Rearrangements of vector valued functions, with application to atmospheric and oceanic flows

R. J. Douglas
Department of Mathematical and Computing Sciences,
University of Surrey,
Guildford, Surrey GU2 5XH, U.K.

Abstract

This paper establishes the equivalence of four definitions of two vector valued functions being rearrangements, and gives a characterisation of the set of rearrangements of a prescribed function. The theory of monotone rearrangement of a vector valued function is used to show the existence and uniqueness of the minimiser of an energy functional arising from a model for atmospheric and oceanic flow. At each fixed time solutions are shown to be equal to the gradient of a convex function, verifying the conjecture of Cullen, Norbury and Purser.

Key words Rearrangement of functions, semigeostrophic, variational problems, generalised solution.
AMS (MOS) subject classification 35D05, 46E30, 76C15.
Newton Institute preprint number NI96009.

1 Introduction

This paper studies properties of rearrangements of vector valued functions, and gives an application to atmospheric and oceanic flow. We say that two vector valued functions $f, g \in L^p(\Omega \subset \mathbf{R}^n, \lambda, \mathbf{R}^d)$, where $1 \leq p < \infty$ and Ω is bounded, are *rearrangements* if

$$\lambda\left(f^{-1}(B)\right) = \lambda\left(g^{-1}(B)\right)$$

for each Borel subset B of $\mathbf{R}^d$. (We restrict our definition of rearrangement to functions defined on measure spaces (Ω, λ) with certain properties, see section 2.1.) This is equivalent to the definition of rearrangement for scalar valued functions when $d = 1$. Rearrangement can be viewed as an equivalence relation on the space of L^p functions, therefore we can define the set of rearrangements (or equivalence class) of a given vector valued function. For a prescribed f_0 we write $R(f_0)$ to denote the set of rearrangements of f_0.

Different definitions have been given for two vector valued functions being rearrangements. Brenier [**2**] defined vector valued functions f and g (belonging to $L^p(\Omega \subset \mathbf{R}^n, \lambda, \mathbf{R}^d)$) to be rearrangements if $\int_\Omega F(f) = \int_\Omega F(g)$ for each F in a subclass of continuous functions from $\mathbf{R}^d$ to $\mathbf{R}$. In contrast, Cullen, Norbury and Purser [**5**] made a direct extension of the definition of scalar valued rearrangement, requiring that $\lambda\{x : f(x) \geq c\} = \lambda\{x : g(x) \geq c\}$ for each $c \in \mathbf{R}^d$, where the inequalities are calculated component by component. Section 2 unifies these concepts, establishing that both are equivalent to the definition in the opening paragraph. We establish a fourth equivalent property which yields a characterisation for the set of rearrangements of a prescribed vector valued function. This is a vector valued extension of the real valued characterisation of Eydeland, Spruck and Turkington [**6**]. These results are stated in Theorem 1. The proofs require a technical lemma from the theory of analytic sets.

Section 3 studies a variational problem arising from a model for atmospheric and oceanic flow. The equations are the three dimensional Boussinesq equations of semigeostrophic flow, a standard model for slowly varying flows constrained by rotation and stratification. (They are recalled in section 3.2.) Cullen, Norbury and Purser [**5**] interpreted solutions as a sequence of minimum energy states: at each time t, the particles arrange themselves so that geostrophic energy is minimised. The state of the fluid is known on particles, therefore we minimise geostrophic energy over the set of rearrangements of a possible state of the fluid at time t, a vector valued function. Cullen, Norbury and Purser conjectured the existence of a unique minimiser, equal to the gradient of a convex function, which is the actual state of the fluid. We make the physically reasonable assumption that the fluid configuration belongs to L^p, so that we may use the theory of monotone rearrangement of vector valued functions, which was developed by Brenier [**2**]. If the fluid configuration satisfies a non–degeneracy condition, (see section 3.3,) the Cullen–Norbury–Purser conjecture follows easily by the results of Brenier [**2**]. However the non–degeneracy condition is severe, as it does not allow the function to have level sets of positive measure. Our main result is a proof of the Cullen–Norbury–Purser conjecture in Theorem 2: we make no restriction on the fluid configuration. We approximate functions which fail the non–degeneracy condition by a sequence of functions which satisfy it, and take appropriate limits. Uniqueness of the energy minimiser is recovered by properties of the monotone rearrangement.

The theory of rearrangements of vector valued functions is a new research area: recent advances have been made by Brenier [**2**]. In comparison the theory of rearrangements of scalar valued functions is well developed: for example see Burton [**3**] and Alvino, Trombetti and Lions [**1**]. Some results for scalar valued rearrangements do not have vector valued equivalents. For example the monotone rearrangement of a vector valued function does not satisfy some of the inequalities which hold for the increasing rearrangement of a real valued function. (See Brenier [**2**] for details.)

We consider the relationship between vector valued functions which are rearrangements, and vector valued functions for which corresponding components are rearrangements in the scalar valued sense. Let f, g be as in the opening paragraph. Define, for $i = 1, ..., d$, $\Pi_i : \mathbf{R}^d \to \mathbf{R}$ to be the projection of the ith component of an element of $\mathbf{R}^d$. Write $f = (f_1, ..., f_d)$ and $g = (g_1, ..., g_d)$ where $f_i = \Pi_i \circ f$, $g_i = \Pi_i \circ g$ for $i = 1, ..., d$. The definition of vector valued rearrangement yields that if $f \in R(g)$ we have $f_i \in R(g_i)$ for each $i = 1, ..., d$ in the scalar valued sense. However the converse is false in general. Let $f : [0,1]^2 \to \mathbf{R}^2$ be defined by

$$f(x) = \begin{cases} (1,1) & \text{if } x \in [1/2,1] \times [1/2,1], \\ (0,0) & \text{if } x \notin [1/2,1] \times [1/2,1]. \end{cases}$$

Then

$$f_1(x) = f_2(x) = \begin{cases} 1 & \text{if } x \in [1/2,1] \times [1/2,1], \\ 0 & \text{if } x \notin [1/2,1] \times [1/2,1]. \end{cases}$$

Define

$$g(x) = \begin{cases} (1,0) & \text{if } x \in [1/2,1] \times [1/2,1], \\ (0,1) & \text{if } x \in [0,1/2] \times [1/2,1], \\ (0,0) & \text{otherwise.} \end{cases}$$

Then $g_1 = f_1$ and

$$g_2(x) = \begin{cases} 1 & \text{if } x \in [0,1/2] \times [1/2,1], \\ 0 & \text{if } x \notin [0,1/2] \times [1/2,1]. \end{cases}$$

It is easily seen that $f_i \in R(g_i)$ for $i = 1, 2$, but $f \notin R(g)$. Consequently in general we cannot apply scalar valued rearrangement theorems to components of vector valued functions and hope to obtain results pertaining to vector valued rearrangements.

2 Equivalent definitions of rearrangement of vector valued functions

2.1 Introduction

In this section we establish four equivalent definitions of rearrangement for vector valued functions, and give a characterisation of the set of rearrangements of a prescribed vector valued function. We define rearrangement for vector valued functions on finite measure spaces (U, μ) which are isomorphic to $(0, \mu(U))$ endowed with Lebesgue measure λ. By isomorphic, we mean there exists a measure preserving transformation $T : U \to (0, \mu(U))$. We recall the definition of measure preserving transformation in the next section. The restriction to finite measure spaces (U, μ) isomorphic to $(0, \mu(U))$ with Lebesgue measure is not severe: Royden [**12**] yields that any separable complete metric space U, equipped with a Borel measure μ such that $\mu(U) < \infty$ and $\mu(\{x\}) = 0$ for each $x \in U$, is isomorphic to $((0, \mu(U)), \lambda)$.
Definition Let (U, μ) be a measure space which is isomorphic to $((0, \mu(U)), \lambda)$. Let $f, g \in L^p(U, \mu, \mathbf{R}^d)$, for $1 \le p < \infty$. Then f is a *rearrangement* of g if

$$\mu\left(f^{-1}(B)\right) = \mu\left(g^{-1}(B)\right)$$

for every Borel subset B of $\mathbf{R}^d$.
We prove the following theorem.

Theorem 1 *Let (U, μ) be as above. Let $f, g \in L^p(U, \mu, \mathbf{R}^d)$, for $1 \le p < \infty$. Then the following are equivalent.*
(i) f is a rearrangement of g.
(ii) For each $F \in C(\mathbf{R}^d)$ such that $|F(\xi)| \le K(1 + |\xi|_2^p)$ (where $|.|_2$ denotes Euclidean distance on $\mathbf{R}^d$, and K is a constant), the following equation is satisfied:

$$\int_U F(f(x))d\mu(x) = \int_U F(g(x))d\mu(x).$$

(iii) $\mu(f^{-1}(C)) = \mu(g^{-1}(C))$ for each set $C \in \{\prod_{i=1}^d [\alpha_i, \infty) : \alpha_i \in \mathbf{R}$ for each $i = 1, ..., d\} \bigcup \{\emptyset, \mathbf{R}^d\}$.
(iv) For each $\sigma \in \mathbf{R}^d$, $\alpha > 0$,

$$\int_U (|g - \sigma|_\infty - \alpha)_+ d\mu = \int_U (|f - \sigma|_\infty - \alpha)_+ d\mu$$

where $|.|_\infty$ denotes the infinity norm on $\mathbf{R}^d$, and the + subscript denotes the positive part of the function.

Brenier [**2**] used property (ii) to define rearrangement of vector valued functions, whilst Cullen, Norbury and Purser used property (iii). This theorem shows that their definitions are equivalent. Property (iv) is a vector valued extension of the characterisation of the set of rearrangements of a given real valued function by Eydeland, Spruck and Turkington [**6**]: for non–negative $f_0 \in L^p(U, \mu)$,

$$R(f_0) = \{w \text{ measurable}, w \ge 0 : \int_U (w - \alpha)_+ = \int_U (f_0 - \alpha)_+, \forall \alpha > 0\}.$$

It follows from (iv) that for $f_0 \in L^p(U, \mu, \mathbf{R}^d)$, where $1 \le p < \infty$ and (U, μ) is as in Theorem 1

$$\begin{aligned} R(f_0) &= \{ w\ \mu\text{–measurable } : \int_U (|w - \sigma|_\infty - \alpha)_+ \, d\mu \\ &= \int_U (|f_0 - \sigma|_\infty - \alpha)_+ \, d\mu, \forall \sigma \in \mathbf{R}^d, \forall \alpha > 0\}. \end{aligned}$$

It may be shown that $R(f_0)$ is closed, and using the characterisation above, that for $w \in R(f_0)$, $||w||_p = ||f_0||_p$, where

$$||w||_p = \left\{ \int_U |w|_\infty^p d\mu \right\}^{\frac{1}{p}}.$$

We omit the proofs, which are elementary.

2.2 Measure preserving mappings and transformations

We recall the concept of a measure preserving mapping.

Definition A *measure preserving mapping* from a finite measure space (U, μ) to a measure space (V, ν) with $\mu(U) = \nu(V)$ is a mapping $s : U \to V$ such that for each ν–measurable set $A \subset V$, $\mu(s^{-1}(A)) = \nu(A)$.

Halmos [**7**, Theorem 2, page 163] yields that this is equivalent to requiring that for every ν–integrable function f, $f \circ s$ is μ–integrable and

$$\int_U f \circ s d\mu = \int_V f d\nu.$$

Measure preserving mappings are surjective (up to sets of measure zero), but not necessarily injective. If a measure preserving mapping s is injective, and s maps μ–measurable sets to ν–measurable sets, then s^{-1} exists and is a measure preserving mapping. Such an s is called a *measure preserving transformation.*

2.3 Analytic set theory

We proceed with the proof of Theorem 1 in stages. We require a result from the theory of analytic sets. As a preliminary, we establish some notation. Let H be a family of subsets of a given set X. Define

$$\begin{aligned} H_{\sigma\delta} &= \{ \text{ countable disjoint unions of elements of } H\} \\ H_C &= \{ \text{ complements (relative to } X) \text{ of elements of } H\} \end{aligned}$$

$B_{cd}(H)$ will denote the smallest family H^*, with $H \subset H^*$, such that $H^*_C = H^*_{\sigma\delta} = H^*$.

Kechris [**9**, page 65, Theorem 10.1 (iii)] yields the following result.

Theorem *Let H be a family of subsets of X such that (i) $X \in H$ (ii) $H_1 \bigcap H_2 \in H$ whenever $H_1, H_2 \in H$. Then $B_{cd}(H)$ is a σ–algebra.*

Lemma 1 *Let f, g be as in Theorem 1. Define*

$$\begin{aligned} \mathcal{M} &= \{A \subset \mathbf{R}^d : \mu(f^{-1}(A)) = \mu(g^{-1}(A))\} \\ H &= \{\prod_{i=1}^{d} [a_i, b_i] : a_i, b_i \in \mathbf{R}, a_i \le b_i, \text{ for } i = 1, ..., d\} \bigcup \{\emptyset, \mathbf{R}^d\} \end{aligned}$$

Suppose $H \subset \mathcal{M}$. Then $\mathcal{M}$ contains the Borel sets of $\mathbf{R}^d$.

Proof H is closed under finite intersection, therefore the above theorem yields that $B_{cd}(H)$ is a σ–algebra. H generates the Borel sets, therefore it follows that the Borel sets are contained in $B_{cd}(H)$. $\mathcal{M}$ is closed under countable disjoint union and complementation (relative to $\mathbf{R}^d$). Given that $H \subset \mathcal{M}$ we have $B_{cd}(H) \subset B_{cd}(\mathcal{M}) = \mathcal{M}$, so $\mathcal{M}$ contains the Borel sets. This completes the proof.

2.4 Proof of Theorem 1

We begin by showing that (i) implies (ii). Let $F \in C(\mathbf{R}^d)$ satisfy $|F(\xi)| \leq K\{1 + |\xi|_2^p\}$ for each $\xi \in \mathbf{R}^d$, where K is some constant. We assume that F is non–negative. (If not we work with the positive and negative parts of F.) F is measurable, therefore the fundamental approximation lemma yields the existence of a sequence of simple functions (φ_n) such that
(i) $0 \leq \varphi_n(\xi) \leq \varphi_{n+1}(\xi)$ for each $\xi \in \mathbf{R}^d$.
(ii) $\varphi_n(\xi) \to F(\xi)$ for each $\xi \in \mathbf{R}^d$.

We demonstrate that

$$\int_U \varphi_n(f(x))d\mu(x) = \int_U \varphi_n(g(x))d\mu(x). \tag{1}$$

A simple function is a finite linear combination of indicator (characteristic) functions of measurable sets, therefore it is sufficient to show

$$\int_U 1_A(f(x))d\mu(x) = \mu(f^{-1}(A)) = \mu(g^{-1}(A)) = \int_U 1_A(g(x))d\mu(x), \tag{2}$$

for each Lebesgue measurable set $A \subset \mathbf{R}^d$, where 1_A denotes the indicator function of A. Noting that a Lebesgue set is the disjoint union of a Borel set and a Lebesgue negligible set, we need only show (2) for Borel sets. This is immediate from (i). Thus we have verified (1).

We have that $\varphi_n \circ f(x) \to F \circ f(x)$ for each $x \in U$, and that $|\varphi_n \circ f(x)| \leq K\{1 + |f(x)|_2^p\}$ for each $x \in U$ and $n \in \mathbf{N}$, and analogous statements hold if we replace f with g. Applying the Dominated Convergence theorem we obtain

$$\begin{aligned}
\int_U F(f(x))d\mu(x) &= \lim_{n\to\infty} \int_U \varphi_n(f(x))d\mu(x) \\
&= \lim_{n\to\infty} \int_U \varphi_n(g(x))d\mu(x) \\
&= \int_U F(g(x))d\mu(x).
\end{aligned}$$

This verifies (ii).

We show that (ii) implies (i). Let families of sets H and $\mathcal{M}$ be as in Lemma 1. Let $H_1 \in H$. There exists a sequence $(\varphi_n) \subset C(\mathbf{R}^d)$ such that $|\varphi_n(y)| \leq 1 + |y|_2^p$ for each $y \in \mathbf{R}^d$ and $n \in \mathbf{N}$, with $\varphi_n(y) \to 1_{H_1}(y)$ for each $y \in \mathbf{R}^d$. It follows that $\varphi_n \circ f(x) \to 1_{H_1} \circ f(x)$ for each $x \in U$ and $|\varphi_n \circ f(x)| \leq 1 + |f(x)|_2^p$ for each $x \in U$ and $n \in \mathbf{N}$, with analogous statements holding if we replace f by g. Noting that (ii) holds, we apply the Dominated Convergence theorem to obtain

$$\begin{aligned}
\mu(f^{-1}(H_1)) &= \int_U 1_{H_1} \circ f(x)d\mu(x) \\
&= \lim_{n\to\infty} \int_U \varphi_n \circ f(x)d\mu(x) \\
&= \lim_{n\to\infty} \int_U \varphi_n \circ g(x)d\mu(x) \\
&= \int_U 1_{H_1} \circ g(x)d\mu(x) = \mu(g^{-1}(H_1)).
\end{aligned}$$

Thus $H_1 \in \mathcal{M}$. It follows that $H \subset \mathcal{M}$. Lemma 1 yields that $\mathcal{M}$ contains the Borel sets of $\mathbf{R}^d$, therefore f and g are rearrangements.

All elements of the family $\{\prod_{i=1}^{d}[\alpha_i, \infty) : \alpha_i \in \mathbf{R} \text{ for each } i = 1, ..., d\}$ are Borel sets of $\mathbf{R}^d$, therefore (i) implies (iii). To see the converse, we show that $H \subset \mathcal{M}$, given that (iii) holds. We proceed by induction. Let $\mathcal{P}(k)$ be the proposition that all sets of the form $\prod_{i=1}^{k}[a_i, b_i] \times \prod_{i=k+1}^{d}[a_i, \infty) \in \mathcal{M}$, where $a_i, b_i \in \mathbf{R}$. We demonstrate $\mathcal{P}(1)$. Now

$$[a_1, b_1] \times \prod_{i=2}^{d}[a_i, \infty) = \prod_{i=1}^{d}[a_i, \infty) \backslash \left(\bigcup_{n=1}^{\infty} \left([b_1 + 1/n, \infty) \times \prod_{i=2}^{d}[a_i, \infty) \right) \right),$$

and noting that $\mathcal{M}$ is closed under countable increasing union, and differences of two ordered elements (with respect to the partial order $\subset$), we obtain that $[a_1, b_1] \times \prod_{i=2}^{d}[a_i, \infty) \in \mathcal{M}$. This shows $\mathcal{P}(1)$. We demonstrate that $\mathcal{P}(k+1)$ is true given that $\mathcal{P}(k)$ holds. We have that

$$\begin{aligned} \prod_{i=1}^{k+1}[a_i, b_i] \times \prod_{i=k+2}^{d}[a_i, \infty) &= \prod_{i=1}^{k}[a_i, b_i] \times \prod_{i=k+1}^{d}[a_i, \infty) \\ &\backslash \left(\bigcup_{n=1}^{\infty} \left(\prod_{i=1}^{k}[a_i, b_i] \times [b_{k+1} + 1/n, \infty) \times \prod_{i=k+2}^{d}[a_i, \infty) \right) \right). \end{aligned}$$

We are given that $\mathcal{P}(k)$ holds, therefore $\prod_{i=1}^{k}[a_i, b_i] \times \prod_{i=k+1}^{d}[a_i, \infty) \in \mathcal{M}$, and $\prod_{i=1}^{k}[a_i, b_i] \times [b_{k+1} + 1/n, \infty) \times \prod_{i=k+2}^{d}[a_i, \infty) \in \mathcal{M}$ for each $n \in \mathbf{N}$. Noting that $\mathcal{M}$ is closed under countable increasing union and differences of ordered elements, we obtain that $\prod_{i=1}^{k+1}[a_i, b_i] \times \prod_{i=k+2}^{d}[a_i, \infty) \in \mathcal{M}$. This verifies $\mathcal{P}(k+1)$. By induction $\mathcal{P}(d)$ holds, that is all sets of the form $\prod_{i=1}^{d}[a_i, b_i] \in \mathcal{M}$ for $a_i, b_i \in \mathbf{R}$, $i = 1, ..., d$. It is immediate that $\emptyset, \mathbf{R}^d \in \mathcal{M}$, therefore $H \subset \mathcal{M}$. Lemma 1 yields that $\mathcal{M}$ contains the Borel sets of $\mathbf{R}^d$. This shows (i).

Let (iv) hold. The characterisation of the set of rearrangements of a scalar valued function by Eydeland, Spruck and Turkington [**6**] yields that $|g - \sigma|_\infty \in R(|f - \sigma|_\infty)$ in the scalar valued sense for each $\sigma \in \mathbf{R}^d$. Therefore we have

$$\mu\{x : |g(x) - \sigma|_\infty \geq \alpha\} = \mu\{x : |f(x) - \sigma|_\infty \geq \alpha\}$$

for each positive $\alpha \in \mathbf{R}$ or equivalently,

$$\mu\{x : |g(x) - \sigma|_\infty < \alpha\} = \mu\{x : |f(x) - \sigma|_\infty < \alpha\}.$$

Therefore we have $\mu(g^{-1}(C_\alpha(\sigma))) = \mu(f^{-1}(C_\alpha(\sigma)))$, where $C_\alpha(\sigma)$ denotes the open cube of side 2α about $\sigma \in \mathbf{R}^d$. Let K denote the set of all d–dimensional open cubes. We have shown that $K \subset \mathcal{M}$. We now demonstrate that this implies that all open subsets of $\mathbf{R}^d$ belong to $\mathcal{M}$. Recall that $\mathcal{M}$ is closed under countable decreasing intersections, increasing countable unions, and differences of ordered elements of $\mathcal{M}$. For $j = 0, ..., d$ every j–dimensional closed cube is a countable decreasing intersection of j–dimensional open cubes. Further, for $j = 1, .., d$ every j–dimensional open cube with one $(j-1)$–dimensional open face attached is an increasing countable union of j– dimensional closed cubes. Now, for $j = 1, ..., d$, every $(j-1)$–dimensional open cube is the difference of a set of the type described in the preceding sentence, and a j dimensional open cube contained in it. It follows by induction that open and closed cubes of dimensions $0, ..., d$ belong to $\mathcal{M}$. Every open subset of $\mathbf{R}^d$ is a countable disjoint union of open cubes of dimensions $0, ..., d$, therefore such sets belong to $\mathcal{M}$. The methods of Lemma 1, (noting that the intersection of two open sets is open,) yield that $\mathcal{M}$ contains the Borel sets. Thus (iv) implies (i). The converse follows because (i) implies that $\mu(g^{-1}(C_\alpha(\sigma))) = \mu(f^{-1}(C_\alpha(\sigma)))$ for each positive $\alpha \in \mathbf{R}$, $\sigma \in \mathbf{R}^d$. This completes the proof.

3 Energy minimising solutions of atmospheric and oceanic flow

3.1 Introduction

This section studies a variational problem over the set of rearrangements of a prescribed vector valued function, which arises from an energy minimising principle. We study the semigeostrophic equations, (recalled in the next section,) a standard model for slowly varying flows constrained by rotation and stratification, using the methods of Cullen, Norbury and Purser [**5**]. At any given time, $\mathbf{X}$, which describes the state of the fluid, is known on particles. The *Cullen–Norbury–Purser* principle states that for a solution, the particles are arranged to minimise geostrophic energy. This yields a variational problem, minimise energy over the set of rearrangements of a prescribed fluid configuration. We verify the conjecture of Cullen, Norbury and Purser [**5**, Section 5] that the energy minimum is uniquely attained, and that the minimiser is equal to the gradient of a convex function. We prove the following theorem.

Theorem 2 *Let Ω be a bounded connected closed subset of $\mathbf{R}^3$, with smooth boundary. Define, for $\mathbf{X} = (X,Y,Z) \in L^p(\Omega,\mu,\mathbf{R}^3)$, where $2 \le p \le \infty$ and μ denotes 3–dimensional Lebesgue measure,*

$$E(\mathbf{X}) = \frac{1}{2}\int_\Omega X^2 + x^2 + Y^2 + y^2 d\mu(\mathbf{x}) - \int_\Omega \mathbf{X}.\mathbf{x} d\mu(\mathbf{x})$$

where $\mathbf{x} = (x,y,z) \in \Omega$. Suppose $\mathbf{X_0} \in L^p(\Omega,\mu,\mathbf{R}^3)$, for p as above. Then there exists ${\mathbf{X_0}}^* \in R(\mathbf{X_0})$ such that
(i) $E({\mathbf{X_0}}^*) < E(\mathbf{X})$ for each $\mathbf{X} \in R(\mathbf{X_0})\backslash\{{\mathbf{X_0}}^*\}$.
(ii) ${\mathbf{X_0}}^* = \nabla\Psi$ for some convex function $\Psi \in W^{1,p}(\Omega)$.
(iii) ${\mathbf{X_0}}^*$ is a cyclically monotone function.

The functional E represents the Geostrophic energy of the fluid. We define E and $\mathbf{X}$ in the next section. The unique energy minimiser is the monotone rearrangement of the prescribed function: this concept was introduced by Brenier [**2**], and is recalled in section 3.3. The proof uses an approximation argument, with the strict inequality following by the uniqueness of the monotone rearrangement.

3.2 The semigeostrophic equations, and the Cullen–Norbury–Purser principle

We state the three dimensional Boussinesq equations of semigeostrophic theory on an f plane. These are a standard model for slowly varying flows constrained by rotation and stratification, and are used to study front formation in meteorology. We state the equations in the form used by Hoskins [**8**].

$$\frac{Du_g}{Dt} - fv_{ag} = 0, \frac{Dv_g}{Dt} + fu_{ag} = 0, \tag{3}$$

$$\frac{D\theta}{Dt} = 0, \tag{4}$$

$$\nabla.\mathbf{u} = 0,$$

$$\nabla\phi = \left(fv_g, -fu_g, \frac{g\theta}{\theta_0}\right) \tag{5}$$

where

$$\begin{aligned}
\mathbf{u} &\equiv (u, v, w) \equiv \mathbf{u_g} + \mathbf{u_{ag}}, \\
\mathbf{u_g} &\equiv (u_g, v_g, 0), \\
\frac{D}{Dt} &\equiv \frac{\partial}{\partial t} + \mathbf{u}.\nabla
\end{aligned}$$

f is the Coriolis parameter, assumed constant, g denotes the acceleration due to gravity, θ_0 is a reference value of the potential temperature θ, and ϕ is a pressure variable. Subscripts g and ag denote geostrophic and ageostrophic velocity (or wind) components respectively, where the geostrophic velocity is defined to be the horizontal component of velocity in balance with the pressure gradient. This definition is included in equation (5), as is the statement of hydrostatic balance. We solve the equations (for the velocity $\mathbf{u}$) in a closed bounded connected set $\Omega \subset \mathbf{R}^3$, with normal velocity $\mathbf{u}.\mathbf{n}$ given on $\partial\Omega$. For $\mathbf{x} = (x, y, z) \in \Omega$, by making the the substitution

$$\mathbf{X} \equiv (X, Y, Z) \equiv (x + v_g/f, y - u_g/f, (g/f^2\theta_0)\theta),$$

it is shown in Purser and Cullen [**11**] that we may replace (3) and (4) by

$$\frac{D\mathbf{X}}{Dt} = \mathbf{u_g}.$$

We think of $\mathbf{X}$ as a function of the physical space co–ordinates $\mathbf{x}$. Rewriting in terms of $\mathbf{X}$ and $\mathbf{x}$, we have

$$\frac{DX}{Dt} = f(y - Y) \tag{6}$$

$$\frac{DY}{Dt} = f(X - x) \tag{7}$$

$$\frac{DZ}{Dt} = 0. \tag{8}$$

The geostrophic energy E is defined as

$$\begin{aligned}
E &= \int_\Omega \frac{1}{2}u_g^2 + \frac{1}{2}v_g^2 - \frac{g\theta z}{\theta_0} d\mu(\mathbf{x}) \\
&= f^2\frac{1}{2}\int_\Omega X^2 + x^2 + Y^2 + y^2 d\mu(\mathbf{x}) - f^2\int_\Omega \mathbf{x}.\mathbf{X} d\mu(\mathbf{x})
\end{aligned}$$

Henceforth we ignore the constant f^2. At any time t, $\mathbf{X}$ is found on particles by predicting (X, Y, Z) on particles using the equations (6), (7) and (8). The *Cullen–Norbury–Purser* principle states that for a solution, the particles are arranged to minimise geostrophic energy. Suppose one possible state of the fluid is described by values $\mathbf{X_0} = (X_0, Y_0, Z_0)$ which are known on particles. The Cullen–Norbury–Purser principle yields the energy minimisation problem

$$\inf_{\mathbf{X} \in R(\mathbf{X_0})} E(\mathbf{X}),$$

where the energy minimiser (if it exists and is unique) gives the actual state of the fluid. In this way, solutions can be viewed as a sequence of minimum energy states.

We make some (physically reasonable) assumptions to enable us to use vector valued rearrangement theory. Let Ω be a closed, bounded, connected subset of $\mathbf{R}^3$, with smooth boundary. Suppose the possible fluid configuration $\mathbf{X_0} \in L^p(\Omega, \mu, \mathbf{R}^3)$, for $2 \leq p < \infty$, where μ denotes 3–dimensional Lebesgue measure. (Choosing $p \geq 2$ ensures finite geostrophic energy.)

3.3 Monotone rearrangement of vector valued functions

We recall the concept of the monotone rearrangement of a vector valued function: essentially, this is the vector valued analogue of the increasing rearrangement of a real valued function. Let Ω and μ be as in the last paragraph of the previous section. The following theorem is due to Brenier [**2**, section 1.2, theorem 1.1].

Theorem 1.1 *For each $u \in L^p(\Omega, \mu, \mathbf{R}^3)$, where $1 \leq p < \infty$, there is a unique $u^* \in R(u)$ such that*

$$u^* \in \{\nabla\Psi : \Psi \in W^{1,p}(\Omega, \mu), \Psi \text{ convex}\},$$

and the mapping $u \to u^*$ is continuous.

When Ω is not convex, Ψ is understood to be the restriction to Ω of a convex function defined on $\mathbf{R}^3$. We call u^* the *monotone rearrangement of* u. The name comes from the fact that u^* is a cyclically monotone function. We note that McCann [**10**] has generalised the first part of this result (concerning the existence of an essentially unique rearrangement equal to the gradient of a convex function) to more general measures than Lebesgue measure.

Definition A function $u \in L^p(\Omega, \mu, \mathbf{R}^3)$ is *non–degenerate* if $\mu(u^{-1}(E)) = 0$ for each set $E \subset \mathbf{R}^3$ with Lebesgue measure zero. We say that a function which fails to be non–degenerate is *degenerate.*

Brenier established further properties of the monotone rearrangement of a non–degenerate function in the following theorem [**2**, section 1.2, theorem 1.2]

Theorem 1.2 *For each non–degenerate $u \in L^p(\Omega, \mu, \mathbf{R}^3)$ there exists a unique pair (u^*, s), where u^* is the monotone rearrangement of u, and s is a measure preserving mapping from (Ω, μ) to (Ω, μ), such that*

(i) $u = u^* \circ s$.

(ii) *s is the unique measure preserving mapping that maximises $\int_\Omega u(\mathbf{x}).s(\mathbf{x})d\mu(\mathbf{x})$.* Note that Theorem 1.2 is not true if u is degenerate: the measure preserving mapping is not unique, nor do we have uniqueness in property (ii). The author is not aware of any corresponding result for degenerate functions.

3.4 Existence and uniqueness of energy minimiser

Recall that we are studying the energy minimisation problem

$$\inf_{\mathbf{X} \in R(\mathbf{X_0})} \int_\Omega x^2 + X^2 + y^2 + Y^2 d\mu(\mathbf{x}) - \int_\Omega \mathbf{x}.\mathbf{X} d\mu(\mathbf{x}),$$

where $\mathbf{X_0} \in L^p(\Omega, \mu, \mathbf{R}^3)$ for $2 \leq p < \infty$, and $\mathbf{X} = (X, Y, Z)$. We show that the first integral is conserved under rearrangements.

Lemma 2 *Let $\mathbf{X_0}$ be as in Theorem 2. Let $\mathbf{X_1} \in R(\mathbf{X_0})$. Then*

$$\int_\Omega x^2 + X_1^2 + y^2 + Y_1^2 d\mu(\mathbf{x}) = \int_\Omega x^2 + X_0^2 + y^2 + Y_0^2 d\mu(\mathbf{x})$$

where $\mathbf{X_0} = (X_0, Y_0, Z_0)$ and $\mathbf{X_1} = (X_1, Y_1, Z_1)$.

Proof $\mathbf{X_1} \in R(\mathbf{X_0})$ implies that $X_1 \in R(X_0)$. It follows that

$$\int_\Omega X_1^2 d\mu(\mathbf{x}) = \int_\Omega X_0^2 d\mu(\mathbf{x}).$$

A similar result holds for Y_0 and Y_1. The result follows.

To show that there is a unique energy minimiser, it remains to show that

$$\sup_{\mathbf{X} \in R(\mathbf{X_0})} \int_\Omega \mathbf{x}.\mathbf{X} d\mu(\mathbf{x})$$

is uniquely attained. If $\mathbf{X_0}$ is non–degenerate, the result follows easily using Theorem 1.2. Our method of proof is to approximate degenerate functions with a sequence of non–degenerate functions. This shows that the monotone rearrangement is an energy minimiser. We demonstrate that an energy minimiser is the gradient of a convex function: the monotone rearrangement is the unique such amongst the set of rearrangements, therefore the result follows.

Lemma 3 *Let $\mathbf{X} \in L^p(\Omega, \mu, \mathbf{R}^3)$ (where Ω, μ and p are as in section 3.2). Then there exists a sequence of non–degenerate functions $(\mathbf{X_n})$ such that $\mathbf{X_n} \to \mathbf{X}$ in $L^p(\Omega, \mu, \mathbf{R}^3)$.*

Proof For each $n \in \mathbf{N}$, choose a simple function φ_n such that $||\mathbf{X} - \varphi_n||_p \leq 1/n$. Now for each $n \in \mathbf{N}$, define $\mathbf{X_n}$ by $\mathbf{X_n}(\mathbf{x}) = \varphi_n(\mathbf{x}) + (1/n)\mathbf{x}$ for $\mathbf{x} \in \Omega$. It is immediate that $\mathbf{X_n} \to \mathbf{X}$ in $L^p(\Omega, \mu, \mathbf{R}^3)$. It remains to show that $\mathbf{X_n}$ is non–degenerate for each $n \in \mathbf{N}$. Fix $n \in \mathbf{N}$. φ_n is a simple function, therefore it takes finitely many values which we enumerate $\{\mathbf{b_1}, \mathbf{b_2}, ..., \mathbf{b_m}\}$. Define $A_i = {\varphi_n}^{-1}(\mathbf{b_i})$ for each $i = 1, ..., m$. Write ${\mathbf{X_n}}^i$ for $\mathbf{X_n}|_{A_i}$. For a given i, ${\mathbf{X_n}}^i = \mathbf{b_i} + (1/n)\mathbf{x}$. Let E be a Lebesgue negligible subset of $\mathbf{R}^3$. Then

$$\begin{aligned} \mu\left(({\mathbf{X_n}}^i)^{-1}(E)\right) &= \mu\left(A_i \bigcap (nE - n\mathbf{b_i})\right) \\ &\leq \mu(nE - n\mathbf{b_i}) \\ &= \mu(nE) = 0. \end{aligned} \tag{9}$$

By way of explanation, we have used translation invariance of Lebesgue measure to obtain the first equality in (9), and properties of Lebesgue measure to obtain the second. This demonstrates that ${\mathbf{X_n}}^i$ is non–degenerate (as an element in $L^p(A_i, \mu, \mathbf{R}^3)$), for each $i = 1, ..., m$.

Let E be a Lebesgue negligible subset of $\mathbf{R}^3$. Then

$$\begin{aligned} \mu\left({\mathbf{X_n}}^{-1}(E)\right) &= \mu\left(\bigcup_{i=1}^m ({\mathbf{X_n}}^i)^{-1}(E)\right) \\ &= \sum_{i=1}^m \mu\left(({\mathbf{X_n}}^i)^{-1}(E)\right) = 0. \end{aligned} \tag{10}$$

To obtain (10) we have used the countable additivity of μ, and the fact that ${\mathbf{X_n}}^i$ is non–degenerate for each $i = 1, ..., m$. This shows that $\mathbf{X_n}$ is non–degenerate, and completes the proof.

Lemma 4 *Let $\mathbf{X_0}$ be as in Theorem 2. Then*

$$\int_\Omega {\mathbf{X_0}}^*(\mathbf{x}).\mathbf{x} d\mu(\mathbf{x}) \geq \int_\Omega \mathbf{X}(\mathbf{x}).s(\mathbf{x}) d\mu(\mathbf{x})$$

for each $\mathbf{X} \in R(\mathbf{X_0})$ and each $s : \Omega \to \Omega$ a measure preserving mapping.

Proof Let $\mathbf{X} \in R(\mathbf{X_0})$ and let $s : \Omega \to \Omega$ be a measure preserving mapping. From the previous lemma we may choose a sequence $(\mathbf{X_n})$ of non–degenerate functions such that $\mathbf{X_n} \to \mathbf{X}$ in $L^p(\Omega, \mu, \mathbf{R}^3)$. For each $n \in \mathbf{N}$, Theorem 1.2 (i) yields the existence of a unique measure preserving mapping $s_n : \Omega \to \Omega$ such that $\mathbf{X_n} = \mathbf{X_n}^* \circ s_n$. Applying Theorem 1.1 we have $\mathbf{X_n}^* \to \mathbf{X}^* = \mathbf{X_0}^*$. Now

$$\begin{aligned}
\int_\Omega \mathbf{X_0}^*(\mathbf{x}).\mathbf{x} d\mu(\mathbf{x}) &= \lim_{n\to\infty} \int_\Omega \mathbf{X_n}^*(\mathbf{x}).\mathbf{x} d\mu(\mathbf{x}) \\
&= \lim_{n\to\infty} \int_\Omega \mathbf{X_n}^* \circ s_n(\mathbf{x}).s_n(\mathbf{x}) d\mu(\mathbf{x}) \qquad (11) \\
&= \lim_{n\to\infty} \int_\Omega \mathbf{X_n}(\mathbf{x}).s_n(\mathbf{x}) d\mu(\mathbf{x}) \\
&\geq \lim_{n\to\infty} \int_\Omega \mathbf{X_n}(\mathbf{x}).s(\mathbf{x}) d\mu(\mathbf{x}) \qquad (12) \\
&= \int_\Omega \mathbf{X}(\mathbf{x}).s(\mathbf{x}) d\mu(\mathbf{x})
\end{aligned}$$

as required. By way of explanation, (11) holds because s_n is a measure preserving map, and (12) follows because Theorem 1.2(ii) yields that

$$\int_\Omega \mathbf{X_n}(\mathbf{x}).s_n(\mathbf{x}) d\mu(\mathbf{x}) \geq \int_\Omega \mathbf{X_n}(\mathbf{x}).s(\mathbf{x}) d\mu(\mathbf{x})$$

for each measure preserving mapping $s : \Omega \to \Omega$, and for each $n \in \mathbf{N}$. This completes the proof.

Lemma 5 *Let* $\mathbf{X_0}$ *be as in Theorem 2. Then*

$$\int_\Omega \mathbf{X_0}^*(\mathbf{x}).\mathbf{x} d\mu(\mathbf{x}) > \int_\Omega \mathbf{X}(\mathbf{x}).\mathbf{x} d\mu(\mathbf{x})$$

for each $\mathbf{X} \in R(\mathbf{X_0})\backslash\{\mathbf{X_0}^*\}$.

Proof Applying the previous lemma for the identity mapping, we have

$$\int_\Omega \mathbf{X_0}^*(\mathbf{x}).\mathbf{x} d\mu(\mathbf{x}) \geq \int_\Omega \mathbf{X}(\mathbf{x}).\mathbf{x} d\mu(\mathbf{x})$$

for each $\mathbf{X} \in R(\mathbf{X_0})\backslash\{\mathbf{X_0}^*\}$. It remains to show strict inequality. Suppose there exists $\mathbf{X_1} \in R(\mathbf{X_0})$ such that $\int_\Omega \mathbf{X_1}.\mathbf{x} d\mu = \int_\Omega \mathbf{X_0}^*.\mathbf{x} d\mu$. Applying the previous lemma to $\mathbf{X_1} \in R(\mathbf{X_0})$ we obtain

$$\begin{aligned}
\int_\Omega \mathbf{X_1}(\mathbf{x}).\mathbf{x} d\mu(\mathbf{x}) &= \int_\Omega \mathbf{X_0}^*(\mathbf{x}).\mathbf{x} d\mu(\mathbf{x}) \\
&\geq \int_\Omega \mathbf{X_1}(\mathbf{x}).s(\mathbf{x}) d\mu(\mathbf{x})
\end{aligned}$$

for each measure preserving mapping $s : \Omega \to \Omega$. Brenier [**2**, Proposition 2.1] yields that $\mathbf{X_1} \in \{\nabla\Psi : \Psi \in W^{1,2}(\Omega), \Psi \text{ convex}\}$. However Theorem 1.1 states that $\mathbf{X_0}^*$ is the unique member of $R(\mathbf{X_0})$ belonging to $\{\nabla\Psi : \Psi \in W^{1,2}(\Omega), \Psi \text{ convex}\}$, therefore $\mathbf{X_1} = \mathbf{X_0}^*$. This completes the proof.

Proof of Theorem 2

Follows from Lemmas 2 and 5.

Acknowledgements

The author would like to acknowledge the insight and help of G.R. Burton. The research was carried out under the guidance of M.J.P. Cullen, and the meteorological significance of these results is discussed in Cullen and Douglas [**4**]. The author would also like to thank J. Norbury, Y. Brenier and J-D. Benamou for stimulating discussions. The author is grateful for the hospitality and support of the Isaac Newton Institute for Mathematical Sciences.

References

1. A. ALVINO, P-L. LIONS and G. TROMBETTI, *On optimization problems with prescribed rearrangements.* Nonlinear Analysis, Theory, Methods and Applications 13, (1989) pp. 185–220.
2. Y. BRENIER, *Polar factorisation and monotone rearrangement of vector–valued functions.* Comm. Pure Appl. Math. 44, (1991) pp. 375–417.
3. G.R. BURTON, *Rearrangement of functions, maximisation of convex functionals, and vortex rings.* Math. Ann. 276, (1987) pp. 225–253.
4. M.J.P. CULLEN and R.J. DOUGLAS, *New Mathematical tools for the Lagrangian description of atmospheric dynamics.* U.K. Meteorological Office Forecasting Research Scientific Paper No. 39.
5. M.J.P. CULLEN, J. NORBURY, and R.J. PURSER, *Generalised lagrangian solutions for atmospheric and oceanic flows.* Siam J. Appl. Math. 51, (1991) pp. 20–31.
6. A. EYDELAND, J. SPRUCK and B. TURKINGTON, *Multiconstrained variational problems of nonlinear eigenvalue type: new formulations and algorithms.* Math. Comp. 55, (1990) pp. 509–535.
7. P.R. HALMOS, *Measure Theory.* Springer–Verlag, New York (1974).
8. B.J. HOSKINS, *The geostrophic momentum approximation and the semigeostrophic equations.* J. Atmospheric Sci. 32, (1975) pp. 233–242.
9. A.S. KECHRIS, *Classical Descriptive Set Theory.* Springer–Verlag, New York (1995).
10. R.J. MCCANN, *Existence and uniqueness of monotone measure preserving maps.* Duke Math. J. 80, (1995) pp. 309–323.
11. R.J. PURSER and M.J.P. CULLEN, *A duality principle in semigeostrophic theory.* J. Atmospheric Sci. 44, (1987) pp. 3449–3468.
12. H.L. ROYDEN, *Real Analysis.* 2nd Edition. Collier–Macmillan Limited, London 1963.