ADVANCES IN MATHEMATICS **128**, 153–179 (1997)
ARTICLE NO. AI971634

A Convexity Principle for Interacting Gases*

Robert J. McCann

Department of Mathematics, Brown University, Providence, Rhode Island 02912

Received June 24, 1995; accepted December 30, 1996

A new set of inequalities is introduced, based on a novel but natural interpolation between Borel probability measures on $\mathbf{R}^d$. Using these estimates in lieu of convexity or rearrangement inequalities, the existence and uniqueness problems are solved for a family of attracting gas models. In these models, the gas interacts with itself through a force which increases with distance and is governed by an equation of state $P = P(\varrho)$ relating pressure to density. $P(\varrho)/\varrho^{(d-1)/d}$ is assumed non-decreasing for a d-dimensional gas. By showing that the internal and potential energies for the system are convex functions of the interpolation parameter, an energy minimizing state—unique up to translation—is proven to exist. The concavity established for $\|\rho_t\|_q^{-p/d}$ as a function of $t \in [0, 1]$ generalizes the Brunn–Minkowski inequality from sets to measures. © 1997 Academic Press

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. The study of an interacting gas model in which the force of attraction increases with distance has led us to the discovery of a new convexity principle. It is based upon a novel but natural interpolation between pairs of probability measures on $\mathbf{R}^d$. The current manuscript develops this theory, and exploits it to prove existence and uniqueness results for the attracting gas. In a subsequent article (or see [17]), the same technique will be used to settle the uniqueness question for the equilibrium shape of a two-dimensional crystal in a convex potential. The underlying estimates—which include a generalization of the Brunn–Minkowski inequality from sets to measures—appear to be both general and powerful: they bring tools of convex analysis to bear on problems in which they have not formerly been thought to apply.

* This work is based on a Princeton University Ph.D. thesis.

0001-8708/97 $25.00
Copyright © 1997 by Academic Press
All rights of reproduction in any form reserved.

Consider a d-dimensional gas of particles. The state of the gas is represented by its mass density $\rho(x) \geqslant 0$ on $\mathbf{R}^d$. Since the total amount of gas should be finite, $\rho \in L^1(\mathbf{R}^d)$, and a suitable normalization ensures $\int \rho = 1$. Thus $\rho \in \mathscr{P}_{ac}(\mathbf{R}^d)$, the space of absolutely continuous probability measures on $\mathbf{R}^d$. An attraction between the particles which increases with distance is represented by a strictly convex interaction potential $V(x)$ on $\mathbf{R}^d$. Resistance of the gas to compression is modelled by an equation of state in which the pressure $P(\varrho)$ depends on the local density $\varrho = \rho(x)$ only. The question is then: is it possible for these two forces to balance each other, and if they do, must the system be in a uniquely determined, stable equilibrium state?

Of course, this problem can be formulated variationally. To each state of the gas corresponds an energy $E(\rho) = U(\rho) + G(\rho)/2$ consisting of an internal energy due to compression plus a potential energy due to the interaction, and one wants to know whether the competition between these two terms results in a unique ground state. The energy of the gas is given by

$$E(\rho) := \int_{\mathbf{R}^d} A(\rho(x))\, dx + \tfrac{1}{2} \iint d\rho(x)\, V(x-y)\, d\rho(y), \tag{1}$$

where the first integral is the internal energy $U(\rho)$. Its density $A(\varrho)$ is derived from the pressure through (24). To be physical, $P(\varrho) \geqslant 0$ should be non-decreasing, in which case $A(\varrho)$ is convex. Under slightly stronger assumptions—notably $P(\varrho)/\varrho^{(d-1)/d}$ non-decreasing—$E(\rho)$ will be shown to admit a minimizer in $\mathscr{P}_{ac}(\mathbf{R}^d)$ which is 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 is the semi-classical approximation to the quantum kinetic energy of a gas of fermions: in three dimensions $q = 5/3$—see e.g. [13].

The energy $E(\rho)$ is not convex. However, for each pair of measures $\rho, \rho' \in \mathscr{P}_{ac}(\mathbf{R}^d)$ and $t \in (0, 1)$, we show that it is possible to define—see (7) below—an interpolant $\rho_t \in \mathscr{P}_{ac}(\mathbf{R}^d)$ for which both internal and potential energies satisfy estimates of the form

$$E(\rho_t) \leqslant (1-t)\, E(\rho) + tE(\rho'); \tag{2}$$

they are convex functions of the interpolation parameter t. Unlike $(1-t)\,\rho + t\rho'$, the interpolant ρ_t will be a translate of ρ when ρ' is; strict inequality fails in (2) precisely when this is the case. Uniqueness of the energy minimizer follows immediately. For potentials which are spherically symmetric, this uniqueness of ground state persists even if the *strict* convexity of $V(x)$ is relaxed, but it is a strength of the method that this symmetry not be required.

The estimate (2) also facilitates the continuity-compactness argument which assures that a ground state exists. Since the energy $E(\rho)$ is translation invariant, it necessary to prevent the escape of mass to infinity when extracting a limit from a minimizing sequence of states. Even without spherical symmetry, Newton's Third Law or the symmetry in (1) show that $V(x)$ may be taken to be even $V(x)=V(-x)$: it can always be replaced by $\frac{1}{2}[V(x)+V(-x)]$. Thus both $\rho \in \mathscr{P}_{ac}(\mathbf{R}^d)$ and $\rho'(x) := \rho(-x)$ share the same energy. Inequality (2) shows that, for the purpose of energy minimization, ρ may be replaced by the symmetrical configuration $\rho_{1/2}(x)=\rho_{1/2}(-x)$ which interpolates between ρ and ρ'. After the sequence has been centered in this way, an elementary estimate precludes the escape of any mass to infinity.

However, the uniqueness result remains more remarkable: the loss of compactness might be surmounted through Lions' concentration compactness lemma [15], but there are very few tools for addressing uniqueness when convexity fails. Even for a spherically symmetric potential $V(x)$, the alternative would be to use a sharp rearrangement inequality to reduce the problem to one-dimension, and then to attempt an analysis of the associated ordinary differential equation. Such an approach has been successfully exploited by Lieb and Yau [14] to handle the important case of Coulomb attraction $V(x)=-|x|^{-1}$ with the Chandrasekhar equation of state.

Formally, the minimizers of $E(\rho)$ are solutions of

$$\rho \nabla \int V(x-y)\, d\rho(y) = -\nabla P. \tag{3}$$

This equation, which expresses the balance of forces (Newton's Second Law) is obtained as the gradient of the Euler–Lagrange equation for $E(\rho)$ [2]. Since (3) is formally equivalent to $(d/dt)|_{t=0}\, E(\rho_t)=0$, the convexity of $E(\rho_t)$ could presumably be used to show that energy minimizers are the only solutions to (3). However, apart from this heuristic remark, we do not consider equation (3) further, being content to establish existence and uniqueness results at the level of the energy functional.

The estimates (2) may be of some interest apart from the application. Convexity of the internal energy $U(\rho_t)$ is a generalization of the Brunn–Minkowski inequality from sets to measures: the classical inequality is recovered from the case $A(\varrho)=-\varrho^{(d-1)/d}$ by interpolating between uniform probability measures on two given sets. For the $L^q(\mathbf{R}^d)$ norm, related inequalities are derived: $\|\rho_t\|_q^{-p/d}$ proves concave as a function of t where $q \geqslant (d-1)/d$ and p is its Hölder conjugate $p^{-1}+q^{-1}=1$. This assertion is sharp in the sense that $\|\rho_t\|_q^{-p/d}$ will be an affine function of t when ρ' is a dilate of ρ. Finally, it should be remarked that the monotonicity assumption required of $P(\varrho)/\varrho^{(d-1)/d}$ merely states that the internal energy $U(\rho)$ be convex non-increasing as a function of dilation factor for mass preserving dilations of ρ.

The organization of this manuscript is as follows. In the next section, the interpolant ρ_t is defined; its elementary properties, including convexity of $G(\rho_t)$, are set forth. Section 2 proves and discusses the deeper result—convexity of the internal energy $U(\rho_t)$—although technical details underlying the proof are relegated to Section 4. The existence and uniqueness theorems for the attracting gas comprise Section 3. An appendix establishes some notation and facts of life regarding differentiability properties of convex functions.

1. INTERPOLATION OF PROBABILITY MEASURES

The current section is devoted to defining and establishing the basic properties of the convex structure on $\mathscr{P}_{ac}(\mathbf{R}^d)$ which is here introduced. A brief digression on the interaction energy $G(\rho)$ motivates the definitions and theorems. For simplicity, the key definition is given on the line $d=1$ before being extended to measures on $\mathbf{R}^d$ through a theorem of Brenier [4, 5].

The energy $G(\rho)$ may be defined (23) on the space $\mathscr{P}(\mathbf{R}^d)$ of all Borel probability measures on $\mathbf{R}^d$. The attractive potential $V(x)=V(-x)$ precludes convexity of $G(\rho)$: a Dirac point mass δ_x at $x\in\mathbf{R}^d$ will minimize $G(\rho)$ while $(1-t)\,\delta_x+t\,\delta_y$ will not. (For the potential $V(x)=x^2$, it is even true that $G(\rho)$ is concave when restricted to $\mathscr{P}(\mathbf{R}^d)$.) However. if $\rho_t=\delta_{(1-t)\,x+ty}$ is used instead of $(1-t)\,\delta_x+t\,\delta_y$ to interpolate between two Dirac measures, then the potential energy $G(\rho_t)$ will be t-independent as a reflection of its translation invariance. Moreover, for a positive linear combination of such point masses

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

convexity of $V(x)$ implies $G(\rho_t)$ convex as a function of t. This point of view, which emphasizes the linear structure of $\mathbf{R}^d$ over that of the measure space, is reminiscent of the Lagrangian formulation in fluid mechanics. It indicates how ρ_t must be defined.

For measures $\rho, \rho'\in\mathscr{P}_{ac}(\mathbf{R})$ on the line, the definition is as follows. Given $x\in\mathbf{R}$, there exists $y(x)\in\mathbf{R}\cup\{\pm\infty\}$ such that

$$\rho[(-\infty, x)]=\rho'[(-\infty, y(x))]. \tag{4}$$

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 $(-\infty, (1-t)\,x + ty(x))$. This condition characterizes ρ_t.

To define ρ_t more generally requires a few notions from measure theory. A measure $\rho \in \mathscr{P}(\mathbf{R}^d)$ together with a Borel transformation $y: \mathbf{R}^d \to \mathbf{R}^n$ defined ρ almost everywhere, induce a measure $y_{\#}\,\rho$ on $\mathbf{R}^n$ given by

$$y_{\#}\,\rho[M] := \rho[y^{-1}(M)] \tag{5}$$

for Borel $M \subset \mathbf{R}^n$. $y_{\#}\,\rho$ is called the *push-forward* of ρ through y; it is a Borel probability measure, though it may not be absolutely continuous with respect to Lebesgue. The change of variables theorem states that if f is a (Borel) measurable function on $\mathbf{R}^n$, then

$$\int_{\mathbf{R}^n} f\, dy_{\#}\,\rho = \int_{\mathbf{R}^d} f(y(x))\, d\rho(x). \tag{6}$$

d-dimensional Lebesgue measure will play a frequent role; it is denoted by *vol*.

Given $\rho, \rho' \in \mathscr{P}_{ac}(\mathbf{R}^d)$, we require a transformation y which pushes ρ forward to ρ'. Although there are many such y, two further properties will prove essential:

(Y1) y must be locally irrotational;

(Y2) globally, y must not involve crossings,

$$(1-t)\,x + ty(x) = (1-t)\,x' + ty(x') \qquad \text{implies} \quad x = x' \quad \text{if } t \in [0, 1).$$

Such a transformation y may be constructed through a recursive procedure [17, Appendix C]—and indeed our results were originally obtained in this way—but the construction suffers from a serious flaw: the resulting map y is quite ugly, being grossly discontinuous. Brenier's theorem [4, 5] offers a beautiful alternative: as extended (see Theorem A.3 below) in [16], it states that y may be taken to be the gradient of a convex function $\psi: \mathbf{R}^d \to \mathbf{R} \cup \{+\infty\}$. Although ψ need not be unique, the map $\nabla\psi$ is uniquely determined ρ-almost everywhere. This theorem is used to define the *displacement interpolation* between ρ and ρ':

Definition 1.1 (Displacement interpolation). Given probability measures $\rho, \rho' \in \mathscr{P}_{ac}(\mathbf{R}^d)$, there exists ψ convex on $\mathbf{R}^d$ such that $\nabla\psi_{\#}\,\rho = \rho'$. Let *id* denote the identity mapping on $\mathbf{R}^d$. At time $t \in [0, 1]$, the displacement interpolant $\rho_t \in \mathscr{P}(\mathbf{R}^d)$ between ρ and ρ' is defined by

$$\rho_t := [(1-t)\, id + t\, \nabla\psi]_{\#}\, \rho. \tag{7}$$

This definition works equally well for all $t \in \mathbf{R}$, but such values of t will be irrelevant here and therefore suppressed. On the line $d = 1$, the monotone function $y = \nabla\psi(x)$ is readily seen to satisfy (4), and the characterization given for ρ_t follows rapidly.

What may not yet be clear is the absolute continuity of ρ_t with respect to Lebesgue; this shall be proved in a moment. Another consequence of Definition 1.1 is verified first: the convexity of the interaction energy $G(\rho)$ in (23) along the *lines* of the displacement interpolation. We say that the functional $G(\rho)$ is *displacement convex*.

PROPOSITION 1.2 (Displacement convexity of potential energy $G(\rho)$). *Given probability measures $\rho, \rho' \in \mathscr{P}_{ac}(\mathbf{R}^d)$, let ρ_t be the displacement interpolant between them (7). Then the interaction energy $G(\rho_t)$ is a convex function of t on $[0, 1]$. If the convexity of $V(x)$ is strict, then $G(\rho_t)$ fails to be strictly convex only when ρ' is a translate of ρ.*

Proof. By the change of variables theorem (6)

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

Since $V(x)$ is a convex function on $\mathbf{R}^d$, 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

$$\nabla\psi(x) - \nabla\psi(y) = x - y. \tag{8}$$

The *integral* will be strictly convex unless (8) 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 $\nabla\psi(x) - x$. ∎

The displacement convexity of the internal energy $U(\rho)$ is a deeper result. There the convexity of ψ, not used in the preceding proof, enters crucially. Before attacking this issue, it will be worthwhile to illuminate some of the elementary properties of the displacement interpolation. The next propositions show that it induces a bona fide convex structure on

$\mathscr{P}_{ac}(\mathbf{R}^d)$ and explore the relationship between this structure and the symmetries of $\mathbf{R}^d$—translation, dilation, reflection, rotation. The proofs are postponed until the end of this section. 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 1.3. *Let $\rho, \rho' \in \mathscr{P}_{ac}(\mathbf{R}^d)$. For $t \in [0, 1]$, the displacement interpolant $\rho_t = \rho \xrightarrow{t} \rho'$ from (7) satisfies*

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

(ii) ρ_t *is absolutely continuous with respect to Lebesgue*;

(iii) $\rho \xrightarrow{t} \rho' = \rho' \xrightarrow{1-t} \rho$;

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

Remark 1.4. In order to verify the displacement convexity of a functional $W\colon \mathscr{P}_{ac}(\mathbf{R}^d) \to \mathbf{R} \cup \{+\infty\}$ it is enough to show that for $\rho, \rho' \in \mathscr{P}_{ac}(\mathbf{R}^d)$, $W(\rho_t) \leqslant (1-t)\, W(\rho) + tW'(\rho')$. For $\lambda = (1-s)\, t + st'$, Proposition 1.3(iv) then implies that $W(\rho_\lambda) \leqslant (1-s)\, W(\rho_t) + sW(\rho_{t'})$.

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

Proposition 1.5. *Let $\rho, \rho' \in \mathscr{P}_{ac}(\mathbf{R}^d)$ with displacement interpolant $\rho_t = \rho \xrightarrow{t} \rho'$. Denote by T_μ the translation $T_\mu(x) = x + \mu$ for $x, \mu \in \mathbf{R}^d$ and by λ the dilation $\lambda(x) = \lambda x$ on $\mathbf{R}^d$ by a factor $\lambda \geqslant 0$. Λ denotes either T_μ, λ or a member of the orthogonal group on $\mathbf{R}^d$. If $v \in \mathbf{R}^d$, $\alpha, \beta > 0$ and $s, t \in [0, 1]$ then*

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

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

(iii) $\lambda_{\#}\, \rho_t = (\alpha_{\#}\, \rho) \xrightarrow{s} (\beta_{\#}\, \rho')$ *if* $\lambda(1-t) = \alpha(1-s)$ *and* $\lambda t = \beta s$.

Example 1.6 (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 1.5(ii) shows that for $\rho' = T_v \rho$ a translate of ρ, the displacement interpolant is $\rho_t = T_{tv}\rho$. For a dilate $\rho' = \beta_{\#}\, \rho$, the displacement interpolant is $\rho_t = \lambda_{\#}\, \rho$ with $\lambda = (1-t) + t\beta$.

Example 1.7 (Gaussian measures). Let $\rho_0, \rho_1 \in \mathscr{P}_{ac}(\mathbf{R}^d)$ 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 $\mu_i \in \mathbf{R}^d$ $(i=0, 1)$ and denote its covariance by Σ_i:

$$\Sigma_{ijk} := \int_{\mathbf{R}^d} x_j x_k \, d\rho_i(x) \qquad j, k = 1, ..., d;$$

$\Sigma_i > 0$ is a *positive* matrix, meaning positive definite and self-adjoint. It suffices to find ρ_t when $\mu_0 = \mu_1 = 0$, since Proposition 1.5(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 (6), 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 $\Lambda > 0$ be matrix positive. Although the matrix equation $\Lambda\Sigma_0\Lambda^\dagger = \Sigma_1$ has many solutions, the uniqueness part of Theorem A.3 shows that only one can be positive; it is $\Lambda = \Sigma_1^{1/2}(\Sigma_1^{1/2}\Sigma_0\Sigma_1^{1/2})^{-1/2}\,\Sigma_1^{1/2}$, as has previously been noted by several authors [7, 18, 9, 11]. Here $\Sigma^{1/2}$ denotes the positive square root of Σ. By uniqueness, ρ_t must be the Gaussian measure $[(1-t)\, id + t\Lambda]_\# \rho_0$.

Remark 1.8 (Singular measures and continuity). The displacement interpolation may be extended to the case where the endpoints ρ, ρ' lie in the set $\mathscr{P}(\mathbf{R}^d)$ of all Borel probability measures on $\mathbf{R}^d$. Let ψ be a convex function on $\mathbf{R}^d$. As a subset of $\mathbf{R}^d \times \mathbf{R}^d$, the graph of $\nabla\psi$ is characterized by a property (33) known as *cyclical monotonicity*. Here $\langle\ ,\ \rangle$ denotes the Euclidean inner product, so the two-point inequality

$$\langle \nabla\psi(x) - \nabla\psi(y), x - y\rangle \geqslant 0 \tag{9}$$

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. Theorem A.3 asserts the existence of a joint probability measure $\gamma \in \mathscr{P}(\mathbf{R}^d \times \mathbf{R}^d)$ with cyclically monotone support having ρ and ρ' as its *marginals* $\Pi_\# \gamma = \rho$ and $\Pi'_\# \gamma = \rho'$ where $\Pi(x, y) = x$ and $\Pi'(x, y) = y$. Let $t \in [0, 1]$ and define

$$\Pi_t(x, y) := (1-t)\, x + ty \tag{10}$$

on $\mathbf{R}^d \times \mathbf{R}^d$. Then $\rho_t := \Pi_{t\#}\gamma$. Since $\Pi(x, \nabla\psi(x)) = (1-t)\, x + t\, \nabla\psi(x)$, Theorem A.3 also ensures equivalence of this definition with (7). As a caveat, we note that unless ρ or ρ' vanishes on all sets of Hausdorff dimension $d-1$, the interpolant ρ_t may fail to be unique.

A second fact, also true but not required, is that the map from $\mathscr{P}_{ac}(\mathbf{R}^d) \times [0, 1] \times \mathscr{P}(\mathbf{R}^d)$ to $\mathscr{P}(\mathbf{R}^d)$ which takes (ρ, t, ρ') to $\rho \xrightarrow{t} \rho'$ is continuous. As in Section 3 the measure spaces are topologized using convergence against $C_\infty(\mathbf{R}^d)$ test functions.

Proof of Proposition 1.3. Let ψ be convex with $\rho' = \nabla\psi_\# \rho$. Then (i) is obvious. To see (ii), let $\phi(x) := (1-t)\, x^2/2 + t\, \nabla\psi(x)$ denote the function whose gradient pushes forward ρ to ρ_t. The claim is that if $M \subset \mathbf{R}^d$ 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

$$
\begin{aligned}
&|\nabla\phi(x) - \nabla\phi(y)|\; |x-y| \\
&\qquad \geqslant \langle \nabla\phi(x) - \nabla\phi(y), x-y\rangle &(11)\\
&\qquad = (1-t)\, |x-y|^2 + t\, \langle \nabla\psi(x) - \nabla\psi(y), x-y\rangle, &(12)
\end{aligned}
$$

(9) shows that $(\nabla\phi)^{-1}$ is Lipschitz with constant no greater than $(1-t)^{-1}$. (ii) is then a consequence of a standard measure theoretic result: $\mathrm{vol}(\nabla\phi)^{-1}\, M \leqslant (1-t)^{-d}\, \mathrm{vol}\, M$ [8].

The alternative definition of ρ_t given in Remark 1.8 provides the easiest way to see (iii). Let $\gamma \in \mathscr{P}(\mathbf{R}^d \times \mathbf{R}^d)$ be the joint probability measure with cyclically monotone support and ρ and ρ' as its marginals. Let $\Pi_t(x, y)$ be the map (10) pushing γ forward to $\rho \xrightarrow{t} \rho'$. If $*$ denotes the involution $*(x, y) = (y, x)$ on $\mathbf{R}^d \times \mathbf{R}^d$, then $*_\# \gamma$ has cyclically monotone support, and ρ' and ρ as its marginals; it pushes forward to $\rho' \xrightarrow{1-t} \rho$ under Π_{1-t}. Since $\Pi_{1-t}(y, x) = \Pi_t(x, y)$, (iii) is proved.

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

$$
\rho \xrightarrow{st} \rho' = \rho \xrightarrow{s} \rho_t \tag{13}
$$

to prove (iv). ϕ as above satisfying $\rho_t = \nabla\phi_\# \rho$ is used to define $\rho \xrightarrow{s} \rho_t$; (13) follows from $(1-s)\, id + s\, \nabla\phi = (1-st)\, id + st\, \nabla\psi$. Now let $\lambda = (1-s)\, t + st'$, and noting (iii) take $t' \leqslant t$ without loss of generality. Then (13) and (iii) imply $\rho_{t'} = \rho \xrightarrow{t'/t} \rho_t = \rho_t \xrightarrow{1-t'/t} \rho$ and also $\rho_\lambda = \rho \xrightarrow{\lambda/t} \rho_t = \rho_t \xrightarrow{1-\lambda/t} \rho$. Since $(t-\lambda)/(t-t') = s \leqslant 1$, (13) once more yields $\rho_\lambda = \rho_t \xrightarrow{s} \rho_{t'}$. ▮

Proof of Proposition 1.5. This proposition is most easily seen via the alternative definition of ρ_t given in Remark 1.8. There ρ_t is defined using the measure $\gamma \in \mathscr{P}(\mathbf{R}^d \times \mathbf{R}^d)$ with cyclically monotone support and ρ and ρ' as its marginals. The relevant observation is that a cyclically monotone subset of $\mathbf{R}^d \times \mathbf{R}^d$ remains cyclically monotone under any of the transformations $\Lambda \times \Lambda$, $T_\mu \times T_\nu$ or $\alpha \times \beta$. The result (i), (ii) or (iii) is then obtained by pushing γ forward through one of these transformations: the push-forward has cyclically monotone support, and the correct marginals by the change of variables theorem (6). Defining $\Pi_t(x, y)$ as in (10), the results follow from

$$\Lambda \Pi_t(x, y) = \Pi_t(\Lambda x, \Lambda y),$$
$$(1-t)\,\mu + tv + \Pi_t(x, y) = \Pi_t(x+\mu, y+v), \qquad \text{and}$$
$$\lambda \Pi_t(x, y) = \Pi_s(\alpha x, \beta y). \quad \blacksquare$$

2. DISPLACEMENT CONVEXITY OF $\int A(\rho)$

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

$$U(\rho) := \int_{\mathbf{R}^d} A(\rho(x))\, dx \tag{14}$$

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

To any $\rho \in \mathscr{P}_{ac}(\mathbf{R}^d)$ is associated the family of dilates $\lambda_{\#}\rho$ which may be obtained as the push-forward of ρ through dilation of $\mathbf{R}^d$ by some factor $\lambda > 0$. The condition for displacement convexity of $U(\rho)$ is merely this: $U(\lambda_{\#}\rho)$ should be convex non-increasing as a function of λ. The necessity of the convexity is obvious; its sufficiency is the content of Theorem 2.2. 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 $A\colon [0, \infty) \to \mathbf{R} \cup \{+\infty\}$, 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 the following observation. Consider mass m of a gas whose internal energy is given by (14). 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. Then $U(t)$ is a convex function of time. Properties (Y1) and (Y2) of $\nabla\psi$, and the linearity of $(1-t)\, id + t\, \nabla\psi$, allow this observation to be used as a local inequality which may be integrated to yield Theorem 2.2. The underlying intuition is the following: the displacement interpolation transfers a small mass of gas with near constant density $\rho(x)$ from a neighbourhood of x

to a neighbourhood of $\nabla\psi(x)$. Here $\nabla\psi$ may be linearly approximated through the non-negative matrix $\nabla^2\psi(x)$, so the neighbourhood can be chosen to be a small cube with sides parallel to the eigenvectors of $\nabla^2\psi(x)$. The contribution of this bit of gas to $U(\rho_t)$ is then $U(t)$. Property (Y2), which follows from (12), ensures that two cubes, initially disjoint, do not interfere with each other during their subsequent motion.

Before proving the theorem, a standard lemma is stated without proof.

LEMMA 2.1. *Let Λ be a non-negative $d \times d$ matrix and* $v(t) := \det[(1-t)I + t\Lambda]$ *where I is the identity matrix. Then $v^{1/d}(t)$ is concave on $t \in [0, 1]$, and the concavity will be strict unless $\Lambda = \lambda I$.*

In the basis for which Λ is diagonal, this lemma is seen to result from the domination of the geometric by the arithmetic mean [10]. It underlies Hadwiger and Ohmann's proof of the Brunn–Minkowski theorem.

THEOREM 2.2 (Displacement convexity of internal energy $U(\rho)$). *Let $\rho, \rho' \in \mathscr{P}_{ac}(\mathbf{R}^d)$, and define the displacement interpolant $\rho_t = \rho \xrightarrow{t} \rho'$ 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 the Aleksandrov second derivative* (35) *coincides ρ-a.e. with the identity matrix, $\nabla^2\psi(x) = I$, in which case $U(\rho_t) = U(\rho)$.*

Proof. Proposition 1.3(ii) shows that ρ_t is absolutely continuous (with respect to Lebesgue). By Theorem 4.4 (monotone change of variables), the set X on which $\nabla^2\psi(x) > 0$ and its inverse exist has full measure for ρ, and moreover

$$U(\rho_t) = \int_X A\left(\frac{\rho(x)}{\det[(1-t)I + t\nabla^2\psi(x)]}\right)\det[(1-t)I + t\nabla^2\psi(x)]\,dx. \quad (15)$$

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 $x \in X$, and let $\Lambda := \nabla^2\psi(x)$ and $v(t) := \det[(1-t)I + t\Lambda]$. Since Λ is positive, Lemma 2.1 shows that $v^{1/d}(t)$ is concave. Composing the convex non-increasing function $h(\lambda) := \lambda^d A(\rho(x)/\lambda^d)$ with the concave function $g(t) := v^{1/d}(t)$ yields a convex function $h \circ g(t)$—the integrand of (15). Strict convexity of $h \circ g$ follows from that of h unless $g(t)$ is a constant. The latter implies $\nabla^2\psi(x) = \lambda I$ with $\lambda = 1$. Integrating proves the result. ∎

In the special case $A(\varrho) = \varrho^q$, a scaling argument strengthens Theorem 2.2 considerably. For the displacement interpolation, convexity of the $L^q(\mathbf{R}^d)$ norm turns out to be better than *logarithmic*: if p is Hölder conjugate to q then $\|\rho\|_q^{-p/d}$ is displacement concave. This inequality is sharp in the sense that $\|\lambda_{\#}\rho\|_q^{-p/d}$ depends linearly on $\lambda > 0$ for a mass preserving dilation of ρ.

THEOREM 2.3 (Displacement concavity of $\|\rho\|_q^{-p/d}$). *Let $\rho_t = \rho \xrightarrow{t} \rho'$ be the displacement interpolant between $\rho, \rho' \in \mathscr{P}_{ac}(\mathbf{R}^d)$. Let $0 < q \leqslant \infty$ satisfy $q \geqslant (d-1)/d$ and define $\alpha := -(1-1/q)^{-1}/d$. Then*

$$\|\rho_t\|_q^\alpha \geqslant (1-t)\ \|\rho\|_q^\alpha + t\ \|\rho'\|_q^\alpha. \tag{16}$$

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 $\rho, \rho' \in L^q(\mathbf{R}^d)$. Letting $\lambda_{\#}$ denote dilation by $\lambda > 0$, it is possible to choose factors $\lambda, \lambda' > 0$ such that $\|\lambda_{\#}\rho\|_q = \|\lambda'_{\#}\rho'\|_q$ and $(1-t)/\lambda + t/\lambda' = 1$. Setting $s = t/\lambda' \in (0, 1)$. Proposition 1.5(iii) shows that $\rho_t = \lambda_{\#}\rho \xrightarrow{s} \lambda'_{\#}\rho'$. Because $A(\varrho) = \varrho^q$ satisfies (A1), Theorem 2.2 shows $\|\rho_t\|_q^q$ to be convex as long as $q < \infty$:

$$\|\rho_t\|_q^q \leqslant (1-s)\ \|\lambda_{\#}\rho\|_q^q + s\ \|\lambda'_{\#}\rho'\|_q^q \tag{17}$$

$$= \|\lambda_{\#}\rho\|_q^q. \tag{18}$$

Since t^α is decreasing for $q > 1$,

$$\|\rho_t\|_q^\alpha \geqslant \|\lambda_{\#}\rho\|_q^\alpha \tag{19}$$

$$= (1-s)\ \|\lambda_{\#}\rho\|_q^\alpha + s\ \|\lambda'_{\#}\rho'\|_q^\alpha. \tag{20}$$

In the case $q = \infty$, (19) follows immediately from Theorem 2.2 with $A(\varrho) = 0$ where $\varrho \leqslant \|\lambda_{\#}\rho\|_\infty$ and $A(\varrho) = \infty$ otherwise. Either way, the case $q > 1$ is established for $\rho, \rho' \in L^q(\mathbf{R}^d)$ by the scaling relation $\|\lambda_{\#}\rho\|_q = \lambda^{1/\alpha}\ \|\rho\|_q$ in (20). If $\|\rho'\|_q = \infty$, a separate argument is required: $\|\rho_t\|_q \leqslant \|(1-t)_{\#}\rho\|_q = (1-t)^{1/\alpha}\ \|\rho\|_q$ follows directly from (15), $\det[(1-t)I + t\nabla^2\psi] \geqslant (1-t)^d$ and monotonicity in (A1). When $q < 1$ the argument is the same, except that the inequality in (17) is reversed because it is $A(\varrho) = -\varrho^q$ which satisfies (A1); on the other hand, the inequality in (19) is restored because $\alpha > 0$. Taking the logarithm of (16), the convexity or concavity of $\log \|\rho_t\|_q$ follows according to the sign of α. Remark 1.4 has been noted. ∎

Remark 2.4 (Brunn–Minkowski inequality). This classical geometric inequality [25] compares the Lebesgue measures of two sets $K, K' \subset \mathbf{R}^d$ and their vector sum $K + K' = \{k + k' \mid (k, k') \in K \times K'\}$; for non-empty sets, it states that

$$\mathrm{vol}^{1/d}[K + K'] \geqslant \mathrm{vol}^{1/d} K + \mathrm{vol}^{1/d} K'. \tag{21}$$

The inequality follows from Theorem 2.2 when $U(\rho)$ is chosen to saturate condition (A1). Alternately, it may be recovered from Theorem 2.3 with $q \neq 1$ [17].

Proof of (21). 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. (21) is equivalent to the concavity of $\mathrm{vol}^{1/d}(1-t)\,K + tK'$ on $[0, 1]$. Therefore, let $\rho \in \mathscr{P}_{ac}(\mathbf{R}^d)$ 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 $(1-t)\,K + tK'$. Choosing $A(\varrho) = -\varrho^{(d-1)/d}$ implies $U(\rho) = -\mathrm{vol}^{1/d}\,K$ and $U(\rho') = -\mathrm{vol}^{1/d}\,K'$. Then Theorem 2.2 gives $-U(\rho_t) \geqslant (1-t)\,\mathrm{vol}^{1/d}\,K + t\,\mathrm{vol}^{1/d}\,K'$, while Jensen's inequality yields

$$U(\rho_t) \geqslant A\left(\frac{1}{\mathrm{vol}[\operatorname{spt} \rho_t]} \int \rho_t\right) \mathrm{vol}[\operatorname{spt} \rho_t] = -\mathrm{vol}^{1/d}[\operatorname{spt} \rho_t].$$

The theorem follows from these two inequalities and $(1-t)\,K + tK' \supseteq \operatorname{spt} \rho_t$. ∎

Remark 2.5. It is interesting to note that the inclusion $(1-t)\,K + tK' \supseteq \operatorname{spt} \rho_t$ 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. Considerations like those of Example 1.7 (see [7] also) 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 generally be an ellipsoid, as is easily seen when K is the unit ball and K' is highly eccentric (even degenerate).

Remark 2.6 (Inequalities of Prékopa, Leindler, Brascamp, and Lieb). Other generalizations of the Brunn–Minkowski inequality to functions on $\mathbf{R}^d$ are due to Brascamp and Lieb [3]. The simplest case of their result dates back to Prékopa [19, 20] and Leindler [12], and asserts that the interpolant

$$h(x) := \sup_{y \in \mathbf{R}^d} f\left(\frac{y}{1-t}\right)^{1-t} g\left(\frac{x-y}{t}\right)^{t} \tag{22}$$

defined between non-negative measurable functions f, g on $\mathbf{R}^d$ satisfies the inequality $\|h\|_1 \geqslant \|f\|_1^{1-t}\,\|g\|_1^{t}$ if $t \in (0, 1)$. By scaling, the case $\|f\|_1 = \|g\|_1 = 1$ is quite general. The displacement interpolant $f \xrightarrow{t} g \in \mathscr{P}_{ac}(\mathbf{R}^d)$ between f and g can then be defined, and the Prèkopa–Leindler theorem becomes a trivial consequence of the observation that $h \geqslant f \xrightarrow{t} g$: the inequality $\|h\|_1 \geqslant 1$ is saturated with the displacement interpolant in place of h! As it turns out, this observation can be parlayed [17] into a transparent proof not only of Prékopa and Leindler's result but also of the stronger inequalities due to Brascamp and Lieb.

3. 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 (1). In this model, the configuration of the gas is given by its mass density $\rho \in \mathscr{P}_{ac}(\mathbf{R}^d)$, and the interaction is through a convex potential V on $\mathbf{R}^d$. This leads to a potential energy

$$G(\rho) := \iint d\rho(x)\, V(x-y)\, d\rho(y). \tag{23}$$

Since $V(x) = V(-x)$ is minimized at the origin, taking $V(0) = 0$ costs no generality.

The gas is also assumed to satisfy an equation of state $P(\varrho)$ relating pressure to density, which leads to an internal energy $U(\rho)$ of the form (14). The local density $A(\varrho)$ of $U(\rho)$ is obtained by integrating $dU = -P\, dv$:

$$A(\varrho) := \int_1^\infty P(\varrho/v)\, dv. \tag{24}$$

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

(P1) $P\colon [0, \infty) \to [0, \infty]$ 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 (24) shows the equivalence of (P1) to the convexity of $U(\rho)$ under dilations, (A1) of the previous section. Strict monotonicity in the former is equivalent to strict convexity in the latter. 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, we show that the total energy $E(\rho) = U(\rho) + G(\rho) \geqslant 0$ attains a unique minimum up to translation, unless $E(\rho) = \infty$.

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 $C_\infty(\mathbf{R}^d)$ be the Banach space of continuous functions vanishing at ∞, under the sup norm. By the Riesz–Markov theorem, its dual $C_\infty(\mathbf{R}^d)^*$ consists of Borel measures of finite total variation. The relevant topology on $\mathscr{P}_{ac}(\mathbf{R}^d) \subset C_\infty(\mathbf{R}^d)^*$ will be the weak-$*$ topology, the topology of convergence against $C_\infty(\mathbf{R}^d)$ test functions.

THEOREM 3.1 (Existence and uniqueness of ground state). *Assume* (P1–P2) *and* $V\colon \mathbf{R}^d \to \mathbf{R} \cup \{+\infty\}$ *to be strictly convex. Let the energy* $E(\rho) = U(\rho) + G(\rho)$ *be given by* (1) *with* $A(\varrho)$ *from* (24), *and* $E_g := \inf E(\rho)$ *over* $\rho \in \mathscr{P}_{ac}(\mathbf{R}^d)$. *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 $\rho_g, \rho'_g \in \mathscr{P}_{ac}(\mathbf{R}^d)$ exist. Fix $t \in (0, 1)$ and consider the displacement interpolant $\rho_t = \rho_g \xrightarrow{t} \rho'_g$ between them. Since $U(\rho)$ and $G(\rho)$ are displacement convex (by Theorem 2.2 and Proposition 1.2), $E(\rho_t) \leqslant (1-t)\, E(\rho_g) + tE(\rho'_g) = E_g$. Strict inequality holds by Proposition 1.2 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 $E(\rho) \geqslant 0$, choose an energy minimizing sequence $\rho_n \in \mathscr{P}_{ac}(\mathbf{R}^d) \subset C_\infty(\mathbf{R}^d)^*$. By Lemmas 3.4 and 3.6 any weak-* limit point $\rho_g \in \mathscr{P}_{ac}(\mathbf{R}^d)$ of ρ_n must minimize $E(\rho)$. In fact, Corollary 3.5 applies because of (P2), and shows that ρ_g need only satisfy the mass constraint $\rho_g \in \mathscr{P}(\mathbf{R}^d)$: finiteness of E_g implies absolute continuity of ρ_g. The Banach–Alaoglu Theorem provides weak-* compactness of the unit ball in $C_\infty(\mathbf{R}^d)^*$ 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 $\mathbf{R}^d$. Propositions 1.5(i) and 1.3(iv) show the displacement interpolant $\rho \xrightarrow{1/2} \Lambda\rho$ 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: $V(x) \geqslant k > 0$ for $|x| = 1$. For $|x| > 1$ convexity yields $V(x) > k\,|x|$, in which case

$$\int_{\mathbf{R}^d} V(x-y)\, d\rho_n(y) \geqslant \int_{\langle y, x\rangle \leqslant 0} V(x-y)\, d\rho_n(y) \geqslant k\,|x|/2$$

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

$$G(\rho_n) \geqslant \frac{kR}{2} \int_{|x| > R} d\rho_n(x). \tag{25}$$

$E(\rho_n)$ may be assumed to be bounded above, so $G(\rho_n) \leqslant L$. Thus (25) controls the mass of ρ_n outside of any large ball, uniformly in n. If $0 \leqslant \varphi \leqslant 1$ is a $C_\infty(\mathbf{R}^d)$ test function with $\varphi = 1$ on $|x| \leqslant R$, weak-* convergence yields $\int \varphi \, d\rho_g \geqslant 1 - 2L/Kr$. Since R was arbitrary, $\rho_g[\mathbf{R}^d] = 1$. ∎

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 3.1 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 3.3 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 $\mathbf{R}^d$.

LEMMA 3.2. *Let ψ and ϕ be convex functions on $\mathbf{R}^d$, and $\Omega \subset \mathbf{R}^d$ an open convex set on which both ψ and ϕ are finite. Suppose ϕ is differentiable on Ω with a gradient $\nabla\phi\colon \Omega \to \mathbf{R}^d$ which is locally Lipschitz. If the Aleksandrov derivatives $\nabla^2\phi = \nabla^2\psi$ 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 $\Omega \subset \mathbf{R}$; 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 $\omega_{sing} \geqslant 0$. 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 Aleksandrov derivatives $\nabla^2\psi$ 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. ∎

THEOREM 3.3 (Uniqueness without strict convexity of $V(x)$). *Assume that $P(\varrho)$ satisfies the monotonicity condition* (P1) *strictly, and that the convex function $V\colon \mathbf{R}^d \to \mathbf{R}^d \cup \{+\infty\}$ is spherically symmetric, not constant. If the energy $E(\rho) = U(\rho) + G(\rho)$ given by* (1) *with $A(\varrho)$ from* (24) *attains a finite minimum at $\rho_g \in \mathscr{P}_{ac}(\mathbf{R}^d)$, 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

$$\mathrm{vol}\{\rho_g^* > k\} = \mathrm{vol}\{\rho_g > k\} \tag{26}$$

for all $k > 0$. The internal energy $U(\rho_g^*) = U(\rho_g)$ by (26), while a rearrangement inequality due to Riesz [21] states that $G(\rho_g^*) \leqslant G(\rho_g)$ since the potential $V(x)$ is symmetric non-decreasing. Thus ρ_g^* also minimizes $E(\rho)$. Suppose $\rho_g' \in \mathscr{P}_{ac}(\mathbf{R}^d)$ 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 2.2 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 3.2 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, $y \in \mathbf{R}^d$, then the monotonicity inequality (9) 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

$$|\nabla\psi(x) - \nabla\psi(y)| > |x - y|. \tag{27}$$

Unless $\psi(x) - x^2/2$ is affine on Ω, (27) will be satisfied at some x, $y \in \Omega$. By the continuity properties of $\nabla\psi$, (27) will continue to hold in a small neighbourhood of $(x, y) \in \mathbf{R}^d \times \mathbf{R}^d$—which is to say, on a set of positive measure $\rho_g^* \times \rho_g^*$. The change of variables formula (6) 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 (27) 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 [8]. 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 (27), 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 $\lambda_{\#}\rho^*_g$ of ρ^*_g by factor $\lambda\geqslant 1$. Defining $G(\lambda):=G(\lambda_{\#}\rho^*_g)$, it will be shown that $G(\lambda)$ grows sublinearly with small $\lambda-1$ while $U(\lambda):=U(\lambda_{\#}\rho^*_g)$ 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 $\lambda\geqslant 1$ 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 $m(\lambda)\leqslant k(\lambda-1)$. By continuity of $V(x)$, λ near 1 ensures $V(x)<\varepsilon$ for $x<\lambda r$, implying $G(\lambda)\leqslant\varepsilon m^2(\lambda)$. 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. ∎

The lower semi-continuity results, although far from novel, are included for completeness.

Lemma 3.4 (Weak-$*$ lower semi-continuity of $U(\rho)$). *Assume A: $[0, \infty)\to[0, \infty]$ is convex and lower semi-continuous. Define $U(\rho)$ by* (14). *Then $U(\rho)$ is weak-$*$ lower semi-continuous on $\mathscr{P}_{ac}(\mathbf{R}^d)\subset C_\infty(\mathbf{R}^d)^*$.*

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

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

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

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

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 $\rho * \varphi_\varepsilon \to \rho$ as $\varepsilon \to 0$. Another application of Fatou's Lemma and the lower semi-continuity of $A(\varrho)$ yields $\underline{\lim}_\varepsilon\, U(\rho * \varphi_\varepsilon) \geqslant U(\rho)$. ∎

Corollary 3.5. *In addition to the hypotheses of Lemma* 3.4, *suppose $A(\varrho)/\varrho$ diverges as $\varrho \to \infty$. Then $U(\rho)$ remains weak-$*$ lower semi-continuous if it is extended to $\mathscr{P}(\mathbf{R}^d) \subset C_\infty(\mathbf{R}^d)^*$ by taking $U(\rho) = \infty$ unless $\rho \in \mathscr{P}_{ac}(\mathbf{R}^d)$.*

Proof. The only case to check is that $\underline{\lim}_n\, U(\rho_n) = \infty$ when a sequence $\rho_n \in \mathscr{P}_{ac}(\mathbf{R}^d)$ tends to a limit $\rho \in \mathscr{P}(\mathbf{R}^d)$ 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 $\operatorname{vol} K = 0$, and an open set $N \supset K$ with arbitrarily small Lebesgue measure. Choose a $C_\infty(\mathbf{R}^d)$ test function $0 \leqslant \varphi \leqslant 1$ 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

$$\int_N A(\rho_n) \geqslant A\left(\frac{m}{\operatorname{vol}[N]}\right) \operatorname{vol}[N].$$

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

Lemma 3.6 (Weak-$*$ lower semi-continuity of $G(\rho)$). *Assume $V\colon \mathbf{R}^d \to [0, \infty]$ convex and define $G(\rho)$ by* (23). *Then $G(\rho)$ is weak-$*$ lower semi-continuous on $\mathscr{P}_{ac}(\mathbf{R}^d) \subset C_\infty(\mathbf{R}^d)$.*

Proof. Let $\rho_n \to \rho$ weak-$*$ in $\mathscr{P}_{ac}(\mathbf{R}^d)$. The claim is that $G(\rho) \leqslant \underline{\lim}_n\, G(\rho_n)$. Certainly the product measure $\rho_n \times \rho_n \to \rho \times \rho$ weak-$*$ in $C_\infty(\mathbf{R}^d \times \mathbf{R}^d)^*$. 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 $C_\infty(\mathbf{R}^d \times \mathbf{R}^d)$, 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, $G_r(\rho) = \lim_n G_r(\rho_n) < \underline{\lim}_n G(\rho_n)$. By Lebesgue's Monotone Convergence Theorem, $G_r(\rho)$ increases to $G(\rho)$ and the result is proved. ∎

4. MONOTONE CHANGE OF VARIABLES THEOREM

Let ψ be a convex function on $\mathbf{R}^d$, and denote the interior of the convex set $\{\psi < \infty\}$ by $\Omega := \operatorname{int} \operatorname{dom} \psi$. As the gradient of a convex function, $\nabla\psi : \Omega \to \mathbf{R}^d$ represents the generalization to higher dimensions of a monotone map on the line. It is a measurable transformation, defined and differentiable (in the sense (35) of Aleksandrov) almost everywhere, and will be casually referred to as a monotone map. Sundry notions related to ψ, including its subdifferential $\partial\psi$, Legendre transform ψ^* and Aleksandrov second derivative $\nabla^2\psi$ are defined in Appendix A. The goal of the current section is to establish Theorem 4.4, which contains the change of variables theory for monotone transformations required in the proof of Theorem 2.2. Although $\nabla\psi$ may not be Lipschitz, the Jacobian factor $\det[\nabla^2\psi(x)]$ appearing in (31) is just what one expects from the standard theory of Lipschitz transformations.

As before, $\nabla\psi$ will be used to push-forward certain positive Radon measures ρ from Ω to $\mathbf{R}^d$. That is, ρ may no longer have unit or even finite mass, but its mass will be finite on compact subsets. The set of such measures will be denoted $\mathscr{M}(\Omega)$. $\rho \in \mathscr{M}(\Omega)$ 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 $\mathscr{M}_{ac}(\Omega)$ 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 [24] is exploited to identify the pushed-forward measure: If $\rho \in \mathscr{M}(\Omega)$ for some domain Ω, its *symmetric derivative* $D\rho$ at $x \in \Omega$ is defined to be

$$D\rho(x) := \lim_{r \to 0} \frac{\rho[B_r(x)]}{\operatorname{vol} B_r}, \tag{28}$$

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 full measure for ρ_{sing}. Thus, knowing its symmetric derivative ρ-a.e., it is possible to reconstruct ρ_{ac} and determine whether $\rho_{sing} = 0$. At the Lebesgue points of $\rho \in \mathscr{M}_{ac}(\Omega)$, hence almost everywhere, the limit (28) remains unchanged if the balls $B_r(x)$ are replaced by a sequence of Borel sets M_n which *shrink nicely to* x, meaning there is a sequence $r(n) \to 0$ such that $M_n \subset B_{r(n)}(x)$ and the ratio vol M_n/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 section.

LEMMA 4.1. *Let ψ be a convex function with $\Omega := \operatorname{int} \operatorname{dom} \psi$, and denote the subdifferential of its Legendre transform by $\partial\psi^*$. Let $\rho \in \mathscr{M}_{ac}(\Omega)$. Under the push-forward $\nabla\psi_{\#}\rho$, the measure of a Borel set $M \subset \mathbf{R}^d$ 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. ∎

PROPOSITION 4.2. *Let ψ be convex on $\mathbf{R}^d$, $\Omega := \operatorname{int} \operatorname{dom} \psi$ and $\rho \in \mathscr{M}_{ac}(\Omega)$. Assume x is a Lebesgue point for ρ at which the Aleksandrov derivative $\Lambda := \nabla^2\psi(x)$ exists and is invertible. At $\nabla\psi(x)$, the symmetric derivative* (28) *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 Aleksandrov derivative. Proposition A.2 then shows that $\partial\psi^*(B_r(y))$ shrinks nicely to x. Since x is a Lebesgue point of ρ,

$$\frac{\rho[\partial\psi^*(B_r(y))]}{\operatorname{vol}[\partial\psi^*(B_r(y))]} \to \rho(x), \tag{29}$$

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

$$\frac{\operatorname{vol}[\partial\psi^*(B_r(y))]}{\operatorname{vol} B_r} \to \det \Lambda^{-1}. \tag{30}$$

Taking the product of these two limits and observing Lemma 4.1 proves the result. ∎

COROLLARY 4.3. *Let ψ be convex on $\mathbf{R}^d$ and $\Omega := \operatorname{int} \operatorname{dom} \psi$. Then the function $\det[\nabla^2\psi(x)] \geqslant 0$ 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 consists of the points for which the Aleksandrov derivative $\nabla^2\psi$ is defined and invertible which are also Lebesgue points for $\det[\nabla^2\psi]$.*

Proof. Consider the convex Legendre transform ψ^* and Lebesgue measure on int dom ψ^*. The first claim is that $\det[\nabla^2\psi] \in L^1_{loc}(\Omega)$; in fact, it is the absolutely continuous part of $\omega := \nabla\psi^*_{\#}$ 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 Aleksandrov's theorem implies existence of $\nabla^2\psi$ almost everywhere there. Where $\det[\nabla^2\psi(x)] > 0$, Proposition 4.2 (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 4.1 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.

A second application of Proposition 4.2, 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 $\partial\psi(M)$. ∎

THEOREM 4.4 (Monotone change of variables theorem). *Let $\rho, \rho' \in \mathscr{P}_{ac}(\mathbf{R}^d)$, and ψ be a convex function on $\mathbf{R}^d$ with $\nabla\psi_{\#}\rho = \rho'$. Let $X \subset \mathbf{R}^d$ denote the set of points where the Aleksandrov derivative $\nabla^2\psi$ is defined and invertible. Then X has full measure for ρ. If $A(\varrho)$ is a measurable function on $[0, \infty)$ with $A(0) = 0$ then*

$$\int A(\rho'(y))\,dy = \int_X A\left(\frac{\rho(x)}{\det[\nabla^2\psi(x)]}\right)\det[\nabla^2\psi(x)]\,dx. \tag{31}$$

(*Either both integrals are undefined or both take the same value in* $\mathbf{R} \cup \{\pm\infty\}$).

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 $Z = \{y \mid \nabla^2\psi^*(y)$ does not exist$\}$. By absolute continuity of ρ' and Lemma 4.1, $\rho[\partial\psi^*(Z)] = \rho'[Z] = 0$. 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 4.3, the change of variables theorem (6) 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 4.2 shows that $\rho'(\nabla\psi(x)) = \rho(x)/\det[\nabla^2\psi(x)]$ at the Lebesgue points of ρ in M. Noting $A(0) = 0$, (31) follows immediately. ∎

Remark 4.5 (The Monge–Ampère equation). With restrictions on ρ', Brenier argued formally in [4, 5] that the convex function ψ for which $\nabla\psi_{\#}\rho = \rho'$ represents a generalized solution to the Monge–Ampère equation

$$\rho'(\nabla\psi(x)) \det[\nabla^2\psi(x)] = \rho(x). \tag{32}$$

A regularity theory for these solutions has been developed by Caffarelli in [6]. Without any assumptions, Proposition 4.2 and the first part of Theorem 4.4 show (32) to be satisfied almost everywhere on $\operatorname{dom}\psi$, provided $\nabla^2\psi(x)$ is interpreted in the sense of Aleksandrov.

APPENDIX: ON CONVEX FUNCTIONS AND THEIR DERIVATIVES

This appendix establishes some facts of life regarding convex functions and notation from convex analysis. Rockafellar's text [23] is a standard reference, while Schneider's [25, Notes to Section 1.5] summarizes Aleksandrov's theorem on second differentiability of convex functions.

By a *convex function* ψ on $\mathbf{R}^d$, we shall mean what is sometimes called a *proper* convex function: ψ takes values in $\mathbf{R} \cup \{+\infty\}$, is not identically $+\infty$, and $\psi((1-t)x + tx') \leqslant (1-t)\psi(x) + t\psi(x')$ when the latter is finite. If ψ is convex, its *domain* $\operatorname{dom}\psi := \{x \mid \psi(x) < \infty\}$ will be convex and ψ will be continuous on the interior Ω of $\operatorname{dom}\psi$. 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. in Ω. It is also useful to consider the *subdifferential* $\partial\psi$ of ψ: this parameterizes the supporting hyperplanes of ψ, and consists of pairs $(x, y) \in \mathbf{R}^d \times \mathbf{R}^d$ such that $\psi(z) \geqslant \langle y, z - x\rangle + \psi(x)$ for all $z \in \mathbf{R}^d$. Here $\langle\ ,\ \rangle$ denotes the usual inner product. $\partial\psi$ should be thought of as a multivalued mapping from $\mathbf{R}^d$ to $\mathbf{R}^d$: the image of a point x is denoted by $\partial\psi(x) := \{y \mid (x, y) \in \partial\psi\}$, and of a set X by $\partial\psi(X) := \bigcup_X \partial\psi(x)$. $\partial\psi(x)$ is a closed convex set, bounded precisely when $x \in \Omega$; it is empty for x outside $\operatorname{dom}\psi$, 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 $\mathbf{R}^d \times \mathbf{R}^d$ 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 $S \subset \mathbf{R}^d \times \mathbf{R}^d$ is said to be *cyclically monotone* if for any n points $(x_i, y_i) \in S$,

$$\langle y_1, x_2 - x_1 \rangle + \langle y_2, x_3 - x_2 \rangle + \cdots + \langle y_n, x_1 - x_n \rangle \leqslant 0. \tag{33}$$

The subdifferential of any convex function ψ will be cyclically monotone: if one linearly approximates the change in ψ around a cycle $x_1, x_2, ..., x_n, x_1$, a deficit must result since the approximation underestimates each step; the deficit will be finite, and the inequality in (33) strict, unless $y_i \in \partial\psi(x_{i+1})$ for each i. A theorem of Rockafellar [22] provides a converse: any cyclically monotone set is contained in the subdifferential 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 $(\partial\psi)^* := \{(y, x) \mid (x, y) \in \nabla^2\psi\}$, 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

$$\psi^*(y) := \sup_{x \in \mathbf{R}^d} \langle y, x \rangle - \psi(x). \tag{34}$$

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

A theorem of Aleksandrov [1] guarantees that a convex function ψ will be twice differentiable almost everywhere on its domain in the following sense: ψ is *twice differentiable* at x_0 with *Aleksandrov derivative* $\nabla^2\psi(x_0)$ if $\nabla\psi(x_0)$ exists, and if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|x - x_0| < \delta$ and $\Lambda = \nabla^2\psi(x_0)$ imply

$$\sup_{y \in \partial\psi(x)} |y - \nabla\psi(x_0) - \Lambda(x - x_0)| < \varepsilon\, |x - x_0|. \tag{35}$$

The Aleksandrov derivative $\nabla^2\psi(x_0)$ will be 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 $\mathbf{R}^d$ has an Aleksandrov derivative $\Lambda := \nabla^2\psi(x_0)$ at $x_0 \in \mathbf{R}^d$, so that $\nabla\psi(x_0)$ exists and $\psi < \infty$ in a neighbourhood of x_0. If Λ is* *invertible, then ψ^* has Λ^{-1} as its Aleksandrov derivative at $\nabla\psi(x_0)$; if Λ is not invertible then ψ^* fails to be twice differentiable at $\nabla\psi(x_0)$ in the sense of Aleksandrov.*

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 Legendre 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 $\varepsilon > 0$, taking t small enough in (35) implies $|\Lambda x| < \varepsilon\,|x|$. Because Λ is invertible, this forces $x = 0$. Thus $\nabla\psi^*(0) = 0$.

To show twice differentiability of ψ^* at 0, let $\varepsilon > 0$ 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 (35) to hold: $|y - \Lambda x| < \varepsilon\,|x|$. The inequality

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

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

Finally, the case Λ non-invertible must be addressed. Some $x \in \mathbf{R}^d$ is annihilated by Λ. From (35), there is a sequence $x_n \to 0$ of multiples of x and $(x_n, y_n) \in \partial\psi$ such that $|y_n| \leqslant n^{-1}\,|x_n|$. For any matrix Λ' and $\varepsilon > 0$, taking n large violates $|x_n - \Lambda' y_n| < \varepsilon\,|y_n|$. Thus ψ^* fails to be twice differentiable at 0. ∎

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 for monotone maps). *Assume ψ is convex on $\mathbf{R}^d$ and has an Aleksandrov derivative $\Lambda := \nabla^2\psi(x_0)$ at $x_0 \in \mathbf{R}^d$. If $B_r(x_0)$ is the ball of radius r centered at x_0, then as $r \to 0$,*

$$\frac{\operatorname{vol}[\partial\psi(B_r(x_0))]}{\operatorname{vol} B_r(x_0)} \to \det[\nabla^2\psi(x_0)]. \tag{36}$$

For Λ invertible, $\partial\psi(B_r(x_0))$ shrinks nicely to $\nabla\psi(x_0)$ in the sense of (28).

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 $\varepsilon > 0$, for $r < \delta$ from (35) it is immediate that

$$\partial\psi(B_r) \subset (1 + \varepsilon\,\|\Lambda^{-1}\|)\,\Lambda B_r. \tag{37}$$

On the other hand, ψ^* has Aleksandrov derivative 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 $\partial\psi^*(\Lambda B_r) \subset (1+\varepsilon \|\Lambda\|) B_r$. Taking r smaller if necessary, so that $(1+\varepsilon \|\Lambda\|)^{-1} \Lambda B_r$, lies in the interior of dom ψ^*, duality yields

$$(1+\varepsilon \|\Lambda\|)^{-1} \Lambda B_r \subset \partial\psi(B_r). \tag{38}$$

Since $\varepsilon > 0$ was arbitrary, (36) follows from (37–38) 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 (37–38) 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 $\mathbf{R}^d$. Given $\varepsilon > 0$, if $(x, y) \in \partial\psi$ for small enough $|x|$, then (35) implies that $|y - \Lambda x| < \varepsilon |x|$. Thus $\mathrm{vol}\, \partial\psi[B_r] \leqslant 2\varepsilon(\|\Lambda\| + \varepsilon)^{d-1} c r^d$, where c is the measure of the unit ball in $\mathbf{R}^{d-1}$. Since $\varepsilon > 0$ was arbitrary, the limit (36) vanishes. ∎

Finally, our results from [16], which extend work of Brenier [4, 5], are summarized by:

THEOREM A.3 (Monotone measure-preserving maps [16]). *Given $\rho, \rho' \in \mathscr{P}(\mathbf{R}^d)$, there exists $\gamma \in \mathscr{P}(\mathbf{R}^d \times \mathbf{R}^d)$ with cyclically monotone support having ρ and ρ' as its marginals. If ρ vanishes on sets of Hausdorff dimension $d-1$, then a convex function ψ exists on $\mathbf{R}^d$ whose gradient $\nabla\psi$ pushes ρ forward to ρ'; in fact, $\gamma = (id \times \nabla\psi)_{\#}\, \rho$. Although ψ may not be unique, the map $\nabla\psi$ is uniquely determined ρ-a.e.*

ACKNOWLEDGMENTS

It is the author's pleasure to express his deep gratitude to his advisor and mentor, Elliott Lieb, as well as to the many others who contributed their suggestions and support: Michael Aizenman, Frederick Almgren, Eric Carlen, Charles Fefferman, Michael Loss, Jan-Philip Solovej, and John Stalker. 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.

REFERENCES

1. A. D. Aleksandrov, Almost everywhere existence of the second differential of a convex function and some properties of convex surfaces connected with it [in Russian], *Uchen. Zap. Leningrad Gos. Univ. Math. Ser.* **6** (1939), 3–35.
2. J. F. G. Auchmuty and R. Beals, Variational solutions of some non-linear free boundary problems, *Arch. Rational Mech. Anal.* **43** (1971), 255–271.

3. H. J. Brascamp and E. H. Lieb, On extensions of the Brunn–Minkowski and Prékopa–Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation, *J. Funct. Anal.* **22** (1976), 366–389.
4. Y. Brenier, Décomposition polaire et réarrangement monotone des champs de vecteurs, *C. R. Acad. Sci. Paris Sér. I Math.* **305** (1987), 805–808.
5. Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions, *Comm. Pure Appl. Math.* **44** (1991), 375–417.
6. L. A. Caffarelli, The regularity of mappings with a convex potential, *J. Amer. Math. Soc.* **5** (1992), 99–104.
7. D. C. Dowson and B. V. Landau, The Fréchet distance between multivariate normal distributions, *J. Multivariate Anal.* **12** (1982), 450–455.
8. H. Federer, "Geometric Measure Theory," Springer-Verlag, New York, 1969.
9. C. R. Givens and R. M. Shortt, A class of Wasserstein metrics for probability distributions, *Michigan Math. J.* **31** (1984), 231–240.
10. H. Hadwiger and D. Ohmann, Brunn–Minkowskischer Satz und Isoperimetrie, *Math. Z.* **66** (1956), 1–8.
11. M. Knott and C. S. Smith, On the optimal mappings of distributions, *J. Optim. Theory Appl.* **43** (1984), 39–49.
12. L. Leindler, On a certain converse of Hölder's inequality II, *Acta Sci. Math. (Szeged)* **33** (1972), 217–223.
13. E. H. Lieb, Thomas–Fermi and related theories of atoms and molecules, *Rev. Modern Phys.* **53** (1981), 603–641.
14. E. H. Lieb and H.-T. Yau, The Chandrasekhar theory of stellar collapse as the limit of quantum mechanics, *Comm. Math. Phys.* **112** (1987), 147–174.
15. 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** (1984), 109–145.
16. R. J. McCann, Existence and uniqueness of monotone measure-preserving maps, *Duke Math. J.* **80** (1995), 309–323.
17. R. J. McCann, "A Convexity Theory for Interacting Gases and Equilibrium Crystals," Ph.D. thesis, Princeton University, 1994.
18. I. Olkin and F. Pukelsheim, The distance between two random vectors with given dispersion matrices, *Linear Algebra Appl.* **48** (1982), 257–263.
19. A. Prékopa, Logarithmic concave measures with application to stochastic programming, *Acta Sci. Math. (Szeged)* **32** (1971), 301–315.
20. A. Prékopa, On logarithmic concave measures and functions, *Acta Sci. Math. (Szeged)* **34** (1973), 335–343.
21. F. Riesz, Sur une inégalité intégrale, *J. London Math. Soc.* **5** (1930), 162–168.
22. R. T. Rockafellar, Characterization of the subdifferentials of convex functions, *Pacific J. Math.* **17** (1966), 497–510.
23. R. T. Rockafellar, "Convex Analysis," Princeton Univ. Press, Princeton, NJ, 1972.
24. W. Rudin, "Real and Complex Analysis," McGraw–Hill, New York, 1987.
25. R. Schneider, "Convex Bodies: The Brunn–Minkowski Theory," Cambridge Univ. Press, Cambridge, UK, 1993.