{\mathrm{Lip}} = \sup |\zeta(x) - \zeta(y)| / |x-y|.$$

We conclude by showing that

$$\max \left\{ \int_{\mathbb{R}^3} \zeta d\varrho \,|\, \zeta \in K \right\} = L \tag{1.10}$$

with the help of a theorem of Kantorovich [20] and Birkhoff's theorem [2, 26] on doubly stochastic matrices.

In general, there is no minimizer for the φ problem (1.4) [or (1.5)] and thus we are led in Sect. VI to investigate the behavior of minimizing sequences. However, the D problem defined by (1.9) and its analogue for the other examples does have a minimum as a vector-valued *measure*. Some properties of these D minimizers are described in Sect. V; for example we prove that $\operatorname{supp} D \subset G$, the union of the minimal connections. Our main result, in the context of Example 1, is that a minimizing sequence φ^n tends (modulo a subsequence) to a constant a.e. and $|\nabla\varphi^n|^2$ tends to a uniform measure distributed on a minimal connection. This is a striking fact since, if there is more than one minimal connection, a D minimizer can be supported by the union of two (or more) connections. This quantization phenomenon is based on the analysis in Appendix E.

A very different problem, one with a more classical flavor, is the subject of Sect. VII. Instead of specifying singularities we investigate the problem of minimizing $E(\varphi)$ on a domain $U \subset \mathbb{R}^3$ when $\varphi = g$ is specified on ∂U and we allow as admissible functions all H^1 maps from U into S^2. Clearly,

$$E(g) = \min\left\{\int_U |\nabla\varphi|^2 | \varphi \in H^1(U; S^2),\ \varphi = g \text{ on } \partial U\right\}$$

is achieved and it is known from the work of Schoen and Uhlenbeck [31, 32] that any minimizing φ has only point singularities and there are only finitely many of these. Our main result is

Theorem 1.2. *These singularities always have degree* ± 1 *and more precisely, near a singularity* x_0,

$$\varphi(x) \simeq \pm R(x - x_0)/|x - x_0|,$$

where R *is a rotation.*

This is a consequence of another result proved in Sect. VII, that if U is a ball, then $g(x/|x|)$ is a minimizer if and only if $\pm g$ is a rotation.

It is obvious that in the foregoing results one can replace the domain of φ by three dimensional manifolds other than $\mathbb{R}^3$, but we have not investigated these extensions. However other extensions are considered in Sect. VIII, for example we have replaced $\mathbb{R}^3, S^2$ by $\mathbb{R}^N, S^{N-1}$ and by $\mathbb{R}^N, \mathbb{R}P^{N-1}$. This replacement does not change the conclusions in any significant way. The $\mathbb{R}P^2$ extension is important for liquid crystals as explained below. We also touch upon a minimization problem where the minimum energy is proportional to an area (and not a length). A simple example of this kind of problem is to consider a closed Jordan curve $\Gamma \subset \mathbb{R}^3$ and $\varphi \in C(\mathbb{R}^3 \backslash \Gamma; S^1)$ having unit circulation around Γ. The energy to be minimized is $E(\varphi) = \int |\nabla\varphi|$. We conjecture that the minimum E is $2\pi A$, where A is the area of a minimal area surface spanning Γ.

In order not to interrupt the main thread of the paper, we have placed many of the technical facts in appendices. Some of these are of independent interest. For example, Appendix D contains a proof of the uniqueness of a divergence free vector-field supported on a curve. In Appendix E we present some noteworthy properties of certain nonlinear expressions involving weakly convergent sequences.

The mathematical analysis in this paper, summarized above, may be relevant to certain problems in physics.

A. Liquid Crystals

A nematic liquid crystal can be described by a vector field φ on a domain U in $\mathbb{R}^3$ (the container). The direction (optic axis) of the rod-like molecules at x is $\varphi(x)$ (called the director), so $|\varphi(x)| = 1$, and therefore we can view $\varphi(x)$ as a point in S^2. Normally, the ends of the molecules cannot be distinguished, so $\varphi(x)$ should really take values in $\mathbb{R}P^2$, i.e. the quotient of S^2 by the equivalence relation $\varphi \simeq -\varphi$.

Except for defects, which are points or curves in Ω, $\varphi(x)$ varies continuously. Frequently the liquid crystal energy is taken to be [7, 9, 13, 14, 17, 18, 21]:

$$\tilde{E}(\varphi) = K_1 \int_U (\operatorname{div}\varphi)^2 + K_2 \int_U (\varphi \cdot \operatorname{curl}\varphi)^2 + K_3 \int_U |\varphi \wedge \operatorname{curl}\varphi|^2. \tag{1.11}$$

A special case that has been frequently studied is the one-constant approximation $K_1 = K_2 = K_3 \equiv K$. Then the integrand on the right side of (1.11) is

$$K\{(\operatorname{div}\varphi)^2 + |\operatorname{curl}\varphi|^2\} = K\{|\nabla\varphi|^2 + 2D \cdot \varphi\} = K\{|\nabla\varphi|^2 + \operatorname{div} W\} \tag{1.12}$$

with D given by (1.7) and

$$W = \varphi \operatorname{div}\varphi - (\varphi \cdot \nabla)\varphi = \varphi \operatorname{div}\varphi + \varphi \wedge \operatorname{curl}\varphi. \tag{1.13}$$

Both (1.12) and (1.13) hold in the sense of distributions for all φ with $\nabla\varphi \in L^2$. Taking $K = 1$, and integrating (1.12) we find

$$\tilde{E}(\varphi) - E(\varphi) = \int_U \operatorname{div} W = \int_{\partial U} W \cdot n. \tag{1.14}$$

It is easy to check that $W \cdot n$ depends only on φ and its tangential derivatives on ∂U. Therefore, in all problems in which φ is prescribed on the boundary (such as Example 4 or the problems in Sect. VII) the boundary integral, $\int W \cdot n$, plays no role; the minimization of $\tilde{E}$ and E are the same problem. However, in Example 3, φ is not prescribed on the boundary and the two minimization problems are different. We shall discuss only the $E(\varphi)$ problem in this paper. It would be interesting to analyze the $\tilde{E}$ problem.

It is to be noted that $\varphi \to |\nabla\varphi|^2$ is $SO(3)$ invariant, namely if $R \in SO(3)$ and $\varphi'(x) \equiv R\varphi(x)$, then $|\nabla\varphi'|^2 = |\nabla\varphi|^2$. Also, D is $SO(3)$ invariant, i.e. $D(x) = D'(x)$, where D' is the D field of φ'. On the other hand, $H(\varphi) \equiv (\operatorname{div}\varphi)^2 + |\operatorname{curl}\varphi|^2$ is not $SO(3)$ invariant; it is only invariant under the simultaneous action of $SO(3)$ on φ and on x, i.e. $\varphi(x) \to R\varphi(Rx)$. From these observations one can conclude that $\tilde{E} \leq E$ in Example 3. Indeed, let $d\mu$ be Haar measure on $SO(3)$ so that $\int d\mu(R) D \cdot R\varphi = 0$. Thus, for all φ

$$\int d\mu(R) \int_U H(R\varphi) = \int_U |\nabla\varphi|^2, \tag{1.15}$$

so $\int H(R\varphi) \leq \int |\nabla\varphi|^2$ for some R.

Long lived point singularities are observed in nature [6] and have degree one, consistent with our Theorem 1.2.

B. The Classical O(3) Nonlinear Sigma Model

The Euler-Lagrange equation corresponding to (1.1) is

$$-\Delta\varphi = \varphi|\nabla\varphi|^2, \tag{1.16}$$

which is the equation of harmonic maps. It is also the equation of the classical nonlinear sigma model, but in the physics literature this is usually studied in $\mathbb{R}^2$, namely $\varphi:\mathbb{R}^2\to S^2$. Our analysis suggests that the $O(3)$ nonlinear sigma model from $\mathbb{R}^3\to S^2$ may be interesting, when singularities are included, although it is known that the quantized version of such a field theory is non-renormalizable. In any event, the expression for the energy needed to create two singularities separated by a distance L, namely $8\pi L$, is amusing. This is precisely the energy expression used in the semiclassical theory of quark confinement. Also, the fact that $\operatorname{supp}|\nabla\varphi^n|^2$ converges to a "string" is consistent with some pictures of quark-quark interactions.

Previously, Parisi [28] described a classical, relativistic field theory having some features in common with our φ field. In the static limit it reduces to monopoles embedded in a superconductor. However, to obtain strict linearity for the effective monopole-monopole interaction potential it seems to be necessary to take the limit of infinite critical field for the superconductor. For our Example 1, on the other hand, no limits are needed.

II. Minimal Connections

This section is concerned with defining some geometric quantities associated with a configuration of points or holes (disjoint compact subsets of $\mathbb{R}^N$) in certain domains in $\mathbb{R}^N$. From this construction we derive a number (with the dimension of a length) which, it will turn out, is proportional to the minimum energy.

A common feature of all the cases of interest to us is that we are given k disjoint holes in $\mathbb{R}^N$, $H_1,\dots,H_k$. According to the case, a certain distance function $D(H_i,H_j)$ will be defined between pairs of holes. D will satisfy the usual properties of a metric ($D(H_i,H_j)+D(H_j,H_k)\geqq D(H_i,H_k)$ and $D(H_i,H_j)>0$ for $i\neq j$ and $=0$ for $i=j$). The different choices of D will be defined subsequently.

Associated with each H_i is a degree $d_i\in\mathbb{Z}$. We assume that

$$\sum_{i=1}^{k} d_i=0. \tag{2.1}$$

The holes with $d_i>0$ (respectively $d_i<0$) are called *positive* (respectively *negative*) *holes*. Let

$$Q=\sum_{d_i>0} d_i=-\sum_{d_i<0} d_i \tag{2.2}$$

be the total positive degree.

Definition of a Connection and Its Length. List the positive holes with each H_i repeated d_i times in the list. Write this list as $P_1,\dots,P_Q$, with each P_j being some H_i. Likewise, list the negative holes, with each one repeated $|d_i|$ times. Write this as $N_1,\dots,N_Q$. Note that the holes of degree zero are omitted from these two lists. A *connection*, C, is a pairing of the two lists $(P_1,N_{\sigma1}),(P_2,N_{\sigma2})\dots(P_Q,N_{\sigma Q})$, where σ is a permutation in S_Q.

The *length* of this connection is defined to be

$$L(C)=\sum_{i=1}^{Q} D(P_i,N_{\sigma i}). \tag{2.3}$$

The *minimal length* is

$$L=\min_C L(C), \tag{2.4}$$

and a *minimal connection* is a connection (which may not be unique) such that

$$L(C)=L.$$

Example 1. The holes are k distinct points, $a_1,\dots,a_k$ in $\mathbb{R}^N$. $D(a_i,a_j)\equiv|a_i-a_j|$ = Euclidean distance. Note that in this case, holes of degree zero play no role whatsoever. We denote the minimal length by $L(\mathbb{R}^N,\{a\},\{d\})$.

Example 2. $H_1,\dots,H_k$ are k disjoint compact subsets of $\mathbb{R}^N$. (H_i could be a point or an object of any "dimension" from 1 to N.) $D(H_i,H_j)$ is defined as follows. First, let $\operatorname{dist}(H_i,H_j)$ be the usual Euclidean distance (i.e. $\min\{|x-y|\,|\,x\in H_i, y\in H_j\}$). Consider a chain $K=(i_0,i_1,\dots,i_p)$ with each $1\leqq i_m\leqq k$ and $i_0=i$, $i_p=j$ and let $\Delta(K)=\sum_{m=1}^{p}\operatorname{dist}(H_{i_{m-1}},H_{i_m})$. Then

$$D(H_i,H_j)\equiv\min_K \Delta(K). \tag{2.5}$$

Note that holes of degree zero that are not points may now play a role in the definition of D since their presence may reduce D (see Fig. 1). Also, one only has to consider chains K without repetition, so the minimum in (2.5) is over a finite set of chains. We denote the minimal length by $L(\mathbb{R}^N,\{H\},\{d\})$. If all the H_i are points this notation is consistent with Example 1.

Example 3. Let $U\neq\mathbb{R}^N$ be an open set in $\mathbb{R}^N$. Let $H_1,\dots,H_k$ be disjoint compact subsets of U with degrees $d_1,\dots,d_k$ but we *do not assume* (2.1). Introduce one more hole, $H_0\equiv\mathbb{R}^N\backslash U$ (which is closed but not necessarily bounded), and let $d_0\equiv-\sum_{i=1}^{k}d_i$. We repeat the construction of D and L in Example 2 (on $H_0,H_1,\dots,H_k$). Note that even though H_0 may not be compact, $D(H_0,H_i)>0$ for $i\neq0$. Also note that even if $d_0=0$, the presence of H_0 influences D and therefore L. We call the minimal length $L(U,\{H\},\{d\})$.

Example 4. Let $U\neq\mathbb{R}^N$ be a connected open set in $\mathbb{R}^N$. Let $H_1,\dots,H_k$ be disjoint compact subsets of U with degrees satisfying (2.1). For $x,y\in U$ let $\operatorname{dist}_G(x,y)$ be the geodesic distance within U, which will be defined in a moment. $\operatorname{Dist}_G(H_i,H_j)$ is defined as in Example 2, but with the Euclidean distance $|x-y|$ being replaced by $\operatorname{dist}_G(x,y)$. Then $D(H_i,H_j)$ is given by (2.5), using dist_G in $\Delta(K)$. The minimal

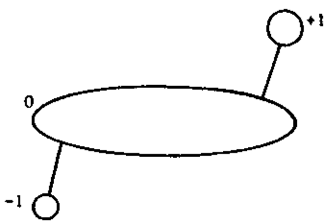

Fig. 1

length in this case will be denoted by $L_G(U,\{H\},\{d\})$. The geodesic distance $\mathrm{dist}_G(x,y)$ is defined as follows. Let $\kappa=\{x_1,\ldots,x_m\}$, with $x_i\in U$ and $x_0=x$, $x_m=y$ be a chain with the property that every line segment $[x_i,x_{i+1}]\subset U$. Note that such chains always exist since U is connected and hence arcwise connected. Let $\Lambda(\kappa)=\sum_{i=1}^{m}|x_{i-1}-x_i|$ and

$$\mathrm{dist}_G(x,y)\equiv\inf_{\kappa}\Lambda(\kappa). \tag{2.6}$$

Note that there exists a function $X:[0,1]\to\bar{U}$ with the properties that $X(0)=x$, $X(1)=y$, and $|X(t)-X(s)|\leqq|t-s|\,\mathrm{dist}_G(x,y)$ for all $t,s\in[0,1]$. This follows easily from Ascoli's theorem. Furthermore the length of the curve $X([0,1])$ equals $\mathrm{dist}_G(x,y)$ $\left(\text{the length is }\int_0^1|\dot{X}(t)|dt\right)$.

If U is convex then $\mathrm{dist}_G(x,y)=|x-y|$ and therefore $L_G(U,\{H\},\{d\})=L(\mathbb{R}^N,\{H\},\{d\})$.

Properties of Minimal Connections. In each example we introduce a distance $D(H_i,H_j)$. It is to be noted that this distance can be realized as the length of a finite union of continuous paths (which may or may not be unique). In Example 1 the path is the line segment $[a_i,a_j]$. In Examples 2 and 3, there is always a certain minimizing chain K in (2.5) and the paths are just line segments which realize $\mathrm{dist}(H_{i_{m-1}},H_{i_m})$. In Example 4 the line segments are replaced by curves in $\bar{U}$ of length $\mathrm{dist}_G(H_{i_{m-1}},H_{i_m})$.

Definition. A *string* is a continuous curve $X(t):[0,1]\to\bar{U}$ with the following properties:

i) $X(0)$ belongs to some hole H, $X(1)$ belongs to some hole H'.

ii) The length of the curve is $D(H,H')$.

iii) For $t\in(0,1)$, $X(t)$ belongs to none of the holes. The string carries an *orientation* from H to H'.

In Examples 1–3 a string is just a directed line segment running from H to H'. Given an arbitrary pair (H,H') there need not be a string from H to H', but $D(H,H')$ can always be realized as a finite chain of strings with the obvious consistent sequence of orientations.

Let C be a minimal connection: it has a pairing of the positive and negative holes and a length L given by (2.3). In an obvious way we can associate a finite union of strings with C, namely, first realize $D(P_i,N_{\sigma i})$ as a union of strings as above, and then take the union of all those strings including multiplicity. The sum of the lengths of all the strings is just L.

For descriptive purposes we can think of putting an arrow on each string in the direction of the orientation of the string. Some properties of the strings are the following:

a) For each hole H_i the number of arrows pointing out minus the number of arrows pointing in is just d_i.

b) If more than one string runs between H and H', all these strings are oriented in the same direction.

c) Given two strings in a minimal connection in Examples 1, 2 or 3, either
1) they are identical, or
2) they do not intersect, or
3) they intersect precisely at one point x with x in some hole.

The reason for this is the triangle inequality. Suppose S_1 (respectively S_2) is a string running from H_1 to H_1' (respectively H_2 to H_2'). Possibly some of these four holes are identical. Suppose $z\in S_1\cap S_2$ and z does not belong to a hole. We claim that $S_1=S_2$. Let x_1,y_1 (respectively x_2,y_2) be the end points of S_1 (respectively S_2) on H_1,H_1' (respectively H_2,H_2'). Consider the two paths $T_1=[x_1,z]\cup[z,y_2]$ and $T_2=[x_2,z]\cup[z,y_1]$. S_1 (respectively S_2) is part of a path joining some P_1 (respectively P_2) to some N_1 (respectively N_2). If we replace S_1 (respectively S_2) by T_1 (respectively T_2) we obtain a new connection in which P_1 (respectively P_2) is paired with N_2 (respectively N_1). The length is the same since $|T_1|+|T_2|=|S_1|+|S_2|$. But T_1 and T_2 are not line segments unless $S_1=S_2$.

In Example 4 the situation is more complicated. Two different strings can have a non-empty intersection.

III. Upper Bound to the Energy

For simplicity we restrict our attention to $\mathbb{R}^3$. In each of the four examples we have:

Theorem 3.1. *$E\leqq 8\pi L$ with L given by* (2.4).

The proof requires a construction, which we call the basic dipole. Take two distinct points a_+,a_- in $\mathbb{R}^3$ and some positive integer d. Given any $\varepsilon>0$ we construct a function $\varphi\in C(\mathbb{R}^3\setminus\{a_+,a_-\};S^2)$ such that:

a)
$$E(\varphi)\leqq 8\pi d|a_+-a_-|+\varepsilon. \tag{3.1}$$

b) φ is constant outside some set $N_\varepsilon(a_+,a_-)$, which we will henceforth call the support of φ, and which will be defined later.

c)
$$\deg(\varphi,\{a_\pm\})=\pm d. \tag{3.2}$$

Without loss of generality take $a_\pm=(0,0,\pm l)$. Given $\varepsilon>0$ we fix a smooth map $\omega:\mathbb{R}^2\to S^2$ such that:

$$\int_{\mathbb{R}^2}|\nabla\omega|^2\leqq 8\pi d+\varepsilon/2, \tag{3.3}$$

$$\omega\equiv\text{const}=e\text{ outside the unit disc}, \tag{3.4}$$

$$\deg\omega=-d. \tag{3.5}$$

Here, $\deg\omega$ is defined to be the degree of ω considered as a map from $S^2\simeq\mathbb{R}^2\cup\{\infty\}$ (by stereographic projection) to S^2. The existence of such a map is standard (see e.g. [4, proof of Theorem 2, Part C] used with $g\equiv$const). The idea for constructing ω is the following:

(i) Let $v(x,y)=(\mathrm{Re}(x+iy)^{-d},\ \mathrm{Im}(x+iy)^{-d})$. (ii) Let $\omega(x,y)=(\Pi\circ v)(x,y)$, where Π is stereographic projection from $\mathbb{R}^2$ to S^2. One finds that (3.3) and (3.5) are satisfied with $\varepsilon=0$. (iii) Replace v by $\chi v=\tilde{v}$, where $0\leqq\chi\leqq 1$ and χ has compact support and $\chi=1$ on a large disc D. Equations (3.3) and (3.5) are satisfied if D is chosen large enough. (iv) Now replace $\tilde{v}(x,y)$ by $\tilde{v}(\lambda x,\lambda y)\equiv\hat{v}$ with λ large enough so that supp $\hat{v}\subset$ unit disc. The left side of (3.3) is independent of λ.

Next, define $\varphi:\mathbb{R}^3\to S^2$ by

$$\varphi(x,y,z)=\begin{cases} e & \text{if } |z|\geq l \\ \omega\left(\dfrac{x}{l^2-z^2},\dfrac{y}{l^2-z^2}\right) & \text{if } |z|<l \end{cases} \tag{3.6}$$

and then set

$$\varphi_n(x,y,z)=\varphi(nx,ny,z). \tag{3.7}$$

φ_n is smooth on $\mathbb{R}^3\backslash\{a_+,a_-\}$ and satisfies (3.1) (if n is large enough) and (3.2). Finally, $\varphi_n=e$ outside the set where $z^2+n(x^2+y^2)^{1/2}\leq l^2$. This set (for n large enough) is the N_ε in (b) above. Note that the opening angle of N_ε at a_+ and a_- goes to zero as $\varepsilon\to 0$.

Proof of Theorem 3.1 for Examples 1–3. Let C be a minimal connection. As explained in Sect. II, C can be thought of as a finite collection of strings, each of which is a directed line segment running between pairs of holes and which carries some multiplicity, m. Suppose a string runs between $x\in H$ and $y\in H'$ and has length l. Then the open ball of radius l centered at y does not intersect H and, similarly, the open ball of radius l centered at x does not intersect H'. Thus, for small enough ε, we can insert a basic dipole (of degree m) between H and H'. If two or more different strings intersect at a common point $x\in H$ we can insert the required number of disjoint dipoles if ε is small enough. Inside each N_ε we take φ to be given by (3.7), and we take $\varphi=e$ outside $(\cup N_\varepsilon)$. Then $E(\varphi)\leq 8\pi L+\varepsilon\cdot$(the number of strings in C). □

Proof of Theorem 3.1 for Example 4. The difference with the previous case is that the strings are now curves instead of line segments and, moreover, they can intersect each other outside of the holes. However, any string between H and H' can be approximated (in length) by a polygonal path in $U\backslash(\cup H_i)$ (not $\bar{U}$). Moreover, we can also assume that any two such polygonal paths intersect at most only at their end points. To imitate the above construction we have to find the analogue of the basic dipole construction for a polygonal path, Γ, with end points $a_\pm$. That is, we want to construct a function φ satisfying (a) $E(\varphi)\leq 8\pi d|\Gamma|+\varepsilon$; (b) $\varphi=e$ outside $N_\varepsilon(\Gamma)$; (c) $\deg(\varphi,a_\pm)=\pm d$. Here, $N_\varepsilon(\Gamma)$ is contained in an ε neighborhood of Γ and has an ε opening angle at $a_\pm$. Let Γ be the union of line segments $[x_{i-1},x_i]$ with $x_0=a_+$, $x_p=a_-$ and all $x_i\in U$. We can, by passing to a refinement if necessary, assume that all $|x_{i-1}-x_i|$ are equal and have the common value $2l$. Think of the points x_i, $i=1,\dots,p-1$ as holes of degree zero and construct the function φ as in the end of the above proof, i.e. construct disjoint basic dipoles of degree d, one for each segment $[x_{i-1},x_i]$. Use the same n in (3.7) for all the intervals. Unfortunately, this function φ is not continuous at the points x_i, $i=1,\dots,p-1$. However, φ has degree zero at each x_i, $i=1,\dots,p-1$. To remedy the lack of continuity we proceed as follows. Let B_i, $i=1,\dots,p-1$ be balls of radius $R<l$ at the x_i and with R small enough so that there are only two basic dipoles in each B_i. We shall modify φ inside the B_i. On ∂B_i there are two disjoint circular caps in which $\varphi\neq e$. These are the intersections with ∂B_i of the two N_ε's of the two dipoles that intersect at x_i. Call the caps K_1 and K_2. There is a unique cylinder, C, with elliptical cross-section, whose intersection with ∂B_i is precisely $K_1\cup K_2$. If λ is a line in C parallel to the axis of C, then $\varphi(\lambda\cap K_1)=\varphi(\lambda\cap K_2)$. The function $\tilde{\varphi}$, which is the modification of φ and which is continuous, is defined by $\tilde{\varphi}(\lambda\cap B_i)=\varphi(\lambda\cap K_1)$. Outside $\cup B_i$, $\tilde{\varphi}=\varphi$. It is easy to see $E(\tilde{\varphi})\to E(\varphi)$ as $R\to 0$. □

IV. Lower Bound to the Energy

Again, as in Sect. III, we restrict our attention to $\mathbb{R}^3$. In each of the four examples we have:

Theorem 4.1. *$E\geq 8\pi L$ with L given by* (2.4).

Proof. Let $H=\bigcup_{i=1}^k H_i$ and $\Omega=U\backslash H$. Let φ satisfy the appropriate conditions, namely $\varphi\in C(\Omega;S^2)$, $\nabla\varphi\in L^2(\Omega)$, $\deg(\varphi,H_i)=d_i$ and, in Example 4 only, $\varphi\in C(\bar{U}\backslash H)$ and $\varphi=\text{constant}=e$ on ∂U. As explained in Appendix A, we can also assume that $\varphi\in C^\infty(\Omega)$. We shall show that $E(\varphi)\geq 8\pi L$.

Construct the vector field $D\in C^\infty(\Omega;\mathbb{R}^3)$ as in Appendix B, namely

$$D=(\varphi\cdot\varphi_y\wedge\varphi_z,\varphi\cdot\varphi_z\wedge\varphi_x,\varphi\cdot\varphi_x\wedge\varphi_y)$$

with $\varphi_x=\partial\varphi/\partial x$, etc.

We claim that a.e. on Ω:

$$|D|\leq\tfrac{1}{2}|\nabla\varphi|^2. \tag{4.1}$$

To see this, suppose that $\varphi=(0,0,1)$, $\varphi_x=(a_1,b_1,0)$, $\varphi_y=(a_2,b_2,0)$, $\varphi_z=(a_3,b_3,0)$, using the fact that $\varphi\cdot\varphi_x=0$, etc. Then $D=A\wedge B$ with $A=(a_1,a_2,a_3)$ and $B=(b_1,b_2,b_3)$. Therefore $|D|\leq|A||B|\leq\frac{1}{2}(A^2+B^2)=\frac{1}{2}|\nabla\varphi|^2$. Equality in (4.1) holds if and only if $A\cdot B=A^2-B^2=0$. Let $\zeta\in C(U)$ with $|\nabla\zeta|\leq 1$ (in $\mathcal{D}'$) and $\zeta=\zeta_i$ is a constant on each H_i. In Example 3 we also assume that $\zeta\in C(\bar{U})$ and $\zeta=0$ on ∂U. By (B.16) in Appendix B,

$$E(\varphi)\geq 2\int_\Omega|D|\geq -2\int_\Omega D\cdot\nabla\zeta=8\pi\sum_{i=1}^k\zeta_i d_i. \tag{4.2}$$

Our goal is to show that with $I(\zeta)=\sum\zeta_i d_i$, $I\equiv\sup_{\zeta\in Z}I(\zeta)=L$, where Z denotes the appropriate above-mentioned class. We only require $I\geq L$, but it is easily seen that $I(\zeta)\leq L$. Indeed, in Examples 1–3 (respectively 4), $|\zeta(x)-\zeta(y)|\leq|x-y|$, [respectively $\text{dist}_G(x,y)$] for all $x,y\in U$ and $\zeta\in Z$, since $|\nabla\zeta|\leq 1$. Consequently, in all cases $|\zeta_i-\zeta_j|\leq D(H_i,H_j)$. Since $\sum\zeta_i d_i=\sum_{j=1}^Q\zeta(P_j)-\zeta(N_j)$ for any pairing (see Sect. II for notation), $I(\zeta)\leq L$. Therefore we need only construct some $\zeta\in Z$ with $I(\zeta)=L$.

First, suppose there are k numbers $\{\zeta_i\}$ such that, with $\zeta_0\equiv 0$ (for Example 3),

$$|\zeta_i-\zeta_j|\leq D(H_i,H_j),\quad\text{for all } i,j. \tag{4.3}$$

Then we can construct $\zeta\in Z$ such that $\zeta=\zeta_i$ on each H_i. One choice is

$$\begin{aligned}\zeta(x)&=\max_i\{\zeta_i-\text{dist}(x,H_i)\},\quad\text{Examples 1–3}\\ &=\max_i\{\zeta_i-\text{dist}_G(x,H_i)\},\quad\text{Example 4}.\end{aligned} \tag{4.4}$$

Here $\operatorname{dist}(x,H)$ [respectively $\operatorname{dist}_G(x,H)] = \inf_{y\in H}|x-y|$ [respectively $\inf_{y\in H}\operatorname{dist}_G(x,y)$]. To see that this $\zeta\in Z$, note that $f_i(x)\equiv\operatorname{dist}(x,H_i)$ [respectively $\operatorname{dist}_G(x,H_i)$] satisfies $|\nabla f_i|\leqq 1$, and hence $|\nabla\zeta|\leqq 1$. Clearly $\zeta(H_i)\geqq\zeta_i$, so we have to check that $\zeta_i\geqq\zeta_j-\operatorname{dist}(x,H_j)$ [respectively $\operatorname{dist}_G(x,H_j)$] for all j and all $x\in H_i$. But $\zeta_j-\zeta_i\leqq D(H_i,H_j)\leqq\operatorname{dist}(x,H_j)$ [respectively $\operatorname{dist}_G(x,H_j)$].

To summarize, we merely have to find k numbers satisfying (4.3) and $\sum\zeta_i d_i=L$. Since $D(H_i,H_j)$ satisfies the triangle inequality, the following lemma establishes the existence of $2Q$ numbers $\{\alpha_i\}$ and $\{\beta_i\}$ such that $\alpha_i=\alpha_j$ (respectively $\beta_i=\beta_j$) if $D(P_i,P_j)=0$ [respectively $D(N_i,N_j)=0$]. With the P's and N's corresponding to holes repeated according to multiplicity, as in Sect. II, we can simply take ζ_i to be the common value of α_i (or β_i) on that hole. □

Lemma 4.2. *Let $P_1,P_2,\ldots,P_Q$ and $N_1,N_2,\ldots,N_Q$ be $2Q$ points and let X be their union. Let D be a semi-metric on X (i.e. a metric without the condition that $D(x,y)=0\Rightarrow x=y$). Let $L=\operatorname{Min}_{\sigma\in S_Q}\sum D(P_i,N_{\sigma i})$, where S_Q is the set of permutations. Then there exist real numbers $\alpha_1,\alpha_2\ldots\alpha_Q$ and $\beta_1,\beta_2\ldots\beta_Q$ such that*

$$\sum_{i=1}^{Q}(\alpha_i-\beta_i)=L, \tag{4.5}$$

and for all i,j

$$|\alpha_i-\alpha_j|\leqq D(P_i,P_j),\quad |\alpha_i-\beta_j|\leqq D(P_i,N_j),\quad |\beta_i-\beta_j|\leqq D(N_i,N_j). \tag{4.6}$$

Proof. This is a consequence of the Kantorovich theorem (see [10, 15, 20, 29]) and the Birkhoff theorem on doubly stochastic matrices (see [2, 26]). The Kantorovich theorem states that if X is a compact metric space with metric D and μ,ν are two non-negative measures on X such that $\int d\mu=\int d\nu$. Then

$$\operatorname*{Max}_{f\in\mathscr{L}}\left(\int f d\mu-\int f d\nu\right)=\operatorname*{Min}_{m}\iint D(x,y)\,dm(x,y), \tag{4.7}$$

where $\mathscr{L}=\{f:X\to\mathbb{R}\,|\,|f(x)-f(y)|\leqq D(x,y)\}$, and where m is a non-negative measure on $X\times X$ whose marginals are μ and ν. We apply this to our X and D with $\mu=\sum_{i=1}^{Q}\delta_{P_i}$ and $\nu=\sum_{i=1}^{Q}\delta_{N_i}$.

The measures m whose marginals are μ and ν are precisely of the form

$$m=\sum_{i,j=1}^{Q}a_{ij}\delta_{P_i}\otimes\delta_{N_j},$$

where $A=(a_{ij})$ is a doubly stochastic matrix (denoted by DS), i.e. $a_{ij}\geqq 0$ and

$$\sum_{j=1}^{Q}a_{ij}=\sum_{i=1}^{Q}a_{ij}=1,$$

for all i,j. The left side of (4.7) is $\operatorname{Max}_{\alpha,\beta}\sum_{i=1}^{Q}(\alpha_i-\beta_i)$, where α,β satisfy (4.6). The right side of (4.7) is $\operatorname{Min}_{A\in DS}\sum a_{ij}D(P_i,N_j)$. Birkhoff's theorem states that every $A\in DS$ is a convex combination of permutation matrices. Therefore the right side of (4.7) is $\operatorname{Min}_{\sigma\in S_Q}\sum D(P_i,N_{\sigma i})=L$. □

V. The D Problem

If we look back at Sect. IV we see that the lower bound for E was obtained by analyzing a problem that, in principle, is different from the original φ problem about S^2-valued vector fields. In this section we shall explore that auxiliary problem – which will be called the D problem – in more detail. Although the two problems give rise to the same minimal energy E in various cases (which fortunately include the cases of interest to us), the vector fields involved are different. At the end of this section we shall remark about the interrelation of φ of D.

The D problem is defined as follows. It will be defined in $\mathbb{R}^N$ instead of just $\mathbb{R}^3$ because the analysis is independent of N. As before we are given an open set $U\subset\mathbb{R}^N$ and k holes H_i (disjoint compact subsets of U). Let $H=\bigcup_{i=1}^{k}H_i$ and $\Omega=U\backslash H$. Associated with each H_i is a real number d_i (which now need not be an integer). We shall be concerned with L^1 vector fields, D, on Ω and distinguish two cases which we call A and B. Let Q_A denote the linear space of all functions $\zeta\in C(\bar U)$ with $\nabla\zeta\in L^\infty(U)$, $\zeta=0$ on ∂U (no condition if $U=\mathbb{R}^N$) and ζ is constant on each H_i. Let Q_B denote the linear space of all functions $\zeta\in C(U)$ with $\nabla\zeta\in L^\infty(U)$ and ζ is constant on each H_i.

Let $\mathscr{A}_A$ (respectively $\mathscr{A}_B$) denote the class of all vector fields $D\in L^1(\Omega;\mathbb{R}^N)$ satisfying

$$-\int_\Omega D\cdot\nabla\zeta=\sigma_N\sum_{i=1}^{k}d_i\zeta(H_i)\quad\text{for all }\zeta\in Q_A\text{ (respectively }Q_B). \tag{5.1}$$

Here σ_N denotes the area of S^{N-1} in $\mathbb{R}^N(\sigma_3=4\pi)$.

Note that if $U=\mathbb{R}^N$, then $\mathscr{A}_A$ is not empty if and only if $\sum_{i=1}^{k}d_i=0$. If $U\neq\mathbb{R}^N$ then $\mathscr{A}_A$ is always non-empty $\left(\text{even if }\sum_{i=1}^{k}d_i\neq 0\right)$. $\mathscr{A}_B$ is non-empty (for any U) if and only if $\sum_{i=1}^{k}d_i=0$. In this section we shall be concerned with minimizing the energy

$$E(D)\equiv\int_\Omega|D|. \tag{5.2}$$

Let E_A (respectively E_B) denote the infimum of $E(D)$ with D in the class $\mathscr{A}_A$ (respectively $\mathscr{A}_B$). Formally, Case A consists of minimizing $\int|D|$ over vector fields D such that $\operatorname{div}D=0$ in Ω and $\int_{\partial H_i}D\cdot\nu=\sigma_N d_i$ for each i, where ν is the normal to the surface ∂H_i. Case B consists of minimizing $\int|D|$ over vector fields D such that $\operatorname{div}D=0$ in Ω, $D\cdot\nu=0$ on ∂U and $\int_{\partial H_i}D\cdot\nu=\sigma_N d_i$ for each i. (If the holes H_i are points a_i, we have, as in Remarks B.2 and B.3, $\operatorname{div}D=\sigma_N\sum d_i\delta_{a_i}$.)

Case A is relevant for Examples 1–3 of Sect. II while Case B is relevant for Example 4. In the following we shall refer to the distance between holes $D(H,H')$ and we shall adopt the convention that for Case A (respectively Case B), $D(H,H')$ is defined as in Examples 1–3 (respectively 4) of Sect. II. If $N=3$ and the d_i's are integers, the analysis of Sects. II and IV shows that

$$E_{A,B}=\max\{\sigma_N\textstyle\sum d_i\zeta(H_i)\,|\,\|\nabla\zeta\|_{L^\infty}\leqq 1,\ \zeta\in Q_A\text{ (respectively }Q_B)\}. \tag{5.3}$$

In fact, (5.3) is always correct for all N and d_i.

Theorem 5.1. *Equation* (5.3) *holds in all cases. Moreover, if* $d_1, \dots, d_p > 0$ *and* $d_{p+1}, \dots, d_q < 0$ *and* $\sum d_i = 0$, *then*

$$E_{A,B} = \sigma_N \inf\left\{\sum_{i=1}^{p} \sum_{j=p+1}^{q} a_{ij} D(H_i, H_j)\right\}, \tag{5.4}$$

with $a_{ij} \geqq 0$ *and* $\sum_{j=p+1}^{q} a_{ij} = d_i$, $\sum_{i=1}^{p} a_{ij} = |d_j|$.

Proof. Equation (5.3) follows from the duality principles given in Appendix C. Equation (5.4) follows from (5.3) as in Lemma 4.2. □

It is intuitively evident from the variational construction in Sect. II, that a minimizing D for (5.2) often does not exist as an L^1 function. This will be clarified later. However, a minimizing sequence $\{D_n\}$ for (5.2) does have a limit in the sense of measures on $\bar{\Omega}$. More precisely there is a subsequence (which we continue to denote by D_n) such that $D_n \to D$ in the weak * topology of measures on $\bar{\Omega}$. This measure D satisfies

$$\int_{\Omega} |D| \leqq E_{A,B}. \tag{5.5}$$

Moreover D satisfies (5.1) except that we have to change the linear spaces Q_A (respectively Q_B) into

$$Q'_{A,B} = \{\zeta \in Q_A \text{ (respectively } \zeta \in Q_B) | \nabla\zeta \in C_c(\bar{U})\},$$

so that, in particular, the expression $\int D \cdot \nabla\zeta$ makes sense. We denote by $\mathscr{A}'_A$ (respectively $\mathscr{A}'_B$) the class of all vector valued measures on $\bar{\Omega}$, $D = (D_1, \dots, D_N)$ such that $\int |D| < \infty$ and

$$-\int_{\Omega} D \cdot \nabla\zeta = \sigma_N \sum_{i=1}^{k} d_i \zeta(H_i) \quad \text{for all } \zeta \in Q'_A \text{ (respectively } Q'_B). \tag{5.6}$$

Our problem is twofold: to establish equality in (5.5) and to identify these limiting measures.

Definition. An open set U is said to be *regular* if the following holds. Let

$$U_\varepsilon = U + B_\varepsilon = \{x + y | x \in U, |y| \leqq \varepsilon\}.$$

We suppose that for any two points $x, y \in U$, their geodesic distance relative to U_ε tends to their geodesic distance relative to U as $\varepsilon \to 0$.

Theorem 5.2.

$$\min\left\{\int_{\Omega} |D| \,\Big|\, D \in \mathscr{A}'_A\right\} = E_A. \tag{5.7}$$

If U is regular then

$$\min\left\{\int_{\Omega} |D| \,\Big|\, D \in \mathscr{A}'_B\right\} = E_B. \tag{5.8}$$

Proof. In view of (5.5) it suffices to prove $\geqq$ in (5.7) [or (5.8)]. Let $D \in \mathscr{A}'_A$; we have

$$\int_{\Omega} |D| \geqq -\int_{\Omega} D \cdot \nabla\zeta = \sigma_N \sum_{i=1}^{k} d_i \zeta(H_i) \tag{5.9}$$

for all $\zeta \in Q'_A$ with $\|\nabla\zeta\|_{L^\infty} \leqq 1$. Therefore we have to show that the supremum of the right side of (5.9) with $\zeta \in Q'_A$ and $\|\nabla\zeta\|_{L^\infty} \leqq 1$ is given by (5.3). We note in passing that, in general, the supremum is not achieved in the class Q'_A. The situation here is "dual" to that of Theorem 5.1 where E_A is not achieved while the right side of (5.3) is achieved. We recall (see Lemma B.5) that given any $\zeta \in Q_A$ with $\|\nabla\zeta\|_{L^\infty} \leqq 1$ there is a sequence ζ_n such that $\zeta_n \in C_c^\infty(U)$, ζ_n is constant on every H_i, $\|\nabla\zeta_n\|_{L^\infty} \leqq 1$ and $\zeta_n \to \zeta$ uniformly on every compact subset of U. This completes the proof of (5.7).

We turn now to the proof of (5.8). Let $U_\varepsilon = U + B_\varepsilon$ and let $H_{i,\varepsilon} = H_i + B_\varepsilon$. Let E_{B_ε} be the right side of (5.3) for this ε problem. Since U is regular $E_{B_\varepsilon} \to E_B$ as $\varepsilon \to 0$ by Remark 5.1. Let ζ_ε be a maximizer of (5.3) for the ε problem. Without loss of generality we may assume that $\zeta_\varepsilon \in L^\infty(U_\varepsilon)$ (otherwise truncate ζ_ε). Let $\zeta' = J_{\varepsilon/2} * \zeta_\varepsilon$. Clearly $\zeta' \in C^1(\bar{U})$, $\|\nabla\zeta'\|_{L^\infty(U)} \leqq 1$ and $\zeta'(H_i) = \zeta_\varepsilon(H_{i,\varepsilon})$. Finally consider $\zeta_n = (1 + C/n)^{-1} \chi_n \zeta'$, where $\chi_n(x) = \chi(x/n)$ and $\chi \in C_c^\infty(\mathbb{R}^N)$ is any function such that $\chi(x) = 1$ for $|x| < 1$ with $\|\chi\|_{L^\infty} \leqq 1$ and $C = \|\nabla\chi\|_{L^\infty} \|\zeta'\|_{L^\infty}$. Note that $\zeta_n \in Q'_B(U)$, $\|\nabla\zeta_n\|_{L^\infty(U)} \leqq 1$ and $\zeta_n \to \zeta'$ uniformly on every compact subset of $\bar{U}$. □

Remark 5.1. Case B of Theorem 5.2 may fail if U is not regular. Take for example

$$U = \mathbb{R}^3 \setminus \{(x, y, 0) | x \geqq 0, \ y \in \mathbb{R}\}.$$

For this U, the requirement that $\nabla\zeta \in C(\bar{U})$ implies that $Q'_B(U) = Q'_B(\mathbb{R}^3)$, and therefore the supremum of the right side of (5.5) can be less than the right side of (5.3).

We now turn to properties of the minimizing D measures.

1. Properties of the Support

Theorem 5.3. *Let D be any one of the following vector valued measures*

i) *a weak * limit of an* L^1 *minimizing sequence for* (5.2),

ii) *one of the minima referred to in* (5.7) *or* (5.8).

Let G be the union (which is closed) of all geodesics running between holes (see Sect. II). Then

$$\operatorname{supp} D \subset G. \tag{5.10}$$

Moreover, if all the d_i*'s are integers, then* $\operatorname{supp} D \subset G'$, *where* $G' \subset G$ *is the union of all minimal connections. Note that the definition of* G' *depends on the* $\{d_i\}$.

Proof. Let B be a closed set in U such that $B \cap G$ and $B \cap H$ are empty. Consider $V = U \setminus B$. The geodesic distance between holes for the V problem is obviously the same as for the U problem. Consider D restricted to $\bar{V}$ (respectively the $D_n \in L^1$ restricted to V) and call it $\hat{D}$ (respectively $\hat{D}_n$). These vector fields satisfy all the right conditions (for V), so

$$\int_{\Omega(V)} |\hat{D}| \geqq E(V) = E(U) = \int_{\Omega} |D| = \int_{\Omega(V)} |\hat{D}| + \int_{B} |D|,$$

and thus $D = 0$ on B, which proves (5.9). Note that $E(U) = E(V)$ by virtue of (5.4). A similar reasoning works for the sequence D_n, as well as for the case of integral d's (see Lemma 4.2). □

Remark 5.2. Note that i) holds even if U is not regular.

Remark 5.3. Consider Case A and assume D is a minimizer in (5.7). Then the $|D|$ measure of $(\cup H_i)$ is zero.

Indeed
$$\int_{\mathbb{R}^N\setminus(\cup H_i)} |D| \geqq -\int_{\mathbb{R}^N\setminus(\cup H_i)} D\cdot\nabla\zeta = \sigma_N \sum d_i\zeta(H_i)$$
for all $\zeta\in C^1(\mathbb{R}^N)$ with $|\nabla\zeta|\leqq 1$, $\zeta=0$ on ∂U and ζ is constant near each H_i. Therefore
$$\int_{\mathbb{R}^N\setminus(\cup H_i)} |D| \geqq \sigma_N \operatorname{Sup}_{\zeta} \sum d_i\zeta(H_i),$$
where ζ runs in the above mentioned class. It follows that
$$\int_{\mathbb{R}^N\setminus(\cup H_i)} |D| \geqq E_A = \int_{\mathbb{R}^N} |D|.$$

Remark 5.4. We conjecture that, for any U (regular or not) we have
$$\min\left\{\int_{\Omega} |D|\,\Big|\, D\in\mathscr{A}'_B,\ \operatorname{supp} D\subset G\right\} = E_B. \tag{5.11}$$

2. The Two Hole Problem

We investigate here two simple cases:

a) Case A with $U=\mathbb{R}^N$ and two disjoint holes H_1 and H_2 with $d_1+d_2=0$.

b) Case B with $U\neq\mathbb{R}^N$ and again two disjoint holes H_1 and H_2 with $d_1+d_2=0$.

In both cases $D(H_1,H_2)=L$.

Let us first analyze the case where there is precisely one geodesic, g, between H_1 and H_2. Let D be as in Theorem 5.3 so that $\operatorname{supp} D\subset g$. By Appendix D we know that D must be a measure of the form
$$D = cD_g. \tag{5.12}$$
On the other hand,
$$\int |D| = \sigma_N |d_1| L, \tag{5.13}$$
where L is the length of g. On the other hand,
$$\int |D_g| = L, \tag{5.14}$$
so that $c=\sigma_N|d_1|$. In particular $|D|$ is the uniform Hausdorff measure on g and the "direction of D is tangent to g." We shall establish similar properties in the general case where there are many geodesics between H_1 and H_2. As before we denote by G the union of all geodesics. In Case A, G is simply a union of line segments of length L which are disjoint except possibly for the end points. In particular every point, x, in $G\setminus(H_1\cup H_2)$ has precisely one geodesic passing through it. We denote its direction (going from H_1 to H_2) by $\mathbf{n}(x)$.

In Case B the situation can be much more complicated. Many geodesics can pass through a single point and the tangent need not be defined at every point of a geodesic. There could also be many geodesics connecting two points.

Theorem 5.4. *Under the assumptions of Theorem 5.3 and in Case A, the vector-valued measure D and the scalar-valued measure |D| are related by*
$$D = \mathbf{n}|D|. \tag{5.15}$$

Note that Eq. (5.12) relied only on the fact that $\operatorname{div} D=0$. But in Theorem 5.4 we really need the fact that D is minimizing. Consider, for example, the case where H_1 and H_2 are two cubes with parallel faces so that G is a cylinder, C, with a square base. Let g be any curve going from H_1 to H_2 in G. Then $\operatorname{div} D_g=0$, but (5.15) fails. The requirement that D is minimizing forces the "integral curves" of D to be geodesics.

We believe that a similar result holds in Case B.

Conjecture 5.1. Under the assumptions of Theorem 5.3, and in Case B, D is such that $\mathbf{n}(x)$ is well defined a.e. $|D|$ and $D=\mathbf{n}(x)|D|$. Here $\mathbf{n}(x)$ is the common tangent – when it exists – to all geodesics through x.

Proof of Theorem 5.4. Let $H_{1,\delta}=\{x|\operatorname{dist}(x,H_1)\leqq\delta\}$, and similarly for H_2, with $\delta>0$ and small. Note that $\operatorname{dist}(H_{1,\delta},H_{2,\delta})=L-2\delta$. Let
$$f(x)=\min\{\operatorname{dist}(x,H_{1,\delta}),\ L-2\delta\}. \tag{5.16}$$
Let g be a nonnegative C^∞ function with support in a ball of radius one around 0 and $\int g=1$. Let $g_\varepsilon(x)=\varepsilon^{-N}g(x/\varepsilon)$. Set
$$f_\varepsilon = g_\varepsilon * f \quad\text{for}\quad \varepsilon<\delta, \tag{5.17}$$
so that f_ε is a smooth function which is zero on H_1 and $f_\varepsilon=L-2\delta$ on H_2. We claim that
$$\nabla f_\varepsilon(x)\to\mathbf{n}(x) \text{ everywhere on } G\setminus(H_{1,\delta}\cup H_{2,\delta}). \tag{5.18}$$
Assuming that (5.18) holds, (5.15) follows easily. Indeed, since f is Lipschitz with constant one, $|\nabla f_\varepsilon|\leqq 1$ and hence
$$\sigma_N L=\int|D|\geqq\int D\cdot\nabla f_\varepsilon=(L-2\delta)\sigma_N. \tag{5.19}$$
From (5.18), and dominated convergence, we have
$$\int_{G\setminus(H_{1,\delta}\cup H_{2,\delta})} D\cdot\nabla f_\varepsilon \xrightarrow[\varepsilon\to 0]{} \int_{G\setminus(H_{1,\delta}\cup H_{2,\delta})} D\cdot\mathbf{n}, \tag{5.20}$$
and thus, combining (5.19) and (5.20) we obtain
$$\int_{G\setminus(H_{1,\delta}\cup H_{2,\delta})} D\cdot\mathbf{n}\geqq(L-2\delta)\sigma_N-\int_{(H_{1,\delta}\cup H_{2,\delta})\setminus(H_1\cup H_2)}|D| \tag{5.21}$$
(note that $f=0$ near H_1 and H_2). Passing to the limit in (5.21) as $\delta\to 0$ we find
$$\int_{G\setminus(H_1\cup H_2)} D\cdot\mathbf{n}\geqq L\sigma_N.$$
By Radon-Nikodym, we may write $D=\mathbf{F}|D|$ for some function $\mathbf{F}\in L^\infty(|D|)$ and $|\mathbf{F}|=1$ a.e. $|D|$. Thus we have $\mathbf{F}\cdot\mathbf{n}=1$ a.e. $|D|$, and so $\mathbf{F}=\mathbf{n}$ a.e. $|D|$. We turn now to the proof of (5.18). Let $\mathbf{e}$ be any unit vector in $\mathbb{R}^N$. We have
$$\frac{1}{t}[f_\varepsilon(x_0+t\mathbf{e})-f_\varepsilon(x_0)]=\frac{1}{t}\int g_\varepsilon(x_0-y)[f(y+t\mathbf{e})-f(y)]dy$$
$$=\frac{1}{t}\int g_\varepsilon(x_0-y)[\operatorname{dist}(y+t\mathbf{e},H_1)-\operatorname{dist}(y,H_1)]dy$$

for $x_0 \in G \backslash (H_{1,\delta} \cup H_{2,\delta})$ and ε small enough. Given a point z we denote by $a(z)$ any measurable projection of z on H_1. We have

$$\mathrm{dist}(y+te, H_1) - \mathrm{dist}(y, H_1) \geqq |y+te-a(y+te)| - |y-a(y+te)|,$$

and thus

$$\frac{1}{t}[f_\varepsilon(x_0+te) - f_\varepsilon(x_0)] \geqq \frac{1}{t}\int g_\varepsilon(x_0-y+te)[|y-a(y)| - |y-te-a(y)|].$$

On the other hand

$$|y-te-a(y)| = [|y-a(y)|^2 - 2te\cdot(y-a(y)) + t^2]^{1/2}$$
$$\leqq |y-a(y)|\left\{1 - t\frac{e\cdot(y-a(y))}{|y-a(y)|^2}\right\} + Ct^2.$$

Therefore as $t \to 0$ (and fixed ε) we have

$$\nabla f_\varepsilon \cdot e \geqq \int g_\varepsilon(x_0-y)\frac{e\cdot(y-a(y))}{|y-a(y)|}dy.$$

Finally we observe that

$$\lim_{y\to x_0}\frac{y-a(y)}{|y-a(y)|} = n(x_0),$$

since $a(y) \to a(x_0)$ because x_0 has a unique projection on H_1. We conclude that

$$\liminf_{\varepsilon\to 0} \nabla f_\varepsilon \cdot e \geqq e \cdot n(x_0).$$

Changing e into $-e$ we obtain (5.18). □

As we remarked earlier, when there is only one geodesic g between H_1 and H_2, then $|D|$ is a uniform measure on g. The analogue of this when there are many geodesics is the following

Theorem 5.5. *Let D be a measure as in Theorem 5.3 (for either Case A and Case B). For $0 \leqq \alpha < \beta \leqq L$ consider the slice*

$$S(\alpha, \beta) = \{x \mid \alpha < \mathrm{dist}(x, H_1) \leqq \beta\} \tag{5.22}$$

(with dist = *geodesic distance in Case B). Then*

$$\int_{S(\alpha,\beta)} |D| = \sigma_N |d_1|(\beta - \alpha). \tag{5.23}$$

Proof. Replace H_1 by $H_1 \cup S(0,\alpha)$. For this new problem $E' = \sigma_N|d_1|(L-\alpha)$. But we can use D restricted to $\bar{U} \backslash S(0,\alpha)$ as a variational measure for E' and obtain $E' \leqq \int_{S(\alpha,L)} |D|$. Likewise, replacing H_2 by $H_2 \cup S(\alpha, L)$,

$$\sigma_N |d_1| = E'' \leqq \int_{S(0,\alpha)} |D|.$$

Adding these inequalities we obtain:

$$\sigma_N|d_1|L \leqq \int_{S(\alpha,L)} |D| + \int_{S(0,\alpha)} |D| = \sigma_N |d_1| L.$$

Therefore

$$\int_{S(0,\alpha)} |D| = \sigma_N |d_1| \alpha. \quad \square$$

3. The Many Hole Problem

We now turn to the description of all minimizing D fields in the general case with many holes.

Suppose there are k holes $H_1, \ldots, H_k$ with degrees $d_1, \ldots, d_k$ (not necessarily integral). Some of these may be zero. We assume $\sum d_i = 0$ because, in Case A, we can assume that the complement of U is also a hole with the appropriate degree. First, let us consider the minimal energy $E_{A,B}$. If $d_1, \ldots, d_p > 0$ are the positive d's and $d_{p+1}, \ldots, d_q < 0$ are the negative d's, (5.4) gives us $E_{A,B}$ in terms of a $p \times r$ matrix $A = \{a_{ij}\}$ (with $r = q - p$). The set of minimizing A's (call it $\mathscr{A}$) is convex, as is the set of minimizing D's.

Recall that $D(H_i, H_j)$ is the geodesic distance (different for Cases A and B). It is realized by a finite sequence of strings (see Sect. II) running between a sequence of holes. More than one sequence may be possible. To be more specific, let $\chi_{ij} = 1$ if there exists a string between H_i and H_j and $\chi_{ij} = 0$ otherwise. For $\chi_{ij} = 1$, define $\mathscr{G}_{ij}$ to be the set of all strings between H_i and H_j, and L_{ij} their common length. Likewise, for $\chi_{ij} = 1$, let $\mathscr{D}_{ij}$ be the set of minimizing D fields (with $d_i = 1$, $d_j = -1$) constructed in the preceding section (the two-hole problem). If $\chi_{ij} = 0$, $\mathscr{G}_{ij}$ (respectively $\mathscr{D}_{ij}$) is a union (respectively sum) of the strings (respectively minimizing D fields) connecting H_i to H_j (respectively with $d_i = +1$, $d_j = -1$).

Now given an $A \in \mathscr{A}$ we can construct a minimizing D field as follows:

$$D = \sum_{i=1}^{p} \sum_{j=p+1}^{q} a_{ij} D^{ij}, \tag{5.24}$$

where $D^{ij} \in \mathscr{D}_{ij}$. Recall that when the d_i's are integral the extreme elements of $\mathscr{A}$ are given by Birkhoff's theorem, namely by a minimal connection in which each $a_{ij} \in \mathbb{Z}^+$.

Theorem 5.6. *For Case A, every minimizing D field is given by* (5.24).

We conjecture that the same is true for Case B.

Proof. Let $M = \{(i,j) \mid \chi_{ij} = 1\}$. A little thought shows that we can rewrite (5.4) as follows:

$$E_A = \tfrac{1}{2}\sigma_N \min \sum_M |\mu_{ij}| L_{ij}, \tag{5.25}$$

where the minimum is over $\mu_{ij} = -\mu_{ji}$ and $\sum_j \mu_{ij} = d_i$ for $i = 1, \ldots, k$. Pictorially, μ_{ij} can be thought of as the flux from i to j; it is not required that μ_{ij} has any definite sign.

Suppose that $(a,b) \in M$ and $(c,d) \in M$ (all points being distinct) and that geodesics $g_{ab} \in \mathscr{G}_{ab}$, $g_{cd} \in \mathscr{G}_{cd}$ (these are line segments). Suppose also that g_{ab} and g_{cd} intersect at a point P. In this case, we claim that either $\mu_{ab} = 0$ or $\mu_{cd} = 0$. If not, we can assume that $\mu_{cd} \geqq \mu_{ab} > 0$. Clearly, $D(H_a, H_d) < \mathrm{dist}(H_a, P) + \mathrm{dist}(P, H_d)$ and

$D(H_c, H_b) < \operatorname{dist}(H_c, P) + \operatorname{dist}(P, H_b)$. If (a,d) and $(c,b) \in M$, then $D(H_a, H_d) = L_{ad}$ and $D(H_c, H_b) = L_{bc}$, and we can replace the four numbers $\mu_{ab}, \mu_{cd}, \mu_{ad}, \mu_{cb}$ by 0, $\mu_{cd} - \mu_{ab}, \mu_{ad} + \mu_{ab}, \mu_{ab} + \mu_{cb}$ and thereby strictly lower the energy. This construction has to be modified in an obvious way if (a,d) or $(c,b) \notin M$. The conclusion we reach is that whenever $\mathcal{G}_{ab} \cap \mathcal{G}_{cd}$ is not empty, then every minimum of (5.25) has $\mu_{ab} = 0$ or $\mu_{cd} = 0$. The choice $[(a,b)$ or $(c,d)]$ is universal; if $\mu_{ab} \neq 0$ in one minimum and $\mu_{cd} \neq 0$ in another, then by taking the mean (which is still a minimum) we would have a contradiction.

Let $N \subset M$ be the set of (i,j) such that $\mu_{ij} \neq 0$ for some minimizer in (5.25). The families of geodesics $\{\mathcal{G}_{ab}\}$ for $(a,b) \in N$ are disjoint except possibly for the endpoints. Let $G = \bigcup_N \mathcal{G}_{ab}$. Now let D be a minimizer. We claim that $\operatorname{supp} D \subset G$. The proof of this is the same as the earlier proof (5.10) that $\operatorname{supp} D$ is contained in the geodesics between the positive and negative holes. If $x \notin G$ then remove a small ball around x (thereby creating a degree zero hole). If the ball is small enough nothing changes in (5.25) (recall the strict inequality of the preceding paragraph). Thus E_A does not change, but $D' = (1 - \chi_B)D$, with χ_B being the characteristic function of the removed ball, is an allowed vector field for the new problem, whence $\int_B |D| = 0$.

For $(a,b) \in N$, consider $D_{ab} \equiv F_{ab} D$, where F_{ab} is the characteristic function of $\mathcal{G}_{ab}$. If $\zeta \in C^1(\mathbb{R}^N)$ and $\zeta = 1$ on H_a, $\zeta = 0$ on H_b then, as is easily seen

$$-\int D_{ab} \cdot \nabla \zeta = \sigma_N \alpha_{ab}, \tag{5.26}$$

where α_{ab} is some constant that is independent of ζ. From the defining condition (5.1) on D we see that $\alpha_{ab} = -\alpha_{ba}$ and $\sum_b \alpha_{ab} = d_a$. By (5.25), $_A \leqq \frac{1}{2}\sigma_N \sum_{(a,b)\in N} |\alpha_{ab}| L_{ab}$. On the other hand, $\int |D| = \frac{1}{2} \sum_{(a,b)\in N} \int |D_{ab}|$ [since the $|D|$ measure of the holes is zero (see Remark 5.3)]. Thus $\int |D| \geqq \frac{1}{2}\sigma_N \sum_{(a,b)\in N} |\alpha_{ab}| L_{ab}$ [by (5.26)]. □

Remark 5.5 on the Relation of φ to D. Let Ω be an open set in $\mathbb{R}^3$. In Sect. IV, to every $\varphi \in C(\Omega; S^2)$ with $\nabla\varphi \in L^2(\Omega)$ we have associated a D field with the property that $\operatorname{div} D = 0$ in $\mathcal{D}'$ [for the generalization to $\mathbb{R}^N$, see (B.7)]. It is a natural question whether any vector field D with $\operatorname{div} D = 0$ comes from a φ. The answer is no, as the following, based on a remark of D. Sullivan, shows. Let S be a smooth closed surface in Ω. Let D a smooth vector field with the property that some integral curve of D is dense in S (for example S could be a two-torus, and D restricted to S is an irrational twist of the torus). Then this D can not come from a φ, as we shall now show.

From the definition of D in Sect. IV it follows that

$$(D \cdot \nabla)\varphi = 0, \tag{5.27}$$

since $|\varphi|^2 = 1$ [and thus $\det(\varphi_x, \varphi_y, \varphi_z) = 0$]. It follows that φ is constant on the integral curves of D and in particular φ is constant on S. Therefore $D = 0$ on S since $D = 0$ whenever $\nabla\varphi$ vanishes in two orthogonal directions. This contradicts the fact that $D \not\equiv 0$ on S.

VI. Behavior of Minimizing Sequences for the φ Problem

As before we are given an open set $U \subset \mathbb{R}^3$ and k holes H_i (disjoint compact subsets of U). Let $H = \bigcup_{i=1}^k H_i$ and $\Omega = U \backslash H$. Associated with each H_i is an *integer* d_i. We are concerned with the behavior of minimizing sequences for the problem

$$E = \inf \int |\nabla\varphi|^2 \tag{6.1}$$

under the appropriate conditions on φ, namely $\varphi \in C(\Omega; S^2)$, $\nabla\varphi \in L^2(\Omega)$, $\deg(\varphi, H_i) = d_i$ and, in Example 4 only, $\varphi \in C(\bar{U} \backslash H)$ and $\varphi = \text{constant}$ on ∂U. In Example 4 we also assume that U is regular.

Let φ^n be a minimizing sequence for (6.1) and let D^n be the field corresponding to φ^n. By passing to a subsequence we may assume that $D^n \rightharpoonup D$ weakly in the sense of measures. We claim that

$$|\nabla\varphi^n|^2 \rightharpoonup 2|D| \tag{6.2}$$

weakly in the sense of measures. Indeed we have $|D^n| \leqq \frac{1}{2}|\nabla\varphi^n|^2$; let us assume that $|\varphi^n|^2 \rightharpoonup \nu$ weakly in the sense of measures. Then we have

$$2|D| \leqq \nu. \tag{6.3}$$

On the other hand, by Theorem 5.2

$$\int |D| \geqq \tfrac{1}{2}E. \tag{6.4}$$

Since $\int \nu = E$, we conclude that $\nu = 2|D|$. Again, by Theorem 5.2, D is a minimizer for E_A or E_B, and thus we have the description of D given in Sect. V.

The conclusion of all this is that any minimizing sequence for the φ problem inherits all properties of minimizing sequences for the D problem that we studied in Sect. V. In particular, since D is supported on G, the union of all geodesics running between holes, (6.2) implies that φ^n converges strongly in H^1 to a constant on each connected component of the complement of G. However, the fact that D comes from a D^n, which comes from a φ^n, leads to some additional properties for D beyond those implied by the fact that D is a minimizer for the D problem. To derive these additional properties, Appendix E will play an essential role.

For simplicity we shall restrict our investigation to Examples 1–3 and with the additional assumption that there are only finitely many strings between any two holes.

Theorem 6.1. *There is a minimal connection, C, which we write $C = \cup g_i$, where the g_i's are strings (which are repeated according to their multiplicity in C), such that*

$$D = 4\pi \sum_i D_{g_i}. \tag{6.4}$$

D_g is defined in Appendix D. In particular,

$$|D| = 4\pi \sum_i \delta_{g_i}, \tag{6.5}$$

where δ_g is the one-dimensional Hausdorff measure on g. Consequently [by (6.2), (6.5)],

$$|\nabla\varphi^n|^2 \rightharpoonup 8\pi \sum_i \delta_{g_i}. \tag{6.6}$$

Remark 6.1. The point of this theorem is the following. If there is only one minimal connection, then the D problem has a unique minimizer and Theorem 6.1 does not give any additional information beyond that contained in (6.2). The interesting case is where there are several minimal connections, say C_1 and C_2 for example. Let D_1 (respectively D_2) be a minimizer with support in C_1 (respectively C_2). Any convex combination of D_1 and D_2 is also a minimizer, but this cannot happen for the φ problem. $|\nabla\varphi^*|^2$ must converge either to $2|D_1|$ or $2|D_2|$ but cannot converge to $|D_1|+|D_2|$, for example. This is a consequence of the quantization condition of Appendix E.

Proof. We recall that

$$\operatorname{supp} D \subset \bigcup_{(i,j)\in N} g_{ij},$$

where g_{ij} is a string running between H_i and H_j (i.e. it is a line segment). The set N is described in the proof of Theorem 5.6. N has the property that two distinct strings in N can intersect only at a common endpoint. We can write D as

$$D = (\tfrac{1}{2})4\pi \sum_{(i,j)\in N} v_{ij} D_{g_{ij}},$$

where g_{ij} is oriented from H_i to H_j and $v_{ij} = -v_{ji}$. By Theorem E.5 we know that $v_{ij}\in\mathbb{Z}$. Moreover the divergence condition implies that $\sum_j v_{ij} = d_i$ for each $i = 1, 2, \ldots, k$. The energy is given by

$$\int |D| = \tfrac{1}{2}4\pi \sum_{(i,j)\in N} |v_{ij}| L_{ij},$$

where L_{ij} is the length of g_{ij}. Since the energy is minimal, it follows that v_{ij} is a minimizer for (5.25). We claim that this set of v_{ij} defines a connection (which must be minimal since the energy is minimal). Take any positive hole, say H_i. By the divergence condition there must be at least one $v_{ij} > 0$. Go to H_j. If this is a negative hole, then stop and replace v_{ij} by $v_{ij} - 1$. If H_j is a zero or positive hole, then keep going until a negative hole is reached. Along this path replace all the v's by $v - 1$. By repeating this construction Q times (where Q is the sum of the positive degrees), we obviously have a connection. We claim that the remaining v's are all zero. This follows from the fact that v is a minimizer for (5.25) and that replacing the residual v's by zero would lower the energy in (5.25) and preserve the divergence condition. □

VII. Minimizing the Energy with Specified Boundary Conditions

A problem that we have so far not addressed in this paper is the minimization of the energy when φ is specified on the boundary of a domain (except for the special case where φ is constant on the boundary). Our analysis of the D problem in Sect. V is a useful guide to understanding the solution to certain open problems of this genre. In particular we shall answer the following questions.

Let B be the open unit ball in $\mathbb{R}^3$ and let

$$C_1 = \{\varphi \in H^1(B; S^2) | \varphi(x) = x \text{ on } \partial B\}. \tag{7.1}$$

Let

$$E(\varphi) = \int |\nabla\varphi|^2, \tag{7.2}$$

and

$$E_1 = \inf_{\varphi\in C_1} E(\varphi). \tag{7.3}$$

Question 1. Is $\psi(x) = x/|x|$ a minimizer for E_1?

Answer. Yes (see Theorem 7.1).

Next, let

$$C_2 = \{\varphi \in H^1(B; S^2) | \varphi(x) = g(x) \text{ on } \partial B\},$$

where $g: S^2 \to S^2$ is a given smooth map. Let

$$E_2 = \inf_{\varphi\in C_2} E(\varphi).$$

Question 2. Is $g(x/|x|)$ a minimizer for E_2?

Answer. No, unless g is an isometry (i.e. $+g$ or $-g$ is a rotation) or g is a constant (see Theorems 7.3 and 7.4).

In other words, if g is any smooth map from S^2 to S^2, and if g is extended radially to B, the extension is unstable unless g happens to be the constant map (degree zero) or g is the identity map modulo an isometry (degree ± 1).

We recall that a (smooth) map g from S^2 to S^2 is called harmonic if it satisfies the equation, $-\Delta g = g|\nabla g|^2$, where Δ is the Laplace-Beltrami operator on S^2. Harmonic maps from S^2 to S^2 have been classified (see e.g. [22, 35]) and their form is given in the proof below of Theorem 7.4. They all have the property that they minimize $\int_{S^2} |\nabla g|^2$ subject to the condition that the degree d of g is prescribed. In particular this integral is $8\pi|d|$.

We also recall a result of Schoen-Uhlenbeck [31, 32] that if we take an arbitrary domain Ω and minimize $E(\varphi)$ in $H^1(\Omega; S^2)$ with specified boundary condition, then any minimizing φ has at most finitely many point singularities. Our result implies that these singularities always have degrees ± 1 (see Corollary 7.12). In an earlier work Hardt et al. [19] showed that the degrees of these singularities are bounded by some universal constant. This confirms the observations on liquid crystals that stable point singularities have degree ± 1 [6]. It also confirms numerical studies by Cohen et al. [8] showing that singularities of degree two or more are unstable. Our first result is the following:

Theorem 7.1. *$\psi(x) = x/|x|$ is a minimizer for E_1; in fact, it is the unique minimizer.*

An obvious consequence of Theorem 7.1 is the following:

Corollary 7.2. *Suppose $g(x) = Rx$, where R is a rotation in $SO(3)$. Then $\psi(x) = Rx/|x|$ uniquely minimizes E_2.*

Our second result is:

Theorem 7.3. *If g has degree $+1$, then $g(x/|x|)$ is not a minimizer for E_2 unless $g(x) = Rx$, where R is a rotation in $SO(3)$.*

Our last main result is:

Theorem 7.4. *If g has degree d with $|d| \geqq 2$, then $g(x/|x|)$ is not a minimizer for E_2.*

Proof of Theorem 7.1. Clearly we have

$$E_1 \leqq E(x/|x|) = 8\pi. \tag{7.4}$$

In order to show that $\psi(x) = x/|x|$ is the unique minimizer for E_1 it suffices to show that

$$E(\varphi) > 8\pi \quad \text{for every } \varphi \in C_1,\ \varphi \neq \psi. \tag{7.5}$$

This leads us to the question of finding lower bound for the energy.

A. Lower Bounds for the Energy

There is always a minimizer for E_1 and, if φ_0 is a minimizer, we know from [32] that φ_0 is smooth on $\bar{B}$ except at most at a finite number of points in B. Therefore it suffices to prove (7.5) for φ in the class

$$\tilde{C} = \{\varphi \in C_1 | \varphi \text{ is continuous on } \bar{B}, \text{ except at a finite number of points in } B\}.$$

This will be achieved using the D field associated to φ. [An alternative to using the result of [32] about minimizers is to use a theorem of Bethuel-Zheng (in preparation) which states that $\tilde{C}$ is dense in C_1 for the H^1 norm.]

Let $\varphi \in H^1(B; S^2)$ be smooth on $\bar{B}$ except at a finite number of points in B (we do not assume that $\varphi(x) = x$ on ∂B). Let D be the D field associated with φ as in Sect. IV.

We have

$$\tfrac{1}{2}\int |\nabla\varphi|^2 \geqq \int |D| \geqq \int D \cdot \nabla\zeta = \int_{\partial B} (D \cdot n)\zeta - \int_B (\operatorname{div} D)\zeta$$

for every $\zeta \in C(\bar{B})$ with $\|\nabla\zeta\|_{L^\infty} \leqq 1$. Recall that $D \cdot n$ depends only on the values of φ restricted to ∂B and, more precisely, $D \cdot n = \varphi \cdot \varphi_x \wedge \varphi_y$, where x, y are orthonormal coordinates on S^2. On the other hand

$$\operatorname{div} D = 4\pi \sum_{i=1}^{p} d_i \delta_{a_i}$$

with $d_i \in \mathbb{Z}$ and $a_i \in B$. Consequently

$$\int_{\partial B} (D \cdot n) = 4\pi \deg(\varphi, S^2) \quad \text{and} \quad \sum_{i=1}^{p} d_i = \deg(\varphi, S^2).$$

Therefore we have

$$\int |\nabla\varphi|^2 \geqq 8\pi \max \left\{ \frac{1}{4\pi} \int_{\partial B} (D \cdot n)\zeta - \sum_{i=1}^{p} d_i \zeta(a_i) \middle| \zeta \in C(\bar{B}) \text{ with } \|\nabla\zeta\|_{L^\infty} \leqq 1 \right\}. \tag{7.6}$$

A basic lower bound for the right side of (7.6) is given by the following

Theorem 7.5. *Let M be a compact metric space with distance $\delta(x, y)$, let μ be a probability measure on M and let $\nu = \sum_{i=1}^{p} d_i \delta_{a_i}$, where $d_i \in \mathbb{Z}$ and $\sum_{i=1}^{p} d_i = 1$.*

Then

$$I(\nu) = \max\{\int \zeta d\mu - \int \zeta d\nu \,|\, \|\zeta\|_{\mathrm{Lip}} \leqq 1\} \geqq \min_{c \in M} \int \delta(x, c) d\mu(x), \tag{7.7}$$

where $\|\zeta\|_{\mathrm{Lip}} = \sup_{x \neq y} |\zeta(x) - \zeta(y)|/\delta(x, y)$.

Note that the right side in (7.7) is independent of ν and that (7.7) is obvious if $p = 1$, namely $\nu = \delta_a$ [take $\zeta(x) = \delta(x, a)$]. It follows that

$$\min_\nu \max_\zeta \{\int \zeta d\mu - \int \zeta d\nu\} = \min_c \int \delta(x, c) d\mu(x).$$

Combining (7.6) and Theorem 7.5 we obtain

Corollary 7.6. *Assume φ restricted to ∂B has degree one and satisfies $D \cdot n \geqq 0$ on ∂B, then*

$$\int |\nabla\varphi|^2 \geqq 2 \min_{c \in \bar{B}} \int_{\partial B} |\sigma - c|(D \cdot n) d\sigma. \tag{7.8}$$

A generalization of (7.8) is given in Remark 7.5 below.

Proof of Theorem 7.5. An easy approximation argument shows that it suffices to prove (7.7) in the case that $\mu = \sum_{i=1}^{q} \alpha_i \delta_{b_i}$ with $\alpha_i \geqq 0$, $\sum_{i=1}^{q} \alpha_i = 1$, and $b_i \in M$. Write $\nu = \sum_{i=1}^{k} \delta_{p_i} - \sum_{i=1}^{k-1} \delta_{n_i}$ (some points are repeated according to their multiplicity d_i). We shall use induction on k. As we have already indicated, the conclusions is obvious for $k = 1$.

As in Sect. IV and V it follows from the Kantorovich theorem [see (5.4)] that

$$I = \min \left\{ \sum_{i=1}^{k} \sum_{j=1}^{k-1} t_{ij} \delta(p_i, n_j) + \sum_{i=1}^{k} \sum_{j=1}^{q} s_{ij} \delta(p_i, b_j) \right\}, \tag{7.9}$$

the minimum being taken over the set of constraints $t_{ij} \geqq 0$, $s_{ij} \geqq 0$, $\sum_{i=1}^{k} t_{ij} = 1$ for all j $1 \leqq j \leqq k-1$, $\sum_{j=1}^{k-1} t_{ij} + \sum_{j=1}^{q} s_{ij} = 1$ for all i, $1 \leqq i \leqq k$ and $\sum_{i=1}^{k} s_{ij} = \alpha_j$ for all j, $1 \leqq j \leqq q$. Fixing the matrix $S = (s_{ij})$, consider the set τ of all matrices $T = (t_{ij})$ satisfying the above constraints. The set τ is compact and convex, therefore

$$\min_\tau \sum \sum t_{ij} \delta(p_i, n_j) \tag{7.10}$$

is achieved by some extremal point of τ. The following lemma, which is a variation of Birkhoff's theorem, gives a useful property of the extremal points of τ.

Lemma 7.7. *Let $\gamma = (c_1, \ldots, c_n)$ and $\varrho = (r_1, \ldots, r_m)$ be $n + m$ given nonnegative numbers satisfying $\sum c_i = \sum r_i$. Assume $m \leqq 2n$ and let $M_{m,n}(\gamma, \varrho)$ be the set of $m \times n$ matrices with nonnegative entries and having the c_i and r_i as column and row sums. I.e. $T \in M_{m,n}(\gamma, \varrho)$ means $T = \{t_{ij}\}$, $t_{ij} \geqq 0$, $\sum_{i=1}^{m} t_{ij} = c_j$, $\sum_{j=1}^{n} t_{ij} = r_i$. $M_{m,n}(\gamma, \varrho)$ is clearly a closed convex subset of $(\mathbb{R}^+)^{mn}$. If T is an extreme point of $M_{m,n}(\gamma, \varrho)$, then some column of T has $m - 2$ zeros, i.e. for some $j \in \{1, \ldots, n\}$, $t_{ij} = 0$ for at least $m - 2$ different i's.*

Proof. We can assume that $m = 2n$ simply by adding $2n - m$ rows of zeros. The lemma is trivially true for $n = 1$, $m = 2$ and we shall use induction on n. Let $n \geqq 2$. If T does not satisfy the lemma then each column of T has at least 3 positive entries. Since T is extremal, it is obvious that every submatrix, A, of T must be extremal (with respect to fixed row and column sums for A). Our goal will be to show that T

has a $k \times n$ submatrix, A, that is not extremal, for some $k \geq 2$. Let R denote the number of rows of T having $n-1$ zeros. The total number of positive elements of T, call it Σ, satisfies $\Sigma \geq 3n$. Then $R+(2n-R)n \geq \Sigma \geq 3n$. This implies that $R \leq 2n - [n/(n-1)] < 2n-1$, so $R \leq 2n-2$. Hence T has $k \geq 2$ rows with the property that there are at least 2 positive elements in the row. A will be the submatrix of T consisting of those k rows.

We claim that each column of A has at least 2 positive entries. Let $j \in \{1, \ldots, n\}$ label some column of T. Suppose there are 2 rows of T with the property that each row has one positive entry and that entry occurs at a common position j. If this is true we are done, for it suffices to consider the $(2n-2) \times (n-1)$ submatrix, B, of T obtained by deleting those 2 rows and the j^{th} column. By induction, B is extremal and thus has a column with at most 2 positive entries. If s labels this column then column s in T has the same property (because column s had zeros in the 2 deleted rows). This contradicts our assumption that every column of T has ≥ 3 positive entries.

Thus, we have found a $k \times n$ submatrix, A, of T with the property that every row and column of A has at least 2 positive entries. This matrix cannot be extremal as we now show. Pick some positive entry of A, walk along the row to another positive entry, walk along that column to another positive entry, and so on until a point (I, J) that has been previously visited is reached. We thus obtain a closed path, starting at (I, J) through positive entries of A. Let F be the matrix that is $+1$ at (I, J), (-1) at the next point in the path and so on. Off the path, $F_{ij} \equiv 0$. Clearly all the row and column sums of F are zero. Moreover, for small ε, $T_\pm \equiv T \pm \varepsilon F \in M_{m,n}(\gamma, \varrho)$, so $T = \frac{1}{2}(T_+ + T_-)$. □

Proof of Theorem 7.5 Completed. Let $T=(t_{ij})$ be an extreme point of τ that achieves the minimum in (7.10). By Lemma 7.7, there is some j, $1 \leq j \leq k-1$ such that $t_{ij} \neq 0$ for at most two values of i. Suppose, for example, $j=1$, $t_{i,1}=0$ when $i \neq 1, 2$, and $t_{21} \leq t_{11}$. Now fix T and S in (7.9), but replace the point n_1 by p_1. By the triangle inequality, (7.9) is not increased by this replacement. This means that $I(\nu) \geq I(\bar{\nu})$, where $\bar{\nu}$ is the measure with n_1 replaced by p_1, namely $\bar{\nu}$ has only $k-1$ positive terms and $k-2$ negative terms. The conclusion follows by induction.

Remark 7.1. One may give an alternative proof of Theorem 7.5 using Graph Theory – more specifically a result of Hamidoune-Las Vergnas [16]. By approximation, we can always assume that $\mu = \sum_j \alpha_j \delta_{b_j}$ with $\alpha_j \geq 0$, $\sum \alpha_j = 1$, $b_j \in M$, and also $\alpha_j \in \mathbb{Q}$. Therefore, it suffices to consider the case where $\mu = \frac{1}{q} \sum_{j=1}^{q} \delta_{b_j}$ (where the points b_j are not necessarily distinct). As above, write $\nu = \sum_{i=1}^{k} \delta_{p_i} - \sum_{i=1}^{k-1} \delta_{n_i}$ so that the left side of (7.7) becomes

$$\frac{1}{q} \max \{ \int \zeta d\mu' - \int \zeta d\nu' \mid \|\zeta\|_{\mathrm{Lip}} \leq 1 \},$$

where $\mu' = \sum_{j=1}^{q} \delta_{b_j} + q \sum_{i=1}^{k-1} \delta_{n_i}$ and $\nu' = q \sum_{i=1}^{k} \delta_{p_i}$. Using the Kantorovich and Birkhoff theorems as in Sect. IV we find that this maximum equals

$$\min_\sigma \sum_{l=1}^{kq} \delta(P_l, N_{\sigma l}),$$

where the system (P_l) consists of the points $(p_i)_{1 \leq i \leq k}$ each repeated q times, and the system (N_l) consists of the points $(n_i)_{1 \leq i \leq k-1}$ each repeated q times together with the points $(b_j)_{1 \leq j \leq q}$ (counted with multiplicity one). It follows from the result of [16], that in any connection σ, there exists some point p_{i_0} which is joined to *every* $(b_j)_{1 \leq j \leq q}$ by disjoint paths. (Two paths are disjoint if they have no strings in common.) In particular we have

$$\sum_{l=1}^{kq} \delta(P_l, N_{\sigma l}) \geq \sum_{j=1}^{q} \delta(p_{i_0}, b_j) = q \int \delta(x, p_{i_0}) d\mu(x) \geq q \min_{c \in M} \int \delta(x, c) d\mu(x),$$

which leads to (7.7). □

Remark 7.2. Suppose $M = \bar{B}$ (the unit ball). It is easy to see by going back to the proof of Theorem 7.5 that (7.7) is a *strict* inequality if $\operatorname{Supp}\mu$ is not contained in a single line and ν has at least three atoms.

Proof of Theorem 7.1. From Corollary 7.6 we obtain

$$E_1 \geq 2 \min_{|c| \leq 1} \int_{\partial B} |\sigma - c| d\sigma = 8\pi$$

(the minimum is achieved when $c=0$). Next we claim that φ is the unique minimizer. Let φ_0 be a minimizer for (7.3) and let D_0 be the corresponding D field. In view of Remark 7.2 we know that $\operatorname{div} D_0$ consists of a single Dirac δ_c and c must be zero $\left(\text{otherwise } 2 \int_{\partial B} |\sigma - c| d\sigma > 8\pi\right)$. Therefore, φ_0 has only one singularity with a nonzero degree, and that singularity is at $x=0$. Finally, we have $\partial\varphi_0/\partial r = 0$ because

$$8\pi = \int |\nabla \varphi_0|^2 = \int |\nabla_T \varphi_0|^2 + \int \left|\frac{\partial \varphi_0}{\partial r}\right|^2 \geq 8\pi + \int \left|\frac{\partial \varphi_0}{\partial r}\right|^2$$

(since φ_0 restricted to every sphere, rS^2, has degree one). □

Corollary 7.8. *Assume $\varphi: \bar{B} \to S^2$ has the following properties:*

$$\varphi(-x) = -\varphi(x) \quad \text{on } \partial B, \tag{7.11}$$

$$D \cdot n = J_\varphi = \varphi \cdot \varphi_x \wedge \varphi_y \geq 0 \quad \text{on } \partial B, \tag{7.12}$$

and

$$\deg(\varphi, S^2) = 1. \tag{7.13}$$

Then $\int |\nabla \varphi|^2 \geq 8\pi$.

Proof. We already know, by Corollary 7.6 that

$$\tfrac{1}{2} \int |\nabla \varphi|^2 \geq \int_{\partial B} |\sigma - c| (D \cdot n) d\sigma \tag{7.14}$$

for some $c \in \bar{B}$. Thus, we also have [by (7.11)]

$$\tfrac{1}{2} \int |\nabla \varphi|^2 \geq \int_{\partial B} |-\sigma - c| (D \cdot n)(-\sigma) d\sigma = \int_{\partial B} |\sigma + c| (D \cdot n)(\sigma) d\sigma. \tag{7.15}$$

By adding (7.14) and (7.15) we find

$$\tfrac{1}{2} \int |\nabla \varphi|^2 \geq \int_{\partial B} (D \cdot n) d\sigma = 4\pi. \quad \square$$

Remark 7.3. We conjecture that the conclusion of Corollary 7.8 holds without assumption (7.12).

Corollary 7.9 (Extension of Theorem 7.1). *Let Ω be any bounded domain in $\mathbb{R}^3$, then $\psi(x)=x/|x|$ is the unique minimizer for $\int_\Omega |\nabla\varphi|^2$ under the constraint that $\varphi=\psi$ on $\partial\Omega$.*

Proof. Let B_R be any large ball containing Ω and consider the problem of minimizing $E(\varphi)$ subject to $\varphi(x)=x/|x|$ on ∂B_R. By Theorem 7.1, the minimizing φ for B_R is uniquely $x/|x|$. Now let $\tilde\varphi$ be the minimizer for the Ω problem. If $\tilde\varphi$ differs from ψ in Ω, then there would be an alternative minimizer for the B_R problem, namely $f(x)=\tilde\varphi(x)$ for $x\in\Omega$ and $f(x)=x/|x|$ for $x\notin\Omega$. This would contradict uniqueness. □

Theorem 7.10 (Extension of Theorem 7.5). *Let M be a compact metric space and let μ be a positive measure with total mass $d\in\mathbb{N}$, and let $\nu=\sum_{i=1}^{p} d_i\delta_{a_i}$, where $d_i\in\mathbb{Z}$ and $\sum_{i=1}^{p} d_i=d$.*

Let

$$I(\nu)=\max\{\int\zeta d\mu-\int\zeta d\nu\,|\,\|\zeta\|_{\mathrm{Lip}}\leqq 1\}.$$

Then $\inf I(\nu)$ (where the infimum runs over all p's, a_i's, and d_i's) is achieved by a measure ν of the form

$$\nu=\sum_{i=1}^{k+d}\delta_{p_i}-\sum_{j=1}^{k}\delta_{n_j}$$

for some $0\leqq k\leqq d-1$.

Proof. Follow the same argument as in the proof of Theorem 7.5. □

Remark 7.4. For the purpose of Theorem 7.5, it would suffice to have Lemma 7.7 only for the case $m=n+1$. The reason we proved it for $m\leqq 2n$ was that this extended version is needed for Theorem 7.10.

Remark 7.5. Theorem 7.10 gives us a way to compute a lower bound for the problem

$$\min_{\varphi=\varphi_0 \text{ on } \partial B}\int_B|\nabla\varphi|^2,$$

provided $D\cdot n\geqq 0$ on ∂B, but without the assumption that φ_0 has degree 1. As far as the D problem is concerned, it can happen that when $d=2$, for example, the minimum of $I(\nu)$ occurs for three plus points and one minus point (i.e. $k=1$). Just take $D\cdot n$ to be three Dirac masses of strength $\frac{2}{3}$ placed in an equilateral triangle around the equator. The minimizing ν consists of three positive unit masses at the vertices of the triangle and one negative unit mass at the origin.

B. Proof of Theorem 7.3

First note that if $v(x)=g(x/|x|)$ is a minimizer for E_2, then g must be harmonic. Indeed v satisfies the equation $-\Delta v=v|\nabla v|^2$ in B and since v is independent of r, we have $-\Delta g=g|\nabla g|^2$. We shall construct explicitly a map, u, which coincides with g on ∂B and whose energy is less than 8π. Let $0<a<1$ and let $A=(0,0,a)$. We introduce polar coordinates centered at A with the direction $(0,0,1)$ as the pole. Thus a point x in B has coordinates (r,θ,φ), where r is the distance to A, θ is the polar angle and φ is the azimuthal angle. For a given angle $\theta\in[0,\pi]$, let $R(\theta)$ denote the maximum allowed radius (in $\bar B$). The points $(R(\theta),\theta,\varphi)$ with θ fixed and $\varphi\in[0,2\pi)$ all have a common polar angle $\psi(\theta)$ relative to the origin 0. We easily compute that

$$R(\theta)\sin\theta=\sin[\psi(\theta)],\qquad \tan\theta=\sin[\psi(\theta)](\cos[\psi(\theta)]-a)^{-1}. \tag{7.16}$$

Our choice for u is

$$u(r,\theta,\varphi)=g(\psi(\theta),\varphi), \tag{7.17}$$

so that its energy is

$$E(u)=\int_0^\pi\sin\theta d\theta\int_0^{2\pi}d\varphi\int_0^{R(\theta)}r^2dr\{r^{-2}|\nabla_{\theta,\varphi}u|^2\}, \tag{7.18}$$

where $\nabla_{\theta,\varphi}=(\partial/\partial\theta,(\sin\theta)^{-1}\partial/\partial\varphi)$. The r integration gives $R(\theta)\sin\theta$ which equals $\sin[\psi(\theta)]$, so (7.18) becomes

$$E(u)=\int_0^\pi\sin[\psi(\theta)]d\theta\int_0^{2\pi}d\varphi\{|g_2(\psi(\theta),\varphi)|^2[\sin\theta]^{-2}+|g_1(\psi(\theta),\varphi)|^2|\psi'(\theta)|^2\}. \tag{7.19}$$

Here g_1 and g_2 mean derivatives with respect to the first and second arguments.

Using (7.16) it is easy to compute

$$\psi'(\theta)=[1-2a\cos\psi(\theta)+a^2]/[1-a\cos\psi(\theta)], \tag{7.20}$$

$$\sin^2[\psi(\theta)]/\sin^2\theta=1-2a\cos\psi(\theta)+a^2. \tag{7.21}$$

Inserting this in (7.19) and changing variables from θ,φ to ψ,φ (with Jacobian $|\psi'|^{-1}$), we have

$$\begin{aligned}E(u)=\int_0^\pi\sin\psi\,d\psi\int_0^{2\pi}d\varphi\{&|g_2(\psi,\varphi)|^2[\sin\psi]^{-2}(1-a\cos\psi)\\&+|g_1(\psi,\varphi)|^2[(1-2a\cos\psi+a^2)/(1-a\cos\psi)]\}.\end{aligned} \tag{7.22}$$

If we set $a=0$ in (7.22) we obtain $E(g(x/|x|))=8\pi$. To prove the theorem, it therefore suffices to show that $E(u)<8\pi$ for small a. Expanding (7.22) in a near $a=0$, we need to show that

$$\int_0^\pi\sin\psi\,d\psi\int_0^{2\pi}d\varphi\{|g_2(\psi,\varphi)|^2(\sin\psi)^{-2}+|g_1(\psi,\varphi)|^2\}\cos\psi\neq 0. \tag{7.23}$$

However, (7.23) can be expressed in coordinate free form as follows. Let $e(\sigma)=|\nabla_T g(\sigma)|^2$, where $\sigma\in S^2$ and ∇_T is the tangential gradient. Then the left side of (7.23) is

$$I(\alpha)=\int_{S^2}e(\sigma)(\alpha\cdot\sigma)d\sigma, \tag{7.24}$$

where $\alpha=(0,0,1)$ and $d\sigma$ is the uniform measure on S^2 with $\int d\sigma=4\pi$. It is now clear that $I(\tilde\alpha)$ is the change of E if we replace $A=a\alpha$ by $\tilde A=a\tilde\alpha$, $\tilde\alpha\in S^2$. Thus, to complete the proof we must show that every harmonic map $g: S^2 \to S^2$, other than $g(x) = \pm Rx$, has the property that, for some $\alpha \in S^2$, $I(\alpha) \neq 0$. In other words, we have to show that for some $i \in \{1, 2, 3\}$,

$$N_i \equiv \int e(\sigma)\sigma_i d\sigma \neq 0. \tag{7.25}$$

Equation (7.25) means that the center of mass of $e(\sigma)$ is not at the origin.

Let Π denote stereographic projection from $\mathbb{C} \to S^2$. If $z = x + iy$,

$$\Pi(z) = (1+|z|^2)^{-1}(2x, 2y, 1-|z|^2). \tag{7.26}$$

Clearly we have

$$d\sigma = 4(1+|z|^2)^{-2} dx\, dy, \tag{7.27}$$

and if $h: S^2 \to \mathbb{C}$ and $H \equiv h \circ \Pi : \mathbb{C} \to \mathbb{C}$, then

$$|\nabla_T h|^2 = \tfrac{1}{4}|\nabla H|^2 (1+|z|^2)^2. \tag{7.28}$$

If $g: S^2 \to S^2$ and $f \equiv \Pi^{-1} \circ g \circ \Pi : \mathbb{C} \to \mathbb{C}$, then

$$|\nabla_T g|^2 = |\nabla f|^2 (1+|f|^2)^{-2}(1+|z|^2)^2. \tag{7.29}$$

If f happens to be holomorphic, then

$$|\nabla f|^2(z) = 2|f'(z)|^2; \tag{7.30}$$

g is harmonic of degree one if and only if

$$f(z) \equiv (\Pi^{-1} \circ g \circ \Pi)(z) = (az+b)/(cz+d) \tag{7.31}$$

for $a, b, c, d \in \mathbb{C}$, see e.g. [22, 35]. By a rotation of S^2 we can assume that $\infty \to \infty$, i.e. $c = 0$, $d = 1$. By a further rotation, $z \to ze^{iw}$, we can assume $a = \lambda b$ with $\lambda > 0$. Thus we may assume $f(z) = b(z + \lambda)$.

From the above formulas

$$N_i = 8\int dx\, dy |f'(z)|^2 [1+|f(z)|^2]^{-2}[1+|z|^2]^{-1} W_i(z) \tag{7.32}$$

with

$$W_1(z) = 2x, \quad W_2(z) = 2y, \quad W_3(z) = 1 - |z|^2.$$

By symmetry, $N_2 = 0$. If $\lambda > 0$ then $N_1 \neq 0$. To see this, let $K(x, y)$ denote the integrand, and note that for $x > 0$, $K(x, y) < K(-x, y)$ for all y when $\lambda > 0$. Thus, $N_1 = 0$ implies $\lambda = 0$. Finally, it is easy to see that $N_3 = 0$ if and only if $|b| = 1$. But $f(z) = e^{iw}z$ corresponds to $g(x) = Rx$ with R being a rotation by the angle w about the north pole. □

Remark 7.6. The proof of Theorem 7.3 shows something about harmonic maps generally (even those of degree $\neq \pm 1$). If $g(x)$, for $|x| = 1$, is given on the boundary, then $g(x/|x|)$ can never be a minimizer if the center of mass of $e(\sigma)$ is not at the origin, $x = 0$. Here, $e = |\nabla_T g|^2$.

C. *Proof of Theorem 7.4*

Let d be the degree of g and assume $v(x) = g(x/|x|)$ is a minimizer for E_2. As we remarked, g must be harmonic, and this is the case if and only if $f(= \Pi^{-1} \circ g \circ \Pi)$ is $P(z)/Q(z)$ if $d \geq 0$ or $P(\bar{z})/Q(\bar{z})$ if $d < 0$, with P and Q being polynomials and with

$|d| = \max\{\deg P, \deg Q\}$. By assumption we have

$$\int_B |\nabla v|^2 \leq \int_B |\nabla \varphi|^2, \quad \forall \varphi \in C_2. \tag{7.33}$$

We have clearly

$$\int_B |\nabla v|^2 = \int_{S^2} |\nabla_T g|^2 = 8\pi|d|. \tag{7.34}$$

In order to prove that $|d| \leq 1$ we shall choose special functions φ of the form described below. Let $\varepsilon \in (0, 1)$ and let $\theta: [0, 1] \to [0, \infty)$ be any smooth function such that $\theta(1) = 1$, $\theta(t) = 0$ for $t \in [0, \varepsilon]$ and $\theta(t) > 0$ for $t \in (\varepsilon, 1]$. Let

$$\varphi(x) = \Pi\left\{\frac{1}{\theta(|x|)} f[\Pi^{-1}(x/|x|)]\right\}$$

(with the convention that $0/0 = \infty$). Note that φ equals $N = (0, 0, 1)$ on the ball $B(0, \varepsilon)$. Moreover φ is smooth on B except at the points εx_i with $g(x_i) = S = (0, 0, -1)$. Also $\varphi(x) = g(x)$ for $|x| = 1$. We claim that

$$E(\varphi) = 8\pi|d|(1-\varepsilon) + 16\int_\varepsilon^1 dr \int_{\mathbb{R}^2} \frac{|\theta'(r)|^2 |f(\zeta)|^2 r^2 d\xi\, d\eta}{(\theta^2(r) + |f(\zeta)|^2)^2(1+|\zeta|^2)^2}, \tag{7.35}$$

where $\zeta = \xi + i\eta$.

Indeed we have

$$E(\varphi) = \int_{r>\varepsilon} \left(|\nabla_T \varphi|^2 + \left|\frac{\partial\varphi}{\partial r}\right|^2\right) = 8\pi|d|(1-\varepsilon) + \int_{r>\varepsilon} \left|\frac{\partial\varphi}{\partial r}\right|^2. \tag{7.36}$$

A direct computation shows that

$$\left|\frac{\partial\varphi}{\partial r}\right|^2 = \frac{4|\theta'(r)|^2 |f(\zeta)|^2}{(\theta^2(r) + |f(\zeta)|^2)^2}, \tag{7.37}$$

where $\zeta = \Pi^{-1}(x/|x|)$ and $r = |x|$. In order to compute $\int_{r>\varepsilon} \left|\frac{\partial\varphi}{\partial r}\right|^2$ we change variables and instead of $x = (x_1, x_2, x_3)$ we use the new variables (r, ξ, η), i.e.

$$r = \sqrt{x_1^2 + x_2^2 + x_3^2}, \quad \xi = \frac{x_1}{r - x_3}, \quad \eta = \frac{x_2}{r - x_3}.$$

Therefore we obtain

$$\int_{r>\varepsilon} \left|\frac{\partial\varphi}{\partial r}\right|^2 dx_1 dx_2 dx_3 = \int_{r>\varepsilon} dr \int_{\mathbb{R}^2} \left|\frac{\partial\varphi}{\partial r}\right|^2 J d\xi\, d\eta, \tag{7.38}$$

where J is the Jacobian determinant, i.e.

$$J = \frac{\partial(x_1, x_2, x_3)}{\partial(r, \xi, \eta)} = \frac{4r^2}{(1+|\xi|^2+|\eta|^2)^2}. \tag{7.39}$$

Combining (7.36), (7.37), (7.38), and (7.39) we obtain (7.35). Going back to (7.33) and (7.34) with (7.35) we obtain

$$8\pi|d| \leq 8\pi|d|(1-\varepsilon) + 16\int_\varepsilon^1 dr \int_{\mathbb{R}^2} \frac{|\theta'|^2 |f|^2 r^2 d\xi\, d\eta}{(\theta^2 + |f|^2)^2(1+|\zeta|^2)^2}$$

[for simplicity we write f instead of $f(\zeta)$] that is

$$\frac{\pi}{2}|d|\varepsilon \leqq \int_\varepsilon^1 dr \int_{\mathbb{R}^2} \frac{|\theta'(r)|^2|f|^2 r^2 d\xi\, d\eta}{(\theta^2(r)+|f|^2)^2(1+|\zeta|^2)^2}. \tag{7.40}$$

We change variable and set $t=\varepsilon/r$, $\alpha(t)=\theta\left(\frac{\varepsilon}{r}\right)$, $t\in[\varepsilon,1]$. From (7.40) we have

$$\pi\frac{|d|}{2} \leqq \int_\varepsilon^1 dt \int_{\mathbb{R}^2} \frac{|\alpha'(t)|^2|f|^2 d\xi\, d\eta}{(\alpha^2(t)+|f|^2)^2(1+|\zeta|^2)^2}. \tag{7.41}$$

Note that (7.41) holds for any function $\alpha:[\varepsilon,1]\to[0,\infty)$ such that $\alpha(\varepsilon)=1$, $\alpha(1)=0$. Passing to the limit in (7.41) we find

$$\pi\frac{|d|}{2} \leqq \int_0^1 dt \int_{\mathbb{R}^2} \frac{|\alpha'(t)|^2|f|^2 d\xi\, d\eta}{(\alpha^2(t)+|f|^2)^2(1+|\zeta|^2)^2} \tag{7.42}$$

for any function $\alpha:[0,1]\to[0,\infty)$ such that $\alpha(0)=1$, $\alpha(1)=0$. Set

$$F(s)=\int_0^s da\left\{\int_{\mathbb{R}^2}\frac{|f|^2 d\xi\, d\eta}{(a^2+|f|^2)^2(1+|\zeta|^2)^2}\right\}^{1/2}.$$

(It will follow from later computations that $F<\infty$.) We choose now $\alpha(t)=F^{-1}(F(1)(1-t))$ and so we obtain from (7.42), $\pi\frac{|d|}{2}\leqq F(1)^2$, and thus

$$\left[\pi\frac{|d|}{2}\right]^{1/2} \leqq F(1)=\int_0^1 ds\left\{\int_{\mathbb{R}^2}\frac{|f|^2 d\xi\, d\eta}{(s^2+|f|^2)^2(1+|\zeta|^2)^2}\right\}^{1/2}. \tag{7.43}$$

Let $R\in SO(3)$ be a rotation. Set $g_R=R\circ g:S^2\to S^2$, $f_R=\Pi^{-1}\circ g_R\circ\Pi$. Since $u(x)=g(x/|x|)$ is a minimizer, it follows that $u_R(x)=g_R(x/|x|)$ is also a minimizer for the boundary condition g_R, and therefore we have [from (7.34)]

$$\left[\pi\frac{|d|}{2}\right]^{1/2} \leqq \int_0^1 ds\left\{\int_{\mathbb{R}^2}\frac{|f_R|^2 d\xi\, d\eta}{(s^2+|f_R|^2)^2(1+|\zeta|^2)^2}\right\}^{1/2} \tag{7.44}$$

for every $R\in SO(3)$.

We shall average (7.44) over all rotations in $SO(3)$. Let m be the Haar measure left invariant on $SO(3)$. We have by (7.44)

$$\begin{aligned}\left[\pi\frac{|d|}{2}\right]^{1/2} &\leqq \int_0^1 ds \int_{SO(3)} dm(R)\left\{\int_{\mathbb{R}^2}\frac{|f_R|^2 d\xi\, d\eta}{(s^2+|f_R|^2)^2(1+|\zeta|^2)^2}\right\}^{1/2}\\ &\leqq \int_0^1 ds\left\{\int_{\mathbb{R}^2} d\xi\, d\eta \int_{SO(3)}\frac{|f_R|^2 dm(R)}{(s^2+|f_R|^2)^2(1+|\zeta|^2)^2}\right\}^{1/2}.\end{aligned} \tag{7.45}$$

Note that for every function $k:S^2\to\mathbb{R}$ and every $a\in S^2$, we have

$$\int_{SO(3)} k(Ra)dm(R)=\frac{1}{4\pi}\int_{S^2}k(\sigma)d\sigma \tag{7.46}$$

[clearly the left side of (7.46) is independent of a and so it equals its average on S^2, i.e. $\frac{1}{4\pi}\int_{SO(3)} dm(R)\int_{S^2} k(Ra)da$]. Also note that by changing variables we have

$$\frac{1}{4\pi}\int_{S^2}k(\sigma)d\sigma=\frac{1}{\pi}\int_{\mathbb{R}^2}k(\Pi(Z))\frac{dX\,dY}{(1+|Z|^2)^2} \tag{7.47}$$

$\left(\text{recall that the Jacobian determinant } \frac{\partial\Pi(Z)}{\partial(X,Y)}=\frac{4}{(1+|Z|^2)^2}\right)$. We use (7.46) and (7.47) with

$$k(\sigma)=\frac{|\Pi^{-1}(\sigma)|^2}{(s^2+|\Pi^{-1}(\sigma)|^2)^2} \quad \text{and} \quad a=\Pi\circ f(\zeta),$$

and we obtain, for every ζ,

$$\int_{SO(3)}\frac{f_R(\zeta)^2 dm(R)}{(s^2+|f_R(\zeta)|^2)^2}=\frac{1}{\pi}\int_{\mathbb{R}^2}\frac{|Z|^2 dX\,dY}{(s^2+|Z|^2)^2(1+|Z|^2)^2}. \tag{7.48}$$

A direct computation shows that the right side of (7.48) equals

$$2\frac{(s^2+1)}{(s^2-1)^3}(\ln s)-\frac{2}{(s^2-1)^2}\equiv G(s). \tag{7.49}$$

Going back to (7.45) we obtain

$$\left[\pi\frac{|d|}{2}\right]^{1/2}\leqq \pi^{1/2}\int_0^1[G(s)]^{1/2}ds,$$

and thus

$$|d|\leqq 2\left\{\int_0^1 G(s)^{1/2}ds\right\}^2.$$

We conclude with the next lemma that $|d|<2$. □

Lemma 7.11. *With $G(s)$ defined by* (7.49),

$$\int_0^1 G(s)^{1/2}ds<1. \tag{7.50}$$

Proof. Note that

$$\int_0^1 G(s)^{1/2}ds=\int_1^\infty G(s)^{1/2}ds$$

[since $G(1/s)=s^4G(s)$]. Set

$$b(s)=\left\{\frac{s^2+1}{s^2-1}\ln s-1\right\}^{1/2} \quad \text{for} \quad s>1,$$

so that, for $s>1$,

$$G(s)^{1/2}=\frac{\sqrt{2}}{s^2-1}b(s). \tag{7.51}$$

We claim that the function

$$s\to b(s)/(\ln s) \text{ is decreasing on } (1,\infty). \tag{7.52}$$

Letting $t=s^2$ we have to check that

$$\left(\frac{(t+1)}{2(t-1)}\ln t-1\right)\Big/(\ln t)^2 \text{ is decreasing},$$

that is

$$(\ln t)^2 \geqq \frac{(1-t^2)}{2t}\ln t + \frac{2}{t}(t-1)^2 . \tag{7.53}$$

Differentiating both sides of (7.53) it suffices to verify that

$$\ln t \geqq \frac{3(t^2-1)}{t^2+4t+1},$$

which holds since

$$\left[\ln t - \frac{3(t^2-1)}{t^2+4t+1}\right]' = \frac{(t-1)^4}{t(t^2+4t+1)^2} \geqq 0 .$$

Thus we have proved (7.52). In particular, we deduce from (7.52) that

$$\frac{b(s)}{\ln s} \leqq b'(1) = \frac{1}{\sqrt{3}} \quad \text{for all } s>1$$

[since $b(1)=0$] and also that

$$\frac{b(s)}{\ln s} \leqq \frac{1}{\sqrt{3}} - \left(\frac{1}{\sqrt{3}} - \frac{b(a)}{\ln a}\right) H(s-a) \tag{7.54}$$

for all $s>1$ and all $a>1$, where H is the Heaviside function [$H(t)=1$ for $t\geqq 0$ and $H(t)=0$ for $t<0$]. It follows from (7.54) that, for all $a>1$,

$$\begin{aligned}\int_1^\infty \frac{b(s)}{(s^2-1)}ds &\leqq \frac{1}{\sqrt{3}}\int_1^\infty \frac{\ln s}{(s^2-1)}ds - \left(\frac{1}{\sqrt{3}} - \frac{b(a)}{\ln a}\right)\int_a^\infty \frac{\ln s}{(s^2-1)}ds \\ &\leqq \frac{1}{\sqrt{3}}\int_1^\infty \frac{\ln s}{(s^2-1)}ds - \left(\frac{1}{\sqrt{3}} - \frac{b(a)}{\ln a}\right)\int_a^\infty \frac{\ln s}{s^2}ds \\ &= \frac{1}{\sqrt{3}}\int_1^\infty \frac{\ln s}{(s^2-1)}ds - \left(\frac{1}{\sqrt{3}} - \frac{b(a)}{\ln a}\right)\frac{(1+\ln a)}{a} .\end{aligned}$$

Finally, we recall that

$$\int_1^\infty \frac{\ln s}{(s^2-1)}ds = \frac{\pi^2}{8}$$

[which may be obtained by applying Fubini to $\int_0^\infty \int_0^\infty \frac{dx\,dy}{(1+y)(x^2+y)}$]. Thus we find, for all $a>1$,

$$\int_1^\infty G(s)^{1/2}ds \leqq \sqrt{\tfrac{2}{3}}\,\frac{\pi^2}{8} - \sqrt{2}\left(\frac{1}{\sqrt{3}} - \frac{b(a)}{\ln a}\right)\left(\frac{1+\ln a}{a}\right),$$

and we conclude that $\int_1^\infty G(s)^{1/2}ds<1$ by choosing for example $a=e^2$. □

Remark 7.7. Theorem 7.4 shows that if g has degree $|d|\geqq 2$, then $u(x)=g(x/|x|)$ is not a minimizer. In fact the construction above shows that it is not even a local minimizer.

Corollary 7.12. *Let u be a minimizer for $E(\varphi)$ in a domain Ω with specified boundary condition. Then each point singularity of u has degree ± 1. Moreover, for every singularity x_0 in Ω we have*

$$\lim_{\varepsilon\to 0} u(\varepsilon(x-x_0)) = \pm R(x-x_0)/|x-x_0| ,$$

where R is a rotation.

Proof. Without loss of generality we may assume that u has a singularity at $x=0$. We know from [31, Theorem III] and [33, Sect. 8] that $u(\varepsilon x)\to u_0(x)$ in $H^1(B)$ and uniformly on every compact subset of $B\backslash\{0\}$, where $u_0(x)=g(x/|x|)$ is a non-constant minimizer for E_2. It follows from Theorems 7.3 and 7.4 that g has degree ± 1 and that $\pm g$ is a rotation.

Remark 7.8. The fact that $x/|x|$ is a minimizer for E_1 (but not uniqueness) could also be deduced from Theorems 7.3 and 7.4 and the Schoen-Uhlenbeck result. Indeed let u be any minimizing harmonic map that happens to have a singularity, say at $x=0$. By [31] we know that $u(\varepsilon x)$ converges (modulo a subsequence) as $\varepsilon\to 0$ to a map $\varphi(x)$ with the properties that: (i) φ is a minimizing harmonic map with a singularity at $x=0$, (ii) $\varphi(x)=g(x/|x|)$ for some g. Our Theorems 7.3 and 7.4 eliminate all possibilities except $g(x)=\pm Rx$. This shows that $Rx/|x|$ is a minimizing harmonic map and therefore so is $x/|x|$.

VIII. Various Extensions

A. The N-Dimensional Case

A natural generalization is to replace $\mathbb{R}^3$ by $\mathbb{R}^N$ with $N\geqq 2$ and S^2 by S^{N-1}. The quantity which has the homogeneity of a length is now

$$E(\varphi)=\int |\nabla\varphi|^{N-1} \tag{8.1}$$

(and not $|\nabla\varphi|^2$) where φ is a map defined on a subset of $\mathbb{R}^N$ with values into S^{N-1} and

$$|\nabla\varphi|^2 = \sum_{i,j}\left(\frac{\partial\varphi_i}{\partial x_j}\right)^2 . \tag{8.2}$$

The analogue of Theorem 1.1 is

Theorem 8.1. *In all four examples*

$$E=\sigma_N(N-1)^{(N-1)/2}L, \tag{8.3}$$

where

$$\sigma_N = 2\pi^{N/2}\Gamma(N/2)^{-1} \tag{8.4}$$

is the area of S^{N-1} in $\mathbb{R}^N$. L is defined in Sect. II.

Proof. As before we construct upper and lower bounds for E. For the lower bound we define D as in (B.7) and note that

$$|D| \leqq (N-1)^{-(N-1)/2} |\nabla \varphi|^{N-1}. \tag{8.5}$$

Indeed, suppose that $\varphi=(0,0,\ldots,1)^t$, then $\varphi_{x_i}=(\alpha_{i,1},\alpha_{i,2}\ldots\alpha_{i,N-1},0)^t$, since φ_{x_i} is orthogonal to φ. The matrix

$$(\varphi_{x_1},\varphi_{x_2},\ldots,\varphi_{x_N})$$

has its last row zero. Replace the last row by $(\alpha_1,\alpha_2,\ldots,\alpha_N)$, and call M this new $(N\times N)$ matrix. We have $\det M=\alpha\cdot D$, so that

$$|D| = \sup_{|\alpha|\leqq 1} |\det M|.$$

On the other hand

$$|\det M| \leqq |\alpha| \prod_{j=1}^{N-1} \beta_j,$$

where $\beta_j^2 = \sum_{i=1}^{N-1} \alpha_{i,j}^2$, and thus

$$|\det M|^2 \leqq |\alpha|^2 \prod_{j=1}^{N-1} \beta_j^2 \leqq |\alpha|^2 \left[\frac{1}{(N-1)} \sum_{j=1}^{N-1} \beta_j^2\right]^{N-1} = |\alpha|^2 \frac{1}{(N-1)^{(N-1)}} |\nabla\varphi|^{2(N-1)}.$$

In dimension N, inequality (8.5) replaces the $\mathbb{R}^3$ inequality $|D| \leqq \frac{1}{2}|\nabla\varphi|^2$, and for the remainder of the proof of the lower bound we proceed as in Sect. IV.

For the upper bound we imitate the dipole construction of Sect. III. Let $\Pi:\mathbb{R}^{N-1}\to S^{N-1}$ be stereographic projection, namely

$$\Pi(x)=(\Pi_1(x),\ldots,\Pi_N(x)),$$

$$\Pi_i(x)=2x_i(1+|x|^2)^{-1} \quad \text{for} \quad i=1,\ldots,N-1 \quad \text{and} \quad \Pi_N(x)=(1-|x|^2)(1+|x|^2)^{-1}.$$

A straightforward computation yields

$$|\nabla \Pi| = 2(N-1)^{1/2}(1+|x|^2)^{-1}. \tag{8.6}$$

Recalling (8.4) we obtain from (8.6)

$$\int_{\mathbb{R}^{N-1}} |\nabla\Pi|^{N-1} = \sigma_N (N-1)^{(N-1)/2}. \tag{8.7}$$

Given $\varepsilon>0$ we first construct a smooth map $\omega:\mathbb{R}^{N-1}\to S^{N-1}$ such that

$$\int_{\mathbb{R}^{N-1}} |\nabla\omega|^{N-1} \leqq \sigma_N(N-1)^{(N-1)/2}+\varepsilon, \tag{8.8}$$

$$\omega \equiv \text{Const} = e \quad \text{outside the unit ball}, \tag{8.9}$$

$$\deg\omega = 1. \tag{8.10}$$

The idea for constructing ω is the following. Let $v(x)=x/|x|^2$ so that $\Pi\circ v$ satisfies (8.8) with $\varepsilon=0$ and (8.10). Next, replace v by $\chi v=\tilde{v}$, where $0\leqq\chi\leqq 1$ and χ has compact support and $\chi=1$ on a large ball. Finally, replace $\tilde{v}(x)$ by $\tilde{v}(\lambda x)$ with λ large enough. In the general case, $d>1$, we glue together d maps ω as above (with disjoint supports) and then rescale x. Analogously, for degree -1 we take $v(x)=|x|^{-2}(-x_1,x_2,\ldots,x_{N-1})$. Finally, having constructed ω, the basic dipole is constructed as in (3.6) and (3.7).

Another consequence of the construction in the proof of Theorem 8.1 is the following striking fact.

Theorem 8.2. *For maps* $\varphi: S^{N-1}\to S^{N-1}$, *let*

$$E(\varphi) = \int_{S^{N-1}} |\nabla_T \varphi|^{N-1}.$$

Then

$$\inf_{\deg\varphi=d} E(\varphi) = |d|\sigma_N(N-1)^{(N-1)/2}. \tag{8.11}$$

When $N\geqq 3$, the behavior of minimizing sequences for (8.1) is the same as for $N=3$ as given in Sect. VI, namely if there are only finitely many strings between any two holes and if φ^n is a minimizing sequence then, for a subsequence, $|\nabla\varphi^n|^{N-1}$ converges in the sense of measures to $\sigma_N(N-1)^{(N-1)/2}\delta_C$, where C is a (single) minimal connection. However when $N=2$ the situation is different, as shown by the following example.

Consider four points

$$a_1=(1,1), \quad a_2=(1,0), \quad a_3=(0,0), \quad a_4=(0,1)$$

with the degrees $d_i=(-1)^i$. Here, we have $E=2\pi L=4\pi$ and two minimal connections C_1,C_2 given by $C_1=[a_2,a_1]\cup[a_4,a_3]$ and $C_2=[a_2,a_3]\cup[a_4,a_1]$. There exist minimizing sequences φ^n such that, for example, $|\nabla\varphi^n|\rightharpoonup 2\pi(\delta_{C_1}+\delta_{C_2})$. Such a sequence can be obtained as follows. Let $\omega_\pm:\mathbb{R}\to S^1$ be any two maps such that $\omega_\pm(-\infty)=(\pm 1,0)$, $\omega_\pm(+\infty)=(\mp 1,0)$, $\omega_\pm$ constant far out and $\int|\nabla\omega_\pm|=\pi$. With $\omega_\pm$ we can associate "half dipoles" which we glue in an appropriate way on each of the intervals $[a_1,a_2]$, $[a_2,a_3]$, $[a_3,a_4]$, $[a_4,a_1]$. The corresponding sequence φ^n has the property that $\varphi^n\to(1,0)$ outside the square $[0,1]\times[0,1]$ and $\varphi^n\to(-1,0)$ inside $[0,1]\times[0,1]$. This lack of quantization in two dimensions is also discussed at the end of Appendix E.

B. Replacing S^{N-1} by $\mathbb{R}P^{N-1}$

For physical reasons as explained in Sect. I, it is interesting to replace S^{N-1} by $\mathbb{R}P^{N-1}$ which is the quotient of S^{N-1} by the equivalence relation $x\simeq -x$. The metric on $\mathbb{R}P^{N-1}$ is that induced by S^{N-1}. The energy is still given by (8.1).

The problem we face is to define the degree of a continuous map $\varphi:\Omega\to\mathbb{R}P^{N-1}$ (with $\Omega\subset\mathbb{R}^N$) around a hole in Ω. Unfortunately, $\mathbb{R}P^{N-1}$ is orientable if and only if N is even and therefore the problem will be more difficult when N is odd. The orientability of a manifold implies that the degree can be defined as an integral of a Jacobian. However, the degree for N even (as we shall define it) is in $\frac{1}{2}\mathbb{Z}$, and we shall be able to solve the minimum energy problem only when the given d_i's are integral, except for $N=2$ in which case $\mathbb{R}P^1$ is homeomorphic to S^1 and a special trick allows us to handle all d_i's.

a) N even. Suppose $\Omega\subset\mathbb{R}^N$ and S is a smooth surface in Ω without boundary. Let φ be C^1 in a neighborhood of S with values in $\mathbb{R}P^{N-1}$. The vector field D can be defined as in (B.7). Note that D in (B.7) is uniquely defined for N even because D is not changed by $\varphi\to-\varphi$. We define.

$$d=\sigma_N^{-1}\int_S D\cdot n. \tag{8.12}$$

Since $D\cdot n$ is the Jacobian, $\int_S D\cdot n$ must be an integer times the area of $\mathbb{R}P^{N-1}$, which is $\frac{1}{2}\sigma_N$. Therefore $d\in\frac{1}{2}\mathbb{Z}$.

Thus, given $\varphi\in C(\Omega;\mathbb{R}P^{N-1})$ and $\nabla\varphi\in L^{N-1}(\Omega)$, with $\Omega=U\backslash(\cup H_i)$, we can (by modifying the analysis in Appendix B) define the D field and $\deg(\varphi,H_i)\in\frac{1}{2}\mathbb{Z}$.

For the lower bound to E the D field analysis goes through as before and hence

$$E(\varphi)\geqq\tfrac{1}{2}\sigma_N(N-1)^{(N-1)/2}L(U,\{H_i\},\{2d_i\}), \tag{8.13}$$

where L is the length of a minimal connection (with d_i replaced by $2d_i$). Note the factor $\frac{1}{2}$ in (8.13).

For the upper bound we can reproduce the dipole construction of Sect. III when all the $d_i\in\mathbb{Z}$ as will be explained. In this case (8.13) becomes an equality for the infimum, and our problem is solved. Also, the obvious analogue of the results in Sect. VI go through. If some $d_i\notin\mathbb{Z}$ the problem is open.

The reason that $d_i\in\mathbb{Z}$ is special is the following topological fact.

Fact. Let ψ be a continuous map from $X\to\mathbb{R}P^{N-1}$, where X is a simply connected topological space. Then there exists a map $\tilde\psi:X\to S^{N-1}$ such that $\psi=P\circ\tilde\psi$, where P is the canonical projection of $S^{N-1}\to\mathbb{R}P^{N-1}$. (See the lifting theorem in [34, p. 76].) If X is also connected, there are exactly two choices for $\tilde\psi$ related by $\tilde\psi_1=-\tilde\psi_2$.

To construct the dipole when $d\in\mathbb{Z}$, first construct the S^{N-1} dipole as in Sect. III and then compose this with P. However, if $d\notin\mathbb{Z}$ we cannot do this because by taking $X=S^{N-1}$ in the above, we would end up with a continuous map $\tilde\psi:S^{N-1}\to S^{N-1}$ of degree $d\notin\mathbb{Z}$; this is impossible.

The topological fact also allows us to conclude that if a hole H_i has a neighborhood $\omega\subset U$ such that $\omega\backslash H_i$ is simply connected then necessarily $d_i\in\mathbb{Z}$. Simply take $X=\omega\backslash H_i$. In particular, if H_i is a point and if $N\geqq4$, $d_i\in\mathbb{Z}$.

b) N=2. In this case every hole, even a point hole, can have $d_i\notin\mathbb{Z}$. However $\mathbb{R}P^1$ is homeomorphic to S^1 and we can take advantage of this fact to solve the problem in *all* cases. We identify S^1 with $\{z\in\mathbb{C}\,|\,|z|=1\}$. Define $Q:S^1\to\mathbb{R}P^1$ as follows:

$$Q(z)=P(z'),\quad\text{where}\quad z'^2=z \tag{8.14}$$

with P being the canonical projection as before.

Clearly $Q(z)$ is independent of the choice of z'. Define $R:\mathbb{R}P^1\to S^1$ to be

$$R(P(z))=z^2 \tag{8.15}$$

(again, R is well defined). Note that $R=Q^{-1}$.

Let $\varphi:S^1\to\mathbb{R}P^1$ be a continuous map. We have

$$\deg\varphi=\tfrac{1}{2}\deg(R\circ\varphi). \tag{8.16}$$

Given a map φ from Ω into $\mathbb{R}P^1$ (respectively S^1) we have

$$|\nabla(R\circ\varphi)|=2|\nabla\varphi|\quad(\text{respectively }|\nabla(Q\circ\varphi)|=\tfrac{1}{2}|\nabla\varphi|). \tag{8.17}$$

Given reals $d_1,d_2,\ldots,d_k\in\frac{1}{2}\mathbb{Z}$, then

$$E=\pi L(U,\{H_i\},\{2d_i\}). \tag{8.18}$$

c) N odd. Here we shall confine our attention to cases in which $\Omega=U\backslash(\cup H_i)$ is connected and simply connected. This includes the case in which all H_i are points. Given a continuous $\varphi:\Omega\to\mathbb{R}P^{N-1}$, there exists a continuous $\tilde\varphi:\Omega\to S^{N-1}$ with $\varphi=P\circ\tilde\varphi$. Since there are exactly two choices for $\tilde\varphi$ ($\tilde\varphi_1=-\tilde\varphi_2$), we can define

$$\deg(\varphi,H_i)\equiv|\deg(\tilde\varphi,H_i)|\in\mathbb{N}. \tag{8.19}$$

(The need for the absolute value is that $\deg(\tilde\varphi_1,H_i)=-\deg(\tilde\varphi_2,H_i)$ when N is odd.) We also have that $|\nabla\tilde\varphi|=|\nabla\varphi|$.

Given nonnegative integers $d_1,\ldots,d_k$, we easily conclude from the above that the infimum satisfies

$$E=\sigma_N(N-1)^{(N-1)/2}\hat L, \tag{8.20}$$

where $\hat L$ is to be computed as follows:

$$\hat L=\min_{\{\varepsilon_i\}}L(U,\{H_i\},\{\varepsilon_i d_i\}), \tag{8.21}$$

where $\varepsilon_i=\pm1$, all i. In particular, we emphasize that (8.21) solves the minimum energy problem for liquid crystals with point defects and with the simplified energy given by (8.1).

C. Energies with the Homogeneity of an Area

Let $\Gamma\subset\mathbb{R}^3$ be an oriented, rectifiable Jordan curve. Consider the class of maps $\varphi:\mathbb{R}^3\backslash\Gamma\to S^1$ (not S^2) which are continuous. Associated with each φ in this class is an integer $d\in\mathbb{Z}$ defined as follows. Let C be any small circle which links with Γ. On C there is a natural orientation which is consistent with the orientation of Γ. Define

$$d=\deg(\varphi,\Gamma)=\deg(\varphi\text{ restricted to }C).$$

The right side is the usual degree of a map from S^1 to S^1. Note that $\deg(\varphi,\Gamma)$ is independent of the choice of C. The energy

$$E(\varphi)=\int_{\mathbb{R}^3\backslash\Gamma}|\nabla\varphi| \tag{8.22}$$

now has the homogeneity of an area (and not a length).

By analogy with the results of Sects. III and IV we expect that given $d\in\mathbb{Z}$

$$\inf_{\deg(\varphi,\Omega)=d}E(\varphi)=2\pi|d|A, \tag{8.23}$$

where A is the area of a minimal area surface spanned by Γ.

More generally, if M is an oriented manifold without boundary, of dimension m, imbedded in $\mathbb{R}^N$, and $\varphi:\mathbb{R}^N\backslash M\to S^{N-m-1}$ is a continuous map, then one can define (in the same way as above) $\deg(\varphi,M)$. The energy

$$E(\varphi)=\int_{\mathbb{R}^N\backslash M}|\nabla\varphi|^{N-m-1} \tag{8.24}$$

has homogeneity $(m+1)$ and we expect that given $d\in\mathbb{Z}$,

$$\inf_{\deg(\varphi,M)=d} E(\varphi)=c(N,m)|d|V, \tag{8.25}$$

where V is the volume of a "minimal" manifold of dimension $(m+1)$ whose boundary is M. Note that the case $m=0$ corresponds to two point holes and (8.25) reduces to (8.3). We could also consider a finite number of such manifolds $M_1,\dots,M_k$ and maps $\varphi:\mathbb{R}^N\backslash(\cup M_i)$ into S^{N-m-1} which are continuous except on M_i. It is a natural question to look for $\inf E(\varphi)$ in the class of maps φ such that $\deg(\varphi,M_i)=d_i$ is prescribed. Presumably, the answer is a formula similar to (8.25) where V is a kind of "minimal volume connection" associated with the M_i's and the d_i's.

We have not investigated the validity of (8.23) (or (8.25)) in full generality, and we shall discuss here only the case of a planar curve $\Gamma=\partial U$, where U is some open set in $\mathbb{R}^2$. Again, we split the argument in two parts: the upper bound and the lower bound.

1. *The Upper Bound.* Let ω be any continuous map from $\mathbb{R}$ to S^1 such that

$$\int_{\mathbb{R}}|\omega'|=2\pi|d|, \tag{8.26}$$

$$\deg\omega=d, \tag{8.27}$$

$$\omega=e \quad \text{outside} \quad [-1,+1]. \tag{8.28}$$

Let $\varphi_n:\mathbb{R}^3\backslash\Gamma\to S^1$ be defined as follows:

$$\begin{aligned}\varphi_n(x,y,z)&=\omega(nz/l) && \text{if} \quad (x,y)\in U,\\ \varphi_n(x,y,z)&=e && \text{if} \quad (x,y)\notin U,\end{aligned} \tag{8.29}$$

where l denotes the distance of (x,y) to ∂U. Clearly $\deg(\varphi_n,\Gamma)=d$ and $\int|\nabla\varphi_n|\to 2\pi A|d|$, where A is the area of U.

2. *The Lower Bound.* The divergence-free vector field D is now replaced by a curl-free vector field H as follows. To every map φ we associate H defined by

$$H=(\varphi\wedge\varphi_x,\varphi\wedge\varphi_y,\varphi\wedge\varphi_z).$$

An easy computation shows that if φ is smooth on $\mathbb{R}^3\backslash\Gamma$, then $\operatorname{curl}H=0$ on $\mathbb{R}^3\backslash\Gamma$ and, moreover, if $\int|\nabla\varphi|<\infty$, then

$$\operatorname{curl}H=2\pi dD_\Gamma \quad \text{in} \quad \mathscr{D}'(\mathbb{R}^3), \tag{8.30}$$

where D_Γ is the basic divergence-free vector field over the curve Γ defined in Appendix D. The proof of (8.30) is similar to that of the analogous formula (B.10) for the D field. Moreover, (8.30) extends (by density) to maps φ which are continuous on $\mathbb{R}^3\backslash\Gamma$ and with $\int|\nabla\varphi|<\infty$. Evidently, we have the inequality

$$|H|\le|\nabla\varphi|, \tag{8.31}$$

which plays the same role as $2|D|\le|\nabla\varphi|^2$. Therefore we have

$$\int_{\mathbb{R}^3}|\nabla\varphi|\ge\int_{\mathbb{R}^3}|H|\ge-\int_{\mathbb{R}^3}H\cdot\operatorname{curl}\zeta=2\pi d\int D_\Gamma\cdot\zeta \tag{8.32}$$

for every smooth ζ such that $|\operatorname{curl}\zeta|\le1$. On the other hand, by Stokes' theorem

$$\int_{\mathbb{R}^3}D_\Gamma\cdot\zeta=\int_\Sigma\operatorname{curl}\zeta\cdot n\,d\sigma \tag{8.33}$$

for any surface Σ spanned by Γ, where n is the unit normal to Σ. Choosing $\Sigma=U\times\{0\}$ and $\zeta(x,y,z)=\pm(0,x,0)$ we obtain $\operatorname{curl}\zeta=\pm(0,0,1)$ and from (8.32), (8.39) and the fact that $n=(0,0,1)$,

$$\int_{\mathbb{R}^3}|\nabla\varphi|\ge2\pi|d|A, \tag{8.34}$$

where A is the area of U. □

Remark 8.1. The upper bound construction presumably extends to nonplanar Γ, at least if the minimal area surface has no self-intersection. M. Gromov has suggested that the lower bound construction might also extend by using Whitney's duality theorem [37].

Appendix A: Approximation by Smooth Functions

Let $\Omega\subset\mathbb{R}^N$ be *any* open set. For the purpose of this paper we are interested in knowing whether we can approximate continuous S^k-valued functions on Ω with derivatives in L^2 by C^∞ S^k-valued functions, both for the uniform norm and energy norm. We present here a result more general than we need.

Lemma A.1. *Assume $u\in C(\Omega;\mathbb{R})$. Then for any $\varepsilon>0$ there is some $g\in C^\infty(\Omega;\mathbb{R})$ such that*

$$\|g-u\|_{L^\infty}<\varepsilon. \tag{A.1}$$

Moreover if we also assume $\nabla u\in L^{p_i}(\Omega)$ for some finite set $1\le p_1<p_2<\dots<p_m<\infty$ (in the distribution sense), then the above g can also be chosen to satisfy

$$\|\nabla(g-u)\|_{L^{p_i}}<\varepsilon \tag{A.2}$$

for all i.

Proof. This is essentially the same as the Meyers-Serrin theorem (see [25] or [1, p. 52]). The only variation is to note, in the notation of [1], that $\psi_k u\in C_c(\Omega)$ and, therefore, we may choose ε_k such that

$$\|J_{\varepsilon_k}*(\psi_k u)-\psi_k u\|_{L^\infty}\le\varepsilon/2^k. \quad □$$

Lemma A.2. *Assume u satisfies the hypotheses of Lemma A.1 and, moreover, $u\in C(\Omega;S^k)$. Then there is a $g\in C^\infty(\Omega;S^k)$ satisfying* (A.1) *and, if appropriate,* (A.2).

Proof. By Lemma A.1 (applied to each component of u) there is a sequence $\{h_n\}$ in $C^\infty(\Omega;\mathbb{R}^{k+1})$ such that

$$\|h_n-u\|_{L^\infty}\to0 \quad [\text{and } \|\nabla(h_n-u)\|_{L^{p_i}}\to0].$$

Assume that $\|h_n-u\|_{L^\infty}<1/2$, all n. Let $F:\mathbb{R}^{k+1}\to S^k$ be the radial projection, that is $F(x)=x/|x|$. Note that F is smooth for $x\neq0$. Let $g_n(x)=F(h_n(x))$. Since $h_n\to u$ uniformly, so does g_n [and $\nabla g_n=F'(h_n)\cdot\nabla h_n\to F'(u)\cdot\nabla u$ in L^p, since $F'(h_n)\to F'(u)$ uniformly and $\nabla h_n\to\nabla u$ in L^{p_i}]. □

Remark. In Lemma A.2 it is essential that u is continuous. Suppose that $\Omega=\{x\in\mathbb{R}^3 \,|\, |x|<1\}$ and $k=2$ and $u(x)=\frac{x}{|x|}$. This u has $\nabla u\in L^2$. However, there is no sequence $\{g_n\}$ with $g_n\in C(\Omega;S^2)\cap H^1(\Omega;S^2)$ such that $g_n\to u$ a.e. and $\nabla g_n\to\nabla u$ in L^2. See [32].

Appendix B: Generalities About Degrees of Maps

Let $U\subset\mathbb{R}^N$ be an open set and let $H\subset U$ be a compact subset (called a *hole*). Let $\varphi: U\backslash H\to\mathbb{R}^N$ be a continuous map such that $\varphi(x)\neq 0$, all $x\in U\backslash H$. We shall define $\deg(\varphi,H)$ as follows. Let

$$H_t=\{x\,|\,\mathrm{dist}(x,H)<t\}\,, \tag{B.1}$$

and assume ε is small enough so that $H_{4\varepsilon}\subset U$.

First, let ψ be any function in $C(U;\mathbb{R}^N)$ such that $\psi=\varphi$ on $U\backslash H_{3\varepsilon}$. (Such functions certainly exist. For example let $\chi\in C(U)$ be such that

$$\chi=\begin{cases}0 & \text{on } H_\varepsilon\\ 1 & \text{on } U\backslash H_{3\varepsilon},\end{cases}$$

then take $\psi=\chi\varphi$).

From the general theory of degrees of maps (see e.g. Nirenberg [27] or Lloyd [24]) the integer

$$d=\deg(\psi,H_{3\varepsilon},0) \tag{B.2}$$

is well defined. Part of this general theory is that d depends only on ψ restricted to $\partial H_{3\varepsilon}$, but this is independent of the choice of ψ (by construction). Conceivably d could depend on ε. However, it does not depend on ε (because if $\varepsilon_1<\varepsilon_2$ and ψ_1 corresponds to ε_1 we may take $\psi_2=\psi_1$).

Hence we are entitled to define

$$\deg(\varphi,H)\equiv\deg(\psi,H_{3\varepsilon},0)\,. \tag{B.3}$$

It follows from standard properties of degrees of maps that if $\varphi_n\to\varphi$ uniformly on every compact subset of $U\backslash H$, then $\deg(\varphi_n,H)\to\deg(\varphi,H)$.

Let us note some explicit formulas for d in (B.2). We can easily construct ψ such that $\psi\in C^1(H_{2\varepsilon};\mathbb{R}^N)$ and $\psi\neq 0$ in $U\backslash H_\varepsilon$. For such ψ,

$$d=\int_{H_\varepsilon} f(\psi(x))J_\psi(x)dx\,, \tag{B.4}$$

where $f:\mathbb{R}^N\to\mathbb{R}$ is any continuous function with compact support contained in the connected component of 0 in $\mathbb{R}^N\backslash\psi(\partial H_\varepsilon)$ and such that $\int_{\mathbb{R}^N} f(y)dy=1$. Here

$$J_\psi(x)=\det(\partial\psi_j/\partial x_i) \tag{B.5}$$

is the Jacobian determinant of ψ. Another formula for d can be obtained if one chooses ψ with the aforementioned properties and additionally $\psi=0$ at only finitely many points $x_1,\dots,x_m$ in H_ε and $J_\psi\neq 0$ at these points. (Such a ψ exists by Sard's lemma.) Then

$$d=\sum_{i=1}^{m}\mathrm{sgn}\,J_\psi(x_i)\,. \tag{B.6}$$

Examples. $U=\{x/|x|<1\}$ and $H=\{0\}$. Let $\varphi_1(x)=x$ and $\varphi_2(x)=x/|x|$. Then $\deg(\varphi_1,\{0\})=\deg(\varphi_2,\{0\})=1$.

Now suppose that $\varphi\in C(U\backslash H;\mathbb{R}^N)$ and $\nabla\varphi\in L^{N-1}(U\backslash H)$ (in $\mathscr{D}'$). To such a φ we associate a vector field $D\in L^1(U\backslash H;\mathbb{R}^N)$, with components D_j, as follows.

$$D_j=\det\left(\frac{\partial\varphi}{\partial x_1},\dots,\frac{\partial\varphi}{\partial x_{j-1}},\varphi,\frac{\partial\varphi}{\partial x_{j+1}},\dots,\frac{\partial\varphi}{\partial x_N}\right), \tag{B.7}$$

which is obviously in $L^1(U\backslash H)$. If, in addition, we assume that $\nabla\varphi\in L^N(U\backslash H)$, then J_φ, given by (B.5), is in $L^1(U\backslash H)$ and

$$\mathrm{div}\,D=NJ_\varphi \quad\text{in}\quad \mathscr{D}'(U\backslash H)\,. \tag{B.8}$$

(This is clear when φ is C^2; the general case follows by density, using Appendix A.)

Now suppose that $\varphi\in C(U\backslash H;S^{N-1})$ and $\nabla\varphi\in L^{N-1}(U\backslash H)$, but we do not assume $\nabla\varphi\in L^N(U\backslash H)$. Then

$$\mathrm{div}\,D=0 \quad\text{in}\quad \mathscr{D}'(U\backslash H)\,. \tag{B.9}$$

[Reason: By Appendix A, we can approximate φ by C^2 functions φ_n with $\|\varphi_n-\varphi\|_{L^\infty}\to 0$, $\|\nabla(\varphi_n-\varphi)\|_{L^{N-1}}\to 0$ and $\varphi_n(x)\in S^{N-1}$. Note that $J_{\varphi_n}=0$, since $\varphi\cdot\varphi=1 \Rightarrow \varphi\cdot\partial\varphi/\partial x_i=0 \Rightarrow$ the N vectors $\partial\varphi/\partial x_i$ are linearly dependent. By (B.8), $\mathrm{div}\,D(\varphi_n)=0$, but $D(\varphi_n)\to D(\varphi)=D$ in L^1.]

Theorem B.1. *Assume $\varphi\in C(U\backslash H;S^{N-1})$ and $\nabla\varphi\in L^{N-1}(U\backslash H)$, (in $\mathscr{D}'$). Then*

$$-\int_{U\backslash H} D\cdot\nabla\zeta=\sigma_N\deg(\varphi,H) \tag{B.10}$$

for every $\zeta\in\mathrm{Lip}(U)$ with compact support in U and $\zeta\equiv 1$ on some neighborhood of H. Here σ_N denotes the area of S^{N-1} in $\mathbb{R}^N$ ($\sigma_3=4\pi$).

Proof. By Lemma A.2 we can assume that $\varphi\in C^\infty(U\backslash H;S^{N-1})$. Clearly we may also assume that $\zeta\in C^\infty(U)$. With $I(\zeta)$ denoting the left side of (B.10) we first prove that $I(\zeta)$ is independent of ζ, and thus that it suffices to prove (B.10) for one ζ. Indeed,

$$I(\zeta_1)-I(\zeta_2)=-\int_{U\backslash H} D\cdot\nabla(\zeta_1-\zeta_2)=\int_{U\backslash H}(\mathrm{div}\,D)(\zeta_1-\zeta_2)=0 \tag{B.11}$$

(because $\zeta_1-\zeta_2$ has compact support in $U\backslash H$).

Now observe that for all $\theta\in C^\infty(U\backslash H)$

$$NJ_{\theta\varphi}=D\cdot\nabla\theta^N+N\theta^N J_\varphi=D\cdot\nabla\theta^N\,, \tag{B.12}$$

which follows from a trivial calculation. Hence

$$\begin{aligned}N\int_U J_{(1-\zeta)\varphi}&=\int_U D\cdot\nabla(1-\zeta)^N\\&=\int_U D\cdot\nabla[(1-\zeta)^N-(1-\zeta)]+\int_U D\cdot\nabla(1-\zeta)=-\int_{U\backslash H} D\cdot\nabla\zeta\,,\end{aligned} \tag{B.13}$$

where we have used that $(1-\zeta)^N-(1-\zeta)$ is a C^∞ function of compact support in $U\backslash H$ [cf. (B.9)]. Take ζ with the properties that $0\leqq\zeta\leqq 1$ and $\zeta=0$ on $U\backslash H_\varepsilon$. Then the left side of (B.13) is $I\equiv\sigma_N\int_{H_\varepsilon} f(\psi)J_\psi$, where $\psi=(1-\zeta)\varphi$ and $f(x)=N/\sigma_N$ for $|x|\leqq 1$ and $f(x)=0$ for $|x|>1$. (Recall that $J_\varphi\equiv 0$ on $U\backslash H_\varepsilon$.) Since $\int f=1$ and $|\psi|=1$ on ∂H_ε, we can apply (B.4) together with an approximation argument using dominated convergence, to conclude that $I=\sigma_N\deg(\varphi,H)$. □

Remark B.1. Let $U\subset\mathbb{R}^N$ be open, let $H\subset U$ be compact and let $\varphi\in C^1(U\backslash H;S^{N-1})$. Let V be open with $\bar V\subset U$ and with $H\subset V$. Assume that V is bounded and that ∂V is (piecewise) smooth. Then

$$\int_{\partial V} D\cdot\nu=\sigma_N\deg(\varphi,H), \tag{B.14}$$

where ν is the outward normal to ∂V. [To prove this, apply (B.10) to any $\zeta\in C_c^\infty(U)$ with $\zeta\equiv 1$ on V. Integrate by parts and use (B.9).] Equation (B.14) is the classical formula for the degree. Note that

$$D\cdot\nu=\det(\varphi,\varphi_{x_1},\dots,\varphi_{x_{N-1}}), \tag{B.15}$$

where $x_1,\dots,x_{N-1}$ are orthonormal coordinates in the tangent space to ∂V. On the other hand, we can think of φ restricted to ∂V as a map from the $N-1$ dimensional manifold $M\equiv\partial V$ to S^{N-1}. This map has a Jacobian determinant, which is nothing other than the right side of (B.15). Thus $\int D\cdot\nu$ can be identified as the right side of (B.4) [with $f(\psi)=1$ for $|\psi|\leqq 1$] with the integrating being over M, and not over V. Alternatively, $\int D\cdot\nu/\sigma_N$ is the number of times (including sign) that φ covers S^{N-1}.

Here are some consequences of Theorem B.1:

Theorem B.2. *Let $U\subset\mathbb{R}^N$ be open and let $H_1,H_2,\dots,H_k$ be disjoint holes in U and let $H=\bigcup_{i=1}^k H_i$. Let $\varphi\in C(U\backslash H;S^{N-1})$ with $\nabla\varphi\in L^{N-1}(U\backslash H)$. Then*

$$-\int_{U\backslash H} D\cdot\nabla\zeta=\sigma_N\sum_{i=1}^k\zeta(H_i)\deg(\varphi,H_i) \tag{B.16}$$

for every $\zeta\in C(\bar U)$ with $\nabla\zeta\in L^\infty(U)$ (in the distributional sense), $\zeta=0$ on ∂U and $\zeta=\zeta(H_i)$ is a constant on each H_i.

Theorem B.3. *Let U, H_i, and H be as in Theorem B.2. Let $\varphi\in C(\bar U\backslash H;S^{N-1})$ with $\nabla\varphi\in L^{N-1}(U\backslash H)$. Assume also that φ is constant on ∂U. Then (B.16) holds for every $\zeta\in C(U)$ with $\nabla\zeta\in L^\infty(U)$ (in the distributional sense) and $\zeta=\zeta(H_i)$ is a constant on each H_i. Note that here we do not assume that $\zeta=0$ on ∂U.*

The proofs rely on the following lemma.

Lemma B.4. *Let $V\subset\mathbb{R}^N$ be open and let $F\subset\mathbb{R}^N$ be closed with $F\subset V$. F need not be compact. Let $\varphi\in C(\bar V\backslash F;S^{N-1})$ with $\nabla\varphi\in L^{N-1}(V\backslash F)$. Assume that φ is constant on ∂V (no assumption is made if $V=\mathbb{R}^N$). Then*

$$\int_{V\backslash F} D\cdot\nabla\zeta=0 \tag{B.17}$$

for every $\zeta\in C(V)$ with $\nabla\zeta\in L^\infty(V)$ and $\zeta=0$ on F.

Proof. The intuitive reason that (B.17) holds is clear. Indeed, set $\Omega=V\backslash F$; we write

$$\int_\Omega D\cdot\nabla\zeta=\int_{\partial\Omega}(D\cdot\nu)\zeta-\int_\Omega(\operatorname{div}D)\zeta,$$

where ν denotes the outward normal on $\partial\Omega$. However, $\partial\Omega$ consists of two disjoint parts, namely ∂V and ∂F. On ∂V we have $D\cdot\nu=0$ (since φ is constant on ∂V), while on ∂F we have $\zeta=0$. On the other hand, $\operatorname{div}D=0$ on Ω [by (B.9)].

Since, in general, we do not assume that ∂V and ∂F are regular, the integration by parts is not justified and the proof becomes more delicate. First, without loss of generality, we can assume that $\zeta\in L^\infty(V)$. Otherwise, consider

$$\zeta_n(x)=\begin{cases}\zeta(x) & \text{if } |\zeta(x)|\leqq n\\ n\operatorname{sgn}\zeta(x) & \text{if } |\zeta(x)|>n.\end{cases}$$

Clearly $\int_\Omega D\cdot\nabla\zeta_n\to\int_\Omega D\cdot\nabla\zeta$ (by dominated convergence).

Second, we can also assume that ζ vanishes outside a large ball. Otherwise, consider a sequence $\zeta_n=\alpha_n\zeta$, where $\alpha_n(x)=1$ for $|x|\leqq n$, $\alpha_n(x)=2-(|x|/n)$ for $n\leqq|x|\leqq 2n$ and $\alpha_n(x)=0$ for $|x|\geqq 2n$. Again, $\int_\Omega D\cdot\nabla\zeta_n\to\int_\Omega D\cdot\nabla\zeta$ since $D\in L^1(\Omega)$.

Next, we can also assume that $\zeta=0$ on a neighborhood of F and that φ is constant on a neighborhood of ∂V. Indeed let $g:\mathbb{R}\to\mathbb{R}$ be a smooth function such that $g(t)=0$ for $|t|\leqq 1$ and $g(t)=t$ for $|t|\geqq 2$. Consider $\zeta_n(x)=\frac{1}{n}g(n\zeta(x))$. It is clear that $\zeta_n\in C(V)\cap L^\infty(V)$, ζ_n vanishes outside a large ball, $\|\nabla\zeta_n\|_{L^\infty}\leqq C\|\nabla\zeta\|_{L^\infty}$, $\zeta_n=0$ on some neighborhood of F (namely $\{x\,|\,|\zeta(x)|<1/n\}$) and $\nabla\zeta_n\to\nabla\zeta$ a.e. on V. We proceed in the same way with φ. Let $G:\mathbb{R}^N\to\mathbb{R}^N$ be defined by $G(v)_i=g(v_i)$ for all i (g as above). Let e be the value of φ on ∂V. Consider $\psi_n(x)=\frac{1}{n}G(n(\varphi(x)-e))+e$. It is easy to check that $\psi_n\in C(\bar V\backslash F;\mathbb{R}^N)$, $\psi_n=e$ in a neighborhood of ∂V [namely $\{x\,|\,|\varphi(x)-e|<1/n\}$], $\|\psi_n-\varphi\|_{L^\infty}\leqq C/n$ and $\nabla\psi_n\to\nabla\varphi$ in $L^{N-1}(V\backslash F)$.

Finally, we choose $\varphi_n=\psi_n/|\psi_n|$ (for n large enough), so that φ_n satisfies the same properties as ψ_n and, moreover, φ_n takes its values in S^{N-1}. Clearly $D_n=D(\varphi_n)\to D(\varphi)=D$ in $L^1(V\backslash F)$ and therefore $\int_\Omega D_n\cdot\nabla\zeta_n\to\int_\Omega D\cdot\nabla\zeta$ (by dominated convergence).

In conclusion, it suffices to establish (B.17) with the additional assumptions that $\zeta=0$ outside a large ball, $\zeta=0$ on a neighborhood of F and $D=0$ on a neighborhood of ∂V. Since $K=\operatorname{Supp}D\cap\operatorname{Supp}\zeta$ is a compact subset of Ω we may fix a function $\alpha\in C_c^\infty(\Omega)$ such that $\alpha\equiv 1$ on some neighborhood of K. By (B.9) we have $\int_\Omega D\cdot\nabla(\alpha\zeta)=0$, and on the other hand, $D\cdot\nabla(\alpha\zeta)=D\cdot\nabla\zeta$ a.e. on Ω (from the definition of α). □

Proofs of Theorems B.2 and B.3. If ζ_1 and ζ_2 are two admissible functions with the same values $\zeta(H_i)$ for every i, then by Lemma B.4, applied to $\zeta=\zeta_1-\zeta_2$ we have $\int_{U\backslash H} D\cdot\nabla\zeta_1=\int_{U\backslash H} D\cdot\nabla\zeta_2$ [choose $V=\mathbb{R}^3$ and $F=H\cup({}^cU)$ for Theorem B.2 and $V=U$, $F=H$ for Theorem B.3]. Thus, it suffices to prove (B.16) for one admissible ζ. Take $\zeta=\sum_{i=1}^k\zeta_i$ with each $\zeta_i=\zeta(H_i)$ near H_i and $\operatorname{Supp}\zeta_i$ is contained in a small neighborhood of H_i. Then apply Theorem B.1. □

Remark B.2. Let $U \subset \mathbb{R}^N$ be open. Assume that all the holes H_i are points a_i in U. Let D be *any* vector field in $L^1(U;\mathbb{R}^N)$. Let d_i be *any* real numbers. Then the relation

$$-\int D\cdot\nabla\zeta=\sigma_N\sum_{i=1}^{k} d_i\zeta(a_i) \tag{B.18}$$

for every $\zeta\in C(\bar{U})$ with $\nabla\zeta\in L^\infty(U)$ (in the distributional sense), $\zeta=0$ on ∂U, is equivalent to the relation

$$\operatorname{div}D=\sigma_N\sum_{i=1}^{k} d_i\delta_{a_i}\quad\text{in}\quad\mathscr{D}'(U), \tag{B.19}$$

where δ_a is the Dirac measure at $a\in\mathbb{R}^N$. [In particular the D field in (B.16) satisfies (B.19) for point holes.] Equation (B.19) looks weaker than (B.18) because the class of testing functions for (B.19) is more restrictive, namely $C_c^\infty(U)$. The equivalence of (B.18) and (B.19) follows from the following general density lemma.

Lemma B.5. *Suppose ζ is a function in $C(\bar{U})$ with $\nabla\zeta\in L^\infty(U)$ (in $\mathscr{D}'(U)$), $\zeta=0$ on ∂U and ζ is a constant on each H_i. Then there exists a sequence ζ_n in $C_c^\infty(U)$ such that $\zeta_n\to\zeta$ uniformly on every compact subset of $\bar{U}$, $\|\nabla\zeta_n\|_{L^\infty}\leqq\|\nabla\zeta\|_{L^\infty}$, $\nabla\zeta_n\to\nabla\zeta$ a.e. on U and ζ_n is a constant on each H_i.*

The proof uses the same techniques as in the proof of Lemma B.4 and therefore we shall omit it.

Remark B.3. Assume the same conditions as in Remark B.2 except that (B.18) holds for every $\zeta\in C(U)$ with $\nabla\zeta\in L^\infty(U)$, as in the setup of Theorem B.3. Then, the analogue of Remark B.2 is that (B.18) is equivalent to

$$\operatorname{div}D=\sigma_N\sum_{i=1}^{k} d_i\delta_{a_i}\quad\text{in}\quad\mathscr{D}'(U),\qquad D\cdot\nu=0,\quad\text{on}\quad\partial U, \tag{B.20}$$

where ν is the normal to ∂U. The relation $D\cdot\nu=0$ has to be interpreted in a formal sense since ∂U need not be smooth and since D is only L^1.

Appendix C: Duality for Vector Fields

We recall a classical abstract duality principle (see [12, 30, 36]).

Theorem C.1. *Let E be a Banach space and E^* its dual. Let $M\subset E$ be a linear subspace (not necessarily closed) and let Φ be a convex function from E into $(-\infty,+\infty]$ such that $\Phi(0)\neq+\infty$ and Φ is continuous at 0. Let Φ^* be the conjugate function on E^*, namely*

$$\Phi^*(f)=\sup\{\langle f,u\rangle-\Phi(u)\,|\,u\in E\}. \tag{C.1}$$

Then

$$\inf_M\Phi=-\min_{M^\perp}\Phi^*, \tag{C.2}$$

where

$$M^\perp=\{f\in E^*\,|\,\langle f,u\rangle=0\text{ for all }u\in M\}. \tag{C.3}$$

The following lemma will also be used.

Lemma C.2. *Let E be a separable Banach space and let N be a linear subspace in E^* that is sequentially closed in the weak * topology. Let*

$$N^\perp=\{u\in E\,|\,\langle f,u\rangle=0\text{ for all }f\in N\}. \tag{C.4}$$

Then $(N^\perp)^\perp=N$.

Proof. It follows easily from the Hahn-Banach theorem that $(N^\perp)^\perp$ is the weak * closure of N. To prove the lemma, therefore, it suffices to show that N is weak * closed. In view of the theorem of Banach, Dieudonné, Krein, and Smulian (see e.g. [11, Theorem V.5.7]) we have only to check that $\tilde{N}=N\cap B$ is weak * closed, where B is the unit ball in E^*. But $\tilde{N}$ is metrizable for the weak * topology (see e.g. [3, Theorem III.25]) so it suffices to note that $\tilde{N}$ is sequentially weak * closed. □

Theorem C.1 will be applied in the following two cases (A and B). In the notation of Sect. V, we take

$$E=L^1(\Omega;\mathbb{R}^N),\qquad E^*=L^\infty(\Omega;\mathbb{R}^N), \tag{C.5}$$

and

$$M_{A,B}=\{D\in E\,|\int D\cdot\nabla\zeta=0\text{ for all }\zeta\in Q_A\text{ (respectively }Q_B)\}. \tag{C.6}$$

Fix any $D^0\in\mathscr{A}_A$ (respectively $\mathscr{A}_B$) and let

$$\Phi(D)=\int_\Omega|D+D^0|. \tag{C.7}$$

Clearly,

$$E_{A,B}=\inf\{\Phi(D)\,|\,D\in M_A\text{ (respectively }M_B)\}, \tag{C.8}$$

and, for every $f\in E^*$,

$$\Phi^*(f)=-\int f\cdot D^0+\sup_{D\in E}\{\int f\cdot D-\int|D|\}=\begin{cases}-\int f\cdot D^0 & \text{if}\quad\|f\|_{L^\infty}\leqq1\\ +\infty & \text{if}\quad\|f\|_{L^\infty}>1.\end{cases} \tag{C.9}$$

Lemma C.3.

$$M_{A,B}^\perp=\{\nabla\zeta\,|\,\zeta\in Q_A\ (\textit{respectively } Q_B)\}. \tag{C.10}$$

Proof. We shall omit the A,B subscript. Let $N\subset E^*$ be the right side of (C.10). By the definition of M, $N^\perp=M$ so that $(N^\perp)^\perp=M^\perp$. We claim that N is sequentially weak * closed, whence, by Lemma C.2, $N=(N^\perp)^\perp=M^\perp$, which is precisely (C.10). To check that N is sequentially weak * closed, let ζ_n be a sequence in Q such that $\nabla\zeta_n\to f\in E^*$ in the weak * topology. We want to prove that $f=\nabla\zeta$ for some $\zeta\in Q$. By the uniform boundedness principle we know that $\|\nabla\zeta_n\|_{L^\infty}\leqq C$. We can always assume $\zeta_n(x_0)=0$ for some fixed $x_0\in U$. By Ascoli's theorem $\zeta_{n_k}\to\zeta$ uniformly on compact subsets of $\bar{U}$ (respectively U) in case A (respectively B). Clearly, $\zeta\in Q$ and $f=\nabla\zeta$. □

Applying Theorem C.1 and Lemma C.3, we find that

$$\begin{aligned}E_{A,B}&=\max\{\int\nabla\zeta\cdot D^0\,|\,\|\nabla\zeta\|_{L^\infty}\leqq1,\ \zeta\in Q_A\text{ (respectively }Q_B)\}\\&=\max\{\sigma_N\sum d_i\zeta(H_i)\,|\,\|\nabla\zeta\|_{L^\infty}\leqq1,\ \zeta\in Q_A\text{ (respectively }Q_B)\}.\end{aligned} \tag{C.11}$$

This is precisely the statement of Theorem 5.1.

Appendix D: The Basic Divergence-Free Vector Field on a Curve

Let g be a rectifiable curve in $\mathbf{R}^N$ with no self-intersection and end points a and b, $a \neq b$. Let L be its length. To be more precise the curve can be parametrized by a Lipschitz function $X(t):[0,1]\to\mathbf{R}^N$ and we can always assume that $\dot{X}(t)\neq 0$ a.e. Among the choices for $X(t)$ there is a canonical constant speed choice denoted by $X_0(t)$, so that $|\dot{X}_0(t)|=L$ a.e.

Now consider the problem of finding an $\mathbf{R}^N$-valued measure, D, on $\mathbf{R}^N$ such that

$$\operatorname{supp} D \subset g\,, \tag{D.1}$$

$$\operatorname{div} D=\delta_a-\delta_b \quad \text{in} \quad \mathscr{D}'(\mathbf{R}^N)\,. \tag{D.2}$$

Theorem D.1. *There is precisely one solution to the above problem, namely*

$$\langle D_g,\varphi\rangle \equiv \int_0^1 \varphi(X(t))\cdot \dot{X}(t)dt \tag{D.3}$$

for all $\varphi\in C_c(\mathbf{R}^N;\mathbf{R}^N)$. Here $X(t)$ denotes any parametrization of g and (D.3) *is independent of the choice of the parametrization. Moreover, $|D_g|$ is the one-dimensional Hausdorff measure of g, denoted by δ_g. In particular*

$$\int_{\mathbf{R}^N} |D_g|=L\,. \tag{D.4}$$

Proof. It is obvious that D_g given by (D.3) is independent of parametrization and satisfies (D.1).

Let us check that D_g satisfies (D.2). Choose $\zeta\in C_c^\infty(\mathbf{R}^N)$. We have

$$\langle D_g,\nabla\zeta\rangle=\int_0^1 \nabla\zeta(X(t))\cdot\dot{X}(t)dt=\int_0^1 \frac{d}{dt}\zeta(X(t))dt=\zeta(b)-\zeta(a)\,. \tag{D.5}$$

The last equality follows from the fact that Lipschitz functions are absolutely continuous. Next, we establish uniqueness. Consider $D-D_g$ and call it D, so that D satisfies

$$\operatorname{supp} D\subset g\,, \tag{D.6}$$

$$\operatorname{div} D=0 \quad \text{in} \quad \mathscr{D}'(\mathbf{R}^N)\,. \tag{D.7}$$

We have to show that $D=0$. It follows from (D.6) that there is an $\mathbf{R}^N$-valued measure, m, on $[0,1]$, such that

$$\langle D,\varphi\rangle=\int_0^1 \varphi(X(t))\cdot dm(t) \tag{D.8}$$

for all $\varphi\in C_c(\mathbf{R}^N;\mathbf{R}^N)$. The existence of m follows from the fact that for any continuous function, α, on $[0,1]$ there exists some $\varphi\in C_c(\mathbf{R}^N;\mathbf{R}^N)$ with $\varphi(X(t))=\alpha(t)$ and $\|\varphi\|=\|\alpha\|$. Thus, D can be viewed as an element of the dual of $C([0,1];\mathbf{R}^N)$, but these are measures.

Next, we claim that

$$\int_0^T \nabla\zeta(X(t))\cdot dm(t)=0 \tag{D.9}$$

for every $\zeta\in C_c^\infty(\mathbf{R}^N;\mathbf{R})$ and a.e. T. Assuming that (D.9) holds we conclude easily that $m=0$ (and so $D=0$). Indeed, by differentiating (D.9) in the sense of distributions we find

$$\nabla\zeta(X(t))\cdot m(t)=0 \quad \text{in} \quad \mathscr{D}'(0,1)\,.$$

Choosing $\zeta(x)=x_i\theta(x)$, where $\theta\in C_c^\infty(\mathbf{R}^N)$ and $\theta\equiv 1$ on some neighborhood of g, we see that $m=0$.

To establish (D.9) we fix $T\in(0,1)$ such that

$$m(\{T\})=0\,, \tag{D.10}$$

$$\text{and} \quad \dot{X}(T) \quad \text{exists and} \quad \dot{X}(T)\neq 0\,. \tag{D.11}$$

For any $\varepsilon>0$ (small enough) let $A=X([0,T])$ and $B_\varepsilon=X([T+\varepsilon,1])$. Set $d_\varepsilon=\operatorname{dist}(A,B_\varepsilon)$. There exists a function $F_\varepsilon\in C_c^\infty(\mathbf{R}^N)$ such that

$$F_\varepsilon=1 \quad \text{near } A\,, \quad F_\varepsilon=0 \quad \text{near } B\,, \tag{D.12}$$

$$|\nabla F_\varepsilon|\leq C/d_\varepsilon\,, \tag{D.13}$$

where C is a constant independent of ε. By (D.7) we have

$$0=\langle D,\nabla(\zeta F_\varepsilon)\rangle=\langle D,\zeta\nabla F_\varepsilon\rangle+\langle D,F_\varepsilon\nabla\zeta\rangle=I_1+I_2 \tag{D.14}$$

with

$$\begin{aligned}I_1=\langle D,\zeta\nabla F_\varepsilon\rangle&=\int_T^{T+\varepsilon}\zeta(X(t))\nabla F_\varepsilon(X(t))\cdot dm(t)\\ &=\int_T^{T+\varepsilon}(\zeta(X(t))-\zeta(X(T)))\nabla F_\varepsilon(X(t))\cdot dm(t)+\zeta(X(T))\int_T^{T+\varepsilon}\nabla F_\varepsilon(X(t))\cdot dm(t)\,.\end{aligned}$$

The last integral is $\langle D,\nabla F_\varepsilon\rangle=0$. We claim that $I_1\to 0$ as $\varepsilon\to 0$. Indeed

$$|I_1|\leq C\int_T^{T+\varepsilon}|X(t)-X(T)|\,|\nabla F_\varepsilon(X(t))|\,|dm(t)|\leq C\frac{\varepsilon}{d_\varepsilon}\int_T^{T+\varepsilon}|dm(t)|\,.$$

Since $\int_T^{T+\varepsilon}|dm(t)|\to 0$ as $\varepsilon\to 0$ (by dominated convergence) it suffices to check that $\varepsilon/d_\varepsilon$ remains bounded as $\varepsilon\to 0$. Suppose not. Then there exists a sequence $\varepsilon_n\to 0$ such that $d_{\varepsilon_n}/\varepsilon_n\to 0$. Thus, there are sequences $t_n\geq T+\varepsilon_n$ and $s_n\leq T$ such that $|X(t_n)-X(s_n)|/\varepsilon_n\to 0$. Clearly $t_n\to T$ and $s_n\to T$ since X is one to one. Observe that $X(t_n)-X(T)=(t_n-T)\dot{X}(T)+o(t_n-T)$ and similarly $X(s_n)-X(T)=(s_n-T)\dot{X}(T)+o(T-s_n)$. Thus,

$$\frac{1}{\varepsilon_n}(X(t_n)-X(s_n))=\frac{t_n-s_n}{\varepsilon_n}\left(\dot{X}(T)+\frac{o(t_n-T)+o(T-s_n)}{(t_n-T)+(T-s_n)}\right).$$

Since $t_n-s_n\geq\varepsilon_n$ and $\dot{X}(T)\neq 0$ we have a contradiction. Therefore $I_1\to 0$.

Next,

$$I_2=\langle D,F_\varepsilon\nabla\zeta\rangle=\int_0^T \nabla\zeta(X(t))\cdot dm(t)+\int_T^{T+\varepsilon}F_\varepsilon(X(t))\nabla\zeta(X(t))\cdot dm(t)\,.$$

The last integral is bounded by $C\int_T^{T+\varepsilon}|dm(t)|$ which goes to zero. This establishes (D.9) and hence (D.3). To prove (D.4) we use $X_0(t)$ in (D.3). First, we have

$$|\langle D_g, \varphi\rangle| \leq \int_0^1 |\varphi(X_0(t))| L dt = \langle \delta_g, |\varphi|\rangle,$$

and hence $|D_g| \leq \delta_g$. On the other hand, $\int |D_g| = L$. Indeed we have

$$\int |D_g| = \mathrm{Sup}\left\{\int_0^1 \varphi(X_0(t))\cdot \dot{X}_0(t)dt \mid \varphi \in C_c(\mathbf{R}^N;\mathbf{R}^N) \text{ with } \|\varphi\|_{L^\infty} \leq 1\right\}.$$

Let $Y_n \in C([0,1];\mathbf{R}^N)$ be a sequence of functions such that $\|Y_n\|_{L^\infty} \leq L$ and $Y_n \to \dot{X}_0$ in $L^2(0,1)$. There exists a sequence of functions, $\psi_n \in C_c(\mathbf{R}^N;\mathbf{R}^N)$ such that $\psi_n(X_0(t)) = Y_n(t)$ and $\|\psi_n\|_{L^\infty} \leq L$. Letting $\varphi_n = (1/L)\psi_n$ we have

$$\int_0^1 \varphi_n(X_0(t))\cdot \dot{X}_0(t)dt = (1/L)\int_0^1 Y_n(t)\cdot \dot{X}_0(t)dt \to L,$$

and therefore $\int |D_g| \geq L$. □

Corollary D.2. *Let everything be as in Theorem D.1 except that hypothesis* (D.2) *is replaced by*

$$\mathrm{div}\, D = 0 \quad in \quad \mathcal{D}'(\mathbf{R}^N \backslash \{a,b\}). \tag{D.15}$$

Then, there exists a constant c and two vectors A and B in $\mathbf{R}^N$ *such that*

$$D = cD_g + A\delta_a + B\delta_b \quad in \quad \mathcal{D}'(\mathbf{R}^N). \tag{D.16}$$

Proof. From (D.15) and a standard result about distributions with support on a point we have

$$\mathrm{div}\, D = \sum_\alpha c_\alpha \partial^\alpha \delta_a + \sum_\alpha c'_\alpha \partial^\alpha \delta_b \quad \text{in} \quad \mathcal{D}'(\mathbf{R}^N), \tag{D.17}$$

where the sums are finite. Since D is a measure, the right side of (D.17) contains only zeroth and first order derivatives. Since $\int \mathrm{div}\, D = 0$, the zeroth order terms have to be equal and opposite, namely $c(\delta_a - \delta_b)$. Therefore,

$$\mathrm{div}(D - cD_g) = \mathrm{div}(A\delta_a) + \mathrm{div}(B\delta_a) \tag{D.18}$$

for some vectors A and B. Transposing the right side of (D.18) to the left side and then using the uniqueness part of Theorem (D.1) we derive (D.16). □

Finally, we mention another corollary which will be used in Appendix E. Let g be a rectifiable curve in $\mathbf{R}^N$ without self-intersection and end points a and b, $a \neq b$. Let Ω be an open set such that $g\backslash\{a,b\} \subset \Omega$. Let D be an $\mathbf{R}^N$-valued bounded measure on Ω such that $\mathrm{supp} D \subset g$, and $\mathrm{div}\, D = 0$ in $\mathcal{D}'(\Omega)$.

Corollary D.3. *Under the above assumptions there exists a constant c such that*

$$D = cD_g \quad in \quad \mathcal{D}'(\Omega).$$

Proof. Extend D to all of $\mathbf{R}^N$ by 0 outside Ω. Let $\tilde{D}$ be the extension. We claim that $\mathrm{div}\,\tilde{D} = 0$ in $\mathcal{D}'(\mathbf{R}^N\backslash\{a,b\})$. Let $\zeta \in C_c^\infty(\mathbf{R}^N\backslash\{a,b\})$ and let $\theta \in C_c^\infty(\Omega)$ with $\theta = 1$ on a neighborhood of $g \cap \mathrm{Supp}\zeta$. We have

$$\int \tilde{D}\cdot \nabla\zeta = \int_\Omega D\cdot\nabla\zeta = \int_\Omega D\cdot\nabla(\theta\zeta) = 0,$$

since $\theta\zeta \in C_c^\infty(\Omega)$. We may now apply Corollary D.2 to $\tilde{D}$. □

Appendix E: Quantization and Weak Limits of Vector Fields

Let V be an open set in $\mathbf{R}^N$ with $N \geq 1$. Let φ be a map from V into S^N such that $\nabla\varphi \in L^N(V)$. Set

$$\Delta = \det(\varphi, \varphi_{x_1}, \ldots, \varphi_{x_N}). \tag{E.1}$$

Similarly let φ^n be a sequence of such maps and set

$$\Delta^n = \det(\varphi^n, \varphi^n_{x_1}, \ldots, \varphi^n_{x_N}). \tag{E.2}$$

We are concerned with the following situation. Suppose $\varphi^n \to \varphi$ a.e. and $\nabla\varphi^n$ is *bounded* in $L^N(V)$. Then Δ^n is bounded in $L^1(V)$ so that, by passing to a subsequence, we can assume that Δ^n tends to some measure μ in the weak * topology of measures. In general $\mu \neq \Delta$ unless $\nabla\varphi^n \to \nabla\varphi$ *strongly* in L^N.

If one merely assumes that $\varphi^n \to \varphi$ a.e. and $\nabla\varphi^n$ is bounded in $L^N(V)$ and if one replaces Δ^n by $|\nabla\varphi^n|^N$, for example, then we may still assume that $|\nabla\varphi^n|^N$ tends weakly to some measure ν. However, in this case one can say virtually nothing about $\nu - |\nabla\varphi|^N$. It is a striking fact that despite the lack of strong convergence it is possible to say something precise about $\mu - \Delta$. This is due to the fact that Δ has a geometric significance. Lions [23] considered maps φ with values in $\mathbf{R}^{N+1}$ instead of S^N and proved that $\mu - \Delta$ is a sum (possibly infinite) of Dirac masses but with arbitrary weights. Our result, Theorem E.1, uses the geometry of S^N and shows that there can only be finitely, many Dirac masses and that they have integer weights. Our proof is completely different from that of Lions.

A typical example is the following. Let ψ be a smooth map from $\mathbf{R}^N$ into S^N which is a constant C far out. Let $\varphi^n(x) = \psi(nx)$ so that $\varphi^n \to \varphi = C$ a.e. and $\nabla\varphi^n$ is bounded in L^N. Note that $\Delta^n \rightharpoonup \alpha\delta_0$, where $\alpha = \int_{\mathbf{R}^N} \det(\psi, \psi_{x_1}, \ldots, \psi_{x_N})dx$ and α/σ_{N+1} belongs to $\mathbf{Z}$, since α/σ_{N+1} is the degree of ψ (cf. Appendix B). This example displays a quantization feature which holds in the general setting.

Theorem E.1. *Assume* $\varphi^n \to \varphi$ a.e., $\nabla\varphi^n$ *is bounded in* $L^N(V)$ *and* $\Delta^n \rightharpoonup \mu$. *Then there exist p integers* $d_1, d_2, \ldots, d_p \in \mathbf{Z}$ *and p points* $a_1, a_2, \ldots, a_p$ *in V such that*

$$\mu - \Delta = \sigma_{N+1}\sum_{i=1}^p d_i\delta_{a_i}. \tag{E.3}$$

The proof relies on three lemmas.

Lemma E.2. *Assume Q is a cube in* $\mathbf{R}^N$ *and let* $\varphi, \bar\varphi$ *be two maps from Q to* S^N *such that* $\nabla\varphi, \nabla\bar\varphi \in L^N(Q)$ *and* $\varphi, \bar\varphi$ *restricted to* ∂Q *belong to* $W^{1,N}(\partial Q)$, *so that, in particular,* φ *and* $\bar\varphi \in C(\partial Q)$. *Then there is an integer d such that*

$$\left|\int_Q (\Delta - \bar\Delta) - \sigma_{N+1} d\right| \leq C\|\varphi - \bar\varphi\|_{L^\infty(\partial Q)}, \tag{E.4}$$

where C depends only on the norms of φ *and* $\bar\varphi$ *in* $W^{1,N-1}(\partial Q)$.

Proof. Consider the cylinder $Q \times (0,1)$ in $\mathbb{R}^{N-1}$ and its boundary Γ. Γ consists of three pieces $\Gamma_0 = \bar{Q} \times \{0\}$, $\Gamma_1 = \bar{Q} \times \{1\}$, and $\Gamma_2 = \partial Q \times [0,1]$. We recall that it follows from the density result of [32] that if $\theta \in W^{1,N}(\Gamma, S^N)$, then

$$\frac{1}{\sigma_{N+1}} \int_\Gamma \det(\theta, \theta_{x_1}, \ldots, \theta_{x_N}) \in \mathbb{Z}, \tag{E.5}$$

where $x_1, x_2, \ldots, x_N$ are orthonormal coordinates in the tangent space to Γ [cf. (B.14)]. Let $\tilde{\theta}(x,t) = t\varphi(x) + (1-t)\bar{\varphi}(x)$, $x \in \bar{Q}$, $t \in [0,1]$, and let $\theta = \tilde{\theta}/|\tilde{\theta}|$. Note that θ is well defined, at least if $\|\varphi - \bar{\varphi}\|_{L^\infty(\partial Q)} < 1/2$; otherwise, the conclusion is trivial. Also, $|\tilde{\theta}| > 1/2$ everywhere on Γ. Clearly,

$$\int_{\Gamma_0 \cup \Gamma_1} \det(\theta, \theta_{x_1}, \ldots, \theta_{x_N}) = \int_Q (\Delta - \bar{\Delta}).$$

Now we estimate $\int_{\Gamma_2} \det(\theta, \theta_{x_1}, \ldots, \theta_{x_N})$. Observe that

$$\det(\theta, \theta_{x_1}, \ldots, \theta_{x_N}) = \frac{1}{|\tilde{\theta}|^{N+1}} \det(\tilde{\theta}, \tilde{\theta}_{x_1}, \ldots, \tilde{\theta}_{x_N})$$

and $\tilde{\theta}_t = \varphi - \bar{\varphi}$. Since we are now on Γ_2, one of the x_j may be taken to be t. Therefore,

$$\left| \int_{\Gamma_2} \det(\theta, \theta_{x_1}, \ldots, \theta_{x_N}) \right| \le C \|\varphi - \bar{\varphi}\|_{L^\infty(\partial Q)},$$

where C depends only on $W^{1,N-1}(\partial Q)$. □

Remark E.1. Clearly, Lemma E.2 extends to domains other than cubes under appropriate assumptions on the regularity of the boundary.

For every $h > 0$, set

$$Q_h = \{x \in \mathbb{R}^N \mid |x_i| < h/2,\ i = 1, \ldots, N\}.$$

Lemma E.3. *Let f_n be a sequence of functions on V which is bounded in $L^1(V)$. Let $h > 0$. Then, for* a.e. $a \in \mathbb{R}^N$ *there is a subsequence f_{n_k} (depending on a) such that f_{n_k} restricted to $(a + \partial Q_h) \cap V$ is bounded in $L^1((a + \partial Q_h) \cap V)$.*

Proof. We consider only the case where $N = 2$ since the argument is the same in the general case. Extend f_n by zero outside V and for a.e. $y \in \mathbb{R}$ set

$$g_n(y) = \int_{\mathbb{R}} (|f_n(x,y)| + |f_n(x, y+h)|)\,dx.$$

Note that

$$\int_0^h g_n(y)\,dy \le \int_V |f_n(x,y)|\,dx\,dy \le C.$$

Applying Fatou's lemma we deduce that $\liminf g_n(y) < \infty$ for a.e. $y \in \mathbb{R}$. Similarly, if we reverse x and y. Therefore, for a.e. $a \in \mathbb{R}^2$, there is a subsequence f_{n_k} (depending on a) such that f_{n_k} restricted to $a + \partial Q_h$ is bounded in $L^1(a + \partial Q_h)$. □

Lemma E.4. *Let λ_n be a sequence of measures on V such that $\lambda_n \rightharpoonup \lambda$ and $|\lambda_n| \rightharpoonup \nu$ weakly in the sense of measures. Let Q be an open cube such that $\bar{Q} \subset V$ and $\nu(\partial Q) = 0$. Then $\lambda_n(Q) \to \lambda(Q)$.*

The proof is straightforward; approximate characteristic functions by continuous functions. □

Proof of Theorem E.1. Without loss of generality we may assume that $|\Delta^n| \rightharpoonup \nu$ weakly in the sense of measures. We shall say that an open cube Q is a *good* cube if $\bar{Q} \subset V$ and Q satisfies the following properties:

(i) there is a subsequence φ^{n_k} (depending on Q) such that $\nabla\varphi^{n_k}$ restricted to ∂Q is bounded in $L^N(\partial Q)$,

(ii) $\nu(\partial Q) = 0$,

(iii) $\varphi^n \to \varphi$ a.e. on ∂Q.

The proof consists of three steps.

Step 1: For every good cube Q one has

$$\frac{1}{\sigma_{N+1}} \left(\mu(Q) - \int_Q \Delta \right) \in \mathbb{Z}.$$

Indeed, $\nabla\varphi^{n_k}$ is bounded in $L^N(\partial Q)$ and therefore $\varphi^{n_k} \to \varphi$ in $L^\infty(\partial Q)$ (by the Morrey-Sobolev imbedding theorem). Applying Lemma E.2 we see that there exists a sequence of integers d_k such that

$$\left| \int_Q (\Delta^{n_k} - \Delta) - \sigma_{N+1} d_k \right| \to 0.$$

The conclusion follows since, by Lemma E.4, we have $\int_Q \Delta^{n_k} \to \mu(Q)$.

Step 2: $\frac{1}{\sigma_{N+1}} \mu(\{a\}) \in \mathbb{Z}$ for every $a \in V$. Let Q_j be a sequence of good cubes such that $a \in Q_j$ for all j and $|Q_j| \to 0$. Such a sequence exists by Lemma E.3 applied to $f_n = |\nabla\varphi^n|^N$ [for (ii) and (iii) the argument is standard]. We know from Step 1 that, for all j,

$$\frac{1}{\sigma_{N+1}} \left(\mu(Q_j) - \int_{Q_j} \Delta \right) = d_j \in \mathbb{Z}.$$

Finally, we let $j \to \infty$ and conclude, using the fact that $\int_{Q_j} \Delta \to 0$.

It follows from Step 2 that μ has only finitely many atoms. The atomic part of μ will be denoted by $\sigma_{N+1} \sum_{i=1}^p d_i \delta_{a_i}$ with $d_i \in \mathbb{Z}$ and $a_i \in V$.

Step 3. Let $m = \mu - \Delta - \sigma_{N+1} \sum_{i=1}^p d_i \delta_{a_i}$. We claim that $m = 0$.

Indeed, by Step 1, we know that $\sigma_{N+1}^{-1} m(Q) \in \mathbb{Z}$ for every good cube Q. Let V' be an open set with compact closure in V. Since m has no atoms there is some $\varepsilon > 0$ such that $m(Q) = 0$ for every good cube, Q, with $|Q| < \varepsilon$ and $Q \cap V' \neq \emptyset$ (the argument is by contradiction). Let $h > 0$ be such that $h^N < \varepsilon$ and $h < \operatorname{dist}(V', \partial V)$. Then $\chi_{Q_h} * m = 0$ in $\mathcal{D}'(V')$, since $m(x - Q_h) = 0$ for a.e. $x \in V'$ (note that $x - Q_h$ is a good cube for a.e. $x \in V'$, by Lemma E.3).

On the other hand, $h^{-N} \chi_{Q_h} * m \rightharpoonup m$ as $h \to 0$ and therefore $m = 0$ in V'. □

Corollary E.5. *Let φ^n be a sequence of maps from S^N into S^N satisfying the same assumptions as in Theorem E.1. Then the same conclusion,* (E.3), *holds. (In* (E.1) *and* (E.2) *one has to interpret the x_i as orthonormal coordinates on S^N.)*

Proof. Use two stereographic projections (for example north and south poles) and note that the measure Δdx is invariant under diffeomorphisms. □

Finally, we consider the situation in which there is a sequence of continuous maps φ^n from $\Omega\subset\mathbb{R}^N$ to S^{N-1} $(N\geqq 3)$ with $\nabla\varphi^n\in L^{N-1}(\Omega)$. Associated with each φ^n is a vector field D^n given by (B.7). Let us suppose that $\nabla\varphi^n$ remains bounded in $L^{N-1}(\Omega)$ so that D^n is bounded in $L^1(\Omega)$, and thus we may assume that $D^n\rightharpoonup D$ weakly in the sense of measure. Let us suppose that

$$\operatorname{supp} D\subset g, \tag{E.6}$$

where g is a rectifiable curve in Ω without self-intersections. I.e. there is a Lipschitz map $X:[0,1]\to\bar\Omega$ which is injective and such that $X((0,1))\subset\Omega$. Since $\operatorname{div}D^n=0$ [see (B.9)] it follows that $\operatorname{div}D=0$ in $\mathcal{D}'(\Omega)$ and thus, by Corollary D.3,

$$D=cD_g, \tag{E.7}$$

where D_g is given by (D.3). Appendix D only tells us that c in (E.7) is some constant, but the fact that φ^n takes values in S^{N-1} leads to the following

Theorem E.5. *Under the conditions on φ^n just stated, the constant c in* (E.7) *is an integer multiple of σ_N.*

Proof. Without loss of generality we may assume that $|D^n|\rightharpoonup\nu$ weakly in the sense of measures (in general, $\operatorname{supp}\nu$ need not be contained in g). Consider, as in Appendix D, the canonical parametrization, $X(t)$, of g and fix some $T\in(0,1)$ such that $\dot X(T)$ exists, $v\equiv\dot X(T)\neq 0$ and also $\nu(\{X(T)\})=0$. Set $a=X(T)$.

We wish to find a hyperplane Π through a with the following properties:

(i) $v\notin\Pi-a$,

(ii) $|\nabla\varphi^n|$ restricted to Π is uniformly bounded in $L^{N-1}(\Pi\cap\Omega)$,

(iii) $\nu(\Pi)=0$.

This construction is possible – indeed (i), (ii), and (iii) hold for almost every Π. Using (i) we can find $r>0$ (small enough) so that

$$g\cap\Pi\cap B(a,r)=\{a\}. \tag{E.8}$$

Indeed suppose not; then there exists a sequence $t_n\in(0,1)$ such that $X(t_n)\in\Pi$, $X(t_n)\neq a$, and $X(t_n)\to a$. We may always assume that $t_n\to t\in[0,1]$ and, since X is injective, we must have $t=T$. On the other hand, $(t_n-T)^{-1}(X(t_n)-X(T))\in\Pi-a$, and at the limit we find $v\in\Pi-a$; this contradicts (i). Further, we may also assume that $B(a,r)\subset\Omega$. Let ζ be a smooth function such that $\zeta=1$ on $B(a,r/2)$ with support in $B(a,r)$. Let H be the open half-space determined by Π and which contains $a-v$, and let $\tilde v$ be the outward normal to H. We have

$$\int_H D^n\cdot\nabla\zeta=\int_\Pi(D^n\cdot\tilde v)\zeta. \tag{E.9}$$

Using (ii) and Theorem E.1, we know that (for some subsequence still denoted D^n)

$$D^n\cdot\tilde v\rightharpoonup f+\sigma_N\sum d_i\delta_{a_i} \tag{E.10}$$

with $f\in L^1(\Pi\cap\Omega)$ and $d_i\in\mathbb{Z}$. The reason that we can apply Theorem E.1 is the following. Since $|\nabla\varphi^n|$ restricted to Π is bounded in $L^{N-1}(\Pi\cap\Omega)$ and $N\geqq 3$, it follows that, for some subsequence, φ^n converges a.e. (on $\Pi\cap\Omega$) to some limit ψ and $\nabla\psi\in L^{N-1}(\Pi\cap\Omega)$. Note that this may fail when $N=2$. (The case $N=2$ is special and will be examined subsequently.) We may always choose r so small that $B(a,r)$ contains at most one a_i, namely a. Let d be the coefficient of δ_a in (E.10). From (E.9) we have

$$\int_H D^n\cdot\nabla\zeta\to\sigma_N d+\int_\Pi f\zeta. \tag{E.11}$$

On the other hand, by (iii) and Lemma E.4 we see that

$$\int_H D^n\cdot\nabla\zeta\to\int_H D\cdot\nabla\zeta. \tag{E.12}$$

We claim that

$$\int_H D\cdot\nabla\zeta=c, \tag{E.13}$$

where c is the constant introduced in (E.7). To prove this, let us assume there exists a $0<T_1<T$ and a radius r such that

$$\begin{aligned}&\text{(i)} && X(t)\in H \quad\text{for}\quad T_1\leqq t<T,\\ &\text{(ii)} && X(t)\notin B(a,r)\cap H \quad\text{if}\quad t\notin[T_1,T).\end{aligned} \tag{E.14}$$

If this is so then, with $\tau=\{t\mid X(t)\in B(a,r)\cap H\}$, it is easy to see that

$$\begin{aligned}\int_H D\cdot\nabla\zeta&=c\int_\tau\nabla\zeta(X(t))\cdot\dot X(t)dt\\ &=c\int_{T_1}^{T}\nabla\zeta(X(t))\cdot\dot X(t)dt=c[\zeta(X(T))-\zeta(X(T_1))]=c.\end{aligned} \tag{E.15}$$

The theorem follows from (E.15) and (E.11) by letting $r\to 0$, so that the integral in (E.11) goes to zero.

Now to prove that (E.14) can be satisfied observe that X is differentiable at T so that $X(t)=X(T)+v(t-T)+o(t-T)$, so that (i) is satisfied for $t<T$ and $T-t<\alpha$ for some α. Likewise, if $\beta>t-T>0$ then $X(t)\notin H$. The curve $X(t)$ for $1\geqq t\geqq\beta+T$ is closed and therefore has a positive distance from the point a. Call it δ_+. Likewise $|X(t)-a|\geqq\delta_->0$ for $0\leqq t\leqq T-\alpha$. Choose $r<\min(\delta_+,\delta_-)$. For $t\geqq T$ either $X(t)\notin H$ or $|X(t)-a|>r$. For $t<T$, either $X(t)\in H$ or $|X(t)-a|>r$. This accomplishes (E.14). □

We turn now to the case $N=2$ which is not covered by Theorem E.5. Suppose φ^n is a sequence of continuous maps from $\Omega\subset\mathbb{R}^2$ to S^1 with $\nabla\varphi^n\in L^1(\Omega)$. Let us suppose that $\nabla\varphi^n$ remains bounded in $L^1(\Omega)$ so that D^n is bounded in $L^1(\Omega)$, and thus we may assume that $D^n\rightharpoonup D$ weakly in the sense of measures. Let us suppose, as above that $\operatorname{Supp}D\subset g$, and therefore, for some constant, c, we have

$$D=cD_g. \tag{E.16}$$

Theorem E.6. *Under the conditions on φ^n just stated, and also that $\varphi^n\to C$* a.e. *on Ω, where C is a constant, then the constant c in* (E.16) *is an integer multiple of $\sigma_2=2\pi$.*

The proof is the same as the proof of Theorem E.5 and we shall omit it. The assumption $\varphi^n \to C$ a.e. is essential, as the following simple case shows. Let Ω be the disk $\{x \in \mathbb{R}^2 \mid |x| < 1\}$ and let $g = \{(x_1, x_2) \mid x_1 = 0, |x_2| \leq 1\}$ be a diameter. Let $f: \mathbb{R} \to \mathbb{R}$ be any smooth function with $f' \in L^1(\mathbb{R})$. The sequence $\varphi^n((x_1, x_2)) = (\cos f(nx_1), \sin f(nx_1))$ has all the right properties except that φ^n converges to two *different* constants for $x_1 > 0$ and $x_1 < 0$ [provided $f(+\infty) - f(-\infty)$ is not an integer multiple of 2π]. On the other hand, the limiting D field is cD_g with $c = f(+\infty) - f(-\infty)$.

Acknowledgements. The authors are grateful to J. Ericksen for suggesting the problem which we call Example 1 and which led us to undertake this study. They are also grateful to D. Kinderlehrer for useful discussions and especially for drawing our attention to the problems considered in Sect. VII. Thanks are also due to C. Berge, P. Cladis, R. Dudley, M. Gromov, R. Gulliver, R. Hardt, M. Las Vergnas, G. Strang, and D. Sullivan. E.H.L. gratefully acknowledges the hospitality of I.H.E.S. as does H.B. for the hospitality of the IMA at the University of Minnesota and of MIT, where part of this work was carried out.

The main results of this paper were announced in [5].

References

1. Adams, R.A.: Sobolev spaces. New York: Academic Press 1975
2. Birkhoff, G.: Tres observaciones sobre el algebra lineal. Univ. Nac. Tucuman Revista A. **5**, 147–151 (1946); Math. Rev. **8**, 561 (1947)
3. Brezis, H.: Analyse fonctionnelle. Paris: Masson 1983
4. Brezis, H., Coron, J.-M.: Large solutions for harmonic maps in two dimensions. Commun. Math. Phys. **92**, 203–215 (1983)
5. Brezis, H., Coron, J.-M., Lieb, E.H.: Estimations d'énergie pour des applications de $\mathbb{R}^3$ à valeurs dans S^2. C.R. Acad. Sci. Paris **303**, 207–210 (1986)
6. Brinkman, W.F., Cladis, P.E.: Defects in liquid crystals. Phys. Today, May 1982, pp. 48–54
7. Chandrasekhar, S.: Liquid crystals. Cambridge: Cambridge University Press 1977
8. Cohen, R., Hardt, R., Kinderlehrer, D., Lin, S.-Y., Luskin, M.: Minimum energy configurations for liquid crystals: Computational results, to appear in ref. [14]
9. De Gennes, P.G.: The physics of liquid crystals. Oxford: Clarendon Press 1974
10. Dudley, R.M.: Probabilities and metrics. Aarhus Universitet, Matematisk Institut Lecture Notes Series n° 45 (1976)
11. Dunford, N., Schwartz, J.T.: Linear operators, Vol. I. New York: Interscience 1964
12. Ekeland, I., Temam, R.: Analyse convexe et problèmes variationnels. Paris: Dunod, Gauthier-Villars 1974
13. Ericksen, J.L.: Equilibrium theory of liquid crystals. In: Advances in liquid crystals, Vol. 2. Brown, G.H. (ed.). New York: Academic Press 1976, pp. 233–299
14. Ericksen, J.L., Kinderlehrer, D. (ed.): Proceedings I.M.A. workshop on the theory and applications of liquid crystals (to appear)
15. Fernique, X.: Sur le théorème de Kantorovitch-Rubinstein dans les espaces polonais. In: Sem. Probabilités XV. Lecture Notes in Mathematics, No. 850, Azema-Yor (ed.). Berlin, Heidelberg, New York: Springer 1981
16. Hamidoune, Y.O., Las Vergnas, M.: Local edge-connectivity in regular bipartite graphs (to appear)
17. Hardt, R., Kinderlehrer, D.: Mathematical questions of liquid crystal theory. To appear in ref. [14]
18. Hardt, R., Kinderlehrer, D., Lin, F.H.: Existence and partial regularity of static liquid crystal configurations. Commun. Math. Phys. **105**, 541–570 (1986)

19. Hardt, R., Kinderlehrer, D., Lin, F.H.: In preparation
20. Kantorovich, L.V.: On the transfer of masses. Dokl. Akad. Nauk SSSR **37**, 227–229 (1942)
21. Kléman, M.: Points, lignes, parois, Vol. I and II. Orsay: Les editions de physique 1977
22. Lemaire, L.: Applications harmoniques de surfaces Riemanniennes. J. Differ. Geom. **13**, 51–78 (1978)
23. Lions, P.L.: The concentration-compactness principle in the calculus of variations. The limit case. Riv. Mat. Iberoamericana **1**, 45–121 and 145–201 (1985)
24. Lloyd, N.G.: Degree theory. Cambridge: Cambridge University Press 1978
25. Meyers, N., Serrin, J.: $H = W$. Proc. Nat. Acad. Sci. USA **51**, 1055–1056 (1964)
26. Minc, H.: Permanents, Encyclopedia of Math. and Appl., Vol. 6. Reading, MA: Addison-Wesley 1978
27. Nirenberg, L.: Topics in nonlinear functional analysis. New York: New York University Lecture Notes 1974
28. Parisi, G.: Quark imprisonment and vacuum repulsion. Phys. Rev. D **11**, 970–971 (1975)
29. Rachev, S.T.: The Monge-Kantorovich mass transference problem and its stochastic applications. Theory Probab. Appl. **29**, 647–676 (1985)
30. Rockafellar, R.T.: Extension of Fenchel's duality theorem for convex functions. Duke Math. J. **33**, 81–90 (1966)
31. Schoen, R., Uhlenbeck, K.: A regularity theory for harmonic maps. J. Differ. Geom. **17**, 307–335 (1982)
32. Schoen, R., Uhlenbeck, K.: Boundary regularity and the Dirichlet problem for harmonic maps. J. Differ. Geom. **18**, 253–268 (1983)
33. Simon, L.: Asymptotics for a class of non-linear evolution equations, with applications to geometric problems. Ann. Math. **118**, 525–571 (1983)
34. Spanier, E.H.: Algebraic topology. New York: McGraw-Hill 1966
35. Springer, G.: Introduction to Riemann surfaces. Reading, MA: Addison-Wesley 1957
36. Strang, G.: L^1 and L^∞ approximation of vector fields in the plane. In: Lecture Notes in Num. Appl. Anal. **5**, 273–288 (1982)
37. Whitney, H.: Geometric integration theory. Princeton, NJ: Princeton University Press 1957

Communicated by A. Jaffe

Received July 14, 1986

EQUATIONS AUX DERIVEES PARTIELLES.

Une remarque sur l'application $x/|x|$. Note à langue dominante anglaise, de Fang Hua LIN

présentée par Haïm BREZIS.

On établit que $\psi(x) = x/|x|$ est une application harmonique minimisante de $\mathbb{R}^n$ dans S^{n-1} pour tout $n \geq 3$.

PARTIAL DIFFERENTIAL EQUATIONS.

A remark on the map $x/|x|$ (mostly in English language).

We prove that $\psi(x) = x/|x|$ is a minimizing harmonic map from $\mathbb{R}^n$ into S^{n-1} for all $n \geq 3$.

Il est facile de vérifier que l'application $\psi(x) = x/|x| : \mathbb{R}^n \to S^{n-1}$ est une application harmonique pour tout $n \geq 3$, c'est-à-dire vérifie l'équation $-\Delta\psi = \psi|\nabla\psi|^2$. La question était posée de savoir si cette application est harmonique minimisante, c'est-à-dire si ψ réalise le minimum de $\int|\nabla U|^2$ parmi toutes les applications $U : B^n \to S^{n-1}$ telles que $U(x) = x$ sur ∂B^n (où B^n désigne la boule unité de $\mathbb{R}^n$). La réponse affirmative était connue pour $n \geq 7$ (voir [4] où il est même prouvé que ψ réalise le minimum de l'énergie parmi toutes les applications de B^n dans S^n) et pour $n = 3$ (voir [1]). Nous montrons que la réponse est affirmative pour tout $n \geq 3$. Notre méthode est basée sur l'inégalité ponctuelle suivante

$$|\nabla U|^2 + \frac{1}{n-2}\,[\mathrm{tr}(\nabla U)^2 - (\mathrm{div}\, U)^2] \geq 0$$

pour toute application $U : \mathbb{R}^n \to S^{n-1}$. On utilise ensuite l'identité

$$\int_{B^n} [(\mathrm{div}\, U)^2 - \mathrm{tr}(\nabla U)^2]dx = (n-1) \text{ Volume de } S^{n-1}$$

pour toute application $U : B^n \to S^{n-1}$ telle que $U(x) = x$ sur ∂B^n.

It is not hard to see that $\psi(x) = x/|x| : \mathbb{R}^n \to S^{n-1}$ for $n \geq 3$, defines a harmonic map from $\mathbb{R}^n$ to S^{n-1}. The question we shall discuss is whether it is an energy minimizing harmonic map or not.

W. Jäger and H. Kaul [4] have shown that the map $x/|x| : \mathbb{R}^n \to S^n$ is an energy minimizing harmonic map if and only if $n \geq 7$. In particular ψ is an energy minimizing harmonic map from $\mathbb{R}^n$ into S^{n-1} for $n \geq 7$. We should also mention an interesting work of R. Schoen and K. Uhlenbeck ([5], [6]) and, independently, of M. Giaquinta and J. Soucek [2], on the minimizing maps into the sphere or a hemisphere. Recently, H. Brezis, J.M. Coron and E. Lieb [1] have shown that any nonconstant minimizing harmonic map from $\mathbb{R}^3$ to S^2 which is homogeneous of degree zero must be of the form $\pm R(x/|x|)$ for some rotation R of $\mathbb{R}^3$. In particular, the map $x/|x| : \mathbb{R}^3 \to S^2$ is energy minimizing.

Here we prove the following:

Theorem: *The map $x/|x| : \mathbb{R}^n \to S^{n-1}$ is an energy minimizing harmonic map.*

The technique we use in the proof of the above theorem is so called "adding a null Lagrangian". This has been used already in [3] in proving the main existence theorem there. This was also widely used in finding optimal bounds for effective coefficients in the theory of homogenization and composite materials. The author wishes to thank M. Avellaneda and G. Milton for interesting discussions on several problems in finding optimal bounds of composite materials.

Lemma 1: *Let $U : \mathbb{R}^n \to S^{n-1}$ $(n \geq 3)$ be a C^1 map in a neighborhood of $Q \in \mathbb{R}^n$. Then we have*

$$P(\nabla U)(Q) = |\nabla U|^2(Q) + \frac{1}{n-2}\,[\mathrm{tr}(\nabla U)^2(Q) - (\mathrm{div}\, U)^2(Q)] \geq 0$$

Proof. Since $P(\nabla U)$ is frame invariant, i.e. $P(\nabla U) = P(\nabla V)$ if $V = R^t \circ U \circ R$ where R is a rotation in $\mathbb{R}^n$, we can assume that $U(Q) = (U^1(Q), U^2(Q), \ldots, U^n(Q)) = (0,0,\ldots,1)$. Hence $\nabla U^n(Q) = 0$ as U maps $\mathbb{R}^n$ into S^{n-1}. Let $a_{ij} = \partial U^i/\partial x_j(Q)$ so that $a_{nn} = 0$. We have

$$(\operatorname{div} U)^2(Q) = \left(\sum_{i=1}^{n-1} a_{ii}\right)^2 \le (n-1)\sum_{i=1}^{n-1} a_{ii}^2$$

$$\operatorname{tr}(\nabla U)^2(Q) = \sum_{i,j=1}^{n} a_{ij}a_{ji} \ge \sum_{i=1}^{n-1} a_{ii}^2 - \sum_{i\neq j} a_{ij}^2$$

and

$$|\nabla U|^2(Q) = \sum_{i=1}^{n-1} a_{ii}^2 + \sum_{i\neq j} a_{ij}^2 .$$

It follows that

$$(n-2)\,|\nabla U|^2(Q) + \operatorname{tr}(\nabla U)^2(Q) - (\operatorname{div} U)^2(Q) \ge (n-3)\sum_{i\neq j} a_{ij}^2 .$$

Lemma 2. *Let* B^n *be the unit ball of* $\mathbb{R}^n$ *and let* $U \in H^1(B^n, S^{n-1})$ *with* $U(x) = x$ *on* $\partial B^n = S^{n-1}$. *Then*

$$\int_{B^n} [(\operatorname{div} U)^2 - \operatorname{tr}(\nabla U)^2]\,dx = (n-1)\,|S^{n-1}|$$

where $|S^{n-1}|$ *denotes the volume of* S^{n-1}.

Proof. By the identity

$$(\operatorname{div} U)^2 - \operatorname{tr}(\nabla U)^2 = \operatorname{div}[(\operatorname{div} U)U - (\nabla U)U]$$

and the divergence theorem one has

$$\int_{B^n} [(\operatorname{div} U)^2 - \operatorname{tr}(\nabla U)^2]\,dx = \int_{S^{n-1}} (\operatorname{div} U)U.\nu - \int_{S^{n-1}} (\nabla U)U.\nu .$$

Since $U(x) = x$ on S^{n-1} we have $(\nabla U)U.\nu = U_\nu.U = 0$ as U maps into S^{n-1}. Also $\operatorname{div} U = n-1$ on S^{n-1}. This yields the conclusion of Lemma 2.

Proof of the Theorem. From Lemma 1 and Lemma 2 we conclude that, for any $U \in H^1(B^n, S^{n-1})$, with $U(x) = x$ on ∂B^n and U is C^1 a.e. in B^n that

$$\int_{B^n} |\nabla U|^2 \ge \frac{n-1}{n-2}\,|S^{n-1}| .$$

On the other hand

$$\int_{B^n} |\nabla(x/|x|)|^2 = \frac{n-1}{n-2}\,|S^{n-1}| .$$

It follows that $x/|x|$ is an energy minimizing map.

Remark. If $\psi(x) = x/|x|$, then $P(\nabla\psi) = 0$. The energy minimizing property of ψ can be also easily seen from this fact.

References

[1] H. Brezis, J.M. Coron and E. Lieb, Comm. Math. Phys. 107 (1986), pp. 649-705.

[2] M. Giaquinta and J. Soucek, Ann. Sc. Norm. Sup. Pisa, 12 (1985), pp. 81-90.

[3] R. Hardt, D. Kinderlehrer and F.H. Lin, Comm. Math. Phys. 105 (1985), pp. 547-570.

[4] W. Jäger and H. Kaul, J. Reine Angew. Math. 343 (1983), pp. 146-161.

[5] R. Schoen and K. Uhlenbeck, Invent. Math. 78 (1984), pp. 89-100.

[6] R. Schoen and K. Uhlenbeck, J. Diff. Geom. 17 (1982), pp. 307-335.

Courant Institute
New York University
251 Mercer Street
New York, NY 10012
U.S.A.

CALCUL DES VARIATIONS. — *Minima de la fonctionnelle énergie libre des cristaux liquides.* Note de Frédéric Hélein.

On cherche à minimiser la fonctionnelle $E(u) = 1/2 \int_B \{ K_1 (\operatorname{div} u)^2 + K_2 (u.\operatorname{rot} u)^2 + K_3 (u \wedge \operatorname{rot} u)^2 \} dx$ où u est une application de la boule unité B de R^3 dans la sphère unité S^2 de R^3 qui vérifie $u(x) = x/|x|$ sur ∂B, et où K_1, K_2, K_3 sont des constantes réelles strictemnt positives. On montre que $u_*(x) = x/|x|$ ne minimise pas E si $8(K_2 - K_1) + K_3 < 0$. On exhibe une application qui minimise E dans le cas dégénéré $K_2 = K_3 = 0$.

CALCULUS OF VARIATIONS. — Minima of the free energy functional of liquid crystals.

We try to minimize the functional $E(u) = 1/2 \int_B \{ K_1 (\operatorname{div} u)^2 + K_2 (u.\operatorname{rot} u)^2 + K_3 (u \wedge \operatorname{rot} u)^2 \} dx$ where u is a map from the unit ball B of R^3 into the unit sphere S^2 of R^3 which verifies $u(x) = x/|x|$ on ∂B, and where K_1, K_2, K_3 are strictly positiv constants. We show that $u_*(x) = x/|x|$ does not minimize E if $8(K_2 - K_1) + K_3 < 0$.We exhibit a map which minimizes E in the degenerate case $K_2 = K_3 = 0$.

$B = \{ x \in R^3 / |x| = R \leq 1 \}$ est la boule unité de R^3, $S^2 = \{ x \in R^3 / |x| = 1 \}$ la sphère unité de R^3. On représente un point x de B par ses coordonnées orthonormées (x_1, x_2, x_3).On se place dans l'espace $H^1(B, S^2)$ des applications u de $H^1(B, R^3)$ qui vérifient $u(x) \in S^2$ p.p. .On note u_* l'élément de $H^1(B, S^2)$ défini par $u_*(x) = x/|x| = x/R$.

On pose $\mathcal{A}(u_*) = \{ u \in H^1(B, S^2) / u(x) = u_*(x)$ pour p. t. $x \in \partial B \}$.

Pour K_1, K_2, K_3 réels strictement positifs, on définit la fonctionnelle

$$E(u) = 1/2 \int_B [K_1 (\operatorname{div} u)^2 + K_2 (u.\operatorname{rot} u)^2 + K_3 (u \wedge \operatorname{rot} u)^2] dx$$

Si on modélise un cristal liquide nématique occupant un volume B de l'espace par une application u de $H^1(B, S^2)$, E représente l'énergie libre de ce cristal. On cherche l'application qui minimise E dans $\mathcal{A}(u_*)$. On pose $I(K_1, K_2, K_3) = \operatorname{Inf} \{ E(u) / u \in \mathcal{A}(u_*) \}$.

On sait d'après R. Hardt, D. Kinderlehrer, F. H. Lin [2] qu'une telle application existe et est analytique réelle en dehors d'un ensemble de mesure de Hausdorf monodimensionnelle nulle. H. Brézis, J.-M. Coron, E. Lieb ont montré dans [1] que dans le cas $K_1 = K_2 = K_3$, u_* était l'unique solution de ce problème, et F. H. Lin a étendu ce résultat au cas $K_2 \geq K_1$ [4]. Nous allons montrer que cela n'est pas toujours le cas.

THEOREME 1.- *Si $8(K_2 - K_1) + K_3 < 0$, alors u_* n'est pas minimisant.*

Démonstration. — Pour $v \in H^1_0(B, R^3) \cap L^\infty(B, R^3)$ et $\lambda > 0$ suffisamment petit, on pose

$$u_\lambda = (u_* + \lambda v) / |u_* + \lambda v| \in H^1(B, S^2).$$

Alors $u_\lambda = u_* + \lambda w_1 + \lambda^2 w_2 + o(\lambda^2)$ avec, en notant v_∞ la composante orthogonale à u_* de v, et $u_R = v.u_*$,

$$w_1 = v - (u_*.v) u_* = v_\infty$$
$$w_2 = [(3 (u_*.v)^2 - v^2) u_* - 2 (u_*.v) v] / 2 = [(3 v_R^2 - v^2) u_* - 2 v_R v] / 2.$$

D'où $E(u_\lambda) = E(u_*) + \lambda E_1(u_*).v + \lambda^2/2\, E_2(u_*).(v, v) + o(\lambda^2)$ avec

$$E_1(u_*).v = \int_B [- K_1 \operatorname{grad}(\operatorname{div} u_*).w_1 + K_2 (u_*.\operatorname{rot} u_*)(u_*.\operatorname{rot} w_1 + w_1.\operatorname{rot} u_*)$$
$$+ K_3 (u_* \wedge \operatorname{rot} u_*).(u_* \wedge \operatorname{rot} w_1 + w_1 \wedge \operatorname{rot} u_*)] dx$$

$$E_2(u_*).(v, v) = \int_B \{ K_1 [(\operatorname{div} w_1)^2 - 2 \operatorname{grad}(\operatorname{div} u_*).w_2]$$
$$+ K_2 [(u_*.\operatorname{rot} w_1 + w_1.\operatorname{rot} u_*)^2 + 2 (u_*.\operatorname{rot} u_*)(u_*.\operatorname{rot} w_2 + w_2.\operatorname{rot} u_*)]$$
$$+ K_3 [(u_* \wedge \operatorname{rot} w_1 + w_1 \wedge \operatorname{rot} u_*)^2 + 2 (u_* \wedge \operatorname{rot} u_*).(u_* \wedge \operatorname{rot} w_2 + w_2 \wedge \operatorname{rot} u_*)] dx$$

Utilisant $\operatorname{grad}(\operatorname{div} u_*) = -2 R^{-2} u_*$, $u_*.w_2 = - v_\infty^2/2$, et $\operatorname{rot} u_* = 0$, il vient

$$E_1(u_*).v = 0$$
$$E_2(u_*).(v, v) = \int_B [K_1 [(\operatorname{div} v_\infty)^2 - 2 v_\infty^2 R^{-2}]$$
$$+ K_2 (u_*.\operatorname{rot} v_\infty)^2 + K_3 (u_* \wedge \operatorname{rot} v_\infty)^2] dx$$

Il suffit donc de trouver v tel que $E_2(u_*).(v, v) < 0$ pour obtenir le résultat.Soit $\chi \in C^0([0, 1], R)$, C^1 par morceaux, telle que $\chi(1) = 0$. Alors $v(x) = \chi(R)(x_2, -x_1, 0)$ définit une application v de $H^1_0(B, R^3) \cap L^\infty(B, R^3)$. Des calculs simples montrent alors que

$$E_2(u_*).(v, v) = \int_B \{ K_1 [-2 (x_1^2 + x_2^2) \chi(R)^2 R^{-2}] + K_2 (4 x_3^2 \chi(R)^2 R^{-2})$$
$$+ K_3 (2\chi(R) R^{-1} + \chi'(R))^2 (x_1^2 + x_2^2) \} dx$$
$$= 4/3 \int_B [(K_2 - K_1) \chi(R)^2 + K_3/2 (2\chi(R) + \chi'(R) R)^2] dx$$
$$= 8\pi/3 \int_0^1 [(K_2 - K_1) \chi(R)^2 + K_3/2 (2\chi(R) + \chi'(R) R)^2] R^2 dR$$

Posons $f(R) = R^2 \chi(R)$, alors

$$E_2(u_*).(v, v) = 8\pi/3 \int_0^1 [(K_2 - K_1) f(R)^2 R^{-2} + K_3/2\, f'(R)^2] dR$$

Pour A > 0, on prend

$$f_A(R) = 0 \qquad \text{sur } [0, \exp(-A\pi)]$$
$$f_A(R) = \sqrt{R} \sin [\operatorname{Log}(R) / A] \qquad \text{sur } [\exp(-A\pi), 1]$$

Alors sur $]\exp(-A\pi), 1[$, $f_A''(R) = -(1/4 + 1/A^2)\, f_A(R)\, R^{-2}$, et

$$\int_{\exp(-A\pi)}^{1} f_A'(R)^2\, dR = -\int_{\exp(-A\pi)}^{1} f_A(R)\, f_A''(R)\, dR = (1/4 + 1/A^2)\int_{\exp(-A\pi)}^{1} f_A(R)^2\, R^{-2}\, dR$$

D'où

$$E_2(u_*).(v, v) = 8\pi/3 \int_{\exp(-A\pi)}^{1} [(K_2 - K_1) + (1/8 + 1/2A^2) K_3]\, f_A(R)^2\, R^{-2}\, dR$$

Et comme $8(K_2 - K_1) + K_3 < 0$, il suffit de prendre A assez grand pour que $E_2(u_*).(v, v) < 0$.

Remarque. –Comme pour tout u dans $A(u_*)$, on a

$$\begin{aligned} 2E(u) \geq K_1 \int_B (\operatorname{div} u)^2\, dx &= K_1 \int_B [9 + 6 \operatorname{div}(u - x) + (\operatorname{div}(u - x))^2]\, dx \\ &= K_1 [12\pi + \int_B [\operatorname{div}(u - x)]^2\, dx \\ &\geq 12\pi K_1 \end{aligned}$$

On a toujours $I(K_1, K_2, K_3) \geq 6\pi K_1$. De plus $E(u_*) = 8\pi K_1$, on en conlut qu'on a toujours

$$8\pi K_1 = E(u_*) \geq I(K_1, K_2, K_3) \geq 6\pi K_1$$

Réciproquement,utilisant la continuité de $(K_2, K_3) \longmapsto I(1, K_2, K_3)$, une conséquence du corollaire du théorème 2 qui suit est que pour chaque valeur d'énergie e de $]6\pi, 8\pi[$, il éxiste K_2, $K_3 > 0$ telles que $e = I(1, K_2, K_3)$.

Introduisons maintenant la fonction $u_0 \in L^\infty(B, S^2)$ définie par

$$u_0 = x + r^{-1}\sqrt{1 - R^2}\,(x_2, -x_1, 0)$$

Où $r = \sqrt{x_1^2 + x_2^2}$. Notons $\Sigma = \{x \in B / r \neq 0\}$, alors sur $B \setminus \Sigma$, u_0 est C^∞, $u_0(x) \in S^2$, et $\operatorname{div} u_0 = 3$. Sur ∂B, $u_0(x) = u_*(x)$. Donc $\int_B (\operatorname{div} u_0)^2\, dx = 12\pi$, et u_0 minimise E dans le cas dégénéré $K_2 = K_3 = 0$. Mais u_0 n'appartient à aucun $W^{1,p}(B, R^3)$ pour tout $p \in [1, +\infty[$. Toutefois, on a:

THEOREME 2. – *Il éxiste une famille u_ε de $A(u_*)$ pour $\varepsilon > 0$ qui tend vers u_0 pour la topologie $L^\infty_{loc}(B \setminus \Sigma, S^2)$ quand $\varepsilon \longrightarrow 0$, et qui vérifie*

$$\int_B (\operatorname{div} u_\varepsilon)^2\, dx \longrightarrow \int_B (\operatorname{div} u_0)^2\, dx = 12\pi$$

COROLLAIRE. – $\lim_{(K_2, K_3) \to 0} I(1, K_2, K_3) = 6\pi$

Démonstration du théorème 2. – On considère les fonctions ρ_ε, α_ε de $C^0([0, 1], [0, 1])$, C^1 par morceaux, définies par

$$\begin{array}{ll} \rho_\varepsilon(R) = R & \text{sur } [0, 1 - 2\varepsilon] \\ \rho_\varepsilon'(R) = 2 - \varepsilon & \text{sur } [1 - 2\varepsilon, 1 - \varepsilon] \\ \rho_\varepsilon(R) = 1 - (R - 1)^2 & \text{sur } [1 - \varepsilon, 1] \end{array}$$

$$\begin{array}{ll} \alpha_\varepsilon(r) = 1 & \text{sur } [0, \varepsilon] \\ \alpha_\varepsilon'(r) = -1/\varepsilon & \text{sur } [\varepsilon, 2\varepsilon] \\ \alpha_\varepsilon(r) = 0 & \text{sur } [2\varepsilon, 1] \end{array}$$

On pose alors $u_\varepsilon(x) = [\alpha_\varepsilon(r) + (1 - \alpha_\varepsilon(r))\rho_\varepsilon(R)]\, x/R + \sqrt{1 - [\alpha_\varepsilon(r) + (1 - \alpha_\varepsilon(r))\rho_\varepsilon(R)]^2}\, r^{-1}(x_2, -x_1, 0)$

Il est immédiat que u_ε tend vers u_0 dans $L^\infty_{loc}(B \setminus \Sigma, S^2)$.

Pour tout $\varepsilon > 0$, sur $B(0, \varepsilon)$, $u_\varepsilon(x) = x/R$, et sur $B \setminus B(0, \varepsilon)$, u_ε est continue et C^1 par morceaux avec une dérivée bornée dans $L^\infty(B \setminus B(0, \varepsilon), R^9)$. Donc $u_\varepsilon \in H^1(B, S^2)$. Il ne reste donc plus qu'à calculer $\operatorname{div} u_\varepsilon$.

$$\begin{aligned} \operatorname{div} u_\varepsilon &= \operatorname{grad}\{[\alpha_\varepsilon + (1 - \alpha_\varepsilon)\rho_\varepsilon] R^{-1}\}.x + [\alpha_\varepsilon + (1 - \alpha_\varepsilon)\rho_\varepsilon] R^{-1} \operatorname{div} x + 0 \\ &= \Big\{ -[\alpha_\varepsilon + (1 - \alpha_\varepsilon)\rho_\varepsilon] R^{-3} x + [\alpha_\varepsilon'(1 - \rho_\varepsilon)] R^{-1} r^{-1}(x_1, x_2, 0) \\ &\quad + [(1 - \alpha_\varepsilon)\rho_\varepsilon'] R^{-2} x \Big\}.x + 3[\alpha_\varepsilon + (1 - \alpha_\varepsilon)\rho_\varepsilon] R^{-1} \\ &= 2[\alpha_\varepsilon + (1 - \alpha_\varepsilon)\rho_\varepsilon] R^{-1} + [\alpha_\varepsilon'(1 - \rho_\varepsilon')] R^{-1} + (1 - \alpha_\varepsilon)\rho_\varepsilon' \end{aligned}$$

Sur $\Omega_1 = \{x \in B / r \in [2\varepsilon, 1], R \in [1, 1 - 2\varepsilon]\}$, $\operatorname{div} u_\varepsilon = 3$

Sur $\Omega_2 = \{x \in B / r \in [0, \varepsilon]\}$, $\operatorname{div} u_\varepsilon = 2R^{-1}$

Sur $\Omega_3 = \{x \in B / r \in [\varepsilon, 2\varepsilon]\}$, $|\operatorname{div} u_\varepsilon| \leq 2R^{-1} + r\varepsilon^{-1}R^{-1} + 2 \leq 2 + 4R^{-1}$

Sur $\Omega_4 = \{x \in B / r \in [2\varepsilon, 1], R \in [1 - 2\varepsilon, 1]\}$, $|\operatorname{div} u_\varepsilon| = |2\rho_\varepsilon R^{-1} + \rho_\varepsilon'| \leq 6$

Et $\left|\int_B (\operatorname{div} u_\varepsilon)^2\, dx - \int_B (\operatorname{div} u_0)^2\, dx\right| \leq \int_{\Omega_2 \cup \Omega_3 \cup \Omega_4} |9 - (\operatorname{div} u_\varepsilon)^2|\, dx$

Comme $| 9 - (\operatorname{div} u_\varepsilon)^2 |$ est intégrable, et comme la mesure de $\Omega_2 \cup \Omega_3 \cup \Omega_4$ tend vers 0, on conclut en utilisant le théorème de Lebesgue.

Démonstration du corollaire. — Pour tout $\beta \in]0, 1[$, en prenant ε suffisamment petit, on peut avoir

$$\int_B (\operatorname{div} u_\varepsilon)^2 dx \le (12 + \beta)\pi \quad \text{et} \quad \| u_\varepsilon \|_{H^1(B)} < +\infty$$

Donc en choisissant K_2, K_3 suffisament petit, on a pour $(1, K_2, K_3)$

$$E(u_\varepsilon) \le (6 + \beta)\pi \Rightarrow I(1, K_2, K_3) \le (6 + \beta)\pi$$

Remarque. — On constate que dans le cas dégénéré $K_2 = K_3 = 0$, il n'y a plus unicité du minimum puisque pour tout $\mathfrak{R} \in SO(3)$, $\mathfrak{R}(u_0)$ minimise également E. De plus, toute ces fonctions possèdent une symétrie cylindrique. On peut donc se demander si les toutes les fonctions minimisant E dans le cas $8(K_2 - K_1) + K_3 < 0$ ne sont pas des applications qui se déduisent les unes des autres par des rotations, et qui possèdent une symétrie cylindrique. On peut se demander également si u_* est minimisante dans le cas $8(K_2 - K_1) + K_3 \ge 0$, et $K_2 < K_1$.

Je remercie J.M. Coron qui a attiré mon attention sur cette question.

REFERENCES BIBLIOGRAPHIQUES

[1] H. BREZIS, J. M. CORON, E. LIEB, *Harmonic maps with defects*, Commun. Math. Phys.,107, 1986, p 649-705.

[2] R. HARDT, D. KINDERLEHRER, F. H. LIN, *Existence and partial regularity of static liquid crystal configurations*, Commun. Math. Phys.,105, 1986, p 541-570.

[3] M. KLEMAN, *Points, lignes, parois*, Les éditions de physique, Orsay, 1977.

[4] F. H. LIN, En préparation.

F. H. : *Centre de Mathématiques, Ecole Polytechnique, 91128 Palaiseau Cedex.*

Minimizing p-Harmonic Maps
into Spheres

by

Jean-Michel Coron and Robert Gulliver

Max-Planck-Institut
für Mathematik
Gottfried-Claren-Straße 26
D-5300 Bonn 3

Germany

Permanent addresses:

(J.-M.C.) Université de Paris-Sud
Département de Mathématiques
Bat. 425
F-91405 Orsay

France

(R.G.) School of Mathematics
University of Minnesota
Minneapolis, MN 55455

USA

MPI/87- 13

0. Introduction

Consider two Riemannian manifolds M^m and $N^n \subset \mathbb{R}^d$, where M is compact, possibly with boundary, and $m \geq 3$. A map $f : M \longrightarrow N$ is harmonic if it is stationary for Dirichlet's integral ("energy")

$$E_2(f) = \int_M |\nabla f|^2 d\mathrm{Vol},$$

where $|\nabla f|^2 = \sum_{i=1}^{d} \sum_{\alpha,\beta=1}^{m} \gamma^{\alpha\beta} \frac{\partial f_i}{\partial x_\alpha} \frac{\partial f_i}{\partial x_\beta}$, and where $\gamma_{\alpha\beta}(x) = (\gamma^{\alpha\beta}(x))^{-1}$ represents the metric of M. In a fundamental paper ([SU1]), Schoen and Uhlenbeck showed that near any singularity, a minimizing harmonic map $f : M^m \longrightarrow N^n$ converges strongly to a minimizing tangent map $u : \mathbb{R}^m \longrightarrow N^n$, which is harmonic and homogeneous of degree zero. The investigation of minimizing tangent maps $u : B^m \longrightarrow N^n$ is therefore an important aspect of current research into minimizing harmonic maps.

We restrict our attention in this paper to the case $N = S^n$, the unit sphere in $\mathbb{R}^{n+1}$. Even in this case, surprisingly few examples are known of maps $u : B^m \longrightarrow S^n$, homogeneous of degree zero, which minimize energy for given Dirichlet boundary conditions. The first nonconstant example was given by Jäger and Kaul in 1983, who proved that the map $u_0 : B^m \longrightarrow S^m$ defined by $u_0(x) = (x/|x|, 0)$ minimizes energy if $m \geq 7$ ([JK]; see also [SU3]). Recently, Brézis, Coron and Lieb have shown that the map $u_0(x) = x/|x|$ from B^3 to S^2 minimizes E_2 ([BCL]). A proof

was communicated to us by Lin that $u_0(x) = x/|x|$ from B^m to S^{m-1} has minimum energy, for all m ([L]). In a related result, Hélein has shown that $E_2(u) \geq E_2(u_0) + \alpha E_2(u-u_0)$ for some $\alpha > 0$, provided that $n = m-1$ and $m \geq 9$ ([H]).
In contrast, it is shown in [SU3] that any minimizing tangent map $u : B^m \longrightarrow S^n$ is constant if $m \leq d(n)$, where $d(3) := 3$ and $d(n) := 1 + \min\{n/2, 5\}$ otherwise.

A natural generalization of the functional E_2 is the p-energy

$$E_p(u) = \int_{B^m} |\nabla u|^p dx ,$$

which is finite if and only if u belongs to the Sobolev class $W^{1,p}(B^m, S^n) := \{u \in W^{1,p}(B^m, \mathbb{R}^{n+1}) : |u| = 1 \text{ a.e.}\}$. Mappings which are stationary for E_p are called p-<u>harmonic</u> maps. Note that regularity theorems analogous to results for $p = 2$ in [SU1] have not yet been proved for general p (uniform ellipticity is lost). One may well expect, however, that minimizing tangent maps will play a role similar to their role in the theory for $p = 2$.

One result of the present paper concerns the homogeneous mapping $u_0 : B^m \longrightarrow S^n$ defined by $u_0(y,z) = y/|y|$, where $y \in \mathbb{R}^{n+1}$ and $z \in \mathbb{R}^{m-n-1}$. We have

<u>Theorem 2.4</u>. <u>If</u> $p \leq n \leq m-1$, <u>then</u> $E_p(u_0) \leq E_p(u)$ <u>for any</u> $u \in W^{1,p}(B^m, S^n)$ <u>with</u> $u = u_0$ <u>on</u> ∂B^m.

If $p = n = m-1$, then this result may be proved by the methods of [BCL]. If $p = 2$ and $n = m-1$, then this is exactly Lin's result. Our proof was discovered later than Lin's and independently, and is of a quite different nature.

Two interesting examples of mappings from B^{2n} to S^n are provided by the homogeneous extension

$$u_0(x) = H(x/|x|)$$

of the Hopf maps $H : S^{2n-1} \longrightarrow S^n$ related to the multiplication of complex numbers $(n = 2)$ and the quaternions $(n = 4)$. We shall prove that both are minimizing maps for E_2 (Theorems 5.1 and 6.1).

Using similar techniques, we shall prove a sharp lower bound

$$E_n(u) \geq n^{n/2} \text{ Volume } (S^n)$$

for $u \in W^{1,n}(B^{n+1}, S^n)$ such that $u(-x) = -u(x)$ for all $x \in \partial B^{n+1}$ (Theorem 4.1).

Finally, we give a theorem with general hypotheses on a mapping $u_0 : B^m \longrightarrow S^n$ which allow us to conclude that u_0 minimizes E_p for its boundary data. The hypotheses are similar to the conditions for a harmonic morphism (compare p. 123 of [B]).

We would like to point out that the results of the present paper do not include a classification of all minimizing tangent maps into S^n. For example, up to an orthogonal motion, $u_0(x) = x/|x|$ is the only known example of a minimizing tangent map from B^4 to S^3; it is not known whether any others exist. It was proved in [BCL] that $u_0(x) = x/|x|$ is the unique minimizing tangent map from B^3 to S^2 modulo $\mathbb{O}(3)$.

One idea in our proof is to bound the p-energy of a map $v : B^m \longrightarrow S^p$ from below by a coarea formula. The usefulness of the coarea formula in the context of the functional E_p for mappings into a p-dimensional manifold was made clear in the paper of Almgren, Browder and Lieb [ABL]. An analogous framework of ideas had been constructed in [BCL] for the case $p = n = m-1$.

A new idea, which plays a central role in our proof, is to estimate the p-energy of a map $u : B^m \longrightarrow S^n$ by averaging a related functional of the composition of u with all nearest-point projections π_Y of S^n onto its totally geodesic p-spheres (Lemma 2.2). This averaging method is simplest in the classical case $p = 2$: the energy of any map $u : B^m \longrightarrow S^n$ is a constant times the average of $E_2(\pi_Y \circ u)$ over all 3-planes Y in $\mathbb{R}^{n+1}$. Here $\pi_Y : S^n \longrightarrow Y \cap S^n$ maps $s \in S^n$ to the nearest point in the 2-sphere $Y \cap S^n$ (Lemma 1.2).

An important technical tool in our proof is a new approximation result for mappings into the p-sphere of class $W^{1,p}$ (Theorem 3.2), which is based on methods of Hardt-Lin and of Bethuel-Zheng. Note that smooth mappings are not dense ([SU2], p. 267 for $p = 2$). However, we construct a dense class $\mathcal{R}$ of mappings whose singularities form submanifolds of codimension $p + 1$, with

simple structure near the singularities. Of course, the slicing theorems of Federer ([F], 4.3.1), which are relevant to the coarea formula, are valid only for Lipschitz-continuous mappings; in effect, the singular set of a mapping of class $\mathcal{R}$ contributes to the boundary of each slice. This difficulty is overcome by considering the difference of the slices at two distinct points in S^p; the difference is a current having no boundary in the interior of the domain.

The authors would like to thank Martin Guest and Uwe Abresch for useful references to the literature. They gratefully acknowledge the hospitality of the Max-Planck-Institut für Mathematik, where this work was carried out.

1. Projection to lower-dimensional spheres $(p = 2)$

Consider $n \le m-1$ and an integer p, $1 \le p \le n$. For this section and the following one, we define boundary data $g : \partial B^m \longrightarrow S^n$ by

$$g(y,z) = y/|y| ,$$

where $y \in \mathbb{R}^{n+1}$ and $z \in \mathbb{R}^{m-n-1}$. The class of admissible mappings is

$$\mathcal{E}_p(g) = \{u \in W^{1,p}(B^m,S^n) : u = g \text{ on } \partial B^m\} .$$

The homogeneous extension of g is $u_0(y,z) = y/|y|$, which is singular on $\{0\} \times \mathbb{R}^{m-n-1} \subset \mathbb{R}^m$. Note that $E_p(u_0)$ is the integral of $|y|^{-p}$, which is finite since $p < n+1$. This shows that $u_0 \in \mathcal{E}_p(g)$, and the admissible class is not empty.

In this section, we shall consider only the case $p = 2$, which is simpler than the general case (compare the averaging Lemmas 1.2 and 2.2). Our result is

Theorem 1.1. *$E_2(u_0) \le E_2(u)$ for any $u \in \mathcal{E}_2(g)$.*

Given a 3-plane $Y \subset \mathbb{R}^{n+1}$, we define $\pi_Y : S^n \longrightarrow S^n \cap Y$ by $\pi_Y(u) = u'/|u'|$, where u' is the orthogonal projection of u onto Y. The singular set of π_Y is the $(n-3)$-sphere $S^n \cap Y^\perp$.

Lemma 1.2. *There is a constant $c = c(n)$ such that for any $u \in W^{1,2}(B^m,S^n)$,*

$$(1.1) \quad c\, E_2(u) = \int_{Y \in G_3(\mathbb{R}^{n+1})} E_2(\pi_Y \circ u)\, dG(Y) .$$

Here dG is the bi-invariant volume form on the Grassmann manifold $G_3(\mathbb{R}^{n+1})$.

Proof. For any tangent vector V to S^n, we have

$$(1.2) \quad c|V|^2 = \int_{Y \in G_3(\mathbb{R}^{n+1})} |D\pi_Y(V)|^2 dG(Y) ,$$

since $\mathbb{O}(n+1)$ acts transitively on the unit tangent vectors to S^n and leaves $dG(Y)$ invariant on $G_3(\mathbb{R}^{n+1})$. Note that π_Y is singular along a totally geodesic $(n-3)$-sphere of S^n, and $|D\pi_Y(V)| \le C|V|/r$, where r is the distance to the singular set; therefore, the integral in equation (1.2) is finite. Since $|\nabla u|^2 = \sum_{\alpha=1}^{m} \left|\frac{\partial u}{\partial x_\alpha}\right|^2$, this formula applied to $V = \frac{\partial u}{\partial x_\alpha}$ yields

$$c|\nabla u|^2 = \int_{Y \in G_3(\mathbb{R}^{n+1})} |\nabla(\pi_Y \circ u)|^2 \, dG(Y) .$$

We integrate both sides over B^m to obtain (1.1) by Fubini's theorem.

q.e.d.

Corollary 1.3. *Let $v_0 : B^m \longrightarrow S^2$ be defined by $v_0(x,y) = \frac{x}{|x|}$, where $x \in \mathbb{R}^3$, $y \in \mathbb{R}^{m-3}$. If $E_2(v) \ge E_2(v_0)$ for every*

$v \in W^{1,2}(B^m,S^2)$ with $v = v_0$ on ∂B, then $E_2(u) \geq E_2(u_0)$ for every $u \in W^{1,2}(B^m,S^n)$ with $u = u_0$ on ∂B.

Proof. Note that $\pi_Y \circ u_0 = v_0$ after performing an appropriate rotation in $\mathbb{R}^m$. Using Lemma 1.2,

$$c\,E_2(u) = \int_{G_3(\mathbb{R}^{n+1})} E_2(\pi_Y \circ u)\,dG(Y) \geq$$

$$\int_{G_3(\mathbb{R}^{n+1})} E_2(\pi_Y \circ u_0)\,dG(Y) = c\,E_2(u_0)\,.$$

q.e.d.

The coarea formula has the serious weakness that it gives a lower bound for energy E_2 only for mappings to a manifold of dimension $n = 2$. The above corollary bypasses this weakness in the case of mappings to the n-sphere.

Lemma 1.4. (Coarea formula, $p = 2$). If $v \in C^{0,1}(\Omega,S^2)$ for Ω open in B^m, then

$$\int_\Omega |\nabla v|^2 dx \geq 2 \int_{S^2} \mathcal{H}^{m-2}(v^{-1}(s))\,dA_{S^2}(s)$$

where $\mathcal{H}^{m-2}$ denotes $(m-2)$-dimensional Hausdorff measure.

Proof. See [F, 3.2.22], with the observation that $|\nabla v|^2 \geq 2\,J(v)$, where $J(v)$ is the determinant of ∇v restricted to the 2-dimensional space orthogonal to $v^{-1}(s)$, and s is any regular value of v.

q.e.d.

In order to use Lemma 1.4, which is only valid for Lipschitz mappings, we need to approximate $W^{1,2}(B^m,S^2)$ by mappings having precisely controlled singularities (recall that Lipschitz functions are not dense for $m \geq 3$: see [SU2], p. 267). Let $\mathcal{R}$ be the class of mappings $v \in W^{1,2}(B^m,S^2)$ such that

(1.3) $v = v_0$ on a neighborhood of ∂B^m (whose size may depend on v) and on a neighborhood of the singular set $\Delta = \{0\} \times \mathbb{R}^{m-3}$ of v_0;

(1.4) $v \in C^\infty(B^m \setminus (\Delta \cup \Sigma))$ for some Lipschitz $(m-3)$-dimensional manifold $\Sigma \subset\subset B^m \setminus \Delta$ ($\partial\Sigma = \phi$); and

(1.5) for a.e. $s \in S^2$, $v^{-1}(s) \cup v^{-1}(-s) \cup \Sigma \cup \Delta$ is a Lipschitz $(m-2)$-dimensional manifold with boundary $\subset \partial B^m$.

Approximation Theorem 1.5. If $v_0 \in \mathcal{R}$, then $\mathcal{R}$ is dense in $E_2(v_0) = \{v \in W^{1,2}(B^m,S^2) : v = v_0 \text{ on } \partial B^m\}$.

We defer the proof of Theorem 1.5 to section 3.

Proof of Theorem 1.1. According to Corollary 1.3 and the Approximation Theorem 1.5, we need only to show that for $v \in \mathcal{R}$, $E_2(v) \geq E_2(v_0)$. We use the coarea formula of Lemma 1.4, with $\Omega = B^m \setminus (\Sigma \cup \Delta)$:

(1.6) $\int_{\Omega} |\nabla v|^2 dx \geq 2 \int_{S^2} H^{m-2}(v^{-1}(s) \cap \Omega)\, dA_{S^2}(s)$.

Since $H^{m-2}(\Sigma \cup \Delta) = 0$, we have $H^{m-2}(v^{-1}(s) \cap \Omega) = H^{m-2}(v^{-1}(s))$. Note that the antipodal map from S^2 to S^2 defined by $s \longmapsto -s$ preserves the volume form $dA_{S^2}(s)$, so that the right-hand side of (1.6) equals

$$\int_{S^2} H^{m-2}(v^{-1}(s) \cup v^{-1}(-s))\, dA_{S^2}(s) .$$

It follows from conditions (1.5) and (1.3) that for almost all s, $v^{-1}(s) \cup v^{-1}(-s) \cup \Sigma \cup \Delta$ is a regular manifold with boundary a totally geodesic sphere of dimension $m-3$. In particular, it has $(m-2)$-dimensional measure $\geq H^{m-2}(B^{m-2}) =: \alpha_{m-2}$. Thus $\int_{B^m} |\nabla v|^2 dx \geq \alpha_{m-2}\, H^2(S^2) = 4\pi\, \alpha_{m-2}$.

Meanwhile, $|\nabla v_0(x,y)|^2 = \frac{2}{|x|^2}$, so that

$$E_2(v_0) = \int_{B^m} \frac{2}{|x|^2}\, dx\, dy = 4\pi\, \alpha_{m-3} \int_0^1 2(1-r^2)^{\frac{m-3}{2}}\, dr$$

$$= 4\pi\, \alpha_{m-2} .$$

q.e.d.

2. Projection to lower-dimensional spheres (general p)

We may now turn our attention to the case of a general integer exponent $1 \leq p \leq n$. Somewhat surprisingly, the counterpart of Lemma 1.2 fails: if the constant c is defined so that

$$c\, E_p(u_0) = \int_{Y \in G_{p+1}(\mathbb{R}^{n+1})} E_p(\pi_Y \circ u_0)\, dG(Y) ,$$

then it is not true that

$$c\, E_p(u) \geq \int_{Y \in G_{p+1}(\mathbb{R}^{n+1})} E_p(\pi_Y \circ u)\, dG(Y) .$$

In order to carry out our program, we will instead compute the average of squares of the Jacobian determinants of $\pi_Y \circ u$ (see Lemma 2.2 below).

Given a linear transformation $L : \mathbb{R}^m \longrightarrow \mathbb{R}^n$, where $m \geq n$, we may write $L = Q_1 \Lambda Q_2$, where $Q_1 \in \mathbb{O}(n)$, $Q_2 \in \mathbb{O}(m)$ and Λ has the form

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n & 0 & \cdots & 0 \end{pmatrix} ,$$

with $\lambda_j \geq \lambda_{j+1} \geq 0$. (For example, $L = \nabla u(x)$.) In fact, $\lambda_1^2, \ldots, \lambda_n^2$ are the eigenvalues of the positive semi-definite symmetric operator

LL^T on $\mathbb{R}^n$. We shall refer to $\lambda_1,\dots,\lambda_n$ as the <u>singular values</u> of L. Observe that λ_k may be given a variational interpretation:

(2.1) $\lambda_k = \max_{Z\in G_k(\mathbb{R}^m)} \min\{|Lx| : x \in Z,\ |x| = 1\}$.

For any p-plane $Z \in G_p(\mathbb{R}^n)$, let π_Z be the orthogonal projection of $\mathbb{R}^n$ onto Z, and write $\mu_1(Z),\dots,\mu_p(Z)$ for the singular values of $\pi_Z \circ L : \mathbb{R}^m \longrightarrow Z \cong \mathbb{R}^p$.

Recall the definition of the elementary symmetric functions $\sigma_1,\dots,\sigma_n$ of $(\alpha_1,\dots,\alpha_n) \in \mathbb{R}^n$:

$$\sigma_k(\alpha_1,\dots,\alpha_n) := \sum_{i_1<i_2<\dots<i_k} \alpha_{i_1}\alpha_{i_2}\cdots\alpha_{i_k}.$$

We have the following formula:

<u>Lemma 2.1</u>. <u>Given any linear transformation</u> $L : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ <u>with singular values</u> $\lambda_1,\dots,\lambda_n$, <u>the average</u>

$$⨍_{G_p(\mathbb{R}^n)} [\mu_1(Z)\dots\mu_p(Z)]^2 dG(Z) = \binom{n}{p}^{-1} \sigma_p(\lambda_1^2,\dots,\lambda_n^2).$$

<u>Proof</u>. Without loss of generality, we may assume $L = \Lambda$. Write $M = \Lambda\Lambda^T$, a symmetric linear operator on $\mathbb{R}^n$ with eigenvalues $\alpha_1 = \lambda_1^2,\dots,\alpha_n = \lambda_n^2$. For any $Z \in G_p(\mathbb{R}^n)$, $\mu_1(Z)^2\dots\mu_p(Z)^2 = \det(\pi_Z M \pi_Z^T)$, which is a homogeneous polynomial of degree p in $\alpha_1,\dots,\alpha_n$. Define

$$f(\alpha_1,\dots,\alpha_n) = ⨍_{G_p(\mathbb{R}^n)} \det(\pi_Z M \pi_Z^T)\, dG(Z);$$

we need only show that

$$f(\alpha_1,\dots,\alpha_n) = \binom{n}{p}^{-1} \sigma_p(\alpha_1,\dots,\alpha_n) =: \binom{n}{p}^{-1} \sigma_p$$

for any $\alpha_1,\dots,\alpha_n \in \mathbb{R}$. Note that f is a homogeneous polynomial of degree p. Further, $f(\alpha_1,\dots,\alpha_n)$ is symmetric under permutations of $(\alpha_1,\dots,\alpha_n)$, since a permutation corresponds to the isometry of $\mathbb{R}^n$ which permutes the eigenvectors, leaving $dG(Z)$ unchanged.

According to the fundamental theorem on symmetric functions (see <u>e.g.</u> [Md], p. 13), any symmetric polynomial $f(\alpha_1,\dots,\alpha_n)$ is equal to a polynomial $P_0(\sigma_1,\dots,\sigma_n)$, with real coefficients; moreover, $\sigma_1,\dots,\sigma_n$ are algebraically independent. In our case $f(\alpha_1,\dots,\alpha_n)$ is homogeneous of degree $p \le n$, and therefore, for some $\gamma \in \mathbb{R}$,

$$P_0(\sigma_1,\dots,\sigma_n) = \gamma\sigma_p + P_1(\sigma_1,\dots,\sigma_{p-1}).$$

Consider the special case $\alpha_p = \dots = \alpha_n = 0$: in this case $\sigma_i(\alpha_1,\dots,\alpha_{p-1}, 0,\dots,0) = \tilde\sigma_i(\alpha_1,\dots,\alpha_{p-1})$, the elementary symmetric function in $p-1$ variables. On the other hand, for each $Z \in G_p(\mathbb{R}^n)$, $\det(\pi_Z M \pi_Z^T) = 0$ since M has rank $\le p-1$, and hence $f(\alpha_1,\dots,\alpha_n) = P_0(\sigma_1,\dots,\sigma_n) = 0$. Clearly, $\sigma_p = 0$ as well. Therefore, for any $(\alpha_1,\dots,\alpha_{p-1}) \in \mathbb{R}^{p-1}$, $P_1(\tilde\sigma_1,\dots,\tilde\sigma_{p-1}) = 0$. But according to the fundamental theorem on symmetric functions, $\tilde\sigma_1,\dots,\tilde\sigma_{p-1}$ are algebraically independent in $p-1$ variables, so that the polynomial P_1 is itself zero. This shows that

$$f(\alpha_1,\ldots,\alpha_n) = \gamma\sigma_p(\alpha_1,\ldots,\alpha_n) .$$

Finally, we may evaluate the constant γ by choosing $M = \mathrm{id}$, which implies $\mu_1(Z) = \ldots = \mu_p(Z) = 1$ for each Z in $G_p(\mathbb{R}^n)$. Since $\sigma_p(1,\ldots,1) = \binom{n}{p}$ we have $\gamma = \binom{n}{p}^{-1}$ as claimed.

q.e.d.

Define, for $v : B^m \longrightarrow S^p$ and for each x in B^m, $J(v)(x) := \lambda_1(x)\ldots\lambda_p(x)$, the product of the p singular values of $L = \nabla v(x) : \mathbb{R}^m \longrightarrow T_{v(x)}S^p$. We recall that $u : B^m \longrightarrow S^n$ is said to be <u>horizontally conformal</u> (see <u>e.g.</u> [B]) if for almost all x in B^m, the singular values of $\nabla u(x) : \mathbb{R}^m \longrightarrow T_{u(x)}S^n$ are equal.

We have the following averaging result:

<u>Lemma 2.2</u>. <u>For</u> $n \geq p$, <u>there is a constant</u> $c = c(n,p)$ <u>such that for any</u> $u \in W^{1,p}(B^m,S^n)$

$$(2.2) \quad c\, E_p(u) \geq \fint_{G_{p+1}(\mathbb{R}^{n+1})} \int_{B^m} J(\pi_Y \circ u)\,dx\, dG(Y) .$$

<u>Moreover, equality holds if</u> u <u>is horizontally conformal</u>.

<u>Proof</u>. We first observe that

$$(2.3) \quad \sigma_p(\alpha_1,\ldots,\alpha_n) \leq \binom{n}{p} \left(\frac{1}{n}\sum_{i=1}^{n}\alpha_i\right)^p$$

for $\alpha_i \geq 0$, with equality if $\alpha_1 = \ldots = \alpha_n$. The case $p = n-1$ was given in inequality (8.5) of [BCL] and A.1.3 of [ABL]. We prove inequality (2.3) by induction on n; for $n = p$ it is the well-known arithmetic-geometric inequality. By reordering, we may assume $\alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_n$; by homogeneity, we may assume $\alpha_1 + \ldots + \alpha_n = 1$. Now consider $(\alpha_1,\ldots,\alpha_n)$ which maximizes σ_p. If $\alpha_n = 0$, we use the induction hypothesis: $\sigma_p(\alpha_1,\ldots,\alpha_{n-1},0) \leq \binom{n-1}{p}(n-1)^{-p} < \binom{n}{p} n^{-p}$. If $\alpha_n > 0$, then by the method of Lagrange, there is $\beta \in \mathbb{R}$ with

$$\beta = \frac{\partial}{\partial\alpha_i}\sigma_p(\alpha_1,\ldots,\alpha_n) = \sigma_{p-1}(\alpha_1,\ldots,\alpha_{i-1},\alpha_{i+1},\ldots,\alpha_n) =: \theta_i$$

for all $1 \leq i \leq n$. But $\theta_{i+1} \geq \theta_i$, and equality implies that $\alpha_i = \alpha_{i+1}$; inequality (2.3) follows.

Applying inequality (2.3) to the eigenvalues $\alpha_1,\ldots,\alpha_n$ of $\nabla u(x)(\nabla u(x))^T$ for some $x \in B^m$, we have

$$(2.4) \quad |\nabla u(x)|^{2p} \geq n^p \binom{n}{p}^{-1} \sigma_p(\alpha_1,\ldots,\alpha_n) =$$

$$= n^p \fint_{G_p(T_{u(x)}S^n)} \det(\pi_Z \nabla u(x)\, \nabla u(x)^T \pi_Z^T)\, dG(Z)$$

by Lemma 2.1. An application of the Cauchy-Schwartz inequality yields

$$(2.5) \quad |\nabla u(x)|^p \geq n^{p/2} \fint_{G_p(T_{u(x)}S^n)} [\det(\pi_Z \nabla u(x) \nabla u(x)^T \pi_Z^T)]^{1/2} dG(Z) .$$

Note that equality holds in the inequalities (2.4) and (2.5) if u is horizontally conformal at x.

Let Y be a (p+1)-plane in $\mathbb{R}^{n+1}$, and write Z for the p-plane in $T_{u(x)}S^n$ parallel to the subspace of Y orthogonal to $\pi_Y(u(x))$. Then up to a parallel translation in $\mathbb{R}^{n+1}$, we have

$$\nabla(\pi_Y \circ u)(x) = \frac{\pi_Z \circ \nabla u(x)}{\cos d(u(x),Y)}$$

where $d(u_1,Y)$ is the distance in S^n from $u_1 \in S^n$ to $Y \cap S^n$. Thus

$$(2.6) \quad J(\pi_Y \circ u)(x) = \frac{[\det(\pi_Z \nabla u(x) \nabla u(x)^T \pi_Z^T)]^{1/2}}{\cos^p d(u(x),Y)} .$$

Integrating formula (2.6) over $Y \in G_{p+1}(\mathbb{R}^{n+1})$, we find that

$$(2.7) \quad ⨍_{G_{p+1}(\mathbb{R}^{n+1})} J(\pi_Y \circ u)(x)\,dG(Y) =$$

$$= c' ⨍_{G_p(T_{u(x)}S^n)} [\det(\pi_Z \nabla u(x) \nabla u(x)^T \pi_Z^T)]^{1/2} dG(Z) ,$$

where $c' = c'(n,p)$ is independent of x and u. Finally, using inequality (2.5) and equation (2.7), and integrating over $x \in B^m$, we find the inequality (2.2) with $c = n^{-p/2} c'$.

q.e.d.

Lemma 2.3. (Coarea formula, general p). *If* $v \in C^{0,1}(\Omega,S^p)$ *for an open set* $\Omega \subset B^m$, *then*

$$\int_\Omega J(v)\,dx = \int_{S^p} H^{m-p}(v^{-1}(s))\,dA_{S^p}(s) .$$

Proof. See [F, 3.2.22].

Let $g : S^{m-1} \longrightarrow S^n$ and $u_0 : B^m \longrightarrow S^n$ be as in Theorem 1.1.

Theorem 2.4. *For any* $1 \le p \le n$, $E_p(u_0) \le E_p(u)$ *for all* $u \in \mathcal{E}_p(g)$.

Proof. According to Lemma 2.2, we have

$$(2.8) \quad c\, E_p(u) \ge ⨍_{G_{p+1}(\mathbb{R}^{n+1})} \int_{B^m} J(\pi_Y \circ u)\,dx\, dG(Y) .$$

Note that for almost all $Y \in G_{p+1}(\mathbb{R}^{n+1})$, the map $v = \pi_Y \circ u \in W^{1,p}(B^m,S^p)$. Write $v_0 = \pi_Y \circ u_0$. We shall show that for all such $Y \in G_{p+1}(\mathbb{R}^{n+1})$,

$$(2.9) \quad \int_{B^m} J(v)\,dx \ge \int_{B^m} J(v_0)\,dx .$$

According to Approximation Theorem 3.2, it is enough to prove inequality (2.9) for v in the class $\mathcal{R}$ of mappings with controlled singularities, since $J(v)$ is dominated by $|\nabla v|^p$. As before, choose $\Omega = B^m \setminus (\Sigma \cup \Delta)$, where $\Sigma \cup \Delta$ is the singular set of v, as in the definition of $\mathcal{R}$; and apply the coarea formula of Lemma 2.3. This yields

$$(2.10)\quad \int_\Omega J(v)dx = \int_{S^p} \mathcal{H}^{m-p}(v^{-1}(s))dA_{S^p} =$$

$$= \frac{1}{2}\int_{S^p} \mathcal{H}^{m-p}(M(s))dA_{S^p} ,$$

where $M(s) := v^{-1}(s) \cup v^{-1}(-s) \cup \Sigma \cup \Delta$ is a regular, oriented Lipschitz manifold having boundary a totally geodesic sphere of dimension $m-p-1$ in ∂B^m, for a.a. $s \in S^p$. In particular,

$$(2.11)\quad \mathcal{H}^{m-p}(M(s)) \geq \mathcal{H}^{m-p}(B^{m-p}) .$$

Note that equality holds in (2.11) for

$$M_0(s) = v_0^{-1}(s) \cup v_0^{-1}(-s) \cup \Delta .$$

Inequality (2.9) now follows from equation (2.10) for v and for v_0.

With u replaced by u_0, we attain equality in (2.8), according to Lemma 2.2, since u_0 is horizontally conformal. Therefore

$$c\, E_p(u) \geq c\, E_p(u_0) .$$

q.e.d.

3. Approximation and slicing

We saw in sections 1 and 2 that it may be useful to approximate mappings in $W^{1,p}(B^m,S^p)$ by mappings with controlled singularities, and particularly, with regular slices $u^{-1}(s)$ for almost all $s \in S^p$. Consider boundary data $g \in W^{1,p}(\partial B^m, S^p)$, such that g is C^∞ except on a Lipschitz submanifold $\Gamma \subset \partial B^m$ of dimension at most $m-p-2$. Write $u_0 \in W^{1,p}(B^m, S^p)$ for the homogeneous mapping $u_0(x) := g(\frac{x}{|x|})$. Observe that u_0 is singular on the cone $\Delta := \{tx : 0 \leq t \leq 1,\ x \in \Gamma\}$. We define $\mathcal{R}$ to be the class of mappings $u \in W^{1,p}(B^m, S^p)$ such that

(3.1) $u = u_0$ on a neighborhood of $\partial B^m \cup \Delta$;

(3.2) u is locally Lipschitz on $B^m \setminus (\Delta \cup \Sigma)$, for some Lipschitz submanifold $\Sigma \subset\subset B^m \setminus \Delta$ $(\partial\Sigma = \phi)$, of dimension $m-p-1$; and

(3.3) for a.a. $s \in S^p$, $u^{-1}(s) \cup u^{-1}(-s) \cup \Sigma \cup \Delta$ is a regular, oriented $(m-p)$-dimensional Lipschitz submanifold of B^m, having boundary only in ∂B^m.

Remark 3.1. The conditions (3.3) and (3.1) are both possible only if the restriction of g to a small p-sphere linking Γ in ∂B^m is one-to-one. In the present paper, this condition is always satisfied; the general case requires methods of geometric measure theory.

__Theorem 3.2__. _If_ $u_0 \in R$, _then_ R _is dense in_

$$\mathcal{E}_p(g) = \{u \in W^{1,p}(B^m, S^p) : u = u_0 \ \text{on} \ \partial B^m\} .$$

__Proof__. We follow ideas of Bethuel and Zheng [BZ]. Consider $u \in \mathcal{E}_p(g)$: we wish to find $u_k \in R$, $u_k \longrightarrow u$ in $W^{1,p}(B^m, S^p)$. First observe that by radially homogeneous extension beyond ∂B and rescaling we may assume that $u = u_0$ on $\{x \in \mathbb{R}^m : |x| \geq 1-\varepsilon\}$. We form $w_k = u*\rho_k$ for some compactly supported mollifier $\rho : \mathbb{R}^m \longrightarrow [0,\infty)$, where $\rho_k(x) := k^m\rho(kx)$. Note that $w_k \longrightarrow u$ in $W^{1,p}(B^m, \bar{B}^{p+1})$. Since Δ has p-capacity zero, we may find a sequence $v_k' \in W^{1,p}(B^m, \bar{B}^{p+1})$ such that each $v_k' = u_0$ on a neighborhood (of size depending on k) of Δ , $v_k' \in C^\infty(B^m \setminus \Delta, \mathbb{R}^{p+1})$ and $v_k' \longrightarrow u$ in $W^{1,p}(B^m, \bar{B}^{p+1})$ and a.e.. Let $\eta \in C^\infty(\mathbb{R}^m, \mathbb{R})$ have support in $B^m_{1-\varepsilon/2}$, such that $\eta(x) = 1$ for $|x| \leq 1-\varepsilon$. Define

$$v_k = \eta \ v_k' + (1-\eta)u_0 ;$$

then $v_k = u_0$ on a neighborhood of $\Delta \cup \partial B^m$, and $v_k \longrightarrow u$ in $W^{1,p}(B^m, \bar{B}^{p+1})$ and a.e.. Let

$$\Omega_k := \{x \in B^m : |v_k(x)| < \tfrac{1}{2}\} ;$$

then $\text{mes}(\Omega_k) \longrightarrow 0$, and hence

$$\int_{\Omega_k} |\nabla v_k|^p \longrightarrow 0 \quad \text{as} \quad k \longrightarrow \infty .$$

Consider a regular value ("center") $a \in B^{p+1}_{1/4}$; let $\Sigma(a,k) = \{x \in B^m : v_k(x) = a\}$. Note that since $|v_k| = |u_0| = 1$ on a neighborhood of $\partial B^m \cup \Delta$, $\Sigma(a,k)$ lies in a compact subset of $B^m \setminus \Delta$. Since $v_k \in C^\infty(B^m \setminus \Delta, \bar{B}^{p+1})$, $\Sigma(a,k)$ is a regular submanifold of dimension $m-p-1$. Define $q_a : \bar{B}^{p+1}_{1/2} \longrightarrow \partial B^{p+1}_{1/2}$ so that x lies in the line segment from a to $q_a(x)$; then $q_a(x) = x$ for $x \in \partial B^{p+1}_{1/2}$. Extend q_a to $\bar{B}^{p+1}$ by defining $q_a(x) = x$ when $|x| \geq 1/2$. We note that

$$|\nabla q_a(x)| \leq C/|x-a| ,$$

where C is independent of a and x . As in the paper of Hardt and Lin ([HL]; see also [HKL], p. 556), we apply Fubini's theorem to show that for every $\varphi \in W^{1,p}(\Omega_k, \bar{B}^{p+1})$,

$$⨍_{B^m_{1/4}} \int_{\Omega_k} |\nabla(q_a \circ \varphi)|^p \, dx \, da \leq C' \int_{\Omega_k} |\nabla\varphi|^p \, dx$$

where C' is independent of φ, a and k . It follows that for a in a subset of $B^m_{1/4}$ of positive measure

$$\text{(3.4)} \quad \int_{\Omega_k} |\nabla(q_a \circ v_k)|^p \, dx \leq C' \int_{\Omega_k} |\nabla v_k|^p \, dx ,$$

which tends to zero.

We define $u_k(x) := \dfrac{q_a(v_k(x))}{|q_a(v_k(x))|}$. From inequality (3.4), we see that $u_k \longrightarrow u$ in $W^{1,p}(B^m, S^p)$. Note that u_k satisfies condition (3.1), and satisfies condition (3.2) for almost all $a \in B^m_{1/4}$.

In order to check condition (3.3), we first observe that for almost all lines $a + \mathbb{R}b$ in $\mathbb{R}^{p+1}$, $v_k^{-1}(a+\mathbb{R}b) \setminus \Delta$ is locally a smooth submanifold of $B^m \setminus \Delta$, as follows from Sard's theorem. Note that for $s \in S^p$, $M(s,a) := u_k^{-1}(s) \cup u_k^{-1}(-s) \cup \Sigma = v_k^{-1}(q_a^{-1}(\mathbb{R}s))$, while $q_a^{-1}(\mathbb{R}s)$ is the union of the four line segments

$$[s, \tfrac{1}{2} s] \cup [\tfrac{1}{2} s, a] \cup [a, -\tfrac{1}{2} s] \cup [-\tfrac{1}{2} s, -s] \ .$$

Observe that $q_a^{-1}(\mathbb{R}s)$ is a Lipschitz 1-manifold with boundary $\{s, -s\}$. In particular, for almost all $a \in B^m_{1/4}$ and $s \in S^p$, $M(s,a)$ is locally a Lipschitz (m-p)-dimensional manifold in $B^m \setminus \Delta$. In a neighborhood of $\partial B^m \cup \Delta$, we have $u_k = u_0$; but since $u_0 \in R$ by hypothesis, $M(s,a) \cup \Delta$ is also a Lipschitz manifold near $\partial B \cup \Delta$, hence everywhere in $\bar{B}^m$. Finally, $M(s,a) \cup \Delta$ is composed of four smooth manifolds-with-boundary, of which two meet at the smooth manifolds $\Sigma = v_k^{-1}(a)$, at $v_k^{-1}(\frac{1}{2} s)$ and at $v_k^{-1}(-\frac{1}{2} s)$; so that its boundary is $(\Delta \cup u_0^{-1}(s) \cup u_0^{-1}(-s)) \cap \partial B^m$.

q.e.d.

4. Odd boundary data

In this section, we consider smooth boundary data $g : S^{m-1} \longrightarrow S^{m-1}$ satisfying the hypothesis

(4.1) $g(-x) = -g(x)$

for all $x \in S^{m-1}$, with $p = m-1$. This includes the specific case of Theorem 1.1 with $n = p = m-1$. Let $u_0 : B^m \longrightarrow S^{m-1}$ be defined by $u_0(x) = \frac{x}{|x|}$. We have the following lower bound for any such g :

Theorem 4.1. *For any* $u \in W^{1,m-1}(B^m, S^{m-1})$ *with* $u = g$ *on* ∂B , $E_{m-1}(u) \geq E_{m-1}(u_0)$.

Remark 4.2. This result settles a conjecture of Brézis-Coron-Lieb [BCL, Remark 7.3]; they proved this theorem under the additional hypothesis that the Jacobian $J(g) \geq 0$ and g has degree 1.

Proof. According to Approximation Theorem 3.2 (see also Theorem 4 of [BZ]), we may assume u belongs to the class R of $W^{1,m-1}$ mappings with controlled singularities. In particular, u is locally Lipschitz continuous on $B^m \setminus \Sigma$, where Σ is a finite set. Further, for almost every $s \in S^{m-1}$, $M(s) := u^{-1}(s) \cup u^{-1}(-s) \cup \Sigma$ is a Lipschitz 1-manifold with boundary $g^{-1}(s) \cup g^{-1}(-s)$. Considered as a one-dimensional integral current, $\partial M = \sum_{i=1}^{k} a_i - \sum_{i=1}^{k} a_{i+k}$, where $\{a_1,\ldots,a_{2k}\} = g^{-1}(s) \cup g^{-1}(-s)$ and a point of $g^{-1}(\pm s)$ is included in the list $\{a_1,\ldots,a_k\}$ provided $\pm J(g)(a_i) > 0$, otherwise in the list $\{a_{k+1},\ldots,a_{2k}\}$. Note that hypothesis (4.1) implies that $J(g)(-a_i) = J(g)(a_i)$. By reordering $\{a_{k+1},\ldots,a_{2k}\}$, we may assume that $a_{i+k} = -a_i$. According to the well-known theorem of Borsuk and Ulam, g has odd degree. Since $s \in S^{m-1}$ is a regular value of g, the number of points in $g^{-1}(s)$ has the same parity as the degree of g. That is, k is odd.

Now each connected component of $M(s)$ has boundary equal to the zero-dimensional integral current $a_i - a_{k+j}$ for some $1 \le i, j \le k$; write $j = \sigma(i)$, and note that σ is a permutation of $\{1,\ldots,k\}$. Clearly, therefore, $M(s)$ has length

$$H^1(M(s)) \ge \sum_{i=1}^{k} |a_i - a_{k+\sigma(i)}| = \sum_{i=1}^{k} |a_i + a_{\sigma(i)}| .$$

Since k is odd, we have $H^1(M(s)) \ge 2$ by Lemma 4.3 below.

From the coarea formula (Lemma 2.3) along with inequality (2.4) with $p = n = m-1$, we have

$$\int_{B^m} |\nabla u|^{m-1} dx \ge (m-1)^{(m-1)/2} \int_{S^{m-1}} H^1(u^{-1}(s))\, dA_{S^{m-1}}(s)$$

$$= \frac{1}{2} (m-1)^{(m-1)/2} \int_{S^{m-1}} H^1(M(s))\, dA_{S^{m-1}}(s)$$

$$\ge (m-1)^{(m-1)/2}\, m\, \alpha_m = E_{m-1}(u_0) .$$

q.e.d.

<u>Lemma 4.3</u>. <u>Consider a set</u> $\{a_1,\ldots,a_k\}$ <u>of points (not necessarily distinct) satisfying</u> $|a_i| \ge 1$. <u>If</u> k <u>is odd, then for any permutation</u> σ <u>of</u> $\{1,2,\ldots,k\}$, <u>the sum</u>

$$S := \sum_{i=1}^{k} |a_i + a_{\sigma(i)}| \ge 2 .$$

<u>Remark 4.4</u>. Note that any even value of k allows counter-examples.

<u>Proof</u>. If $\sigma(j) = j$ for some $1 \le j \le k$, then the term $|a_j + a_{\sigma(j)}| = 2|a_j| \ge 2$, and the conclusion follows. If $k = 1$, then $\sigma(1) = 1$, and the conclusion again follows. Thus we may proceed by induction, with the assumption that $\sigma(j) \ne j$, $1 \le j \le k$.

Since $\sigma(k) \ne k$, we may reorder $\{a_1,\ldots,a_k\}$ so that $\sigma(k) = k-1$. Then a_k appears only in the two terms $|a_k + a_{k-1}|$ and $|a_j + a_k|$, where $\sigma(j) = k$. If $j = k-1$, then we may discard these two terms to form the sum

$$\sum_{i=1}^{k-2} |a_i + a_{\sigma(i)}| \ge 2$$

by the induction hypothesis, since the restriction of σ is a permutation of $\{1,\ldots,k-2\}$. If $j \ne k-1$, then a_{k-1} appears in one additional term $|a_{k-1} + a_i|$, where $i = \sigma(k-1)$. By the triangle inequality,

$$|a_k + a_{k-1}| + |a_j + a_k| + |a_{k-1} + a_i| \geq$$

$$\geq |a_{k-1} - a_j| + |a_{k-1} + a_i| \geq |a_j + a_i| .$$

Now define the permutation $\tilde{\sigma}$ on $\{1,\ldots,k-2\}$ so that $\tilde{\sigma}(j) = i$, and otherwise $\tilde{\sigma} = \sigma$. Then the corresponding sum $\tilde{S} \leq S$. But $\tilde{S} \geq 2$ by the induction hypothesis.

q.e.d.

5. The complex Hopf map

The Hopf map $H : S^3 \longrightarrow S^2$ is defined by the restriction to $S^3 = \{(z,w) \in \mathbb{C}^2 : |z|^2 + |w|^2 = 1\}$ of

$$H(z,w) = (|z|^2 - |w|^2, 2z\bar{w}) \in \mathbb{R} \times \mathbb{C} = \mathbb{R}^3 .$$

Let $u_0 : B^4 \longrightarrow S^2$ be its homogeneous extension of degree zero:

$$u_0(x) = H(\frac{x}{|x|}) .$$

Note that $u_0(z,w)$ is the stereographic projection of $z/w \in \mathbb{C} \cup \{\infty\}$.

Theorem 5.1. For any $u \in E_2(H) = \{u \in W^{1,2}(B^4, S^2) : u = H$ on $\partial B^4\}$, $E_2(u) \geq E_2(u_0) = 8\pi^2$.

Proof. According to Theorem 3.2, we may assume u belongs to the class R of $W^{1,2}$ maps with controlled singularities. In particular, u is locally Lipschitz on $B^4 \setminus \Sigma$, where Σ is a one-dimensional Lipschitz manifold without boundary. Further, for almost every $s \in S^2$,

$$M(s) := u^{-1}(s) \cup u^{-1}(-s) \cup \Sigma$$

is a Lipschitz 2-manifold with boundary $H^{-1}(s) \cup H^{-1}(-s)$.

The energy-minimizing property of u_0 will be proved by first showing that the cone $M_0(s)$ over $H^{-1}(s) \cup H^{-1}(-s)$ has smallest area. In fact, $M_0(s)$ is the union of the disks of radius 1 in the 2-dimensional planes

$$\{z = \xi w\} \quad \text{and} \quad \{z = -w/\bar{\xi}\}$$

where $\xi \in \mathbb{C} \cup \{\infty\}$ corresponds under stereographic projection to $s \in S^2$. Now any vector $(\xi w_1, w_1)$ in the first plane is orthogonal to any vector $(-w_2/\bar{\xi}, w_2)$ in the second plane, which implies that $M_0(s)$ is a complex-analytic variety for some orthogonal complex structure (which depends on s) for $\mathbb{R}^4$. In particular, $M_0(s)$ has minimum area among all surfaces in B^4 having boundary $H^{-1}(s) \cup H^{-1}(-s)$ (including unorientable surfaces: see [M], Corollary 6). Specifically,

$$\mathcal{H}^2(M(s)) \geq \mathcal{H}^2(M_0(s)) = 2\pi .$$

It now follows from Lemma 1.4 that

$$E_2(u) \geq 2 \int_{S^2} \mathcal{H}^2(u^{-1}(s))dA_{S^2}(s) =$$

$$= \int_{S^2} \mathcal{H}^2(M(s))dA_{S^2}(s) \geq 8\pi^2 .$$

Meanwhile $|\nabla u_0(x)|^2 = 8/|x|^2$, so that

$$E_2(u_0) = 4\mathcal{H}^3(S^3) = 8\pi^2 .$$

q.e.d.

6. The quaternionic Hopf map

Quaternionic multiplication in $\mathbb{R}^4$ is defined via an orthonormal basis $\{1,i,j,k\}$ with the properties $i^2 = j^2 = k^2 = ijk = -1$, forming a skew field $\mathbb{H}$. The Hopf map $H : S^7 \longrightarrow S^4$ is defined by identifying $\mathbb{R}^8$ as $\mathbb{H} \times \mathbb{H}$ and setting

$$H(q_1,q_2) = (|q_1|^2-|q_2|^2,\ 2q_1\bar{q}_2) \in \mathbb{R} \times \mathbb{H} = \mathbb{R}^5 .$$

Let $u_0 : B^8 \longrightarrow S^4$ be its homogeneous extension of degree zero:

$$u_0(x) = H(\tfrac{x}{|x|}) .$$

Note that $u_0(q_1, q_2)$ is the stereographic projection of $q_1 q_2^{-1} \in \mathbb{H} \cup \{\infty\}$.

<u>Theorem 6.1</u>. <u>For any</u> $u \in \mathcal{E}_2(H) = \{u \in W^{1,2}(B^8, S^4) : u = H$ <u>on</u> $\partial B^8\}$, <u>we have</u> $E_2(u) \geq E_2(u_0)$.

<u>Remark 6.2</u>. The map u_0 also minimizes E_4, as may be proved by direct analogy with the proof of Theorem 5.1, and with the proof of Corollary 6 of [M]. The case $p = 2$, however, requires averaging over projections $\pi_Y : S^4 \longrightarrow S^2$, and is more interesting.

<u>Proof</u>. For any $Y \in G_3(\mathbb{R}^5)$, write $\pi_Y : S^4 \longrightarrow S^2$ for the nearest-point projection. According to Lemma 1.2, we have

$$(6.1) \quad c\, E_2(u) = \oint_{Y \in G_3(\mathbb{R}^5)} E_2(\pi_Y \circ u)\, dG(Y) .$$

Write $v = \pi_Y \circ u : B^8 \longrightarrow S^2$, and $v_0 = \pi_Y \circ u_0$. We need to show that $E_2(v) \geq E_2(v_0)$ for any $v \in W^{1,2}(B^8, S^2)$ with $v = \pi_Y \circ H$ on ∂B^8. It suffices to prove this for $v \in R$, according to Theorem 3.2. Applying Lemma 1.4 and regrouping s with $-s$ as before, we have

$$(6.2) \quad E_2(v) \geq \int_{S^2} H^6(v^{-1}(s) \cup v^{-1}(-s))\, dA_{S^2} .$$

Now since $v \in R$,

$$M(s) := v^{-1}(s) \cup v^{-1}(-s) \cup \Sigma \cup \Delta$$

is a 6-dimensional oriented Lipschitz submanifold of B^8, with $\partial M(s) = (\pi_Y \circ H)^{-1}(s) - (\pi_Y \circ H)^{-1}(-s)$ as integral currents with the natural slice orientations ([F], 4.3.1). The cone over $\partial M(s)$ is

$$M_0(s) := v_0^{-1}(s) \cup v_0^{-1}(-s) \cup \Delta .$$

We need to show that

$$(6.3) \quad H^6(M_0(s)) \leq H^6(M(s)) .$$

For each $Y \in G_3(\mathbb{R}^5)$ and each $s \in S^2 = Y \cap S^4$, observe that $\pi_Y^{-1}(s) \cup \pi_Y^{-1}(-s) = S^4 \cap Z$, where Z is the 3-plane spanned by s and the orthogonal complement of Y. According to Lemma 6.3 below, it is enough to verify inequality (6.3) for the special case where Y is spanned by $(1,0)$, $(0,j)$ and $(0,k)$, and where $s = (1,0)$. In this case, Z is spanned by $(1,0)$, $(0,1)$ and $(0,i)$. Write the point $(q_1, q_2) \in \mathbb{R}^8$ in terms of complex variables z_1, w_1, z_2, w_2 by defining $q_\alpha = z_\alpha + w_\alpha j$. Then

$$q_1\bar{q}_2 = z_1\bar{z}_2 + w_1\bar{w}_2 + (z_2 w_1 - z_1 w_2) j ,$$

so that $M_0(s)$ is given by

$$\{(z_1, w_1, z_2, w_2) \in B^8 : g(z_1, w_1, z_2, w_2) := z_2 w_1 - z_1 w_2 = 0\} .$$

Note that the orientation induced on $M_0(s) = g^{-1}(0)$ by $g : B^8 \longrightarrow \mathbb{C}$ from the appropriate orientations on $\mathbb{C}$ and $\mathbb{R}^8$ is consistent with the orientation given in the Approximation Theorem 3.2. For the (standard) complex structure on $\mathbb{R}^8$ given by

$$J(z_1, w_1, z_2, w_2) = (iz_1, iw_1, iz_2, iw_2),$$

$M_0(s)$ is a complex variety, and inequality (6.3) follows, since $\partial M_0(s) = \partial M(s)$ as integral currents ([F], pp. 435 and 652).

On the other hand, $u_0 : B^8 \longrightarrow S^4$ is horizontally conformal ([B], Theorem 7.1.1 and Examples 7.2.1, 8.2.1) and therefore $v_0 : B^8 \longrightarrow S^2$ is horizontally conformal for any choice of Y. It follows from Lemma 1.4 that equality holds in (6.2) when v is replaced by v_0. Finally, using inequality (6.3), inequality (6.2) for v and equation (6.1) for u_0 and for u, we conclude that

$$c\, E_2(u) \geq c\, E_2(u_0) .$$

q.e.d.

The following lemma is known, since it is an immediate consequence of the fact that the Hopf map: $S^7 \longrightarrow S^4$ induces the isomorphism of the symplectic group $Sp(2)$ of quaternionic 2×2 matrices in $SO(8)$ with the oriented double cover of $SO(5)$ (which fact may be proved in analogous fashion to p. 38 of [A]). Since the literature may be unfamiliar to many, we prefer to present a direct proof.

Lemma 6.3. Given $Z_0, Z_1 \in G_3(\mathbb{R}^5)$, there exist rotations $R \in SO(5)$ and $Q \in SO(8)$ such that $R(Z_1) = Z_0$ and $H(Q(q_1,q_2)) = R(H(q_1,q_2))$ for all $q_1, q_2 \in \mathbb{H}$.

Proof. Without loss of generality, we may assume $Z_0 \subset \mathbb{R}^5 = \mathbb{R} \times \mathbb{H}$ is spanned by $(1,0)$, $(0,1)$ and $(0,i)$. According to a theorem of Cayley ([C], p. 71), any $\mathcal{R}_1 \in SO(4)$ may be written in terms of quaternionic multiplication as $R_1(q) = q_1 q q_2$ for some

$q_1, q_2 \in \mathbb{H}$ of norm one. This corresponds to $Q_1 \in SO(8)$ given by $Q_1(p,q) = (q_1 p, \bar{q}_2 q)$. Consider R_1 to be in $SO(5)$ by $R_1(t,q) = (t, R_1(q))$; then $H \circ Q_1 = R_1 \circ H$ (recall that $\overline{pq} = \bar{q}\bar{p}$). By choosing R_1 appropriately, we may achieve $Z_2 = R_1(Z_1)$ so that $(0,1)$ and $(0,i)$ are in Z_2. Next, let $R_2 \in SO(5)$ be the rotation which fixes $(0,i)$, $(0,j)$ and $(0,k)$, while $R_2(1,0) = (\cos 2\theta, \sin 2\theta)$ and $R_2(0,1) = (-\sin 2\theta, \cos 2\theta)$. This corresponds to $Q_2 \in SO(8)$ defined by $Q_2(p,q) = ((\cos\theta)p - (\sin\theta)q, (\sin\theta)p + (\cos\theta)q)$; namely, $H \circ Q_2 = R_2 \circ H$. For two choices of θ, we find $Z_0 = R_2(Z_2)$.

q.e.d.

We would like to conclude our paper with a theorem of more general character, whose proof is analogous to the proofs of Theorems 1.1, 2.4, 5.1 and 6.1.

Consider $u_0 \in W^{1,p}(B^m, S^n)$ for some integer p, $1 \leq p \leq n$. For each $Y \in G_{p+1}(\mathbb{R}^{n+1})$, let $v_0 = \pi_Y \circ u_0$. We require that

(6.4) $v_0 \in C^{0,1}(B^m \setminus \Delta, S^p)$ for some Lipschitz $(m-p-1)$-submanifold $\Delta \subset B^m$ with $\partial\Delta \subset \partial B^m$,

(6.5) there exists a measurable and measure-preserving map $h : S^p \longrightarrow S^p$ such that the difference of slices $M_0(s) := v_0^{-1}(s) - v_0^{-1}(h(s))$ defines an $(m-p)$-dimensional integral current of smallest mass for its boundary; and

(6.6) v_0 is horizontally conformal a.e. in B^m.

Theorem 6.4. Suppose that for almost all $Y \in G_{p+1}(\mathbb{R}^{n+1})$, hypotheses (6.4), (6.5) and (6.6) hold. Then $E_p(u_0) \leq E_p(u)$ for all $u \in W^{1,p}(B^m, S^n)$ with $u = u_0$ on ∂B^m.

References

[ABL] Almgren, F., Browder, W., Lieb, E.: Co-area, liquid crystals, and minimal surfaces. Preprint.

[A] Atiyah, M.: Geometry of Yang-Mills Fields (1978 Fermi Lectures). Pisa: Scuola Normale Superiore, 1979.

[B] Baird, P.: Harmonic maps with symmetry, harmonic morphisms and deformations of metrics. Boston-London-Melbourne: Pitman 1983.

[BZ] Bethuel, F., Zheng, X.: Sur la densité des fonctions régulières entre deux variétés dans des espaces de Sobolev, C.R. Acad. Sci. Paris 303, I, 447-449 (1986) and Density of smooth functions between two manifolds in Sobolev spaces, preprint.

[BCL] Brézis, H., Coron, J.-M., Lieb, E.H.: Harmonic maps with defects. Commun. Math. Phys. 107, 649-705 (1986).

[C] Coxeter, H.S.M.: Regular Complex Polytopes. Cambridge: Cambridge University Press 1974.

[F] Federer, H.: Geometric Measure Theory. Berlin - Heidelberg-New York: Springer 1969.

[HKL] Hardt, R., Kinderlehrer, D., Lin, F.-H.: Existence and partial regularity of static liquid crystal configurations. Commun. Math. Phys. 105, 547-570 (1986).

[HL] Hardt, R., Lin, F.-H.: Mappings that minimize the p th power of the gradient. In preparation.

[H] Hélein, : In preparation.

[JK] Jäger, W., Kaul, H.: Rotationally symmetric harmonic maps from a ball into a sphere and the regularity problem for weak solutions of elliptic systems. J. Reine Angew. Math. 343, 146-161 (1983).

[L] Lin, F.-H.: private communication.

[Md] Macdonald, I.G.: Symmetric Functions and Hall Polynomials. Cambridge: Clarendon Press 1979.

[M] Morgan, F.: On the singular structure of two-dimensional area minimizing surfaces in $\mathbb{R}^n$. Math. Ann. 261, 101-110 (1982).

[SU1] Schoen, R., Uhlenbeck, K.: A regularity theory for harmonic maps. J. Diff. Geom. 17, 307-335 (1982).

[SU2] Schoen, R., Uhlenbeck, K.: Boundary regularity and the Dirichlet problem for harmonic maps. J. Differ. Geom. 18, 253-268 (1983).

[SU3] Schoen, R., Uhlenbeck, K.: Regularity of minimizing harmonic maps into the sphere. Invent. Math. 78, 89-100 (1984).

CO-AREA, LIQUID CRYSTALS, AND MINIMAL SURFACES[1]

F. Almgren, W. Browder, and E. H. Lieb
Department of Mathematics, Princeton University
Princeton, New Jersey 08544, USA

Abstract. Oriented n area minimizing surfaces (integral currents) in $\mathcal{M}^{m+n}$ can be approximated by level sets (slices) of nearly m-energy minimizing mappings $\mathcal{M}^{m+n} \to \mathbf{S}^m$ with essential but controlled discontinuities. This gives new perspective on multiplicity, regularity, and computation questions in least area surface theory.

In this paper we introduce a collection of ideas showing relations between co-area, liquid crystals, area minimizing surfaces, and energy minimizing mappings. We state various theorems and sketch several proofs. A full treatment of these ideas is deferred to another paper.

Problems inspired by liquid crystal geometries.[2] Suppose Ω is a region in 3 dimensional space $\mathbf{R}^3$ and f maps Ω to the unit 2 dimensional sphere $\mathbf{S}^2$ in $\mathbf{R}^3$. Such an f is a unit vectorfield in Ω to which we can associate an 'energy'

$$\mathcal{E}(f) = \left(\frac{1}{8\pi}\right) \int_\Omega |Df|^2 \, d\mathcal{L}^3;$$

here Df is the differential of f and $|Df|^2$ is the square of its Euclidean norm—in terms of coordinates,

$$|Df(x)| = \sum_{k=1}^{3} \sum_{i=1}^{3} \left(\frac{\partial f^k}{\partial x_i}(x)\right)^2$$

for each x. The factor $1/8\pi$ which equals 1 divided by twice the area of $\mathbf{S}^2$ is a useful normalizing constant. It is straightforward to show the existence of f's of least energy for given boundary values (in an appropriate function space).

[1] This research was supported in part by grants from the National Science Foundation
[2] The research which led to the present paper began as an investigation of a possible equality between infimums of m-energy and the n area of area minimizing n dimensional area minimizing manifolds in $\mathbf{R}^{m+n}$ suggested in section VIII(C) of the paper, *Harmonic maps with defects* [BCL] by H. Brezis, J-M. Coron, and E. Lieb. Although the specific estimates suggested there do not hold (by virtue of counterexamples [MF][W1][YL]) their general thrust does manifest itself in the results of the present paper.

Such boundary value problems have been associated with liquid crystals.[3] In this context, a "liquid crystal" in a container Ω is a fluid containing long rod like molecules whose directions are specified by a unit vectorfield. These molecules have a preferred alignment relative to each other—in the present case the preferred alignment is parallel. If we imagine the molecule orientations along $\partial\Omega$ to be fixed (perhaps by suitably etching container walls) then interior parallel alignment may not be possible. In one model the system is assumed to have 'free energy' given by our function $\mathcal{E}$ and the crystal geometry studied is that which minimizies this free energy.

If Ω is the unit ball and $f(x) = x$ for $|x| = 1$, then there is no continuous extension of these boundary values to the interior; indeed the unique least energy f is given by setting $f(x) = x/|x|$ for each x. It turns out that this singularity is representative, and the general theorem is that *least energy f's exist and are smooth except at isolated points p of discontinuity where 'tangential structure' is* $\pm x/|x|$ *(up to a rotation), e.g. f has local degree equal to* ± 1 [SU] [BCL VII].

As a further step towards an understanding of the geometry of of energy minimizing f's one might seek estimates on the number of points of discontinuity which such an f can have—e.g. if the boundary values are not to wild must the number of points of discontinuity be not too big?[4] An alternative problem to this is to seek a lower bound on the energy when the points of discontinuity are prescribed together with the local degrees of the mapping being sought. This question has a surprisingly simple answer as follows.

THEOREM. *Suppose* $p_1, \ldots, p_N$ *are points in* $\mathbf{R}^3$ *and* $d_1, \ldots, d_N \in \mathbf{Z}$ *are the prescribed degrees with* $\sum_{i=1}^N d_i = 0$. *Let* $\inf \mathcal{E}$ *denote the infimum of the energies of (say, smooth) mappings from* $\mathbf{R}^3 \sim \{p_1, \ldots, p_N\}$ *to* $\mathbf{S}^2$ *which map to the 'south pole' outside some bounded region in* $\mathbf{R}^3$ *and which, for each* i, *map small spheres around* p_i *to* $\mathbf{S}^2$ *with degree* d_i. *Then* $\inf \mathcal{E}$ *equals the least mass* $\mathbf{M}(T)$ *of integral 1 currents* T *in* $\mathbf{R}^3$ *with*

$$\partial T = \sum_{i=1}^{N} d_i [p_i].$$

This fact (stated in slightly different language) is one of the central results of [BCL]. We would like to sketch a proof in two parts: first by showing that $\inf \mathcal{E} \le \inf \mathbf{M}$ (with

[3] See, for example, the discussion by R. Hardt, D. Kinderlehrer, and M. Luskin in [HKL].
[4] As it turns out, away from the boundary of Ω, the number of these points is bounded *a priori* independent of boundary values.

the obvious meanings) and then by showing that $\inf \mathbf{M} \leq \inf \mathcal{E}$. The proof of the first part follows [BCL] while the second part is new. It is in this second part that the coarea formula makes its appearance.

Proof that $\inf \mathcal{E} \leq \inf \mathbf{M}$. The first inequality is proved by construction as illustrated in Figure 1. We there represent that case in which N equals 2 and p_1 and p_2 are distinct points with $d_1 = -1$ and $d_2 = +1$. We choose and fix a smooth curve C connecting these two points and orient C by a smoothly varying unit tangent vector field ς which points away from p_1 and towards p_2. The associated 1 dimensional integral current is $T = \mathbf{t}(C, 1, \varsigma)$ and its *mass* $\mathbf{M}(T)$ is the length of C since the density specified is everywhere equal to 1.[5] We now choose (somewhat arbitrarily) and fix two smoothly varying unit normal vector fields η_1 and η_2 along C which are perpendicular to each other and for which, at each point x of C, the 3-vector $\eta_1(x) \wedge \eta_2(x) \wedge \varsigma(x)$ equals the orienting 3-vector $\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3$ for $\mathbf{R}^3$. These two vector fields are a 'framing' of the normal bundle of C.

We then construct a mapping γ of $\mathbf{R}^2$ onto the unit 2 sphere $\mathbf{S}^2$ which is a slight modification of the inverse to stereographic projection. To construct such γ we fix a huge radius R in $\mathbf{R}^2$ and require: (i) if $|y| \leq R$ then $\gamma(y)$ is that point in $\mathbf{S}^2$ which maps to y under stereographic projection $\mathbf{S}^2 \to \mathbf{R}^2$ from the south pole q of $\mathbf{S}^2$; (ii) if $|y| \geq 2R$ then $\gamma(y) = \mathrm{q}$; (iii) for $R < |y| < 2R$, $\gamma(y)$ is suitably interpolated. See Appendix A.2.

Next we choose some smoothly varying (and very small) radius function δ on C which vanishes only at the endpoints p_1 and p_2.

Finally, as our mapping f from $\mathbf{R}^3$ to $\mathbf{S}^2$ with which to estimate $\mathcal{E}(f)$ we specify the following. If p in $\mathbf{R}^3$ can be written $p = x + s\eta_1(x) + t\eta_2(x)$ for some x in C and some s and t with $s^2 + t^2 \leq \delta(x)^2$, then

$$f(p) = \gamma\left(\frac{2Rs}{\delta(x)}, \frac{2Rt}{\delta(x)}\right).$$

Otherwise, $f(p) = \mathrm{q}$. We leave it as an exercise to the reader to use the fact that γ is

[5] Formally, a 1 current such as T is a linear functional on smooth differential 1 forms in $\mathbf{R}^3$. If φ is such a 1 form then

$$T(\varphi) = \int_{x \in C} \langle \varsigma(x), \varphi(x) \rangle \, d\mathcal{H}^1 x.$$

To each point p in $\mathbf{R}^3$ is associated the 0 dimensional current $[p]$ which maps the smooth function ψ to the number $\psi(p)$. See Appendix A.4.

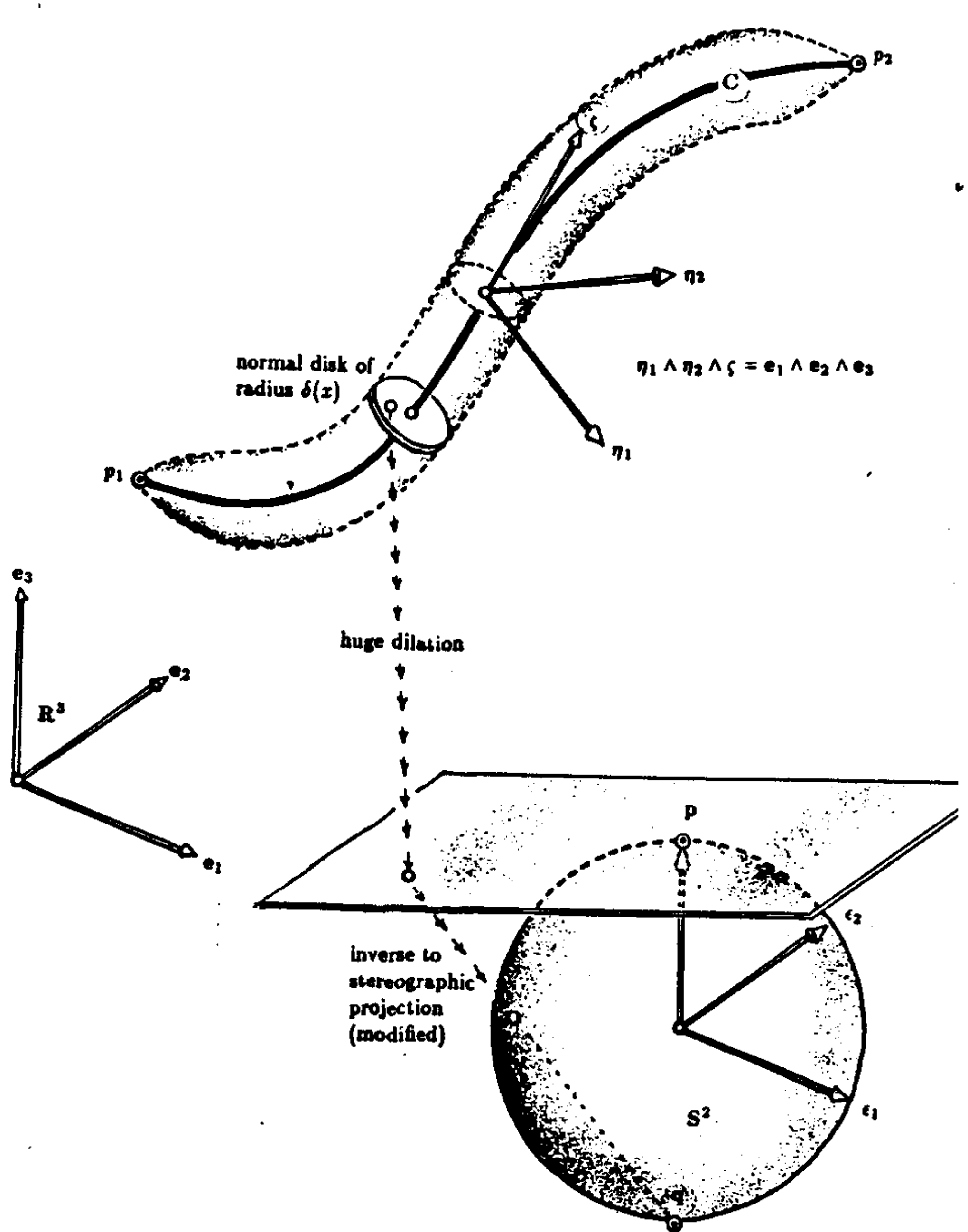

Figure 1. Construction of a mapping f (indicated by dashed arrows) from $\mathbf{R}^3$ to $\mathbf{S}^2$ having energy $\mathcal{E}(f)$ not much greater than the length of the curve C connecting the points p_1 and p_2. Small disks normal to C map by f to cover $\mathbf{S}^2$ once in a nearly conformal way. This implies that small spheres around p_1 map to $\mathbf{S}^2$ with degree -1 while small spheres around p_2 map with degree $+1$. The 1 current $\mathbf{t}(C, 1, \varsigma)$ is the slice $\langle \mathbf{E}^3, f, \mathrm{p} \rangle$ of the Euclidean 3 current $\mathbf{E}^3$ by the mapping f and the 'north pole' p of $\mathbf{S}^2$.

conformal for $|y| < R$ to check that $\mathcal{E}(f)$ very nearly equals $\mathbf{M}(T)$; see Appendix A.2. The remainder of the proof that $\inf \mathcal{E} \leq \inf \mathbf{M}$ is also left to the reader.

Proof that $\inf \mathbf{M} \leq \inf \mathcal{E}$. Suppose that f does map $\mathbf{R}^3$ to $\mathbf{S}^2$, has degree d_i at each p_i, and maps to the south pole outside some bounded region. From dimensional considerations one would expect that for most points w in $\mathbf{S}^2$ the inverse image $f^{-1}\{w\}$ would be a collection of curves connecting the various points $p_1, \ldots, p_N$. H. Federer's *coarea formula* is what enables one to quantify this idea; see Appendix A.5. This formula asserts

$$\int_{w \in \mathbf{S}^2} \mathcal{H}^1(f^{-1}\{w\})\, d\mathcal{H}^2 w = \int_{x \in \mathbf{R}^3} J_2 f(x)\, d\mathcal{L}^3 x;$$

here $\mathcal{H}^1$ and $\mathcal{H}^2$ are Hausdorff's 1 and 2 dimensional measures in $\mathbf{R}^3$ and $\mathcal{L}^3$ is Lebesgue's 3 dimensional measure for $\mathbf{R}^3$. Also $J_2 f(x)$ here denotes the 2 dimensional Jacobian of f at x and a key observation (as noted in [BCL]) is that $J_2 f(x)$ is always less than or equal to half of $|Df(x)|^2$ with equality only if the differential mapping $Df(x): \mathbf{R}^3 \to \operatorname{Tan}(\mathbf{S}^2, f(x))$ is maximally conformal; see Appendix A.1.3. Also central to the present analysis is the manner in which the curves $f^{-1}\{w\}$ connect the various points $p_1, \ldots, p_N$ and how they relate to the prescribed degrees $d_1, \ldots, d_N$. This connectivity is naturally measured by the current structure of these $f^{-1}\{w\}$'s which comes from the slicing theory for currents; see Appendix A.5. To set this up we regard $\mathbf{R}^3$ as the Euclidean current $\mathbf{E}^3$ (oriented by the 3 vector $\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3$). The *slice* of $\mathbf{E}^3$ by the map f at the point w in $\mathbf{S}^2$ is the current

$$\langle \mathbf{E}^3, f, w \rangle = \mathbf{t}(f^{-1}\{w\}, 1, \varsigma);$$

the meanings here are the same as for the current T discussed above. A check of orientations and degrees shows that

$$\partial \langle \mathbf{E}^3, f, w \rangle = \sum_{i=1}^{N} k_i [p_i];$$

compare with our construction of η_1 and η_2 above. It follows immediately that

$$\begin{aligned} 4\pi \inf \mathbf{M}(T) &= \mathcal{H}^2(\mathbf{S}^2) \inf \mathbf{M}(T) \\ &\leq \int_{w \in \mathbf{S}^2} \mathbf{M}(\langle \mathbf{E}^3, f, w \rangle)\, d\mathcal{H}^2 w \\ &= \int_{\mathbf{R}^3} J_2 f\, d\mathcal{L}^3 \\ &= \left(\frac{1}{2}\right) \int_{\mathbf{R}^3} |Df|^2\, d\mathcal{L}^3. \end{aligned}$$

This finishes the proof that $\inf \mathbf{M} \leq \inf \mathcal{E}$.

First Generalization. Since the methods used in the proofs of the two inequalities are quite general one might correctly suspect that considerable generalization is possible. Suppose, for example, we fix $B = \{p_1, \ldots, p_N\}$ as a general boundary set and let $\mathcal{F}_0$ be the family of those mappings f of $\mathbf{R}^3$ to $\mathbf{S}^2$ which are locally Lipschitzian except possibly on B, which map to the southpole outside some bounded region, and which have finite energy. Since deformations of mappings in $\mathcal{F}_0$ do not alter discrete combinatorial structures we are led to study properties of homotopy classes $\Pi(\mathcal{F}_0)$ of mappings in $\mathcal{F}_0$—it is most useful here if our homotopies $[0,1] \times \mathbf{R}^3 \to \mathbf{S}^2$ are permitted to have isolated point discontinuties; see Appendix A.3.

Our conditions about mapping degrees above generalize to requirements about degrees $\mathbf{d}(f, S)$ of f on general integral 2 dimensional cycles S in $\mathbf{R}^3 \sim B$. It turns out that such a degree $\mathbf{d}(f, S)$ depends only on the homotopy class of f and on the homology class of S.

It also turns out that the relative homology classes of the slices $\langle \mathbf{E}^3, f, w \rangle$ depend only on the homotopy class $[f]$ of f. We denote this homology class by $\mathbf{s}[f]$.

The *Kronecker index* is a pairing between 2 dimensional cycles S in $\mathbf{R}^3 \sim B$ and 1 currents T having boundary in B. In general the Kronecker index $\mathbf{k}(S, T)$ is the sum over points of intersection of S and T of an index of relative orientations; see Appendix A.6

These various ideas are related in the following theorem.

THEOREM. *The diagram below is commutative. Furthermore,* $\mathbf{s}$ *is an isomorphism, and* $\mathbf{d}$ *and* $\mathbf{k}$ *are injections.*

$$\begin{array}{ccc} & & \mathbf{H}_1(\mathbf{R}^3, B; \mathbf{Z}) \\ & \nearrow \mathbf{s} & \\ \Pi(\mathcal{F}_0) & & \downarrow \mathbf{k} \\ & \searrow \mathbf{d} & \\ & & \operatorname{Hom}(\mathbf{H}_2(\mathbf{R}^3 \sim B, \mathbf{Z}), \mathbf{Z}) \end{array}$$

Here

$\mathbf{s}[f] =$ "$[f^{-1}\{w\}]$" $= [\langle \mathbf{E}^3, f, w \rangle] =$ the integral homology class of the 1 current slice;

$\mathbf{d}[f][S] = \mathbf{d}(f, S) =$ the degree of f on the 2 cycle S;

$k|T||S| = \mathbf{k}(S,T) =$ the Kronecker Index of the 2 cycle S and the 1 current T.

Our relations between energy minimization and area minimization become the following.

THEOREM. *Suppose that P is an integral 1 current in $\mathbf{R}^3$ with the support of ∂P in B. Suppose also that $T^{\mathbf{Z}}$ has least mass among all integral 1 currents which are homologous to P over the integers $\mathbf{Z}$ and that $T^{\mathbf{R}}$ has least mass among all integral 1 currents which are homologous to P over the real numbers $\mathbf{R}$. Then*

$$\mathbf{M}(T^{\mathbf{Z}}) = \inf\{\mathcal{E}(f) : s|f| = |P|\}$$

and

$$\mathbf{M}(T^{\mathbf{R}}) = \inf\{\mathcal{E}(f) : d|f| = k|P|\}.$$

Moreover, $\mathbf{M}(T^{\mathbf{Z}}) = \mathbf{M}(T^{\mathbf{R}})$ *(because of our special situation).*

Further generalizations. The essential ingredients of the analyses above remain, for example, if $\mathbf{R}^3$ is replaced by a general $m+n$ dimensional manifold $\mathcal{M}$ (without boundary) which is smooth, compact, and oriented (or $\mathcal{M} = \mathbf{R}^{m+n}$), and B is replaced by a sufficiently nice (possibly empty) compact subset of $\mathcal{M}$ of dimension $n-1$. To study n dimensional integral currents in $\mathcal{M}$ having boundary in B we consider mappings f of $\mathcal{M}$ to a sphere of the complementary dimension m. The spaces $\mathcal{F}$ and $\mathcal{F}_0$ of such mappings and the homotopy classes $\Pi(\mathcal{F})$ are specified in sections A.3.1 and A.3.2 of the Appendix. Some discontinuities are essential.[6] It seems worthwhile to consider three different energies $\mathcal{E}_1$, $\mathcal{E}_2$, and $\mathcal{E}_3$ for mappings in $\mathcal{F}_0$. $\mathcal{E}_1$ is a normalization of the usual 'n energy' of mappings, $\mathcal{E}_3$ is a normalized Jacobian integral associated with the coarea formula, and $\mathcal{E}_2$ is an intermediate energy; see Appendix A.3.2. As indicated above, mapping degrees and the Kronecker index have general meanings which are set forth in sections A.6 and A.7 of the Appendix. These various ideas are related as the following theorem shows.

THEOREM. *The diagram of mappings below is well defined and is commutative. In particular, the images of d and k and j in $\mathbf{Hom}(\mathbf{H}_m(\mathcal{M} \sim B, \mathbf{Z}), \mathbf{Z})$ are the same.*

[6] Suppose $m=2$ and $n=5$ and $\mathcal{M} = \mathbf{R}^7$, and B is a smoothly embedded copy of 2 dimensional complex projective space $\mathbf{CP}(2)$. Then there are no continuous mappings f from the complement of B to $\mathbf{S}^2$ such that small 2 spheres S which link B once map to $\mathbf{S}^2$ with degree one. Any f satisfying such a linking condition for general position S's near B must have interior discontinuities of dimension at least 3.

Furthermore, s is an isomorphism.

$$\begin{array}{ccccc}
 & & \mathbf{H}_n(\mathcal{M},B;\mathbf{Z}) & \xrightarrow{c} & \mathbf{H}_n(\mathcal{M},B;\mathbf{R}) \\
 & \nearrow s & & \searrow c & \uparrow i \\
\Pi(\mathcal{F}) & & \downarrow k & & c[\mathbf{H}_n(\mathcal{M},B;\mathbf{Z})] \\
 & \searrow d & & \swarrow j & \\
 & & \mathbf{Hom}(\mathbf{H}_m(\mathcal{M}\sim B,\mathbf{Z}),\mathbf{Z}) & &
\end{array}$$

Here

$s|f| =$ "$|f^{-1}\{p\}|$" $= |\langle [\![\mathcal{M}]\!], f, p\rangle| =$ the integral homology class of the n current slice;

$d|f||S| = \mathbf{d}(f,S) =$ the degree of f on the m cycle S;

$k|T||S| = \mathbf{k}(S,T) =$ the Kronecker index of the m cycle S and the n current T;[7]

c is induced by the coefficient inclusion $\mathbf{Z} \to \mathbf{R}$;

i is the inclusion; and

j is defined by commutivity.

We defer proof of this theorem to our fuller treatment of this subject. The natural setting and generality of such relationships are still under investigation.

The relations between energy minimization and area minimization then become the following.

MAIN THEOREM. *Suppose P is an integral current in $\mathcal{M}$ with the support of ∂P contained in B so that the integral homology class $|P|$ of P belongs to $\mathbf{H}_n(\mathcal{M},B;\mathbf{Z})$.*

[7] Suppose $m=2$ and $n=1$ and $\mathcal{M}$ is a 3 dimensional real projective space $\mathbf{RP}(3)$ and $T = \mathbf{t}(\mathcal{N}, 1, \varsigma)$; here $\mathcal{N}$ is a 1 dimensional real projective space $\mathbf{RP}(1)$ sitting in $\mathbf{RP}(3)$ in the usual way and ς is some orientation function. Since T is not a boundary while $2T$ is, we conclude that the homology class

$$|T| \in \mathbf{H}_1(\mathcal{M}, \emptyset;\ \mathbf{Z}) = \mathbf{Z}_2$$

is not the 0 class although $\mathbf{k}(S,T) = 0$ for each 2 cycle S in $\mathcal{M}$. In particular, the mapping k is generally not an injection.

Let $T^{\mathbf{Z}}$ be an integral current of least mass among all integral currents belonging to the ame integral homology class as P in $\mathbf{H}_n(\mathcal{M},B,\mathbf{Z})$, and let $T^{\mathbf{R}}$ be an integral current of east mass among all integral currents belonging to the same real homology class as P in $\mathbf{H}_n(\mathcal{M},B,\mathbf{R})$. Then

$$\mathbf{M}(T^{\mathbf{Z}}) = \inf\{\mathcal{E}_1(f): s[f] = [P]\} = \inf\{\mathcal{E}_2(f): s[f] = [P]\} = \inf\{\mathcal{E}_3(f): s[f] = [P]\}$$

nd

$$\mathbf{M}(T^{\mathbf{R}}) = \inf\{\mathcal{E}_1(f): d[f] = k[P]\} = \inf\{\mathcal{E}_2(f): d[f] = k[P]\} = \inf\{\mathcal{E}_3(f): d[f] = k[P]\}.$$

In general, of course, $\mathbf{M}(T^{\mathbf{R}}) < \mathbf{M}(T^{\mathbf{Z}})$. Although we again defer complete proofs to ur fuller treatment of this subject, it does seem useful to sketch some of the main ideas.

Proof of the inequality "$\inf \mathcal{E} \leq \inf \mathbf{M}$". The proof here is again by construction. Ve will indicate the main ingredients in a special case. Suppose, say, $\mathcal{M} = \mathbf{R}^{m+n}$, B polyhedral, and T is an integral n current which is mass minimizing subject to some ppropriate constraints as in the Main Theorem above. We will construct a mapping $: \mathbf{R}^{m+n} \to \mathbf{S}^m$ in the relevant homotopy class such that $\mathcal{E}_1(f), \mathcal{E}_2(f)$, and $\mathcal{E}_3(f)$ are early equal and are not much bigger that $\mathbf{M}(T)$. By virtue of the Strong Approximation heorem for integral currents [FH1 4.2.20] we can modify T slightly to become simplicial ith only a slight increase in mass.

Suppose then that we can express

$$T = \sum_{\alpha=1}^{M} \mathbf{t}(\Delta^n_\alpha, z_\alpha, \zeta_\alpha)$$

s a 'simplicial' integral current (with the obvious interpretation). For each $k = 0, \ldots, n$ e denote by K_k the collection of closed k simplexes which occur as k dimensional faces f n simplexes among the Δ^n_α's. We then choose numbers $0 < \delta_n << \delta_{n-1} << \delta_{n-2} << \ldots << \delta_0 << 1$ and define sets $N_0, N_1, \ldots, N_n$ in $\mathbf{R}^{m+n}$ by setting

$$N_0 = \{x: \operatorname{dist}(x, \cup K_0) < \delta_0\}$$

nd, for each $k = 1, \ldots, n$ set

$$N_k = \{x: \operatorname{dist}(x, \cup K_k) < \delta_k\} \sim (N_{k-1} \cup N_{k-2} \cup \ldots \cup N_0).$$

We assume that $\delta_0, \ldots, \delta_n$ have been chosen so that the distinct components of each N_k correspond to distinct k simplexes in K_k.

We now define mappings $f_{n+1}, f_n, \ldots, f_0 = f$ as follows.

First, the mapping $f_{n+1}: \mathbf{R}^{m+n} \sim (N_n \cup \ldots \cup N_0) \to \mathbf{S}^m$ is defined by setting $f_{n+1}(x) = \mathbf{q}$ for each x.

Second, the mapping $f_n: \mathbf{R}^{m+n} \sim (N_{n-1} \cup \ldots \cup N_0) \to \mathbf{S}^m$ is constructed geometrically in virtually the same manner as the mapping g in the example A.8 in the Appendix. Details are left to the reader.

Third, the mapping $f_{n-1}: \mathbf{R}^{m+n} \sim (N_{n-2} \cup \ldots \cup N_0) \to \mathbf{S}^m$ is constructed geometrically in a manner virtually identical with the construction of the mapping $f_{\delta,r}$ of example A.8 of the Appendix (with δ, r replaced by $\delta_n/2, \delta_{n-1}$ respectively there). The mapping f_{n-1} is Lipschitz across parts of $n-1$ simplexes which do not lie in B and is discontinuous on those $n-1$ simplexes which contain part of ∂T.

Assuming $f_{n+1}, f_n, \ldots, f_{k+1}$ have been constructed we define

$$f_k: \mathbf{R}^{m+n} \sim (N_{k-1} \cup \ldots \cup N_0) \to \mathbf{S}^m$$

as follows. Each point v in $N_k \sim (N_{k-1} \cup \ldots \cup N_0)$ can be written uniquely in the form $v = v_0 + (v - v_0)$ where v_0 is the unique closest point in $\cup K_k$ to v and $|v - v_0| < \delta_k$. If $v \neq v_0$ we note that

$$v_1 = v_0 + \delta_k \left(\frac{v - v_0}{|v - v_0|}\right) \in \operatorname{dmn}(f_{k+1})$$

and we set $f_k(v) = f_{k+1}(v_1)$. A direct extension of the estimates used for the example A.8 of the Appendix shows that the energies $\mathcal{E}_1(f), \mathcal{E}_2(f)$, and $\mathcal{E}_3(f)$ very nearly equal $\mathbf{M}(T)$.

Proof of the inequality "$\inf \mathbf{M} \leq \inf \mathcal{E}$". The argument here is a direct extension of the corresponding argument given above and is left to the reader.

Remarks.

(1) One of the main reasons for analyzing relations between the energy of mappings and the area of currents is that it provides a way to study n dimensional area minimizing *integral* currents (whose geometry is not specified ahead of time) by studying functions and integrals over the given ambient manifold. This seems the first such scheme which works in general codimensions. For *real* currents, however, differential forms play a role roughly analogous to that of our function spaces $\mathcal{F}_0$; in this regard see, for example, the paper

of H. Federer, *Real flat chains, cochains, and variational problems* [F2 4.10(4), 4.11(2)]. Incidentally, in the language of [F2 5.12, page 400], examples show that the equation in question there is not always true under the alternative hypotheses of [F2 5.10].

(2) Suppose C consists of smooth simple closed curves in $\mathbf{R}^3$ oriented by ς. Suppose also for positive integers ν we have reasonable mappings f_ν from the complement of C in $\mathbf{R}^3$ to the circle $\mathbf{S}^1$ with the property that small circles which link C once are mapped to $\mathbf{S}^1$ by f_ν with degree ν. Because of the dimensions we have

$$\mathcal{E}_1(f_\nu) = \mathcal{E}_2(f_\nu) = \mathcal{E}_3(f_\nu) = \left(\frac{1}{2\pi}\right) \int |Df_\nu| \, d\mathcal{L}^3.$$

If f_ν is nearly $\mathcal{E}_1$ energy minimizing then for most w's in $\mathbf{S}^1$ the slice

$$T_\nu(w) = \langle \mathbf{E}^3, f_\nu, w \rangle \in \mathbf{I}_2(\mathbf{R}^3)$$

will be defined with $\partial T_\nu(w) = \mathbf{t}(C, \nu, \varsigma)$ and will be nearly mass minimizing. H. Parks, in his memoir, *Explicit determination of area minimizing hypersurfaces, II* [PH], used a similar energy for mappings to the real numbers $\mathbf{R}$ (instead of to $\mathbf{S}^1$) and was able to exhibit an algorithm for finding area minimizing surfaces. The technique used by Parks requires that C be extreme, i.e. that it lie on the boundary of its convex hull. The analysis of our paper on the other hand applies to any collection of curves which, for example, may be knotted or linked in any way. One of our hopes is to develop a method of computation analogous to that of Parks.

(3) Suppose that C and the mappings f_ν have the same meaning as in (2) above. If θ denotes the usual (multiple-valued radian) angle function on $\mathbf{S}^1$ then $d\theta$ as a well defined closed 1-form whose pullbacks $f_\nu^\sharp d\theta$ give closed 1 forms on the complement of C in $\mathbf{R}^3$ with $|f_\nu^\sharp d\theta| = |Df_\nu|$. For fixed x_0 in the complement of C we define functions g_ν mapping the complement of C to $\mathbf{S}^1$ by requiring that

$$\theta \circ g_\nu(x) = \theta \circ f_\nu(x_0) + \int_\gamma f_\nu^\sharp d\theta \pmod{2\pi}$$

for each x (with the obvious meanings); here $\gamma(x)$ denotes any oriented path in the complement of C starting at x_0 and ending at x. It is immediate to check that $g_\nu = f_\nu$ for each ν. If we write $\nu = \lambda \cdot \mu$ for some λ and μ and define $h_\lambda(x)$ in $\mathbf{S}^1$ by requiring

$$\theta \circ h_\lambda(x) = \int_\gamma \left(\frac{1}{\mu}\right) f_\nu^\sharp d\theta \pmod{2\pi}$$

for γ as above. The mapping h_λ maps small circles with the same degrees as does f_λ. Taking $\mu = \nu$ we readily conclude, for example, that

$$\inf\{\mathbf{M}(T) : \partial T = \mathbf{t}(C, \nu, \varsigma)\} = \nu \cdot \inf\{\mathbf{M}(T) : \partial T = \mathbf{t}(C, 1, \varsigma)\}$$

for each ν. This estimate implies that integral and real mass minimizing 2 currents having boundary $\mathbf{t}(C, 1, \varsigma)$ have the same masses [F2 5.8]; although this has been known for some time, the present proof by factoring mappings seems new and simpler. This fact (and our proof) extend to $n-1$ dimensional boundaries in general manifolds $\mathcal{M}$ of dimension $n+1$ with, for example, the property that each 1 cycle is a boundary. There are counterexamples to such equalities in higher codimensions given first by L. C. Young [YL] and later by F. Morgan [MF] and B. White [W1]. How *badly* such an equality can fail remains an important open question. It is not even known, for example, if the number

$$\inf\{\mathbf{M}(S)/\mathbf{M}(T) : S, T \in \mathbf{I}_2(\mathbf{R}^4, \mathbf{R}^4) \text{ are mass minimizing with } 0 \neq \partial S = 2\partial T\}$$

is positive; note, however, the isoperimetric inequality [A1 2.6].

(4) Suppose $\mathcal{M}$ is a complex submanifold of some complex projective space $\mathbf{CP}(n)$ (or, more generally, $\mathcal{M}$ is a Kähler manifold). Then any complex analytic (meromorphic) function f from $\mathcal{M}$ to the Riemann Sphere $\mathbf{CP}(1) = \mathbf{S}^2$ has integral current slices which are absolutely mass minimizing in their integral homology classes [F1 5.4.19]. Such f's are thus necessarily maximally conformal and minimize each of the energies $\mathcal{E}_1$, $\mathcal{E}_2$, and $\mathcal{E}_3$ among functions in the same homotopy classes.

(5) In the context of this paper, if the mass minimizing current T being sought happens to be unique then most slices of nearly minimizing mappings will be close to that current. In a sense this describes the asymptotic behavior of a sequence $\{f_k\}_k$ of mappings in $\mathcal{F}_0$ converging towards energy minimization; in particular, the real currents

$$\left\{ \left(\frac{1}{(m+1)\alpha(m+1)} \right) [\![\mathcal{M}]\!] \llcorner f_k^\sharp \sigma^\cdot \right\}_k$$

must converge to T as $k \to \infty$. If $m = 2$ then the energy $\mathcal{E}_1$ is Dirichlet's integral which is widely studied in the general theory of harmonic mappings between manifolds pioneered by J. Eells and J. Sampson.

In any codimension m each n dimensional mass minimizing integral current is a *regular* minimal submanifold except possibly on a singular set of dimension not exceeding $n-2$ as shown by F. Almgren in [A2]. It is not yet clear to what extent the present new setup will rovide new tools for study of the regularity and singularity properties of mass minimizing itegral currents. This could be one of its most important potential uses.

APPENDIX

When not otherwise specified we follow the general terminology of pages 669-671 of H. 'ederer's treatise, *Geometric Measure Theory* [**F1**] or the newer standardized terminology f the 1984 AMS Summer Research Institute in Geometric Measure Theory and the Calulus of Variations as summarized in pages 124-130 of F. Almgren's paper, *Deformations nd multiple-valued functions* [**A1**].

A.1 Terminology.

A.1.1 We fix positive integers m and n and suppose that M is an $m+n$ dimenional submanifold (without boundary) of $\mathbf{R}^N$ (some N) which is smooth, compact, and riented by the continuous unit $(m+n)$-vectorfield $\xi: M \to \wedge_{m+n}\mathbf{R}^N$; alternatively $\mathbf{l} = \mathbf{R}^{m+n}$ with standard orthonormal basis vectors $\mathbf{e}_1, \ldots, \mathbf{e}_{m+n}$ and orienting $(m+n)$-ector $\mathbf{e}_1 \wedge \ldots \wedge \mathbf{e}_{m+n}$. We also suppose that B is a finite (possibly empty) union of various :urvilinear) $n-1$ simplexes $\Delta_1, \Delta_2, \ldots, \Delta_J$ associated with some smooth triangulation f M.

A.1.2 We denote by $\mathbf{S}^m$ the unit sphere in $\mathbf{R} \times \mathbf{R}^m = \mathbf{R}^{1+m}$ with its usual orintation given by the unit m-vectorfield $\sigma: \mathbf{S}^m \to \wedge_m \mathbf{R}^{1+m}$; in particular, for each $w \in$ $^m \subset \mathbf{R}^{1+m} = \wedge_1 \mathbf{R}^{1+m}$, $\sigma(w) = *w$. It is convenient to let $z, y_1, \ldots, y_m$ denote the sual orthonormal coordinates for $\mathbf{R} \times \mathbf{R}^m$ and also let $\mathbf{p}, \varepsilon_1, \ldots, \varepsilon_m$ be the associated rthonormal basis vectors. In particular, $\sigma(\mathbf{p}) = *\mathbf{p} = \varepsilon_1 \wedge \ldots \wedge \varepsilon_m$. We regard $\mathbf{p}$ as the orth pole' of $\mathbf{S}^m$. The 'south pole' is $\mathbf{q} = -\mathbf{p}$. We denote by σ^* the differential m form the 'volume form') on $\mathbf{S}^m$ dual to σ.

A.1.3 If L is a linear mapping $\mathbf{R}^{m+n} \to \mathbf{R}^m$ then the polar decomposition theorem uarantees the existence of orthonormal coordinates for $\mathbf{R}^{m+n}$ and $\mathbf{R}^m$ with respect to hich L has the matrix representation

$$L = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_m & 0 & \cdots & 0 \end{pmatrix}$$

with $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m \geq 0$. In these coordinates we can express the Euclidean norm $|L|$ of L as

$$|L| = (\lambda_1^2 + \lambda_2^2 + \ldots + \lambda_m^2)^{\frac{1}{2}},$$

express the mapping norm $\|L\|$ of L as

$$\|L\| = \lambda_1,$$

and express the mapping norm $\| \wedge_m L\|$ of the linear mapping $\wedge_m L$ of m-vectors induced by L as

$$\| \wedge_m \| = \lambda_1 \cdot \lambda_2 \cdots \lambda_m.$$

Whenever $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$ we have

$$\lambda_1 \cdot \lambda_2 \cdots \lambda_m \leq \frac{1}{m^{\frac{m}{2}}} (\lambda_1^2 + \lambda_2^2 + \ldots + \lambda_m^2)^{\frac{m}{2}} \leq \lambda_1^m \leq (\lambda_1^2 + \lambda_2^2 + \ldots + \lambda_m^2)^{\frac{m}{2}}.$$

The first two inequalities are equalities if and only if $\lambda_1 = \lambda_2 = \ldots = \lambda_m$. The right hand inequalitiy is an equality if and only if $\lambda_2 = \lambda_3 = \ldots = \lambda_m = 0$.

If f is a mapping and $L = Df(a)$ is the differential of f at a, then $|Df(a)|^2$ is of value of *Dirichlet's integrand* of f at a, and

$$J_m f(a) = \| \wedge_m Df(a)\|$$

is the *m dimensional Jacobian* of f at a.

A.2 Modified Stereographic Projection.

Stereographic projection of $\mathbf{S}^m$ onto $\mathbf{R}^m$ from the south pole $\mathbf{q}$ maps $(z, y) \in \mathbf{S}^m \sim \{\mathbf{q}\}$ to $2y/(1+z) \in \mathbf{R}^m$ while the inverse mapping $\gamma_0: \mathbf{R}^m \to \mathbf{S}^m$ sends $y \in \mathbf{R}^m$ to

$$\gamma_0(y) = \left(\frac{4 - |y|^2}{4 + |y|^2}, \quad \frac{4y}{4 + |y|^2} \right) \in \mathbf{S}^m \sim \{\mathbf{q}\}.$$

γ_0 is an orientation preserving conformal diffeomorphism between $\mathbf{R}^m$ and $\mathbf{S}^m \sim \{\mathbf{q}\}$ as is readily checked.

For convenience we let $\theta: \mathbf{S}^m \to [0, \pi]$ denote angular distance in radians (equivalently, geodesic distance in $\mathbf{S}^m$) to $\mathbf{p}$. General level sets of θ are thus $m-1$ spheres of constant latitude while $\theta(\mathbf{p}) = 0$ and $\theta(\mathbf{q}) = \pi$. Also for $(z, y) \in \mathbf{S}^m$ we have $z = \cos\theta(z, y)$

and $|y| = \sin\theta(x,y)$. Latitude lines on $\mathbf{S}^m$ are level sets of the function ω which maps $(x,y) \in \mathbf{S}^m \sim \{\mathbf{p},\mathbf{q}\}$ to

$$\omega(x,y) = \frac{y}{|y|} \in \mathbf{S}^{m-1} \subset \mathbf{R}^m.$$

Certain mappings derived from γ_0 are important in our constructions. If $0 < \delta << 1/2$ is a given very small number we fix $0 < r = r(\delta) << R < \infty$ by requiring that R be the radius of the sphere in $\mathbf{R}^m$ which γ_0 maps to the latitude sphere $\theta = \pi - \delta$ near q in $\mathbf{S}^m$ and that rR/δ be the radius of the sphere in $\mathbf{R}^m$ which γ_0 maps to the latitude sphere $\theta = \delta$ near p.

We now modify γ_0 to obtain a mapping $\gamma = \gamma_\delta = \gamma_{\delta,1}$ which maps $\mathbf{R}^m$ onto all of $\mathbf{S}^m$ and which maps points y in $\mathbf{R}^m$ with norm less that r^2 to p, maps points y in $\mathbf{R}^m$ with norm greater that 2δ to q, maps points y in $\mathbf{R}^m$ with norm between r and δ to $\gamma_0(Ry/\delta)$ and suitable interpolates in the two remaining annular regions. More precisely, we set

$$\gamma(y) = \begin{cases} \mathbf{p} & \text{if } 0 \le |y| \le r^2 \\ \left(\cos\left(\delta\left(\frac{|y|-r^2}{r-r^2}\right)\right), \sin\left(\delta\left(\frac{|y|-r^2}{r-r^2}\right)\right)\frac{y}{|y|}\right) & \text{if } r^2 \le |y| \le r \\ \gamma_0\left(\frac{Ry}{\delta}\right) & \text{if } r \le |y| \le \delta \\ \left(\cos(\pi + |y| - 2\delta), \sin(\pi + |y| - 2\delta)\frac{y}{|y|}\right) & \text{if } \delta \le |y| \le 2\delta \\ \mathbf{q} & \text{if } 2\delta \le |y| < \infty. \end{cases}$$

In the region $0 \le |y| \le r$ we estimate that the Lipschitz constant of γ does not exceed $\delta/(r - r^2)$ which is less that $2\delta/r$ since $r < 1/2$. Hence

$$\int_{|y|\le r} |D\gamma|^m \, d\mathcal{L}^m \le m^{\frac{1}{2}} \left(\frac{2\delta}{r}\right)^m \alpha(m) r^m = 2^m m^{\frac{1}{2}} \alpha(m) \delta^m$$

which is small if δ is small.

Similarly, in the region $\delta \le |y| \le 2\delta$ we estimate that the local Lipschitz constants do not exceed 1. Hence

$$\int_{\delta\le|y|\le2\delta} |D\gamma|^m \, d\mathcal{L}^m < m^{\frac{1}{2}} \alpha(m)(2\delta)^m = 2^m m^{\frac{1}{2}} \alpha(m) \delta^m$$

which is small is δ is small.

Finally we note that, in the region, $r < |y| < \delta$ the mapping γ is conformal so that

$$\int_{r\le|y|\le\delta} J_m\gamma \, d\mathcal{L}^m = \int_{r\le|y|\le\delta} \|D\gamma\|^m \, d\mathcal{L}^m = \frac{1}{m^{\frac{m}{2}}} \int_{r\le|y|\le\delta} |D\gamma|^m \, d\mathcal{L}^m = \mathcal{H}^m(\mathbf{S}^m \cap \theta^{-1}[\delta, \pi-\delta]). \blacksquare$$

Our mapping $\gamma_{\delta,1}$ from $\mathbf{R}^m$ to $\mathbf{S}^m$ preserves orientations and covers once. It is useful to have mappings $\gamma_{\delta,\nu}$ with similar conformal properties but covering ν times. To do this we fix a ratio $\rho = (r(\delta)^2/\delta)$ and let $\tau(x,y) = (-x, -y_1, y_2, \ldots, y_m)$ for $(x,y) \in \mathbf{S}^m$; the map τ thus interchanges the north and south poles of $\mathbf{S}^m$ while preserving orientation. We then define

$$\gamma_{\delta,\nu}(y) = \begin{cases} \gamma_\delta(y) & \text{if } \rho\delta \le |y| < \infty \\ \tau^k \circ \gamma_\delta(y/\rho^k) & \text{if } k \in \{1, \ldots, \nu-2\} \text{ and } \rho^{k+1}\delta \le |y| \le \rho^k\delta \\ \tau^{\nu-1} \circ \gamma_\delta(y/\rho^{\nu-1}) & \text{if } 0 \le |y| \le \rho^{\nu-1}\delta. \end{cases}$$

A.3. Mappings and homotopies from $\mathcal{M}$ to $\mathbf{S}^m$ with contolled discontinuities.

A.3.1 Whenever $f: \mathcal{M} \to \mathbf{S}^m$ we denote by

$$C_f$$

the closure of the set of points of discontinuity of f. We then let

$$\mathcal{F}$$

be the collection of all functions $f : \mathcal{M} \to \mathbf{S}^m$ such that the closure of $C_f \sim B$ (recall A.1.1) has dimension not exceeding $n - 2$. In case m equals 1 we require that $C_f \subset B$ for functions f in $\mathcal{F}$. Also, if $\mathcal{M}$ is $\mathbf{R}^{m+n}$ we require that $f(x) = \mathbf{q}$ whenever $|x|$ is sufficiently large.

Similarly, whenever $h: [0,1] \times \mathcal{M} \to \mathbf{S}^m$ we denote by

$$C_h$$

the closure of the set of discontinuities of h. We then say that f and g in $\mathcal{F}$ are s-homotopic provided there is a function $h: [0,1] \times \mathcal{M} \to \mathbf{S}^m$ such that $h(0,\cdot) = f$ and $h(1,\cdot) = g$ and also

$$C_h \sim ((\{0\} \times C_f) \cup (\{1\} \times C_g) \cup ([0,1] \times B))$$

lies in $(0,1)\times\mathcal{M}$ and has dimension not exceeding $n-1$ (in case $\mathcal{M}$ is $\mathbf{R}^{m+n}$ we additionally require that $h(t,x)=q$ for all t when $|x|$ is sufficiently large); such a function h is called an s-homotopy between f and g. We then denote by

$$\Pi(\mathcal{F})$$

the s-homotopy equivalence classes of $\mathcal{F}$.

A.3.2 We denote by

$$\mathcal{F}_0$$

those functions f in $\mathcal{F}$ for which $f|(\mathcal{M}\sim C_f)$ is locally Lipschitz and then associate to each such f three energies $\mathcal{E}_1(f)$, $\mathcal{E}_2(f)$, and $\mathcal{E}_3(f)$ given by setting

$$\mathcal{E}_1(f)=\frac{1}{m^{m/2}(m+1)\alpha(m+1)}\int_{\mathcal{M}}|Df|^m\,d\mathcal{H}^{m+n},$$

$$\mathcal{E}_2(f)=\frac{1}{(m+1)\alpha(m+1)}\int_{\mathcal{M}}\|Df\|^m\,d\mathcal{H}^{m+n},$$

$$\mathcal{E}_3(f)=\frac{1}{(m+1)\alpha(m+1)}\int_{\mathcal{M}}J_mf\,d\mathcal{H}^{m+n}.$$

For some analyses (beyond the scope of this present paper) it is important to recognize that

$$J_mf(x)=|\langle\sigma^*(f(x)),\wedge^m Df(x)\rangle|.$$

We also call the reader's attention to the paper *Homotopy classes in Sobolev spaces and the existence of energy minimizing mappings* [W2] by B. White in which p energy minimization is studied in homotopy classes of mappings which are not necessarily continuous.

A.3.3 A basic fact is the following

PROPOSITION.

(1) Each s-homotopy class in $\Pi(\mathcal{F})$ contains a representative f which belongs to $\mathcal{F}_0$ and for which each of the energies $\mathcal{E}_1(f)$, $\mathcal{E}_2(f)$, and $\mathcal{E}_3(f)$ is finite.

(2) Suppose f and g belong to $\mathcal{F}_0$ and are representatives of the same s-homotopy class in $\Pi(\mathcal{F})$. Suppose also that $\mathcal{E}_1(f)$ and $\mathcal{E}_1(g)$ are both finite. Then there is an s-homotopy h between f and g such that $h|([0,1]\times\mathcal{M}\sim C_h)$ is locally Lipschitz and

$$\int_{[0,1]\times\mathcal{M}}|Dh|^m\,d\mathcal{H}^{m+n+1}<\infty.$$

A.4 Currents. A general k (dimensional) current T is a continuous linear functional on an appropriate space of smooth differential k forms in $\mathbf{R}^N$. The boundary of a k current T is the $k-1$ current ∂T which maps a smooth differential $k-1$ form ω to the number $\partial T(\omega)=T(d\omega)$—Stokes's theorem becomes a definition. In this paper we are concerned with currents of the form $T=\mathbf{t}(\Sigma,\theta,\varsigma)$. In writing such an expression we mean that $\mathbf{set}(T)=\Sigma$ is a (bounded) $\mathcal{H}^k$ measurable and $(\mathcal{H}^k,k)$ rectifiable subset of $\mathcal{M}$, and that the *density function* $\theta:\Sigma\to\mathbf{R}^+$ is $\mathcal{H}^k\llcorner\Sigma$ summable, and that the *orientation* ς is an $\mathcal{H}^k\llcorner\Sigma$ measurable function whose simple unit k vector values are compatible with the tangent plane structure of Σ. Such a k current T maps a differential k form φ to the number

$$T(\varphi)=\int_{x\in\Sigma}\langle\varsigma(x),\varphi(x)\rangle\,\theta(x)\,d\mathcal{H}^kx.$$

Associated with $\mathcal{M}$ itself is the $m+n$ current

$$[\![\mathcal{M}]\!]=\mathbf{t}(\mathcal{M},1,\xi);$$

if $\mathcal{M}=\mathbf{R}^{m+n}$ a standard notation is

$$\mathbf{E}^{m+n}=\mathbf{t}(\mathbf{R}^{m+n},1,\xi)$$

with $\xi(x)=\mathbf{e}_1\wedge\ldots\wedge\mathbf{e}_{m+n}$ for each x.

The area of a current $T=\mathbf{t}(\Sigma,\theta,\varsigma)$ weighted with its density gives its *mass*,

$$\mathbf{M}(T)=\int_\Sigma\theta\,d\mathcal{H}^k=\sup\{T(\varphi):\|\varphi\|\le 1\}.$$

The theorems of this paper relate to minimization of this mass rather than, say, the k areas of the underlying set Σ (which is called the *size* of T and is denoted $\mathbf{S}(T)$). The measure $\|T\|$ associated with mass is thus $\mathcal{H}^k\llcorner\Sigma\wedge\theta$ so that $\mathbf{M}(T)=\|T\|(\mathcal{M})=\|T\|(\mathbf{R}^N)$.

A general fact about such a current $T=\mathbf{t}(\Sigma,\theta,\varsigma)$ is that its general current boundary ignores closed sets of zero k-1 measure, e.g. if $U\subset\mathbf{R}^N$ is open and the support of ∂T inside U has zero $\mathcal{H}^{k-1}$ measure, then $\partial T(\omega)=0$ for each ω supported in U [F1 4.1.20].

Suppose that $T=\mathbf{t}(\Sigma,\theta,\varsigma)$ is an n current such that the support of ∂T lies in B. Because of our special assumptions about B in A.1.1 we can use [F1 4.1.31] together with our preceding remark to infer for each $k=1,\ldots J$ the existence of nonnegative real numbers r_k and continuous orientation functions ς_k on Δ_k such that

$$\partial T=\sum_{k=1}^{J}\mathbf{t}(\Delta_k,r_k,\varsigma_k).$$

For general (possibly empty) subsets A and C of $\mathcal{M}$ with $C \subset A$ we denote by $\mathbf{R}_k(A,C)$ the vector space of those k currents $T = \mathbf{t}(\Sigma,\theta,\varsigma)$ with the closure of Σ contained in A such that $\partial T = \mathbf{t}(\Sigma',\theta',\varsigma')$ for some $\Sigma',\theta',\varsigma'$ with the closure of Σ' contained in C. We further let $\mathbf{I}_k(A,C)$ denote the subgroup of those currents $T = \mathbf{t}(\Sigma,\theta,\varsigma)$ in $\mathbf{R}_k(A,C)$ such that θ assumes only positive integer values. It follows from [**F1** 4.2.16(2)] that $\partial T \in \mathbf{I}_{k-1}(C,\emptyset)$ whenever $T \in \mathbf{I}_k(A,C)$.

When convenient we will denote by $\text{spt}T$ the support of a current T.

A.5 The coarea formula and slices of currents. A key ingredient of the present paper is slicing the current $[\![\mathcal{M}]\!]$ by mappings $f:\mathcal{M} \to \mathbf{S}^m$ belonging to $\mathcal{F}_0$ and use of the *coarea formula* to estimate the masses of these slices in terms of the energy $\mathcal{E}_3(f)$. As a consequence of [**F1** 3.2.22, 4.3.8, 4.3.11] we infer that for $\mathcal{H}^m$ almost every $w \in \mathbf{S}^m$ the slice

$$\langle [\![\mathcal{M}]\!], f, w\rangle = \mathbf{t}(f^{-1}\{w\}, 1, \varsigma)$$

is well defined as an n dimensional current. Here, for $\mathcal{H}^n$ almost every $x \in f^{-1}\{w\}$, if $\eta(x)$ is that simple unit m vector associated with the m plane $\ker Df(x)^{\perp}$ in $\text{Tan}(\mathcal{M},x)$ for which

$$\langle \eta(x), \wedge_m Df(x)\rangle \bullet \sigma(w) > 0$$

then we specify $\varsigma(x)$ to be that simple unit n vector associated with $\ker Df(x)$ in $\text{Tan}(\mathcal{M},x)$ for which $\xi(x) = \eta(x) \wedge \varsigma(x)$; we have used the symbol $\bullet$ to denote the inner product in $\wedge_m \mathbf{R}^{m+1}$.

We further infer from the coarea formula [**F1** 3.2.22] that

$$(m+1)\alpha(m+1)\mathcal{E}_3(f) = \int_{w\in\mathbf{S}^m} \mathbf{M}(\langle [\![\mathcal{M}]\!], f, w\rangle)\, d\mathcal{H}^m w.$$

Since $\partial[\![\mathcal{M}]\!] = 0$ we readily infer from [**F1** 4.3.1] together with A.3.2 and A.4 above that for $\mathcal{H}^m$ almost every $w \in \mathbf{S}^m$, $\partial\langle [\![\mathcal{M}]\!], f, w\rangle$ belongs to $\mathbf{I}_{m-1}(B,\emptyset)$.

A.6 Kronecker indices of integral currents. Whenever $S \in \mathbf{I}_m(\mathcal{M},\mathcal{M})$ and $T \in \mathbf{I}_n(\mathcal{M},\mathcal{M})$ with

$$\emptyset = \text{spt}\partial S \cap \text{spt}T = \text{spt}S \cap \text{spt}\partial T,$$

there is naturally defined the *Kronecker index* of S and T in $\mathcal{M}$, denoted

$$\mathbf{k}(S,T) = \mathbf{k}(S,T;\mathcal{M}) \in \mathbf{Z}.$$

which is a direct extension of the definitions in [**F1** 4.3.20]. For 'sufficiently regular' such currents

$$S = \mathbf{t}(\Sigma_1,\theta_1,\varsigma_1) \quad \text{and} \quad T = \mathbf{t}(\Sigma_2,\theta_2,\varsigma_2)$$

in 'general position', we can write

$$\mathbf{k}(S,T) = \sum_{x\in\Sigma_1\cap\Sigma_2} \theta_1(x)\cdot\theta_2(x)\cdot\text{sign}(\varsigma_1(x)\wedge\varsigma_2(x)\bullet\xi(x)).$$

Among the important facts about the Kronecker index is its ability to characterize real homology classes. We have the following.

PROPOSITION. *Suppose $T_1,T_2 \in \mathbf{I}_n(\mathcal{M},B)$ with $\partial T_1 = \partial T_2$ and*

$$\mathbf{k}(S,T_1) = \mathbf{k}(S,T_2)$$

for each $S \in \mathbf{I}_m(\mathcal{M},\emptyset)$ for which both Kronecker indices are defined. Then there is $Q \in \mathbf{R}_{n+1}(\mathcal{M},\mathcal{M})$ such that $\partial Q = T_1 - T_2$.

Proof. In view of [**F1** 4.4.1] it is sufficient to verify the assertion in the context of Lipschitz singular chains of algebraic topology. Moreover it is sufficient to check than an n cycle T in $\mathcal{M}$ is a boundary in case its general position intersections with m cycles S in $\mathcal{M}$ all have Kronecker index zero. This is well known.

A.7 Degrees of mappings of currents. Suppose $f \in \mathcal{F}_0$ and

$$S = \mathbf{t}(\Sigma,\theta,\varsigma) \in \mathbf{I}_m(\mathcal{M} \sim C_f,\emptyset).$$

Then the m current $f_\sharp S$ in $\mathbf{S}^m$ is naturally defined in accordance with [**F1** 4.1.14, 4.1.15] with $\partial f_\sharp S = 0$ since $\partial S = 0$. We then infer from [**F1** 4.1.31] the existence of an integer $\mathbf{d}(f,S)$ such that

$$f_\sharp S = \mathbf{t}(\mathbf{S}^m, \mathbf{d}(f,S), \sigma).$$

We call $\mathbf{d}(f,S)$ the *degree* of f on S. If f and S are 'sufficiently regular' then, for $\mathcal{H}^m$ almost every $w \in \mathbf{S}^m$,

$$\mathbf{d}(f,S) = \sum_{x\in\Sigma\cap f^{-1}\{w\}} \theta(x)\,\text{sign}\left(\langle\varsigma(x),\wedge_m Df(x)\rangle\bullet\sigma(w)\right).$$

Basic properties of degrees are the following.

PROPOSITION.

(1) The degree $d(f,S)$ *depends only on the real homology class of S in* $\mathcal{M} \sim B$. *More precisely, if* $f \in \mathcal{F}_0$, *and* $S_1, S_2 \in \mathbf{I}_m(\mathcal{M} \sim C_f, \emptyset)$, *and* $Q \in \mathbf{R}_{m+1}(\mathcal{M} \sim B, \mathcal{M} \sim B)$ *with* $\partial Q = S_1 - S_2$, *then* $d(f,S_1) = d(f,S_2)$.

(2) The degree $d(f,S)$ *depends only on the s-homotopy class of f. More precisely, if* $f, g \in \mathcal{F}_0$ *are s-homotopic and* $S \in \mathbf{I}_m(\mathcal{M} \sim (C_f \cup C_g \cup B), \emptyset)$, *then* $d(f,S) = d(g,S)$.

A.8 An example showing relations between integral current slices and boundaries, Kronecker indices, and mapping degrees. Suppose, as illustrated in Figure 2, the following.

(a) $\mathcal{M} = \mathbf{R}^{m+n}$ with its usual orthonormal basis, and

$$U = \mathbf{U}^{m+1}(0,1) \times \mathbf{U}^{n-1}(0,1),$$

is an open set, and

$$\Delta = \{0\} \times \mathbf{U}^{n-1}(0,1)$$

is an $n-1$ disk with orientation function

$$\beta : \Delta \to \{\mathbf{e}_{m+2} \wedge \ldots \wedge \mathbf{e}_{m+n}\}.$$

(b) K and $z_1, \ldots, z_K$ are positive integers and $\epsilon_1, \ldots, \epsilon_K \in \{-1, +1\}$.

(c) For each k the vectors

$$p(k), \eta_1(k), \ldots, \eta_m(k) \in \mathbf{S}^m \times \{0\} \subset \mathbf{R}^{m+1} \times \mathbf{R}^{n-1}$$

are an orthonormal family such that

$$\eta_1(k) \wedge \ldots \wedge \eta_m(k) \wedge p(k) = \mathbf{e}_1 \wedge \ldots \wedge \mathbf{e}_{m+1}$$

and also $p(1), \ldots, p(K)$ are distinct.

(d) For each k we let Π_k denote the n plane spanned by $p(k)$ and $\{0\} \times \mathbf{R}^{n-1}$ and define the n half disk

$$\Delta_k = \Pi_k \cap U \cap \{x : x \bullet p(k) \le 0\}$$

with orientation function

$$\varsigma : \Delta_k \to \{\epsilon_k\, p(k) \wedge \mathbf{e}_{m+2} \wedge \ldots \wedge \mathbf{e}_{m+n}\}.$$

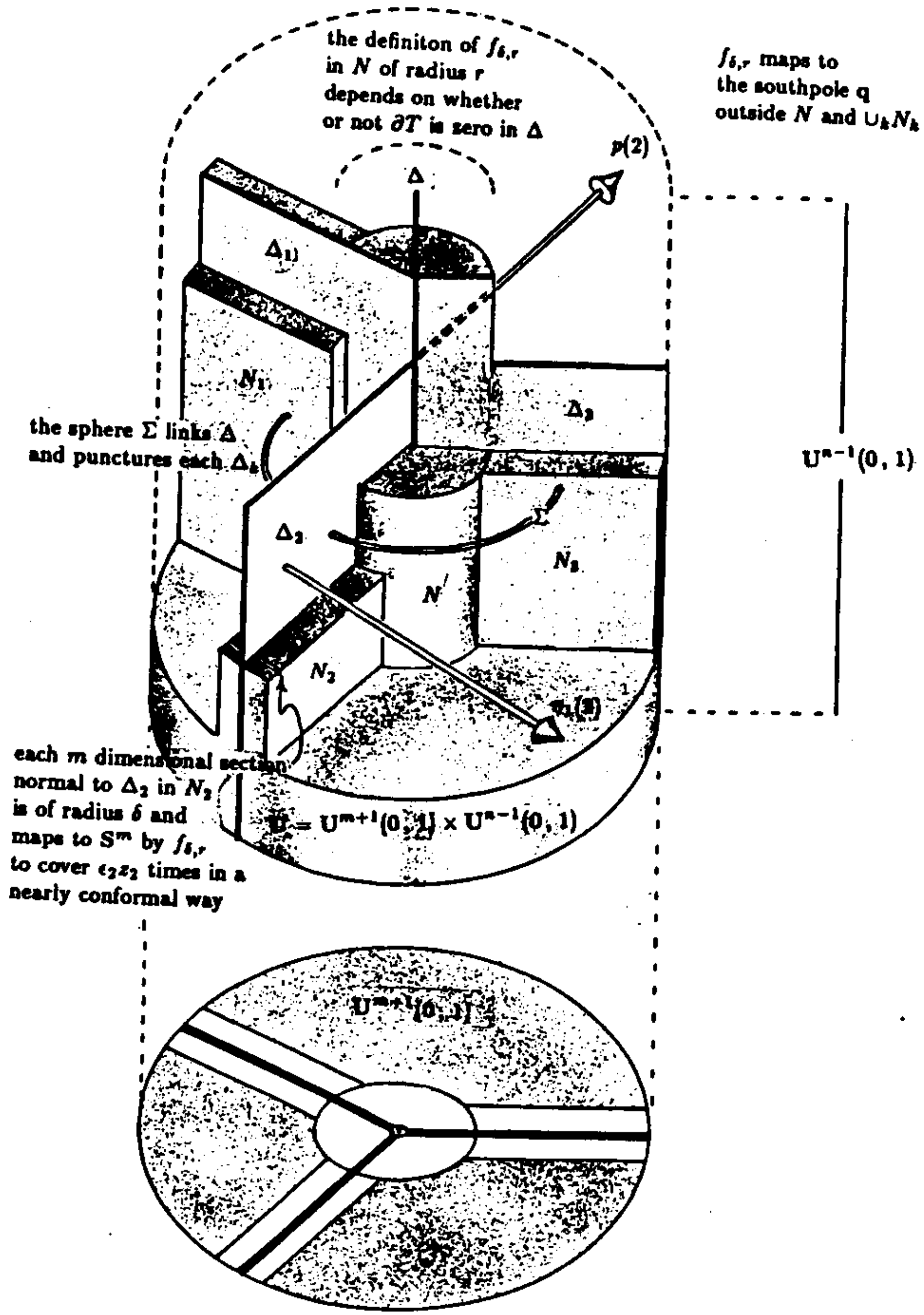

Figure 2. Relations between integral current slices and boundaries, Kronecker indices, and mapping degrees are illustrated by example in Appendix A.8.

(e) $0 < \delta << r \le s << 1$ are very small numbers and

$$N = U \cap \{x: \text{dist}(x, \Delta) < r\} \qquad \text{and} \qquad N_k = (U \sim N) \cap \{x: \text{dist}(x, \Delta_k < 2\delta\}$$

for each k; we assume that δ is small enough so that the sets $N_1, \dots, N_k$ are positive distances apart.

(e) We denote by Σ the small m sphere

$$\Sigma = \partial \mathbf{B}^{m+1}(0, s) \times \{0\}$$

with the standard continuous orientation function $\tau: \Sigma \to \wedge_m \mathbf{R}^{m+n}$ determined by requiring

$$x \wedge \tau(x) = s \cdot \mathbf{e}_1 \wedge \dots \wedge \mathbf{e}_{m+1}$$

for each x in Σ; it follows that

$$\tau(-s \cdot p(k)) = (-1)^{m+1} \eta_1(k) \wedge \dots \wedge \eta_m(k)$$

for each k. Here $\cdot$ denotes scalar multiplication of a vector. The m sphere Σ 'links' the $n-1$ disk Δ in U while 'puncturing' each Δ_k at the point $-s \cdot p(k)$.

We then set

$$T = \sum_{k=1}^{K} \mathbf{t}(\Delta_k, z_k, \epsilon_k \cdot \varsigma_k) \qquad \text{and} \qquad S = \mathbf{t}(\Sigma, 1, \tau)$$

and estimate

(1) The boundary of T inside U is given by

$$\partial T \llcorner U = \sum_{k=1}^{K} \mathbf{t}(\Delta, z_k, \epsilon_k \cdot \beta)$$

[F1 4.1.8] so that $\partial T \llcorner U = 0$ if and only if $\sum_{k=1}^{K} \epsilon_k z_k = 0$.

(2) The Kronecker index of S and T is given by

$$\begin{aligned} \mathbf{k}(S,T) &= \sum_{k=1}^{K} z_k \cdot \tau(-s \cdot p(k)) \wedge \varsigma(-s \cdot p(k)) \bullet \mathbf{e}_1 \wedge \dots \wedge \mathbf{e}_{m+n} \\ &= \sum_{k=1}^{K} z_k (-1)^{m+1} \eta_1(k) \wedge \dots \eta_m(k) \wedge \epsilon_k \cdot p(k) \wedge \mathbf{e}_{m+1} \wedge \dots \wedge \mathbf{e}_{m+n} \bullet \mathbf{e}_1 \wedge \dots \wedge \mathbf{e}_{m+n} \\ &= (-1)^{m+1} \sum_{k=1}^{K} \epsilon_k z_k \end{aligned}$$

so that $\mathbf{k}(S,T) = 0$ if and only if $\partial T \llcorner U = 0$.

We now assume $r = s$ and will construct a mapping $g: U \sim N \to \mathbf{S}^m$. We first set $g(x) = \mathbf{q}$ (the southpole) if x lies outside both N and all the N_k's. Each point in each N_k can be written uniquely in the form

$$x + y_1 \eta_1(k) + \dots + y_m \eta_m(k)$$

where x is the unique closest point in Δ_k and $y \in \mathbf{B}^{m+1}(0, 2\delta)$; for each such point we set

$$g(x + y_1 \eta_1(k) + \dots y_m \eta_m(k)) = \gamma_{\delta, z_k}(\epsilon_k \cdot y_1, y_2, \dots, y_m).$$

Since $r \le s < 1$ our function g is defined on Σ and there is a well defined mapping degree $\mathbf{d}(g, S)$ (with the obvious meaning). Since each γ_{δ, z_k} is orientation preserving (and δ is very small) the orientation of g on Σ near $p(k)$ is determined by ϵ_k and by the inner product

$$\eta_1(k) \wedge \dots \wedge \eta_m(k) \bullet \tau(-s \cdot p(k)),$$

and we compute

(3) The degree of g on S is given by

$$\begin{aligned} \mathbf{d}(g, S) &= \sum_{k=1}^{K} z_k \epsilon_k \cdot \eta_1(k) \wedge \dots \wedge \eta_m(k) \bullet \tau(-s \cdot p(k)) \\ &= (-1)^{m+1} \sum_{k=1}^{K} \epsilon_k z_k \end{aligned}$$

so that $\mathbf{d}(g, S) = 0$ if and only if $\partial T \llcorner U = 0$.

The extension of g to a mapping $f = f_{\delta, r}$ on all of U depends on which of two cases occurs.

Case 1. If $\mathbf{d}(g, S) = 0$ we infer from Hurewicz's theorem the existence of a Lipschitz mapping $h: \mathbf{B}^{m+1}(0, r) \to \mathbf{S}^m$ such that

$$h(w) = \begin{cases} g(w, 0) & \text{if } |w| = r \\ \mathbf{q} & \text{if } |w| \le r/2. \end{cases}$$

We then define our mapping $f: U \to \mathbf{S}^m$ by setting

$$f(x) = \begin{cases} g(x) & \text{if } x \notin N \\ h(x_1, \ldots, x_{m+1}) & \text{if } x \in N. \end{cases}$$

Case 2. If $d(g, S) \neq 0$ we define a discontinuous mapping $h: \mathbf{B}^{m+1} \to \mathbf{S}^m$ by setting

$$h(w) = g\left(\frac{rw}{|w|}, 0\right)$$

for each w and, as above, define $f: U \to \mathbf{S}^m$ by setting

$$f(x) = \begin{cases} g(x) & \text{if } x \notin N \\ h(x_1, \ldots, x_{m+1}) & \text{if } x \in N. \end{cases}$$

With the obvious interpretation of $\mathcal{E}_1, \mathcal{E}_2$, and $\mathcal{E}_3$ for function on U, each of these energies of mappings $f_{\delta,r}$ nearly equals the mass of T when δ and r are small (and reasonable choices are made for h in Case 1). More precisely, we have.

$$\lim_{\delta,r\downarrow 0} \mathcal{E}_1(f_{\delta,r}) = \lim_{\delta,r\downarrow 0} \mathcal{E}_2(f_{\delta,r}) = \lim_{\delta,r\downarrow 0} \mathcal{E}_3(f_{\delta,r}) = \mathbf{M}(T) = \sum_{k=1}^{K} z_k \mathcal{H}^n(\Delta_k).$$

It is also straight forward to check that for $\mathcal{H}^m$ almost every $w \in \mathbf{S}^m$ the slice

$$T_w = \langle \mathbf{E}^{m+n} \llcorner U, f_{\delta,r}, w \rangle$$

exists with

$$\partial T_w \llcorner U = \partial T \llcorner U,$$

and also if a sequence of δ's and r's converging to 0 is fixed then, for $\mathcal{H}^m$ almost every w in $\mathbf{S}^m$,

$$\lim_{\delta,r\downarrow 0} \langle \mathbf{E}^{m+n} \llcorner U, f_{\delta,r}, w \rangle = T.$$

REFERENCES

[A1] F. Almgren, *Deformations and multiple-valued functions*, Geometric Measure Theory and the Calculus of Variations, Proc. Symposia in Pure Math. 44 (1986), 29-130.

[A2] ________, *Q valued functions minimizing Dirichlet's integral and the regularity of area minimizing rectifiable currents up to codimension two*, preprint.

[BCL] H. Brezis, J-M Coron, E. Lieb, *Harmonic maps with defects*, Comm. Math. Physics, 1987; see also C. R. Acad. Sc. Paris 303 (1986), 207-210.

[F1] H. Federer, *Geometric Measure Theory*, Springer-Verlag, 1969, XIV + 676 pp.

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

[HKL] R. Hardt, D. Kinderlehrer, and M. Luskin, *Remarks about the mathematical theory of liquid crystals*, Institute for Mathematics and its Applications, preprint, 1986.

[MF] F. Morgan, *Area-minimizing currents bounded by higher multiples of curves*, Rend. Circ. Matem. Palermo, (II) 33 (1984), 37–46.

[PH] H. Parks, *Explicit determination of area minimizing hypersurfaces, II*, Mem. Amer. Math. Soc. 60, March 1986, iv + 90 pp.

[SU] R. Schoen and K. Uhlenbeck, *A regularity theory for harmonic maps*, J. Diff. Geom. 60 (1982), 307–335.

[W1] B. White, *The least area bounded by multiples of a curve*, Proc. Amer. Math. Soc. 90 (1984), 230-232.

[W2] ________, *Homotopy classes in Sobolev spaces and the existence of energy minimizing maps*, preprint.

[YL] L. C. Young, *Some extremal questions for simplicial complexes V. The relative area of a Klein bottle*, Rend. Circ. Matem. Palermo, (II), 12 (1963), 257–274.