Weather Forecast Error Decomposition using Rearrangements of Functions

by Gareth Lanagan

<small>ABERYSTWYTH UNIVERSITY</small>

2012

<small>THESIS</small>

<small>SUBMITTED TO ABERYSTWYTH UNIVERSITY</small>

<small>BY</small>

<small>GARETH DANIEL EDWARD LANAGAN, M.SC (WALES)</small>

<small>IN CANDIDATURE FOR THE DEGREE OF</small>

<small>PHILOSOPHIAE DOCTOR</small>

Declarations.

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed _____ (Gareth Lanagan)

Date _____

Statement 1

This thesis is the result of my own investigations, except where otherwise stated. Where correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s). Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed _____ (Gareth Lanagan)

Date _____

Statement 2

I hereby give consent for my thesis to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed _____ (Gareth Lanagan)

Date _____

Acknowledgments.

Upon writing a document of acknowledgment, spanning four years of endeavour, individuals are always forgotten. To those people, I am grateful for your help, but just a little forgetful in my old age.

Every PhD student must start with his supervisor, in my case this is Dr. Robert Douglas. I am extremely grateful for the help I have received with my work, during the past four years, and before; having been in Aberystwyth since September 2000. However my biggest debt is for the pastoral care I've received during my time at Aberystwyth. Without this support, especially during my many bouts of illness, it is safe to say I wouldn't have reached the end of this thesis. Sir, I will now leave you in peace and quiet.

The support given to me at the Institute of Mathematics and Physics, has been comforting during my time as a postgraduate student. My room buddy, Miss Hannah Whittle, has been a constant source of fun; always essential to breaking up the day. Hannah fach, dwi'n edrych mlaen i gerdded ar y llwyfan gyda ti yn seremoni graddio 2012! Simon Cox has been a sympathetic ear, and another lecturer who will have some more free time upon my departure. I am grateful to the academic staff in the institute, especially on the mathematics side of the river; special mentions go to Alan Jones, Dave Binding and John Lane. Further sources of entertainment have been the staff in the general office, the recent departees of Sarah Spring, Rebecca Sidford, Elizabeth Hyde and Christine Gilbert are all lovely ladies, and have always been ready with a smile. Finally, the portering staff of Derek and Alan alongside the infamous Dai Chef will be missed upon my departure.

Every man needs to be mothered, and being 120 miles from home, Tegwen Meredith, the recently retired Institute librarian, has acted as my Aberystwyth mother. Tegwen has been a massive help with my literature survey, and is a fountain of knowledge on all matters to do with the library. Furthermore, the fact that Tegwen has kept a motherly eye on me has not gone unnoticed, for this I am eternally grateful. Mwynhewch eich amser hamdden, nawr rydych chi wedi ymddeol.

During such a piece of work, leisure time is important in striking the balance in the search of sanity. Pantycelyn, the students, staff and fellow wardens has been a welcome distraction after a hard day at the grindstone. Furthermore, Barry Wanderers Cricket Club, has provided me with many great days, e.g. the promotions of 2009 and 2010, to Jamie Keir, Jimmy Hassett and the boys, I thank you for the good times.

My father Nicolas and mother Gwenda, both Aberystwyth graduates, have been a constant source of support, financially and otherwise. I'm sure they will both look forward to the day that I cease to be a tax dodger! Words cannot do justice the gratitude that I need to show them both, but I hope they both know how grateful I am. Also, my brother Richard, and his wife Natalie have provided unerring support. Thank you for being a constant source of amusement, and thank you for allowing me to be a part of your greatest day. To my parents, my brother and sister in law, one day, I will repay my debt to you.

Last, but by no means least is my beautiful girlfriend Rhian James. While Rhian had only had to put up with my Ph.D. stress for six months, her unerring support during times of sleep deprivation and thesis stress was essential for my sanity. Her proof reading picked up many a typo and the odd coma and meant my minor changes were just that. Caru ti Rhian fach!

iv

Abstract.

This thesis applies rearrangement and optimal mass transfer theory to weather forecast error decomposition. Errors in weather forecasting are often due to displacement of key features; conventional error scores do not necessarily favour good forecasts, nor are they descriptive of how the forecast failed. We study forecast error decomposition, where error is split into an error due to displacement and an error due to differences in qualitative features. In its simple formulation, we seek rearrangements of the forecast which are a best fit to the actual data, and then find the "least kinetic energy" of a notional velocity transporting the forecast to a best fit. In mathematical terms, we are characterising those elements of a set of rearrangements which are closest (in the sense of L^2) to a prescribed square integrable function, and seeking the least 2-Wasserstein distance squared between the forecast and the closest displaced forecasts. We demonstrate that there are closest rearrangements, and characterise this set; the best fitting rearrangements are determined up to rearrangement on the level sets of positive size of the prescribed function.

Displacement error is calculated by finding the minimum value of an optimal mass transfer problem; we review previous work, demonstrating the connection with transport of the forecast to the best fit. A problem with the simple formulation of forecast error decomposition is that because the qualitative features error is taken first, an error in qualitative features may be penalise as a large displacement error. We conclude this thesis by considering a formulation which minimises both errors simultaneously.

Contents

x

List of Figures

Chapter 1

Introduction.

1.1 Outline.

How do we measure the difference between two quantities? We describe a measure of error which is split into two parts: difference in qualitative features and displacement error. This thesis is concerned with the theoretical underpinning of this idea.

Weather forecast error decomposition measures the difference between the data which is forecasted and the data which is observed. Errors in weather forecasts are often displacements in significant features, either in space or time. This thesis will consider how to apply rearrangements of functions and optimal mass transfer theory to weather forecast error decomposition: splitting the error into a component due to differences in qualitative features and a component due to displacement.

This chapter begins by illustrating the problems which exist in the current forecast error decomposition in section 1.2. The section starts with a discussion of what meteorologists perceive to be a good forecast, which in turn we will use to define our forecast error decomposition. To illustrate the problems with the current error scores we consider

1

some examples.

Having discussed the problems which exist for error scores such as the L^2-norm of the difference in section 1.2, in section 1.3 we suggest a fresh approach. This formulation, based on work by Hoffman, Liu, Louis and Grassotti [47] is presented via an illustrative example; we outline the work of Hoffman and his co-workers.

Error decomposition can be applied to other situations where displacement is a significant contribution to the error; we consider a specific example. Then we build on the ideas of Douglas [30] to show we can use rearrangements and optimal mass transfer to describe weather forecast error decomposition. We define the error decomposition formally and conclude the section by explaining some of the problems which arise with our simple formulation and possible solutions.

The notion of rearrangements in section 1.4 is introduced with an example illustrating the concept, before formalising the definition of rearrangements. Furthermore, we introduce the important notion of the set of rearrangements before giving a formal definition; we state some basic properties of the set. With a view to the work which follows we consider an equivalent definition for rearrangements.

An example of optimal mass transfer is minimising the work done when transferring mass from a pile of rubble to an excavation; the minimisation is based on a cost function which is the cost of transporting the mass. We can use optimal mass transfer along with the notion of rearrangements to define the second part of our forecast error. This is the content of section 1.5. We begin by introducing the problem as considered by Monge [59]. We use the iron ore example of McCann [58] for ease of understanding of this concept. We move onto a discussion regarding the Monge-Kantorovich problem, a relaxation of the Monge problem, discussing certain classes of cost functions. Solutions

of the Monge-Kantorovich problem may be of a form which gives solutions of Monge's problem.

One of the problems with Monge's original minimal work problem was the difficulty in obtaining solutions to the cost function suggested. Section 1.5.1 notes the non-uniqueness of the solution for Monge's original cost function while noting that there is a unique solution for the quadratic cost function. The section on optimal mass transfer is concluded with an historical review of the problem since it was first stated by Monge in 1781. We note the recent developments since Brenier's work in 1987, this illustrated how the field applied to many different mathematical fields.

Section 1.6 includes all the widely used notation throughout the text. Then section 1.7 outlines what is to come for the remainder of the thesis. Finally, section 1.8 is a literature survey on the practical applications of both rearrangements of functions and optimal mass transfer problems.

1.2   Limitations of conventional error scores.

A current hot topic in the field of meteorology is how best to measure the error in weather forecasting. This has applications in rainfall verification (see Douglas [30]), global drought prediction (see Begueria et al. [7]) and many other areas. We consider the question:

<div align="center">What is a good forecast?</div>

The key to successful weather forecast error evaluation is to find a measure of error which agrees with what a meteorologist would say about the forecast. The error decomposition provides us with two things; firstly, how good the forecast is, and secondly, where and

4

how the forecast failed.

We begin with the conventional methods used within weather forecast error decompoistion, of which there are several. For reasons which will become apparent, we concentrate on the L^2-norm of the difference:

$$\|q_1 - q_2\|_2 = \left\{ \int_\Omega (q_1 - q_2)^2 \right\}^{\frac{1}{2}}, \tag{1.1}$$

where in this case q_1 is actual rainfall and q_2 is forecasted rainfall and Ω is the domain on which the forecast error is being carried out. We use rainfall as our primary example, however it could be any other important quantity in weather forecasting, either scalar valued (like rainfall) or vector valued (such as wind velocity). In our error score we consider scalar valued quantities.

If the error is primarily one of displacement, conventional error scores are likely to give a misleading impression of how good the forecast is. We illustrate with an example.
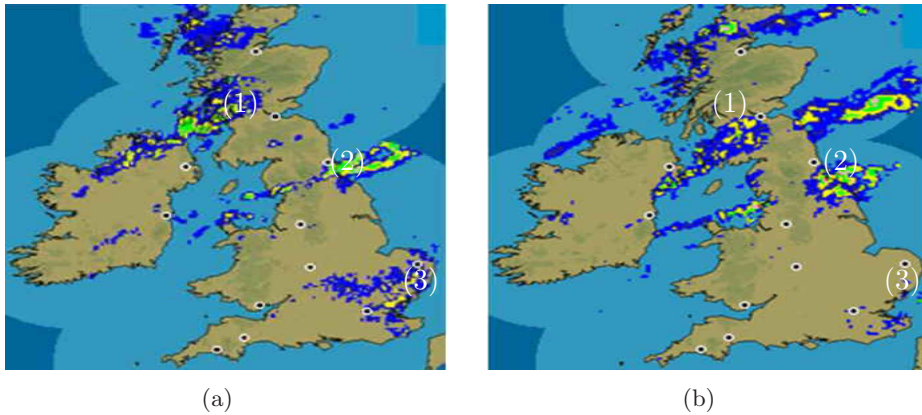
Figure 1.1: An illustration of the (a) Forecasted Data (b) Observed Data for rainfall on a given day

There are three main features in the forecast:

(1) Some heavy rain forecasted to the west of Scotland around the Stirling area;

(2) A band of rain is forecast in Newcastle, with the heaviest of the rain taking place just south of Newcastle;

(3) Light rain is forecast in the south, in between London and Norwich.

By considering the observed data Figure 1.1(b) we note the forecast has erred with all three main features, albeit in different ways:

(1) The heavy rain forecasted for Stirling actually occurs over Edinburgh, keeping similar qualitative features to the forecast;

(2) The band of rain which was due to occur over Newcastle has moved south and takes place over York, but again has similar features to those forecasted;

(3) The light rain which was forecasted above London does not materialise, hence there are very different qualitative features.

Features (1) & (2) would be considered by meteorologists to be well forecast. Although the features are not forecast in the correct place, they are simply displaced a little. The conventional error score would penalise the misplaced rain twice; once for having no rain where there should be rain, and once for forecasting rain where there should be none. However a forecast of no rain would only get penalised once, therefore giving us a better error score. It is not only the meteorologists who would consider this to be a worse forecast. The general public would prefer to know when and if these is rainfall in their region.

We conclude from the above that the error in both (1) and (2) is due to a small displacement and (3) is a difference in qualitative features. Our simple formulation of weather forecast error decomposition describes the error in this form and hence is more descriptive of how the forecast fails. The next subsection describes the mathematical formulation of this idea.

1.3   Mathematical formulation.

A key idea in our formulation is that of a best fitting forecast. Unlike the conventional error score, we search for a best fit of the forecasted data. We move the forecast onto the actual data to get a best fit of the actual data; this is illustrated by figure (1.2(c)). The best fitting data provides two things for us. Firstly, it provides a clear indication of the displacement between the forecasted and the actual data. Secondly, the superimposition of the forecasted data onto the best fitting data allows us to compare how similar the qualitative features are on the forecasted and the actual data.

(a)   (b)   (c)
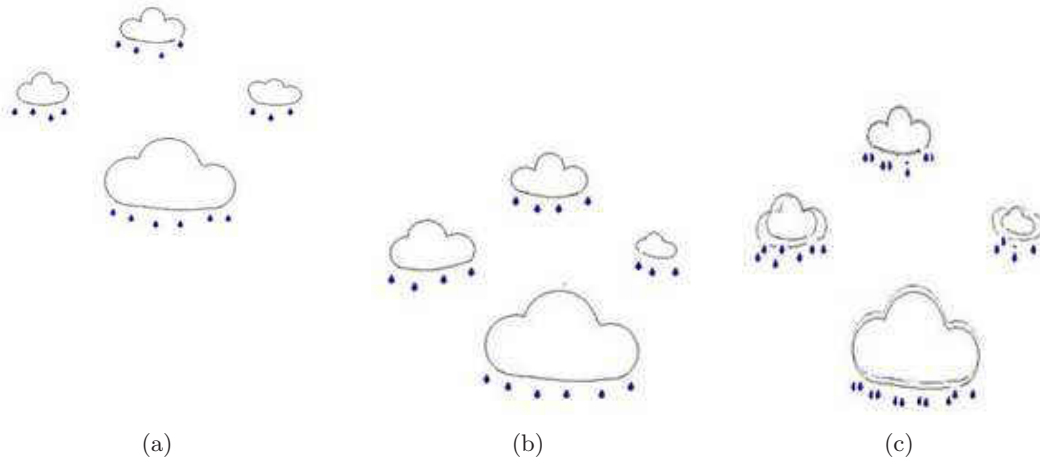
Figure 1.2: An illustration of the (a) Forecasted Data; (b) Observed Data; (c) Best Fitting Data to show the need for a displacement error

Similarly to figures (1.1(a)) and (1.1(b)) the error in figures (1.2(a)) and (1.2(b)) is primarily one of displacement. Figure (1.2(c)) illustrates the similarities in qualitative features. The forecasted data can only be moved in a space of allowable displacements, in our case a set of rearrangements, which we define formally in the next section.

From this idea of a best fitting forecast we can develop our definitions of the displacement and qualitative features errors. The qualitative features error measures the difference in features between the actual data and the best fit. The displacement error measures something similar to the kinetic energy of a notional velocity which transports the forecast to the best fit which is used to rearrange the forecast onto the best fit. To think of the displacement error in another form: it gives us a way of measuring how far we've had to move the forecast onto the actual to get the best fit.

Hoffman et al [47] suggested a solution to some of the problems which have arisen with the conventional error score:

> "Better descriptions of forecast error are often given in terms of displacement."

Hoffman's paper decomposes the error score into three sections.

1. Displacement Error

2. Qualitative Features Error(or Amplitude Error)

3. Residual Error

The displacement error measures how much of the forecast error can be accounted for by moving the forecast to best fit the analysis. The qualitative features error measures how much of the forecast error can be accounted for by changing the amplitude of the

displaced forecast to best fit the analysis. The distortion error is defined to be the sum of the displacement error and the qualitative features error. The residual error is the part of the forecast error which is not accounted for by the distortion error. In this case the best fits are local translations.

While we base our decomposition on this model we believe the error score need only be decomposed into the displacement error and the qualitative features error. It is our belief that the residual error should be zero if the mathematical formulation of the concept is correct; it should only be non-zero for numerical implementation. If the aim is to express the error as a combination of displacement and difference in features, residual error is a measure of the extent to which one has failed in this aim. If the definitions of "displacement" and "difference in features" do not describe all the error these definitions need to be changed. Therefore, our formulation does not consider the residual error further. Our total error is a weighted sum of the displacement and qualitative features error. Our definitions of these terms are inspired by Hoffman [47], but based on different mathematical methods.

One of the advantages of considering the new error score is that it is more descriptive of how the forecast has failed, i.e. how much error is due to differences in features and how much is due to displacement. This sentiment is supported by the following quote [47]:

> "The distortion representation of forecast errors should prove useful for describing forecast skill."

While in this thesis error decomposition has been applied to weather forecasting, the concept can be applied to any problem where there is significant error due to displace-

ment. As an illustration we consider the problem of readings taken from an oscilloscope. We consider the following figures:
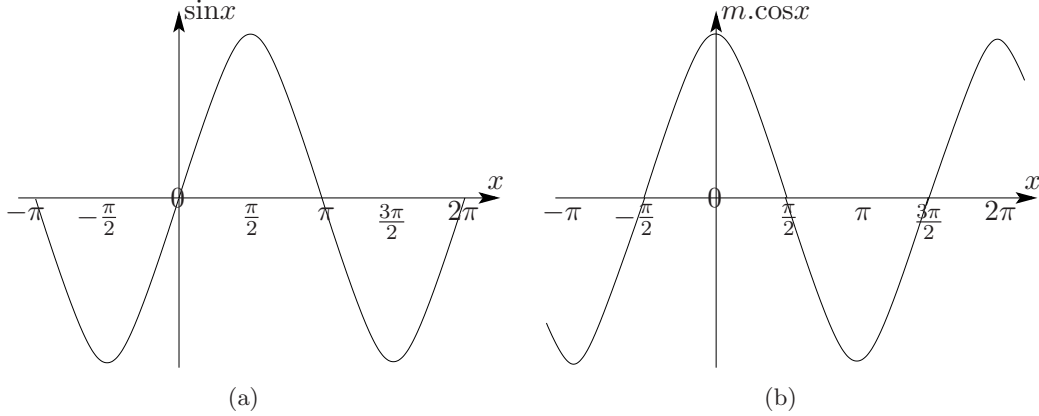
Figure 1.3: An illustration of (a) the sine curve and (b) a multiple m of the cosine curve, where $1 \leq m \leq 1 + \epsilon$

It is not appropriate to calculate the error associated with figure 1.3(b) with the L^2-norm given that the error is primarily one of displacement, and L^2-norm would not describe this well. Therefore, some sort of displacement error is required. The error is due to the phase of the wave and/or the amplification with the phase (which corresponds to the displacement in our case) being of greater significance. This idea is analogous to that of Hoffman et al. [47]. However, both the phase and amplification can change. Our formulation describes these changes and would be descriptive of the big error due to displacement and small error due to change of amplitude. (Note this example is simplified because of the periodicity; this feature will not be present in general.)

Before giving a mathematical description of the forecast error decomposition, we define rearrangements and the set of rearrangements. Two functions are rearrangements if any given set of values are taken on sets of the same size, but the values may be taken in different places. Therefore, the set of rearrangements is the set which contains all such functions; for a function f the set of rearrangements is denoted $\mathcal{R}(f)$. A good example

of rearrangement is a Lagrangian conserved quantity in an incompressible fluid; values are fixed on particles, but the particles are free to move as long as they do not change size.

We give a mathematical description of forecast error decomposition. Let q denote an important quantity in weather forecasting. Let the actual distribution be q_1 and the forecasted distribution be q_2. We note that $q : \Omega \rightarrow \mathbb{R}^d$ and the domain $\Omega \subset \mathbb{R}^n$. From Douglas [30, Section 4.1] we split the forecast error up in the following manner:

(i) Find $\hat{q} \in \mathcal{R}(q_2)$ such that $\hat{q}$ attains

$$\inf_{q \in \mathcal{R}(q_2)} \|q_1 - q\|_2. \tag{1.2}$$

(ii) Calculate the 'length of the shortest path' between q_2 and $\hat{q}$ in $\mathcal{R}(q_2)$. Denote the shortest length by $dist(q_2, \hat{q})$ where $dist(q_2, \hat{q})$ is the "kinetic energy" required to move the forecasted distribution onto the best fit to the actual distribution, the extent to which $\hat{q}$ is displaced from q_2.

The value of (1.2) is the qualitative features error. The total error is

$$TE(q_1, q_2) = (1 - \theta) \inf_{q \in \mathcal{R}(q_2)} \|q_1 - q\|_2 + \theta \inf_{\hat{q} \in \mathcal{M}} dist(\hat{q}, q_2), \tag{1.3}$$

where $0 \leq \theta \leq 1$ and $\mathcal{M}$ denotes the set of $q \in \mathcal{R}(q_2)$ which attains the infimum in (1.2). (One of the main results of this thesis is a characterisation of $\mathcal{M}$; it is non-empty.) θ is

chosen so that neither component of (1.3) dominates the other. (A suitable value of θ could be determined by running a number of test cases.) We note that the first half of the error, $(1-\theta)\inf_{q\in\mathcal{R}(q_2)}\|q_1 - q\|_2$ is the qualitative features part of the error. The second half of the error, $\theta\inf_{\hat{q}\in\mathcal{M}}dist(\hat{q}, q_2)$ is the displacement error.

While we've addressed some problems with our simple formulation of the weather forecast error, there are new problems to consider. We begin by finding the qualitative features error, then search through all the best fitting rearrangements to see which minimises the kinetic energy to yield the displacement error. Therefore, the qualitative features error is taken without anticipating the impact on the displacement error. Consider figures (1.1(a)) and (1.1(b)). In the forecast there are scattered showers forecast for between Norwich and London. With the current formulation, for the best fit, rain could be taken down from Scotland to make up for the lack of the scattered showers without consideration for the displacement error. The result would be an error in qualitative features being penalised as though it were a large displacement error.

One approach would be to minimise forecast and displacement errors simultaneously. This is the content of Chapter 5; similarly to the simple formulation we consider this as a rearrangement problem.

There is a simpler solution to our problem. The idea would be to restrict the area on which we carry out an error decomposition; we would only seek "local best fits". Something like figure 1.4 could be used:

(a)   (b)
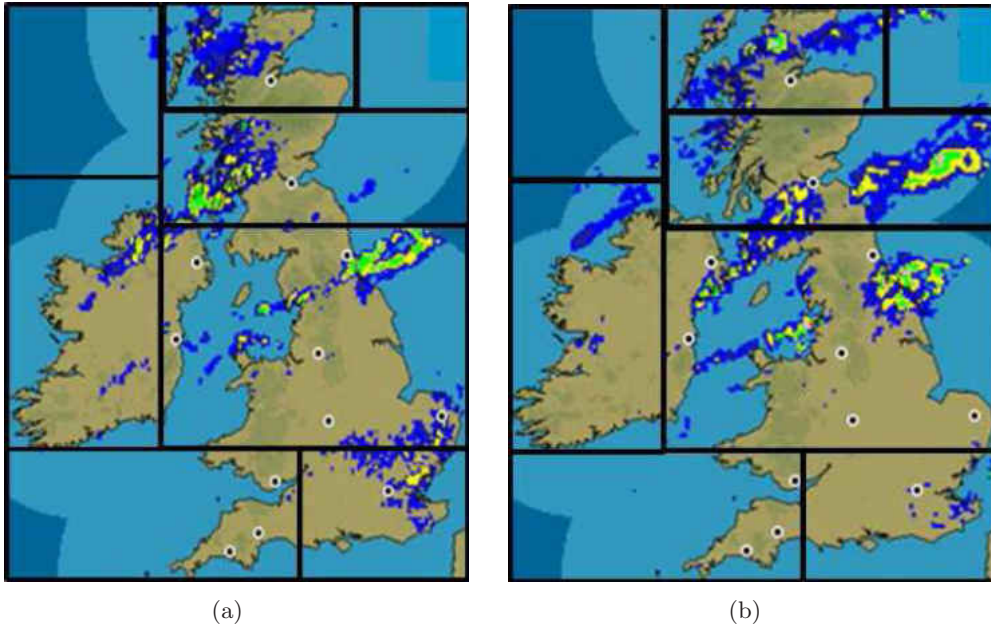
Figure 1.4: An illustration of (a) the forecasted data; and (b) the actual data broken into several sections with in which the weather features must remain

This would be a straightforward of way combating the large displacement accrued from moving the rainfall in the manner discussed previously. However, the choice of the regions could possibly pose problems; one would not want heavy rainfall near the boundaries. However, if we seek advice from meteorologists along with the use of common sense when choosing the regions this should minimise any problems. Different choices of regions could yield different answers and this is another reason to utilise insight from meteorologists when selecting the regions.

1.4   Rearrangements of functions.

The notion of rearrangements of functions is central to forecast error decomposition. We introduce the concept via an illustrative example. Consider the following real valued functions f and g on the domain $x \in [0,1]$, defined by $f(x) = x$ and $g(x) = |1 - 2x|$. We will show that

$$\int_0^1 (f(x) - \alpha)_+ \, \mathrm{d}x = \int_0^1 (g(x) - \alpha)_+ \, \mathrm{d}x, \tag{1.4}$$

for every real number α, where the $_+$ denotes the positive part of the function, i.e.

$h_+ (x) = \max \{ h(x), 0 \}$. Fix $0 < \alpha < 1$
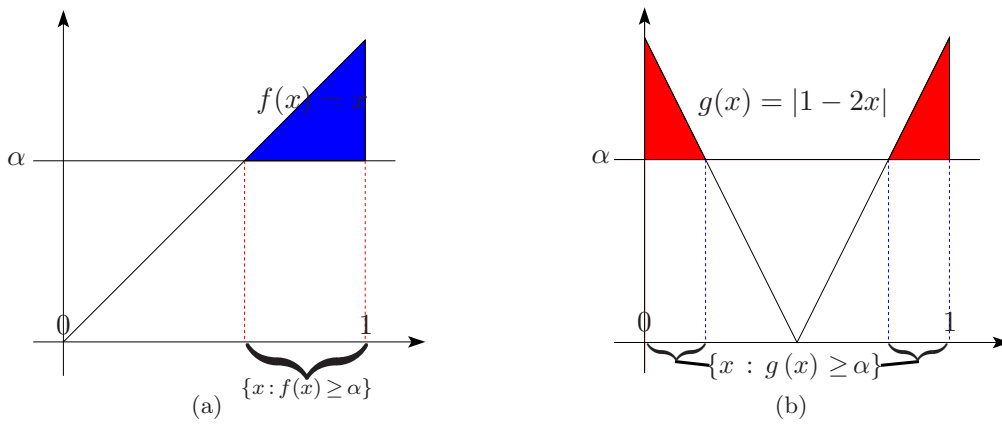
Figure 1.5: An illustration of the following functions (a) $f = f(x)$ and (b) $g = g(x)$ with the line $y = \alpha$ intersecting the graph of the function.

We consider the area trapped between $y = \alpha$ and the graph of the function. This is

illustrated in figures 1.5(a) and 1.5(b). We compare the blue area to the red areas.

They are in fact equal; moreover for whichever α we choose this will still be true. For

$0 < \alpha < 1$ we have

$$\int_0^1 (f(x) - \alpha)_+ \, \mathrm{d}x = \frac{1}{2}(1 - \alpha)^2 = \int_0^1 (g(x) - \alpha)_+ \, \mathrm{d}x. \tag{1.5}$$

If $\alpha \geq 1$ then $y = \alpha$ lies above the graphs of the functions, hence:

$$\int_0^1 (f(x) - \alpha)_+ dx = 0 = \int_0^1 (g(x) - \alpha)_+ dx. \tag{1.6}$$

Furthermore, if $\alpha \leq 0$ then y=α lies below the graphs. We have that

$$\begin{aligned}
\int_0^1 (f(x) - \alpha)_+ dx &= -\alpha + \int_0^1 f(x) dx \\
&= -\alpha + \int_0^1 g(x) dx = \int_0^1 (g(x) - \alpha)_+ dx. \tag{1.7}
\end{aligned}$$

When f and g satisfy (1.4) for every $\alpha \in \mathbb{R}$ we will say that they are rearrangements. Before generalising this definition we define the measure spaces we use in the work that follows. Here μ is a positive measure on Ω, i.e. a function that assigns "size" to subsets of Ω. We assume $\mu(\Omega) < \infty$, that is (Ω, μ) is a finite measure space. We denote n-dimensional Lebesgue measure by λ_n; for $n = 1$, 2 or 3 this corresponds to length, area or volume respectively. We say that a measure μ is absolutely continuous with respect to a measure ν if for every ν-measurable set A satisfying $\nu(A) = 0$ then $\mu(A) = 0$ as well. Typically we will work on a measure space $(\Omega \subset \mathbb{R}^n, \mu)$ where μ is absolutely continuous with respect to λ_n.

Definition 1.4.1. *Let $\Omega \subset \mathbb{R}$ be a bounded set and let μ be a measure as above. Let $f, g : \Omega \to \mathbb{R}$ be two integrable functions(i.e. $\int_\Omega |f|, \int_\Omega |g| < \infty$). f is a rearrangement of g if*

$$\int_{\Omega} (f(x) - \alpha)_+ d\mu(x) = \int_{\Omega} (g(x) - \alpha)_+ d\mu(x), \tag{1.8}$$

for all real α.

To anticipate a discussion on several equivalent definitions for rearrangements we note that f and g satisfying the following:

$$\lambda_1 \{x : f(x) \geq \alpha\} = \lambda_1 \{x : g(x) \geq \alpha\}, \tag{1.9}$$

where λ_1 is a 1-dimensional Lebesgue measure, for every $\alpha \in \mathbb{R}$, are rearrangements. As before we intersect the graphs of the functions f and g with the line $y = \alpha$. But, in this case we consider the length of the line where the graph of the function lies above the line $y = \alpha$. If this length is equal for two functions, f and g, we can say that f and g are rearrangements. For the functions as considered in figure 1.5 we've illustrated the process for a specific α. For both functions the length of the line is 0 when $\alpha \geq 1$, $1 - \alpha$ for $0 \leq \alpha \leq 1$ and 1 for $\alpha < 0$.

In general a function has many rearrangements. We have that f and g are rearrangements of each other on $[0\,,1]$; another rearrangement is $h(x) = 1 - x$ because the area above $y = \alpha$ (for all real α) and below the graph of h is $\frac{1}{2}(1 - \alpha)^2$. This is illustrated below:
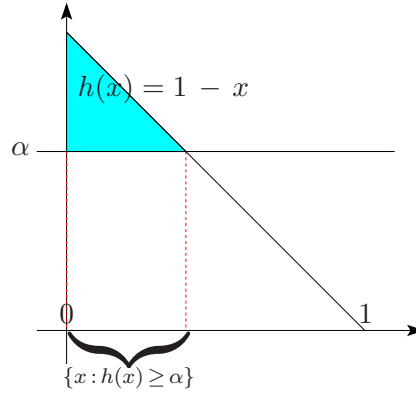
Figure 1.6: The function $h = h(x)$ with the line $y = \alpha$ intersecting the graph of the function

We note that h is a rearrangement of both f and g; it is easily seen that rearrangement is an equivalence relation. Let Ω be a bounded subset of $\mathbb{R}^n$ and μ is absolutely continuous with respect to the n-dimensional Lebesgue measure λ_n. We define an equivalence class, or the set of rearrangements as follows:

Definition 1.4.2. *For a given non-negative square integrable $f : \Omega \to \mathbb{R}$ the set of rearrangements is characterised in the following manner:*

$$\mathcal{R}(f_0) = \left\{ f \geq 0 \; : \; \int_\Omega (f - \alpha)_+ \; \mathrm{d}\mu = \int_\Omega (f_0 - \alpha)_+ \; \mathrm{d}\mu \;\; \forall \alpha > 0 \right\}, \qquad (1.10)$$

where $+$ denotes the non-negative part of the function.

An important point to note is that all elements from the set of rearrangements have the same L^p norm. For functions f and g rearrangements we have that

$$\|f\|_p \;=\; \|g\|_p \ \text{ for } 1 \le p \le \infty, \ \text{where } \|f\|_p \;\;=\;\; \left(\int_\Omega |f|_\infty^p \ \mathrm{d}\mu \right)^{\frac{1}{p}}$$

$$\text{and } \|f\|_\infty \;\;=\;\; ess\,\sup\,\{|f(x)| \,:\, x \in \Omega\}. \quad (1.11)$$

Finally, for non-trivial f_0, the set of rearrangements is closed but non-convex and not compact.

1.5 Optimal Mass Transfer.

Optimal mass transfer theory has been a topic of recent interest in many fields. The theory is used to define the displacement error in our forecast error score using a result by Brenier and Benamou [10]. The problem of optimal mass transfer is to transfer mass from one set to another set in such a way as to minimise cost. The optimal strategy minimises the cost amongst allowable strategies (which will be defined later). The archetypal optimal mass transfer problem is how to minimise cost when transferring mass from a set U to a set V of equal (finite) size. Optimality is measured against a cost function $c = c(x, y)$, where $c(x, y)$ represents the cost per unit mass of transporting material from $x \in U$ to $y \in V$; we have a set of possible strategies, S, with the optimal one (if such exists) minimising the total cost. Optimal mass transfer problems were introduced by Monge [59]; the problem was to minimise the work done when filling in trenches with rubble/soil.
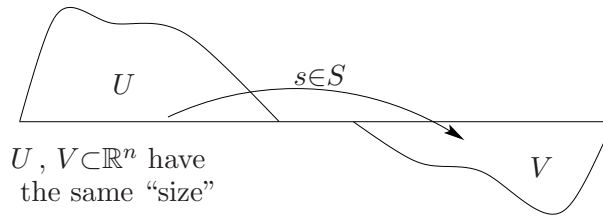
Figure 1.7: Monge's mass transportation problem [71]

Monge's suggestion was that the two equal volumes should be split into infinitesimal particles. Each location in U has a destination in V; the sum of the path lengths multiplied by the volumes of the particles is minimised. Given that Monge was looking to minimise the work done the cost function was chosen to be $c(x, y) = |x - y|$, where $|\cdot|$ denotes the Euclidean distance. Therefore, for this choice of the cost function the work done is proportional to the distance moved.

Optimal mass transfer problems occur naturally in many applications, particularly in economics [39]. A specific example is described by McCann [58]. McCann's problem is how to supply iron ore to factories in order to minimise the cost of transporting the iron ore. The iron mines are defined to be the domain U and the destination of the iron mine, the factories, is V. We need to hire lorries to transport the iron ore, which is at a cost of $c = c(x, y)$ to transport a ton of iron ore from the mine, x, to the factory, y. The amount of iron ore we can extract and the amount of iron ore the factories can receive are fixed, whence the strategies about how to transport the iron ore must be size-preserving. In the problem as considered in [58] the cost function is $c(x, y) = h(|x - y|)$ given by a strictly concave function $h \geq 0$.

Allowable strategies are constrained by the requirement that "sets of equal size are mapped to each other" for Monge's problem. More formally we define S to be a set of measure-preserving mappings $s : U \rightarrow V$. (We give precise definitions later.)

Before defining Monge's problem mathematically, we define the cost associated with $s \in S$. The cost function is as follows

$$C(s) \;=\; \int c\,(x\,,\,s(x))\; \mathrm{d}\mu(x),\tag{1.12}$$

where x is a point in the domain and μ is the measure of the mass being transferred. We write

$$\inf_{s \in S} \int_U c(x\,,\,s(x))\; \mathrm{d}\mu(x) \;=\; \inf_{s \in S} C(s).\tag{1.13}$$

An optimiser $\hat{s}$ satisfies $C(\hat{s}) \;=\; \inf_{s \in S} C(s)$. The quantity $\inf_{s \in S} C(s)$ is itself of interest and can be thought of as measuring distance between the domain and co-domain and their measure of size. (See Wasserstein distance later.)

The interest in optimal mass transfer theory surrounds the cost function, in particular the choice of the cost function and the uniqueness of the solution of the cost function.

There are cost functions for which

(i) there are many minimisers (non-uniqueness);

(ii) there is a unique minimiser;

(iii) there is no minimiser; forward reference the reader to section 4.5.

Given the choice of cost function for Monge's original work problem we consider the five

book problem in section 1.5.1; this will illustrate the non-uniqueness of the solution for Monge's original cost function.

There are three major difficulties with Monge's problem. Firstly, the set S, defined above, is neither convex nor compact (this problem is discussed at the end of the Chapter 2). Secondly, Monge's problem is a nonlinear problem. The Monge-Kantorovich problem generalises to a linear problem but has more possible solutions. Given that the cost function is non-negative, Monge's original minimal work problem has a lower bound and thus an infimum exists; a minimiser need not exist. Therefore for specific cost functions with undesirable properties, e.g. not being smooth, a minimiser need not exist. A further difficulty of Monge's problem is that there aren't simple conditions on the cost function such that a minimiser exists.

The next big development in the area dealt with the lack of convexity and compactness. Kantorovich's problem also stated specific properties required from the cost function to yield a solution to the new problem. This work would be published by Kantorovich in 1942 [50]. Kantorovich [50] made such significant developments in the field that the problem became the Monge-Kantorovich problem. Gangbo [38] stated "We write the Monge-Kantorovich problem as a relaxation of the Monge problem". The relaxed problem implies that Kantorovich's infimum encompasses a larger class of objects than that of Monge.

Kantorovich's relaxed problem made the optimal mass transfer problem an easier one to solve. Given the link between the work, Kantorovich had provided a simpler method to obtain solutions for Monge's problem. Firstly, solutions would need to satisfy the Monge-Kantorovich problem then the solution would need to have specific properties.

Another advantage of the Monge-Kantorovich problem is the dual problem associ-

ated with it. Villani [71] noted that a linear minimization problem with convex constraints, such as the Monge-Kantorovich problem, has a dual formulation. Gangbo and McCann [39] state that duality "has been the principal tool for investigating the Monge-Kantorovich problem". The duality problem was utilised to obtain solutions to the Monge-Kantorovich problem; this was considered to be a simpler problem to obtain solutions for than the original Monge-Kantorovich problem.

In the 1980s, for the quadratic cost function, $c(x, y) = |x - y|^2/2$, several authors made significant advances in the field. Brenier [13], Knott and Smith [53] made developments about the uniqueness of minimisers for certain costs; the quadratic cost being the archetypal example. We will consider both Monge's cost function and the quadratic cost function in the next subsection. The quadratic cost function is an example where we have a unique minimiser. We forward reference the reader to section 4.3 for further details on the Monge-Kantorovich problem.

1.5.1 Five Book Problem.

The five book problem is an illustrative optimal mass transfer problem which demonstrates how different costs yield different outcomes.
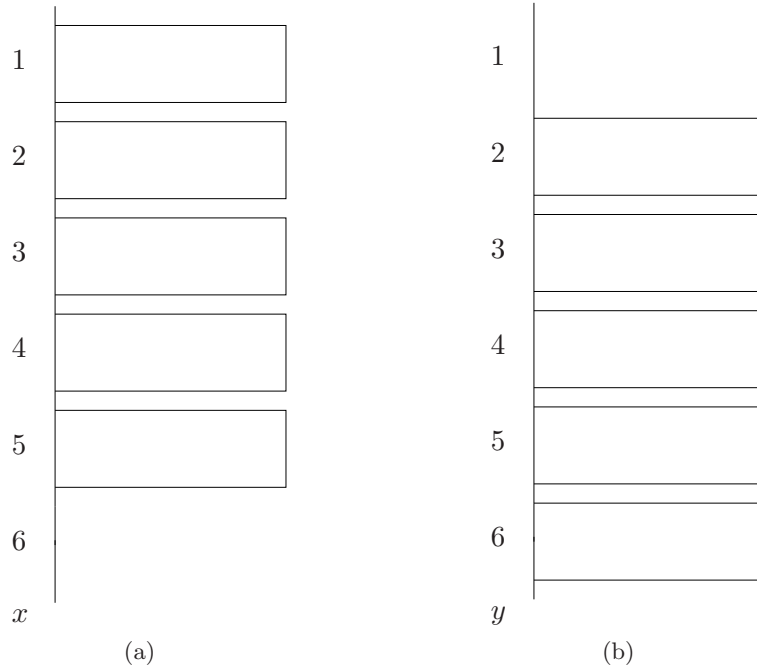
Figure 1.8: An illustration of the five book problem (a) original position of the five books and (b) the final position of the five books

The five book problem is a discrete problem. The optimal mass transfer problem as it has been considered thus far consists of an infinitesimal continuum of particles thus an extremiser need not exist. However, the discrete problem considered here has a finite number of strategies, therefore an extremiser must exist.

As the problem has been defined previously the pile of sand is the original position for the books; the final position of the books is the excavation which the sand must fill. The problem has six positions in which the books can lie, as numbered in figure 1.8; $U = \{1, 2, 3, 4, 5\}$ and $V = \{2, 3, 4, 5, 6\}$. The five books are identical and indistinguishable. We consider two options to transfer the books from U to V:

1. Move each book one place to the right(i.e. $1 \to 2$, $2 \to 3$, $\cdots$); denote this strategy 1.

2. Move the first book to the final position, i.e $1 \to 6$, with the positions of the other

books remaining unchanged; denote this strategy 2.

We have that $x \in U$ and that $y \in V$. In this discrete case the strategies are permutations and the integral in equation (1.13) is replaced by a finite sum, so $C_1(x, y) = \sum_{i=1}^{5} c_1(x, y)$. Each of the five movements from strategy one contributes one to the cost hence for the first strategy, $C_1(x, y) = 1 + 1 + 1 + 1 + 1 = 5$. For the second strategy we have one movement and again $C_2(x, y) = 5 + 0 + 0 + 0 + 0 = 5$. Whence for both strategies the cost function equals 5. In fact these strategies are minimisers. This illustrates the non-uniqueness of the solution for the cost function of Monge's original minimal work problem.

We consider the same strategies for the quadratic cost function. When moving five books one position there is a contribution of $c_1(x, y) = \frac{|1-2|^2}{2}$ from each book yielding $C_1(x, y) = \frac{5}{2}$. For the other strategy the quadratic cost function gives us $C_2(x, y) = \frac{25}{2}$. In fact the first strategy is the unique minimiser relative to the quadratic cost. This simple example illustrates the difficulty surrounding Monge's original minimal work problem and the more desirable qualities associated with the quadratic cost function.

1.5.2 The History of the Optimal Mass Transfer Problem.

The first individual to extend Monge's work was Charles Dupin [33]. Following the extensions of Dupin a prize was offered by the Academy of Paris in 1884 [28]. This prize was taken a few years later by Appell [5] in 1887. Rachev and Ruschendorf [62] stated that the prize was claimed by Appell as he "established some geometric properties of optimal maps in the plane and in $\mathbb{R}^3$". However, it was apparent that the problem was an involved one and neither Dupin [33] or Appell [5] did anything to aid the understanding of (1.13). Villani [72] noted that the arguments used, by current mathematical standards,

24

were flawed. However, Appell made attempts to simplify the work [6][1].

Just over twenty years later came the biggest development since the statement of the original problem. Kantorovich in the 1940's, [50; 51], had such an impact on the problem that it is now known as the Monge-Kantorovich problem. Moreover, Kantorovich along with T.C. Koopmans, received the Nobel Prize in Economics, [66]. One important point that Rachev and Ruschendorf noted [62] was that when Kantorovich presented his first paper on optimal mass transfer he did so completely unaware of Monge's work of the eighteenth century. Kantorovich only learnt of Monge's work later, and linked his work to that of Monge [59] in his second paper, [51].

The next advance in this problem was presented by Sudakov [67]. In this paper, which was almost 200 years after the original statement of the problem, Sudakov claimed to have shown that a solution existed for Monge's original minimal work problem. From previous work it was apparent that this map wasn't to be unique, given the issue of non-linearisation. But, while the outline of the proof was there, there were significant gaps in the proof as presented by Sudakov [67], especially when $n \geq 3$.

Ambrosio [2] dealt with the gaps in the proof of Sudakov and finally proved that an optimal map existed for the Monge original minimal work problem. While Ambrosio was proving the problem outlined by Sudakov, many researchers [24; 35; 70] were proving that an optimal map existed using alternative methods.

The progress in solving optimal mass transfer problems had been slow for several years, but, towards the end of the twentieth century there was a surge of interest in the subject as researchers realised the importance of optimal mass transfer to their own subject. Villani [71] gave the "birthdate for this revival" as Yann Brenier's paper, [13],

---

[1]For a review of Appell's work [35]

with Brenier bringing the subject to the attention of many around him. Brenier's paper ran parallel with another group of authors who used a more probabilistic approach, [53, Theorem 2.1] to prove that for the quadratic cost function that a map existed and it was unique.

1.6   Notation.

Notation.	Description.	First Appears.
q_1	Actual rainfall.	Chapter 1.
q_2	Forecasted rainfall.	Chapter 1.
$\mathcal{R}(f)$	Set of rearrangements of f.	Chapter 1.
$c = c(x, y)$	Cost function associated with the optimal mass transfer problem.	Chapter 1
S	Set of measure-preserving mappings.	Chapter 1.
$s_{\#}\mu$	Push forward of μ through s.	Chapter 1.
λ_n	n-dimensional Lebesgue measure.	Chapter 2.
B	Borel Set.	Chapter 2.
$1_A(x)$	Characteristic/Indicator function of the set A.	Chapter 2.
$f^*(x)$	Monotone increasing rearrangement of f, for f a scalar valued function.	Chapter 2.
$f^{\Delta}(x)$	Monotone decreasing rearrangement of f, for f a scalar valued function.	Chapter 2.
$f^{\#}(x)$	Monotone rearrangement of f, where f is a vector valued function.	Chapter 2.
$\overline{\text{conv}\mathcal{R}(f)}$	Closed convex hull of the set of rearrangements of f.	Chapter 2.

One final point to note is the Theorem numbering system. Results used from literature are numbered in the the following way Theorem 3.2.1; which is the first theorem of section 3.2. Results proven in the thesis have a single number, e.g. Nealers Theorem, which is the first result of this thesis.

1.7 Outline of Thesis.

The thesis will be laid out in the following way. In Chapter 2 we define several equivalent definitions for rearrangements of functions; this is done for both the scalar and vector valued cases in sections 2.2 and 2.3. A specific type of rearrangement, the monotone rearrangement, is defined for both the vector and scalar valued cases in section 2.4. Monotone rearrangements are used extensively when considering rearrangement inequalities. The section on rearrangement inequalities is motivated by section 2.5.1 where we consider an equivalent extremisation problem (i.e. same set of extremisers) to that in (1.2). Then we state the rearrangement inequalities which will be used to prove our main Theorem in Chapter 3.

Another notion which will be extensively used will be measure-preserving mappings, therefore we illustrate its importance noting its relation to rearrangements and push-forward measures. We use measure-preserving mappings to define the concept of polar factorisation (see Ryff [65, Proposition 3]); extending the notion for vector valued rearrangements (see Brenier [13; 14]. We finish by considering the problem of extremising over sets of rearrangements, giving special consideration to the idea of convexity.

The way in which the simple formulation was defined in (1.3) leaves some questions unanswered, particularly the classification of the set of best fitting rearrangements. We answer this question in Chapter 3. The chapter begins by using the methods of

rearrangement inequalities from Burton [17] to prove that a minimiser exists in (1.2) and establishes exact conditions for when the minimiser is unique. The remainder of the chapter establishes when the minimiser is unique. Level sets having positive measure play a major role in these arguments. The concept of level sets is explained and illustrated and the problems surrounding level sets are explained. The main theorem in the thesis is then explained via an illustrative example before breaking the problem down into a series of propositions and lemmas brought together to prove the theorem.

Having briefly introduced the problem in the first chapter the idea of optimal mass transfer is discussed in greater depth in Chapter 4. We begin by introducing the methods which are central to optimal mass transfer theory when considering the Monge-Kantorovich problem. Having defined the methods of optimal mass transfer we consider the problems addressed by Kantorovich's formulation. We then define formally both the Monge-Kantorovich problem and its dual problem. We conclude the chapter by surveying the literature to summarise the results which exist for different cost functions.

For the majority of the thesis we consider error scores for scalar valued rearrangements; it is a natural progression to consider vector valued rearrangements problem as well. In section 5.2 we illustrate that the closest rearrangements problem need not be satisfied for vector valued rearrangements.

In the introduction we alluded to some problems which might arise from our forecast error decomposition. The remainder of Chapter 5 is how we deal with these problems. We begin the chapter by defining the displacement error; we use results from Brenier and Benamou [11]. Further work by Brenier and Benamou [10] contains a possible solution for the problem; they suggest a specific cost function. This cost function is for a time continuous formulation and thus we need to adapt this idea. We state the problem

as a rearrangements problem ensuring that both the qualitative features error and the displacement error are taken at the same time. This rearrangements problem can also be formulated as a partial differential equations problem or an optimal mass transfer problem. Through the work of Haker [73] we consider some of the limitations of the partial differential equations problem. The more sophisticated formulation, in the same manner as the closest rearrangements problem of the simple formulation, has an equivalent formulation; this is the content of Theorem 8. The more sophisticated formulation should have fewer extremisers, Proposition 9 shows some of these extremisers.

With a view to illustrating an extremiser exists for the more sophisticated formulation we also consider a relaxed version of this formulation. The relaxed formulation is considered over the closed convex hull as opposed to the set of rearrangements. Theorem 10 shows the equivalent maximisation problem. Section 5.5 shows via an example the differences between the simple formulation and the more sophisticated formulation.

Chapter 5 is drawn to a close by showing that the relaxed more sophisticated formulation has an extremiser; this result is contained in Theorem 11. The next result of Theorem 11, (iii), shows that for specific values of θ the value of the more sophisticated formulation is larger than that of the value of the relaxed more sophisticated formulation. This means that we cannot generalise the relaxed formulation result to the more sophisticated version. The final result of the Theorem shows that the equivalent maximisation problems are the formulations to consider. An example showing specific values of θ where Theorem 11(iii) is satisfied concludes the chapter.

We summarise the achievements of the thesis in the final chapter, and discuss future directions for forecast error decomposition. One future direction is the numerical implementation of the ideas of the thesis. We finish the thesis with a discussion of the

open problems of the thesis.

1.8   Applications of rearrangements and optimal mass trans-

fer.

The semigeostrophic equations are a model for slowly varying flows which are constrained by both rotation and stratification; this system has been used to model weather fronts forming and moving. The f-plane (constant rotation) semigeostrophic equations are as follows:

$$\frac{D\mathbf{X}}{Dt} = fJ(\mathbf{X} - \mathbf{x}), \tag{1.14}$$

$$\mathbf{X} = \nabla P, \tag{1.15}$$

$$\nabla \cdot \mathbf{u} = 0, \tag{1.16}$$

where

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

There are some things to note: $\mathbf{X} = (X, Y, Z)$, $\mathbf{u} = (u, v, w)$ is the velocity field, $\frac{D}{Dt}$ is the Lagrangian derivative, P is a pressure variable and f is the Coriolis parameter, assumed constant in this simplified model. Equations (1.14), (1.15) and (1.16) are solved

on an open, bounded and connected $\Omega \subset \mathbb{R}^3$ for the boundary condition $\mathbf{u} \cdot \mathbf{n} = 0$ on $\delta\Omega$. We solve the set of equations for $\mathbf{X}$, $\mathbf{u}$ and P.

An extra constraint on these equations is the Cullen-Norbury-Purser minimum energy principle(see Cullen [27]). At each time t values of $\mathbf{X}(t, \cdot)$ are considered to be fixed on particles. The principle states that at each time particles arrange themselves to minimise the geostrophic energy $E = E(\mathbf{X})$. This yields a vector-valued rearrangements minimisation problem:

find those $\mathbf{X}$ attaining

$$\inf_{\mathbf{X} \in \mathcal{R}(\mathbf{X}_t)} E(\mathbf{X}). \tag{1.17}$$

It was shown by Douglas [32] that the infimum is uniquely attained by $\mathbf{X}_t^{\#} = \nabla\psi_t$, where ψ_t is a convex function. (This is an example of a Brenier map, a concept that will be defined later.) We identify $P(t, \cdot)$ with ψ_t at each time t. Tracking the singularities of $P(t, \cdot)$ as time evolves is thought of as tracking weather fronts forming and moving. Convex functions (on $\mathbb{R}^3$) fail to be differentiable on a set of dimension at most 2 which corresponds with the idea that a weather front is "thin".

There is a second rearrangements problem associated with this system. When the mapping $\mathbf{X}(t, \cdot)$ is invertible (almost everywhere) the equations can be rewritten in $\mathbf{X}$ variables as a coupled Monge-Ampère/transport equation system (See for example, Douglas [30] for details). There is a rearrangement conserved quantity and an associated scalar-valued rearrangement minimisation problem. Cullen and Douglas [26] have demonstrated some results for nonlinearly stable flows in a simplified case.

An example related to the second rearrangements problem described above is that of energy extremisation for 2-D ideal fluid flow; the theory is much more developed in this case.

The vorticity/stream function formulation of a two dimensional incompressible inviscid flow

$$-\Delta\psi = \omega \text{ in } \quad \Omega \quad , \quad \psi = 0 \text{ on } \partial\Omega, \tag{1.18}$$

$$\frac{\mathrm{D}\omega}{\mathrm{D}t} = 0, \tag{1.19}$$

$$\mathbf{u} = \left(\frac{\partial\psi}{\partial y}, -\frac{\partial\psi}{\partial x}\right), \tag{1.20}$$

where Ω is a planar region bounded by a simple closed curve $\partial\Omega$, ψ denotes the stream function, ω the vorticity, $\mathbf{u}$ the velocity, and $\mathrm{D}\left(\cdot\right)/\mathrm{D}t = \partial\left(\cdot\right)/\partial t + \mathbf{u}\,.\,\nabla\left(\cdot\right)$, the time derivative following the flow. For a given square integrable function ω we can find (a unique) ψ which satisfies (1.18) in the weak sense; call this "inverse for $-\Delta$ with zero Dirichlet boundary conditions" $K\omega$. Kinetic energy E satisfies

$$E = \frac{1}{2}\int_{\Omega}|\mathbf{u}|^2 \,\mathrm{d}\lambda_2 = \frac{1}{2}\int_{\Omega}|\nabla\psi|^2 \,\mathrm{d}\lambda_2 = \frac{1}{2}\int_{\Omega}\omega K\omega \,\mathrm{d}\lambda_2, \tag{1.21}$$

and from (1.19) and (1.20) we have that vorticity is preserved on particles in an incompressible flow, that is vorticity is rearrangement preserved. We maximise kinetic energy over a family of flows whose vorticities are rearrangements of a prescribed function; there is a principle that a maximiser should yield a steady stable flow, the fluid being

"trapped" in the top energy state. The origins of this idea can be found in Kelvin [69]; the formulation in terms of rearrangements of functions was proposed by Benjamin [12] in the context of steady vortex rings in three-dimensional axisymmetric ideal fluid flow.

This principle has been justified to some extent. Suppose that $\omega_0 \in L^2(\Omega)$. Then Burton [17; 19] yields that there exists $\bar{\omega} \in \mathcal{R}(\omega_0)$ such that $E(\bar{\omega}) \geq E(\omega)$ for each $\omega \in \mathcal{R}(\omega_0)$ and

$$-\Delta\bar{\psi} = \phi \circ \bar{\psi}, \tag{1.22}$$

for some increasing function ϕ, where $\bar{\psi} = K\bar{\omega}$. Now (1.22) is the equation for the stream function of a steady flow. For non-negative $\omega_0 \in L^2(\Omega)$ solutions to (1.18)-(1.20) with $\omega(0, \cdot) = \omega_0$ conserves energy. Suppose that $\bar{\omega}$ is the unique maximiser of E relative to $\mathcal{R}(\omega_0)$. Then Burton [20] demonstrates that the solution associated with $\omega(t, \cdot) = \bar{\omega}$ for all time t is nonlinearly stable, that is for every $\epsilon > 0$ we can choose $\delta > 0$ such that for any solution of (1.18), (1.19) and (1.20) with $\|\omega(0, \cdot) - \bar{\omega}\|_2 < \delta$ we have $\|\omega(t, \cdot) - \bar{\omega}\|_2 < \epsilon$ for all $t > 0$.

Practical applications of optimal mass transfer can be found in Benamou and Brenier [11]. This states we can use the Monge-Kantorovich theory in "various fields of applications such as shape recognition in image processing, computer vision and signal treatment, data assimilation in meteorology and oceanography, quantum chemistry, etc", see [40; 48; 49]. Geman and Geman [41] was an illustration of this. The paper uses the Gibbs Sampler method to restore images such as photographs which are blurred or include nonlinear deformations. We have a photo as it should appear, the photo as it is

and the photo after it has been restored. From our work on forecast error decomposition the forecasted data is the photo as it should appear; the actual data is the degraded image and the best fit is the restored image. The problem has also be studied by Haker et al. [44; 73]. We discuss their results in Chapters 4, 5 and 6.

A further use of optimal mass transfer is shape recognition as considered by Gangbo and McCann [40]. The idea is based on an algorithm suggested by Fry [36] which uses computers to identify leaves from New England using the outline of the leaf. The algorithm compares the outline of the unknown leaf to several known leaves. In the same way that our forecast error decomposition used the best fit this algorithm uses the best fit to decide which well known leaf the original outline is.

Plakhov [60; 61] used the Monge-Kantorovich problem to examine the problem of how air passed over convex and concave bodies. Of interest in this problem was a cost function which included both a convex and concave function. Furthermore Plakhov found, surprisingly, better results for concave bodies and this led to the study of the way particles interact with golf balls. We forward reference the reader to section 4.5 for the formal outline of the problem.

Another application is probability and statistics, more specifically game theory and hypothesis testing which are reviewed by Rachev and Rüschendorf [63]. From Kantorovich's work an obvious application would be mathematical economics [55]. There are too many applications to mention them all and to this end we refer the reader to Rachev and Rüschendorf [62, Preface] where an extensive list is provided.

Chapter 2

Rearrangements.

2.1   Introduction.

The idea of rearrangements of functions is central to this thesis. Section 2.2 formalises the ideas of section 1.4. Section 2.2.1 states four formulations and shows they are equivalent, drawing an analogy to Casimir integrals or conserved quantities for dynamical systems. Measure spaces of infinite size introduce new complications; we discuss the definition of rearrangement with reference to a specific example.

Vector-valued rearrangements are relevant to optimal mass transfer: the content of section 2.3. We begin by noting by formalising the idea of vector valued rearrangements in definition 2.3.1. Then we note the differences between the vector valued and scalar valued case. As was the case with the scalar valued case there are several equivalent definitions for vector valued rearrangements. In section 2.3.1 we state the definitions, showing they are equivalent, comparing them to the scalar valued versions.

Special types of rearrangements give rise to important inequalities; we discuss these in section 2.4. We consider scalar valued monotone rearrangements linking these rear-

rangements to the functions considered in section 1.4. From the methods of, for example Douglas [30], we illustrate how to construct a decreasing or increasing rearrangement. Then we consider monotone rearrangements of vector valued functions.

Rearrangements inequalities involving monotone rearrangements have extensive use throughout the thesis. We begin section 2.5 by considering an equivalent problem to the closest rearrangements problem. Following the methods of Burton [17] we state the rearrangement inequalities in question.

A measure interval is a measure space with the same measure theoretic structure as an interval on the real line with Lebesgue measure; we define this in section 2.6. In section 2.6.1 we define the concept of push-forward measures in definition 2.6.5 before comparing measure-preserving mappings, rearrangements and push forward measures. Increasing, or decreasing, scalar valued rearrangements along with measure-preserving mappings will give rise to a notion known as polar factorisation; we consider this in section 2.6.2. We conclude section 2.6.2 by extending the notion of polar factorisation from scalar valued functions to vector valued functions.

When extremising a functional over a set of rearrangements the non-convexity and non-compactness of the set of rearrangements can be problematic; this is discussed in section 2.7. In section 2.7.1 we recall the idea of convexity. We define a superset of the set of rearrangements in 2.7.2 which is (weakly sequentially) compact. We finish the chapter by considering another theorem of Burton [17, Theorem1] which will be central in the proof of the main theorem of the thesis.

2.2   Scalar Valued Rearrangements.

2.2.1   Equivalent definitions of scalar valued rearrangements.

There are alternative definitions for the rearrangement of a scalar valued function. We consider four such definitions used by different authors (see, for example [30, Theorem 1]).

Theorem 2.2.1. *Let* $\Omega \subset \mathbb{R}^n$ *be a bounded set and let* μ *be a measure, which is absolutely continuous with respect to n-dimensional Lebesgue measure.*

Let $f, g : \Omega \to \mathbb{R}$ *be integrable functions. The following are equivalent:*

  (i) f is a rearrangement of g(as in definition 1.8);

  (ii) For every real α,

$$\mu\left(\{x : f(x) \geq \alpha\}\right) = \mu\left(\{x : g(x) \geq \alpha\}\right). \tag{2.1}$$

(iii) For every Borel set $B \subset \mathbb{R}$,

$$\mu\left(\{x : f(x) \in B\}\right) = \mu\left(\{x : g(x) \in B\}\right). \tag{2.2}$$

(iv) For $F \in C(\mathbb{R})$ with $|F(y)| \leq const(1 + |y|^p)$ for some $1 \leq p < \infty$,

$$\int_\Omega F(f(x))\, d\mu(x) = \int_\Omega F(g(x))\, d\mu(x). \tag{2.3}$$

(Equation (2.3) holds in the sense that if one of the integrals is finite, so is the other and they are equal.)

Proof: Follows by the methods of, for example Douglas [32, Section 2.4]. □

The role of the growth constraint on F, $|F(y)| \leq const(1 + |y|^p)$ in definition (iv), is to ensure that the integrals are finite when f, $g \in L^p$. Generally, (2.3) will also hold for an F which is the (pointwise) limit of continuous functions. One such type of function is the characteristic or the indicator function which is defined as follows:

Definition 2.2.2. *For $A \subset X$, $1_A : X \to \mathbb{R}$, the indicator function of A is defined by*

$$1_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A. \end{cases} \tag{2.4}$$

The four equivalent statements, for different classes of functions F can be written in the form of (iv). For instance, to obtain (1.8) we choose $F(y) = (y - \alpha)_+$. We choose $F(y) = 1_B$ to give (2.2), for $B \subset \mathbb{R}$ where B is a Borel set. Moreover, by choosing $F = 1_{[\alpha, \infty)}$ we have (2.1).

We note that (2.3) reduces to (1.11) for the specific choice of $F(y) = |y|^p$ which illustrates that rearrangements preserve the L^p-norm.

Let ω be a rearrangement conserved quantity, such as vorticity in 2-dimensional ideal fluid flow. We can express definition (iv) from Theorem (2.2.1) in the following way, that is the infinite family of Casimir integrals:

$$C_F \;=\; \int_\Omega F\left(\omega\right)\; \mathrm{d}\lambda_2, \tag{2.5}$$

are conserved for all $\omega \in \mathcal{R}\left(\omega_0\right)$, where F is an arbitrary function that can be approximated by continuous functions.

Rearrangements of functions still makes sense on measure spaces with infinite size. However, the properties in Theorem 2.2.1 are no longer equivalent. We consider the following example(from Douglas [29]):
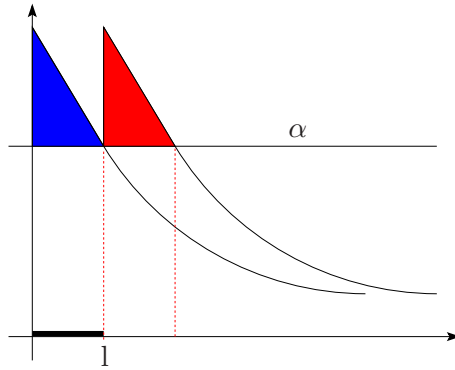
Figure 2.1: The same function separated by a block of zero's.

From figure 2.1, if we denote the function with the blue area by f then the function with the red area, denoted g, is defined as follows:

$$g(x) = \begin{cases} 0, & \text{on } [0\,,l]\,, \\[2mm] f\,(x-l)\,, & \text{on } (l\,,\infty)\,. \end{cases} \tag{2.6}$$

The specific counterexample $B = \{0\}$ does not satisfy Theorem 2.2.1 (iii) but satisfies definition 1.4.1. With $F = 1_B$ for this choice of Borel set has that Theorem 2.2.1 (iv) is not satisfied; Theorem 2.2.1 (i) and (ii) are both defined. We consider the following definition from Douglas [30] which states when function are rearrangements on an infinite measure space:

Definition 2.2.3. *Let Ω be an unbounded subset of $\mathbb{R}^n$ of infinite size, that is $\lambda_n(\Omega) = \infty$ where λ_n denotes n-dimensional Lebesgue measure. Two non-negative integrable functions $f\,,g\,:\,\Omega \rightarrow \mathbb{R}$ are rearrangements if*

$$\lambda_n\left(\{x\,:\,f\,(x) \geq \alpha\}\right) = \lambda_n\left(\{x\,:\,g\,(x) \geq \alpha\}\right), \tag{2.7}$$

for every $\alpha > 0$.

An assumption we should make about the functions we are rearranging is that they vanish at infinity, i.e. that $\lambda_n\left(\{x\,:\,f\,(x) \geq \alpha\}\right) < \infty$ for each $\alpha > 0$, see Lieb and Loss for more information [56, Section 3.2].

2.3 Vector Valued Rearrangements.

Definition 2.3.1. *Let Ω be a bounded set in $\mathbb{R}^n$, let μ be absolutely continuous with respect to n-dimensional Lebesgue measure, and let f , g : $\Omega \to \mathbb{R}^d$ be integrable functions. f and g are rearrangements if*

$$\mu\left(\{x \, : \, f\left(x\right) \in B\}\right) \, = \, \mu\left(\{x \, : \, g\left(x\right) \in B\}\right), \tag{2.8}$$

for every Borel subset B of $\mathbb{R}^d$.

Definition 2.3.1 is a vector valued extension of the Theorem 2.2.1 (iii) with $d = 1$.

A vector valued function can be defined in terms of its components, we use these components to illustrate the differences between the vector and scalar valued cases. The components of a vector valued function, f : $\Omega \to \mathbb{R}^d$, are f_i : $\Omega \to \mathbb{R}$ where $f_i = (f_1 , \dots , f_d)$. From the definition of vector valued rearrangements we have the following:

$$f \in \mathcal{R}(g) \Rightarrow f_i \in \mathcal{R}(g_i). \tag{2.9}$$

Hence, for vector valued rearrangements we require the components of the vector valued functions also to be scalar valued rearrangements. However, the converse to (2.9) is not always true, for a specific counterexample see Douglas [30].

2.3.1 Equivalent definitions of vector valued rearrangements.

Theorem 2.3.2. *Let $\Omega \subset \mathbb{R}^n$ be a bounded set and let μ be absolutely continuous with respect to n-dimensional Lebesgue measure. Let $f, g : \Omega \to \mathbb{R}^d$ be integrable functions. The following are equivalent*

(i) f is a rearrangement of g.

(ii) for each $F \in C\left(\mathbb{R}^d\right)$,

$$\int_\Omega F(f(x))\,\mathrm{d}\mu(x) = \int_\Omega F(g(x))\,\mathrm{d}\mu(x), \tag{2.10}$$

such that $|F\left(\xi\right)| \leq K\left(1 + |\xi|^p\right)$, where each K is a constant.

(Equation (2.10) holds when one of the integrals is finite.)

(iii) For each $c \in \mathbb{R}^d$,

$$\mu\left(\{x \,:\, f(x) \geq c\}\right) = \mu\left(\{x \,:\, g(x) \geq c\}\right). \tag{2.11}$$

where the inequalities are calculated component by component on $\mathbb{R}^d$ (Partial ordering).

(iv) For each $\sigma \in \mathbb{R}^d$, $\alpha > 0$,

$$\int_\Omega \left(|f - \sigma|_\infty - \alpha \right)_+ \, \mathrm{d}\mu = \int_\Omega \left(|g - \sigma|_\infty - \alpha \right)_+ \, \mathrm{d}\mu, \tag{2.12}$$

where $|\cdot|_\infty$ denotes the infinity norm on $\mathbb{R}^d$.

Proof: Follow the methods of Douglas [32, Theorem4]. $\square$

We note that f being a rearrangement of g in the form of (i) is the same as definition (2.8).

Definition (ii) can be thought of in the sense that if one of the integrals is finite then the other integral is also finite. To this end, Brenier [14] restricted the integral to a subclass of $C(\mathbb{R}^d)$ which guarantees that the integrals in equation (2.10) are finite, moreover, if f, $g \in L^p$ then the integrals in (2.10) are finite. Definition (ii) is a vector valued extension of definition (iv) from theorem 2.2.1. As was the case with the scalar valued case this definition can unify the other definitions.

The inequalities in definition (iii) were introduced by Cullen, Norbury and Purser [27] and is a direct extension of definition (ii) in Theorem 2.2.1. Finally, property (iv) is an extension of definition 1.4.1.

2.4 Monotone Rearrangements.

Recalling the functions defined on the unit interval in the opening chapter, $f(x) = x$, $g(x) = |1 - 2x|$ and $h(x) = 1 - x$, where f, g and h are rearrangements of each other. Both f and g are special types of rearrangements which is the content of this section.

From figure 1.5(a), f is a monotonic increasing function and h is a monotonic de-

creasing function (figure 1.6). It is of interest whether f is the unique increasing function, which is a rearrangement of g and whether h the unique decreasing function, which is a rearrangement of g. From the methods of Burton [17, Lemma 1] we have that both the increasing and decreasing rearrangements are essentially unique. They are unique except possibly at points of discontinuity. There are only countably many of these for a monotonic function. Furthermore, we use the methods of Burton [17, Lemma 1] to define the monotone rearrangements for functions defined on a general measure space (Ω, μ).

From any function $f_0 : \Omega \subset \mathbb{R}^n \to \mathbb{R}$, where (Ω, μ) is a measure space satisfying $\mu(\Omega) = 1$ for μ absolutely continuous with respect to Lebesgue measure we can construct a monotone increasing rearrangement. We can construct the increasing function f_0^* on $[0, 1]$ by replacing every set $\{x : f_0(x) \geq \alpha\}$ by $[1 - \lambda_1(\{x : f_0 \geq \alpha\}), 1]$; hence defining the value of $f_0^*(\alpha)$ on an interval of the same size extending leftwards from 1.

2.4.1 Monotone rearrangements of vector valued functions.

We can extend the notion of montone rearrangement to the vector valued functions. Unfortunately, $\mathbb{R}^d$ is not totally well-ordered for $d \geq 2$ thus the notion of increasing does not make sense. However, we have an increasing scalar valued function which is the gradient of a convex function; an idea we can generalise to the vector valued case. Brenier [14, Theorem 1.1] defined the monotone rearrangement for a vector valued function f to be $f^{\#} = \nabla \Psi$, where Ψ is convex and is known in some literature as the Brenier map. He demonstrated its existence and (essential) uniqueness. The Brenier map has been utilised in optimal mass transfer theory. We illustrate with the following example. Let f, defined on the unit square, be defined as follows:

$$
f(\mathbf{x} = (x, y)) = \begin{cases} (1, 1), & \text{if } x \in \left[\frac{1}{2}, 1\right] \times \left[\frac{1}{2}, 1\right], \\ \\ (0, 0), & \text{otherwise.} \end{cases} \tag{2.13}
$$

This is illustrated by the following diagram:
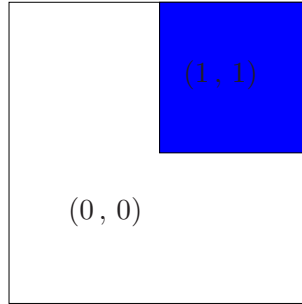
Figure 2.2: The function f as defined in (2.13)

We define the the monotone rearrangement to be $f^{\#} : [0, 1]^2 \to \mathbb{R}^2$:

$$
f^{\#}(\mathbf{x} = (x, y)) = \begin{cases} (1, 1), & \text{if } x + y \geq 2 - \frac{1}{\sqrt{2}}, \\ \\ (0, 0), & \text{otherwise.} \end{cases} \tag{2.14}
$$

The convex function Ψ is:

$$
\Psi(\mathbf{x} = (x, y)) = \begin{cases} x + y - \left(2 - \frac{1}{\sqrt{2}}\right), & \text{if } x + y \geq 2 - \frac{1}{\sqrt{2}}, \\ \\ 0, & \text{otherwise.} \end{cases} \tag{2.15}
$$

We illustrate the monotone rearrangement in the following figure:
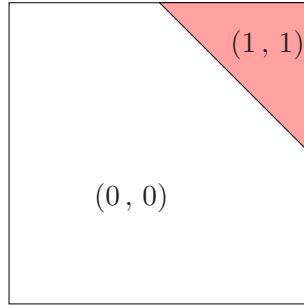
Figure 2.3: The function $f^{\#}$, an example of a Brenier map

We can see easily from figures 2.3 and 2.2 that $f^{\#} \in \mathcal{R}(f)$. Recall that a convex function f is said to be proper if its epigraph(forward reference to section 2.7.1 for definition) is non-empty and contains no vertical lines, i.e. if $f(x) < +\infty$ for at least one x and $f(x) > -\infty$ for every x. (See, for example, Rockafellar [64].) We define lower and upper semi continuity:

Definition 2.4.1. *A function $f(X)$, defined in a metric space X, is said to be lower semi continuous if for any point $x \in X$*

$$\liminf_{y \to x} f(y) \geq f(x). \qquad (2.16)$$

Definition 2.4.2. *A function $f(X)$, defined in a metric space X, is said to be upper semi continuous if for any point $x \in X$*

$$\limsup_{y \to x} f(y) \leq f(x). \qquad (2.17)$$

While we do not formally define the idea of the measure-interval until later we consider

the following until the definition: some measure theoretic properties as an interval of

the same size equipped with one-dimensional Lebesgue measure. We define monotone

rearrangements for vector valued function formally from Douglas [31]

Definition 2.4.3. *Let* $u \in L^1(X, \mu, \mathbb{R}^n)$, *where* (X, μ) *is a measure-interval. Let*
Lebesgue measurable $Y \subset \mathbb{R}^n$ *be such that* $\lambda_n(Y) = \mu(X)$. *The monotone rearrange-*
ment of u *on* Y *is the unique function* $u^{\#} : Y \to \mathbb{R}^n$ *that is a rearrangement of* u *and*
satisfies $u^{\#} = \nabla \psi$ *almost everywhere in* Y *for some proper lower semi continuous*
convex function $\psi : \mathbb{R}^n \to \overline{\mathbb{R}}$.

In the above setting uniqueness almost everywhere of the monotone rearrangement (or

Brenier map) follows from the main theorem of McCann [57]. The convex potential Ψ

is not unique; if Y is connected, Ψ is defined up to additive constant(see Gangbo [37]).

2.5 Rearrangement inequalities.

2.5.1 Closest rearrangements - an equivalent problem.

When we expand the square of the 2-norm in (1.2) we have

$$
\begin{aligned}
\|q_1 - q\|_2^2 &= \int (q_1 - q)^2 \\
&= \int q_1^2 - 2 \int q_1 q + \int q^2.
\end{aligned} \tag{2.18}
$$

q_1 is a constant hence the first term of (2.18) is a constant. Also, given that $\|q\|_2 = \|q_2\|_2$

for every $q \in \mathcal{R}(q_2)$ the final term is also a constant. We are left with the second term

48

and we minimise (1.2). Given the negative sign in (2.18) we are now maximising,

$$\sup_{q \in \mathcal{R}(q_2)} \int_\Omega q \, q_1 \, \mathrm{d}\mu. \tag{2.19}$$

The problem is equivalent in the sense that the minimisers of (1.2) are the maximisers of (2.19) and vice versa.

2.5.2 Maximising pairs of rearrangements.

We consider $f \in \mathcal{R}(f_0)$ and $g \in \mathcal{R}(g_0)$. We pose the question:

Which pairs of functions maximise $\int fg$?

The answer to this question includes the corresponding pairs of monotonic rearrangements. What happens is that the monotonic rearrangement pairs arrange themselves so that the larger values of f_0 and the larger values of g_0 are multiplying each other; similarly the smaller values of f_0 and g_0 arrange themselves so they are multiplying each other. In other words

$$\int_\Omega fg d\mu \le \int_0^{\mu(\Omega)} f_0^* g_0^* d\lambda_1 = \int_0^{\mu(\Omega)} f_0^\Delta g_0^\Delta d\lambda_1, \tag{2.20}$$

where $\Omega \subset \mathbb{R}$ and μ follow above definitions.

Lemma 2.5.1. *Let (Ω, μ) be a finite measure space, where λ_1 is a one-dimensional Lebesgue measure defined as before. Let $g : \Omega \to \mathbb{R}$ be a square integrable function and $U \subset \Omega$ and let $\alpha = \mu(U)$. Then*

$$\int_U g \, d\mu \leq \int_0^\alpha g^\Delta \, d\lambda_1(x). \tag{2.21}$$

Proof: Burton [17, Lemma 2]. □

Theorem 2.5.2. *Let (Ω, μ) be a finite measure space, where λ_1 is a one-dimensional Lebesgue measure defined as before. Let $f_0, g_0 : \Omega \to \mathbb{R}$ be square integrable functions and let $\omega = \mu(\Omega)$. Then for all rearrangements f of f_0 and g of g_0 on Ω we have*

$$\int_\Omega fg \, d\mu \leq \int_0^\omega f_0^\Delta g_0^\Delta \, d\lambda(x).$$

Proof: Burton [17, Theorem 1]. □

The fact that the L^2-norm is preserved under rearrangement leads to many researchers studying equality in Hardy-Littlewood type inequalities; that is for which $q \in \mathcal{R}(q_2)$ does $\int q_1 q = \int q_1^\Delta q_2^\Delta$? Crandall [25] stated (2.20) in the general case, where J incorporates functions such as the L^2-norm:

$$\int_0^{\mu(\Omega)} J\left(f^\Delta(x) - g^\Delta(x)\right) \, d\lambda(x) \leq \int_\Omega J\left(f(x) - g(x)\right) \, d\mu(x), \tag{2.22}$$

where $J : \mathbb{R} \to \mathbb{R}$ is a non-negative convex function satisfying $J(0) = 0$. Burchard and Hajaiej [16; 43] make use of (2.22) with various applications in mathematics, see Hajaiej [43]. There are parallels between the work in this thesis and the work presented by Burchard and Hajaiej. The problems are to find pairs of functions f and g which achieve equality in (2.22). Moreover, results exist when f is fixed to be f^Δ and again seek pairs which achieve equality in (2.22).

2.6 Measure-Preserving Mappings.

For an incompressible planar flow in a bounded domain the trajectory mapping of the fluid particles at each time t is area-preserving. For a subset of the plane the set of points that were mapped to that set has the same area. We recall the definition of inverse image [68]

Definition 2.6.1. *Let $f : X \to Y$ be any map and let C be a subset of Y. The inverse image $f^{-1}(C)$ of C under f is the subset of X given by $\{x \in X : f(x) \in C\}$.*

We define measure-preserving mappings and invertible measure-preserving mappings which map measurable sets to measurable sets; we call the latter measure-preserving transformations, [32, Section 2.2]:

Definition 2.6.2. *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 $A \subset V$, $\mu(s^{-1}(A)) = \nu(A)$.*

Moreover, $s : X \to Y$ is a measure-preserving transformation if

(i) $s : X \setminus L \to Y \setminus M$ is a bijection, where L and M are some sets of zero (respectively μ and ν) measure; and

(ii) s and s^{-1} are measure-preserving mappings.

Under mild hypotheses measure-preserving mappings are surjective, but need not be smooth nor injective. We give an example of the latter in the next section. Before defining a measure interval we define the isomorphism of measure spaces:

Definition 2.6.3. *Two finite measure spaces (U, μ) and (V, ν) are isomorphic if there exists a measure-preserving transformation $T : U \to V$.*

We will state our results on measure intervals, although some could be generalised to finite non-atomic measure spaces. The definition follows the methods of Burton and Douglas [21]:

Definition 2.6.4. *A finite measure space (U, μ) is a measure interval if it is isomorphic to $[0, \mu(U)]$ with Lebesgue measure.*

We have an equivalent definition to definition 2.6.2 from Halmos [45] which states that $s : U \to V$ (where (U, μ) and (V, ν) are measure spaces) is a measure-preserving mapping if:

$$\int_U f \circ s \, d\mu = \int_V f \, d\nu, \tag{2.23}$$

for every ν-integrable function f.[1] We will make extensive use of this form of the definition in our work on optimal mass transfer.

---

[1]For an explanation of the terms ν-measurable and ν-integrable see, for example Halmos [45]

2.6.1 Relationship between measure-preserving mappings, push-forward measures and rearrangements.

Sets of rearrangements can always be written as sets of measure-preserving mappings. Sets of measure-preserving mappings can sometimes be written as a set of rearrangements. We show the connections between measure-preserving mappings and rearrangements. We define S to be the set of measure-preserving mappings from $(U\,,\,\mu)$ to $(V\,,\,\nu)$. We recall definition (2.2) for rearrangements. Let $(U\,,\,\mu)$ be a measure interval and suppose $f\,,\,g\in L^1\,(U\,,\,\mu\,,\,\mathbb{R}^n)$. Then f and g are rearrangements if

$$\mu\left(f^{-1}(B)\right)\,=\,\mu\left(g^{-1}(B)\right),\qquad(2.24)$$

for every Borel set $B\subset\mathbb{R}^n$. Define a measure ν on the Borel field of $\mathbb{R}^n$ by

$$\nu(B)\,=\,\mu\left(f^{-1}(B)\right).\qquad(2.25)$$

Now, for $g\in\mathcal{R}(f)$ we have that

$$\nu(B)\,=\,\mu\left(g^{-1}(B)\right)\,=\,\mu\left(f^{-1}\left(B\right)\right).\qquad(2.26)$$

Hence, we deduce that g is a measure-preserving mapping from $(U\,,\,\mu)\rightarrow(V\equiv\mathbb{R}^n\,,\,\nu)$. it follows that S and $\mathcal{R}\left(f\right)$ are the same set. Conversely, we can characterise the set of measure-preserving mappings as a set of rearrangements in certain circumstances.

Let $U \subset \mathbb{R}^n$, μ a measure on the Borel field of $\mathbb{R}^n$. Let $s : (U, \mu) \to (U, \mu)$ be a measure-preserving mapping and let $B \subset U$ be a Borel set.

$$\mu\left(s^{-1}(B)\right) = \mu(B) = \mu\left((id)^{-1}(B)\right). \tag{2.27}$$

Now, we note that $s \in S$, where S is the set of measure-preserving mappings from $(U, \mu) \to (U, \mu)$ if and only if $s \in \mathcal{R}(id)$. In particular we deduce that all members of S have the same L^p-norms.

From what we've shown about measure-preserving mappings being rearrangements of the identity function we can compare this to three functions considered in section 1.4, $f(x) = x$, $g(x) = |1 - 2x|$ and $h(x) = 1 - x$. Given that $g, h \in \mathcal{R}(f)$ and f is the identity function on $[0, 1]$, then $g, h : [0, 1] \to [0, 1]$ are measure-preserving mappings. Furthermore, we note that g is not an injective mapping and thus the measure-preserving mappings need not be injective. However, on measure spaces as defined in definition 2.6.4 the measure-preserving mapping is essentially surjective.

We define push-forward measures as follows:

Definition 2.6.5. *For a mapping $s : U \to V$ we define for (measurable) $A \subset V$*

$$
\begin{aligned}
s_{\#}\mu(A) &= \mu(s^{-1}(A)), & (2.28) \\
&= \mu(\{x : s(x) \in A\}). & (2.29)
\end{aligned}
$$

$s_{\#}\mu$ is known as the push forward of μ through s or alternatively it can be said that s

pushes μ forward to $s_\# \mu$. It is evident from definition 2.6.2 the similarities to definition 2.6.5. It is immediate that s is a measure-preserving mapping from (U, μ) to (V, ν) if s pushes μ forward to ν, that is $s_\# \mu = \nu$. From (2.25) we have that $\nu = f_\# \mu$ and for a measure-preserving mapping from (U, μ) to $(\mathbb{R}^n, f_\# \mu)$ the set S and $\mathcal{R}(f)$ are the same set.

If we now define the left hand side to be a measure, i.e. $\nu(B) = \mu\left(f^{-1}(B)\right)$, we have that any push forwards from μ to ν are rearrangements of f. Given that the set of rearrangements is an equivalence class we can pick any element of the set of rearrangements to yield the push-forward measure.

We show that both measure-preserving mappings and transformations preserve rearrangements. Let bounded sets $\Omega \subset \mathbb{R}^n$, $\Omega' \subset \mathbb{R}^d$ satisfy $\mu(\Omega) = \nu(\Omega')$. Furthermore, we let $f, g : \Omega \to \mathbb{R}$ be integrable functions such that $f \in \mathcal{R}(g)$ and let $s : \Omega' \to \Omega$ be a measure-preserving mapping. We use the notation from definition 2.6.1 for the inverse image. Now, for each Borel set $B \subset \mathbb{R}$ we have

$$
\begin{aligned}
\nu\left((f \circ s)^{-1}(B)\right) &= \nu\left(s^{-1} \circ f^{-1}(B)\right) = \mu\left(f^{-1}(B)\right) \\
&= \mu\left(g^{-1}(B)\right) = \nu\left(s^{-1} \circ g^{-1}(B)\right) = \nu\left((g \circ s)^{-1}(B)\right).
\end{aligned}
$$

Therefore, given $f \in \mathcal{R}(g)$ we have $f \circ s \in \mathcal{R}(g \circ s)$. Further to this if we have that s is a measure-preserving transformation then the opposite argument holds.

We previously discussed a constructive approach to defining monotone rearrangements; using measure-preserving mappings, if (Ω, μ) is a measure-interval and $s : (\Omega, \mu) \to ([0, \mu(\Omega)], \lambda_1)$ a measure-preserving mapping, we have that the increasing

rearrangement of $f_0 : \Omega \to \mathbb{R}$ is given by $(f_0 \circ s)^*$.

Similarly to the set of rearrangements, the set of measure-preserving mappings is closed in L^2 but it is not convex nor compact(in general).

2.6.2 Polar Factorisation.

From work by Ryff [65, Proposition 3] we have that any real integrable function on a bounded interval can be expressed as the composition of its increasing rearrangement and a measure-preserving mapping. For an integrable function f there exists a measure-preserving mapping $s : \Omega \to [0, \mu(\Omega)]$ such that $f = f^* \circ s$, where f^* is an increasing rearrangement, $f^* : [0, \mu(\Omega)] \to \mathbb{R}$. This notion is known as polar factorisation. To give a simple example we consider (1.6). If we let $f(x) = 1 - x$ our increasing rearrangement is $f^*(x) = x$; where $\mu(\Omega) = 1$. Moreover, $h(x) = 1 - x$ is measure-preserving. We have that $f = \mathrm{id} \circ h$. As a direct result of polar factorisation we have the following theorem; we defer the proof until Theorem 3.2.1 (i).

Proposition 2.6.6. *Let $\Omega \subset \mathbb{R}^n$ be a bounded set and let μ be a measure as described in 2.2.1. Let f, $g : \Omega \to \mathbb{R}$ be square integrable functions. Then there exists $\hat{f} \in R(f)$ such that*

$$\int_\Omega \hat{f}(x)g(x)d\mu(x) = \int_0^{\mu(\Omega)} f^*(x)g^*(x)d\lambda_1(x). \tag{2.30}$$

From this result we have that $\sup_{\tilde{f} \in \mathcal{R}(f)} \int_\Omega \bar{f}g$, is attained by $\hat{f}$. The importance of the specific equality (2.30) will become apparent when considering our characterisation of the minimiser in the qualitative features error in the chapter that follows.

56

Polar Factorisation of Vector Valued Rearrangements.

Brenier [13; 14] extended the polar factorisation of Ryff [65] to vector valued functions. Douglas [31] stated that "A vector valued function has a polar factorisation if it can be written as the composition of its monotone rearrangement which is equal almost everywhere to the gradient of a convex function, with a measure-preserving mapping." More precisely

Definition 2.6.7. *Let $u \in L^1(X, \mu, \mathbb{R}^n)$ where (X, μ) is a measure interval. Let Lebesgue measurable $Y \subset \mathbb{R}^n$ be such that $\lambda_n(Y) = \mu(X)$ and let $u^\#$ denote the monotone rearrangement of u on Y. We say u has a polar factorisation through Y if there exists a measure-preserving mapping s from (X, μ) to (Y, λ_n) such that $u = u^\# \circ s$ almost everywhere.*

Burton and Douglas [21; 22] have established exact conditions (on the monotone rearrangement) for unique existence of a polar factorisation and a sharp theorem of general existence. They also established non-existence results(extended by Douglas [31]) which we will make use of in Chapter 6.

Theorem 2.6.8. *Let integrable $u^\# : Y \to \mathbb{R}^n$ be the restriction of the gradient of a proper lower semi continuous convex function to a set $Y \subset \mathbb{R}^n$ of finite positive Lebesgue measure and suppose that $u^\#$ restricted to the complement of its level sets of positive measure is not almost injective. Let (X, μ) be a measure interval satisfying $\mu(X) = \lambda_n(Y)$. Then $u^\#$ has a rearrangement $u : X \to \mathbb{R}^n$ which does not have a polar factorisation through Y.*

2.7 Properties of the set of rearrangements.

2.7.1 Convexity.

When minimising a functional over some set a key property that the set may possess is convexity. We recall the concept of convexity for a subset of $\mathbb{R}^n$.
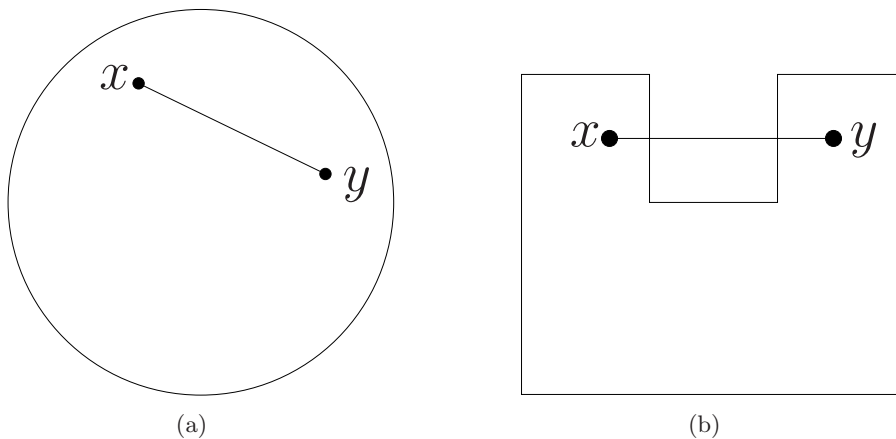
Figure 2.4: An illustration of a (a) convex set (b) non-convex set

A set is convex if you can pick two points anywhere in the set and the line segment that connects the two points remains within the set. The equation for the line joining the two points is given by $(1 - \lambda)x + \lambda y$, where $\lambda \in [0, 1]$ is a parameter. Unlike figure (2.4(a)) the line which joins the two points in figure (2.4(b)) doesn't remain in the set.

We extend the notion of convexity to functions by considering:

The epigraph of a function is the region above or on the graph of the function f. We say that the function f is convex if and only if its epigraph is convex(in the sense described above). We formalise this idea with the following definition:

Definition 2.7.1. *Let V be a vector space. A function $f : V \to \overline{\mathbb{R}}$ is convex if for $x, y \in V$ and $\lambda \in [0, 1]$*
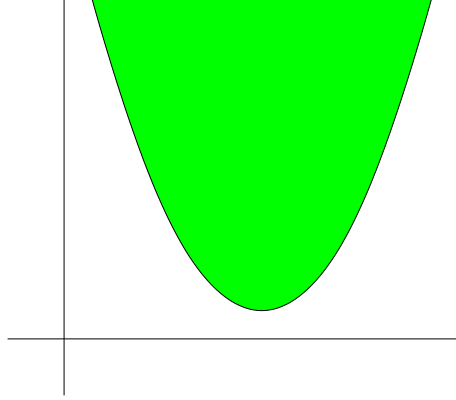
Figure 2.5: An epigraph of a function, denoted $epi\,(f)$ for a given function f, where $epi\,(f) = \{(x, y) : y \geq f(x)\}$

$$f\left((1 - \lambda)x + \lambda y\right) \leq (1 - \lambda)f(x) + \lambda f(y),\qquad(2.31)$$

where the right hand side is defined.

The set of rearrangements is non-convex in general; we consider a specific counterexample. For $f(x) = x$ as illustrated in figure 1.5(a), where $a(x) = f^*$ and $b(x) = f^\Delta = 1 - x$, illustrated by figures 1.5(a) and 1.6 and $\lambda = \frac{1}{2}$, our linear sum gives us the constant function $\frac{1}{2}$. However $1/2 \notin \mathcal{R}\,(f)$. In fact $\mathcal{R}\,(f)$ is only convex when f is constant (and $\mathcal{R}\,(f) = \{f\}$).

To overcome non-convexity we consider the closed convex hull of a set. The convex hull of a set D, which we denote conv D, is the smallest convex set which contains D. It is the intersection of all convex sets that contains D. When D is the set of rearrangements there is a characterisation of the closed convex hull that is of great utility in extremisation problems. We discuss this in the next section.

2.7.2 Do extremisers exist in sets of rearrangements?

In later sections we will be interested in extremising some functional over a set of rearrangements. A classical approach would be to pick a maximising sequence, extract a convergent subsequence and show that the limit it converges to is a maximiser of the functional. The latter claim depends on (some form of) continuity of the functional; we concentrate on extracting a convergent subsequence, a property of compactness. However, we note that $\mathcal{R}(f_0)$ is not weakly compact; for a specific counterexample see Douglas [30, Section 2.4].

The solution is to work with the weak closure of the set of rearrangements as it contains the weak limits, hence, possible extremisers over the set of rearrangements. The weak closure of the set of rearrangements is the smallest weakly closed set that contains the set of rearrangements, denoted $\overline{\mathcal{R}(f_0)}$ for a function f_0. We define the weak closure as follows:

Definition 2.7.2. *A set $D \in L^2(\Omega, \mu)$ is weakly(or sequentially) closed if the weak limit of every weakly convergent sequence in D belongs to D.*

In fact for a square integrable non-negative function f we have that the weak closure of the set of rearrangements is equal to the closed convex hull of the set of rearrangements, see Ryff [65]; this set is weakly compact. For a non-negative square integrable function f_0, the following characterisation of the closed convex hull of the set of rearrangements, denoted $\overline{\text{conv}\mathcal{R}(f_0)}$, is due to Douglas [29, Theorem 2.1]:

$$\overline{\text{conv}\mathcal{R}(f_0)} = \left\{ f \geq 0 : \int_\Omega (f - \alpha)_+ d\mu \leq \int_\Omega (f_0 - \alpha)_+ d\mu \, \forall \alpha > 0, \int_\Omega f d\mu = \int_\Omega f_0 d\mu \right\} (2.32)$$

where, $+$ denotes the positive part of the function. The difference between the closed convex hull of the set of rearrangements and the set of rearrangements is the inequality $\|f\|_p \leq \|f_0\|_p$; the inequality allows us to include the weak limits. Therefore, the L^p-norm is no longer preserved under the closed convex hull of the set of rearrangements. Note that for our earlier example, $1/2 \in \overline{\text{conv}\mathcal{R}(f_0)} \setminus \mathcal{R}(f)$.

We finish this chapter with an important theorem for the work that follows:

Theorem 2.7.3. *Let* (Ω, μ) *be a measure interval. Let* f, $g : \Omega \to \mathbb{R}$ *be square integrable functions. Suppose there is an increasing function* φ *such that* $f^* = \varphi \circ g \in \mathcal{R}(f_0)$. *Then* f^* *is the unique maximiser of the functional* $\langle \cdot, g \rangle$ *relative to* $\overline{\text{conv}\mathcal{R}(f_0)}$.

Proof: Burton [17, Theorem 3]. □

Chapter 3

Closest Rearrangements.

3.1 Introduction.

Our approach to weather forecast error decomposition uses the concept of "best fitting"

rearrangements of the forecast; the closest rearrangements problem characterises the set

of best fits. The chapter studies the problem of characterising those elements of a set of

rearrangements which are closest (in the L^2 space) to a given square integrable function

q_1; that is characterise those $q \in \mathcal{R}(q_2)$ which attain

$$\inf_{q \in \mathcal{R}(q_2)} \|q - q_1\|_2.$$

Noting that L^2 norms are preserved by rearrangement we deduce that

$$\inf_{q \in \mathcal{R}(q_2)} \|q_1 - q\|_2 \;\; = \;\; \left\|q_1^\Delta - q_2^\Delta\right\|_2 = \left\|q_1^* - q_2^*\right\|_2, \tag{3.1}$$

$$= \;\; \left\{ \|q_1\|_2^2 + \|q_2\|_2^2 - 2\int_0^{\mu(\Omega)} q_1^\Delta q_2^\Delta \; \mathrm{d}\lambda_1 \right\}^{1/2}. \tag{3.2}$$

For a practical application of forecast error decomposition we would have to calculate this quantity: we consider possible numerical schemes in Chapter 6.

We saw in section 2.5.1 that an equivalent problem (i.e. same extremisers) is to find $q \in \mathcal{R}(q_2)$ which attain

$$\sup_{q \in \mathcal{R}(q_2)} \int_\Omega q q_1 \; \mathrm{d}\mu(x). \tag{3.3}$$

From equation 3.3 we are lead to the study of equality in Hardy-Littlewood type inequalities which happen as a result of Hardy et al. [46, Theorems 368-370 & Theorem 378]. The study of Hardy-Littlewood inequalities ask for which $q \in \mathcal{R}(q_2)$ does $\int q_1 q = \int q_1^\Delta q_2^\Delta$ (we refer the reader to Lieb and Loss [56] for more details on these inequalities)? Generally, the work on such inequalities have considered them for a strictly decreasing $q_1 = q_1^\Delta$; for generalisations of the Hardy-Littlewood inequalities, for example see Hajaiej [15; 43]. Comparatively to the work of Hajaiej and Burchard amongst others the difficulty with our problem revolves around the fact that q_1 has level sets of positive measure, moreover, that not every element of $\mathcal{R}(f)$ is of the form $f = f^\Delta \circ T$, where T is a measure-preserving transformation, that is a bijection T where both T and T^{-1} are measure-preserving mappings.

In the study of (3.3) we could utilise a result from Burchard and Hajaiej [15, Theorem 1]. But, the idea of the closest rearrangement problem is to get information about the maximiser q on the level set having positive measure of q_1; Burchard and Hajaiej does not provide any further insight into our problem.

We define the set of all minimisers:

Definition 3.1.1. $\mathcal{M} = \left\{ All\ q \in \mathcal{R}(q_2)\ attaining\ \inf_{q \in \mathcal{R}(q_2)} \|q - q_1\|_2 \right\}.$

As discussed previously, $\mathcal{M}$ is equally the set of all maximisers of (3.3).

In section 3.2 we use the methods of Douglas [30, Proposition 2] to show that $\mathcal{M}$ is non-empty and gives a precise condition for when $\mathcal{M}$ is singleton. We repeat the methods of Douglas [30, Proposition 2] for convenience and for ease of understanding in what follows. The content of Proposition 2 is to provide a sufficient, but not necessary condition for when $\mathcal{M}$ is singleton; this is the content of section 3.3. The major result in this section is a characterisation of $\mathcal{M}$. Before stating this characterisation we define notation integral to the main theorem:

Notation. *For a function $q_1 : \Omega \to \mathbb{R}$ let $\alpha \in \mathbb{R}$ and denote the level set of q_1 corresponding to α by $L(\alpha) = \{x \in \Omega : q_1(x) = \alpha\}$.*

Notation. *Let non-negative $q_1 \in L^2(\Omega, \mu)$, where (Ω, μ) is a measure interval. For $n \in I$ write L_n to denote the level sets of q_1 which have positive measure. Then I is countable and*

$$L_n = \{x \in \Omega : q_1(x) = l_n\}, \tag{3.4}$$

for some $(l_n)_{n \in I} \subset \mathbb{R}$. We suppose that $(l_n)_{n \in I}$ are decreasing. Write $L = \bigcup_{n \in I} L_n$,

the union of all level sets of q_1 having positive measure.

Nealers Theorem. *Let non-negative q_1 , $q_2 \in L^2(\Omega , \mu)$, where (Ω , μ) is a measure interval. Suppose $\hat{q} \in \mathcal{M}$. Then $q \in \mathcal{M}$ if and only if $q = \hat{q}$ on $\Omega \setminus L$ and $q\mid_{L_n} \in \mathcal{R}\left(\hat{q}\mid_{L_n}\right)$ for each $n \in I$. We assume that (l_n) can be written as a decreasing sequence.*

We explain the final assumption of Nealers Theorem. There is no difficulty writing (l_n) as a decreasing sequence if the index set I is finite, which would be the case for any practical application of rainfall error forecast decomposition. However I is countable in general. If (l_n) can be arranged as an increasing sequence, analogous arguments demonstrate the result.

We let $\hat{q} \in \mathcal{M}$. We let $(\star)$, which we define formally later, be the condition that maximisers are unique up to level sets of positive measure of q_1 and rearrangements on the level sets of positive measure. Nealers Theorem is an if and only if condition which states that $q \in \mathcal{M} \Leftrightarrow q$ satisfies $(\star)$.

The remainder of the chapter discusses Nealers Theorem. We begin in section 3.3.1 illustrating the difficulties with the theorem via an example. Nealers Theorem reduces to the problem of when we have $q \in \mathcal{R}(q_2)$ which satisfy

$$\int_\Omega q\, q_1 \; \mathrm{d}\mu\,(x) \; = \; \int_0^{\mu(\Omega)} q_2^\Delta\, q_1^\Delta \; \mathrm{d}\lambda_1\,(x)\,. \tag{3.5}$$

We disintegrate the difference between the right hand and left hand side of (3.5) using a layer cake principle (as discussed in Villani [71], for example.). We denote the integrand after the application of the layer cake principle by G_q. Lemma 3 and Proposition 4 establish properties of G_q; section 3.4.1 proves both Lemma 3 and Proposition 4.

Using these properties the problem in Nealers Theorem changes from considering $q \in \mathcal{M}$ which we show is equivalent to showing $\lambda_1 \left(G_q^{-1} (0, \infty) \right) = 0$; this is proved in Lemma 5 and is the content of section 3.4.2.

We are now in a position to prove Nealers Theorem. We begin by proving for $q \in (\star) \Rightarrow q \in \mathcal{M}$, using the fact that L^1-norms are preserved by rearrangement. Section 3.5 begins with this argument, which is Proposition 6.

The difficulty in proving Nealers Theorem comes from showing that $q \in \mathcal{M}$ implies $q \in (\star)$. We reduce this problem to working on level sets having positive measure of q_1 and "non-level" sets between values corresponding to level sets of q_1 of positive measure; we generalise the result using induction. However, there is no priori reason why two elements of $\mathcal{M}$ should be maximisers for the reduced problems nor that they should be rearrangements when restricted to these sets. Completing the proof of Nealers Theorem concludes section 3.5 (and the chapter).

We finish the introduction illustrating what is meant by Nealers Theorem with the following example from Douglas [30] that the minimisers need not be unique. We let $q_1, q_2, q_3 : [-1, 1]^2 \to \mathbb{R}$, where

$$q_1(x) = \begin{cases} 3, & \text{if } x^2 + y^2 \leq 1/16, \\ 2, & \text{if } 1/16 < x^2 + y^2 \leq 9/16, \\ 0, & \text{otherwise.} \end{cases} \tag{3.6}$$

$$q_2(x) = \begin{cases} 3, & \text{if } x^2 + y^2 \leq 1/4 \\ 0, & \text{otherwise,} \end{cases} \tag{3.7}$$

$$q_3(x) = \begin{cases} 3, & \text{if } x^2 + y^2 \leq 1/16 \text{ or } 6/16 \leq x^2 + y^2 \leq 9/16, \\ \\ 0, & \text{otherwise,} \end{cases} \tag{3.8}$$

These functions are illustrated with the following figures:

(a)   (b)   (c)
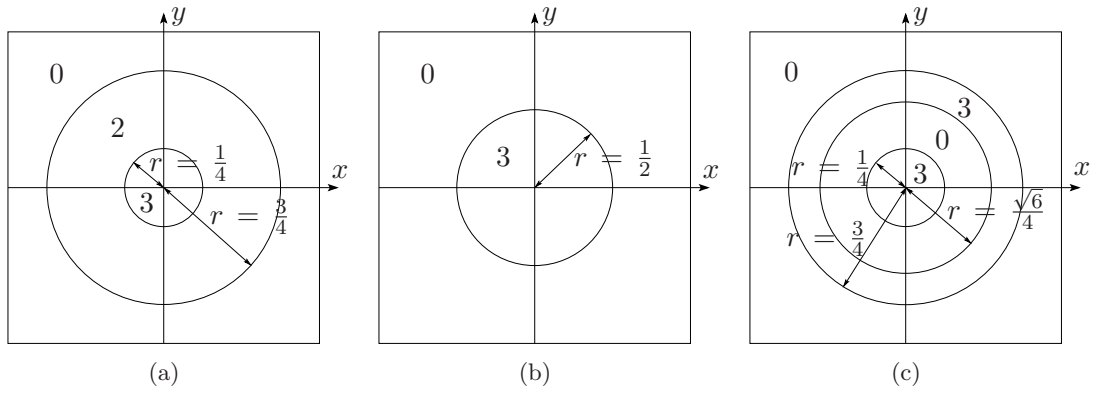
Figure 3.1: An illustration of (a) $q_1 = q_1(x, y)$ (b) $q_2 = q_2(x, y)$ (c) $q_3 = q_3(x, y)$ as defined by (3.6), (3.7) and (3.8), to show that $\mathcal{M}$ need not be singleton.

It is easily seen that we have the following:

$$\|q_1 - q_2\|_2 = \|q_1 - q_3\|_2 = \inf_{q \in \mathcal{R}(q_2)} \|q_1 - q\|_2, \tag{3.9}$$

illustrating that the minimiser need not be unique. In this case elements of $\mathcal{M}$ must be equal to 3 in the inner disc of radius 1/4, zero outside the disc of radius 3/4, and rearrangements in the annulus where the radius varies between 1/4 and 3/4.

3.2 The well posedness of the Closest Rearrangements problem.

The closest rearrangements problem gives rise to two questions;

(1) Do maximisers exist such that (3.3) is satisfied?

(2) If maximisers exist, are they unique?

For the displacement error we "minimise the kinetic energy of moving all best fits". To do this we must all the best fits are. For our error score to be defined it was essential that (3.3) attained its bound; this is the content of this section. When maximisers exist they need not be unique; we illustrate the necessary and sufficient conditions which give us a unique maximiser. We repeat the arguments of Burton [17] for completeness.

Theorem 3.2.1. *Let q_1, q_2: $\Omega \to \mathbb{R}$ be non-negative square integrable functions, where $\Omega \subset \mathbb{R}^n$ is bounded. Then*

(i) *there exists $\hat{q} \in \mathcal{R}(q_2)$ such that $\hat{q}$ attains $\sup_{q \in \mathcal{R}(q_2)} \int_\Omega q q_1 \, d\mu$.*

(ii) *$\hat{q}$ is the unique maximiser of $\int_\Omega q q_1 \, d\mu$ over $q \in \mathcal{R}(q_2)$ if and only if $\hat{q} = \varphi \circ q_1$ for some increasing function φ.*

Proof: (i) We repeat a standard argument from the literature (see, for example, Alvino, Trombetti and Lions [1]). Ryff [65] showed that there exists a measure-preserving mapping $s : \Omega \to [0, \lambda_n(\Omega)]$ such that $q_1 = q_1^* \circ s$, where q_1^* is the increasing rearrangement of q_1. For $q \in \mathcal{R}(q_2)$

$$\int_\Omega q q_1 d\lambda_n \quad \leq \quad \int_0^{\lambda_n(\Omega)} q_2^* q_1^* d\mu \qquad (3.10)$$

$$= \quad \int_\Omega (q_2^* \circ s)(q_1^* \circ s)\, d\mu, \qquad (3.11)$$

$$= \quad \int_\Omega (q_2^* \circ s)\, q_1 d\mu.$$

By way of explanation equation (2.20) gives us (3.10). From the definition of a measure-preserving mapping the value of the integral is preserved (see (2.23)); we have (3.11). Moreover, $q_2^* \circ s \in \mathcal{R}(q_2)$. Therefore, $\hat{q} = q_2^* \circ s$ is one maximiser for $\sup_{q \in \mathcal{R}(q_2)} \int_\Omega q q_1$ but this need not be the only maximiser.

(ii) From part (i) we have that $\hat{q}$ attains $\sup_{q \in \mathcal{R}(q_2)} \int_\Omega q q_1 d\lambda_2$ if and only if

$$\int_\Omega q_1 \hat{q} d\mu = \int_0^{\lambda_n(\Omega)} q_1^* q_2^* d\lambda_1. \qquad (3.12)$$

From Theorem 2.7.3 we have that if $\hat{q} = \varphi \circ q_1 \in \mathcal{R}(q_2)$ then $\hat{q}$ is the unique element of $\mathcal{R}(q_2)$ where equality holds in (3.12). The only if condition is a result of work by Burton [17, Theorem 5]. This states that if $\hat{q}$ is a unique element where equality holds in (3.12) then there is an increasing function φ such that $\hat{q} = \varphi \circ q_1$ almost everywhere. This completes the proof.

$\square$

3.3 When are the maximisers unique?

The necessary and sufficient conditions required for a unique maximiser in Theorem 3.2.1 is a difficult condition to verify. We establish a sufficient (but not necessary) condition for the maximiser to be unique. It is a special case of Nealers Theorem. In this section and for the remainder of the chapter we will answer question as to when maximisers are unique. We begin by illustrating the definition of the level set in figure (3.2); the level set is section circled in red which would give us a level set of positive measure.
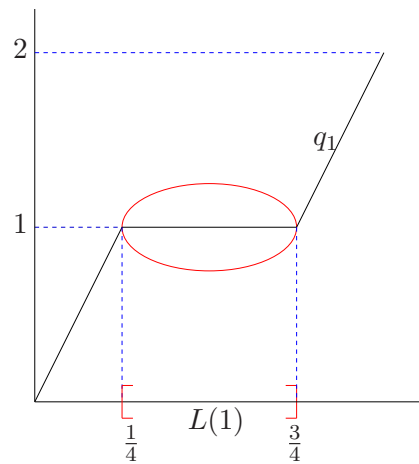
Figure 3.2: The definition of a level set for $q_1 = q_1(x)$ as defined in equation (3.14)

Given the importance of this notion we consider when there are no level sets having positive measure of q_1.

Proposition 2. *Let q_1 and q_2 be defined as in Nealers Theorem. Suppose q_1 has no level sets having positive measure. Then there exists an increasing function, φ, such that:*

$$\varphi \circ q_1 \in \mathcal{R}(q_2). \tag{3.13}$$

Moreover, $\varphi \circ q_1$ is the unique maximiser of $\int_\Omega q q_1 \, \mathrm{d}\mu$ over $q \in \mathcal{R}(q_1)$.

Proof: From the proof of Theorem 3.2.1 we have that (2.19) is achieved by $\hat{q} = q_2^* \circ s$. Also, q_1 has a polar factorisation $q_1 = q_1^* \circ s$. As there are no level sets of q_1^*, q_1^* has no intervals of constancy and is increasing, therefore it is injective. We note that q_1 and q_1^* are rearrangements; we can delete a set of zero measure from Ω and assume the range of q_1 lies in the range of q_1^*. Given $q_1 = q_1^* \circ s$ we have $s = (q_1^*)^{-1} \circ q_1$.

$$\hat{q} = q_2^* \circ (q_1^*)^{-1} \circ q_1 \quad \text{where} \quad \varphi = q_2^* \circ (q_1^*)^{-1}.$$

For $\hat{q}$ we have an increasing function composed with the inverse of an increasing function composed with q_1. The inverse of an increasing function is also increasing; therefore φ is an increasing function. From Theorem 3.2.1 (ii) we have a unique minimiser and this completes the proof. □

3.3.1 Example.

We illustrate what is being described in Nealers Theorem with the following example. Consider the following functions

$$q_1(x) = \begin{cases} 4x, & \text{for } 0 \le x \le \frac{1}{4} \\ 1, & \text{for } \frac{1}{4} \le x \le \frac{3}{4} \\ 4x - 2, & \text{for } \frac{3}{4} \le x \le 1. \end{cases} \tag{3.14}$$

$$q_2(x) = 2x \tag{3.15}$$

$$\overline{q_2}(x) = \begin{cases} 2x, & \text{for } 0 \le x \le \frac{1}{4} \\ 2(1-x), & \text{for } \frac{1}{4} \le x \le \frac{3}{4} \\ 2x & \text{for } \frac{3}{4} \le x \le 1, \end{cases} \tag{3.16}$$

where (3.15) and (3.16) are illustrated below. Now, q_2 and $\overline{q_2}$ are maximisers of $\int_0^1 q_1 q$ over $q \in \mathcal{R}(q_2)$.
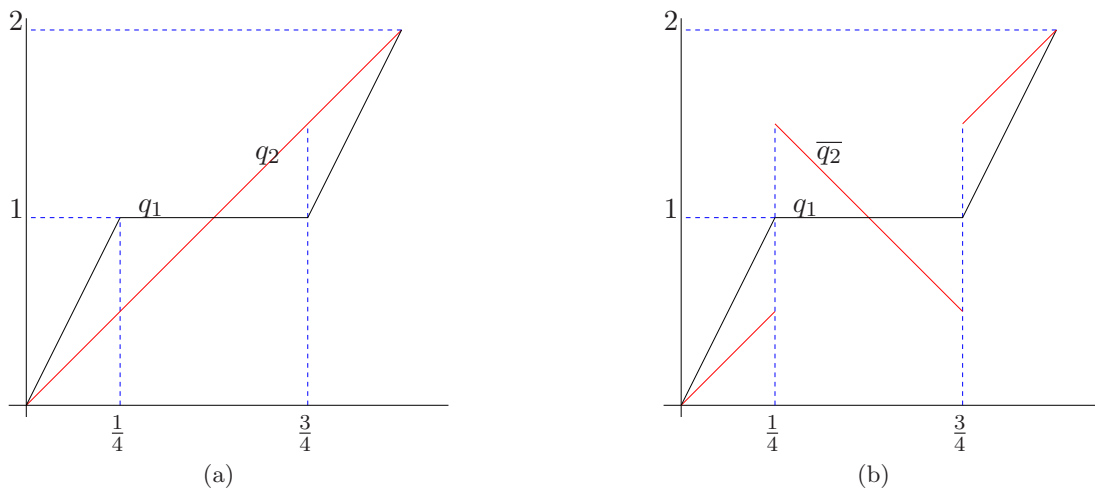
Figure 3.3: An illustration of (a) q_2, (b) $\overline{q_2}$, maximisers of $\int q_1 q$ over $q \in \mathcal{R}(q_2)$. See definitions (3.14),(3.15) and (3.16)

(3.14) has a level set of positive measure between $\frac{1}{4} \le x \le \frac{3}{4}$; this is illustrated in figure (3.3). Nealers Theorem states that the maximisers, (3.15) and (3.16), are equal on $[0, 1/4] \bigcup [3/4, 1]$ and maximisers (3.15) and (3.16) are rearrangements on $[1/4, 3/4]$.

However, on the level set of q_1 of positive measure there can be infinitely many maximisers all belonging to the set of rearrangements of q_1. From the above we can note the following

$$\int_{\frac{1}{4}}^{\frac{3}{4}} 1 \cdot q_2^* \ \mathrm{d}x \ = \ \int_{\frac{1}{4}}^{\frac{3}{4}} 1 \cdot h \ \mathrm{d}x \text{ for any } h \in \mathcal{R}\left(q_2|_{\left[\frac{1}{4}, \frac{3}{4}\right]}\right). \tag{3.17}$$

3.4   Preliminaries to Nealers Theorem.

3.4.1   A layer-cake principle.

Following the methods of Burton [17, Proof of Theorem 1] and the use of the layer-cake principle in Lieb and Loss [56] we consider (3.3) in a different form. Firstly, we define for $s \in [\, 0\, ,\, \infty)$

$$Z\left(s\right) \ = \ \{x \in \Omega \ : \ q_1(x) \geq s\}. \tag{3.18}$$

Moreover, we define σ by $\sigma\left(s\right) \ = \ \mu\left(Z\left(s\right)\right)$. Now:

$$\int_0^{\mu(\Omega)} q_2^\Delta q_1^\Delta \ \ \mathrm{d}\lambda_1\left(x\right) - \int_\Omega qq_1 \ \mathrm{d}\mu\left(x\right)$$
$$= \ \int_0^\infty \left(\int_0^{\sigma(s)} q_2^\Delta\left(x\right) \ \mathrm{d}\lambda_1(x) - \int_{Z(s)} q\left(x\right) \ \mathrm{d}\mu(x)\right) \ \mathrm{d}\lambda_1\left(s\right) \tag{3.19}$$
$$= \ \int_0^\infty G_q\left(s\right) \ \mathrm{d}\lambda_1\left(s\right), \tag{3.20}$$

where $G_q\left(s\right) \ = \ \int_0^{\sigma(s)} q_2^\Delta\left(x\right) \ \mathrm{d}\lambda_1\left(x\right) - \int_{Z(s)} q\left(x\right) \ \mathrm{d}\mu\left(x\right)$. We note that we get (3.19)

from the methods of Lemma 2.5.1. The first integral of $G_q(s)$ is:

$$s \to \int_0^{\sigma(s)} q_2^{\Delta}(x) \, \mathrm{d}\lambda_1(x). \tag{3.21}$$

The question of whether (3.21) is continuous can be broken into whether the following

two mappings are continuous

1. $s \to \sigma(s)$

2. $t \to \int_0^t q_2^{\Delta}(x) \, \mathrm{d}\lambda_1(x).$

If we can prove that both the above mappings are continuous we can deduce that

the composition is continuous [52]. Lemma 3 proves when the mapping $s \to \sigma(s)$ is

continuous.

Lemma 3. *Let q_1 be defined as in Nealers Theorem and $Z(s)$ as before. We define*

$\sigma(s) : [\, 0\,,\, \infty\,) \to \mathbb{R}$ *by* $\sigma(s) = \mu(Z(s)).$

Then σ is non-increasing (decreasing), upper semi continuous and continuous except at

those s corresponding to level sets of q_1 having positive measure.

Preamble to proof of Lemma 3. To map s to $\sigma(s)$ we build up our graph by

considering the size of $Z(s)$ for each s. As an example we consider $s = 0$ and suppose

$(\Omega\,,\mu)$ is $([0\,,1]\,,\lambda_1)$. Now, $\sigma(0) = \mu\{x \in \Omega : q_1(x) \geq 0\} = \lambda_1([0\,,1])$, so $\sigma(0) = 1$.

The set for which $q_1(x) \geq 0$ is the whole of the domain; as s increases $\mu(\sigma(s))$ will

decrease. Therefore, $\sigma(s)$ is a decreasing function. For a q_1 that doesn't have a level set

of positive measure our $\sigma(s)$ is a continuous function. Even with jumps $\sigma(s)$ is upper

semi continuous and is continuous apart from the finite jumps. But, the continuity of

$\sigma(s)$ fails when there are level sets of q_1 of positive measure. This is illustrated with the following example:
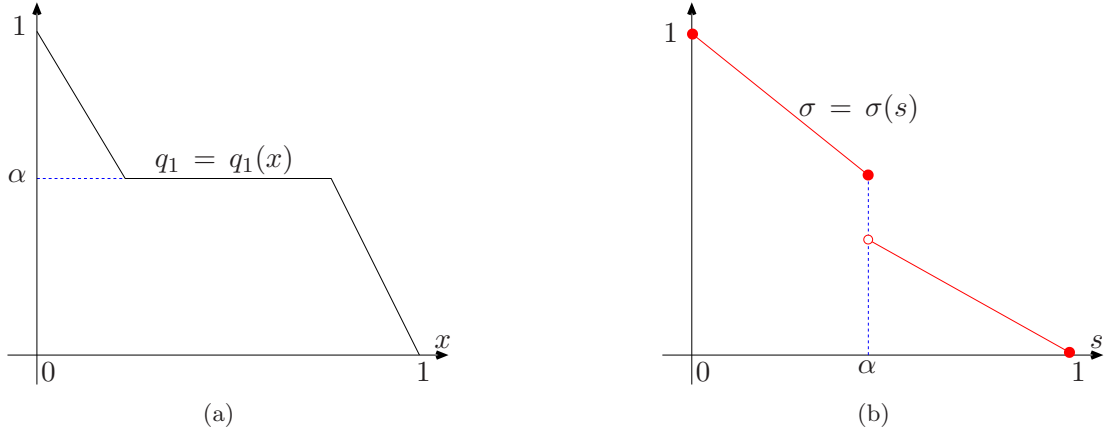
Figure 3.4: An illustration of (a) an example of q_1 (b) σ derived from q_1; equation (3.14)

We consider figure 3.4(b) and where the continuity of $\sigma(s)$ fails. We note that $\sigma(s)$ is left continuous, but the continuity of $\sigma(s)$ fails when we consider the continuity from the right. Definitions 2.4.1 and 2.4.2, which follow, generalises this idea of one sidedness. We note for $\sigma = \sigma(s)$ as shown in figure 3.4(b) that the function is upper semi continuous, but not lower semi continuous.

We finish this preamble by considering a simple example to illustrate the more formal proof later. So, what we are trying to show is $\sigma(\alpha) \geq \limsup_{n \to \infty} \sigma(\alpha_n)$ for every sequence $\alpha_n \to \alpha$. We illustrate this by considering two sequences $\left(\alpha - \frac{1}{n}\right)$ and $\left(\alpha + \frac{1}{n}\right)$, one tending to α from the left and the other tending to α from the right.

$$\bigcap_{n \in \mathbb{N}} \left\{ x : q_1(x) \geq \alpha - \frac{1}{n} \right\} = \{x : q_1(x) \geq \alpha\}, \tag{3.22}$$

which implies from the nested interval theorem that

$$\mu\left(\bigcap_{n\in\mathbb{N}}\left\{x \,:\, q_1(x)\geq\alpha-\frac{1}{n}\right\}\right) \;=\; \lim_{n\to\infty}\mu\left(\left\{x \,:\, q_1(x)\geq\alpha-\frac{1}{n}\right\}\right), \quad (3.23)$$

$$\;=\; \mu\left(\{x \,:\, q_1(x)\geq\alpha\}\right).$$

Similarly,

$$\bigcup_{n\in\mathbb{N}}\left\{x \,:\, q_1(x)\geq\alpha+\frac{1}{n}\right\} \;=\; \{x \,:\, q_1(x)>\alpha\}, \quad (3.24)$$

which implies that

$$\mu\left(\bigcup_{n\in\mathbb{N}}\left\{x \,:\, q_1(x)\geq\alpha+\frac{1}{n}\right\}\right) \;=\; \lim_{n\to\infty}\mu\left(\left\{x \,:\, q_1(x)\geq\alpha+\frac{1}{n}\right\}\right), \quad (3.25)$$

$$\;=\; \mu\left(\{x \,:\, q_1(x)>\alpha\}\right).$$

Given that the problems occur around the level set we consider the functions as they tend to the level set from above and below. From equation (3.23) we have

$$\sigma\left(\alpha-\frac{1}{n}\right) \to \sigma(\alpha) \text{ as } n\to\infty. \quad (3.26)$$

From equation (3.24) we have

$$\sigma\left(\alpha + \frac{1}{n}\right) \to \sigma(\alpha) - |\{x : q_1(x) = \alpha\}|. \qquad (3.27)$$

For this specific example we have that the upper semi continuity condition is satisfied from equation (3.26). However, as we alluded to previously the definition of lower semi continuity is dependent on whether q_1 has a level set of positive measure or not. If $|\{x : q_1(x) = \alpha\}| = 0$ then the lower semi continuity definition is satisfied and the mapping is continuous as required for the proof of the lemma. One final point to note is that for the argument which follows we split the sequence $(\alpha_n)_{n \in I}$ into three parts:

1. one tending towards α from the left,

2. one of constant α's,

3. one tending to α from the right.

If either or both of the subsequences described by (1) and (3) have infinite cardinality, we can pass to monotonic subsequences. We can then obtain analogous results to (3.26) and (3.27).

Proof of Lemma 3: Given the definition of $\sigma(s)$ we have that $q_1(x)$ is fixed and decreasing, hence it is immediate that σ is non-increasing. Fix $\alpha \geq 0$. Let (α_n) be a sequence such that $\alpha_n \to \alpha$ as $n \to \infty$. We consider what happens around α, therefore we split the sequences into three subsequences and define

$$I_1 = \{n : \alpha_n < \alpha\}, \ I_2 = \{n : \alpha_n = \alpha\}, \ I_3 = \{n : \alpha_n > \alpha\}. \qquad (3.28)$$

Given that $n \to \infty$ at least one of I_1, I_2 and I_3 must have infinite cardinality[1]. Trivially, for the index set I_2 as $\alpha_n = \alpha$ it is again immediate that $\sigma(\alpha_n) \to \sigma(\alpha)$; we concentrate on the index sets I_1 and I_3.

Given the definition of convergence and that we have already supposed that $\alpha_n \to \alpha$; if either or both of I_1 and I_3 have infinite cardinality we can extract a monotone subsequence from $(\alpha_n)_{n \in I_1}$ $\left(\text{and/or } (\alpha_n)_{n \in I_3}\right)$ converging to α. Writing I_4 and I_5 for the new index sets let:

$$I_4 = (n_k)_{k=1}^{\infty}, \; I_5 = (m_k)_{k=1}^{\infty}. \tag{3.29}$$

We have

$$\bigcap_{k=1}^{\infty} \{x : q_1 \geq \alpha_{n_k}\} = \{x : q_1(x) \geq \alpha\}, \tag{3.30}$$

and

$$\bigcup_{r=1}^{\infty} \{x : q_1 \geq \alpha_{m_k}\} = \{x : q_1(x) > \alpha\}, \tag{3.31}$$

$$= \{x : q_1(x) \geq \alpha\} \setminus \{x : q_1(x) = \alpha\}. \tag{3.32}$$

The left hand side of equation (3.30) is an intersection of decreasing sets, and the left hand side of equation (3.31) is a union of increasing sets. For a decreasing sequence E_n

---

[1] where cardinality is the number of elements in the set.

of measurable we have the following

$$\mu\left(\bigcap E_n\right) = \lim_{n\to\infty} \mu\left(E_n\right). \tag{3.33}$$

Furthermore, for an increasing sequence of measurable sets E_n we have the following

$$\mu\left(\bigcup E_n\right) = \lim_{n\to\infty} \mu\left(E_n\right). \tag{3.34}$$

Using equations (3.30) and (3.33) we have that

$$\sigma\left(\alpha_{n_k}\right) \to \sigma\left(\alpha\right) \text{ as } k\to\infty, \tag{3.35}$$

and furthermore using equations (3.31) and (3.34) we have that

$$\sigma\left(\alpha_{m_k}\right) \to \sigma\left(\alpha\right) - \mu\left(\{x \,:\, q_1(x) = \alpha\}\right) \text{ as } r\to\infty. \tag{3.36}$$

We deduce from equation (3.35) that

$$\sigma(\alpha) \geq \limsup_{n\to\infty} \sigma(\alpha_n), \tag{3.37}$$

and conclude that σ is upper semi continuous.

If $\mu\left(\{x : q_1(x) = \alpha\}\right) = 0$ then $\sigma(\alpha_{m_k}) \to \sigma(\alpha)$ as $r \to \infty$. Therefore, the lower semi continuity depends on whether q_1 contains level sets having positive measure. Combining the above it follows that σ is continuous when there are no level sets of q_1 having positive measure. This completes the proof. $\square$

The first integral in the definition of G_q is a composition of functions; we have examined the continuity of σ in Lemma 3 and Proposition 4 shows when $t \to \int_0^t q_2^\Delta(x) \, d\lambda_1(x)$ is continuous.

Proposition 4. *Let q_2 be defined as before. If we define $H(t) : [0, \infty) \to \mathbb{R}$ by*

$$H(t) = \int_0^t q_2^\Delta, \qquad (3.38)$$

then H is continuous.

Proof of Proposition 4: Applying the Cauchy-Schwarz inequality for $s < t$

$$|H(t) - H(s)| = \left| \int_s^t q_2^\Delta \right| \le \left\| 1_{(s,t)} \right\|_2 \left\| q_2^\Delta \right\|_2 = (t - s)^{\frac{1}{2}} \cdot \left\| q_2 \right\|_2, \qquad (3.39)$$

where $1_{(s,t)}$ denotes the characteristic function of the interval (s, t). We note that $\left\| 1_{(s,t)} \right\|_2 \to 0$ as $s \to t$ and $\left\| q_2^\Delta \right\|_2$ is a constant and this yields that the mapping is continuous. $\square$

80

3.4.2   Properties of G_q.

We have made use of Lemma 3 and Proposition 4 to illustrate when (3.21) is continuous.
We note that $\mu\left(Z\left(s\right)\right) = \sigma\left(s\right)$, therefore when we establish the properties of σ we can
apply similar arguments to the second integral of $G_q\left(s\right)$. Combining these results we've
shown when G_q is continuous. Lemma 5 establishes when we have a maximiser, utilising
Lemma 3 and Proposition 4 in the proof.

Lemma 5. *Let q_1, q_2 and $\mathcal{M}$ be defined as before. Suppose that $q \in \mathcal{R}\left(q_2\right)$. We define
$G_q\left(s\right) : \left[\,0\,,\,\infty\,\right) \to \mathbb{R}$ by*

$$G_q\left(s\right) = \int_0^{\sigma\left(s\right)} q_2^{\Delta}\left(x\right)\ \mathrm{d}\lambda_1\left(x\right) - \int_{Z\left(s\right)} q\left(x\right)\ \mathrm{d}\mu\left(x\right), \tag{3.40}$$

*where $\sigma\left(s\right)$ and $Z\left(s\right)$ are as defined in Lemma 3. Then $G_q\left(s\right) \geq 0$ for every $s \in \left[\,0\,,\,\infty\,\right)$
$G_q^{-1}\left(0\,,\,\infty\right)$ is measurable and $q \in \mathcal{M}$ if and only if $\lambda_1\left(G_q^{-1}\left(0\,,\,\infty\right)\right) = 0$. Moreover,
G_q is continuous at each s where σ is continuous.*

Preamble to proof of Lemma 5. Using the methods of Burton [17, Theorem 2] we
will demonstrate that $q \in \mathcal{M}$ if and only if $\int_0^{\infty} G_q\left(s\right)\ \mathrm{d}s = 0$. From Burton [17, Lemma
2] we have that $G_q\left(s\right) \geq 0$ for each s. To complete the proof we show:

1. if $\lambda_1\left(G_q^{-1}\left(0\,,\,\infty\right)\right) = 0$ then $q \in \mathcal{M}$

2. if $\lambda_1\left(G_q^{-1}\left(0\,,\,\infty\right)\right) > 0$ then $q \notin \mathcal{M}$

We use the nested interval theorem to show that if $\lambda_1\left(G_q^{-1}\left(0\,,\,\infty\right)\right) > 0$ then $\int G_q\left(s\right) >$
0 and then show that $q \notin \mathcal{M}$. To this end we select a decreasing sequence which tends
towards zero, but never reaches zero. The obvious choice is $\frac{1}{n}$. Now,

$$\{s \,:\, G_q(s) > 0\} = \bigcup_{n \in \mathbb{N}} \left\{s \,:\, G_q(s) > \frac{1}{n}\right\}. \tag{3.41}$$

Equation (3.41) follows because if $G_q(s) > 0$ for some s then $G_q(s) > 1/n$ if we choose $n \in \mathbb{N}$ sufficiently large. We make use of the nested interval theorem to yield

$$\lambda_1 \left(\bigcup_{n \in \mathbb{N}} \{s \,:\, G_q(s) > 0\}\right) = \lambda_1 \left(\bigcup_{n \in \mathbb{N}} \left\{s \,:\, G_q(s) > \frac{1}{n}\right\}\right), \tag{3.42}$$

$$= \lim_{n \to \infty} \lambda_1 \left(\left\{s \,:\, G_q(s) > \frac{1}{n}\right\}\right). \tag{3.43}$$

It follows that $\lambda_1 \left(\{s \,:\, G_q(s) > 1/n\}\right) > 0$ for some $n \in \mathbb{N}$; define $\beta = \lambda_1 \left(\{s \,:\, G_q(s) > \frac{1}{n}\}\right)$. This leads us to the following calculation

$$\int_0^\infty G_q(s) \ \mathrm{d}s \geq \beta \cdot \frac{1}{n} > 0.$$

Therefore, we can conclude that $\lambda_1 \left(G_q^{-1}(0, \infty)\right) > 0$, then $q \notin \mathcal{M}$. We finish the preamble by considering the mapping:

$$s \to \int_{Z(s)} q(x) \ \mathrm{d}\mu(x), \tag{3.44}$$

which we will show is continuous except at those s corresponding to a level set of positive measure of q_1. It is also monotone decreasing and hence we deduce, from work which

will follow, that it is measurable. Combining with Lemma 3 and Proposition 4 this verifies our claims about G_q.

Proof of Lemma 5: The methods of Burton [17, Lemma 2] yield that $G_q(s) \geq 0$ for every s. The first integral of G_q is the composition of an upper semi continuous mapping (see Lemma 3) with a continuous mapping (see Proposition 4). Moreover σ is continuous except at those s corresponding to level sets of q_1 having positive measure. It follows that $s, \to \int^{\sigma(s)} q_2^{\Delta}(x) \, \mathrm{d}\lambda_1(x)$ is continuous except at those s corresponding to level sets of q_1 having positive measure.

The methods of Lieb and Loss [56] show that upper semi continuous function are Borel measurable functions. It follows immediately that continuous functions are Borel measurable. We deduce that $s \to \int_0^{\sigma(s)} q_2^{\Delta}(x) \, \mathrm{d}\lambda_1(x)$ is the composite of Borel measurable functions therefore measurable.

Noting that $\sigma(s) = \mu(Z(s))$ we look to prove that (3.44) is continuous except at those s corresponding to a level set of positive measure of q_1. We consider the following if $s_n \to s$, where $\mu\{x : q_1(x) = s\} = 0$.

$$\int_{Z(s_n)} q(x) \, \mathrm{d}\mu(x) \quad - \quad \int_{Z(s)} q(x) \, \mathrm{d}\mu(x),$$
$$= \int_{Z(s_n) \setminus Z(s)} q(x) \, \mathrm{d}\mu(x) - \int_{Z(s) \setminus Z(s_n)} q(x) \, \mathrm{d}\mu(x). \quad (3.45)$$

We have that $\mu(Z(s_n) \setminus Z(s)) \to 0$ as $s_n \to s$. Now, the Dominated Convergence Theorem yields that $s \to \int_{Z(s)} q$ is continuous at points of continuity of σ. The mapping is monotone decreasing. The inverse image of $[\alpha, \infty)$, some $\alpha \in \mathbb{R}$, will be an interval which is a Borel set whence the mapping is measurable. Now, from the definition of G_q

we have the sum of two integrals, both of which are measurable and continuous at each s where σ is continuous; therefore, we can deduce that G_q is continuous at each s where σ is continuous and measurable. In particular $G_q^{-1}(0, \infty)$ is a measurable set.

Write $Q = \{(x, s) : 0 \leq s \leq q_1(x)\}$ and let 1_Q denote the characteristic function of Q. Using the methods of Burton [17, Theorem 2] we have that:

$$\int_\Omega qq_1 \, \mathrm{d}\mu \;=\; \int_\Omega \int_0^{q_1(x)} q(x) \, \mathrm{d}\lambda_1(s) \, \mathrm{d}\mu(x), \tag{3.46}$$

$$=\; \int_\Omega \int_0^\infty q(x) \, 1_Q(x, s) \, \mathrm{d}\lambda_1(s) \, \mathrm{d}\mu(x), \tag{3.47}$$

$$=\; \int_0^\infty \int_\Omega 1_Q(x, s) q(x) \, \mathrm{d}\mu(x) \, \mathrm{d}\lambda_1(s), \tag{3.48}$$

$$=\; \int_0^\infty \int_{Z(s)} q(x) \, \mathrm{d}\mu(x) \, \mathrm{d}\lambda_1(s), \tag{3.49}$$

$$\leq\; \int_0^\infty \int_0^{\sigma(s)} q_2^\Delta \, \mathrm{d}\lambda_1(x) \, \mathrm{d}\lambda_1(s), \tag{3.50}$$

$$=\; \int_0^{\mu(\Omega)} q_1^\Delta(x) \, q_2^\Delta(x) \, \mathrm{d}\lambda_1(x). \tag{3.51}$$

By way of explanation equation (3.48) follows by Fubini's theorem and equation (3.49) by noting that for fixed $s \geq 0$:

$$\left\{x : q_1^\Delta(x) \geq s\right\} \;=\; [0, \sigma(s)].$$

Equation (3.50) holds by Burton [17, Lemma 2]. To obtain equation (3.51) we reverse the process of equations (3.46)-(3.49). Noting (3.49) (and an analogous result for $\int q_1^\Delta q_2^\Delta$) we have

$$\int_0^\infty G_q(s)\ \mathrm{d}\lambda_1(s) = \int_0^{\mu(\Omega)} q_1^\Delta(x)\ q_2^\Delta(x)\ \mathrm{d}\lambda_1(x) - \int_\Omega q(x)\,q_1(x)\ \mathrm{d}\mu(x). \qquad (3.52)$$

From the proof of Theorem 3.2.1 we have that $q \in \mathcal{M}$ if and only if $\int_0^\infty G_q = 0$. Given

Burton's result [17, Lemma 2] which states that $G_q(s) \geq 0$ for each s, therefore it follows

that $q \in \mathcal{M}$ if and only if $\lambda_1\left(G_q^{-1}(0,\infty)\right) = 0$. To conclude the proof we consider what

happens when

$\lambda_1\left(G_q^{-1}(0,\infty)\right) > 0$ given that $G_q \geq 0$. From the methods discussed in the preamble

there must exist some $\beta > 0$ such that $\lambda_1\left(G_q^{-1}[\,\beta\,,\,\infty\,)\right) > 0$. For this we get the fol-

lowing

$$\int_0^\infty G_q(s)\ \mathrm{d}\lambda_1(s) > \lambda_1\left(G_q^{-1}[\,\beta\,,\,\infty\,)\right).\beta > 0, \qquad (3.53)$$

thus when $\lambda_1\left(G_q^{-1}(0,\infty)\right) > 0$ we have that $q \notin \mathcal{M}$. This finishes the proof. $\qquad \square$

3.5   Characterisation of $\mathcal{M}$.

Before moving onto the proof of Nealers Theorem we formalise $(\star)$ introduced earlier in

the chapter.

Definition 3.5.1. *Given $\hat{q} \in \mathcal{M}$ we can say that $q \in \mathcal{R}(q_2)$ satisfies $(\star)$ if*

$$(\star) = \begin{cases} q = \hat{q} \text{ on } \Omega \backslash L, \\ \\ (q|_{L_\alpha}) \in \mathcal{R}(\hat{q}|_{L_\alpha}). \end{cases} \tag{3.54}$$

In order to prove Nealers Theorem we split the proof into two parts.

1. If a function satisfies $(\star)$ does it belong to $\mathcal{M}$?

2. Are all elements of $\mathcal{M}$ maximisers of the form $(\star)$?

Proposition 6 deals with the proof of the former case.

Proposition 6. *Let q_1, q_2, $\mathcal{M}$, L and L_n for $n \in I$ be as in the statement of Nealers Theorem. Let $\hat{q} \in \mathcal{M}$. Suppose that for $q \in \mathcal{R}(q_2)$ satisfies $q = \hat{q}$ on $\Omega \backslash L$ and that $q|_{L_n} \in \mathcal{R}(\hat{q}|_{L_n})$ for every $n \in I$. Then $q \in \mathcal{M}$.*

Proof of Proposition 6: Let $\hat{q} \in \mathcal{M}$. Suppose that $q \in \mathcal{R}(q_2)$ satisfies $q = \hat{q}$ on $\Omega \backslash L$ and $q|_{L_\alpha} \in \mathcal{R}(\hat{q}|_{L_\alpha})$ for every $\alpha \in I$. Then

$$\begin{align} \int_\Omega qq_1 &= \int_{\Omega \backslash L} qq_1 + \int_L qq_1, \tag{3.55} \\ &= \int_{\Omega \backslash L} \hat{q}q_1 + \sum_{\alpha \in I} \int_{L_\alpha} l_\alpha q, \tag{3.56} \\ &= \int_{\Omega \backslash L} \hat{q}q_1 + \sum_{\alpha \in I} \int_{L_\alpha} l_\alpha \hat{q}, \tag{3.57} \\ &= \int_\Omega \hat{q}q_1. \tag{3.58} \end{align}$$

We split the integral into the parts where there are level sets of q_1 having positive measure and where there are no level sets of q_1 having positive measure then use definition 3.5.1.

On $\Omega \setminus L$ we have $q = \hat{q}$ to give the first integral in (3.56). On L_α we have $q_1 = L_\alpha$ gives us the second integral in (3.56). $\hat{q}\,|_{L_\alpha}$ and $q\,|_{L_\alpha}$ are rearrangements and thus have the same L^1-norm and this gives us the second integral in (3.57). We use definition 3.5.1 to get (3.58) and this completes the proof. $\qquad\square$

We conclude the chapter with the proof of Nealers Theorem.

Preamble to proof of Nealers Theorem. In this proof we prove that elements of $\mathcal{M}$ are of the form of ($\star$). We split this proof up further by considering the cases:

1. there is exactly one level set of q_1^Δ having positive measure.

2. there are countably many level sets of q_1^Δ having positive measure.

Therefore, we look to prove the result for only one level set having positive measure before generalising to countably many level sets having positive measure doing so through proof by induction. We illustrate our problem below:
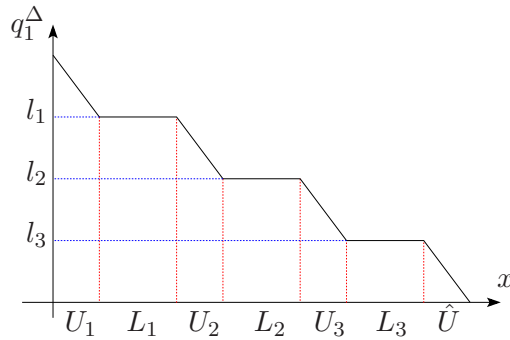
Figure 3.5: A function q_1^Δ with three level sets having positive measure.

We introduce some notation for the "non-level sets" of q_1, i.e. those elements of the domain corresponding to values of q_1 between two level sets having positive measure of q_1.

$$U_1 = \{x \,:\, q_1(x) > l_1\}, \qquad (3.59)$$

$$U_n = \{x \,:\, l_n < q_1(x) < l_{n-1}\} \quad \text{for} \quad n \in \mathbb{N}, n \geq 2, \qquad (3.60)$$

$$\hat{U} = \Omega \setminus \left(\bigcup_{n \in I} L_n \bigcup \bigcup_{n \in I} U_n \right). \qquad (3.61)$$

We note that some of these sets may have zero size; and possibly may be empty.

The formal definition of the mathematical induction problem is as follows:

Induction Problem. *Let $\mathcal{P}(n)$ be the statement that for some $n \in I$*

1. *$q = \hat{q}$ on $U_1, \cdots, U_n$.*

2. *$(q|_{L_i}) \in \mathcal{R}(\hat{q}|_{L_i})$ for $i = 1, \cdots, n$.*

The rest of this proof will go as follows. We will prove $\mathcal{P}(1)$; then we demonstrate $\mathcal{P}(n+1)$ given that $\mathcal{P}(n)$ holds; we will use the methods we used to prove $\mathcal{P}(1)$ to prove $\mathcal{P}(n+1)$. From this we will generalise the result via mathematical induction for every $n \in I$. Finally, we will need to demonstrate that $q = \hat{q}$ on $\hat{U}$ to finish the proof. One point to note that if $\hat{U}$ or U_n for any $n \in I$ has zero size then it is immediate that $\hat{q} = q$(almost everywhere) on that set. In the arguments that follow we will assume that these sets have positive size.

In the latter parts of the proof of Nealers Theorem we make extensive use of the definition of the decreasing rearrangements. The definition of the monotone rearrangement and Lemma 2.5.1(Burton [17, Lemma 2]) gives us the following equation:

$$\int_{U_1} (q - \alpha)_+ \leq \int_{U_1^\Delta} \left(q_2^\Delta - \alpha\right)_+ \text{ for each } \alpha > 0. \tag{3.62}$$

We look to prove equality in (3.62) and hence show that $q_2^\Delta \mid_{U_1^\Delta}$ and $q \mid_{U_1}$ are rearrangements from definition 1.4.1 (When stating $q_2^\Delta \mid_{U_1^\Delta}$ and $q \mid_{U_1}$ are rearrangements we mean $q_2^\Delta \mid_{U_1^\Delta} \circ T$ and $q \mid_{U_1}$ are rearrangements, where $T : U_1 \to U_1^\Delta$ is a measure-preserving transformation.). Unfortunately, we are unable to prove this directly. However, we are able to show that $\int_{U_1} q = \int_{U_1^\Delta} q_2^\Delta$ and noting (3.62) we deduce q belongs to the closed convex hull of $q_2^\Delta \mid_{U_1^\Delta} \circ T$. Similar statements apply to $\hat{q}$ and we can show that both q and $\hat{q}$ are maximisers of $\int_{U_1} f q_1$ relative to the closed convex hull of $q_2^\Delta \mid_{U_1^\Delta} \circ T$. As $q_1 \mid_{U_1}$ has no level sets having positive measure a result of Burton yields that the maximiser is unique, so $q = \hat{q}$ on U_1. Having achieved this result we can deduce that (3.62) holds with U_1 replaced with L_1, U_1^Δ with L_1^Δ. To prove that $q_2^\Delta \mid_{L_1^\Delta}$ and $q \mid_{L_1}$ are rearrangements we show equality for each $\alpha > 0$. We illustrate a key step in our methods with the following figure:
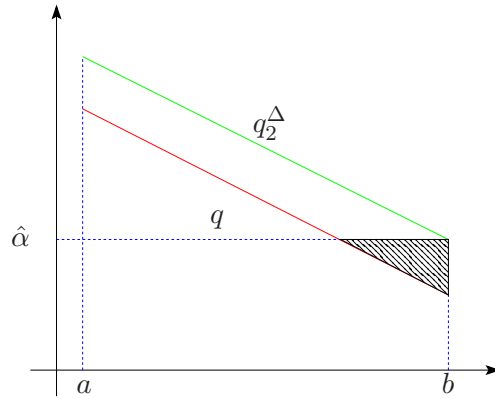
Figure 3.6: An illustration of q_2^Δ and q for $\alpha = \hat{\alpha}$ on the set L_1^Δ.

Let $L_1^\Delta = [a, b]$ and write $\hat{\alpha} = q_2^\Delta(b)$. We are considering the area under the graph

of the function $q_2^\triangle$ and looking to express the area in terms of (3.62). We have the following

$$\int_{L_1^\triangle} q_2^\triangle = \int_{L_1^\triangle} \left(q_2^\triangle - \hat\alpha\right)_+ + \hat\alpha \left(b - a\right). \qquad (3.63)$$

The integral on the right hand side is the area between the line $y = \hat\alpha$ (shown in green in figure 3.6) and the graph of $q_2^\triangle$. The second half of the right hand side is rectangular area. For simplicity suppose $L_1 = [a\,,\,b]$ and q is as shown in figure 3.6. Now,

$$\int_{L_1} \left(q - \hat\alpha\right)_+ + \hat\alpha \left(b - a\right) \geq \int_{L_1} q, \qquad (3.64)$$

as the value of the right hand side of the inequality does not include the area of the shaded triangle. Now, if $\int_{L_1^\triangle} \left(q_2^\triangle - \hat\alpha\right)_+ > \int_{L_1} \left(q - \alpha\right)_+$ then $\int_{L_1^\triangle} q_2^\triangle > \int_{L_1} q$ which contradicts $\int_{L_1^\triangle} q_2^\triangle = \int_{L_1}$. This argument is valid for $\alpha \leq \hat\alpha$ and a different approach deals with $\alpha > \hat\alpha$.

Proof of Nealers Theorem: Let $\mathcal{M}$ be as defined in Definition 3.1.1, and let L, L_n be as described in the notation at the beginning of chapter 3. Suppose that $q \in \mathcal{R}\left(q_2\right)$ satisfies $q = \hat{q}$ on $\Omega \setminus L$, and that $q\,|_{L_n} \in \mathcal{R}\left(\hat{q}\,|_{L_n}\right)$ for every $n \in I$. Then Proposition 6 yields that $q \in \mathcal{M}$.

From the proof of Theorem 3.2.1 we have that

$$\int_0^{\mu(\Omega)} q_1^\Delta \, q_2^\Delta \;=\; \int_\Omega q_1 \hat{q}, \tag{3.65}$$

For the converse, suppose $q \in \mathcal{M}$. Let U_1, U_n, $\hat{U}$ be defined as in (3.59), (3.60) and

(3.61); let $\mathcal{P}(n)$ be the statement of the Induction Problem.

We begin by establishing $\mathcal{P}(1)$. $\mathcal{P}(1)$ is defined as follows:

1. $q = \hat{q}$ on U_1.

2. $(q|_{L_1}) \in \mathcal{R}(\hat{q}|_{L_1})$

We define sets L_n^Δ, L^Δ, U_n^Δ, $\widehat{U^\Delta}$ on $[0, \mu(\Omega)]$ by replacing q_1 with q_1^Δ in the definition

of L_n, L, U_n and $\hat{U}$ respectively. We show $q = \hat{q}$ on U_1. On U_1, q_1 has no level sets of

positive measure and from Lemma 3 $\sigma(s)$ is continuous on (l_1, ∞). Furthermore, from

Lemma 5 the continuity of $\sigma(s)$ implies that $G_q(s)$ is continuous on (l_1, ∞) also. It

must now be the case, given that $G_q(s)$ is continuous, that $G_q(s) = 0$ otherwise we

would be contradicting $\lambda_1 \left(G_q^{-1}(0, \infty)\right) = 0$; we show why this is the case. If $G(\hat{s}) > 0$

for $\hat{s} \in (l_1, \infty)$ then given the continuity of $G_q(s)$ on (l_1, ∞) we have that $G_q(s) > 0$

for $s \in (\hat{s} - \epsilon, \hat{s} + \epsilon)$ for $\epsilon > 0$. We conclude that $q \notin \mathcal{M}$ for $G(s) > 0$ by Lemma 5 and

this illustrates why $G_q(s) = 0$ for $s \in (l_1, \infty)$.

Now, for $s > l_1$ we have the following

$$\int_s^\infty G_q(s) \, \mathrm{d}\lambda_1(s) \;=\; 0 \tag{3.66}$$

We consider the first integral of $\int_s^\infty G_q(s) \, \mathrm{d}\lambda_1(s)$, i.e. $\int_0^{\sigma(s)} q_2^\Delta(x) \, q_1^\Delta(x) \, \mathrm{d}\lambda_1(x)$. We

note that $\lambda_1 \left(U_1^\Delta \setminus [0, \sigma(s)] \right) \to 0$ as $s \to l_1 +$. Using this and the Dominated Convergence Theorem we deduce that

$$\int_0^{\sigma(s)} q_2^\Delta(x) \, q_1^\Delta(x) \, \mathrm{d}\lambda_1(x) \to \int_{U_1^\Delta} q_2^\Delta(x) \, q_1^\Delta(x) \, \mathrm{d}\lambda_1(x). \tag{3.67}$$

Similarly, for the second integral of $G_q(s)$ we note that $\mu(U_1 \setminus Z(s)) \to 0$ as $s \to l_1 +$ then we make further use of the Dominated Convergence Theorem to yield

$$\int_{Z(s)} q(x) \, q_1(x) \, \mathrm{d}\mu(x) \to \int_{U_1} q(x) \, q_1(x) \, \mathrm{d}\mu(x). \tag{3.68}$$

Using (3.67) and (3.68) the definition of G_q, and (3.66), gives us the following

$$\int_{U_1^\Delta} q_2^\Delta \, q_1^\Delta - \int_{U_1} q \, q_1 = 0. \tag{3.69}$$

Considering $G_q(s) = 0$ for $s > l_1$ similar arguments to the above give us

$$\int_{U_1^\Delta} q_2^\Delta - \int_{U_1} q = 0. \tag{3.70}$$

Using the definition of the decreasing rearrangement, and Lemma 2.5.1 yields (3.62). From equations (3.70) and (3.62) we have that our maximiser is an element of the following set, which is the closed convex hull of a set of rearrangements. To this end we

have:

$$q \in \left\{ f \geq 0 : \int_{U_1} (f - \alpha)_+ \leq \int_{U_1^\triangle} \left(q_2^\triangle - \alpha \right)_+ \text{ for each } \alpha > 0, \int_{U_1} f = \int_{U_1^\triangle} q_2^\triangle \right\}. \quad (3.71)$$

Let $T : U_1 \to U_1^\triangle$ be a measure-preserving transformation. Now, from the characterisation of the closed convex hull of the set of rearrangements by Douglas [29, Theorem 2.1] we have that $q \in \overline{\text{conv} \left(\mathcal{R} \left(q_2^\triangle \vert_{U_1^\triangle} \circ T \right) \right)}$. Therefore, from equation (3.69) q is a maximiser of $\int_{U_1} f q_1$ relative to $\overline{\text{conv} \left(\mathcal{R} \left(q_2^\triangle \vert_{U_1^\triangle} \circ T \right) \right)}$. Moreover, so is $\hat{q}$. Given that $q \in \overline{\text{conv} \left(\mathcal{R} \left(q_2^\triangle \vert_{U_1} \circ T \right) \right)}$ and from the methods of Burton [17, Theorem3] there is a unique maximiser over $\overline{\text{conv} \left(\mathcal{R} \left(q_2^\triangle \vert_{U_1} \circ T \right) \right)}$, whence $q = \hat{q}$ on U_1. We've proven the first half of $\mathcal{P} (1)$, we move onto the second half of $\mathcal{P} (1)$.

We already know the following:

$$\int_{l_1}^{\infty} G_q (s) \, d\lambda_1 (s) = \int_{U_1^\triangle \cup L_1^\triangle} q_2^\triangle q_1^\triangle - \int_{U_1 \cup L_1} q \, q_1 = 0. \quad (3.72)$$

Given the equality on U_1 between q and $q_2^\triangle$ on U_1 equation (3.72) reduces to the following

$$\int_{L_1^\triangle} q_2^\triangle q_1^\triangle = \int_{L_1} q \, q_1. \quad (3.73)$$

From the definitions of L_1 and $L_1^\triangle$, we have that on the level set the following holds:

$$l_1 \int_{L_1^\Delta} q_2^\Delta = l_1 \int_{L_1} q. \tag{3.74}$$

We have two options, either $\int_{L_1^\Delta} q_2^\Delta = \int_{L_1} q$ or we have that $l_1 = 0$. If $l_1 = 0$ then $\Omega = U_1 \bigcup L_1$. Now, $q \in \mathcal{R}(q_2)$ and we have that $q\mid_{U_1} = \hat{q}\mid_{U_1}$ so $q\mid_{L_1} \in \mathcal{R}(\hat{q}\mid_{L_1})$. For this special case this is Nealers Theorem.

Write $\hat{\alpha} = q_2^\Delta \left(\lambda_1 \left(U_1^\Delta \bigcup L_1^\Delta \right) \right)$. Let $\alpha \in (\, 0\, , \, \hat{\alpha}\,]$. Suppose for a contradiction that

$$\int_{L_1^\Delta} \left(q_2^\Delta - \alpha \right)_+ > \int_{L_1} \left(q - \alpha \right)_+ , \tag{3.75}$$

then we have:

$$\int_{L_1^\Delta} q_2^\Delta = \int_{L_1^\Delta} \left(q_2^\Delta - \alpha \right)_+ + \alpha\, \lambda_1 \left(L_1^\Delta \right), \tag{3.76}$$

$$> \int_{L_1} \left(q - \alpha \right)_+ + \alpha\, \mu\left(L_1 \right), \tag{3.77}$$

$$\geq \int_{L_1} q. \tag{3.78}$$

Equation (3.76) comes by finding the area of the green function from figure 3.6. Inequality (3.77) comes as a result of equation (3.75). From equation (3.74), equation (3.78) yields a contradiction, whence for $\alpha \in [0\, , \, \hat{\alpha}]$ we have equality in (3.62). We now look to show the following inequality

$$\int_{L_1^\Delta} \left(q_2^\Delta - \alpha\right)_+ > \int_{L_1} \left(q - \alpha\right)_+ , \tag{3.79}$$

for $\alpha \geq \hat{\alpha}$. The methods of Douglas [29, Lemma 2.4] yield that the functions W_1, W_2 : $[0, \infty) \to \mathbb{R}$ defined by:

$$W_1\left(\alpha\right) = \int_{L_1} \left(q - \alpha\right)_+ , \ W_2\left(\alpha\right) = \int_{L_1^\Delta} \left(q_2^\Delta - \alpha\right)_+ , \tag{3.80}$$

are convex and continuous on the interiors of their effective domains, namely $(0, \infty)$. It follows that the set

$$\Theta(\alpha) = \{\alpha \ : \ W_2\left(\alpha\right) > W_1\left(\alpha\right)\} , \tag{3.81}$$

is open in $(0, \infty)$, whence it is a countable union of disjoint open intervals. Take a (possibly unbounded) connected component of Θ, denoted (α_0, α_1) where α_1 may be ∞. Now $W_1\left(\alpha_0\right) = W_2\left(\alpha_0\right)$, and $\alpha_0 \geq \hat{\alpha}$. Moreover $W_2\left(\alpha\right) > W_1\left(\alpha\right)$ for every $\alpha \in (\alpha_0, \alpha_1)$. If for every Borel set $B \subset (\alpha_0, \infty)$ we have

$$\mu\left(\{x \in L_1 \ : \ q\left(x\right) \in B\}\right) \leq \lambda_1 \left(\left\{x \in L_1^\Delta \ : \ q_2^\Delta\left(x\right) \in B\right\}\right) , \tag{3.82}$$

then $W_2\left(\alpha\right) > W_1\left(\alpha\right)$ for $\alpha \in (\alpha_0, \alpha_1)$ implies that $W_2\left(\alpha_0\right) > W_1\left(\alpha_0\right)$. The point is

that $W_1'(\alpha) = -\lambda_1(\{x \mid q(x) > \alpha\})$ for almost every α (and the same is true for W_2, so (3.82) would imply $W_1'(\alpha) \geq W_2'(\alpha)$ if $W_1(\alpha_0) \geq W_2(\alpha_0)$. This contradiction yields a Borel set $B \subset (0, \infty)$ such that

$$\mu(\{x \in L_1 \,:\, q(x) \in B\}) > \lambda_1\left(\{x \in L_1^\Delta \,:\, q_2^\Delta(x) \in B\}\right). \tag{3.83}$$

Now:

$$\mu(\{x \in \Omega \,:\, q(x) \in B\}) \;\geq\; \mu\left(\left\{x \in U_1 \bigcup L_1 \,:\, q(x) \in B\right\}\right) \tag{3.84}$$

$$> \;\lambda_1\left(\left\{x \in U_1^\Delta \bigcup L_1^\Delta \,:\, q_2^\Delta(x) \in B\right\}\right) \tag{3.85}$$

$$= \;\lambda_1\left(\left\{x \in [0, \mu(\Omega)] \,:\, q_2^\Delta(x) \in B\right\}\right), \tag{3.86}$$

which contradicts $q \in \mathcal{R}(q_2)$. By way of explanation, we note that the strict inequality in (3.85) follows because $q\mid_{U_1}$ and $q_2^\Delta\mid_{U_1^\Delta}$ are rearrangements and (3.83) holds, and the equality in (3.86) follows because:

$$\lambda_1\left(\left\{x \in [0, \mu(\Omega)] \setminus \left(U_1^\Delta \bigcup L_1^\Delta\right) \,:\, q_2^\Delta(x) > \hat{\alpha}\right\}\right) = 0. \tag{3.87}$$

It follows that the set Θ is empty, that is (3.62) holds with equality for each $\alpha > 0$. We deduce that $q\mid_{L_1}$ and $q_2^\Delta\mid_{L_1^\Delta}$ are rearrangements. Similar statements are true for $\hat{q}$, so $q\mid_{L_1} \in \mathcal{R}(\hat{q}\mid_{L_1})$, verifying $\mathcal{P}(1)$(ii). Combining the above, $\mathcal{P}(1)$ holds.

Having proved the first inductive step we move onto the main inductive hypothesis.

So, we assume that $\mathcal{P}(n)$ holds for $n \in I$; we look to prove that $\mathcal{P}(n+1)$ holds. Writing $Y = \bigcup_{i=1}^{n} (U_i \bigcup L_i)$, $\mathcal{P}(n)$ implies that $q \mid_{\Omega \setminus Y} \in \mathcal{R}\left(\hat{q} \mid_{\Omega \setminus Y}\right)$ and $\hat{q}$, q restricted to $\Omega \setminus Y$ are maximisers of $\int_{\Omega \setminus Y} f q_1$ relative to $f \in \mathcal{R}\left(\hat{q} \mid_{\Omega \setminus Y}\right)$. $\mathcal{P}(n+1)$ holds from arguments used in the first inductive step. It follows by induction that $\mathcal{P}(n)$ holds for every $n \in I$. Finally, noting that $\hat{q}$, q restricted to $\hat{U}$ are both maximisers of $\int_{\hat{U}} f q_1$ relative to $f \in \mathcal{R}\left(\hat{q} \mid_{\hat{U}}\right)$ and that q_1 restricted to $\hat{U}$ has no level sets of positive measure, Theorem 3.2.1 and Proposition 2 yield that $q = \hat{q}$ on $\hat{U}$. This completes the proof. $\square$

Chapter 4

Optimal Mass Transfer Problems.

4.1   Introduction.

We begin in section 4.2 by considering the methods which are integral to the concept of optimal mass transfer. In section 4.2.1 we introduce the concept of allowable strategies and transference plans, moving onto discuss admissible allowable strategies.

One of the major advances in optimal mass transfer theory is that of the Monge-Kantorovich problem; this is the content of section 4.3. The ' problem is a linear minimisation with convex constraints (or maximisation if you consider the dual problem) ensuring Kantorovich addressed the issues of Monge's problem. We define the Monge-Kantorovich problem formally before stating which properties yield optimal solutions to the Monge-Kantorovich problem.

One of the major advantages of the Monge-Kantorovich problem is the duality associated with the problem. In section 4.4 we discuss the dual formulation and define it formally. We demonstrate some of the advantages of both the relaxed methods and the dual comparatively to the original problem of Monge. Furthermore, we note methods

that have been used to solve the dual problem. We conclude section 4.4 by defining the notions of c-transforms, c-cyclically monotone and conjugate convex functions. The chapter is brought to a close by considering the results which exist for the various cost functions in section 4.5.

4.2   Methods of Optimal Mass Transfer.

We recall that the Monge problem is how to transfer mass from a set U to a set V of equal size, minimising cost. Let (U, μ) and (V, ν) be finite measure spaces with $\mu(U) = \nu(V)$.

4.2.1   Allowable Strategies.

Allowable strategies is a size-preserving way of mapping elements U to elements V. We write S for the set of allowable strategies. Equivalently, we can say that S is the set of measure-preserving mappings from U to V.

If we denote the transference plans by π this is a probability measure on the product space $U \times V$. π measures how much of the mass has been transferred from location x to location y. Allowable strategies can be though of as a particle being split up into several possible destinations. Roughly speaking points are mapped to probability distributions. We restrict to measures on $U \times V$ that have μ and ν as their marginals, thus the following is satisfied:

$$\pi[X \times V] = \mu[X] \qquad\qquad \pi[U \times Y] = \nu[Y], \qquad (4.1)$$

for all measurable subsets X of U and Y of V.

Definition 4.2.1. *The set of all measure-preserving plans is given by*

$$\Pi(\mu, \nu) = \{\pi \in P(U \times V); \ \ (4.1) \ \textit{holds for all measurable} \ X \subset U, Y \subset V\}, \qquad (4.2)$$

where $\mu(U) = \nu(V) = 1$.

N.B. Villani [71] states that the set is always nonempty as the tensor product $\mu \otimes \nu (A \times B) = \mu(A)\,\nu(B)$, where $A \subset V$, $B \subset V$ lies in $\Pi(\mu, \nu)$.

Ambrosio [2, Proposition 2.1] illustrates the difference between measure-preserving plans and measure-preserving maps. This also provides insight into the form required to show that solutions to the Monge-Kantorovich problem are also solutions to the Monge problem. Knowing that the support of a measure is the smallest closed set with measure equal to that of the whole space we illustrate the connection between measure-preserving plans and maps.

100

Proposition 4.2.2. *Any Borel measure-preserving map $s \,:\, (U\,,\,\mu) \to (V\,,\,\nu)$ induces a measure-preserving plan $\gamma_s \in \Pi\,(\mu\,,\,\nu)$ defined by*

$$\gamma_s \,:=\, (id \times s)_{\#}\,\mu. \tag{4.3}$$

Conversely, a measure-preserving plan γ is induced by a measure-preserving map if γ is concentrated on a measurable graph Γ.

Proof: Follow the methods of Ambrosio [2]. □

We illustrate what we mean in Proposition 4.2.2 via the example that follows. If we consider a set $[0\,,\,1] \times [0\,,\,1]$, where $\lambda_1\big|_{[0\,,\,1]}$ are the marginals. Measure-preserving mappings given by $s \,:\, ([0\,,\,1]\,,\,\lambda_1) \to ([0\,,\,1]\,,\,\lambda_1)$ generate measure-preserving plans through $(id \times s)_{\#}\,\lambda_1 \,=\, \pi.$, where $(id \times s\,(x)) \,=\, (x\,,\,s\,(x))$. We can represent this graphically in the following way:
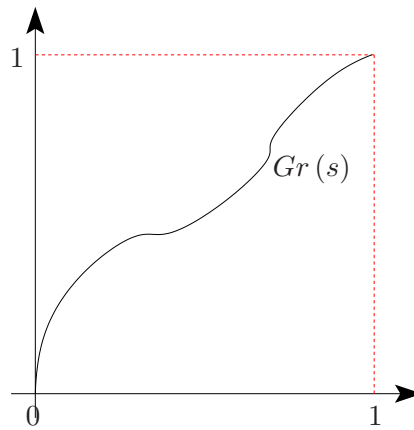
Figure 4.1: The graph of the function $s_1 \,=\, s_1\,(x)$

In this case the support of π is a graph. Thus a measure-preserving mapping s generates a measure-preserving plan $(id \times s)_{\#}\,\lambda_1$. We now consider an example which illustrates

that a measure-preserving plan need not be of the form $(id \times s)_{\#} \lambda_1$.
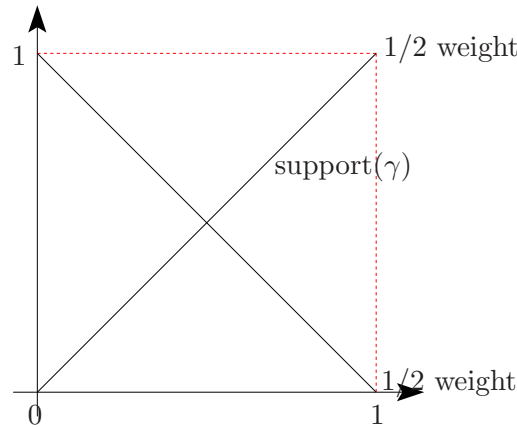
Figure 4.2: A measure-preserving plan which is not supported on a graph.

It is apparent that this is not a graph of a function. For all values, apart from $x = 1/2$, it is not possible to represent $s_1(x)$ as a function of just one x. Whence a particle is sent to x with probability $1/2$ and to $1 - x$ with probability $1/2$. We note that this s is a measure-preserving plan, however, is not of the form $(id \times s_1)_{\#} \lambda_1$.

Finally, for the work that follows we must consider the definition of c-concavity, where $c(x, y)$ is a given cost function.

Definition 4.2.3. *A function* $\varphi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}$ *is* **c-concave** *if there exists some function* $\eta : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}$ *for which*

$$\varphi(x) = \inf_{y \in \mathbb{R}^n} c(x, y) - \eta(y). \tag{4.4}$$

For the cost function $c(x, y) = |x - y|^2 / 2$ the c-concavity of φ is equivalent to the concavity of $\varphi(x) - |x|^2 / 2$ or convexity of $\psi(x) = |x|^2 / 2 - \varphi(x)$ and not to the concavity of φ. We demonstrate this, showing that the Brenier map is a special case of

what Gangbo and McCann proved for strictly convex cost functions. For the quadratic cost function we have that $\nabla h = id$, $h(x) = |x|^2/2$, and the map becomes $s(x) = \nabla \psi(x)$. From the c-concavity of φ we have that ψ becomes

$$\psi(x) = |x|^2/2 - \varphi(x) = |x|^2/2 - \inf_{y \in \mathbb{R}^n}\left(|x-y|^2/2 - \eta(y)\right) \tag{4.5}$$

$$= |x|^2/2 - \inf_{y \in \mathbb{R}^n}\left(|x|^2/2 + |y|^2/2 - x \cdot y - \eta(y)\right) \tag{4.6}$$

$$= \sup_{y \in \mathbb{R}^n} x \cdot y + \eta(y) - \frac{|y|^2}{2}, \tag{4.7}$$

For fixed $y \in \mathbb{R}$, $x \to x \cdot y + \eta(y) - \frac{|y|^2}{2}$ is the sum of a continuous linear function and a constant, thus an affine continuous function and 4.7 Ekeland and Témam [34, Proposition 3.1.] yields that $|x|^2/2 - \varphi(x)$ is both convex and lower semi continuous. Thus the potential ψ of the Brenier map has the required properties. We can think of c-concavity as an adaption of the notion of convexity(or concavity) to the cost function c.

4.3   Monge-Kantorovich Problem.

We define the Monge-Kantorovich problem:

Definition 4.3.1. *Let (U, μ) and (V, ν) be finite Borel measure spaces with $\mu(U) = \nu(V)$. Let $\pi \in P(U \times V)$, then $\pi \in \Pi(\mu, \nu)$ if (4.1) is satisfied. The Monge-Kantorovich optimal transfer problem is*

$$\inf_{\pi \in \Pi(\mu,\nu)} \int_{U \times V} c(x,y)\, d\pi(x,y) \equiv \inf_{\pi \in \Pi(\mu,\nu)} I(\pi). \tag{4.8}$$

We note that if a solution to equation (4.8) is of the form $(id \times s)_{\#}\,\mu$, then s minimises

equation (1.13) and therefore we have a solution to the Monge problem. Further to

this, if the minimiser to the Monge-Kantorovich problem is unique and is of the form

$(id \times s)_{\#}\,\mu$ then the minimiser for Monge's problem is also unique.

Relaxation in this sense means that the infimum that Kantorovich's problem encom-

passes a larger class of objects than the infimum from equation (1.13). If $s\,:\,U \to V$ is

a measure-preserving mapping and $(id \times s)_{\#}\,\mu\,=\,\pi$, we have:

$$\int_{U \times V} c(x\,,\,y)\,\mathrm{d}(id \times s)_{\#}\mu\,=\,\int_{U} c(x\,,\,s(x))\,\mathrm{d}\mu. \tag{4.9}$$

Instead of minimising the cost functions over the set of push forward measures(or

measure-preserving mappings) we minimise over probability measures on $U \times V$ with

prescribed marginals μ and ν or measure-preserving plans. We can use the above infor-

mation to form an upper bound for the Monge problem, as follows:

$$\min_{\pi \in \Pi} \int_{U \times V} c(x\,,\,y)\,\mathrm{d}\pi(x\,,\,y) \leq \inf_{s \in S} \int_{U} c(x\,,\,s(x))\,\mathrm{d}\mu(x). \tag{4.10}$$

We have the following result by Gangbo and McCann [39, Proposition 2.1] which justifies

the use of minimum in (4.10) for non-negative lower semi continuous c:

Proposition 4.3.2. *If $c(x\,,\,y) \geq 0$ is lower semi continuous on $U \times V$ and measures*

$\mu \in \mathcal{P}(U)$ and $\nu \in \mathcal{P}(V)$ then there is at least one optimal measure $\gamma \in \mathcal{P}(U \times V)$ with

marginals μ and ν.

4.4   The duality of the Monge-Kantorovich Problem.

We define the Kantorovich duality problem:

Definition 4.4.1. *Let $\phi \in L^1(U, \mu)$, $\psi \in L^1(V, \nu)$. We define*

$$J(\varphi, \psi) = \int_U \varphi \, d\mu + \int_V \psi \, d\nu. \tag{4.11}$$

Let Φ_c be the set of all measurable functions $(\varphi, \psi) \in L^1(U, \mu) \times L^1(V, \nu)$, where φ and ψ are integrable pairs satisfying

$$\varphi(x) + \psi(y) \leq c(x, y), \tag{4.12}$$

for μ almost every $x \in U$, ν almost every $y \in V$. Then, the Monge-Kantorovich dual problem is

$$\sup_{(\varphi, \psi) \in \Phi_c} J(\varphi, \psi). \tag{4.13}$$

The importance of the dual problem to optimal mass transfer is that it gives us a simpler way to find solutions. To this end we have the following duality theorem:

Duality Theorem. $c : U \times V \to \mathbb{R}_+ \bigcup \{\infty\}$ *be a lower semi continuous cost function.*

Then, the duality theorem states that

$$\inf_{\pi \in \Pi(\mu,\nu)} I[\pi] = \sup_{(\varphi,\psi) \in \Phi_c} J(\varphi,\psi), \tag{4.14}$$

where Φ_c is as defined in Definition 4.4.1.

Proof: See, for example, Villani [71, Section 1.1.7]. □

A key concept in the study of the dual problem is the c-transform of a function:

Definition 4.4.2. *For $\varphi : U \to \mathbb{R} \bigcup \{-\infty\}$ we define the c-transform of φ by*

$$\varphi^c(y) = \inf_{x \in U} c(x,y) - \varphi(x). \tag{4.15}$$

We can rearrange equation (4.15) to show the following

$$\varphi(x) + \varphi^c(y) \le c(x,y), \tag{4.16}$$

which we can see satisfies equation (4.12) and therefore we have that $(\varphi,\varphi^c) \in \Phi_c$ which provides us with information on how to solve the Monge-Kantorovich problem. Moreover, the functions φ and φ^c are known as c-conjugate. Furthermore, we have $\psi^c = \psi^{ccc}$ from Villani [71] and $\psi^{cc} = \psi$ if and only if ψ is c-concave. We note that ψ^c is c-concave by its definition.

If an optimal solution exists, we have for any c that: for (x, y) and (x', y') from

$$c(x, y) + c(x', y') \leq c(x, y') + c(x', y),\qquad(4.17)$$

if this was not the case it would be more efficient to pair x with y' and x' with y and the optimality would be compromised. For an optimiser, from Gangbo and McCann [39], we use (4.17) to deduce c-cyclic monotonicity; as stated by Villani in [72] "a c-cyclically monotone plan is a plan that cannot be improved" .

Definition 4.4.3. *A subset $\Gamma \subset U \times V$ is said to be c-cyclically monotone if, for any $N \in \mathbb{N}$, and any family $(x_1, y_1), \ldots, (x_N, y_N)$ of points in Γ, holds the following in-equality*

$$\sum_{i=1}^{N} c(x_i, y_i) \leq \sum_{i=1}^{N} c(x_i, y_{i+1}).\qquad(4.18)$$

where $y_{N+1} = y_1$

To solve the dual problem we need to show that problem (4.13) has a solution in the form of c-concave conjugate functions (φ, ψ), i.e. $\varphi^c = \psi$, with ψ c-concave. To then show that the solution of the dual problem generalises to the original Monge-Kantorovich problem we must show that a minimiser π is concentrated on the set

$$\{(x, y) \in U \times V : \varphi(x) + \psi(y) = c(x, y)\}.\qquad(4.19)$$

Villani notes [71] that for the quadratic cost the support of minimiser π is cyclically monotone. We again call on Villani [71] to formalise the process of yielding a solution for the dual Monge-Kantorovich problem. We introduce the following notation:

Definition 4.4.4. $\tilde{\Phi}$ *is the set of all pairs* $(\varphi\,,\,\psi)$ *in* $L^1\,(\,\mathrm{d}\mu)\,\times\,L^1\,(\,\mathrm{d}\nu)$ *(with values in* $\mathbb{R}\bigcup\{+\infty\}$*) such that all* $x,\,y,$

$$x\cdot y\leq\varphi\,(x)\,+\,\psi\,(y)\,.\qquad\qquad(4.20)$$

We get (4.20) by unwrapping $|x\,-\,y|^2\,/2$ and by utilising (4.12) and (4.16). We use definition 4.4.4 in illustrating how to solve the dual Monge-Kantorovich problem.

Theorem 4.4.5. *Let* $\mu\,,\,\nu$ *be two probability measures on* $\mathbb{R}^n$ *with finite second order moments. Let* $\tilde{\Phi}$ *be defined as in definition 4.4.4. Then there exists a pair* $(\varphi\,,\,\varphi^*)$ *of lower semi continuous proper conjugate convex functions on* $\mathbb{R}^n$*, such that*

$$\inf_{\tilde{\Phi}}J\,=\,J\,(\varphi\,,\,\varphi^*)\,.\qquad\qquad(4.21)$$

Now the question is whether (4.19) (or (4.20) with equality) defines a graph? For the quadratic cost the answer is yes.

4.5   Existence of results for specific cost functions.

The first result comes from Sudakov [67] who stated that for Monge's original minimal problem there was an optimal map which existed. It was apparent that the solution of

the cost function would not be unique and while other researchers, such as Trudinger, Wang, Evans and Gangbo [35; 70], used other methods to make Sudakov's work [67] more rigorous; there was a limit to what could be proved. Sudakov used probability theory to prove that the dual problem of the Monge-Kantorovich problem had a solution for the cost function chosen by Monge. Furthermore, Brenier [14], stated the solution took the form, for a cost function of the form $c(x, y) = |x - y|^r$ for a given $r > 0$ and for a solution γ, $\mathrm{d}\gamma = \delta(y - s(x)) \ \mathrm{d}\mu(x)$ where s is a one-to-one measure-preserving mapping. This implies that it can be generalised from the dual problem to the Monge problem and we have existence of a solution. The claim by Sudakov was that he had a solution for any distance cost function. However, Ambrosio [4] noted we have that one of the important parts of the proof was not true in general. The approach was to decompose $\mathbb{R}^n$ into regions and then solve the optimal mass transfer problem in any one of these regions. The map would be obtained by joining these regions together. However, this work was dependent on the measure in the original set X, μ being absolutely continuous and this was not the case. This problem was dealt with by Ambrosio [4, Theorem 6.5] and Monge's problem now had a solution.

The cost function $c(x, y) = |x - y|^2 /2$ was the first where a unique minimiser was proven; there was a surge of interest after Sudakov's original work. Kantorovich's relaxed formulation allowed researchers to give more precise detail on the quadratic cost function. Furthermore, work by Brenier [13] and Knott and Smith [53, Theorem 2.1] characterised the mapping as the gradient of a convex function. Brenier's work came when he was studying incompressible fluid mechanics; the link between the topics lead to links between optimal mass transfer and other topics.

The two approaches when proving results about the quadratic cost functions can be

decomposed into those that use Kantorovich's duality theorem and those that do not. We concentrate on the use of Kantorovich's duality theorem. To this end we state the work proved by Brenier, Knott & Smith [71, Theorem 2.12]:

Theorem 4.5.1. *Let μ, ν be probability measures on $\mathbb{R}^n$, with finite second order moments in the sense of Villani [71]. We consider the Monge-Kantorovich transportation problem with a quadratic cost function $c(x, y) = |x - y|^2$. Then,*

1. **(Knott-Smith optimality criterion)** *$\pi \in \Pi(\mu, \nu)$ is optimal if and only if there exists a convex lower semi continuous function φ such that*

$$Supp(\pi) \subset Graph(\partial\varphi). \tag{4.22}$$

2. **(Brenier's theorem)** *If μ does not give mass to small sets then there is a unique optimal π, which is*

$$\mathrm{d}\pi(x, y) = \mathrm{d}\mu(x)\,\delta\left[y = \nabla\varphi(x)\right], \tag{4.23}$$

where δ is the Dirac mass.

We now see from section 2.4.1 why the gradient of a convex function is known as Brenier's map. Moreover, from section 2.4.1 we see its applications to vector valued rearrangements, specifically to the monotone rearrangements of vector valued functions. We cite the same section for an example of a Brenier map.

One corollary of Theorem 4.5.1 is that, given the form of the solution in equation (4.23), $\nabla\varphi$ is also the unique solution of the Monge transportation problem (for the quadratic cost).

Moving forward it is now pertinent to consider whether there are specific features of cost functions which yield existence and optimality of the solution. To this end we consider both strictly convex and strictly concave cost functions. We begin with a cost function of the form $c(x\,,\,y)\,=\,h(x\,-\,y)$ where h is strictly convex.

We use definition 2.7.1 and the methods of Ekeland and Témam [34, Definition 2.1] to define strict convexity:

Definition 4.5.2. *For $\varphi\,,\,x\,,\,y$ and λ as defined in definition 2.7.1, φ is said to be strictly convex when we have strict inequality in equation (2.31), i.e.*

$$\varphi\left(\lambda x\,+\,(1\,-\,\lambda)y\right) < \lambda\varphi(x)\,+\,(1\,-\,\lambda)\varphi(y), \tag{4.24}$$

for $0 < \lambda < 1$ and for $x \neq y$, where the right hand side is defined.

We need to consider another definition from work by De Lellis [54, Definition 4.1]

Definition 4.5.3. *A k-dimensional Borel set $E \subset \mathbb{R}^n$ is called rectifiable if there exists a countable family $\{\Gamma_i\}_i$ of k-dimensional Lipschitz graphs such that $\mathcal{H}^k\left(E\backslash\bigcup\Gamma_i\right)\,=\,0$, where $\mathcal{H}^k$ is a k-dimensional Hausdorff measure.*

We state a result parallel to Theorem 4.5.1 from Gangbo and McCann [39, Theorem 1.2]:

Theorem 4.5.4. *Fix $c(x, y) = h(x - y)$ with h satisfying the following from Gangbo and McCann [39]*

(H1) $h : \mathbb{R}^d \to [\, 0\, , \infty\,)$ *is strictly convex.*

To handle measures with unbounded support, we also assume that the cost grows superlinearly at large $|x|$ while the curvature of its level sets decays:

(H2) given height $r > \infty$ and angle $\theta \in (0\, , \pi)$: whenever $\mathbf{p} \in \mathbb{R}^{\mathbf{d}}$ is far enough from the origin, one can find a cone

$$K\,(r\, , \theta\, , \hat{\mathbf{z}}\, , \mathbf{p})\, =\, \left\{ \mathbf{x} \in \mathbb{R}^{\mathbf{d}} : |\mathbf{x} - \mathbf{p}|\, |\mathbf{z}| \cos\,(\theta/2) \leq \langle \mathbf{z}\, , \mathbf{x} - \mathbf{p} \rangle \leq r\, |\mathbf{z}| \right\}$$

with vertex at $\mathbf{p}$ (and $\mathbf{z} \in \mathbb{R}^d \setminus \{0\}$) on which $h\,(\mathbf{x})$ assumes its maximum at $\mathbf{p}$;

(H3) $\lim h\,(\mathbf{x})\,/\,|\mathbf{x}|\, =\, +\infty$ *as* $|\mathbf{x}| \to \infty$,

and Borel probability measures μ and ν on $\mathbb{R}^d$. If μ is absolutely continuous with respect to a d-dimensional Lebesgue measure then

1. *there is a c-concave function φ on $\mathbb{R}^d$ for which the map $s := id - (\nabla h)^{-1} \circ \nabla \varphi$ pushes forward μ to ν;*

2. *this map $s(x)$ is uniquely determined (μ-almost everywhere) by 1.;*

3. *the joint measure $\gamma := (id \times s)_{\#}\, \mu$ is optimal;*

4. *γ is the only optimal measure in $\Gamma(\mu\, , \nu)$ - (except trivially when $C(\gamma) = \infty$).*

If μ fails to be absolutely continuous with respect to Lebesgue but vanishes on rectifiable sets of dimension $d - 1$, then (1)-(4) continue to hold provided $h \in C_{loc}^{1,1}\left(\mathbb{R}^d\right)$.

Theorem 4.5.4 states that γ is both an optimal and unique solution to the Monge-Kantorovich problem. Furthermore, the optimal solution takes the form as noted in (4.23), whence the Monge problem has a solution. The map s depends on $(\nabla h)^{-1}$ which illustrates that strict convexity is essential for the uniqueness of the map.

Gangbo and McCann [39] stated that "concave functions of the distance $|x - y|$ form the most interesting class of costs"; the concavity ensures the invertibility of ∇h. Unfortunately, there are some distinct differences between the result proven for strictly convex and strictly concave cost functions.

The strictly concave cost functions take the form $c(x\,,\,y) \,=\, l\,(|x\,-\,y|)$ where $l \geq 0$ and is strictly concave. Gangbo and McCann [39] noted that we have that optimal maps will not be smooth and the optimal maps will have an intricate and unexpected structure, where McCann [58] explores this structure on the line. The setting provides a model for applications where paths don't cross, e.g. a railway route. McCann [58] notes that the concavity "reflects a shipping tariff that increases with the distance, even while the cost per mile shipped goes down". Hence, the concavity prefers a long trip and a short trip over two trips of average length.

We consider the following result for strictly concave cost functions [39, Theorem 1.4]:

Theorem 4.5.5. *Use $l \,:\, [0\,,\,\infty) \to [0\,,\,\infty)$ strictly concave to define $c(x\,,\,y) \,:=\, l\,(|x\,-\,y|)$. Let μ and ν be Borel probability measures on $\mathbb{R}^d$ and define $\mu_0 \,:=\, [\mu\,-\,\nu]_+$ (where $[\,\cdot\,]_+$ denotes the positive part of the measure) and $\nu_0 \,:=\, [\nu\,-\,\mu]_+$. If μ_0 vanishes on $\mathrm{spt}\,\nu_0$ and on rectifiable sets of dimension $d\,-\,1$ then*

1. *the c-transform $\psi \,:\, \mathbb{R}^d \to \mathbb{R}$ of some function on $\mathrm{spt}\,\nu_0$ induces a map $s \,:=\, id\,-\,(\nabla h)^{-1} \circ \nabla \psi$ which pushes μ_0 forward to ν_0;*

2. *the map $s(x)$ is uniquely determined μ_0-almost everywhere by 1.*

3. *there is a unique optimal measure γ in $\Gamma(\mu, \nu)$ - except where $C(\gamma) = \infty$*

4. *the restriction of γ to the diagonal is given by $\gamma_d = (id \times id)_\#(\mu - \mu_0)$;*

5. *the off diagonal part of $\gamma = \gamma_d + \gamma_0$ is given by $\gamma_0 = (id \times s)_\# \mu_0$.*

The one major difference between Theorem 4.5.4 and 4.5.5 is that for the strictly concave cost function the solution of the Monge-Kantorovich problem does not generalise as a solution for the Monge problem. For us to have a solution to the Monge problem we require, from Villani [71] that the mass which is common between μ and ν has to stay in the same place. However, as we consider the Monge-Kantorovich problem , this does not concern us; for our work we always generalise a solution to Monge's problem.

Plakhov [61] reduces the various problems of minimal and maximal resistance bodies to a special one-dimensional Monge-Kantorovich problem. The problem is of a body which is bounded by a curve which comprises convex parts and cavities, which are concave. Any particle hitting the body either reflects once only from the convex part of the curve or reflects into the cavity before leaving the body forever. The angles in and out of any cavity are defined to be φ and φ^+ respectively. Assigned to any given cavity is a measure defined by the in and out angles, it is this measure which defines the resistance in the problems of minimal and maximal resistance.

Let $B \subset \mathbb{R}^2$ be a compact, connected subset of Euclidean space $\mathbb{R}^2$ with piecewise smooth boundary. We consider a billard in $\mathbb{R}^2 \setminus B$ and a particle of the billiard with a trajectory which intersects the convex hull of B, denoted $convB$. Initially the particle moves freely in $\mathbb{R}^2 \setminus B$ with a unit velocity v. We let the point of the first intersection of the particle with $\partial(convB)$ be x and denote the unit outer normal vector to $\partial(convB)$

by n_x and denote the angle between n_x and $-v$ by $\varphi \in [-\pi/2, \pi/2]$. For each particle motion a value will be assigned, $(x, \varphi) \in \partial(conv\, B) \times [-\pi/2, \pi/2]$.

After intersecting $\partial(convB)$, the particle moves inside $convB \setminus B$, elastically reflecting off the boundary ∂B, then intersects $\partial(convB)$ again and moves freely afterwards. The particle can make an infinite number of reflections. We let $x^+ = x_B^+(x, \varphi)$ be the point of the second intersection. Denote the angle between the normal n_{x^+} and the velocity of final free motion v^+ by $\varphi^+ = \varphi_B^+(x, \varphi) \in [-\pi/2, \pi/2]$.

The characterisation of the resistance is how the Monge-Kantorovich problem is defined. The mean resistance for a generally non-convex body B is as follows:

$$\overline{R}(B) = |\partial(convB)| \cdot F(\nu_B), \tag{4.25}$$

with

$$F(\nu) = \int\int \left(1 + \cos(\varphi - \varphi^+)\right) \, \mathrm{d}\nu(\varphi, \varphi^+). \tag{4.26}$$

The mean resistance can be thought of as the impulses per unit of time on the bodies. Therefore, the Monge-Kantorovich problem is

$$\inf_{\nu \in \mathcal{M}} F(\nu), \tag{4.27}$$

where $\mathcal{M} = \Pi(\varphi, \varphi^+)$ and the cost function being used is $c(\varphi, \varphi^+) = 1 + \cos(\varphi - \varphi^+)$.

The existence results are not straight forward. For $d = 2, 3$ we have existence of solutions for the Monge-Kantorovich problem. However, for $d = 2$ the optimiser isn't the form $(id \times s)_{\#} \mu$ and therefore we cannot generalise to a Monge minimiser. To the contrary, for $d = 3$ the optimiser is in the form $(id \times s)_{\#} \mu$.

We conclude this chapter by showing that cost functions exist which do not yield solutions to the Monge-Kantorovich problem. We consider the work of Beiglböck et al. [8]. Consider the measure spaces $([\,0\,,\,1\,)\,,\,\lambda_1)$ and $([\,1\,,\,2\,)\,,\,\lambda_1)$. Define $\hat{s}$: $[\,0\,,\,1\,) \to [\,1\,,\,2\,)$ by $\hat{s}(x) = x + 1 = \nabla\left((x^2/2) + x\right)$, the Brenier map between $([\,0\,,\,1\,)\,,\,\lambda_1)$ and $([\,1\,,\,2\,)\,,\,\lambda_1)$. Then $\hat{s}$ is the unique solution of the Monge problem with respect to the cost function $c_1(x\,,\,y) = (x - y)^2/2$, $(id \times \hat{s})_{\#}\,\lambda_1$. If we modify the cost function to penalise shifts by 1, that is define

$$c_2(x\,,\,y) = \begin{cases} 1, & \text{if } y = x + 1 \\ (x - y)^2/2, & \text{otherwise,} \end{cases} \tag{4.28}$$

it can be shown that

$$\inf_{\pi \in \Pi(\lambda_1\,,\,\lambda_1)} \int_{([\,0\,,\,1\,)\,,\,\lambda_1) \times ([\,1\,,\,2\,)\,,\,\lambda_1)} c_2(x\,,\,y)\ \mathrm{d}\Pi(x\,,\,y) = \frac{1}{2}. \tag{4.29}$$

By way of explanation define s_ϵ to be the mapping that shifts $[\,0\,,\,1 - \epsilon\,)$ to $[\,1 + \epsilon\,,\,2\,)$ and $[\,1 - \epsilon\,,\,1\,)$ to $[\,1\,,\,1 + \epsilon\,)$ for a given $\epsilon > 0$. Then $(id \times s_\epsilon)_{\#}\,\lambda_1 \in \Pi(\lambda_1\,,\,\lambda_1)$ and

$$\int_{([\,0\,,\,1\,)\,,\,\lambda_1)\,\times\,([\,1\,,\,2\,)\,,\,\lambda_1)} c_2\left(x\,,\,y\right)\;\mathrm{d}\left(\mathrm{id}\times s_\epsilon\right)_{\#}\lambda_1\left(x\,,\,y\right)\to\frac{1}{2}\quad\text{as}\quad\epsilon\to 0.\qquad(4.30)$$

However, the infimum in (4.29) is not attained by any $\pi\in\Pi\left(\lambda_1\,,\,\lambda_1\right)$: the Monge-Kantorovich problem has no solution.

Chapter 5

Displacement error and a more sophisticated formulation.

5.1 Introduction.

Up to this point in the thesis we have considered our error formulations for scalar valued rearrangements. However, it could be natural to consider errors in vector valued quantities, such as wind velocity. In contrast to the scalar valued case we show in section 5.2 that it need not be the case that a minimiser of the best fit problem exists; this is the content of Proposition 7.

Having considered the methods of optimal mass transfer in the previous chapter we move on to link the displacement error to optimal mass transfer. We use the methods of Brenier and Benamou [10] to show that the definition of the displacement error, i.e the square of 2-Wasserstein distance to be introduced shortly, is equivalent to the Action Integral of transferring the mass between two distributions.

117

A problem with the simple formulation is that the qualitative features error is taken first with the displacement error minimised over best fits of the qualitative features error. Error in qualitative features could be penalised as a large displacement error. Brenier and Benamou [10] outline how minimising both errors simultaneously could be formulated as an optimal mass transfer problem. However, the cost function is in the setting of the time continuous formulation; they also insist that $\int q_1 = \int q_2$, which is not a realistic assumption for our application.

In section 5.4 we formulate a rearrangements problem for minimising the displacement and qualitative features error at the same time. We do not require that q_1 and q_2 have the same 1-norm. In the case of the simple formulation we considered an equivalent maximisation problem to the original minimisation problem. Theorem 8 formulates an equivalent maximisation problem for the more sophisticated problem. One would expect "fewer" extremisers for the more sophisticated problem; we demonstrate in Proposition 9 that if the forecast is a best fit of the original problem, it is the unique extremiser of the more sophisticated version.

It can be difficult to prove that extremisers exist for the more sophisticated formulation directly, therefore we consider a different approach. Section 5.4.1 considers the relaxed more sophisticated formulation which is the more sophisticated formulation functional extremised over the closed convex hull of the set of rearrangements; this deals with the problem of the lack of convexity and compactness of the set of rearrangements. We state the minimisation problem before formulating an equivalent maximisation problem; this is the content of Theorem 10.

Before progressing to the main results of the chapter we consider specific examples which illustrate the main differences between the simple formulation and the more so-

phisticated formulation in section 5.5. Following this are the main results of the chapter.

Theorem 11 shows that the relaxed more sophisticated formulation is attained, but that

it cannot be generalised to the more sophisticated formulation in the way that we'd

hoped for all functions. The theorem finishes by considering the maximisation prob-

lem (5.16). Section 5.6 concludes by considering some properties of the maximisation

problem. The chapter concludes by consider specific values of θ where the minimiser is

bigger for the more sophisticated formulation than the relaxed version.

5.2 Error Decomposition for vector valued functions.

Having proved that the simple formulation is well posed for scalar valued rearrange-

ments, it is a natural progression to ask whether forecast error decomposition can be

generalised to vector-valued mappings. An example of a vector-valued extension would

be to consider an error in wind velocity.

Suppose $q_1 , q_2 \in L^2 (\Omega , \mu , \mathbb{R}^n)$ where (Ω , μ) is a measure interval. For the simple

formulation of error decomposition to be well posed we need to show that

$\inf_{q \in \mathcal{R}(q_2)} \|q_1 - q\|_2$ is attained by some $\hat{q} \in \mathcal{R} (q_2)$. Let Lebesgue measurable $Y \subset \mathbb{R}^n$

satisfy $\lambda_n (Y) = \mu (\Omega)$ and write $q_1^{\#} , q_2^{\#} : Y \to \mathbb{R}^n$ for the monotone rearrangements (or

Brenier maps) of q_1 , q_2 respectively. We can view the Brenier map as the appropriate

generalisation to vector valued mappings of the increasing rearrangement. However,

inequalities satisfied by the increasing rearrangement do not necessarily hold for the

Brenier map: $\int_Y q_1^{\#} \cdot q_2^{\#} \geq \int_\Omega q_1 \cdot q_2$ is not true in general(See Brenier [14]).

We give an example to show that the infimum in $\inf_{q \in \mathcal{R}(q_2)} \|q_1 - q\|_2$ is not neces-

sarily attained. As per Burton and Douglas [21, Section 3] let $D_1 \subset \mathbb{R}^2$ denote the open

disc of the radius 1 centered at the origin and let $u : D_1 \to \mathbb{R}^2$ be defined by

$$u\left(x\right) = \begin{cases} x/|x|, & \text{if } x \neq 0, \\ \\ 0, & \text{if } x = 0. \end{cases} \tag{5.1}$$

Now $u\left(x\right) = \nabla\left(|x|\right)$ for almost every $x \in D_1 : u = u^\#$. Burton and Douglas [21, Lemma 3] gives the existence of an almost injective rearrangement $\hat{u}$ of $u^\#$, $\hat{u} : D_1 \to \mathbb{R}^2$. Choose $q_1 = \hat{u}$, q_2 equal to the identity mapping on D_1. Now $\mathcal{R}\left(q_2\right)$ is the same set as S, the set of measure-preserving mappings from (D_1, λ_2) to (D_1, λ_2).

$$\inf_{q \in \mathcal{R}(q_2)} \|q_1 - q\|_2 = \inf_{s \in S} \|\hat{u} - s\|_2. \tag{5.2}$$

The right hand side of (5.2) is the projection problem of polar factorisation; Burton and Douglas [22, Corollary 1] yields that s attains the infimum in (5.2) if and only if $\hat{u} = u^\# \circ s$, that is $\hat{u}$ has a polar factorisation through D_1. However, the methods of the proof of Theorem 2 of Burton and Douglas [22] yield that $\hat{u}$ has no polar factorisation through D_1. It follows that there is no closest rearrangement in this case.

In the above example the level sets of u are the radii of D_1 (and the origin). It follows that u is a monotone rearrangement which is not almost injective: no matter what set of zero size is ignored, the remaining mapping will not be injective. Furthermore its level sets all have zero area. We generalise this counterexample in Proposition 7. Before stating this result we need some preliminary definitions.

Definitions. *A mapping $f : U \to V$, where (U, μ) is a finite positive measure space, is almost countable to 1 if there exists a set U_0 of full measure such that the intersection of any level set of f with U_0 is countable. For $m \in \mathbb{N}$ we say that f is almost m to 1 if there exists a set U_0 of full measure such that the intersection of any level set of f with U_0 (whenever it is non-empty) has exactly m elements. In this notation almost injective functions are called almost 1 to 1.*

Proposition 7. *Let square integrable $u^{\#} : Y \to \mathbb{R}^n$ be the restriction of the gradient of a proper lower semi continuous convex function to a set $Y \subset \mathbb{R}^n$ of finite positive Lebesgue measure and suppose that $u^{\#}$ restricted to the complement of its level sets of positive measure is not almost injective. Let (Ω, μ) be a measure interval satisfying $\mu(\Omega) = \lambda_n(Y)$. Suppose that $q_1 : \Omega \to \mathbb{R}^n$ is a rearrangement of $u^{\#}$ which is almost countable to 1 on the complement of its level sets of positive measure and that $q_2 : \Omega \to Y$ is a measure-preserving mapping. Then*

$$\inf_{q \in \mathcal{R}(q_2)} \|q_1 - q\|_2, \tag{5.3}$$

is not attained by any $q \in \mathcal{R}(q_2)$.

Proof: We begin by noting that almost m to 1 rearrangements of $u^{\#}$ exist by the construction of Douglas [31, Proposition 1]. Now, $\mathcal{R}(q_2)$ and S are equal, where S is the set of measure-preserving mappings from (Ω, μ) to (Y, λ_n). It follows that

$$\inf_{q \in \mathcal{R}(q_2)} \|q_1 - q\|_2 = \inf_{s \in S} \|q_1 - s\|_2. \tag{5.4}$$

Now, the right hand side of (5.4) is the projection problem of polar factorisation; Burton and Douglas [22, Corollary 1] yields that $s \in S$ attains the infimum in (5.4) if and only if $q_1 = u^\# \circ s$, that is q_1 has a polar factorisation through Y. However, Douglas [31, Theorem 1] proves that q_1 does not have a polar factorisation through Y. We deduce that the infimum in the left hand side of (5.4) is not attained $\qquad\square$

Recall (see Ryff [65], Burton and Douglas [21, Section 3] for a summary) that there is always a polar factorisation in the scalar-valued case; our earlier results are not contradicted.

5.3 Formulation of the displacement error.

The concept of Wasserstein distance is central to our definition of the displacement error. The optimal value corresponding to the quadratic cost function is defined to be the quadratic Wasserstein distance. We define the Wasserstein distance formally and explain its relation to the Monge problem; then we consider the specific Wasserstein distance that we require.

Definition 5.3.1. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain where (Ω , λ_n) is a measure interval. μ is absolutely continuous with respect to n-dimensional Lebesgue measure. The Wasserstein distance between two density functions $q(x) \geq 0$ and $q_2(x) \geq 0$ satisfying $\int_\Omega q = \int_\Omega q_2$ is*

$$W_p(q , q_2)^p = \inf_{s \in Y} \int |s(x) - x|^p q(x) \, \mathrm{d}x, \tag{5.5}$$

where $p \geq 1$ is fixed, $|\cdot|$ denotes the Euclidean norm and Y is the set of measure-preserving

mappings from $(\Omega, q\lambda_n)$ to $(\Omega, q_2\lambda_n)$.

The measures in definition 5.3.1 are $q\lambda_n$ and $q_2\lambda_n$, where $q\mu(A) = \int_A q(x) \, \mathrm{d}\mu(x)$ and

$q_2\mu(A) = \int_A q_2(x) \, \mathrm{d}\mu(x)$. In this sense q and q_2 are weights (or densities); alternatively

they are known as Radon-Nikodym derivatives. We introduce the notation of $\nabla\varphi_q$ which

is the Brenier map dependent on q between $(\Omega, q\mu)$ and $(\Omega, q_2\mu)$. For the quadratic

cost we consider $p = 2$ in definition 5.5.

Using the work of Benamou and Brenier [10; 11] we explain the link to the displace-

ment error; this was made rigorous for our assumption of non-negative square integrable

q_2 by Guittet [42]. Suppose $\Omega \subset \mathbb{R}^n$ is a set of finite Lebesgue measure, i.e. $\lambda_n(\Omega) < \infty$.

We assume the following:

$$\int_\Omega |x|^4 \, \mathrm{d}\lambda_n < \infty, \tag{5.6}$$

which would certainly be the case if Ω is bounded. Assumption (5.6) allows us to assume

that, by use of the Cauchy-Schwarz inequality, that the following is true

$$\int_\Omega \frac{|x|^2}{2} q_2(x) \, \mathrm{d}\lambda_n(x) \, , \, \int_\Omega \frac{|x|^2}{2} q(x) \, \mathrm{d}\lambda_n(x) < \infty. \tag{5.7}$$

We say that $q_2\lambda_n$ and $q\lambda_n$ have finite second moments.

Condition (5.7) ensures that $W_2^2(q_2)$ is finite. For the work in this chapter it is sufficient

that $\Omega \subset \mathbb{R}^n$, $\mu(\Omega) < \infty$, where μ is absolutely continuous with respect to Lebesgue

measure λ_n. Furthermore, for our problem we assume that q_1, $q_2 \geq 0$, as do Brenier and Benamou [11], Guittet [42] and Ambrosio et al. [3]. Another point to note from Brenier and Benamou's work is that they assume that q_1 and q_2 are bounded, merely square integrable; and thus q_1, $q_2 \in L^\infty(\Omega)$. However, in our work, we do not assume that q_1 and q_2 are bounded, this is where we make use of work by Guittet [42] and Ambrosio et al. [3]. Under these assumptions Definition 5.5 is still true and we will see later that Proposition 5.3.2 remains true also.

From the optimal mass transfer problem the initial mass must equal the whole where the mass is being transferred to. Therefore, if the initial domain is taken to be Ω and the final domain is taken to be Ω', it must be the case that $\mu(\Omega) = \nu(\Omega')$. To ensure that this holds for our functions, $q \in \mathcal{R}(q_2)$, it must be the case that $\int q = \int q_2$.

Benamou and Brenier [10; 11] provide a reformulation of the mass transfer problem. They use a continuous time variable $t \in [0, 1]$, where the time interval is arbitrarily chosen and fixed. We state the following transport equation for smooth, time dependent, density and velocity fields, $\rho = \rho(t, x)$, $v = v(t, x) \in \mathbb{R}^d$,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho v) = 0 \text{ with } \nabla \cdot v = 0, \tag{5.8}$$

for $0 < t < 1$ and $x \in \mathbb{R}^d$, subject to the initial and final conditions,

$$\rho(0, \cdot) = q, \rho(1, \cdot) = q_2. \tag{5.9}$$

As we are considering $W_2^2(\hat{q}, q_2)$ we are transporting best fits $\hat{q}$ to q_2(equivalently we

could consider $W_2^2 (q_2 , \hat{q})$. (ρ , v) defines a path if it is a solution of (5.8) and (5.9). What we are doing here is transporting q to q_2 with velocity fields. These are transported over "possible paths" and the problem is to minimise the kinetic energy over these "possible paths". The following result from Brenier and Benamou [11] shows us that this is equivalent to the 2-Wasserstein distance.

Proposition 5.3.2. *The square of the L^2 Wasserstein distance is equal to the infimum of*

$$\int_{\mathbb{R}^d} \int_0^1 \rho\,(t\,,\,x)\,|v\,(t\,,\,x)|^2 \ \mathrm{d}x\,\mathrm{d}t, \tag{5.10}$$

among all (ρ , v) satisfying (5.8) and (5.9), where $\rho \in L^2 (\Omega)$.

(5.10) is known as the action integral and is denoted $A(\rho , v)$. There are several possible definitions for the action integral; we consider two of these definitions.

$$
\begin{aligned}
A_1(\rho, v) &= \int_{\mathbb{R}^d} \int_0^1 |v(t,x)|^2 \ \mathrm{d}x \ \mathrm{d}t, &\tag{5.11}\\
A_2(\rho, v) &= \int_{\mathbb{R}^d} \int_0^1 \rho(t,x)|v(t,x)|^2 \ \mathrm{d}x \ \mathrm{d}t. &\tag{5.12}
\end{aligned}
$$

The choice of the action integral is dependent on what we are looking to penalise. $A_1(\rho, v)$ will penalise the magnitude of the velocity which transports best fits to the forecasted distribution. Thinking of $\rho(\text{t},\cdot)$ as a density it is evident that $A_2(\rho, v)$ will penalise the kinetic energy during transportation from the forecasted to the best fitting distribution; by way of explanation, the integrand (5.12) integrated with respect to x

gives "kinetic energy" at time t. Our choice from Proposition 5.3.2, is (5.12). While (5.11) is another possible choice, it is not known to be directly related to an optimal mass transfer problem. This is an open question and could be part of our further work.

Brenier and Benamou stated that the time-continuous formulation was "implicitly contained" in the original problem of Monge [59]. In this problem eliminating the time variable was a way to reduce the dimension of the problem. Benamou and Brenier [10] present several reasons as to why the time-continuous formulation should be kept. One reason is that it leads to the numerical scheme suggested in Benamou and Brenier [10]. Of major significance to us is the section on "Interpolation of the L^2 and the Wasserstein distance". They note in some situations that it is natural to consider both the Wasserstein distance and the L^2 distances at the same time, especially in data assimilation for meteorological forecasting. Therefore, for practical applications it might be desirable to have a weighted combination of the L^2 and Wasserstein distances between q and q_2 for $\theta \in [0\,,\,1]$. The cost function becomes

$$\int_{\mathbb{R}^d} \int_0^1 \left[(1 - \theta)\, \rho\,(t\,,\,x) \,|v\,(t\,,\,x)|^2 \,+\, \theta \left(\frac{\partial \rho}{\partial t}\,(t\,,\,x) \right)^2 \right] \; \mathrm{d}x\, \mathrm{d}t. \tag{5.13}$$

This is the content of the next section as we seek a more sophisticated forecast error decomposition where the errors are taken simultaneously; solving the problem of a forecast with an error in qualitative features being penalised as if it was a displacement error. One of the major problems is that Brenier and Benamou's problem had the assumption that $\int q = \int q_2$ which is fine for $q \in \mathcal{R}\,(q_2)$. However, there is no reason to assume that $\int q_1 = \int q_2$ and in general, for our applications to weather forecasting, this need not

be the case. This is not appropriate for the more sophisticated formulation and some adaptations need to be made. This is the content of the next section.

5.4 More sophisticated formulation of the weather forecast error.

From the introductory chapter we have that while the simple formulation addressed problems from the conventional error scores there are new problems to consider. Given that the displacement error is minimised over the minimisers of the qualitative features error, the qualitative features error is minimised without consideration for the displacement error. One suggestion to address this problem, based on (5.13), is as follows:

$$\inf_{q \in \mathcal{R}(q_2)} (1 - \theta) \, \|q_1 - q\|_2 \, + \, \theta \, W_2^2 \, (q \, , \, q_2) \, . \tag{5.14}$$

From the definition of the Wasserstein distance, specifically for $p = 2$, we have to take the square root of the minimiser hence we square the whole function to counteract this. In the same manner and given the square root which comes from the definition of the L^2-norm, it might be more appropriate to consider the following

$$\inf_{q \in \mathcal{R}(q_2)} (1 - \theta) \, \|q_1 - q\|_2^2 \, + \, \theta \, W_2^2 \, (q \, , \, q_2) \, . \tag{5.15}$$

(5.15) is the natural generalisation of the work of Brenier and Benamou [10].

Remark: *Whilst the simple formulation of $E(q_1, q_2)$ is not well defined in general for vector-valued mappings, it is an open question whether the more sophisticated version makes sense. One would need to define $W_2^2(q, q_2)$ appropriately for $q \in \mathcal{R}(q_2)$.*

There are three methods which could be used to tackle this problem. We could study the rearrangements problem. Alternatively we could use the similarities between push-forward measures, rearrangements and measure-preserving mappings from section 2.6.1 to turn the problem into an optimal mass transfer problem, for a suitable cost function. Finally, we could consider the problem as a partial differential equations problem.

The partial differential equations problem has been considered by Haker, Zhu, Tannenbaum and Angenent [44, Section 3.4]; we note the similarities to (5.15). However, for the partial differential equations problem to work Haker and his co-workers assume smoothness and invertibility of the measure-preserving mapping to ensure that the derivatives exist. Unlike Haker [44], in our problem there is no need to assume that the measure-preserving mapping is invertible. Therefore, there are obvious limitations to how much we can use this numerical scheme. Haker et al's numerical scheme for calculating Brenier maps for the Wasserstein distance converges to an optimal map. But, in the L^2 and Wasserstein distance problem, it is only conjecture that the numerical scheme converges to an optimal map. In fact there may be multiple solutions. Furthermore, in the same way as Brenier and Benamou's problem, Haker's problem also assumes that $\int q_1 = \int q_2$ which is an unrealistic assumption for our application. These problems are the reasons we do not consider the more sophisticated formulation as a partial differential equations problem.

Having established the assumptions of the problem we reformulate (5.15) as a maximisation problem.

Theorem 8. *Let $\Omega \subset \mathbb{R}^n$ be a set of finite n-dimensional Lebesgue measure, and suppose that $\int_\Omega |x|^4 \ \mathrm{d}\lambda_n(x) < \infty$. Let non-negative $q_1, q_2 \in L^2(\Omega, \lambda_n)$. Then (5.15) is equivalent to*

$$\sup_{q \in \mathcal{R}(q_2)} \int_\Omega \left(2(1-\theta)q_1(x) + \theta x \cdot \nabla\phi_q - \frac{\theta}{2}|x|^2 \right) q(x) \ \mathrm{d}\lambda_n(x), \qquad (5.16)$$

where $\nabla\phi_q$ denotes the Brenier map between $(\Omega, q\lambda_n)$ and $(\Omega, q_2\lambda_n)$. The problems are equivalent in the sense that minimisers attaining the infimum in (5.15) are maximisers of the supremum in (5.16) and vice versa.

Proof: We begin by noting that $\int_\Omega |x|^2 q_2(x) \ \mathrm{d}\lambda_n(x)$, $\int_\Omega |x|^2 q(x) \ \mathrm{d}\lambda_n(x) < \infty$ (by the Cauchy-Schwarz inequality).

Let $\nabla\phi_q$ denote the Brenier map between $(\Omega, q\lambda_n)$ and $(\Omega, q_2\lambda_n)$. Then the definition of 2-Wasserstein distance (see definition 5.3.1) yields:

$$\theta W_2^2(q, q_2) = \frac{\theta}{2} \int_\Omega |x - \nabla\phi_q|^2 q(x) \ \mathrm{d}\lambda_n(x). \qquad (5.17)$$

For $q \in \mathcal{R}(q_2)$,

$$(1 - \theta) \|q_1 - q\|_2^2 + \theta W_2^2(q, q_2)$$

$$= (1 - \theta) \int_\Omega \left(q_1^2(x) - 2q_1(x) q(x) + q^2(x) \right) \, d\lambda_n(x)$$

$$+ \frac{\theta}{2} \int_\Omega \left(|x|^2 - 2x \cdot \nabla \phi_q(x) + |\nabla \phi_q(x)|^2 \right) q(x) \, d\lambda_n(x), \qquad (5.18)$$

$$= (1 - \theta) \int_\Omega q_1^2(x) - 2q_1(x) q(x) + q_2^2(x) \, d\lambda_n(x),$$

$$+ \frac{\theta}{2} \int_\Omega \left(|x|^2 - 2x \cdot \nabla \phi_q(x) \right) q(x) \, d\lambda_n(x),$$

$$+ \frac{\theta}{2} \int_\Omega |x|^2 q_2(x) \, d\lambda_n(x). \qquad (5.19)$$

$$= \int_\Omega \left((1 - \theta) \left(q_1^2(x) + q_2^2(x) \right) + \theta |x|^2 q_2(x) \right) \, d\lambda_n(x)$$

$$+ \int_\Omega \left(-2(1 - \theta) q_1(x) + \frac{\theta}{2} |x|^2 - \theta x \cdot \nabla \phi_q(x) \right) q(x) \, d\lambda_n(x). \qquad (5.20)$$

By way of explanation the first integral in (5.19) follows by noting that rearrangement preserves L^2-norms and that (5.18) holds. We achieve the last integral in (5.19) by noting that $\nabla \phi_q : (\Omega, q\lambda_n) \to (\Omega, q_2\lambda_n)$ is measure-preserving. Now, the first integral in (5.20) is constant; the result follows. $\qquad \square$

The modified functional we are extremising penalises the movement of qualitative features; we would expect "smaller" sets of extremisers compared with the original formulation. The following proposition demonstrates that if the forecast is a best fit of the original formulation, then it is the unique extremiser of the more sophisticated version.

Proposition 9. *If $q_2 \in \mathcal{M}$, then q_2 is the unique minimiser/maximiser of (5.15)/(5.16) for $0 < \theta \leq 1$.*

Proof: Suppose $q \in \mathcal{R}(q_2)$ is a minimiser of (5.15). Then

$$\theta W_2^2(q, q_2) \le (1 - \theta) \|q_2 - q_1\|_2^2 - (1 - \theta) \|q - q_1\|_2^2 + \theta W_2^2(q_2, q_2) \le 0,$$

where we have used the fact that $q_2 \in \mathcal{M}$ to obtain the second inequality. We deduce that $W_2^2(q, q_2) = 0$ and that $q\lambda_n$ and $q_2\lambda_n$ are equal (as measures). Noting that q, q_2 are square integrable, almost every point of Ω is a Lebesgue point of q and q_2. We deduce $q = q_2$. The result follows. $\square$

Superficially Theorem 8 appears similar to the "best fits" problem of Nealers Theorem. However, the dependence of the Brenier map $\nabla \phi_q$ on $q \in \mathcal{R}(q_2)$ makes (5.16) more challenging to study. We conclude this section by noting that if we define $\nu = q_{2\#}\mu$, S, the set of measure-preserving mappings from (Ω, μ) to $(\mathbb{R}, \nu)$, is $\mathcal{R}(q_2)$. Writing (5.15) as a Monge problem, that is determining the cost function, is however a more challenging problem to formulate.

5.4.1 The relaxed more sophisticated problem.

We consider the more sophisticated problem over the weak closure of the set of rearrangements, known as the relaxed more sophisticated problem. Through considering the problem over the weak closure we hope to generalise this result over the set of rearrangements. As stated in section 2.7 the weak closure of the set of rearrangements is equivalent to the closed convex hull of the set of rearrangements and thus solves the problem of the set of rearrangements not being weakly compact.

From Theorem 8 we have that the more sophisticated formulation, equation (5.15), is equivalent to problem (5.16). To this end we consider the maximisation problem over the closed convex hull of the set of rearrangements. We begin by considering the minimisiation problem over the weak closure of the set of rearrangements which is as follows:

$$\inf_{q \in \overline{\text{conv}(\mathcal{R}(q_2))}} (1 - \theta) \|q_1 - q\|_2^2 + \theta W_2^2 (q, q_2) \tag{5.21}$$

From the reformulation of the more sophisticated formulation it was evident that part of the result was dependent on the fact that rearrangement preserves the L^2-norm. This no longer holds for the closed convex hull of the set of rearrangement, thus, there are additional terms to be considered in the relaxed more sophisticated formulation. The reformulation of (5.21) into a maximisation problem is the content of the next theorem.

Theorem 10. *Let $\Omega \subset \mathbb{R}^n$ be a set of finite n-dimensional Lebesgue measure and suppose that $\int_\Omega |x|^4 \, d\lambda_n(x) < \infty$. Let non-negative $q_1 \, q_2 \in L^2(\Omega, \lambda_n)$. Then (5.21) is equivalent to:*

$$\sup_{q \in \overline{\text{conv}(\mathcal{R}(q_2))}} \int_\Omega (2(1 - \theta)q_1(x) + \theta x \cdot \nabla \phi_q(x)$$
$$- \frac{\theta}{2} |x|^2 - (1 - \theta) q(x)) q(x) \, d\lambda_n(x) \tag{5.22}$$

Proof: Following the methods of the proof of Theorem 8 we omit some detail and refer the reader to the proof of Theorem 8 for this detail. For $q \in \overline{\mathrm{conv}\left(\mathcal{R}\left(q_2\right)\right)}$,

$$
\begin{aligned}
& \left(1-\theta\right)\left\|q_1-q\right\|_2^2 + \theta W_2^2\left(q,\,q_2\right) \\
= \quad & \left(1-\theta\right)\int_\Omega \left(q_1^2\left(x\right) - 2q_1\left(x\right)q\left(x\right) + q^2\left(x\right)\right)\,\mathrm{d}\lambda_n\left(x\right) \\
+ \quad & \frac{\theta}{2}\int_\Omega \left(|x|^2 - 2x\cdot\nabla\phi_q\left(x\right) + |\nabla\phi_q\left(x\right)|^2\right)q\left(x\right)\,\mathrm{d}\lambda_n\left(x\right) \qquad (5.23) \\
= \quad & \left(1-\theta\right)\int_\Omega q_1^2\left(x\right)\,\mathrm{d}\lambda_n\left(x\right) + \frac{\theta}{2}\int_\Omega |x|^2 q_2\left(x\right)\,\mathrm{d}\lambda_n\left(x\right) \\
+ \quad & \left(1-\theta\right)\int_\Omega \left(q^2\left(x\right) - 2q_1\left(x\right)q\left(x\right)\right)\,\mathrm{d}\lambda_n\left(x\right) \\
+ \quad & \frac{\theta}{2}\int_\Omega \left(|x|^2 - 2x\cdot\nabla\phi_q\left(x\right)\right)q\left(x\right)\,\mathrm{d}\lambda_n\left(x\right) \qquad (5.24)
\end{aligned}
$$

We note that as we are extremising over the closed convex hull of the set of rearrangements, that rearrangement no longer preserves the L^2-norm. The first two integrals of (5.24) are constant, and the result follows. $\qquad\qquad\square$

5.5   The More Sophisticated Problem - An example.

Before we prove results about the more sophisticated formulation we consider the following example which illustrates the differences between the simple and the more sophisticated formulation. We work on the measure space $\left([0,\,1]^2,\,\lambda_2\right)$ where $R_1 = [0,\,1/2] \times [0,\,1]$ and $R_2 = [1/2,\,1] \times [0,\,1]$. We let q_1 be the actual data and q_2, q_3 and q_4 be all the possible forecasts. We evaluate the error, $E\left(q_1,\,q_i\right)$ for $i = 2,\,3$ and 4. We define $q_1 = 2\times 1_{R_2}$, $q_2 = 1_{R_2}$, $q_3 = 1_{R_1}$ and $q_4 = 2\times 1_{R_1}$ as illustrated in the following figure:
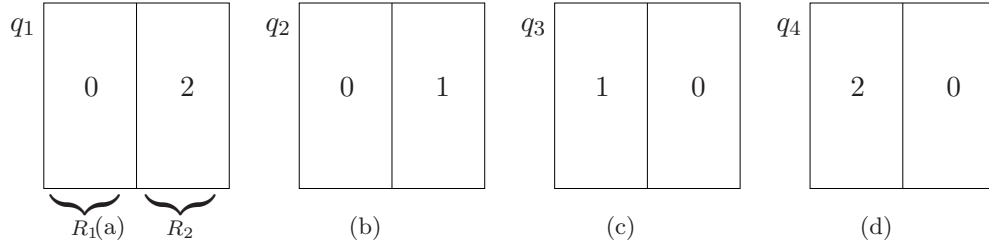
Figure 5.1: An illustration of the (a) outcome q_1 and three cases of forecast q_2, q_3 and q_4 corresponding to (b) Case 1 (c) Case 2 and (d) Case 3

We consider cases 1 and 2. The set of rearrangements is as follows

$\mathcal{R}(q_2) = \{q = 1_A : A \subset \mathbb{R}^2 \text{ such that } \lambda_2(A) = 1/2\}$. Let $q = 1_A$ for some $A \subset \mathbb{R}^2$. The set of measure-preserving mappings from $\left([0,1]^2, q\lambda_2\right) \to \left([0,1]^2, q_2\lambda_2\right)$ is (almost everywhere) the set of measure-preserving mappings (A, λ_2) to (R_2, λ_2). The Brenier map from $\left([0,1]^2, q_3\lambda_2\right)$ to $\left([0,1]^2, q_2\lambda_2\right)$ is

$$(x, y) + (1/2, 0) = \nabla\left(\frac{|\mathbf{x}|^2}{2} + \frac{1}{2}x\right) \quad \text{where} \quad \mathbf{x} = (x, y). \tag{5.25}$$

For case one we treat the error as a purely qualitative features error. Therefore, we have the following:

$$\inf_{q \in \mathcal{R}(q_2)} \|q_1 - q\|_2 = \|q_1 - q_2\|_2 = \|2 \times 1_{R_2} - 1_{R_2}\|_2 = 1/\sqrt{2}. \tag{5.26}$$

(q_2 is the unique minimiser - see Nealers Theorem.) The purely qualitative features is illustrated with $W_2^2(q_2, q_2) = 0$.

In case 2 we have that both the qualitative features and displacement errors contribute to the overall error. The error is as follows:

$$\inf_{q \in \mathcal{R}(q_3)} \|q_1 - q\|_2 \quad = \quad \|2 \times 1_{R_2} - 1_{R_2}\|_2 = 1/\sqrt{2}. \tag{5.27}$$

$$\inf_{q \in \mathcal{M}} W_2^2 (q, q_3) \quad = \quad W_2^2 (1_{R_2}, 1_{R_1})$$

$$= \quad \int_{R_2} |\mathbf{x} + (1/2, 0) - \mathbf{x}|^2 \; \mathrm{d}\lambda_2 (\mathbf{x}) = 1/8. \tag{5.28}$$

Again q_2 is the unique minimiser. For the simple formulation we finish with the third case where the error could be formulated as a displacement error only, as shown below:

$$\inf_{q \in \mathcal{R}(q_4)} \|q_1 - q\|_2 \quad = \quad \|q_1 - q_1\|_2 = 0. \tag{5.29}$$

$$\inf_{q \in \mathcal{M}} W_2^2 (q, q_4) \quad = \quad W_2^2 (2 \times 1_{R_2}, 2 \times 1_{R_1})$$

$$= \quad 2 . \int_{R_2} |\mathbf{x} + (1/2, 0) - \mathbf{x}|^2 \; \mathrm{d}\lambda_2 (\mathbf{x}) = 1/4. \tag{5.30}$$

But, when we consider the more sophisticated formulation the errors compete against each other. However, for case 1 the formulations are similar. They are not equal given that the more sophisticated formulation considers the square of the L^2 norm. But, it is the same rearrangement which yields the unique minimiser.

We consider the error associated with the three cases for the more sophisticated formulation. We begin with case 1:

$$E(q_1, q_2) \quad = \quad (1 - \theta) \|q_1 - q_2\|_2^2 + \theta W_2^2 (q_2, q_2)$$

$$= \quad (1 - \theta)/2. \tag{5.31}$$

q_2 is the unique minimiser (over $q \in \mathcal{R}(q_2)$). Before moving on to case 2 we consider the following figure:
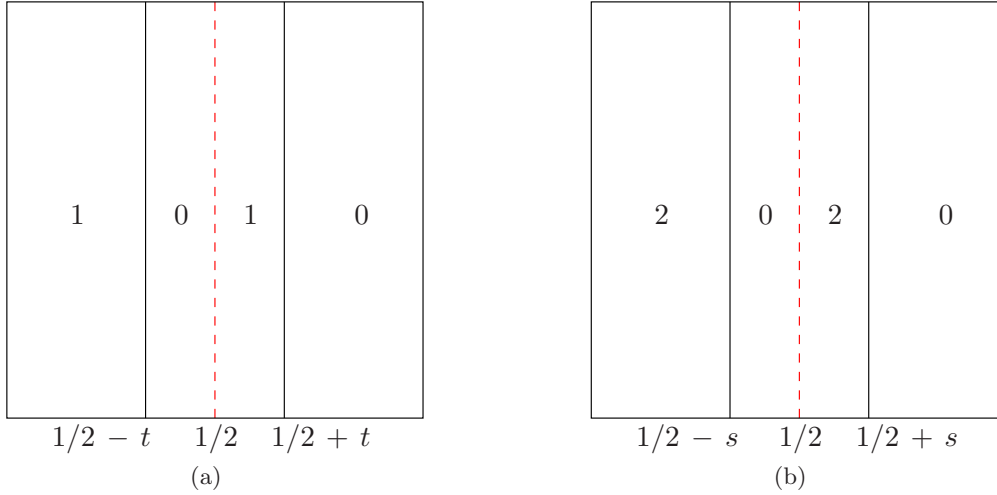

Figure 5.2: An illustration of possible error formulations for (a) case 2, which we denote by q_t and (b) case 3, which we denote by p_t.

q_t is a rearrangement of q_3. We consider the following:

$$
\begin{aligned}
W_2^2(q_t, q_3) = W_2^2(q_3, q_t) &= \inf_{s \in S} \int |s(x) - x|^2 \, q_3 \\
&\leq \int |\hat{s}(x) - x|^2 \, q_3 \text{ for a particular } \hat{s} \\
&= \int_{R_1} |\hat{s}(x) - x|^2 \\
&= \int_{R_3} t^2 = t^3.
\end{aligned}
$$

The first equality happens given that q_3 and q_t are rearrangements of each other. Given that it is the infimum over all s, for a particular $\hat{s}$ we yield the inequality. Now, for fixed θ (θ close to 1 when the displacement error is dominant), for some $0 \leq t < 1/2$,

$$(1 - \theta) \left\| q_1 - q_t \right\|_2^2 + \theta W_2^2 (q_t , q_3)$$

$$\leq (1 - \theta) \left(\frac{5}{2} - 4t \right) + \theta t^3$$

$$< (1 - \theta) / 2 + \theta/8 \text{ for appropriate choices of } \theta , t.$$

$$= (1 - \theta) \left\| q_1 - q_2 \right\|_2^2 + \theta W_2^2 (q_2 , q_3) \tag{5.32}$$

So, q_2 is not a minimiser. (We are not claiming that q_t is necessarily a minimiser.) Case 3 is similar to case 2. From the simple formulation we have that the error is entirely displacement error with q_1 being the unique minimiser. For the more sophisticated formulation we let the function from figure 5.2(b) be denoted by q_s. q_s will score better than q_1 for some θ and s. For the more sophisticated formulation, $E (q_1 , q_4)$ will have non-zero contributions from the qualitative features and displacement terms. So, the error is no longer considered to be purely displacement. It is not now beneficial to move the whole block, in both cases, and is more beneficial to move parts of the level set.

5.6   Well-posedness of the more sophisticated formulation.

To illustrate that the relaxed more sophisticated formulation is well posed we follow a classic approach; this is to pick an extremising sequence and show that the limit it converges to is a maximiser of the functional. For this we require some form of continuity of the functional. Therefore, we must firstly show some continuity results for both the qualitative features and the displacement error parts of the more sophisticated problem.

We state the following theorem from [71], which we will make use of later in this section:

Theorem 5.6.1 (Partial.). *Let $p \in (0, \infty)$, let $(\mu_k)_{k \in \mathbb{K}}$ be a sequence of probability measures in $P_p(X)$, the set of probability measures on X with finite p^{th} moments. Then, the following two statements are equivalent:*

(i) $W_p(\mu_k, \mu) \to 0$ *as $k \to \infty$.*

(ii) *Whenever a continuous function φ on X satisfies the growth condition*

$$|\varphi(x)| \leq C\left[1 + \mathrm{d}(x_0, x)^p\right] \text{ for some } x_0 \in X, C \in \mathbb{R}, \text{ then}$$

$$\int \varphi \, \mathrm{d}\mu_k \to \int \varphi \, \mathrm{d}\mu \quad as \quad k \to \infty. \tag{5.33}$$

It would be sufficient to prove the well posedness of the problem that $q \to \|q_1 - q\|_2^2$ is weakly sequentially lower semi continuous(see definition 2.4.1) and that $q \to W_2^2(q, q_2)$ is weakly sequentially continuous. The following theorem illustrates continuity of the Wasserstein distance, using this fact to show that relaxed more sophisticated formulation has a minimiser. From this further results are proved about the relationship between the relaxed more sophisticated formulation and the more sophisticated formulation.

Theorem 11. *Let $q_1, q_2 : \Omega \to \mathbb{R}$ be as above, where $\Omega \subset \mathbb{R}^n$ is such that $\int_\Omega |x|^4 \, \mathrm{d}x < \infty$. Then:*

(i) *If $q_n \rightharpoonup \hat{q}$, where $\rightharpoonup$ denotes weak convergence, for a sequence $(q_n) \subset \overline{\mathrm{conv}(\mathcal{R}(q_2))}$, then $W_2^2(q_n, q_2) \to W_2^2(\hat{q}, q_2)$.*

(ii) *(5.21) is attained by some $q \in \overline{\mathrm{conv}(\mathcal{R}(q_2))}$.*

(iii) *Suppose that $q_2 \in \mathcal{M}$ and that $q_1 \in \overline{\mathrm{conv}(\mathcal{R}(q_2))} \setminus \mathcal{R}(q_2)$. Choose $0 < \theta < 1$ such that:*

$$\theta W_2^2 \left(q_1, q_2\right) < (1 - \theta) \left\|q_1 - q_2\right\|_2^2. \tag{5.34}$$

Then we have the value of $(5.21) <$ the value of (5.15).

(iv) (5.16) is attained by some $\tilde{q} \in \overline{\mathrm{conv}\left(\mathcal{R}\left(q_2\right)\right)}$. Moreover, the value of (5.16) is the same for $q \in \mathcal{R}\left(q_2\right)$ and $q \in \overline{\mathrm{conv}\left(\mathcal{R}\left(q_2\right)\right)}$.

Remarks:

(i) *From the definition of the closed convex hull, we note that $\int_\Omega q \, \mathrm{d}\mu = \int_\Omega q_2 \, \mathrm{d}\mu$.*
It follows that $(\Omega, q\lambda_n)$ and $(\Omega, q_2\lambda_n)$ are measure spaces of the same size, so
$W_2^2 \left(q, q_2\right)$ is well-defined.

(ii) *Part (iii) represents a case where qualitative features error is emphasised over the*
displacement error. One would choose θ sufficiently small.

Proof: (i) Suppose that $(q_n) \subset \overline{\mathrm{conv}\left(\mathcal{R}\left(q_2\right)\right)}$ is such that $q_n \rightharpoonup \hat{q}$ as $n \to \infty$ for some $\hat{q} \in \overline{\mathrm{conv}\left(\mathcal{R}\left(q_2\right)\right)}$. Let $\varphi : \Omega \to \mathbb{R}$ be continuous and satisfy $|\varphi\left(x\right)| \le C \left(1 + |x_0 - x|^2\right)$ for some $x_0 \in \Omega$, $C \in \mathbb{R}$. Now, $\int_\Omega |x|^4 \, \mathrm{d}x < \infty$, thus we have that $\varphi \in L^2 \left(\Omega, \lambda_n\right)$. The definition of weak convergence yields:

$$\int_\Omega \varphi q_n \to \int_\Omega \varphi \hat{q} \quad \text{as } n \to \infty \tag{5.35}$$

This proves statement (iv) of Theorem 5.6.1 and therefore given the equivalence of the four statements we can use statement (i) to state the following:

$$W_2\left(q_n\lambda_n\,,\,\hat{q}\lambda_n\right)\to 0 \quad\text{as}\quad n\to\infty. \tag{5.36}$$

From the triangle inequality we have that

$$\left|W_2\left(q_n\,,\,q_2\right)-W_2\left(\hat{q}\,,\,q_2\right)\right|\le W_2\left(q_n\,,\,\hat{q}\right), \tag{5.37}$$

as the quadratic Wasserstein distance is a metric on Borel measures with $\mu\left(\Omega\right)=\int_\Omega q_2\,\mathrm{d}\lambda_n$. From this we can deduce that

$$W_2^2\left(q_n\,,\,q_2\right)\to W_2^2\left(\hat{q}\,,\,q_2\right)\quad\text{as}\quad n\to\infty. \tag{5.38}$$

This illustrates the weak continuity of the Wasserstein distance; this result will be used extensively in the work that follows.

(ii) We deal with the case of $q_2=0$ separately. In this case we have that $\overline{\mathrm{conv}\left(\mathcal{R}\left(q_2\right)\right)}=\{0\}$ and the result follows immediately.

The definition of weak convergence yields that

$$\int q_n\,\frac{\hat{q}}{\|\hat{q}\|_2}\to\int\hat{q}\,\frac{\hat{q}}{\|\hat{q}\|_2}\quad\text{as}\quad n\to\infty \tag{5.39}$$

We have from (5.39) that

$$\left(\int q_n \frac{\hat{q}}{\|\hat{q}\|_2} \right)^2 \to \left(\int \hat{q} \frac{\hat{q}}{\|\hat{q}\|_2} \right)^2 = \frac{1}{\|\hat{q}\|_2^2} \left(\int \hat{q}^2 \right)^2 = \frac{\|\hat{q}\|_2^4}{\|\hat{q}\|_2^2} = \|\hat{q}\|_2^2 \qquad (5.40)$$

Making use of the Cauchy-Schwarz inequality we manipulate the left hand side of equation (5.39)

$$\int q_n \frac{\hat{q}}{\|\hat{q}\|_2} \leq \|q_n\|_2 \left\| \frac{\hat{q}}{\|\hat{q}\|_2} \right\|_2 = \|q_n\|_2 \qquad (5.41)$$

noting $\left\| \frac{\hat{q}}{\|\hat{q}\|_2} \right\|_2 = 1$. From (5.40) and (5.41) we can deduce that

$$\|\hat{q}\|_2^2 \leq \liminf_{n \to \infty} \|q_n\|_2^2 \qquad (5.42)$$

Given that $q_n \rightharpoonup \hat{q}$ as $n \to \infty$ we note that

$$\int q_n q_1 \to \int \hat{q} q_1 \quad \text{as} \ \ n \to \infty \qquad (5.43)$$

We deduce that

$$\|\hat{q} - q_1\|_2^2 \leq \liminf_{n \to \infty} \|q_n - q_1\|_2^2 \tag{5.44}$$

From definition 2.4.1 we have that the qualitative features part of the more sophisticated formulation is weakly sequentially lower semi continuous. Utilising this and the continuity of the Wasserstein distance we have the following:

$$(1 - \theta) \quad \|q_1 - \hat{q}\|_2^2 + \theta W_2^2 (\hat{q}, q_2)$$

$$\leq \quad \liminf_{n \to \infty} (1 - \theta) \|q_1 - q_n\|_2^2 + \lim_{n \to \infty} \theta W_2^2 (q_n, q_2) \tag{5.45}$$

$$\leq \quad \liminf_{n \to \infty} \left((1 - \theta) \|q_1 - q_n\|_2^2 + \theta W_2^2 (q_n, q_2) \right) \tag{5.46}$$

$$\leq \quad \inf_{q \in \overline{\text{conv}(\mathcal{R}(q_2))}} (1 - \theta) \|q_1 - q\|_2^2 + \theta W_2^2 (q, q_2). \tag{5.47}$$

The first part of the right hand side of (5.45) is due to (5.44). Inequality (5.47) follows given that (q_n) is a minimising sequence.

(iii) We have that:

$$\inf_{q \in \overline{\text{conv}(\mathcal{R}(q_2))}} \quad (1 - \theta) \|q_1 - q\|_2^2 + \theta W_2^2 (q, q_2) \tag{5.48}$$

$$\leq \qquad \theta W_2^2 (q_1, q_2) \tag{5.49}$$

$$< \qquad (1 - \theta) \|q_1 - q_2\|_2^2 \tag{5.50}$$

$$= \qquad \inf_{q \in \mathcal{R}(q_2)} (1 - \theta) \|q_1 - q\|_2^2 \tag{5.51}$$

$$\leq \qquad \inf_{q \in \mathcal{R}(q_2)} (1 - \theta) \|q_1 - q\|_2^2 + \theta W_2^2 (q, q_2). \tag{5.52}$$

We explain the steps above. If we choose $q = q_1$ in the functional of (5.48) the qualitative features error is zero and hence we obtain inequality (5.49). Inequality (5.34) yields inequality (5.50). Noting that $q_2 \in \mathcal{M}$ yields (5.51). Finally, we are adding the displacement error to the qualitative features error, hence we yield the final inequality (5.52). This illustrates that the relaxed problem has a lower error score.

(iv) We choose a maximising sequence, $(q_n) \subset \mathcal{R}(q_2)$ (or $\overline{\text{conv}(\mathcal{R}(q_2))}$). We expand the following:

$$
\begin{aligned}
W_2^2 (q_n, q_2) &= \min_{s_\# q \lambda_d = q_2 \lambda_d} \int_\Omega \frac{|x - s(x)|^2}{2} q(x) \, \mathrm{d}\lambda_d(x) \\
&= \min_{s_\# q \lambda_d = q_2 \lambda_d} \int_\Omega \left(\frac{|x|^2}{2} - x \cdot s(x) \right) q(x) \, \mathrm{d}\lambda_d(x) + \int_\Omega \frac{|y|^2}{2} q_2(y) \, \mathrm{d}\lambda_d(y) \\
&= \int_\Omega \frac{|x|^2}{2} q(x) \, \mathrm{d}\lambda_d(x) + \int_\Omega \frac{|y|^2}{2} q_2(y) \, \mathrm{d}\lambda_d(y) \\
&\quad - \max_{s_\# q \lambda_d = q \lambda_d} \int_\Omega x \cdot s(x) q(x) \, \mathrm{d}\lambda_d(x).
\end{aligned}
$$

Noting (5.38) and $\frac{|x|^2}{2}$, $q(x) \in L^2(\Omega)$, from our assumption that $\int_\Omega |x|^4 \, d\lambda_d(x) < \infty$, then we have the following:

$$q \to \int_\Omega \frac{|x|^2}{2} q(x) \, d\lambda_d(x)$$

is weakly continuous. Thus the following mapping is weakly continuous

$$q \to \max_{s_\# q\lambda_d = q_2\lambda_d} \int_\Omega x \cdot s(x) q(x) \, d\lambda_d(x) = \int_\Omega x \cdot \nabla\phi_q(x) q(x) \, d\lambda_d(x),$$

where $\nabla\phi_q$ denotes the Brenier map between $(\Omega, q\lambda_d)$ and $(\Omega, q_2\lambda_d)$. (By way of explanation we have used weak continuity of $q \to W_2(q, q_2)$.) It follows that for $q \in \overline{\text{conv}(\mathcal{R}(q_2))}$, the mapping:

$$q \to \int_\Omega \left(2(1-\theta) q_1(x) + 2\theta x \cdot \nabla\phi_q - \theta |x|^2 \right) q(x) \, d\lambda_d(x)$$

is also weakly continuous. Let (q_n) be a maximising sequence for (5.16). Passing to a subsequence if necessary, $q_n \rightharpoonup \tilde{q}$ say as $n \to \infty$, for some $\tilde{q} \in \overline{\text{conv}(\mathcal{R}(q_2))}$. Now

$$\sup_{q \in \mathcal{R}(q_2)} \int_\Omega \left(2 \left(1 - \theta \right) q_1 \left(x \right) + 2\theta x \cdot \nabla \phi_q - \theta \left| x \right|^2 \right) q \left(x \right) \; \mathrm{d}\lambda_d \left(x \right)$$

$$= \int_\Omega \left(2 \left(1 - \theta \right) q_1 \left(x \right) + 2\theta x \cdot \nabla \phi_{\tilde{q}} - \theta \left| x \right|^2 \right) \tilde{q} \left(x \right) \; \mathrm{d}\lambda_d \left(x \right).$$

Moreover, $\mathcal{R}\left(q_2 \right)$ is weakly sequentially dense in $\overline{\mathrm{conv} \left(\mathcal{R} \left(q_2 \right) \right)}$ so the supremum is the same if we consider the larger constraint set $\overline{\mathrm{conv} \left(\mathcal{R} \left(q_2 \right) \right)}$.

$\square$

There are a few points to note. Firstly, having proved that a minimiser exists for the relaxed more sophisticated formulation it would have been our hope to show existence of minimisers for the rearrangements problem. However, the next part of the theorem illustrates that for a class of functions, the value of the infimum of the more sophisticated formulation is greater than that of the value for a minimiser of the relaxed more sophisticated formulation. Minimisers of the relaxed version cannot be rearrangements. Thus we must consider a different approach and that is the content of the final section. As discussed previously the difference between the formulations is that $\|q\|_2^2 \le \|q_2\|_2^2$ for $q \in \overline{\mathrm{conv} \left(\mathcal{R} \left(q_2 \right) \right)}$, whereas for $\mathcal{R}\left(q_2 \right)$ there is equality. Noting that:

$$\sup_{q \in \overline{\mathrm{conv}(\mathcal{R}(q_2))}} \int_\Omega \left(2 \left(1 - \theta \right) q_1 \left(x \right) + \theta x \cdot \nabla \phi_q - \frac{\theta}{2} \left| x \right|^2 \right) q \left(x \right) \; \mathrm{d}\lambda_n \left(x \right)$$

$$= \sup_{q \in \mathcal{R}(q_2)} \int_\Omega \left(2 \left(1 - \theta \right) q_1 \left(x \right) + \theta x \cdot \nabla \phi_q - \frac{\theta}{2} \left| x \right|^2 \right) q \left(x \right) \; \mathrm{d}\lambda_n \left(x \right) \qquad (5.53)$$

(5.16) seems the more tractable version of the problem: maximisers of the relaxed problem may yield maximisers of (5.16). Suppose $\tilde{q} \in \overline{\mathrm{conv} \left(\mathcal{R} \left(q_2 \right) \right)}$ attains the supremum

on the left hand side of (5.53). Our wish is to find $\hat{q} \in \mathcal{R}(q_2)$ which attains the supremum on the right-hand side of (5.53). Noting that $\left(2(1-\theta)q_1(x) - \theta|x|^2\right) \in L^2(\Omega)$, and making use of the results in Chapter 2 and properties of $\overline{\text{conv}(\mathcal{R}(q_2))}$, we can choose $\hat{q} \in \mathcal{R}(q_2)$ such that

$$\int_\Omega \left(2(1-\theta)q_1(x) - \theta|x|^2\right)\hat{q}(x) \, \mathrm{d}\lambda_n(x)$$
$$\geq \int_\Omega \left(2(1-\theta)q_1(x) - \theta|x|^2\right)\tilde{q}(x) \, \mathrm{d}\lambda_n(x) \tag{5.54}$$

Using this we get the following:

$$\int_\Omega \left(2(1-\theta)q_1(x) + 2\theta x{\cdot}\nabla\phi_{\hat{q}} - \theta|x|^2\right)\hat{q}(x) \, \mathrm{d}\lambda_n(x)$$
$$\leq \int_\Omega \left(2(1-\theta)q_1(x) + 2\theta x{\cdot}\nabla\phi_{\tilde{q}} - \theta|x|^2\right)\tilde{q}(x) \, \mathrm{d}\lambda_n(x) \tag{5.55}$$
$$\leq \int_\Omega \left(2(1-\theta)q_1(x) - \theta|x|^2\right)\hat{q}(x) \, \mathrm{d}\lambda_n(x) + \int_\Omega (2\theta x{\cdot}\nabla\phi_{\tilde{q}})\tilde{q}(x) \, \mathrm{d}\lambda_n(x) \tag{5.56}$$

By way of explanation, the first inequality in (5.55) follows because $\hat{q} \in \overline{\text{conv}(\mathcal{R}(q_2))}$ and $\tilde{q}$ attains the supremum on the left hand side of (5.53), and the second from (5.54). From Theorem 11 we know that the mapping $q \to \int_\Omega (2\theta x \cdot \nabla\phi_q)q(x)$ is weakly continuous; one needs to establish further properties to make progress.

5.6.1 Example.

Given the hypotheses of Theorem 11 (iii) we have specific values of θ for which the value of $(5.21) <$ the value of (5.15). Our original hope was to use the classical approach of

considering the problem over the closed convex hull, before extracting minimisers of the rearrangements problem; this is not possible for this class of examples. We demonstrate this class is non-empty, and exhibit a value of θ. Define the following:

$$q_1(x) = \frac{1}{2} \text{ for } x \in [0,\,1] \tag{5.57}$$

$$q_2(x) = x \text{ for } x \in [0,\,1] \tag{5.58}$$
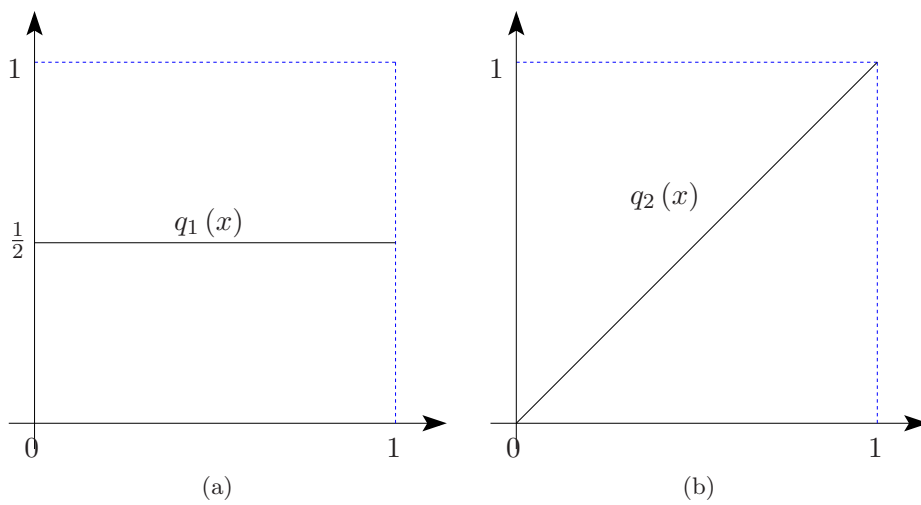
which are illustrated below:

Figure 5.3: An illustration of the (a) $q_1(x)$; (b) $q_2(x)$.

For both functions we consider the qualitative features error and the displacement error for both examples. Example 1 considers functions $q_1(x) = \frac{1}{2}$ and $q_2(x) = x$. The qualitative features error is as follows:

$$\|q_1 - q_2\|_2^2 = \int_0^1 \left(\frac{1}{2} - x\right)^2 \, \mathrm{d}x = \frac{1}{12}$$

In order to evaluate the displacement error we must ascertain what the Brenier map is. In this example we're considering the Brenier map between $\left([0,\,1]\,,\,\frac{\lambda_1}{2}\right)$ and $([0,\,1]\,,\,x\,\mathrm{d}\lambda_1\,(x))$.

We define the "length of $[0, r]$" functions:

$$\varphi(r) = \int_0^r \frac{1}{2} \, d\lambda_1(x) = \frac{r}{2}; \tag{5.59}$$

$$\psi(r) = \int_0^r x \, d\lambda_1(x) = \frac{r^2}{2}. \tag{5.60}$$

Noting that a measure-preserving mapping $s : \left([0, 1], \frac{\lambda_1}{2}\right) \to ([0, 1], x \, d\lambda_1(x))$ satisfies $\frac{1}{2}\lambda_1(s^{-1}(A)) = x \, d\lambda_1(A)$, we want $\varphi(r) = \psi(g(r))$, so $s(r) = \sqrt{r}$. Therefore, the displacement error is:

$$
\begin{aligned}
W_2^2(q_1\lambda_1, q_2\lambda_2) &= \frac{1}{4}\int_0^1 \left|\sqrt{x} - x\right|^2 = \frac{1}{4}\int_0^1 \left(x - 2x^{3/2} + x^2\right) \, d\lambda_1(x) \\
&= \frac{1}{4}\left[\frac{x^2}{2} - \frac{4x^{5/2}}{5} + \frac{x^3}{3}\right]_0^1 \\
&= \frac{1}{120}
\end{aligned}
$$

We use this example to illustrate specific occurrences when inequality (5.34) is satisfied. We seek θ such that $\theta W_2^2(q_1, q_2) < (1 - \theta)\|q_1 - q_2\|_2^2$ that is:

$$
\begin{aligned}
\frac{\theta}{120} &< \frac{1 - \theta}{12} \\
\theta &< \frac{10}{11}.
\end{aligned}
$$

Chapter 6

Conclusions.

6.1   Summary.

This thesis has considered the decomposition of weather forecast error. In the opening

chapter we discussed the problems with using conventional error scores. Noting that

such forecast errors are not descriptive of how a forecast failed and may unduly penalise

displacement errors, we consider a new formulation. Based on work by Hoffman, Liu,

Louis and Grassotti we suggest a simple formulation which we split into two parts; an

error due to qualitative features and an error due to displacement, noted in (1.3). The

simple formulation is more descriptive of how the forecast fails.

The qualitative features error is defined to be the differences between the actual and

best fitting data measured by L^2-norm over the set of rearrangements of the forecasted

data. The displacement error is defined to be the Wasserstein distance between the fore-

casted and best fitting data; it is minimised over the set of extremisers of the qualitative

features error. Given that best fitting rearrangements are calculated first, problems can

arise where an error in qualitative features can be penalised as a large displacement

error. Solutions to these problems are outlined in section 1.3 and considered in Chapter 5. One solution would be to split the forecast into a series of regions, therefore not allowing features to move large distances. Another solution which was considered later in the thesis was to minimise the errors simultaneously.

The early part of the thesis is concerned with the theory behind the simple formulation. From the definition of the error of our simple formulation some questions were left unanswered. In the closest rearrangements chapter we show that the problem is well posed. From Douglas [30] we have that a maximiser (or minimiser) exists and we give a precise condition for when the maximiser is unique; this is the content of Theorem 3.2.1. The precise condition is not easy to verify; we give a sufficient condition. The main theorem of this thesis, Nealers Theorem, is the characterisation of the set of closest rearrangements, $\mathcal{M}$. Nealers Theorem characterises the set $\mathcal{M}$ in the following way:

> Maximisers (or minimisers) are unique up to level sets having positive measure of q_1. On such level sets having positive measure the maximisers are rearrangements of each other.

Therefore, we note the importance of level sets having positive measure and prove the uniqueness on the complement of the level sets having positive measure in Proposition 2. The proof of Nealers Theorem is split into various lemmas and propositions. Nealers Theorem can be written as those $q \in \mathcal{R}(q_2)$ which satisfies (3.5). We use the layer cake principle, as used by Villani [71], to disintegrate in the way shown in (3.19) and (3.20), which allows us to define G_q.

Lemma 3 and Proposition 4 establishes properties about G_q, more specifically about the continuity of G_q. We use these properties in Lemma 5 to simplify the problem in

Nealers Theorem. Using the simplified problem we now break down Nealers Theorem into an if and only if condition. We let (⋆) be the condition that maximisers are unique up to level sets having positive measure and are rearrangements on level sets having positive measure. Proposition 6 deals with one side of this argument, i.e. given an element of (⋆) do we have a maximiser? Nealers Theorem uses this and proves the other side of the if and only if condition.

We then consider the theory of optimal mass transfer defining the Monge-Kantorovich problem before moving onto its dual formulation. Having defined these we consider the results for specific cost functions. This leads us directly into our final chapter. As we discussed previously to solve one of the problems of the simple formulation we turn the simple formulation into an optimal mass transfer problem to ensure that both the qualitative features and displacement errors are taken simultaneously. This was based on methods considered by Brenier and Benamou [10] who suggested turning the whole error into an optimal mass transfer problem with a suitable choice for the cost function. Using this work we turn the problem into a rearrangements problem. However, we can turn the problem into an optimal mass transfer problem or a projection problem as considered by Burton & Douglas [22]. The new error score is known as the more sophisticated formulation. This problem can also be reformulated as a partial differential equations problem or an optimal mass transfer problem. We use Haker's work [73] to show the limitations to that problem.

Before considering the more sophisticated formulation we make the natural progression from the scalar version of the simple formulation to consider the vector version of the simple formulation. While its scalar counterpart was well posed, the vector version need not be. We show via an example before generalising the example in Proposition 7

which illustrates that the closest rearrangement problem need not be attained for vector valued functions.

We progress to consider the more sophisticated formulation where the qualitative features error and the displacement are taken at the same time; a problem highlighted in Chapter 1. For the simple formulation it was beneficial to consider an equivalent maximisation problem to the original minimisation problem. We do this for the more sophisticated formulation and this is the content of Theorem 8.

We would expect the new formulation to have less extremisers and we begin to illustrate this in Proposition 9, where it is illustrated that if the forecast is a best fit of the simple formulation it is the unique extremiser of the more sophisticated formulation.

In Chapter 3 we prove that the simple formulation is well posed and we characterise the extremisers. The aim of Chapter 5 was to prove similar results for the more sophisticated formulation. From section 2.7.2 it is a classical approach to consider the problem over the closed convex hull of the set of rearrangements before generalising back to the set of rearrangements. This deals with the lack of compactness and convexity which exist for the set of rearrangements. Thus in section 5.4.1 we consider the relaxed more sophisticated problem, which is the more sophisticated problem over the closed convex hull. Again we consider an equivalent maximisation for this formulation, this is the content of Theorem 10. The difference between the maximisation problems of the relaxed more sophisticated problem and the more sophisticated formulation is that in the closed convex hull rearrangement no longer preserves the L^2-norm. While we prove that an extremiser attains the relaxed more sophisticated formulation in Theorem 11 (ii) the next part of the theorem illustrates that we cannot generalise in the way in which we had hoped. We finish the theorem and the chapter by showing that the maximisation

problems of Theorem 8 and 10 are the most likely to yield results regarding the more sophisticated formulation.

6.2 Future Work.

6.2.1 Numerical Implementation.

An obvious practical question is how to implement the theoretical ideas of this thesis numerically. For the simple formulation the qualitative features error is given by:

$$
\inf_{q \in \mathcal{R}(q_2)} \|q_1 - q\|_2 \;=\; \|q_1^* - q_2^*\|_2 \;=\; \left\|q_1^\Delta - q_2^\Delta\right\|_2, \tag{6.1}
$$

$$
=\; \left\{ \|q_1\|_2^2 - 2 \int q_1^* q_2^* + \|q_2\|_2^2 \right\}^{1/2}, \tag{6.2}
$$

$$
=\; \|q_1 - q_2^* \circ s\|_2, \tag{6.3}
$$

where s is a polar factor of q_1, that is a measure-preserving mapping satisfying $q_1 = q_1^* \circ s$. Computing increasing (or decreasing) rearrangements should be straightforward, then one would calculate $\|q_1^* - q_2^*\|_2$ or $\int q_1^* q_2^*$ according to preference. Alternatively, one could calculate a polar factorisation of q_1 and find q_2^*. Benamou [9] described a scheme for calculating monotone rearrangements and polar factorisations.

The work of Benamou and Brenier [10; 11] gives a numerical scheme for calculating $W_2^2(q, q_2)$ and carries out some numerical experiments. The challenge here would be to find an efficient way to optimise over the "closest rearrangements" as characterised in Nealers Theorem. It appears that the numerical calculation of the simple formulation could be carried out with existing methods.

We consider stochastic effects. Our primary motivating example has been the distribution of rainfall; the rainfall data has been obtained from a satellite and treated to deal with any stochastic effects. However, one might want to carry out error decomposition for a quantity with random noise. In this case the residual error of Hoffman et al. [47] makes more sense.

As we have alluded to previously Haker et al. [44] discusses numerical schemes for the more sophisticated formulation. However, Haker and his co-workers made the assumption that $\int q_1 = \int q_2$; this need not be the case in our work. Therefore, we would need to modify this numerical scheme to use it in our case. Moreover, to utilise this numerical scheme, we would need to use the Monge formulation of the more sophisticated error; this is an open question of this thesis.

6.2.2   Open questions.

The questions left unanswered from Chapter 5 surround illustrating that the more sophisticated formulation is attained and how to characterise the extremisers. After Theorem 11 we are left with (5.53). From this we have established that a maximiser attains the bound of the more sophisticated formulation over the closed convex hull of the set of rearrangements. From (5.53) we know that the value of the supremum over the closed convex hull is equal to the value of the supremum over the set of rearrangements. The question left unanswered is whether a maximiser exists for maximisation problem of the more sophisticated formulation over the set of rearrangements.

We discussed the study of energy extremising isovortical planar ideal fluids flows in Chapter 1 (see also Burton [23]). There may be similarities to the maximisation problem of the more sophisticated formulation. In Burton's problem, he was able to

demonstrate the existence of an energy maximiser relative to the weak closure of the set of rearrangements by standard methods and then used rearrangement inequalities to find a rearrangement with at least the same energy. This demonstrates the existence of energy maximisers relative to the set of rearrangements. It is our hope to be able to utilise these methods for our work. We can split (5.16) on the following way:

$$\sup_{q \in \mathcal{R}(q_2)} \left[\iint_{\Omega} \left(2\left(1-\theta\right) q_1\left(x\right) - \frac{\theta}{2}\left|x\right|^2 \right) q\left(x\right) \, \mathrm{d}\lambda_n\left(x\right) + \int_{\Omega} \left(\theta x \cdot \nabla \phi_q\right) q\left(x\right) \, \mathrm{d}\lambda_n\left(x\right) \right] (6.4)$$

The first integral is linear and weakly continuous. For this integral we have already proven that a rearrangement does just as well as the maximiser with respect to the weak closure, (5.54). However, the Brenier map which is dependent on q adds extra complexity to our problem, and while we have that the second integral is weakly continuous we require further properties to make use of the arguments of Burton [18]. Something like convexity would be sufficient. Even if we could show a property such as this, it might still be difficult to work with both integrals simultaneously; the way in which we deal with the coupling of the integrals could be central to solving the problem.

Numerical implementation, as we've described above, is a natural next step. Given this, Haker's work is important to us and to use Haker's work, we require a formulation of the more sophisticated problem as a Monge problem over a set of measure-preserving mappings. The cost function involves a Brenier map dependent on the measure-preserving mapping and seems challenging to manipulate.

Finally we wish to remove the assumption in Nealers Theorem that the values taken by the level set of positive measure of q_1 can be written as a decreasing sequence.

156

Whilst unlikely to be of practical importance to forecast error decomposition, it would

be a better theoretical result.

Bibliography

[1] A. Alvino, G. Trombetti, and P.L. Lions. On optimization problems with prescribed rearrangements. *Nonlinear Anal-Theor*, 13:185–220, 1989.

[2] L. Ambrosio. Lecture Notes on Optimal Transport Problems. 1990.

[3] L. Ambrosio, N. Gigli, and G. Savaré. Gradient flows with metric and differentiable structures and applications to the Wasserstein space. *Mat. Appl.*, 15:327–343, 2004.

[4] L. Ambrosio and A. Pratelli. Existence and stability results in the L^1 theory of optimal transportation. *Lecture Notes in Mathematics "Optimal transportation and applications" (CIME Series, Martina Franca, 2001)* **1813**, *L.A. Caffarelli and S. Salsa Eds.*, pages 123–160, 2003.

[5] P. Appell. Mémoire sur les déblais et les remblais des systémes continus ou discontinus. *Mémoires présentés apr divers Savantas à l'Académie des Sciences de l'Institut de France*, 29:1–208, 1887.

[6] P. Appell. Le probleme geometrique des déblais et remblais. *Memorial de Sc. Math. Fasc.*, 29:1–208, 1928.

[7] S. Begueria, S.M. Vicente-Serrano, and M. Angulo-Martinez. A multiscalar global

drought dataset: The speibase *a new gridded product for the analysis of drought variability and impacts. American Meteorological Society*, 10:1351–1356, 2010.

[8] M. Beiglböck, M. Goldstern, G. Maresch, and W. Schachermayer. Optimal and better transport plans. *J. Funct. Anal.*, 256:1907–1927, 2009.

[9] J.D. Benamou. A Domain Decomposition Method for the Polar Factorization of Vector Valued Mappings. *SIAM J. Numer. Anal.*, 32:1808–1838, 1995.

[10] J.D. Benamou and Y. Brenier. A numerical method for the Optimal Time-Continuous Mass Transport Problem and related problems. Monge Ampere Equation : Applications to Geometry and Optimization, Contemporary Mathematics. *American Mathematical Society*, 226:1–11, 1998b.

[11] J.D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84:375–393, 2000.

[12] T. B. Benjamin. The alliance of practical and analytical insights into the nonlinear problems of fluid mechanics. Applications of methods of functional analysis to problems in mechanics. Lecture notes in mathematics. *Springer*, 503:8–29, 1976.

[13] Y. Brenier. Décomposition polaire et réarrangment monotone des champs de vecteurs. *C. R. Acad. Sci. Paris*, 305:805–808, 1987.

[14] Y. Brenier. Polar Factorization and Monotone Rearrangement of Vector-Valued Functions. *Comm. Pure. Appl. Math.*, 44:375–417, 1991.

[15] A. Burchard. Cases of Equality in the Riesz Rearrangement Inequality. *Ann. Math.*, 143:499–527, 1996.

[16] A. Burchard and H. Hajeiej. Rearrangement inequalities for functionals with monotone integrands. *J. Funct. Anal.*, 233:225–253, 1996.

[17] G.R. Burton. Rearrangement of functions, maximisation of convex functionals, and vortex rings. *Math. Ann.*, 276:225–253, 1987.

[18] G.R. Burton. Rearrangements and Ideal Fluids. *Ann. I. H. Poincar Anal.*, 6:295–319, 1989.

[19] G.R. Burton. Variational problems on classes of rearrangements and multiple configurations for steady vortices. *Ann. I. H. Poincar Anal.*, 6:295–319, 1989.

[20] G.R. Burton. Global nonlinear stability for steady ideal fluid flow in bounded planar domains. *Arch. Rational Mech. Anal.*, 176:149–163, 2005.

[21] G.R. Burton and R. J. Douglas. Rearrangements and polar factorisation of countably degenerate functions. *Proc. Roy. Edin.*, 128A:671–681, 1998.

[22] G.R. Burton and R. J. Douglas. Uniqueness of the Polar Factorisation and Projection of a Vector-Valued Mapping. *Ann. I. H. Poincar*, 3:405–418, 2003.

[23] G.R. Burton and J.B. McLeod. Maximisation and minimisation on classes of rearrangements. *Proc. Roy. Edin.*, 119A:287–300, 1991.

[24] L. A. Caffarelli, M. Feldman, and R. J. McCann. Constructing optimal maps for Monge's transport problem as a limit of strictly convex costs. *J. Amer. Math. Soc.*, 15:1–26, 2001.

[25] M. G. Crandall and L. Tartar. Some relations between nonexpansive and order preserving mappings. *Proc. Amer. Math. Soc.*, 78:385–390, 1980.

[26] M. J. P. Cullen and R. J. Douglas. Large-Amplitude nonlinear stability results for atmospheric circulations. *Q.J.R. Meteorol. Soc.*, 129:1969–1988, 2003.

[27] M. J. P. Cullen, J. Norbury, and R. J. Purser. Generalised Langrangian solutions for atmospheric and oceanic flows. *SIAM J. Appl. Math.*, 51:20–31, 1991.

[28] G. Darboux. Prix Bordin (géométrie). *C.R. Acad. Sci. Paris*, 101:1312–1316, 1885.

[29] R. J. Douglas. Rearrangements of functions on unbounded domains. *Proc. Roy. Edin.*, 124A:621–644, 1994.

[30] R. J. Douglas. Rearrangements of functions with application to meteorology and ideal fluid flow-the mathematics of atmosphere-ocean dynamics. *Cambridge University Press*, 1:288–341, 2002.

[31] R. J. Douglas. Non-existence of polar factorisations and polar inclusion of a vector-valued mapping. *IJPAM*, 41:363–374, 2007.

[32] R.J. Douglas. Rearrangements of vector valued functions, with application to atmospheric and oceanic flows. *SIAM J. Math. Anal.*, 29:891–902, 1998.

[33] C. Dupin. Applications de la géométrie et de la mécanique. *Bachelier, Paris*, 1822.

[34] I. Ekeland and R. Téman. Convex Analysis and Variational Problems. *SIAM*, 1999.

[35] L.C. Evans and W. Gangbo. Differential Equations Methods for the Monge-Kantorovich Mass Transfer Problem. *Memoires Amer. Math. Soc.*, 653, 1999.

[36] D. S. Fry. Shape Recognition Using Metrics on the Space of Shapes. *PhD thesis, Harvard University*, 1993.

[37] W. Gangbo. An elementary proof of the polar factorization of vector-valued functions. *Arch. Rat. Mech. Anal.*, 128:380–399, 1995.

[38] W. Gangbo. The Monge mass transfer problem and its applications. *Contemp. Maths*, 226:79–105, 1999.

[39] W. Gangbo and R.J. McCann. The geometry of optimal transportation. *Acta. Math.*, 177:113–161, 1996.

[40] W. Gangbo and R.J. McCann. Shape recognition via Wasserstein Distance. *Q. Appl. Math*, 58:705–737, 2000.

[41] S. Geman and D. Geman. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE*, pages 721–741, 1984.

[42] K. Guittet. On the Time-Continuous Mass Transport Problem and its Approximation by Augmented Lagrangian Techniques. *SIAM J. Numer.*, 41:382–399, 2003.

[43] H. Hajaiej. Cases of equality and strict inequality in extended Hardy-Littlewood inequalities. *Proc. Roy. Soc. Edin.*, 135A:643–661, 2005.

[44] S. Haker, L. Zhu, A. Tannenbaum, and S. Angenent. Optimal Mass Transport for Registration and Warping. *International Journal of Computer Vision*, 60:225–240, 2004.

[45] P. R. Halmos. Measure Theory. *D. Van Nostrand Company*, 1950.

[46] G.E. Hardy, J.E. Littlewood, and G. Pólya. Inequalities. *Cambridge University Press*, 1934.

[47] R.N. Hoffman, Z. Liu, J-F. Louis, and C. Grassotti. Distortion Representation of Forecast Errors. *Month. Weather Rev.*, 123:2758–2770, 1994.

[48] F. Lindgren J. Lindström, E. Gilleland. Image Warping for Forecast Verification. *Manuscript in Preparation.*

[49] T. Kaijser. Computing the Kantorovich Distance for Images. *J. Math. Imag. Vis.*, 9:173–191, 1998.

[50] L. V. Kantorovich. On a translocation of masses. *Dokl. Akad. Nauk*, 37:227–229, 1942.

[51] L. V. Kantorovich. On a problem of Monge. *Uspekhi Mat. Nauk*, 3:225–226, 1948.

[52] V. Khatskevich and D. Shoiykhet. Differentiable Operators and Nonlinear Equations(Operator Theory: Advances and Applications). *Birkhauser*, 1993.

[53] M. Knott and C. S. Smith. On the optimal mapping of distributions. *J. Optimiz Theory App.*, 43:39–49, 1984.

[54] C. De Lellis. Rectifiable sets, densities and tangent measures. *European Mathematical Society*, 2008.

[55] V. Levin. Reduced cost functions and their applications. *J. Math. Econ.*, 18:155–186, 1997.

[56] E.H. Lieb and M. Loss. Analysis. *American Mathematical Society*, 2001.

[57] R. J. McCann. Existence and Uniqueness of Monotone measure-preserving maps. *Duke Math. J.*, 80:309–323, 1995.

[58] R. J. McCann. Exact solutions to the transportation problem on the line. *Proc. R. Soc. Lond.*, 455:1341–1380, 1999.

[59] G. Monge. Mémoire sur las théorie des déblais et de remblais. *Histoire de L'Académie Royale des Sciences de Paris, avec les Mémoires de Mathématique et de Physique pour la même année*, pages 666–704, 1781.

[60] A. Plakhov. Newton's problem of the body of minimal mean resistance. *Sb. Math.*, 195:1017–1037, 2004.

[61] A. Plakhov. Billiards and Two-Dimensional Problems of Optimal Resistance. *Arch. Rational Mech. Anal.*, 194:349–381, 2009.

[62] S. Rachev and L. Rüschendorf. Mass Transportation Problems. Vol I: Theory. Probability and its applications. *Springer-Verlag, New York*, 1998.

[63] S. Rachev and L. Rüschendorf. Mass Transportation Problems. vol II: Applications. Probability and its applications. *Springer-Verlag, New York*, 1998.

[64] R. T. Rockafellar. Convex Analysis. *Princeton University Press*, 1970.

[65] J.V. Ryff. Measure preserving transformations and rearrangements. *J. Math. Anal. and Applics.*, 31:449–458, 1970.

[66] H. Scarf. The 1975 Nobel Prize in Economics: Resource allocation. *Science*, 190:649, 710, 712, 1975.

[67] V.N. Sudakov. Geometric problems in the theory of infinite-dimensional probability distributions. *Proc. Steklov Inst. Math.*, 141:1–178, 1979.

164

[68] W. A. Sutherland. Introduction to Metric & Topological Spaces. *Oxford University Press*, 2009.

[69] Lord Kelvin(Sir William Thomson). Mathematical and physical papers. *Cambridge University Press.*, IV, 1910.

[70] N. S. Trudinger and X-J. Wang. On the Monge mass transfer problem. *Calc. Var.*, 13:19–31, 2001.

[71] C. Villani. Topics in Optimal Transportation. *American Mathematical Society*, 2000.

[72] C. Villani. Optimal Transport, old and new. *Springer*, 2008.

[73] L. Zhu, Y. Yang, S. Haker, and A. Tannenbaum. An Image Morphing Technique Based on Optimal Mass Preserving Mapping. *IEEE Transactions on Image Processing*, 16:1481–1494, 2007.