GLOBAL MINIMIZERS OF THE INTERACTION ENERGY AND THE BEHAVIOUR OF NEARBY SOLUTIONS TO THE AGGREGATION EQUATION

by

cameron davies

A thesis submitted in conformity with
the requirements for the degree of

Master of Science

Graduate Department of Mathematics
University of Toronto

ABSTRACT

Global Minimizers of the Interaction Energy and the Behaviour of Nearby Solutions to the Aggregation Equation

Cameron Davies

Master of Science

Graduate Department of Mathematics

University of Toronto

2021

Over the past decade, the aggregation equation has become a popular area of research within mathematics. While much of this interest has been driven by the equation's scientific applications, the equation also raises significant mathematical questions. Steady states of the aggregation equation are known to correspond to critical points — including global and local minima — of an associated interaction energy. In this thesis, we partially classify global energy minimizers, when working with the aggregation equation endowed with a specific class of 'power-law' potentials. We then explain this partial classification, as well as its limitations.

After this, we follow in the footsteps of Simione's recent PhD thesis to query the dynamics of solutions to the aggregation equation and, in particular, the behaviour of solutions which begin 'close to' a global minimizer of the interaction energy. In doing so, we outline a promising future avenue of research.

To my wonderful parents,
Glen and Tammy Davies.

ACKNOWLEDGEMENTS

I would like to begin by thanking my supervisor, Professor Robert McCann. Professor McCann has always been an incredibly kind and understanding supervisor, with dozens upon dozens of intriguing ideas and suggestions, clear and detailed explanations, and I relish both the productive and casual discussions I've had with him throughout this past year. I do not think I could have asked for a better supervisor, and am immensely grateful for all that he has done for my research career. I am particularly thankful for Professor McCann's numerous and thoughtful pieces of feedback on the initial draft of this thesis, which have made what follows much more clear, concise, and (I hope!) readable.

Additionally, I wish to thank my informal supervisor and mentor, Professor Tongseok Lim. Though I was not his grad student, Professor Lim took the time to meet with me almost every week and share his ideas and conjectures with me, which I deeply appreciate. Without Professor Lim's help, I doubt I would have been able to pursue many of the ideas in this thesis in such depth or clarity.

I would also like to thank many of my professors and TAs from my undergraduate career for providing me with the nudges I needed to stick with mathematics, and eventually settle in to this field. I will always remember back in 2016 when, after receiving what I thought was a terrible grade on my first midterm of university, my TA, Dr. Yvon Verberne saw that I had potential and convinced me to stay in my first year analysis course. Professor Joe Repka, who taught and coordinated the same course, also inspired my love of mathematics, and was one of the main factors why I stayed in math past my first year. I am also proud to call Professor Repka a mentor and a friend, who has always been there for me when I've had any questions or doubts about mathematics, and who has provided me with numerous reference letters over the years. Next, I want to thank Professor Almut Burchard, who I first met when she taught me introductory PDEs in my third year of undergraduate studies. Professor Burchard has always been a caring and energetic presence, who propelled me deeper into the study of PDEs and analysis. I am also incredibly grateful to her for recommending Professor McCann as an NSERC-USRA supervisor, and for, on numerous occasions, offering to write me a reference letter before I could even get the question out of my mouth. Finally, I want to thank Professor Ila Varma, for writing me a reference letter for graduate school, and for providing me with seemingly endless feedback and support during the application process.

I also want to thank some of my middle school and high school math teachers for showing me that, in spite of some necessary ennui when learning the basics, mathematics is an interesting and exciting field. In particular, I want to thank Ms. Karen Malone, Mr. John Munro, and Ms. Kelly O'Connell for showing me that math a fun, worthwhile pursuit.

I would also like to thank my friend and colleague, Alex Rodriguez for meeting with me more or less every week this summer to work on, and talk about, math. I am glad that we were able to make a platform for this sort of informal discussion. I always came away from our chats with a clearer vision of not just Alex's research, but also my own.

I am also immensely grateful to NSERC, for funding my research throughout the summer of 2020 with a USRA grant.

Last but far from least, I want to thank my parents, Glen and Tammy, for their support throughout my life - from teaching me basic addition and my times tables as soon as I could read, to putting up with me when I went through the more angsty phases of high school, to letting me move back to their house for seven months during the pandemic.

PUBLICATIONS

Much of the material in Sections 2.1 and 2.2 appears in my joint work with Robert McCann and Tongseok Lim, found in [14] and [15]. Much of the rest of Chapter 2 consists of additional discussion of and elaboration on that work.

CONTENTS

1 SETTING: THE AGGREGATION EQUATION AND BASIC OPTIMAL TRANSPORT

At its core, the purpose of this thesis is to explain some crucial methods of studying the aggregation equation, (1.2). To do so at an appropriate level of detail, we use this chapter to build up some necessary background. We begin by introducing a few basic concepts in optimal transportation, a theory which provides us with invaluable perspectives and tools for studying spaces of probability measures. After this, we introduce the aggregation equation, our key object of study, and introduce the reader to some of that equation's properties and applications. We then discuss the interaction energy associated with the aggregation equation, especially when equipped with what is termed a 'power-law potential.' After this, we explicitly define two classes of probability measures — simplicial measures, and spherical shell measures — which are crucially important to the work in Chapter 2. To cap off this introductory chapter, we derive an Euler-Lagrange equation satisfied by any global minimizer of the interaction energy within the space of probability measures.

The second purpose of this chapter is to serve as an informal literature review, wherein we provide the reader with an overview of the work which led up to and, in many ways, inspired the present work. Thus, upon reading this chapter, a reader should have a full picture of the context for the work done in later chapters.

1.1 PRELIMINARIES ON OPTIMAL TRANSPORTATION

In this section, we introduce the reader to some basic preliminaries on optimal transportation which will be necessary throughout this thesis. This is not by any means meant to provide a complete introduction to the field, and for a quick and understandable introduction, we refer readers to [33]. Additionally, for readers who wish to learn about the discipline in more depth, an excellent general reference for optimal transport is provided in

[37] — in fact, the definitions in this section have generally been either sourced from or cross-referenced with that text.

We begin with some notation on spaces of probability measures.

Definition 1.1 (Spaces of probability measures)**.** We let $\mathcal{P}(\mathbb{R}^n)$ denote the space of Borel probability measures on $\mathbb{R}^n$. For $0 < p < \infty$, we define

$$\mathcal{P}_p(\mathbb{R}^n) := \left\{ \mu \in \mathcal{P}(\mathbb{R}^n) \mid \int |x|^p d\mu(x) < \infty \right\}. \tag{1.1}$$

to be the subspace of probability measures with finite pth moment. In the limiting case, as $p \to \infty$, we define

$$\mathcal{P}_\infty(\mathbb{R}^n) := \{\mu \in \mathcal{P}(\mathbb{R}^n) \mid spt\, \mu \text{ is bounded}\}.$$

Each of these spaces of probability measures comes endowed with a metric structure which, as we will see, is provided by a Wasserstein distance. To study these distances in their natural environment, we will need to first define pushforward measures and transport plans:

Definition 1.2 (Pushforward measure)**.** Given a measure μ on the space X, and a measurable function $F : X \to Y$, we define a probability measure on Y, called **pushforward,** $F_\#\mu$**, of** μ **by** F, by

$$F_\#\mu(A) = \mu(F^{-1}(A))$$

for measurable subsets $A \subset Y$.

Definition 1.3 (Transport plans)**.** We define the space $\Gamma(\mu, \nu)$ of transport plans between the measures $\mu, \nu \in \mathcal{P}(\mathbb{R}^n)$ by

$$\Gamma(\mu, \nu) := \{\gamma \in \mathcal{P}(\mathbb{R}^{2n}) | \pi_1\#\gamma = \mu \text{ and } \pi_{2\#}\gamma = \nu\}.$$

Here, we think of $\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$, and of π_i as the projection map onto the ith copy of $\mathbb{R}^n$.

With these definitions out of the way, we now define the Wasserstein d_p distance

Definition 1.4 (Kantorovich-Rubenstein-Wasserstein distance)**.** Fix $p \in [1, \infty]$ and let $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^n)$. We define

$$d_p(\mu, \nu) := \inf_{\gamma \in \Gamma(\mu,\nu)} \|x - y\|_{L^p(d\gamma)} = \inf_{\gamma \in \Gamma(\mu,\nu)} \left[\int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|^p d\gamma(x, y) \right]^{1/p},$$

where the last identity only holds for $p < \infty$.

Lemma 1.5. *For any $p \in [1, \infty]$, d_p is a metric on $\mathcal{P}_p(\mathbb{R}^n)$.*

Proof. Symmetry and the fact that $d_p(\mu, \nu) = 0$ if and only if $\mu = \nu$ are immediate from the definition of the Wasserstein metric. For the proof that d_p satisfies the triangle inequality, we refer the reader to Section 5.1 of [37] in the $1 \leq p < \infty$ case and Section 5.5.1 of [37] in the $p = \infty$ case. □

Interestingly enough, the d_∞ distance lends itself to a rather simple heuristic interpretation — $d_\infty(\mu, \nu)$ can be thought of as the maximum amount of distance which a non-negligible amount of mass is required to move in order to transport the mass distribution represented by the measure μ to the mass distribution represented by ν. In contrast, for $p < \infty$, while $d_p(\mu, \nu)$ can be interpreted as the lowest cost to transport μ to ν subject to the cost function $c(x, y) = |x - y|^p$, $d_p(\mu, \nu)$ does not admit a simple heuristic interpretation.

We will later consider an energy functional on $\mathcal{P}(\mathbb{R}^n)$ which is translation-invariant. As such, we often find it convenient to work exclusively with probability measures whose barycentres lie at the origin. We define the space of such measures as follows:

Definition 1.6 (Centred Probability Measures)**.** We define the space, $\mathcal{P}_0(\mathbb{R}^n)$, of centred probability measures on $\mathbb{R}^n$ by

$$\mathcal{P}_0(\mathbb{R}^n) := \left\{ \mu \in \mathcal{P}(\mathbb{R}^n) \mid \int x d\mu(x) = 0 \right\}.$$

Finally, although we suspect that the notion of convolution of measures and functions is known to most readers, we provide the definition for reference. A more thorough discussion can be found in [20, Chapter 8.6].

Definition 1.7 (Convolutions of Measures and Functions)**.** Let $p \in [1, \infty]$, let $f \in L^p(\mathbb{R}^n)$, and let μ be a signed measure. Then we define the convolution $f * \mu$ of f and μ by

$$(f * \mu)(x) = \int_{\mathbb{R}^n} f(x - y) d\mu(y).$$

Remark 1.8. As seen in [20, Chapter 8.6], the assumptions in the preceding definition ensure that $f * \mu$ is an L^p function.

1.2 THE AGGREGATION EQUATION

At this point, we introduce one of the key players in this thesis, the aggregation equation, which is given by:

$$\frac{\partial \mu}{\partial t} = \nabla \cdot (\mu \nabla (W * \mu)) \tag{1.2}$$

For the aggregation equation, and indeed throughout this thesis, we assume that $\mu \in \mathcal{P}(\mathbb{R}^n)$. In the specific case of the aggregation equation, we think of $\mu = \mu(t)$ as being a t-parametrized family of such measures, lying in $\mathcal{P}(\mathbb{R}^n) \times [0, \infty)$. We call such a parametrized family of Borel measures a **probability curve** [8]. We typically interpret Borel probability measures as mass distributions, and the parameter t as time. Under this viewpoint, the aggregation equation describes the time evolution of a mass distribution, hinting that it is useful to study the equation and its solutions through optimal transport theory and techniques.

On the other hand, W is simply a Borel-measurable function on $\mathbb{R}^n$. We tend to treat it as a parameter of the aggregation equation, and the diversity of choices for W gives rise to a vast landscape of possible behaviours for solutions to the aggregation equation, which we will discuss in the next section. Depending on the application in the literature, W may be required to be radial or sufficiently regular [4].

It is often interesting to study the steady-states of the aggregation equation — that is, probability measures for which $\frac{\partial \mu}{\partial t} = \nabla \cdot (\mu \nabla W * \mu) = 0$. These steady states often act as attractors for the dynamics of the aggregation equation — if $\overline{\mu}$ is a steady state, and $\mu(0)$ is 'close' to $\overline{\mu}$, in a sense which is to be made precise later with appropriate choice of Wasserstein distance, it is often possible to quantify how 'close' $\mu(t)$ will be to $\overline{\mu}$ for later times t. In some instances it can be shown that $\mu(t)$ converges to $\overline{\mu}$ at an exponential rate [39], whereas in others all we can say is that $\mu(t)$ cannot stray too far away from the steady state [14].

1.3 APPLICATIONS OF THE AGGREGATION EQUATION IN THE SCIENCES

In this section, we introduce the reader to a variety of applications of the aggregation equation in the physical and social sciences. Many of the most interesting applications arise through biological phenomena, but we also discuss applications in physics and game theory. The aim of this section is to convince readers outside of mathematics that the study of the aggregation equation is a valuable and interesting pursuit, with much to say about the sciences. On the other hand, readers who are already satisfied with the aggregation equation's intrinsic mathematical significance and value may freely skip this section without compromising their understanding of the remainder of this thesis.

Depending on the choice of potential, the aggregation equation has many applications in the natural and social sciences. In particular, this equation has often been applied to the swarming, schooling, and flocking behaviours of insects, fish, and birds, respectively. Heuristically, the aggregation equation (equipped with an attractive-repulsive potential W) provides a plausible model of these phenomena for the following reasons. First, such organisms have biological incentives to remain close together - such as safety from predators - balanced with very strong biological incentives to remain a safe distance apart - in particular, movement impediments. Insect swarming is a particularly interesting example of this, as it can be easily modelled through both the two-dimensional aggregation equation (as in the case of ants, roaches, or other flightless insects) or by the three-dimensional aggregation equation (for flying insects such as locusts) [41]. As Topaz and Bertozzi discuss in the introduction of [41], there are two methods of modelling large swarms. The first such method represents each insect as a Dirac mass, which interacts pairwise with other insect Dirac masses through its senses. Of course, such models become complicated when working with a large number of insects, meaning that modelling large numbers of insects through a continuous density function often yields better results. The models involving the aggregation equation which we have just outlined are considered to be some of the simplest models which still generate the emergent phenomena found so often in nature [29]. We refer readers to the text and references of [3, 29] for a full discussion of the applications of the aggregation equation in modelling macroscopic organisms.

Of course, microscopic organisms such as bacteria also have the ability to interact with each other through releasing chemicals, in a process known as chemotaxis [16, 26]. Thus, the heuristic considerations motivating the application of the aggregation equation to macroscopic life also apply to bacteria, and such considerations are formalized through the Patlak-Keller-Segel model, which was pioneered in [28, 36]. There are well-known examples of both attractive and repulsive chemotaxis in both microorganisms and their macroscopic counterparts [26] and this, combined with physical constraints such as overcrowding, means that the aggregation equation is well-suited to describe the behaviour of such organisms. For example, under certain regimes, the aggregation equation has been used to model the process of chemotactic collapse, whereby microorganisms interacting via chemotaxis converge to a single point [24].

Another interesting application of the aggregation equation comes in the domain of game theory and opinion formation. For example, by using a potential which is attractive at short distances and repulsive at long distances, we can model political party formation. That is, imagine that people start out with a broad range of opinions, say, equidistributed over a metric space such as $\mathbb{R}$. Allowing people's opinions to evolve according to the aggregation equation, we find that people who start with rather similar opinions wind up converging to a partial consensus — whereas people who are outside of this basin of ideological attraction will diverge from this partial consensus. In effect, depending on the shape of a potential, people with a wide range of initial opinions can wind up coming to a consensus, or coalescing into a number of ideologically opposed camps [23]. Conversely, Robert McCann has informally suggested that the aggregation equation, when equipped with a potential which is repulsive at short distances and attractive at long distances, could be used to model academic integrity offences. Imagine a Discord channel full of students who are all asked to complete a specific math problem, and who are all aware of each other's solutions. In this situation, there is an attractive force acting to make students' assignments somewhat similar, since no student wants to submit an assignment which is outright incorrect, or radically different from the others' assignments. On the other hand, if any pair of students submits overly similar assignments, then this will surely make the grader suspicious. Thus, in this sense, assignments are both attracted to and repelled from each other.

Of course, the examples provided from above are far from the only applications of the aggregation equation and, as such, we will briefly provide references for some additional selected applications. To begin, most readers of this thesis will also have an application of the aggregation equation on hand — quite literally, as the equation has been used to model fingerprint formation during embryonic development [17]! Additionally, the aggregation equation has been used to model vortex density evolution in superconductors, as appears in [2], and the references of [2, 3, 12]. The aggregation equation has also shown to be a physically plausible model for clumping behaviour of particles, as discussed in [25]. The final interesting application of the aggregation equation to note comes from the study of granular media, which is treated in detail in [9, 10].

1.4 THE INTERACTION ENERGY

One of the most fruitful ways in which we study the aggregation equation and its steady states is through its interaction energy, which is given by

$$\mathcal{E}_W[\mu] = \frac{1}{2} \iint_{\mathbb{R}^n \times \mathbb{R}^n} W(x-y)d\mu(x)d\mu(y). \tag{1.3}$$

The relationship between the aggregation equation and its interaction energy can be viewed through a number of lenses. First, the interaction energy is a Lyapunov functional for the aggregation equation [9], meaning that, if $\mu(t)$ is a probability curve which solves the aggregation equation, then $f(t) := \mathcal{E}_W[\mu(t)]$ is non-increasing.

Another perspective on the relationship between the aggregation equation and its interaction energy comes from the theory of Wasserstein gradient flows on spaces of probability measures, which is explained quickly and readably in [37, Chapter 8], and in full detail in [1]. In particular, the aggregation equation can be thought of as the Wasserstein d_2 gradient flow of $\mathcal{E}_W$, given that W is balanced (i.e. $W(x) = W(-x)$ for all $x \in \mathbb{R}^n$) [37, Chapter 8].

This interaction energy also provides us with a new interpretation of the interaction potential W, in the case that W is balanced with $W(0) = 0$. Namely, $W(x_0 - x_1)$ can be viewed as twice the interaction energy of two-equal mass particles at positions x_0 and x_1. To see why, we use Fubini's Theorem and symmetry to deduce that

$$
\begin{aligned}
2\mathcal{E}_W\left[\frac{1}{2}(\delta_{x_0} + \delta_{x_1})\right] =& \frac{1}{2}\iint_{\mathbb{R}^n \times \mathbb{R}^n} W(x-y)d\delta_{x_0}(x)d\delta_{x_0}(y) \\
&+ \iint_{\mathbb{R}^n \times \mathbb{R}^n} W(x-y)d\delta_{x_0}(x)d\delta_{x_1}(y) \\
&+ \frac{1}{2}\iint_{\mathbb{R}^n \times \mathbb{R}^n} W(x-y)d\delta_{x_1}(x)d\delta_{x_1}(y) \\
=& W(x_0 - x_1),
\end{aligned}
$$

as desired.

We often assume that the potential W is radial or, in other words, that there exists some function $w : [0, \infty) \to \mathbb{R}$ such that $W(x) = w(|x|)$. In most applications, this is not an unreasonable assumption to make, given that the attractive and repulsive forces between any two particles are often taken to depend only on the distance between the two objects. However, in the Appendix, we introduce a rudimentary theory of $k-$welled potentials which might be robust enough to allow for anisotropic, or non-radial, potentials.

1.5 POWER-LAW POTENTIALS AND REGIMES

As mentioned earlier, the qualitative behaviour of solutions of (1.2) is governed by the choice of potential W. Perhaps the most widely-studied class of such potentials are the power-law potentials, which take the form

$$
W_{\alpha,\beta}(x) = \frac{|x|^\alpha}{\alpha} - \frac{|x|^\beta}{\beta}, \tag{1.4}
$$

where $-n < \beta < \alpha < \infty$, with the convention that the terms of the form $\frac{|x|^0}{0}$ are interpreted as $\log|x|$ whenever they appear [6]. While, depending on the values of α and β, these potentials can give rise to a wide range of behaviours, the requirement that $\alpha > \beta > 0$ ensures that such potentials are negative near the origin and positive for large enough $|x|$. We notice that these power law-potentials are radially symmetric since, if we define the function $w_{\alpha,\beta} : (0,\infty) \to \mathbb{R}$ by

$$w_{\alpha,\beta}(r) = \frac{r^\alpha}{\alpha} - \frac{r^\beta}{\beta},$$

then we recognize that

$$W_{\alpha,\beta}(x) = w_{\alpha,\beta}(|x|).$$

In many relevant cases, this allows us to study the aggregation equation with potential $W_{\alpha,\beta}$ by studying the single variable function $w_{\alpha,\beta}$. As such, we now point out some basic properties of $w_{\alpha,\beta}$. By factoring out r^β, and setting $w_{\alpha,\beta}(r) = 0$, we can see that, if $\alpha > \beta > 0$, $w_{\alpha,\beta}(r)$ has precisely two zeroes on $[0,\infty)$, one at $r = 0$ and the other given by $z_{\alpha,\beta} := (\frac{\alpha}{\beta})^{\frac{1}{\alpha-\beta}}$. Looking at derivatives, we see that

$$w'_{\alpha,\beta}(r) = r^{\alpha-1} - r^{\beta-1}$$

is zero only at $r = 1$ and $r = 0$ (provided that $\beta > 1$). Combining this with our knowledge of the end behaviour of $w_{\alpha,\beta}$ shows that $w_{\alpha,\beta}$ has a unique minimum at $r = 1$.

Power law potentials tend to be simple enough that it is feasible to analyze their behaviour, yet complex enough that, depending on the attractive exponent α and repulsive exponent β, steady states of the aggregation equation with $W = W_{\alpha,\beta}$ can have radically different appearances as is seen, for example, in [4, 13, 14, 15, 19, 31, 41, 42]. For a full picture, we direct the reader to the references of [29].

Throughout the literature, there has been a consistent distinction made between the mildly repulsive regime where $\beta > 2$ and the strongly repulsive regime where $\beta < 2$, separated by the centrifugal line $\beta = 2$ [31]. While it is possible to make further distinctions than this, this one is particularly relevant. This is because it has been known since the work of Balagué, Carrillo, Laurent, and Raoul that the supports of energy minimizers in the mildly repulsive regime have Hausdorff dimension zero, whereas supports of energy minimizers in the strongly repulsive regime have nonzero Hausdorff measure [4]. In effect, the line $\beta = 2$ represents a phase transition, from the strongly repulsive regime where the repulsive potential is, as the name suggests, strong enough to counteract the attractive potential and prevent minimizers from concentrating on sets of Hausdorff dimension zero, and the mildly repulsive regime where the repulsive potential is not strong enough to do so. We should also note that, in the case of power-law potentials, numerical experiments have failed to find evidence of energy minimizers whose support has non-integer Hausdorff dimension [4]. Moreover, a recent preprint by Carrillo and Shu partially validates these numerics by showing that, in one dimension, the support of any energy minimizer necessarily has integer Hausdorff dimension (i.e. Hausdorff dimension 0 or 1), although for more exotic potentials, it is possible for energy minimizers to be supported on a set of fractal dimension [11].

The distinction between the mildly and strongly repulsive regimes was further strengthened by Carrillo, Figalli, and Patacchini, who showed that, in the mildly repulsive regime, the support of any global energy minimizer must have finite cardinality [8]. Following up on this work, Lim and McCann used Γ-convergence to show that the unit simplex, a specific measure with only a finite number of points in its support, is the unique minimizer of $\mathcal{E}_{W_{\alpha,\beta}}[\cdot]$ for large enough α in the mildly repulsive regime [31]. Recently, Lim, McCann, and myself were able to show that the unit simplex is indeed the unique energy minimizer on all but a small sliver of the mildly-repulsive regime [15], and this work features prominently in Chapter 2 of the present thesis. To prepare for this, we will define simplices and discuss their properties in the following section.

One outstanding consideration, however, is which the regime the centrifugal line should be considered as a part of. Without going into too much detail, and using the classification based on the Hausdorff dimension of the support of the global minimizer as in [4], my recent work with Lim and McCann indicates that, if $n \geq 2$, energy minimizers have Hausdorff dimension 0 for $(\alpha, \beta) \in (4, \infty) \times \{2\}$ and Hausdorff dimension $n-1$ for $(\alpha, \beta) \in (2, 4) \times \{2\}$ [14, 15]. Due to a uniqueness failure at $(\alpha, \beta) = (4, 2)$, energy minimizers of $\mathcal{E}_{W_{4,2}}$ can have Hausdorff dimension 0 or $n-1$, or possibly Hausdorff dimensions in between. This indicates that parts of the centrifugal line are perhaps most at home with the mildly repulsive regime, and other parts act more similarly to the strongly repulsive regime.

1.6 SIMPLICES AND SPHERICAL SHELLS

One key class of measures which appears in the following work is the simplicial measures, i.e. those which distribute their mass uniformly over the vertices of a regular simplex. More precisely,

Definition 1.9 (Unit Simplex)**.** A unit k-simplex (where $k \leq n$) is any collection of $k+1$ points $\{x_0, ..., x_k\}$ such that $|x_i - x_j| = 1 - \delta_{ij}$ for all $i, j = 0, ..., k$.

Notice that the energy (1.3) is translation invariant. As such, throughout this paper, we will assume any simplices which appear have barycentre at the origin, such that they lie in $\mathcal{P}_0(\mathbb{R}^n)$ as discussed in Definition 1.6. In this case, each point of the simplex is found at distance

$$r_n := \sqrt{\frac{n}{2n+2}}$$

from the origin [30].

Additionally, for computational purposes, it is often convenient to work with one of two canonical versions of the unit n-simplex, as defined below in definitions (1.10) and (1.12):

Definition 1.10 (Canonical unit n-simplex in $\mathbb{R}^n$)**.** The canonical unit simplex in $\mathbb{R}$ is given by the set of points $\{x_0^1, x_1^1\}$, where $x_0^1 := \frac{1}{2}$ and $x_1^1 = -\frac{1}{2}$. The canonical unit simplex in $\mathbb{R}^n$ for $n \geq 2$ is given by the set of points $\{x_0^n, ..., x_n^n\}$ where

$$x_0^n = r_n e_1 \text{ and } x_i^n = -\frac{r_n}{n} e_1 + \sum_{j=1}^{n-1} (x_{i-1}^{n-1} \cdot e_j^{n-1}) e_{j+1}^n \text{ for } 1 \leq i \leq n.$$

Here, $\{e_j^n\}_{j=1}^n$ is the standard orthonormal basis for $\mathbb{R}^n$, and likewise for $\{e_j^{n-1}\}_{j=1}^{n-1}$.

Example 1.11. The canonical unit simplex in $\mathbb{R}^2$ has vertices given by $(\frac{1}{\sqrt{3}}, 0)$, $(-\frac{1}{2\sqrt{3}}, \frac{1}{2})$, and $(-\frac{1}{2\sqrt{3}}, -\frac{1}{2})$. In $\mathbb{R}^3$, these vertices are given by $(\sqrt{\frac{3}{8}}, 0, 0)$, $(-\frac{1}{\sqrt{24}}, \frac{1}{\sqrt{3}}, 0)$, $(-\frac{1}{\sqrt{24}}, -\frac{1}{2\sqrt{3}}, \frac{1}{2})$, and $(-\frac{1}{\sqrt{24}}, -\frac{1}{2\sqrt{3}}, -\frac{1}{2})$. This choice of canonical unit simplex in n dimensions is often useful in calculations, as it maximizes the number of zeroes present in the expression of each vertex, while still working within the ambient space $\mathbb{R}^n$.

Of course, we can express the unit $n-$simplex in an even simpler form by considering it as a subset of $\mathbb{R}^{n+1}$, as is done in [31]:

Definition 1.12 (Canonical unit n-simplex in $\mathbb{R}^{n+1}$)**.** The vertices of the unit $n-$simplex in $\mathbb{R}^{n+1}$ are given by $\{x_0, ..., x_n\}$, where $x_i = \frac{1}{\sqrt{2}} e_{i+1}^{n+1}$.

This definition is also often very useful for calculations, provided that those calculations are unaffected by working in a higher-dimensional ambient space.

The first indication that simplices would be relevant to the study of the aggregation equation may have come from the numerical experiments of Sun, Uminsky, and Bertozzi [40]. These numerics were validated by some recent work of Lim and McCann in [31]. In that paper, the authors first show that, for any $\beta \geq 2$, the simplex minimizes the interaction energy at the **hard confinement limit**, $\alpha = +\infty$ (where the potential $\frac{|x|^\infty}{\infty}$ is interpreted as the pointwise limit as $\alpha \to \infty$ of $\frac{|x|^\alpha}{\alpha}$, i.e. 0 if $|x| \leq 1$ and ∞ if $|x| > 1$). As the name suggests, working with this limit ensures that interaction energy minimizers $\overline{\mu}$ of $\mathcal{E}_{\infty,\beta}$ are confined by a diameter bound on their suppports, which requires that $|x - y| \leq 1$ for all points $x, y \in$spt $\overline{\mu}$. Lim and McCann show that this effect, coupled with the still extant repulsive potential $\frac{|x|^\beta}{\beta}$ ensures that the simplex uniquely minimizes $\mathcal{E}_{\infty,\beta}$, up to translations and rotations.

In the same paper, Lim and McCann build on their work at the hard confinement limit by using techniques of gamma convergence to work perturbatively around the hard confinement limit. In doing so, they show that, for any $\beta \geq 2$, there exists some number $\alpha_\Delta(\beta)$ such that the simplex uniquely minimizes $\mathcal{E}_{\alpha,\beta}$ for all $\alpha > \alpha_\Delta(\beta)$. As the notation suggests, $\alpha_\Delta(\beta)$ can be thought of as a function of β, known as the **threshold function.** For the most part, Lim and McCann left the behaviour of the threshold function as an open question, only providing the explicit lower bound $\alpha_\Delta(2) \geq 4$ [31]. As a preview of what follows, in my recent joint work with Lim and McCann [14, 15], we have explicitly calculated $\alpha_\Delta(2) = 4^*$, where

$$4^* = \begin{cases} 3 & \text{if } n = 1 \\ 4 & \text{if } n \geq 2 \end{cases}.$$

In particular, we must treat the case $n = 1$ separately due to the fact that, in $\mathbb{R}$, there is no distinction between the uniform distribution over the unit 1-simplex and the uniform distribution over a spherical shell of radius $\frac{1}{2}$. Moreover, we have shown that there exists some $\overline{\beta} \leq 4^*$ such that $\alpha_\Delta(\beta) = \beta$ for $\beta \geq \overline{\beta}$. For $\beta \in (2, \overline{\beta})$ we also provide some reasonably strong estimates on $\alpha_\Delta(\beta)$ and, in general, we derive some properties of this threshold function. All of this work is either reproduced or summarized in Chapter 2.

1.7 THE EULER-LAGRANGE EQUATION

We now derive the Euler-Lagrange equation for the interaction energy. Our derivation is analogous to the standard derivation found in Chapter 8 of [18], albeit with some tweaks, as generously suggested by Robert McCann and inspired by [22], to account for the fact that we are applying variational techniques to a functional which takes arguments from $\mathcal{P}(\mathbb{R}^n)$, rather than from a function space.

Proposition 1.13. *Let $\overline{\mu}$ be a global minimizer for (1.3) in $\mathcal{P}(\mathbb{R}^n)$, where $W \in L^1_{loc}(d\mu)$ is balanced in the sense that $W(x) = W(-x)$. Then*

$$\overline{\mu} * W(x) \geq 2\mathcal{E}_W[\overline{\mu}], \text{ with equality } \overline{\mu} - a.e.. \tag{1.5}$$

Proof. We begin by fixing $p \in \operatorname{spt} \overline{\mu}$, and $r > 0$. Since $p \in \operatorname{spt} \overline{\mu}$, we have that $\overline{\mu}(B_r(p)) \neq 0$, and hence we may define a probability measure $\overline{\mu}_{p,r}$ by

$$\overline{\mu}_{p,r}(A) = \frac{\overline{\mu}(A \cap B_r(p))}{\overline{\mu}(B_r(p))}.$$

In turn, for any $q \in \mathbb{R}^n$, we can define the a signed measure $\rho_{p,q,r}$ by

$$\rho_{p,q,r} = \overline{\mu}(B_r(p))(\delta_q - \overline{\mu}_{p,r}),$$

noting that $\rho_{p,q,r}(\mathbb{R}^n) = 0$, and moreover that $\overline{\mu} + t\rho_{p,q,r}$ is a probability measure for any $t \in [0, 1]$.

Thus, we consider the variation of $\mathcal{E}_W$ in the direction of $\rho_{p,q,r}$. That is, we consider the function given by

$$f(t) := \mathcal{E}_W[\overline{\mu} + t\rho_{p,q,r}].$$

As we assumed that $\overline{\mu}$ is a global minimizer of $\mathcal{E}_W[\cdot]$ in $\mathcal{P}(\mathbb{R}^n)$, and as $\overline{\mu} + t\rho_{p,q,r} \in \mathcal{P}(\mathbb{R}^n)$ for $t \in [0, 1]$, we may deduce that $f'(0) \geq 0$, where $f'(0)$ is interpreted as the right-sided derivative. In other words,

$$\lim_{t \to 0^+} \frac{f(t) - f(0)}{t} \geq 0 \tag{1.6}$$

Notice that the condition in Equation (1.6) differs from the equivalent condition in Evans' standard derivation of the Euler-Lagrange equation [18, Chapter 8] because $\overline{\mu} + t\rho_{p,q,r}$ is not, in general a probability measure for $t < 0$, with the only exception occurring when $\overline{\mu}(\{p\}) > 0$. This means that, while $\overline{\mu} + t\rho_{p,q,r}$ is a well-defined path in the space of signed measures $\mathcal{M}(\mathbb{R}^n)$, and although we can easily extend $\mathcal{E}_W$ along this path such that f is differentiable on $\mathbb{R}$, the path, in general, exits $\mathcal{P}(\mathbb{R}^n)$ for negative t. Thus, since we are only concerned with minimizing $\mathcal{E}_W$ over $\mathcal{P}(\mathbb{R}^n)$, and the above discussion shows that it is possible for $\overline{\mu}$ to be a minimizer with $f'(0) > 0$, we need to allow for cases of inequality in Equation (1.6). With this distinction covered, we may directly compute that:

$$\begin{aligned} 2\mathcal{E}_W[\overline{\mu} + t\rho_{p,q,r}] = &\iint_{\mathbb{R}^n \times \mathbb{R}^n} W(x-y) d\overline{\mu}(x) d\overline{\mu}(y) \\ &+ 2t \iint_{\mathbb{R}^n \times \mathbb{R}^n} W(x-y) d\rho_{p,q,r}(x) d\overline{\mu}(y) \\ &\quad + t^2 \iint_{\mathbb{R}^n \times \mathbb{R}^n} W(x-y) d\rho_{p,q,r}(x) d\rho_{p,q,r}(y), \end{aligned}$$

and hence

$$\begin{aligned} 0 \leq f'(0) &= 2 \iint_{\mathbb{R}^n \times \mathbb{R}^n} W(x-y) d\overline{\mu}(y) d\rho_{p,q,r}(x) \\ &= \int_{\mathbb{R}^n} \overline{\mu} * W(x) d\rho_{p,q,r}(x) \\ &= \overline{\mu}(B_r(p)) \left[\overline{\mu} * W(q) - \frac{1}{\overline{\mu}(B_r(p))} \int_{B_r(p)} \overline{\mu} * W(x) d\overline{\mu}(x) \right]. \end{aligned}$$

Rearranging, we deduce that

$$c := \frac{1}{\overline{\mu}(B_r(p))} \int_{B_r(p)} \overline{\mu} * W(x) d\overline{\mu}(x) \leq \overline{\mu} * W(q) \tag{1.7}$$

for our fixed $p \in \operatorname{spt} \overline{\mu}$ and $r > 0$, and for any $q \in \mathbb{R}^n$. Averaging the inequality (1.7) over $B_r(p)$ with respect to $d\overline{\mu}(q)$ yields the same quantity, c, on both sides, and hence we can conclude that (1.7) is saturated for $\overline{\mu}-$a.e. $q \in B_r(p)$. As r is arbitrary, we may take $r \to \infty$ to see that $\overline{\mu} * W(x) = c$ for $\overline{\mu}$-a.e. $x \in \mathbb{R}^n$.

To conclude, we compute the value of c. Since $\overline{\mu} * W(x) = c$ for $\overline{\mu}$-a.e. x, we find that

$$\begin{aligned} 2\mathcal{E}_W[\overline{\mu}] &= \int_{\mathbb{R}^n} \overline{\mu} * W(x) d\overline{\mu}(x) \\ &= \int_{\mathbb{R}^n} c d\overline{\mu}(x) \\ &= c, \end{aligned}$$

finishing the proof.

□

2

GLOBAL MINIMIZERS OF THE INTERACTION ENERGY IN THE MILDLY REPULSIVE REGIME

In this chapter of the thesis, I will discuss some recent joint work with Tongseok Lim and Robert McCann on the global minimizers of the interaction energy (1.3) for the mildly repulsive regime $\alpha > \beta > 2$ and its lower boundary $\alpha > \beta = 2$ [14, 15]. The primary goal of this chapter is to provide readers with a state-of-the-art view of what is known about minimizers of the interaction energy in the mildly repulsive regime.

In Section 2.1, I state some of our key results — in particular, the classification of minimizers on the centripetal line, and the northeast comparison principle. The northeast comparison principle is a powerful tool, which states that, if the unit n-simplex minimizes $\mathcal{E}_{\alpha_0,\beta_0}$ for some pair (α_0, β_0), it minimizes $\mathcal{E}_{W_{\alpha,\beta}}$ for all points (α, β) to the 'northeast' of (α_0, β_0), in a sense to be defined later in this section. We conclude Section 2.1 by showing that lower-dimensional simplices are saddle points of $\mathcal{E}_{W_{\alpha,\beta}}$ with respect to the Wasserstein d_p metric, for any $p \in [1, \infty]$.

Next, in Section 2.2, we introduce the threshold function $\alpha_{\Delta^n}(\beta)$, which has the key property that, for all $\alpha > \alpha_{\Delta^n}(\beta)$, the unit n-simplex is the unique minimizer of $\mathcal{E}_{W_{\alpha,\beta}}$. We then derive a dimension-dependent lower bound and a dimension-independent upper bound for α_{Δ^n}, and show that, as the dimension n tends to ∞, the lower bound tends to the upper bound — meaning that, when working in high-dimensional spaces, we have a very good hold on the behaviour of α_{Δ^n}.

Finally, Section 2.3 serves as a discussion, wherein we speculate about how to improve the bounds found in Section 2.2, and provide numerical evidence supporting this speculation.

2.1 MINMIZERS ON THE CENTRIFUGAL LINE ($n \geq 2$)

In this section, we first give a complete characterization of minimizers of $\mathcal{E}_{W_{\alpha,\beta}}$ along the centrifugal line $\beta = 2$, provided that $n \geq 2$. Then we introduce the northeast comparison principle, which states that, if $\mathcal{E}_{W_{\alpha,\beta}}$ is (possibly non-uniquely) minimized by a simplex at some point (α, β) of the mildly repulsive regime $\alpha > \beta \geq 2$, then both $\mathcal{E}_{W_{\alpha',\beta}}$ and $\mathcal{E}_{W_{\alpha,\beta'}}$ are uniquely minimized by the simplex for any $\alpha' > \alpha$ and $\beta' > \beta$. Then, with this northeast comparison principle in hand, we discuss the one-dimensional case which, as we will see, exhibits substantially different behaviour to that found in higher dimensions. Finally, we treat the case of lower-dimensional simplices in $\mathbb{R}^n$, and show that these are always saddle points of $\mathcal{E}_{W_{\alpha,\beta}}$.

The $n \geq 2$ Case.

Along the centrifugal line ($\beta = 2,\ \alpha > 2$), which serves as the lower boundary of the mildly repulsive regime, we provide a complete characterization of the global minimizers of the interaction energy, provided that $n \geq 2$:

Theorem 2.1 (Energy Minimizers on the Centrifugal Line in Higher Dimensions)**.** *Let $n \geq 2$, and let $\beta = 2$. Then:*

1. *if $\alpha \in (2, 4)$, $\mathcal{E}_{W_{\alpha,2}}$ is uniquely minimized in $\mathcal{P}_0(\mathbb{R}^n)$ by a spherical shell of radius*
$$R_{\alpha,2} := \left[\frac{\Gamma(\frac{n+1}{2})\Gamma(\frac{\alpha}{2} + n - 1)}{\Gamma(n)\Gamma(\frac{\alpha+n-1}{2})} \right]. \tag{2.1}$$
2. *if $\alpha = 4$, $\mathcal{E}_{W_{4,2}}$ is minimized in $\mathcal{P}_0(\mathbb{R}^n)$ precisely by those measures concentrated on sphere of radius $r_n = \sqrt{\frac{n}{2n+2}}$ which have second moment tensor given by*
$$\int x \otimes x d\mu(x) = \left(\int x_i x_j d\mu(x) \right)_{1 \leq i,j \leq n} = \frac{1}{2n+2} Id,$$
where Id is the $n \times n$ identity matrix.
3. *if $\alpha > 4$, $\mathcal{E}_{W_{\alpha,2}}$ is minimized in $\mathcal{P}_0(\mathbb{R}^n)$, uniquely up to rotations, by the uniform distribution over the vertices of a unit n-simplex.*

We sketch the proofs of parts (1) and (2), then postpone the proof of part (3) until after we have discussed the northeast comparison principle, as (3) is a direct corollary of (2) and the northeast comparison principle. The proof sketches of the first two parts are meant to give the reader an intuitive sense of why each of the results holds, and for the full, rigorous proofs, we refer the reader to [14] (for (1)) and [15] (for (2)).

Proof Sketch of Theorem 2.1(1). There are two key steps in this proof: first, we show that any minimizers of $\mathcal{E}_{W_{\alpha,2}}$ must be spherically symmetric, and second, we show that such minimizers are indeed spherical shells.

The first step is to show uniqueness of minimizers in this regime. In particular, we show that the functional $\mathcal{E}_{W_{\alpha,2}}$ is strictly convex in a variational sense. In other words, if μ_0 and μ_1 are probability measures in $\mathcal{P}_0(\mathbb{R}^n)$, and if we define $\mu_t := (1-t)\mu_0 + t\mu_1$, we show that

$$\frac{d^2}{dt^2}\mathcal{E}_{W_{\alpha,2}}[\mu_t] = \iint_{\mathbb{R}^n\times\mathbb{R}^n} W_{\alpha,2}(x-y)d(\mu_1-\mu_0)(x)d(\mu_1-\mu_0)(y) > 0.$$

In order to do so, we consider the functional

$$F_\alpha[\rho] := \iint_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^\alpha d\rho(x)d\rho(y),$$

where ρ is thought of as the difference of two probability measures in $\mathcal{P}_0(\mathbb{R}^n)$. Notice that, in this case,

$$\frac{d^2}{dt^2}\mathcal{E}_{W_{\alpha,2}}[\mu_t] = \frac{1}{\alpha}F_\alpha[\mu_1-\mu_0] - \frac{1}{2}F_2[\mu_1-\mu_0]. \tag{2.2}$$

Moreover, by the definition of the Euclidean inner product,

$$F_2[\mu_1-\mu_0] = \iint_{\mathbb{R}^n\times\mathbb{R}^n}(|x|^2 - 2x\cdot y + |y|^2)d(\mu_1-\mu_0)(x)d(\mu_1-\mu_0)(y) = 0, \tag{2.3}$$

since $\int_{\mathbb{R}^n} d(\mu_1-\mu_0)(x) = \int_{\mathbb{R}^n} xd(\mu_1-\mu_0)(x) = 0$. This means that we need only to show that $F_\alpha[\mu_1-\mu_0] > 0$ for any $\alpha \in (2,4)$ and any two distinct probability measures $\mu_0, \mu_1 \in \mathcal{P}_0(\mathbb{R}^n)$.

This follows from generalizing an argument of Lopes [32] from the case of L^1 probability densities to the case of more general probability measures, as was originally done in [7, Theorem 27]. As we were unaware of this work until Rupert Frank generously informed us of it, Lim, McCann, and I formulated an independent argument in [14]. In short, our argument proceeds by mollifying to get a sequence approximating $\mu_0 - \mu_1$, taking the Fourier transforms of the mollified sequences as in [32], then taking the limit on both sides to recover the identity in general. It should be mentioned that our proof relies on the fact that the first and second moments of $\mu_0 - \mu_1$ vanish, as this, through Laurent Schwartz' Paley-Wiener Theorem for distributions [38], is what ensures that the mollified Fourier transform vanishes to sufficiently high order at 0.

This is enough to show strict convexity and hence uniqueness of the minimizer of $\mathcal{E}_{W_{\alpha,2}}$ for $\alpha \in (2,4)$. By a standard convexity argument, if $\overline{\mu}_0$

and $\overline{\mu}_1$ are distinct candidate minimizers, then strict convexity ensures that

$$\mathcal{E}_{W_{\alpha,2}}\left[\frac{\overline{\mu}_0 + \overline{\mu}_1}{2}\right] < \mathcal{E}_{W_{\alpha,2}}[\overline{\mu}_0] = \mathcal{E}_{W_{\alpha,2}}[\overline{\mu}_1],$$

and hence neither $\overline{\mu}_0$ nor $\overline{\mu}_1$ minimize $\mathcal{E}_{W_{\alpha,2}}[\cdot]$. Moreover, as $\mathcal{E}_{W_{\alpha,2}}[\cdot]$ is rotation-invariant, we may conclude that the minimizer of $\mathcal{E}_{W_{\alpha,2}}[\cdot]$ is spherically symmetric.

The second step of the proof is to show that the unique, spherically symmetric minimizer is indeed a spherical shell. To show this, we use elementary methods to examine the radial profile of $W_{\alpha,2} * \overline{\mu}$, where $\overline{\mu}$ is taken to be the unique minimizer of $\mathcal{E}_{W_{\alpha,2}}$. In particular, we show that this radial profile has positive third derivative and hence, by the Euler-Lagrange equation (1.5), the support of $\overline{\mu}$ must coincide with a spherical shell. Finally, by considering spherical shells σ_r and optimizing $\mathcal{E}_{W_{\alpha,2}}(\sigma_r)$ over r, we can derive the optimal radius in equation (2.1), which we note was originally derived by [5] in an analogous context. □

We now move on to the next case:

Proof Sketch of Theorem 2.1(2). Notice that, in this case, $\mathcal{E}_{W_{4,2}}$ is still a convex functional on $\mathcal{P}_0(\mathbb{R}^n)$. This can be seen by allowing $\alpha \nearrow 4$. However, this argument is not enough to yield strict convexity. Thus, in order to derive necessary and sufficient conditions for $\mathcal{E}_{W_{4,2}}[\cdot]$ to be convex along lines in $\mathcal{P}(\mathbb{R}^n)$, we employ the notation of the previous proof sketch to derive that:

$$\begin{aligned} F_4[\mu_1 - \mu_0] &= \iint_{\mathbb{R}^n \times \mathbb{R}^n} (|x|^2 - 2x \cdot y + |y|^2)^2 d(\mu_1 - \mu_0)(x) d(\mu_1 - \mu_0)(y) \\ &= 2 \iint_{\mathbb{R}^n \times \mathbb{R}^n} |x|^2 |y|^2 d(\mu_1 - \mu_0)(x) d(\mu_1 - \mu_0)(y) \\ &\quad + 4 \iint_{\mathbb{R}^n \times \mathbb{R}^n} (x \cdot y)^2 d(\mu_1 - \mu_0)(x) d(\mu_1 - \mu_0)(y) \\ &= 2 \left[\int_{\mathbb{R}^n} |x|^2 d(\mu_1 - \mu_0)(x) \right]^2 + 4 \sum_{i,j=1}^{n} \left[\int_{\mathbb{R}^n} x_i x_j d(\mu_1 - \mu_0)(x) \right]^2. \end{aligned}$$

We recognize the first integral which appears above as the trace of the second moment tensor $I(\mu_1 - \mu_0)$ of $\mu_1 - \mu_0$, and the second sum as the trace of $I(\mu_1 - \mu_0)^2$, allowing us to write

$$F_4[\mu_1 - \mu_0] = 2(\operatorname{Tr} I(\mu_1 - \mu_0))^2 + 4 \operatorname{Tr}(I(\mu_1 - \mu_0)^2).$$

Here, we notice that, since the second moment tensor is linear, $F_4[\mu_1 - \mu_0]$ is zero if and only if $I(\mu_1) = I(\mu_0)$. Since, in light of equation (2.3), $F_2[\mu_1 - \mu_0]$ is identically zero, we can conclude that $\frac{d^2}{dt^2}\mathcal{E}_{W_{4,2}}(\mu_t) = 0$ if and only if $I(\mu_1) = I(\mu_0)$. This means that, if $\overline{\mu}_0$ and $\overline{\mu}_1$ are candidate minimizers of $\mathcal{E}_{W_{4,2}}[\cdot]$ with $I(\overline{\mu}_1) \neq I(\overline{\mu}_0)$, we can apply the uniqueness argument from the proof of Theorem 2.3(1) to show that neither $\overline{\mu}_0$ nor $\overline{\mu}_1$ actually minimize $\mathcal{E}_{W_{4,2}}$. Thus, since [13, Theorem 2.3] shows that there exists a minimizer of $\mathcal{E}_{W_{4,2}}$ in $\mathcal{P}_0(\mathbb{R}^n)$, and since $\mathcal{E}_{W_{\alpha,\beta}}$ is translation-invariant, we find that all global minimizers of $\mathcal{E}_{W_{4,2}}$ have the same second moment tensor.

To determine what this second moment tensor is, we notice that $\mathcal{E}_{W_{4,2}}$ is still convex, even if not strictly so, and hence we may take any given minimizer $\overline{\mu}$ and average over its rotations, in order to get a radially symmetric, centred global minimizer of $\mathcal{E}_{W_{4,2}}$, which we will denote by $\overline{\sigma}$. It is clear to see that, by radial symmetry,

$$\int x_i^2 d\overline{\mu}(x) = \int x_j^2 d\overline{\sigma}(x)$$

for any $i, j = 1, ..., n$. Moreover, a quick application of Fubini's Theorem implies that, for any $i \neq j$,

$$\int x_i x_j d\overline{\sigma}(x) = 0.$$

Hence, we can write the shared second moment tensor of minimizers of $\mathcal{E}_{W_{4,2}}$ as $\lambda\mathrm{Id}$ for some constant $\lambda > 0$.

However, only some measures with second moment tensor $\lambda\mathrm{Id}$ will actually minimize $\mathcal{E}_{W_{4,2}}[\cdot]$. To see why this is the case, we let μ be an arbitrary probability measure with second moment tensor given by $\lambda\mathrm{Id}$ and directly calculate

$$\begin{aligned}
\mathcal{E}_{W_{4.2}}[\mu] &= \frac{1}{8}F_4[\mu] - \frac{1}{4}F_2[\mu] \\
&= \frac{1}{4}\int_{\mathbb{R}^n} |x|^4 d\mu(x) + \frac{1}{4}\left[\int |x|^2 d\mu(x)\right]^2 \\
&\quad + \frac{1}{2}\sum_{i,j=1}^{n}\left[\int_{\mathbb{R}^n} x_i x_j d\mu(x)\right]^2 - \frac{1}{2}\int_{\mathbb{R}^n} |x|^2 d\mu(x) \\
&= \frac{1}{4}\int_{\mathbb{R}^n} |x|^4 d\mu(x) + \frac{(\mathrm{Tr}\, I(\mu))^2}{4} + \frac{\mathrm{Tr}(I(\mu))^2}{2} - \frac{\mathrm{Tr}\, I(\mu)}{2}. \qquad (2.4)
\end{aligned}$$

While the last three terms in (2.4) depend only μ only through $I(\mu)$, the first one can vary, even among measures with the same second moment tensor. Hence, the problem of finding the minimizers of $\mathcal{E}_{W_{4,2}}$ is equivalent to that of minimizing $\int_{\mathbb{R}^n} |x|^4 d\mu(x)$ over the set of measures μ with second moment tensor given by $\lambda \mathrm{Id}$. Of course, knowing that $I(\mu) = \lambda \mathrm{Id}$ implies that $\int_{\mathbb{R}^n} |x|^2 d\mu(x) = \mathrm{Tr}\, I(\mu) = n\lambda$, we may use the (less restrictive) constraint that $\int_{\mathbb{R}^n} |x|^2 d\mu(x) = n\lambda$. This allows us to apply [15, Lemma 2.4], which tells us that, in order to minimize $\int_{\mathbb{R}^n} |x|^4 d\mu(x)$, and hence $\mathcal{E}_{W_{4,2}}[\mu]$, subject to the constraint $\int_{\mathbb{R}^n} |x|^2 d\mu(x) = n\lambda$, μ must concentrate all of its mass on a spherical shell. Moreover, [15, Lemma 2.4] provides the radius of this spherical shell in terms of λ — in particular, it must be given by $r(\lambda) = \sqrt{n\lambda}$.

All that remains is to solve for the optimal value of λ. Notice that, since any global minimizer $\overline{\mu}$ is a probability measure supported on a sphere of radius $\sqrt{n\lambda}$, we find that $\int_{\mathbb{R}^n} |x|^4 d\overline{\mu}(x) = \lambda^2 n^2$. Likewise, since $\overline{\mu}$ has second moment tensor given by $I(\overline{\mu}) = \lambda \mathrm{Id}$, we find that $\mathrm{Tr}\, I(\overline{\mu}) = n\lambda$ and $\mathrm{Tr}(I(\overline{\mu}))^2 = n\lambda^2$. Hence, Equation (2.4) simplifies to

$$\mathcal{E}_{W_{4,2}}[\overline{\mu}] = \frac{1}{2}(n^2\lambda^2 + n\lambda^2 - n\lambda).$$

By basic calculus, the expression on the right hand side is minimized when $2(n^2 + n)\lambda = n$ or, in other words, when $\lambda = \frac{1}{2n+2}$, precisely as desired. To conclude, we use this value as λ in [15, Lemma 2.4] to see that $\overline{\mu}$ is supported on a sphere of radius $r_n := \sqrt{n\lambda} = \sqrt{\frac{n}{2n+2}}$, precisely as desired. □

The Northeast Comparison Principle and Immediate Applications

As discussed in the introduction to this chapter, we will now discuss the northeast comparison principle [15], which is vital for determining for which values of α and β the simplex minimize $\mathcal{E}_{W_{\alpha,\beta}}$. The statement of the comparison principle is as follows:

Theorem 2.2 (Northeast Comparison Principle)**.** *Assume that the unit simplex ν minimizes $\mathcal{E}_{W_{\alpha,\beta}}$ in $\mathcal{P}_0(\mathbb{R}^n)$. Then, for any $\alpha' \geq \alpha$ and $\beta' \geq \beta$ such that $(\alpha', \beta') \neq (\alpha, \beta)$, $\mathcal{E}_{W_{\alpha',\beta'}}$ is uniquely minimized in $\mathcal{P}_0(\mathbb{R}^n)$ by ν and its rotates.*

We refer the reader to [15] for a proof of the comparison principle, noting only that the proof relies on looking at the α and β derivatives of the normalized radial profile of the potential

$$\overline{w}_{\alpha,\beta}(r) := \frac{\alpha\beta}{\alpha - \beta} w_{\alpha,\beta}(r) = \frac{\beta r^\alpha - \alpha r^\beta}{\alpha - \beta}. \tag{2.5}$$

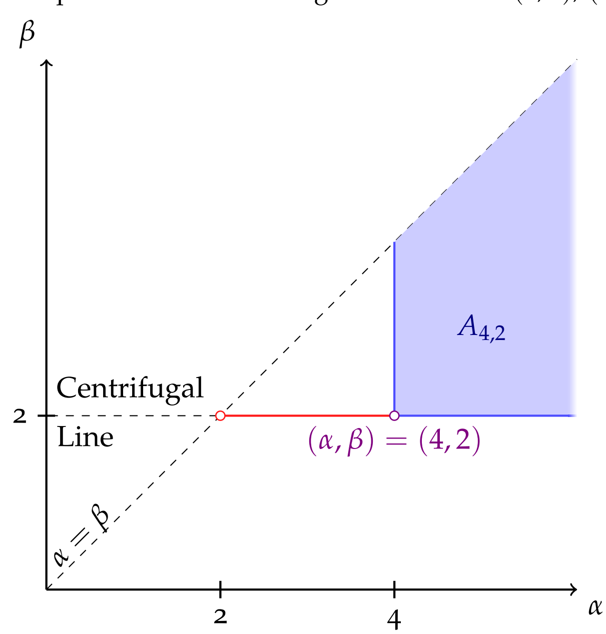
The northeast comparison principle allows us to provide a succinct proof of Theorem 2.1(3):

Proof of Theorem 2.1(3). Notice that the unit $n-$simplex ν has second moment tensor given by $\frac{1}{2n+2}\mathrm{Id}$. Thus, by Theorem 2.1(2), ν minimizes $\mathcal{E}_{W_{4,2}}$ and hence, by Theorem 2.2, ν uniquely minimizes $\mathcal{E}_{W_{\alpha,2}}$ for any $\alpha > 4$. □

Remark 2.3 ('North' Comparison)*.* The preceding proof only uses part of the power of the comparison principle, as it is only comparing energies for changing α (i.e. an 'east' comparison). Using the full power of the comparison principle to do a 'north' comparison as well, we see that $\mathcal{E}_{W_{\alpha,\beta}}$ is uniquely minimized (up to rotations) in $\mathcal{P}_0(\mathbb{R}^n)$ by the unit simplex for any pair of parameters (α, β) such that $\alpha > \beta$, $\alpha \geq 4$, $\beta \geq 2$, and $(\alpha, \beta) \neq (4, 2)$.

To conclude this subsection, we summarize the results of Theorem 2.1 and Remark 2.3 in the following diagram:

Figure 2.1: A phase diagram, reproduced from [15], of the mildly repulsive regime (for dimension $n \geq 2$) which summarizes the results of Theorem 2.1. Theorem 2.1(1) shows that $\mathcal{E}_{W_{\alpha,2}}$ is uniquely minimized by a spherical shell along the red line segment, and Theorem 2.1(2) characterizes the minimizers of $\mathcal{E}_{W_{\alpha,\beta}}$ at the purple point $(\alpha, \beta) = (4, 2)$. Since Theorem 2.1(2) implies that the unit simplex ν is one of the minimizers of $\mathcal{E}_{W_{4,2}}$, it follows from the northeast comparison principle that ν uniquely minimizes $\mathcal{E}_{W_{\alpha,2}}$ on the entire blue region labeled $A_{4,2}$. As we will see in the next section, the unit simplex ν does in fact minimize $\mathcal{E}_{W_{\alpha,\beta}}$ on at least part of the white triangle with vertices $(2, 2)$, $(4, 2)$, and $(4, 4)$.



Minimizers on the Centrifugal Line ($n = 1$)

As mentioned in the introduction to this section, energy minimizers for the interaction energy (1.3) behave quite differently when working on $\mathbb{R}$ when compared to the $\mathbb{R}^n$ for $n \geq 2$ case. Heuristically, this can be explained by noticing that, if we are given a configuration of k particles on $\mathbb{R}$, which we will denote by $\{x_1, ..., x_k\}$ and wish to move one such particle, say x_{k_0}, then inevitably we have to move x_{k_0} either closer to or further away from each of the other particles x_j. In other words, particles do not have a great degree of freedom in how they move. On the other hand, even in two dimensions, particles can move much more freely, and this, as a comparison of our earlier results with those we will discuss in this subsection shows, markedly changes the behaviour of the steady states of the aggregation equation.

It should also be noted that, in one dimension, the unit 1-simplex and the sphere of radius $\frac{1}{2}$ are the same shape, and hence the uniform distributions over each of these are precisely the same measure. This means that the northeast comparison principle applies equally well to spherical shell measures. In particular, we have shown the following result in [14]:

Theorem 2.4 (Energy Minimizers on the Centrifugal Line in One Dimension)**.** *Let $n = 1$, $\beta = 2$, and $\alpha \geq 3$. Then the centred unit simplex ν is the unique minimizer of $\mathcal{E}_{W_{\alpha,\beta}}$ in $\mathcal{P}_0(\mathbb{R})$.*

We leave the tried-and-true proof to [14], and instead briefly outline a potential alternative strategy:

Alternative strategy. By the northeast comparison principle, it suffices to prove the theorem in the case $(\alpha, \beta) = (3, 2)$. It seems feasible that this can be done relatively easily through a direct proof, which would complete the alternative argument. □

While our results have little to say about the portion of the centrifugal line where $\alpha \in (2, 3)$, shortly after we posted our preprint [15] on arxiv, Rupert Frank resolved this case, and we reproduce his result below, with some minor tweaks to account for different notation conventions:

Theorem 2.5. *[Theorem 1 of [21]] Let $2 < \alpha < 3$, and set*

$$R_\alpha := \left(\frac{\sqrt{\pi}}{2} \frac{\Gamma(\frac{3-\alpha}{2})}{\Gamma(\frac{4-\alpha}{2})} \frac{\sin((\alpha-1)\frac{\pi}{2})}{(\alpha-1)\frac{\pi}{2}} \right)^{\frac{1}{\alpha-2}}.$$

Then

$$\inf_{\mu \in \mathcal{P}(\mathbb{R})} \mathcal{E}_{W_{\alpha,2}}[\mu] = -\frac{\alpha-2}{2\alpha(4-\alpha)} R_\alpha^2.$$

Moreover, the infimum is attained if and only if, for some $a \in \mathbb{R}$,

$$d\mu(x) = \frac{\Gamma(\frac{4-\alpha}{2})}{\sqrt{\pi}\Gamma(\frac{3-\alpha}{2})} R_\alpha^{\alpha-2}(R_\alpha^2 - (x-a)^2)^{-\frac{\alpha-1}{2}} \mathbb{1}(|x-a| < R_\alpha)dx.$$

Of course, throughout this thesis, we have used translation invariance of the interaction energy to restrict our focus to $\mathcal{P}_0(\mathbb{R}^n)$. Thus, for convenience, we specialize Frank's result to this case below:

Corollary 2.6 (Application to $\mathcal{P}_0(\mathbb{R}^n)$)**.** *Let $2 < \alpha < 3$, and let R_α be defined as in the statement of Theorem 2.5. Then $\mathcal{E}_{W_{\alpha,2}}$ is uniquely minimized in $\mathcal{P}_0(\mathbb{R}^n)$ by the measure $\overline{\mu}$, defined by*

$$d\overline{\mu}(x) = \frac{\sin((\alpha-1)\frac{\pi}{2})}{\pi(\alpha-1)}(R_\alpha^2 - x^2)^{-\frac{\alpha-1}{2}} \mathbb{1}(|x| < R_\alpha)dx.$$

A Few Words on Lower-Dimensional Simplices

Given that the unit n-simplex is the global minimizer of $\mathcal{E}_{W_{\alpha,\beta}}$ over most of the mildly repulsive regime when considering the aggregation equation in $\mathbb{R}^n$, it is natural to ask whether the unit n-simplex continues to minimize $\mathcal{E}_{W_{\alpha,\beta}}$ in $\mathbb{R}^m$, for $m > n$. While Theorems 2.1 and 2.2 imply that such measures cannot be global minimizers of $\mathcal{E}_{W_{\alpha,\beta}}$, we may still ask whether lower-dimensional simplices are local minimizers with respect to d_p for some $p \in [1, \infty]$.

This question turns out to be answered in the negative - lower dimensional simplices are always saddle points, as the following proposition shows:

Proposition 2.7. *Fix $m \in \mathbb{N}$, $\alpha > \beta > -m$, $n < m$, and $p \in [1, \infty]$. Then the unit n-simplex ν_n is a d_p saddle point of $\mathcal{E}_{W_{\alpha,\beta}}$ in $\mathcal{P}(\mathbb{R}^m)$.*

Proof. Fix $\varepsilon > 0$. We will find measures μ_n^+ and μ_n^- in $\mathcal{P}(\mathbb{R}^m)$ such that $d_p(\mu_n^+, \nu_n) < \varepsilon$, $d_p(\mu_n^-, \nu_n) < \varepsilon$, $\mathcal{E}_{W_{\alpha,\beta}}[\mu_n^+] > \mathcal{E}_{W_{\alpha,\beta}}[\nu_n]$ and $\mathcal{E}_{W_{\alpha,\beta}}[\mu_n^-] < \mathcal{E}_{W_{\alpha,\beta}}[\nu_n]$.

We will denote the set of vertices of the simplex ν_n as $\{x_0, ..., x_n\}$, and let $x_0' := x_0 + \frac{\varepsilon}{n}\sum_{i=1}^n x_i$. Define μ_n^+ to be the uniform distribution over $\{x_0', x_1, ..., x_n\}$. We notice that

$$|x_0' - x_0| = |\frac{\varepsilon}{n}\sum_{i=1}^n x_i| \leq \frac{\varepsilon}{n}\sum_{i=1}^n |x_i| < \varepsilon,$$

since we have that $|x_i| = r_n < 1$. To check that $d_p(\nu_n, \mu_n^+) < \varepsilon$, as we consider the transport plan $\pi_n^+ := \frac{1}{n+1}\delta_{(x_0,x_0')} + \frac{1}{n+1}\sum_{i=1}^n \delta_{(x_i,x_i)}$ between ν_n and μ_n^+, and notice that

$$d_p(\nu_n, \mu_n^+) \leq \left[\iint_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^p d\pi_n^+(x,y)\right]^{1/p} = \left[\frac{1}{n+1}|x_0 - x_0'|^p\right]^{1/p} < \varepsilon.$$

On the other hand, we may directly compute

$$\mathcal{E}_{W_{\alpha,\beta}}[\mu_n^+] - \mathcal{E}_{W_{\alpha,\beta}}[\nu_n] = \frac{2}{n+1}\sum_{i=1}^n \left[w_{\alpha,\beta}(|x_0' - x_i|) - w_{\alpha,\beta}(|x_0 - x_i|)\right]. \quad (2.6)$$

Now, notice that $\frac{\varepsilon}{n}\sum_{i=1}^n x_i = -cx_0$ for some constant $c = c(\varepsilon) > 0$. Hence,

$$|x_0' - x_i| = |x_0 - x_i + \frac{\varepsilon}{n}\sum_{i=1}^n x_i| = |(1-c)x_0 - x_i|.$$

Furthermore, notice that, for small $\varepsilon > 0$ (and hence $c > 0$) $(1-c)x_0$ lies in the interior of the simplex with vertices $x_0, ..., x_n$ and hence

$$|x_0' - x_i| = |(1-c)x_0 - x_i| < 1 = |x_0 - x_i|.$$

Thus, since $w_{\alpha,\beta}(r)$ is uniquely minimized at $r = 1$, we find that each term in (2.6) is strictly positive, and hence $\mathcal{E}_{W_{\alpha,\beta}}[\mu_n^+] > \mathcal{E}_{W_{\alpha,\beta}}[\nu_n]$, as desired.

Next, we notice that, since $m > n$, the set $X := \{x \in \mathbb{R}^m \mid |x - x_i| = 1 \text{ for } i = 1, ..., n\}$ can also be characterized as the intersection of the hypersphere

$$S := \left\{x \in \mathbb{R}^m \;\middle|\; \left|x - \frac{1}{n}\sum_{i=1}^n x_i\right| = \left|x_0 - \frac{1}{n}\sum_{i=1}^n x_i\right|\right\}$$

and the linear space $P :=\text{span}(\{x_1, ...x_n\})^\perp$. This means $S \cap P$ is a sphere of dimension at least one. Thus, we may find two distinct points, a and b in $S \cap P$ such that $|x_0 - a| < \varepsilon$, $|x_0 - b| < \varepsilon$, and $|a - b| > 0$. Thus, we define

$$\mu_n^- := \frac{1}{2(n+1)}(\delta_a + \delta_b) + \frac{1}{n+1}\sum_{i=1}^n \delta_{x_i}.$$

Defining the transport plan

$$\pi_n^- := \frac{1}{2(n+1)}(\delta_{x_0,a} + \delta_{x_0,b}) + \frac{1}{n+1}\sum_{i=1}^n \delta_{x_i},$$

we see that

$$\begin{aligned} d_p(\nu_n, \mu_n^-) &\leq \left[\iint_{\mathbb{R}^n \times \mathbb{R}^n} |x-y|^p d\pi_n^-(x-y) \right]^{1/p} \\ &= \left[\frac{1}{2(n+1)} (|x_0 - a|^p + |x_0 - b|^p) \right]^{1/p} \\ &= \left[\frac{1}{(n+1)} \varepsilon^p \right]^{1/p} < \varepsilon. \end{aligned}$$

Finally, we may calculate

$$\mathcal{E}_{W_{\alpha,\beta}}[\mu_n^-] - \mathcal{E}_{W_{\alpha,\beta}}[\nu_n] = \frac{1}{n+1} w_{\alpha,\beta}(|a-b|), \tag{2.7}$$

as all other terms are zero (in the case of self-interaction), or cancel out (in the case of repeated terms). Thus, since we can readily take $\varepsilon < \frac{1}{2}$, and since this implies that $0 < |a-b| < 1$, we find that the difference in (2.7) is negative, finishing the proof.

□

To help the reader gain more intuition behind the preceding proof, we include a diagram portraying the supports of μ_n^+ and μ_n^- in the case $(m,n) = (2,1)$ (i.e. the unit 1−simplex embedded in $\mathbb{R}^2$):

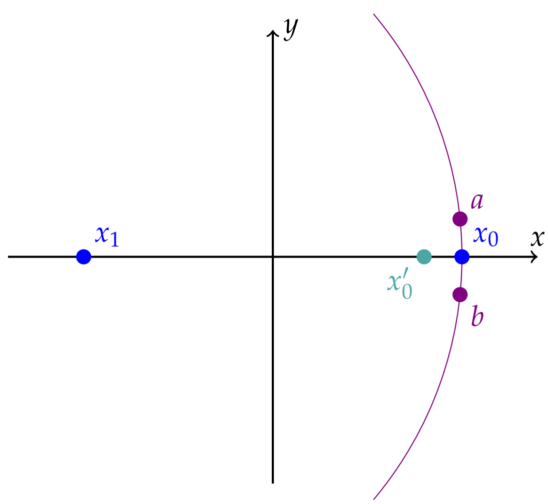


Figure 2.2: Images of spt $\nu_1 = \{x_0, x_1\}$, spt $\mu_1^+ = \{x_0', x_0\}$, and spt $\mu_1^- = \{a, b, x_1\}$ in $\mathbb{R}^2$. When defining μ_1^+, we move one vertex ν_1 closer to the other, and when defining μ_1^-, we split the mass into two pieces, and move each piece along the violet circle, without changing the distance between each piece and x_1.

2.2 AN UNFINISHED PORTRAIT OF THE MILDLY REPULSIVE REGIME

We now turn our attention to a thorough treatment of the threshold function $\alpha_{\Delta^n}(\beta)$, whose graph forms the boundary between the portion of the mildly repulsive regime where the unit n-simplex uniquely minimizes $\mathcal{E}_{W_{\alpha,\beta}}$ and the portion where it is not a minimizer. While we are not able to provide an explicit formula for this function, we do provide explicit, usable bounds and derive certain properties of the threshold function. Most of the work in this section is found in [15], and as such represents the current state of knowledge on minimizers of the interaction energy — although, as we emphasize, little is known about minimizers of $\mathcal{E}_{W_{\alpha,\beta}}$ for $\alpha < \alpha_{\Delta^n}(\beta)$.

Introduction to the Threshold Function

Since the work of Lim and McCann in [31, Remark 1.5], it has been known that, for each $\beta \geq 2$, there exists a phase transition $\alpha_{\Delta^n}(\beta)$ such that $\mathcal{E}_{W_{\alpha,\beta}}$ is uniquely minimized (up to rotations and translations) in $\mathcal{P}(\mathbb{R}^n)$ by the unit simplex ν_n. More precisely,

Definition 2.8 (Threshold Function)**.** We define the **threshold function** $\alpha_{\Delta^n} : [2, \infty) \to [2, \infty)$ by

$$\alpha_{\Delta^n}(\beta) = \sup\{\alpha \mid \text{ there exists a minimizer of } \mathcal{E}_{W_{\alpha,\beta}} \\ \text{which is not a rotated translate of } \nu_n\}.$$

Until our joint work in [15], very little was known about the behaviour of α_{Δ^n}, as the pioneering work in [31] was done by perturbing around the hard confinement limit $\alpha = \infty$, and as such, for the most part, could only prove existence of α_{Δ^n}. The only exception to this is Lim and McCann's proof that $\alpha_{\Delta^n}(2) \geq 4$ [31, Remark 1.5], wherein they adapted an argument of Lopes [32] analogous to the adaptation found in the proof of Theorem 2.1(1).

However, the discovery of the northeast comparison principle (Theorem 2.2) has allowed us to derive many additional properties of α_{Δ^n}. Even work in the previous section sheds a lot of light on this matter, as is illustrated by the following corollary to Theorems 2.1 and 2.2:

Corollary 2.9 (Immediate Results on the Threshold Function)**.** *Let $\beta \geq 2$, and let $\alpha_{\Delta^n}(\beta)$ be as defined above. Then:*

- *if $\beta = 2$, then $\alpha_{\Delta^n}(\beta) = 4$*
- *if $2 < \beta < 4$, then $\alpha_{\Delta^n}(\beta) \leq 4$*

- *if* $\beta \geq 4$*, then* $\alpha_{\Delta^n(\beta)} = \beta$*.*

In particular, this means that α_{Δ^n} is totally determined when β lies outside of the interval $(2, 4)$. However, using the results from the previous section, as well as the notion of Γ-convergence, it is possible to deduce even more about the behaviour of the threshold function, as is done in [15, Theorem 4.1]. We reproduce this result here, but refer the reader to the original work for the proof:

Theorem 2.10 (Behaviour of the Threshold Function)**.** *Let* α_{Δ^n} *be as defined above. Then, for any* $\alpha < \alpha_{\Delta^n}(\beta)$*, the unit simplex does not minimize* $\mathcal{E}_{W_{\alpha,\beta}}$ *in* $\mathcal{P}(\mathbb{R}^n)$*. Moreover, if* $\alpha = \alpha_{\Delta^n}(\beta)$ *and* ν *is a unit simplex, then either there are non-simplicial minimizers of* $\mathcal{E}_{W_{\alpha,\beta}}$ *or* $\operatorname{spt} \nu \subsetneq \operatorname{argmin}(W_{\alpha,\beta} * \nu)$*. Finally, there exists some* $\beta_n \leq 4$ *such that* $\alpha_{\Delta^n}(\beta) = \beta$ *for all* $\beta \geq \beta_n$*, and such that* $\alpha_{\Delta^n} : [2, \beta_n] \to [\beta_n, 4]$ *is continuous and strictly decreasing.*

Other than the bounds which will be discussed in the remainder of this section, Theorem 2.10 represents the current state of knowledge on α_{Δ^n}.

Additionally, the reader should recall that, if $\beta > 2$ and α lies in the non-trivial interval $(\beta, \alpha_{\Delta^n}(\beta))$, then Carrillo, Figalli, and Patacchini [8] have shown that any global minimizer of $\mathcal{E}_{W_{\alpha,\beta}}[\cdot]$ is supported on a finite point set. Moreover, in the case $n = 1$, the trio derived an explicit upper bound on the cardinality of this finite point set [8].

An Upper Bound for the Threshold Function

We now establish an upper bound for α_{Δ^n} which is almost independent of the dimension n, in the sense that this upper bound has different formulas for the $n = 1$ case and $n \geq 2$ case, but that the $n \geq 2$ formula is independent of the exact value of n.

Definition 2.11. Let $\beta \in [2, \beta_\infty]$, where $\beta_\infty := \frac{1}{\log(3/2)}$ if $n = 1$ and $\beta_\infty := \frac{2}{\log 2}$ if $n \geq 2$. We define $\alpha_\infty = \alpha_\infty(\beta)$ as the largest solution of

$$\frac{e^{\alpha/\beta_\infty}}{\alpha} = \frac{e^{\beta/\beta_\infty}}{\beta}. \tag{2.8}$$

Remark 2.12 (Number of solutions)*.* For any given $\beta \geq 2$, there are at most two solutions to equation (2.8), which follows from the fact that $-\frac{e^{t/\beta_\infty}}{t}$ is unimodal on $(0, \infty)$, in the sense which we will later discuss in Lemma 2.18. In particular, we see that

$$-t^2 \beta_\infty e^{-t/\beta_\infty} \frac{d}{dt} \frac{e^{t/\beta_\infty}}{t} = t - \beta_\infty$$

is positive on $(0, \beta_\infty)$, zero at β_∞, and negative on (β_∞, ∞).

Remark 2.13 (Alternative interpretation). In the context of (2.5), we understand

$$\overline{w}_{\beta,\beta}(r) := \lim_{\alpha\to\beta} \overline{w}_{\alpha,\beta}(r) = r^{\beta}(\beta \log r - 1),$$

and let $z_{\alpha,\beta}$ denote the positive zero of $\overline{w}_{\alpha,\beta}$, where $z_{\alpha,\beta} = (\frac{\alpha}{\beta})^{\frac{1}{\alpha-\beta}}$ for $\alpha \neq \beta$ and $z_{\beta,\beta} := e^{1/\beta}$. Notice that $z_{4^*,2} = \frac{3}{2}$, if $n = 1$, and $z_{4^*,2} = \sqrt{2}$, if $n \geq 2$. Hence, after some rearranging, we may define β_∞ by the equation $z_{\beta_\infty,\beta_\infty} = z_{4^*,2}$ and α_∞ as the largest solution of $z_{\alpha,\beta} = z_{4^*,2}$, or rather $w_{\alpha,\beta}(z_{4^*,2}) = 0$.

The following proposition and corollary demonstrate that α_∞ is indeed an upper bound for the threshold function:

Proposition 2.14. *Let $2 < \beta < \alpha < 4^*$. Then $\overline{w}_{4^*,2}(r) \leq \overline{w}_{\alpha,\beta}(r)$ for all $r \in [0, z_{\alpha,\beta}]$ if and only if $z_{\alpha,\beta} \leq z_{4^*,2}$.*

Proof. One direction is trivial, as if $\overline{w}_{4^*,2}(r) \leq \overline{w}_{\alpha,\beta}(r)$ for all $r \in [0, z_{\alpha,\beta}]$, then, in particular, $\overline{w}_{4^*,2}(z_{\alpha,\beta}) \leq \overline{w}_{\alpha,\beta}(z_{\alpha,\beta}) = 0$, and hence $z_{\alpha,\beta} \leq z_{4^*,2}$.

The proof of the other direction is somewhat more involved. We begin by defining

$$g(r) := \overline{w}_{4^*,2}(r) - \overline{w}_{\alpha,\beta}(r) = \frac{2r^{4^*} - 4^* r^2}{4^* - 2} - \frac{\beta r^{\alpha} - \alpha r^{\beta}}{\alpha - \beta}.$$

In particular, we divine the behaviour of g from that of its fifth derivative, which is given by

$$g^{(5)}(r) = -\frac{\alpha\beta r^{\beta-5}}{\alpha-\beta}(\alpha-1)(\alpha-2)(\alpha-3)(\alpha-4)r^{\alpha-\beta} + \frac{\alpha\beta r^{\beta-5}}{\alpha-\beta}(\beta-1)(\beta-2)(\beta-3)(\beta-4),$$

for $r \in (0, \infty)$. Written in this form, we see that $g^{(5)}(r)$ is monotone and hence has at most one sign change. Thus, $g'''(r)$ is either convex-concave, concave-convex, or strictly convex on $(0, \infty)$. Moreover, we may write

$$g'''(r) = 2 \cdot 4^*(4^*-1)r^{4^*-3} + \frac{\alpha\beta}{\alpha-\beta}\left[-(\alpha-1)(\alpha-2)r^{\alpha-3} + (\beta-1)(\beta-2)r^{\beta-3}\right].$$

Here, both the highest order term, corresponding to r^{4^*-3}, and the lowest order term, corresponding to $r^{\beta-3}$, have positive coefficients, which implies that g''' is positive outside a compact subinterval of $(0, \infty)$. This, combined with the convex/concave structure of g''', implies that g''' can have at most two zeroes on $(0, \infty)$ and, in particular, may change signs at most twice - from positive to negative to positive.

Thus, at worst, g' is convex-concave-convex on $(0, \infty)$. That is, we will assume g' is convex-concave-convex as, if it is simply convex, an easier argument than what follows will yield the desired conclusion. Notice that

$$g'(r) = \frac{4^* \cdot 2}{4^* - 2}(r^{4^*-1} - r) - \frac{\alpha\beta}{\alpha - \beta}(r^{\alpha-1} - r^{\beta-1})$$

is negative near zero and hence, the fact that it is convex-concave-convex implies that it changes sign at most thrice on $(0, \infty)$. It is clear that $g'(1) = 0$. Moreover, since $g(0) = g(1) = 0$, and since g' is negative near zero, we see that g' necessarily changes from negative to positive somewhere on the interval $(0, 1)$, implying the existence of a zero of g' on this interval. This leaves only one (potential) zero of g' unaccounted for, and hence g' has at most one zero on $(1, \infty)$. Now notice that there exists some $\varepsilon > 0$ such that either g' is positive on $(1, 1 + \varepsilon)$ or g' is negative on $(1, 1 + \varepsilon)$. In particular, if g' is positive on some interval of the form $(1, 1 + \varepsilon)$, it must remain non-negative on the entirety of $(1, \infty)$. In this case, since $g(1) = 0$, we know that $g(r)$ is positive on the entire interval $(0, \infty)$. Hence, we may rule out this case as it only occurs if $z_{\alpha,\beta} > z_{4^*,2}$.

Thus, we know that, if $z_{\alpha,\beta} \leq z_{4^*,2}$, there exists some $c > 0$ such that $g' < 0$ on $(1, 1 + c)$ and $g' > 0$ on $(1 + c, \infty)$. Since $g(1) = 0$, and since the leading term of g has positive coefficient, this implies that g has a unique zero on $(1, \infty)$. Of course, since $z_{\alpha,\beta} \leq z_{4^*,2}$, we know that $g(z_{\alpha,\beta}) \leq 0$, as $w^1_{4^*,2}(z_{\alpha.\beta}) \leq 0$ and $w^1_{\alpha,\beta}(z_{\alpha.\beta}) = 0$. Thus, since there is only one zero of g on $(1, \infty)$, and as g is positive for sufficiently large r, our assumption that $z_{\alpha,\beta} \leq z_{4^*,2}$ allows us to conclude that $g \leq 0$ on $[1, z_{\alpha,\beta}]$.

All that remains is to show that, in this case, $g \leq 0$ on $[0, 1]$. By way of contradiction, assume that g is positive at some point on $p \in (0, 1)$. then g' would have to change signs (at least) twice on the interval $(0, 1)$ — from negative to positive to negative again. Combining this with the fact that $g'(0) = 0$ and the inferred zero of g' on $(1, \infty)$ would yield a total of at least four zeroes of g' on $(0, \infty)$, whereas our earlier argument implies that g' can have at most three zeroes on this interval, thus concluding the proof. □

Corollary 2.15. *For any $\beta \geq 2$, $\alpha_\infty(\beta) \geq \alpha_{\Delta^n}(\beta)$.*

Proof. By [13, Theorem 2.3], there exists (at least one) minimizer of $\mathcal{E}_{W_{\alpha_\infty(\beta),\beta}}[\cdot]$ in $\mathcal{P}_0(\mathbb{R}^n)$. Moreover, by [27, Lemma 1], any such minimizer has support of diameter at most $z_{\alpha_\infty(\beta),\beta}$. In light of Remark 2.13, we know that $\alpha_\infty(\beta)$

is the largest solution of $z_{\alpha,\beta} = z_{4^*,2}$, and hence, by Proposition 2.14, $\overline{w}_{4^*,2}(r) \leq \overline{w}_{\alpha_\infty(\beta),\beta}(r)$ for all $r \in [0, z_{\alpha_\infty(\beta),\beta}]$.

Integrating, we find that, for any candidate minimizer μ of $\mathcal{E}_{W_{\alpha_\infty(\beta),\beta}}$,

$$\frac{2\cdot 4^*}{4^*-2}\mathcal{E}_{W_{4^*,2}}(\mu) \leq \frac{\beta\alpha_\infty(\beta)}{\alpha_\infty(\beta)-\beta}\mathcal{E}_{W_{\alpha_\infty(\beta),\beta}}(\mu). \tag{2.9}$$

On the other hand, if ν is the uniform distribution over the vertices of a unit simplex, then $|x-y| = 0$ or 1 for any pairs of points $x, y \in \operatorname{spt}\nu$. Thus, since $\overline{w}_{4^*,2}(0) = \overline{w}_{\alpha_\infty(\beta),\beta}(0) = 0$ and $\overline{w}_{4^*,2}(1) = \overline{w}_{\alpha_\infty(\beta),\beta}(1) = -1$, we have that ν saturates inequality (2.9). Moreover, since Theorems 2.1(2) and 2.4 imply that $\mathcal{E}_{W_{4^*,2}}(\nu) \leq \mathcal{E}_{W_{4^*,2}}(\mu)$, we find that

$$\begin{aligned}\frac{\beta\alpha_\infty(\beta)}{\alpha_\infty(\beta)-\beta}\mathcal{E}_{W_{\alpha_\infty(\beta),\beta}}(\nu) &= \frac{2\cdot 4^*}{4^*-2}\mathcal{E}_{W_{4^*,2}}(\nu)\\ &\leq \frac{2\cdot 4^*}{4^*-2}\mathcal{E}_{W_{4^*,2}}(\mu)\\ &\leq \frac{\beta\alpha_\infty(\beta)}{\alpha_\infty(\beta)-\beta}\mathcal{E}_{W_{\alpha_\infty(\beta),\beta}}(\mu),\end{aligned}$$

and hence $\mathcal{E}_{W_{\alpha_\infty(\beta),\beta}}(\nu) \leq \mathcal{E}_{W_{\alpha_\infty(\beta),\beta}}(\mu)$ for any candidate minimizer μ. Thus, we may conclude that the simplex ν is a global minimizer of $\mathcal{E}_{W_{\alpha_\infty(\beta),\beta}}$ in $\mathcal{P}_0(\mathbb{R}^n)$, and hence that $\alpha_\infty(\beta) \geq \alpha_{\Delta^n}(\beta)$, as desired. □

Lower Bounds for the Threshold Function

While $\underline{\alpha}^+_{\Delta^n}(\beta)$, as defined in [15], is the strongest lower bound for $\alpha_{\Delta^n}(\beta)$ apparent to us, it does not lend itself particularly well to computation. As such, we will provide a somewhat weaker lower bound for α_{Δ^n}, which can be easily computed directly. Before doing so, it will be useful to introduce the following family of functions and prove their unimodality:

Definition 2.16. Define $f_n : (0,\infty) \to \mathbb{R}$ by

$$f_n(t) := \begin{cases} (2^{-1} - 2^{-t})/t & \text{if } n = 1 \\ (n - (\frac{2n}{n+1})^{t/2} - n(\frac{n-1}{n+1})^{t/2})/t & \text{if } n \geq 2. \end{cases}$$

Through this family of functions, we are able to properly define a new family of lower bounds:

Definition 2.17. For $\beta \geq 2$, define $\underline{\alpha}_{\Delta^n}(\beta)$ by

$$\underline{\alpha}_{\Delta^n}(\beta) = \max\{\alpha \geq 2 \mid f_n(\alpha) = f_n(\beta)\}.$$

In particular, the set over which we take the maximum in the previous definition has at most two elements, as the following lemma shows:

Lemma 2.18 (Unimodality of f_n). *For any $n \geq 1$, the function f_n is unimodal. By this, it is meant that there exists a unique global maximum $\underline{\beta}_n := \operatorname{argmax}_{t>0} f_n$, and moreover, $f_n'(t) > 0$ for $t \in (0, \underline{\beta}_n)$ and $f_n'(t) < 0$ for $t \in (\underline{\beta}_n, \infty)$.*

Proof. We first treat the case $n = 1$ separately. Here, notice that $t^2 f_1'(t) = (t \log 2 + 1)2^{-t} - 2^{-1}$, and hence $f_1'(t)$ has the same sign as $g_1(t) := (t \log 2 + 1)2^{-t} - 2^{-1}$. Since $g_1'(t) = -t2^{-t} \log^2 2$ is always negative, and since $g_1(0) = \frac{1}{2}$ and $\lim_{t\to\infty} g_1(t) = -\frac{1}{2}$, we conclude that f_1' switches sign from positive to negative at its unique zero in $(0, \infty)$, and has no other sign changes. As such, we denote the unique zero of f_1' by $\underline{\beta}_1$.

The $n \geq 2$ case proceeds in a similar manner. Here, we notice that

$$g_n(t) := t^2 f_n'(t) = -\frac{t}{2}\left(\left(\frac{2n}{n+1}\right)^{t/2} \log \frac{2n}{n+1} + n\left(\frac{n-1}{n+1}\right)^{t/2} \log \frac{n-1}{n+1}\right) - n + \left(\frac{2n}{n+1}\right)^{t/2} + n\left(\frac{n-1}{n+1}\right)^{t/2},$$

and compute

$$g_n'(t) = -\frac{t}{4}\left[\left(\frac{2n}{n+1}\right)^{t/2} \log^2 \frac{2n}{n+1} + n\left(\frac{n-1}{n+1}\right)^{t/2} \log^2 \frac{n-1}{n+1}\right].$$

Since $g_n'(t)$ is negative everywhere, $g_n(0) = 1$, and $\lim_{t\to\infty} g_n(t) = -\infty$, we may apply an identical argument to the one employed in the $n = 1$ case to show the existence of $\underline{\beta}_n$ with all desired properties. □

Remark 2.19. Notice that $\underline{\alpha}_{\Delta^n}(\beta) > \beta$ if and only if $\beta < \underline{\beta}_n$. In other words, the graph of $\underline{\alpha}_n$ intersects the line $\alpha = \beta$ at the point $(\underline{\beta}_n, \underline{\beta}_n)$.

Proposition 2.20 (Comparison of bounds). *For any $n \geq 1$, and $\beta \geq 2$. $\underline{\alpha}_{\Delta^n}(\beta) \leq \alpha_{\Delta^n}(\beta)$.*

Proof. We proceed by deriving the defining equations for $\underline{\alpha}_{\Delta^n}$ directly from the Euler-Lagrange equation (1.5) for a unit simplex $\nu \in \mathcal{P}(\mathbb{R}^n)$. As in the introduction, we denote the vertices of the unit $n-$simplex by $\{x_0, ..., x_n\}$. We will prove this result in two cases, $n = 1$ and $n \geq 2$. Notice that, in either case, the inequality is trivial for any β for which $\underline{\alpha}_{\Delta^n}(\beta) = \beta$, so we are free to assume that $\underline{\alpha}_{\Delta^n}(\beta) > \beta$.

If $n = 1$, notice that the Euler-Lagrange equation (1.5) ensures that

$$(W_{\alpha,\beta} * \nu)(x_0) \le (W_{\alpha,\beta} * \nu)(0).$$

More explicitly, since $\nu = \frac{\delta_{x_0}+\delta_{x_1}}{2}$, this inequality reads:

$$\frac{1}{2}\left[\frac{1}{\alpha} - \frac{1}{\beta}\right] \le \frac{1}{\alpha 2^{\alpha}} - \frac{1}{\beta 2^{\beta}},$$

or, after some minor rearranging,

$$f_1(\alpha) = \frac{2^{-1} - 2^{-\alpha}}{\alpha} \le \frac{2^{-1} - 2^{-\beta}}{\beta} = f_1(\beta).$$

By definition, $\alpha = \underline{\alpha}_{\Delta^1}(\beta)$ saturates this inequality. Thus, assuming that $\underline{\alpha}_{\Delta^1}(\beta) > \beta$, unimodality of f_1, as in Lemma 2.18, ensures that for any $\gamma \in (\beta, \underline{\alpha}_{\Delta^1}(\beta))$,

$$f_1(\gamma) > f_1(\beta) = f_1(\underline{\alpha}_{\Delta^1}(\beta)).$$

This implies that the simplex ν violates the Euler-Lagrange equation for $W_{\beta,\gamma}$, and hence that $\alpha_{\Delta^1}(\beta) \ge \gamma$. Of course, since this inequality holds for all $\gamma \in (\beta, \underline{\alpha}_{\Delta^1}(\beta))$, this proves that $\underline{\alpha}_{\Delta^1}(\beta) \le \alpha_{\Delta^1}(\beta)$ for any $\beta \ge 2$.

Our proof proceeds analogously for $n \ge 2$, with the key difference being that the formula for f_n is derived from the inequality

$$(W_{\alpha,\beta} * \psi)(x_0) \le (W_{\alpha,\beta} * \psi)(-x_0),$$

which again is a necessary condition for the Euler-Lagrange equation to hold for ν. Exploiting the geometry of the simplex, which ensures that $|x_0| = \left(\frac{n}{2n+2}\right)^{1/2}$ and $|x_0 + x_1| = \left(\frac{n-1}{n+1}\right)^{1/2}$, this equation can be expressed as:

$$\frac{n}{n+1}\left(\frac{1}{\alpha} - \frac{1}{\beta}\right) \le \frac{1}{n+1}\left(\frac{\left(\frac{2n}{n+1}\right)^{\alpha/2}}{\alpha} - \frac{\left(\frac{2n}{n+1}\right)^{\beta/2}}{\beta}\right) + \frac{n}{n+1}\left(\frac{\left(\frac{n-1}{n+1}\right)^{\alpha/2}}{\alpha} - \frac{\left(\frac{n-1}{n+1}\right)^{\beta/2}}{\beta}\right),$$

or, after multiplying by $n+1$ and rearranging,

$$f_n(\alpha) = \frac{n - \left(\frac{2n}{n+1}\right)^{\alpha/2} - n\left(\frac{n-1}{n+1}\right)^{\alpha/2}}{\alpha} \le \frac{n - \left(\frac{2n}{n+1}\right)^{\beta/2} - n\left(\frac{n-1}{n+1}\right)^{\beta/2}}{\beta} = f_n(\beta).$$

Since Lemma 2.18 ensures that f_n is still unimodal for $n \geq 2$, the proof proceeds in an identical manner to the one-dimensional proof, and hence we omit it. □

We summarize our findings in the $n = 2$ case with the following graph, which is an analogue of one of the figures of [15]

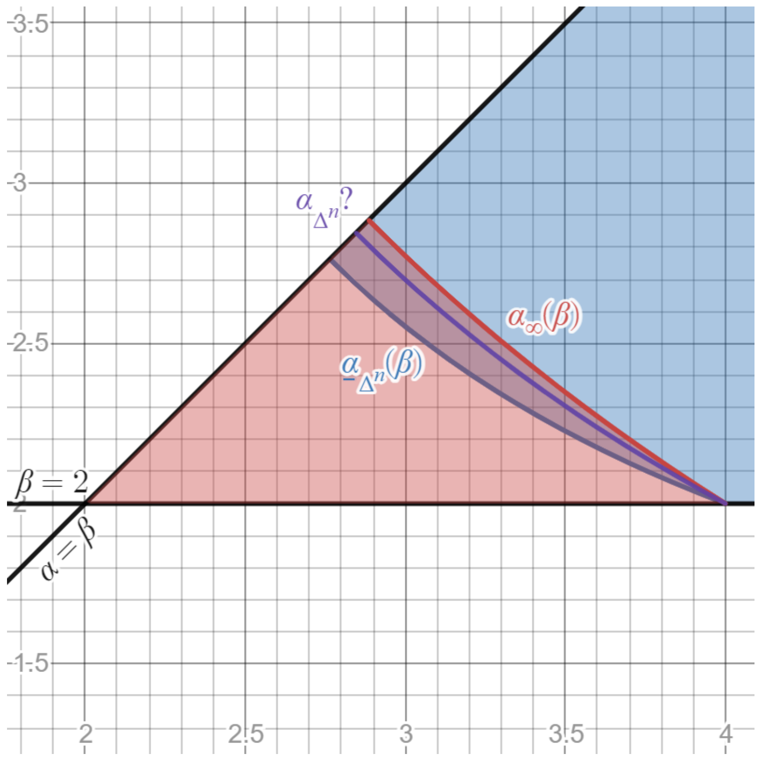


Figure 2.3: The mildly repulsive regime for $n = 2$. In the red region to the left of the blue curve $\alpha = \underline{\alpha}_{\Delta^2}(\beta)$, the simplex does not minimize $\mathcal{E}_{W_{\alpha,\beta}}$. Conversely, in the rightmost blue region, the simplex uniquely minimizes $\mathcal{E}_{W_{\alpha,\beta}}$. In the intermediate region, it is not entirely known where the simplex minimizes $\mathcal{E}_{W_{\alpha,\beta}}$ and, in fact, the graph of the threshold function α_{Δ^2} must lie entirely in this region.

Relationship Between Upper and Lower Bounds

Interestingly enough, even this weak lower bound tends to the upper bound α_∞ as $n \to \infty$.

Proposition 2.21 (Dimensional asymptotics)**.** *As $n \to \infty$, $\underline{\alpha}_{\Delta^n}(\beta) \to \alpha_\infty(\beta)$ for all $\beta \in [2, 4]$.*

Proof. Recall that the defining equation of $\underline{\alpha}_{\Delta^n}(\beta)$ is given by

$$\frac{n - \left(\frac{2n}{n+1}\right)^{\alpha/2} - n\left(\frac{n-1}{n+1}\right)^{\alpha/2}}{\alpha} = \frac{n - \left(\frac{2n}{n+1}\right)^{\beta/2} - n\left(\frac{n-1}{n+1}\right)^{\beta/2}}{\beta}.$$

Taking the limit as $n \to \infty$ of this equation yields:

$$\frac{2^{\alpha/2}}{\alpha} - 1 = \frac{2^{\beta/2}}{\beta} - 1,$$

or rather that $w_{\alpha,\beta}(\sqrt{2}) = 0$, which is precisely the equation which appears in Remark 2.13. □

Remark 2.22 (Monotonicity). Our numerical experiments indicate that $\underline{\alpha}_{\Delta^n}(\beta) \nearrow \alpha_\infty(\beta)$ for all $\beta \in [2, 4]$ and that $\underline{\alpha}_{\Delta^n}(\beta) \searrow \alpha_\infty(\beta)$ for $\beta \notin (2, 4)$. To prove this, it would suffice to show that, for any $n \geq 2$, the difference $f_{n+1} - f_n$ is a unimodal function on $(0, \infty)$. This is because, for $n \geq 2$, $f_n(t)$ has zeroes only at $t = 2$ and $t = 4$, and hence, assuming unimodality, these are the only two zeroes of $f_{n+1} - f_n$. Since $\lim_{t\to\infty} (f_{n+1}(t) - f_n(t)) = -\infty$, this implies that $f_{n+1} > f_n$ on $(2, 4)$, $f_{n+1} = f_n$ on $\{2, 4\}$, and $f_{n+1} < f_n$ outside of $[2, 4]$.

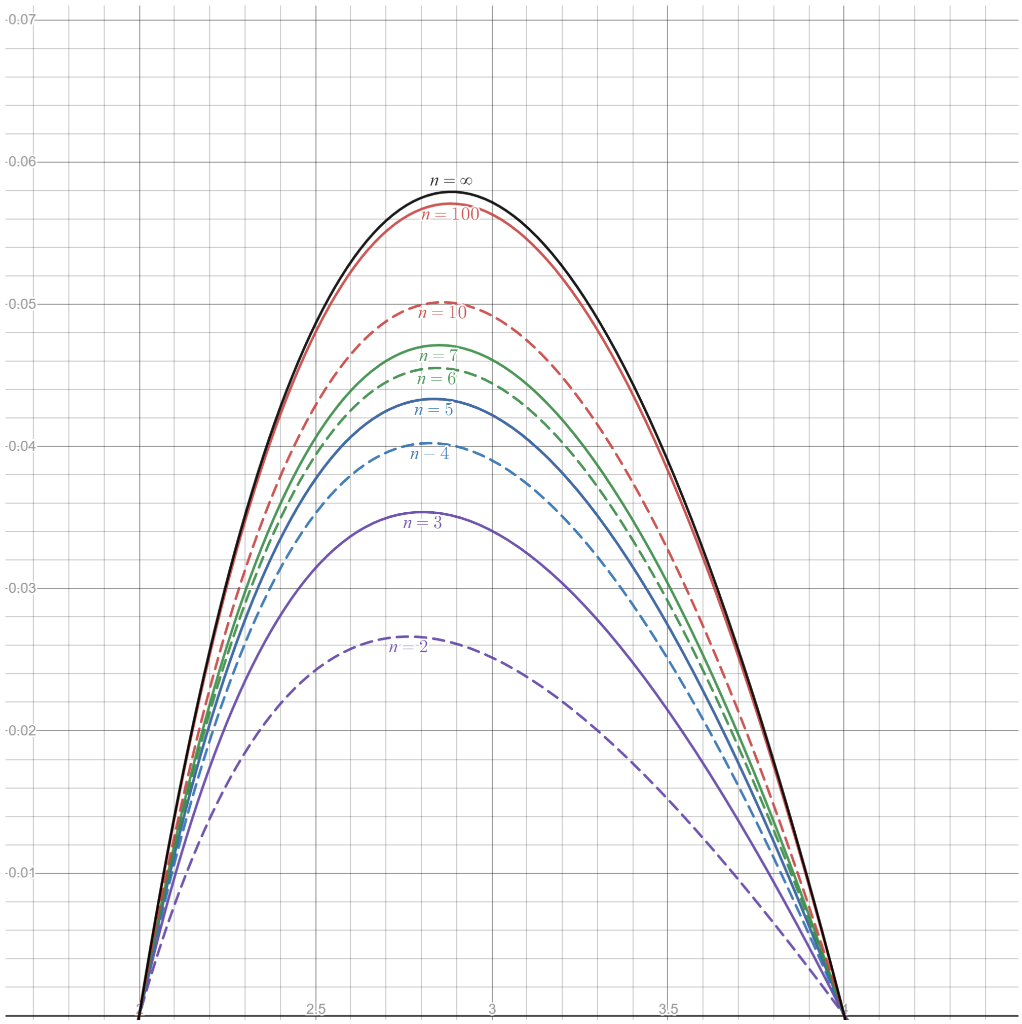


Figure 2.4: Graphs of $f_n(t)$ for selected values of n. Our numerical experiments indicate that, for all $t \in [2,4]$, $f_n(t)$ increases monotonically to $f_\infty(t) := 1 - \frac{2^{t/2}}{t}$.

2.3 DISCUSSION: INTUITION FOR LOWER BOUNDS

In this section of the present work, we study the values of α and β for which the unit simplices in one and two dimensions violate the Euler-Lagrange equations. In particular, we detail partial progress made toward simplifying the checking of such violations. Much of what follows will be mere speculation about where such violations are most likely to occur but, nevertheless, may be useful in practice - as any violation of the Euler-Lagrange equation at (α_0, β_0) provides a lower bound for the threshold function $\alpha_{\Delta^n}(\beta_0)$, and better guesses of course give better bounds.

The One-Dimensional Case

In the proof of Proposition 2.20, we derived the lower bound by checking whether or not the Euler-Lagrange equation for the simplex is violated at the origin. Our rationale for this stems from the work of Kang, Kim, Lim, and Seo in [27]. In this work, the quartet showed that for $(\alpha, \beta) = (2.5, 2.1)$ the simplex $\nu = \frac{1}{2}(\delta_0 + \delta_1)$ does not minimize $\mathcal{E}_{W_{\alpha,\beta}}$ over $\mathcal{P}(\mathbb{R})$, by comparing ν directly to the measure $\nu' = 0.420137\delta_0 + 0.159726\delta_{0.548674} + 0.420137\delta_{1.09735}$. In addition, in unpublished notes, I showed that, if $\beta = 2$ and $\alpha \in (2, 3)$, then at least one measure of the form

$$M\delta_0 + \frac{1-M}{2}\delta_r + \frac{1-M}{2}\delta_{-r},$$

where $M > 0$, has strictly lower interaction energy than the simplex ν. As such, it seems that moving even a small amount of mass to the origin can often decrease the energy of a particle configuration and. This means that, in the one-dimensional case, it makes sense to start by checking for Euler-Lagrange violations at the origin.

The Two-Dimensional Case

The lower bounds $\underline{\alpha}_{\Delta^n}$ for $n \geq 2$ were derived primarily using intuition and symmetry, motivated by numerics, without much in the way of rigorous justification. Nevertheless, I provide some intuition in this subsection, as well as some open conjectures. Throughout this subsection, we will assume, for the sake of simplifying notation, that any simplices which appear are canonical in the sense of Definition 1.10. We begin with a conjecture:

Conjecture 2.23. *Let $n \geq 2$, ν_n be the canonical unit n-simplex, and $\alpha > \beta \geq 2$. Then, if there exists some point in $x \in \mathbb{R}^n \setminus \operatorname{spt} \nu_n$ such that the Euler-Lagrange equation (1.5) is violated (where $W = W_{\alpha,\beta}$), there also exists some $r > 0$ such that (1.5) is violated at $-re_1$.*

Remark 2.24. As in the proof of Proposition 2.20, we suspect that it will be useful to express the fact that x violates the Euler-Lagrange equation by the inequality

$$(W_{\alpha,\beta} * \nu_n)(x) < (W_{\alpha,\beta} * \nu_n)(x_0).$$

Assuming this conjecture holds, checking if the Euler-Lagrange equation is violated reduces to a single variable optimization problem, i.e.

Corollary 2.25. *Assume Conjecture 2.23 holds. Then ν_n satisfies the Euler-Lagrange equation if and only if:*

$$2\mathcal{E}_{W_{4,2}}[\nu_n] = (W_{\alpha,\beta} * \nu_n)(x_0) \leq \inf_{r>0}(W_{\alpha,\beta} * \nu_n)(-re_1). \tag{2.10}$$

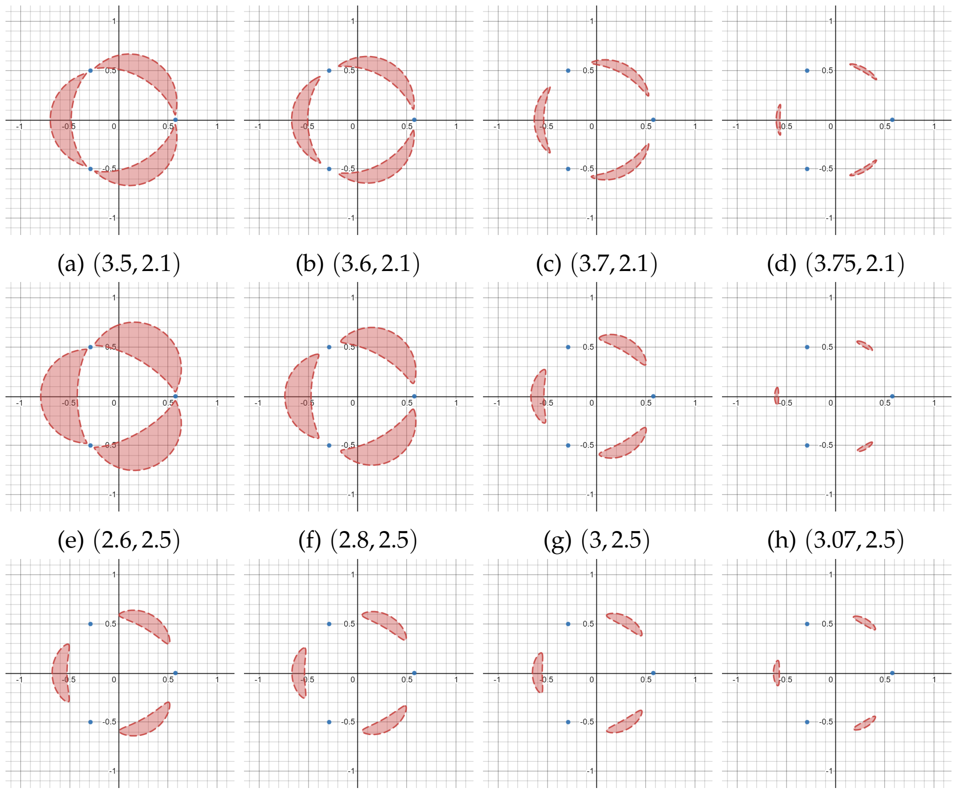
Proof. Clearly, (2.10) is a necessary condition for the Euler-Lagrange equation to hold, as it can be derived directly from (1.5).

By the conjecture, the inequality in (2.10) is enough to guarantee that $(W_{\alpha,\beta} * \nu_n)(x_0) \leq (W_{\alpha,\beta} * \nu_n)(x)$ for all $x \in \mathbb{R}^n \setminus \operatorname{spt} \nu_n$, so all that remains is to show that $(W_{\alpha,\beta} * \nu_n)(x) \equiv 2\mathcal{E}_{W_{4,2}}[\nu_n]$ on $\operatorname{spt} \nu_n$. This follows from symmetry, as $|x_i - x_j| = 1$ for all $i \neq j$, and hence

$$\begin{aligned}(W_{\alpha,\beta} * \nu_n)(x_i) &= \frac{1}{n+1}\sum_{j=0}^{n} W_{\alpha,\beta}(x_i - x_j) \\ &= \frac{1}{n+1}\sum_{j=0}^{n} w_{\alpha,\beta}(|x_i - x_j|) \\ &= \frac{1}{n+1} w_{\alpha,\beta}(0) + \frac{n}{n+1} w_{\alpha,\beta}(1).\end{aligned} \tag{2.11}$$

This implies that $(W_{\alpha,\beta} * \nu_n)(x_i) = (W_{\alpha,\beta} * \nu_n)(x_j)$ for any i, j. Hence, we can average (2.11) over i to find that $(W_{\alpha,\beta} * \nu_n)(x_i) = 2\mathcal{E}_{W_{4,2}}[\nu_n]$, as desired. □

Now that we have explained the motivation for the conjecture, we discuss some numerical experiments which lend support to its veracity. First, for every pair of (α, β) tested, the region on which the Euler-Lagrange equation is violated tends to look something like the union of three lens-shaped regions opposite each vertex of the simplex, provided that $\beta \neq 2$ and α is sufficiently larger than β, as illustrated in the figure below:



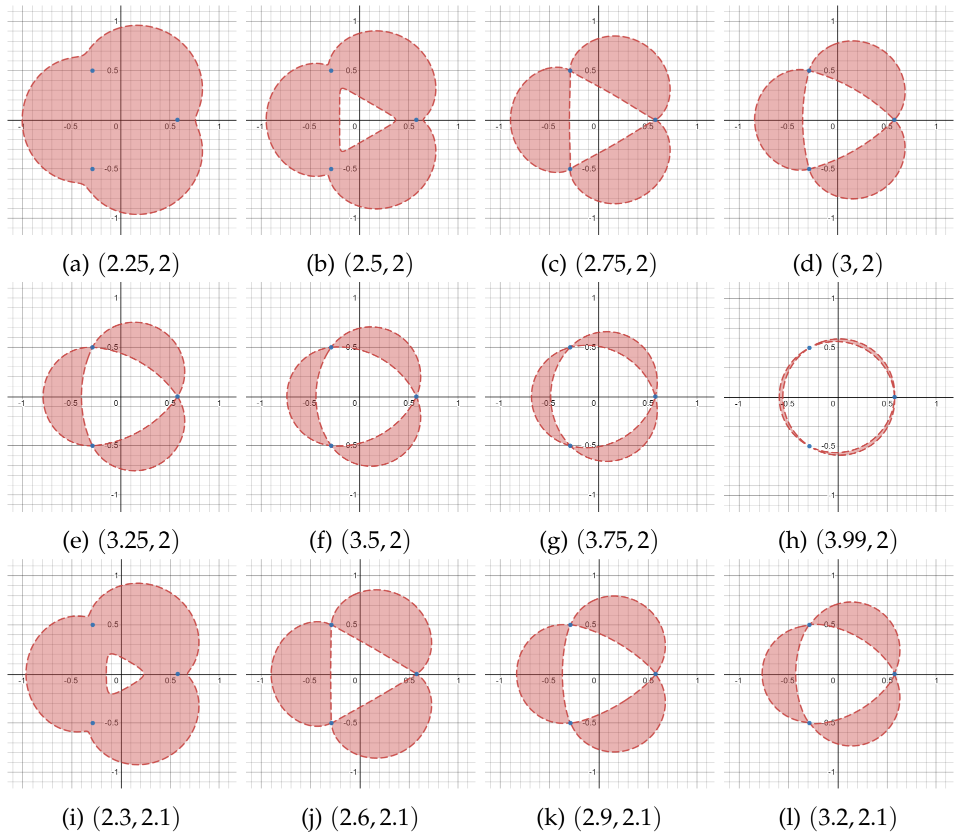
(a) $(3.5, 2.1)$ (b) $(3.6, 2.1)$ (c) $(3.7, 2.1)$ (d) $(3.75, 2.1)$

(e) $(2.6, 2.5)$ (f) $(2.8, 2.5)$ (g) $(3, 2.5)$ (h) $(3.07, 2.5)$

(i) $(2.75, 2.7)$ (j) $(2.775, 2.7)$ (k) $(2.8, 2.7)$ (l) $(2.825, 2.7)$

Figure 2.5: The shaded region of $\mathbb{R}^2$ on which the Euler-Lagrange equation is violated, plotted for selected pairs (α, β) where this region is clearly partitioned into three disjoint subregions.

In the case where $\beta = 2$, or in the case where α and β are close enough, these regions often fuse into a single region, as indicated in the following figure:



(a) $(2.25, 2)$ (b) $(2.5, 2)$ (c) $(2.75, 2)$ (d) $(3, 2)$

(e) $(3.25, 2)$ (f) $(3.5, 2)$ (g) $(3.75, 2)$ (h) $(3.99, 2)$

(i) $(2.3, 2.1)$ (j) $(2.6, 2.1)$ (k) $(2.9, 2.1)$ (l) $(3.2, 2.1)$

Figure 2.6: The shaded region of $\mathbb{R}^2$ on which the Euler-Lagrange equation is violated, for selected pairs (α, β), where $\beta = 2$ or 2.1. This illustrates that the region on which the Euler-Lagrange equation is violated is sometimes connected.

Of course, in both Figures 2.5 and 2.6, the Euler-Lagrange equation is violated at some point on the negative x-axis, lending credence to Conjecture 2.23. Moreover, symmetry alone does not guarantee that any Euler-Lagrange violations must happen (at least partially) on the negative $x-$axis. For example, one could imagine that the Euler-Lagrange equation is violated on a region composed of six lobes, one on either side of each vertex of the simplex.

Additionally, it is important to notice that Conjecture 2.23 **does not** require that the global minimum of $W_{\alpha,\beta} * \nu$ lie on the negative $x-$axis. It is entirely possible that this global minimum could be attained at six (or more) separate points, two of which straddle the negative $x-$axis, and where the other four are rotations of the first two by $\frac{2\pi}{3}$ and $\frac{4\pi}{3}$ about the origin.

Finally, it should be noted that, for generic pairs of (α, β), $-x_0$ is **not** the best point against which to check for Euler-Lagrange violations and, in fact, the plot in the following figure indicates that it is entirely possible for the Euler-Lagrange equation to be satisfied at $-x_0$, but violated elsewhere on the negative real-axis:

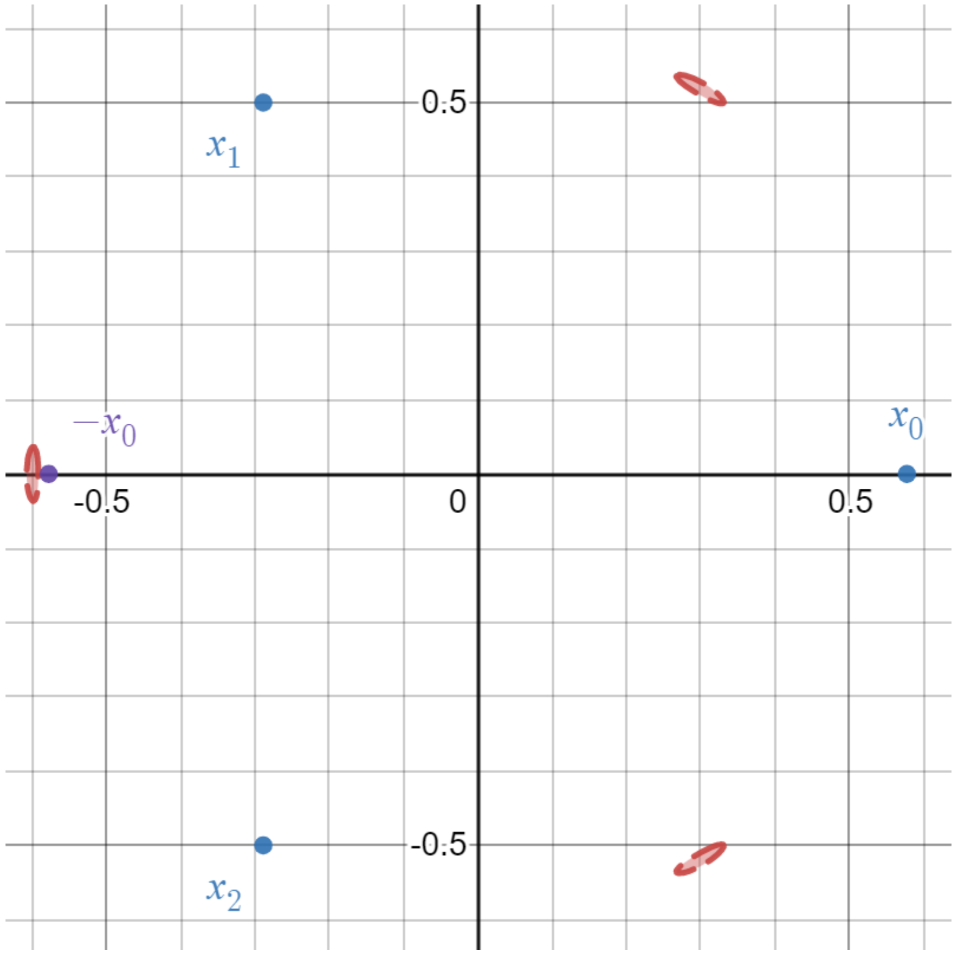


Figure 2.7: The red region on which the simplex ν_2 violates the Euler-Lagrange equation for $\mathcal{E}_{W_{\alpha,\beta}}$, where $(\alpha, \beta) = (3.076, 2.5)$. Notice that, according to this plot, the purple point $-x_0$ is not contained in such a region, and hence the Euler-Lagrange equation is satisfied at that point.

3 A TASTE OF THE DYNAMICS

In this section, we begin by explaining the distinction between two important types of stabiilty — Lyapunov and asymptotic, inspired by the discussion in [14, Section 5]. Then we use this framework to contextualize and elaborate on Simione's treatment of asymptotic stability in [39], explaining in particular his distinction between spreading stability and displacement stability as they apply to the aggregation equation. After this, given that our previous arguments show that the simplex uniquely minimizes $\mathcal{E}_{W_{\alpha,\beta}}$ on most of the mildly repulsive regime, we examine what Simione's arguments have to say about measures which uniformly distribute their mass over the vertices of a unit $n-$simplex. In one dimension, Simione's spreading stability condition is strictly more restrictive than his displacement stability condition and, as such, governs whether or not the unit simplex is asymptotically stable. In higher dimensions, we can hope that such a result holds, especially given that it tends to be less computationally demanding to quantify the spreading stability of a measure.

3.1 TYPES OF STABILITY

We now explain the distinction between two key types of stability for steady states of general PDEs, and in particular steady states of the aggregation equation (1.2), in a discussion heavily influenced by that in [14, Section 5]. The first notion of stability is asymptotic stability which, heuristically, means that configurations which start out sufficiently 'close' to a given steady state will, in either finite or infinite time, become arbitrarily close to the steady state. More precisely,

Definition 3.1 (Asymptotic Stability for the Aggregation Equation)**.** Let d be a metric on (some subset of) $\mathcal{P}(\mathbb{R}^n)$, and let $\overline{\mu}$ be a steady state of the aggregation equation (1.2). We say that $\overline{\mu}$ is asymptotically stable with respect to d if there exists some $\delta > 0$ such that, for any $\varepsilon > 0$ and any probability curve $\mu(t)$ with $d(\mu(0), \overline{\mu}) < \delta$, there exists some $T > 0$ such that $d(\mu(t), \overline{\mu}) < \varepsilon$ for any $t > T$.

When working with an asymptotically stable steady state, it is often possible to quantify the rate at which nearby probability curves converge to the steady state — for example, in the domain of application of Simione's arguments in the next subsection, nearby probability curves converge to the steady state at an exponential rate, and in infinite time. Nevertheless, not every steady state — and indeed, not every global minimizer of (1.3) — is asymptotically stable. For example, this failure occurs when there are d-local minimizers of $\mathcal{E}_{W_{\alpha,\beta}}$ arbitrarily close to the steady state $\overline{\mu}$. For example, in the context of Theorem 2.1(1), it is known that discrete particle rings are steady states of the aggregation equation [14] and, as the number of particles tends to ∞, such particle ring configurations get arbitrarily close to the global minimizer $\overline{\sigma}$ in the d_p metric, for any $p \in [1, \infty]$.

On the other hand, we have a strong theoretical incentive to be able to say that such steady states are stable in some sense. This leads us to the notion of Lyapunov stability, which is defined as follows for the aggregation equation:

Definition 3.2 (Lyapunov Stability for the Aggregation Equation)**.** Let d be a metric on (some subset of) $\mathcal{P}(\mathbb{R}^n)$ and let $\overline{\mu}$ be a steady state of the aggregation equation. We say that $\overline{\mu}$ is Lyapunov stable if, for any $\varepsilon > 0$, there exists some $\delta > 0$ such that, if $\mu(t)$ is a probability curve with $d(\mu(0), \overline{\mu}) < \delta$, then $d(\mu(t), \overline{\mu}) < \varepsilon$ for all $t \geq 0$.

The notion of Lyapunov stability is useful in the context of the aggregation equation since [14] showed that minimizers of (1.3) for $\alpha > \beta > 0$ and $\alpha \geq 1$ are d_α Lyapunov-stable.

A useful toy model to illustrate the difference between asymptotic and Lyapunov stability is the continuous function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) := \begin{cases} |x|(1 + \sin \frac{1}{|x|}) & x \neq 0 \\ 0 & x = 0 \end{cases}, \tag{3.1}$$

as illustrated in the graph below:

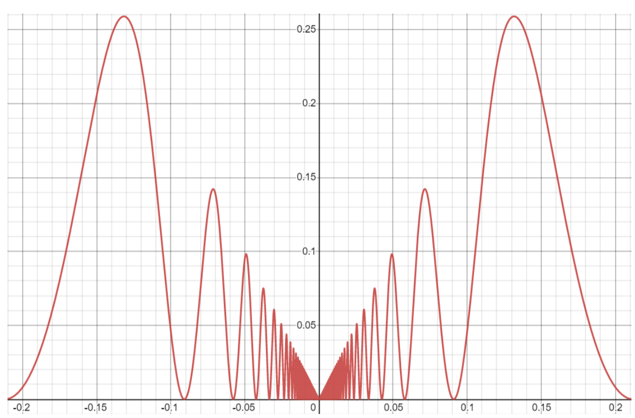


Figure 3.1: The graph of $f(x)$ as defined in Equation (3.1)

Clearly, f attains its minimum value of 0 precisely at 0 and points of the form $\pm\frac{1}{\pi(2k+3/2)}$, for integers $k \geq 0$. Now, imagine we place a ball on the graph of f, subject to the effects of gravity and friction. For each minimum of the form $\frac{1}{\pi(2k+3/2)}$, so long as we place the ball close enough to $x_k = \frac{1}{\pi(2k+3/2)}$, the ball will eventually come to rest exactly at the point x_k, meaning that, in some sense, x_k is asymptotically stable. On the other hand, no matter how close we place our ball to 0, there are always infinitely many points of the form x_k between its initial position and the origin, each of which comes with its own well. This means that, unless our ball is placed directly at $x_0 = 0$, it will come to rest at some point of the form x_k, rather than exactly at 0. Thus, the point 0 is Lyapunov stable, but not asymptotically stable, as there are many nearby local minima which are more attractive to the particle dynamics.

Remark 3.3 (Two Views on Stability)*.* Given that the interaction energy is both rotation and translation invariant, any rotated translates of the steady state $\overline{\mu}$ are also steady states of the interaction equation, which means that it is difficult to directly apply any arguments involving either type of stability. There are two principal methods of resolving this difficulty and allowing a stability analysis. First, we can restrict our focus from $\mathcal{P}(\mathbb{R}^n)$ to a subset of $\mathcal{P}_0(\mathbb{R}^n)$ where, in some sense, all probability measures have the same orientation as $\overline{\mu}$. However, this method has its flaws, as it is possible for a probability measure to rotate as it evolves under the aggregation equation. The second, more robust, method of treating this problem is to define a manifold $\overline{\mathcal{M}} \subset \mathcal{P}(\mathbb{R}^n)$ which consists of $\overline{\mu}$ and its rotated translates, and instead ask questions about the stability of this manifold. Thankfully, Definitions 3.1 and 3.2 can be generalized in a natural manner to describe the stability of manifolds. Additionally, we will outline Simione's treatment of this issue, which contains aspects of both approaches, in the next section.

3.2 Simione's arguments

We are now ready to discuss Simione's results on the asymptotic stability of minimizers of $\mathcal{E}_{W_{\alpha,\beta}}$, as formulated in his PhD Thesis [39]. In particular, we will outline the background material for [39, Theorem 25], apply this result to the one-dimensional case, discuss Simione's application to the two-dimensional case, and in general discuss some ideas on how to optimize and simplify Simione's arguments.

Basic Definitions and Preliminaries

We begin by providing some background, as well as some definitions which are necessary to understand Simione's work. We note that Simione works exclusively on $\mathcal{P}_2(\mathbb{R}^n)$, which we recall is the space of Borel probability measures with finite second moment. While this is at first glance a non-trivial restriction, we recall that from, e.g., [27] it is known that any global minimizer $\overline{\mu}$ of $\mathcal{E}_{W_{\alpha,2}}$ has compact support, and hence lies in $\mathcal{P}_\infty(\mathbb{R}^n) \subseteq \mathcal{P}_2(\mathbb{R}^n)$. Likewise, by the triangle inequality, if $\mu \in \mathcal{P}(\mathbb{R}^n)$ is such that $d_2(\mu, \overline{\mu}) < \infty$, then $\mu \in \mathcal{P}_2(\mathbb{R}^n)$, and hence Simione's restriction does not affect our work around global minimizers.

Thus, we begin in earnest by defining the tangent plane to a given probability measure in $\mathcal{P}_2(\mathbb{R}^n)$:

Definition 3.4 (Tangent Plane at $\mu \in \mathcal{P}_2(\mathbb{R}^n)$)**.** We define the tangent plane $\mathcal{T}_\mu$ to $\mu \in \mathcal{P}_2(\mathbb{R}^n)$ as the $L^2(d\mu)$ closure of the set of vector fields which can be realized as gradients of functions in $\mathcal{C}_c^\infty$.

In other words, a typical element of $\mathcal{T}_\mu$ is a vector field v on $\mathbb{R}^d$ such that there exists a sequence of $\mathcal{C}_c^\infty$ functions $\{\varphi_n\}_{n=1}^\infty$ with the property that $\left[\int_{\mathbb{R}^n} |\nabla\varphi - v|^2 d\mu\right]^{1/2} \to 0$. It should also be noted that Definition 3.4 is motivated by Otto's formal Riemannian structure on $\mathcal{P}_2(\mathbb{R}^n)$ [34, 43]. Heuristically, and with some caveats (see Definition 3.7 and the preceding discussion), we can think of such vector fields v as describing the evolution of the probability measure μ over an infinitesimal amount of time. That is, for each point $x \in \operatorname{spt}\mu$, we imagine that the mass μ assigns to x moves an infinitesimal amount in the direction of the vector $v(x)$. While this intuition is valuable, we will often only treat $\mathcal{T}_\mu$ as a useful space with which to formulate other definitions. One of the most important such definitions is that of the Hessian of the energy functional $\mathcal{E}_W$:

Definition 3.5 (Hessian of the Energy Functional)**.** Let $W : \mathbb{R}^n \to \mathbb{R}$ be a potential function of the type discussed in Section 1.2. Then we define the Hessian $\operatorname{Hess}\mathcal{E}_{W,\mu}$ at μ by its action on $\mathcal{T}_\mu$ as follows. We say that, for each $v \in \mathcal{T}_\mu$,

$$\begin{aligned}&\operatorname{Hess}\mathcal{E}_{W,\mu}[v,v]\\&\quad= \iint_{\mathbb{R}^n\times\mathbb{R}^n} (v(x)-v(y))\cdot \operatorname{Hess}W(x-y)(v(x)-v(y))d\mu(x)d\mu(y).\end{aligned}$$

Simione derives this definition using a variational approach in [39, Chapter 1.5]. In particular, this Hessian quantifies the convexity of $\mathcal{E}_W$ near the measure $\mu \in \mathcal{P}_2(\mathbb{R}^n)$ - for example, if $\overline{\mu}$ is a local minimizer of the aggregation equation with potential μ, then $\operatorname{Hess}\mathcal{E}_{W,\mu} \geq 0$.

However, matters are complicated due to the fact that, for many potentials of interest, including power-law potentials, $\mathcal{E}_W$ has a number of symmetries - for example, if $W_{\alpha,\beta}$ is a power-law potential, then $\mathcal{E}_{W_{\alpha,\beta}}[\cdot]$ is both rotation and translation invariant. Of course, for such vector fields, any rotation or translation of an energy minimizer will also be a minimizer, which significantly dilutes the notion of uniqueness which we must work with. In Chapter 2, we handled the issue of translation invariance by considering minimizers in $\mathcal{P}_0(\mathbb{R}^n)$, and largely skirted around rotation invariance by adding appropriate language to our assumptions and conclusions. Nevertheless, when treating the dynamics, it makes sense to handle rotations and translations in a somewhat different and more careful manner.

More precisely, a minimizer $\overline{\mu}$ of $\mathcal{E}_{W_{\alpha,\beta}}$ can only be unique up to translation and rotation. Hence, heuristically, if we consider the Hessian along a path consisting exclusively of rotated translates of $\overline{\mu}$, it will be identically zero, even though $\overline{\mu}$ is the unique minimizer of $\mathcal{E}_{W_{\alpha,\beta}}$ up to rotation and translation. As such, Simione defines a notion of *admissible vector fields* to account for this nonuniqueness.

Definition 3.6. We say that $v \in \mathcal{T}_\mu$ is an admissible vector field if

- $\int v(x)d\mu(x) = 0$ (orthogonality to translation).
- $\int v(x) \cdot (Ax)d\mu(x) = 0$ for any skew-symmetric matrix $A \in so(n)$ (orthogonality to rotation).

Using the heuristic intuition which we discussed before, the first condition means that, after each point x in the support of μ moves an infinitesimal amount in the direction $v(x)$, the resulting measure has the same centre of mass as the original measure μ. Likewise, the second condition ensures that the resulting measure is, in some sense, oriented in the same way as the original - i.e. it has not rotated.

Of course, as we mentioned in the discussion after Definition 3.4, we will need to introduce an additional complication to deal with the fact that we are working with probability measures. Namely, the space $\mathcal{P}_2(\mathbb{R}^n)$ allows for Dirac δ masses to split, or even spread out into non-singular measures. For example, for any point $x \in \mathbb{R}^n$, both δ_0 and $\frac{\delta_x+\delta_{-x}}{2}$ lie in $\mathcal{P}_0(\mathbb{R}^n)$. However, vector fields are limited in that they can only prescribe one direction for mass at a given point to travel, and do not, for example, allow mass which is originally located at the origin to end up at two distinct locations. Thus, in order for our theory to reflect our intuitive idea of tangent vectors as modelling the infinitesimal-in-time evolution of probability measures into other probability measures, we will need to make it somewhat more robust. To do so, Simione introduces the notion of the full tangent plane, which we provide a revised interpretation of as follows:

Definition 3.7 (Full Tangent Plane)**.** Let $\mu \in \mathcal{P}_2(\mathbb{R}^n)$. We define the full tangent plane, $\mathcal{FT}_\mu$ of μ by:

$$\mathcal{FT}_\mu := \{(v, \pi) \mid v \in \mathcal{T}_\mu,\ \pi \in \Gamma(\mu, \nu) \text{ for some } \nu \in \mathcal{P}_2(\mathbb{R}^n), \\ \text{and } v(x,s) : \mathbb{R}^n \times \operatorname{spt} \nu \to \mathbb{R}^n,\ v \in L^2(d\pi)\}.$$

In effect, the choice of the measure ν reflects the 'end goal' of μ, i.e. the measure which we wish for it to change into. This definition allows our (full) tangent vector field v to prescribe direction vectors to mass based not just on its starting location, but also based on where the mass eventually ends up. In particular, this allows for mass to split and spread, as we would expect in $\mathcal{P}_2(\mathbb{R}^n)$. This treatment of the full tangent plane is the last of the background material needed to understand Simione's thesis, and as such, we move on to his first key idea — the decomposition of Hess $\mathcal{E}_W$ into spreading and displacement components.

Spreading and Displacement Hessians

Simione's work from here on is concerned with the behaviour of probability curves $\mu(t)$ which start d_∞ close to a steady state $\overline{\mu}$ (which we will think of as a global minimizer of $\mathcal{E}_W[\cdot]$).

In particular, fix $\varepsilon > 0$ and assume that $d_\infty(\mu(0), \overline{\mu}) < \varepsilon$. Simione's key result in Chapter 4 of his thesis [39, Theorem 16] states that, under certain spreading and displacement assumptions on $\overline{\mu}$ which we will discuss in what follows, we may estimate

$$\operatorname{Hess} \mathcal{E}_{W,\mu}[v, v] \geq r \int |v|^2 d\mu,$$

where $r = r(\varepsilon, \overline{\mu}, W) > 0$. Of course, we will need to build up some additional background in order to rigorously discuss the appropriate spreading and displacement assumptions.

The first thing to note is that, since the measures we are working with are all found in close proximity to $\overline{\mu}$, it is natural to consider the subspace of $\mathcal{FT}_\mu$ given by

$$\{(v, \pi) \mid v \in \mathcal{T}_\mu,\ \pi \in \Gamma_{opt}(\mu, \overline{\mu}) \text{ and } v(x,s) : \mathbb{R}^n \times \operatorname{spt} \overline{\mu} \to \mathbb{R}^n,\ v \in L^2(d\pi)\},$$

i.e. where the splitting plan is an optimal transport plan between μ and $\overline{\mu}$.

In this context, Simione decomposes the vector field $v \in \mathcal{T}_\mu$ into two pieces, each of which lies in $\mathcal{FT}_\mu$. In order to do this, he first trivially extends v to an element of $\mathcal{FT}_\mu$, by setting $v(x, \overline{x}) = v(x)$ and, as above, letting π be a d_∞-optimal transport plan between μ and $\overline{\mu}$. In effect, this assigns the initial velocity $v(x)$ to all the mass which has the initial position of x, regardless of how that mass is split, or its ultimate location after transporting μ to $\overline{\mu}$ through the transport plan π.

Using this, Simione defines a vector field $\overline{v}$ on $\operatorname{spt}\overline{\mu}$ by

$$\overline{v}(\overline{x}) = \int_{\mathbb{R}^n} v(x, \overline{x}) d\pi_{\overline{x}}(x),$$

where the transition kernel $\pi_{\overline{x}}$ is defined by the equation $d\pi(x, \overline{x}) = d\pi_{\overline{x}}(x) d\overline{\mu}(x)$ (see, for example, [35, Chapter 1.4] for a general discussion on transition kernels). Heuristically, this assigns the point $\overline{x}$ an initial velocity which is a weighted average of the initial velocities of the particles of μ which get transported to the point $\overline{x}$ by the optimal transport plan π. In effect, $\overline{v}$ represents the displacement, i.e. the effects of the velocity field v which translate over to a velocity field $\overline{v}$ on $\operatorname{spt}\overline{\mu}$. Simione then defines the remainder, $\tilde{v}$ by

$$\tilde{v}(x, \overline{x}) = v(x) - \overline{v}(\overline{x})$$

so that we may decompose

$$v(x) = \overline{v}(\overline{x}) + \tilde{v}(x, \overline{x}). \tag{3.2}$$

We interpret $\tilde{v}(x, \overline{x})$ as the spreading tangent vector, which represents how mass must spread out when transporting μ to $\overline{\mu}$.

Using this decomposition, Simione [39, Section 4.3.1] decomposes $\operatorname{Hess}\mathcal{E}_{W,\mu}$ as

$$\begin{aligned}
&\operatorname{Hess}\mathcal{E}_{W,\mu}[v,v] \\
&= \iint_{\mathbb{R}^n\times\mathbb{R}^n} (v(x) - v(y)) \cdot \operatorname{Hess} W(x-y)(v(x) - v(y)) d\mu(x) d\mu(y) \\
&= \iint (\overline{v}(\overline{x}) - \overline{v}(\overline{y})) \cdot \operatorname{Hess} W(x-y)(\overline{v}(\overline{x}) - \overline{v}(\overline{y})) d\pi(x,\overline{x}) d\pi(y,\overline{y}) \\
&\qquad + 2\iint \tilde{v}(x,\overline{x}) \operatorname{Hess} W(x-y) \tilde{v}(x,\overline{x}) d\pi(x,\overline{x}) d\pi(y,\overline{y}).
\end{aligned}$$

As the decomposition (3.2) suggests, Simione uses the first term to quantify the displacement stability of a steady state, and the second term to quantify its spreading stability. In practice, Simione provides alternative criteria which are easier to check, although we will need to establish yet more definitions in order to make use of these criteria. Most of the technical difficulties arise as a result of the fact that the displacement vector field $\overline{v}$ need not be orthogonal to rotations, in the sense of Definition 3.6. This means Simione's framework requires a further decomposition of each vector field $\overline{v}$ into vector fields $\overline{v}_{\mathcal{R}}$ and $\overline{v}_{\mathcal{R}^\perp}$, by defining

$$\overline{v}_{\mathcal{R}}(x) = \underset{Ax, A \in so(n)}{\operatorname{argmax}} \int_{\mathbb{R}^n} (Ax) \cdot \overline{v}(\overline{x}) d\overline{\mu}(\overline{x}), \tag{3.3}$$

and by defining $\overline{v}_{\mathcal{R}^\perp}$ by

$$\overline{v}_{\mathcal{R}^\perp}(x) = \overline{v}(x) - \overline{v}_{\mathcal{R}}(x).$$

We believe that the definition of $\overline{v}_{\mathcal{R}}$ in [39], as reproduced in (3.3), contains a clerical error of some sort. This is because the algebra, $so(n)$, of skew-symmetric matrices is not compact, and hence it is not clear why $\int_{\mathbb{R}^n} (Ax) \cdot \overline{v}(\overline{x}) d\overline{\mu}(\overline{x})$ can be optimized over $so(n)$. Notwithstanding this potential clerical error (or misinterpretation), we proceed with our exposition of Simione's work.

The vector field $\overline{v}_{\mathcal{R}}$ is called the rotational part of $\overline{v}$ and, as Simione's arguments show, does not affect the displacement stability of a steady state, meaning that Simione can limit his focus to the component $\overline{v}_{\mathcal{R}^\perp}$.

Finally, Simione defines a notion of 'general position' for finite sets of particles:

Definition 3.8. We say the particles $\{x_1, ..., x_k\} \subset \mathbb{R}^n$ are in **general position** if there exists some $c > 0$ such that,

$$\max_{i \in \{1,2,...,n\}} |Ax_i| \geq c||A||$$

for every skew-symmetric matrix $A \in so(n)$, and where the norm $|| \cdot ||$ is given by $||A|| = \sqrt{\operatorname{Tr}(A^t A)}$.

Remark 3.9. Due to the equivalence of matrix norms, it is possible to work with any matrix norm, although this will affect the exact value of the constant c.

With this background out of the way, we now can quantify the displacement and spreading stability of a steady state:

Definition 3.10 (Coefficient of Displacement Stability)**.** Let $\overline{\mu}$ be a steady state of the aggregation equation with potential W, and let v be a vector field which is not identically zero in $\operatorname{spt}\overline{\mu}$. We define the quotient $Q(\overline{\mu}, v)$ of spreading stability by:

$$Q(\overline{\mu}, v) := \frac{\iint (v(\overline{x}) - v(\overline{y})) \cdot (\text{Hess}\, W)(\overline{x} - \overline{y})(v(\overline{x}) - v(\overline{y}))d\overline{\mu}(\overline{x})d\overline{\mu}(\overline{y})}{\int |v(\overline{x})|^2 d\overline{\mu}(\overline{x})}$$

Subsequently, we define the coefficient λ_d^W of spreading stability by:

$$\lambda_d^W(\overline{\mu}) := \inf\{Q(\overline{\mu}, \overline{v}_{\mathcal{R}^\perp}) |\ \overline{v}_{\mathcal{R}^\perp} \text{ is as in (3.3) and } \overline{\mu}(\overline{v}_{\mathcal{R}^\perp} = 0) = 0\},$$

with the convention that $\lambda_d^W(\overline{\mu}) = \infty$ and $\infty \cdot 0 = 0$ in the event that vector fields $\overline{v}_{\mathcal{R}^\perp}$ vanish except on a set of $\overline{\mu}$ measure zero.

Remark 3.11 (Alternative Characterization). Notice that, if c is such that, for any vector field $\overline{v}_{\mathcal{R}^\perp}$ as defined above,

$$\iint (\overline{v}_{\mathcal{R}^\perp}(\overline{x}) - \overline{v}_{\mathcal{R}^\perp}(\overline{y})) \cdot (\text{Hess}\, W(\overline{x} - \overline{y}))(\overline{v}_{\mathcal{R}^\perp}(\overline{x}) - \overline{v}_{\mathcal{R}^\perp}(\overline{y}))d\overline{\mu}(\overline{x})d\overline{\mu}(\overline{y}) \geq c \int |\overline{v}_{\mathcal{R}^\perp}(\overline{x})|^2 d\overline{\mu}(\overline{x}), \quad (3.4)$$

then $c \leq \lambda_d^W(\mu)$.

Definition 3.12 (Coefficient of Spreading Stability)**.** If $\overline{\mu}$ is a steady state of the aggregation equation with potential W, we define the coefficient $\lambda_s^W(\overline{\mu})$ of spreading stability by:

$$\lambda_s^W(\overline{\mu}) := 2 \min\{\text{eigenvalues of Hess}\, W * \overline{\mu}(\overline{x}) \mid \overline{x} \in \operatorname{spt}(\overline{\mu})\}.$$

Simione's Key Results

Using the assumptions discussed in the previous subsection, Simione proves some of his key results. His first result, which we reproduce here, shows that if $\overline{\mu}$ is both spreading and displacement stable, then the Hessian is a positive definite bilinear form on the space of admissible vector fields. More precisely,

Theorem 3.13 (Theorem 16 of [39])**.** *Let $\overline{\mu} \in \mathcal{P}_2$ be a finite particle steady state ($\overline{\mu} = \sum_{i=1}^k m_i \delta_{\overline{x}_i}$) of the aggregation equation with potential W, and let there be some $\lambda > 0$ such that $\min(\lambda_s^W(\overline{\mu}), \lambda_d^W(\overline{\mu})) \geq \lambda$. Moreover, assume that c is such that the particles $\{\overline{x}_1, ..., \overline{x}_k\}$ in $\operatorname{spt}\overline{\mu}$ are in general position with constant c, or, in other words, that*

$$\max_{i=1,...,k} |A\overline{x}_i| \geq c||A||.$$

Let $\underline{m} = \min_{i=1,\dots,k} m_i$, *and let* c_W *be the Lipschitz constant of* $Hess\, W$ *(i.e.* $|(Hess\, W(x) - Hess\, W(y))z| \leq c_W|x-y||z|$ *for all* $x, y, z \in \mathbb{R}^n$*).*

Then, if $\mu \in \mathcal{P}_2$ *is such that* $d_\infty(\mu, \overline{\mu}) < \varepsilon$, *and* v *is admissible as defined in Definition 3.6, we find that*

$$Hess\, \mathcal{E}_{W,\mu}[v,v] \geq \left(\lambda - 2c_W\varepsilon - \frac{\lambda}{c\underline{m}}\varepsilon^2\right) \int |v(x)|^2 d\mu(x).$$

Additionally, if $r \in (0,1)$ *and* ε *satisfies*

$$\varepsilon < -\frac{c_W cm}{\lambda} + \sqrt{\left(\frac{c_W c\underline{m}}{\lambda}\right)^2 + (1-r)c\underline{m}},$$

then

$$Hess\, \mathcal{E}_{W,\mu}[v,v] \geq r\lambda \int |v(x)|^2 d\mu(x).$$

Simione uses this theorem to prove that, under sufficient spreading and displacement stability assumptions at the steady state $\overline{\mu}$, the dynamics of the aggregation equation are both d_∞-Lyapunov stable and d_2-asymptotically stable near $\overline{\mu}$. To show this, we will need to define a few additional quantities:

Definition 3.14 (Various quantities)**.**

- Let B be an orthonormal basis of $so(n)$, with respect to the inner product $\langle A, A' \rangle = \mathrm{Tr}(A^t A')$ such that $\{B = A_1, \dots, A_{n(n-1)/2}\}$. We take c_1 to be an arbitrary number such that $||A_i|| \leq \frac{1}{\sqrt{c_1}}$ for all $i = 1, \dots, \frac{n(n-1)}{2}$, i.e. we can take $c_1 = 1$. If $\overline{\mu}$ is in general position in the sense of Definition 3.8, then such a c_1 is guaranteed to exist.

- If $\mathrm{spt}\, \overline{\mu} = \{\overline{x}_1, \dots, \overline{x}_k\}$, then we define $\ell := \min_{i \neq j} |\overline{x}_i - \overline{x}_j|$ and $L := \max_i |\overline{x}_i|$.

- Given a steady state $\overline{\mu}$, we define its manifold of rotations, $\mathcal{M}_{\overline{\mu}}$ by

$$\mathcal{M}_{\overline{\mu}} := \{O\#\overline{\mu} \mid O \in SO(n)\},$$

  where $SO(n)$ refers to the special orthogonal group of rotation matrices.

- Given a potential W, we define L_W to be the Lipschitz constant of ∇W, in the sense that

$$|\nabla W(x) - \nabla W(y)| \leq L_W|x-y|.$$

We are now ready to state Simione's main result:

Theorem 3.15 (Theorem 25 of [39] — d_∞-Lyapunov and d_2-asymptotic stability near steady states). *Let $\overline{\mu} := \sum_{i=1}^{k} m_i \delta_{\overline{x}_i}$ be a finite particle steady state of the aggregation equation with potential function W, and let $\lambda > 0$ satisfy $\min(\lambda_d^W(\overline{\mu}), \lambda_s^W(\overline{\mu})) \geq \lambda$ Let c_1, ℓ, and L be as defined in Definition 3.14, let $\underline{m} := \min_{1,\dots,n} m_i$, and assume $0 < \delta < \frac{1}{8}\underline{m}\sqrt{c_1}$. Moreover, assume that $\mu(t)$ is a solution to the aggregation equation such that*

$$d_\infty(\mu(0), \mathcal{M}_{\overline{\mu}}) < \delta_0 := \frac{\delta}{4} \min \left\{ \frac{1}{(\frac{4L}{\underline{m}\sqrt{c_1}} + 1)}, \frac{\lambda}{2(L_W + \frac{4n(n-1)}{\underline{m}}\lambda)} \right\}.$$

Then, for all $t > 0$

1. *$d_\infty(\mu(t), \mathcal{M}_{\overline{\mu}}) < \frac{3\delta}{4}$ (d_∞-Lyapunov stability), and*
2. *$d_2(\mu(t), \mathcal{M}) \leq \delta d_2(\mu(0), \mathcal{M}_{\overline{\mu}}) e^{-\frac{\lambda}{4}t}$ (d_2-asymptotic stability).*

In what follows, we will apply Theorem 3.15 to simplicial measures on $\mathbb{R}$ and discuss Simione's application on $\mathbb{R}^2$.

Remark 3.16. In the next section, we will use Theorem 3.15 to derive some rather surprising stability results. We take this to mean that either (i) the scope of that theorem is in some way narrower than what we envisioned in this thesis or (ii) the theorem needs to be further adapted and specialized in order to apply in one dimension.

3.3 APPLICATION TO THE ONE-DIMENSIONAL CASE

We now specialize Simione's results to the one-dimensional case and, in particular, show that, for a simplicial steady state $\nu_m := m\delta_{-1/2} + (1 - m)\delta_{1/2}$ and a power-law potential $W = W_{\alpha,\beta}$, $\lambda_s^{W_{\alpha,\beta}}(\nu_m) \geq \lambda_d^{W_{\alpha,\beta}}(\nu_m)$. We proceed by directly computing each quantity for a general potential W which is radial and sufficiently regular, and then specialize to the case of power law potentials $W_{\alpha,\beta}$, where our results can be made more explicit.

Proposition 3.17 (Calculation of λ_s^W). *Let ν_m be as defined above, and let W be twice-differentiable on $\{-\frac{1}{2}, \frac{1}{2}\} + \{-\frac{1}{2}, \frac{1}{2}\}$. Then*

$$\lambda_s^W(\nu_m) = 2\min(mW''(-1) + (1-m)W''(0), mW''(0) + (1-m)W''(1))$$

In particular, if $W = W_{\alpha,\beta}$ for $\alpha > \beta \geq 2$, then

$$\lambda_s^W(\nu_m) = \begin{cases} 2\min(m(\alpha-\beta) - (1-m), -m + (1-m)(\alpha-\beta)) & \beta = 2 \\ 2\min(m(\alpha-\beta), (1-m)(\alpha-\beta)) & \beta > 2 \end{cases} \tag{3.5}$$

Proof. Since we are working in $\mathbb{R}$, and since $-\frac{1}{2}$ and $\frac{1}{2}$ are the only two points in spt ν_m, it is immediate that

$$\lambda_s^W(\nu_m) = 2\min((W'' * \nu_m)(-\frac{1}{2}), (W'' * \nu_m)(\frac{1}{2}))$$
$$= 2\min(mW''(-1) + (1-m)W''(0), mW''(0) + (1-m)W''(1)).$$

Now notice that, if $W = W_{\alpha,\beta}$ for $\alpha > \beta \geq 2$, then we may directly compute:

$$W''_{\alpha,\beta}(x) = \begin{cases} (\alpha-1)|x|^{\alpha-2} - (\beta-1)|x|^{\beta-2}, & \beta > 2 \\ (\alpha-1)|x|^{\alpha-2} - 1 & \beta = 2 \end{cases}. \tag{3.6}$$

Hence, $W''_{\alpha,\beta}(1) = W''_{\alpha,\beta}(-1) = \alpha - \beta$, and $W''_{\alpha,\beta}(0) = 0$ if $\beta > 2$ and $W''_{\alpha,\beta}(0) = -1$ if $\beta = 2$. Substituting these values into the formula for $\lambda_s^W(\nu_m)$ yields Equation (3.5), finishing the proof. □

Proposition 3.18 (Calculation of λ_d^W)**.** *Let ν_m be as defined above, and let W be differentiable on $\{-\frac{1}{2}, \frac{1}{2}\} + \{-\frac{1}{2}, \frac{1}{2}\}$, with $W''(0) = 0$. Then*

$$\lambda_d^W(\nu_m) = W''(1) + W''(-1).$$

In particular, if $W = W_{\alpha,\beta}$ for $\alpha > \beta \geq 2$, then

$$\lambda_d^W(\nu_m) = 2(\alpha - \beta)$$

Proof. We first consider the admissibility conditions in Definition 3.6. Notice that, since we are working on $\mathbb{R}$, any vector field v is trivially orthogonal to rotations, and hence $\overline{v} = \overline{v}_{\mathcal{R}^\perp}$. On the other hand, the requirement that $\overline{v}$ be orthogonal to translation simplifies to the condition

$$0 = \int \overline{v}(\overline{x}) d\nu_m(\overline{x}) = m\overline{v}(-\frac{1}{2}) + (1-m)v(\frac{1}{2}),$$

and hence

$$(1-m)\overline{v}(\frac{1}{2}) = -m\overline{v}(-\frac{1}{2}).$$

Thus, we explicitly calculate:

$$\iint (\overline{v}(\overline{x}) - \overline{v}(\overline{y}))W''(\overline{x} - \overline{y})(\overline{v}(\overline{x}) - \overline{v}(\overline{y}))d\overline{\mu}(\overline{x})d\overline{\mu}(\overline{y})$$
$$= \iint W''(\overline{x} - \overline{y})(\overline{v}(\overline{x}) - \overline{v}(\overline{y}))^2 d\overline{\mu}(\overline{x})d\overline{\mu}(\overline{y})$$
$$= m(1-m)\left[W''(1)(\overline{v}(\frac{1}{2}) - \overline{v}(-\frac{1}{2}))^2 + W''(-1)(\overline{v}(-\frac{1}{2}) - \overline{v}(\frac{1}{2}))^2\right]$$

$$= m(1-m)(W''(1)+W''(-1))(\overline{v}(-\frac{1}{2}) - \overline{v}(\frac{1}{2}))^2. \tag{3.7}$$

Employing the fact that $\overline{v}_{\mathcal{R}^\perp}$ is admissible, we see that

$$m(\overline{v}(-\frac{1}{2}) - \overline{v}(\frac{1}{2})) = -((1-m)\overline{v}(\frac{1}{2}) + m\overline{v}(\frac{1}{2})) = -\overline{v}(\frac{1}{2}),$$

and likewise

$$(1-m)(\overline{v}(-\frac{1}{2}) - \overline{v}(\frac{1}{2})) = (1-m)(\overline{v}(-\frac{1}{2}) + m\overline{v}(-\frac{1}{2})) = \overline{v}(-\frac{1}{2}).$$

Thus, we may rewrite the expression in (3.7) as:

$$-\overline{v}(\frac{1}{2})\overline{v}(-\frac{1}{2})(W''(1)+W''(-1)).$$

On the other hand, we can express

$$\begin{aligned}\int (\overline{v}(\overline{x}))^2 d\overline{\nu}_m(\overline{x}) &= m\overline{v}(-\frac{1}{2})^2 + (1-m)\overline{v}(\frac{1}{2})^2 \\ &= -(1-m)v(-\frac{1}{2})v(\frac{1}{2}) - mv(-\frac{1}{2})v(\frac{1}{2}) \\ &= -v(-\frac{1}{2})v(\frac{1}{2}),\end{aligned}$$

which means the inequality (3.4) simplifies to

$$-\overline{v}(\frac{1}{2})\overline{v}(-\frac{1}{2})(W''(1)+W''(-1)) \geq -\lambda\overline{v}(\frac{1}{2})\overline{v}(-\frac{1}{2})$$

for all admissible vector fields $\overline{v}$. Moreover, notice that since $\overline{v}$ is orthogonal to translation, $-\overline{v}(\frac{1}{2})\overline{v}(-\frac{1}{2}) \geq 0$. Hence, since there exist non-trivial admissible vector fields v, we can conclude that

$$W''(1) + W''(-1) \geq \lambda,$$

or rather, that the largest possible choice of λ is $W''(1) + W''(-1)$.

In the case of the power-law potential $W_{\alpha,\beta}$, we notice that $W''_{\alpha,\beta}(1) = W''_{\alpha,\beta}(-1) = \alpha - \beta$, and hence, in this case, the optimal choice of λ is given by $2(\alpha-\beta)$. □

Remark 3.19 (Comparison of $\lambda_s^{W_{\alpha,\beta}}(\nu_m)$ and $\lambda_d^{W_{\alpha,\beta}}(\nu_m)$). Regardless of the value of β, the quantity in (3.5) is, at most, $\alpha - \beta$, and hence

$$\lambda_d^{W_{\alpha,\beta}}(\nu_m) = 2(\alpha-\beta) > \alpha - \beta \geq \lambda_s^{W_{\alpha,\beta}}(\nu_m).$$

This implies that, at least in one dimension, any simplex ν_m which is spreading stable with coefficient $\lambda_s^{W_{\alpha,\beta}}(\nu_m)$ will also be displacement stable with coefficient $\lambda_s^{W_{\alpha,\beta}}(\nu_m)$. Thus, in our following application of Simione's argument, we need only to concern ourselves with calculating $\lambda_s^{W_{\alpha,\beta}}(\nu_m)$.

We now turn our attention to applying Theorem 3.15 in the case of power-law potentials $W_{\alpha,\beta}$ and simplicial measures ν_m such that $\lambda_s^{W_{\alpha,\beta}}(\nu_m) > 0$. Of course, if $\beta > 2$, then Equation (3.5) implies that $\lambda_s^{W_{\alpha,\beta}}(\nu_m) > 0$. However, for $\beta = 2$, it is possible for $\lambda_{s,W_{\alpha,\beta}}(\nu_m)$ to be non-positive. Thus, for the sake of convenience, we provide an equivalent condition for $\lambda_s^{W_{\alpha,2}}(\nu_m)$ to be positive:

Lemma 3.20. *$\lambda_s^{W_{\alpha,2}}(\nu_m)$ is positive if and only if $m \in (\frac{1}{\alpha-1}, \frac{\alpha-2}{\alpha-1})$.*

Proof. Recall that

$$\lambda_s^{W_{\alpha,2}}(\nu_m) = 2\min(m(\alpha-2) - (1-m), -m + (1-m)(\alpha-2))$$

$$= 2\min(m(\alpha-1) - 1, \alpha - 2 - m(\alpha-1)).$$

Clearly, $m(\alpha-1)-1$ is positive precisely on $(\frac{1}{\alpha-1}, \infty)$, and likewise, $\alpha - 2 - m(\alpha-1)$ is positive precisely on $(-\infty, \frac{\alpha-2}{\alpha-1})$, and hence their minimum is positive only on the intersection of those intervals, precisely as desired. □

Now we may properly apply Theorem 3.15. Notice that, in the one-dimensional context, where $\bar{\nu}_n = m\delta_{-1/2} + (1-m)\delta_{1/2}$, we have that $\ell = 1$, $L = \frac{1}{2}$, and we can choose c_1 to be arbitrary. As such, we will take $c_1 = \infty$, as this simplifies the result. Moreover, we notice that $\nabla W_{\alpha,\beta}(x) = W'_{\alpha,\beta}(x) = x(|x|^{\alpha-2} - |x|^{\beta-2})$ is not globally Lipschitz. However, we can resolve this issue by noticing that ν_m is compactly supported, and hence, we may restrict our potential to a compact interval of the form $[-R, R]$, and define a new potential $\tilde{W}_{\alpha,\beta}$ by changing the values of $W_{\alpha,\beta}$ outside of $[-R, R]$ such that the Lipschitz constant of $W_{\alpha,\beta}$ on $[-R, R]$ coincides with the global Lipschitz constant of $\tilde{W}_{\alpha,\beta}$. Notice that the Lipschitz constant of $W'_{\alpha,\beta}$ on $[-R, R]$ is simply the maximum modulus of its derivative, $W''_{\alpha,\beta}$, on $[-R, R]$. Thus, in light of Equation (3.6), we may assume that R is large enough that

$$L_{\tilde{W}_{\alpha,\beta}} = \begin{cases} (\alpha-1)R^{\alpha-2} - (\beta-1)R^{\beta-2}, & \beta > 2 \\ (\alpha-1)R^{\alpha-2} - 1, & \beta = 2. \end{cases} \tag{3.8}$$

Thus, we are at last able to apply Theorem 3.15:

Corollary 3.21 (Theorem 25 of [39] for Simplicial Measures and power-law potentials.)**.** *Assume that the simplex* $\nu_m = m\delta_{-1/2} + (1-m)\delta_{1/2}$ *is a steady state of the aggregation equation with potential* $W_{\alpha,\beta}$*, for some* $\alpha > \beta \geq 2$*. Moreover, assume that either* $\beta > 2$ *or* $m \in (\frac{1}{\alpha-1}, \frac{\alpha-2}{\alpha-1})$*, such that*

$$\lambda := \begin{cases} 2\min(m(\alpha-1)-1, (\alpha-2)-m(\alpha-1)), & \beta = 2 \\ 2\min(m(\alpha-\beta), (1-m)\alpha-\beta), & \beta > 2 \end{cases}$$

is positive. By symmetry, assume $m < \frac{1}{2}$*. Let* $\delta > 0$*, and let* $\mu(t)$ *be a solution to the aggregation equation such that*

$$d_\infty(\mu(0), \nu_m) < \frac{\delta}{4}\min\left\{1, \frac{\lambda}{2(L_W + \frac{4n(n-1)}{m}\lambda)}\right\}.$$

Then for all $t > 0$*,*

1. $d_\infty(\mu(t), \nu_n) < \frac{3\delta}{4}$*, and*
2. $d_2(\mu(t), \nu_n) \leq \delta d_2(\mu(0), \nu_n) e^{-\frac{\lambda}{4}t}$.

Remark 3.22. We now clarify our earlier discussion in Remark 3.16. In particular, our concern is that Corollary 3.21 applies for any $\delta > 0$ and, as such, guarantees that, for each potential $W_{\alpha,\beta}$, there exists a single steady state $\overline{\mu}$ to which all other probability measures in $\mathcal{P}_2(\mathbb{R}^n)$ converge asymptotically with respect to the d_2 metric. More work needs to be done to determine the cause and implications of this unexpected result.

3.4 APPLICATION TO THE TWO-DIMENSIONAL CASE

In [39, Chapter 6], Simione explores the consequences of his argument in the context of simplicial steady states of the aggregation equation with power-law potential $W_{\alpha,\beta}$. We summarize his results in the following, and then attempt to build on them.

Proposition 3.23 (Spreading Stability; Section 6.2.1.2 of [39])**.** *Let* $m = (m_1, m_2)$*, and consider the simplicial steady state* $\nu_m := m_1\delta_{x_1} + m_2\delta_{x_2} + (1 - m_1 - m_2)\delta_{x_3}$ *where, without loss of generality,* $m_1 \leq m_2 \leq 1 - m_1 - m_2$*. Then for any radially symmetric potential* $W(x) = w(|x|)$*, which is sufficiently regular,*

$$\lambda_s^W(\nu_m) = w''(0)(1-m_1-m_2) + \frac{w''(\frac{1}{\sqrt{3}})}{2}\left(m_1 + m_2 - \sqrt{m_1^2 - m_1m_2 + m_2^2}\right)$$

Remark 3.24. In the case of power-law potentials $W_{\alpha,\beta}$ with $\beta > 2$, we will have that $w''(0) < 0$ and that $w''(\frac{1}{\sqrt{3}}) > 0$. Moreover, notice that

$$\lim_{m_1 \to 0} \left[m_1 + m_2 - \sqrt{m_1^2 - m_1 m_2 + m_2^2} \right] = 0,$$

and hence, by Proposition 3.23,

$$\lim_{m_1 \to 0} \lambda_s^W(\nu_m) < 0.$$

This implies that, as in the one-dimensional case, a simplicial probability measure which fails to allocate a certain minimum amount of mass to each vertex of the simplex will necessarily fail to be spreading stable.

Turning to displacement stability, Simione does not explicitly compute the coefficient of displacement stability in [39]. While this thesis does not have much to add to the work done in Simione's thesis, one can hope that, as in the one-dimensional case, the simplicial measure ν_m satisfies

$$\lambda_s^W(\nu_m) \leq \lambda_d^W(\nu_m) \tag{3.9}$$

While we do not know how to prove (3.9) in general dimension, or indeed even in two dimensions, we hope to compute $\lambda_d^W(\nu_m)$ in the future and, by comparing this to the formula for $\lambda_s^W(\nu_m)$ from Proposition 3.23, readjust our expectations for whether or not (3.9) holds.

Remark 3.25 (Stability in higher dimensions). If (3.9) indeed holds for simplices in general dimension, then this makes Simione's arguments, and in particular Theorems 3.13 and 3.15, much more useful for computations. In particular, while it becomes harder to compute both $\lambda_s^W(\nu_m)$ and $\lambda_d^W(\nu_m)$ as the dimension of the ambient space increases, the computation of $\lambda_d^W(\nu_m)$ involves computing the eigenvalues of a larger matrix. For example, in two dimensions, Simione computes $\lambda_s^W(\nu_m)$ as the minimum eigenvalue of a 2×2 matrix, whereas in order to compute $\lambda_d^W(\nu_m)$ using Simione's method, one would have to compute the minimum eigenvalue of a 6×6 matrix.

A

MULTI-WELLED POTENTIALS

We now take a brief excursion into the theory of k-welled potentials, both in $\mathbb{R}$ and $\mathbb{R}^n$. The purpose of this theory, as suggested by Robert McCann, is to allow for certain results on power-law potentials to be extended to account for more general potentials, as the definitions we will provide in this section do not require any assumptions of radiality or symmetry. One potential use of this theory is to show stability results, by considering perturbations of probability measures. In this section, we begin by defining k-welled potentials in $\mathbb{R}$, show a key recursive criterion for a potential to be k-welled, and then proceed to generalize the theory to $\mathbb{R}^n$, by discussing potentials which have k-welled restrictions to any given line.

A.1 MOTIVATION AND DEFINITION IN ONE DIMENSION

Although we have not yet defined k-welled potentials, we may view them as a generalization of attractive-repulsive potentials, at least those which have no singularities at zero. We define k-welled potentials as follows:

Definition A.1 (k-welled potential)**.** A function $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is said to be k-welled if it has at most $k-1$ local maxima, is C^{2k-2}_{loc} smooth except possibly at the local maxima, has at most k other critical points and, when it possesses $k-1$ local maxima as well as k other critical points, is non-degenerate, in the sense that its second derivative is nonzero, at each of the latter.

For example, any potential of the form $W_{\alpha,\beta}(x) = \frac{|x|^\alpha}{\alpha} - \frac{|x|^\beta}{\beta}$ is a 2-welled potential, provided that $\alpha > \beta$. Notice that our definition is broad enough to include even the case where $\beta < 0$, and hence $W_{\alpha,\beta}$ has a singularity at the origin - in this case, we simply say that $W_{\alpha,\beta}(0) = \infty$.

The theory of k-welled potentials is potentially quite powerful - while it loses some of the specific advantages of working with power-law potentials, the class is broad enough to include sums of power law potentials, as well as their perturbations. This means that, while there have not yet been many substantive results proved for k-welled potentials, they provide a promising new avenue for generalizing our work in the first three chapters.

A.2 RECURSIVE CRITERION IN ONE DIMENSION

The theory of k-welled potentials seems to be a promising avenue of study because this class of potentials admits a useful recursive criterion:

Proposition A.2 (Recursive criterion to be k-welled)**.** *Fix $k \geq 2$. If the second derivative of $f \in C^{2k-2}(\mathbb{R})$ is $(k-1)$-welled, then f is k-welled.*

Proof. We first show that the set $N := \{t | f''(t) \leq 0\}$ may have at most $k-1$ connected components. This is because f'' is continuous, and hence the Extreme Value Theorem guarantees a local maximum of f'' between any two connected components of N. Thus, were there k connected components, f'' would have $k-1$ local maxima, whereas $(k-1)$-welled potentials may have at most $k-2$ local maxima.

Moreover, f has at most one critical point (and hence local maximum) per connected component of N: were there to be two critical points, say $t_0 < t_1$, on the same connected component of N then we would have that $0 = f'(t_1) - f'(t_0) = \int_{t_0}^{t_1} f''(t)dt$. Hence, as $f'' \leq 0$ on $[t_0, t_1]$ and it integrates to 0 over this interval, f'' would have to be be identically zero on $[t_0, t_1]$, a contradiction to the fact that $(k-1)$-welled functions can only have finitely many critical points. Given that any local maximum of f must occur on N, f has at most $k-1$ local maxima - at most one on each of the at most $k-1$ connected components of N.

To show that f has at most k other critical points, we assume by way of contradiction that it has $k+1$ non-maximal critical points, with, say ℓ of them local minima. Since the remaining $k-\ell+1$ critical points of f are not local minima, they must lie in N. On the other hand, the presence of ℓ local minima implies the existence of $\ell-1$ local maxima on f which, again, must lie in N. Thus, f has $k-\ell+1$ non-maximal critical points on N and $\ell-1$ maxima on N, for a grand total of k critical points on N. However, our earlier argument again shows that f may have at most one critical point per connected component of N, and that there are at most $k-1$ connected components of N, allowing us to draw a contradiction.

Finally, if f has $k-1$ local maxima and k other critical points, our previous argument demonstrates that each local maximum must be the only critical point on a connected component of N. Thus, all other critical points must occur on $\mathbb{R} \setminus N = \{t | f''(t) > 0\}$, and hence must be non-degenerate local minima of f. □

In addition, in the case that $k = 2$, we may prove an alternative version of Proposition A.2, wherein we replace the condition that f'' is $(k-1)$-welled, with the condition that f'' is convex. Moreover, in this case, we may relax the requirement that $f \in C^2(\mathbb{R})$ and replace it with the requirement that $f \in C^2_{loc}(\mathbb{R})$.

Lemma A.3. *If $f \in C^2_{loc}(\mathbb{R})$ has a strictly convex second derivative, then f is $2-$welled.*

The proof of Lemma A.3 proceeds in an identical manner to that of Proposition A.2, and hence we omit it. The following corollary is an immediate consequence of applying Lemma A.3 once and Proposition A.2 $k-2$ times:

Corollary A.4 (Convexity criterion to be k-welled)**.** *For $k \geq 2$, if $f \in C^{2k-2}_{loc}(\mathbb{R})$ has a strictly convex $(2k-2)$-th derivative, then f is k-welled.*

A.3 MULTI-WELLED POTENTIALS IN GENERAL DIMENSION

We now take some first steps to generalize the theory of k-welled potentials to higher dimensions. In light of Corollary A.4, and the emphasis it places on convexity, it is natural to look at functions whose restriction to any given line is k-welled. More precisely:

Definition A.5 (k-welled Restrictions)**.** We say that the function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is said to have k-welled restrictions if, for every line $\ell : \mathbb{R} \to \mathbb{R}^n$, thought of as a parametrized curve, the function $f \circ \ell : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ is k-welled.

The class of potentials which have k-welled restrictions is significantly more general than the class of power-law potentials - for example, it is clear that power-law potentials are 2-welled, and that any finite linear combination of power-law potentials is k-welled for some k. Moreover, this class is clearly quite robust under smooth perturbations — heuristically, along any line, a perturbation can only add a finite number of wells. As such, while there have not yet been any significant results proven for k-welled potentials, such results will be considerably more general than those discussed in the preceding chapters of the present thesis.

BIBLIOGRAPHY

[1] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures,* Second edition, Lectures in Mathematics ETH Zurich, Birkhäuser Verlag, Basel, 2008. MR2401600

[2] L. Ambrosio, E. Mainini, and S. Serfaty, *Gradient flow of the Chapman–Rubinstein–Schatzman model for signed vortices,* Ann. Inst. H. Poincaré, **28** (2011), no. 2, 217-246, DOI: 10.1016/j.anihpc.2010.11.006. MR2784070

[3] D. Balagué, A. Barbaro, J. A. Carrillo, and R. Volkin, *Analysis of Spherical Shell Solutions for the Radially Symmetric Aggregation Equation*, SIAM J. Appl. Dyn. Syst. **19** (2020), no. 4, 2628–2657, DOI: 10.1137/20M1314549. MR4176887

[4] D. Balagué, J. A. Carrillo, T. Laurent, and G. Raoul, *Dimensionality of Local Minimizers of the Interaction Energy*, Arch. Ration. Mech. Anal. **209** (2013), no. 3, 1055-1088, DOI: 10.1007/s00205-013-0644-6. MR3067832

[5] D. Balagué, J. A. Carrillo, T. Laurent, and G. Raoul, *Nonlocal interactions by repulsive-attractive potentials: radial ins/stability*, Phys. D. **260** (2013), 5-25, DOI: 10.1016/j.physd.2012.10.002. MR3143991

[6] A. Burchard, R. Choksi, and I. Topaloglu, *Nonlocal shape optimization via interactions of attractive and repulsive potentials*, Indiana Univ. Math. J., **67** (2018), no. 1, 375-395, DOI: 10.1512/iumj.2018.67.6234. MR3776026

[7] J. A. Carrillo, M. G. Delgadino, J. Dolbeault, R. L. Frank, and F. Hoffman, *Reverse Hardy-Littlewood-Sobolev inequalities,* J. Math. Pures Appl., **132** (2019), 133-165, DOI: 10.1016/j.matpur.2019.09.001. MR4030251

[8] J. A. Carrillo, A. Figalli, and F. S. Patacchini, *Geometry of minimizers for the interaction energy with mildly repulsive potentials*, Ann. Inst. H. Poincaré, **34** (2017), no. 5, 1299-1308, DOI: 10.1016/j.anihpc.2016.10.004. MR3742525

[9] J. A. Carrillo, R. McCann, and C. Villani, *Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates*, Revista Mat. Iberoamericana **19** (2003), no. 3, 971-1018, DOI: 10.4171/RMI/376. MR2053570

[10] J. A. Carrillo, R. McCann, and C. Villani, *Contractions in the 2-Wasserstein length space and thermalization of granular media*, Arch. Ration. Mech. Anal. **179** (2006), no. 2, 217–263, DOI: 10.1007/s00205-005-0386-1. MR2209130

[11] J. A. Carrillo and R. Shu, *From radial symmetry to fractal behavior of aggregation equilibria for repulsive-attractive potentials*. Preprint at arxiv: https://arxiv.org/abs/2107.05079

[12] K. Craig and A. Bertozzi, *A blob method for the aggregation equation,* Math. Comput., **35** (2016), no. 300, 1681-1717, DOI: 10.1090/mcom3033. MR3471104

[13] R. Choksi, R. C. Fetecau, and I. Topaloglu, *On minimizers of interaction functionals with competing attractive and repulsive potentials*, Ann, Inst. H. Poincaré, **32** (2015) no. 6, 1283-1305, DOI: 10.1016/j.anihpc.2014.09.004. MR3425263

[14] C. Davies, T. Lim and R. McCann. *Classifying minimum energy states for interacting particles (I) — spherical shells.* Preprint (2021), arXiv:2107.11718

[15] C. Davies, T. Lim and R. McCann. *Classifying minimum energy states for interacting particles (II) — regular simplices.* Preprint (2021), arXiv:2109.07091

[16] J. Dolbeault and B. Perthame, *Optimal critical mass in the two dimensional Keller-Segel model in* $\mathbb{R}^2$, C. R. Math. Acad. Sci. Paris, **339** (2004), no. 9, 611-616, DOI: 10.1016/j.crma.2004.08.011. MR2103197

[17] B. Düring, C. Gottschlich, S. Huckemann, L. M. Kreusser, and C.-B. Schönlieb, *An anisotropic interaction model for simulating fingerprints*, J. Math. Biol. **78** (2019), no. 7, 2171-2206, DOI: 10.1007/s00285-019-01338-3. MR3954377

[18] L. C. Evans, *Partial Differential Equations Second Edition,* Graduate Studies in Mathematics, Vol. 19, American Mathematical Society, Providence, 2010. MR2597943

[19] R. C. Fetecau, Y. Huang, and T. Kolokolnikov, *Swarm dynamics and equilibria for a nonlocal aggregation model,* Nonlinearity, **24** (2011), no. 10, 2681-2716, DOI: 10.1088/0951-7715/24/10/002. MR2834242

[20] G. B. Folland, *Real Analysis: Modern Techniques and their Applications,* Second edition, Pure and Applied Mathematics: A Wiley-Interscience Publication, John Wiley & Sons, New York, 1999. MR1681462

[21] R. Frank, *Minimizers for a one-dimensional interaction energy,* Preprint (2021) arXiv:2109.08929 .

[22] W. Gangbo and R. McCann, *The geometry of optimal transportation,* Acta. Math. **177** (1996), no. 2, 113-161, DOI: 10.1007/BF02392620. MR1440931

[23] R. Hegselmann and U. Krause, *Opinion dynamics and bounded confidence: Models, analysis, and simulation,* J. Artif. Soc. Soc. Simul. **5** (2002), no. 3, available at `https://www.jasss.org/5/3/2.html`

[24] M. A. Herrero and J. J. L. Velásquez, *Chemotactic collapse for the Keller-Segel model,* J. Math. Biol. **35** (1996), no. 2, 177-194, DOI: 10.1007/s002850050049. MR1478048

[25] D. D. Holm and V. Putkaradze, *Formation of clumps and patches in self-aggregation of finite-size particles,* Physica D, **220** (2006), no. 2, 183-196, DOI: 10.1016/j.physd.2006.07.010. MR2256869

[26] D. Horstmann, *From 1970 until present: the Keller-Segel model in chemotaxis and its consequences,* I. Jahresber. Deutsch. Math.-Verein. **105** (2003), no. 3, 103–165. MR2013508

[27] K. Kang, H. K. Kim, T. Lim, and G. Seo, *Uniqueness and characterization of local minimizers for the interaction energy with mildly repulsive potentials*, Calc. Var. Partial Differential Equations **60** (2021), no. 1, DOI: 0.1007/s00526-020-01882-7. MR4201638

[28] E. F. Keller and L. A. Segel, *Model for Chemotaxis*, J. Theor. Biol. **30** (1971), no. 2, 225-234, DOI: 10.1016/0022-5193(71)90050-6

[29] T. Kolokolnikov, J. A. Carrillo, A. Bertozzi, R. Fetecau, and M. Lewis, *Emergent behaviour in multi-particle systems with non-local interactions [Editorial]*, Phys. D, **260** (2013), 1–4, DOI: 10.1016/j.physd.2013.06.011. MR3143990

[30] T. Lim and R. McCann, *Geometrical bounds for the variance and recentered moments*, Math. Oper. Res. Published online in Articles in Advance 24 May 2021, DOI: https://doi.org/10.1287/moor.2021.1125.

[31] T. Lim and R. McCann, *Isodiametry, variance, and regular simplices from particle interactions*, Arch. Ration. Mech. Anal. **241** (2021), no. 2, 553-576, DOI: 10.1007/s00205-021-01632-9. MR4275740

[32] O. Lopes, *Uniqueness and radial symmetry of minimizers for a nonlocal variational problem*, Comm. Pure. Appl. Anal. **18** (2019), no. 5, 2265-2282, DOI: 10.3934/cpaa.2019102. MR3962176

[33] R. McCann and N. Guillen, *Five lectures on optimal transportation: geometry, regularity and applications*, In Analysis and Geometry of Metric-Measure Spaces: Lecture Notes of the Séminaire Mathématiques Supériuere (SMS) Montréal (2011), eds Dafni G, et al, 145–180. MR3060502

[34] F. Otto, *The geometry of dissipative evolution equations: the porous medium equation*, Comm. Partial Differential Equations, **26** (2001), no. 1-2, 101-174, DOI: 10.1081/PDE-100002243. MR1842429

[35] D. Panchenko, *Lecture Notes on Probability Theory,* First Edition, self-published, (2019), available at: `https://drive.google.com/file/d/0B6JeBUquZ5BwVlV1dEpsTHVVbTA/view?resourcekey=0-5xbvKhDXZjrYRLfqppsHuQ`

[36] C. S. Patlak, *Random walk with persistence and external bias*, Bull. Math. Biophys. **15** (1953), no. 3, 311-338, DOI: 10.1007/BF02476407. MR0081586

[37] F. Santambrogio, *Optimal Transport for Applied Mathematicians*, Progress in Nonlinear Differential Equations and their Applications, Vol. 87, Birkhäuser/Springer, Cham, 2015. MR3409718

[38] L. Schwartz, *Transformation de Laplace des distributions*, Comm. Sém. Math. Univ. Lund, (1952), Tome supplementaire, p. 196-206. MR0052555

[39] R. Simione, *Properties of Energy Minimizers of Nonlocal Interaction Energy*, PhD Thesis, Carnegie Mellon University and Instituto Superior Técnico, 2014. MR3337754

[40] H. Sun, D. Uminsky, and A. L. Bertozzi, *Stability and clustering of self-similar solutions of aggregation equations,* J. Math. Phys. **53** (2012), no. 11, 115610, 18 pp. DOI: 10.1063/1.4745180. MR3026555

[41] C. M. Topaz and A. L. Bertozzi, *Swarming patterns in a two-dimensional kinematic model for biological groups,* SIAM J. Appl. Math. **65** (2004), no. 1, 152-174, DOI: 10.1137/S0036139903437424. MR2111591

[42] C. M. Topaz, A. L. Bertozzi, and M. A. Lewis, *A nonlocal continuum model for biological aggregation,* Bull. Math. Biol. **68** (2006), no. 7, 1601-1623, DOI: 10.1007/s11538-006-9088-6. MR2257718

[43] C. Villani, *Optimal Transport: Old and New*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Vol. 338, Springer, New York, 2009. MR2459454

COLOPHON

This thesis was typeset using the typographical look-and-feel `classicthesis` developed by André Miede and Ivo Pletikosić.

The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*".

Figures were created using the tikz package and Desmos graphing software.

Please send any feedback or corrections to me via email at cameron.davies@mail.utoronto.ca.