A CONVEXITY THEORY FOR INTERACTING GASES AND EQUILIBRIUM CRYSTALS

Robert John McCann

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
MATHEMATICS

November 1994

To my parents
John Joseph and Daisy McCann

ABSTRACT

A new set of variational inequalities is introduced, based on a novel but natural interpolation between Borel probability measures on $\mathbb{R}^d$. Used in lieu of convexity or rearrangement inequalities, these estimates lead to existence and uniqueness results concerning equilibrium states for (i) attracting gases; and (ii) plane crystals in an external field.

Consider a d-dimensional gas of particles interacting through a force which increases with distance and obeying an equation of state $P = P(\varrho)$ relating pressure to density. For $P(\varrho)/\varrho^{1-1/d}$ non-decreasing, a unique energy minimizing state is shown to exist up to translation.

For a two-dimensional crystal in a convex potential, the equilibrium shape was known to consist of a countable disjoint union of closed convex sets. Here it is shown that each convex component minimizes the energy uniquely among convex sets of its area. Assuming symmetry under $x \leftrightarrow -x$, the crystal formed must be unique, convex and connected. The last result leads to a new proof that a convex crystal $C = -C$ away from equilibrium remains convex and balanced under curvature-driven flow.

Incidental results include new generalizations of the Brunn-Minkowski inequality from sets to measures, and new derivations of inequalities due to Prékopa, Leindler, Brascamp and Lieb. A theorem of Brenier is improved to yield existence of a unique measure-preserving mapping with a convex potential between any pair of $L^1(\mathbb{R}^d)$ probability measures on $\mathbb{R}^d$; the potential satisfies a Monge-Ampère equation almost everywhere.

A separate section considers compressible fluid models for a rotating star. For fixed mass and large angular momentum, stable uniformly rotating solutions to the associated Navier-Stokes-Poisson system are constructed in the form of binary stars with specified mass ratio. A one-dimensional toy model admitting explicit solution is also introduced: to any specified num-

ber of components and their masses corresponds a single family of solutions, parameterized by angular velocity up to the point of equatorial break-up.

Acknowledgements

I am deeply indebted to my advisor, Elliott Lieb, for his guidance and inspiration. He led me into the study of variational problems, and taught me to penetrate the heart of an issue. His generosity with his time and ideas will not be forgotten.

Many others provided fruitful discussions and vital remarks. It is a pleasure to thank Ivan Blank, Luis Caffarelli, Eric Carlen, Michael Loss, Andrew Mayer, Kate Okikiolu, Peter Osvath, Jean Taylor, Horng-Tzer Yau and especially John Stalker.

Frederick Almgren guided my forays into the mathematics of crystals, and was an ongoing source of advice and encouragement. Elliott Lieb suggested the problem of rotating stars; its formulation in terms of angular momentum and the reduction to uniform rotation are his. I thank Michael Aizenman and Charles Fefferman for their interest and support.

Funding was provided by a 1967 Scholarship of the Natural Sciences and Engineering Research Council of Canada, and a Proctor Fellowship of the Princeton Graduate School.

Finally, I wish to express my gratitude to Jan-Philip Solovej. He has always been there for me, and his example and insight continue to inspire.

Contents

1 Introduction

The analysis of energy functionals plays a crucial role both in mathematical physics and in partial differential equations. Here the central issues are to determine the existence of stationary configurations, particularly optimizers, and their properties: uniqueness, stability, symmetry.... Convexity, when present, is a powerful tool for resolving these questions.

This thesis develops a new convex structure on $\mathcal{P}(\mathbb{R}^d)$, the space of Borel probability measures on $\mathbb{R}^d$. That is, for $\rho, \rho' \in \mathcal{P}(\mathbb{R}^d)$ and $t \in (0,1)$, an interpolant $\rho_t \in \mathcal{P}(\mathbb{R}^d)$ is defined — see (3) below — which is for some purposes more natural than $(1-t)\rho + t\rho'$. Interesting estimates of the form $E(\rho_t) \leq (1-t)E(\rho) + tE(\rho')$ are established. With these estimates in lieu of convexity or rearrangement inequalities for $E(\rho)$, the tools of convex analysis are brought to bear on problems in which they were not formerly thought to apply.

Two problems from mathematical physics are solved in this way. The first models an interacting gas in which the attractive force increases with separation, while the second involves the shape of an equilibrium crystal grown in a convex potential. The third part of the thesis is essentially independent; it presents some results concerning rotating stars.

Main Results

The first part of this thesis concerns a d-dimensional gas of particles interacting through a convex potential $V(x)$ on $\mathbb{R}^d$, and obeying an equation of state $P = P(\varrho)$ relating pressure to density. Normalizing the mass of the gas to be one, the state of the system is given by a mass density $\rho \in \mathcal{P}(\mathbb{R}^d)$ absolutely continuous with respect to Lebesgue: $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$. The corresponding energy is

$$E(\rho) := \int_{\mathbb{R}^d} A(\rho(x))dx + \iint d\rho(x)V(x-y)d\rho(y). \tag{1}$$

Here $A(\varrho)$ is a convex function determined by the pressure through (27); examples include $A(\varrho) = \varrho^q$ for $q > 1$. If $P(\varrho) \geq 0$ with $P(\varrho)/\varrho^{1-1/d}$ non-decreasing, then $E(\rho_t)$ will be convex as a function of t. Under a slightly stronger assumption[1], such as strict convexity of $V(x)$, existence of a unique energy minimizer up to translation is proved.

The same technique yields a uniqueness result for a two-dimensional crystal in an external field. The field is assumed to be the negative gradient of a convex potential $Q(x)$, and to vanish only on a bounded set of measure zero. The crystal configuration is given by a set $K \subset \mathbb{R}^2$ of unit area, while the energy to be minimized is

$$\varepsilon(K) := \int_{\partial K} F(\hat{\tau}_x)\, d\mathcal{H}^1(x) + \int_K Q(x)d^2x. \tag{2}$$

Here $\mathcal{H}^1$ is one-dimensional Hausdorff measure on the boundary ∂K of K; $F(\hat{\tau}_x) > 0$ is the surface tension, which depends on the oriented unit tangent $\hat{\tau}_x$ to K and satisfies a triangle inequality[2]. For the isotropic case $F = 1$ the surface energy is the length of ∂K. If $\nabla Q = 0$, the shape of the minimizer is given by the Wulff construction [1], but when ∇Q is non-zero, little is known beyond the case of a sessile crystal in a uniform field [2]. Only an unpublished result of Okikiolu [3] shows that each connected component of the minimizing crystal must be convex. Here it is proved that any such component uniquely minimizes $\varepsilon(K)$ among convex sets of its area. If W and F are symmetric under $x \leftrightarrow -x$, it follows that the minimizing crystal will be unique, convex and connected. Whether there is freedom to translate depends on ∇Q.

[1] $P(\varrho)/\varrho^2$ non-integrable at ∞ to ensure $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$.

[2] F is convex when extended to $x \in \mathbb{R}^2$ by $F(\lambda x) := \lambda F(x)$ for $\lambda \geq 0$.

Finally, the problem of rotating stars is examined. Instead of prescribing a rotation law [4] or uniform angular velocity [5], it is formulated as a variational minimization over all densities $\rho \in \mathcal{P}_{ac}(\mathbb{R}^3)$ and velocity vector fields on $\mathbb{R}^3$ which are consistent with a specified linear momentum and angular momentum about the center of mass. The energy of such a configuration consists of (1) with $V(x) = -1/|x|$ (Newtonian gravity), plus a kinetic term. For suitable $P(\varrho)$, this energy will be bounded below, but it never assumes a global minimum. Even for *local* minima to exist, we show that it is necessary to put a very strong, physically motivated topology on the configuration space: that induced by the Wasserstein L^∞ metric [6]. Local minima in this topology turn out to be stable, uniformly rotating solutions to the Euler- or Navier-Stokes-Poisson system (61-62), but the stationarity condition they satisfy is slightly weaker than for a global energy minimizer: disconnected components of the star need not have the same chemical potential. For large angular momentum, such minima are proved to exist in the form of binary stars with arbitrary mass ratio. This is our only existence result for the real problem. However, we also introduce a one-dimensional model which captures some of the complexity of the full problem, while retaining the virtue of being explicitly solvable. For a given mass, the solutions ρ come in uncountably many disjoint families, distinguished by the number of connected components in $\{\rho > 0\}$ and their masses. Each family is continuously parameterized by velocity of rotation, and terminates with equatorial break-up of the lightest component.

Outline of Methods

If $\rho \in \mathcal{P}(\mathbb{R}^d)$ and $T : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ is a measurable transformation of $\mathbb{R}^d$, a new probability measure $T_\# \rho$ is defined by $T_\# \rho[M] = \rho[T^{-1}(M)]$ for $M \subset \mathbb{R}^d$. $T_\# \rho$ is called the *push-forward* of ρ through T. A slight extension of a theorem of Brenier [7] (also Theorem B.1 below) yields:

Theorem: *Given $\rho, \rho' \in \mathcal{P}(\mathbb{R}^d)$ with $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$, there exists a convex function ψ on $\mathbb{R}^d$ whose gradient $\nabla\psi$ pushes forward ρ to ρ'. $\nabla\psi$ is uniquely determined almost everywhere with respect to ρ.*

Our interpolant ρ_t between ρ and ρ' may now be defined in terms of ψ:

$$\rho_t := [(1-t)id + t\nabla\psi]_{\#}\rho \tag{3}$$

where $id : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ is the identity transformation. A trivial example, to be contrasted with the usual interpolation, occurs when ρ' is a translate of ρ: $\rho'(\,\cdot\,) = \rho(\,\cdot\, - x)$ implies $\rho_t(\,\cdot\,) = \rho(\,\cdot\, - tx)$, also a translate of ρ.

Control of the internal energy in (1) depends crucially on convexity of ψ. For the $L^q(\mathbb{R}^d)$ norm, rather strong inequalities are implied: $\|\rho_t\|_q^{-q'/d}$ is a concave function of t when $q^{-1}+q'^{-1} = 1$. This result is a generalization of the Brunn-Minkowski inequality from sets to measures: the classical inequality is recovered by interpolating between the uniform probability measures on two given sets. Convexity of ψ also plays a role in the estimates for the energy in (2): here the endpoints of the interpolation are characteristic functions of two convex sets with unit area; for the interpolating measure $\|\rho_t\|_\infty \leq 1$.

Organization

The three parts of this thesis focus on the three problems here discussed. Properties of the interpolation (3) are developed alongside the first application. Four appendices are also provided. The first summarizes facts of life regarding differentiability of convex functions, while the second gives the refinement of Brenier's theorem described above. A different interpolation sharing the convexity properties of (3) may be based on certain explicitly constructed measure-preserving maps; Appendix C explores this alternative.

Finally, Appendix D exploits these ideas to provide a new proof of inequalities due to Prékopa-Leindler [8, 9, 10] and Brascamp-Lieb [11]. The error term in these inequalities is obtained exactly!

Part I
Interacting Gases

2 Attraction Through a Convex Potential

Consider a d-dimensional gas of particles interacting through a convex potential $V(x)$ on $\mathbb{R}^d$ and obeying an equation of state in which the pressure $P(\varrho)$ is a function of the density only. The state of the system is described by an $L^1(\mathbb{R}^d)$ mass density $\rho(x) \geq 0$; since the total mass of gas is fixed it may be normalized to $\int \rho = 1$. Thus $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$, the space of absolutely continuous probability measures on $\mathbb{R}^d$. The corresponding energy $E(\rho) = U(\rho) + G(\rho)$ is a sum of the internal energy due to compression and the potential energy due to the interaction; one would like to show that the competition between these two terms results in a unique ground state. $E(\rho)$ is given by

$$E(\rho) := \int_{\mathbb{R}^d} A(\rho(x))dx + \iint d\rho(x)V(x-y)d\rho(y) \tag{4}$$

where the first term is the internal energy $U(\rho)$. Its local density $A(\varrho)$ is determined by the pressure through (27); to be physical, $P(\varrho)$ should be non-decreasing and $A(\varrho)$ convex. Under slightly stronger assumptions — notably $P(\varrho)/\varrho^{1-1/d}$ non-decreasing, and $V(x)$ either strictly convex or spherically symmetric — Theorems 5.1 and 5.3 show that $E(\rho)$ admits a minimizer in $\mathcal{P}_{ac}(\mathbb{R}^d)$, unique up to translation. Examples satisfying these assumptions include the polytropic equations of state $P(\varrho) = (q-1)A(\varrho) = \varrho^q$ with $q > 1$; for particular q, this approximates (semi-classically) the quantum kinetic energy of a gas of fermions: in three dimensions $q = 5/3$ [12].

The existence result is obtained by a continuity-compactness argument, but in the absence of convexity there are few general tools for proving uniqueness. For a spherically symmetric potential $V(x)$, one might use a sharp rearrangement inequality to reduce the problem to one dimension and then try to study the associated ordinary differential equation. This approach has been used successfully [13] in the important case of the Coulomb potential $V(x) = -|x|^{-1}$ with the Chandrasekhar equation of state. The strategy pursued in the sequel is rather different: it is based on the presence of a peculiar sort of convexity in the functional (4).

Although $V(x)$ need not be spherically symmetric, Newton's Third Law or the symmetry in (4) show $V(x) = V(-x)$ to be completely general. Thus $V(x)$ must be minimized at the origin, and $G(\rho)$ cannot be convex: the Dirac point mass δ_x at $x \in \mathbb{R}^d$ is a minimum for $G(\rho)$, but $(1-t)\delta_x + t\delta_y$ is not. However, if $\rho_t = \delta_{(1-t)x+ty}$ were used instead of $(1-t)\delta_x + t\delta_y$ to interpolate between the two Dirac measures, then the potential energy $G(\rho_t)$ would be t-independent as a reflection of its translation invariance. Moreover, for a positive linear combination of such point masses

$$\rho_t = \sum_i m_i \delta_{(1-\mathrm{t})\mathrm{x_i}+\mathrm{ty_i}},$$

$G(\rho_t)$ is a convex function of t. This point of view, which emphasizes the linear structure of $\mathbb{R}^d$ over that of the measure space, is reminiscent of the Lagrangian formulation in fluid mechanics. It is developed in the following chapter, where we introduce a new convex structure on $\mathcal{P}_{ac}(\mathbb{R}^d)$: between arbitrary measures $\rho, \rho' \in \mathcal{P}_{ac}(\mathbb{R}^d)$, for $t \in (0,1)$ an interpolant $\rho_t \in \mathcal{P}_{ac}(\mathbb{R}^d)$ is defined. (This structure extends to the space of all Borel probability measures $\mathcal{P}(\mathbb{R}^d)$). Moreover, both the internal and potential energies satisfy estimates of the form

$$E(\rho_t) \leq (1-t)E(\rho) + tE(\rho'); \tag{5}$$

they are convex functions of the interpolation parameter t. The existence and uniqueness results follow rapidly. The estimates (5) may be of some interest apart from the application: when $A(\varrho) = \varrho^q$ for $q > 1$, scaling of $U(\rho)$ implies logarithmic convexity of the $L^q(\mathbb{R}^d)$ norm. In fact, $\|\rho_t\|_q^{-q'/d}$ is concave as a function of t when $q^{-1} + q'^{-1} = 1$. That this result generalizes the Brunn-Minkowski inequality from sets to measures is most readily seen when $q = \infty$. The assumption on $P(\varrho)$ leading to (5) was that, as a function of dilation factor, $U(\rho)$ be convex non-increasing under mass preserving dilations of ρ.

In the following chapter, the interpolant ρ_t is introduced and its basic properties are described. Convexity of the internal energy $U(\rho_t)$ is established in Chapter 4, but the technical details underlying the proof are relegated to Chapter 6. The existence and uniqueness theorems for the attracting gas may be found in Chapter 5.

3 Interpolation of Probability Measures

The current chapter is devoted to defining and establishing the basic properties of the convex structure on $\mathcal{P}(\mathbb{R}^d)$ which is here introduced. Apart from Remark 3.9, the discussion is restricted to the case in which one of the measures is absolutely continous with respect to Lebesgue, in order to streamline the exposition. Before considering measures on $\mathbb{R}^d$, the interpolation is characterized in the simplest case: for measures on the line $d = 1$.

Let $\rho, \rho' \in \mathcal{P}_{ac}(\mathbb{R})$. For $x \in \mathbb{R}$, there exists $y(x) \in \mathbb{R} \cup \{\pm\infty\}$ such that

$$\rho[\,(-\infty, x)\,] = \rho'[\,(-\infty, y(x))\,]. \tag{6}$$

Although $y(x)$ may not be one-to-one or single-valued, its value will be uniquely determined ρ-a.e. At the remaining points, a choice may be made for which $y(x)$ will be non-decreasing. As the *time* t is varied between 0 and 1, the idea of the interpolation is to linearly displace the mass lying under ρ at x towards the corresponding point $y(x)$ for ρ', so that the interpolant ρ_t assigns mass $\rho[(-\infty, x)]$ to the interval $(-\infty,\, (1-t)\,x + t\,y(x)\,)$. This condition turns out to characterize ρ_t. A simple example occurs when ρ and ρ' are Gaussian measures with means μ and μ' and standard deviations σ and σ' respectively: $y(x)$ is an affine function with slope σ'/σ, while ρ_t is the Gaussian measure with mean $(1-t)\mu + t\mu'$ and deviation $(1-t)\sigma + t\sigma'$.

To define ρ_t more generally requires a few notions from measure theory. Let $(X, \mathcal{X})$ denote a *measurable space*, meaning $\mathcal{X}$ is a σ-algebra of subsets of X. Let $(Y, \mathcal{Y})$ be another measurable space, $T : X \longrightarrow Y$ a measurable transformation and ω a measure on $(X, \mathcal{X})$. Then T and ω induce a measure $T_{\#}\omega$ on $(Y, \mathcal{Y})$ defined by

$$T_{\#}\omega[M] := \omega[T^{-1}(M)] \tag{7}$$

for $M \in \mathcal{Y}$. $T_{\#}\omega$ is called the *push-forward* of ω through T; it is a probability measure if ω is. Observe that T need only be defined ω-a.e. The change of variables theorem states that if f is a measurable function on Y, then

$$\int_Y f(y)\, dT_{\#}\omega(y) = \int_X f(T(x))\, d\omega(x). \tag{8}$$

Unless otherwise indicated, the measurable spaces X and Y will both be $\mathbb{R}^d$ with the σ-algebra of Borel sets. d-dimensional Lebesgue measure will play a frequently role; it is denoted by *vol.*

Given $\rho, \rho' \in \mathcal{P}(\mathbb{R}^d)$, ρ absolutely continuous with respect to Lebesgue, we require a transformation T which will push forward ρ to ρ'. There are many such T, several of which are suited for the present purposes. One suitable map is constructed explicitly in Appendix C, while a more elegant transformation is shown to exist by a result of Brenier [7]. It states that T can be realized as the gradient of a convex function $\psi : \mathbb{R}^d \longrightarrow \mathbb{R} \cup \{+\infty\}$. Brenier proved this under mild restrictions on the measures ρ and ρ', but the restrictions are lifted in Theorem B.6. Using this theorem, we define the displacement interpolation between ρ and ρ':

Definition 3.1 (Displacement Interpolation)
Given probability measures $\rho, \rho' \in \mathcal{P}(\mathbb{R}^d)$ with $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$, there exists ψ convex on $\mathbb{R}^d$ such that $\nabla\psi_{\#}\rho = \rho'$. Let id *denote the identity mapping on $\mathbb{R}^d$. At time $t \in [0,1]$, the displacement interpolant $\rho_t \in \mathcal{P}(\mathbb{R}^d)$ between ρ and ρ' is defined by*

$$\rho_t := [(1-t)id + t\nabla\psi]_{\#}\rho. \tag{9}$$

The extension of this definition to $t \in \mathbb{R}$ is suppressed. On the line $d = 1$, the monotone function $y = \nabla\psi(x)$ is readily seen to satisfy (6), and the characterization given for ρ_t follows rapidly.

From Definition 3.1, it is almost immediate that the interaction energy $G(\rho)$ in (26) will be convex function along the *lines* of the displacement interpolation. We say that the functional $G(\rho)$ is *displacement convex.*

Proposition 3.2 (Displacement Convexity of Potential Energy G) *Given probability measures $\rho, \rho' \in \mathcal{P}_{ac}(\mathbb{R}^d)$, let ρ_t be the displacement interpolant between them (9). Then the interaction energy $G(\rho_t)$ is a convex function of t on $[0,1]$. Strict convexity of $G(\rho_t)$ will follow from that of the potential $V(x)$, except when ρ' is a translate of ρ.*

Proof: By the change of variables theorem (8)

$$\begin{aligned} G(\rho_t) &:= \iint d\rho_t(x)\, V(x-y)\, d\rho_t(y) \\ &= \iint d\rho(x)\, V\Big((1-t)(x-y) + t(\nabla\psi(x) - \nabla\psi(y))\Big)\, d\rho(y). \end{aligned}$$

Since $V(x)$ is a convex function on $\mathbb{R}^d$, the integrand above is manifestly convex as a function of t. This proves the initial assertion. If the convexity of $V(x)$ is strict, the integrand will be strictly convex unless

$$\nabla\psi(x) - \nabla\psi(y) = x - y. \tag{10}$$

The *integral* will be strictly convex unless (10) holds almost everywhere $\rho \times \rho$, in which case $\nabla\psi(x) - x$ is x-independent ρ-a.e. This would imply that ρ' is ρ translated by $\nabla\psi(x) - x$. QED.

The displacement convexity of the internal energy $U(\rho)$ is a deeper result. There the convexity of ψ, not used in the proof of Proposition 3.2, enters crucially. Before attacking this issue, it will be helpful to illuminate some of the elementary properties of the displacement interpolation. The next propositions show that it induces a bona fide convex structure on $\mathcal{P}_{ac}(\mathbb{R}^d)$ and explore the relationship between this structure and the symmetries of $\mathbb{R}^d$ — translation, dilation, reflection, rotation. The proofs are postponed until the end of this chapter. Wherever ambiguity seems likely to arise, $\rho \xrightarrow{t} \rho'$ is used instead of ρ_t to indicate explicit dependence on the endpoints ρ and ρ'.

Proposition 3.3 *Let $\rho, \rho' \in \mathcal{P}(\mathbb{R}^d)$ be probability measures with $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$. For $t \in [0,1]$, the displacement interpolant $\rho_t = \rho \xrightarrow{t} \rho'$ from (9) satisfies*

(i) $\rho_0 = \rho$ *and* $\rho_1 = \rho'$*;*

(ii) ρ_t *is uniquely determined by* ρ *and* ρ'*;*

(iii) ρ_t *is absolutely continuous with respect to Lebesgue for* $t < 1$*;*

(iv) $\rho \xrightarrow{t} \rho' = \rho' \xrightarrow{1-t} \rho$ *(when the latter is defined);*

(v) if $s, t' \in [0,1]$*, then* $\rho_t \xrightarrow{s} \rho_{t'} = \rho \xrightarrow{(1-s)t+st'} \rho'$.

Remark 3.4 Proposition 3.3(iii) may be interpreted to mean that the absolutely continuous measures $\mathcal{P}_{ac}(\mathbb{R}^d)$ form a *displacement convex subset* of $\mathcal{P}(\mathbb{R}^d)$, with the remaining measures lying on its 'boundary'.

Remark 3.5 In order to verify the displacement convexity of a functional $W : \mathcal{P}_{ac}(\mathbb{R}^d) \longrightarrow \mathbb{R} \cup \{+\infty\}$ it is enough to show that for $\rho, \rho' \in \mathcal{P}_{ac}(\mathbb{R}^d)$, $W(\rho_t) \leq (1-t)W(\rho) + tW(\rho')$. For $\lambda = (1-s)t + st'$, Proposition 3.3(v) then implies that $W(\rho_\lambda) \leq (1-s)W(\rho_t) + sW(\rho_{t'})$.

In the next proposition, $\Lambda : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ denotes a translation, dilation, or orthogonal transformation of $\mathbb{R}^d$. In the usual way, the action of Λ on a measure $\rho \in \mathcal{P}(\mathbb{R}^d)$ is defined to be $\Lambda\rho := \Lambda_{\#}\rho$.

Proposition 3.6 *Let $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$ and $\rho' \in \mathcal{P}(\mathbb{R}^d)$ with displacement interpolant $\rho_t = \rho \xrightarrow{t} \rho'$. Denote by T_μ the translation $T_\mu(x) = x + \mu$ for $x, \mu \in \mathbb{R}^d$, and by S_λ the dilation $S_\lambda(x) = \lambda x$ on $\mathbb{R}^d$ by a factor $\lambda \geq 0$. Λ denotes either T_μ, S_λ or a member of the orthogonal group on $\mathbb{R}^d$. If $\nu \in \mathbb{R}^d$, $\alpha, \beta > 0$ and $s, t \in [0,1]$ then*

(i) $\Lambda\rho_t = \Lambda\rho \xrightarrow{t} \Lambda\rho'$ *;*

(ii) $T_{(1-t)\mu+t\nu}\,\rho_t = T_\mu\rho \xrightarrow{t} T_\nu\rho'$ *;*

(iii) $S_\lambda\rho_t = S_\alpha\rho \xrightarrow{s} S_\beta\rho'$ *if* $\lambda(1-t) = \alpha(1-s)$ *and* $\lambda t = \beta s$.

Example 3.7 (Translates and Dilates) In the trivial case $\rho' = \rho$, the convex function ψ may be taken to be $\psi(x) = x^2/2$ since $\nabla\psi = id$ pushes forward ρ to itself. The displacement interpolant is $\rho_t = \rho$ independent of t. Having made this observation, Proposition 3.6(ii) shows that for $\rho' = T_\nu \rho$ a translate of ρ, the displacement interpolant is $\rho_t = T_{t\nu}\rho$. For a dilate $\rho' = S_\beta \rho$, the displacement interpolant is $\rho_t = S_\lambda \rho$ with $\lambda = (1-t) + t\beta$.

Example 3.8 (Gaussian Measures) Let $\rho_0, \rho_1 \in \mathcal{P}_{ac}(\mathbb{R}^d)$ be Gaussian measures. At time $t \in (0,1)$ the displacement interpolant ρ_t will also be a Gaussian; its mean and covariance interpolate between those of ρ_0 and ρ_1. More specifically, let ρ_i be centered at $\mu_i \in \mathbb{R}^d$ $(i = 0, 1)$ and denote its covariance by Σ_i. Then Σ_i is the $d \times d$ matrix whose entries are

$$\int_{\mathbb{R}^d} x_j x_k \, d\rho_i(x) \qquad j, k = 1 \ldots d;$$

it is a *positive* matrix $\Sigma_i > 0$, meaning positive definite and self-adjoint. It suffices to find ρ_t when $\mu_0 = \mu_1 = 0$, since Proposition 3.6(ii) shows that the general interpolant is then obtained by translating ρ_t to $(1-t)\mu_0 + t\mu_1$. By the change of variables theorem (8), the push-forward of a Gaussian ρ_0 through a linear transformation Λ yields another Gaussian with covariance $\Lambda\Sigma_0\Lambda^\dagger$. For the transformation Λ to be the gradient of a convex function, it is necessary and sufficient that $\Lambda \geq 0$ be matrix positive. Although the matrix equation $\Lambda\Sigma_0\Lambda^\dagger = \Sigma_1$ has many solutions, the uniqueness part of Theorem B.6 shows that only one can be positive; it is $\Lambda = \Sigma_1^{1/2}(\Sigma_1^{1/2}\Sigma_0\Sigma_1^{1/2})^{-1/2}\Sigma_1^{1/2}$, computed previously in a related context [14]. Here $\Sigma^{1/2}$ denotes the positive square root of Σ. By uniqueness, ρ_t must be the Gaussian $[(1-t)id + t\Lambda]_\# \rho_0$.

Remark 3.9 (Singular Measures) A displacement interpolant may still be defined even if neither of the endpoints $\rho, \rho' \in \mathcal{P}(\mathbb{R}^d)$ is in $\mathcal{P}_{ac}(\mathbb{R}^d)$. Let ψ be a convex function on $\mathbb{R}^d$. As a subset of $\mathbb{R}^d \times \mathbb{R}^d$, the graph of $\nabla\psi$

is characterized by a property (118) known as *cyclical monotonicity*. Here $\langle\ ,\ \rangle$ denotes the Euclidean inner product, so the two-point inequality

$$\langle \nabla\psi(x) - \nabla\psi(y),\, x - y\rangle \geq 0 \tag{11}$$

has a clear geometrical interpretation: it states that the directions of the displacement vectors between x and y and between their images under $\nabla\psi$ differ by no more than $90°$; on the line this reduces to monotonicity. Corollary B.2 provides a joint probability measure $p \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ with cyclically monotone support having ρ and ρ' as its marginals. Let $t \in [0,1]$ and define

$$\Pi_t(x,y) := (1-t)x + ty \tag{12}$$

on $\mathbb{R}^d \times \mathbb{R}^d$. Then $\rho_t := \Pi_{t\#}p$. Corollary B.4 shows that this definition coincides with (9) when one of the endpoints is absolutely continuous; if neither is absolutely continuous, Remark B.7 shows ρ_t need not be unique. Even when uniqueness fails, Propositions 3.2, 3.3(iv)-(v) and 3.6 will continue to hold along each of the non-unique paths.

Remark 3.10 (Continuity) Although it is not required here, Lemma B.3 and the uniqueness part of Theorem B.6 combine to show continuity of the map from $\mathcal{P}_{ac}(\mathbb{R}^d) \times [0,1] \times \mathcal{P}(\mathbb{R}^d)$ to $\mathcal{P}(\mathbb{R}^d)$ taking (ρ, t, ρ') to $\rho \xrightarrow{t} \rho'$. As in Chapter 5, the measure spaces are topologized using convergence against $C_\infty(\mathbb{R}^d)$ test functions.

Remark 3.11 For comparison with the convexity of Proposition 3.2, we note that the potential energy $G(\rho)$, restricted to $\mathcal{P}(\mathbb{R}^d)$, may well be concave in the usual sense: $G(\,(1-t)\,\rho + t\,\rho'\,) \geq (1-t)\,G(\rho) + t\,G(\rho')$. For example, take $V(x) = |x|^2$. To see concavity of $G(\rho) = Q(\rho,\rho)$ on $\mathcal{P}(\mathbb{R}^d)$, expand

$$G(\,(1-t)\,\rho + t\,\rho'\,) = (1-t)^2\, G(\rho) + 2(1-t)t\, Q(\rho,\rho') + t^2\, G(\rho')$$

where $Q(\rho, \rho')$ is the implicit quadratic form, and examine the coefficient of t^2. Non-positivity of this coefficient is transparent in the following probablistic setting. Consider indepedent random variables $x, \tilde{x}, y, \tilde{y} : X \longrightarrow \mathbb{R}^d$ on some measure space $(X, \mathcal{X})$ with probability measure ω. Assume $\rho = x_{\#}\omega = \tilde{x}_{\#}\omega$ and $\rho' = y_{\#}\omega = \tilde{y}_{\#}\omega$, so that $Q(\rho, \rho') = \mathbf{E}[|x-y|^2]$ where $\mathbf{E}[\,\cdot\,]$ denotes expectation. Concavity of $G(\rho)$ is equivalent to

$$\mathbf{E}\left[|x-y|^2\right] \geq \mathbf{E}\left[|x-\tilde{x}|^2\right]/2 + \mathbf{E}\left[|y-\tilde{y}|^2\right]/2. \tag{13}$$

Taking expectations, (13) is immediate from the identities

$$\begin{aligned} |x-\tilde{y}|^2 + |y-\tilde{x}|^2 &= |x-\tilde{x}|^2 + |y-\tilde{y}|^2 + 2\langle\, x-y,\, \tilde{x}-\tilde{y}\rangle \quad \text{and} \\ \mathbf{E}[\langle\, x-y,\, \tilde{x}-\tilde{y}\rangle] &= \langle\, \mathbf{E}[x-y]\,, \mathbf{E}[\tilde{x}-\tilde{y}]\rangle = |\mathbf{E}[x-y]\,|^2 \geq 0. \end{aligned}$$

Proof of Proposition 3.3: Let ψ be convex with $\rho' = \nabla\psi_{\#}\rho$. Then (i) is obvious. Theorem B.6 shows $\nabla\psi$ is uniquely determined ρ-a.e., which implies the uniqueness (ii) of ρ_t. To see (iii), let $\phi(x) := (1-t)\,x^2/2 + t\,\psi(x)$ denote the function whose gradient pushes forward ρ to ρ_t. The claim is that if $M \subset \mathbb{R}^d$ is (a Borel set) of Lebesgue measure zero, so is $(\nabla\phi)^{-1}(M)$; ρ_t then vanishes on the former because ρ vanishes on the latter. Convexity of ψ implies strict convexity of ϕ, so that $(\nabla\phi)^{-1}$ must be a single-valued function on its domain. Moreover, since

$$\begin{aligned} |\nabla\phi(x) - \nabla\phi(y)|\,|x-y| &\geq \langle\, \nabla\phi(x) - \nabla\phi(y),\, x-y\rangle \\ &= (1-t)\,|x-y|^2 + t\langle\, \nabla\psi(x) - \nabla\psi(y),\, x-y\rangle, \end{aligned}$$

(11) shows that $(\nabla\phi)^{-1}$ is Lipshitz with constant no greater than $(1-t)^{-1}$. (iii) is then a consequence of a standard measure theoretic result [15] stating that $\mathrm{vol}\,(\nabla\phi)^{-1}M \leq (1-t)^{-d}\mathrm{vol}\,M$.

The alternative definition of ρ_t given in Remark 3.9 provides the easiest way to see (iv). Let $p \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ be the joint probability measure with cyclically monotone support and ρ and ρ' as its marginals, i.e. connecting ρ to ρ' in the sense of Corollary B.2. Let $\Pi_t(x,y)$ the map (12) pushing p forward to $\rho \stackrel{t}{\rightarrow} \rho'$. If $*$ denotes the involution $*(x,y) = (y,x)$ on $\mathbb{R}^d \times \mathbb{R}^d$, then $*_\# p$ connects ρ' to ρ and is pushed forward to $\rho' \stackrel{1-t}{\rightarrow} \rho$ by Π_{1-t}. Since $\Pi_{1-t}(y,x) = \Pi_t(x,y)$, (iv) is proved.

Finally, (iv) is used along with the special case

$$\rho \stackrel{st}{\rightarrow} \rho' \; = \; \rho \stackrel{s}{\rightarrow} \rho_t \tag{14}$$

to prove (v). ϕ as above satisfying $\rho_t = \nabla\phi_\#\rho$, is used to define $\rho \stackrel{s}{\rightarrow} \rho_t$; (14) follows from $(1-s)\,id + s\,\nabla\phi = (1-st)\,id + st\,\nabla\psi$. Now let $\lambda = (1-s)t + st'$, and noting (iv) take $t' \leq t$ without loss of generality. Then (14) and (iv) imply $\rho_{t'} = \rho \stackrel{t'/t}{\rightarrow} \rho_t = \rho_t \stackrel{1-t'/t}{\rightarrow} \rho$ and also $\rho_\lambda = \rho \stackrel{\lambda/t}{\rightarrow} \rho_t = \rho_t \stackrel{1-\lambda/t}{\rightarrow} \rho$. Since $(t-\lambda)/(t-t') = s \leq 1$, (14) once more yields $\rho_\lambda = \rho_t \stackrel{s}{\rightarrow} \rho_{t'}$. QED.

Proof of Proposition 3.6: Remark 3.9 gives a definition of ρ_t through the measure $p \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ with cyclically monotone support and ρ and ρ' as marginals. The relevant observation is that a cyclically monotone subset of $\mathbb{R}^d \times \mathbb{R}^d$ remains cyclically monotone under any of the transformations $\Lambda \times \Lambda$, $T_\mu \times T_\nu$ or $S_\alpha \times S_\beta$. The result (i), (ii) or (iii) is then obtained by pushing p forward through one of these transformations: the push-forward has cyclically monotone support, and the correct marginals by the change of variables theorem (8). Defining $\Pi_t(x,y)$ as in (12), the results follow from

$$\begin{aligned} \Lambda\Pi_t(x,y) &= \Pi_t(\Lambda x, \Lambda y), \\ (1-t)\,\mu + t\,\nu + \Pi_t(x,y) &= \Pi_t(x+\mu, y+\nu), \text{ and} \\ \lambda\Pi_t(x,y) &= \Pi_s(\alpha x, \beta y). \end{aligned}$$

QED.

4 Displacement Convexity of $\int A(\rho)$

In the sequel it is shown that for suitable convex functions $A(\varrho)$, the functional

$$U(\rho) := \int_{\mathbb{R}^d} A(\rho(x))\, dx \tag{15}$$

will be displacement convex on $\mathcal{P}_{ac}(\mathbb{R}^d)$, i.e. $U(\rho_t)$ will be a convex function of t along the path of the displacement interpolation $\rho \xrightarrow{t} \rho'$. For the $L^q(\mathbb{R}^d)$ norm rather more can be said: $\|\rho_t\|_q^{-q'/d}$ is concave provided $q^{-1} + q'^{-1} = 1$, and linear when ρ and ρ' are dilates. The Brunn-Minkowski inequality is recovered as a special case of this result.

To any $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$ is associated the family of dilates $S_\lambda \rho$ which may be obtained as the push-forward of ρ through dilation of $\mathbb{R}^d$ by some factor $\lambda > 0$. The condition for displacement convexity of $U(\rho)$ is merely this: $U(S_\lambda \rho)$ should be convex non-increasing as a function of λ. The necessity of the convexity is obvious; its sufficiency is the content of Theorem 4.2. The hypothesis is also physically reasonable: as a gas expands, its internal energy must certainly decrease; it should vanish as $\lambda \to \infty$ and diverge as $\lambda \to 0$. In terms of $A : [0, \infty) \longrightarrow \mathbb{R} \cup \{+\infty\}$, the condition is:

(A1) $\lambda^d A(\lambda^{-d})$ be convex non-increasing on $\lambda \in (0, \infty)$, with $A(0) = 0$.

Having made this assumption, the displacement convexity of $U(\rho)$ hinges on two observations. Consider mass m of a gas whose internal energy is given by (15). If the gas is uniformly distributed uniformly throughout a box of volume v, $U = A(m/v)v$. Imagine then that the side lengths of the (d-dimensional) box are varied linearly with time, so that the volume, density, and internal energy $U(t)$ become functions of time. The first observation is that $U(t)$ is a convex function of time. The second observation is that the form of the displacement interpolation makes it possible to use convexity of $U(t)$ locally to obtain a global inequality — Theorem 4.2. The idea of the interpolation is to transfer the small mass of gas with approximately constant density $\rho(x)$ from a neighbourhood of x to a neighbourhood of $\nabla\psi(x)$. The mapping $\nabla\psi$ may be linearly approximated almost everywhere by its derivative $\nabla^2\psi(x)$, a non-negative matrix. Choosing a basis which diagonalizes $\nabla^2\psi(x)$ and replacing the neighbourhood of x by a small cube, the t-linearity of the interpolating map $(1-t)id+t\nabla\psi$ throws us back to the first observation. The monotonicity condition (11) ensures that two disjoint cubes initially at points x and y will not interfere with each other during subsequent motion for $t \in [0,1]$.

The proof is facilitated by an elementary lemma:

Lemma 4.1 *Suppose* $h : (0,\infty) \longrightarrow \mathbb{R} \cup \{+\infty\}$ *is convex non-increasing while* $g : [0,1] \longrightarrow (0,\infty)$ *is concave. Then the composition* $h \circ g$ *is convex. Except on intervals where* $g(t)$ *is constant,* $h \circ g$ *will be strictly convex if* h *is.*

Proof: For the analysis of cases of equality, assume the convexity — hence the monotonicity — of h to be strict. Let $s, t, t' \in [0,1]$. Then

$$\begin{aligned} g\Big((1-s)t+st'\Big) &\geq (1-s)\,g(t)+s\,g(t') \\ \text{so} \qquad h\circ g\left((1-s)t+st'\right) &\leq h\Big((1-s)\,g(t)+s\,g(t')\Big) \\ &\leq (1-s)\,h\circ g\,(t)+s\,h\circ g\,(t'). \end{aligned}$$

The first two inequalities are strict unless g is affine on $[t,t']$, while the last is strict unless $g(t)=g(t')$. QED.

Theorem 4.2 (Displacement Convexity of Internal Energy $U(\rho)$) *Let* $\rho, \rho' \in \mathcal{P}_{ac}(\mathbb{R}^d)$*, and define the displacement interpolant* $\rho_t = \rho \stackrel{t}{\rightarrow} \rho'$ *using the convex function* ψ *for which* $\nabla\psi_{\#}\rho = \rho'$*. Assuming (A1), the internal energy* $U(\rho_t)$ *will be a convex function of* t *on* $[0,1]$*. Strict convexity follows* *from that of* $\lambda^d A(\lambda^{-d})$ *unless* $\nabla^2\psi(x) = I$ *holds* ρ*-a.e.*[3] *In the latter case* $U(\rho_t)$ *will be* t*-independent.*

Proof: Proposition 3.3(iii) shows that ρ_t is absolutely continuous (with respect to Lebesgue). By the Monotone Change of Variables Theorem 6.4, the set X on which $\nabla^2\psi(x) > 0$ and its inverse exist has full measure for ρ, and moreover

$$U(\rho_t) = \int_X A\left(\frac{\rho(x)}{\det\left[(1-t)I + t\nabla^2\psi(x)\right]}\right) \det\left[(1-t)I + t\nabla^2\psi(x)\right]\, dx. \quad (16)$$

Actually, for $t < 1$, one should integrate over all points at which $\nabla^2\psi(x)$ exists, but the distinction is moot because $A(0) = 0$ and X is full measure for ρ. Fix $x \in X$, letting $\Lambda := \nabla^2\psi(x)$ and $v(t) := \det\left[(1-t)I + t\Lambda\right]$. Since Λ is matrix positive, $v^{1/d}(t)$ is known to be a concave function of t, strictly concave unless Λ is a multiple of the identity. This is the kernel of any proof of the Brunn-Minkowski inequality [15, 16]: in the basis diagonalizing Λ, concavity may be seen to result from the domination of the geometric by the arithmetic mean. With $g(t) := v^{1/d}(t)$ and $h(\lambda) := \lambda^d A\big(\rho(x)/\lambda^d\big)$, Lemma 4.1 shows the integrand of (16) to be a convex function of t, strictly convex (unless $v(t)$ is constant) if the convexity in (A1) is strict. $v(t)$ constant implies $\nabla^2\psi(x) = I$. Integrating proves the result. QED.

Remark 4.3 In the proof of Theorem 4.2, a crucial role is played by the fact that the transfer of mass defining the displacement interpolation is *irrotational*; the matrix positivity of $\nabla^2\psi(x)$ may be interpreted to mean that under $\nabla\psi$, small neighbourhoods of x are stretched or shrunk in various directions, but never turned or inverted. For comparison, consider a mass transfer algorithm in which a rotation by $\pi/2$ radians replaces $\nabla\psi$. Now

[3] The Hessian $\nabla^2\psi$ is defined as in (120).

$U(0) = U(1)$ for a global rotation; but for a transfer which is linear in time the volume of any small set in spt ρ has been reduced by a factor $1/\sqrt{2}$ at time $t = 1/2$. The contribution to $U(1/2)$ is increased accordingly. Obviously, $U(1/2) > U(0) = U(1)$ is not compatible with convexity. The same argument applies locally to a transfer procedure which need not be a global rotation.

In the special case $A(\varrho) = \varrho^q$, a scaling argument strengthens Theorem 4.2 considerably. For the displacement interpolation, convexity of the $L^q(\mathbb{R}^d)$ norm turns out to be better than *logarithmic*: if q' is Hölder conjugate to q then $\|\rho\|_q^{-q'/d}$ is displacement concave. This inequality is sharp in the sense that $\|S_\lambda \rho\|_q^{-q'/d}$ depends linearly on $\lambda > 0$ for a mass-preserving dilation of ρ. The formulation of the next two remarks has been selected to facilitate comparison with inequalities from [11], recapitulated here in Corollary D.4.

Corollary 4.4 (Logarithmic Convexity of the $L^q(\mathbb{R}^d)$ Norm)
Let $\rho_t = \rho \xrightarrow{t} \rho'$ be the displacement interpolant between $\rho, \rho' \in \mathcal{P}_{ac}(\mathbb{R}^d)$. Let $0 < q \leq \infty$ satisfy $q \geq 1 - 1/d$ and define $\alpha := -(1 - 1/q)^{-1}/d$. Then

$$\|\rho_t\|_q^\alpha \geq (1-t)\,\|\rho\|_q^\alpha + t\,\|\rho'\|_q^\alpha. \tag{17}$$

As a result, $\log \|\rho_t\|_q$ is convex on $t \in [0,1]$ for $q > 1$ and concave for $q < 1$.

Proof: Unless $q \neq 1$ and $t \in (0,1)$, the assertion is vacuous. To begin, assume $q > 1$ and $\rho, \rho' \in L^q(\mathbb{R}^d)$. Letting S_λ denote dilation by $\lambda > 0$, it is possible to choose factors $\lambda, \lambda' > 0$ such that $\|S_\lambda \rho\|_q = \|S_{\lambda'} \rho'\|_q$ and $(1-t)/\lambda + t/\lambda' = 1$. Setting $s = t/\lambda' \in (0,1)$, Proposition 3.6(iii) shows that $\rho_t = S_\lambda \rho \xrightarrow{s} S_{\lambda'} \rho'$. Because $A(\varrho) = \varrho^q$ satisfies (A1), Theorem 4.2 shows $\|\rho_t\|_q^q$ to be convex as long as $q < \infty$:

$$\begin{aligned} \|\rho_t\|_q^q &\leq (1-s)\|S_\lambda \rho\|_q^q + s\|S_{\lambda'}\rho'\|_q^q && (18)\\ &= \|S_\lambda \rho\|_q^q. && (19) \end{aligned}$$

Since t^α is decreasing for $q > 1$,

$$\|\rho_t\|_q^\alpha \geq \|S_\lambda \rho\|_q^\alpha \tag{20}$$

$$= (1-s)\|S_\lambda \rho\|_q^\alpha + s\|S_{\lambda'} \rho'\|_q^\alpha. \tag{21}$$

In the case $q = \infty$, (20) follows immediately from Theorem 4.2 with $A(\varrho) = 0$ where $\varrho \leq \|S_\lambda \rho\|_\infty$ and $A(\varrho) = \infty$ otherwise. Either way, the case $q > 1$ is established for $\rho, \rho' \in L^q(\mathbb{R}^d)$ by the scaling relation $\|S_\lambda \rho\|_q = \lambda^{1/\alpha}\|\rho\|_q$ in (21). If $\|\rho'\|_q = \infty$, a separate argument is required: $\|\rho_t\|_q \leq \|S_{1-t}\,\rho\|_q = (1-t)^{1/\alpha}\|\rho\|_q$ follows directly from (16), $\det\,[(1-t)I + t\nabla^2\psi] \geq (1-t)^d$ and monotonicity in (A1). When $q < 1$ the argument is the same, except that the inequality in (18) is reversed because it is $A(\varrho) = -\varrho^q$ which satisfies (A1); on the other hand, the inequality in (20) is restored because $\alpha > 0$. Taking the logarithm of (17), the convexity or concavity of $\log \|\rho_t\|_q$ follows according to the sign of α. Remark 3.5 has been noted. QED.

Remark 4.5 (Brunn-Minkowski Inequality)
This classical geometric inequality [17] compares the Lebesgue measures of two sets $K, K' \subset \mathbb{R}^d$ and their vector sum $K+K' = \{k+k' \mid (k,k') \in K \times K'\}$; for non-empty sets, it states that

$$\mathrm{vol}^{1/d}[K+K'] \geq \mathrm{vol}^{1/d} K + \mathrm{vol}^{1/d} K'. \tag{22}$$

The inequality is recovered from Corollary 4.4 with $q \neq 1$.

Proof of (22): Assume K, K' are compact, since the general case will follow by regularity of Lebesgue measure; unless both sets are of positive measure, there is little to prove. (22) is equivalent to the concavity of $\mathrm{vol}^{1/d}(1-t)K + tK'$ on $[0,1]$. Therefore, let $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$ be the restriction of Lebesgue measure to K, normalized to have unit mass, and let ρ'

be the analogous measure on K'. Let ψ be the convex function for which $\nabla\psi_{\#}\rho = \rho'$; then $\nabla\psi(x) \in K'$ ρ-a.e. Therefore the *support* $\operatorname{spt}\rho_t$ of the displacement interpolant between ρ and ρ' must lie in the (compact) set $(1-t)K + tK'$. With $q^{-1} + q'^{-1} = 1$, Corollary 4.4 asserts that

$$\|\rho_t\|_q^{-q'/d} \geq (1-t)\|\rho\|_q^{-q'/d} + t\|\rho'\|_q^{-q'/d}. \tag{23}$$

By construction, $\operatorname{vol} K = \|\rho\|_q^{-q'}$ and the same holds for K'. Applied to $\rho_t \in \mathcal{P}_{ac}(\mathbb{R}^d)$, Jensen's inequality yields $\operatorname{vol}[\operatorname{spt}\rho_t] \geq \|\rho_t\|_q^{-q'}$ for $q < \infty$; this estimate is trivial when $q = \infty$. The theorem then follows from the inclusion

$$(1-t)K + tK' \supseteq \operatorname{spt}\rho_t. \tag{24}$$

QED.

Remark 4.6 It is interesting to note that the inclusion (24) will typically be strict; $\operatorname{spt}\rho_t$ interpolates more efficiently between K and K' than the Minkowski combination $(1-t)K + tK'$. As an example, take both K and K' to be ellipsoids — affine images of the unit ball. The same considerations as in Example 3.8 show the mass of the displacement interpolant ρ_t to be uniformly distributed over a third ellipsoid. On the other hand, $(1-t)K + tK'$ will not be an ellipsoid except in special cases, a fact which is easily appreciated when K is the unit ball and K' is highly eccentric (even degenerate).

Remark 4.7 (Prékopa-Leindler Theorems)
A theorem of Prékopa and Leindler [8, 9, 10] gives another generalization of the Brunn-Minkowski inequality to functions on $\mathbb{R}^d$. For non-negative measurable functions f, g on $\mathbb{R}^d$ and $t \in (0,1)$, it states that the interpolant

$$h(x) := \sup_{y \in \mathbb{R}^d} f\left(\frac{y}{1-t}\right)^{1-t} g\left(\frac{x-y}{t}\right)^t \tag{25}$$

satisfies the inequality $\|h\|_1 \geq \|f\|_1^{1-t}\|g\|_1^t$. By scaling, the case $\|f\|_1 = \|g\|_1 = 1$ is quite general. The displacement interpolant $f \xrightarrow{t} g \in \mathcal{P}_{ac}(\mathbb{R}^d)$ between f and g can then be defined, and the Prèkopa-Leindler theorem becomes a trivial consequence of the observation that $h \geq f \xrightarrow{t} g$: the inequality $\|h\|_1 \geq 1$ is saturated with the displacement interpolant in place of h! A stronger assertion is verified in Appendix D, and parlayed into a transparent proof of both the Prékopa-Leindler theorem and related inequalities due to Brascamp and Lieb [11].

5 Existence and Uniqueness of Ground State

Armed with the estimates of the two preceding sections, we return to the existence and uniqueness questions regarding the ground state of the attracting gas model described by (4). In this model, the configuration of the gas is given by its mass density $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$, and the interaction is through a convex potential V on $\mathbb{R}^d$. This leads to a potential energy

$$G(\rho) := \iint d\rho(x)V(x-y)d\rho(y). \tag{26}$$

Although V need not be spherically symmetric, it is clear from (26) or by Newton's third law that it may be taken to be *even*: V(x)=V(-x). Thus V is minimized at the origin, and $V(0) = 0$ without loss of generality.

The gas is also assumed to satisfy an equation of state $P(\rho)$ relating pressure to density, which leads to an internal energy $U(\rho)$ of the form (15). The local density $A(\varrho)$ of $U(\rho)$ is obtained by integrating $dU = -Pdv$:

$$A(\varrho) := \int_1^\infty P(\varrho/v)\, dv. \tag{27}$$

To be physical, the pressure $P(\varrho) \geq 0$ should be non-decreasing; we make the stronger assumptions

(P1) $P : [0,\infty) \longrightarrow [0,\infty]$ with $P(\varrho)/\varrho^{1-1/d}$ non-decreasing;

(P2) $P(\varrho)/\varrho^2$ not integrable at ∞.

For fixed λ, changing variables to $s = \lambda v^{1/d}$ in (27) shows the equivalence of (P1) to the convexity of $U(\rho)$ under dilations, (A1) of the previous section. Strict monotonicity in (P1) is equivalent to strict convexity in (A1). Thus $U(\rho)$ will be displacement convex. $A(\varrho)$ is also seen to be convex and lower semi-continuous. (P2) implies that $A(\varrho)/\varrho$ diverges with ϱ, and excludes the possibility that the energy minimizing measure might have a singular part with respect to Lebesgue. Under these assumptions, if $V(x)$ is strictly convex we show that the total energy $E(\rho) = U(\rho) + G(\rho) \geq 0$ attains a unique minimum up to translation, unless $E(\rho) = \infty$. If convexity of $V(x)$ is not strict, the conclusions will still hold provided the monotonicity in (P1) is strict and $V(x) = V(|x|)$ is spherically symmetric, not identically zero.

Uniqueness is proved by combining the displacement convexity of $G(\rho)$ and $U(\rho)$. Displacement convexity also plays a role in the existence proof, which relies on a compactness argument. Let $C_\infty(\mathbb{R}^d)$ be the Banach space of continuous functions vanishing at ∞, under the sup norm. By the Riesz-Markov Theorem, its dual $C_\infty(\mathbb{R}^d)^*$ consists of Borel measures of finite total variation. The relevant topology on $\mathcal{P}_{ac}(\mathbb{R}^d) \subset C_\infty(\mathbb{R}^d)^*$ will be the weak-$*$ topology, the topology of convergence against $C_\infty(\mathbb{R}^d)$ test functions.

Theorem 5.1 (Existence and Uniqueness of Ground State) *Assume (P1-P2) and $V : \mathbb{R}^d \longrightarrow \mathbb{R} \cup \{+\infty\}$ to be strictly convex. Let the energy $E(\rho) = U(\rho) + G(\rho)$ be given by (4) with $A(\varrho)$ from (27), and $E_g :=$ $\inf E(\rho)$ over $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$. If $E_g < \infty$, the infimum is uniquely attained up to translation. The minimizer ρ_g may be taken to be even: $\rho_g(x) = \rho_g(-x)$.*

Proof: Uniqueness is proven first: suppose two minimizers $\rho_g, \rho'_g \in \mathcal{P}_{ac}(\mathbb{R}^d)$ exist. Fix $t \in (0,1)$ and consider the displacement interpolant $\rho_t = \rho_g \overset{t}{\rightarrow} \rho'_g$ between them. Since $U(\rho)$ and $G(\rho)$ are displacement convex (Theorem 4.2 and Proposition 3.2), $E(\rho_t) \leq (1-t)E(\rho_g) + tE(\rho'_g) = E_g$. Strict inequality holds by Proposition 3.2 unless ρ_g is a translate of ρ'_g. Since no configuration can have energy less than E_g, uniqueness is established.

For the existence proof, replace $V(x)$ by $(V(x) + V(-x))/2$, adding a constant so the minimum $V(0) = 0$; the effect on $E(\rho)$ is a shift by the same constant. Noting that $E(\rho) \geq 0$, choose an energy minimizing sequence $\rho_n \in \mathcal{P}_{ac}(\mathbb{R}^d) \subset C_\infty(\mathbb{R}^d)^*$. By Lemmas 5.4 and 5.6 any weak-$*$ limit point $\rho_g \in \mathcal{P}_{ac}(\mathbb{R}^d)$ of ρ_n must minimize $E(\rho)$. In fact, Corollary 5.5 applies because of (P2), and shows that ρ_g need only satisfy the mass constraint $\rho_g \in \mathcal{P}(\mathbb{R}^d)$:

finiteness of E_g implies absolute continuity of ρ_g. The Banach-Alaoglu Theorem provides weak-$*$ compactness of the unit ball in $C_\infty(\mathbb{R}^d)^*$, but because $E(\rho)$ is translation invariant, precautions must be taken to ensure that no mass escapes to ∞.

Consider the reflection $\Lambda(x) := -x$ on $\mathbb{R}^d$. Propositions 3.6(i) and 3.3(iv) show the displacement interpolant $\rho \stackrel{1/2}{\to} \Lambda\rho$ to be invariant under Λ; it should be thought of as a symmetrization of ρ. Moreover $E(\Lambda\rho) = E(\rho)$, so by displacement convexity this symmetrization can only lower the energy of ρ. The minimizing sequence ρ_n may therefore be replaced by one for which $\rho_n(x) = \rho_n(-x)$. Extracting a subsequence if necessary, ρ_n may be taken to converge to a limit ρ_g weak-$*$. ρ_g is a positive Borel measure; it is even and has total mass no greater than unity.

Since V is strictly convex with minimum $V(0) = 0$, it is bounded away from zero on the unit sphere: $V(x) \geq k > 0$ for $|x| = 1$. For $|x| > 1$ convexity yields $V(x) > k|x|$, in which case

$$\int_{\mathbb{R}^d} V(x-y)d\rho_n(y) \;\geq\; \int_{\langle y,\,x\rangle \leq 0} V(x-y)d\rho_n(y) \quad\geq\quad k|x|/2$$

since half of the mass of ρ_n lies on either side of the hyperplane $\langle y,\, x\rangle = 0$. Integrating this inequality against $\rho_n(x)$ over $|x| > R > 1$ yields a lower bound

$$G(\rho_n) \geq \frac{kR}{2}\int_{|x|>R} d\rho_n(x). \tag{28}$$

$E(\rho_n)$ may be assumed to be bounded above, so $G(\rho_n) \leq L$. Thus (28) controls the mass of ρ_n outside of any large ball, uniformly in n. If $0 \leq \varphi \leq 1$ is a $C_\infty(\mathbb{R}^d)$ test function with $\varphi = 1$ on $|x| \leq R$, weak-$*$ convergence yields $\int \varphi \, d\rho_g \geq 1 - 2L/kR$. Since R was arbitrary, $\rho_g[\mathbb{R}^d] = 1$. QED.

In the event that the potential $V(x)$ is not strictly convex, it may yet be possible to prove existence of a unique energy minimizer using a more delicate argument. This will be true if the monotonicity in (P1) is strict and the convex potential $V(x) = V(|x|)$ is spherically symmetric but not identically zero. The existence argument of Theorem 5.1 requires only the slightest modification: $V(x)$ might vanish on $|x| = 1$, but it is non-zero on some sphere of finite radius. On the other hand, the uniqueness argument fails, because the displacement convexity of the interaction energy need not be strict. However, Theorem 5.3 shows that the condition for strict displacement convexity of the internal energy can be used instead to force two minimizers to be translates of each other. It is necessary to state a preliminary lemma regarding the decomposition of convex functions on $\mathbb{R}^d$.

Lemma 5.2 *Let ψ and ϕ be convex functions on $\mathbb{R}^d$, and $\Omega \subset \mathbb{R}^d$ an open convex set on which both ψ and ϕ are finite. Suppose ϕ is differentiable on Ω with locally Lipschitz derivative $\nabla\phi : \Omega \longrightarrow \mathbb{R}^d$. If the Hessians $\nabla^2\phi = \nabla^2\psi$ agree almost everywhere there, then $\psi - \phi$ is convex on Ω.*

Proof: First, consider functions on the line $d = 1$. ψ may be viewed as a distribution on $\Omega \subset \mathbb{R}$; its convexity is characterized by the fact that its distributional second derivative is a positive Radon measure ω on Ω. Lebesgue decompose $\omega = \omega_{ac} + \omega_{sing}$. Integrating ω_{ac} twice from some base point in Ω yields a differentiable convex function v. Its derivative v' is a monotone function, absolutely continuous on compact subsets, hence v'' exists and coincides with ω_{ac} both pointwise almost everywhere and in the distributional sense. ϕ'' also coincides with ω_{ac}, thus $v' - \phi'$ — being absolutely continuous — is constant, and $v - \phi$ is affine. On the other hand, $\psi - v$ is convex since its distributional second derivative is $\omega_{sing} \geq 0$. Thus $\psi - \phi$ is convex.

The higher dimensional case $d > 1$ is reduced to the case $d = 1$ as follows. Suppose convexity of $\psi - \phi$ were violated along some line segment with endpoints $x', y' \in \Omega$. Continuity of ψ and ϕ shows that convexity is also violated along any line segment with endpoints x and y sufficiently close to x' and y'. Since $\nabla^2\psi$ and $\nabla^2\phi$ exist and coincide almost everywhere on Ω, Fubini's Theorem shows that for some such x and y, $\nabla^2\psi = \nabla^2\phi$ almost everywhere along $(1-t)x+ty$ (with respect to the one dimensional Lebesgue measure). Viewing ψ and ϕ as functions of $t \in [0,1]$ along this segment, their second derivatives are determined by $y-x$ and the Hessians $\nabla^2\psi$ and $\nabla^2\phi$ wherever the latter exist. A contradiction with the $d=1$ result would be reached, forcing the conclusion that convexity of $\psi - \phi$ cannot be violated. QED.

Theorem 5.3 (Uniqueness Without Strict Convexity of V(x))
Assume that $P(\varrho)$ satisfies the monotonicity condition (P1) strictly, and that the convex function $V : \mathbb{R}^d \longrightarrow \mathbb{R} \cup \{+\infty\}$ is spherically symmetric, not constant. If the energy $E(\rho) = U(\rho) + G(\rho)$ given by (4) with $A(\varrho)$ from (27) attains a finite minimum at $\rho_g \in \mathcal{P}_{ac}(\mathbb{R}^d)$, then ρ_g is unique up to translation.

Proof: Denote by ρ_g^* the symmetric decreasing rearrangement of ρ_g: that is, the spherically symmetric, radially non-increasing function satisfying

$$\mathrm{vol}\,\{\rho_g^* > k\} = \mathrm{vol}\,\{\rho_g > k\} \tag{29}$$

for all $k > 0$. The internal energy $U(\rho_g^*) = U(\rho_g)$ by (29), while a rearrangement inequality due to Riesz [18] states that $G(\rho_g^*) \leq G(\rho_g)$ since the potential $V(x)$ is symmetric non-decreasing. Thus ρ_g^* also minimizes $E(\rho)$. Suppose $\rho_g' \in \mathcal{P}_{ac}(\mathbb{R}^d)$ is another energy minimizer, and define the displacement interpolant ρ_t between ρ_g^* and ρ_g' via the convex ψ for which $\nabla\psi_{\#}\rho_g^* = \rho_g'$. $U(\rho)$ and $G(\rho)$ are displacement convex as before. Since strict convexity of $U(\rho_t)$ would imply a contradiction, Theorem 4.2 shows that $U(\rho_g^*) = U(\rho_g')$ and $\nabla^2\psi(x) = I$ a.e. on the interior Ω of spt ρ_g^*. Lemma 5.2 shows that $\psi(x) - x^2/2$ must be convex on Ω. Unless $\psi(x) - x^2/2$ is affine on this ball — so that ρ_g' is a translate of ρ_g^* — we show $G(\rho_g^*) < G(\rho_g')$, a contradiction. If $\nabla\psi$ exists at $x, y \in \mathbb{R}^d$, then the monotonicity inequality (11) will be strict unless $\nabla\psi(x) = \nabla\psi(y)$. Applied to the function $\psi(x) - x^2/2$ rather than $\psi(x)$, this shows that if $\nabla\psi(x) - x \neq \nabla\psi(y) - y$, then

$$|\nabla\psi(x) - \nabla\psi(y)| > |x - y|. \tag{30}$$

Unless $\psi(x) - x^2/2$ is affine on Ω, (30) will be satisfied at some $x, y \in \Omega$. By the continuity properties of $\nabla\psi$, (30) will continue to hold in a small neighbourhood of $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ — which is to say, on a set of positive measure $\rho_g^* \times \rho_g^*$. The change of variables formula (8) yields

$$G(\rho_g') = \iint d\rho_g^*(x)\; V(\nabla\psi(x) - \nabla\psi(y))\; d\rho_g^*(y).$$

If the convex potential $V(x)$ assumes a unique minimum at $x = 0$ — so that it is strictly attractive — then $V(\nabla\psi(x) - \nabla\psi(y)) > V(x - y)$ wherever (30) holds. The contradiction $G(\rho_g') > G(\rho_g^*)$, and therefore the theorem, is established in this case.

The remaining case — $V(x)$ constant on a ball of radius r about 0 — requires an additional argument. Take $V(0) = 0$. If $V(x) = \infty$ for $|x| > r$, all of the mass of the minimizer must lie in a set of diameter r; in fact it must be uniformly distributed over a ball of diameter r by Jensen's inequality and the isodiametric inequality [15]. This case aside, it is necessary to show that the diameter of Ω is greater than r; then the argument of the preceding paragraph will apply: unless $\nabla\psi$ is affine, it will be possible to choose $x, y \in \Omega$ with $|x| > r/2$ and $y = -x$ to satisfy (30), and $V(\nabla\psi(x) - \nabla\psi(y)) > V(x - y)$ will hold on a neighbourhood of (x, y). The possibility that $\Omega \subset B_{r/2}(0)$ is precluded by contradiction. Assume $G(\rho_g^*) = 0$, and consider the dilation $S_\lambda\rho_g^*$ of ρ_g^* by factor $\lambda \geq 1$. Defining $G(\lambda) := G(S_\lambda\rho_g^*)$, it will be shown that $G(\lambda)$ grows sublinearly with small $\lambda - 1$ while $U(\lambda) := U(S_\lambda\rho_g^*)$ decreases linearly; the contradiction is obtained since ρ_g^* is allegedly a minimizer. In fact, $G(\lambda) = o(\lambda - 1)^2$ as $\lambda \to 1^+$. To see this, note that for $\lambda \geq 1$ the only contribution to $G(\lambda)$ comes from the self-interaction of the mass $m(\lambda)$ lying within a spherical shell of thickness $(\lambda - 1)r$ around the surface $|x| = r/2$. Since ρ_g^* is symmetric decreasing, its density is bounded except at the origin; thus $m(\lambda) \leq k(\lambda - 1)$. By continuity of $V(x)$, λ near 1 ensures $V(x) < \epsilon$ for $x < \lambda r$, implying $G(\lambda) \leq \epsilon m^2(\lambda)$. Certainly $G'(1^+) = 0$. On the other hand, strict convexity of the decreasing function $U(\lambda)$ follows from strict convexity in (A1) or strict monotonicity in (P1). Thus $U'(1^+) < 0$. In combination, these estimates preclude $G(\rho_g^*) = 0$, and conclude the proof. QED.

Lemma 5.4 (Weak-$*$ Lower Semi-Continuity of $U(\rho)$)
Assume $A : [0, \infty) \longrightarrow [0, \infty]$ is convex and lower semi-continuous. Define $U(\rho)$ by (15). Then $U(\rho)$ is weak-$*$ lower semi-continuous on $\mathcal{P}_{ac}(\mathbb{R}^d) \subset C_\infty(\mathbb{R}^d)^*$.

Proof: Let $\rho_n \longrightarrow \rho$ weak-$*$ in $\mathcal{P}_{ac}(\mathbb{R}^d)$. The claim is that $\underline{\lim}_n U(\rho_n) \geq U(\rho)$. Let $\varphi \geq 0$ be a continuous (spherically symmetric) function of compact support such that $\int \varphi = 1$. Convolving with the mollifier $\varphi_\epsilon(x) := \epsilon^{-d}\varphi(x/\epsilon) \in C_\infty(\mathbb{R}^d)$, one has pointwise convergence of $\rho_n * \varphi_\epsilon$ to $\rho * \varphi_\epsilon$ as $n \to \infty$. Jensen's inequality with the convex function $A(\varrho)$ yields

$$\int A(\rho(y))\varphi_\epsilon(x-y)\, dy \geq A\left(\int \rho(y)\varphi_\epsilon(x-y)\, dy\right).$$

$U(\rho) \geq U(\rho * \varphi_\epsilon)$ is obtained by integrating over $x \in \mathbb{R}^d$. For fixed $\epsilon > 0$,

$$\begin{aligned} \underline{\lim}_n U(\rho_n) &\geq \underline{\lim}_n\, U(\rho_n * \varphi_\epsilon) \\ &\geq \int \underline{\lim}_n A(\rho_n * \varphi_\epsilon) \\ &\geq \int A(\rho * \varphi_\epsilon) \\ &= U(\rho * \varphi_\epsilon). \end{aligned}$$

The second inequality is Fatou's Lemma while the third is the lower semi-continuity of $A(\varrho)$. At the Lebesgue points of ρ, hence almost everywhere, it can be shown that $\rho * \varphi_\epsilon \longrightarrow \rho$ as $\epsilon \to 0$. Another application of Fatou's Lemma and the lower semi-continuity of $A(\varrho)$ yields $\underline{\lim}_\epsilon U(\rho * \varphi_\epsilon) \geq U(\rho)$.
QED.

Corollary 5.5 *In addition to the hypotheses of Lemma 5.4, suppose $A(\varrho)/\varrho$ diverges as $\varrho \to \infty$. Then $U(\rho)$ remains weak-$*$ lower semi-continuous if it is extended to $\mathcal{P}(\mathbb{R}^d) \subset C_\infty(\mathbb{R}^d)^*$ by taking $U(\rho) = \infty$ unless $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$.*

Proof: The only case to check is that $\underline{\lim}_n U(\rho_n) = \infty$ when a sequence $\rho_n \in \mathcal{P}_{ac}(\mathbb{R}^d)$ tends to a limit $\rho \in \mathcal{P}(\mathbb{R}^d)$ not absolutely continuous. Lebesgue decompose $\rho = \rho_{ac} + \rho_{sing}$. The singular part ρ_{sing} is a positive Borel measure with finite mass. By regularity, there is a compact set K and $m > 0$ such that $\rho_{sing}[K] > m$ but $\mathrm{vol}\, K = 0$, and an open set $N \supset K$ with arbitrarily small Lebesgue measure. Choose a $C_\infty(\mathbb{R}^d)$ test function $0 \leq \varphi \leq 1$ vanishing outside N with $\varphi = 1$ on K. For n large $\rho_n[N] > m$ and Jensen's inequality together with the monotonicity of $A(\varrho)$ yields

$$\int_N A(\rho_n) \geq A\left(\frac{m}{\mathrm{vol}\,[N]}\right) \mathrm{vol}\,[N].$$

Since $A(\varrho)/\varrho$ diverges, starting with $\mathrm{vol}\,[N]$ very small forces $U(\rho_n) \longrightarrow \infty$ with n.
QED.

Lemma 5.6 (Weak-$*$ Lower Semi-Continuity of $G(\rho)$)
Assume $V : \mathbb{R}^d \longrightarrow [0, \infty]$ convex and define $G(\rho)$ by (26). Then $G(\rho)$ is weak-$*$ lower semi-continuous on $\mathcal{P}_{ac}(\mathbb{R}^d) \subset C_\infty(\mathbb{R}^d)^*$.

Proof: Let $\rho_n \longrightarrow \rho$ weak-$*$ in $\mathcal{P}_{ac}(\mathbb{R}^d)$. The claim is that $G(\rho) \leq \underline{\lim}_n G(\rho_n)$. Certainly the product measure $\rho_n \times \rho_n \longrightarrow \rho \times \rho$ weak-$*$ in $C_\infty(\mathbb{R}^d \times \mathbb{R}^d)^*$.

Being convex, $V(x,y) := V(x-y)$ agrees with a lower semi-continuous function except on a set of measure zero. Although $V(x,y)$ is not $C_\infty(\mathbb{R}^d \times \mathbb{R}^d)$, it can therefore be approximated pointwise a.e. by an increasing sequence of positive functions $V_r(x,y)$ which are. Define $G_r(\rho)$ analogously to $G(\rho)$ but with V_r replacing V. For fixed r, $G_r(\rho) = \lim_n G_r(\rho_n) \leq \underline{\lim}_n G(\rho_n)$. By Lebesgue's Monotone Convergence Theorem, $G_r(\rho)$ increases to $G(\rho)$ and the result is proved. QED.

6 Monotone Change of Variables Theorem

Let ψ be a convex function on $\mathbb{R}^d$, and denote the interior of the convex set $\{\psi < \infty\}$ by $\Omega := \operatorname{int}\operatorname{dom}\psi$. As the gradient of a convex function, $\nabla\psi : \Omega \longrightarrow \mathbb{R}^d$ represents the generalization to higher dimensions of a monotone map on the interval. It is a measurable transformation, defined and differentiable[4] almost everywhere, and will be casually referred to as a monotone map. Sundry notions related to ψ, including the subgradient $\partial\psi$ and the Legendre transform ψ^* may be found in Appendix A. The goal of the current chapter is to establish Theorem 6.4, which contains the change of variables theory for monotone transformations required in the proof of Theorem 4.2. Although $\nabla\psi$ may not be Lipshitz, the Jacobian factor $\det[\nabla^2\psi(x)]$ appearing in (34) is exactly what one expects from the standard theory of Lipshitz transformations.

As before, $\nabla\psi$ will be used to push-forward certain positive Radon measures ρ from Ω to $\mathbb{R}^d$. That is ρ may no longer have unit or even finite mass, but will be finite on compact subsets. The set of such measures will be denoted $\mathcal{M}(\Omega)$. $\rho \in \mathcal{M}(\Omega)$ is well known to decompose as $\rho = \rho_{ac} + \rho_{sing}$, where ρ_{ac} is absolutely continuous with respect to Lebesgue and ρ_{sing} vanishes except on a set of Lebesgue measure zero. The set $\mathcal{M}_{ac}(\Omega)$ of absolutely continuous measures is just the positive cone in $L^1_{loc}(\Omega)$; thus ρ_{ac} may be viewed simultaneously as a measure and a function. Differentiation of measures [19] is exploited to identify the pushed-forward measure: If $\rho \in \mathcal{M}(\Omega)$ for some domain Ω, its *symmetric derivative* $D\rho$ at $x \in \Omega$ is defined to be

$$D\rho(x) := \lim_{r\to 0} \frac{\rho B_r(x)}{\operatorname{vol} B_r}, \tag{31}$$

where $B_r(x)$ is the ball of radius r centered on x. $D\rho(x)$ exists — and agrees with $\rho_{ac}(x)$ — Lebesgue almost everywhere; $D\rho(x) = \infty$ on a set of

[4]Its derivative $\nabla^2\psi$ is defined in the sense of (120).

full measure for ρ_{sing}. From the symmetric derivative ρ-a.e., it is therefore possible to reconstruct ρ_{ac} and determine whether $\rho_{sing} = 0$. At the Lebesgue points of $\rho = \rho_{ac}$, hence almost everywhere, the limit (31) remains unchanged if the balls $B_r(x)$ are replaced by a sequence of Borel sets M_n *shrinking nicely to* x, meaning there is a sequence $r(n) \to 0$ such that $M_n \subset B_{r(n)}(x)$ and the ratio $\mathrm{vol}\, M_n / \mathrm{vol}\, B_{r(n)}$ is bounded away from zero.

The first lemma provides an alternative definition of $\nabla\psi_{\#}\rho$ which is exploited freely throughout this chapter.

Lemma 6.1 *Let ψ be a convex function with $\Omega := \mathrm{int}\,\mathrm{dom}\,\psi$, and denote the subgradient of its Legendre transform by $\partial\psi^*$. Let $\rho \in \mathcal{M}_{ac}(\Omega)$. Under the push-forward $\nabla\psi_{\#}\rho$, the measure of a Borel set $M \subset \mathbb{R}^d$ is equal to $\rho[\partial\psi^*(M)]$.*

Proof: The $\nabla\psi_{\#}\rho$ measure of M is $\rho[(\nabla\psi)^{-1}M]$. But $(\nabla\psi)^{-1}M \subset \partial\psi^*(M)$, and the difference is a set of zero measure for ρ. The containment is obvious: if $\nabla\psi(x) = y \in M$ then $(x, y) \in \partial\psi$ or $x \in \partial\psi^*(y)$. On the other hand, one can have $x \in \partial\psi^*(y)$ without $\nabla\psi$ being uniquely determined at x; however, this happens only for x in a set of Lebesgue (a fortiori ρ) measure zero. QED.

Proposition 6.2 *Let ψ be convex on $\mathbb{R}^d$, $\Omega := \mathrm{int}\,\mathrm{dom}\,\psi$ and $\rho \in \mathcal{M}_{ac}(\Omega)$. Assume x is a Lebesgue point for ρ at which ψ is twice differentiable*[5] *with invertible Hessian $\Lambda := \nabla^2\psi(x)$. At $\nabla\psi(x)$, the symmetric derivative (31) of the measure $\nabla\psi_{\#}\rho$ exists and equals $\rho(x)/\det\Lambda$.*

Proof: The Inverse Function Theorem (Proposition A.1) yields ψ^* twice differentiable at $y = \nabla\psi(x)$ with Λ^{-1} as its Hessian. Proposition A.2 then

[5] In the sense of (120).

shows that $\partial\psi^*(B_r(y))$ shrinks nicely to x. Since x is a Lebesgue point of ρ,

$$\frac{\rho[\partial\psi^*(B_r(y))]}{\mathrm{vol}\,[\partial\psi^*(B_r(y))]} \longrightarrow \rho(x), \tag{32}$$

as $r \to 0$. For the same limit, Proposition A.2 also shows

$$\frac{\mathrm{vol}\,[\partial\psi^*(B_r(y))]}{\mathrm{vol}\, B_r} \longrightarrow \det \Lambda^{-1}. \tag{33}$$

Taking the product of these two limits and observing Lemma 6.1 proves the result. QED.

Corollary 6.3 *Let ψ be convex on $\mathbb{R}^d$, $\Omega := \mathrm{int\,dom}\,\psi$. Then the function*[6] $\det[\nabla^2\psi(x)] \geq 0$ *is $L^1_{loc}(\Omega)$. Moreover, the push-forward of* $\det[\nabla^2\psi]$ *through $\nabla\psi$ is Lebesgue measure restricted to $\partial\psi(M)$, where M is the set of points where $\nabla^2\psi$ and its inverse exist which are also Lebesgue points for* $\det[\nabla^2\psi]$.

Proof: Consider the convex Legendre transform ψ^* and Lebesgue measure on $\mathrm{int\,dom}\,\psi^*$. The first claim is that $\det[\nabla^2\psi] \in L^1_{loc}(\Omega)$; in fact, it is the absolutely continuous part of $\omega := \nabla\psi^*_\#\mathrm{vol}$. Although the push-forward ω may have infinite mass near the boundary of Ω, its restriction to Ω is a Radon measure: if $K \subset \Omega$ is compact, so is $\partial\psi(K)$, whence $\omega[K] = \mathrm{vol}\,[\partial\psi(K)] < \infty$. The result is proven if $\det[\nabla^2\psi]$ agrees with the symmetric derivative $D\omega$ Lebesgue-a.e. on Ω. Recall that $\nabla^2\psi$ exists almost everywhere there. Where $\det[\nabla^2\psi(x)] > 0$, Proposition 6.2 (applied to ψ^* with $\rho := \mathrm{vol}$) and the Inverse Function Theorem (Proposition A.1) yield $D\omega(x) = \det[\nabla^2\psi(x)]$. On the other hand, $D\omega$ must vanish almost everywhere on the set Z where $\det[\nabla^2\psi(x)] = 0$: noting Lemma 6.1 and Proposition A.1, $0 < \omega[Z] = \mathrm{vol}\,\partial\psi(Z)$ would be incompatible with the fact that $\nabla^2\psi^*$ does not exist on $\partial\psi(Z)$. This establishes the first claim.

[6]Here the Hessian $\nabla^2\psi$ is defined in the sense of (120).

A second application of Proposition 6.2, but to ψ and with $\rho := \det[\nabla^2\psi]$, shows that the symmetric derivative of $\nabla\psi_{\#}\rho$ equals 1 on $\partial\psi(M)$. $\partial\psi(M)$ is of full measure for $\nabla\psi_{\#}\rho$, since M is for ρ. Thus $\nabla\psi_{\#}\rho$ can be nothing but Lebesgue measure on $\partial\psi(M)$. QED.

Theorem 6.4 (Monotone Change of Variables Theorem)
Let $\rho, \rho' \in \mathcal{P}_{ac}(\mathbb{R}^d)$, and ψ be a convex function on $\mathbb{R}^d$ with $\nabla\psi_{\#}\rho = \rho'$. With $\nabla^2\psi$ defined as in (120), the set $X := \{x \mid \nabla^2\psi(x) \text{ exists and invertible}\}$ has full measure for ρ. If $A(\varrho)$ is a measurable function on $[0,\infty)$ with $A(0) = 0$ then

$$\int A(\rho'(y))\, dy = \int_X A\left(\frac{\rho(x)}{\det[\nabla^2\psi(x)]}\right) \det[\nabla^2\psi(x)]\, dx. \tag{34}$$

(If $A(\varrho)$ is not single signed, either both integrals are undefined or both take the same value in $\mathbb{R} \cup \{\pm\infty\}$).

Proof: Since $\nabla\psi$ pushes ρ forward to ρ', ψ must be finite on a set of full measure for ρ. Thus $\nabla^2\psi(x)$ exists ρ-a.e., and by Proposition A.1 can only fail to be invertible on the set $\partial\psi^*(Z)$ where $Z = \{y | \nabla^2\psi^*(y) \text{ does not exist}\}$. By absolute continuity of ρ' and Lemma 6.1, $\rho[\partial\psi^*(Z)] = \rho'[Z] = 0$. Thus X is of full measure for ρ. Let $M \subset X$ consist of Lebesgue points for $\det[\nabla^2\psi]$, which is $L^1_{loc}(\Omega)$ by the preceding corollary. M differs from X by a set of Lebesgue (a fortiori ρ) measure zero. Thus $\partial\psi(M)$ is of full measure for ρ'. Since $\nabla\psi$ pushes forward $\det[\nabla^2\psi]$ to Lebesgue on $\partial\psi(M)$ by Corollary 6.3, the change of variables theorem (8) yields

$$\int_{\partial\psi(M)} A(\rho'(y))\, dy = \int_M A(\rho'(\nabla\psi(x))) \det[\nabla^2\psi(x)]\, dx.$$

Taking ρ' to coincide with its symmetric derivative, Proposition 6.2 shows that $\rho'(\nabla\psi(x)) = \rho(x)/\det[\nabla^2\psi(x)]$ at the Lebesgue points of ρ in M. Noting $A(0) = 0$, (34) follows immediately. QED.

Remark 6.5 (The Monge-Ampère Equation)
With restrictions on ρ', it was argued formally in [7] that the convex function ψ for which $\nabla\psi_{\#}\rho = \rho'$ represents a generalized solution to the Monge-Ampère equation

$$\rho'\Big(\nabla\psi(x)\Big) \det\left[\nabla^2\psi(x)\right] = \rho(x). \tag{35}$$

A regularity theory for these solutions has been developed by Caffarelli in [20, 21, 22]. Without any assumptions, Proposition 6.2 and the first part of Theorem 6.4 show (35) to be satisfied almost everywhere on $\operatorname{dom}\psi$, provided the Hessian $\nabla^2\psi(x)$ is interpreted in the sense of (120).

Part II
2-d Equilibrium Crystals in an External Field

In which the displacement interpolation is used to determine aspects of the shape of a crystal in a convex background potential.

7 Plane Crystals in a Convex Potential

The purpose of this section is to provide a characterization of the equilibrium shape of a crystal in an external field. When the field-strength is neglible, the crystal will form a convex set given by Wulff's construction [1] of 1901. The effect of a uniform field has been investigated in [2], but regarding the shape of formation in non-uniform fields, little is known. Even when the field is the (negative) gradient of a convex potential, the equilibrium crystal is not known to be connected, much less convex or unique. This problem, formulated as a variational minimization in d dimensions, was proposed for investigation by Almgren — whose interest may have been stimulated by its connection with curvature-driven flow [23]. Here it is addressed in the plane $d = 2$, where an unpublished result of Okikiolu [3] shows that any energy minimizing solution consists of a countable disjoint union of closed convex sets. Our main result — also limited to the plane — states that each convex set in this union is the *unique* energy minimizer among *convex sets* of its area. If the system is reflection symmetric under $x \leftrightarrow -x$, it follows that the crystal formed must be unique, convex and connected. In the context of curvature-driven flow, the last result leads to a new proof that a non-equilibrium crystal $C = -C$, initially convex, will remain convex and balanced as it melts or relaxes.

A crystal in equilibrium with its vapor or melt may be modeled as a subset $K \subset \mathbb{R}^2$ having area determined by the ambient thermodynamic variables. In the absence of competing effects, the shape of the crystal will minimize its interfacial surface energy, which is typically non-isotropic: an initial quantity of the condensate breaks the symmetry of the underlying space, establishing certain preferred directions. This non-isotropy is modeled by a *surface tension* $F : S^1 \longrightarrow (0, \infty)$, which depends continuously on the exterior unit normal to a boundary point x of K, or equivalently (in $d = 2$) on the oriented unit tangent $\hat{\tau}_x$ at x. The *measure theoretic boundary* $\partial_* K$ of K is defined to consist of precisely those points x at which K enjoys an exterior normal in the measure theoretic sense [15, §4.5.5], while the surface energy $\Phi(\partial_* K)$ to be minimized is

$$\Phi(\partial_* K) := \int_{\partial_* K} F(\hat{\tau}_x) d\mathcal{H}^1(x). \tag{36}$$

Here $\mathcal{H}^1$ denotes one-dimensional Hausdorff measure. The set K is assumed to be bounded and measurable, and to have finite perimeter $\mathcal{H}^1(\partial_* K) < \infty$ so its surface energy is finite; the collection of such sets is denoted by $\mathcal{K}$.

If K is dilated by $\lambda \geq 0$, its surface energy scales: $\Phi\Big(\partial_*(\lambda K)\Big) = \lambda \Phi(\partial_* K)$. Among sets of a fixed area, Φ is uniquely minimized by the *Wulff shape* $W \subset \mathbb{R}^2$ and its translates. W is conveniently constructed by Legendre transforming the surface tension F, after extending F from the unit circle $S^1 = \{|x| = 1\} \subset \mathbb{R}^2$ to the entire plane by taking it to be *positively homogeneous*: $F(\lambda x) = \lambda F(x)$ whenever $x \in \mathbb{R}^2$ and $\lambda \geq 0$. Positive homogeneity implies that the Legendre transform

$$F^*(y) := \sup_{x \in \mathbb{R}^2} \langle\, y,\, x \rangle - F(x)$$

takes only the values 0 and ∞; W is the compact convex set on which F^* vanishes. That W is the unique minimizer of $\Phi(\partial_* K)$ among sets of its area is demonstrated in [24]; a history of this fact is there given. When the surface tension F is convex on $\mathbb{R}^2$, the most economical interface connecting two points will be a straight line: [23, §3.1.9] or Lemma 9.5. From the duality theory of Legendre transformations it is clear that any surface tension F may be replaced by a convex one F^{**} which shares the same Wulff shape.

If the crystalline material interacts with an external potential Q on $\mathbb{R}^2$, then — in addition to its surface energy — $K \in \mathcal{K}$ carries potential energy

$$\Psi(K) := \int_K Q(x)\, d\mathcal{H}^2(x). \tag{37}$$

Here $\mathcal{H}^2$ denotes two-dimensional Lebesgue measure. At equilibrium, the shape of the crystal in the field $-\nabla Q$ should minimize the energy

$$\varepsilon(K) := \Phi(\partial_* K) + \Psi(K), \tag{38}$$

subject to the constraint of fixed area, normalized so that $\mathcal{H}^2(K) = 1$. One therefore wants to know that with this constraint, a minimum on $\mathcal{K}$ is attained by (38), and to deduce what properties one can of the minimizer. For the remainder of this paper, both energy integrands F and Q are assumed to be convex. These are the conditions under which Okikiolu's result is known to hold: any minimizer coincides — up to sets of measure zero which contribute neither to $\Psi(K)$ nor $\partial_* K$ — with a countable disjoint union of closed convex sets[7]. The potential $Q : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is also assumed to attain a minimum, and to do so only on a bounded set of measure zero in $\mathbb{R}^2$.

Existence of a minimizer $K_g \in \mathcal{K}$ follows from a continuity-compactness argument as in [23, §3.1.5]. Here the salient features of that argument are recounted without the machinery of integral currents: $\mathcal{H}^2(K \Delta L)$ is a metric on $\mathcal{K}$, where $K \Delta L := (K \sim L) \cup (L \sim K)$ is the symmetric difference of $K, L \in \mathcal{K}$; K is not distinguished from L if $\mathcal{H}^2(K \Delta L) = 0$, so $\mathcal{K}$ really consists of equivalence classes. For $\lambda, R < \infty$, the subset $\{K \in \mathcal{K}_R \mid \mathcal{H}^1(\partial_* K) \leq \lambda\}$

[7]Lemmas 9.5 and 9.6 provide essential elements of a proof.

is compact in this metric, where

$$\mathcal{K}_R := \{K \in \mathcal{K} \mid \mathcal{H}^2(K) = 1,\ \sup_K |x| \leq R\}; \tag{39}$$

the potential energy $\Psi(K)$ is continuous on $\mathcal{K}_R$ while the surface energy $\Phi(\partial_* K)$ is lower semi-continuous. Therefore, $\varepsilon(K)$ attains its minimum on $\mathcal{K}_R$ at some K_R: for any minimizing sequence, $\Phi(\partial_* K_n)$ must be bounded above; it controls $\mathcal{H}^1(\partial_* K_n)$ since F is bounded away from zero on S^1. It remains to show that for R sufficiently large, the infimum of $\varepsilon(K)$ over $\mathcal{K}_R$ is independent of R; the corresponding K_R will be a global energy minimizer K_g among sets with unit area. This last argument is provided by Proposition 11.1. Regularity of $\partial_* K_g$ is addressed in [23].

The main results of this paper apply to the equilibrium crystal K_g, now known to exist and to consist of countably many closed convex components. Stated precisely in the following chapter, they are proved by constructing an interpolation between pairs of convex sets which satisfies suitable energy estimates. This construction is described in Chapter 9; it is the *displacement interpolation* of Part I, tailored to match the background potential $Q(x)$. Chapter 10 sketches an application of the results to the dynamical problem of curvature-driven flow, while a final chapter contains estimates required for the size and number of components of K_g.

Since crystal formation and evolution are of the most physical interest in dimension $d = 3$, it is regretable that the key intermediate result — Theorem 9.7 — is restricted to the plane. In higher dimensions, the argument breaks down in two places. The estimate (49), which controls the surface energy of the interpolant, is false in dimension $d > 2$. Lemma 9.6 due to Okikiolu, which indicates that the surface energy of a connected open set dominates that of its convex hull, also fails when $d > 2$. Counterexamples to both are easily constructed in the isotropic case $F(x) = |x|$.

8 Statement of Results

The results are most clearly formulated by first considering a simpler variational problem: that of minimizing $\varepsilon(C)$ among *convex* sets of fixed area. Therefore, let $\mathcal{C}$ denote the collection of all closed convex sets of area one, a complete metric space under the *Hausdorff distance*

$$\mathrm{HD}(B,C) := \sup_{c\in C} \inf_{b\in B} |c-b| + \sup_{b\in B} \inf_{c\in C} |b-c| \tag{40}$$

for $B, C \in \mathcal{C}$. The topologies induced on $\mathcal{C}$ by $\mathrm{HD}(B,C)$ and $\mathcal{H}^2(B\Delta C)$ coincide [25], thus $\mathcal{C}$ is a closed subset of $\mathcal{K}$. Existence of a minimizer on $\mathcal{C}$ then follows from a compactness lower semi-continuity argument as above. Its uniqueness will be proved up to translation. Whether or not there is freedom to translate depends on the background potential $Q(x)$: strict convexity of $Q(x)$ forces the minimizer to be absolutely unique; other cases are easily resolved on an individual basis. All results in this paper pertain to the energy $\varepsilon(K)$ from (38), defined through energy integrands $F, Q : \mathbb{R}^2 \longrightarrow \mathbb{R}$ which satisfy

(E1) F convex with $F(\lambda x) = \lambda F(x)$ for $\lambda > 0$; $F > 0$ unless $x = 0$;

(E2) Q is convex, assuming its minimum on a bounded set of area zero.

Theorem 8.1 (Uniqueness of Minimizer Among Convex Sets)
Let $m \in (0, \infty)$ and assume (E1-E2). Among closed convex sets $C \subset \mathbb{R}^2$ with area $\mathcal{H}^2(C) = m$, the energy $\varepsilon(C)$ is minimized by a set $C_g(m)$ uniquely determined up to possible translations.

Proof: It is enough to consider $m = 1$; the other cases are equivalent after a new choice of length scale and corresponding modification of the integrands F and Q. Let $C \in \mathcal{C}$ be a convex set with $\varepsilon(C) \leq \lambda < \infty$. The convex potential $Q(x)$ assumes a minimum λ_0. Thus the surface energy $\Phi(\partial_* C) \leq \lambda - \lambda_0$ and the diameter of C is bounded (55) in terms of λ. C is contained in a ball $B_r(y)$ whose radius depends only on λ; it cannot be centered too far from the origin since $Q(x)$, being convex, grows linearly away from the bounded set $\{Q = \lambda_0\}$ and

$$\inf_{|x-y|\leq r} Q(x) \leq \Psi(C) \leq \lambda.$$

The conclusion is that $C \in \mathcal{K}_R$ for some R which depends on λ alone. Existence of a minimizer on $\mathcal{C}$ follows from lower semi-continuity of $\varepsilon(K)$ on $\mathcal{K}_R$ and the compactness result previously stated, equivalent here to the Blashke Selection Theorem.

To prove uniqueness, suppose there are two convex sets C and C' of minimum energy on $\mathcal{C}$. Obviously $\varepsilon(C) = \varepsilon(C')$. Unless C' is a translate of C, Theorem 9.7 provides a path $C(t) \in \mathcal{C}$ joining C to C' along which inequality (42) is strict for $t \in (0,1)$. $\varepsilon(C(t)) < \varepsilon(C)$ contradicts the assumption that C minimizes $\varepsilon(K)$ on $\mathcal{C}$. QED.

This result combines nicely with Okikiolu's to yield a description of the energy minimizer K_g among all sets of fixed area.

Theorem 8.2 (Classification of Connected Crystal Components) *Assume (E1-E2), and suppose K_g minimizes $\varepsilon(K)$ among $K \in \mathcal{K}$ with unit area. Then K_g is a finite disjoint union of closed convex sets $C_g(m)$, each with distinct area m and minimizing $\varepsilon(C)$ among convex sets C of its area.*

Proof: K_g is already known to consist of a countable disjoint union of closed convex components [3]; Proposition 11.2 bounds the number of such components. Let C be a convex component of K_g, and with $m := \mathcal{H}^2(C)$ define $C' := C_g(m)$ from Theorem 8.1. Then $\varepsilon(C) \geq \varepsilon(C')$ and equality holds only when $C = C'$ or possibly a translate. Otherwise, $\varepsilon(C) > \varepsilon(C')$. Choose a length scale so that $m = 1$. By Theorem 9.7 it is possible to define a continuous curve $C(t) \in \mathcal{C}$ joining C to C' along which the energy satisfies (42) on $[0,1]$. Thus $\varepsilon(C(t)) < \varepsilon(C)$ for $t > 0$. Since C is one of finitely many compact convex components of K_g, it enjoys a neighbourhood which is disjoint from $K_g \sim C$. Continuity of the curve $C(t)$ in the Haudorff metric ensures that for $t' > 0$ small enough, $C(t')$ lies in this neighbourhood; it can be substituted for C without interfering with the remainder of K_g. The energy of K_g would be lowered and its area unchanged, contradicting the fact that K_g is a minimizer and proving the main assertion.

The proof is concluded by showing that even if translates of the minimizer $C_g(m)$ share its energy, no two such translates C and C' occur as components in K_g. Otherwise, C may be translated toward C' using $C(t) := (1-t)C + tC'$; for $t \in [0,1]$ the energy $\varepsilon(C(t))$ is convex and therefore constant. As long as $C(t)$ remains disjoint from $K_g \sim C$, the set

$$K := C(t) \cup (K_g \sim C) \tag{41}$$

is a minimizer sharing the area and energy of K_g. As soon as $C(t)$ touches C' or some other component of K_g, a contradiction is reached: either $C(t)$ and C' share an edge, in which case the surface energy has been reduced and $\varepsilon(K) < \varepsilon(K_g)$, or else $C(t)$ and C' meet at a point, in which case K has a non-convex component $C(t) \cup C'$ violating Okikiolu's theorem. QED.

A corollary to the preceding theorems gives a sufficient condition for the equilibrium crystal to consist of a single convex component. This will be the case if the energy integrands $F(x) = F(-x)$ and $Q(x) = Q(-x)$ are both even. When the minimizers $C_g(m)$ among convex sets with areas m are truly unique, this result follows immediately from Theorem 8.2: each $C_g(m)$ is convex and balanced, hence contains the origin; no two of these sets are disjoint.

Corollary 8.3 (A Convex Equilibrium Crystal in the Even Case) *Assume $\varepsilon(K) = \varepsilon(-K)$ and (E1-E2). Then among $K \in \mathcal{K}$ with unit area, the minimizer K_g of $\varepsilon(K)$ is convex: it is unique up to translation, and may be taken to be balanced $K_g = -K_g$.*

Proof: If K_g is convex, the first part of the theorem is proved. If not, choose a convex component C of K_g not containing the origin; it is a minimizer among convex sets of its area by Theorem 8.2. Since $-C$ has the same energy and area, Theorem 8.1 forces $-C$ to be a translate of C. Let $-C = C - x$ and define $C(t) := C - t\,x$ for $t \in [0, 1]$. The energy $\varepsilon(C(t))$ is independent of t. As in the proof of the preceding theorem, a contradiction will be reached if $C(t)$ intersects $K_g \sim C$ at any t. Thus $C(t)$ and $K_g \sim C$ are disjoint for all t, which could not happen if $0 \in K_g$: being convex and balanced, $C(1/2)$ contains the origin. In any case, the minimizer K defined by (41) with $t = 1/2$ contains the origin. Applying the preceding argument to K instead of K_g leads to a contradiction. Thus K_g must have consisted of a single convex component C.

Thus $K_g = C$ coincides with the minimizer $C_g(1)$ of Theorem 8.1, which was uniquely determined up to translation. The translate $C(1/2)$ defined as above also minimizes $\varepsilon(K)$ and is balanced. QED.

9 Uniqueness of Convex Crystals

Let $\mathcal{C}$ denote the collection of compact convex sets in $\mathbb{R}^2$ which have unit area, metrized by the Hausdorff distance (40). Choose two sets C and C' from $\mathcal{C}$. In the current section, the displacement interpolation is used to construct a continuous curve $C(t) \in \mathcal{C}$ joining $C(0) = C$ to $C(1) = C'$, along which the energy satisfies the convexity estimate

$$\varepsilon\Big(C(t)\Big) \leq (1-t)\,\varepsilon(C) + t\,\varepsilon(C'). \tag{42}$$

The inequality is strict for $t \in (0,1)$ unless C' and C are translates. The existence of such a path forces C and C' to be translates if both minimize $\varepsilon(K)$ on $\mathcal{C}$. Theorem 8.2 exploits the continuity of $C(t)$ as well.

A result of Brenier [7] guarantees the existence of a unique[8] convex function $\psi : C \longrightarrow \mathbb{R}$ whose gradient $\nabla\psi$ is (an $\mathcal{H}^2$) measure-preserving map between C and C'. Because the sets are convex, $\nabla\psi$ is a homeomorphism from C to C': it is Hölder continuously differentiable ($C^{1,\alpha}$) on the interior of C and Hölder continuous (C^α) up to the boundary [20, 21]. The idea developed in Part I was to use the mapping $\nabla\psi$ to define an interpolant between C and C', viewed not as sets but as measures. That is, the characteristic functions $\rho := \chi_C$ and $\rho' := \chi_{C'}$ lie in the space $\mathcal{P}_{ac}(\mathbb{R}^2)$ of non-negative functions with integral one, and may therefore be regarded as probability measures. Being measure-preserving, $\nabla\psi$ may be said to *push-forward* ρ to ρ':

$$\int f\Big(\nabla\psi(x)\Big)\,\rho(x)\,d\mathcal{H}^2(x) = \int f(y)\,\rho'(y)\,d\mathcal{H}^2(y) \tag{43}$$

for measurable f on $\mathbb{R}^2$. The measure of a set M under ρ' may be recovered from ρ and $\nabla\psi$ by taking $f = \chi_M$ in (43). For $t \in [0,1]$, the *displacement interpolant* $\rho_t \in \mathcal{P}_{ac}(\mathbb{R}^2)$ is defined to be the push-forward of ρ through

[8] Up to an additive constant.

$(1-t)id + t\nabla\psi$; here id is the identity map on $\mathbb{R}^2$. ρ_t is again a non-negative function with integral one; for x in the interior of C its value at $(1-t)x + t\nabla\psi(x)$ is given by

$$\rho_t\Big((1-t)x + t\nabla\psi(x)\Big) = \det\left[(1-t)I + t\nabla^2\psi(x)\right]^{-1}; \tag{44}$$

here I is the identity matrix while $\nabla^2\psi(x)$ is the derivative of $\nabla\psi$ at x — matrix positive by convexity of ψ. Unlike ρ and ρ', ρ_t will not generally be the characteristic function of any set; however by Proposition 9.4 it is true is that $\rho_t(x) \leq 1$. The notation $\rho \xrightarrow{t} \rho'$ may replace ρ_t in order to emphasize explicitly the dependence on endpoints ρ and ρ'.

Although the surface and potential energies Φ and Ψ have been defined for sets rather than measures, the next two results show why one one can hope to use ρ_t to construct a set $C(t)$ whose energy satisfies the convexity estimate (42). To state these results, it is convenient to extend the definition of Φ to rectifiable curves, and of Ψ to measures $\rho \in \mathcal{P}_{ac}(\mathbb{R}^2)$ by

$$\Psi(\rho) := \int Q(x)\,\rho(x)\,d\mathcal{H}^2(x). \tag{45}$$

A continuous curve $\sigma : [a,b] \longrightarrow \mathbb{R}^2$ is said to be *rectifiable* if has finite *arc length* $\mathcal{L}(\sigma)$. The latter is defined as a supremum over finite partitions $\Pi = \{s_0 < s_1 < \cdots < s_n\} \subset [a,b]$:

$$\mathcal{L}(\sigma) := \sup_{\Pi\subset[a,b]} \sum_{i=1}^{n} |\sigma(s_i) - \sigma(s_{i-1})|. \tag{46}$$

In its arc length reparameterization $\tau : [0, \mathcal{L}(\sigma)] \longrightarrow \mathbb{R}^2$, the rectifiable curve σ is seen to be Lipschitz. As a consequence, the tangent $\tau'(s)$ exists for almost all s, and it is natural to define the surface energy $\Phi(\sigma)$ by

$$\Phi(\sigma) := \int_0^{\mathcal{L}(\sigma)} F(\tau'(s))ds. \tag{47}$$

If the measure theoretic boundary $\partial_* K$ of a set $K \in \mathcal{K}$ coincides (up to sets of $\mathcal{H}^1$ measure zero) with a positively oriented, rectifiable simple closed curve σ, then the definitions (36) and (47) coincide; this is certainly the case when K is convex. $\Phi(\sigma)$ may also be computed directly from σ in a manner analogous to the arc length (46):

$$\Phi(\sigma) := \sup_{\Pi \subset [a,b]} \sum_{i=1}^{n} F\Big(\sigma(s_i) - \sigma(s_{i-1})\Big). \tag{48}$$

(48) is manifestly invariant under reparameterization; in the arc length parameterization it can be seen to coincide with (47) by the dominated convergence theorem. Note that finiteness of (48) is equivalent to the rectifiability of σ: the surface tension $F(x)$ is positively homogeneous, bounded away from zero and infinity on the unit circle.

Lemma 9.1 (Displacement Convexity of Potential Energy $\Psi(\rho_t)$) *Let $C, C' \in \mathcal{C}$ be convex sets, and define the measure $\rho_t := \chi_C \overset{t}{\rightarrow} \chi_{C'}$ to be the displacement interpolant between them. Then the potential energy $\Psi(\rho_t)$ is convex as a function of t on $[0,1]$. Strict convexity of $\Psi(\rho_t)$ follows from strict convexity in (E2) unless $C = C'$.*

Proof: Let ψ be the convex function whose gradient is a measure preserving map from C to C'. The change of variables formula (43) is used to define ρ_t. Taking $f(y) = Q(y)$, the potential energy (45) is given by

$$\Psi(\rho_t) = \int Q\Big((1-t)x + t\nabla\psi(x)\Big)\, \chi_C(x)\, d\mathcal{H}^2(x).$$

The integrand is manifestly convex as a function of t, therefore the integral must be as well. If the convexity of $Q(x)$ is strict, then the integral will be strictly convex unless $\nabla\psi(x) = x$ almost everywhere on C. In the latter case, C and C' differ only by a set of measure zero; $C = C'$ since both are compact and convex. QED.

Remark 9.2 The fact that the endpoints ρ_0 and ρ_1 were characteristic functions of convex sets played no role in Lemma 9.1. The displacement interpolant ρ_t may in fact be defined between any pair of probability measures $\rho_0, \rho_1 \in \mathcal{P}_{ac}(\mathbb{R}^d)$ and in any dimension. *Displacement convexity* of Ψ — convexity of $\Psi(\rho_t)$ as a function of t — holds true in this general framework.

Proposition 9.3 (Displacement Convexity of the Surface Energy) *Let $C, C' \in \mathcal{C}$ be convex sets, $t \in [0,1]$, and the measure $\rho_t := \chi_C \xrightarrow{t} \chi_{C'}$ be the displacement interpolant between C and C'. The set $\{\rho_t > 0\}$ has a (positively oriented) rectifiable simple closed curve $\sigma_t : S^1 \longrightarrow \mathbb{R}^2$ as its boundary, and the surface energy $\Phi(\sigma_t)$ satisfies*

$$\Phi(\sigma_t) \leq (1-t)\ \Phi(\partial_* C) + t\ \Phi(\partial_* C'). \tag{49}$$

Proof: Let ψ be the convex function whose gradient is a measure-preserving homeomorphism from C to C'; by the invariance of domain, $\nabla\psi$ is a homeomorphism of their boundaries as well. Therefore, choose a positively oriented simple closed curve $\sigma : S^1 \longrightarrow \mathbb{R}^2$ parameterizing the boundary of C. For $s \in S^1$, define $\sigma_t(s) := T_t(\sigma(s))$ where $T_t := (1-t)\, id + t\, \nabla\psi$. The boundary of C' is parameterized by σ_1, and $\sigma_t(s)$ is a homotopy between σ_0 and σ_1. For $t < 1$, the continuous map $T_t : C \longrightarrow \mathbb{R}^2$ is the gradient of a *strictly* convex function, hence one-to-one; σ_t must be a simple closed curve with the same orientation for all $t \in [0,1]$. The image of C under the map T_t must be simply connected; it therefore covers the region enclosed by σ_t. From (44) it is clear that $\rho_t > 0$ in the interior of this region.

To prove that ρ_t vanishes outside of σ_t, it is enough to show that for x from the interior of C, the point $T_t(x)$ is enclosed by σ_t; the support of ρ_t will then be encircled. This is obviously true at $t = 0$ and $t = 1$. The homeomorphic image of $S^1 \times [0,1]$ in $\mathbb{R}^3$ under the map $(s,t) \longrightarrow (\sigma_t(s), t)$ forms a "tube" having $C \times \{0\}$ and $C' \times \{1\}$ as its ends. Both ends of the line segment $(T_t(x), t)$ lie in this tube; the entire segment must also lie within unless it crosses through the tube at some $t \in (0,1)$. In the event of a crossing, $T_t(x) = T_t(\sigma(s))$ with $s \in S^1$. Since x lies in the interior of C, $\sigma(s) \neq x$ and a contradiction is reached since T_t is one-to-one.

It remains to show (49); rectifiability of σ_t is a consequence. From (48),

$$\begin{aligned}
\Phi(\sigma_t) &= \sup_{\Pi \subset S^1} \sum_i F\Big(\sigma_t(s_i) - \sigma_t(s_{i-1})\Big) \\
&= \sup_{\Pi \subset S^1} \sum_i F\Big((1-t)\Big(\sigma_0(s_i) - \sigma_0(s_{i-1})\Big) + t\Big(\sigma_1(s_i) - \sigma_1(s_{i-1})\Big)\Big) \\
&\leq (1-t)\,\Phi(\sigma_0) + t\,\Phi(\sigma_1).
\end{aligned}$$

The final inequality was obtained using the convexity of the surface tension F. The proof is completed by the observation that $\Phi(\sigma_0) = \Phi(\partial_* C)$ and $\Phi(\sigma_1) = \Phi(\partial_* C')$ by construction. QED.

It remains to construct a convex set $C(t) \in \mathcal{C}$ to replace the measure ρ_t in the preceding estimates. The construction, given in Theorem 9.7, is facilitated by three preliminary results. The first of these is a refinement of Theorem 4.2, easily proved directly in this context. It shows that ρ_t will not be the characteristic function of any set unless C and C' are translates.

Proposition 9.4 *Let $C, C' \in \mathcal{C}$ be convex, and the measure $\rho_t := \chi_C \xrightarrow{t} \chi_{C'}$ be the displacement interpolant between them. Then $\rho_t(x) \leq 1$ on $\mathbb{R}^2$. Unless C and C' are translates, $\rho_t(x) \in (0,1)$ holds on a set of positive measure.*

Proof: As in Theorem 4.2, the fundamental observation is that for a positive $d \times d$ matrix $\Lambda > 0$, $\det\,[(1-t)I + t\Lambda]^{1/d}$ is a concave function of t, strictly concave unless Λ is a multiple of the identity I. In the basis diagonalizing Λ this is seen to result from the domination of the geometric by the arithmetic mean. Observing (44) and taking $\Lambda := \nabla^2\psi(x)$, from $t = 1$ it is clear that $\det\,[\Lambda] = 1$. Therefore, $\rho_t \leq 1$ for $t \in (0,1)$ with equality only if $\Lambda = I$.

Either ρ_t takes values in $(0,1)$ on a set of positive measure, or $\nabla^2\psi(x) = I$ almost everwhere in the interior of C. ψ is continuously twice differentiable there, so the latter implies $\nabla\psi(x) = x + y$ for some $y \in \mathbb{R}^2$ and $C' = C + y$ a translate of C. QED.

The following lemmas show that the surface energy $\Phi(\partial_* C(t))$ is decreased by convex intersections or by taking convex hulls. The second lemma is based on an idea from Okikiolu's proof that the connected components of an equilibrium crystal are convex.

Lemma 9.5 (Convex Set Comparisons) *Suppose $K \in \mathcal{K}$ and let $C \subset \mathbb{R}^2$ be convex. Then the surface energy $\Phi\Big(\partial_*(K \cap C)\Big) \leq \Phi(\partial_* K)$.*

Proof: C is assumed to be closed, since the difference is a set of measure zero and therefore irrelevant. If C is a half-plane $H = \{x \mid \langle x, y \rangle \leq \lambda\}$ with $y \in \mathbb{R}^2$ and $\lambda \in \mathbb{R}$, the desired inequality was noted in [23, §3.1.9]. The observation underlying it is that the surface energy of a piecewise linear curve with vertices at $x_0, x_1, \ldots, x_n \in \mathbb{R}^2$ dominates the energy of the segment joining x_0 to x_n:

$$F(x_n - x_0) \leq \sum_{i=1}^{n} F(x_i - x_{i-1}) \tag{50}$$

by convexity and positive homogeneity of the surface tension F. This generalizes through (48) to rectifiable curves, and thence to $\partial_* K$.

If C is an arbitrary closed convex set, choose a countable dense set of points from its boundary. C is the intersection of its supporting half-planes $H_n \supset C$ at these points. Define $K_n := H_n \cap K_{n-1}$ inductively from $K_0 := K$. $\Phi(K_n)$ will be non-increasing. Moreover, $\mathcal{H}^2(K_n \Delta (C \cap K)) \to 0$ with n since $\mathcal{H}^2(K) < \infty$. By the lower semi-continuity of Φ in this metric, $\Phi(\partial_*(K \cap C)) \leq \lim_n \Phi(\partial_* K_n) \leq \Phi(\partial_* K)$. QED.

Lemma 9.6 (Okikiolu [3]) *Let $\sigma : S^1 \longrightarrow \mathbb{R}^2$ be a simple closed curve and $\Omega := \operatorname{conv}\sigma$ the convex hull of the domain it encloses. Then $\Phi(\partial_*\Omega) \leq \Phi(\sigma)$.*

Proof: Let K denote the domain enclosed by the curve σ. By the preceding lemma or inequality (50), $\Phi(\sigma)$ is decreased whenever any portion of the curve is replaced by a straight line. If P is a convex polygon having vertices in K, then $\Phi(\partial_* P) \leq \Phi\Big(\partial_*(K \cup P)\Big) \leq \Phi(\partial_* K)$. Ω is approximated up to a set of area $1/n$ by a convex polygon $P_n \subset \Omega$. Exploiting lower semi-continuity of Φ as before, $\Phi(\partial_*\Omega) \leq \underline{\lim}_n \Phi(\partial_* P_n) \leq \Phi(\partial_* K) = \Phi(\sigma)$. QED.

Theorem 9.7 (Interpolation Between Convex Crystals)
Let $C, C' \in \mathcal{C}$ be convex. Then there is a continuous curve $C(t) \in \mathcal{C}$ joining $C(0) = C$ to $C(1) = C'$ along which the energy $\varepsilon(C(t))$ satisfies (42) on $[0,1]$, with strict inequality when $t \in (0,1)$ unless C' is a translate of C.

Proof: If $C' = C + x$ is a translate of C, then $C(t) := C + t\,x$ satisfies the theorem. Otherwise, define the displacement interpolant $\rho_t := \chi_C \stackrel{t}{\rightarrow} \chi_{C'}$ between C and C', and let $\Omega(t) := \operatorname{conv}\{\rho_t > 0\}$. By Proposition 9.3 and Lemma 9.6 the surface energy

$$\Phi\Big(\partial_*\Omega(t)\Big) \leq (1-t)\,\Phi(\partial_* C) + t\Phi(\partial_* C').$$

Proposition 9.4 shows $\rho_t \leq 1$ on $\Omega(t)$, with strict inequality on a set of positive measure. $\mathcal{H}^2(\Omega(t)) > 1$ since $\int \rho_t = 1$. To replace $\Omega(t)$ by a set of unit area, consider its intersection with the nested convex sets $Q_\lambda = \{Q(x) \leq \lambda\}$ indexed by $\lambda \in \mathbb{R}$. Since $Q(x)$ is convex and attains its minimum λ_0 on a set of measure zero, the area of $\{Q = \lambda\}$ vanishes for all λ. $\mathcal{H}^2(Q_\lambda)$ is continuously increasing on $[\lambda_0, \infty)$ as a result, having $[0, \infty)$ as its image. For each $t \in (0,1)$, choose λ depending on t so that $\mathcal{H}^2(\Omega(t) \cap Q_\lambda) = 1$, and set $C(t) := \Omega(t) \cap Q_\lambda$. Then $C(t) \in \mathcal{C}$, and by Lemma 9.5 the surface energy has not been increased:

$$\Phi\Big(\partial_* C(t)\Big) \leq \Phi\Big(\partial_* \Omega(t)\Big).$$

The potential energies are compared by observing that $0 \leq (\rho_t - \chi_{C(t)})(Q - \lambda)$, with strict inequality on a set of positive measure. Integrating yields

$$\Psi\Big(C(t)\Big) < \Psi(\rho_t).$$

These estimates, together with Lemma 9.1, imply strict inequality in (42).

It remains to show that $C(t)$ is a continuous function of t in the Hausdorff metric. This metric (40) applies equally well to compact sets which need not be convex. Since the homeomorphism $(1-t)id + t\nabla\psi$ of C onto the closure of $\{\rho_t > 0\}$ varies continuously with t in the sup norm, its image varies continuously in the Hausdorff metric. Taking convex hulls of compact sets is a continuous operation, thus $\Omega(t)$ is a continuous function of $t \in [0,1]$. At the same time $\mathcal{H}^2(Q_\lambda \Delta Q_{\lambda'}) = \mathcal{H}^2(Q_\lambda) - \mathcal{H}^2(Q_{\lambda'})$ for $\lambda > \lambda'$, so the nested sets Q_λ depend continuously on λ in the area metric, or equivalently in the Hausdorff distance. $\Omega(t) \cap Q_\lambda$ is jointly continuous as a function of (t, λ) since

$$\Big(\Omega(t) \cap Q_\lambda\Big) \,\Delta\, \Big(\Omega(t') \cap Q_{\lambda'}\Big) \subset \Big(\Omega(t) \Delta \Omega(t')\Big) \cup \Big(Q_\lambda \Delta Q_{\lambda'}\Big).$$

The proof will be completed by showing that in order to define $C(t)$, $\lambda(t) := \sup\{Q(x) \mid x \in C(t)\}$ was chosen continuously as a function of t.

The function $h(t,\lambda) := \mathcal{H}^2(\Omega(t) \cap Q_\lambda)$ is continuous on $[0,1] \times (\lambda_0, \infty)$ and $h(t, \lambda(t)) = 1$. For t fixed, $h(t,\lambda)$ is a monotone non-decreasing function of λ, strictly monotone for λ near $\lambda(t)$ when $t \in (0,1)$ since the convex set $\Omega(t)$ has area larger than one. At the endpoints $t = 0$ and $t = 1$, strict monotonicity holds only in the one-sided neighbourhood $\lambda < \lambda(t)$. For all t and $\epsilon > 0$,

$h(t, \lambda(t) - \epsilon) < 1$. For t' near t, it will still be true that $h(t', \lambda(t) - \epsilon) < 1$; thus $\lambda(t') > \lambda(t) - \epsilon$. For $t \in (0, 1)$ the reverse inequalities hold with $\epsilon < 0$. As a result, $\lambda(t)$ is continuous on $(0, 1)$ and lower semi-continuous at the endpoints. Finally, choose $\lambda > \lambda(0)$. Q_λ contains a neighbourhood of the compact set C; by continuity in the Hausdorff metric $\Omega(t) \subset Q_\lambda$ for small enough t, hence $\lambda(t) \leq \lambda$. The same argument applies at $t = 1$ and shows $\lambda(t)$ to be continuous on the interval $[0, 1]$. Thus $C(t) := \Omega(t) \cap Q_{\lambda(t)}$ is a continuous curve in the Hausdorff metric on $\mathcal{C}$. QED.

10 Curvature-Driven Flow

Curvature-driven flow, or *motion by weighted mean curvature*, is a dynamical model for the time evolution of a non-equilibrium crystal under the influence of its surface tension. In this model, the normal velocity of each point on the crystalline interface is presumed proportional to the local change of surface energy with volume (area in $\mathbb{R}^2$) [23, §2.2]. If the interface and its evolution are both smooth, the flow may be referred to as *classical*. For curves in the plane, the resulting motion has been studied by Angenent and Gurtin [26]. Here a connection will be established between curvature-driven flow and the statical problem we have considered. Corollary 8.3 will be used to provide an alternative proof of a result announced in [26, §7.3]: that a convex crystal $K_0 \in \mathcal{K}$ away from equilibrium, remains convex for all time under curvature-driven flow. This result was known earlier in the isotropic case $F(x) = |x|$ both for curves in the plane [27] and surfaces in higher dimensions [28], but [26] is more general, addressing a wide class of surface energies as well as non-convex crystals. Our approach — currently limited to the case $K_0 = -K_0$ — is of interest because it reduces the non-isotropic question in higher dimensions to the study of a statical problem.

The connection is established through results of Almgren, Taylor and Wang [23]. There it was shown that the curvarture-driven flow starting from $K_0 \in \mathcal{K}$ was approximated by a discrete time flow, in which the evolved crystal K after time Δt is the minimizer on $\mathcal{K}$ of a functional (38). The potential $Q(x)$, which need not be convex, represents the tendency of the flow to remain near its initial condition K_0 for short times: it is proportional to the signed distance to the boundary $\partial_* K_0$ and decays with elapsed time Δt:

$$Q(x)\,\Delta t := \mathrm{dist}_{\pm}(x, K_0) := \begin{cases} \mathrm{dist}(x, \partial_* K_0) \text{ if } x \notin K_0 \\ -\mathrm{dist}(x, \partial_* K_0) \text{ if } x \in K_0. \end{cases} \tag{51}$$

A discrete evolution is generated by repeated minimization, replacing K_0 by K at each step. A continuous time flow, called a *flat Φ curvature flow*, may be extracted in the limit $\Delta t \to 0$. Under additional restrictions this flat flow coincides with classical curvature flow when the latter exists.

If the initial configuration K_0 is a convex set, Lemma 10.2 shows that the corresponding potential (51) is convex. If K_0 is balanced and $F(x) = F(-x)$ is even, an application of Corollary 8.3 then shows that the crystal remains balanced and convex at all subsequent times. It is of interest to note that the distance function appearing in (51) need not be Euclidean distance: it is sufficient that

$$\mathrm{dist}(x, \partial_* K) := \inf_{k \in \partial_* K} \mathrm{M}(x - k) \tag{52}$$

for any norm $\mathrm{M}(x)$ on $\mathbb{R}^2$. A non-Euclidean norm corresponds to a non-isotropic *mobility*: a direction dependent response of the crystalline interface to applied force.

Theorem 10.1 (Curvature Flow Preserves Balanced Convex Sets) *Let $K(t) \in \mathcal{K}$ be a flat Φ curvature flow [23] on some interval $t \in [0, T]$. If the initial condition $K(0) = -K(0)$ is convex and the surface tension $F(x) = F(-x)$ satisfying (E1) is even, then the crystal $K(t)$ will be convex at each subsequent time. $K(t)$ will also have reflection symmetry through some point $x_t \in \mathbb{R}^2$: $K(t) - x_t = x_t - K(t)$.*

Proof: Since $K_0 = K(0)$ is convex and balanced, Lemma 10.2 shows the potential $Q(x)$ from (51) to be convex and balanced, and to assume its minimum on a bounded set of measure zero. Let K minimize $\varepsilon(K)$ over $\mathcal{K}$, i.e. among sets of all areas. K exists, though it may be the null set; it is a fortiori an energy minimizer among sets of its area. Corollary 8.3 shows K to be convex, and if not balanced, then reflection symmetric through some other

$x \in \mathbb{R}^2$. K is an approximant to $K(\Delta t)$. The approximant to $K(2\Delta t)$ is obtained by repeating the procedure, starting from K instead of K_0. Since the problem is translation invariant, and a translate of K satisfies the hypotheses on K_0, the approximants to $K(n\Delta t)$ must all be convex and symmetric for $n > 1$. The flat Φ curvature flow $K(t)$ at time t is obtained [23, 2.6] as a limit of such approximants in the metric $\mathcal{H}^2(\,\cdot\,\Delta\,\cdot\,)$. $K(t)$ is convex since convex sets form a closed subset of $\mathcal{K}$ in this metric; it has a balanced translate for a similar reason. QED.

Lemma 10.2 (Convexity of the Signed Distance to a Convex Set) *Let $C \subset \mathbb{R}^d$ be a convex set, and $\mathrm{M}(x)$ a norm on $\mathbb{R}^d$. The signed distance $\mathrm{dist}_{\pm}(x, C)$ from (51) and (52) is a convex function of x on $\mathbb{R}^d$. If C is bounded, $\mathrm{dist}_{\pm}(x, C)$ assumes its minimum on a bounded set of measure zero in $\mathbb{R}^d$; if $C = -C$ is balanced, $\mathrm{dist}_{\pm}(x, C) = \mathrm{dist}_{\pm}(-x, C)$.*

Proof: Choose any supporting hyperplane to C, and let $H \supset C$ be the corresponding half-space. The first observation

$$\mathrm{dist}_{\pm}(x, C) \geq \mathrm{dist}_{\pm}(x, H) \tag{53}$$

is seen from three cases: if $x \notin H$, the boundary of H lies M-closer to x than the boundary of C; if $x \in C$ the situation is reversed, and (53) holds since both distances are negated; if $x \in H \sim C$, (53) holds on the basis of sign.

Now fix $x \in \mathbb{R}^d$. There is some c in the boundary of C such that $\mathrm{dist}(x, \partial_* C) = M(x - c)$. A supporting half-space $H \supset C$ exists with $c \in \partial_* H$ and with $\mathrm{dist}(x, H) = M(x - c)$: if $x \in C$ this is obvious, while if $x \notin C$ the hyperplane $\partial_* H$ must be slipped between the convex sets C and $\{y \mid M(x - y) < \mathrm{dist}(x, C)\}$. Thus (53) will be saturated for this H, and

$$\mathrm{dist}_{\pm}(x, C) = \sup_{H \supset C} \mathrm{dist}_{\pm}(x, H),$$

where the supremum is over half-spaces $H \supset C$. Convexity of $\text{dist}_{\pm}(x, C)$ is manifest since $\text{dist}_{\pm}(x, H)$ is linear (or at least affine).

The signed distance is positive for $x \notin C$, thus attains its minimum on a compact convex set $K \subset C$ when C is bounded. Any interior point of K would be M-farther from $\partial_* C$ than some point on the boundary of K is, contradicting the fact that $\text{dist}_{\pm}(x, C)$ is constant on K. Thus K is measure zero. $\text{dist}_{\pm}(x, C)$ is obviously even if $C = -C$. QED.

11 Support and Number of Components

In this section, two a priori estimates are proved regarding energy minimizing crystals K_g of $\varepsilon(K)$ with constrained area; the first localizes the support of K_g, while the second provides a lower bound for the area of its convex components, implying an upper bound on their number.

The next proposition implies that the minimizers of $\varepsilon(K)$ among crystals with unit area must all lie in a single large ball. To prove this, it is useful to have scaled copies of the Wulff shape W available for comparison; for $m > 0$, W_m is defined as the dilate of W having $\mathcal{H}^2(W_m) = m$. If $K \in \mathcal{K}$,

$$\Phi(\partial_* K) \geq \sqrt{\mathcal{H}^2(K)}\, \Phi(\partial_* W_1). \tag{54}$$

In two dimensions, a trivial estimate shows that the diameter of a connected open set $U \in \mathcal{K}$ is controlled by its surface energy:

$$\operatorname{diam} U := \sup_{x,y \in U} |x - y| \leq (2F_0)^{-1}\, \Phi(\partial_* U); \tag{55}$$

here $F_0 > 0$ is the minimum of the surface tension $F(x)$ on $|x| = 1$.

Proposition 11.1 (Bound on the Radius of a Minimizing Crystal) *Let K' be a minimizer of $\varepsilon(K)$ on $\mathcal{K}_R$ for $R < \infty$. There is some radius R' — given by (56) — depending only on the integrands F and Q such that $K' \in \mathcal{K}_{R'}$ whenever $R \geq R'$.*

Proof: The proof uses the fact that K' is expressible as a disjoint union of connected open sets U_n with $\Phi(\partial_* K') = \sum_n \Phi(\partial_* U_n)$. In fact, the U_n are convex by Okikiolu's argument, which applies even though K' may only minimize $\varepsilon(K)$ on $\mathcal{K}_R$.

Assume the Wulff shape W_1 is translated so that $\Psi(W_1) \leq \Psi(W_1 + x)$ for all $x \in \mathbb{R}^2$, and define $\lambda := \sup\{Q(x) \mid x \in W_1\}$ and $Q_\lambda := \{Q(x) \leq \lambda\}$. This set is bounded since the convex potential $Q(x)$ attains its minimum λ_0 on a bounded set. Since R' will be taken from (56), $\varepsilon(K') \leq \varepsilon(W_1)$ when $R \geq R'$. Thus the surface energy $\Phi(K')$ is controlled, leading through (55) to a bound $r := (2F_0)^{-1}(\varepsilon(W_1) - \lambda_0)$ on the diameter of each connected component U_n. Enlarging Q_λ by radius r, the result will be proved by showing that unless $K' \subset Q_\lambda + B_r(0)$, some $L \in \mathcal{K}_R$ has lower energy. It will therefore be sufficient to take

$$R' := r + \sup_{x \in Q_\lambda} |x|. \tag{56}$$

Suppose that a connected component U of K' intersects the complement of $Q_\lambda + B_r(0)$; it must be disjoint from Q_λ by construction. This observation, together with (54), shows that the energy gained by removing U from K' to leave $K := K' \sim U$, is at least

$$\varepsilon(K') - \varepsilon(K) > \sqrt{\mathcal{H}^2(U)}\, \Phi(\partial_* W_1) + \lambda\, \mathcal{H}^2(U). \tag{57}$$

K will not satisfy the area constraint, but there is room inside $W_1 \sim K$ to restore the excess mass $\mathcal{H}^2(U)$ since $\mathcal{H}^2(K) + \mathcal{H}^2(U) = 1 = \mathcal{H}^2(W_1)$. Because $W_1 \subset Q_\lambda$, the potential energy cost for introducing this mass will be

$$\Psi(L) - \Psi(K) \leq \lambda \mathcal{H}^2(U);$$

if L can be formed without paying too great a price in surface energy, the gain (57) will dominate. Choose a scaled copy of the Wulff shape $W_m \subset W_1$ for which $\mathcal{H}^2(W_m \sim K) = \mathcal{H}^2(U)$, and define $L := W_m \cup K$. Certainly $L \in \mathcal{K}_R$ if $R \geq R'$. The surface energy of L is controlled by an inclusion-exclusion estimate [23, §3.1.4], (54) and $m \geq \mathcal{H}^2(U)$:

$$\begin{aligned}
\Phi(\partial_* L) - \Phi(\partial_* K) &\leq \Phi(\partial_* W_m) - \Phi\Big(\partial_*(W_m \cap K)\Big) \\
&\leq \Phi(\partial_* W_m)\left(1 - \sqrt{1 - \mathcal{H}^2(U)/m}\right) \\
&\leq \Phi(\partial_* W_m)\sqrt{\mathcal{H}^2(U)/m}.
\end{aligned}$$

The three preceding inequalities yield $\varepsilon(L) < \varepsilon(K')$, contradicting the fact that K' is an energy minimizer. QED.

Proposition 11.2 (Lower Bound on Area of Convex Components) *Let K_g minimize $\varepsilon(K)$ among $K \in \mathcal{K}$ with unit area. If C is one of the disjoint convex components of K_g, then $\mathcal{H}^2(C) \geq m' > 0$; the area bound m' depends only on the integrands F and Q.*

Proof: Choose the origin of $\mathbb{R}^2$ to lie somewhere in C. Since C is convex, it may be contracted by a factor $0 < \eta < 1$ without intersecting $K := K_g \sim C$ or indeed any dilation of K by factor $\lambda > 1$. Moreover, η and λ may be chosen to depend on each other in such a way that $\eta C \cup \lambda K$ has unit area. Then the energy of this configuration is bounded below by $\varepsilon(K_g)$, which will lead to a lower bound on $\mathcal{H}^2(C)$.

Before the origin was shifted, K was contained in the ball $B_{R'}(0)$ by Proposition 11.1; R' depended only on F and Q. It will still be true that K is contained in a ball of radius $2R'$ about the new origin. The infinitesmal increase in $\varepsilon(\lambda K) = \lambda\Phi(\partial_* K) + \lambda^2 \int_K Q(\lambda x)\, d\mathcal{H}^2(x)$ with λ is given by

$$\left.\frac{d}{d\lambda}\right|_{\lambda=1} \varepsilon(\lambda K) = \Phi(\partial_* K) + \int_K (2Q + \langle\, x,\, \nabla Q\rangle)\, d\mathcal{H}^2(x); \tag{58}$$

being convex, $Q(x)$ is Lipschitz on $|x| < 2R'$ and the dominated convergence theorem has been applied. The cost (58) of dilating K is controlled by a constant depending only on F and Q. On the other hand, both the surface and potential energy of C will decrease as it is contracted ...the latter because $\eta C \subset C$. The proposition is proved by the next estimate, which shows that unless $\mathcal{H}^2(C)$ is bounded below, the gain in $\Phi(\eta C)$ with a small change in λ outweighs the cost (58); this would be inconsistent with $\varepsilon(K_g)$ a global minimum. The estimate relies on (54) and the area constraint

$\eta^2\mathcal{H}^2(C) + \lambda^2(1 - \mathcal{H}^2(C)) = 1$; when $\mathcal{H}^2(C)$ is small, a slight change in λ results in a huge change in η. Thus

$$-\frac{d}{d\lambda}\bigg|_{\lambda=1} \Phi(\partial_*(\eta C)) = -\Phi(\partial_* C)\frac{d\eta}{d\lambda}\bigg|_{\eta=\lambda=1} \tag{59}$$

$$\geq \mathcal{H}^2(C)^{-1/2}(1 - \mathcal{H}^2(C))\,\Phi(W_1) \tag{60}$$

diverges with $\mathcal{H}^2(C) \to 0$. The proposition is concluded by choosing m' small enough so $\mathcal{H}^2(C) < m'$ implies the gain (60) in surface energy alone outweighs the K_g independent bound on the cost (58). QED.

Part III
Rotating Stars

12 The Stability of Rotating Stars

In a simple class of astrophysical models, a star is represented as a fixed mass of gravitating fluid, obeying an equation of state in which the pressure $P(\varrho)$ depends on the density only; with an appropriate equation of state this is a reasonable model for a cold white dwarf star [13]. The investigation of rotating equilibria for such a fluid has been of great mathematical and physical interest since the time of Newton: the homogeneous incompressible case alone has a venerable history chronicled elsewhere [29]. More recently, compressible fluid models have enjoyed a revival of interest [30, 31, 32, 33, 34, 35, 5, 36] since Auchmuty and Beals [4, 37] demonstrated the existence of axisymmetric equilibria in which infinitesmal concentric cylinders of fluid rotate differentially. Either the angular velocity or angular momentum profile of differential rotation was specified a priori, and satisfied decay conditions precluding the possibility of uniform rotation [5]. Thus the equilibria of Auchmuty and Beals, though they solve the inviscid Euler equations, do not represent ground states of the physical system: differential rotation implies that there is excess energy waiting to be dissipated through viscous friction.

A more fundamental problem is to determine the stable equilibrium states of the system, subject only to the physical constraints of specified fluid mass, linear momentum and angular momentum $\mathbf{J}$ about the center of mass. This is the problem addressed here. It is formulated as a variational minimization of the energy $E(\rho, \mathbf{v})$, which depends on the fluid density $\rho(\mathbf{x}) \geq 0$ and velocity field $\mathbf{v}(\mathbf{x})$ on $\mathbb{R}^3$. The problem is peculiar in that the energy — although bounded below — does not attain its constrained minimum except in the non-rotating case $\mathbf{J} = 0$. As a result, one is forced to settle for *local* energy minimizers, where local must be suitably defined. Such minimizers prove to be stable, uniformly rotating solutions of the Euler or Navier-Stokes-Poisson system:

$$\begin{aligned} \nabla P(\rho) &= \rho \left\{ \nabla (V\rho) + \omega^2 r(\mathbf{x})\, \hat{\mathbf{r}} \right\}; & (61) \\ -\Delta V\rho &= 4\pi\rho. & (62) \end{aligned}$$

Here cylindrical co-ordinates have been chosen for the center of mass frame; the angular momentum $\mathbf{J} = J\hat{\mathbf{z}}$ selects the z-axis. This axis will be a principal axis of inertia for ρ, and the corresponding moment of inertia $I(\rho)$ determines the angular velocity $\omega := J/I(\rho)$. The gravitational potential $V\rho$ is given by (68). Regarding the astrophysical relevance of this formulation, we concede that for many applications relaxation to uniform rotation takes place on unreasonably long time scales. Nevertheless, there are contexts in which it may be a dominant effect. For example, observational evidence indicates that in ancient close binary systems, the rotational periods of the component stars coincide with the system's orbital period; the two stars rotate as a solid body [38].

On physical grounds, it is evident that the system (61-62) should have solutions for any prescribed fluid mass M and angular momentum J. However, solutions have been proven to exist only for J small [5]; the analysis there is formulated in terms of angular velocity ω and assumes axisymmetry. In the following pages, the existence of solutions is demonstrated in the opposite regime: for large angular momentum J. These solutions take the form of binary stars, in which the fluid mass is divided into two disjoint regions widely separated relative to their size. The ratio of masses between the two regions may be specified a priori. It is clear that these solutions will not be axisymmetric, but they do have $z = 0$ as a symmetry plane.

Since they are constructed as local energy minimizers, these binary stars will be stable. However, the stationarity condition they satisfy (75) differs slightly from the Euler-Lagrange equation for a global energy minimizer, in that the *chemical potential* — usually thought of as a Lagrange multiplier conjugate to the constraint of fixed mass — need not be constant throughout the set $\{\rho > 0\}$; instead, each connected component of $\{\rho > 0\}$ has its own chemical potential. This possibility is of particular relevance if one is interested in counting connected components of a solution as in [32, 36]. It also makes perfect sense physically: one would not expect mass at the earth's surface to be as tightly bound as at the surface of the sun, even if the system were in equilibrium.

Unfortunately, intermediate values of the angular momentum J remain inaccessible to us. However, some global features of the problem may be demonstrated in the context of a one-dimensional toy model proposed in Chapter 15. This model represents an interacting compressible fluid, constrained to live in a long light tube, and rotating end-over-end about its center of mass. It has the virtue of being exactly solvable: for a given mass, the solutions come in an uncountable number of disjoint families, each parameterized continuously by the angular velocity ω. The solutions with connected support — single stars — form a family which persists as long as J is not too large. The remaining families persist for J not too small, and represent binary stars or stellar systems in the astrophysical analogy. All families terminate with equatorial break-up of the lightest star in the system.

In the following chapter, the three-dimensional problem and results are formulated precisely. Chapter 14 collects results which, although not original, are required for the analysis; the reduction to uniform rotation is due to Elliott Lieb [39]. In Chapter 15 the one-dimensional model is introduced and analyzed, while the stationarity and regularity properties of local energy minimizers for the real problem are discussed in Chapter 16. The last chapter contains the proof that for large angular momentum, local minimizers exist in the form of binary stars.

13 Formulation of the Problem

The state of a fluid may be represented by its mass density $\rho(\mathbf{x}) \geq 0$ and velocity vector field $\mathbf{v}(\mathbf{x})$ on $\mathbb{R}^3$. If the fluid interacts with itself through Newtonian gravity and satisfies an equation of state in which the pressure $P(\varrho)$ is an increasing function of the density only, then its energy $E(\rho, \mathbf{v})$ is given as the sum of three terms: the internal energy $U(\rho)$, gravitational interaction energy $G(\rho, \rho)$, and kinetic energy $T(\rho, \mathbf{v})$. Each is expressed as an integral over $\mathbf{x} \in \mathbb{R}^3$:

$$
\begin{aligned}
E(\rho, \mathbf{v}) &:= U(\rho) - G(\rho, \rho)/2 + T(\rho, \mathbf{v}); && (63) \\
U(\rho) &:= \int A(\rho)\, d^3\mathbf{x}; && (64) \\
G(\sigma, \rho) &:= \int V\sigma \, d\rho(\mathbf{x}); && (65) \\
T(\rho, \mathbf{v}) &:= \frac{1}{2}\int |\mathbf{v}|^2 d\rho(\mathbf{x}). && (66)
\end{aligned}
$$

Here $A(\varrho)$ is a convex function obtained from the equation of state by integrating $dU = -P dv$ from infinite to unit volume,

$$A(\varrho) := \int_1^\infty P(\varrho/v)\, dv, \tag{67}$$

while $V\rho$ represents the gravitational potential of the mass density $\rho(\mathbf{x})$:

$$V\rho(\mathbf{x}) := \int \frac{d\rho(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|}. \tag{68}$$

Units are chosen so that the total mass of fluid $M = 1$ and the gravitational constant $G = 1$, and a frame of reference is chosen in which the center of mass

$$\overline{\mathbf{x}}(\rho) := \left(\int d\rho\right)^{-1} \int \mathbf{x}\, d\rho(\mathbf{x}) \tag{69}$$

is at rest. One is then interested in finding minimum energy configurations subject to constraints of fixed mass and angular momentum $\mathbf{J}$ about the center of mass $\overline{\mathbf{x}}(\rho)$. The fluid angular momentum is given by $\mathbf{J}(\rho, \mathbf{v})$:

$$\mathbf{J}(\rho, \mathbf{v}) := \int \Big(\mathbf{x} - \overline{\mathbf{x}}(\rho)\Big) \times \mathbf{v} \; d\rho(\mathbf{x}). \tag{70}$$

Before addressing the rotating problem $\mathbf{J} \neq 0$, further assumptions and results will be required of the non-rotating energy $E_0(\rho) := U(\rho) - G(\rho, \rho)/2$. The pressure $P(\varrho)$ may take a quite general form, including the polytropic equations of state $P(\varrho) = \varrho^q$ with $q > 4/3$ and the Chandrasekhar equation of state [13]. Following Auchmuty and Beals [4], the tacit assumptions on $P(\varrho)$ will be:

(F1) $P : [0, \infty) \longrightarrow [0, \infty)$ continuous and strictly increasing;

(F2) $\lim_{\varrho \to 0} P(\varrho)\varrho^{-4/3} = 0$;

(F3) $\liminf_{\varrho \to \infty} P(\varrho)\varrho^{-4/3} > K(M)$.

In (F3), the constant $K(M) > 0$ must be sufficiently large to prevent gravitational collapse: $U(\rho)$ must control $G(\rho, \rho)$ at large energies so that the Chandrasekhar mass for the model is greater than $M = 1$. (F1-F2) ensure that $A(\varrho)$ is C^1 and strictly convex with $A'(\varrho)\varrho - A(\varrho) = P(\varrho)$ on $[0, \infty)$. (F2) also ensures that a diffuse gas does not disperse to ∞. Under these assumptions, $E_0(\rho)$ will be bounded below on

$$\mathcal{R}(\mathbb{R}^3) = \{\rho \in L^{4/3}(\mathbb{R}^3) \mid \rho \geq 0 \quad \int \rho = 1\}, \tag{71}$$

assuming its minimum there [4]. The problem is formulated in $L^{4/3}(\mathbb{R}^3)$ because $E_0(\rho) \leq C$ and (F3) imply a bound on $\|\rho\|_{4/3}$. Results are also required regarding the non-rotating minimizer σ_m of $E_0(\rho)$ among configurations of mass $m < 1$, and the corresponding minimum energy

$$e_0(m) := \inf_{\rho \in \mathcal{R}(\mathbb{R}^3)} E_0(m\rho) = E_0(\sigma_m). \tag{72}$$

Drawn from [4, 13], these are summarized in Theorem 14.5 below.

In the presence of rotation, it is convenient to formulate the variational problem on a subset $\mathcal{R}_0(\mathbb{R}^3)$ of $\mathcal{R}(\mathbb{R}^3)$:

$$\mathcal{R}_0(\mathbb{R}^3) := \{\rho \in \mathcal{R}(\mathbb{R}^3) \mid \overline{\mathbf{x}}(\rho) = 0 \quad \text{spt}\, \rho \text{ is bounded}\}. \tag{73}$$

Here spt ρ denotes the *support* of ρ, the smallest closed set carrying the full mass of ρ. Bounded support ensures that ρ has a center of mass and finite moments of inertia. It is not implausible that solutions to (61) will have bounded support, since that equation is trivially satisfied where ρ vanishes. The velocity fields $\mathbf{v}$ will be taken to lie in $\mathcal{V}(\mathbb{R}^3) := \{\mathbf{v} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 \text{ measurable}\}$.

For prescribed angular momentum $\mathbf{J} \neq 0$, the energy $E(\rho, \mathbf{v})$ is bounded below by the non-rotating energy $e_0(1)$. However, Example 14.6 demonstrates that this bound — although approached — will not be attained. Thus the search for a global energy minimizer will be futile, and one is forced to settle for *local minimizers* of $E(\rho, \mathbf{v})$ in an appropriate topology. However, the choice of topology on $\mathcal{R}(\mathbb{R}^3)$ is quite delicate: Remark 14.7 shows that local energy minimizers will not exist if this topology is inherited from a topological vector space. Instead, $\mathcal{R}(\mathbb{R}^3)$ is topologized via the Wasserstein L^∞ metric of the probability literature. This metric is defined in Chapter 16 and denoted by W_∞. The velocity fields $\mathbf{v}$ may be topologized in any way which makes $\mathcal{V}(\mathbb{R}^3)$ a topological vector space. *Local* and *continuous* refer to these topologies hereafter.

With these definitions, Theorem 13.1 may be stated; it collects the results of Chapter 14 and 16. Its conclusions apply to energy minimizers subject only to a constraint on the z-component $J_z(\rho, \mathbf{v}) := \hat{\mathbf{z}} \cdot \mathbf{J}(\rho, \mathbf{v})$ of the angular momentum, but are extended to the case of physical interest by the corollary and remark following. That local minimizers exist in the form of binary stars for large $\mathbf{J}$ is the content of Theorem 17.1 and its corollary. Both theorems are proved by adapting the approach of [4] to the context of W_∞-local energy minimizers.

Two final definitions are required: let $[\lambda]_+ := \max\{\lambda, 0\}$; and for $\delta > 0$ define the *δ-neighbourhood* of $\Omega \subset \mathbb{R}^3$ to be the set

$$\Omega + B_\delta(0) := \bigcup_{y \in \Omega} \{\mathbf{x} \in \mathbb{R}^3 \mid |\mathbf{x} - \mathbf{y}| < \delta\}. \tag{74}$$

Theorem 13.1 (Properties of W_∞-Local Energy Minimizers)
Let $J > 0$. If $(\rho, \mathbf{v})$ minimizes $E(\rho, \mathbf{v})$ locally on $\mathcal{R}_0(\mathbb{R}^3) \times \mathcal{V}(\mathbb{R}^3)$ subject to the constraint $J_z(\rho, \mathbf{v}) = J$ then:

(i) the z-axis is a principal axis of inertia for ρ, with a moment of inertia $I(\rho)$ which is maximal and non-degenerate;

(ii) the rotation is uniform: $\mathbf{v}(\mathbf{x}) := (J\hat{\mathbf{z}} \times \mathbf{x})/I(\rho)$;

(iii) ρ is continuous on $\mathbb{R}^3$;

(iv) on each connected component Ω_i of $\{\rho > 0\}$, ρ satisfies

$$A'\big(\rho(\mathbf{x})\big) \;=\; \left[\frac{J^2}{2I^2(\rho)} r^2(\mathbf{x}) + V\rho(\mathbf{x}) + \lambda_i\right]_+ \tag{75}$$

for some chemical potential $\lambda_i < 0$ depending on the component;

(v) the equations (75) continue to hold on a δ-neighbourhood of the Ω_i;

(vi) where ρ is positive, it has as many derivatives as the inverse of $A'(\varrho)$;

(vii) if $P(\varrho)$ is continuously differentiable on $[0, \infty)$ then ρ satisfies (61);

(viii) this solution is stable with respect to L^∞-small perturbations of the Lagrangian fluid variables.

Corollary 13.2 (Local Minimizers with $J_z = J$ have $\mathbf{J}(\rho, \mathbf{v}) = J\hat{\mathbf{z}}$)
Let $J > 0$. Suppose $(\rho, \mathbf{v})$ minimizes $E(\rho, \mathbf{v})$ locally on $\mathcal{R}_0(\mathbb{R}^3) \times \mathcal{V}(\mathbb{R}^3)$ subject to the constraint $J_z(\rho, \mathbf{v}) = J$. Then $(\rho, \mathbf{v})$ also minimizes $E(\rho, \mathbf{v})$ locally subject to the constraints $\mathbf{J}(\rho, \mathbf{v}) = J\hat{\mathbf{z}}$.

Proof: Theorem 13.1(i-ii) shows that the angular momentum of $(\rho, \mathbf{v})$ satisfies the constraints $\mathbf{J}(\rho, \mathbf{v}) = J\hat{\mathbf{z}}$ of the more restricted minimization. QED.

Remark 13.3 Although the proof is not given, the converse to Corollary 13.2 is also true provided the topology on $\mathcal{V}(\mathbb{R}^3)$ enjoys a little more structure: the map taking $\mathbf{w} \in \mathbb{R}^3$ to $\mathbf{v}(\mathbf{x}) := \mathbf{w} \times \mathbf{x} \in \mathcal{V}(\mathbb{R}^3)$ should be continuous. The proof requires Remark 14.3, and the observation that a local energy minimizer subject to the constraint $\mathbf{J}(\rho, \mathbf{v}) = \mathbf{J}$ must rotate about a principal axis with maximal moment of inertia. Otherwise a slight rotation would lower its energy. To exploit this observation, it is necessary to know that slight rotations are *local* perturbations in $\mathcal{R}_0(\mathbb{R}^3) \times \mathcal{V}(\mathbb{R}^3)$, but this follows from the topology on $\mathcal{V}(\mathbb{R}^3)$ and Lemma 16.1(iii).

Remark 13.4 (Stationarity Conditions for Energy Minimizers) The Euler-Lagrange equation for a global energy minimizer, or indeed any critical point of the functional $E(\rho, \mathbf{v})$, differs from Theorem 13.1(iv) in that (75) would be satisfied on all of $\mathbb{R}^3$ for a single chemical potential λ_i. Since the Navier-Stokes equation (61) follows from (75) by taking a gradient and multiplying by ρ, it will be satisfied whether or not $\lambda_i = \lambda_j$ on different connected components of $\{\rho > 0\}$. Conversely, for $\rho \in \mathcal{R}_0(\mathbb{R}^3)$ to be a solution of (61), an integration shows that the conclusion of Theorem 13.1(iv) is necessary as well as sufficient.

14 Uniform Rotation

This chapter recounts several results which, although not original, will be required for the analysis. In particular, it is shown that the problem of minimizing $E(\rho, \mathbf{v})$ is equivalent to a minimization in which the fluid rotates uniformly about its center of mass; this reduction is due to Elliott Lieb [39]. Results regarding the minimization of the non-rotating energy are also recalled. Used here to demonstrate that the energy of a rotating star — though bounded below — cannot attain its minimum, they will also be required in Chapter 17.

Since the z-component of the angular momentum is specified, the moment of inertia $I(\rho)$ of $\rho \in \mathcal{R}_0(\mathbb{R}^3)$ in the direction of $\hat{\mathbf{z}}$ will be relevant; in cylindrical co-ordinates $(r(\mathbf{x}), \phi(\mathbf{x}), z(\mathbf{x}))$ it is given by

$$I(\rho) := \int r^2\Big(\mathbf{x} - \overline{\mathbf{x}}(\rho)\Big)\; d\rho(\mathbf{x}). \tag{76}$$

Proposition 14.1 (Uniform Rotation around Center of Mass [39]) *Fix a fluid density $\rho \in \mathcal{R}_0(\mathbb{R}^3)$ and $J \geq 0$. Among all velocities $\mathbf{v} \in \mathcal{V}(\mathbb{R}^3)$ for which $T(\rho, \mathbf{v}) < \infty$ and satisfying the constraint $J_z(\rho, \mathbf{v}) = J$, the kinetic energy $T(\rho, \mathbf{v})$ is uniquely minimized by a uniform rotation $\mathbf{v}(\mathbf{x}) := \omega\hat{\mathbf{z}} \times \mathbf{x}$ with angular velocity $\omega := J/I(\rho)$.*

Proof: Let $\mathcal{H} := L^2(\mathbb{R}^3, d\rho(\mathbf{x})) \subset \mathcal{V}(\mathbb{R}^3)$ denote the Hilbert space of vector fields on $\mathbb{R}^3$, with inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ determined by $\langle \mathbf{v}, \mathbf{v} \rangle_{\mathcal{H}} := 2T(\rho, \mathbf{v})$. The uniform rotation $\hat{\mathbf{z}} \times \mathbf{x} \in \mathcal{H}$, while the velocities $\mathbf{v}$ of interest lie in the affine subspace $\mathcal{G} \subset \mathcal{H}$ where the constraint $\langle \mathbf{v}, \hat{\mathbf{z}} \times \mathbf{x} \rangle_{\mathcal{H}} = J$ is satisfied. Minimizing the norm $\langle \mathbf{v}, \mathbf{v} \rangle_{\mathcal{H}}$ over $\mathcal{G}$ yields $\mathbf{v}_g := \omega(\hat{\mathbf{z}} \times \mathbf{x})$: any other $\mathbf{v} \in \mathcal{G}$ differs from $\mathbf{v}_g$ by a vector orthogonal to $\hat{\mathbf{z}} \times \mathbf{x}$ in $\mathcal{H}$. QED.

Corollary 14.2 (Local Energy Minimizers Rotate Uniformly)
Let $J \geq 0$. If $(\rho, \mathbf{v})$ minimizes $E(\rho, \mathbf{v})$ locally on $\mathcal{R}_0(\mathbb{R}^3) \times \mathcal{V}(\mathbb{R}^3)$ subject to the constraint $J_z(\rho, \mathbf{v}) = J$, then $\mathbf{v}(\mathbf{x}) = \omega \hat{\mathbf{z}} \times \mathbf{x}$ with $\omega := J/I(\rho)$.

Proof: The curve $(1-t)\mathbf{v}(\mathbf{x}) + t\,\omega(\hat{\mathbf{z}} \times \mathbf{x})$ is continuous in the topological vector space $\mathcal{V}(\mathbb{R}^3)$, and the linear constraint is satisfied along it. Moreover, $T(\rho, \mathbf{v})$ and hence $E(\rho, \mathbf{v})$ is a quadratic function of t along this curve, assuming its minimum at $t = 1$ by Proposition 14.1. Thus $(\rho, \mathbf{v})$ cannot be a local minimum unless $\mathbf{v}(\mathbf{x}) = \omega \hat{\mathbf{z}} \times \mathbf{x}$. QED.

Remark 14.3 (Uniform Rotation when $\mathbf{J}(\rho, \mathbf{v})$ is Prescribed)
The proofs of Proposition 14.1 and its corollary extend to the case where the linear constraint $J_z(\rho, \mathbf{v}) = J$ is replaced by three linear constraints $\mathbf{J}(\rho, \mathbf{v}) = J\hat{\mathbf{z}}$. The conclusion then is that $\mathbf{v}(\mathbf{x}) = \mathbf{w} \times \mathbf{x}$, where $\mathbf{w} \in \mathbb{R}^3$ is the unique angular velocity compatible with the given density ρ and angular momentum $\mathbf{J}$. Of course, the axis $\mathbf{w}$ of rotation may not coincide with the z-axis.

If $\rho \in \mathcal{R}_0(\mathbb{R}^3)$ rotates with velocity $\mathbf{v}(\mathbf{x}) = (J\hat{\mathbf{z}} \times \mathbf{x})/I(\rho)$, then its kinetic energy $T(\rho, \mathbf{v})$ is given by

$$T_J(\rho) := \frac{J^2}{2I(\rho)}. \tag{77}$$

A second corollary shows that the minimization of Theorem 13.1 is equivalent to the minimization of

$$E_J(\rho) := U(\rho) - G(\rho, \rho)/2 + T_J(\rho). \tag{78}$$

Corollary 14.4 (Velocity Free Reformulation)
Let $J \geq 0$ and $\rho \in \mathcal{R}_0(\mathbb{R}^3)$, and define $\omega := J/I(\rho)$. Then $(\rho, \mathbf{v})$ minimizes $E(\rho, \mathbf{v})$ locally on $\mathcal{R}_0(\mathbb{R}^3) \times \mathcal{V}(\mathbb{R}^3)$ subject to the constraint $J_z(\rho, \mathbf{v}) = J$ if and only if ρ minimizes $E_J(\rho)$ locally on $\mathcal{R}_0(\mathbb{R}^3)$ and $\mathbf{v}(\mathbf{x}) = \omega \hat{\mathbf{z}} \times \mathbf{x}$.

Proof: Assume $(\rho, \mathbf{v})$ minimizes $E(\rho, \mathbf{v})$ locally on $\mathcal{R}_0(\mathbb{R}^3)$ subject to the constraint $J_z(\rho, \mathbf{v}) = J$. By Corollary 14.2, $\mathbf{v}(\mathbf{x}) = \omega\hat{\mathbf{z}} \times \mathbf{x}$, whence $T(\rho, \mathbf{v}) = T_J(\rho)$. Lemma 16.1(v) shows that $I(\rho)$ is continuous on $\mathcal{R}_0(\mathbb{R}^3)$, therefore ρ' sufficiently close to ρ ensures that $\omega' := J/I(\rho')$ differs little from ω. Because $\mathcal{V}(\mathbb{R}^3)$ is a topological vector space, $\mathbf{v}'(\mathbf{x}) := \omega'\hat{\mathbf{z}} \times \mathbf{x}$ can be made close to $\mathbf{v}$. Since $(\rho, \mathbf{v})$ is a local energy minimum, taking ρ' closer to ρ if necessary ensures $E(\rho, \mathbf{v}) \leq E(\rho', \mathbf{v}') = E_J(\rho')$, establishing one implication.

The other implication is easier. Assume ρ minimizes $E_J(\rho)$ locally, and define $\mathbf{v}(\mathbf{x}) := \omega\hat{\mathbf{z}} \times \mathbf{x}$. For ρ' near ρ and any $\mathbf{v}' \in \mathcal{V}(\mathbb{R}^3)$, Proposition 14.1 yields $E(\rho', \mathbf{v}') \geq E_J(\rho') \geq E_J(\rho) = E(\rho, \mathbf{v})$. QED.

The analysis will henceforth be devoted to $E_J(\rho)$. Some results regarding the non-rotating problem $J = 0$ are required. Implications of [4, Theorems A and B] and [13, Theorem 3(b,d,e)] are summarized here. Results from the latter are statedly explicitly for the Chandrasekhar equation of state, but apply equally well to all $A(\varrho)$ consistent with (F1-F3). If, in addition, $A'(\varrho^3)$ is convex, uniqueness of minimizer up to translation is also known [13, Lemma 11 and remark following].

Theorem 14.5 (Non-rotating Stars [4, 13])
For $E_0(\rho)$ from (78), $e_0(m)$ from (72) and $m \in [0, 1]$:

(i) $E_0(\rho)$ attains its minimum $e_0(m)$ among ρ such that $m^{-1}\rho \in \mathcal{R}(\mathbb{R}^3)$;

(ii) $e_0(m)$ decreases continuously from $e_0(0) = 0$ and is strictly concave;

There are bounds $R_0(m)$ and $C_0(m)$ on the radius and central density, such that any mass m minimizer σ_m of $E_0(\rho)$ satisfies

(iii) σ_m is spherically symmetric and radially decreasing after translation;

(iv) $\|\sigma_m\|_\infty \leq C_0(m)$;

(v) $\operatorname{spt} \sigma_m$ is contained in a ball of radius $R_0(m)$;

(vi) σ_m is continuous; where positive it has as many derivatives as the inverse of $A'(\varrho)$;

(vii) σ_m *satisfies (75) on all of* $\mathbb{R}^3$ *for* $J = 0$ *and a single* $\lambda < 0$*;*

(viii) the left and right derivatives of $e_0(m)$ *bound* λ*:* $e_0'(m^+) \leq \lambda \leq e_0'(m^-)$*.*

For a rotating star $J > 0$, it has already been asserted that the lower bound $E_J(\rho) \geq e_0(1)$ is approached but not attained on $\mathcal{R}(\mathbb{R}^3)$. That it cannot be attained is now clear: $E_0(\rho) \geq e_0(1)$ while $T_J(\rho) \geq 0$; when the first inequality is saturated, Theorem 14.5(v) forces the second inequality to be strict. The following example uses Theorem 14.5(ii) to construct $\rho \in \mathcal{R}_0(\mathbb{R}^3)$ with $E_J(\rho)$ arbitarily close to $e_0(1)$.

Example 14.6 (No Constrained Minimum of $E(\rho, \mathbf{v})$ is Attained)
Let $J > 0$ and σ_m and $\sigma_{(1-m)}$ be the non-rotating energy minimizers of masses m and $1 - m$ respectively. From Theorem 14.5(ii), $e_0(m) + e_0(1-m)$ approximates $e_0(1)$ for $m > 0$ sufficiently small. Since σ_m has a finite radius, $|\mathbf{y}|$ sufficiently large yields a trial function $\rho(\mathbf{x}) := \sigma_m(\mathbf{x}) + \sigma_{(1-m)}(\mathbf{x} - \mathbf{y})$ with energy

$$E_J(\rho) = e_0(m) + e_0(1-m) - G(\sigma_{(1-m)}, \sigma_m) + T_J(\rho). \qquad (79)$$

Taking $|\mathbf{y}|$ larger if necessary forces $T_J(\rho)$ to be small since

$$I(\rho) = I(\sigma_m) + I(\sigma_{(1-m)}) + m(1-m)|\mathbf{y}|^2. \qquad (80)$$

Thus $E_J(\rho)$ can be made to approach the energy $e_0(1)$ of the non-rotating minimizer.

Remark 14.7 (No Local Minimizers in a Vector Space Topology)
The preceding example showed that the search for a global minimizer will be fruitless. More is true: for $E_J(\rho)$ to have even local minimizers, the topology on $\mathcal{R}_0(\mathbb{R}^3)$ must not be inherited from a topological vector space. Otherwise, a local minimum $\rho \in \mathcal{R}_0(\mathbb{R}^3)$ would be stable with respect to

all perturbations $\rho + t\sigma \in \mathcal{R}_0(\mathbb{R}^3)$; that is, $t > 0$ sufficiently small would imply $E_J(\rho + t\sigma) \geq E_J(\rho)$. The resulting stationarity condition would be (75), satisfied on on all of $\mathbb{R}^3$ for a fixed λ_i. But this is absurd: it implies $\rho(\mathbf{x}) \to \infty$ as $r(\mathbf{x}) \to \infty$. Stated physically, it is energetically favorable to slow down a rotating star by removing a small bit of mass to a far away orbit, where it carries little kinetic energy but great angular momentum.

15 Fluid in a Tube: A Toy Model

Before proceeding with the analysis of the three-dimensional problem, a one-dimensional toy model is introduced which illustrates a number of subtleties. This model represents an interacting fluid, constrained to live in a long light tube, and rotating end-over-end about its center of mass. The interaction is one-dimensional Coulomb attraction — force independent of distance — while the equation of state is taken to be $P(\varrho) = c\varrho^2$. As in the three-dimensional problem, the energy (81) of a mass of fluid carrying angular momentum J assumes its minimum only in the non-rotating case $J = 0$. However, the (one-dimensional) Euler-Poisson system (88) is explicitly soluble for this model, and a complete catalog of solutions may be obtained. These fall into an uncountable number of disjoint families or *sequences*, each parameterized continuously by the angular velocity $\omega = J/I(\rho) > 0$ up to some *critical value* ω_c. Beyond ω_c the sequence fails to exist. The solutions with connected support — single stars — begin with the non-rotating minimizer and persist as long as J is not too large. Each remaining family persists for J not too small, and consists of configurations in which a number of components with fixed masses 'orbit' each other; these represent binary stars or stellar systems in the astrophysical analogy.

The absence of bifurcations in this model should be emphasized. In problems of stellar evolution, bifurcations along equilibrium sequences raise interesting cosmological possibilities. For example, a theory of formation of double stars known as the fission hypothesis [40, 41] asserts that as ω is increased by gravitational contraction, a single star may deform *quasi-statically* into a binary system. Proposed by Kelvin and Tait before the turn of the century, this conjecture has not yet been rigorously resolved even in the context of the homogeneous incompressible model in $\mathbb{R}^3$. On the other hand, numerical studies [42] show that bifurcations do not occur in axisymmetric uniformly rotating models with polytropic equations of state $P(\varrho) = \varrho^q$ in which $q < 2.24$. Instead, the axisymmetric equilibria remain stable up to a point of 'equatorial break-up'. This is also the case in our toy model. There cooling or contraction may be represented by decreasing c at fixed J, which, after rescaling units, is equivalent to increasing J at fixed c. For a single star, ω increases with J; the radius grows, and the atmosphere near the surface becomes thinner and thinner until it is no longer gravitationally bound. For larger J there is no nearby equilibrium and the family ends. The same mechanism is responsible for the demise of the other equilibrium sequences as well. In these sequences however, ω varies *inversely* with J at large angular momentum: $\omega \to 0$ as $J \to \infty$. In this limit, the components approximate non-rotating minimizers of the same masses, placed so far apart that the system rotates very slowly. For larger ω, the stars draw closer together and the stellar material becomes less concentrated; equilibrium persists only as long as the atmosphere of the lightest star (or planet) continues to be bound.

In our one-dimensional model, the state of the fluid is represented by its mass density $\rho(x) \geq 0$ on the line; its total mass is is taken to be M and its center of mass to lie at the origin. If the whole tube rotates about this center of mass, the energy of the fluid is given by

$$E_J(\rho) := \int_{\mathbb{R}} \rho^2(x)\, dx \;+\; \frac{1}{2} \iint d\rho(x)\, |x-y|\, d\rho(y) \;+\; \frac{J^2}{2I(\rho)}. \qquad (81)$$

Units of mass, length and energy may be fixed to ensure $c = M = G = 1$, where G is the 'gravitational' constant — the coefficient of the potential energy. The angular momentum J scales with $(M^6 c^3/G)^{1/4}$, and the moment of inertia $I(\rho)$ is given by

$$I(\rho) := \int_{\mathbb{R}} x^2 d\rho(x). \qquad (82)$$

The energy $E_J(\rho)$ is defined on the space $\mathcal{R}_0(\mathbb{R}) \subset L^2(\mathbb{R})$ of densities $\rho(x)$ with bounded support and satisfying the constraints

$$\rho(x) \geq 0, \tag{83}$$

$$\int d\rho(x) = 1, \tag{84}$$

$$\int x\, d\rho(x) = 0. \tag{85}$$

Defining $[x]_+ := \max\{x, 0\}$, any minimizer of $E_J(\rho)$ on $\mathcal{R}_0(\mathbb{R})$ must be a pointwise a.e. solution to the Euler-Lagrange equation

$$2\rho(x) = \left[\frac{J^2 x^2}{2I^2(\rho)} - V\rho(x) + \lambda\right]_+ . \tag{86}$$

Here λ is the Lagrange multiplier conjugate to the mass constraint, while $V\rho$ is the gravitational potential

$$V\rho(x) := \int |x - y|\, d\rho(y). \tag{87}$$

For the real model, (86) is established rigorously in [4] (see Chapter 16); for the toy model, the proof would be similar. However, unless $J = 0$, (86) can have no solutions in $\mathcal{R}_0(\mathbb{R})$, thus $E_J(\rho)$ is not minimized there: since $V\rho(x)$ grows no faster than linearly for $\rho \in \mathcal{R}_0(\mathbb{R})$, any solution of (86) would diverge quadratically as $|x| \to \infty$, violating the mass constraint.

On the other hand, the equations corresponding to the Euler-Poisson system (61) are quite easy to solve; they are obtained by differentiating (86-87) to yield:

$$2\rho'(x) + 2M(x) - M(\infty) - \frac{J^2 x}{I^2(\rho)} = \xi \quad \text{where } \rho(x) > 0 \tag{88}$$

and $M(x) := \int_{-\infty}^{x} \rho$. Here $\xi = 0$ if ρ has its center of mass at the origin. The sections below classify all continuous solutions $\rho \in \mathcal{R}_0(\mathbb{R})$ to these equations; their properties are immediate from the exact solutions.

Solutions with Connected Support

All continuous solutions $\rho \in \mathcal{R}_0(\mathbb{R})$ of (88), differentiable where positive, must be C^∞ there: $M(x)$ gains regularity from ρ and the result follows by a bootstrap. Thus ρ satisfies

$$\rho''(x) + \rho(x) = \omega^2/2 \quad \text{where } \rho(x) > 0, \tag{89}$$

for some ω. Conversely, any solution ρ to (89) with $\{\rho > 0\}$ connected also solves (88) for $J = \omega I(\rho)$ and some ξ. Such a star, if it has radius r and center of mass at the origin, can only be of the form

$$\rho_r(x) := \frac{\eta(r)}{2}\left(1 - \frac{\cos(x)}{\cos(r)}\right) \ \text{ if } x \in [-r, r], \qquad 0 \text{ otherwise.} \tag{90}$$

r must lie in $[\pi/2, \pi]$, while the normalization constant $\eta(r) := (r - \tan r)^{-1}$ for unit mass. The angular velocity required to sustain ρ_r is related to r by $\omega^2 = 2\rho_r''(r) + 2\rho_r(r) = \eta(r)$. (88) is satisfied for $J = \omega I(\rho_r)$ and some ξ, while $\rho_r'(0) = 0$ and $M(0) = 1/2$ imply $\xi = 0$. Finally, $\eta(r)$ increases from 0 to π^{-1} on $[\pi/2, \pi]$, so ω parameterizes the sequence as it ranges from 0 to $\omega_c(1) = \pi^{-1/2}$.

For single stars it remains to demonstrate that $J = I(\rho_r)\omega$ varies directly with ω. Since ω increases with r, it suffices to show that $I(\rho_r)$ is also increasing. For $r < r'$, $\rho_r(x) = \rho_{r'}(x)$ is solved at a unique value of $|x| < r'$; $I(\rho_r) < I(\rho_r')$ therefore follows from (82). Thus J attains its maximal value for $r = \pi$. At this value, the density gradient at the star's surface vanishes: $\rho_r'(r) = \eta(\pi)\tan(\pi) = 0$; the pressure gradient must vanish as well, so the fluid at the surface is 'in orbit'. For larger angular momentum this outermost fluid cannot remain contiguous with the star in equilibrium.

Solutions with Disconnected Support

Having enumerated the solutions corresponding to single stars, it remains to consider solutions ρ to (88) for which $\{\rho > 0\}$ is disconnected. These must also satisfy (89). If the interval $(z-r, z+r)$ is a connected component of $\{\rho > 0\}$, then the retriction of ρ to this interval must be $m\rho_r(x-z)$ for some mass $m < 1$. As before, $r \in [\pi/2, \pi]$. One can then ask: given an ordered n-tuple of masses satisfying $m_1 + \ldots + m_n = 1$, for which angular velocities will there be a solution $\rho \in \mathcal{R}_0(\mathbb{R})$ given by

$$\rho(x) = \sum m_i \rho_{r_i}(x - z_i) \tag{91}$$

for some radii r_i and centers z_i, ordered so that $z_i + r_i \leq z_{i+1} - r_{i+1}$. All solutions to (88) in $\mathcal{R}_0(\mathbb{R})$ must be of this form: a star cannot have infinitely many planets with radii $r \geq \pi/2$ and also have bounded support. Below it is demonstrated that exactly one such solution exists for each $\omega > 0$ up to some critical value $\omega_c(m_1, m_2, \ldots, m_n) < \infty$. The sizes of the components vary inversely with their masses, and it is easiest to parameterize the sequence in terms of the radius r of the lightest component. $r = \pi$ at the critical value $\omega = \omega_c(m_1, \ldots, m_n)$, while $r \to \pi/2$ (the radius of the non-rotating minimizer) and $J \to \infty$ as $\omega \to 0$.

If (91) is to satisfy (89), it is necessary that $\rho''(z_i + r_i) = \omega^2/2$ independently of i. Thus the radii must satisfy

$$\eta(r_i) = \omega^2/m_i. \tag{92}$$

Since $\eta(r)$ increases from 0 to π^{-1} on $[\pi/2, \pi]$, these equations are soluble provided $\omega^2/m \leq \pi^{-1}$ for the lightest mass m. Conversely, $\omega > 0$ may selected by precribing the radius $r \in (\pi/2, \pi]$ of the lightest component, in which case the remaining radii are uniquely determined. (89) will be satisfied, provided the centers z_i are chosen far enough apart so that the components do not overlap. This will be verified a posteriori. With $J/I(\rho)$ replaced by ω, (88) will also be satisfied on each component separately if the constant of integration ξ is allowed to depend on the component. The trick is to choose the centers so that $\xi = 0$ for all i. Computing (88) at $x = z_i$ where $\rho'(z_i) = 0$, it is clear that $\xi = 0$ is equivalent to

$$z_i := \omega^{-2} \left(\sum_{j<i} m_j - \sum_{j>i} m_j \right). \tag{93}$$

$\sum m_i z_i = 0$ follows, proving that ρ has its center of mass at the origin. At this point $I(\rho)$ may be determined, and (88) will be satisfied with $J = I(\rho)\omega$. A posteriori, one notes that $z_{i+1} - z_i = \omega^{-2}(m_{i+1} + m_i) \geq 2\pi$; since $r_i \leq \pi$ there is no danger of overlapping components. However, if the lightest component has radius π and the *same mass* as one of its neighbours, these components will just touch.

The foregoing is summarized by:

Theorem 15.1 (Catalog of One-dimensional Equilibria)
Choose the number of components $n \geq 1$ and their masses $(m_1, \ldots, m_n)$, ordered from left to right and with $\sum m_i = 1$. The radius $r \in (\pi/2, \pi]$ of the lightest component may also be specified. Then there is unique solution (91) to (88) in $\mathcal{R}_0(\mathbb{R})$ with the given parameters. The angular velocity ω and radii r_i of any heavier components are determined by (92), while the locations z_i of the components are determined by (93). $\xi = 0$ and $J = \omega I(\rho)$. All continuous $\rho \in \mathcal{R}_0(\mathbb{R})$ which solve (88) with $J > 0$ are of this form. The r_i and ω increase continously with r while the $|z_i|$ decrease. As $r \to \pi$, all tend to finite limiting values determined by the masses.

Having shown these solutions to exist, it is natural to remark upon their relationship to the energy functional $E_J(\rho)$. Fix a solution ρ to (88) from Theorem 15.1. While (86) cannot be satisfied for a global choice of λ, it is satisfied on each component of ρ separately: permitting λ to vary from component to component, differentiating makes this clear. Thus the mass within each component of ρ is in equilibrium. There will be perturbations $\rho + t\sigma$ in $\mathcal{R}_0(\mathbb{R})$ which lead to a linear decrease in $E_J(\rho)$, but these involve either a transfer of mass between components, or from some component(s) into the vacuum $\{\rho = 0\}$. Such perturbations involve 'tunneling' of mass from one region to another, and as such are unphysical. If they could be precluded, ρ would be a critical point for the functional (81). This is also the nature of the local minima for the three dimensional model which are investigated in the following chapters.

16 W_∞-Local Energy Minimizers

If a stability analysis is to explore local minima, a topology must be specified which determines a precise meaning for *local*. The examples of the preceding chapters illustrate that for rotating stars this choice will be delicate: for $E_J(\rho)$ to have local minima the topology must be strong enough to preclude tunneling of mass; for such minima to be meaningful, it must be weak enough so that physical flows are continuous. A topology enjoying these properties can be found in the probability literature [6]: it is induced by the Wasserstein L^∞ metric on $\mathcal{R}(\mathbb{R}^3)$. This metric is defined in the sequel, where results from [4] are applied to show that a local minimizer $\rho \in \mathcal{R}_0(\mathbb{R}^3)$ of $E_J(\rho)$ must be continuous everywhere, smooth where positive, and satisfy the stationarity condition of Theorem 13.1(iv). It follows by taking a gradient that ρ represents a stable solution to the Navier-Stokes-Poisson system (61).

Viewed as a measure, $\rho \in \mathcal{R}(\mathbb{R}^3)$ has unit mass. It may be represented in many ways as the probability distribution of a vector-valued random variable $\mathbf{x} : S \longrightarrow \mathbb{R}^3$ on a probability space $(S, \mathcal{S}, \nu)$. The relationship between $\mathbf{x}$ and ρ, here denoted by $\mathbf{x}_{\#}\nu = \rho$, is that $\nu[\mathbf{x}^{-1}(\Omega)] = \rho[\Omega]$ for Borel $\Omega \subset \mathbb{R}^3$; $\mathbf{x}$ is said to *push-forward* the measure ν to ρ. The Wasserstein L^∞ distance between two measures $\rho, \kappa \in \mathcal{R}(\mathbb{R}^3)$ may now be defined as an infimum over all random variable representations of ρ and κ on a space $(S, \mathcal{S}, \nu)$:

$$W_\infty(\rho, \kappa) := \inf_{\substack{\mathbf{x}_{\#}\nu = \rho \\ \mathbf{y}_{\#}\nu = \kappa}} \|\mathbf{x} - \mathbf{y}\|_{\infty,\nu}. \tag{94}$$

Here $\|\mathbf{x} - \mathbf{y}\|_{\infty,\nu}$ denotes the supremum of $|\mathbf{x} - \mathbf{y}|$ over S, discarding sets of ν-measure zero. Whether the infimum in (94) ranges over all probability spaces $(S, \mathcal{S}, \nu)$, or is restricted to (say) $S = [0, 1]$ with Lebesgue measure, is irrelevant. That W_∞ is a metric follows from Strassen's Theorem [6]. Note that although $W_\infty(\rho, \kappa)$ may be infinite on $\mathcal{R}(\mathbb{R}^3)$, it is finite whenever ρ and κ are of bounded support.

It is clear from the Lagrangian formulation of fluid mechanics that the Wasserstein L^∞ metric is not unphysically strong. In that formulation, the state of a fluid system is specified by its original density profile $\rho \in \mathcal{R}(\mathbb{R}^3)$, together with the positions of the fluid particles as a function of time. At time t, $\mathbf{Y}_t(\mathbf{y}) \in \mathbb{R}^3$ represents the position of the fluid which originated at $\mathbf{Y}_0(\mathbf{y}) = \mathbf{y}$. The density profile ρ_t after time t is obtained as the push-forward of ρ through $\mathbf{Y}_t$. From its definition,

$$W_\infty(\rho_s, \rho_t) \leq \|\mathbf{Y}_s - \mathbf{Y}_t\|_{\infty,\rho}. \tag{95}$$

If the fluid particles move with bounded velocities, then $\mathbf{Y}_t(\mathbf{y})$ will be a Lipshitz function of t uniformly in $\mathbf{y}$, and it is evident that (95) will be controlled by a multiple of $|s-t|$. Thus $\rho_t \in \mathcal{R}(\mathbb{R}^3)$ evolves continuously as a function of time, at least for bounded fluid velocities. The same argument shows Lemma 16.1(iii): an L^∞-small perturbation of the Lagrangian fluid variables produces only a W_∞-small perturbation of the density: a local energy minimum $\rho \in \mathcal{R}_0(\mathbb{R}^3)$ must be physically stable.

The next lemma collects elementary properties required of W_∞. The proofs are immediate from the definition (94). Here $\operatorname{spt}(\rho - \kappa) \subset \mathbb{R}^3$ denotes the support of the signed measure $\rho - \kappa$, while a *δ-neighbourhood* is defined as in (74).

Lemma 16.1 (Simple Properties of the Wasserstein L^∞ Metric)
Let $\rho, \kappa \in \mathcal{R}(\mathbb{R}^3)$. Then

(i) $W_\infty(\rho,\kappa)$ *does not exceed the diameter of* $\operatorname{spt}(\rho-\kappa)$;

(ii) if $W_\infty(\rho,\kappa) < \delta$, *each connected component of the δ-neighbourhood of* $\operatorname{spt}\rho$ *has the same mass for κ as for ρ;*

(iii) $W(\rho, \mathbf{y}_{\#}\rho) \leq \|\mathbf{y} - id\|_{\infty,\rho}$ *for* $\mathbf{y} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *measurable and* $id(\mathbf{x}) := \mathbf{x}$;

(iv) the centers of mass $|\overline{\mathbf{x}}(\rho) - \overline{\mathbf{x}}(\kappa)| \leq W_\infty(\rho,\kappa)$;

(v) the moment of inertia $I(\rho)$ depends continuously on ρ.

Lagrange multipliers conjugate to the center of mass constraint do not appear in (75). This is because a local energy minimizer ρ on $\mathcal{R}_0(\mathbb{R}^3)$ is also stable under perturbations which shift its center of mass:

Corollary 16.2 *If ρ minimizes $E_J(\rho)$ locally on $\mathcal{R}_0(\mathbb{R}^3)$, then it minimizes $E_J(\rho)$ locally on $\mathcal{R}(\mathbb{R}^3)$.*

Proof: There exists $\delta > 0$ such that $E_J(\rho) \leq E_J(\kappa)$ whenever $\kappa \in \mathcal{R}_0(\mathbb{R}^3)$ with $W_\infty(\rho,\kappa) < 2\delta$. Now, suppose $\kappa \in \mathcal{R}(\mathbb{R}^3)$ with $W_\infty(\rho,\kappa) < \delta$. Part (iv) of the lemma shows that $|\overline{\mathbf{x}}(\kappa)| < \delta$; part (iii) then shows that the translate of κ by $-\overline{\mathbf{x}}(\kappa)$ lies within 2δ of ρ in $\mathcal{R}_0(\mathbb{R}^3)$. By translation invariance, $E_J(\rho) \leq E_J(\kappa)$. QED.

Therefore, suppose ρ minimizes $E_J(\rho)$ locally on $\mathcal{R}_0(\mathbb{R}^3)$ and let $\sigma \in L^\infty(\mathbb{R}^3)$. Even if the perturbation $\rho + t\sigma \in \mathcal{R}(\mathbb{R}^3)$ for $t \in [0,1]$, it may not be W_∞-continuous as function of t; nevertheless, Lemma 16.1(i) guarantees that $E_J(\rho+t\sigma)$ is minimized by ρ provided σ is supported on a small enough set. σ will then be a useful variation of $E_J(\rho)$.

The *variational derivative* $E'_J(\rho)$ of the energy $E_J(\rho)$ is formally given by

$$E'_J(\rho)(\mathbf{x}) := A'(\rho(\mathbf{x})) - V\rho(\mathbf{x}) - \frac{J^2}{2I^2(\rho)} r^2\Big(\mathbf{x} - \overline{\mathbf{x}}(\rho)\Big). \tag{96}$$

For $J = 0$ and a restricted class of perturbations $\sigma \in P_0$, a more general result [4] not quite including the kinetic energy $T_J(\rho)$ shows that

$$\lim_{t\to 0} t^{-1}\left(E_J(\rho + t\sigma) - E_J(\rho)\right) = \int E'_J(\rho)\sigma. \tag{97}$$

The admissable perturbations depend on ρ:

$$P_0 := \bigcup_{R<\infty} \left\{ \sigma \in L^\infty(\mathbb{R}^3) \;\middle|\; \begin{array}{c} \sigma = 0 \text{ where } \rho > R \text{ or } |\mathbf{x}| > R \\ \sigma \geq 0 \text{ where } \rho < R^{-1} \end{array} \right\}. \tag{98}$$

(97) may be immediately extended to positive angular momentum $J > 0$ by the following observation: even though σ may not have its center of mass at the origin, $I(\rho)$ is cubic in ρ; a direct computation shows that $\lim_{t\to 0} t^{-1}(I(\rho + t\sigma) - I(\rho)) = I(\sigma)$.

The next pair of propositions show that local minimizers of $E_J(\rho)$ satisfy the stationarity conditions of Theorem 13.1(iv-v).

Proposition 16.3 (Locally Constant Chemical Potential)
Let $\rho \in \mathcal{R}_0(\mathbb{R}^3)$ minimize $E_J(\kappa)$ among $\kappa \in \mathcal{R}(\mathbb{R}^3)$ for which $W_\infty(\rho, \kappa) < 2\delta$. Let M be an open set with diameter no greater than 2δ which intersects spt ρ*. There is a unique $\lambda_i \in \mathbb{R}$ depending on M such that (75) holds on M a.e.*

Proof: The proof is an application of a constrained minimization argument as may be found in [4]. We rely on intermediate results formulated there. Therefore, define the convex cone $P_{loc} := \{\sigma \in P_0 \mid \operatorname{spt}\sigma \subset M\}$, and let $\mathcal{U} = \{\rho + \sigma \geq 0 | \sigma \in P_{loc}\}$, so that P_{loc} is the tangent cone of $\mathcal{U}$ at ρ. It is noted above that $E_J(\rho)$ is differentiable at ρ in the directions σ of P_{loc}. On $\mathcal{W}_{loc} = \mathcal{R}(\mathbb{R}^3) \cap \mathcal{U}$, where the mass constraint is satisfied, Lemma 16.1(i) shows that ρ minimizes $E_J(\kappa)$. Moreover, since the open set M intersects $\operatorname{spt}\rho$, it must carry positive mass under ρ. Thus there is a smaller subset $C \subset M$ of positive measure on which $\rho(\mathbf{x})$ is bounded away from zero and infinity. If χ_C is the characteristic function of this set, both $\pm\chi_C \in P_{loc}$, although $\rho \pm \chi_C \notin \mathcal{W}_{loc}$. These conditions imply that there is a unique Lagrange multiplier $\lambda \in \mathbb{R}$ such that

$$\int E_J'(\rho)\sigma \geq \lambda \int \sigma \tag{99}$$

for all $\sigma \in P_{loc}$ [4, Proposition 2]. If $E_J'(\rho) < \lambda$ on a subset $K \subset M$ which had positive measure, this subset may be taken slightly smaller so that ρ is bounded on K; $\chi_K \in P_{loc}$ would then contradict (99). On the other hand, if

$E'_J(\rho) > \lambda$ on a subset $K \subset M$ with positive measure and where $\rho > 0$, then K may be taken slightly smaller so that ρ is bounded away from zero and infinity on K; in this case $-\chi_K \in P_{loc}$ contradicts (99). Since $A'(\rho)$ in (96) vanishes precisely where ρ does, these two inequalities show that (75) holds for almost all $\mathbf{x} \in M$ with $\lambda_i := \lambda$. QED.

Proposition 16.4 (Component-wise Constant Chemical Potential) *Let $\rho \in \mathcal{R}_0(\mathbb{R}^3)$ minimize $E_J(\kappa)$ among $\kappa \in \mathcal{R}(\mathbb{R}^3)$ for which $W_\infty(\rho, \kappa) < 2\delta$. Choose one of the connected components Ω_i of the δ-neighbourhood (74) of* spt ρ. *Then there is a constant $\lambda_i < 0$ such that (75) holds a.e. on Ω_i.*

Proof: For $\mathbf{y} \in \Omega_i$ the ball $B_\delta(\mathbf{y})$ intersects spt ρ. Thus Proposition 16.3 guarantees a unique $\lambda(\mathbf{y})$ such that (75) holds a.e. on $B_\delta(\mathbf{y})$ when $\lambda_i := \lambda(\mathbf{y})$. The claim is that $\lambda(\mathbf{y})$ is independent of $\mathbf{y}$. Therefore, fix $\mathbf{y} \in \Omega_i$. Since $B_\delta(\mathbf{y})$ is open, it will also be true that a slightly smaller ball $B_{\delta-\epsilon}(\mathbf{y})$ intersects spt ρ. If $|\mathbf{x} - \mathbf{y}| < \epsilon$, then $M = B_\delta(\mathbf{x}) \cap B_\delta(\mathbf{y})$ intersects spt ρ. In Proposition 16.3, the uniqueness of λ corresponding to M forces $\lambda(\mathbf{x}) = \lambda(\mathbf{y})$. Thus $\lambda(\mathbf{y})$ is locally constant. As a result, the disjoint sets $C = \{\mathbf{x} \in \Omega_i \mid \lambda(\mathbf{x}) = \lambda(\mathbf{y})\}$ and $D = \{\mathbf{x} \in \Omega_i \mid \lambda(\mathbf{x}) \neq \lambda(\mathbf{y})\}$ are both open. Since $\Omega_i = C \cup D$ is connected, $C = \Omega_i$. Defining $\lambda_i := \lambda(\mathbf{y})$, (75) must be satisfied a.e. on Ω_i.

An additional argument shows $\lambda < 0$. Any point on the boundary of Ω_i cannot lie within δ of spt ρ. Since spt ρ is bounded, Ω_i has non-empty boundary, and it follows that $\rho(x) = 0$ on a set of positive measure in Ω_i. On the other hand, $A(\varrho)$ is strictly convex so $A'(\rho(\mathbf{x}))$ vanishes only if $\rho(\mathbf{x}) = 0$. $\lambda \geq 0$ in (75) would imply $\rho > 0$ a.e. on Ω_i, a contradiction. QED.

Arguments from [4] now apply to local minimizers on $E_J(\rho)$, yielding:

Proposition 16.5 (Regularity of W_∞ Local Energy Minimizers)
Let ρ minimize $E_J(\rho)$ locally on $\mathcal{R}_0(\mathbb{R}^3)$. Then ρ is continuous everywhere; where positive it has as many derivatives as the inverse of $A'(\varrho)$.

Proof: Proposition 16.4 applies by Corollary 16.2. The stationarity condition (75) must be used to control ρ with $V\rho$ at large densities. The chemical potential $\lambda_i < 0$ may be discarded, while $r^2(\mathbf{x})$ cannot be too large on the bounded support of ρ, so $A'(\rho(\mathbf{x})) \leq V\rho(\mathbf{x}) + C$ for $C < \infty$ depending on ρ but independent of $\mathbf{x}$. Wherever $A'(\rho) \geq 2C$, the bound $A'(\rho) \leq 2V\rho$ holds. Thus $V\rho$ is continuous on $\mathbb{R}^3$ as in [4, Lemma 3 and Theorem A].

Continuity of ρ on Ω_i follows from that of $V\rho$ through (75) because $A'(\varrho)$ is continuously invertible. Ω_i was a component of some δ-neighbourhood of spt ρ, so it is clear that ρ will be compactly supported on it. Because $V\rho$ gains a derivative from ρ, smoothness of ρ where positive follows from a bootstrap in (75). QED.

Only Theorem 13.1(i) remains to be proven:

Lemma 16.6 (Principal Axis of Inertia)
Let ρ minimize $E_J(\rho)$ locally on $\mathcal{R}_0(\mathbb{R}^3)$. Then the z-axis is a principal axis of inertia for ρ, with a moment of inertia $I(\rho)$ which is maximal and non-degenerate.

Proof: Let $I_{ij}(\rho) := \int(\delta_{ij}|\mathbf{x}|^2 - x_j x_i)d\rho(\mathbf{x})$ denote the moment of inertia tensor $\underline{\mathrm{I}}(\rho)$ of ρ, and $\hat{\mathbf{l}} \in \mathbb{R}^3$ denote the eigenvector of $\underline{\mathrm{I}}(\rho)$ corresponding to its maximal eigenvalue. Then $I(\rho) = \langle\, \hat{\mathbf{z}}, \underline{\mathrm{I}}(\rho)\hat{\mathbf{z}}\rangle \leq \langle\, \hat{\mathbf{l}}, \underline{\mathrm{I}}(\rho)\hat{\mathbf{l}}\rangle$. The first claim is that the inequality is saturated. If not, a slight rotation of ρ bringing $\hat{\mathbf{l}}$ toward the z-axis would increase $I(\rho)$: letting $\hat{\mathbf{k}}(\theta) := \cos(\theta)\hat{\mathbf{l}} + \sin(\theta)\hat{\mathbf{k}}$ where $\hat{\mathbf{k}}$ and $\hat{\mathbf{l}}$ are orthonormal, either $\langle\, \hat{\mathbf{k}}(\theta), \underline{\mathrm{I}}(\rho)\, \hat{\mathbf{k}}(\theta)\rangle$ is constant or it attains a unique local maximum at $\theta = 0$. Since $E_0(\rho)$ is rotation invariant, $E_J(\rho)$ would be decreased. But ρ minimizes $E_J(\rho)$ locally. A contradiction is produced since for ρ with bounded support, a slight rotation is a W_∞-local perturbation by Lemma 16.1(iii).

Now suppose that $\langle \hat{\mathbf{z}}, \underline{\mathrm{I}}(\rho)\hat{\mathbf{z}}\rangle$, although maximal, is not unique. Then a slight rotation of ρ (about an axis other than $\hat{\mathbf{z}}$) is also a W_∞ local minimizer of $E_J(\rho)$ on $\mathcal{R}_0(\mathbb{R}^3)$. By Propositions 16.4 and 16.5, $A'(\rho) - V\rho$ must be constant along line segments parallel to the z-axis where $\rho > 0$, and cannot be constant along line segments with other orientations. This cannot be true for both ρ and its rotate. QED.

Proof of Theorem 13.1 Let $(\rho, \mathbf{v})$ locally minimize $E(\rho, \mathbf{v})$ subject to the constraint $J_z(\rho, \mathbf{v}) = J$. Corollary 14.4 proves (ii) and implies that ρ locally minimizes $E_J(\rho)$. Parts (i), (iii, vi) and (iv-v) then follow from Lemma 16.6, and Propositions 16.5 and 16.4 respectively. If $P(\varrho)$ is continuously differentiable, then $A''(\varrho) = P'(\varrho)/\varrho$ and (vii) follows by taking the gradient of (75). By Lemma 16.1(iii), the energy cannot be decreased by perturbations of ρ which result from L^∞-small perturbations in the Lagrangian fluid variables. Perturbations of the velocity field $\mathbf{v}$ are irrelevant: if consistent with the constraint, Proposition 14.1 shows that $E(\rho, \mathbf{v})$ can only increase relative to $E_J(\rho)$. QED.

17 Existence of Binary Stars

This chapter is devoted to establishing the existence of local minimizers for $E_J(\rho)$ carrying large angular momentum J. Such minimizers represent stable, uniformly rotating solutions to the Navier-Stokes-Poisson system (61). They are constructed in form of binary stars, which is to say that the fluid mass is divided into two disjoint regions Ω^- and Ω^+, widely separated relative to their size. The mass ratio $m : 1 - m$ between the two regions is specified a priori.

The $\Omega^\pm \subset \mathbb{R}^3$ will be closed balls centered on the plane $z = 0$, whose size and separation scale with J^2 (103); the relevant fluid configurations are

$$\mathcal{W}_J := \{\rho^- + \rho^+ \in \mathcal{R}(\mathbb{R}^3) \mid \int \rho^- = m, \quad \operatorname{spt} \rho^\pm \subseteq \Omega^\pm\}. \qquad (100)$$

The following theorem will be proved:

Theorem 17.1 (Existence of Binary Stars)
Given $m \in (0,1)$, choose the angular momentum J to be sufficiently large depending on m. Then any global minimizer of $E_J(\rho)$ on $\mathcal{W}_J$ will, after a rotation about the z-axis and a translation, have support contained in the interior of $\Omega^- \cup \Omega^+$. It will also be symmetric about the plane $z = 0$ and a decreasing function of $|z|$.

Since a global energy minimizer on $\mathcal{W}_J$ exists by arguments [4, 5] summarized below, this theorem has as its consequence:

Corollary 17.2 *Given $m \in (0,1)$, let $J > J(m)$ as in Theorem 17.1. Then the energy $E_J(\rho)$ admits a local minimizer ρ on $\mathcal{R}(\mathbb{R}^3)$ in the form of a global energy minimizer on $\mathcal{W}_J$. Uniformly rotating, ρ minimizes $E(\rho, \mathbf{v})$ locally on $\mathcal{R}(\mathbb{R}^3) \times \mathcal{V}(\mathbb{R}^3)$ subject to the constraint $J_z(\rho, \mathbf{v}) = J$ or $\mathbf{J}(\rho, \mathbf{v}) = J\hat{\mathbf{z}}$.*

Proof: Let ρ be the minimizer on $\mathcal{W}_J$. The theorem shows that $\operatorname{spt}\rho$ is compact in the interior of $\Omega^- \cup \Omega^+$, therefore separated from the boundary by a positive distance δ. Lemma 16.1(ii) shows that if $\kappa \in \mathcal{R}(\mathbb{R}^3)$ with $W_\infty(\rho, \kappa) < \delta$, then κ in $\mathcal{W}_J$. Thus $E_J(\rho) \le E_J(\kappa)$. Since $E_J(\rho)$ is locally minimized, Corollary 14.4 provides a local minimizer $(\rho, \mathbf{v})$ of $E(\rho, \mathbf{v})$ subject to the constraint on J_z. Corollary 13.2 shows that $(\rho, \mathbf{v})$ satisfies the constraint on the vector angular momentum as well. QED.

The separation of the domains $\Omega^\pm$ is determined by the Kepler problem for two point masses m and $1-m$, rotating with angular momentum $J > 0$ about their fixed center of mass. The reduced mass of that system is denoted by $\mu := m(1-m)$. As a function of the radius of separation d, the gravitational plus kinetic energy

$$-\frac{\mu}{d} + \frac{J^2}{2\mu d^2} \ge -\frac{\mu^3}{2J^2} \tag{101}$$

assumes its minimum at separation $\eta := \mu^{-2}J^2$. This is the radius of the circular orbit. Therefore, choose two points $\mathbf{y}^\pm \in \mathbb{R}^3$ from the plane $z = 0$, separated by η, to be the centers of $\Omega^\pm$:

$$\eta := \mu^{-2}J^2 = |\mathbf{y}^- - \mathbf{y}^+|; \tag{102}$$

$$\Omega^\pm := \left\{ \mathbf{x} \in \mathbb{R}^3 \mid |\mathbf{x} - \mathbf{y}^\pm| \le \eta/4 \right\}. \tag{103}$$

Here and throughout the following, the superscripts $\pm$ denote an implicit dependence on J, or equivalently η. When η is large, one expects a stable, slowly rotating equilibrium to exist in which fluid components with masses m and $1-m$ lie near $\mathbf{y}^-$ and $\mathbf{y}^+$. The distance separating $\Omega^\pm$ and the diameter of their union is given by:

$$\operatorname{dist}(\Omega^-, \Omega^+) = \eta/2; \tag{104}$$

$$\operatorname{diam}(\Omega^- \cup \Omega^+) = 3\eta/2. \tag{105}$$

It follows that for $\rho \in \mathcal{W}_J$ rotating uniformly with angular momentum J, the fluid velocities will not be too large:

Lemma 17.3 (Velocity Bound)
Fix $m \in (0,1)$ and let $\epsilon > 0$. For $J \geq \epsilon$, there is a maximum velocity $v(m,\epsilon)$ which does not depend on J, such that if $\rho \in \mathcal{W}_J$ and $\mathbf{x} \in \Omega^- \cup \Omega^+$ then $Jr(\mathbf{x} - \overline{\mathbf{x}}(\rho))/I(\rho) \leq v(m,\epsilon)$. Moreover $v(m,\epsilon) \to 0$ as $\epsilon \to \infty$.

Proof: Let $\rho = \rho^- + \rho^+ \in \mathcal{W}_J$. The centers of mass $\overline{\mathbf{x}}(\rho^\pm) \in \Omega^\pm$, or rather their projections onto $z = 0$, are separated by at least $\eta/2$ (104). The moment of inertia $I(\rho)$ is bounded below by that of two point masses m and $1-m$ at this separation:

$$
\begin{aligned}
I(\rho) &= \mu\, r^2\Big(\overline{\mathbf{x}}(\rho^-) - \overline{\mathbf{x}}(\rho^+)\Big) + I(\rho^-) + I(\rho^+) && (106)\\
&\geq \mu\, \eta^2/4. && (107)
\end{aligned}
$$

At the same time $\mathbf{x} \in \Omega^\pm$ implies $r(\mathbf{x} - \overline{\mathbf{x}}(\rho)) \leq 3\eta/2$. Since $\eta = \mu^{-2}J^2$, these two estimates show $r(\mathbf{x} - \mathbf{x}(\rho))/I(\rho) \leq O(J^{-2})$ as $J \to \infty$, proving the lemma. QED.

Before addressing the proof of Theorem 17.1, the existence of a global energy minimizer on $\mathcal{W}_J$ is sketched following [4, 5]. For constants $\lambda^\pm$, such a minimizer satisfies the Euler-Lagrange equations

$$
A'\Big(\rho(\mathbf{x})\Big) = \left[\frac{J^2}{2I^2(\rho)} r^2\Big(\mathbf{x} - \overline{\mathbf{x}}(\rho)\Big) + V\rho(\mathbf{x}) + \lambda^\pm\right]_+ \text{ a.e. on } \Omega^\pm, \qquad (108)
$$

much like (75) before setting $\overline{\mathbf{x}}(\rho) = 0$. To prove existence of a minimizer, one first imposes a large bound $\|\rho\|_\infty \leq R$ on the configurations in $\mathcal{W}_J$. $E_J(\rho)$ is then lower semi-continuous in the weak topology on $\mathcal{W}_J \subset L^{4/3}(\mathbb{R}^3)$; the kinetic term $T_J(\rho)$ is continuous. Since $E_J(\rho)$ diverges with $\|\rho\|_{4/3}$, the Banach-Alaoglu compactness theorem guarantees a minimizer ρ_R. Because

ρ_R was constrained to be bounded, it satisfies a version of (108) in which the truncation $[x]_+$ is modified so that $[x]_+ = A'(R)$ when $x > A'(R)$. From this equation, an additional argument [5, Proposition 1.4] using the bound on $Jr/I(\rho_R)$ from Lemma 17.3 shows

$$\|\rho_R\|_\infty < C(m) \tag{109}$$

independent of R. This R-independent ρ_R is the desired minimizer. As in Lemma 17.3, the constant $C(m)$ is J-independent for J bounded away from zero.

Theorem 17.1 controls the support of the global minimizer $\rho = \rho^- + \rho^+$ on $\mathcal{W}_J$. Its proof begins with a series of estimates on $\rho^\pm$, the components of ρ supported in $\Omega^\pm$ respectively. Using the symmetry in m and $1-m$, it is sufficient to establish these estimates for ρ^- only. The first proposition relies on an energetic comparison with the configuration $\sigma^- + \sigma^+$ obtained from suitable translations of non-rotating minimizers σ_m from Theorem 14.5:

$$\sigma^-(\mathbf{x}) := \sigma_m(\mathbf{x} - \mathbf{y}^-); \quad \sigma^+(\mathbf{x}) := \sigma_{(1-m)}(\mathbf{x} - \mathbf{y}^+). \tag{110}$$

Proposition 17.4 (Energy Converges to Non-rotating Minimum) *Given $\epsilon > 0$, if J is sufficiently large and $\rho^- + \rho^+$ minimizes $E_J(\rho)$ on $\mathcal{W}_J$, then $E_0(\rho^-) \le e_0(m) + \epsilon$. Here $e_0(m)$ is the mass m infimum of $E_0(\rho)$.*

Proof: By Theorem 14.5(v), taking J large enough will ensure that $\sigma^\pm$ is supported in $\Omega^\pm$. Then $\sigma^- + \sigma^+ \in \mathcal{W}_J$, so its energy decomposes as

$$E_J(\sigma^- + \sigma^+) - E_0(\sigma^-) - E_0(\sigma^+) = -G(\sigma^-, \sigma^+) + T_J(\sigma^- + \sigma^+). \tag{111}$$

The gravitational interaction and kinetic energy may be estimated by comparison with the point masses (101): $G(\sigma^-, \sigma^+) = \mu\eta^{-1}$ by Newton's Theorem and (102), while $I(\sigma^- + \sigma^+) \ge \mu\eta^2$ as in (106). Thus the right side of

(111) is less than $-\mu^3 J^{-2}/2$, yielding

$$\begin{aligned} E_0(\sigma^-) + E_0(\sigma^+) &> E_J(\sigma^- + \sigma^+) \\ &\geq E_J(\rho^- + \rho^+) \\ &> E_0(\rho^-) + E_0(\rho^+) - G(\rho^-, \rho^+) \end{aligned}$$

The last inequality follows from $T_J(\rho) > 0$, since $E_J(\rho^- + \rho^+)$ also decomposes as in (111). Taking J large forces the separation $\eta/2$ between Ω^- and Ω^+ to diverge. Taking $G(\rho^-, \rho^+) \leq 2\mu\eta^{-1} < \epsilon$ proves the proposition because $E_0(\rho^+) \geq E_0(\sigma^+)$. QED.

Thus $E(\rho^-)$ converges to the minimum energy for a non-rotating mass m as $J \to \infty$. In this case [43, Theorem II.2 and Corollary II.1] a subsequence of the ρ^- may be extracted, which, after translation, converges strongly in $L^{4/3}(\mathbb{R}^3)$ to a minimizer for the non-rotating problem. The next two results exploit this convergence.

Lemma 17.5 (Bound for the Chemical Potential)
Given $\epsilon > 0$ and J large enough, if $\rho^- + \rho^+$ minimizes $E_J(\rho)$ on $\mathcal{W}_J$, then the chemical potential λ^- in (108) satisfies $\lambda^- \leq e'(m^-) + \epsilon$. Here $e'(m^-) < 0$ is the bound for the non-rotating chemical potential from Theorem 14.5(viii).

Proof: The proposition can only fail if there exists a sequence of angular momenta $J_n \to \infty$ together with minimizers $\rho_n^- + \rho_n^+$ for $E_{J_n}(\rho)$ on $\mathcal{W}_{J_n}$, for which the chemical potentials λ_n^- have a limit greater than $e'(m^-)$. The Euler-Lagrange equation (108) implies

$$A'(\rho_n^-) \geq V\rho_n^- + \lambda_n^- \text{ a.e. on } \Omega^-, \tag{112}$$

an invariant statement under translations of ρ_n^-. Proposition 17.4 and [43] imply — after translating each ρ_n^- and extracting a subsequence also denoted

ρ_n^- — that one has $L^{4/3}(\mathbb{R}^3)$ convergence to a non-rotating minimizer σ_m for $E_0(\rho)$. Since $V\rho$ is the convolution of ρ with a weak $L^3_w(\mathbb{R}^3)$ function, the Generalized Young's Inequality shows that $V\rho_n^- \to V\sigma_m$ strongly in $L^{12}(\mathbb{R}^3)$ (here $3/4 + 1/3 = 1 + 1/12$). Extracting another subsequence, one has pointwise convergence a.e. of both ρ_n^- and $V\rho_n^-$. A contradiction follows from (112) on the set $\{\sigma_m > 0\}$, where by Theorem 14.5(vii-viii):

$$A'(\sigma_m) - V\sigma_m = \lambda_m \leq e'(m^-).$$

QED.

Proposition 17.6 (Bound on the Radius of Support)
There exists a radius $R(m)$ independent of J, such that if $\rho^- + \rho^+$ minimizes $E_J(\rho)$ on $\mathcal{W}_J$ for J sufficiently large, then spt ρ^- *is contained in a ball of radius $R(m)$.*

Proof: Take J large enough that $\lambda > e_0'(m^-)$ bounds λ^- by Proposition 17.5, while the velocity bound $v(m)$ of Proposition 17.3 satisfies $v^2(m) \leq -\lambda$. In the Euler-Lagrange equation (108) these estimates yield

$$A'(\rho^-) \leq \left[V\rho^- + V\rho^+ + \lambda/2\right]_+ \text{ a.e. on } \Omega^-. \tag{113}$$

Strict convexity of $A(\varrho)$ forces $\rho = 0$ where $A'(\rho) = 0$, so ρ^- must vanish where the gravitational potential is less than $-\lambda/2$. $V\rho^+$ is easily controlled: for J large enough, $V\rho^+ < -\lambda/6$ on Ω^- since the distance to spt $\rho^+ \subset \Omega^+$ will be large (104). Therefore, consider $V\rho^-$. For $\rho \in L^1(\mathbb{R}^3) \cap L^\infty(\mathbb{R}^3)$, there is a pointwise bound

$$\|V\rho\|_\infty \leq k\|\rho\|_1^{2/3}\|\rho\|_\infty^{1/3} \tag{114}$$

saturated when ρ is supported on the smallest ball consistent with $\|\rho\|_\infty$. Since $\|\rho^-\|_\infty \leq C(m)$ from (109), choose $\delta > 0$ such that $\|\rho\|_1 \leq \delta$ and

$\|\rho\|_\infty \le C(m)$ imply $\|V\rho\|_\infty \le -\lambda/6$. Now, let $R_0(m)$ from Theorem 14.5(v) bound the support radii of all mass m non-rotating minimizers σ_m, and choose $R(m) \ge R_0(m)$ large enough so that $m/(R(m) - R_0(m)) \le -\lambda/6$. Using Proposition 17.4 and [43] once again, J large enough implies that ρ^- is $L^{4/3}(\mathbb{R}^3)$ close to a translate of some σ_m; in particular, all but mass δ of ρ^- is forced into a ball of radius $R_0(m)$. Neither the restriction of ρ^- to this ball, nor the remaining mass δ, contributes more than $-\lambda/6$ to $V\rho^-$ outside the larger ball of radius $R(m)$. Thus (113) establishes the proposition. QED.

The following lemma and proposition essentially prove Theorem 17.1.

Lemma 17.7 *For $\epsilon > 0$ define $g_\epsilon(x) := (x^2 + \epsilon^2/\mu)^{-1} - 2(x - 2\epsilon)^{-1}$. If ϵ is sufficiently small, the function $g_\epsilon(x)$ is uniquely minimized on the interval $(1/2, 3/2)$ and has no local maxima there.*

Proof: For ϵ sufficiently small, the functions $g_\epsilon(z)$ are analytic and uniformly bounded on $\{z \in \mathbb{C} \mid |z| > 1/4\}$. It follows that $g_\epsilon(z)$ converges uniformly to $g_0(z) := z^{-2} - 2z^{-1}$ as $\epsilon \to 0$ on $|z| \ge 1/2$. The derivatives converge also. $g_0'(x)$ vanishes on $(0, \infty)$ only at $x = 1$, while $g_0''(x) > 0$ for $x < 3/2$. Therefore, if $\delta < 1/2$, sufficiently small ϵ ensures: $g_\epsilon''(x) > 0$ where $|x-1| < \delta$, while $g_\epsilon'(x) < 0$ for $1/2 \le x \le 1 - \delta$ and $g_\epsilon'(x) > 0$ for $x \ge 1 + \delta$. The lemma is proved. QED.

Proposition 17.8 (Estimate for the Center of Mass Separation) *Let $0 < \delta < 1/2$. For J sufficiently large, if $\rho = \rho^- + \rho^+$ minimizes $E_J(\rho)$ on $\mathcal{W}_J$ then the ratio $|\overline{\mathbf{x}}(\rho^-) - \overline{\mathbf{x}}(\rho^+)| \,/\, \eta$ lies within δ of 1. Here $\eta = \mu^{-2} J^2$.*

Proof: Take J large enough so that Proposition 17.6 provides bounds $R(m)$ and $R(1-m)$ for the support of $\rho^\pm$. Taking J larger if necessary ensures

$$R := 2 \max\{R(m), R(1-m)\} < \eta/4. \tag{115}$$

Since spt $\rho^{\pm}$ must be contained within radius R of $\overline{\mathbf{x}}(\rho^{\pm})$, there is room in $\Omega^{\pm}$ to translate ρ^- and ρ^+ independently so that $\overline{\mathbf{x}}(\rho^{\pm}) = \mathbf{y}^{\pm}$ lie at separation η. Denote these translates by κ^- and κ^+, so that $\kappa = \kappa^- + \kappa^+ \in \mathcal{W}_J$. As in (111), $E(\kappa)$ differs only from $E(\rho)$ by terms of the form $-G(\rho^-, \rho^+) + T_J(\rho)$. These terms may be estimated using the center of mass separation d between the translates of ρ^- and ρ^+; with an abuse of notation, they are denoted by $G(d)$ and $T_J(d)$, and the moment of inertia by $I(d)$:

$$\begin{aligned} \frac{\mu}{d+2R} &< G(d) < \frac{\mu}{d-2R} \\ \mu d^2 &< I(d) < \mu d^2 + R^2 \\ \frac{J^2}{2(\mu d^2 + R^2)} &< T_J(d) < \frac{J^2}{2\mu d^2}. \end{aligned}$$

If ρ minimizes $E_J(\rho)$ on $\mathcal{W}_J$, comparison with κ forces $d := |\overline{\mathbf{x}}(\rho^-) - \overline{\mathbf{x}}(\rho^+)|$ to satisfy

$$-G(d) + T_J(d) \leq -G(\eta) + T_J(\eta). \tag{116}$$

Using the preceding estimates and $J^2 = \mu^2\eta$, the implication of (116) for the dimensionless parameter $x := d/\eta$ in terms of $\epsilon := R/\eta$ is

$$-\frac{2}{x-2\epsilon} + \frac{1}{x^2 + \epsilon^2\mu^{-1}} \leq -\frac{2}{1+2\epsilon} + 1. \tag{117}$$

This condition is satisfied for $x = 1$. However, it fails to be satisfied at $x = 1 \pm \delta$ for large J, because it does not hold in the $\epsilon \to 0$ limit. Lemma 17.7 then guarantees that for large J, (117) can hold on $x \in [1/2, 3/2]$ only when $|x-1| < \delta$. This range includes all relevant separations by (104), thus proving the proposition. QED.

Proof of Theorem 17.1 First it is shown that any minimizer $\rho = \rho^- + \rho^+$ for $E_J(\rho)$ on $\mathcal{W}_J$ may be translated so that both $\overline{\mathbf{x}}(\rho^{\pm})$ lie in the plane $z = 0$. Since the $\Omega^{\pm}$ are convex and symmetric about $z = 0$, it is enough to know that ρ enjoys a plane of symmetry $z = c$. This follows from a strong rearrangement inequality [44, Lemma 3] and Fubini's Theorem: the symmetric decreasing rearrangement of ρ along lines parallel to the z-axis leaves $U(\rho)$ and $I(\rho)$ unchanged; however, since the potential $(r^2 + z^2)^{-1/2}$ is strictly decreasing as a function of $|z|$, the rearrangement increases $G(\rho, \rho)$ unless ρ is already symmetric decreasing about a plane $z = c$. Since ρ minimizes $E_J(\rho)$ and its rearrangement is in $\mathcal{W}_J$, $G(\rho, \rho)$ cannot be increased.

Now, take J large enough so that Proposition 11.1 provides a bound R such that spt $\rho^{\pm} \subset B_R(\overline{\mathbf{x}}(\rho^{\pm}))$ if $\rho^- + \rho^+$ minimizes $E_J(\rho)$ on $\mathcal{W}_J$. Translate ρ so that its symmetry plane is $z = 0$ and let $d := |\overline{\mathbf{x}}(\rho^-) - \overline{\mathbf{x}}(\rho^+)|$. Then (103-105) show that if $d - 2R > \eta/2$ and $d + 2R < 3\eta/2$, a translation and rotation of ρ yields a minimizer in $\mathcal{W}_J$ supported away from the boundary of $\Omega^- \cup \Omega^+$. By Proposition 17.8, this is certainly true when J and hence η is sufficiently large. QED.

Part IV
Appendices

A Differentiability of Convex Functions

This appendix establishes some facts of life regarding convex functions and notation from convex analysis. Rockafellar's text [45] is the standard reference, while [17, Notes to §1.5] contains a brief synopsis of the differentiability almost everywhere of convex functions.

By a *convex function* ψ on $\mathbb{R}^d$, we shall mean what is technically called a *proper* convex function: ψ takes values in $\mathbb{R} \cup \{+\infty\}$, is not identically $+\infty$, and is convex along any line in $\mathbb{R}^d$. If ψ is convex, its *domain* $\mathrm{dom}\,\psi :=$ $\{x \mid \psi(x) < \infty\}$ will be convex and ψ will be continuous on the interior Ω of $\mathrm{dom}\,\psi$. ψ may be taken to be lower semi-continuous by modifying its values on the boundary of Ω, in which case ψ is said to be *closed*.

The convex function ψ will be differentiable ($\nabla\psi$ exists) Lebesgue-a.e. on Ω. It is also useful to consider the *subgradient* $\partial\psi$ of ψ: this parameterizes the supporting hyperplanes of ψ, and consists of pairs $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ such that $\psi(z) \geq \langle y,\, z - x\rangle + \psi(x)$ for all $z \in \mathbb{R}^d$. Here $\langle\ \ ,\ \ \rangle$ denotes the usual inner product. $\partial\psi$ should be thought of as a multivalued mapping from $\mathbb{R}^d$ to $\mathbb{R}^d$: the image of a point x is denoted by $\partial\psi(x) := \{y | (x, y) \in \partial\psi\}$, and of a set X by $\partial\psi(X) := \cup_X \partial\psi(x)$. $\partial\psi(x)$ is a closed convex set, bounded precisely when $x \in \Omega$; it is empty for x outside $\mathrm{dom}\,\psi$, and possibly for some of the boundary points as well. Differentiability of ψ at x is equivalent to the existence of a unique $y \in \partial\psi(x)$, in which case $\nabla\psi(x) = y$. $\partial\psi$ will be closed as a subset of $\mathbb{R}^d \times \mathbb{R}^d$ if ψ is a closed convex function; this property can frequently be used in lieu of continuity of $\nabla\psi$. Related expressions of the continuity of $\partial\psi$ include: compactness of $\partial\psi(K)$ when $K \subset \Omega$ is compact, and convergence of y_n to $\nabla\psi(x)$ when the latter exists and $x_n \to x$ with $y_n \in \partial\psi(x_n)$.

A subset $S \subset \mathbb{R}^d \times \mathbb{R}^d$ is said to be *cyclically monotone* if for any n points $(x_i, y_i) \in S$,

$$\langle y_1, x_2 - x_1 \rangle + \langle y_2, x_3 - x_2 \rangle + \cdots + \langle y_n, x_1 - x_n \rangle \leq 0. \tag{118}$$

The subgradient of any convex function ψ will be cyclically monotone: if one linearly approximates the change in ψ around a cycle $x_1, x_2, \ldots x_n, x_1$, a deficit must result since the approximation underestimates each step; the deficit will be finite, and the inequality in (118) strict, unless $y_i \in \partial\psi(x_{i+1})$ for each i. Conversely, any cyclically monotone set is contained in the subgradient of some convex function. This is an integrability result: if the set were known to be the gradient of a potential ψ, the two-point (n=2) inequality alone would guarantee convexity of ψ. Applied to the closure of the set $\partial\psi^* = \{(y, x) | (x, y) \in \partial\psi\}$, it implies the existence of a convex *dual* function ψ^* to ψ. Of course, ψ^* is just the *Legendre transform* of ψ, more commonly defined by

$$\psi^*(y) := \sup_{x \in \mathbb{R}^d} \langle y, x \rangle - \psi(x). \tag{119}$$

ψ^* will be closed, and $\psi^{**} \leq \psi$ with equality if and only if ψ is closed.

A convex function ψ will be twice differentiable almost everywhere on its domain in the following sense: ψ is said to be *twice differentiable* at x_0 with *Hessian* $\nabla^2\psi(x_0)$ if $\nabla\psi(x_0)$ exists, and if for every $\epsilon > 0$ there exists $\delta > 0$ such that $|x - x_0| < \delta$ and $\Lambda = \nabla^2\psi(x_0)$ imply

$$\sup_{y \in \partial\psi(x)} |y - \nabla\psi(x_0) - \Lambda(x - x_0)| < \epsilon |x - x_0|. \tag{120}$$

The Hessian $\nabla^2\psi(x_0)$ is a non-negative (i.e. positive semi-definite and self-adjoint) $d \times d$ matrix. Even though points where $\nabla\psi$ is not uniquely determined may accumulate on x_0, it is not difficult to see that many of the fundamental results pertaining to differentiable transformations remain true in this modified context. Two such results are required herein:

Proposition A.1 (Inverse Function Theorem for Monotone Maps) *Assume ψ convex on $\mathbb{R}^d$, twice differentiable at $x_0 \in \mathbb{R}^d$ in the sense of (120), so that $\nabla\psi(x_0)$ exists and $\psi < \infty$ in a neighbourhood of x_0. If $\Lambda = \nabla^2\psi(x_0)$ is invertible, then ψ^* is twice differentiable at $\nabla\psi(x_0)$ with Hessian Λ^{-1}; if Λ is not invertible then ψ^* fails to be twice differentiable at $\nabla\psi(x_0)$.*

Proof: Denote $\nabla\psi(x_0)$ by y_0. Replacing the functions $\psi(x)$ and $\psi^*(y)$ by $\psi(x + x_0) - \langle y_0, x \rangle$ and its transform $\psi^*(y + y_0) - \langle y + y_0, x_0 \rangle$, the case $y_0 = x_0 = 0$ is seen to be completely general. The first thing to show is that for Λ invertible, ψ^* is differentiable at 0 with $\nabla\psi^*(0) = 0$. This follows if $x \in \partial\psi^*(0)$ implies $x = 0$. Since the convex set $\partial\psi^*(0)$ contains the origin, it is clear that $(tx, 0) \in \partial\psi$ whenever $x \in \partial\psi^*(0)$ and $t \in [0, 1]$. For any $\epsilon > 0$, taking t small enough in (120) implies $|\Lambda x| < \epsilon|x|$. Because Λ is invertible, this forces $x = 0$. Thus $\nabla\psi^*(0) = 0$.

To show twice differentiabilty of ψ^* at 0, let $\epsilon > 0$ be small. By the continuity properties of $\partial\psi^*$ at 0, $(x, y) \in \partial\psi$ and $|y|$ sufficiently small imply $|x|$ will be small enough for (120) to hold: $|y - \Lambda x| < \epsilon|x|$. The inequality

$$\|\Lambda^{-1}\|^{-1}|\Lambda^{-1}y - x| < \epsilon|x - \Lambda^{-1}y| + \epsilon|\Lambda^{-1}y|$$

is immediate. For $\epsilon < (2\|\Lambda^{-1}\|)^{-1}$ one obtains $|x - \Lambda^{-1}y| < 2\epsilon\|\Lambda^{-1}\|^2|y|$, which expresses twice differentiability of ψ^* at 0.

Finally, the case Λ non-invertible must be addressed. Some $x \in \mathbb{R}^d$ is annihilated by Λ. From (120), there is a sequence $x_n \to 0$ of multiples of x and $(x_n, y_n) \in \partial\psi$ such that $|y_n| \leq n^{-1}|x_n|$. For any matrix Λ' and $\epsilon > 0$, taking n large violates $|x_n - \Lambda' y_n| < \epsilon|y_n|$. Thus ψ^* fails to be twice differentiable at 0. QED.

The second proposition states that the local volume distortion under the transformation $\nabla\psi$ at x is given by the determinant of $\nabla^2\psi(x)$, or in other words, that the geometric and arithmetic Jacobians agree.

Proposition A.2 (Jacobian Theorem) *Assume ψ is convex on $\mathbb{R}^d$, twice differentiable at $x_0 \in \mathbb{R}^d$ with Hessian $\Lambda := \nabla^2\psi(x_0)$ in the sense of (120). If $B_r(x_0)$ is the ball of radius r centered at x_0, then as $r \to 0$,*

$$\frac{\mathrm{vol}\,[\partial\psi(B_r(x_0))]}{\mathrm{vol}\,B_r(x_0)} \longrightarrow \det\,[\nabla^2\psi(x_0)]. \tag{121}$$

For Λ invertible, $\partial\psi(B_r(x_0))$ shrinks nicely to $\nabla\psi(x_0)$ in the sense of (31).

Proof: As in the preceding proposition, the case $x_0 = \nabla\psi(x_0) = 0$ is quite general. Assume Λ invertible. Denote $B_r(0)$ by B_r, and its image under Λ by ΛB_r. Given $\epsilon > 0$, for $r < \delta$ from (120) it is immediate that

$$\partial\psi(B_r) \subset (1 + \epsilon\|\Lambda^{-1}\|)\Lambda B_r. \tag{122}$$

On the other hand, ψ^* is twice differentiable with Hessian Λ^{-1} at 0 by Proposition A.1. The same argument, applied to ΛB_r instead of B_r, shows that for r small enough $\partial\psi^*(\Lambda B_r) \subset (1 + \epsilon\|\Lambda\|)B_r$. Taking r smaller if necessary, so that $(1 + \epsilon\|\Lambda\|)^{-1}\Lambda B_r$ lies in the interior of $\mathrm{dom}\,\psi^*$, duality yields

$$(1 + \epsilon\|\Lambda\|)^{-1}\Lambda B_r \subset \partial\psi(B_r). \tag{123}$$

Since $\epsilon > 0$ was arbitrary, (121) follows from (122-123) in the limit $r \to 0$, with the identity $\det\,[\Lambda] = \mathrm{vol}\,[\Lambda B_r]/\mathrm{vol}\,B_r$. For small r, it is evident from (122-123) that $\partial\psi(B_r)$ is nicely shrinking: i.e. it is contained in a family of balls $B_{R(r)}$ for which $R(r) \to 0$ with r, while $\partial\psi(B_r)$ occupies a fraction of $B_{R(r)}$ which is bounded away from zero.

Finally, Λ non-invertible must be dealt with. In this case ΛB_r lies in a $d-1$ dimensional subspace of $\mathbb{R}^d$. Given $\epsilon > 0$, if $(x, y) \in \partial\psi$ for small enough

$|x|$, (120) implies that $|y - \Lambda x| < \epsilon|x|$. Thus $\operatorname{vol} \partial\psi[B_r] \leq 2\epsilon(\|\Lambda\| + \epsilon)^{d-1} c r^d$, where c is the measure of the unit ball in $\mathbb{R}^{d-1}$. Since $\epsilon > 0$ was arbitrary, the limit (121) vanishes.

QED.

B Monotone Measure-Preserving Mappings

Let $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$ and $\rho' \in \mathcal{P}(\mathbb{R}^d)$ be probability measures. In this appendix we recall and refine a result of Brenier which provides a unique measure preserving map between $(\mathbb{R}^d, \rho)$ and $(\mathbb{R}^d, \rho')$ realized as the gradient of a convex function. Such a map might technically be termed *cyclically monotone*, but this will be casually and imprecisely abbreviated to *monotone* here. Brenier proved this result — Proposition 3.1 and Theorem 3.1 of [7] — for a weak-$*$ dense set of ρ and ρ' in $\mathcal{P}(\mathbb{R}^d) \subset C_\infty(\mathbb{R}^d)^*$. The proof was based on a lovely duality argument requiring moment conditions as follows; a variant of the proof is sketched in [21].

Theorem B.1 (Brenier's Monotone Mapping [7]) *Let $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$ be supported on $\overline{\Omega}$ for some bounded, smooth and connected open set Ω. Assume $\rho(x)$ to be bounded away from 0 and ∞ on $\overline{\Omega}$, and the boundary of Ω to be measure zero for ρ. If $\rho' \in \mathcal{P}(\mathbb{R}^d)$ with $\int |y|\, d\rho'(y) < \infty$, then there exists a convex function ψ on $\mathbb{R}^d$ whose gradient is a measure preserving map between $(\mathbb{R}^d, \rho)$ and $(\mathbb{R}^d, \rho')$. On Ω, ψ is unique up to an additive constant.*

In the language of Chapter 3, $\nabla\psi$ pushes forward ρ to ρ': $\nabla\psi_{\#}\rho = \rho'$. The gradient $\nabla\psi$ is defined Lebesgue almost everywhere on $\{\psi < \infty\}$. It is the pointwise limit of a sequence of continuous approximants (finite differences), hence Borel measurable, and enjoys an irrotationality property which has been emphasized by Caffarelli [22].

To extend Theorem B.1 to arbitrary $\rho, \rho' \in \mathcal{P}(\mathbb{R}^d)$, it is necessary to reformulate the result slightly and extract a limit. This limit will be unique as long as one of the measures is absolutely continuous with respect to Lebesgue, but since the duality argument will not be of use without moment conditions we rely on earlier, geometrical ideas of Aleksandrov [46] to prove it.

Notions from convex analysis established in Appendix A will be freely employed. Cyclically monotone subsets S of $\mathbb{R}^d \times \mathbb{R}^d$, and the subgradient $\partial\psi$ of a convex function ψ play a central role; the cyclical monotonicity of S is equivalent to the existence of a convex function whose subgradient contains S. The *support* $\operatorname{spt}\rho$ of a measure ρ refers to the smallest closed set which is of full measure in ρ. Using this language, our extension — Theorem B.6 — of Brenier's result begins with:

Corollary B.2 (Monotone Correlation) *Let $\rho, \rho' \in \mathcal{P}(\mathbb{R}^d)$. Then there exists a joint probability measure p on $\mathbb{R}^d \times \mathbb{R}^d$ with cyclically monotone support having marginals ρ and ρ': for $M \subset \mathbb{R}^d$ Borel, $\rho[M] = p[M \times \mathbb{R}^d]$ and $p[\mathbb{R}^d \times M] = \rho'[M]$.*

The proof proceeds through a lemma regarding weak-$*$ limits of measures with cyclically monotone support. As in Chapter 5, $C_\infty(\mathbb{R}^d)$ denotes the Banach space of continuous functions vanishing at ∞ under the sup norm.

Lemma B.3 *Let $\rho_n \to \rho$ and $\rho'_n \to \rho'$ weak-$*$ in $\mathcal{P}(\mathbb{R}^d) \subset C_\infty(\mathbb{R}^d)^*$. Assume $p_n \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ with cyclically monotone support has marginals ρ_n and ρ'_n. A weak-$*$ limit $p \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$, with cyclically monotone support and ρ and ρ' as its marginals, may be extracted from a subsequence of the p_n.*

Proof: $\mathcal{P}(\mathbb{R}^d)$ lies in the unit ball of $C_\infty(\mathbb{R}^d)^*$. Letting $\overline{\mathbb{R}}^d$ denote the one-point compactification of $\mathbb{R}^d$, it is equivalent to view $\mathcal{P}(\mathbb{R}^d)$ as lying in the Banach space dual of $C(\overline{\mathbb{R}}^d)$ (the continuous functions under the sup norm): the weak-$*$ topologies coincide on $\mathcal{P}(\mathbb{R}^d)$ because $C(\overline{\mathbb{R}}^d) = C_\infty(\mathbb{R}^d) \oplus \mathbb{C}$. Similarly, the p_n lie in $C_\infty(\overline{\mathbb{R}}^d \times \overline{\mathbb{R}}^d)$. By the Banach-Alaoglu Theorem, the p_n admit a weak-$*$ convergent subsequence with limit $p \in \mathcal{P}(\overline{\mathbb{R}}^d \times \overline{\mathbb{R}}^d)$. Since any continuous $f \in C(\overline{\mathbb{R}}^d)$ extends to a function $f \in C(\overline{\mathbb{R}}^d \times \overline{\mathbb{R}}^d)$ which is independent of its second argument, the first marginal of p must coincide with ρ:

$$\int f\, dp = \lim_n \int f\, dp_n = \lim_n \int f\, d\rho_n = \int f\, d\rho.$$

By symmetry, the second marginal of p agrees with ρ'. Moreover, it is clear that none of the mass of p can live "at ∞", so $p \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$. It remains to check that the support of p is cyclically monotone. If not, it would contain m points (x_i, y_i) violating (118). Choosing a sufficiently small neighbourhood in $\mathbb{R}^d \times \mathbb{R}^d$ for each point, the inequality would also be violated when any or all of the (x_i, y_i) were replaced by points from these neighbourhoods. Each neighbourhood would have positive mass for p, hence for p_n when n is large. Cyclical monotonicity of $\operatorname{spt} p_n$ produces a contradiction. QED.

Proof of Corollary B.2: When Theorem B.1 applies, p may be obtained by pushing forward ρ through $id \times \nabla\psi$ where id is the identity map on $\mathbb{R}^d$. From $\nabla\psi_{\#}\rho = \rho'$, p is readily seen to have the correct marginals; it is supported on (the closure of) the subgradient of ψ, so the corollary is proved in this case. To extend to arbitrary $\rho \in \mathcal{P}(\mathbb{R}^d)$, approximate by:

(i) convolving with the uniform probability measure on the ball $B_{1/n}(0)$;

(ii) restricting the resulting measure to a ball $B_r(0)$ with radius chosen to yield total mass $1 - 1/n$; and

(iii) adding $1/n$ of the uniform probability measure on $B_r(0)$.

If ρ_n denotes the approximating measure, then $\int f\, d\rho_n \longrightarrow \int f\, d\rho$ whenever $f \in C_\infty(\mathbb{R}^d)$. Moreover, ρ_n satisfies the hypotheses of Theorem B.1: $\rho_n \in \mathcal{P}_{ac}(\mathbb{R}^d)$ and bounded by (i), it is supported on $\overline{\Omega} = B_r(0)$ by (ii) and bounded away from zero there by (iii). If ρ' is similarly approximated by ρ'_n, then there is a $p_n \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ with cyclically monotone support having ρ_n and ρ'_n as its marginals. Now Lemma B.3 implies the Corollary. QED.

A second corollary recovers the connection with convex functions and measure preserving mappings.

Corollary B.4 *Define ρ, ρ', p as in Corollary B.2. If $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$ then there exists a convex function ψ on $\mathbb{R}^d$ such that* $\mathrm{id} \times \nabla\psi$ *pushes ρ forward to p. If $\rho' \in \mathcal{P}_{ac}(\mathbb{R}^d)$ also, then $\nabla\psi^* \times \mathrm{id}$ pushes forward ρ' to p. Here* id *denotes the identity map on $\mathbb{R}^d$, and ψ^* the Legendre transform of ψ.*

Proof: Corollary B.2 guarantees a joint measure p with cyclically monotone support having ρ and ρ' as its marginals. By cyclical monotonicity there is a convex function ψ whose subgradient contains the support of p. It follows that $\psi < \infty$ on the support of ρ, whence $\nabla\psi$ exists ρ-a.e. Let X be the set on which $\nabla\psi$ exists; $p[X \times \mathbb{R}^d] = \rho[X] = 1$. For $x \in X$ there is at most one y such that $(x, y) \in \mathrm{spt}\, p \subset \partial\psi$. Thus $p\left[\left\{(x, \nabla\psi(x)) \mid x \in X\right\}\right] = 1$. The claim is that $(id, \nabla\psi)_{\#}\rho = p$. It is enough to check that these two measures coincide on sets of the form $M \times N$, where $M, N \subset \mathbb{R}^d$ are Borel sets.

$$\begin{aligned} p[M \times N] &= p\left[\left\{(x,y) \,\middle|\, \begin{array}{c} x \in X \cap M \\ \nabla\psi(x) \in N \end{array}\right\}\right] \\ &= \rho\,[\,M \cap (\nabla u)^{-1} N\,] \\ &= (id, \nabla\psi)_{\#}\rho\,[M \times N]. \end{aligned}$$

The symmetrical statement for ρ' follows from the observation that the involution $*(x, y) = (y, x)$ on $\mathbb{R}^d \times \mathbb{R}^d$ pushes p forward to a measure whose cyclically monotone support is contained in $\partial\psi^*$. QED.

Finally, the monotone mapping of Corollary B.4 (and hence the joint measure p) is shown to be unique. The following lemma and theorem employ variants of ideas of Aleksandrov [46].

Lemma B.5 (Aleksandrov) *Let ϕ and ψ be closed convex functions on $\mathbb{R}^d$, differentiable at x_0 with $\phi(x_0) = \psi(x_0) = 0$ and $\nabla\phi(x_0) \neq \nabla\psi(x_0) = 0$.*

Define $M := \{\phi > \psi\} \subset \operatorname{dom}\psi$ *and* $X := \partial\psi^*(\partial\phi(M))$. *Then* $X \subset M$, *while* x_0 *lies a positive distance from* X.

Proof: To obtain the inclusion, let $x \in X$. There exist $m \in M$ and $y \in \mathbb{R}^d$ with $(m, y) \in \partial\phi$ and $(x, y) \in \partial\psi$. For any $z \in \mathbb{R}^d$

$$\begin{aligned} \phi(z) &\geq \langle y, z - m \rangle + \phi(m) \quad \text{and} \\ \psi(m) &\geq \langle y, m - x \rangle + \psi(x). \end{aligned}$$

Noting that $\phi(m) > \psi(m)$, these inequalities combine to yield

$$\phi(z) > \langle y, z - x \rangle + \psi(x). \tag{124}$$

Taking $z = x$ shows $x \in M$.

Next, suppose a sequence $x_n \in X$ has limit x_0. Again, there exist $m_n \in M$ and $y_n \in \mathbb{R}^d$ with $(m_n, y_n) \in \partial\phi$ and $(x_n, y_n) \in \partial\psi$. Now $\nabla\psi(x_0) = 0$ implies $\psi \geq 0$ and $y_n \to 0$ by the continuity of $\partial\psi$; on the other hand, $\nabla\phi(x_0) \neq 0$ implies $\phi(z) < 0$ for some z near x_0. Making use of (124) once more yields

$$\begin{aligned} 0 > \phi(z) &> \langle y_n, z - x_n \rangle + \psi(x_n) \\ &\geq -|y_n|\,|x - x_n|. \end{aligned}$$

Since $x_n \to x_0$ and $y_n \to 0$, a contradiction is obtained. The conclusion is that x_0 cannot lie in the closure of X. QED.

Theorem B.6 (Monotone Mapping) *Let* $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$ *and* $\rho' \in \mathcal{P}(\mathbb{R}^d)$. *There is a convex function* ψ *whose gradient* $\nabla\psi$ *pushes forward* ρ *to* ρ'. *Moreover,* $\nabla\psi$ *is uniquely determined* ρ*-almost everywhere.*

Proof: Corollary B.4 gives existence of ψ. To establish uniqueness, assume ϕ is another convex function for which $\nabla\phi_{\#}\rho = \nabla\psi_{\#}\rho = \rho'$, but that $\nabla\phi = \nabla\psi$

does not hold ρ-a.e. Choose a point x_0 from the set satisfying

(i) x is a Lebesgue point for ρ with positive density $\rho(x) > 0$, and

(ii) ψ and ϕ are differentiable at x with $\nabla\psi(x) \neq \nabla\phi(x)$.

This set cannot be empty, for it has positive mass under ρ: (i) holds ρ-a.e. while (ii) holds on a set of positive measure by hypothesis.

The properties of ψ and ϕ are insensitive to the addition of arbitrary constants, and it is convenient to take $\psi(x_0) = \phi(x_0) = 0$. A linear function may also be added to ψ and ϕ; this corresponds to a translation of ρ' and allows us to take $\nabla\psi(x_0) = 0$. Finally, the convex functions ψ and ϕ may be assumed to be closed. Define $M := \{x \mid \psi(x) < \phi(x)\}$ and $Y = \partial\phi(M)$. As in [46], the idea is to show that the two push-forwards — given by Lemma 6.1 — cannot agree on Y, because

$$\rho[\partial\psi^*(Y)] < \rho[M] \leq \rho[\partial\phi^*(Y)]. \tag{125}$$

The second inequality follows from the obvious inclusion

$$M \cap \operatorname{int}\operatorname{dom}\phi \subseteq \partial\phi^*(Y);$$

the first inequality follows from Lemma B.5, which shows that $\partial\psi^*(Y) \subset M$ and excludes a neighbourhood of x_0. Strict inequality in (125) will be established by showing that a little bit of the mass of ρ in M must lie near x_0.

This follows from the definition of x_0 and M. Translate ρ, ψ and ϕ so that $x_0 = 0$. Let $0 < \epsilon < 1$ and consider the cone $C := \{x \mid \langle \nabla\phi(x_0), x\rangle \geq \epsilon|x|\}$. Using (ii) to approximate ψ and ϕ near $x_0 = 0$,

$$\phi(x) - \psi(x) = \langle \nabla\phi(x_0), x\rangle + o(|x|).$$

Thus $x \in C$ sufficiently small implies $x \in M$. The average of ρ over $C \cap B_r(0)$ must converge to $\rho(x_0) > 0$ by (i), since $C \cap B_r(0)$ shrinks nicely to $x_0 = 0$ with r in the sense of (31). For small r, this set is contained in M but disjoint from $\partial\psi^*(Y)$. (125) is established and the proof is complete. QED.

Remark B.7 (Non-Uniqueness of Monotone Correlation)
If neither ρ nor ρ' is absolutely continuous, the joint distribution p of Corollary B.2 need not be unique in dimension $d \geq 2$. For example, let $d = 2$ and suppose that ρ is supported in the interval [-1,1] on the x-axis while ρ' is supported in the corresponding interval on the y-axis. Any *p having the correct marginals will be supported on the subgradient of the convex function $\psi(x, y) = |y|$.*

C Alternative Measure-Preserving Maps

This appendix discusses an alternative formulation of the displacement interpolation between two measures $\rho, \rho' \in \mathcal{P}_{ac}(\mathbb{R}^d)$. It is based on an explicit construction of $\mathbb{R}^d$-valued random variables representing ρ and ρ', rather than the existence of a convex function ψ with gradient pushing ρ forward to ρ'. Except in dimension $d = 1$, the interpolant $\rho \xrightarrow{t} \rho'$ here defined will not generally coincide with that of Definition 3.1, but it will satisfy the same convexity estimates. For the applications of Part I and Part II, Definition 3.1 has been preferred only because it appears to be more natural, and provides conditions for strict convexity in Theorem 4.2 and Proposition 9.4. These conditions are required to infer conclusions when the potentials $V(x)$ and $Q(x)$ fail to be *strictly* convex.

Let $\Sigma := \{0, 1\}$ be the probability space with two elements equally weighted, and ν the product measure induced on the Cartesian product $\Sigma^\infty := \bigtimes_1^\infty \Sigma$, the space of binary sequences. Let $\sigma_n \in \Sigma$ denote the n*-th* digit of $\sigma \in \Sigma^\infty$, and define $\pi_n : \Sigma^\infty \longrightarrow \Sigma^n$ to be truncation of σ to n digits: $\pi_n(\sigma) := (\sigma_1 \ldots \sigma_n)$. Then the measure ν assigns mass 2^{-n} to each *cylinder* $\{\sigma \in \Sigma^\infty \mid \pi_n(\sigma) = \tau\}$ with $\tau \in \Sigma^n$. Together with the null set, these cylinders form a semi-algebra of sets; ν is uniquely determined on the σ-algebra they generate. Let $\Pi_n(x) := x_n$ denote the (n mod d)*-th* coordinate of $x \in \mathbb{R}^d$. The construction to follow yields:

Theorem C.1 (Existence and Uniqueness of Dissecting Maps)
Let $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$. There is a measurable map $y : \Sigma^\infty \longrightarrow \mathbb{R}^d$ pushing ν forward to ρ which satisfies: $\Pi_n\big(y(\sigma)\big) \leq \Pi_n\big(y(\tau)\big)$ when $\pi_{n-1}(\sigma) = \pi_{n-1}(\tau)$ but $\sigma_n < \tau_n$. The map y is uniquely determined up to sets of ν-measure zero.

Any map $y : \Sigma^\infty \longrightarrow \mathbb{R}^d$ which satisfies the inequalities in Theorem C.1 will be referred to as a *dissecting map*. Obviously, this definition depends on choice of basis for $\mathbb{R}^d$. The existence part of the theorem can be extended to measures $\rho \in \mathcal{P}(\mathbb{R}^d)$ not absolutely continuous, but uniqueness may fail. Using this theorem, the displacement interpolant (9) is replaced by:

Definition C.2 (Dissection Interpolation)
Given probability measures $\rho, \rho' \in \mathcal{P}_{ac}(\mathbb{R}^d)$, let y, y' be the dissecting maps with $y_{\#}\nu = \rho$ and $y'_{\#}\nu = \rho'$ of Theorem C.1. At time $t \in [0,1]$, the dissection interpolant $\rho_t = \rho \xrightarrow{t} \rho' \in \mathcal{P}(\mathbb{R}^d)$ between ρ and ρ' is defined to be

$$\rho_t := [(1-t)y + ty']_{\#}\nu. \tag{126}$$

With this definition, conclusions between Proposition 3.2 and Example 3.7 (inclusive) — as well as Lemma 9.1 and Remark 9.2 — apply to dissection interpolation as to displacement interpolation; the linear transformations Λ of Proposition 3.6(i) must be diagonal instead of orthogonal. The proofs are essentially the same and will not be repeated; some exploit convexity of the set of dissecting maps. Only the absolute continuity of ρ_t and the *dissection convexity* of $\int A(\rho)$ will be proved in detail.

Dissecting $\mathbb{R}^d$ with a Probability Measure

A *half-open d-cell* is a non-empty subset $D \subseteq \mathbb{R}^d$ of the form

$$D = \{x \in \mathbb{R}^d \mid a_k \leq \Pi_k(x) < b_k\} \tag{127}$$

for some constants $a_k, b_k \in \mathbb{R} \cup \{\pm\infty\}$ and $k = 1, \ldots, d$. A *closed d-cell* is a non-empty subset of the same form, but with the strict inequalities relaxed. In particular, a closed d-cell may be of dimension less than d. Given a probability measure $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$, the following recursive algorithm generates nested partitions of $\mathbb{R}^d$ into d-cells (or cells). At the n-th stage of the construction, there will be 2^n d-cells $D(\tau)$, each of measure 2^{-n} with respect to ρ. These cells will be indexed by n-digit binary sequences $\tau \in \Sigma^n$. The d-cells will taken to be closed, but for absolutely continuous measures the distinction is moot.

At the first ($n = 1$) step of the construction, $\mathbb{R}^d$ is divided into two d-cells by a hyperplane $\Pi_1(x) = const$ perpendicular to the x_1-axis. The d-cells (half-spaces at this stage) may be labeled $D(0)$ and $D(1)$, where $D(0)$ lies on the negative x_1 side of $D(1)$. The dividing hyperplane is chosen so that $\rho[D(0)] = \rho[D(1)] = 1/2$. This choice need not be unique. Here $\rho[D]$ denotes the measure of D with respect to ρ, or *mass* of D. $\mathrm{vol}\,[D]$ will be used to denote the Lebesgue measure, or *volume*, of D.

At the *n-th* step of the construction, each $D(\tau)$ with $\tau \in \Sigma^{n-1}$ is subdivided by a *coordinate* hyperplane $\Pi_n(x) = const$ into two d-cells of mass 2^{-n} under ρ. Using juxtaposition to denote concatenation, so that $\tau 0 = (\tau_1 \tau_2 \ldots \tau_{n-1} 0)$, the two subcells of $D(\tau)$ will be indexed by $\tau 0$ and $\tau 1$ to be consistent with (iii) below. Taking the sole element of Σ^0 to be ϕ by convention, the key properties of the construction are summarized by

(i) $D(\phi) = \mathbb{R}^d$,

(ii) $D(\tau) = D(\tau 0) \bigcup D(\tau 1)$,

(iii) $\Pi_n(x) \leq \Pi_n(y)$ whenever $x \in D(\tau 0)$ and $y \in D(\tau 1)$ for $\tau \in \Sigma^{n-1}$,

together with

$$\rho[D(\tau)] = 2^{-n} \quad \text{for} \quad \tau \in \Sigma^n. \tag{128}$$

(128) is the only property involving the original measure $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$; it has been separated to motivate the following abstract definition.

Definition C.3 *A* **dissection** *of* $\mathbb{R}^d$ *is a map* $D : \bigcup_{n \geq 0} \Sigma^n \longrightarrow$ *closed d-cells* $\subseteq \mathbb{R}^d$ *which satisfies properties (i)–(iii) above. The image of D restricted to* Σ^n *is called the n-th level of the dissection.*

In particular, it is clear that $\bigcup_{\tau \in \Sigma^n} D(\tau) = \mathbb{R}^d$ and that this union is disjoint apart from sets of Lebesgue measure zero. To recapitulate the foregoing:

Proposition C.4 (Dissection of a Probability Measure on $\mathbb{R}^d$) *Associated to $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$ is a dissection $D(\tau)$ of $\mathbb{R}^d$ satisfying (128).*

Recovering a Measure from a Dissection of $\mathbb{R}^d$

The next step is to show that any dissection $D(\tau)$ of $\mathbb{R}^d$ induces a dissecting map $y : \Sigma^\infty \longrightarrow \mathbb{R}^d$, and corresponding measure $y_\# \nu \in \mathcal{P}(\mathbb{R}^d)$. If the dissection was constructed from a measure $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$, then $y_\# \nu = \rho$.

Naturally, one expects the map $y(\sigma)$ to satisfy

$$y(\sigma) \in \bigcap_{n=1}^{\infty} D(\pi_n(\sigma)). \tag{129}$$

Proposition C.7 shows that (129) determines $y(\sigma)$ uniquely almost everywhere in Σ^∞. The key to its proof, as well as the demonstration that $y_\# \nu = \rho$ when $D(\tau)$ is constructed from ρ, is the following lemma. Fix a coordinate hyperplane $\Pi_k(x) = c$. If mass 2^{-n} is assigned to each cell from the *n-th* level of D, the lemma states that the mass straddling $\Pi_k(x) = c$ goes to zero as $n \to \infty$. In fact, it decays like $2 \cdot 2^{-n/d}$.

Lemma C.5 *Let $\Pi_k(x) = c$ be a coordinate hyperplane, and $D(\tau)$ a dissection of $\mathbb{R}^d$. For each integer $n > 0$ define*

$$\Gamma_n := \Big\{ \tau \in \Sigma^n \;\Big|\; \{c\} \subset \Pi_k\big(D(\tau)\big) \;\text{ strictly} \Big\}.$$

Then $\#(\Gamma_n)/2^n \to 0$ as $n \to \infty$, where $\#(\Gamma_n)$ is the cardinality of Γ_n.

Proof: It will be enough to count d-cells $D(\tau)$ which intersect both $H := \{x \mid \Pi_k = c\}$ and $H^+ := \{x \mid \Pi_k > c\}$; at the *n-th* level of D, the number a_n of such cells is dissection independent, and $\#(\Gamma_n) \leq 2a_n$. At level 0, there is exactly one d-cell $D(\phi) = \mathbb{R}^d$, so that $a_0 = 1$. Properties C.3(ii)–(iii) of D

make it clear that $a_n = 2a_{n-1}$ when $n \bmod d \neq k$. In the case $n \bmod d = k$, the hyperplane separating $D(\tau 0)$ from $D(\tau 1)$ is parallel to H. If it lies in H^+, only one of the subcells of $D(\tau)$ will intersect H; otherwise, only one of the subcells of $D(\tau)$ will intersect H^+. Either way, $a_n = a_{n-1}$ for $n \bmod d = k$. Note that the subcells are closed by definition. For arbitrary n, the recursion yields $a_{n\cdot d} = 2^{n(d-1)}$, so $a_n/2^n \to 0$ as claimed. QED.

Corollary C.6 *If $D(\tau)$ is a dissection of $\mathbb{R}^d$, then as $n \to \infty$*

$$\#\Big(\{\tau \in \Sigma^n \mid D(\tau) \text{ unbounded}\}\Big)/2^n \longrightarrow 0.$$

Proof: The argument is the same as that of the lemma. Letting a_n be the number of d-cells at the *n-th* level of D which extend to $+\infty$ in the x_k direction: $a_0 = 1$ and

$$a_{n+1} = \begin{cases} a_n, & \text{if } n \bmod d = k-1; \\ 2a_n, & \text{otherwise.} \end{cases}$$

Thus the total number of cells which extend to infinity at the (nd)*-th* level of D must be less than $2d \cdot 2^{n(d-1)}$. QED.

Proposition C.7 *If $D(\tau)$ is a dissection of $\mathbb{R}^d$, then $\cap_n D(\pi_n(\sigma))$ consists of a single point for ν almost every $\sigma \in \Sigma^\infty$.*

Proof: Let $\operatorname{diam} R := \sup_{y,z\in R} |y - z|$ denote the diameter of $R \subseteq \mathbb{R}^d$. For $\sigma \in \Sigma^\infty$, when $\operatorname{diam} D(\pi_n(\sigma)) \to 0$ as $n \to \infty$ the intersection in (129) is uniquely determined; it is also non-empty since the closed nested d-cells are eventually compact. Let $Z \subseteq \Sigma^\infty$ be the set for which $\operatorname{diam} D(\pi_n(\sigma)) \not\to 0$, so that

$$Z = \bigcup_k \bigcap_n \{\sigma \in \Sigma^\infty \mid \operatorname{diam} D(\pi_n(\sigma)) > 1/k\}.$$

For fixed k and n, the set on the right is a union of cylinders in Σ^∞, showing Z to be measurable. To see that $\nu[Z] = 0$, fix an integer k and $\epsilon > 0$. Corollary C.6 shows that $\Gamma := \{\tau \in \Sigma^N | \text{diam}\, D(\tau) = \infty\}$ satisfies $\#(\Gamma)/2^N < \epsilon/2$ for some $N < \infty$. Choose a bounded d-cell R which contains all bounded $D(\tau)$ from the N-*th* level of D. For $n > N$ define

$$\Gamma_n := \{\tau \in \Sigma^n \mid \text{diam}\, D(\tau) > 1/k\}.$$

A finite number of coordinate hyperplanes, spaced sufficiently close, divide R into cubes of diameter less than $1/k$. For $\tau \in \Gamma_n$, either $\pi_N(\tau) \in \Gamma$, which holds for at most $2^{n-1}\epsilon$ elements $\tau \in \Sigma^n$, or $D(\tau)$ crosses one of these hyperplanes. Applying Lemma C.5, it is clear that for n sufficiently large $\#(\Gamma_n) < 2^n\epsilon$. Then $\nu\Big[\{\sigma \in \Sigma^\infty \mid \pi_n(\sigma) \in \Gamma_n\}\Big] < \epsilon$. Because $\epsilon > 0$ was arbitrary, the intersection over all n has mass zero for ν. A countable union of such sets, $\nu[Z] = 0$. QED.

Corollary C.8 *A dissection $D(\tau)$ of $\mathbb{R}^d$ determines a dissecting map $y : \Sigma^\infty \longrightarrow \mathbb{R}^d$ satisfying (129) uniquely up to sets of ν-measure zero. Any dissecting map $y(\sigma)$ must be measurable.*

Proof: A map $y(\sigma)$ which satisfies (129) is dissecting because of property (iii) of Definition C.3. Thus the first assertion is immediate from the proposition. It remains to show that any dissecting map must be measurable. Therefore, assume a dissecting map $y(\sigma)$ to be given. Construct a dissection $D(\tau)$ satisfying (129) recursively from $y(\sigma)$, by using the hyperplane

$$\Pi_n(x) = \sup_{\pi_n(\sigma)=\tau 0} \Pi_n(y(\sigma)). \tag{130}$$

to subdivide the d-cell $D(\tau)$ into $D(\tau 0)$ and $D(\tau 1)$ when $\tau \in \Sigma^{n-1}$. The recursion is initiated from $D(\phi) = \mathbb{R}^d$. The supremum (130) is finite from the definition of a dissecting map; it lies in $\Pi_n(D(\tau))$ from the preceding level of the recursion. Using $D(\tau)$ and the proposition, y will be shown to be the pointwise limit of a sequence y_n of measurable transformations almost everywhere, hence measurable. For each $\tau \in \Sigma^n$, choose $x(\tau) \in D(\tau)$. If $y_n(\sigma) := x(\pi_n(\sigma))$, then y_n is obviously measurable: it takes at most 2^n values and its level sets are cylinders. Proposition C.7 shows that $y_n(\sigma) \to y(\sigma)$ almost everywhere as $n \to \infty$. QED.

Proposition C.9 *Let $D(\tau)$ be a dissection of $\mathbb{R}^d$, and $y(\sigma)$ the map which satisfies (129). For a d-cell $R \subset \mathbb{R}^d$ define $\gamma_n := \{\tau \in \Sigma^n \mid D(\tau) \subset R\}$. Then*

$$y_\#\nu[R] = \lim_{n\to\infty} \#(\gamma_n)/2^n. \tag{131}$$

Proof: The *n-th* level of the dissection D may be used to approximate R from the inside and the outside via γ_n and $\Gamma_n := \{\tau \in \Sigma^n \mid D(\tau) \cap R \neq \emptyset\}$: from (129) it follows that

$$\{\sigma \in \Sigma^\infty \mid \pi_n(\sigma) \in \gamma_n\} \;\subset\; y^{-1}(R) \;\subset\; \{\sigma \in \Sigma^\infty \mid \pi_n(\sigma) \in \Gamma_n\},$$

whence $\#(\gamma_n)/2^n \leq y_\#\nu[R] \leq \#(\Gamma_n)/2^n$. Clearly $\gamma_n \subset \Gamma_n$. For $\tau \in \Gamma_n \sim \gamma_n$ the cell $D(\tau)$ straddles one of the hyperplanes bounding R. Lemma C.5 implies $(\#(\Gamma_n) - \#(\gamma_n))/2^n \to 0$. Both must tend to the limit $y_\#\nu[R]$.

QED.

Proof of Theorem C.1: Proposition C.4 associates to $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$ a dissection $D(\tau)$ of $\mathbb{R}^d$. Let $y(\sigma)$ be the dissecting map of Corollary C.8. It induces a measure $y_\#\nu \in \mathcal{P}(\mathbb{R}^d)$. Since the σ-algebra of Borel sets in $\mathbb{R}^d$ is generated by the semi-algebra of half-open d-cells R, it suffices to verify $y_\#\nu[R] = \rho[R]$. From Proposition C.9 and (128) it follows that $y_\#\nu[R] \leq \rho[R]$. Since both are probability measures, equality holds, and $y_{\#}\nu = \rho$. This establishes the existence part of the theorem.

To establish uniqueness, let y be a second dissecting map with $y_{\#}\nu = \rho$. By Corollary C.8, it is enough to show that y satisfies (129) almost everywhere in Σ^∞. This follows if $y(\sigma) \in D(\tau)$ holds almost everywhere in each cylinder $\{\sigma \in \Sigma^\infty \mid \pi_n(\sigma) = \tau\}$. The latter is proved by induction; the case $\tau = \phi$ is obvious from property (i) of Definition C.3. Assuming the condition to be satisfied for τ, it is clear from property (ii) that $\pi_{n-1}(\sigma) = \tau$ implies either $y(\sigma) \in D(\tau 0)$ or $y(\sigma) \in D(\tau 1)$. Suppose that $y(\sigma) \in D(\tau 1)$ on a set of positive measure with $\pi_n(\sigma) = \tau 0$. Since y is a dissecting map, it follows from property (iii) of the same definition that $y(\sigma) \in D(\tau 1)$ whenever $\pi_n(\sigma) = \tau 1$. Thus $\rho[D(\tau 1)] = y_{\#}\nu[D(\tau 1)] > 2^n$, contradicting (128). QED.

Dissection Convexity of $\int A(\rho)$

Finally, it remains to prove the analog of Theorem 4.2 for the dissection interpolation. Such an estimate becomes more plausible in view of the fact that dissection resembles the idea underlying the Hadwiger-Ohmann proof of the Brunn-Minkowski Theorem [16]. Unlike the displacement convexity of $U(\rho)$, dissection convexity will be proved by first establishing the estimate when the interpolant $\rho \xrightarrow{t} \rho'$ and its endpoints are suitably chosen simple functions.

Definition C.10 (Dissection Approximants)
Let $D(\tau)$ be a dissection of $\mathbb{R}^d$ and $y_{\#}\nu \in \mathcal{P}(\mathbb{R}^d)$ the measure it induces through Corollary C.8. For a bounded d-cell $R \subset \mathbb{R}^d$, let μ_R denote the uniform probability measure on R: $\mu_R(x) := \chi_R(x)/\mathrm{vol}\,[R]$ when $\mathrm{vol}\,[R] > 0$; define $\mu_R = 0$ if R is unbounded. The measures $\rho_n := 2^{-n} \sum_{\tau \in \Sigma^n} \mu_{D(\tau)}$ will be referred to as dissection approximants to $y_{\#}\nu$.

Because of unbounded d-cells, the measures ρ_n will not be normalized, but $\lim_n \rho_n[\mathbb{R}^d] = 1$ by Corollary C.6. The resulting approximation lemma is:

Lemma C.11 (Weak-$*$ Convergence of Dissection Approximants) *Let $D(\tau)$ be a dissection of $\mathbb{R}^d$, and ρ_n be dissection approximants to the induced measure $y_\# \nu$. For continuous functions φ vanishing at infinity on $\mathbb{R}^d$,*

$$\lim_n \int \varphi \, d\rho_n = \int \varphi \, dy_\# \nu. \tag{132}$$

Proof: Since φ vanishes at ∞ it can be approximated in $\| \cdot \|_\infty$ by a simple function $\sum_k a_k \chi_{R_k}$ in which the R_k are d-cells. It is therefore enough to prove (132) when φ is the characteristic function of a d-cell R. For a bounded d-cell, $\int_R d\rho_n \geq \#(\gamma_n)/2^n$ where γ_n is from Proposition C.9. The same proposition implies

$$\underline{\lim}_n \int_R d\rho_n \geq \int_R dy_\# \nu. \tag{133}$$

This inequality continues to hold for unbounded d-cells R by Corollary C.6. Since the complement of a d-cell in $\mathbb{R}^d$ is a finite union of d-cells, strict inequality in (133) for some R and any subsequence of the ρ_n would violate $\rho_n(\mathbb{R}^d) \leq y_\# \nu(\mathbb{R}^d) = 1$. Thus the lim inf in (133) may be replaced by a limit, and the inequality by equality, concluding the proof. QED.

Lemma C.12 (Dissection Convexity of $U(\rho)$ Between d-Cells) *Let $R, R' \subset \mathbb{R}^d$ be d-cells with finite (non-zero) measure. For $t \in [0,1]$, define the dissection interpolant $\rho_t := \mu_R \xrightarrow{t} \mu_{R'}$ between the uniform probability measures on R and R'. Then $\rho_t = \mu_{(1-t)R+tR'}$. If $U(\rho)$ is defined by (15) and $A(\varrho)$ satisfies (A1), then $U(\rho_t)$ will be convex as a function of t on $[0,1]$.*

Proof: Let $T(x) := \Lambda x + k$ be the affine transformation of $\mathbb{R}^d$ taking R to R' for which Λ is a positive matrix and $k \in \mathbb{R}^d$. Since R and R' are both d-cells,

Λ is diagonal. If $y(\sigma)$ is the unique dissecting map of Theorem C.1 with $y_{\#}\nu = \mu_R$, then the map $y'(\sigma) := T(y(\sigma))$ is dissecting; it satisfies $y'_{\#}\nu = \mu_{R'}$ since $T_{\#}\mu_R = \mu_{R'}$. Thus the dissection interpolant $\rho_t := [(1-t)y + ty']_{\#}\nu$ is the push-forward of μ_R through the affine map $(1-t)x + t\,T(x)$, which is also the gradient of a convex function. The dissection and displacement interpolants (126) and (9) coincide in this case: both are given by the uniform probability measure on $(1-t)R + tR'$. Convexity of $U(\rho_t)$ is a special case of Theorem 4.2, or may be verified directly. QED.

Proposition C.13 (Absolute Continuity of $\rho \xrightarrow{t} \rho'$)
For $t \in [0,1]$, define the dissection interpolant $\rho_t := \rho \xrightarrow{t} \rho'$ (126) between $\rho, \rho' \in \mathcal{P}_{ac}(\mathbb{R}^d)$. Then $\rho_t \in \mathcal{P}_{ac}(\mathbb{R}^d)$. Moreover, if D and D' are dissections of ρ and ρ' as in Proposition C.4, then Corollary C.8 associates a map to the dissection $(1-t)D + t\,D'$. This map pushes ν forward to ρ_t.

Proof: For $\tau \in \Sigma^n$, the d-cell $(1-t)D(\tau) + t\,D'(\tau)$ is defined through Minkowski addition; it is then clear that $(1-t)D + tD'$ satisfies the properties (i)–(iii) of a dissection. If the dissecting maps $y, y' : \Sigma^\infty \longrightarrow \mathbb{R}^d$ satisfy (129) for D and D', then the map $(1-t)y + ty'$ satisfies (129) for the dissection $(1-t)D + t\,D'$. Since $y_{\#}\nu = \rho$ and $y_{\#}\nu = \rho'$ from the proof of Theorem C.10, $[(1-t)y + ty']_{\#}\nu = \rho_t$ by definition. This establishes the second part of the proposition.

The first part depends on Proposition C.9 and the observation that

$$\mathrm{vol}\,[(1-t)R + tR'] \geq (1-t)^d \mathrm{vol}\,[R] \tag{134}$$

for two d-cells $R, R' \subset \mathbb{R}^d$. Suppose ρ_t not absolutely continuous with respect to Lebesgue. For some $m > 0$, a subset $Z \subset \mathbb{R}^d$ has $\mathrm{vol}\,[Z] = 0$ and $\rho_t[Z] > m$. Since $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$, there is a $\delta > 0$ such that $\rho[M] < m$ for any measurable set M with $\mathrm{vol}\,[M] < \delta$. By the regularity of Lebesgue measure, there is an open set of volume less than $(1-t)^d\delta$ containing Z. Any open set is countable disjoint union of half-open d-cells, so there must be a finite disjoint union U of such cells for which $\rho_t[U] > m$ but $\rho_t[U] < (1-t)^d\delta$. Defining $\gamma_n := \{\tau \in \Sigma^n \mid (1-t)D(\tau) + t\,D'(\tau) \subset U\}$, it is possible to choose n large enough so that $\#(\gamma_n)/2^n > m$ by Proposition C.9. Taking $M := \cup_{\tau\in\gamma_n} D(\tau)$ leads to a contradiction: (134) implies $\mathrm{vol}[M] < \delta$ while (128) implies $\rho[M] > m$. QED.

Theorem C.14 (Dissection Convexity of Internal Energy $U(\rho)$) *Let $\rho, \rho' \in \mathcal{P}_{ac}(\mathbb{R}^d)$. Define $U(\rho)$ through (15) and assume $A(\varrho)$ satisfies (A1). Then $U(\rho \xrightarrow{t} \rho')$ will be a convex function of $t \in [0,1]$ for the dissection interpolation (126).*

Proof: Using Proposition C.4 and Definition C.10, define the dissection $D(\tau)$ of $\mathbb{R}^d$ and dissection approximants ρ_n associated to ρ. Let y_n be the dissecting maps of Theorem C.1 which push-forward ν to ρ_n. Let $D'(\tau)$, ρ'_n and y'_n be similarly associated to ρ'. For $\tau \in \Sigma^n$, the restriction of y_n to the cylinder $C := \{\sigma \in \Sigma^\infty \mid \pi_n(\sigma) = \tau\}$ is itself a dissecting map. Up to its *n-th* level, D coincides with the dissection of ρ_n. Therefore (129) shows that apart from a set of ν-measure zero, $y_n(\sigma) \in D(\tau)$ if and only if $\sigma \in C$. A similar statement holds for y'_n. Noting the definition of ρ_n, $(1-t)y_n + ty'_n$ must push forward the restricted measure $\nu|_C$ to $2^{-n}\mu_{(1-t)D(\tau)+tD'(\tau)}$ as in Lemma C.12. It follows that

$$\rho_n \xrightarrow{t} \rho'_n = 2^{-n} \sum_{\tau\in\Sigma^n} \mu_{(1-t)D(\tau)+tD'(\tau)}, \tag{135}$$

As in Proposition C.13, $(1-t)D + t\,D'$ is a dissection so the sets in (135) are disjoint (up to sets of Lebesgue measure zero). Lemma C.12 then implies

$$U(\rho_n \xrightarrow{t} \rho'_n) \leq (1-t)U(\rho_n) + tU(\rho'_n). \tag{136}$$

Noting Proposition C.13 and Definition C.10, the dissection approximants $(\rho \xrightarrow{t} \rho')_n$ to $\rho \xrightarrow{t} \rho'$ coincide with (135). Since $\rho \in \mathcal{P}_{ac}(\mathbb{R}^d)$ was quite general, the conclusion of the theorem follows from (136) if it can be shown that

$$U(\rho) = \lim_n U(\rho_n). \tag{137}$$

Jensen's inequality, together with (128) and convexity of $A(\varrho)$ yield

$$A\Big(2^{-n}\mathrm{vol}\,[D(\tau)]^{-1}\Big)\ \mathrm{vol}\,[D(\tau)] \leq \int_{D(\tau)} A(\rho). \tag{138}$$

Summing (138) over $\tau \in \Sigma^n$ implies $U(\rho_n) \leq U(\rho)$. On the other hand, Lemma 5.4 combines with Lemma C.11 to yield $U(\rho) \leq \underline{\lim}_n U(\rho_n)$. These two inequalities imply (137), completing the theorem. QED.

D Prékopa-Leindler and Brascamp-Lieb Inequalities from Monge-Ampère Equation

In this appendix the displacement interpolation $\rho \xrightarrow{t} \rho'$ is exploited to provide a new proof of a theorem due to Brascamp and Lieb [11], which generalized earlier results of Prékopa and Leindler [8, 9, 10]. The theorem applies to an interpolant $h_\alpha(x)$ defined between non-negative measurable functions $f(x)$ and $g(x)$ on $\mathbb{R}^d$ for $t \in (0,1)$:

$$h_\alpha(x; f, g, t) := \sup_{y \in \mathbb{R}^d} \left\{ (1-t) f\left(\frac{y}{1-t}\right)^\alpha \oplus t\, g\left(\frac{x-y}{t}\right)^\alpha \right\}^{1/\alpha}. \tag{139}$$

Here $\alpha \in [-1/d, \infty)$, while $\oplus$ is distinguished from ordinary addition $+$ only in that $\{(1-t)f^\alpha \oplus tg^\alpha\}^{1/\alpha} = 0$ if either $f = 0$ or $g = 0$. The case $\alpha = 0$ is defined in the limit, so that $h_0(x)$ coincides with the interpolant (25) of Prékopa and Leindler. The notation $h_\alpha(x; f, g, t)$ will be replaced by $h_\alpha(x)$ when the dependence on f, g and t is quite clear.

The central result concerning h_α pertains to its mass $\|h_\alpha\|_1 := \int h_\alpha$. Stated as Corollary D.4, it reduces to the assertion that $\|h_\alpha\|_1 \geq 1$ in the case $\|f\|_1 = \|g\|_1 = 1$. In this case our displacement interpolant $\rho_t := f \xrightarrow{t} g$ may also be defined. The next proposition shows $h_\alpha \geq \rho_t$ almost everywhere, so that a fortiori $\|h_\alpha\|_1 \geq \|\rho_t\|_1 = 1$; the discrepancy between h_α and ρ_t accounts for the full error in the Prékopa-Leindler and Brascamp-Lieb inequalities. The proposition follows formally from the Monge-Ampère equation (141) satisfied almost everywhere by ρ_t, together with an elementary lemma. The results for $\|f\|_1 \neq \|g\|_1$ are recovered by scaling x, f, g and t.

Lemma D.1 *For constants $f, g > 0$, $t \in (0,1)$ and a positive $d \times d$ matrix Λ with determinant* $\det[\Lambda] = f/g$*,*

$$f \det[(1-t)I + t\Lambda]^{-1} \leq \left((1-t)f^{-1/d} + tg^{-1/d}\right)^{-d}. \tag{140}$$

Proof: In the basis which diagonalizes Λ, concavity of $\det\,[(1-t)I+t\Lambda]^{1/d}$ is seen using the domination of the geometric by the arithmetic mean; (140) is equivalent to

$$\det\,[(1-t)I+t\Lambda]^{1/d} \geq (1-t)\det\,[I]^{1/d} + t\,(f/g)^{1/d}.$$

QED.

Proposition D.2 ($h_\alpha(x)$ Dominates the Displacement Interpolant) *Let $f, g \in \mathcal{P}_{ac}(\mathbb{R}^d)$ and $\alpha \geq -1/d$. For $t \in (0,1)$ define $h_\alpha(x)$ as in (139). Then the displacement interpolant $\rho_t := f \xrightarrow{t} g$ of (9) satisfies $\rho_t(x) \leq h_\alpha(x)$ for almost all x.*

Proof: Since (139) is non-decreasing in α, it is enough to establish the proposition when $\alpha = -1/d$. Let ψ be the convex function satisfying $\nabla\psi_{\#}f = g$ through which ρ_t is defined. As noted in Theorem 6.4, there is a set $X \subset \mathbb{R}^d$ with $f[X] = 1$ on which the positive matrix $\nabla^2\psi(x)$ and its inverse exist. Both $f, g \in \mathcal{P}_{ac}(\mathbb{R}^d)$ have Lebesgue points almost everywhere, and since $f[(\nabla\psi)^{-1}M] = g[M]$, taking X smaller if necessary ensures that any $x \in X$ is a Lebesgue point of f at which $f(x) > 0$, while $\nabla\psi(x)$ is a Lebesgue point for g; the image of X under $\nabla\psi$ remains of full measure for g. Fix $x_0 \in X$ and define $x_t = (1-t)x_0 + t\nabla\psi(x_0)$. Proposition 6.2 verifies that $g(x_1) = f(x_0)\det\,[\nabla^2\psi(x_0)]^{-1}$. Similarly, the values of ρ_t are specified by

$$\rho_t\big(x_t\big) = f(x_0)\,\det\,[(1-t)I + t\nabla^2\psi(x_0)]^{-1} \tag{141}$$

ρ_t-almost everywhere. Identifying $x = x_t$ and $y = (1-t)x_0$ in (139), the conclusion $\rho_t(x_t) \leq h_\alpha(x_t)$ follows immediately from Lemma D.1; the supremum is gratuitous. Thus $\rho_t \leq h_\alpha$ holds almost everywhere ρ_t, and therefore Lebesgue almost everywhere. QED.

Remark D.3 As was pointed out by Brascamp and Lieb, $h_\alpha(x)$ may depend on the values of f and g everywhere on $\mathbb{R}^d$, not just almost everywhere. This annoyance is remedied [11] by replacing the supremum in (139) by an essential supremum, yielding a modified interpolant $k_\alpha(x) \le h_\alpha(x)$. Moreover, for a particular choice of $f^\#$ and $g^\#$, which differ from f and g only on a set of measure zero, the interpolants $h_\alpha(x; f^\#, g^\#, t) = k_\alpha(x; f^\#, g^\#, t) = k_\alpha(x; f, g, t)$ coincide. Thus Proposition D.2 applies equally well to $k_\alpha(x)$ as to $h_\alpha(x)$. It is not difficult to argue this directly, using the fact that x_0 and x_1 are Lebesgue points for f and g (respectively) in the preceding proof.

Corollary D.4 (Brascamp-Lieb [11])
Let f, g be non-negative measurable functions on $\mathbb{R}^d$. For $t \in (0,1)$, define h_α as in (139). Let $\|f\|_1 > 0$, $\|g\|_1 > 0$. If $\alpha > -1/d$ then $\|h_\alpha\|_1 \ge C$, where

$$C := \Big((1-t)\, \|f\|_1^\gamma + t\, \|g\|_1^\gamma \Big)^{1/\gamma}$$

and $\gamma := \alpha/(1+d\alpha)$. In particular, $\|h_0\|_1 \ge \|f\|_1^{1-t} \|g\|_1^t$.

Proof: When $\|f\|_1$ and $\|g\|_1$ are both finite, define $\rho(x) := f(x)/\|f\|_1$ and $\rho'(x) = g(x)/\|g\|_1$. Denote the mass preserving dilation S_λ by factor $\lambda > 0$ as in Proposition 3.6. The scaling

$$h_\alpha(\, x;\, f,\, g,\, t\,) = C\, h_\alpha(\, x;\, S_\lambda \rho\,,\, S_{\lambda'} \rho'\,,\, t/\lambda'\,) \tag{142}$$

follows directly from (139) provided $\lambda := (C/\|f\|_1)^\gamma$ and $\lambda' := (C/\|g\|_1)^\gamma$. Since $\|S_\lambda \rho\|_1 = \|S_{\lambda'}\rho'\|_1 = 1$, using Proposition D.2 to integrate (142) over x yields the desired inequality. The case $\|g\|_1 = \infty$ follows from the Monotone Convergence Theorem, after noting that (139) increases with g. QED.

Remark D.5 The result of [11] extends to the case $\alpha = -1/d$, but since $\gamma \to -\infty$ the extension cannot be obtained from the scaling relation (142).

References

[1] G. Wulff. Zur Frage der Geschwindigkeit des Wachsthums und der Auflösung der Krystallflachen. *Z. Krist.*, 34:449–530, 1901.

[2] J.E. Avron, J.E. Taylor and R.K.P. Zia. Equilibrium shapes of crystals in a gravitational field: Crystals on a table. *J. Statist. Phys.*, 33:493–522, 1983.

[3] K. Okikiolu. Personal communication.

[4] J.F.G. Auchmuty and R. Beals. Variational solutions of some non-linear free boundary problems. *Arch. Rational Mech. Anal.*, 43:255–271, 1971.

[5] Y.Y. Li. On uniformly rotating stars. *Arch. Rational Mech. Anal.*, 115:367–393, 1991.

[6] C.R. Givens and R.M. Shortt. A class of Wasserstein metrics for probability distributions. *Michigan Math. J.*, 31:231–240, 1984.

[7] Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44:375–417, 1991.

[8] A. Prékopa. Logarithmic concave measures with application to stochastic programming. *Acta Sci. Math. (Szeged)*, 32:301–315, 1971.

[9] L. Leindler. On a certain converse of Hölder's inequality II. *Acta Sci. Math. (Szeged)*, 33:217–223, 1972.

[10] A. Prékopa. On logarithmic concave measures and functions. *Acta Sci. Math. (Szeged)*, 34:335–343, 1973.

[11] H.J. Brascamp and E.H. Lieb. On extensions of the Brunn-Minkowksi and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *J. Funct. Anal.*, 22:366–389, 1976.

[12] E.H. Lieb. Thomas-Fermi and related theories of atoms and molecules. *Rev. Modern Phys.*, 53:603–641, 1981.

[13] E.H. Lieb and H.-T. Yau. The Chandrasekhar theory of stellar collapse as the limit of quantum mechanics. *Comm. Math. Phys.*, 112:147–174, 1987.

[14] D.C. Downson and B.V. Landau. The Fréchet distance between multivariate normal distributions. *J. Multivariate Anal.*, 12:450–455, 1982.

[15] H. Federer. *Geometric Measure Theory.* Springer-Verlag, New York, 1969.

[16] H. Hadwiger and D. Ohmann. Brunn-Minkowskischer Satz und Isoperimetrie. *Math. Z.*, 66:1–8, 1956.

[17] R. Schneider. *Convex bodies: the Brunn-Minkowski theory.* Cambridge University Press, Cambridge, 1993.

[18] F. Riesz. Sur une inégalité intégrale. *J. London Math. Soc.*, 5:162–168, 1930.

[19] W. Rudin. *Real and Complex Analysis.* McGraw-Hill Book Company, New York, 1987.

[20] L. Caffarelli. The regularity of mappings with a convex potential. *J. Amer. Math. Soc.*, 5:99–104, 1992.

[21] L. Caffarelli. Boundary regularity of maps with convex potentials. *Comm. Pure Appl. Math.*, 45:1141–1151, 1992.

[22] L. Caffarelli. Boundary regularity of maps with convex potentials - II. Preprint.

[23] F. Almgren, J.E. Taylor and L. Wang. Curvature-driven flows: A variational approach. *SIAM J. Control Optim.*, 31:387–438, 1993.

[24] J.E. Taylor. Crystalline variational problems. *Bull. Amer. Math. Soc.*, 84:568–588, 1978.

[25] G.C. Shephard and R.J. Webster. Metrics for sets of convex bodies. *Mathematika*, 12:73–88, 1965.

[26] S. Angenent and M.E. Gurtin. Multiphase thermomechanics with interfacial structure 2. Evolution of an isothermal interface. *Arch. Rational Mech. Anal.*, 108:323–391, 1989.

[27] M. Gage and R.S. Hamilton. The heat equation shrinking convex plane curves. *J. Differential Geom.*, 23:69–96, 1986.

[28] G. Huisken. Flow by mean curvature of convex surfaces into spheres. *J. Differential Geom.*, 20:237–266, 1984.

[29] S. Chandrasekhar. Ellipsoidal figures of equilibrium — an historical account. *Comm. Pure Appl. Math.*, 20:251–265, 1967.

[30] J.F.G. Auchmuty. Existence of axisymmetric equilibrium figures. *Arch. Rational Mech. Anal.*, 65:249–261, 1977.

[31] G. Auchmuty. The global branching of rotating stars. *Arch. Rational Mech. Anal.*, 114:179–194, 1991.

[32] L. Caffarelli and A. Friedman. The shape of axisymmetric rotating fluid. *J. Funct. Anal.*, 35:109–142, 1980.

[33] A. Friedman and B. Turkington. Asymptotic estimates for an axisymmetric rotating fluid. *J. Funct. Anal.*, 37:136–163, 1980.

[34] A. Friedman and B. Turkington. The oblateness of an axisymmetric rotating fluid. *Indiana Univ. Math. J.*, 29:777–792, 1980.

[35] A. Friedman and B. Turkington. Existence and dimensions of a rotating white dwarf. *J. Differential Equations*, 42:414–437, 1981.

[36] S. Chanillo and Y.Y. Li. On diameters of uniformly rotating stars. To appear in *Comm. Math. Phys.*

[37] J.F.G. Auchmuty and R. Beals. Models of rotating stars. *Astrophys. J.*, 165:79–82, 1971.

[38] E.C. Olson. Rotational velocities in early-type binaries. *Publ. Astron. Soc. Pacific*, 80:185–191, 1968.

[39] E.H. Lieb. Personal communication.

[40] R.A. Lyttleton. *The Stability of Rotating Liquid Masses.* Cambridge University Press, Cambridge, 1953.

[41] J.-L. Tassoul. *Theory of Rotating Stars.* Princeton University Press, Princeton, 1978.

[42] R.A. James. The structure and stability of rotating gas masses. *Astrophys. J.*, 140:552–582, 1964.

[43] P.L. Lions. The concentration-compactness principle in the calculus of variations. The locally compact case: Part 1. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 1:109–145, 1984.

[44] E.H. Lieb. Existence and uniqueness of the minimizing solution of Choquard's nonlinear equation. *Stud. Appl. Math.*, 57:93–105, 1977.

[45] R.T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, 1972.

[46] A.D. Aleksandrov. Existence and uniqueness of a convex surface with a given integral curvature. *C.R. (Doklady) Acad. Sci. URSS (N.S.)*, 35:131–134, 1942.