Fine Properties of Geodesics and Geodesic λ-Convexity for the Hellinger–Kantorovich Distance

Matthias Liero⬤, Alexander Mielke⬤ & Giuseppe Savaré⬤

Communicated by A. Figalli

Abstract

We study the fine regularity properties of optimal potentials for the dual formulation of the Hellinger–Kantorovich problem (HK), providing sufficient conditions for the solvability of the primal Monge formulation. We also establish new regularity properties for the solution of the Hamilton–Jacobi equation arising in the dual dynamic formulation of HK, which are sufficiently strong to construct a characteristic transport-growth flow driving the geodesic interpolation between two arbitrary positive measures. These results are applied to study relevant geometric properties of HK geodesics and to derive the convex behaviour of their Lebesgue density along the transport flow. Finally, exact conditions for functionals defined on the space of measures are derived that guarantee the geodesic λ-convexity with respect to the Hellinger–Kantorovich distance. Examples of geodesically convex functionals are provided.

Contents

1. Introduction

In [26,27] the Hellinger–Kantorovich distance (in [10,11,21] it is also called Wasserstein–Fisher–Rao distance or Kantorovich–Fisher–Rao distance in [17]) was introduced to describe the interaction between optimal transport and optimal creation and destruction of mass in a convex domain of $\mathbb{R}^d$. Here we further investigate the structure of (minimal) geodesics, and we fully analyze the question of geodesic λ-convexity of integral functionals with respect to this distance.

The Hellinger–Kantorovich distance can be considered as a combination, more precisely the inf-convolution, of the Hellinger–Kakutani distance on the set of all measures (cf. e.g. [33]) and the L^2 Kantorovich–Wasserstein distance, which is well-known from the theory of optimal transport, see e.g. [2,34]. Throughout this text, we denote by $\mathcal{M}(\mathbb{R}^d)$ all nonnegative and finite Borel measures endowed with the weak topology induced by the canonical duality with the continuous functions $C_0(\mathbb{R}^d)$ decaying at infinity. While the L^2 Kantorovich–Wasserstein distance $W(\mu_0, \mu_1)$ of measures $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^d)$ requires μ_0 and μ_1 to have the same mass to be finite, the Hellinger–Kakutani distance, which is defined via

$$H(\mu_0, \mu_1)^2 = \int_{\mathbb{R}^d} \left(\sqrt{\theta_0} - \sqrt{\theta_1}\right)^2 \mathrm{d}(\mu_0 + \mu_1), \text{ where } \theta_j = \frac{\mathrm{d}\mu_j}{\mathrm{d}(\mu_0 + \mu_1)},$$

has the upper bound $H(\mu_0, \mu_1) \leqq \mu_0(\mathbb{R}^d) + \mu_1(\mathbb{R}^d)$, with equality if μ_0 and μ_1 are mutually singular.

As a generalization of the dynamical formulation of the Kantorovich–Wasserstein distance (see [6]), the Hellinger–Kantorovich distance $HK_{\alpha,\beta}$ can be defined in a dynamic way via

$$HK_{\alpha,\beta}(\mu_0, \mu_1)^2 = \inf \left\{ \int_{t=0}^1 \int_{\mathbb{R}^d} \left(\alpha |\Upsilon(t,x)|^2 + \beta \xi(t,x)^2\right) \mathrm{d}\mu_t(x)\,\mathrm{d}t \right.$$
$$\left. \bigg| \mu \in C([0,1]; \mathcal{M}(\mathbb{R}^d)), \ \mu_{t=0} = \mu_0, \ \mu_{t=1} = \mu_1, \ \text{(gCE) holds} \right\},$$
$$(1.1)$$

where $\Upsilon : (0,1) \times \mathbb{R}^d \to \mathbb{R}^d$ and $\xi : (0,1) \times \mathbb{R}^d \to \mathbb{R}$ are Borel maps characterizing the generalized continuity equation

$$\text{(gCE)} \qquad \frac{\partial}{\partial t}\mu + \alpha \operatorname{div}(\mu \Upsilon) = \beta \xi \mu,$$

formulated in a distributional sense. The parameters $\alpha > 0$ and $\beta > 0$ allow us to control the relative strength of the Kantorovich–Wasserstein part and the Hellinger–Kakutani part, i.e. $\mathsf{HK}_{\alpha,\beta}$ is the inf-convolution of $\mathsf{HK}_{\alpha,0} = \frac{1}{\sqrt{\alpha}}\mathsf{W}$ and $\mathsf{HK}_{0,\beta} = \frac{1}{\sqrt{\beta}}\mathsf{H}$, see [27, Rem. 8.19]. Subsequently, we will restrict to the standard case $\alpha = 1$ and $\beta = 4$, since the general case can easily be obtained by scaling the underlying space $\mathbb{R}^d$. We will shortly write HK instead of $\mathsf{HK}_{1,4}$.

It is a remarkable fact, deeply investigated in [27], that the HK distance has many interesting equivalent characterizations, which highlight its geometric and variational character. A first one arises from the dual dynamic counterpart of (1.1) in terms of subsolutions of a suitable Hamilton–Jacobi equation:

$$\frac{1}{2}\mathsf{HK}^2(\mu_0, \mu_1) = \sup \left\{ \int_{\mathbb{R}^d} \xi(\tau, \cdot)\,\mathrm{d}\mu_1 - \int_{\mathbb{R}^d} \xi(0, \cdot)\,\mathrm{d}\mu_0 \,\Big|\, \xi \in C_c^\infty([0,1] \times \mathbb{R}^d), \right.$$
$$\left. \frac{\partial}{\partial t}\xi + \frac{1}{2}|\nabla\xi|^2 + 2\xi^2 \leqq 0 \quad \text{in } [0,1] \times \mathbb{R}^d \right\}. \tag{1.2}$$

By expressing solutions of (1.2) in terms of a new formula of Hopf–Lax type, one can write a static duality representation

$$\mathsf{HK}^2(\mu_0, \mu_\tau) = \sup \left\{ \int_{\mathbb{R}^d} (1 - e^{-2\varphi_\tau})\,\mathrm{d}\mu_\tau - \int_{\mathbb{R}^d} (e^{2\varphi_0} - 1)\,\mathrm{d}\mu_0 \,\Big| \right.$$
$$\left. \varphi_0, \varphi_\tau \in C_b(\mathbb{R}^d), \ \varphi_\tau(x_\tau) - \varphi_0(x_0) \leqq \mathsf{L}_1(x_\tau - x_0) \right\} \tag{1.3}$$

associated with the convex cost function $\mathsf{L}_1(z) := \frac{1}{2}\log(1 + \tan^2(|z|))$ which forces $|z| < \pi/2$. Notice that it is possible to write (1.3) in a symmetric form with respect to φ_0, φ_1 just by changing the sign of φ_1.

It is remarkable that (1.3) can be interpreted as the dual problem of the static Logarithmic Entropy Transport (LET) variational formulation of HK. By introducing the logarithmic entropy density $F : [0, \infty[\to [0, \infty[$ via

$$F(s) := s\log s - s + 1 \quad \text{for } s > 0 \quad \text{and} \quad F(0) := 1, \tag{1.4}$$

we get

$$\mathsf{HK}^2(\mu_0, \mu_\tau) = \min \left\{ \int_{\mathbb{R}^d} F(\sigma_0)\,\mathrm{d}\mu_0 + \int_{\mathbb{R}^d} F(\sigma_1)\,\mathrm{d}\mu_1 + \iint_{\mathbb{R}^d \times \mathbb{R}^d} 2\mathsf{L}_1(x_0 - x_1)\,\mathrm{d}\boldsymbol{\eta} \right\} \tag{1.5}$$

where the minimum is taken over all positive finite Borel measures $\boldsymbol{\eta}$ in $\mathbb{R}^d \times \mathbb{R}^d$ whose marginals $(\pi_i)_\sharp \boldsymbol{\eta} = \sigma_i \mu_i$ are absolutely continuous with respect to μ_i.

The subdifferential

$$\mathsf{DL}_1(z) = \partial\mathsf{L}_1(z) = \mathbf{tan}(z) := \tan(|z|)\frac{z}{|z|}$$

and its inverse $w \mapsto \mathbf{arctan}(w)$ will play an important role. We continue to use bold function names for vector-valued functions constructed from real-valued ones as follows:

for a map $f : \mathbb{R} \to \mathbb{R}$ with $f(0) = 0$ we set $\boldsymbol{f} : \mathbb{R}^d \to \mathbb{R}^d$ via $\boldsymbol{f}(x) := f(|x|)\frac{x}{|x|}$.

$$\tag{1.6}$$

A fourth crucial formula, which we will extensively study in the present paper, is related to the primal Monge formulation of Optimal Transport, and clarifies the two main components of HK arising from transport and growth or decay effects. Its main ingredient is the notion of transport-growth pair $(\boldsymbol{T}, q) : \mathbb{R}^d \to \mathbb{R}^d \times [0, \infty)$ acting on measures $\mu \in \mathcal{M}(\mathbb{R}^d)$ as

$$(\boldsymbol{T}, q)_\star \mu := \boldsymbol{T}_\sharp (q^2 \cdot \mu),$$

$$((\boldsymbol{T}, q)_\star \mu)(A) := \int_{\boldsymbol{T}^{-1}(A)} q^2 \, \mathrm{d}\mu \quad \text{for every Borel set } A \subset \mathbb{R}^d. \quad (1.7)$$

The Monge formulation of HK then looks for the optimal pair $(\boldsymbol{T}, q)$ among the ones transforming μ_0 into μ_1 by $(\boldsymbol{T}, q)_\star \mu_0 = \mu_1$ which minimizes the conical cost

$$\mathcal{C}(\boldsymbol{T}, q; \mu_0) := \int_{\mathbb{R}^d} \left(1 + q^2(x) - 2q(x) \cos_{\pi/2} \left(|\boldsymbol{T}(x) - x| \right) \right) \mathrm{d}\mu_0(x), \quad (1.8)$$

where $\cos_{\pi/2}(r) := \cos \left(\min\{r, \pi/2\} \right)$. As for the usual Monge formulation of optimal transport, the existence of an optimal transport-growth pair $(\boldsymbol{T}, q)$ minimizing (1.8) requires more restrictive properties on μ_0, μ_1 which we will carefully study. It is worth noticing that the integrand in (1.8) has a relevant geometric interpretation as the square distance $\mathsf{d}_{\pi, \mathfrak{C}}^2$, where $\mathsf{d}_{\pi, \mathfrak{C}}$ is the distance on the cone space $\mathfrak{C}$ over $\mathbb{R}^d$ (cf. (2.5)) between the points $[x, 1]$ and $[\boldsymbol{T}(x), q(x)]$ and suggests that HK induces a distance in $\mathcal{M}(\mathbb{R}^d)$ which plays a similar role than the L^2 Kantorovich–Rubinstein–Wasserstein distance in $\mathcal{P}_2(\mathbb{R}^d)$. The dynamic formulation (1.1), moreover, suggests that its minimizers $(\mu_t)_{t \in [0,1]}$ should provide minimal geodesics in $(\mathcal{M}(\mathbb{R}^d), \mathsf{HK})$ which behave like transport-growth interpolations between μ_0 and μ_1.

Inspired by the celebrated paper [28], we want to study the structure of such minimizers and to characterize integral functionals which are convex along such kind of interpolations.

1.1. Improved Regularity of Potentials and Geodesics

In the first part of the paper we will exploit the equivalent formulations of HK in order to obtain new information on the regularity and on the fine structure of the solutions to (1.3), (1.2), and (1.8).

More precisely, we will initially prove in Section 3 that the optimal HK potential φ_0 is locally semi-convex outside a closed $(d-1)$-rectifiable set, so that when $\mu_0 \ll \mathcal{L}^d$ and μ_1 is concentrated in a neighborhood of $\mathrm{supp}(\mu_0)$ of radius $\pi/2$ the Monge formulation (1.8) has a unique solution.

After the transformation $\xi_0 := \frac{1}{2}(\mathrm{e}^{\varphi_0} - 1)$ (which linearizes the second integrand in the duality formula (1.3)), we also obtain a family of maps, for $t \in [0, 1]$,

$$\boldsymbol{T}_{0 \to t}(x) = x + \mathbf{arctan}\left(\frac{t \nabla \xi_0}{1 + 2t \xi_0(x)} \right), \quad q_{0 \to t}^2(x) := (1 + 2t \xi_0(x))^2 + t^2 |\nabla \xi_0(x)|^2,$$

$$(1.9)$$

with the following properties:

1. $(T_{0\to1}, q_{0\to1})$ is the unique solution of (1.8) and provides the beautiful formula

$$\mathsf{HK}^2(\mu_0, \mu_1) = \int_{\mathbb{R}^d} \left(4\xi_0^2 + |\nabla\xi_0|^2 \right) d\mu_0, \tag{1.10}$$

showing that the (closure of the) space of $C_c^1(\mathbb{R}^d)$ functions with respect to the Hilbertian norm

$$\|\xi\|_{H^{1,2}(\mathbb{R}^d,\mu)}^2 = \int_{\mathbb{R}^d} \left(4\xi^2 + |\nabla\xi|^2 \right) d\mu \tag{1.11}$$

provides the natural notion of tangent space $\mathrm{Tan}_\mu \mathcal{M}(\mathbb{R}^d)$ and a nonsmooth Riemannian formalism in $(\mathcal{M}(\mathbb{R}^d), \mathsf{HK})$ as for the Otto calculus in $(\mathcal{P}_2(\mathbb{R}^d), W_2)$.

2. The curve $\mu_t = (T_{0\to t}, q_{0\to t})_\star \mu_0$ is an explicit characterization of the geodesic interpolation solving (1.1). A crucial fact is that for μ_0-a.e. x the curve $[T_{0\to t}(x), q_{0\to t}(x)]$ is a geodesic in the cone space $\mathfrak{C}$ interpolating the points $[x, 1]$ and $[T_{0\to1}(x), q_{0\to1}(x)]$.

It is then natural to investigate if the potential ξ_0 can be used to build an optimal solution ξ_t of (1.2), which should at least formally solve the Hamilton-Jacobi equation

$$\partial_t \xi_t + \frac{1}{2}|\nabla\xi_t|^2 + 2\xi_t^2 = 0 \quad \text{on the support of } \mu \text{ in } (0, 1) \times \mathbb{R}^d. \tag{1.12}$$

This problem will be investigated in Section 4, by a detailed analysis of the regularity of the forward solutions to (1.2) provided by the generalized Hopf–Lax formula (see (4.2))

$$\xi_t(x) = \xi(t, x) = (\mathscr{P}_t\xi_0)(x) = \frac{1}{t} \; \mathscr{A}\left(t\xi_0(\cdot)\right)(x) = \inf_{y\in\mathbb{R}^d} \frac{1}{2t}\left(1 - \frac{\cos_{\pi/2}^2(|x-y|)}{1+2t\xi_0(y)}\right). \tag{1.13}$$

It is well known that one cannot expect smoothness of such a solution; however, the particular structure of transport duality suggests that the final value ξ_1 given by (1.13) corresponds to the optimal potential φ_1 of the dual formulation (1.3) via the transformation $\xi_1 = \frac{1}{2}(1 - e^{-2\varphi_1})$, so that the initial and final optimal potentials ξ_0 and ξ_1 are simultaneously linked by the forward-backward relation

$$\xi_1 = \mathscr{P}_1\xi_0, \quad \xi_0 = \mathscr{R}_1(\xi_1) \quad \text{where } \mathscr{R}_t(\eta) := -\mathscr{P}_t(-\eta) \text{ is the backward flow.} \tag{1.14}$$

Following the approach of [34, Cha. 7] (see also [27, Sec. 8]) and using the reversibility in time of geodesics, we can add to the family of forward potentials ξ_t given by (1.13) the crucial information provided by the backward solutions $\bar{\xi}_t$ starting from ξ_1:

$$\bar{\xi}_t := \mathscr{R}_{1-t}\xi_1 = -\mathscr{P}_{1-t}\left(-\xi_1\right) \quad \text{for } t \in [0, 1]. \tag{1.15}$$

In general, ξ_t and $\bar{\xi}_t$ do not coincide for $t \in (0, 1)$ but still satisfy

$$\xi_t(x) \geqq \bar{\xi}_t(x) \quad \text{in } (0, 1) \times \mathbb{R}^d, \quad \xi_0 = \bar{\xi}_0, \quad \xi_1 = \bar{\xi}_1. \tag{1.16}$$

The crucial fact arising from the optimality condition (1.14), and the geometric property of the geodesic $(\mu_t)_{t\in[0,1]}$ is that for every $t \in [0, 1]$

the support of μ_t is contained in the *contact set* $\Xi_t := \left\{ x \in \mathbb{R}^d \,\big|\, \xi_t(x) = \bar{\xi}_t(x) \right\}$.

On the contact set $(\Xi_t)_{t\in[0,1]}$, we can combine the (delicate) first- and second-order super-differentiability properties of ξ_t arising from the inf-convolution structure of (1.13) with the corresponding sub-differentiability properties exhibited by $\bar{\xi}_t$.

Using tools from nonsmooth analysis, we are then able to give a rigorous meaning to the characteristic flow associated with (1.12), i.e. to the maps $t \mapsto \boldsymbol{T}(t, \cdot) = \boldsymbol{T}_{s\to t}(\cdot), t \mapsto q(t, \cdot) = q_{s\to t}(\cdot)$ solving (we omit to write the explicit dependence on x when not needed)

$$\begin{cases} \dot{\boldsymbol{T}}(t) = \nabla\xi_t(\boldsymbol{T}(t)), \\ \dot{q}(t) = 2\xi_t(\boldsymbol{T}(t))q(t), \end{cases} \quad \text{in } (0, 1), \quad \boldsymbol{T}(s, x) = x, \ q(s, x) = 1. \quad (1.17)$$

Moreover, we will prove that $\boldsymbol{T}_{s\to t}$ is a family of bi-Lipschitz maps on the contact sets obeying a natural concatenation property. As can be expected, the maps $\boldsymbol{T}_{s\to t}, q_{s\to t}$ provide a precise representation of the geodesics via $\mu_t = (\boldsymbol{T}_{s\to t}, q_{s\to t})_\star\mu_s$ for all $s, t \in (0, 1)$. In particular $(\boldsymbol{T}_{s\to t}, q_{s\to t})$ is an optimal transport-growth pair between μ_s and μ_t minimizing the cost of (1.8).

Using this valuable information, in Section 5 we obtain various relevant structural properties of geodesics in $(\mathcal{M}(\mathbb{R}^d), \mathsf{HK})$ such as *non-branching, localization, and regularization effects*. In particular, independently of the regularity of μ_0 and μ_1, we will show that for $s \in (0, 1)$ the Monge problem between μ_s and μ_0 or between μ_s and μ_1 always admit a unique solution, a property which is well known in the Kantorovich–Wasserstein framework.

Surprisingly enough, despite the lack of global regularity, we will also establish precise formulae for the first and second derivative of the differential of $\boldsymbol{T}_{s\to t}$ (and thus the second order differential of ξ_t) along the flow, which coincides with the equations that one obtains by formally differentiation using the joint information of the Hamilton–Jacobi equation (1.12) and (1.17) assuming sufficient regularity. For instance, differentiating in time the first equation of (1.17) and differentiating in space (1.12), one finds that

$$\ddot{\boldsymbol{T}}(t) = \partial_t\nabla\xi_t(\boldsymbol{T}(t)) + \mathrm{D}^2\xi_t\nabla\xi_t(\boldsymbol{T}(t)), \quad \partial_t\nabla\xi_t = -\mathrm{D}^2\xi_t\nabla\xi_t + 4\xi_t\nabla\xi_t,$$

which yield

$$\ddot{\boldsymbol{T}}(t) = 4\xi_t(\boldsymbol{T}(t))\nabla\xi_t(\boldsymbol{T}(t)). \quad (1.18a)$$

For $q(t)$, $\mathsf{B}(t) := \mathrm{D}\boldsymbol{T}_{s\to t}$, and its determinant $\delta(t) := \det \mathsf{B}(t)$ similar, just more involved, calculations yield the crucial second order equations

$$\ddot{q}(t) = |\nabla\xi_t(\boldsymbol{T}(t))|^2 q(t), \quad (1.18b)$$

$$\ddot{\mathsf{B}}(t) = -4\Big(\nabla\xi_t \otimes \nabla\xi_t + \xi_t\mathrm{D}^2\xi_t\Big) \circ \boldsymbol{T}(t) \cdot \mathsf{B}(t), \quad (1.18c)$$

$$\ddot{\delta}(t) = \Big((\Delta\xi_t)^2 - |\mathrm{D}^2\xi_t|^2 - 4|\nabla\xi_t|^2 - 4\xi_t\Delta\xi_t\Big) \circ \boldsymbol{T}(t) \cdot \delta(t). \quad (1.18d)$$

In our case, even though we do not have enough regularity to justify the above formal computations, we can still derive them rigorously by a deeper analysis using the variational properties of the contact set. Even if our discussion is restricted to the Hellinger–Kantorovich case and uses the particular form of the Hopf–Lax semigroup (1.13) and its characteristics (1.9), we think that our argument applies to more general cases and may provide new interesting estimates also in the typical balanced case of Optimal Transport.

Such regularity and the related second order estimates are sufficient to express the Lebesgue density c_t of the measures μ_t and thus to obtain crucial information on its behavior along the flow. In particular, Corollary 5.5 shows that $c(t, \cdot)$ is given by

$$c(t, y)\big|_{y=T_{s\to t}(x)} = c(s, x)\frac{\alpha_s(t, x)}{\delta_s(t, x)} \quad \text{with} \tag{1.19a}$$

$$\alpha_s(t, x) = (1+2(t-s)\xi_s(x))^2 + (t-s)^2|\nabla\xi_s(x)|^2 = q_{s\to t}(x) \tag{1.19b}$$

$$\delta_s(t, x) := \det(DT_{s\to t}(x)), \tag{1.19c}$$

and the time-dependent transport-growth mapping $T_{s\to t}, q_{s\to t}$ are given in terms of ξ via (1.17) and the analog of (1.9). In particular, we will show that if $\mu_s \ll \mathcal{L}^d$ for some $s \in (0, 1)$ then $\mu_t \ll \mathcal{L}^d$ for every $t \in (0, 1)$ and combining (1.18b), (1.18c), and (1.19a) we will also prove that c_t is a convex function along the flow maps $T_{s\to t}$.

1.2. Geodesic λ-Convexity of Functionals

The second part of the paper is devoted to establish necessary and sufficient conditions for geodesic λ-convexity of energy functionals $\mathscr{E}$ defined for a closed and convex domain $\Omega \subset \mathbb{R}^d$ with non-empty interior in the form

$$\mathscr{E}(\mu) = \int_\Omega E(c(x))\,dx + E'_\infty\mu^\perp(\Omega) \quad \text{for} \quad \mu = c\mathcal{L}^d + \mu^\perp \text{ with } \mu^\perp \perp \mathcal{L}^d, \tag{1.20}$$

where $E'_\infty := \lim_{c\to\infty} E(c)/c \in \mathbb{R}\cup\{+\infty\}$ is the recession constant and $E(0) = 0$ holds.

In [26, Prop. 19] it was shown that the total-mass functional $\mathscr{M} : \mu \mapsto \mu(\mathbb{R}^d)$ has the surprising property that it is exactly quadratic along HK geodesics $\gamma : [0, 1] \to \mathcal{M}(\mathbb{R}^d)$, namely

$$\mathscr{M}(\gamma(t)) = (1-t)\mathscr{M}(\gamma(0)) + t\mathscr{M}(\gamma(1)) - t(1-t)\mathsf{HK}(\gamma(0), \gamma(1))^2 \text{ for } t \in [0, 1]. \tag{1.21}$$

Thus, as a first observation we see that a density function E generates a geodesically λ-convex functional $\mathscr{E}$ if and only if $E_0 : c \mapsto E(c) - \lambda c$ generates a geodesically convex functional (i.e. geodesically 0-convex). Hence, subsequently we can restrict to $\lambda = 0$.

To explain the necessary and sufficient conditions on E for $\mathscr{E}$ to be geodesically convex, we first look at the differentiable case, and we define the shorthand notation

$$\varepsilon_0(c) = E(c), \quad \varepsilon_1(c) = cE'(c), \quad \varepsilon_2(c) = c^2 E''(c).$$

For the Kantorovich–Wasserstein distance W the necessary and sufficient conditions are the so-called McCann conditions [28]:

$$\varepsilon_2(c) \;\geqq\; \frac{d-1}{d}\big(\varepsilon_1(c) - \varepsilon_0(c)\big) \geqq 0 \text{ for all } c > 0$$
$$\Longleftrightarrow \begin{cases} r \mapsto r^d E(r^{-d}) \text{ is lower semi-continuous and convex and} \\ r \mapsto (d-1)r^d E(r^{-d}) \text{ is non-increasing on }]0, \infty[, \end{cases} \quad (1.22)$$

(see also [2, Prop. 9.3.9]). For the Hellinger–Kakutani distance we simply need the condition

$$2\varepsilon_2(c) + \varepsilon_1(c) \geqq 0 \quad \Longleftrightarrow \quad \Big(r \mapsto E(r^2) \text{ is convex}\Big). \quad (1.23)$$

In the case of differentiable E, our main result yields the following necessary and sufficient conditions for geodesic convexity of $\mathscr{E}$ on $(\mathcal{M}(\mathbb{R}^d), \mathsf{HK})$, see Proposition 6.1,

$$(d-1)\big(\varepsilon_1(c) - \varepsilon_0(c)\big) \geqq 0 \quad \text{and} \quad \mathbb{B}(c) \geqq 0 \quad \text{for all } c > 0, \quad (1.24)$$

where the matrix $\mathbb{B}(c) \in \mathbb{R}^{2 \times 2}_{\text{sym}}$ is given by

$$\mathbb{B}(c) := \begin{pmatrix} \varepsilon_2(c) - \frac{d-1}{d}\big(\varepsilon_1(c)-\varepsilon_0(c)\big) & \varepsilon_2(c) - \frac{1}{2}\big(\varepsilon_1(c)-\varepsilon_0(c)\big) \\ \varepsilon_2(c) - \frac{1}{2}\big(\varepsilon_1(c)-\varepsilon_0(c)\big) & \varepsilon_2(c) + \frac{1}{2}\varepsilon_1(c) \end{pmatrix}.$$

We immediately see that the non-negativity of the diagonal element $\mathbb{B}_{11}(c)$ gives the first McCann condition in (1.22), and $\mathbb{B}_{22}(c) \geqq 0$ gives (1.23). However, the condition $\mathbb{B}(c) \geqq 0$ is strictly stronger, since e.g. it implies that the additional condition $(d+2)\varepsilon_1(c) - 2\varepsilon_0(c) \geqq 0$ holds, see (6.2). This condition means that $c \mapsto c^{-2/(d+2)} E(c)$ has to be non-decreasing, which will be an important building block for the main geodesic convexity result.

Indeed, our main result in Theorem 7.2 is formulated for general lower semi-continuous and convex functions $E : [0, \infty[\to \mathbb{R} \cup \{\infty\}$ without differentiability assumptions. The conditions on E can be formulated most conveniently in terms of the auxiliary function $N_E :]0, \infty[^2 \to \mathbb{R} \cup \{\infty\}$ defined via

$$N_E(\rho, \gamma) = \Big(\frac{\rho}{\gamma}\Big)^d E\Big(\frac{\gamma^{2+d}}{\rho^d}\Big). \quad (1.25a)$$

Then, $\mathscr{E}$ defined in (1.20) is geodesically convex if and only if N_E satisfies

$$N_E :]0, \infty[^2 \to \mathbb{R} \cup \{\infty\} \text{ is convex, and} \quad (1.25b)$$
$$\rho \mapsto (d-1)N_E(\rho, \gamma) \text{ is non-increasing.} \quad (1.25c)$$

The McCann conditions (1.22) are obtained by looking at $N_E(\cdot, \gamma)$ for fixed γ, while the Hellinger–Kakutani condition (1.23) follows by looking at $s \mapsto N_E(s\rho, s\gamma)$ for fixed (ρ, γ).

The proof of the sufficiency and necessity of condition (1.25) for geodesic convexity of $\mathscr{E}$ is based on the explicit representation (1.19) of the geodesic curves giving

$$\mathscr{E}(\mu(t)) = \int_\Omega E(c(t,y))\,\mathrm{d}y = \int_\Omega e(t,x)\,\mathrm{d}x$$
$$\text{where } e(t,x) := \delta_s(t,x)\,E\Big(c_s(x)\frac{\alpha_s(t,x)}{\delta_s(t,x)}\Big).$$

By definition, we have $\alpha_s(t,x) \geqq 0$, and Corollary 5.5 guarantees $\delta_s(t,x) > 0$. Hence, we can introduce the two functions

$$\gamma(t,x) = \big(c_s(x)\alpha_s(t,x)\big)^{1/2} \quad \text{and} \quad \rho(t,x) = \big(c_s(x)\alpha_s(t,x)\big)^{1/2}\delta_s(t,x)^{1/d},$$

which connect the densities $e(t,x)$ with the function N_E defined in (1.25a) in the form

$$e = \delta\, E\big(c\,\frac{\alpha}{\delta}\big) = N_E(\rho,\gamma).$$

For smooth E we have smooth N_E and may show convexity of $t \mapsto e(t,x)$ via

$$\partial_t^2 e(t,x) =: \ddot{e} = \Big\langle \mathrm{D}^2 N_E(\rho,\gamma)\begin{pmatrix}\dot{\rho}\\\dot{\gamma}\end{pmatrix}, \begin{pmatrix}\dot{\rho}\\\dot{\gamma}\end{pmatrix}\Big\rangle + \Big\langle \mathrm{D}N_E(\rho,\gamma), \begin{pmatrix}\ddot{\rho}\\\ddot{\gamma}\end{pmatrix}\Big\rangle \geqq 0.$$

By convexity of N_E, the term involving $\mathrm{D}^2 N_E$ is non-negative, so it remains to show

$$\partial_\rho N_E(\rho,\gamma)\ddot{\rho} + \partial_\gamma N_E(\rho,\gamma)\ddot{\gamma} \geqq 0. \tag{1.26}$$

To establish this, we use first that the scaling property $N_E(s^{1+d/2}\rho, s\gamma) = s^2 N_E(\rho,\gamma)$ for all $s > 0$ (which follows from the definition of N_E via E) and the convexity of N_E imply

$$(1-4/d^2)\,\rho\,\partial_\rho N_E(\rho,\gamma) + \gamma\,\partial_\gamma N_E(\rho,\gamma) \geqq 0, \tag{1.27}$$

see Proposition 6.2. Second, we rely on a nontrivial curvature estimate for (ρ,γ), namely

$$\frac{\ddot{\gamma}(t,x)}{\gamma(t,x)} \geqq 0 \quad \text{and} \quad \frac{\ddot{\rho}(t,x)}{\rho(t,x)} \leqq \Big(1 - \frac{4}{d}\Big)\frac{\ddot{\gamma}(t,x)}{\gamma(t,x)}. \tag{1.28}$$

Estimates (1.28) are provided in Proposition 5.7 and strongly rely on the explicit representation and the regularity properties of the geodesics developed in Sects. 4 and 5.

Combining (1.28) with $\partial_\rho N_E(\rho,\gamma) \leqq 0$, the desired relation (1.26) easily follows, see Section 7. Finally, a simple integration over $\mathbb{R}^d$ provides the convexity of $t \mapsto \mathscr{E}(\mu(t))$. Note that we have indeed the larger factor $(1-4/d^2)$ in (1.27) while the curvature estimate in (1.28) has the smaller and hence "better" factor $(1-4/d)$.

As a consequence, we find that the power functionals $\mathscr{E}_m$ with $E_m(c) = c^m$ with $m > 1$ are all geodesically convex, see Corollary 7.3. This result was already exploited in [13, Thm. 2.14]. We can study the discontinuous "Hele–Shaw case" $E(c) = -\lambda c$ for $c \in [0, 1]$ and $E(c) = \infty$ for $c > 1$. Moreover, in dimensions $d = 1$ or 2 the densities $E_q(c) = -c^q$ with $q \in [\frac{d}{d+2}, \frac{1}{2}]$ also lead to geodesically convex functionals $\mathscr{E}_q$, see again Corollary 7.3.

Two important differences with the balanced Kantorovich–Wasserstein case are worth noting. First, the Boltzmann logarithmic entropy functional corresponding to $E(c) = c \log c$ is not geodesically λ-convex for any value of λ, see Example 6.5. Second, if the space dimension d is larger than or equal to 3, then there are no geodesically convex power functionals of the form $E(x) = -c^m$ with exponent $m < 1$, see Example 6.4. Some of these statements follow easily by observing that $\mu_t = t^2 \mu_1$ is the unique geodesic connecting $\mu_0 = 0$ and μ_1.

1.3. Applications and Outlook

In [15,23], the JKO scheme (minimizing movement scheme) for a gradient system $(\mathcal{M}(\Omega), \mathsf{HK}_{\alpha,\beta}, \mathscr{E})$ is considered, i.e., for $\tau > 0$ we iteratively define

$$\mu_{k\tau} \in \operatorname{Arg\,Min}\left\{ \frac{1}{2\tau} \mathsf{HK}_{\alpha,\beta}^2(\mu_{(k-1)\tau}, \mu) + \mathscr{E}(\mu) \;\middle|\; \mu \in \mathcal{M}(\Omega) \right\} \qquad (1.29)$$

and consider the limit $\tau \downarrow 0$ (along subsequences) to obtain *generalized minimizing movements* (GMM) (cf. [2]). Under suitable conditions, including the assumption $\mathscr{E}(\mu) = \int_\Omega \left(E(c) + cV \right) dx$ with $\mu = c\mathcal{L}^d$ and E superlinear, it is shown in [15, Thm. 3.4] that all GMM μ have the form $\mu(t) = c(t)\mathcal{L}^d$, and the density c is a weak solution of the reaction-diffusion equation

$$\partial_t c = \alpha \operatorname{div}\left(c \nabla(E'(c)+V) \right) - \beta\, u\left(E'(c)+V \right) \text{ in } \Omega,$$
$$c \nabla(E'(c)+V) \cdot \mathrm{n} = 0 \text{ on } \partial\Omega.$$

In [24], the equation $u_t = 0 = \Delta u + au \log u + bu$ is studied, whose solutions are steady states for HK gradient flows for $\mathscr{E}(u) = \int_{\mathbb{R}^d} u \log u \, dx$. We also refer to [13,32], where equation (1.29) was studied for $E(c) = \frac{1}{m} c^m - \lambda c$ and $V \equiv 0$. The linear functional $\Phi(\mu) = \int_{\mathbb{R}^d} V(x) \, d\mu$ for a given potential $V \in C^0(\mathbb{R}^d)$ can easily be added, as its geodesic λ-convexity is characterized in [26, Prop. 20]. Note that our main convexity result, proved here for the first time, plays an important role in the existence and uniqueness results of [13], cf. Thm. 2.14 there.

In [23] it is shown that the GMM for the gradient system $(\mathcal{M}(\Omega), \mathsf{HK}_{\alpha,\beta}, \mathscr{E})$ are EVI_λ solutions in the sense of [30]. Again the main ingredient is the geodesic λ-convexity of $\mathscr{E}$ in the form (1.20) contained in our main Theorem 7.2.

2. The Hellinger–Kantorovich Distance

In this section, we recall a few properties and equivalent characterizations of the Hellinger–Kantorovich distance from [26,27], that will turn out to be crucial in the following.

Main Notation	
$\mathcal{M}(X)$, $\mathcal{M}_2(X)$	finite positive Borel measures on X (with finite quadratic moment)
$\mathcal{P}(X)$, $\mathcal{P}_2(X)$	Borel probability measures on X (with finite quadratic moment)
$T_\sharp \mu$	push forward of $\mu \in \mathcal{M}(X)$ by a map $T : X \to Y$: (2.1)
$\mu = c\mathcal{L}^d + \mu^\perp$	Lebesgue decomposition of a measure $\mu \mathcal{M}(\mathbb{R}^d)$
$C_b(X)$	continuous and bounded real functions on X
$\cos_a(r)$	truncated function $\cos\left(\min\{a, r\}\right)$, $a > 0$ (typically $a = \pi/2$)
$W_X(\mu_1, \mu_2)$	Kantorovich–Wasserstein distance in $\mathcal{P}_2(X)$
$\sin$, $\tan$, $\arctan$, $\cdots$	vector-valued version of the usual scalar functions, see (1.6)
$\mathsf{HK}(\mu_1, \mu_2)$	Hellinger–Kantorovich distance in $\mathcal{M}(X)$: Section 2
$(\mathfrak{C}, \mathsf{d}_{a,\mathfrak{c}})$, $\mathfrak{o}$	metric cone on $\mathbb{R}^d$ and its vertex, see Subsection 2.1.2
$W_{a,\mathfrak{c}}$	L^2-Kantorovich–Wasserstein distance on $\mathcal{P}_2(\mathfrak{C})$ induced by $\mathsf{d}_{a,\mathfrak{c}}$
x, r	coordinate maps on $\mathfrak{C}$, see Subsection 2.1.2
π^0, π^1	coordinate maps on a Cartesian product $X_0 \times X_1$, $\pi^i(x_0, x_1) = x_i$
$\mathfrak{h}$	homogeneous projection from $\mathcal{M}_2(\mathfrak{C})$ to $\mathcal{M}(\mathbb{R}^d)$, see (2.8)
$S_i, S_i', S_i'', S_i^{\pi/2}, \mu_i', \mu_i''$	see (2.12)–(2.13)
$(T, q)_\star$	action of a transport-growth map, Definition 2.7
$AC^p([0, 1]; X)$	space of curves $\mathsf{x} : [0, 1] \to X$ with p-integrable metric speed
$\varphi_0^{\mathbb{L}\to}, \varphi_1^{\leftarrow\mathbb{L}}$	forward and backward $\mathbb{L}$-transform for cost function $\mathbb{L}$, see (3.3)
D_i', D_i''	domains of $\nabla\varphi_i$ and $\mathrm{D}^2\varphi_i$, see Theorem 3.2
$\xi_s = \mathscr{P}_s\xi, \bar{\xi}_s = \mathscr{R}_s\bar{\xi}$	for- and backward solution of Hamilton–Jacobi equation, (4.2), (4.3)
Ξ_s	contact set of forward and backward solutions $\xi_s, \bar{\xi}_s$, see (4.39)
$(T_{s\to t}, q_{s\to t})$	transport-growth map induced by for/backward solutions, Theorem 4.3

First, we fix some notation that we will extensively use: Let (X, d_X) be a complete and separable metric space. In the present paper X will typically be $\mathbb{R}^d$ with the Euclidean distance, a closed convex subset thereof, the cone space $\mathfrak{C}$ on $\mathbb{R}^d$ (see Subsection 2.1.2), product spaces of the latter two, etc. We will denote by $\mathcal{M}(X)$ the space of all non-negative and finite Borel measures on X endowed with the weak topology induced by the duality with the continuous and bounded functions of $C_b(X)$. The subset of measures with finite quadratic moment will be denoted by $\mathcal{M}_2(X)$. The spaces $\mathcal{P}(X)$ and $\mathcal{P}_2(X)$ are the corresponding subsets of probability measures.

If $\mu \in \mathcal{M}(X)$ and $T : X \to Y$ is a Borel map with values in another metric space Y, then $T_\sharp\mu$ denotes the push-forward measure on $\mathcal{M}(Y)$, defined by

$$T_\sharp\mu(B) := \mu(T^{-1}(B)) \quad \text{for every Borel set } B \subset Y. \tag{2.1}$$

We will often denote elements of $X \times X$ by (x_0, x_1) and the canonical projections by $\pi^i : (x_0, x_1) \to x_i$, $i = 0, 1$. A coupling on X is a measure $\boldsymbol{\gamma} \in \mathcal{M}(X \times X)$ with marginals $\gamma_i := \pi^i_\sharp\boldsymbol{\gamma}$.

Given two measures $\mu_0, \mu_1 \in \mathcal{M}_2(X)$ with equal mass $\mu_0(X) = \mu_1(X)$, their (quadratic) Kantorovich–Wasserstein distance W_X is defined by

$$W_X(\mu_0, \mu_1)^2 := \min \left\{ \iint \mathsf{d}_X(x_0, x_1)^2 \, d\boldsymbol{\gamma}(x_0, x_1) \, \Bigg| \right.$$
$$\left. \boldsymbol{\gamma} \in \mathcal{M}(X \times X), \ \pi^i_\sharp\boldsymbol{\gamma} = \mu_i, \ i = 0, 1 \right\}. \tag{2.2}$$

We refer to [2] for a survey on the Kantorovich–Wasserstein distance and related topics.

2.1. Equivalent Formulations of the Hellinger–Kantorovich Distance

The Hellinger–Kantorovich distance was introduced in [26,27] and independently in [21] and [10,11]. It is a generalization of the Kantorovich–Wasserstein distance to arbitrary non-negative and finite measures by taking creation and annihilation of mass into account. Indeed, the latter can be associated with a different notion of distance, namely the Hellinger–Kakutani distance, see [19] and [33]. In this sense, the Hellinger–Kantorovich distance should be viewed as an infimal convolution of the Kantorovich–Wasserstein and the Hellinger–Kakutani distance, cf. [27, Rem. 8.19].

In [27], five different equivalent formulations of the Hellinger–Kantorovich distance are given: (i) the dynamical formulation, (ii) the cone space formulation, (iii) the optimal entropy-transport problem, (iv) the dual formulation in terms of Hellinger–Kantorovich potentials, and (v) the formulation using Hamilton–Jacobi equations. We will present and briefly discuss each of them below, as all are useful for our analysis of geodesic convexity.

In the follows, we consider the Hellinger–Kantorovich distance for measures on the domain $\mathbb{R}^d$. However, it is easy to see that all arguments also work in the case of a closed and convex domain $\Omega \subset \mathbb{R}^d$. In particular, the latter is a complete, geodesic space.

2.1.1. Dynamic Approach   A first approach to the Hellinger–Kantorovich distance is related to the dynamic formulation, which naturally depends on two positive parameters $\alpha, \beta > 0$; these control the relative strength of the Kantorovich–Wasserstein part and of the Hellinger-Kakutani part (see [27, Section 8.5]).

Definition 2.1. (*The dynamic formulation*) For every $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^d)$ we set

$$\mathsf{HK}_{\alpha,\beta}(\mu_0, \mu_1)^2 = \min \left\{ \int_0^1 \int_{\mathbb{R}^d} \left(\alpha \, |\Upsilon(t,x)|^2 + \beta \xi(t,x)^2 \right) \mathrm{d}\mu_t(x) \, \mathrm{d}t \, \right| \tag{2.3}$$
$$\left. \mu \in \mathrm{C}([0,1]; \mathcal{M}(\mathbb{R}^d)), \ \mu_{t=i} = \mu_i, \ \text{(gCE) holds} \right\},$$

where the generalized continuity equation for the Borel vector and scalar fields $\Upsilon : (0,1) \times \mathbb{R}^d \to \mathbb{R}^d$ and $\xi : (0,1) \times \mathbb{R}^d \to \mathbb{R}$ reads as

$$\text{(gCE)} \qquad \frac{\partial}{\partial t} \mu + \alpha \, \mathrm{div}(\mu \, \Upsilon) = \beta \, \xi \mu \quad \text{in } \mathcal{D}'((0,1) \times \mathbb{R}^d).$$

Notice that (2.3) yields in particular that $\mu \, \Upsilon$ and $\xi \mu$ are (vector and scalar) measures with finite total mass, so that the canonical formulation of (gCE) in $\mathcal{D}'((0,1) \times \mathbb{R}^d)$ makes sense. For optimal solutions one has $\Upsilon(t,x) = \nabla \xi(t,x)$ and the dual potential solves the generalized Hamilton–Jacobi equation

$$\partial_t \xi + \frac{\alpha}{2} |\nabla \xi|^2 + \frac{\beta}{2} \xi^2 = 0 \tag{2.4}$$

in a suitable sense [27, Theorem 8.20].

A simple rescaling technique shows that it is sufficient to restrict ourselves to a specific choice of the parameters α and β. In fact, it is easy to see that for every $\theta > 0$ we have

$$\mathsf{HK}_{\alpha,\beta}(\mu_0,\mu_1)^2 = \theta \mathsf{HK}_{\theta\alpha,\theta\beta}(\mu_0,\mu_1)^2.$$

Moreover, if $\lambda > 0$ and we consider the spatial dilation $H : x \mapsto \lambda x$ in $\mathbb{R}^d$, we find

$$\mathsf{HK}_{\alpha,\beta}(\mu_0,\mu_1)^2 = \mathsf{HK}_{\alpha/\lambda^2,\beta}(H_\sharp\mu_0, H_\sharp\mu_1)^2.$$

Choosing $\lambda := \sqrt{4\alpha/\beta}$, $\theta = 4/\beta$, and setting $\mathsf{HK} := \mathsf{HK}_{1,4}$ we get

$$\mathsf{HK}_{\alpha,\beta}(\mu_0,\mu_1)^2 = \frac{4}{\beta}\mathsf{HK}_{4\alpha/\beta,4}(\mu_0,\mu_1)^2 = \frac{4}{\beta}\mathsf{HK}(H_\sharp\mu_0, H_\sharp\mu_1)^2.$$

Therefore, in order to keep simpler notation, in the remaining paper we will mainly consider the case $\alpha = 1$ and $\beta = 4$.

2.1.2. Cone Space Formulation

There is a second characterization that connects HK with the classic Kantorovich–Wasserstein distance on the extended cone $\mathfrak{C} := (\mathbb{R}^d \times [0,\infty[)/\sim$, where $\sim$ is the equivalence relation which identifies all the points $(x, 0)$ with the vertex $\mathfrak{o}$ of $\mathfrak{C}$. More precisely, we write $(x_0, r_0) \sim (x_1, r_1)$ if and only if $x_0 = x_1$ and $r_0 = r_1$ or $r_0 = r_1 = 0$ and introduce the notation $[x, r]$ to denote the equivalence class associated with $(x, r) \in \mathbb{R}^d \times [0, \infty[$. The cone $\mathfrak{C}$ is a complete metric space endowed with the cone distances

$$\mathsf{d}_{a,\mathfrak{C}}(\mathfrak{z}_0, \mathfrak{z}_1)^2 := r_0^2 + r_1^2 - 2r_0 r_1 \cos_a(|x_1-x_0|), \quad \mathfrak{z}_i = [x_i, r_i], \ a \in (0, \pi], \tag{2.5}$$

see e.g. [4, Sect. 3.6.2], where we use the abbreviation $\cos_a(r) := \cos\big(\min\{a, r\}\big)$. Notice that the projection map $(x, r) \mapsto [x, r]$ is bijective from $\mathbb{R}^d \times (0, \infty)$ to $\mathfrak{C}_* := \mathfrak{C} \setminus \{\mathfrak{o}\}$; we will denote by (x, r) its inverse, which we extend to $\mathfrak{o}$ by setting $\mathsf{x}(\mathfrak{o}) = 0$, $\mathsf{r}(\mathfrak{o}) = 0$.

The most natural choice of the parameter a in (2.5) is $a := \pi$: in this case the cone $(\mathfrak{C}, \mathsf{d}_{\pi,\mathfrak{C}})$ is a geodesic space, i.e., given $\mathfrak{z}_i = [x_i, r_i]$, $i = 0, 1$, there exists a curve $\mathfrak{z}_t = [x_t, r_t] = \mathrm{geo}_t(\mathfrak{z}_0, \mathfrak{z}_1)$, $t \in [0, 1]$, connecting $\mathfrak{z}_0$ to $\mathfrak{z}_1$ and satisfying

$$\forall 0 \leqq s, t \leqq 1: \quad \mathsf{d}_{\pi,\mathfrak{C}}(\mathfrak{z}_s, \mathfrak{z}_t) = |t-s|\,\mathsf{d}_{\pi,\mathfrak{C}}(\mathfrak{z}_0, \mathfrak{z}_1). \tag{2.6}$$

If one of the two points coincides with $\mathfrak{o}$, e.g. for $\mathfrak{z}_0 = \mathfrak{o}$, it is immediate to check that $\mathfrak{z}_t = [x_1, tr_1]$. If $r_0, r_1 > 0$ and $|x_1-x_0| < \pi/2$ then the unique geodesic curve reads (recall the convention in (1.6))

$$r_t := r_0\Big((1+tu)^2 + t^2|\mathbf{v}|^2\Big)^{1/2}, \quad x_t := x_0 + \mathbf{arctan}\Big(\frac{t\mathbf{v}}{1+tu}\Big),$$

$$\text{where } u = \frac{r_1}{r_0}\cos(|x_1-x_0|) - 1 \text{ and } \mathbf{v} := \frac{r_1}{r_0}\sin(x_1-x_0). \tag{2.7}$$

For example, if we operate the same construction starting from the one-dimensional set $\Omega = [0, L] \subset \mathbb{R}$ with $0 < L \leqq \pi$ we can isometrically identify the cone

space over Ω with the two-dimensional sector $\Sigma_\Omega = \{ y = (r \cos x, \ r \sin x) \in \mathbb{R}^2 \,|\, r \geqq 0, \ x \in [0, L] \}$ endowed with the Euclidean distance. For $L \in {]}\pi, 2\pi{[}$ the identification with the sector still holds, but the sector Σ_Ω is no more convex and for $x_0, x_1 \in \Omega$ with $|x_0 - x_1| \geqq \pi$ the cone distance corresponds to the geodesic distance on the sector Σ_Ω, i.e. the length of the shortest path in Σ_Ω connecting two points.

On the one hand, we can define a homogeneous projection $\mathfrak{h} : \mathcal{M}_2(\mathfrak{C}) \to \mathcal{M}(\mathbb{R}^d)$, via

$$\mathfrak{h}\lambda := \mathsf{x}_\sharp(\mathsf{r}^2\lambda) = \int_{r=0}^\infty r^2 \, \lambda(\cdot, \mathrm{d}r), \tag{2.8}$$

i.e. for every $\lambda \in \mathcal{M}_2(\mathfrak{C})$ and $\zeta \in C_b(\mathbb{R}^d)$ we have

$$\int_{\mathbb{R}^d} \zeta(x) \, \mathrm{d}(\mathfrak{h}\lambda) = \int_{\mathfrak{C}} r^2 \zeta(x) \, \mathrm{d}\lambda(x, r).$$

On the other hand, measures in $\mathcal{M}(\mathbb{R}^d)$ can be "lifted" to measures in $\mathcal{M}_2(\mathfrak{C})$, e.g. by considering the measure $\mu \otimes \delta_1$ for $\mu \in \mathcal{M}(\mathbb{R}^d)$. More generally, for every Borel map $r : \mathbb{R}^d \to {]}0, \infty{[}$ and constant $m_0 \geqq 0$, the measure $\lambda = m_0 \delta_{\mathfrak{o}} + \mu \otimes \frac{1}{r(\cdot)^2}\delta_{r(\cdot)}$ gives $\mathfrak{h}\lambda = \mu$.

Now, the cone space formulation of the Hellinger–Kantorovich distance between two measures $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^d)$ is given as follows, (see [26, Sec. 3]);

Theorem 2.2. (Optimal transport formulation on the cone) *For $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^d)$ we have*

$$\mathsf{HK}(\mu_0, \mu_1)^2 = \min \left\{ \mathsf{W}_{\pi, \mathfrak{C}}(\lambda_0, \lambda_1)^2 \,\Big|\, \lambda_i \in \mathcal{P}_2(\mathfrak{C}), \ \mathfrak{h}\lambda_i = \mu_i \right\} \tag{2.9a}$$

$$= \min \left\{ \iint_{\mathfrak{C} \times \mathfrak{C}} \mathsf{d}_{\pi, \mathfrak{C}}(z_0, z_1)^2 \mathrm{d}\boldsymbol{\lambda}(z_0, z_1) \,\Big|\, \mathfrak{h}_i \boldsymbol{\lambda} = \mu_i \right\}, \tag{2.9b}$$

where $\mathfrak{h}$ is defined in (2.8) and $\mathfrak{h}_i\boldsymbol{\lambda} := \mathfrak{h}(\pi_\sharp^i\boldsymbol{\lambda})$ for $\boldsymbol{\lambda} \in \mathcal{M}_2(\mathfrak{C} \times \mathfrak{C})$ and $i = 0, 1$.

The cone space formulation is reminiscent of classical optimal transport problems. Here, however, the marginals λ_i of the transport plan $\boldsymbol{\lambda} \in \mathcal{M}(\mathfrak{C} \times \mathfrak{C})$ are not fixed, and it is part of the problem to find an optimal pair of measures λ_i satisfying the constraints $\mathfrak{h}\lambda_i = \mu_i$ and having minimal Kantorovich–Wasserstein distance on the cone space.

Remark 2.3. (**Hellinger–Kantorovich space as cone**) In [22] it is shown that the space $(\mathcal{M}(\mathbb{R}^d); \mathsf{HK})$ can be understood as a cone space over the geodesic space $(\mathcal{P}(\mathbb{R}^d), \mathsf{SHK})$ where the *spherical Hellinger–Kantorovich distance* in $\mathcal{P}(\mathbb{R}^d)$ reads $\mathsf{SHK}(\nu_0, \nu_1) := \arccos\left(1 - \frac{1}{2}\mathsf{HK}(\nu_0, \nu_1)^2\right)$. It would be interesting to analyze geodesic convexity properties of functionals $\mathscr{E}$ as in (1.20) on this space; see [23] for a first result.

The cone space formulation in (2.9) reveals many interesting geometric properties of the Hellinger–Kantorovich distance, e.g. Hellinger–Kantorovich geodesics are directly connected to geodesic curves in the cone space $\mathfrak{C}$, see below. Moreover, it can be deduced that a sharp threshold exists, which distinguishes between transport of mass and pure growth (i.e. creation or destruction) of mass.

Remark 2.4. The link between the dynamical formulation in (2.3) and the cone-space formulation in (2.9) of the Hellinger–Kantorovich distance can be best seen from a Lagrangian point of view. Let $\mathsf{Lag}_{\alpha,\beta}(X, r; V, \varrho) = \frac{r^2}{\alpha}|V|^2 + \frac{4}{\beta}\varrho^2$ denote the rescaled Lagrangian in the definition of the dynamical functional (2.3) corresponding to a curve of the form $\mu_t := r^2(t)\delta_{X(t)}$ and consider for fixed $r_0, r_1 > 0$ and $x_0, x_1 \in \mathbb{R}^d$ the minimization problem

$$M_{\alpha,\beta}(x_0, r_0; x_1, r_1) := \min \left\{ \int_0^1 \mathsf{Lag}_{\alpha,\beta}\big(X(s), r(s); \dot{X}(s), \dot{r}(s)\big) \mathrm{d}s \, \Big| \right.$$
$$\left. (X, r) \in \mathrm{C}^1\big([0, 1]; \mathbb{R}^d \times \mathbb{R}_+\big), \, X(i) = x_i, \, r(i) = r_i, \right\}.$$

It is not hard to check [26, Sec. 3.1] that we obtain for $(\alpha, \beta) = (1, 4)$ the explicit formula

$$\mathsf{HK}(\mu_0, \mu_1)^2 = M_{1,4}(x_0, r_0; x_1, r_1) = \mathsf{d}_{\pi,\mathfrak{C}}([x_0, r_0], [x_1, r_1])^2,$$

which is the Hellinger–Kantorovich distance of the two Dirac measures $\mu_0 = r_0^2\delta_{x_0}$ and $\mu_1 = r_1^2\delta_{x_1}$ in the case that $|x_0 - x_1| < \pi/2$.

When $|x_0 - x_1| \geq \pi/2$, one can always connect μ_0 to μ_1 by the curve $\mu_t := \big((1-t)r_0\big)^2\delta_{x_0} + t^2 r_1^2\delta_{x_1}$ (whose support is no longer concentrated on a single point) obtaining

$$\mathsf{HK}(\mu_0, \mu_1)^2 = 2 = \mathsf{d}_{\pi/2,\mathfrak{C}}([x_0, r_0], [x_1, r_1])^2,$$

and showing the role of the threshold $\pi/2$ instead of π in the computation of HK.

The explicit computation of the previous remark is in fact a particular case of a general result [27, Lem. 7.9+7.19].

Theorem 2.5. (Effective $\pi/2$-threshold in the cone distance) *Let $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^d)$, if $\lambda \in \mathcal{M}_2(\mathfrak{C} \times \mathfrak{C})$ is an optimal plan for the cone-space formulation (2.9) then $\lambda \llcorner (\mathfrak{C} \times \mathfrak{C}) \setminus \{(\mathfrak{o}, \mathfrak{o})\}$ is still optimal and*

$$\lambda\left(\left\{([x_0, r_0], [x_1, r_1]) \in \mathfrak{C} \times \mathfrak{C} \,\Big|\, r_0 r_1 > 0 \text{ and } |x_0 - x_1| > \frac{\pi}{2}\right\}\right) = 0, \quad (2.10)$$

so that

$$\mathsf{HK}(\mu_0, \mu_1)^2 = \min\left\{\mathsf{W}_{\pi/2,\mathfrak{C}}(\lambda_0, \lambda_1)^2 \,\Big|\, \lambda_i \in \mathcal{P}_2(\mathfrak{C}), \, \mathfrak{h}\lambda_i = \mu_i\right\} \quad (2.11a)$$

$$= \min\left\{\iint_{\mathfrak{C} \times \mathfrak{C}} \mathsf{d}_{\pi/2,\mathfrak{C}}(z_0, z_1)^2 \mathrm{d}\lambda(z_0, z_1) \,\Big|\, \mathfrak{h}_i\lambda = \mu_i\right\}. \quad (2.11b)$$
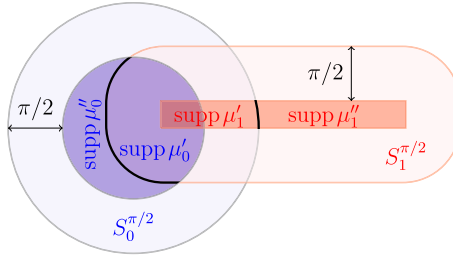
Fig. 1. The decomposition of the closed supports $S_i = \operatorname{supp} \mu_i$ of the measures $\mu_i = \mu_i' + \mu_i''$ as given in (2.13) with cut-off at $\pi/2$. The open sets $S_0^{\pi/2}$ and $S_1^{\pi/2}$ denote the $\pi/2$-neighborhoods of the supports S_1 and S_0, respectively, and $\mu_i' = \mu_i \llcorner (S_{1-i}^{\pi/2} \cap S_i)$, $\mu_i'' = \mu_i \llcorner (S_i \backslash S_{1-i}^{\pi/2})$ are the corresponding restrictions of the measures μ_i

Moreover, setting for $i = 0, \tau$

$$S_i := \operatorname{supp}(\mu_i), \quad S_i^{\pi/2} := \left\{ x \in \mathbb{R}^d \,\middle|\, \operatorname{dist}(x, S_i) < \pi/2 \right\},$$
$$S_i' := S_i \cap S_{1-i}^{\pi/2}, \quad and \; S_i'' := S_i \setminus S_{1-i}^{\pi/2}, \tag{2.12}$$

(see Fig. 1) with the related decomposition

$$\mu_i := \mu_i' + \mu_i'', \quad \mu_i' := \mu_i \llcorner S_i' = \mu_i \llcorner S_{1-i}^{\pi/2}, \quad and \; \mu_i'' := \mu_i \llcorner S_i'', \tag{2.13}$$

then we have that

$$\mathsf{HK}(\mu_0, \mu_\tau)^2 = \mathsf{HK}(\mu_0', \mu_\tau')^2 + \mathsf{HK}(\mu_0'', \mu_\tau'')^2, \tag{2.14a}$$
$$\mathsf{HK}(\mu_0'', \mu_\tau'')^2 = \mu_0''(\mathbb{R}^d) + \mu_\tau''(\mathbb{R}^d) = \mu_0(\mathbb{R}^d \setminus S_0') + \mu_\tau(\mathbb{R}^d \setminus S_\tau'). \tag{2.14b}$$

Note that (2.14a) shows that the decomposition in (2.13) is extremal with respect to the subadditivity property in Lemma 7.8 of [27], and (2.14b) shows that the computation of HK^2 between μ_0'' and μ_1'' is trivial, so that no information is lost if one restricts the evaluation of HK^2 to $\mu_0' = \mu_0 \llcorner S_0'$ and $\mu_1' = \mu_1 \llcorner S_1'$. Motivated by the above properties, we introduce the following definition of reduced pairs, which will play a crucial role in our analysis of geodesic curves;

Definition 2.6. (*Reduced pairs*) A pair $(\mu_0, \mu_1) \in \mathcal{M}(\mathbb{R}^d)^2$ is called *reduced* (resp. *strongly reduced*) if $\mu_i(S_i'') = 0$, i.e. $\mu_i = \mu_i'$ for $i = 0$ and 1 (resp. if $S_i \subset S_{1-i}^{\pi/2}$).

By definition, the sets $S_i = \operatorname{supp}(\mu_i)$ are closed and $S_i^{\pi/2}$ are open, so that $S_i'' = S_i \setminus S_{1-i}^{\pi/2}$ is closed as well, but $S_i' = S_i \cap S_{1-i}^{\pi/2}$ may be neither closed nor open. In the strongly reduced case the condition $S_i \subset S_{1-i}^{\pi/2}$ means that, at least locally, the closed set S_i has a positive distance to the boundary of the open set $S_{1-i}^{\pi/2}$.

Notice that for every $(\mu_0, \mu_1) \in \mathcal{M}(\mathbb{R}^d)^2$ the corresponding pair (μ_0', μ_1') defined according to (2.12)–(2.13) is reduced by construction. In fact, if $x \in S_0'$ then there exists $y \in \operatorname{supp}(\mu_1)$ with $|x-y| < \pi/2$: clearly $y \in S_1'$ so that $\operatorname{dist}(x, \operatorname{supp}(\mu_1')) \leqq \operatorname{dist}(x, S_1') < \pi/2$.

2.1.3. Transport-growth Maps   It is useful to express (2.11b) in an equivalent way, which extends the notion of transport maps to the unbalanced case. This relies on special families of plans in $\lambda \in \mathcal{M}_2(\mathfrak{C}^2)$ with $\mathfrak{h}_i \lambda = \mu_i$ generated by transport-growth systems.

Definition 2.7. (*Transport-growth maps*) Let $v \in \mathcal{M}(Y)$, where Y is some Polish space. A transport-growth map is a v-measurable map $(\boldsymbol{T}, q) : Y \to X \times [0, \infty)$ with $q \in L^2(Y, v)$. It acts on v according to this rule:

$$(\boldsymbol{T}, q)_\star v := \boldsymbol{T}_\sharp(q^2 v) = \mathfrak{h}((\boldsymbol{T}, q)_\sharp v). \tag{2.15}$$

there the last identity involves the obvious generalization of the definition (2.8) of homogeneous projection $\mathfrak{h}$ from $\mathcal{M}_2(X \times [0, \infty))$ to $\mathcal{M}(X)$.

We notice that transport-growth maps obey the composition rule

$$(\boldsymbol{T}_2, q_2)_\star(\boldsymbol{T}_1, q_1)_\star v = (\boldsymbol{T}, q)_\star v \quad \text{where} \quad \boldsymbol{T} := \boldsymbol{T}_2 \circ \boldsymbol{T}_1, \ q := (q_2 \circ \boldsymbol{T}_1)q_1. \tag{2.16}$$

Transport-growth maps provide useful upper bounds for the HK metric, playing a similar role of transport maps for the Kantorovich–Wasserstein distance. In fact, for every choice of maps $(\boldsymbol{T}_i, q_i) : Y \to \mathbb{R}^d \times [0, \infty)$, $i = 0, 1$, associated with the measure $v \in \mathcal{M}(Y)$, we have

$$\mathsf{HK}^2(\mu_0, \mu_1) \leqq \int_Y \left(q_0^2 + q_1^2 - 2q_0q_1 \cos_{\pi/2}(|\boldsymbol{T}_0 - \boldsymbol{T}_1|) \right) \mathrm{d}v \quad \mu_i := (\boldsymbol{T}_i, q_i)_\star v. \tag{2.17}$$

In order to show (2.17) it is sufficient to check that the measure $\lambda \in \mathcal{M}_2(\mathfrak{C}^2)$ defined by

$$\lambda := (\boldsymbol{T}_0, q_0; \boldsymbol{T}_1, q_1)_\sharp v, \quad \text{satisfies} \quad \mathfrak{h}_i \lambda = \mu_i$$

so that (2.17) follows from (2.11b) and the identity

$$\int_{\mathfrak{C}^2} \mathsf{d}_{\pi/2, \mathfrak{C}}(z_0, z_1)^2 \, \mathrm{d}\lambda = \int_Y \left(q_0^2 + q_1^2 - 2q_0q_1 \cos_{\pi/2}(|\boldsymbol{T}_0 - \boldsymbol{T}_1|) \right) \mathrm{d}v. \tag{2.18}$$

On the other hand, choosing $Y = \mathfrak{C} \times \mathfrak{C}$ and an optimal plan $v = \lambda \in \mathcal{M}_2(\mathfrak{C} \times \mathfrak{C})$ for (2.11b)

and setting $\boldsymbol{T}_i([x_0, r_0], [x_1, r_1]) := x_i$ and $q_i([x_0, r_0], [x_1, r_1]) = r_i$, we immediately find

$$\mathsf{HK}^2(\mu_0, \mu_1) = \int_{\mathfrak{C} \times \mathfrak{C}} \left(q_0^2 + q_1^2 - 2q_0q_1 \cos_{\pi/2}(|\boldsymbol{T}_0 - \boldsymbol{T}_1|) \right) \mathrm{d}\lambda, \quad \mu_i := (\boldsymbol{T}_i, q_i)_\star \lambda, \tag{2.19}$$

and therefore equality holds in (2.17).

Corollary 2.8. (HK *via transport-growth maps*) *For every* $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^d)$ *we have*

$$\mathsf{HK}^2(\mu_0, \mu_1) = \min \left\{ \int_{\mathfrak{C} \times \mathfrak{C}} \left(q_0^2 + q_1^2 - 2q_0 q_1 \cos_{\pi/2}(|\boldsymbol{T}_0 - \boldsymbol{T}_1|) \right) d\boldsymbol{\lambda} \, \middle| \, \boldsymbol{\lambda} \in \mathcal{M}(Y), \right.$$
$$\left. Y \text{ Polish}, \ (\boldsymbol{T}_i, q_i) : Y \to \mathbb{R}^d \times [0, +\infty), \ \mu_i := (\boldsymbol{T}_i, q_i)_\star \boldsymbol{\lambda} \right\}; \tag{2.20}$$

moreover, it is not restrictive to choose $Y = \mathfrak{C} \times \mathfrak{C}$ *in* (2.20).

Inspired by the so-called Monge formulation of Optimal Transport, it is natural to look for similar improvement of (2.20), when $Y = \mathbb{R}^d$, $\nu = \mu_0$, $\boldsymbol{T}_0(x_0) = x_0$ is the identity map, and $q(x_0) \equiv 1$.

Problem 2.9. *(Monge formulation of* HK *problem) Given* $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^d)$ *such that* $\mu_1 = \mu_1'$, $\mu_1'' = 0$ *(recall* (2.12) *and* (2.13)*), find an optimal transport-growth pair* $(\boldsymbol{T}, q) : \mathbb{R}^d \to \mathbb{R}^d \times [0, \infty)$ *minimizing the cost*

$$\mathcal{C}(\boldsymbol{T}, q; \mu_0) := \int_{\mathbb{R}^d} \left(1 + q^2(x) - 2q(x) \cos_{\pi/2}(|\boldsymbol{T}(x) - x|) \right) d\mu_0(x) \tag{2.21}$$

among all the transport-growth maps satisfying $(\boldsymbol{T}, q)_\star \mu_0 = \mu_1$

By (2.17) we have the bound

$$\mathsf{HK}(\mu_0, \mu_1)^2 \leqq \inf \left\{ \mathcal{C}(\boldsymbol{T}, q; \mu_0) \, \middle| \, (\boldsymbol{T}, q)_\star \mu_0 = \mu_1 \right\}. \tag{2.22}$$

When $\mu_0 \ll \mathcal{L}^d$ and the support of μ_1 is contained in the closed neighborhood of radius $\pi/2$ of the support of μ_0, the results of the next section (cf. Corollary 3.5), which are a consequence of the optimality conditions in Theorem 2.14, show that the minimum of Problem 2.9 is attained and realizes the equality in (2.22).

2.1.4. Entropy-Transport Problem   A third point of view, typical of optimal transport problems, characterizes the Hellinger–Kantorovich distance via the static Logarithmic Entropy Transport (LET) variational formulation.

We define the logarithmic entropy density $F : [0, \infty[\to [0, \infty[$ via

$$F(s) := s \log s - s + 1 \quad \text{for } s > 0 \quad \text{and} \quad F(0) := 1,$$

and the cost function $\mathsf{L}_1 : \mathbb{R}^d \to [0, \infty]$ via

$$\mathsf{L}_1(x) := \frac{1}{2} \ell(|x|), \quad \ell(r) := \begin{cases} -\log(\cos^2(r)) = \log\left(1 + \tan^2(r)\right) & \text{for } r < \pi/2, \\ +\infty & \text{otherwise.} \end{cases} \tag{2.23}$$

For given $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^d)$ the entropy-transport functional $\mathscr{ET}(\cdot \, ; \mu_0, \mu_1) : \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d) \to [0, \infty]$ reads as

$$\mathscr{ET}(\eta; \mu_0, \mu_1) := \int_{\mathbb{R}^d} F(\sigma_0)\, d\mu_0 + \int_{\mathbb{R}^d} F(\sigma_1)\, d\mu_1 + \iint_{\mathbb{R}^d \times \mathbb{R}^d} 2L_1(x_0 - x_1)\, d\eta,$$

(2.24)

with $(\pi_i)_\sharp \eta = \sigma_i \mu_i \ll \mu_i$. As usual, we set $\mathscr{ET}(\eta; \mu_0, \mu_1) := +\infty$ if one of the marginals $(\pi_i)_\sharp \eta$ of η is not absolutely continuous with respect to μ_i. With this definition, the equivalent formulation of the Hellinger–Kantorovich distance as entropy-transport problem reads as follows:

Theorem 2.10. (LET formulation) *For every $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^d)$ we have*

$$\mathsf{HK}(\mu_0, \mu_1)^2 = \min\left\{\mathscr{ET}(\eta; \mu_0, \mu_1) \,\middle|\, \eta \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d)\right\}. \qquad (2.25)$$

Moreover, recalling the decomposition (2.12)–(2.13)*,*

(1) *the pairs (μ_0, μ_τ) and (μ_0', μ_τ') share the same optimal plans η*
(2) *if we set $g_0(x_0) := ([x_0, 1], \mathfrak{o})$ and $g_1(x_1) := (\mathfrak{o}, [x_1, 1])$, every optimal plan $\eta \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d)$ for the entropy-transport formulation in* (2.25) *induces optimal plans β (resp. β') in $\mathcal{M}(\mathfrak{C} \times \mathfrak{C})$ for the pair (μ_0, μ_1) (resp. the reduced pair (μ_0', μ_1')) via*

$$\beta' := (x_0, \sigma_0^{-1/2}; x_1, \sigma_1^{-1/2})_\sharp \eta, \quad \beta := \beta' + (g_0)_\sharp \mu_0'' + (g_1)_\sharp \mu_1''. \tag{2.26}$$

An optimal transport plan η, which always exists, gives the effective transport of mass. Note, in particular, that the finiteness of $\mathscr{ET}$ only requires $(\pi_i)_\sharp \eta = \eta_i \ll \mu_i$ (which is considerably weaker than the usual transport constraint $(\pi_i)_\sharp \eta = \mu_i$) and the cost of a deviation of η_i from μ_i is given by the entropy functionals associated with F. Moreover, the cost function ℓ is finite in the case $|x_0 - x_1| < \pi/2$, which highlights the sharp threshold between transport and pure creation/destruction. Notice that we could equivalently use the truncated function $\cos^2_{\pi/2}(r) = \cos^2(\min\{r, \pi/2\})$ instead of $\cos^2(r)$ in (2.23). As we have already seen, the function $r \mapsto \cos^2_{\pi/2}(r)$ plays an important role in many formulae.

In general, optimal entropy-transport plans $\eta \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d)$ are not unique. However, due to the strict convexity of F, their marginals η_i are unique so that the non-uniqueness of the plan η is solely a property of the optimal transport problem associated with the cost function $(x_0, x_1) \mapsto 2L_1(x_1 - x_0) = \ell(|x_1 - x_0|)$.

Remark 2.11. Besides (2.26), the connection between the cone-space formation and the logarithmic entropy-transport problem is given by the homogeneous marginal perspective function, namely

$$\mathsf{d}_{\pi/2, \mathfrak{C}}([x_0, r_0], [x_1, r_1])^2 = \inf\left\{r_0^2 F(\tfrac{\theta}{r_0^2}) + r_1^2 F(\tfrac{\theta}{r_1^2}) + 2\theta L_1(x_0 - x_1) \,\middle|\, \theta > 0\right\},$$

where r_i^2 plays the role of the reverse densities $1/\sigma_i$ and θ is a scaling parameter, see [27, Sec. 5].

We highlight that the logarithmic entropy-transport formulation (2.25) can be easily generalized by considering convex and lower semi-continuous functions F_0 and F_1 and cost functions ℓ, see [27, Part I].

Applying the previous Theorem 2.10 we can refine formula (2.18) by providing an optimal pair of transport-growth maps solving (2.20) in the restricted set $Y = S_0 \times S_1 \subset \mathbb{R}^d \times \mathbb{R}^d$. Indeed, we can choose arbitrary points $\bar{x}_i \in S_i$ and

$$\boldsymbol{\nu} := \boldsymbol{\eta} + \mu_0'' \otimes \delta_{\bar{x}_1} + \delta_{\bar{x}_0} \otimes \mu_1'',$$

$$\boldsymbol{T}_i(x_0, x_1) := x_i, \quad q_i(x_0, x_1) := \begin{cases} \sigma_i^{-1/2}(x_i) & \text{if } (x_0, x_1) \in S_0' \times S_1', \\ 1 & \text{if } (x_0, x_1) \in (S_0 \times S_1) \setminus (S_0' \times S_1'), \end{cases}$$

$$(2.27)$$

which satisfies

$$(\boldsymbol{T}_i, q_i)_\star \boldsymbol{\nu} = \mu_i, \quad \mathsf{HK}^2(\mu_0, \mu_1) = \int_Y \left(q_0^2 + q_1^2 - 2q_0 q_1 \cos_{\pi/2}(|\boldsymbol{T}_0 - \boldsymbol{T}_1|) \right) \mathrm{d}\boldsymbol{\nu}.$$

$$(2.28)$$

2.1.5. Dual Formulation with Hellinger–Kantorovich Potentials In analogy to the Kantorovich–Wasserstein distance, we can give a dual formulation in terms of Hellinger–Kantorovich potentials. We slightly modify the notation of [27], in order to be more consistent with the approach by the Hamilton–Jacobi equations (and the related Hopf–Lax solutions) of Section 4 and to deal with rescaled distances. As we will study segments of constant-speed geodesics $t \to \mu_t$ of length $\tau = t - s$ for $0 \leqq s < t \leqq 1$, it will be convenient to introduce a scaling parameter $\tau > 0$ that in certain parts will be replaced by 1, namely if we consider a whole geodesic. With this parameter, we set

$$F_\tau(s) := \frac{1}{2\tau} F(s), \quad \mathsf{L}_\tau(x) = \frac{1}{2\tau} \ell(|x|), \quad \mathscr{T}_\tau(\boldsymbol{\eta}; \mu_0, \mu_\tau) = \frac{1}{2\tau} \mathscr{T}(\boldsymbol{\eta}; \mu_0, \mu_\tau),$$

$$(2.29)$$

and the corresponding

$$\frac{1}{2\tau} \mathsf{HK}^2(\mu_0, \mu_\tau) = \min \left\{ \mathscr{T}_\tau(\boldsymbol{\eta}; \mu_0, \mu_\tau) \,\Big|\, \boldsymbol{\eta} \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d) \right\}. \quad (2.30)$$

It is clear that minimizers $\boldsymbol{\eta}$ of (2.30) are independent of the coefficient $\frac{1}{2\tau}$ in front of HK and coincide with solutions to (2.25) if $\mu_\tau = \mu_1$. The role of τ just affects the rescaling of the potentials φ and ξ we will introduce below.

We also introduce the Legendre transform of F_τ

$$\check{\mathsf{G}}_\tau(\varphi) := F_\tau^*(\varphi) = \sup_{s > 0} \varphi s - F_\tau(s) = \frac{\mathrm{e}^{2\tau\varphi} - 1}{2\tau},$$

$$\mathsf{G}_\tau(\varphi) := \frac{1 - \mathrm{e}^{-2\tau\varphi}}{2\tau} = -\check{\mathsf{G}}_\tau(-\varphi), \quad (2.31)$$

extended to $[-\infty, +\infty]$ by

$$\mathsf{G}_\tau(+\infty) = -\check{\mathsf{G}}_\tau(-\infty) = \frac{1}{2\tau}, \quad \mathsf{G}_\tau(-\infty) = -\check{\mathsf{G}}_\tau(+\infty) = +\infty, \quad (2.32)$$

and their inverses

$$\check{\mathsf{G}}_\tau^{-1}(\xi) := \frac{1}{2\tau}\log(1+2\tau\xi), \quad \mathsf{G}_\tau^{-1}(\xi) := -\frac{1}{2\tau}\log(1-2\tau\xi) = -\check{\mathsf{G}}_\tau^{-1}(-\xi),$$

(2.33)

defined for $\xi \in [-\frac{1}{2\tau}, +\infty]$ and $\xi \in [-\infty, \frac{1}{2\tau}]$ respectively, with the obvious convention induced by (2.32). With Theorem 6.3 in [27] (see also Section 4 therein), we have the equivalent characterization of HK via the dual formulation

$$\frac{1}{2\tau}\mathsf{HK}(\mu_0, \mu_\tau)^2 = \sup\left\{ \int_{\mathbb{R}^d} \mathsf{G}_\tau(\varphi_\tau)\,\mathrm{d}\mu_\tau - \int_{\mathbb{R}^d} \check{\mathsf{G}}_\tau(\varphi_0)\,\mathrm{d}\mu_0 \;\middle|\; \right.$$
$$\left. \varphi_0, \varphi_\tau \in \mathrm{C}_b(\mathbb{R}^d), \; \varphi_\tau(x_\tau)-\varphi_0(x_0) \le \mathrm{L}_\tau(x_\tau-x_0)\right\}$$

$$= \sup\left\{ \int_{\mathbb{R}^d} \xi_\tau\,\mathrm{d}\mu_\tau - \int_{\mathbb{R}^d} \xi_0\,\mathrm{d}\mu_0 \;\middle|\; \xi_i \in \mathrm{C}_b(\mathbb{R}^d),\right.$$
$$\left. \sup_{\mathbb{R}^d}\xi_\tau < \frac{1}{2\tau}, \; \inf_{\mathbb{R}^d}\xi_0 > -\frac{1}{2\tau} \right. \tag{2.34a}$$
$$\left. \left(1-2\tau\xi_\tau(x_\tau)\right)\left(1+2\tau\xi_0(x_0)\right) \ge \cos^2_{\pi/2}(|x_0-x_\tau|)\right\}. \tag{2.34b}$$

Note that the formulations in (2.34a) and (2.34b) are connected by the transformation $\xi_\tau = \mathsf{G}_\tau(\varphi_\tau)$, $\xi_0 = \check{\mathsf{G}}_\tau(\varphi_0)$ and the last condition in (2.34b) is equivalent to

$$\mathsf{G}_\tau^{-1}\left(\xi_\tau(x_\tau)\right)-\check{\mathsf{G}}_\tau^{-1}\left(\xi_0(x_0)\right) \le \mathrm{L}_\tau(x_\tau-x_0). \tag{2.35}$$

It is not difficult to check that one can also consider Borel functions in (2.34a) and (2.34b), e.g. for all Borel functions $\varphi_i : \mathbb{R}^d \to [-\infty, +\infty]$ with

$$\int_{\mathbb{R}^d} \mathrm{e}^{-2\tau\varphi_\tau}\,\mathrm{d}\mu_\tau < \infty, \quad \int_{\mathbb{R}^d} \mathrm{e}^{2\tau\varphi_0}\,\mathrm{d}\mu_0 < \infty,$$
$$\varphi_\tau(x_1) \le \mathrm{L}_\tau(x_1-x_0) + \varphi_0(x_0) \quad \text{for all } x_0, x_\tau \in \mathbb{R}^d \text{ with } |x_0-x_\tau| < \pi/2,$$

(2.36)

we have

$$\frac{1}{2\tau}\mathsf{HK}(\mu_0, \mu_\tau)^2 \ge \int_{\mathbb{R}^d} \mathsf{G}_\tau(\varphi_\tau)\,\mathrm{d}\mu_\tau - \int_{\mathbb{R}^d} \check{\mathsf{G}}_\tau(\varphi_0)\,\mathrm{d}\mu_0. \tag{2.37}$$

If we allow extended valued Borel functions, the supremum in (2.34a) and (2.34b) are attained.

Theorem 2.12. (Existence of optimal dual pairs) *For all $\mu_0, \mu_\tau \in \mathcal{M}(\mathbb{R}^d)$ and $\tau > 0$ there exists an optimal pair of Borel potentials $\varphi_0, \varphi_\tau : \mathbb{R}^d \to [-\infty, +\infty]$ which is admissible according to (2.36) and realizes equality in (2.37), namely*

$$\frac{1}{2\tau}\mathsf{HK}(\mu_0, \mu_\tau)^2 = \int_{\mathbb{R}^d} \mathsf{G}_\tau(\varphi_\tau)\,\mathrm{d}\mu_\tau - \int_{\mathbb{R}^d} \check{\mathsf{G}}_\tau(\varphi_0)\,\mathrm{d}\mu_0. \tag{2.38}$$

The transformations $\xi_0 := \check{G}_\tau(\varphi_0) : \mathbb{R}^d \to [-1/(2\tau), +\infty]$, and $\xi_\tau := G_\tau(\varphi_\tau) : \mathbb{R}^d \to [-\infty, 1/(2\tau)]$, give an optimal pair for (2.34b) (dropping $\xi_i \in C_b(\mathbb{R}^d)$) satisfying

$$\int_{\mathbb{R}^d} |\xi_i| \, d\mu_i < \infty, \quad i = 0, \tau, \tag{2.39}$$

$$(1-2\tau\xi_\tau(x_\tau))(1+2\tau\xi_0(x_0)) \geqq \cos^2_{\pi/2}(|x_0-x_\tau|) \quad \text{if } \xi_0(x_0) < \infty, \ \xi_\tau(x_\tau) > -\infty, \tag{2.40}$$

$$\frac{1}{2\tau}\mathsf{HK}(\mu_0, \mu_\tau)^2 = \int_{\mathbb{R}^d} \xi_\tau \, d\mu_\tau - \int_{\mathbb{R}^d} \xi_0 \, d\mu_0. \tag{2.41}$$

Remark 2.13. Denoting by $S_i := \text{supp}(\mu_i)$ the support of μ_i for $i = 0$ and 1, we remark that it is always sufficient to find Borel potentials $\varphi_i : S_i \to [-\infty, +\infty]$ satisfying (2.36) on $S_0 \times S_1$ instead of $\mathbb{R}^d \times \mathbb{R}^d$. By setting $\tilde{\varphi}_1 := -\infty$ in $\mathbb{R}^d \setminus S_1$ and $\tilde{\varphi}_0 := +\infty$ in $\mathbb{R}^d \setminus S_0$ we obtain a pair still satisfying (2.36) and (2.38). This freedom will be useful in Theorem 2.14 below.

Moreover, notice that (2.34b) can be rewritten as

$$\frac{1}{2\tau}\mathsf{HK}(\mu_0, \mu_\tau)^2 = \sup\left\{ \int_{\mathbb{R}^d} \mathscr{P}_\tau\xi_0 \, d\mu_\tau - \int_{\mathbb{R}^d} \xi_0 \, d\mu_0 \,\Big|\, \xi_0 \in C_b(\mathbb{R}^d), \ \xi_0 > -\frac{1}{2\tau} \right\},$$

where $\mathscr{P}_\tau\xi$ is defined in (1.13). In particular, the operator $\mathscr{P}_\tau$ is directly connected to the dynamical formulation in (2.3), and we will thoroughly study its properties in Section 4.

2.2. First Order Optimality for HK

From the above discussion, we have already seen that there is never any transport over distances larger than $\pi/2$. This transport bound will also be seen in the following optimality conditions for the marginal densities σ_i defined in (2.24).

Theorem 2.14. (Optimality conditions [27, Thm. 6.3]) *Let $\mu_0, \mu_\tau \in \mathcal{M}(\mathbb{R}^d)$ and let S_i, S_i', S_i'', μ_i' be defined as in (2.12)–(2.13). The following holds:*

(1) A plan $\eta \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d)$ is optimal for the logarithmic entropy-transport problem in (2.30) if and only if
  – $\iint \ell \, d\eta < \infty$
  – its marginals η_i are absolutely continuous with respect to μ_i' (equivalently, η_i are absolutely continuous with respect to μ_i and $\eta_i(S_i'') = 0$),
  – there exist Borel densities $\sigma_i : \mathbb{R}^d \to [0, \infty]$ such that $\eta_i = \sigma_i \mu_i'$ and

$$\sigma_i = 0 \quad \text{on } S_i'', \tag{2.42a}$$

$$0 < \sigma_i < \infty \quad \text{on } S_i', \tag{2.42b}$$

$$\sigma_i = +\infty \quad \text{on } \mathbb{R}^d \setminus S_i, \tag{2.42c}$$

$$\sigma_0(x_0)\sigma_\tau(x_\tau) \geqq \cos^2_{\pi/2}(|x_0-x_\tau|) \quad \text{on } S_0 \times S_\tau, \tag{2.42d}$$

$$\sigma_0(x_0)\sigma_\tau(x_\tau) = \cos^2_{\pi/2}(|x_0-x_\tau|) \quad \eta\text{-a.e. on } S_0 \times S_\tau. \tag{2.42e}$$

In particular, the marginals η_i are unique and the densities σ_i are unique μ_i'-a.e.

(2) If η is optimal and S_i, S_i', S_i'' and σ_i are defined as above, the pairs of potentials defined by

$$\varphi_\tau := \begin{cases} -\frac{1}{2\tau}\log\sigma_\tau & in\ S_\tau', \\ +\infty & in\ S_\tau'', \\ -\infty & in\ \mathbb{R}^d \setminus S_\tau; \end{cases} \qquad \varphi_0 := \begin{cases} \frac{1}{2\tau}\log\sigma_0 & in\ S_0', \\ -\infty & in\ S_0'', \\ +\infty & in\ \mathbb{R}^d \setminus S_0; \end{cases} \qquad (2.43)$$

$$\xi_\tau := \begin{cases} \frac{1-\sigma_\tau}{2\tau} & in\ S_\tau', \\ \frac{1}{2\tau} & in\ S_\tau'', \\ -\infty & in\ \mathbb{R}^d \setminus S_\tau; \end{cases} \qquad \xi_0 := \begin{cases} \frac{\sigma_0-1}{2\tau} & in\ S_0', \\ -\frac{1}{2\tau} & in\ S_0'', \\ +\infty & in\ \mathbb{R}^d \setminus S_0; \end{cases} \qquad (2.44)$$

are optimal in the respective dual relaxed characterizations of Theorem 2.12 and satisfy η-a.e. in $\mathbb{R}^d \times \mathbb{R}^d$

$$\varphi_i(x_i) \in \mathbb{R}, \qquad\qquad \varphi_\tau(x_\tau) - \varphi_0(x_0) = \mathrm{L}_\tau(x_\tau - x_0), \qquad (2.45\mathrm{a})$$

$$-\xi_0(x_0), \xi_\tau(x_\tau) \in \Big(\frac{1}{2\tau}, \infty\Big), \ (1+2\tau\xi_0(x_0))(1-2\tau\xi_\tau(x_\tau)) = \cos^2_{\pi/2}(|x_0 - x_\tau|). \qquad (2.45\mathrm{b})$$

(3) Conversely, if η is optimal and $(\varphi_0, \varphi_\tau)$ (resp. (ξ_0, ξ_τ)) is an optimal pair according to Theorem 2.12, then $(2.45\mathrm{a})$ (resp. $(2.45\mathrm{b})$) holds η-a.e. and

$$\sigma_\tau = \mathrm{e}^{-2\tau\varphi_\tau} = 1 - 2\tau\xi_\tau \quad \mu_\tau\text{-a.e. in } S_\tau', \quad \varphi_\tau = +\infty, \ \xi_\tau = \frac{1}{2\tau} \quad \mu_\tau\text{-a.e. in } S_\tau'',$$

$$\sigma_0 = \mathrm{e}^{2\tau\varphi_0} = 1 + 2\tau\xi_0 \quad \mu_0\text{-a.e. in } S_0', \quad \varphi_0 = -\infty, \ \xi_0 = -\frac{1}{2\tau} \quad \mu_0\text{-a.e. in } S_0''.$$

$$(2.46)$$

3. Regularity of Static ℍK Potentials φ_0 and φ_1

In this section, we will carefully study the regularity of a pair (φ_0, φ_1) of optimal ℍK potentials arising in (2.43) of Theorem 2.14. We will improve the previous approximate differentiability result of [27, Thm. 6.6(iii)] (see also [2, Thm. 6.2.7]) by adapting the argument of [14] and extending the classical result of [16] to the ℍK setting. In fact, this section is largely independent of the specific ℍK setting but relies purely on the theory of $\mathbb{L}$-transforms. As we are interested in the special case of continuous, extended values cost functions $\mathbb{L} = \mathrm{L}_\tau = \frac{1}{\tau}\mathrm{L}_1 : \mathbb{R}^d \to [0, +\infty]$ which attain the value $+\infty$ outside a ball, we cannot rely on existing results and have to provide a careful analysis of this case (but see also [7,8,18,20,29] for different situations of discontinuous costs taking the value $+\infty$).

We will use the notion of locally semi-concave and semi-convex functions; recall that a function $\varphi : U \to \mathbb{R}$ defined in some open set U of $\mathbb{R}^d$ is locally semi-concave if for every point $\bar{x} \in U$ there exists $\rho > 0$ and a constant $C > 0$ with

$$x \mapsto \varphi(x) - \frac{C}{2}|x|^2 \quad \text{is concave in } B_\rho(\bar{x}). \qquad (3.1)$$

A function φ is locally semi-convex if $-\varphi$ is locally semi-concave. Let us recall that locally semi-concave functions are locally Lipschitz and thus differentiable almost everywhere. We will denote by $\mathrm{dom}(\nabla\varphi)$ the domain of their differential. By Alexandrov's Theorem (see [2, Thm. 5.5.4]), there exists for almost every $x \in \mathrm{dom}(\nabla\varphi)$ a symmetric matrix $\mathsf{A} =: \mathrm{D}^2\varphi(x)$ such that

$$\lim_{y \to x} \frac{\varphi(y) - \varphi(x) - \langle \nabla\varphi(x), y-x \rangle - \frac{1}{2}\langle \mathsf{A}(y-x), y-x \rangle}{|y-x|^2} = 0, \qquad (3.2\mathrm{a})$$

$$\text{and} \quad \lim_{\substack{y \to x \\ y \in \mathrm{dom}(\nabla\varphi)}} \frac{\nabla\varphi(y) - \nabla\varphi(x) - \mathsf{A}(y-x)}{|y-x|} = 0. \qquad (3.2\mathrm{b})$$

We will denote by $\mathrm{dom}(\mathrm{D}^2\varphi)$ the subset of density points in $\mathrm{dom}(\nabla\varphi)$ where (3.2a) and (3.2b) hold.

As the optimality of potential pairs (φ_0, φ_1) is closely related to the theory of $\mathbb{L}$-transforms, we give the basic definitions first and then derive the associated regularity properties under additional smoothness assumptions.

For simplicity, we restrict the analysis of the remaining text to continuous functions $\mathbb{L} : \mathbb{R}^d \to [0, \infty]$ satisfying $\mathrm{dom}(\mathbb{L}) = \left\{ z \in \mathbb{R}^d \,\middle|\, \mathbb{L}(z) \in \mathbb{R} \right\} = B_{\mathsf{R}}(0)$ for some $\mathsf{R} > 0$, i.e. $\mathbb{L}(z) < \infty$ for $|z| < \mathsf{R}$ and $\mathbb{L}(z) = +\infty$ for $|z| \geqq \mathsf{R}$. By continuity of $\mathbb{L}$ this behavior implies $\mathbb{L}(z_k) \to +\infty$ if $\liminf_{k \to \infty} |z_k| \geqq \mathsf{R}$.

We define the *forward $\mathbb{L}$-transform* $\varphi_0^{\mathbb{L}\to}$ of a l.s.c. function φ_0 and the *backward $\mathbb{L}$-transform* $\varphi_1^{\leftarrow\mathbb{L}}$ of an u.s.c. function φ_1 via

$$\begin{aligned}\varphi_0^{\mathbb{L}\to}(x_1) &:= \inf_{x_0 \in B_{\mathsf{R}}(x_1)} \varphi_0(x_0) + \mathbb{L}(x_1-x_0) \quad \text{and} \\ \varphi_1^{\leftarrow\mathbb{L}}(x_0) &:= \sup_{x_1 \in B_{\mathsf{R}}(x_0)} \varphi_1(x_1) - \mathbb{L}(x_1-x_0),\end{aligned} \qquad (3.3)$$

where the restriction of the infimum and supremum in (3.3) to the balls $B_{\mathsf{R}}(x_i)$, corresponding to the shifted proper domain of $\mathbb{L}$, is important to avoid the expression "$\infty - \infty$". It will turn out that $\varphi_0^{\mathbb{L}\to}$ is u.s.c. and $\varphi_1^{\leftarrow\mathbb{L}}$ is l.s.c. Of course, these transformations are related by

$$\varphi_0^{\mathbb{L}\to}(x) = -(-\varphi_0)^{\leftarrow\mathbb{L}}(x), \qquad (3.4)$$

and for arbitrary functions $\psi_i : \mathbb{R}^d \to [-\infty, +\infty]$ we have the general relations

$$\psi_0^{\mathbb{L}\to} = \left(\left(\psi_0^{\mathbb{L}\to} \right)^{\leftarrow\mathbb{L}} \right)^{\mathbb{L}\to} \quad \text{and} \quad \psi_1^{\leftarrow\mathbb{L}} = \left(\left(\psi_1^{\leftarrow\mathbb{L}} \right)^{\mathbb{L}\to} \right)^{\leftarrow\mathbb{L}}, \qquad (3.5)$$

see [34, Ch. 5]. For later usage, we consider the following elementary example.

Example 3.1. (Forward and backward $\mathbb{L}$-transform) We consider the potentials

$$\varphi_0(x_0) = \begin{cases} a_0 & \text{for } x_0 = y_0, \\ +\infty & \text{otherwise,} \end{cases} \quad \text{and} \quad \varphi_1(x_1) = \begin{cases} a_1 & \text{for } x_1 = y_1, \\ -\infty & \text{otherwise,} \end{cases}$$

where $-\infty \leqq a_0 < +\infty$, $-\infty < a_1 \leqq +\infty$ and $y_0, y_1 \in \mathbb{R}^d$ are fixed. For $a_0, a_1 \in \mathbb{R}$ we find the transforms

$$\varphi_0^{\mathbb{L}\to}(x_1) = \begin{cases} a_0 + \mathbb{L}(x_1-y_0) & \text{for } x_1 \in B_{\mathsf{R}}(y_0), \\ +\infty & \text{otherwise,} \end{cases}$$
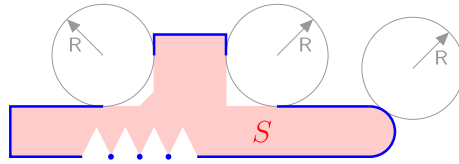
Fig. 2. Visualization of $\mathrm{bdry}_{\mathsf{R}}(S)$ (thick) as subset of the boundary ∂S of the set S (light red)

$$\varphi_1^{\leftarrow \mathbb{L}}(x_0) = \begin{cases} a_1 - \mathbb{L}(y_1 - x_0) & \text{for } x_0 \in B_{\mathsf{R}}(y_1), \\ -\infty & \text{otherwise.} \end{cases}$$

For $a_0 = -\infty$ and $a_1 = +\infty$, we obtain the transforms

$$\varphi_0^{\mathbb{L}\rightarrow}(x_1) = \begin{cases} -\infty & \text{for } x_1 \in B_{\mathsf{R}}(y_0), \\ +\infty & \text{otherwise,} \end{cases} \quad \text{and } \varphi_1^{\leftarrow \mathbb{L}}(x_0) = \begin{cases} +\infty & \text{for } x_0 \in B_{\mathsf{R}}(y_1), \\ -\infty & \text{otherwise,} \end{cases}$$

As $B_{\mathsf{R}}(y_i)$ is open, we see that $\varphi_0^{\mathbb{L}\rightarrow}$ is u.s.c. and $\varphi_1^{\leftarrow \mathbb{L}}$ is l.s.c. Moreover, observe that $\left(\varphi_0^{\mathbb{L}\rightarrow}\right)^{\leftarrow \mathbb{L}} = \varphi_0$ and $\left(\varphi_1^{\leftarrow \mathbb{L}}\right)^{\mathbb{L}\rightarrow} = \varphi_1$, so that (3.5) is true for $\psi_0 \in \left\{\varphi_0, \varphi_1^{\leftarrow \mathbb{L}}\right\}$ and $\psi_1 \in \left\{\varphi_1, \varphi_0^{\mathbb{L}\rightarrow}\right\}$, respectively.

For $\mathsf{R} > 0$ and sets $S \subset \mathbb{R}^d$, we introduce the notation

$$S^{\mathsf{R}} := \left\{ x \in \mathbb{R}^d \, \middle| \, \mathrm{dist}(x, S) < \mathsf{R} \right\},$$

$$\mathrm{ext}_{\mathsf{R}}(S) := \bigcup_{x: \, \mathrm{dist}(x,S) > \mathsf{R}} B_{\mathsf{R}}(x), \quad \mathrm{bdry}_{\mathsf{R}}(S) := \partial S \cap \partial \left(\mathrm{ext}_{\mathsf{R}}(S) \right). \quad (3.6)$$

In particular, $\mathrm{ext}_{\mathsf{R}}(S)$ is the open subset of $\mathbb{R}^d \setminus S$ obtained by taking the union of all the open balls of radius R that do not intersect S. If S is closed and satisfies an exterior sphere condition of radius R at every point of its boundary (e.g. if S is convex) then $\mathrm{ext}_{\mathsf{R}}(S)$ coincides with $\mathbb{R}^d \setminus S$ and $\mathrm{bdry}_{\mathsf{R}}(S) = \partial S$.

In general, $\mathrm{bdry}_{\mathsf{R}}(S)$ is a subset of the boundary of S, precisely made by all points of ∂S satisfying an exterior sphere condition of radius R with respect to S:

$$x \in \mathrm{bdry}_{\mathsf{R}}(S) \iff x \in \partial S \text{ and } \exists \, y \in \mathbb{R}^d : \, |x - y| = \mathsf{R}, \, B_{\mathsf{R}}(y) \cap S = \emptyset. \quad (3.7)$$

In fact, if $x \in \mathrm{bdry}_{\mathsf{R}}(S)$ then there exist sequences x_n, y_n such that $x_n \to x$, $|x_n - y_n| < \mathsf{R}$ and $B_{\mathsf{R}}(y_n) \cap S = \emptyset$. Possibly extracting a subsequence, we can assume that $y_n \to y$, $B_{\mathsf{R}}(y) \cap S = \emptyset$, and $|x - y| \leqq \mathsf{R}$. Since $x \in \partial S$, it is not possible that $|x - y| < \mathsf{R}$, so that the left-to-right implication of (3.7) holds. On the other hand, if $x \in \partial S$, $|x - y| = \mathsf{R}$, and $B_{\mathsf{R}}(y) \cap S = \emptyset$, it is immediate to check that $x \in \partial(\mathrm{ext}_{\mathsf{R}}(S))$, see also Fig. 2.

In Theorem 3.3(2) we will use that for arbitrary sets S the boundary part $\mathrm{bdry}_{\mathsf{R}}(S)$ is countably $(d-1)$-rectifiable, see [34, Th. 10.48(ii)], and hence has $\mathcal{L}^d$ measure 0.

The next result shows how the properties of $\mathbb{L}$ provide regularity of the backward transform $\varphi^{\leftarrow \mathbb{L}}$. Of course, an analogous statement holds for the forward transform using (3.4). The important fact is that the *upper bounds* on the second derivatives of $\mathbb{L}$ generate semi-convexity of φ_0 (i.e. *lower bounds* on $D^2\varphi_0$), see Assertions 5 and 6. As $D^2\mathbb{L}(z)$ blows up at the boundary of $B_{\mathsf{R}}(0)$, it is essential to use the fact that $\mathbb{L}(z_k) \to +\infty$ for $|z_k| \uparrow \mathsf{R}$.

Theorem 3.2. (Regularity of the $\mathbb{L}$-transform) *Let* $\mathbb{L} : \mathbb{R}^d \to [0, +\infty]$ *satisfy*

$$\mathbb{L} : \mathbb{R}^d \to [0, +\infty] \text{ is continuous and } \mathbb{L}(0) = 0, \tag{3.8a}$$

$$\mathbb{L}\big|_{B_{\mathsf{R}}(0)} \in C^2(B_{\mathsf{R}}(0)) \text{ and } \mathbb{L}(z) = +\infty \text{ if } |z| \geqq \mathsf{R}, \tag{3.8b}$$

$$\mathbb{L} \text{ is uniformly convex, i.e. } \exists\, \lambda_* > 0 \,\forall\, z \in B_{\mathsf{R}}(0) : \ D^2\mathbb{L}(z) \geqq \lambda_* I. \tag{3.8c}$$

For an u.s.c. function $\varphi_1 : \mathbb{R}^d \to [-\infty, +\infty]$, *we consider the backward* $\mathbb{L}$-*transform* $\varphi_0 = \varphi_1^{\leftarrow \mathbb{L}}$ *and set*

$$\begin{aligned} &O_0 = \{\varphi_0 > -\infty\}, \quad Q_0 = \{\varphi_0 < +\infty\}, \\ &O_1 = \{\varphi_1 < +\infty\}, \quad Q_1 = \{\varphi_1 > -\infty\}, \quad \text{and} \quad \Omega_0 = O_0 \cap \text{int}(Q_0). \end{aligned} \tag{3.9}$$

Then, the following assertions hold:

(1) *The function* φ_0 *is l.s.c. and satisfies*

$$\inf \varphi_0 \geqq \inf \varphi_1 \quad \text{and} \quad \sup \varphi_0 \leqq \sup \varphi_1, \tag{3.10}$$

$$(Q_1)^{\mathsf{R}} \subset O_0 \quad \text{and} \quad \Big(\varphi_0(x_0) = -\infty \ \Leftrightarrow \ B_{\mathsf{R}}(x_0) \subset \text{ext}_{\mathsf{R}}(Q_1) \subset \{\varphi_1 = -\infty\} \Big). \tag{3.11}$$

　　The sets O_0, O_1, *and* Ω_0 *are open.*

(2) *The set* Q_0 *satisfies an external sphere condition of radius* R, *namely*

$$\mathbb{R}^d \setminus \text{cl}(Q_0) = \text{ext}_{\mathsf{R}}(Q_0) \quad \text{and} \quad \partial Q_0 = \text{bdry}_{\mathsf{R}}(Q_0), \tag{3.12}$$

　　so that the topological boundary of Q_0 *is countably* $(d-1)$-*rectifiable.*

(3) *The "contact set"* $M := M_{-\infty} \cup M_{+\infty} \cup M_{\text{fin}} \subset \mathbb{R}^d \times \mathbb{R}^d$ *defined via*

$$\begin{aligned} M_{\text{fin}} &:= \big\{ (x_0, x_1) \,\big|\, \varphi_i(x_i) \in \mathbb{R}, \ \varphi_1(x_1) = \mathbb{L}(x_0 - x_1) + \varphi_0(x_0) \big\}, \\ M_{-\infty} &:= \big\{ (x_0, x_1) \,\big|\, \varphi_0(x_0) = -\infty, \ |x_1 - x_0| \geqq \mathsf{R} \big\}, \\ M_{+\infty} &:= \big\{ (x_0, x_1) \,\big|\, \varphi_1(x_1) = +\infty, \ |x_1 - x_0| \geqq \mathsf{R} \big\}, \end{aligned} \tag{3.13}$$

　　is closed.

(4) *For every* $\bar{x}_0 \in \Omega_0$, *the section* $M_{0 \to 1}[\bar{x}_0] := \big\{ x_1 \,\big|\, (\bar{x}_0, x_1) \in M_{\text{fin}} \big\}$ *of* M_{fin} *is nonempty, compact, and included in* Q_1. *Moreover, for every compact* $K \subset \Omega_0$ *there exists* $\theta \in (0, \mathsf{R})$ *and* $a', a'' \in \mathbb{R}$ *such that*

$$|x_1 - \bar{x}_0| \leqq \theta \text{ and } a' \leqq \varphi_1(x_1) \leqq a'' \quad \text{whenever } \bar{x}_0 \in K \text{ and } x_1 \in M_{0 \to 1}[\bar{x}_0]. \tag{3.14}$$

(5) *The restriction of* φ_0 *to the open set* Ω_0 *is locally semi-convex, and in particular locally Lipschitz and thus continuous.*

(6) *If $D_0' := \mathrm{dom}(\nabla\varphi_0) \subset \Omega_0$, $D_0'' = \mathrm{dom}(\mathrm{D}^2\varphi_0) \subset D_0'$, then D_0'' has full Lebesgue measure in Ω_0. For every $x \in D_0'$, the set $M_{0\to 1}[x]$ contains a unique point $y = T_{0\to 1}(x)$. The induced map $T_{0\to 1} : D_0' \to \mathbb{R}^d$ is differentiable according to (3.2b) in D_0'' and satisfies the following properties:*

(a) $|x - T_{0\to 1}(x)| < \mathsf{R}$ *and* $\nabla\varphi_0(x) = (\nabla\mathbb{L})(x - T_{0\to 1}(x))$ *for all* $x \in D_0'$,

$$(3.15)$$

(b) $\mathrm{D}^2\varphi_0(x) \geqq -\mathrm{D}^2\mathbb{L}(x - T_0(x))$ *for all* $x \in D_0''$, (3.16)

(c) $\mathrm{D}T_{0\to 1}(x)$ *is diagonalizable with nonnegative eigenvalues on* D_0''.

$$(3.17)$$

Proof. We divide the proof in various steps, corresponding to each assertion.
Assertion (1). To check that φ_0 is l.s.c. we assume $\varphi_0(x_0) > a$ for some $a \in [-\infty, +\infty)$, then there exists $y \in B_{\mathsf{R}}(x_0)$ such that $\varphi_1(y) - \mathbb{L}(y - x_0) > a$. As $\mathbb{L}$ is continuous, we can find $\delta \in (0, \mathsf{R} - |y - x_0|)$ such that $\varphi_1(y) - \mathbb{L}(y - x) > a$ for every $x \in B_\delta(x_0)$. By definition of φ_0 this estimate implies $\varphi_0(x) > a$ on $B_\delta(x_0)$, and lower semi-continuity is shown.

The estimates in (3.10) are elementary following from $\mathbb{L}(0) = 0$ and $\mathbb{L}(z) \geqq 0$, respectively. The relation in (3.11) follows from the fact that $\varphi_0(x_0) = -\infty$ implies $\varphi_1(y) \equiv -\infty$ in $B_{\mathsf{R}}(x_0)$. The openness of O_0 and O_1 follows because φ_0 is l.s.c. and φ_1 is u.s.c. This property in turn implies that $\Omega_0 = O_0 \cap \mathrm{int}(Q_0)$ is open.
Assertion (2). Recalling $Q_0 = \{\varphi_0 < +\infty\}$ it is sufficient to notice that

$$\bar{x} \in \mathbb{R}^d \setminus Q_0 \Leftrightarrow \varphi_0(\bar{x}) = +\infty \Rightarrow \exists\, \bar{y} : |\bar{x} - \bar{y}| \leqq \mathsf{R} \text{ and } \varphi_1(\bar{y}) = +\infty,$$

$$(3.18\mathrm{a})$$

where we used $\mathbb{L} \geq 0$ and the upper semicontinuity of φ_1. However, using $\mathrm{dom}(\mathbb{L}) = B_{\mathsf{R}}(0)$ we obtain

$$\varphi_1(\bar{y}) = +\infty \quad \Rightarrow \quad \varphi_0(x) = +\infty \text{ for all } x \in B_{\mathsf{R}}(\bar{y}).$$ $(3.18\mathrm{b})$

This implication means that if $\bar{x} \in \mathbb{R}^d \setminus Q_0$ then $\bar{x} \in \mathrm{cl}(\mathrm{ext}_{\mathsf{R}}(Q_0))$, so that $\partial Q_0 = \partial(\mathbb{R}^d \setminus Q_0) = \partial\,\mathrm{cl}(\mathrm{ext}_{\mathsf{R}}(Q_0)) = \partial\,\mathrm{ext}_{\mathsf{R}}(Q_0)$.
Assertion (3). The closedness of $M_{\pm\infty}$ follows easily by the semi-continuities of φ_i. For M_{fin} we consider a sequence $(x_{0,n}, x_{1,n}) \in M_{\mathrm{fin}}$ to (x_0, x_1). If $|x_0 - x_1| < \mathsf{R}$, then we have $\varphi_1(x_1) \geqq \mathbb{L}(x_1 - x_0) + \varphi_0(x_0)$ by the semi-continuities. As the opposite inequality is always satisfied, we obtain the equality. We can also exclude that $\varphi_0(x_0) = \varphi_1(x_1) = +\infty$ (resp. $-\infty$), since otherwise $\varphi_0(x) \equiv +\infty$ in $B_{\mathsf{R}}(x_1)$ by (3.18b) which contains a neighborhood of x_0 (resp. $\varphi_1(x) \equiv -\infty$ in $B_{\mathsf{R}}(x_0)$ by (3.11), which contains a neighborhood of x_1), so that $(x_0, x_1) \in M_{\mathrm{fin}}$. If $|x_1 - x_0| \geqq \mathsf{R}$ and (x_0, x_1) does not belong to $M_{-\infty}$ then we have $\liminf_{n\to\infty} \varphi_0(x_{0,n}) \geqq \varphi_0(x_0) > -\infty$ so that

$$\varphi_1(x_1) \geqq \limsup_{n\to\infty} \varphi_1(x_{1,n}) = \limsup_{n\to\infty} \mathbb{L}(x_{1,n} - x_{0,n}) + \varphi_1(x_{0,n}) = +\infty$$

and $(x_0, x_1) \in M_{+\infty}$. Hence, $M = M_{\mathrm{fin}} \cup M_{+\infty} \cup M_{-\infty}$ is closed.

Assertion (4). Let us first show that φ_0 *is locally bounded from above in the interior of* Q_0, *i.e. the open set* $Q_0 \setminus \partial Q_0$. In fact, if a sequence x_n is converging to $\bar{x} \in Q_0 \setminus \partial Q_0$ with $\varphi_0(x_n) \uparrow +\infty$, by arguing as before and using $\varphi_0(x_n) = \sup_{y \in B_R(x_n)} \varphi_1(y) - \mathbb{L}(y - x_n)$, we find $\bar{y} \in \overline{B_R(\bar{x})}$ with $\varphi_1(\bar{y}) = +\infty$. Now (3.18b) gives $\varphi_0(x) = +\infty$ for all $x \in B_R(\bar{y})$, which contradicts the fact that $\varphi_0(x) < +\infty$ in a neighborhood of $\bar{x}$, because of $|\bar{x} - \bar{y}| \leqq R$.

We fix now a compact subset K of the open set Ω_0, a point $\bar{x} \in K$, and consider the section $M_{0 \to 1}[\bar{x}]$ of the contact set M_{fin}. Let $\eta > 0$ be sufficiently small so that $K_\eta := \{ x \in \mathbb{R}^d \mid \mathrm{dist}(x, K) \leqq \eta \} \subset \Omega_0$ and let $\bar{a} := \sup_{K_\eta} \varphi_0$, where $a < +\infty$ by the previous claim. By l.s.c. of φ_0, we also have $\underline{a} := \inf_{K_\eta} \varphi_0 > -\infty$.

By the definition of $\varphi_0 = \varphi_1^{\leftarrow \mathbb{L}}$, for every $\varepsilon \in (0, 1]$ the sets

$$M^\varepsilon(\bar{x}) := \left\{ y \in B_R(\bar{x}) \mid \varphi_1(y) \geqq \mathbb{L}(y - \bar{x}) + \varphi_0(\bar{x}) - \varepsilon \right\}, \qquad (3.19)$$

are non-empty. We choose $y \in M^1(\bar{x})$ and set $x_\vartheta := \vartheta \bar{x} + (1 - \vartheta) y$ with $\vartheta = 1 - \eta/R$, which implies $|x_\vartheta - \bar{x}| \leqq \eta$, and hence $x_\vartheta \in K_\eta$. Moreover, we have $|x_\vartheta - y| \leqq R - \eta$. Therefore, for $y \in M^1(\bar{x}) \subset B_R(\bar{x})$ we find

$$\begin{aligned} \varphi_1(y) &\leqq \mathbb{L}(y - x_\vartheta) + \varphi_0(x_\vartheta) \leqq a'' := \bar{a} + \widehat{\ell}(R - \eta) < \infty, \\ \varphi_1(y) &\geqq \varphi_0(\bar{x}) + \mathbb{L}(y - \bar{x}) - 1 \geqq a' := \underline{a} > -\infty, \end{aligned} \qquad (3.20)$$

where $\widehat{\ell}(\varrho) := \sup_{z \in B_\varrho(0)} \mathbb{L}(z)$. Combining the last two estimates we additionally find

$$\mathbb{L}(y - \bar{x}) \leqq \varphi_1(y) - \varphi_0(\bar{x}) \leqq a'' - \underline{a} =: \widehat{\ell}(\theta) \quad \text{with} \quad \theta \in (0, R). \quad (3.21)$$

Hence, all elements $y \in M^1(\bar{x})$ satisfies $|\bar{x} - y| \leqq \theta$ and (3.20).

We now consider a sequence $y_\varepsilon \in M^\varepsilon(\bar{x}) \subset M^1(\bar{x})$, then a standard compactness argument and the upper semi-continuity of φ_1 show that any limit point $\bar{y}$ is an element of $M_{0 \to 1}[\bar{x}]$, which is therefore not empty. The compactness of $M_{0 \to 1}[\bar{x}]$ and (3.14) again follow by (3.21)

Assertion (5). Let us now fix $\bar{x}_0 \in \Omega_0$ and $\delta > 0$ such that $K := \overline{B_\delta(\bar{x}_0)} \subset \Omega_0$. The previous assertion yields $\theta < R$ and $a', a'' \in \mathbb{R}$ such that $|x' - x| \leqq \theta$ and $a' \leqq \varphi_1(x') \leqq a''$ whenever $x \in K$ and $x' \in M_{0 \to 1}[x]$. By possibly reducing δ, we can also assume that $3\delta + \theta < R$. For every $x \in K$, we now have by construction

$$\varphi_0(x) = \max_{x' \in B_{\delta + \theta}(\bar{x}_0)} \varphi_1(x') - \mathbb{L}(x' - x) \qquad (3.22)$$

which is bounded and semi-convex in K because it is a supremum over a family of uniformly semi-convex functions, where we use $|x' - x| \leqq |x' - \bar{x}_0| + |\bar{x}_0 - x| \leqq 2\delta + \theta$ and that $-\mathbb{L}$ is semi-convex on $\overline{B_{2\delta + \theta}(\bar{x}_0)}$ by (3.8b).

Assertion (6). This assertion follows in the standard way by using the extremality conditions in the contact set, see e.g. [2, Thm. 6.2.4 and 6.2.7]. We give the main argument to show how the assumptions in (3.8) enter. By Alexandrov's theorem and Assertion (5) the set D_0'' has full Lebesgue measure. To obtain the optimality conditions, we fix $x_0 \in Q_0 \cap D_0''$ and know from (3.22) that there exists $\bar{x}_1$ such that $\varphi_0(x_0) = \varphi_1(\bar{x}_1) - \mathbb{L}(\bar{x}_1 - x_0)$. However, for all $x \in B_\delta(x_0)$ we have $\varphi_0(x) +$

$\mathbb{L}(\bar{x}_1 - x) \geqq \varphi_1(\bar{x}_1)$ with equality for $x = x_0$. Thus, we obtain the optimality conditions

$$\nabla\varphi_0(x_0) - \nabla\mathbb{L}(\bar{x}_1 - x_0) = 0 \text{ in } \mathbb{R}^d \quad \text{and} \quad D^2\varphi_0(x_0) + D^2\mathbb{L}(\bar{x}_1 - x_0) \geqq 0 \text{ in } \mathbb{R}^{d \times d}_{\text{sym}}.$$

This result gives the conditions (a) to (c), if we observe that $\bar{x}_1$ is unique. But this property follows from the first optimality condition by using (3.8c) which allows us to write

$$\bar{x}_1 = \boldsymbol{T}_{0 \to 1}(x_0) := x_0 + (\nabla\mathbb{L})^{-1}(\nabla\varphi_0(x_0)),$$

i.e. $\bar{x}_1$ is uniquely determined by x_0. Moreover, $D\boldsymbol{T}_{0 \to 1}(x_0)$ exists and satisfies $D^2\varphi_0(x_0) = (D^2\mathbb{L})(\boldsymbol{T}_{0 \to 1}(x_0) - x_0)(D\boldsymbol{T}_{0 \to 1}(x_0) - I)$, which implies the diagonalization result. □

The previous result can now be applied to the solution of the LET problem in Theorem 2.10 using $\mathbb{L} = \mathbb{L}_1$; thus in this case $\mathsf{R} = \pi/2$. Using the notations for $\text{supp}(\mu_i) = S_i = S'_i + S''_i$ and $\mu_i = \mu'_i + \mu''_i$ from Theorem 2.5 we can compare these to the sets O_i, Q_i, D'_i, and D''_i defined for an optimal pair (φ_0, φ_1) as in Theorem 3.2. So far we constructed optimal pairs (φ_0, φ_1) satisfying

$$\varphi_0 \geqq \varphi_1^{\leftarrow L_1} \text{ on } \mathbb{R}^d, \quad \varphi_0^{L_1 \to} \geqq \varphi_1 \text{ on } \mathbb{R}^d, \quad \varphi_0 = \varphi_1^{\leftarrow L_1} \ \mu_0\text{-a.e.}, \quad \varphi_0^{L_1 \to} = \varphi_1 \ \mu_1\text{-a.e.}$$
(3.23)

However, following [34, Ch. 5], we will show that it is possible to restrict to "tight optimal pairs" satisfying $\varphi_0 = \varphi_1^{\leftarrow L_1}$ and $\varphi_1 = \varphi_0^{L_1 \to}$, which implies that φ_0 is l.s.c. and φ_1 is u.s.c. This possibility leads to the following refinement of the results in [27, Thm. 6.6(iii)].

Theorem 3.3. (Regularity of optimal **HK** potentials) *Let μ_0, μ_1 be nontrivial measures in $\mathcal{M}(\mathbb{R}^d)$ with decompositions given by (2.12)–(2.13).*

(1) *There exists an optimal pair of potentials $\varphi_0, \varphi_1 : \mathbb{R}^d \to [-\infty, +\infty]$ with φ_0 being l.s.c. and φ_1 u.s.c., solving the dual problem of Theorem 2.12 and*

$$\varphi_0 = \varphi_1^{\leftarrow L_1} \quad \text{and} \quad \varphi_1 = \varphi_0^{L_1 \to} \quad \text{on } \mathbb{R}^d, \tag{3.24}$$

$$S_i \subset Q_i, \quad S'_0 \subset S_1^{\pi/2} \subset O_0, \quad S'_1 \subset S_0^{\pi/2} \subset O_1, \tag{3.25}$$
$$\varphi_0 = -\infty \text{ on } S''_0, \quad \text{and} \quad \varphi_1 = +\infty \text{ on } S''_1, \tag{3.26}$$

where the sets O_i and Q_i are as in (3.9).

(2) *If $\boldsymbol{\eta}$ is an optimal solution of the LET problem (2.25), the functions $\sigma_0 := e^{2\varphi_0}$ and $\sigma_1 := e^{-2\varphi_1}$ provide lower semi-continuous representatives of the densities of the marginals $\eta_i = \pi^i_\sharp \boldsymbol{\eta}$ with respect to μ_i, i.e., $\eta_i = \sigma_i \mu_i$, and $\boldsymbol{\eta}$ is concentrated on the contact set M_{fin} so that $\text{supp}(\boldsymbol{\eta}) \subset M$ (see Theorem 3.2). The marginals η_i are concentrated on the open sets O_i. Conversely, if $\widetilde{\boldsymbol{\eta}}$ satisfies $\text{supp}(\widetilde{\boldsymbol{\eta}}) \subset M$ and $\widetilde{\eta}_i = \sigma_i \mu_i$, then $\widetilde{\boldsymbol{\eta}}$ is an optimal solution of the LET problem (2.25).*

(3) *If μ_0 (resp. μ_0') does not charge $(d-1)$-rectifiable sets, e.g. in the case that $\mu_0 \ll \mathcal{L}^d$ or if $\mu_0(\mathrm{bdry}_{\pi/2}(S_0)) = 0$ (resp. $\mu_0'(\mathrm{bdry}_{\pi/2}(S_0)) = 0$), then for every optimal pair (φ_0, φ_1) with $\varphi_0 = \varphi_1^{\leftarrow L_1}$ and φ_1 u.s.c., the measure μ_0 is concentrated on the open set $\mathrm{int}(Q_0)$ (resp. μ_0' is concentrated on the open set Ω_0).*

(4) *If μ_0' is concentrated on $D_0' = \mathrm{dom}(\nabla\varphi_0)$ (in particular if $\mu_0' \ll \mathcal{L}^d$) then the optimal transport plan $\boldsymbol{\eta}$ solving the LET formulation is unique, it is concentrated on $D_0' \times S_0^{\pi/2}$, and it is induced by the graph of $\boldsymbol{T}_{0\to 1}$, i.e. $\boldsymbol{\eta} = (\mathrm{Id}, \boldsymbol{T}_{0\to 1})_\sharp \eta_0$ with $\boldsymbol{T}_{0\to 1}$ from Theorem 3.2(6).*

(5) *If $\mu_0', \mu_1' \ll \mathcal{L}^d$ then μ_0' is concentrated on $D_0'' \cap \boldsymbol{T}_{0\to 1}^{-1}(D_1'')$, where $D_i'' = \mathrm{dom}(D^2\varphi_i)$, and $\boldsymbol{T}_{0\to 1}$ is μ_0'-essentially injective with $\det D\boldsymbol{T}_{0\to 1} > 0$ μ_0-a.e. in D_0''.*

Proof. <u>Assertion (1)</u>. Let (ϕ_0, ϕ_1) be an optimal Borel pair according to Theorem 2.14(2), see (2.43), satisfying

$$\phi_i \in \mathbb{R} \ \mu_i\text{-a.e. in } S_i', \quad \phi_0 = -\infty \ \mu_0\text{-a.e. in } S_0'', \quad \phi_1 = +\infty \ \mu_1\text{-a.e. in } S_1''. \tag{3.27}$$

With this pair, we set $\varphi_0 := \phi_1^{\leftarrow L_1}$, and recalling (3.3) we easily obtain

$$\varphi_0 \leqq \phi_0 \ \text{ in } \mathbb{R}^d, \quad \phi_1(x_1) \leqq L_1(x_1 - x_0) + \varphi_0(x_0) \ \text{ if } x_0, x_1 \in \mathbb{R}^d, \ |x_1 - x_0| < \pi/2. \tag{3.28}$$

Looking at the dual problem (2.34a) with the more general admissible set of Borel pairs as described in (2.36), we see that (φ_0, ϕ_1) is still optimal.

Repeating the argument, we can set $\varphi_1 = \varphi_0^{L_1 \to}$ to find a new optimal pair satisfying $\varphi_1 \geqq \phi_1$. However, exploiting (3.5) we see that the tightness relation (3.24) holds for the optimal pair (φ_0, φ_1). This fact implies that φ_0 is l.s.c. and φ_1 is u.s.c.

By the construction of ϕ_i in Theorem 2.14(2) we have

$$\{\phi_i \in \mathbb{R}\} = S_i', \quad \{\phi_0 = -\infty\} = S_0'', \quad \text{and} \quad \{\phi_1 = +\infty\} = S_1''.$$

Together with $\phi_0 \geqq \varphi_0$ and $\phi_1 \leqq \varphi_1$ we find

$$S_0'' = \{\phi_0 = -\infty\} \subset \{\varphi_0 = -\infty\} \ \text{ and } \ S_0 = \{\phi_0 < +\infty\} \subset \{\varphi_0 < +\infty\} = Q_0,$$
$$S_1'' = \{\phi_1 = +\infty\} \subset \{\varphi_1 = +\infty\} \ \text{ and } \ S_1 = \{\phi_1 > -\infty\} \subset \{\varphi_1 > -\infty\} = Q_1.$$

Clearly, $S_0' = S_0 \cap S_1^{\pi/2} \subset S_1^{\pi/2}$. Moreover, for $x_0 \in S_1^{\pi/2}$ we find $y_1 \in S_1$ with $|y_1 - x_0| < \pi/2$, i.e. $L_1(y_1 - x_0) < \infty$. With this we have $\varphi_0(x_0) = \varphi_1^{\leftarrow L_1}(x_0) \geqq \varphi_1(y_1) - L_1(y_1 - x_0) > -\infty$ and conclude $x_0 \in O_0$. Thus, $S_0' \subset S_1^{\pi/2} \subset O_0$ is shown and $S_1' \subset S_0^{\pi/2} \subset O_1$ follows similarly. Hence, (3.25) and (3.26) are established.

<u>Assertion (2)</u>. The claim follows immediately from Theorem 2.14.

Assertion (3). We just consider the case of μ_0, since the argument for μ_0' is completely analogous and eventually uses the fact that $\Omega_0 = O_0 \cap \text{int}(Q_0)$ and μ_0' is also concentrated on O_0 by (3.25).

By Theorem 3.2 (cf. (3.12)) we know that $\partial Q_0 = \text{bdry}_{\pi/2}(Q_0)$. Since ∂Q_0 is $(d-1)$-rectifiable and μ_0 does not charge $(d-1)$-rectifiable sets, we conclude $\mu_0(\partial Q_0) = 0$.

If $\mu_0(\text{bdry}_{\pi/2}(S_0)) = 0$, we also obtain $\mu_0(\partial Q_0) = 0$ via the following arguments: By (3.25) we have $S_0 \subset Q_0$, which implies that a point $x \in \partial S_0 \cap \text{bdry}_{\pi/2}(Q_0)$ also lies $\text{bdry}_{\pi/2}(S_0)$. Using $\partial Q_0 = \text{bdry}_{\pi/2}(Q_0)$ we obtain $\partial S_0 \cap \partial Q_0 \subset \text{bdry}_{\pi/2}(S_0)$ and find

$$\mu_0(\partial Q_0) \overset{\text{(i)}}{=} \mu_0(\partial Q_0 \cap S_0) \overset{\text{(ii)}}{=} \mu_0(\partial Q_0 \cap \partial S_0) \leqq \mu_0\big(\text{bdry}_{\pi/2}(S_0)\big) = 0,$$

where we used $S_0 = \text{sppt}(\mu_0)$ in $\overset{\text{(i)}}{=}$ and $S_0 \subset Q_0$ in $\overset{\text{(ii)}}{=}$. Thus, we have shown that μ_0 is concentrated on $\text{int}(Q_0)$.

Assertion (4). If $\mu_0' \ll \mathcal{L}^d$ then μ_0' is concentrated on Ω_0 by Claim 3 and $\mu_0(\Omega_0 \setminus D_0') = 0$ by 3.2(6). By the previous claim (2), we know that the first marginal η_0 of $\boldsymbol{\eta}$ is given by $e^{2\varphi_0}\mu_0 = e^{2\varphi_0}\mu_0' = e^{2\varphi_0}\mu_0' \llcorner D_0'$ (in particular $\eta_0(\mathbb{R}^d \setminus D_0') = 0$) so that $\boldsymbol{\eta}$ is concentrated on $M_{\text{fin}} \cap (D_0' \times \mathbb{R}^d)$ which is the graph of the map $\boldsymbol{T}_{0\to1}$ given by Theorem 3.2(6).

Assertion (5). Let us first recall that for $i = 0, 1$ the marginal η_i of $\boldsymbol{\eta}$ and the measure μ_i' are mutually absolutely continuous. Since $\mu_i' \ll \mathcal{L}^d$ we know by Theorem 3.2(6) and the third claim that $\mu_i'(\mathbb{R}^d \setminus D_i'') = \mu_i'(\Omega_i \setminus D_i'') = 0$, so that $\eta_i(\mathbb{R}^d \setminus D_i'') = 0$ and $\eta_0(\boldsymbol{T}_{0\to1}^{-1}(\mathbb{R}^d \setminus D_1'')) = \eta_1(\mathbb{R}^d \setminus D_1'') = 0$; we deduce that η_0 and μ_0' are concentrated on $D_0'' \cap \boldsymbol{T}_{0\to1}^{-1}(D_1'')$.

We can apply Theorem 3.2(6), inverting the order of the pair (φ_0, φ_1) and obtaining that for every $x_1 \in D_1'$ there is a unique element $x_0 \in \mathbb{R}^d$ in the section $M_{1\to0}(x_1)$, i.e. such that $(x_0, x_1) \in M_{\text{fin}}$. This result precisely shows that the restriction of $\boldsymbol{T}_{0\to1}$ to $D_0' \cap \boldsymbol{T}_{0\to1}^{-1}(D_1') \supset D_0'' \cap \boldsymbol{T}_{0\to1}^{-1}(D_1'')$ is injective. Since $(\boldsymbol{T}_{0\to1})_\sharp \eta_0 = \eta_1 \ll \mathcal{L}^d$, we can eventually apply [2, Lemma 5.5.3] which shows that $\det D\boldsymbol{T}_{0\to1} > 0$ μ_0-a.e. in D_0''. $\qquad\square$

It is important to realize that the tightness condition (3.24) is strictly stronger than the optimality conditions (3.23). However, even for tight optimal pairs there is some freedom outside the supports of the measures μ_0 and μ_1, as is seen in the following simple case.

Example 3.4. (**Tight optimal pairs for two Diracs**) This example lies in-between Examples 3.1 and 4.5. For two points $z_0, z_1 \in \mathbb{R}^d$ with $\varrho = |z_1 - z_0| = \pi/3$, such that $\cos_{\pi/2}(\varrho) = 1/2$. We consider two measures $\mu_i = \delta_{z_i}$. With $s_i = S_i' = \{z_i\}$ we easily find the two optimal potential (ϕ_0, ϕ_1) according to Theorem 2.14, see (2.43):

$$\phi_0(x_0) = \begin{cases} -\frac{\log 2}{2} & \text{for } x_0 = z_0, \\ +\infty & \text{otherwise,} \end{cases} \quad \text{and} \quad \phi_1(x_1) = \begin{cases} \frac{\log 2}{2} & \text{for } x_1 = z_1, \\ -\infty & \text{otherwise,} \end{cases}$$

In particular, we have $\phi_1(z_1) - \phi_0(z_0) = \log 2 = \mathrm{L}_1(z_1 - z_0) = \frac{1}{2}\ell(\varrho)$.

Proceeding as in Step 1 of the above proof with $\varphi_0 = \phi_1^{\leftarrow L_1}$ and taking into account the calculations of Example 3.1, we obtain a first tight optimal pair

$$(\varphi_0^{(1)}, \varphi_1^{(1)}) \quad \text{with } \varphi_0^{(1)}(x_0)$$

$$= \begin{cases} \frac{\log 2}{2} - L_1(z_1 - x_0) & \text{for } x_0 \in B_{\pi/2}(z_1), \\ -\infty & \text{otherwise,} \end{cases} \quad \text{and } \varphi_1^{(1)} = \phi_1.$$

Interchanging the roles of ϕ_0 and ϕ_1 we arrive at a second tight optimal pair

$$(\varphi_0^{(2)}, \varphi_1^{(2)}) \quad \text{with } \varphi_0^{(2)} = \phi_0 \quad \text{and}$$

$$\varphi_1^{(2)}(x_1) = \begin{cases} -\frac{\log 2}{2} + L_1(x_1 - z_0) & \text{for } x_1 \in B_{\pi/2}(z_0), \\ \infty & \text{otherwise.} \end{cases}$$

A third case is obtained by choosing $z_1^* \neq z_1$ and considering an optimal pair $(\phi_0, \widetilde{\phi}_1)$ with ϕ_0 from above and

$$\widetilde{\phi}_1(x_1) = \begin{cases} \frac{\log 2}{2} & \text{for } x_1 = z_1, \\ a_1 & \text{for } x_1 = z_1^*, \quad \text{where } a_1 \lessgtr -\frac{\log 2}{2} + L_1(z_1^* - z_0). \\ -\infty & \text{otherwise.} \end{cases}$$

We obtain $\varphi_0^{(3)} : x_0 \mapsto \max\{\frac{\log 2}{2} - L_1(z_1 - x_0), a_1 - L_1(z_1^* - x_0)\}$ and the tight optimal pair $\left(\varphi_0^{(3)}, (\varphi_0^{(3)})^{L_1 \rightarrow}\right)$.

With the notation of Theorem 3.2 we have $O_0^{(3)} = \{\varphi_0^{(3)} > -\infty\} = B_{\pi/2}(z_1) \cup B_{\pi/2}(z_1^*) = \widetilde{Q}_1^{\pi/2}$, since $\widetilde{Q}_1 = \{\widetilde{\phi}_1 > -\infty\} = \{z_1, z_1^*\}$, i.e. (3.11) holds. Because of $Q_0^{(3)} = \{\varphi_1^{(3)} < +\infty\} = \mathbb{R}^2$, also (3.12) is true.

The following corollary shows that in the case of an absolutely continuous reduced pair (μ_0, μ_1) the density of μ_1 can be written in terms of the optimal pair (σ_0, σ_1), the transport map T, and the density of μ_0, and vice versa:

Corollary 3.5. (Monge solutions) *Let $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^d)^2$ with $\mu_1'' = 0$, and let (φ_0, φ_1) be a tight optimal pair of potentials according to Theorem 3.3. If μ_0' is concentrated on $D_0' = \mathrm{dom}(\nabla \varphi_0)$ (in particular if $\mu_0' \ll \mathcal{L}^d$), then there exists a "unique" (up to μ_0-negligible sets) optimal transport-growth pair (T, q) attaining the minimum for the Monge Problem 2.9, namely*

$$(T, q)_\star \mu_0 = \mu_1 \quad \text{and} \quad \mathcal{C}(q, T; \mu_0) = \mathsf{HK}^2(\mu_0, \mu_1). \tag{3.29}$$

If $\sigma_i, \varphi_i, D_i', D_i'', \eta, T_{0\to1}, T_{1\to0}$ are given as in Theorem 3.2 and 3.3, the pair (T, q) can be obtained in the following way:

(1) *The restriction of T to D_0' coincides with the map $T_{0\to1}$ (and the plan η) as in Theorem 3.2, whereas $T(x) := x$ for every $x \in \mathbb{R}^d \setminus D_0'$ (in particular in S_0'').*
(2) *$q(x) \equiv 0$ for $x \in \mathbb{R}^d \setminus D_0'$ (in particular in S_0'') and*

$$q^2(x) = \frac{\sigma_0(x)}{\sigma_1(T_{0\to1}(x))} = \sigma_0^2(x) + \frac{1}{4}|\nabla \sigma_0(x)|^2 \text{ for } x \in D_0'. \tag{3.30}$$

Moreover, $\boldsymbol{T}$ *satisfies*

$$|\boldsymbol{T}(x)-x| < \pi/2 \quad \text{and} \quad \sigma_0(x)\sigma_1(\boldsymbol{T}(x)) = \cos(|x-\boldsymbol{T}(x)|)^2 \quad \text{in } D_0'. \quad (3.31)$$

If $\mu_0 \ll \mathcal{L}^d$, *then* $\mu_1 \ll \mathcal{L}^d$ *if and only if* $\det \mathrm{D}\boldsymbol{T}(x) > 0$ *for* μ_0-*a.e.* $x \in D_0''$. *In this case, setting* $\mu_i = c_i \mathcal{L}^d \ll \mathcal{L}^d$ *we have*

$$c_1 = \left(c_0 \frac{q^2}{\det \mathrm{D}\boldsymbol{T}}\right) \circ \boldsymbol{T}^{-1} \quad \mathcal{L}^d\text{-a.e. in } \boldsymbol{T}(D_0'') \subset \Omega_1. \quad (3.32)$$

To obtain the second identity in (3.30), we exploit the first-order optimality (3.15) and $\sigma_0 = \mathrm{e}^{2\varphi_0}$ giving $\frac{1}{2\sigma_0}\nabla\sigma_0 = \nabla\varphi_0(x) = \tan(x-\boldsymbol{T}(x))$ by (3.15). Thus, using the optimality condition (3.31) (coming from (2.42e)) we find

$$q^2(x) = \frac{\sigma_0^2(x)}{\cos^2(|x-\boldsymbol{T}(x)|)} = \sigma_0^2(x)(1+\tan^2(|x-\boldsymbol{T}(x)|)) = \sigma_0^2(x) + \frac{1}{4}|\nabla\sigma_0(x)|^2. \quad (3.33)$$

We can also rephrase the above results in terms of the optimal Kantorovich potential ξ_0 in (2.34b). This potential, which satisfies the relations $\xi_0 = \frac{1}{2}(\sigma_0-1) = \mathsf{G}_1(\varphi_0) = \frac{1}{2}(\mathrm{e}^{2\varphi_0}-1)$, will be the best choice for characterizing the densities of the Hellinger–Kantorovich geodesic curves. Indeed, the transport map $\boldsymbol{T}$ on D_0'' takes the form

$$\boldsymbol{T}(x) = x + \arctan\left(\frac{\nabla\xi_0(x)}{1+2\xi_0(x)}\right) = x + \arctan\left(\nabla\varphi_0(x)\right),$$
$$q^2(x) = (1+2\xi_0(x))^2 + |\nabla\xi_0(x)|^2. \quad (3.34)$$

If μ_0, μ_1 have full support $S_0 = S_1 = \mathbb{R}^d$, then Theorem 3.3 immediately yields $\Omega_i = O_i \cap \mathrm{int}(Q_i) = \mathbb{R}^d$, so that φ_0 and φ_1 take values in $\mathbb{R}$, are locally Lipschitz, and locally semi-convex and semi-concave, respectively. Another important case where the properties of φ_0, φ_1 can be considerably refined is when μ_0, μ_1 are *strongly reduced* (cf. Definition 2.6) and have compact support.

Theorem 3.6. (Improved regularity in case of strongly reduced pairs) *Let us assume that the supports* S_0, S_1 *of* μ_0, μ_1 *are compact and satisfy* $S_i \subset S_{1-i}^{\pi/2}$, *so that* μ_0, μ_1 *is a strongly reduced pair (cf. Definition 2.6). Then it is possible to find a pair of optimal potentials* φ_0, φ_1 *as in Theorem 3.3 satisfying the following additional properties:*

(1) φ_i *are uniformly bounded (in particular* $\Omega_i = \mathbb{R}^d$ *and* $M = M_{\mathrm{fin}}$): *there exist constants* $\phi_{\min} < \phi_{\max} \in \mathbb{R}$ *such that*

$$\phi_{\min} \leqq \varphi_i \leqq \phi_{\max} \quad \text{in } \mathbb{R}^d. \quad (3.35)$$

(2) *If* $\theta \in [0, \pi/2[$ *satisfies* $\cos^2(\theta) = \mathrm{e}^{2(\phi_{\min}-\phi_{\max})}$ *then for every* $x_0, x_1 \in \mathbb{R}^d$

$$(x_0, x_1) \in M \quad \Rightarrow \quad |x_1-x_0| \leqq \theta. \quad (3.36)$$

(3) φ_i *are Lipschitz,* φ_0 *is semi-convex,* φ_1 *is semi-concave.*

Proof. Assertion (1). Let φ_0', φ_1' be an optimal pair as in Theorem 3.3. Since φ_1' is u.s.c and $\varphi_1' < +\infty$ on $S_0^{\pi/2}$, we have $\phi_{\max} := \max_{S_1} \varphi_1' < +\infty$. We can then define $\zeta_1 := \min\{\varphi_1', \phi_{\max}\}$ observing that $\zeta_1 \leqq \phi_{\max}$ in $\mathbb{R}^d$ and (φ_0', ζ_1) is still optimal since $\zeta_1 = \varphi_1'$ on S_1.

Arguing as in the proof of Theorem 3.3, we define $\zeta_0 := (\zeta_1)^{\leftarrow L_1}$, observing that $\zeta_0 \leqq \phi_{\max}$ as well. On the other hand, ζ_0 is l.s.c. and $\zeta_0 > -\infty$ on $S_1^{\pi/2} \supset S_0$, so that $\phi_{\min} := \min_{S_0} \zeta_0' > -\infty$. Setting $\zeta_0' := \max\{\zeta_0, \phi_{\min}\}$ we obtain a new optimal pair (ζ_0', ζ_1) with $\phi_{\min} \leqq \zeta_0 \leqq \phi_{\max}$. Hence, with $\zeta_1' := (\zeta_0')^{L_1 \rightarrow}$ we get the desired optimal pair (ζ_0', ζ_1') satisfying $\phi_{\min} \leqq \zeta_i' \leqq \phi_{\max}$ as well.

Assertion (2). This assertion is now an easy consequence of the definition of contact set (3.13) and the fact that $\varphi_1(x_1) - \varphi_0(x_0) \leqq \phi_{\max} - \phi_{\min}$. Assertion (3). The last assertion follows as Theorem 3.2(5). $\quad\square$

4. Dynamic Duality and Regularity Properties of the Hamilton–Jacobi Equation

In the previous section, the regularity properties of the optimal HK pairs (φ_0, φ_1) were studied, which can be understood via the static formulations of HK as only the measures μ_0 and μ_1 are involved. Now, we consider the dual potentials $\xi_t(x) = \xi(t, x)$ along geodesics $(\mu_t)_{t \in [0,1]}$. At this stage, the present Section 4 is completely independent of the previous Section 3. Only in the upcoming Section 5, we will combine the two results to derive the finer regularity properties of the geodesics μ_t.

In [27, Sect. 8.4], it is shown that the optimal dual potentials ξ in the dynamic formulation in (2.3) (but now for $\alpha = 1$ and $\beta = 4$) are subsolutions to a suitable Hamilton–Jacobi equation, namely,

$$\frac{1}{2\tau}\mathsf{HK}(\mu_0, \mu_\tau)^2 = \sup\left\{\int_{\mathbb{R}^d} \xi(\tau, \cdot)\,\mathrm{d}\mu_\tau - \int_{\mathbb{R}^d} \xi(0, \cdot)\,\mathrm{d}\mu_0 \,\Big|\, \xi \in C_c^\infty([0, \tau] \times \mathbb{R}^d),\right.$$
$$\left.\frac{\partial}{\partial t}\xi + \frac{1}{2}|\nabla\xi|^2 + 2\xi^2 \leqq 0 \quad \text{in } [0, \tau] \times \mathbb{R}^d\right\}. \tag{4.1}$$

Theorem 8.11 in [27] shows that the maximal subsolutions of the generalized Hamilton–Jacobi equation (2.4) for $t \in (0, \tau)$ are given by the following *generalized Hopf–Lax formula*

$$\xi_t(x) = \xi(t, x) = (\mathscr{P}_t \xi_0)(x) = \frac{1}{t}\mathscr{A}\left(t\xi_0(\cdot)\right)(x) = \inf_{y \in \mathbb{R}^d} \frac{1}{2t}\left(1 - \frac{\cos_{\pi/2}^2(|x-y|)}{1 + 2t\xi_0(y)}\right), \tag{4.2}$$

where $\xi_0 \in C^1(\mathbb{R}^d)$ is fixed and such that $\inf_{\mathbb{R}^d} \xi_0(\cdot) > -\frac{1}{2\tau}$, compare with (1.13).

In the spirit of the previous section, it is possible to derive some semi-concavity properties of ξ_t from this formula. However, these are not enough as we need more precise second order differentiability. To obtain the latter, we use the fact that a geodesic curve is not oriented, meaning that $t \mapsto \mu_{1-t}$ is still a geodesic, or in other words that $t \mapsto \xi_{1-t}$ has to also solve a Hamilton–Jacobi equation. Thus, our

strategy will be the following: For an optimal pair $(\xi_0, \bar{\xi}_1)$ in (2.34b), we construct a forward solution ξ_t starting from ξ_0 and backward solutions starting from $\bar{\xi}_1$ via

$$\xi_t = \mathscr{P}_t \xi_0 \text{ for } t \in (0, 1] \quad \text{and} \quad \bar{\xi}_t = \mathscr{R}_t \bar{\xi}_1 := -\mathscr{P}_t(-\bar{\xi}_1) \text{ for } t \in [0, 1). \quad (4.3)$$

In Section 5, optimality will be used to guarantee that ξ_t and $\bar{\xi}_t$ are essentially the same so that semi-concavity of ξ_t and semi-convexity of $\bar{\xi}_t$ provide the desired smoothness.

4.1. Exploiting the Generalized Hopf–Lax Formula for Regularity

In this section, we study in detail the regularity properties of the function ξ_t arising in (4.2). Assuming that $\inf_{x \in \mathbb{R}^d} \xi_0(x) \geqq -\frac{1}{2\tau}$ we see that $\mathscr{P}_t \xi_0$ is well-defined for $t \in (0, \tau)$ and can be equivalently characterized by

$$\left(\mathscr{P}_t \xi_0\right)(x) = \inf_{y \in B_{\pi/2}(x)} \frac{1}{2t} \left(1 - \frac{\cos^2_{\pi/2}(|x-y|)}{1 + 2t\xi_0(y)}\right). \quad (4.4)$$

We can extend (4.4) at $t = \tau$ if we define the quotients $a/0 := +\infty$, $a/(+\infty) := 0$ for every $a > 0$. Moreover, since $t \mapsto \mathscr{P}_t \xi_0(x)$ is decreasing, we easily get

$$\xi_t(x) = \left(\mathscr{P}_t \xi_0\right)(x) = \lim_{s \uparrow t} \left(\mathscr{P}_s \xi_0\right)(x) \quad \text{for every } x \in \mathbb{R}^d, \ t \in (0, \tau] \quad (4.5)$$

so that many properties concerning the limiting case $t = \tau$ can be easily derived by continuity as $t \uparrow \tau$.

If ξ_0 is l.s.c. and $\left(\mathscr{P}_t \xi_0\right)(x) < \frac{1}{2t}$, the infimum in (4.4) it attained at a compact set denoted by

$$\mathsf{M}_t \xi_0(x) := \operatorname{argmin}_y \frac{1}{2t} \left(1 - \frac{\cos^2_{\pi/2}(|x-y|)}{1 + 2t\xi_0(y)}\right) \subset B_{\pi/2}(x). \quad (4.6)$$

Notice that $\left(\mathscr{P}_t \xi_0\right)(x) = \frac{1}{2t}$ only if ξ_0 is identically $+\infty$ in $B_{\pi/2}(x)$ and in this case any element of $\overline{B_{\pi/2}(x)}$ is a minimizer. For later usage we also define $\mathsf{M}_0 \xi(x) = \{x\}$.

We also observe that if $\xi_0(x) = a$ is constant then $\mathscr{P}_t \xi_0$ is constant in x, namely

$$\mathscr{P}_t \xi_0(x) = \mathscr{P}_t a(x) = P_a(t) := \frac{a}{1 + 2at}, \quad \text{with } P_\infty(t) := \frac{1}{2t}. \quad (4.7)$$

A crucial property of (4.2) is the link with the classical Hopf–Lax formula on the cone $\mathfrak{C}$ for a function $\zeta : \mathfrak{C} \to \mathbb{R}$ satisfying $\zeta([x, r]) \geqq -\frac{1}{2\tau} r^2$. For $t \in (0, \tau)$ the Hopf–Lax formula on $\mathfrak{C}$ reads as

$$\mathscr{Q}_t \zeta([x, r]) := \inf_{[x', r'] \in \mathfrak{C}} \zeta([x', r']) + \frac{1}{2t} \mathsf{d}^2_{\mathfrak{C}}([x, r], [x', r']). \quad (4.8)$$

For ξ_0 satisfying $\xi_0 \geqq -\frac{1}{2\tau}$ and $t \in (0, \tau)$ we set $\zeta([x, r]) := \xi_0(x)r^2$ and find (cf. [27, Thm. 8.11])

$$\xi_t = \mathscr{P}_t \xi_0 \quad \Longleftrightarrow \quad \xi_t(x)r^2 = \mathscr{Q}_t \zeta([x, r]) \quad \text{for all } x \in \mathbb{R}^d; \quad (4.9)$$

Moreover, if ξ_0 is lower semi-continuous the infimum in (4.8) is attained and we have

$$\xi_t(x)r^2 = \zeta([x', r']) + \frac{1}{2t}\mathsf{d}_{\mathfrak{C}}^2([x, r], [x', r'])$$

$$\Longleftrightarrow \begin{cases} x' \in \mathsf{M}_t\xi_0(x) \text{ and} \\ (1+2t\xi_0(x'))(r')^2 = (1-2t\xi_t(x))r^2 \end{cases} \tag{4.10}$$

(where $[x', r'] = \mathfrak{o}$ if $r' = 0$, corresponding to the case $1-2t\xi_t(x) = 0$). From (4.8) and (4.9) we also deduce the estimate

$$\left(1-2t\xi_t(x)\right)r^2 + \left(1+2t\xi_0(x')\right)(r')^2 \geqq 2rr'\cos_{\pi/2}(|x-x'|) \tag{4.11}$$

for every $x, x' \in \mathbb{R}^d$ and $r, r' \geqq 0$. Optimizing with respect to r, r' we find

$$\left(1-2t\xi_t(x)\right)\left(1+2t\xi_0(x')\right) \geqq \cos_{\pi/2}^2\left(|x-x'|\right) \quad \text{for every } x, x' \in \mathbb{R}^d \tag{4.12}$$

and arrive at the following characterization: For all $x \in \mathbb{R}^d$ with $1-2t\xi_t(x) > 0$ we have

$$x' \in \mathsf{M}_t\xi_0(x) \quad \Longleftrightarrow \quad \left(1-2t\xi_t(x)\right)\left(1+2t\xi_0(x')\right) = \cos_{\pi/2}^2\left(|x-x'|\right). \tag{4.13}$$

To treat the factor of r and r' in (4.11) efficiently, we define the function

$$Z_t(u', u) := \frac{1-2tu}{1+2tu'} \quad \text{for } 1+2tu', \ 1-2tu \geqq 0 \quad \text{and } Z_t(+\infty, u) \equiv 0. \tag{4.14}$$

Using (4.13), the optimal r' in (4.10) can now be equivalently characterized by

$$(r')^2 = Z_t(\xi_0(x'), \xi_t(x)) = \frac{(1-2t\xi_t(x))^2}{\cos_{\pi/2}^2(|x-x'|)} = (1-2t\xi_t(x))^2(1 + \tan^2(|x-x'|)). \tag{4.15}$$

The following result collects the properties of $\mathscr{P}_t$ that will be needed in the sequel:

Proposition 4.1. (Properties of the generalized Hopf–Lax operator $\mathscr{P}_t$) *Let $\xi_0 : \mathbb{R}^d \to [a, b]$ with $-1/2 \leqq a \leqq b \leqq +\infty$ be lower semi-continuous and set $\xi_t := \mathscr{P}_t\xi_0$ for $t \in [0, 1]$.*

(1) *Lower/upper bounds. The functions ξ_t are well defined and satisfy (cf. (4.7) for P_a)*

$$-\frac{1}{2(1-t)} \leqq P_a(t) \leqq \xi_t \leqq P_b(t) \leqq \frac{1}{2t} \quad \text{for every } t \in (0, 1), \ x \in \mathbb{R}^d. \tag{4.16}$$

Moreover, it holds that

$$\xi_0(x) = -1/2 \quad \Leftrightarrow \quad \xi_t(x) = -\frac{1}{2(1-t)}. \tag{4.17}$$

(2) Semi-concavity. *Setting* $\Lambda_a(t) := \frac{1}{t(1+2at)} \leq \frac{1}{t(1-t)}$ *the functions* ξ_t *are* $\Lambda_a(t)$-*Lipschitz and* $\Lambda_a(t)$ *semi-concave, i.e.* $x \mapsto \xi_t(x) - \frac{\Lambda_a(t)}{2}|x|^2$ *is concave.*

(3) Semigroup property. *For every* $0 \leq s < t \leq 1$ *we have*

$$\xi_t = \mathscr{P}_{t-s}\xi_s \tag{4.18}$$

(4) Concatenation of optimal points. *For* s, t *with* $0 \leq s < t < 1$ *and* $x \in \mathbb{R}^d$ *we define the set-valued function* $\mathsf{M}_{t \to s}$ *via* $\mathsf{M}_{t \to s}(x) := \mathsf{M}_{t-s}\xi_s(x)$. *For all* $0 \leq t_0 < t_1 < t_2 < 1$ *and all* $x_0, x_1, x_2 \in \mathbb{R}^d$ *we have:*

If $x_1 \in \mathsf{M}_{t_2 \to t_1}(x_2)$ *and* $x_0 \in \mathsf{M}_{t_1 \to t_0}(x_1)$, *then* $x_0 \in \mathsf{M}_{t_2 \to t_0}(x_2)$ *and*

$$Z_{t_2-t_0}(\xi_{t_0}(x_0), \xi_{t_2}(x_2)) = Z_{t_1-t_0}(\xi_{t_0}(x_0), \xi_{t_1}(x_1))\, Z_{t_2-t_1}(\xi_{t_1}(x_1), \xi_{t_2}(x_2)). \tag{4.19}$$

(5) Geodesics on $\mathfrak{C}$. *If* $0 \leq t_0 < t_1 < t_2 < 1$, $x_0 \in \mathsf{M}_{t_2 \to t_0}(x_2)$, $r_0 = Z_{t_2-t_0}(\xi_{t_0}(x_0), \xi_{t_2}(x))r_2$, *and* $[x_1, r_1] = \mathrm{geo}_\theta\big([x_0, r_0], [x_2, r_2]\big)$ *for* $\theta = \frac{t_1-t_0}{t_2-t_0}$, *then* $x_1 \in \mathsf{M}_{t_2 \to t_1}(x_2)$.

(6) Characterization of optimality. *For all* $x, y \in \mathbb{R}^d$ *and* $0 \leq s < t < 1$ *with* $\tau := t - s$ *we have*

$$(1 - 2\tau\xi_t(x))(1 + 2\tau\xi_s(y)) \geq \cos^2_{\pi/2}(|x-y|), \tag{4.20}$$

$$y \in \mathsf{M}_{t \to s}(x),\ \xi_t(x) < \frac{1}{2\tau} \ \Leftrightarrow\ (1 - 2\tau\xi_t(x))(1 + 2\tau\xi_s(y)) = \cos^2_{\pi/2}(|x-y|). \tag{4.21}$$

Proof. Assertion (1). The first assertion follows by the monotonicity property of $\mathscr{P}_t$ and (4.7). Note that (4.17) is a simple consequence of the property

$$\frac{1}{2t}\left(1 - \frac{\cos^2_{\pi/2}(|x-y|)}{1 + 2t\xi(y)}\right) \geq -\frac{1}{2(1-t)}$$

with equality if and only if $x = y$ and $\xi(y) = -1/2$.

Assertion (2). It is sufficient to observe that for every $y \in \mathbb{R}^d$

$$x \mapsto \cos^2_{\pi/2}(|x-y|) \text{ is } 2\text{-Lipschitz}, \quad x \mapsto \cos^2_{\pi/2}(|x-y|) - |x|^2 \text{ is concave}, \tag{4.22}$$

so that

$$x \mapsto \frac{1}{2t}\left(1 - \frac{\cos^2_{\pi/2}(|x-y|)}{1 + 2t\xi_0(y)}\right) \text{ is } \Lambda_a(t)\text{-Lipschitz} \tag{4.23}$$

and

$$x \mapsto \frac{1}{2t}\left(1 - \frac{\cos^2_{\pi/2}(|x-y|)}{1 + 2t\xi_0(y)}\right) - \frac{\Lambda_a(t)}{2}|x|^2 \text{ is concave.} \tag{4.24}$$

Assertion (3). If $t < 1$ the semigroup property for $\mathscr{P}_t$ can be derived by the link with the Hopf-Lax semigroup in $\mathfrak{C}$ given by (4.9) and the fact that $(\mathfrak{C}, \mathsf{d}_{\pi,\mathfrak{C}})$ is a geodesic space. The case $t = 1$ follows by approximation and (4.5).

<u>Assertion (4).</u> We set $\tau_0 := t_1 - t_0$, $\tau_1 := t_2 - t_1$, $r > 0$,

$$r_1 = Z_{\tau_1}(\xi_{t_1}(x_1), \xi_{t_2}(x))r, \quad r_0 = Z_{\tau_0}(\xi_{t_0}(x_0), \xi_{t_1}(x_1))r_1$$

and use (4.10) and (4.18):

$$
\begin{aligned}
\xi_{t_2}(x)r^2 &= \xi_{t_1}r_1^2 + \frac{1}{2\tau_1}\mathsf{d}_{\pi,\mathfrak{C}}^2([x_1, r_1], [x, r]) \\
&= \xi_{t_0}r_0^2 + \frac{1}{2\tau_0}\mathsf{d}_{\pi,\mathfrak{C}}^2([x_0, r_0], [x_1, r_1]) + \frac{1}{2\tau_1}\mathsf{d}_{\pi,\mathfrak{C}}^2([x_1, r_1], [x, r]).
\end{aligned}
$$

On the other hand,

$$\xi_{t_2}(x)r^2 \leqq \xi_{t_0}r_0^2 + \frac{1}{2\tau}\mathsf{d}_{\pi,\mathfrak{C}}^2([x_0, r_0], [x, r]), \tag{4.25}$$

so that we obtain

$$\frac{1}{2\tau_0}\mathsf{d}_{\pi,\mathfrak{C}}^2([x_0, r_0], [x_1, r_1]) + \frac{1}{2\tau_1}\mathsf{d}_{\pi,\mathfrak{C}}^2([x_1, r_1], [x, r]) \leqq \frac{1}{2\tau}\mathsf{d}_{\pi,\mathfrak{C}}^2([x_0, r_0], [x, r]); \tag{4.26}$$

since $\tau = \tau_0 + \tau_1$ the opposite inequality always hold in (4.26), and we deduce the equality, which implies that the equality holds in (4.25) as well, showing (4.19) thanks to (4.10).

<u>Assertion (5).</u> We can argue as in the previous assertion, starting from the characterization of x_0, r_0

$$\xi_{t_2}(x)r^2 = \xi_{t_0}r_0^2 + \frac{1}{2\tau}\mathsf{d}_{\pi,\mathfrak{C}}^2([x_0, r_0], [x, r]) \tag{4.27}$$

and using the identity along the geodesic in $\mathfrak{C}$ connecting $[x_0, r_0]$ to $[x, r]$, namely

$$\frac{1}{2\tau_0}\mathsf{d}_{\pi,\mathfrak{C}}^2([x_0, r_0], [x_1, r_1]) + \frac{1}{2\tau_1}\mathsf{d}_{\pi,\mathfrak{C}}^2([x_1, r_1], [x, r]) = \frac{1}{2\tau}\mathsf{d}_{\pi,\mathfrak{C}}^2([x_0, r_0], [x, r]). \tag{4.28}$$

<u>Assertion (6).</u> The final assertion follows from (4.12) and (4.13). $\qquad\square$

4.2. Backward Generalized Hopf–Lax Flow and Contact Sets

Let us now consider the backward version of the generalized Hopf–Lax semigroup. By the simple structure of the generalized Hamilton–Jacobi equation (2.4), we immediately see that time reversal leads to the same effect as the sign reversal $\xi \rightsquigarrow -\xi$. Hence, the backward semigroup $\mathscr{R}_t$ is defined for $\bar{\xi}$ with $\bar{\xi} \leqq 1/(2\tau)$ via

$$\mathscr{R}_t\bar{\xi} := -\mathscr{P}_t(-\bar{\xi}) \quad \text{for } t \in (0, \tau]. \tag{4.29}$$

The corresponding properties of $\mathscr{R}_t$ follow easily from Proposition 4.1, but observe that we use $\bar{\xi}_t = \mathscr{R}_{1-t}\bar{\xi}_1$ to go backward in time.

Corollary 4.2. (Properties of $\mathscr{R}_t$) *Let* $\bar{\xi}_1 : \mathbb{R}^d \to [-\bar{b}, -\bar{a}]$ *with* $-\infty \leqq -\bar{b} \leqq -\bar{a} \leqq 1/2$ *be upper semi-continuous and set*

$$\bar{\xi}_t := \mathscr{R}_{1-t}\bar{\xi}_1 \quad for \, t \in [0, 1]. \tag{4.30}$$

(1) *Lower/upper bounds.* The functions $\bar{\xi}_t$ are well-defined and satisfy

$$-\frac{1}{2t} \leqq P_{\bar{b}}(1-t) \leqq \bar{\xi}_t \leqq P_{\bar{a}}(1-t) \leqq \frac{1}{2(1-t)} \quad \text{for all } t \in (0, 1), \; x \in \mathbb{R}^d. \tag{4.31}$$

Moreover, we have the equivalence

$$\bar{\xi}_1(x) = 1/2 \quad \Leftrightarrow \quad \bar{\xi}_t(x) = \frac{1}{2t}. \tag{4.32}$$

(2) *Semi-convexity.* The functions $\bar{\xi}_t$ are $\Lambda_{\bar{a}}(1-t)$-Lipschitz and $\Lambda_{\bar{a}}(1-t)$ semi-convex, i.e. $x \mapsto \bar{\xi}_t(x) + \frac{\Lambda_{\bar{a}}(1-t)}{2}|x|^2$ is convex (cf. Proposition 4.1(2) for Λ_a).

(3) *Time-reversed semigroup property.* For every $0 \leqq s < t \leqq 1$ we have

$$\bar{\xi}_s = \mathscr{R}_{t-s}\bar{\xi}_t. \tag{4.33}$$

(4) *Concatenation of optimal points.* Setting $\overline{\mathsf{M}}_{s\leftarrow t}(x) := \mathsf{M}_{t-s}(-\bar{\xi}_t)(x)$ for every $0 < s < t \leqq 1$ and $x \in \mathbb{R}^d$, the set-valued function $\overline{\mathsf{M}}_{s\leftarrow t}$ satisfies the concatenation property for $0 < t_0 < t_1 < t_2 \leqq 1$ and $x_0, x_1, x_2 \in \mathbb{R}^d$:

If $x_1 \in \overline{\mathsf{M}}_{t_0\leftarrow t_1}(x_0)$ and $x_2 \in \overline{\mathsf{M}}_{t_1\leftarrow t_2}(x_1)$, then $x_2 \in \overline{\mathsf{M}}_{t_0\leftarrow t_2}(x_0)$ and

$$Z_{t_2-t_0}(\bar{\xi}_{t_0}(x_0), \bar{\xi}_{t_2}(x_2)) = Z_{t_1-t_0}(\bar{\xi}_{t_0}(x_0), \bar{\xi}_{t_1}(x_1)) \cdot Z_{t_2-t_1}(\bar{\xi}_{t_1}(x_1), \bar{\xi}_{t_2}(x_2)). \tag{4.34}$$

(5) *Characterization of optimality.* For all $x, y \in \mathbb{R}^d$ and $0 < s < t \leqq 1$ with $\tau := t - s$

$$(1-2\tau\bar{\xi}_t(x))\,(1+2\tau\bar{\xi}_s(y)) \geqq \cos^2_{\pi/2}(|x-y|), \tag{4.35}$$

$$x \in \mathsf{M}_{s\to t}(y), \; \bar{\xi}_s(y) > -\frac{1}{2\tau} \; \Leftrightarrow \; (1-2\tau\bar{\xi}_t(x))\,(1+2\tau\bar{\xi}_s(y)) = \cos^2_{\pi/2}(|x-y|). \tag{4.36}$$

Proof. We just observe that the second statement in (4.34) follows by the corresponding statement in (4.19) which now reads as

$$Z_{t_2-t_0}(-\bar{\xi}_{t_2}(x_2), -\bar{\xi}_{t_0}(x_0)) = Z_{t_1-t_0}(-\bar{\xi}_{t_1}(x_1), -\bar{\xi}_{t_0}(x_0)) \cdot Z_{t_2-t_1}(-\bar{\xi}_{t_2}(x_2), -\bar{\xi}_{t_1}(x_1)), \tag{4.37}$$

and the property $Z_\tau(-u', -u) = Z_{-\tau}(u', u) = Z_\tau^{-1}(u, u')$. Equations (4.35) and (4.36) follow by (4.20) and (4.21) changing $\xi_s(y)$ with $-\bar{\xi}_t(x)$ and $\xi_t(x)$ with $-\bar{\xi}_s(y)$. □

We are now in the position to compare the forward solution ξ_t and the backward solution $\bar{\xi}_t$. The main philosophy is that in general we only have $\mathcal{R}_t \mathcal{P}_t \xi_0 \leq \xi_0$ (cf. (4.38) below), but equality holds μ_t-a.e. if $(\xi_0, \mathcal{P}_t \xi_0)$ is an optimal pair. In the following result, we still stay in the general case comparing arbitrary forward solutions $\xi_t = \mathcal{P}_t \xi_0$ and backward solutions $\bar{\xi}_t = \mathcal{R}_{1-t} \bar{\xi}_1$ only assuming $\xi_1 \geq \bar{\xi}_1$. Along the contact set Ξ_t where ξ_t and $\bar{\xi}_t$ coincide, we can then derive differentiability and optimality properties of ξ_t and $\bar{\xi}_t$.

Theorem 4.3. (Contact set Ξ_t) *Let $\xi_0 : \mathbb{R}^d \to [a, +\infty]$ be l.s.c. with $a \geq -1/2$ and $\bar{\xi}_1 : \mathbb{R}^d \to [-\infty, -\bar{a}]$ u.s.c. with $\bar{a} \leq 1/2$. Assume $\mathcal{P}_1 \xi_0 \geq \bar{\xi}_1$ and set*

$$\xi_t := \mathcal{P}_t \xi_0 \quad and \quad \bar{\xi}_t := \mathcal{R}_{1-t} \bar{\xi}_1 \quad for \ t \in [0, 1]. \tag{4.38}$$

Then, the following assertions hold:

(1) *For every $t \in [0, 1]$ we have $\xi_t \geq \bar{\xi}_t$ and the contact set*

$$\Xi_t := \left\{ x \in \mathbb{R}^d : \bar{\xi}_t(t) = \xi_t(x) \right\} \quad is \ closed. \tag{4.39}$$

(2) *For every $t \in (0, 1)$ and $x \in \Xi_t$ there exists a unique $p = \boldsymbol{g}_t(x)$ satisfying*

$$\xi_t(y) - \xi_t(x) - \frac{1}{2} \Lambda_{\bar{a}}(1-t)|x-y|^2 \leq \langle p, y - x \rangle \leq \bar{\xi}_t(y) - \bar{\xi}_t(x) + \frac{1}{2} \Lambda_a(t)|x-y|^2 \tag{4.40}$$

so that in particular ξ_t and $\bar{\xi}_t$ are differentiable at x with gradient $\boldsymbol{g}_t(x)$ (cf. Proposition 4.1(2) for Λ_a).

(3) *The map $x \mapsto \boldsymbol{g}_t(x)$ is bounded and $C(t)$-Lipschitz with $C(t) \leq 2(\Lambda_a(t) + \Lambda_{\bar{a}}(1-t)) \leq \frac{4}{t(1-t)}$ on Ξ_t. Moreover, the sets*

$$\Xi_t^- := \left\{ x \in \mathbb{R}^d \ \Big| \ \xi_0 = -\frac{1}{2} \right\} = \left\{ x \in \Xi_t \ \Big| \ \xi_t = \frac{-1}{2(1-t)} \right\}$$
$$\Xi_t^+ := \left\{ x \in \mathbb{R}^d \ \Big| \ \bar{\xi}_1 = \frac{1}{2} \right\} = \left\{ x \in \Xi_t \ \Big| \ \bar{\xi}_t = \frac{1}{2t} \right\} \tag{4.41}$$

are independent of t, are contained in Ξ_t for every $t \in [0, 1]$, and the critical set $\Xi_t^0 := \{x \in \Xi_t : \boldsymbol{g}_t(x) = 0\}$ of $\boldsymbol{g}_t$ contains $\Xi_t^{\pm}$:

$$\Xi_t^0 \supset \Xi_t^- \cup \Xi_t^+ \quad for \ every \ t \in (0, 1). \tag{4.42}$$

(4) *Let $s \in (0, 1)$, $t \in [0, 1]$, and $\tau := t - s \neq 0$. Then, for every $x_s \in \Xi_s$ with $1 + 2\tau \xi_s(x_s) > 0$ the set $\mathsf{M}_{s \to t}(x_s)$ consists of a unique element $x_t =: \boldsymbol{T}_{s \to t}(x_s)$ satisfying*

$$x_t \in \Xi_t \ and \ x_s \in \overline{\mathsf{M}}_{t \leftarrow s}(x_t),$$
$$x_t = \boldsymbol{T}_{s \to t}(x) = x + \mathbf{arctan}\left(\frac{\tau \boldsymbol{g}_s(x)}{1 + 2\tau \xi_s(x)} \right), \tag{4.43}$$
$$\left(1 - 2\tau \xi_t(\boldsymbol{T}_{s \to t}(x)) \right) \left(1 + 2\tau \xi_s(x) \right) = \cos^2_{\pi/2}\left(|x - \boldsymbol{T}_{s \to t}(x)| \right).$$

(5) *For every $x \in \Xi_s^0 \supset \Xi_s^{\pm}$ we have $T_{s \to t}(x) = x$ (and thus we set $T_{s \to t}(x) := x$ also for $t = 0$ or $t = 1$). Let $s \in (0,1)$ and define $T_{s \to s}(x) = x$, then for all $x \in \Xi_s$ the mappings $t \mapsto T_{s \to t}(x)$ are analytic in $[0,1]$. For $s, t \in (0,1)$ the mappings $T_{s \to t} : \Xi_s \to \mathbb{R}^d$ are Lipschitz. If $t = 0$ (resp. $t = 1$) then $T_{s \to t}$ is locally Lipschitz in $\Xi_s \setminus \Xi_s^+$ (resp. in $\Xi_s \setminus \Xi_s^-$).*

(6) *Setting*

$$q_{s \to t}^2(x) := \frac{1 + 2\tau \xi_s(x)}{1 - 2\tau \xi_t(T_{s \to t}(x))} = (1 + 2\tau \xi_s(x))^2 + \tau^2 |g_s(x)|^2 \quad (4.44)$$

for every $x \in \Xi_s$, the map $t \mapsto q_{s \to t}(x)$ is analytic in $[0,1]$, $q_{t \to s}$ is bounded and Lipschitz with respect to x, and $q_{s \to t}(x) > 0$ for $t \in (0,1)$ or $t = 0$ and $x \notin \Xi_s^+$ (resp. $t = 1$ and $x \notin \Xi_s^-$). Moreover, $q_{s \to t}(x) = 1 + 2(t - s)\xi_s(x)$ for $x \in \Xi_s^{\pm}$.

(7) *For all $t_0, t_1 \in (0,1)$, $t_2 \in [0,1]$, the maps $T_{t_i \to t_j}$ are Lipschitz on Ξ_{t_i} for $i \in \{0, 1\}$, and we have*

$$T_{t_1 \to t_2} \circ T_{t_0 \to t_1} = T_{t_0 \to t_2}, \quad q_{t_1 \to t_2}(T_{t_0 \to t_1}(x)) \cdot q_{t_0 \to t_1}(x) = q_{t_0 \to t_2}(x). \tag{4.45}$$

Proof. <u>Assertion (1)</u>. The inequality

$$\xi_s \geqq \mathscr{R}_{t-s}\left(\mathscr{P}_{t-s}\xi_s\right) = \mathscr{R}_{t-s}\xi_t \quad \text{for } 0 < s < t < 1 \tag{4.46}$$

can be derived by the link with the Hopf-Lax semigroup in $\mathfrak{C}$ given by [27, Theorem 8.11] and arguing as in [34, Thm. 7.36]. We prove it by a direct computation as follows: Set $\tau = t - s$, observe that $\inf \mathscr{P}_\tau \xi_s = \inf \xi_t \leqq \frac{1}{2t} < \frac{1}{2\tau}$, and use $\xi_t = \mathscr{P}_\tau \xi_s$ to obtain

$$\frac{1}{1 - 2\tau \xi_t(y)} = \inf_{z \in B_{\pi/2}(y)} \frac{1 + 2\tau \xi_s(z)}{\cos_{\pi/2}^2(|y-z|)} \leqq \frac{1 + 2\tau \xi_s(x)}{\cos_{\pi/2}^2(|y-x|)} \quad \text{if } |x-y| < \pi/2. \tag{4.47}$$

With this estimate, we find

$$\mathscr{R}_\tau \xi_t(x) \overset{\text{def}}{=} \sup_{y \in B_{\pi/2}(x)} \frac{1}{2\tau}\left(\frac{\cos_{\pi/2}^2(|x-y|)}{1 - 2\tau \xi_s(y)} - 1\right)$$

$$\overset{(4.47)}{\leqq} \sup_{y \in B_{\pi/2}(x)} \frac{1}{2\tau}\left(\frac{\cos_{\pi/2}^2(|x-y|)}{\cos_{\pi/2}^2(|x-y|)}(1 + 2\tau \xi_s(x)) - 1\right) = \xi_s(x).$$

Using $\xi_t \geqq \xi_1 \geqq \bar{\xi}_1$ we thus get (4.46). Passing to the limit as $t \uparrow 1$ in (4.46), we arrive at $\xi_s \geqq \mathscr{R}_{1-s}\bar{\xi}_1 = \bar{\xi}_s$.

The closedness of Ξ_t follows from the semi-continuities of ξ_t and $\bar{\xi}_t$ and the estimate $\xi_t \geqq \bar{\xi}_t$. Indeed, assume $y_k \to y$ with $y_k \in \Xi_t$, then we have $y \in \Xi_t$ because of

$$\xi_t(y) \overset{\text{l.s.c.}}{\leqq} \liminf_{k \to \infty} \xi_t(y_k) = \liminf_{k \to \infty} \bar{\xi}_t(y_k) \leqq \limsup_{k \to \infty} \bar{\xi}_t(y_k) \overset{\text{u.s.c.}}{\leqq} \bar{\xi}_t(y) \leqq \xi_t(y).$$

Assertion (2). Let us fix $x \in \Xi_t$, $\Lambda := \Lambda_a(t)$ and $\bar{\Lambda} := \Lambda_{\bar{a}}(1-t)$, and let p (resp. p') be an element of the superdifferential of $x \mapsto \xi_t(x) - \frac{1}{2}\Lambda|x|^2$ (resp. of the subdifferential of $x \mapsto \bar{\xi}_t(x) + \frac{1}{2}\bar{\Lambda}|x|^2$). The superdifferential (subdifferential) is not empty, since the function is concave (convex) and finite everywhere. For every $x, y \in \mathbb{R}^d$ with $x \in \Xi_t$ we have

$$\langle p, y-x \rangle \geq \xi_t(y) - \xi_t(x) - \frac{1}{2}\Lambda|x-y|^2 \text{ and } \langle p', y-x \rangle$$
$$\leq \bar{\xi}_t(y) - \bar{\xi}_t(x) + \frac{1}{2}\bar{\Lambda}|x-y|^2.$$

Subtracting the two inequalities and using $\xi_t(x) = \bar{\xi}_t(x)$ and $\bar{\xi}_t(y) \leq \xi_t(y)$ yields

$$\langle p' - p, y - x \rangle \leq \bar{\xi}_t(y) - \xi_t(y) + \frac{1}{2}(\Lambda+\bar{\Lambda})|y-x|^2 \leq \frac{1}{2}(\Lambda+\bar{\Lambda})|y-x|^2$$
$$\text{for every } y \in \mathbb{R}^d,$$

so that $p = p'$ is uniquely determined and (4.40) holds.

Assertion (3). The fact that $\Xi_t^{\pm}$ are independent of t and contained in Ξ_t follows from (4.17) and (4.32). Moreover, (4.42) follows easily since ξ_t takes its minimum at Ξ_t^- and its maximum at Ξ_t^+.

Let us now fix $t \in (0, 1)$, $x_0, x_1 \in \Xi_t$, $p_i = \boldsymbol{g}_t(x_i) + \bar{\Lambda}x_i$, and set $\bar{\zeta}(x) := \bar{\xi}_t(x) + \frac{1}{2}\bar{\Lambda}|x|^2$, $\zeta(x) := \xi_t(x) + \frac{1}{2}\bar{\Lambda}|x|^2$. Notice that $\bar{\zeta}(x)$ is convex and $\zeta(x)$ is $C = \Lambda + \bar{\Lambda}$ semi-concave with $\bar{\zeta}(x) \leq \zeta(x)$. We get

$$\bar{\zeta}(x_0) \leq \bar{\zeta}(x) - \langle p_0, x - x_0 \rangle \leq \zeta(x) - \langle p_0, x - x_0 \rangle$$
$$\leq \zeta(x_1) + \langle p_1, x - x_1 \rangle - \langle p_0, x - x_0 \rangle + \frac{C}{2}|x-x_1|^2$$
$$= \bar{\zeta}(x_1) + \langle p_1 - p_0, x - x_1 \rangle - \langle p_0, x_1 - x_0 \rangle + \frac{C}{2}|x-x_1|^2.$$

Minimizing with respect to x we find $\bar{\zeta}(x_0) \leq \bar{\zeta}(x_1) - \langle p_0, x_1-x_0 \rangle - \frac{1}{4C}|p_1-p_0|^2$. Inverting the role of x_0 and x_1 and summing up gives $\frac{1}{2C}|p_1-p_0|^2 \leq \langle p_1-p_0, x_1-x_0 \rangle$ and therefore

$$|p_1-p_0| \leq 2C\,|x_1-x_0|. \tag{4.48}$$

The boundedness of $\boldsymbol{g}_t$ on Ξ_t follows by the fact that ξ_t is Lipschitz.

Assertion (4). Let us first consider the case $s > t$ with $\tau := s-t$ and let $y \in \mathsf{M}_{s \to t}(x)$. If $\xi_s(x) > -\frac{1}{2\tau}$ then y satisfies the identity (4.21). Since $\bar{\xi}_t(y) \leq \xi_t(y)$, (4.35) and $\bar{\xi}_s(x) = \xi_s(x)$ yields $\bar{\xi}_t(y) = \xi_t(y)$ so that $y \in \Xi_t$ as well with $x \in \mathsf{M}_{t \leftarrow s}(y)$ since $\bar{\xi}_t(y) \geq -\frac{1}{2(1-t)} > -\frac{1}{2\tau}$.

Since the function $x' \mapsto (1+2\tau\xi_s(x'))\,(1-2\tau\xi_t(y)) - \cos^2_{\pi/2}\big(|x'-y|\big)$ has a global minimizer at x, we arrive at the Euler–Lagrange equations

$$2\tau(1-2\tau\xi_t(y))\,\boldsymbol{g}_s(x) + 2\cos_{\pi/2}(|x-y|)\,\boldsymbol{\sin}(x-y) = 0$$

Since we can assume $|x-y| < \pi/2$ we obtain

$$x - y = -\mathbf{arctan}\left(\frac{\tau \boldsymbol{g}_s(x)}{1+2\tau \xi_s(x)}\right), \tag{4.49}$$

which characterizes y uniquely and establishes (4.43).

The case $t > s$ follows by the same arguments.

Assertion (5). This assertion is an immediate consequence of (4.43) and (4.42).

Assertion (6). The claims are simple consequences of the identity (4.15) and the definition of $q_{s\to t}$ of (4.44).

Assertion (7). The final assertion follows by (4.19) (and the corresponding (4.34)). $\quad\square$

Remark 4.4. (Strongly reduced pairs) It is worth noticing that if $\inf \xi_0 > -\frac{1}{2}$ and $\sup \bar{\xi}_1 < \frac{1}{2}$, then the sets $\Xi_t^{\pm}$ in (4.41) are empty and many properties of ξ_t, $\boldsymbol{T}_{s\to t}$ and $q_{s\to t}$ become considerably simpler. This situation is, e.g., the case of the solution induced by a strongly reduced pair with compact support, see Theorem 3.6.

We close this subsection by giving a small example for ξ_t and $\bar{\xi}_t$ and their contact set Ξ_t derived from an optimal pair $(\xi_0, \bar{\xi}_1)$ for the transport between two Dirac measures.

Example 4.5. (The contact set for two Dirac measures) For points $z_0, z_1 \in \mathbb{R}^d$ and $r_0, r_1 > 0$ we consider the Dirac measures $\mu_j = r_j^2 \delta_{z_j}$. We have

$$\mathsf{HK}^2(\mu_0, \mu_1) = r_0^2 + r_1^2 - 2r_0 r_1 \cos_{\pi/2}(\varrho) \quad \text{with } \varrho = |z_1 - z_0|,$$

and all geodesic curves are known, see [26, Sec. 5.2]. For $\varrho < \pi/2$ we have a unique geodesic $\mu_t = r(t)^2 \delta_{z(t)}$ defined by transport, and for $\varrho > \pi/2$ the unique geodesic $\mu_t = (1-t)^2 r_0^2 \delta_{z_0} + t^2 r_1^2 \delta_{z_1}$ consists of growth (annihilation and decay) only. For $\varrho = \pi/2$ there is an infinite-dimensional convex set of geodesics, and we will see that this property is also reflected by a larger contact set.

Using the simple one-point supports of μ_j it is easy to calculate the optimal potentials and the transport plan $\boldsymbol{\eta}$ in Theorem 2.14(ii). We obtain

$$s_0 := \sigma_0(z_0) = \frac{r_1}{r_0} \cos_{\pi/2}(\varrho), \quad s_1 := \sigma_1(z_1) = \frac{r_0}{r_1} \cos_{\pi/2}(\varrho),$$

$$\boldsymbol{\eta} = r_0 r_1 \cos_{\pi/2}(\varrho)\, \delta_{(z_0, z_1)}.$$

Thus, we will distinguish the case $\cos_{\pi/2}(\varrho) > 0$ and $\cos_{\pi/2}(\varrho) = 0$.

<u>Case $\varrho < \pi/2$</u> : By (2.44) the optimal pair $(\xi_0, \bar{\xi}_1)$ reads as

$$\xi_0(x) = \begin{cases} \frac{s_0-1}{2} & \text{for } x = z_0, \\ +\infty & \text{for } x \neq z_0; \end{cases} \quad \text{and} \quad \xi_1(x) = \begin{cases} \frac{1-s_1}{2} & \text{for } x = z_1, \\ -\infty & \text{for } x \neq z_1. \end{cases}$$

From these identities, we obtain the forward and backward solutions $\xi_t = \mathscr{P}_t \xi_0$ and $\bar{\xi}_t = \mathscr{R}_{1-t} \bar{\xi}_1$:

$$\xi_t(x) = \frac{1-t+ts_0 - \cos_{\pi/2}^2(|x-z_0|)}{2t\,(1-t+ts_0)} \quad \text{and} \quad \bar{\xi}_t(x) = \frac{\cos_{\pi/2}^2(|x-z_1|) - t - (1-t)s_1}{2\,(1-t)\,(t+(1-t)s_1)}.$$
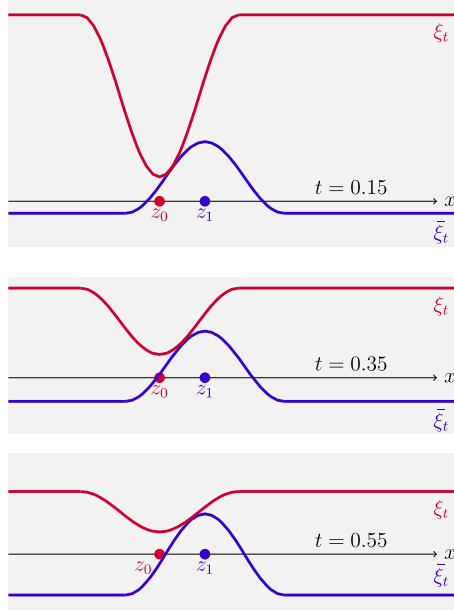
Fig. 3. For the case $\varrho = |z_1 - z_0| = 0.9 < \pi/2$ the functions $\xi_t(x)$ (red) and $\bar{\xi}_t(x)$ (blue) from (4.50) are displayed for the different times $t = 0.15,\ 0.35$, and 0.55 (with parameter $r_1/r_0 = 2$). We always have $\xi_t(x) \geqq \bar{\xi}_t(x)$ with equality at the one-point contact set $\Xi_t = \{z(t)\}$, where $z(t) = T_{0 \to t}(z_0)$ moves continuously from z_0 to z_1

$$(4.50)$$

The following optimality conditions can be checked by direct computation:

(a) $\xi_0 \geqq \bar{\xi}_0$ and $\xi_1 \geqq \bar{\xi}_1$ on $\mathbb{R}^d$

(b) $\xi_0 = \bar{\xi}_0$ μ_0-a.e. and $\xi_1 = \bar{\xi}_1$ μ_1-a.e.

As $\xi_0(x) = +\infty$ for $x \neq z_0$ and $\xi_1(x) = -\infty$ for $x \neq z_1$ statement (a) follows from (b). For (b) observe

$$\bar{\xi}_0(z_0) = \frac{\cos^2 \varrho - s_1}{2s_1} = \frac{\cos^2 \varrho - (r_0/r_1)\cos \varrho}{2(r_0/r_1)\cos \varrho} = \frac{1}{2}\left(\frac{r_1}{r_0}\cos \varrho - 1\right) = \frac{s_0 - 1}{2}$$
$$= \xi_0(z_0).$$

Similarly, $\xi_1(z_1) = \bar{\xi}_1(z_1)$ follows, which provides a first result on the contact sets $\Xi_t := \{x \in \mathbb{R}^d \mid \xi_t(x) = \bar{\xi}_t(x)\}$, namely $\Xi_0 = \{z_0\}$ and $\Xi_1 = \{z_1\}$.

The general theory in Theorem 4.3(i) guarantees $\xi_t \geqq \bar{\xi}_t$. A lengthy computation shows that Ξ_t is a singleton also for $t \in (0,1)$, i.e. $\Xi_t = \{a(t)\}$ from $\mu_t = r(t)^2 \delta_{z(t)}$ and $\Xi^{\pm} = \emptyset$. We refer to Fig. 3, where $x \mapsto (\xi_t(x), \bar{\xi}_t(t))$ is plotted.

$\underline{Case \varrho \geqq \pi/2}$: Now we have $s_0 = s_1 = 0$ and ξ_t and $\bar{\xi}_t$ simplify accordingly:

$$\xi_t(x) = \frac{1 - t - \cos^2_{\pi/2}(|x - z_0|)}{2t(1-t)} \quad \text{and} \quad \bar{\xi}_t(x) = \frac{\cos^2_{\pi/2}(|x - z_1|) - t}{2t(1-t)}. \quad (4.51)$$
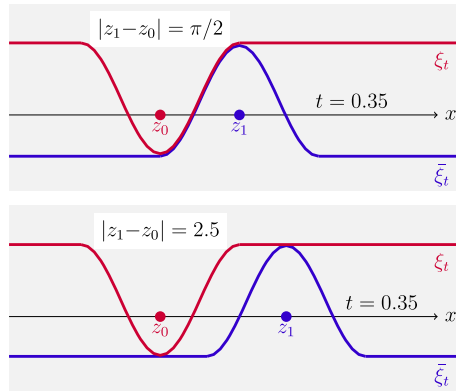
Fig. 4. For $\varrho = |z_1 - z_0| \geqq \pi/2$ the contact set Ξ_t for the functions $\xi_t(x)$ (red) and $\bar{\xi}_t(x)$ (blue) from (4.51) is no longer a singleton. For $\varrho = \pi/2$ (upper figure) we obtain $\Xi_t = [z_0, z_1]$. For $\varrho > \pi/2$ (lower figure), we have $\Xi_t = \Xi_t^+ \cup \Xi^-$ with $\Xi_t^- = \{z_0\}$ and $\Xi_t^+ = \{z_1\}$

The contact sets are easily found depending on $\varrho = \pi/2$ or $\varrho > \pi/2$, namely

$$\varrho > \pi/2: \quad \Xi_t = \Xi^- \cup \Xi_t^+ \text{ with } \Xi^- = \{z_0\} \text{ and } \Xi_t^+ = \{z_1\},$$
$$\varrho = \pi/2: \quad \Xi_t = [z_0, z_1] \text{ and } \Xi^- = \{z_0\} \text{ and } \Xi^+ = \{z_1\},$$

where $[z_0, z_1]$ denotes the segment $\{(1-\theta)z_0 + \theta z_1 \mid \theta \in [0, 1]\}$, see Fig. 4.

The interesting fact that for $\varrho = |z_1 - z_0| = \pi/2$ the contact set Ξ_t is constant and consists of a full segment reflects the observation in [26, Sec. 5.2] that μ_0 and μ_1 can be connected by geodesics satisfying $\mathrm{sppt}(\mu_t) = [z_0, z_1]$ for all $t \in [0, 1]$.

4.3. Geodesic Flow and Characteristics

Finally, we study the differentiability of $g_s = \nabla \xi_s$ and $T_{t \to s}$ on Ξ_s. Let us denote by $\widetilde{\Xi}_t$ the subset of density points of the contact set Ξ_t, which is closed by (4.39):

$$x \in \widetilde{\Xi}_t \quad \Leftrightarrow \quad \lim_{\varrho \downarrow 0} \frac{\mathcal{L}^d(\Xi_t \cap B_\varrho(x))}{\mathcal{L}^d(B_\varrho(x))} = 1. \tag{4.52}$$

Notice that $\widetilde{\Xi}_t$ is just the set of Lebesgue points of the characteristic functions of Ξ_t, so that [1] $\mathcal{L}^d(\Xi_t \setminus \widetilde{\Xi}_t) = 0$. By [9, Thm. 1], the family of sets $(\widetilde{\Xi}_t)_{t \in (0,1)}$ is invariant with respect to the action of the bi-Lipschitz maps $T_{s \to t}$, i.e., $T_{s \to t}(\widetilde{\Xi}_s) = \widetilde{\Xi}_t$ for every $s, t \in (0, 1)$.

Given a locally Lipschitz function $F : \Xi_t \to \mathbb{R}^d$ and $x \in \widetilde{\Xi}_t$, we say that F is differentiable at x if there exists a matrix $\mathsf{A} = DF(x) \in \mathbb{R}^{d \times d}$ such that

$$|F(y) - F(x) - \mathsf{A}(y - x)| = o(|y - x|) \quad \text{as } y \to x, \ y \in \Xi_t. \tag{4.53}$$

Since x belongs to the set $\widetilde{\Xi}_t$ of density points of Ξ_t, the matrix A is unique and every (locally) Lipschitz extension of F is differentiable at x with the same differential A (e.g. one can argue as in the proof of [1, Thm. 2.14]).

We call $\mathrm{dom}_t(\mathrm{D}F)$ the set of differentiability points $x \in \widetilde{\Xi}_t$ of F. If F is locally Lipschitz in Ξ_t, considering an arbitrary Lipschitz extension of F and applying Rademacher's theorem, we know that $\mathcal{L}^d(\Xi_t \setminus \mathrm{dom}_t(\mathrm{D}F)) = 0$. We will use the simple chain-rule property that if $y = F(x)$ is a density point of $F(\Xi_t)$ and $H : F(\Xi_t) \to \mathbb{R}^k$ is differentiable at y, then

$$\mathrm{D}(H \circ F)(x) = \mathrm{D}H(F(x)) \cdot \mathrm{D}F(x). \tag{4.54}$$

In the proof of the following lemma we will denote by $\partial \xi_s$ the Fréchet subdifferential of ξ_s, which coincides with $\nabla \xi_s$ whenever ξ_s is differentiable, in particular in $x \in \Xi_s$.

Lemma 4.6. *Let $s \in (0, 1)$ and let $x \in \widetilde{\Xi}_s$ be a density point of Ξ_s where $\boldsymbol{g}_s = \nabla \xi_s$ is differentiable in the sense of* (4.53) *with $p = \boldsymbol{g}_s(x)$ and $\mathsf{A} = \mathrm{D}\nabla \xi_s(x)$. Then*

$$\mathsf{A} = \mathrm{D}\nabla \xi_s(x) \quad \text{is symmetric,} \tag{4.55a}$$

$$\sup_{z \in \partial \xi_s(y)} |z - p - \mathsf{A}(y-x)| = o\left(|y-x|\right) \quad \text{as } y \to x, \tag{4.55b}$$

$$\xi_s(y) - \xi_s(x) - \langle p, y-x \rangle - \frac{1}{2}\langle \mathsf{A}(y-x), y-x \rangle = o\left(|y-x|^2\right) \quad \text{as } y \to x, \tag{4.55c}$$

Analogous results hold for $\bar{\xi}_s$. We will denote $\mathrm{D}\nabla \xi_s$ by $\mathrm{D}^2 \xi_s$.

Notice that the points y in the limits in (4.55b) and (4.55c) are not restricted to Ξ_s.

Proof. We adapt some ideas of [3,5] to our setting, and we consider the case of $\bar{\xi}_s$ (to deal with a semi-convex function, instead of semi-concave). We will assume $x = 0$ and will shortly write $\bar{\xi}$ and Ξ for $\bar{\xi}_s$ and Ξ_s omitting the explicit dependence on the parameter s. For $h > 0$ we define the blowup set $\Xi^h := h^{-1}\Xi$. Up to an addition of a quadratic term, it is also not restrictive to assume that $\bar{\xi}$ is convex.

For $h > 0$ we set $\omega_h(y) := \frac{1}{h^2}\left(\bar{\xi}(hy) - \bar{\xi}(x) - h\langle p, y \rangle\right)$ so that ω_h is a convex and nonnegative function. By (4.40) there exists a positive constant C such that

$$0 \leqq \omega_h(y) \leqq C|y|^2 \quad \text{for every } y \in \Xi^h. \tag{4.56}$$

Since $x = 0$ is a density point of Ξ, $\mathcal{L}^d(B_r(0) \setminus \Xi^h) \to 0$ as $h \downarrow 0$ so that every point of $z \in B_r(0)$ is a limit of a sequence in $z_h \in \Xi^h \cap B_r(0)$. Therefore, for h sufficiently small we can find points $y_{h,i} \in \Xi^h \cap B_{4d}(0)$, $i = 1, \cdots, 2d$, such that $\overline{B_2(0)} \subset \mathrm{conv}(\{ y_{h,i} \mid i = 1, \cdots, 2d \})$. For this it is sufficient to approximate the (rescaled) elements of the canonical basis $\pm e_i$, $i = 1, \cdots, d$. If $y \in B_2(0)$ we then find coefficients $\alpha_{h,i} \geqq 0$, $\sum_i \alpha_{h,i} = 1$ such that

$$\omega_h(y) \leqq \sum_i \alpha_{h,i}\omega_h(y_{h,i}) \leqq C \sum_i \alpha_{h,i}|y_{h,i}|^2 \leqq 2dC$$

so that ω_h is uniformly bounded in $B_2(0)$ and therefore is also uniformly Lipschitz in $\overline{B_1(0)}$. Every infinitesimal sequence $h_n \downarrow 0$ has a subsequence $m \mapsto h_{n(m)}$ such that $\omega_{h_{n(m)}}$ is uniformly convergent to a nonnegative, convex Lipschitz function $\omega : \overline{B_1(0)} \to \mathbb{R}$. We want to show that any limit point ω coincides with the quadratic function induced by the differential A, namely $\omega(y) = \omega_\mathsf{A}(y) = \frac{1}{2}\langle \mathsf{A}y, y \rangle$

Let ω be the uniform limit of ω_h along a subsequence $h_n \downarrow 0$. If $y_n \in \Xi^{h_n} \cap B_1(0)$ is converging to $y \in B_1(0)$ we know that any limit point of $p_n = \nabla \omega_{h_n}(y_n)$ belongs to $\partial \omega(y)$. On the other hand, $p_n = \frac{1}{h_n}(\nabla \bar{\xi}(h_n y_n) - p) = \mathsf{A}y_n + o(1)$ thanks to the differentiability assumption, so that $\mathsf{A}y \in \partial \omega(y)$. Since we can approximate every point of $B_1(0)$ we conclude that $\mathsf{A}y \in \partial \omega(y)$ for every $y \in B_1(0)$. On the other hand, ω is Lipschitz, so that it is differentiable a.e. in $B_1(0)$ with $\nabla \omega(y) = \mathsf{A}y$ and therefore the distributional differential of $\nabla \omega$ coincides with A. We conclude that A is symmetric and $\omega(y) = \frac{1}{2}\langle \mathsf{A}y, y \rangle$. The fact that ω_h uniformly converges to ω eventually yields (4.55b) and (4.55c). □

We now use the second-order differentiability of ξ_s to derive differentiability of $\boldsymbol{T}_{s \to t}$ by using the formula (4.43) with $\boldsymbol{g}_s(x) = \nabla \xi_s(x)$. For $s \in (0, 1)$ we define

$$\mathfrak{D}_s = \mathrm{dom}_s(D\nabla \xi_s)) \cap \widetilde{\Xi}_s = \mathrm{dom}_s(D^2 \xi_s) \cap \widetilde{\Xi}_s. \tag{4.57}$$

As we already observed, since $\boldsymbol{g}_s$ is Lipschitz on Ξ_s, $\mathcal{L}^d(\Xi_s \backslash \mathfrak{D}_s) = 0$ for every $s \in (0, 1)$.

For $t \in (0, 1)$ and $\tau = t - s$ we also have $1 + 2\tau \xi_s \geqq (1-t)/(1-s) > 0$ so that

$$x \mapsto \frac{\tau}{1 + 2\tau \xi_s(x)} \nabla \xi_s(x) = \nabla \phi_{s,t}(x) \text{ with } \phi_{s,t}(x) = \frac{1}{2} \log \left(1 + 2\tau \xi_s(x)\right)$$

is again Lipschitz on Ξ_s. Thus, Lemma 4.6 can be applied and $\phi_{s,t}$ is differentiable in the sense of (4.53) on $\mathfrak{D}_s$. Finally, we exploit the explicit representation of $\boldsymbol{T}_{s \to t}$ via (4.43), namely for all $x \in \Xi_s$ we have

$$\boldsymbol{T}_{s \to t}(x) = x + \mathbf{arctan}\left(\frac{\tau \nabla \xi_s(x)}{1 + 2\tau \xi_s(x)}\right) = x + \mathbf{arctan}\left(\nabla \phi_{s,t}(x)\right). \tag{4.58}$$

Now the chain rule (4.54) guarantees the differentiability of $\boldsymbol{T}_{s \to t}$ on the set $\mathfrak{D}_s$.

Lemma 4.7. (Differentiability of $\boldsymbol{T}$) *For all $s, t \in (0, 1)$ the mapping $\boldsymbol{T}_{s \to t}$ is differentiable on $\mathfrak{D}_s$, and we have*

$$D\boldsymbol{T}_{s \to t}(x) = \mathbb{T}\big(t - s, \xi_s(x), \nabla \xi_s(x), D^2 \xi_s(x)\big) \text{ with}$$

$$\mathbb{T}(\tau, \xi, \boldsymbol{g}, \mathsf{A}) := \mathbb{I} + \big(D^2 \mathsf{L}_1(z)\big)^{-1}\Big|_{z=\mathbf{arctan}(\frac{\tau \boldsymbol{g}}{1+2\tau \xi})} \left(\frac{\tau \mathsf{A}}{1 + 2\tau \xi} - \frac{2\tau^2 \boldsymbol{g} \otimes \boldsymbol{g}}{(1 + 2\tau \xi)^2}\right). \tag{4.59a}$$

$$\boldsymbol{T}_{s \to t}(\mathfrak{D}_s) = \mathfrak{D}_t \text{ and } D\boldsymbol{T}_{t \to s}(\boldsymbol{T}_{s \to t}(x))D\boldsymbol{T}_{s \to t}(x) = \mathbb{I} \text{ for } x \in \mathfrak{D}_s. \tag{4.59b}$$

For every $t_0, t_1 \in (0, 1)$, $t_2 \in [0, 1]$ we also have

$$D\boldsymbol{T}_{t_1 \to t_2}(\boldsymbol{T}_{t_0 \to t_1}(x))D\boldsymbol{T}_{t_0 \to t_1}(x) = D\boldsymbol{T}_{t_0 \to t_2}(x) \text{ for } x \in \mathfrak{D}_{t_0}. \tag{4.59c}$$

Proof. Recall $\tau = t - s$, then the explicit formula (4.59a) follows from differentiating $\nabla L_1(x - \boldsymbol{T}_{s \to t}(x)) = -\frac{\tau}{1 + 2\tau \xi_s} \nabla \xi_s$. Since $\boldsymbol{T}_{s \to t}^{-1} = \boldsymbol{T}_{t \to s}$ there exists a constant L such that

$$L^{-1}|x - x'| \leqq |\boldsymbol{T}_{s \to t}(x) - \boldsymbol{T}_{s \to t}(x')| \leqq L|x - x'| \quad \text{for every } x, x' \in \Xi_s. \quad (4.60)$$

If $x \in \mathfrak{D}_s$ and $A = \mathrm{D}\boldsymbol{T}_{s \to t}(x)$, choosing $\varepsilon > 0$ we can find $\varrho > 0$ such that

$$|\boldsymbol{T}_{s \to t}(x') - \boldsymbol{T}_{s \to t}(x) - A(x' - x)| \leqq \varepsilon |x' - x| \quad \text{for every } x' \in \Xi_t \cap B_\varrho(x), \tag{4.61}$$

so that choosing $\varepsilon < \frac{1}{2L}$ and $x' = x + v$ we get

$$|Av| \geqq |\boldsymbol{T}_{s \to t}(x + v) - \boldsymbol{T}_{s \to t}(x)| - \varepsilon |v| \geqq \frac{1}{2L}|v| \quad \text{for every } v \in B_\varrho(0) \cap (\Xi_t - x)$$

Using the fact that 0 is a density point of $\Xi_t - x$ we conclude that A is invertible with $|A^{-1}| \leqq 2L$. For every $y' \in \Xi_t$ with $L|y' - y| < \varrho$ and $x' = \boldsymbol{T}_{t \to s}(y')$, we get $|x' - x| < \varrho$ and (4.61) yields

$$|\boldsymbol{T}_{t \to s}(y') - \boldsymbol{T}_{t \to s}(y) - A^{-1}(y' - y)| = \left| A^{-1}\left(A(x' - x) - \boldsymbol{T}_{s \to t}(x') + \boldsymbol{T}_{s \to t}(x) \right) \right|$$
$$\leqq 2L\varepsilon |x' - x| \leqq 2L^2 \varepsilon |y' - y|$$

showing that $y \in \mathfrak{D}_t$ and $A^{-1} = \mathrm{D}\boldsymbol{T}_{t \to s}(y)$. Hence, (4.59b) is established.

Equation (4.59c) then follows by the concatenation property (4.45). $\quad \square$

The explicit formula (4.59a) shows that $\mathrm{D}\boldsymbol{T}_{s \to t}$ is the product of the positive matrix $\mathrm{D}^2 \mathrm{L}_1(z)^{-1}$ and a symmetric matrix, hence it is always real diagonalizable. The following result shows that the determinant and hence all eigenvalues stay positive for $s, t \in (0, 1)$. In fact, we now derive differential equations with respect to $t \in (0, 1)$ for the transport-growth pairs $(\boldsymbol{T}_{s \to t}(x), q_{s \to t}(x)) \in \mathbb{R}^d \times (0, +\infty)$ as well as for $\mathrm{D}\boldsymbol{T}_{s \to t}(x) \in \mathbb{R}^{d \times d}$ and $\det \mathrm{D}\boldsymbol{T}_{s \to t}(x)$. Recall that $t \mapsto (\boldsymbol{T}_{s \to t}(x), q_{s \to t}(x))$ is analytic for $t \in (0, 1)$ by Theorem 4.3(5) and (6).

The following relations will be crucial to derive the curvature estimate needed for our main result on geodesic HK-convexity.

Theorem 4.8. (*The characteristic system on the contact set* Ξ_s) *We fix $s \in (0, 1)$, $x \in \Xi_s$, and $y \in \mathfrak{D}_s$ (cf. (4.57)) and define the maps*

$$\boldsymbol{T}(t) := \boldsymbol{T}_{s \to t}(x), \quad q(t) := q_{s \to t}(x), \quad \mathsf{B}(t) := \mathrm{D}\boldsymbol{T}_{s \to t}(y), \quad \text{and} \quad \delta(t) := \det \mathsf{B}(t).$$

Then, we have the initial conditions $\boldsymbol{T}(s) = x$, $q(s) = 1$, $\mathsf{B}(s) = \mathbb{I}$, and $\delta(s) = 1$, and for $t \in (0, 1)$ the following differential equations are satisfied:

$$\dot{\boldsymbol{T}}(t) = \nabla \xi_t(\boldsymbol{T}(t)) \qquad \text{and} \quad \ddot{\boldsymbol{T}}(t) = -4\xi_t(\boldsymbol{T}(t)) \nabla \xi_t(\boldsymbol{T}(t)), \tag{4.62a}$$

$$\dot{q}(t) = 2\xi_t(\boldsymbol{T}(t)) q(t) \qquad \text{and} \quad \ddot{q}(t) = |\nabla \xi_t(\boldsymbol{T}(t))|^2 q(t), \tag{4.62b}$$

$$\dot{\mathsf{B}}(t) = \mathrm{D}^2 \xi_t(\boldsymbol{T}(t)) \mathsf{B}(t) \quad \text{and} \quad \ddot{\mathsf{B}}(t) = -4\left(\nabla \xi_t \otimes \nabla \xi_t + \xi_t \mathrm{D}^2 \xi_t \right) \circ \boldsymbol{T}(t) \cdot \mathsf{B}(t), \tag{4.62c}$$

$$\dot{\delta}(t) = \Delta\xi_t(\boldsymbol{T}(t))\,\delta(t), \quad \frac{\ddot{\delta}(t)}{\delta(t)} = \left((\Delta\xi_t)^2 - |\mathrm{D}^2\xi_t|^2 - 4|\nabla\xi_t|^2 - 4\xi_t\Delta\xi_t\right)\circ\boldsymbol{T}(t).$$

(4.62d)

there $\Delta\xi_t(z) = \mathrm{tr}\left(\mathrm{D}^2\xi_t(z)\right)$ *and* $|\mathrm{D}^2\xi_t(z)|^2 = \sum_{i,j}\left(\partial_{x_i}\partial_{x_j}\xi_t(z)\right)^2$.

Proof. We use (4.45) and the Taylor expansion

$$\mathbf{arctan}\left(\frac{h\boldsymbol{g}}{1+2h\xi}\right) = h\boldsymbol{g} - 2h^2\xi\boldsymbol{g} + O(h^3) \quad\text{as } h \to 0.$$

Setting $y = \boldsymbol{T}(t) = \boldsymbol{T}_{s\to t}(x)$ and using the fact that $y \in \Xi_t$, (4.58) yields

$$\boldsymbol{T}_{t\to t+h}(y) = y + h\nabla\xi_t(y) - 2h^2\xi_t(y)\nabla\xi_t(y) + O(|h|^3) \quad\text{as } h \to 0.$$

With the composition rule (4.45) we have $\boldsymbol{T}_{s\to t+h}(x) = \boldsymbol{T}_{t\to t+h}(y)$ and compute

$$\dot{\boldsymbol{T}}(t) = \lim_{h\to 0}\frac{\boldsymbol{T}_{s\to t+h}(x) - \boldsymbol{T}_{s\to t}(x)}{h} = \lim_{h\to 0}\frac{\boldsymbol{T}_{t\to t+h}(y) - y}{h} = \nabla\xi_t(y).$$

This identity yields the first equation in (4.62a). For the second relation in (4.62a) we use

$$\ddot{\boldsymbol{T}}(t) = \lim_{h\to 0}\frac{\boldsymbol{T}_{s\to t+h}(x) - 2\boldsymbol{T}_{s\to t}(x) + \boldsymbol{T}_{s\to t-h}(x)}{h^2}$$
$$= \lim_{h\to 0}\frac{\boldsymbol{T}_{t\to t+h}(y) - 2y + \boldsymbol{T}_{t\to t-h}(y)}{h^2} = -4\xi_t(y)\nabla\xi_t(y).$$

The relations (4.62b) for $q(t) = q_{s\to t}$ follow similarly, using the scalar product rule for $q_{s\to t}$ in (4.45) and by taking the square root of (4.44), namely,

$$q_{t\to t+h}(y) = 1 + 2h\xi_t(y) + \frac{h^2}{2}|\nabla\xi_t(y)|^2 + o(h^2) \quad\text{as } h \to 0.$$

To show that $\mathsf{B}(t)$ satisfies (4.62c), we exploit the matrix product rule (4.59c) and expand $\mathrm{D}\boldsymbol{T}_{t\to t+h}(y)$ in (4.59a) to obtain

$$\mathrm{D}\boldsymbol{T}_{t\to t+h}(y) = \mathbb{I} + h\mathrm{D}^2\xi_t - 2h^2\left(\nabla\xi_t\otimes\nabla\xi_t + \xi_t\mathrm{D}^2\xi_t\right) + o(h^2) \quad\text{as } h \to 0.$$

(4.63)

For this note that $y - \boldsymbol{T}_{t\to t+h}(y) = O(|h|)$ so that $\mathrm{D}^2\mathrm{L}_1\left(y - \boldsymbol{T}_{t\to t+h}(y)\right) = \mathbb{I} + O(|h|^2)$ as L_1 is even. Thus, (4.62c) follows as in the previous two cases.

For the determinant $\delta(t)$ we again have a scalar product rule, and it suffices to expand $\det(\mathrm{D}\boldsymbol{T}_{t\to t+h}(x))$ at $h = 0$. For this we can use the classical expansion $\det(\mathbb{I}+h\mathbb{A}) = 1 + h\,\mathrm{tr}\,\mathbb{A} + \frac{1}{2}h^2\left((\mathrm{tr}\,\mathbb{A})^2 - \mathrm{tr}(\mathbb{A}^2)\right) + O(h^3)$, and obtain

$$\det\mathrm{D}\boldsymbol{T}_{t\to t+h} = 1 + h\Delta\xi_t + \frac{1}{2}h^2\left((\Delta\xi_t)^2 - |\mathrm{D}^2\xi_t|^2 - 4|\nabla\xi_t|^2 - 4\xi\Delta\xi_t\right) + o(h^2).$$

(4.64)

As before this shows (4.62d), and the theorem is proved. $\quad\square$

In this section, we have studied the forward solutions $t \mapsto \xi_t$ for $t \in (0, 1)$ and its contact sets Ξ_t with a corresponding backward solution $\bar{\xi}_t$. We obtained differentiability properties in these sets or in the slightly smaller sets D_t and derived transport relations for important quantities such as $q_{s \to t}$ and $\delta_s(t) = \det \mathrm{D} \boldsymbol{T}_{s \to t}(x)$. In the following section, we still have to show that the contact sets Ξ_t are sufficiently big, if we define $\xi_t = \mathscr{P}_t \xi_0$ and $\bar{\xi}_t = \mathscr{R}_{1-t} \xi_1$ for an optimal pair (ξ_0, ξ_1). This will be done in Theorem 5.1.

5. Geodesic Curves

In this section, we improve the characterization of Hellinger–Kantorovich geodesic curves as discussed already in [27, Sec. 8.6]. More precisely, we consider constant-speed geodesics $\mu : [0, 1] \to \mathcal{M}(\mathbb{R}^d)$ that satisfy

$$\forall s, t \in [0, 1]: \quad \mathsf{HK}(\mu(s), \mu(t)) = |s-t| \, \mathsf{HK}(\mu_0, \mu_1).$$

We first show the optimality of potentials ξ_t and $\bar{\xi}_t$ obtained from the forward or backward Hamilton–Jacobi equation in Theorem 5.1. With this, we are able to show in Theorem 5.2 that for subparts $(s, t) \subset [0, 1]$ with $\tau = t-s < 1$ the corresponding LET problem has a unique solution in Monge form, which implies that $(\mathcal{M}(\mathbb{R}^d), \mathsf{HK})$ has the strong non-branching property. Finally, in Theorem 5.4 and Corollary 5.5 we provide restrictions and splittings of geodesic curves needed for the main theorem in Section 7.

5.1. Geodesics and Hamilton–Jacobi Equation

The next result clarifies the connection with the forward and backward Hopf–Lax flows ξ_t and $\bar{\xi}_t$ studied in Theorem 4.3 and the importance of the contact set Ξ_t defined in (4.39) (see also [27, Thm. 8.20] and [34, Chap. 7] for a similar result in the framework of Optimal Transport and displacement interpolation). We emphasize that despite the non-uniqueness of the geodesics $(\mu_t)_{t \in [0,1]}$ (see [26, Sec. 5.2]) in the following result, ξ_t and $\bar{\xi}_t$ only depend on μ_0 and μ_1 and the optimal potentials φ_0 and φ_1.

The result brings together the results of Sects. 3 and 4 by starting with an optimal pair (φ_0, φ_1) from Section 3 and considering the corresponding solutions ξ_t and $\bar{\xi}_t$ of the forward and backward Hamilton–Jacobi equation starting with $\xi_0 = \check{\mathsf{G}}_1(\varphi_0)$ and $\bar{\xi}_1 = \mathsf{G}_1(\varphi_1)$, respectively. First, we observe that "intermediate" pairs (ξ_s, ξ_t) or $(\bar{\xi}_s, \bar{\xi}_t)$ are optimal for connecting the intermediate points μ_s and μ_t on an arbitrary geodesic connecting μ_0 and μ_1. Second, we observe that certain results obtained in Section 4 for $s, t \in (0, 1)$ also hold in the limit points $s, t \in \{0, 1\}$. Finally, we show that the contact set Ξ_t is large enough in the sense that it contains $\mathrm{supp}(\mu_t)$ (see Example 4.5 for some instructive case with $\varrho = \pi/2$).

Theorem 5.1. *For $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^d)$ consider a tight optimal pair (φ_0, φ_1) of (lower,upper) semi-continuous potentials as in Theorem 3.3. With $\xi_0 := \check{\mathsf{G}}_1(\varphi_0) = \frac{1}{2}(\mathrm{e}^{2\varphi_0}-1)$ and $\bar{\xi}_1 := \mathsf{G}_1(\varphi_1) = \frac{1}{2}(1-\mathrm{e}^{-2\varphi_1})$ we define $\xi_t = \mathscr{P}_t \xi_0$ and $\bar{\xi}_t = \mathscr{R}_{1-t} \bar{\xi}_1$*

as in (4.30) *and the contact sets* $\Xi_t = \{\xi_t = \bar{\xi}_t\}$ *as in* (4.39). *Finally, consider an arbitrary geodesic* $(\mu_t)_{t \in [0,1]}$ *connecting* μ_0 *to* μ_1. *Then, the following holds:*

(1) *For all* $s, t \in [0,1]$ *with* $s < t$ *both pairs* (ξ_s, ξ_t) *and* $(\bar{\xi}_s, \bar{\xi}_t)$ *are optimal for* (2.40) *and* (2.41) *for connecting* μ_s *to* μ_t, *viz.*

$$\frac{1}{2(t-s)}\mathsf{HK}^2(\mu_s, \mu_t) = \int \xi_t \, d\mu_t - \int \xi_s \, d\mu_s = \int \bar{\xi}_t \, d\mu_t - \int \bar{\xi}_s \, d\mu_s \quad (5.1)$$

(2) $S_t = \operatorname{supp}(\mu_t) \subset \Xi_t$ *for every* $t \in [0,1]$.

Proof. Assertion (1). It is sufficient to consider the forward flow $\mathscr{P}_t$. Fixing $t \in (0,1)$ we have

$$\frac{1}{2t}\mathsf{HK}^2(\mu_0, \mu_t) \geqq \int \xi_t \, d\mu_t - \int \xi_0 \, d\mu_0 \text{ and } \frac{1}{2(1-t)}\mathsf{HK}^2(\mu_t, \mu_1)$$

$$\geqq \int \xi_1 \, d\mu_1 - \int \xi_t \, d\mu_t. \qquad (5.2)$$

On the other hand, the geodesic property and the optimality of (ξ_0, ξ_1) yield

$$\int \xi_1 \, d\mu_1 - \int \xi_0 \, d\mu_0 = \frac{1}{2}\mathsf{HK}^2(\mu_0, \mu_1) = \frac{1}{2t}\mathsf{HK}^2(\mu_0, \mu_t) + \frac{1}{2(1-t)}\mathsf{HK}^2(\mu_t, \mu_1)$$

showing that the inequalities in (5.2) are in fact equalities, in particular (5.1) with $s = 0$. For $s > 0$ we still get (5.1) since $\frac{1}{2(t-s)}\mathsf{HK}^2(\mu_s, \mu_t) = \frac{1}{2t}\mathsf{HK}^2(\mu_0, \mu_t) - \frac{1}{2s}\mathsf{HK}^2(\mu_0, \mu_s)$ if $0 < s < t \leqq 1$.

Assertion (2). Equation (5.1) for $s = 0$ yields $\int (\xi_t - \bar{\xi}_t) \, d\mu_t = 0$ for all $t \in (0,1)$, so that $\xi_t \leqq \bar{\xi}_t$ and the continuity of $\xi_t, \bar{\xi}_t$ yield $\xi_t = \bar{\xi}_t$ on $S_t = \operatorname{supp}\mu_t$. The cases $t = 0$ and 1 follow by the relations between ξ_i and φ_i and the fact that $\varphi_0 = \varphi_1^{\leftarrow L_1}$, $\varphi_1 = \varphi_0^{\rightarrow L_1}$. □

Note that the inclusion $S_t = \operatorname{supp}(\mu_t) \subset \Xi_t$ is in general a strict inclusion. This can be seen for the case $|z_1 - z_0| = \pi/2$ in Example 4.5, where $\Xi_t = [z_0, z_1]$, however, there exists a pure Hellinger geodesic with $\operatorname{supp}(\mu_t) = \{z_0, z_1\}$ for $t \in (0,1)$.

We can now exploit all the regularity features of the maps $\mathbf{T}_{s \to t}$ and $q_{s \to t}$ on the contact set Ξ_t (cf. Theorem 4.3). A first important consequence is that, given an HK geodesic $(\mu_t)_{t \in [0,1]}$ and $s \in (0,1)$, the HK problem between μ_s and μ_t for any $t \in [0,1]$ has only one solution, which can be expressed in Monge form (see [2, Lem. 7.2.1] for the corresponding properties for the L^2-Wasserstein distance in $\mathbb{R}^d$).

Theorem 5.2. (Regularizing effect along geodesics) *Under the assumptions of Theorem* 5.1, *if* $s \in (0,1)$ *and* $t \in [0,1]$, *then the transport-growth pair* $(\mathbf{T}_{s \to t}, q_{s \to t})$ *of Theorem* 4.3 *is the unique solution of the Monge formulation* (2.21) *of the Entropy-Transport problem between* μ_s *and* μ_t. *In particular, the optimal Entropy-Transport problem between* μ_s *and* μ_0 *or between* μ_s *and* μ_1 *has a unique solution, and this solution is in Monge form.*

Proof. Let us consider the case $0 < s < t \leqq 1, \tau = t - s < t$. By Theorem 5.1, the pair $(\bar{\xi}_s, \bar{\xi}_t)$ is optimal for (μ_s, μ_t) and $\text{supp}(\mu_s) \subset \Xi_s$. Using the transformations

$$\varphi_0 := \frac{1}{2\tau} \log(1 + 2\tau \bar{\xi}_s) \quad \text{and} \quad \varphi_\tau := -\frac{1}{2\tau} \log(1 - 2\tau \bar{\xi}_t), \qquad (5.3)$$

we see that $(\varphi_0, \varphi_\tau)$ is a pair of potentials satisfying the assumptions of Theorem 3.3(2). Since $1 - 2\tau \xi_t \geqq 1 - \tau/t > 0$ we deduce that φ_τ is bounded from above, so that $\mu_t'' = 0$ thanks to (3.26) (where the measures μ_t' and μ_t'' are defined as in (2.13)).

Moreover, we know that μ_s' is concentrated on $\{\varphi_0 > -\infty\}$; since it is also concentrated on Ξ_s we deduce that μ_s' is concentrated on $D_0' = \text{dom}(\nabla \varphi_0)$, so that we can apply Corollary 3.5, recalling the expression of T, q given by (3.34). $\quad \square$

The above theorem allows us to deduce the fact that $(\mathcal{M}(\mathbb{R}^d), \mathsf{HK})$ has a strong non-branching property. It is shown in [26, Sec. 5.2] that the set of geodesics connecting two Dirac measures δ_{y_0} and δ_{y_1} is very large if $|y_1 - y_0| = \pi/2$: it is convex but does not lie in a finite-dimensional space. The following result shows that all these geodesics are mutually disjoint except for the two endpoints μ_0 and μ_1:

Corollary 5.3. (Strong non-branching) *If for some $s \in (0, 1)$ we have $\mathsf{HK}(\mu_0, \mu_s) = s\mathsf{HK}(\mu_0, \mu_1)$ and $\mathsf{HK}(\mu_s, \mu_1) = (1-s)\mathsf{HK}(\mu_0, \mu_1)$, then there exists a unique geodesic curve $t \mapsto \mu(t)$ such that $\mu(0) = \mu_0, \mu(s) = \mu_s$, and $\mu(1) = \mu_1$.*

The next result shows that from a given geodesic we may construct new geodesics by multiplying the measures μ_t by a suitably transported function. This will be useful in the proof of the main Theorem 7.2.

Theorem 5.4. (Restriction of geodesics) *Let $(\mu_t)_{t \in [0,1]}$ be an HK geodesic. For a given $s \in (0, 1)$ let $\nu_s \in \mathcal{M}(\mathbb{R}^d)$ with $\text{supp}(\nu_s) \subset \text{supp}(\mu_s)$. Then the curve $[0, 1] \ni t \mapsto \nu_t := (T_{s \to t}, q_{s \to t})_\star \nu_s$ is also an HK geodesic. If in addition $\nu_s = \varrho_s \mu_s$ for some Borel function $\varrho_s : \text{supp}(\mu_s) \to [0, +\infty]$, then $\nu_t' = \varrho_t \mu_t$ with $\varrho_t(y) = \varrho_s(T_{t \to s}(y))$ for every $t \in (0, 1)$.*

Proof. We keep the same notation of Theorem 5.1, let $0 < t_1 < s < t_2 < 1$, and set $\tau_1 := s - t_1, \tau_2 := t_2 - s$, and $\tau = \tau_1 + \tau_2$. We clearly have

$$\frac{1}{2\tau} \mathsf{HK}^2(\nu_{t_2}, \nu_{t_1}) \geqq \int \xi_{t_2} \, d\nu_{t_2} - \int \xi_{t_1} \, d\nu_{t_1}$$

$$= \left(\int \xi_{t_2} \, d\nu_{t_2} - \int \xi_s \, d\nu_s \right) + \left(\int \xi_s \, d\nu_s - \int \xi_{t_1} \, d\nu_{t_1} \right)$$

The conclusion then follows, if we show that $\int \xi_{t_2} \, d\nu_{t_2} - \int \xi_s \, d\nu_s \geqq \frac{1}{2\tau_2} \mathsf{HK}^2(\nu_{t_2}, \nu_s)$ and $\int \xi_s \, d\nu_s - \int \xi_{t_1} \, d\nu_{t_1} \geqq \frac{1}{2\tau_2} \mathsf{HK}^2(\nu_s, \nu_{t_1})$. We check the first inequality, the second follows similarly.

Define $q_2 := q_{s \to t_2}$ and $T_2 := T_{s \to t_2}$. Using the fact that (i) $\nu_{t_2} = (T_2, q_2)_\star \nu_s$ and (ii) identity (4.44) we obtain

$$\int (1 - 2\tau_2 \xi_{t_2}) \, d\nu_{t_2} \overset{(i)}{=} \int \left(1 - 2\tau_2 \xi_{t_2}(T_{s \to t_2}(x)) \right) q_{s \to t_2}^2(x) \, d\nu_s(x) \overset{(ii)}{=} \int (1 + 2\tau_2 \xi_s) \, d\nu_s.$$

Combining (4.43) and (4.44), we arrive at

$$\int \left(1 - 2\tau \xi_{t_2}\right) d\nu_{t_2} = \int \left(1 + 2\tau_2 \xi_s\right) d\nu_s = \int q_2 \cos(|x - \boldsymbol{T}_2(x)|) \, d\nu_s. \quad (5.4)$$

With this, we find

$$\mathsf{HK}^2(\nu_{t_2}, \nu_s) \overset{(2.17)}{\leq} \int \left(q_2^2 + 1 - 2q_2 \cos(|x - \boldsymbol{T}_2(x)|)\right) d\nu_s$$

$$\overset{(5.4)}{=} \nu_{t_2}(\mathbb{R}^d) + \nu_s(\mathbb{R}^d) - \int (1 - 2\tau \xi_{t_2}) \, d\nu_{t_2} - \int (1 + 2\tau_2 \xi_s) \, d\nu_s$$

$$= 2\tau \left(\int \xi_{t_2} \, d\nu_{t_2} - \int \xi_s \, d\nu_s \right).$$

Hence, we have shown $\frac{1}{2\tau} \mathsf{HK}^2(\nu_{t_2}, \nu_{t_1}) = \int \xi_{t_2} \, d\nu_{t_2} - \int \xi_s \, d\nu_s$, which implies that $(\nu_t)_{t \in (0,1)}$ is a geodesic as well.

We can then pass to the limits $t_1 \downarrow 0$ and $t_2 \uparrow 1$ as follows. Notice that the curve $t \mapsto \nu_t, t \in (0, 1)$, is converging in $(\mathcal{M}(\mathbb{R}^d), \mathsf{HK})$ to a limit ν_0 and ν_1 for $t \downarrow 0$ and $t \uparrow 1$, since (ν_t) is a geodesic. Moreover, for every $\zeta \in C_b(\mathbb{R}^d)$ we can pass to the limit $t \uparrow 1$ in

$$\int \zeta \, d\nu_t = \int \zeta(\boldsymbol{T}_{s \to t}(x)) q^2_{s \to t}(x) \, d\nu_s(x), \quad (5.5)$$

since $\lim_{t \uparrow 1} \boldsymbol{T}_{s \to t}(x) = \boldsymbol{T}_{s \to 1}(x)$ and $\lim_{t \uparrow 1} q_{s \to t}(x) = q_{s \to 1}(x)$ and q is uniformly bounded. A similar argument holds for the case $t \downarrow 0$.

In order to check the identity concerning the density ϱ'_t of ν_t, we use (5.5) and find

$$\int \zeta \, d\nu_t = \int \zeta(\boldsymbol{T}_{s \to t}(x)) q^2_{s \to t}(x) \, d\nu_s = \int \zeta(\boldsymbol{T}_{s \to t}(x)) q^2_{s \to t}(x) \varrho_s(x) \, d\mu_s$$

$$= \int \zeta(\boldsymbol{T}_{s \to t}(x)) q^2_{s \to t}(x) \varrho_t(\boldsymbol{T}_{s \to t}(x)) \, d\mu_s(x) = \int \zeta(y) \varrho_t(y) \, d\mu_t(y).$$

The case $t \in [0, s]$ is analogous. ☐

The next result provides the fundamental formula for the representation of densities along geodesics. Generalizing the celebrated formulas for the Kantorovich–Wasserstein geodesics, the densities are again obtained by transport along geodesics, but now with non-constant speed and an additional growth factor $a_s(t, x) = q^2_{s \to t}(x)$ to account for the annihilation and creation of mass. Recall that $\mathfrak{D}_s = \text{dom}(\mathrm{D}^2 \xi_s) \subset \Xi_s$ has full Lebesgue measure in Ξ_s, i.e. $\mathcal{L}^d(\Xi_s \backslash \mathfrak{D}_s) = 0$.

Corollary 5.5. (Representation of densities along geodesics) *For $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^d)$ consider a geodesic $(\mu_t)_{t \in [0,1]}$ connecting μ_0 to μ_1. Assume that at least one of the following properties holds:*

(a) there exists $s \in (0, 1)$ such that $\mu_s = c_s \mathcal{L}^d \ll \mathcal{L}^d$;
(b) $\mu_0 = c_0 \mathcal{L}^d \ll \mathcal{L}^d$ and $\mu''_1 \ll \mathcal{L}^d$.

Then, we have

(1) $\mu_t \ll \mathcal{L}^d$ *for every* $t \in (0, 1)$, *viz.* $\mu_t = c(t, \cdot)\mathcal{L}^d$.
(2) *For every* $s \in (0, 1)$ *the density* $c(t, \cdot)$ *can be expressed via the formula*

$$c(t, y)\big|_{y=T_{s \to t}(x)} = c(s, x)\frac{\alpha_s(t, x)}{\delta_s(t, x)} \quad \text{for every } x \in \mathfrak{D}_s, \quad t \in (0, 1), \quad (5.6a)$$

with $\mathfrak{D}_s = \mathrm{dom}_s(D\nabla\xi_s)) = \mathrm{dom}_s(D^2\xi_s)$ *(cf.* (4.57)*) and*

$$\alpha_s(t, x) := (1 + 2(t-s)\xi_s(x))^2 + (t-s)^2|\nabla\xi_s(x)|^2, \quad \delta_s(t, x) := \det DT_{s \to t}(x).$$
(5.6b)

Moreover, we have $D^2\xi_s(x) = 0$ *and* $\delta_s(t, x) = 1$ *for* $\mathcal{L}^d$-*a.e.* $x \in \Xi_s^0 \supset \Xi^\pm$; *in particular*

$$c(t, x) = \frac{t^2}{s^2}c(s, x) \text{ for } x \in \Xi^+ \quad \text{and} \quad c(t, x) = \frac{(1-t)^2}{(1-s)^2}c(s, x) \text{ for } x \in \Xi^-.$$
(5.7)

(3) *If* $\mu_0 \ll \mathcal{L}^d$ *(resp.* $\mu_1 \ll \mathcal{L}^d$*)* (5.6) *and* (5.7) *hold up to* $t = 0$ *(resp. up to* $t = 1$*).*
(4) *If* $\mu_1'' = 0$ *the representations in* (5.6) *also hold for* $s = 0$ *by restricting* x *in* $D_0'' = \mathrm{dom}(D^2\varphi_0)$, *and we have the formula*

$$DT_{0 \to t}(x) = \mathbb{T}(t, \xi_0(x), \nabla\xi_0(x), D^2\xi_0(x)) \quad \text{for every } x \in D_0'', \quad (5.8)$$

where $\mathbb{T}$ *is defined in* (4.59a).

Proof. Assertion (1). In the case (a) holds for $s \in (0, 1)$, there exists a bi-Lipschitz map $T_{s,t} : \Xi_s \to \Xi_t$ and bounded growth factors $q_{s,t} : \Xi_s \to [a, b]$ with $0 < a < b < \infty$ such that $\mu_t = (T_{s \to t}, q_{s \to t})_\star \mu_s$. In particular, for every Borel set A we have

$$\mu_t(A) \leqq b^2\mu_s(T_{s \to t}^{-1}(A)) = b^2\mu_s(T_{t \to s}(A)). \quad (5.9)$$

If $\mathcal{L}^d(A) = 0$ then $\mathcal{L}^d(T_{t \to s}(A)) = 0$ because $T_{t \to s}$ is Lipschitz. Hence, using $\mu_s \ll \mathcal{L}^d$ we find $\mu_s(T_{t \to s}(A) = 0$, such that (5.9) gives $\mu_t(A) = 0$. With this we conclude $\mu_s \ll \mathcal{L}^d$.

In the case of assumption (b), we argue as before but with $\mu_0 = c_0\mathcal{L}^d$ for $s = 0$. Using the fact that $q_{t \to 0}$ is locally bounded from below and that $T_{t \to 0}$ is locally Lipschitz on $A_t := \Xi_t \setminus \Xi^+$, we deduce that $\mu_t \llcorner A_t \ll \mathcal{L}^d$. On the other hand we have $\mu_1'' \ll \mathcal{L}^d$ and the restriction of $T_{t \to 1}$ to Ξ^+ coincides with the identity and $q_{t \to 1}$ is bounded from below thanks to (4.44). Thus, we obtain $\mu_t \ll \mathcal{L}^d$.

Assertion (2). The representation (5.6a) follows by Theorem 5.2 and Corollary 3.5.

Relation (5.7) can be deduced directly by Theorem 5.2. In order to prove that $D^2\xi_s = 0$ μ_s-a.e. in $\Xi^\pm$ it is sufficient to consider density points of $\Xi^\pm$, since $\mu_s \ll \mathcal{L}^d$, and to compute the differential of $\nabla\xi_s$ on $\Xi^\pm$, where it is constant.

Assertions (3) and (4). Both assertions follow from Corollary 3.5. □

As a last application, we will also discuss the propagation of the singular part with respect to $\mathcal{L}^d$, which will be needed in the proof of the main result in Theorem 7.2.

Corollary 5.6. (Propagation of the singular part) *Let* $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^d)$ *and let* $(\mu_t)_{t \in [0,1]}$ *be a geodesic connecting* μ_0 *to* μ_1 *and let* $\mu_s = \mu_s^a + \mu_s^\perp$ *be the decomposition of* μ_s *with respect to the Lebesgue measure* $\mathcal{L}^d$ *at some point* $s \in (0, 1)$. *For every* $t \in [0, 1]$ *we set*

$$\widetilde{\mu}_t := (\boldsymbol{T}_{s \to t}, q_{s \to t})_\star \mu_s^a \quad and \quad \widehat{\mu}_t := (\boldsymbol{T}_{s \to t}, q_{s \to t})_\star \mu_s^\perp. \tag{5.10}$$

Then, the curves $(\widetilde{\mu}_t)_{t \in (0,1)}$ *and* $(\widehat{\mu}_t)_{t \in (0,1)}$ *are* HK *geodesics, we have* $\widehat{\mu}_t \perp \mathcal{L}^d$ *for* $t \in [0, 1]$ *and* $\mu_t = \widetilde{\mu}_t + \widehat{\mu}_t$ *provides the Lebesgue decomposition for* $t \in (0, 1)$, *viz.* $\mu_t^a = \widetilde{\mu}_t$ *and* $\mu_t^\perp = \widehat{\mu}_t$.

Proof. Let us decompose Ξ_s in the disjoint union of two Borel sets A, B such that $\mu_s^a = \mu_s \llcorner A$ and $\mu_s^\perp = \mu_s \llcorner B$ with $\mathcal{L}^d(B) = 0$. By Theorem 5.4 we clearly have $\mu_t = \widetilde{\mu}_t + \widehat{\mu}_t$. On the one hand, $\widetilde{\mu}_t \ll \mathcal{L}^d$ by Corollary 5.5 for all $t \in (0, 1)$. On the other hand, for all $t \in [0, 1]$ the measure $\widehat{\mu}_t$ is concentrated on the set $\boldsymbol{T}_{s \to t}(B)$ which is $\mathcal{L}^d$-negligible, since $\boldsymbol{T}_{s \to t}$ is Lipschitz. If follows that $\widehat{\mu}_t \perp \mathcal{L}^d$, so that $\widehat{\mu}_t = \mu_t^a$ and $\widehat{\mu}_t = \mu_t^\perp$ for all $t \in (0, 1)$.

The fact that (μ_t^a) and $(\mu_t^\perp)$ are geodesics follows by Theorem 5.4 as well.   □

5.2. Convexity of the Lebesgue Density Along HK-Geodesics

In this subsection, we consider geodesics $(\mu_t)_{t \in [0,1]}$ such that $\mu_s \ll \mathcal{L}^d$ for some, and thus for all, $s \in (0, 1)$. We fix s and introduce the functions α_s, δ_s as in (5.6b) and the functions

$$\left. \begin{aligned} \gamma_s(t, x) &:= \alpha_s^{1/2}(t, x) = q_{s \to t}(x), \\ \rho_s(t, x) &:= \alpha_s^{1/2}(t, x) \delta_s^{1/d}(t, x) = q_{s \to t}(x) \big(\det \mathrm{D} \boldsymbol{T}_{s \to t}(x) \big)^{1/d} \end{aligned} \right\} \quad \text{for } x \in D_s. \tag{5.11}$$

We now exploit the explicit differential relations for $\gamma_s(t, x) = q_{s \to t}(x)$ and $\delta_s(t, x) = \det \mathrm{D} \boldsymbol{T}_{s \to t}(x)$ provided in Theorem 4.8 and derive lower estimates for $\ddot{\gamma}_s$ and $\ddot{\rho}_s$. It remains unclear whether the given choice for γ_s and ρ_s is the only possible, however it turns out that for these variables the following curvature estimates are relatively simple and hence the final convexity calculus goes through. For comparison, we mention that in the Kantorovich–Wasserstein case we have $\gamma_{\mathsf{K}}\mathsf{W}(t) \equiv 1$ and $\rho_{\mathsf{K}}\mathsf{W}(t) = \big(\delta_{\mathsf{K}}\mathsf{W}(t)\big)^{1/d}$ with $\delta_{\mathsf{K}}\mathsf{W}(t) = \det((1-t)\mathbb{I} + t\mathrm{D}\boldsymbol{T}_{\mathsf{K}}\mathsf{W}(x))$, such that $\ddot{\rho}_{\mathsf{K}}\mathsf{W}(t) \leqq 0$ since $\mathrm{D}\boldsymbol{T}_{\mathsf{K}}\mathsf{W}(x)$ is diagonalizable with nonnegative real eigenvalues, see [2, Eqn. (9.3.12)].

Proposition 5.7. (Curvature estimates for (ρ, γ)) *Let* $(\rho_s, \gamma_s) : (0, 1) \times \mathfrak{D}_s \to [0, \infty[^2$ *be defined as above along a geodesic. Then, we have for all,* $t \in (0, 1)$,

the relations

$$\frac{\ddot{\gamma}_s(t,x)}{\gamma_s(t,x)} \geqq 0 \quad and \quad \begin{cases} \dfrac{\ddot{\rho}_s(t)}{\rho_s(t)} \leqq \left(1-\dfrac{4}{d}\right)\dfrac{\ddot{\gamma}_s(t)}{\gamma_s(t)} & for\ d \geqq 2, \\[2ex] \dfrac{\ddot{\rho}_s(t)}{\rho_s(t)} = \left(1-\dfrac{4}{d}\right)\dfrac{\ddot{\gamma}_s(t)}{\gamma_s(t)} & for\ d = 1. \end{cases} \tag{5.12}$$

Proof. As $s \in (0,1)$ and $x \in \mathfrak{D}_s$ are fixed, we will simply write $\rho(t)$ instead of $\rho_s(t,x)$, and do similarly for the other variables. Using the specific definition of ρ we obtain

$$\frac{\ddot{\rho}}{\rho} = \frac{\ddot{\gamma}}{\gamma} + \frac{2}{d}\frac{\dot{\gamma}}{\gamma}\frac{\dot{\delta}}{\delta} + \frac{1}{d}\frac{\ddot{\delta}}{\delta} + \frac{1}{d}\left(\frac{1}{d}-1\right)\left(\frac{\dot{\delta}}{\delta}\right)^2.$$

We can now use the formulas provided in (4.62a)–(4.62d) giving $\dot{\gamma} = 2\xi_t\gamma$ and $\ddot{\gamma} = |\nabla\xi_t|^2\gamma$, where ξ_t and its derivatives are evaluated at $y = \boldsymbol{T}_{s\to t}(x)$. Inserting this and (4.62d) for $\dot{\delta}$ and $\ddot{\delta}$ into the above relation for $\ddot{\delta}/\delta$ we observe significant cancellations and obtain

$$\frac{\ddot{\gamma}}{\gamma} = |\nabla\xi_t|^2 \quad and \quad \frac{\ddot{\rho}}{\rho} = \frac{1}{d^2}\left((\Delta\xi_t)^2 - d|\mathrm{D}^2\xi_t|^2\right) + \left(1 - \frac{4}{d}\right)|\nabla\xi_t|^2. \tag{5.13}$$

For $d = 1$ we have $\mathrm{D}^2\xi = \Delta\xi$, while for $d \geq 2$ all matrices $A \in \mathbb{R}^{d\times d}$ satisfy $d|A|^2 = d\sum_{i,j=1}^d A_{ij}^2 \geqq (\operatorname{tr} A)^2 = \left(\sum_1^d A_{ii}\right)^2$. Thus, the curvature estimates (5.12) follow. $\quad\square$

The above curvature estimates will be crucial in Section 7 for deriving our main result on geodesic convexity. We remark that for $d \geq 2$ they are even slightly better that the "sufficient curvature estimates" given in (7.3) because of $1-4/d \leq 1-4/d^2$ (with equality only for $d = 1$).

We finally derive a useful result concerning the convexity of the density $t \mapsto c(t,x)$ along geodesics. This provides a direct proof of the fact, which was used in [13] that the L^∞-norm along geodesics is bounded by the L^∞-norm of the two endpoints. Indeed, we show more, namely that the function $t \mapsto c(t,\boldsymbol{T}_t(x))$ is either trivially constant or it is strictly convex.

Theorem 5.8. (Convexity of densities along geodesics)

(1) *Under the assumption of Corollary 5.5, for every $s \in (0,1)$ and $x \in \mathfrak{D}_s \cup \Xi^\pm$ the function $c_s(t) = c(t,\boldsymbol{T}_{s\to t}(x))$ given by (5.6a) or (5.7), respectively, is convex and positive in $(0,1)$; moreover, with a possible $\mathcal{L}^d$-negligible exception, it is either constant or strictly convex.*

(2) *If moreover $\mu_0 \ll \mathcal{L}^d$ (resp. $\mu_1 \ll \mathcal{L}^d$) then for μ_s-a.e. x their limit as $t \downarrow 0$ (resp. as $t \uparrow 1$) coincides with $c_0 \circ \boldsymbol{T}_{s\to 0}$ (resp. $c_1 \circ \boldsymbol{T}_{s\to 1}$).*

Proof. Assertion (1). Since $x \in \mathbb{R}^d$ and $s \in (0,1)$ play no role, we drop them for notational simplicity. We simply calculate the second derivative of the function $t \mapsto c(t) = \gamma(t)^{d+2}c_s/\rho(t)^d$. If $c_s = c(s,x) = 0$ then $c(t,\boldsymbol{T}_{s\to t}(x)) = 0$ and the

result is obviously true. Hence, we may assume $c_s > 0$ and obtain after an explicit calculation

$$\ddot{c} = c\left((d+2)\frac{\ddot{\gamma}}{\gamma} - d\frac{\ddot{\rho}}{\rho} + (d+1)(d+2)\left(\frac{\dot{\gamma}}{\gamma}\right)^2 - 2d(d+2)\frac{\dot{\gamma}}{\gamma}\frac{\dot{\rho}}{\rho} + d(d+1)\left(\frac{\dot{\rho}}{\rho}\right)^2\right). \tag{5.14}$$

The quadratic form involving the first derivatives is positive definite, and for the terms involving the second derivatives we can use the curvature estimates in (5.12) to obtain

$$\ddot{c} \geqq c\left((d+2)\frac{\ddot{\gamma}}{\gamma} - d\left(1 - \frac{4}{d}\right)\frac{\ddot{\gamma}}{\gamma} + 0\right) = 6\,c\,\frac{\ddot{\gamma}}{\gamma}.$$

Notice that $t \mapsto \gamma(t)$ is the square root of the non-negative (and strictly positive in $(0, 1)$) quadratic polynomial $\alpha(\cdot, x)$ given by (5.6b), so that $\gamma'' \geqq 0$ and we conclude that $\ddot{c}(t) \geqq 0$ as well due to $c(t) > 0$.

Moreover, if $x \notin \Xi_s^0$ then $|\nabla \xi_s(x)| > 0$, have $\ddot{\gamma}(t) > 0$, and we deduce that $\ddot{c}(t) > 0$ obtaining the strict convexity of c.

If $x \in \Xi_s^0$ where $\nabla \xi_s(x) = 0$, we can use the representation (5.7) for c up to a $\mathcal{L}^d$-negligible set.

Assertion (2). If $\mu_0 = c_0 \mathcal{L}^d \ll \mathcal{L}^d$, then $\delta_s(0, x) > 0$ for μ_s-a.a. $x \in \mathfrak{D}_s$ thanks to the last statement of Corollary 3.5 (which is a direct consequence of Theorem 3.3(5)) and both $\delta_s(0, x)$ and $\alpha_s(0, x)$ coincides with their limit as $t \downarrow 0$. A further application of Corollary 5.5(3) yields the result. The case $t = 1$ is completely analogous.   □

The above result easily provides the following statement on convexity of L^∞ norms along HK-geodesics. This can be generalized to a corresponding result for the Kantorovich–Wasserstein geodesics (which might been known, but the authors were not able to identify a reference, see the Remark 5.10 below):

Corollary 5.9. (Convexity of the L^∞ norm along geodesics) *Let $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^d)$ be absolutely continuous with respect to $\mathcal{L}^d$ with densities $c_i \in L^\infty(\mathbb{R}^d)$ and let $(\mu_t)_{t \in [0,1]}$ be a HK geodesic connecting μ_0 to μ_1. Then $\mu_t = c_t \mathcal{L}^d$ and $\|c_t\|_{L^\infty} \leqq (1-t)\|c_0\|_{L^\infty} + t\|c_1\|_{L^\infty}$.*

Proof. The result for $(\mathcal{M}(\mathbb{R}^d), \text{HK})$ follows directly from Theorem 5.8.   □

Remark 5.10. Let $(\mu_t^{\text{W}})_{t \in [0,1]}$ be the Kantorovich–Wasserstein geodesic connection between two probability measures $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$ with $\mu_i = c_i \mathcal{L}^d$ and $c_0, c_1 \in L^\infty(\mathbb{R}^d)$. Similar to the previous result, $\mu_t^{\text{W}} = c_t^{\text{W}} \mathcal{L}^d$ is absolutely continuous w.r.t. $\mathcal{L}^d$ and $\|c_t^{\text{W}}\|_{L^\infty} \leqq (1-t)\|c_0^{\text{W}}\|_{L^\infty} + t\|c_1^{\text{W}}\|_{L^\infty}$.

In fact, for $(\mathcal{P}_2(\mathbb{R}^d), \text{W}_2)$ we replace (5.6) by the simpler formula for the Kantorovich–Wasserstein transport

$$c^{\text{W}}(t, \mathbf{T}_{s \to t}(x)) = \frac{c_s^{\text{W}}(x)}{\delta_s(t, x)} \quad \text{with } \delta_s(x) = \det \mathbf{T}_{s \to t}^{\text{W}}(x);$$

see [2, Prop. 9.3.9]. Using $\mu_0 = c_0 \mathcal{L}^d$ we can choose $s = 0$ and have $T^{\mathrm{W}}_{0 \to t}(x) = x + t(\nabla\varphi(x) - x)$ for a convex Kantorovich potential. Since for every symmetric positive semidefinite matrix D the function $t \mapsto 1/\det\big((1-t)I + tD\big)$ is convex, the desired result follows with the same arguments as for Theorem 5.8.

6. Preliminary Discussion of the Convexity Conditions

In this section, we discuss the equivalence of two formulations of the convexity conditions and give a few examples. The proof of sufficiency and necessity of these conditions is then given in the following Section 7.

For most parts of this section, we assume that $E : [0, \infty[\to \mathbb{R} \cup \{\infty\}$ is lower semi-continuous and convex, satisfies $E(0) = 0$, and is twice continuously differentiable on the interior of its domain $D(E) := \{c \geq 0 \mid E(c) < \infty\}$. The following result gives a characterization of the conditions (1.25) on $N_E : (\rho, \gamma) \mapsto (\rho/\gamma)^d E(\gamma^{d+2}/\rho^d)$ in terms of the derivatives of E, namely $\varepsilon_j(c) = c^j E^{(j)}(c)$ for $j = 0, 1$, and 2, which appear in

$$\mathbb{B}(c) := \begin{pmatrix} \varepsilon_2(c) - \frac{d-1}{d}\big(\varepsilon_1(c) - \varepsilon_0(c)\big) & \varepsilon_2(c) - \frac{1}{2}\big(\varepsilon_1(c) - \varepsilon_0(c)\big) \\ \varepsilon_2(c) - \frac{1}{2}\big(\varepsilon_1(c) - \varepsilon_0(c)\big) & \varepsilon_2(c) + \frac{1}{2}\varepsilon_1(c) \end{pmatrix}. \quad (6.1)$$

This characterization will then be used to derive a nontrivial monotonicity result in Proposition 6.2, which is a crucial building block of the main geodesic convexity result.

Note that the variables ρ and γ are related to the variable c via $c = c_0 \gamma^{d+2}/\rho^d$.

Proposition 6.1. (Equivalent conditions on E) *Let N_E and $\mathbb{B}$ be defined in terms of E as in (1.25a) and (6.1), respectively. Then the following conditions are equivalent:*

(A) *N_E satisfies (1.25);*
(B) *in the interior of the domain $D(E)$ we have $\mathbb{B}(c) \geq 0$ and $(d-1)\big(\varepsilon_1(c) - \varepsilon_0(c)\big) \geq 0$.*

Proof. We first observe that the desired monotonicity of $\rho \mapsto N_E(\rho, \gamma)$ for $d \geq 2$ is indeed equivalent to the condition $\varepsilon_1(c) \geq \varepsilon_0(c)$. This follows easily from the relation

$$\begin{aligned} \partial_\rho N_E(\rho, \gamma) &= \frac{d\rho^{d-1}}{\gamma^d} E\Big(\frac{\gamma^{d+2}}{\rho^d}\Big) + \frac{\rho^d}{\gamma^d} E'\Big(\frac{\gamma^{d+2}}{\rho^d}\Big)\Big(-d\frac{\gamma^{d+2}}{\rho^{d+1}}\Big) \\ &= \frac{d\rho^{d-1}}{\gamma^d}\big(\varepsilon_0(c) - \varepsilon_1(c)\big). \end{aligned}$$

It remains to establish the equivalence between the convexity of N_E and the positive semi-definiteness of $\mathbb{B}$. For this we note that N_E is given as a linear function of E, hence the Hessian $\mathrm{D}^2 N_E$ will be a given as a linear combination of E, E', and E''. Indeed, an explicit calculation yields

$$\mathrm{D}^2 N_E(\rho, \gamma) = \frac{\rho^d}{\gamma^d} \begin{pmatrix} d/\rho & -d/\gamma \\ 0 & -2/\gamma \end{pmatrix}^\top \mathbb{B}\Big(\frac{\gamma^{d+2}}{\rho^d}\Big) \begin{pmatrix} d/\rho & -d/\gamma \\ 0 & -2/\gamma \end{pmatrix}.$$

With this, we see that $D^2 N_E$ is positive semidefinite if and only if $\mathbb{B}$ is. Hence, the assertion is proved.   □

From the semi-definiteness of the matrix $\mathbb{B}(c)$, we obtain as necessary conditions the non-negativity of the two diagonal elements which provide the McCann condition $\mathbb{B}_{11} = \varepsilon_2 - \frac{d-1}{d}(\varepsilon_1 - \varepsilon_0) \geq 0$ and the convexity conditions with respect to the Hellinger–Kakutani distance $\mathbb{B}_{22} = \varepsilon_2 + \frac{1}{2}\varepsilon_1 \geqq 0$. Moreover, testing $\mathbb{B}$ with $(1, -1)^\top$ reveals the additional condition

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix} \cdot \mathbb{B}(c) \begin{pmatrix} 1 \\ -1 \end{pmatrix} \geqq 0 \quad \Longleftrightarrow \quad (d+2)\varepsilon_1(c) - 2\varepsilon_0(c) \geqq 0. \tag{6.2}$$

Proposition 6.2. (New necessary monotonicity) *Let E be such that the conditions in Proposition 6.1 hold and let N_E be defined via (1.25a). Then, the following three equivalent conditions hold:*

(A) *The function $]0, \infty[\ni c \mapsto c^{-2/(d+2)} E(c)$ is non-decreasing.*
(B) *For all $\rho, \gamma > 0$ we have the inequality $\left(1 - \frac{4}{d^2}\right)\rho\,\partial_\rho N_E(\rho, \gamma) + \gamma\,\partial_\gamma N_E(\rho, \gamma) \geqq 0$.*
(C) *For all $\rho, \gamma > 0$ the mapping $]0, \infty[\ni s \mapsto N_E(s^{1-4/d^2}\rho, s\gamma)$ is non-decreasing.*

Proof. Expressing $\partial_\rho N_E$ and $\partial_\gamma N_E$ via ε_0 and ε_1 and using $\delta = (\rho/\gamma)^d$ we obtain

$$\rho\,\partial_\rho N_E(\rho, \gamma) = -d\delta(\varepsilon_1 - \varepsilon_0) \quad \text{and} \quad \gamma\,\partial_\gamma N_E(\rho, \gamma) = \delta\big((d+2)\varepsilon_1 - d\varepsilon_0\big).$$

Thus, we conclude $(1 - \frac{4}{d^2})\rho\,\partial_\rho N_E(\rho, \gamma) + \gamma\,\partial_\gamma N_E(\rho, \gamma) = \frac{2\delta}{d}\big((d+2)\varepsilon_1 - 2\varepsilon_0\big)$, which is positive because of (6.2). Thus, (B) is established and the monotonicity of $s \mapsto N_E(s^{1-4/d^2}\rho, s\gamma)$ in (C) follows simply by differentiation.

Statement (A) follows by applying (C) for $\rho = \gamma = 1$ and choosing $s = c^{2(d+2)/d}$.   □

The crucial monotonicity stated at the end of the above proposition means

$$0 \leqq c_1 < c_2 \quad \Longrightarrow \quad E(c_1) \leqq \left(\frac{c_1}{c_2}\right)^{2/(d+2)} E(c_2). \tag{6.3}$$

It implies that if E attains a negative value it cannot be differentiable at $c = 0$: If $E(c_1) < 0$ then $E(c) \leqq (c/c_1)^{2/(d+2)} E(c_1) < 0$, which leads to $E'(c) \searrow -\infty$ for $c \searrow 0$.

In the next examples we investigate which functions E satisfy the above conditions. The following two results will be used in Corollary 7.3 to obtain geodesic convexity for functionals of the form $E(c) = \int_\Omega ac^r \, \mathrm{d}x$. The third example shows that in case of the Boltzmann entropy with $E(c) = c \log c$ the conditions do not hold and hence geodesic convexity fails.

Example 6.3. (**Density function $E(c) = c^m$**) We have $\varepsilon_0(c) = c^m$, $\varepsilon_1(c) = mc^m$, and $\varepsilon_2(c) = m(m-1)c^m$, which gives the matrix

$$\mathbb{B}(c) = c^m \begin{pmatrix} (m-1)\big(m - \frac{d-1}{d}\big) & (m-1)\big(m - \frac{1}{2}\big) \\ (m-1)\big(m - \frac{1}{2}\big) & m\big(m - \frac{1}{2}\big) \end{pmatrix}.$$

The Hellinger condition $\mathbb{B}_{22}(c) \geqq 0$ holds for $m \notin {]}0, \frac{1}{2}[$, while the McCann condition $\mathbb{B}_{11}(c) \geqq 0$ holds for $m \notin {]}\frac{d-1}{d}, 1[$. Moreover, for $d \geqq 2$ the monotonicity condition $\varepsilon_1 \geqq \varepsilon_0$ implies $m \geqq 1$.

Thus, the remaining cases are either $m \geqq 1$ or $d = 1$ and $m \leqq 0$, and it remains to check $\det \mathbb{B}(c) \geqq 0$. An explicit calculation gives

$$\det \mathbb{B}(c) = (m-1)\left(m - \frac{1}{2}\right)\frac{(d+2)m - d}{2d}.$$

Clearly, for $m \geqq 1$ we have $\det \mathbb{B}(c) \geqq 0$ for all space dimensions $d \in \mathbb{N}$. Moreover, $\det \mathbb{B}(c) < 0$ for $m \leqq 0$.

In summary, we obtain geodesic convexity if and only if $m \geqq 1$.

Example 6.4. (**Density function $E(c) = -c^q$**) As in the previous example we have

$$\mathbb{B}(c) = c^q \begin{pmatrix} (1-q)\left(q - \frac{d-1}{d}\right) & (1-q)\left(q - \frac{1}{2}\right) \\ (1-q)\left(q - \frac{1}{2}\right) & q\left(\frac{1}{2} - q\right) \end{pmatrix}.$$

The Hellinger condition $\mathbb{B}_{22}(c) \geqq 0$ holds for $q \in [0, \frac{1}{2}]$, while the McCann condition $\mathbb{B}_{11}(c) \geqq 0$ holds for $q \in [\frac{d-1}{d}, 1]$, which also implies the monotonicity $\varepsilon_1 \geqq \varepsilon_0$. With

$$\det \mathbb{B}(c) = (1-q)\left(\frac{1}{2} - q\right)\frac{(d+2)q - d}{2d} c^{2q}.$$

we obtain the additional condition $q \geqq d/(d+2)$ and summarize that $E(c) = -c^q$ leads to a geodesically convex functional if and only if $q \in \left[\max\{\frac{d-1}{d}, \frac{d}{d+2}\}, \frac{1}{2}\right]$, which has solutions only for $d = 1$ and $d = 2$.

Example 6.5. (**Boltzmann entropy**) As a negative example where the geodesic convexity fails, we consider the Boltzmann function $E(c) = c \log c$. We compute $\mathbb{B}_{22}(c) = \varepsilon_2(c) + \frac{1}{2}\varepsilon_1(c) = \frac{3}{2}c + \frac{1}{2}c \log c$, which shows that the necessary Hellinger condition fails. Moreover, considering the measures $\mu_0 = 0$ and $\mu_1 = c\mathcal{L}^d$ for a non-negative density $c \in L^1(\Omega)$ we find that along the geodesic curve, given by $\mu(s) = s^2\mu_1$, we have

$$\mathscr{E}(\mu(s)) = \int_\Omega E(s^2 c)\,dx = s^2\mathscr{E}(\mu_1) + 2s^2 \log(s)\int_\Omega c\,dx,$$

which is clearly not convex if $\int_\Omega c\,dx = \mu_1(\Omega) > 0$.

Finally, we discuss a few examples where the density function E is not smooth. Note that the conditions in (1.25) form a closed cone. Moreover, as for convex functions, the supremum $E : c \mapsto \sup\{\widetilde{E}_\alpha(c) \mid \alpha \in A\}$ satisfies (1.25) if all $\widetilde{E}_\alpha$ do so.

Example 6.6. (**Nonsmooth** E) In applications one is also interested in cases where E is nonsmooth. For example the case $E_\kappa(c) = \kappa c$ for $c \in [0, c^*]$ and $E(c) = \infty$ for $c > c_*$ is considered in [13]. Clearly, E_0 satisfies our assumptions (1.25) since N_E only takes the values 0 and ∞ and the value 0 is taken on the convex set $\gamma^{d+2} \leqq c^* \rho^d$. Thus, E_κ generates a functional $\mathscr{E}_\kappa = \mathscr{E}_0 + \kappa \mathscr{M}$ that is geodesically 2κ-convex.

A second example is given by $E(c) = \max\{0, c^2 - c\}$. We first observe that $\widetilde{E}_1(c) = c$ and $\widetilde{E}_2(c) = c^2$ satisfy (1.25). Hence, $c \mapsto \max\{\widetilde{E}_1(c), \widetilde{E}_2(c)\} = E(c) + c$ satisfies (1.25) as well. Thus, we know that E generates a functional $\mathscr{E}$ that is at least geodesically (-2)-convex. However, we may inspect the function $c \mapsto c^2 - c$ in the region $c \geqq 1$ directly and find that E itself satisfies (1.25).

In practical applications, in particular for evolutionary variational inequalities as treated in [23], it is desirable to find the optimal λ for the geodesic λ-convexity. So far, we have treated the case of geodesic 0-convexity and now return to the general case, which leads to the conditions

$$\mathbb{B}(c) \geqq \begin{pmatrix} 0 & 0 \\ 0 & \lambda c/2 \end{pmatrix} \quad \text{and} \quad (d-1)\big(\varepsilon_1(c) - \varepsilon_0(c)\big) \geqq 0.$$

The monotonicity condition is clearly independent of λ. The first equation still relies on the necessary McCann condition $\mathbb{B}_{11}(c) \geqq 0$. If this holds with strict inequality we see that the optimal λ is characterized by

$$\lambda_{\text{opt}} = \inf\left\{ \frac{2 \det \mathbb{B}(c)}{c \mathbb{B}_{11}(c)} \,\Big|\, c > 0 \right\}. \tag{6.4}$$

Example 6.7. ($d = 1$ and $E(c) = c^2 - c^{2/5}$) From the previous examples, we know that $E^2(c) = c^2$ and $E^{2/5}(c) = -c^{2/5}$ are both geodesically 0-convex, and we want to show that the sum is geodesically λ-convex for $\lambda > 0$. As $\mathbb{B}$ is linear in E we have $\mathbb{B}(c) = c^2 \mathbb{B}^{(2)} + c^{2/5} \mathbb{B}^{(2/5)}$ with constant matrices $\mathbb{B}^{(2)}$ and $\mathbb{B}^{(2/5)}$ that are both strictly positive definite. Thus,

$$\ell(c) := \frac{2 \det \mathbb{B}(c)}{c \, \mathbb{B}_{11}(c)} > 0 \quad \text{for all } c > 0.$$

Moreover, we find $\ell(c) \sim 2c^{-3/5} \det \mathbb{B}^{(2/5)} / \mathbb{B}_{11}^{(2/5)}$ for $c \approx 0$ and $\ell(c) \sim 2c \det \mathbb{B}^{(2)} / \mathbb{B}_{11}^{(2)}$ for $c \gg 1$. Thus, by compactness $\lambda_{\text{opt}} = \inf\{\ell(c) \,|\, c > 0\}$ is strictly positive.

Numerically, we find $\lambda_{\text{opt}} \approx 0.638$ which is attained at $c_* \approx 0.0319$.

Remark 6.8. (**Geodesic convexity via the Otto calculus**) Following the key ideas in [12,31] a formal calculus for reaction-diffusion systems was developed in [25]. It uses the dynamical formulation in Subsection 2.1.1 and the associated Onsager operator $\mathbb{K}(c)\xi = -\alpha \, \text{div}(c\nabla\xi) + \beta c\xi$ to characterize the geodesic λ-convexity of the functional $\mathscr{E}$ by calculating the quadratic form $M(c, \cdot)$ (contravariant Hessian of $\mathscr{E}$):

$$M(c, \xi) = \langle \xi, \mathbf{D}\mathbf{V}(c)\mathbb{K}(c)\xi \rangle - \frac{1}{2}\mathbf{D}_c\langle \xi, \mathbb{K}(c)\xi \rangle[\mathbf{V}(c)] \quad \text{with } \mathbf{V}(c) = \mathbb{K}(c)\mathbf{D}\mathscr{E}(c).$$

Then, one needs to show the estimate $M(c, \xi) \geq \lambda \langle \xi, \mathbb{K}(c)\xi \rangle$.

Following the methods in [25, Sect. 4], for $c \in C_c^0(\Omega)$ and smooth ξ we obtain

$$
M(c, \xi) = \int_\Omega \left[\alpha^2 \left((A(c) - H(c))(\Delta\xi)^2 + H(c)|D^2\xi|^2 \right) \right.
$$
$$
\left. + \alpha\beta \left(B_1(c)|\nabla\xi|^2 + B_2(c)\xi\Delta\xi \right) + \beta^2 B_3(c)\xi^2 \right] dx,
$$

where $A(c) = \varepsilon_2(c), \quad H(u) = \varepsilon_1(c) - \varepsilon_0(c), \quad B_1(c) = \dfrac{3}{2}\varepsilon_1(c) - \varepsilon_0(c),$

$$
B_2(c) = -2\varepsilon_2(c) + \varepsilon_1(c) - \varepsilon_0(c), \quad B_3(c) = \varepsilon_2(c) + \frac{1}{2}\varepsilon_1(c).
$$

Analyzing the condition $M(c, \xi) \geq \lambda \langle \xi, \mathbb{K}(c)\xi \rangle$, we find the conditions

$$
\forall c \geq 0: \quad (d-1)H(c) \geq 0,
$$
$$
B_1(c) \geq \frac{\lambda}{\beta}c, \quad \begin{pmatrix} A(c) - \frac{d-1}{d}H(c) & \frac{1}{2}B_2(c) \\ \frac{1}{2}B_2(c) & B_3(c) - \frac{\lambda}{\beta}c \end{pmatrix} \geq 0, \tag{6.5}
$$

which for $\lambda = 0$ give the same conditions as $\mathbb{B}(c) \geq 0$, see Proposition 6.1. Note that the middle estimate in (6.5) follows from the first and the third estimates because of

$$
B_1(c) = \tfrac{3}{2}\varepsilon_1 - \varepsilon_0 = \tfrac{d-1}{d}(\varepsilon_1 - \varepsilon_0) + \tfrac{1}{2d}\left((d+2)\varepsilon_2 - 2\varepsilon_0 \right).
$$

7. Proof of Geodesic Convexity of $\mathscr{E}$

In this section, we finally prove the necessity and sufficiency of the conditions for geodesic convexity of functionals $\mathscr{E}$ on $\mathcal{M}(\Omega)$ in (1.25), where we now allow for a general closed and convex domain $\Omega \subset \mathbb{R}^d$. In order to keep the arguments clear, we first restrict ourselves to absolutely continuous measures μ_0 and μ_1. Thus, by Corollary 5.5 the connecting geodesic curves are also absolutely continuous, and we can rewrite $\mathscr{E}$ along the latter in the form

$$
\mathscr{E}(\mu_t) = \int_\Omega E(c(t, y)) \, dy = \int_\Omega e(t, x) \, dx, \quad \text{where}
$$
$$
e(t, x) = \delta(t, x) E\left(c_*(x) \frac{\alpha(t, x)}{\delta(t, x)} \right).
$$

The general case will then be treated by using an approximation argument.

Under the assumption that E is twice differentiable in the interior of its domain, we show that for μ_0-a.a. $x \in \Omega$ the function $t \mapsto e(t, x)$ is convex. Since $\alpha(\cdot, x)$ and $\delta(\cdot)$ are analytic functions on $[0, 1]$, we can show convexity in this case by establishing $\ddot{e}(t, x) \geq 0$. For this, we can fix $x \in \Omega$, drop the dependence on x for notational convenience, and set

$$
e(t) = \delta(t) E\left(c_* \frac{\alpha(t)}{\delta(t)} \right) = N_E\big(\rho(t), \gamma(t) \big) \quad \text{with } \rho := (c_*\alpha)^{1/2} \delta^{1/d},
$$

$$\gamma := (c_*\alpha)^{1/2}, \tag{7.1}$$

and N_E from (1.25a). Now, the classical chain rule implies the relation

$$\ddot{e} = \left\langle \begin{pmatrix} \dot{\rho} \\ \dot{\gamma} \end{pmatrix}, D^2 N_E(\rho, \gamma) \begin{pmatrix} \dot{\rho} \\ \dot{\gamma} \end{pmatrix} \right\rangle + \partial_\rho N_E(\rho, \gamma)\ddot{\rho} + \partial_\gamma N_E(\rho, \gamma)\ddot{\gamma}. \tag{7.2}$$

The aim is to show $\ddot{e}(t) \geqq 0$ for all $t \in [0, 1]$. By the convexity of N_E it suffices to treat the last two terms.

For this we exploit the curvature estimates (5.12) on $\ddot{\gamma}$ and $\ddot{\rho}$ as well as the monotonicities in (1.25c) and Proposition 6.2.

7.1. Usage of the Curvature Estimates

We first show that it is sufficient to use the curvature estimates

$$\frac{\ddot{\gamma}}{\gamma} \geqq 0 \quad \text{and} \quad \begin{cases} \dfrac{\ddot{\rho}}{\rho} \leqq \left(1 - \dfrac{4}{d^2}\right)\dfrac{\ddot{\gamma}}{\gamma} & \text{for } d \geq 2, \\[2mm] \dfrac{\ddot{\rho}}{\rho} = \left(1 - \dfrac{4}{d^2}\right)\dfrac{\ddot{\gamma}}{\gamma} & \text{for } d = 1. \end{cases} \tag{7.3}$$

In particular, the equality condition for $d = 1$ is different from the inequality conditions for $d \geq 2$. This will be used to compensate for the missing monotonicity of N_E in (1.25c) in the case $d = 1$.

Below we will see that the curvature estimates (7.3) are *necessary* to complete our proof. Note that they are implied by the curvature estimates derived in Proposition 5.7. In fact, both coincide for $d = 1$, while for $d \geq 2$ the former are strictly weaker as the latter because of $1 - 4/d < 1 - 4/d^2$.

Proposition 7.1. ($\ddot{e} \geq 0$ via curvature estimates) *Assume that N_E satisfies (1.25) and that $t \mapsto (\rho(t), \gamma(t))$ satisfies (7.3), then $\ddot{e} \geq 0$ in (7.2).*

Proof. As the first term (involving $D^2 N_E$) on the right-hand side of (7.2) is nonnegative, we only have to show that the last two terms have a non-negative sum. For this we rearrange terms as follows:

$$\partial_\rho N_E(\rho, \gamma)\ddot{\rho} + \partial_\gamma N_E(\rho, \gamma)\ddot{\gamma}$$
$$= \left(-\rho \partial_\rho N_E\right)\left(\left(1 - \frac{4}{d^2}\right)\frac{\ddot{\gamma}}{\gamma} - \frac{\ddot{\rho}}{\rho}\right) + \left(\gamma \partial_\gamma N_E + \left(1 - \frac{4}{d^2}\right)\rho \partial_\rho N_E\right)\frac{\ddot{\gamma}}{\gamma}.$$

The right-hand side is the sum of two products, both of which are non-negative. Indeed, the first product equals 0 in the case $d = 1$ independently of the sign of $\partial_\rho N_E$, because the second factor is 0. In the case $d \geq 2$ both factors are nonnegative (using $\partial_\rho N_E \leq 0$ and the second curvature estimate in (7.3)), so the first product is non-negative again.

In the second product both terms are non-negative by Proposition 6.2(B) and the first curvature estimate in (7.3). Thus, $\ddot{e} \geqq 0$ in (7.2) is proved.   □

7.2. The Main Results on Geodesic λ-Convexity

We are now ready to establish our main result on the geodesic convexity of functionals $\mathscr{E}$ given in terms of a density E. We now make our general assumptions of E precise.

$$E : [0, \infty[\to \mathbb{R} \cup \{\infty\} \text{ is lower semi-continuous, convex,}$$
$$E(0) = 0, \text{ and there exists } c_\circ > 0 \text{ such that } E(c_\circ) < \infty. \tag{7.4a}$$

We also want to include the case that E is not necessarily superlinear, so we introduce the recession constant

$$E'_\infty := \lim_{c \to \infty} \frac{1}{c} E(c) \in \mathbb{R} \cup \{\infty\}.$$

The case $E'_\infty = \infty$ is the superlinear case where the functional $\mathscr{E}(\mu)$ is always $+\infty$, if μ has a singular part, i.e. $\mu^\perp \neq 0$ in the decomposition $\mu = c\mathcal{L}^d + \mu^\perp$ with $\mu^\perp \perp \mathcal{L}^d$.

We introduce a closed (convex) domain $\Omega \subset \mathbb{R}^d$, and we consider the set of measures μ with support contained in Ω, which we identify with $\mathcal{M}(\Omega)$. In the case that the right derivative $E'_0 := \lim_{c \downarrow 0} \frac{1}{c} E(c)$ of E at 0 is not finite, we further have to impose that Ω has finite Lebesgue measure. Therefore, we will assume that

$$\Omega \text{ is a closed convex set with nonempty interior and}$$
$$\Omega \text{ is also bounded, i.e. } \mathcal{L}^d(\Omega) < \infty, \text{ if } E'_0 = -\infty. \tag{7.4b}$$

Thus, the functionals $\mathscr{E}$ are defined as follows:

$$\mathscr{E}(\mu) = \int_\Omega E(c(x))\, \mathrm{d}x + E'_\infty \mu^\perp(\Omega) \quad \text{for} \quad \mu = c\mathcal{L}^d + \mu^\perp \text{ with } \mu^\perp \perp \mathcal{L}^d. \tag{7.5}$$

It is well known that (7.4) guarantees that $\mathscr{E}$ is a weakly lower semi-continuous functional on $\mathcal{M}(\Omega)$. In particular, condition (7.4b) is necessary to guarantee that the negative part $x \mapsto \min\{E(c(x)), 0\}$ is integrable, because for $c \in L^1(\Omega)$ the functions $x \mapsto -\sqrt{c(x)}$ may not lie in $L^1(\Omega)$. We refer to Example 7.4 for a case where (7.4b) can be avoided by using a confining potential.

We are now in the position to formulate our main result on the geodesic λ-convexity of integral functionals $\mathscr{E}$ on the Hellinger–Kantorovich space $(\mathcal{M}(\Omega), \mathsf{HK})$. The proof consists of three steps. First, we assume that E is twice continuously differentiable in its domain. Restricting to geodesic curves connecting absolutely continuous measures, we can use the above differentiable theory giving $\ddot{e} \geqq 0$. In Step 2, we generalize to possibly non-differentiable density functions E, but keep absolutely continuous measures. For smoothing a given E, we use that whenever E solves the conditions (1.25) and (7.4) then $c \mapsto E(rc)$ does so for each $r \in [0, 1]$. With a multiplicative convolution we construct a smooth E_δ to which Step 1 applies. Finally, Step 3 handles the case where $\mu_0^\perp$ or $\mu_1^\perp$ are non-zero by a standard approximation argument of general measures using absolutely continuous measures.

Theorem 7.2. (Geodesic convexity of $\mathscr{E}$) *Assume that $E : [0, \infty[\to \mathbb{R} \cup \{\infty\}$ and $\Omega \subset \mathbb{R}^d$ satisfy* (7.4a) *and* (7.4b), *respectively. If for a $\lambda_* \in \mathbb{R}$ the function*

$$N_{\lambda_*, E}(\rho, \gamma) := \left(\frac{\rho}{\gamma}\right)^d E\left(\frac{\gamma^{d+2}}{\rho^d}\right) - \frac{\lambda_*}{2}\gamma^2, \quad \text{for } \rho, \gamma > 0,$$

satisfies the conditions (1.25b) *and* (1.25c), *then the functional $\mathscr{E}$ defined in* (7.5) *is geodesically λ_*-convex on* $(\mathcal{M}(\Omega), \mathsf{HK})$.

Proof. Without loss of generality, we set $\lambda_* = 0$ throughout the proof and shortly write $N_E = N_{\lambda_*, E}$.

Step 1: The smooth and absolutely-continuous case. We first assume that E is twice continuously differentiable in the interior $]0, c_E[$ of its domain and that the measures μ_0 and μ_1 are absolutely continuous with respect to $\mathcal{L}^d$, i.e. $\mu_j = c_j \mathcal{L}^d$ for $c_j \in L^1(\Omega)$.

We fix $s \in (0, 1)$ adopting the notation of Corollary 5.5. Then, the geodesic curve $t \mapsto \mu_t = c(t, \cdot)\mathcal{L}^d$ satisfies

$$\mathscr{E}(\mu_t) = \int_\Omega E(c(t, y)) \, \mathrm{d}y = \int_\Omega E\left(c_s(x)\frac{\alpha_s(t, x)}{\delta_s(t, x)}\right)\delta_s(t, x) \, \mathrm{d}x = \int_\Omega e(t, x) \, \mathrm{d}x$$

with $e(t, x) = N_E(\rho(t, x), \gamma(t, x))$ as above. We want to show that for a.a. $x \in \Omega$ the function $t \mapsto e(t, x)$ is convex.

As shown in Theorem 5.8 the functions $t \mapsto \widetilde{c}(t, x) = c_0(x)\frac{\alpha(t,x)}{\delta(t,x)} = \gamma(t, x)^{d+2}$ $/\rho(t, x)^d$ are either constant or strictly convex. If the function $\widetilde{c}(\cdot, x)$ is constant then either $c_0(x) = 0$ or $(\rho(\cdot, x), \gamma(\cdot, x))$ is constant. In both cases, $e(\cdot, x)$ is constant as well, and hence convex.

In the strictly convex case, the values of $\widetilde{c}(t, x)$ for $t \in]0, 1[$ lie in the interior of the domain of E, where E is twice differentiable. Hence, combining Propositions 7.1 and 5.7 shows that $t \mapsto e(t, x)$ is convex for a.a. $x \in \Omega$. Since integration over Ω maintains convexity we conclude that $t \mapsto \mathscr{E}(\mu_t)$ is convex, too.

Step 2: The nonsmooth but absolutely-continuous case. We still assume $\mu_j = c_j \mathcal{L}^d$, but now consider an E that is not necessarily twice differentiable, but still satisfies (7.4). We choose a function $\chi \in C_c^\infty(\mathbb{R})$ satisfying $\chi(r) \geqq 0$, $\int_{-2}^{-1} \chi(r) \, \mathrm{d}r = 1$, and $\chi(r) = 0$ for $r \notin [-2, -1]$. Now for $\delta \in]0, 1/2[$ we define the smoothings

$$E_\delta(c) = \int_0^1 \chi_\delta(r)E(rc) \, \mathrm{d}r, \quad \text{where } \chi_\delta(r) = \frac{1}{\delta}\chi\left(\frac{1}{\delta}(r-1)\right).$$

Hence, χ_δ has support in $[1-2\delta, 1-\delta]$. If the closure of the domain of E is $[0, c_E]$, then E_δ is well-defined and C^∞ on $]0, c_E/(1-\delta)[$. Moreover, for all $c \in [0, c_E]$ we have $E_\delta(c) \to E(c)$ for $\delta \searrow 0$. We easily check, that E_δ still satisfies the assumption (1.25) and (7.4). Moreover, $E_\delta(c)$ can be estimated by $E(c)$ via

$$\exists \, K > 0 \, \forall \delta \in]0, 1/4[\, \forall c \geqq 0 : \quad |E_\delta(c)| \leqq K\big(c + |E(c)|\big). \qquad (7.6)$$

To see this, we first consider the largest interval $[0, c_1[$ on which E is non-increasing. Then $0 = E(0) \geqq E_\delta(c) \geqq E(c)$ which implies (7.6) with $K = 1$. If $c_1 = \infty$ then

we are done. If $c_1 < \infty$, then E starts to increase and there exists $c_2 \in [c_1, \infty[$ with $E(c) \geqq 0$ for $c \geqq c_2$. Using the construction of E_δ, we obtain for all $c \geqq 2c_2 \geqq c_2/(1-2\delta)$ the lower bound $E_\delta(c) \geqq 0$. Using (6.3) we easily get $E_\delta(c) \leqq E(c)$.

It remains to cover the case $c \in [c_1, 3c_2]$. If $c_1 = 0$ then $E(c) \geqq 0$ for all c, which means $c_2 = 0$ as well, then (7.6) follows immediately from the above arguments. If $c_1 > 0$, a uniform continuity argument gives the estimate $|E_\delta(c) - E(c)| \leqq M$ for $c \in [c_1, 3c_2]$. Then, choosing $K = M/c_1$ provides (7.6).

With this preparation, Lebesgue's dominated convergence theorem implies

$$\mu = c\mathcal{L}^d \text{ with } c \in L^1(\Omega) \text{ and } \mathscr{E}(\mu) < \infty$$
$$\implies \int_\Omega E_\delta(c(x))\,dx \to \mathscr{E}(\mu) \text{ as } \delta \downarrow 0.$$

Taking any constant-speed geodesic $[0,1] \ni t \mapsto \mu_t = c(t,\cdot)\mathcal{L}^d$, we know by Step 1 that the curves

$$\overline{e}_\delta : t \mapsto \int_\Omega E_\delta(c(t,x))\,dx$$

are convex. As $\overline{e}_\delta(t) \to \mathscr{E}(\mu_t)$ we conclude that $t \mapsto \mathscr{E}(\mu_t)$ is convex on $[0,1]$.

Step 3: Pure growth. The curve $t \mapsto t^2\mu_1$ is the unique geodesic connecting $\mu_0 = 0$ and μ_1. Using the Lebesgue decomposition $\mu_1 = c_1\mathcal{L}^d + \mu^\perp$ we see that

$$t \mapsto \mathscr{E}(\mu_t) = \mathscr{E}(t^2c_1\mathcal{L}^d) + t^2 E'_\infty\mu^\perp,$$

is convex on $[0,1]$ by Step 2 for the first term and by $E'_\infty \geqq 0$. The nonnegativity of $E'_\infty = \lim_{c\to\infty} E(c)/c$ follows from (7.4a) and Proposition 6.2(A), namely for $c \geqq c_\circ$ we have

$$\frac{1}{c}E(c) = \frac{1}{c^{d/(d+2)}}c^{-2/(d+2)}E(c) \overset{(i)}{\geqq} \frac{1}{c^{d/(d+2)}}c_\circ^{-2/(d+2)}E(c_\circ) \to 0 \text{ for } c \to \infty.$$

Step 4: The general case allowing for singular measures. Singular measures can only occur for E with sublinear growth. Hence, we assume $E'_\infty \in \mathbb{R}$ from now on. In particular $\mathscr{E}$ is finite everywhere, and using $E(c) \leqq E'_\infty c$ we have $\mathscr{E}(\mu) \leqq E'_\infty\mu(\Omega)$.

As in Corollary 5.6, we consider an arbitrary geodesic $(\mu_t)_{t\in[0,1]}$ connecting μ_0 and μ_1. For a fixed $s \in (0,1)$, we decompose μ_s as $\mu_s^a + \mu_s^\perp$. Then, $\mu_t = \widetilde{\mu}_t + \widehat{\mu}_t$ splits into two geodesics with disjoint supports and $\widetilde{\mu}_s = \mu_s^a$ and $\widehat{\mu}_s = \mu_s^\perp$, see Corollary 5.6. Moreover, we have $\widehat{\mu}_t \perp \mathcal{L}^d$ and $\widetilde{\mu}_t \ll \mathcal{L}^d$ for all $t \in (0,1)$. This implies the relation

$$\mathscr{E}(\mu_t) = \mathscr{E}(\widetilde{\mu}_t) + \mathscr{E}(\widehat{\mu}_t) = \mathscr{E}(\widetilde{\mu}_t) + E'_\infty\widehat{\mu}_t(\Omega).$$

Since $(\widehat{\mu}_t)_{t\in[0,1]}$ is a geodesic and the total mass functional $\mathscr{M}(\mu) = \mu(\Omega)$ is convex (see (1.21)) and $E'_\infty \geqq 0$, the last term $t \mapsto E'_\infty\widehat{\mu}_t(\Omega)$ is convex. Hence, it is sufficient to check the convexity of $t \mapsto \mathscr{E}(\widetilde{\mu}_t)$.

Since $\widetilde{\mu}_t \ll \mathcal{L}^d$ for all $t \in (0, 1)$, the function $t \mapsto \mathcal{E}(\widetilde{\mu}_t)$ is convex in the open interval $(0, 1)$ by Step 2. Hence, to show convexity on $[0, 1]$ it is sufficient to check that

$$\limsup_{t\downarrow 0} \mathcal{E}(\widetilde{\mu}_t) \leqq \mathcal{E}(\widetilde{\mu}_0) \quad \text{and} \quad \limsup_{t\uparrow 1} \mathcal{E}(\widetilde{\mu}_t) \leqq \mathcal{E}(\widetilde{\mu}_1),$$

because HK convergence implies weak convergence and $\mathcal{E}$ is weakly l.s.c.

Let us focus on the limit $t \downarrow 0$ as the limit $t \uparrow 1$ is completely analogous. The problem is that $\widetilde{\mu}_t \ll \mathcal{L}^d$ for $t \in (0, 1)$ only, but $\widetilde{\mu}_0$ may have a singular part. Hence, we forget the decomposition $\mu_t = \widetilde{\mu}_t + \widehat{\mu}_t$ and use a different one. Before that, we restrict to the case $\mu_0(\Xi^+) = 0$ because on Ξ^+ we have pure growth and this case is covered by Step 3.

Now, we exploit the Lebesgue decomposition of $\mu_0 = \mu_0^a + \mu_0^\perp$ at $t = 0$ and consider two disjoint Borel sets $A, B \subset \Omega \setminus \Xi^+$ such that $\mu_0^a = \mu_0 \llcorner A$ and $\mu_0^\perp = \mu_0 \llcorner B$. We define the corresponding disjoints sets $A_t := T_{t\to 0}^{-1}(A)$ and $B_t := T_{t\to 0}^{-1}(B)$ as well as the measures $\nu_t^A := \mu_t \llcorner A_t$ and $\nu_t^B := \mu_t \llcorner B_t$. By Theorem 5.4, we obtain two geodesics ν_t^A, ν_t^B concentrated on disjoint sets giving $\mathcal{E}(\mu_t) = \mathcal{E}(\nu_t^A) + \mathcal{E}(\nu_t^B)$. Since $\nu_t^A \ll \mathcal{L}^d$ for every $t \in [0, 1)$ we deduce that $t \mapsto \mathcal{E}(\nu_t^A)$ is convex up to 0 by Step 2. Concerning $\mathcal{E}(\nu_t^B)$, we use $\mathcal{E}(\mu) \leqq E_\infty' \mu(\Omega)$ and find

$$\limsup_{t\downarrow 0} \mathcal{E}(\nu_t^B) \leqq E_\infty' \limsup_{t\downarrow 0} \nu_t^B(\Omega) = E_\infty' \nu_0^B(\Omega) = \mathcal{E}(\nu_0^B),$$

where we exploited $\nu_0^B \perp \mathcal{L}^d$ in the last identity.

This finishes the proof of the main theorem   □

The next result is a direct consequence of the main result by using the results of Examples 6.3 and 6.4, respectively. In particular, this establishes the result announced in [13, Thm. 2.14].

Corollary 7.3. (Power-law functionals) *Assume that $\Omega \subset \mathbb{R}^d$ and $E : [0, \infty[\to \mathbb{R}$ satisfy (7.4) and let $\mathcal{E}$ be defined via (7.5).*

(1) *If $E(c) = c^m$ with $m \geqq 1$, then $\mathcal{E}$ is geodesically convex on $(\mathcal{M}(\Omega), \mathsf{HK})$.*
(2) *If $\mathcal{L}^d(\Omega) < \infty$, $d \in \{1, 2\}$, and $E(c) = -c^q$ with $d/(d+2) \leqq q \leqq 1/2$, then $\mathcal{E}$ is geodesically convex on $(\mathcal{M}(\Omega), \mathsf{HK})$.*

Example 7.4. We have seen above that the density $E(c) = -\sqrt{c}$ produces a geodesically convex functional in dimensions $d = 1$ and 2, if $\mathcal{L}^d(\Omega) < \infty$. The restriction of finite volume for Ω can be dropped by using a confining potential V as follows: let

$$\mathcal{E}_{1/2,V}(\mu) = \int_{\mathbb{R}^d} \left(-\sqrt{c(x)}\right) dx + \int_{\mathbb{R}^d} V \, d\mu \text{ for } \mu = c\mathcal{L}^d + \mu^\perp,$$

where $V \in C(\mathbb{R}^d)$ satisfies for $m > d$ and $A \in \mathbb{R}$ the lower bound $V(x) \geqq a_0 |x|^m - A$ on $\mathbb{R}^d$. Then it is easy to see that $\mathcal{E}_{1/2,V}$ is well-defined and weakly lower semi-continuous.

Moreover, in [26, Prop. 20] it was shown for a continuous $V : \mathbb{R}^d \to \mathbb{R}$ with $\inf V > -\infty$ that the linear mapping $\mu \mapsto \int_{\mathbb{R}^d} V \, \mathrm{d}\mu$ is geodesically λ_V-convex on $(\mathcal{M}(\Omega), \mathsf{HK})$ if and only if the mapping $\widetilde{V} : [x, r] \mapsto r^2 V(x)$ is geodesically λ_V-convex on the metric cone space $(\mathfrak{C}, \mathsf{d}_{\mathfrak{C}})$. For smooth V, this amounts to the estimate

$$\begin{pmatrix} \nabla^2 V(x) + 2V(x)\mathbb{I}_d & \nabla V(x) \\ \nabla V(x)^\top & 2V(x) \end{pmatrix} \geqq \lambda_V \mathbb{I}_{d+1}.$$

Thus, for V satisfying both of the above assumptions, the functional $\mathscr{E}_{1/2,V}$ is geodesically λ_V-convex on $(\mathcal{M}(\mathbb{R}^2), \mathsf{HK})$ for $d \in \{1, 2\}$. For $d = 1$ we may choose $V(x) = \alpha + \beta|x|^2$ with $\beta > 0$ and obtain $\lambda_V = 2\alpha$.

7.3. Necessity of the Conditions on E

Theorem 7.2 states that the conditions (1.25) and (7.4) on the density $E : [0, \infty) \to (-\infty, +\infty]$ are sufficient for the geodesic convexity of the integral functional $\mathscr{E}$. We finally show that the conditions are also sufficient. To simplify the analysis we restrict ourselves to the smooth case where $E : \mathrm{dom}(E) \to \mathbb{R}$ lies in C^2. Thus, we can obtain conditions by differentiation along suitably chosen geodesic curves. For this, the characteristic equations (4.62) derived in Theorem 4.8 will be the main tool.

Theorem 7.5. (Necessity of conditions on E) *Consider a closed, convex domain $\Omega \subset \mathbb{R}^d$ with nonempty interior and a density function $E : [0, \infty) \to (-\infty, +\infty]$ such that (7.4) holds and that E is C^2 on the interior of its domain. If the induced functional $\mathscr{E} : \mathcal{M}(\Omega) \to (-\infty, +\infty]$ defined in (7.5) is geodesically convex on $(\mathcal{M}(\Omega), \mathsf{HK})$, then E satisfies the conditions (1.25).*

Proof. We first observe that it is sufficient to show that for (ρ_*, γ_*) with $c_* = \gamma_*^{2+d}/\rho_*^d \in \mathrm{int}(\mathrm{dom}(E))$ we have the inequalities

$$\mathrm{D}^2 N_E(\rho_*, \gamma_*) \geqq 0 \quad \text{and} \quad (d-1)\partial_\rho N_E(\rho_*, \gamma_*) \leqq 0. \tag{7.7}$$

By the scaling properties of $N_E(\rho, \gamma) = (\rho/\gamma)^d E(\gamma^{2+d}/\rho^d)$ it is sufficient to look at the case $(\rho_*, \gamma_*) = (c_*^{1/2}, c_*^{1/2})$.

The main idea is to construct suitable geodesic curves μ_t such that the convexity of $t \mapsto \mathscr{E}(\mu_t)$ gives the desired inequality. For this we choose a point $x_* \in \mathrm{int}(\Omega)$ and $r_* > 0$ such that $B_{3r_*}(x_*) \subset \Omega$. Without loss of generality we assume $x_* = 0$ and write B_r in place of $B_r(x_*)$ for $r \in (0, 3r_*]$.

We further choose an $s \in (0, 1)$ and a smooth function $\xi_s \in \mathrm{C}^3(B_{3r_*})$. Then, there exists an $\varepsilon > 0$ such that there is a unique smooth solution $\xi : (s-\varepsilon, s+\varepsilon) \times B_{2r_*} \to \mathbb{R}$ of the Hamilton–Jacobi equation (2.4). With this $\xi_t = \xi(t, \cdot)$ and $r \in (0, r_*)$ we can construct a geodesic curve

$$(s-\varepsilon, s+\varepsilon) \ni t \mapsto \mu_t^{(r)} \quad \text{with } \mu_s^{(r)} = c_* \mathcal{L}^d \llcorner B_r \text{ and } \mathrm{sppt}(\mu_t^{(r)}) \subset B_{2r_*}$$

of absolutely continuous measures $\mu_t = c^{(r)}(t, \cdot)\mathcal{L}^d$, see (5.6a) in Corollary 5.5. If necessary ε needs to be reduced to avoid mass flowing outside B_{2r_*}. For this geodesic we have

$$\mathcal{E}(\mu_t^{(r)}) = \int_\Omega E(c^{(r)}(t, y))\,\mathrm{d}y = \int_{B_r} e(t, x)\,\mathrm{d}x \quad \text{with } e(t, x) = N_E(\rho(t, x), \gamma(t, x)),$$

where $\gamma(t, x) = c_*^{1/2}q_{s\to t}(x)$ and $\rho(t, x) = c_*^{1/2}q_{s\to t}(x)\big(\delta_{s\to t}(x)\big)^{1/d}$ with q and δ from (4.62). Note that q, δ, and e do not depend on r, cf. Theorem 5.4.

By the smoothness of ξ, and hence of ρ and γ, we may pass to the limit $r \downarrow 0$ in the convex functions $t \mapsto \frac{1}{\mathcal{L}^d(B_r)}\mathcal{E}(\mu_t^{(r)}) = \frac{1}{\mathcal{L}^d(B_r)}\int_{B_r} e(t, x)\,\mathrm{d}x$. Thus, the limit

$$t \mapsto e(t, 0) = N_E(\rho(t, 0), \gamma(t, 0)) \quad \text{is convex on } (s{-}\varepsilon, s{+}\varepsilon).$$

In particular, the second derivative is non-negative which means that

$$0 \leqq \ddot{e}(s, 0) = \mathrm{D}^2 N_E(c_*^{1/2}, c_*^{1/2})\Big[\begin{pmatrix}\dot{\rho}\\\dot{\gamma}\end{pmatrix}, \begin{pmatrix}\dot{\rho}\\\dot{\gamma}\end{pmatrix}\Big]$$
$$+\partial_\rho N_E(c_*^{1/2}, c_*^{1/2})\ddot{\rho} + \partial_\gamma N_E(c_*^{1/2}, c_*^{1/2})\ddot{\gamma},$$

where now $\dot{\rho} = \partial_t \rho(s, 0)$, $\dot{\gamma}$, $\ddot{\rho}$, and $\ddot{\gamma}$ are given by (4.62) and (5.13) in terms of ξ only:

$$\frac{\dot{\rho}}{c_*^{1/2}} = 2\xi_s + \Delta\xi_2, \quad \frac{\dot{\gamma}}{c_*^{1/2}} = 2\xi_s,$$
$$\frac{\ddot{\rho}}{c_*^{1/2}} = \frac{(\Delta\xi_s)^2 - d|\mathrm{D}^2\xi_s|^2}{d^2} + \frac{d-4}{d}|\nabla\xi_s|^2,$$
$$\frac{\ddot{\gamma}}{c_*^{1/2}} = |\nabla\xi_s|^2.$$

there ξ_s and its derivatives are evaluated at $x = x_* = 0$.

To obtain the convexity of N_E we can now choose the functions ξ_s such that $\ddot{\rho} = \ddot{\gamma} = 0$, which is the case for $\xi_s(x) = \alpha + \beta|x|^2$, which implies $\nabla\xi_s(0) = 0$, $\Delta\xi_s(0) = 2d\beta$, and $|\mathrm{D}^2\xi_s|^2 = d\beta^2$. Moreover, $\dot{\rho}$ and $\dot{\gamma}$ can be chosen arbitrarily by adjusting α, $\beta \in \mathbb{R}$. Thus, $\mathrm{D}^2 N \geqq 0$ is established.

To prove the second estimate in (7.7) we may assume $d \geqq 2$, as there is nothing to show for $d = 1$. Choosing the function $\xi_s = \alpha(x_1^2 - x_2^2)$ we obtain $\dot{\rho} = \dot{\gamma} = \ddot{\gamma} = 0$ and $\ddot{\rho} = -8c_*^{1/2}\alpha^2/d$. This implies $\partial_\rho N_E \leqq 0$ and the theorem is established. $\square$

7.4. A More Direct Sufficiency Proof for $2 \leqq d \leqq 4$

The above proof of Theorem 7.2 strongly relies on differentiating $e(t, x) = N_E(\rho(t, x), \gamma(t, x))$ with respect to t. In the case $2 \leqq d \leqq 4$, this can be avoided since we have the curvature estimates

$$\text{(a) } t \mapsto \gamma(t, x) \text{ is convex} \quad \text{and} \quad \text{(b) } t \mapsto \rho(t, x) \text{ is concave}, \qquad (7.8)$$

where we used $d \leqq 4$ in (5.12) for (b). With (a) and (b), we can further exploit

(i) the convexity of N_E,
(ii) the monotonicity of $\rho \mapsto N_E(\rho, \gamma)$ (non-increasing, cf. (1.25c) for $d \geqq 2$), and
(iii) the monotonicity of $s \mapsto N_E(s^{1-4/d^2}\rho, s\gamma)$ (non-decreasing, cf. Proposition 6.2).

Choosing $t_0, t_1, \theta \in [0, 1]$ with $t_0 < t_1$, we set $t_\theta := (1-\theta)t_0 + \theta t_1$ and have to show

$$N_E(\rho(t_\theta), \gamma(t_\theta)) \leqq (1-\theta)N_E(\rho(t_0), \gamma(t_0)) + \theta N_E(\rho(t_1), \gamma(t_1)). \quad (7.9)$$

We start with the right-hand side and use convexity (i) first:

$$(1-\theta)N_E(\rho(t_0), \gamma(t_0)) + \theta N_E(\rho(t_1), \gamma(t_1)) \overset{(i)}{\geqq} N_E\big((1-\theta)\rho(t_0)+\theta\rho(t_1), (1-\theta)\gamma(t_0)+\theta\gamma(t_1)\big).$$

With the convexity (a) of γ we have $s := \gamma(t_\theta)/\big[(1-\theta)\gamma(t_0)+\theta\gamma(t_1)\big] \in [0, 1]$ and continue

$$N_E\big((1-\theta)\rho(t_0)+\theta\rho(t_1), (1-\theta)\gamma(t_0)+\theta\gamma(t_1)\big)$$
$$\overset{(iii)}{\geqq} N_E\big(s^{1-4/d^2}\big[(1-\theta)\rho(t_0)+\theta\rho(t_1)\big], s\big[(1-\theta)\gamma(t_0)+\theta\gamma(t_1)\big]\big)$$
$$= N_E\big(s^{1-4/d^2}\big[(1-\theta)\rho(t_0)+\theta\rho(t_1)\big], \gamma(t_\theta)\big).$$

Using the monotonicity (ii) (for $d \geqq 2$) we can increase the first argument using $s^{1-4/d^2} \leq 1$ (because of $s \in [0, 1]$ and $d \leq 4$) and then exploit the concavity in (b) of ρ (i.e. $\rho(t_\theta) \geqq (1-\theta)\rho(t_0)+\theta\rho(t_1)$) giving

$$N_E\big(s^{1-4/d^2}\big[(1-\theta)\rho(t_0)+\theta\rho(t_1)\big], \gamma(t_\theta)\big) \overset{(ii), (a+b)}{\geqq} N_E(\rho(t_\theta), \gamma(t_\theta)).$$

Thus, we have proved the desired convexity (7.9) for the case $d \in \{2, 3, 4\}$.

Data Availability Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. AMBROSIO, L., FUSCO, N., PALLARA, D.: Functions of bounded variation and free discontinuity problems. The Clarendon Press, Oxford University Press, New York, Oxford Mathematical Monographs (2000)
2. AMBROSIO, L., GIGLI, N., SAVARÉ, G.: *Gradient flows in metric spaces and in the space of probability measures*, second ed., Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2008.
3. ALBERTI, G., AMBROSIO, L.: A geometrical approach to monotone functions in $\mathbb{R}^n$. *Math. Z.* **230**(2), 259–316, 1999
4. BURAGO, D., BURAGO, Y., IVANOV, S.: A course in metric geometry, Graduate Studies in Mathematics, vol. 33. American Mathematical Society, Providence, RI (2001)
5. BIANCHI, G., COLESANTI, A., PUCCI, C.: On the second differentiability of convex surfaces. *Geom. Dedicata* **60**(1), 39–48, 1996
6. BENAMOU, J.-D., BRENIER, Y.: A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.* **84**(3), 375–393, 2000
7. BERTRAND, J., PUEL, M.: The optimal mass transport problem for relativistic costs. *Calc. Var. Partial Differential Equations* **46**(1–2), 353–374, 2013
8. BERTRAND, J., PRATELLI, J., PUEL, M.: *Kantorovich potentials and continuity of total cost for relativistic cost functions*. J. Math. Pures Appl. (9) **110** (2018) 93–122.
9. BUCZOLICH, Z.: Density points and bi-Lipschitz functions in $\mathbb{R}^m$. *Proc. Amer. Math. Soc.* **116**(1), 53–59, 1992
10. CHIZAT, L., PEYRÉ, G., SCHMITZER, B., VIALARD, F.-X.: An interpolating distance between optimal transport and Fisher-Rao metrics. *Found. Comput. Math.* **18**(1), 1–44, 2015
11. CHIZAT, L., PEYRÉ, G., SCHMITZER, B., VIALARD, F.-X.: Unbalanced optimal transport: geometry and Kantorovich formulation. *J. Funct. Analysis* **274**(11), 3090–3123, 2018
12. DANERI, S., SAVARÉ, G.: Eulerian calculus for the displacement convexity in the Wasserstein distance. *SIAM J. Math. Analysis* **40**, 1104–1122, 2008
13. DI MARINO, S., CHIZAT, L.: A tumor growth model of Hele-Shaw type as a gradient flow. *ESAIM Control Optim. Calc. Var.* **26**(103), 1–38, 2020
14. FIGALLI, A., GIGLI, N.: Local semiconvexity of Kantorovich potentials on non-compact manifolds. *ESAIM Control Optim. Calc. Var.* **17**(3), 648–653, 2011

15. FLEISSNER, F.: A minimizing movement approach to a class of scalar reaction-diffusion equations. *ESAIM Control Optim. Calc. Var.* **27**(18), 1–29, 2021
16. GANGBO, W., McCANN, R.J.: The geometry of optimal transportation. *Acta Math.* **177**(2), 113–161, 1996
17. GALLOUËT, T.O., MONSAINGEON, L.: A JKO splitting scheme for Kantovorich-Fisher-Rao gradient flows. *SIAM J. Math. Analysis* **49**(2), 1100–1130, 2017
18. GANGBO, W., OLIKER, V.: Existence of optimal maps in the reflector-type problems. *ESAIM Control Optim. Calc. Var.* **13**(1), 93–106, 2007
19. HELLINGER, E.: Neue Begründung der Theorie quadratischer Formen von unendlichvielen Veränderlichen (in German). *J. Reine Angew. Math.* **136**, 210–271, 1909
20. JIMENEZ, C., SANTAMBROGIO, F.: *Optimal transportation for a quadratic cost with convex constraints and applications*. J. Math. Pures Appl. (9) **98**:1, 103–113, 2012
21. KONDRATYEV, S., MONSAINGEON, L., VOROTNIKOV, D.: A new optimal transport distance on the space of finite Radon measures. *Adv. Differ. Eqns.* **21**(11/12), 1117–1164, 2016
22. LASCHOS, V., MIELKE, A.: Geometric properties of cones with applications on the Hellinger-Kantorovich space, and a new distance on the space of probability measures. *J. Funct. Analysis* **276**(11), 3529–3576, 2019
23. LASCHOS, V., MIELKE, A.: *Evolutionary Variational Inequalites on the Hellinger-Kantorovich space and the spherical Hellinger-Kantorovich space*. arXiv:2207.09815 2022 .
24. LI, M.: Gradient estimates for a simple elliptic equation on complete non-compact Riemann manifolds. *J. Funct. Analysis* **241**, 374–382, 2006
25. LIERO, M., MIELKE, A.: *Gradient structures and geodesic convexity for reaction-diffusion systems*. Phil. Trans. Royal Soc. A **371**:2005 (2013) 20120346, 28.
26. LIERO, M., MIELKE, A., SAVARÉ, G.: Optimal transport in competition with reaction - the Hellinger-Kantorovich distance and geodesic curves. *SIAM J. Math. Analysis* **48**(4), 2869–2911, 2016
27. LIERO, M., MIELKE, A., SAVARÉ, G.: Optimal entropy-transport problems and a new Hellinger-Kantorovich distance between positive measures. *Invent. math.* **211**, 969–1117, 2018
28. McCANN, R.J.: A convexity principle for interacting gases. *Adv. Math.* **128**, 153–179, 1997
29. McCANN, R. J., PUEL, M.: *Constructing a relativistic heat flow by transport time steps*. Ann. Inst. H. Poincaré C Anal. Non Linéaire **26**:6 (2009) 2539–2580.
30. MURATORI, M., SAVARÉ, G.: *Gradient flows and evolution variational inequalities in metric spaces. I: structural properties*. J. Funct. Analysis **278**:4 (2020) 108347.
31. OTTO, F., WESTDICKENBERG, M.: Eulerian calculus for the contraction in the Wasserstein distance. *SIAM J. Math. Analysis* **37**, 1227–1255, 2005
32. PERTHAME, B., QUIRÓS, F., VÁZQUEZ, J.L.: The Hele-Shaw asymptotics for mechanical models of tumor growth. *Arch. Rational Mech. Anal.* **212**(1), 93–127, 2014
33. SCHIED, A.: Geometric aspects of Fleming-Viot and Dawson-Watanabe processes. *Ann. Probab.* **25**(3), 1160–1179, 1997
34. VILLANI, C.: Optimal transport. Old and new. Springer, Berlin (2009)

Matthias Liero, Alexander Mielke
Weierstraß-Institut für Angewandte Analysis und Stochastik,
Berlin
Germany.

and

Alexander Mielke
Humboldt Universität zu Berlin,
Berlin
Germany.

and

Giuseppe Savaré
Department of Decision Sciences and BIDSA,
Bocconi University,
Via Roentgen 1,
20136 Milan
Italy.
e-mail: giuseppe.savare@unibocconi.it