SMOOTHING AND NON-SMOOTHING VIA A FLOW TANGENT TO THE RICCI FLOW

MATTHIAS ERBAR AND NICOLAS JUILLET

Abstract. We study a transformation of metric measure spaces introduced by Gigli and Mantegazza consisting in replacing the original distance with the length distance induced by the transport distance between heat kernel measures. We study the smoothing effect of this procedure in two important examples. Firstly, we show that in the case of particular Euclidean cones, a singularity persists at the apex. Secondly, we generalize the construction to a sub-Riemannian manifold, namely the Heisenberg group, and show that it regularizes the space instantaneously to a smooth Riemannian manifold.

1. Introduction

There are many ways to deform a Riemannian manifold into a singular metric space as discussed for instance in the influential essay of Gromov [18]. We are interested in the opposite question whether there exists a deformation, intrinsically defined for a wide class of metric spaces that instantaneously turns the space into a Riemannian manifold. In this paper, we investigate a method that has been introduced by Gigli and Mantegazza [17]. We examine its regularization properties in two important cases: Euclidean cones and the Heisenberg group. These are emblematic examples of Alexandrov spaces and subRiemannian spaces respectively. We also discuss normed vector spaces where the transformation turns out to be the identity as an example of Finsler structures.

Before we state our results we briefly explain the main features of the construction of Gigli and Mantegazza which is based on the interplay of optimal transport and Ricci curvature. The starting point is a metric measure space (X, d, m) on which a reasonable notion of heat kernel can be defined. For $t > 0$ a new distance $d_t(x, y)$ is defined as the length distance induced by the L^2 Wasserstein distance built from d between the heat kernel measures centered at x and y.

The striking feature of this approach is the following main result of [17]: When (X, d, m) is a Riemannian manifold then d_t is induced by a smooth metric tensor g_t that is tangent to the Ricci flow, i.e. $\partial_t|_{t=0} g_t = -2\,\mathrm{Ric}$ in a weak sense. Gigli and Mantegazza then generalize this construction to metric measures spaces satisfying generalized Ricci curvature lower bounds, more precisely the RCD condition, which ensures existence of a well-behaved heat kernel. This can be seen as a first step towards constructing a Ricci flow

2010 *Mathematics Subject Classification.* Primary 53C44; Secondary: 49Q20, 51F99, 51K10, 53C17.

Key words and phrases. Ricci flow, optimal transport, Euclidean cone, Heisenberg group.

for non-smooth initial data. Synthetic characterizations of super-Ricci flows based on optimal transport have been obtained by McCann and Topping [24] and by Sturm [31].
Since d_t can be thought of as a sort of convolution of the original distance with the heat kernel, having the smoothing effect of the heat equation and Ricci flow in mind, one might expect that this procedure gives a canonical way of regularizing the metric measure space.
A first study of the regularizing effects of the Gigli-Mantegazza flow has been performed by Bandara, Lakzian and Munn [6] in the case where the distance d is induced by a metric tensor with low regularity and isolated conic singularities. It is shown that d_t is induced by a metric tensor with at least the same regularity away from the original singular set. The question, what happens at the singularities has been left unanswered.
In the present paper, we give an answer showing that conic singularities can persist under the Gigli-Mantegazza transformation. We analyse in detail the transformation for two specific Euclidean cones of angle π and $\pi/2$. Our results are the following (see Theorem 3.11 and Proposition 3.10 below).

Theorem 1.1. *Let $C(\pi)$ be the two-dimensional Euclidean cone of angle π and d its distance. For every $t > 0$ the convoluted distance d_t has a conic singularity of angle $\sqrt{2}\pi$ at the apex.*
As t goes to zero, the metric space $(C(\pi), d_t)$ tends to $(C(\pi), d)$ pointwise and in the pointed Gromov–Hausdorff topology. As t goes to infinity, it tends to the Euclidean cone of angle $\sqrt{2}\pi$ in the pointed Gromov–Hausdorff topology.

In fact, it turns out that for fixed $\theta > 0$ all spaces $(C(\theta), d_t)$ for $t > 0$ are isometric up to a multiplicative constant. An isometry is induced by the radial dilation $x \in C(\theta) \mapsto t^{-1/2}x$. Our second result shows that for the cone of angle $\pi/2$ the behavior of the singularity is even worse (see Theorem 3.17 and Proposition 3.16 below).

Theorem 1.2. *Let $C(\pi/2)$ be the two dimensional Euclidean cone of angle $\pi/2$ and d its distance. For every $t > 0$, the distance d_t has a cusp singularity at the apex, more precisely the asymptotic angle is zero.*
As t goes to zero, the metric space $(C(\pi/2), d_t)$ tends to $(C(\pi/2), d)$ pointwise and in the pointed Gromov–Hausdorff topology. As t goes to infinity, it tends to $\mathbb{R}^+$ with the Euclidean distance in the pointed Gromov–Hausdorff sense.

The reason why we focus on these two specific cones is that they can be conveniently represented as quotients of $\mathbb{R}^2$ under rotation by π and $\pi/2$ respectively. It turns out that the convoluted distance d_t is the length distance induced by the L^2-Wasserstein distance between a mixtures of two (respectively four) rotated copies of Gaussian measures with variance $2t$.
A corollary of the previous theorem is that the space $(C(\pi/2), d_t)$ is not an Alexandrov space even though $C(\pi/2)$ is. In fact, in Alexandrov spaces a triangle with one angle zero is flat, which is wrong for $(C(\pi/2), d_t)$, see Remark 3.18. This negative result has to be compared to positive results by Takatsu [32], where it is shown that the subspace made of all Gaussian measures in the Wasserstein space over Euclidean space is an Alexandrov space.

Note moreover, that the Wasserstein space over a non-negatively curved Alexandrov space is again a non-negatively curved Alexandrov space [30, Proposition I.2.10] and that many subspaces of finite dimensional Alexandrov spaces are known to be Alexandrov spaces, for instance convex hypersurfaces in Euclidean spaces or Riemannian manifolds of sectional curvature bounded below [1, 8, 25].
It turns out that the metric d_t on the cone is induced by a metric tensor of warped-product form. Conditional on this metric tensor beeing smooth away from the apex we can show, see Remark 3.19, that the metric measure space $\big(C(\pi/2), d_t, m_t\big)$, where m_t is the Hausdorff measure induced by d_t, is not an $\mathrm{CD}(K, N)$ space for any $K \in \mathbb{R}$ and $N \geq 0$. This would give a negative answer to a question raised in [17] whether the flow preserves the class of RCD metric measure spaces.
Given the relation of the Gigli–Mantagazza flow with the Ricci flow, the convergence of d_t to the original cone distance d has to be compared with the fact that any Euclidean cone of dimension 2 can be obtained as the backward limit of classical solutions to the Ricci flow [13, Chapter 4.5]. See also [29, 14] for related results in higher dimension.

Our second contribution in this paper is an investigation of the Gigli–Mantegazza flow applied to the first Heisenberg group equipped with the Carnot–Carathéodory distance. The Heisenberg group is one of the simplest examples of a non trivial Carnot group, i.e. a nilpotent stratified Lie group with a left-invariant metric on the first strata, and of a non trivial subRiemannian manifold. These classes are of course connected: As proved by Bellaïche [7], the tangent cones of subRiemannian spaces are Carnot groups. The differentiable structure of the Heisenberg group is the one of $\mathbb{R}^3$ and the group structure is given in coordinates (x, y, u) by $(x, y, u).(x', y', u') = (x + x', y + y', u + u' + (1/2)(xy' - x'y))$.
The Carnot–Carathéodory distance is obtained by minimizing the length of curves that are tangent to the 2-dimensional horizontal subbundle spanned by $X = \partial_x - \frac{y}{2}\partial_u$ and $Y = \partial_y + \frac{x}{2}\partial_u$. A standard way to approximate this distance is to consider for $\varepsilon > 0$ the Riemannian distance $d_{\mathrm{Riem}(\varepsilon)}$ obtained by considering $X, Y, \varepsilon\partial_u$ as an orthonormal frame. In fact, this penalization principle permits to see any subRiemannian manifold as a limit of Riemannian manifolds. Note that $(\mathbb{H}, d_{cc})$ does not satisfy a generalized lower Ricci curvature bound in the sense of the RCD condition. Therefore we slightly generalize the construction in [17] and obtain the following result (see Theorem 4.6 and Proposition 4.9 below).

Theorem 1.3. *Let $(\mathbb{H}, d_{cc})$ be the first Heisenberg group equipped with the Carnot–Carathéodory distance. For $t > 0$, the convoluted distance d_t coincides with $K d_{\mathrm{Riem}(\kappa\sqrt{t})}$, for some constants K and κ satisfying $K \geq 2$ and $K/\kappa < \sqrt{2}$.*
As t goes to zero the distance d_t converges to $K d_{cc}$ pointwise. In the pointed Gromov–Hausdorff topology the space $(\mathbb{H}, d_t)$ converges to $(\mathbb{H}, d_{cc})$.

The striking part of the theorem is that also non-horizontal curves can have finite length after lifting them to the Wasserstein space built from d_{cc} via the heat kernel and thus d_t becomes a Riemannian distance. We believe that this behavior also holds for more general contact manifolds. However, let us stress the fact that even for the Heisenberg group the distance d_t does not converge pointwise to d_{cc} as t goes to zero. Convergence in pointed Gromov–Hausdorff sense only holds due to the high amount of symmetry of the space, in particular, due to the fact that the dilation $(x, y, u) \mapsto (Kx, Ky, K^2u)$ is an isometry between $(\mathbb{H}, Kd_{cc})$ and $(\mathbb{H}, d_{cc})$. The Gromov–Hausdorff convergence probably does not hold for generic contact manifolds of dimension 3 with a subRiemannian metric on the nonholonomic contact distribution. Finally, note that also the Heisenberg group can be obtained as a backward limit of classical solution to the Ricci flow as was shown by Cao and Saloff-Coste [11].

Three sections follow this introduction. The next section contains the construction of the convoluted distance d_t in a general setting. As a first example we discuss the case of normed spaces. In Section 3 we establish our results on the Euclidean cones $C(\pi)$ and $C(\pi/2)$. Section 4 is devoted to the Heisenberg group.

Acknowledgements. The authors would like to thank Michel Bonnefont, Thomas Richard and André Schlichting for stimulating discussions on this work and related topics and the referee for valuable comments. Part of this work was accomplished while the authors were enjoying the hospitality of the Hausdorff Research Institute for Mathematics in Bonn during the Junior Trimester Program on Optimal Transport. They would like to thank HIM for its support and the inspiring atmosphere. M.E. gratefully acknowledges support by the German Research Foundation through the Collaborative Research Center 1060 *The Mathematics of Emergent Effects* and the Hausdorff Center for Mathematics. N.J. is partially supported by the Programme ANR JCJC GMT (ANR 2011 JS01 011 01).

2. Construction of the flow

In this section we present the construction of the convoluted distance d_t in a general framework. The reason is that the framework of RCD spaces considered in [17] (see subsection 2.3) does not cover the Heisenberg group. Moreover, unlike in [17] the spaces considered here are non-compact

2.1. **Preliminaries.** Let (X, d) be a Polish metric space. Recall that for $p \geq 1$ a curve $(\gamma)_{t\in[0,T]}$ in (X, d) is called p-absolutely continuous, for short $\gamma \in \mathrm{AC}^p\big([0, T], (X, d)\big)$, if there exist a function $m \in L^p(0, T)$ such that for any $0 \leq s \leq t \leq T$:

$$d(\gamma(s), \gamma(t)) \;\leq\; \int_s^t m(r)\,\mathrm{d}r\;.$$

For $p = 1$, we simply say absolutely continuous instead of 1-absolutely continuous. For any absolutely continuous curve γ the metric derivative defined by

$$|\dot{\gamma}_s| \;=\; \lim_{h\to 0} \frac{d(\gamma_{s+h}, \gamma_s)}{h}$$

exists for a.e. $s \in (0,T)$ and is the minimal m as above, see [2, Thm. 1.2.1]. Lipschitz curves with respect to a distance d are called d-Lipschitz curves, they are locally p-absolutely continuous for every $p \geq 1$.
We denote by $\mathscr{P}(X)$ the set of Borel probability measures. The subset of measures μ with finite second moment, i.e. satisfying

$$\int d(x_0, x)^2 \, \mathrm{d}\mu(x) < \infty$$

for some, hence any $x_0 \in X$ will be denoted by $\mathscr{P}_2(X)$. Given $\mu, \nu \in \mathscr{P}_2(X)$ their L^2-Wasserstein distance is defined by

$$W(\mu, \nu) = \inf_\pi \sqrt{\int d(x,y)^2 \, \mathrm{d}\pi(x,y)} \; ,$$

where the infimum is taken over all couplings π of μ and ν. Recall that $\big(\mathscr{P}_2(X), W\big)$ is again a Polish metric space. Sometimes we will write W_X or $W_{(X,d)}$ to avoid confusion about the underlying metric space (X,d).

2.2. **Construction of the flow.** Recall that (X,d) is a metric Polish space. Let us assume in addition that it is proper, i.e. closed balls are compact, and that it is a length space, i.e. we have

$$d(x,y) = \inf_\gamma \int_0^T |\dot\gamma_s| \, \mathrm{d}s \; ,$$

where the infimum is taken over all absolutely continuous curves γ connecting x to y. Notice that (X,d) is in fact geodesic, i.e. the infimum is attained.
The construction of the flow is based on a family of maps from X to $\mathscr{P}_2(X)$ satisfying some properties that we list now.

Assumption 2.1. *There exists a family $(\iota_t)_{t\geq 0}$ of maps $\iota_t : X \to \mathscr{P}_2(X)$ with the following properties:*

- $\iota_0(x) = \delta_x$ *for all* $x \in X$,
- ι_t *is injective for all* $t \geq 0$,
- ι_t *is Lipschitz, more precisely, there exist constants* $C_t > 0$ *such that*

$$W\big(\iota_t(x), \iota_t(y)\big) \;\leq\; C_t d(x,y) \quad \forall x, y \in X \; , \tag{2.1}$$

  and $t \mapsto C_t$ *is locally bounded from above,*
- *the curve* $[0,\infty) \ni t \mapsto \iota_t(x)$ *is continuous with respect to* W *for all* $x \in X$.

In the examples considered later $\iota_t(x)$ will be the heat kernel measure at time t centered at x.
We introduce a new family of distance functions $\widetilde{d}_t : X \times X \to [0,\infty)$ for $t \geq 0$ given by

$$\widetilde{d}_t(x,y) \;=\; W\big(\iota_t(x), \iota_t(y)\big) \; .$$

As W is a distance it follows from the injectivity of ι_t that $\widetilde{d}_t$ is also a distance. It is the *chord* distance induced by the embedding ι_t. The main object of study here will be the corresponding *arc* distance, i.e. the length distance induced by $\widetilde{d}_t$, denoted by d_t. More precisely, we define for $t \geq 0$ and $x, y \in X$:

$$d_t(x,y) \; = \; \inf_{\gamma} \int_0^T |\dot{\gamma}_s|_t \, \mathrm{d}s \; , \tag{2.2}$$

where the infimum is taken over all curves $\gamma \in \mathrm{AC}\big([0,T];(X,\widetilde{d}_t)\big)$ such that $\gamma_0 = x, \gamma_T = y$ and $|\dot{\gamma}_s|_t$ denotes the metric derivative with respect to $\widetilde{d}_t$. Note that (2.1) implies that

$$d_t(x,y) \; \leq \; C_t d(x,y) \qquad \forall x,y \in X \; . \tag{2.3}$$

Indeed, for any curve $(\gamma_s)_s$ that is absolutely continuous with respect to d its metric derivative with respect to d is bounded above as $|\dot{\gamma}_s|_t \leq C_t |\dot{\gamma}_s|$. The claim then follows by integrating in s and taking the infimum over all such curves $(\gamma_s)_s$ noting that they are also absolutely continuous with respect to $\widetilde{d}_t$ and that (X,d) is a length space.

Remark 2.2. This construction is slightly different from the one in [17], where the infimum in the definition of d_t is taken over γ in $\mathrm{AC}\big([0,T];(X,d)\big)$ which is a subset of $\mathrm{AC}\big([0,T];(X,\widetilde{d}_t)\big)$ by the Lipschitz assumption (2.1). Allowing curves in the latter larger class will be crucial when applying the construction in the case of the Heisenberg group in Section 4. In the case of the Euclidean cones $C(\pi), C(\pi/2)$ discussed in Section 3, we show in Lemma 3.6 that the infima over both classes of curves agree so that we are consistent with the construction in [17].

Remark 2.3. Note that the value of the infimum in (2.2) does not change, if we restrict the infimum to $\widetilde{d}_t$-Lipschitz curves. Indeed, the right hand side of (2.2) is invariant by reparametrization and every absolutely continuous curve can be reparametrized as a Lipschitz curve, see for instance [2, Lem. 1.1.4].

We can reformulate the definition of d_t as follows. Given an absolutely continuous curve $(\gamma_s)_{s\in[0,T]}$ in $(X,\widetilde{d}_t)$ we obtain an absolutely continuous curve $(\mu^t_{\gamma_s})_{s\in[0,T]}$ in $\big(\mathscr{P}_2(X),W\big)$ by setting $\mu^t_{\gamma_s} = \iota_t(\gamma_s)$. Then we have

$$d_t(x,y) \; = \; \inf_{\gamma} \int_0^T |\dot{\mu}^t_{\gamma_s}| \, \mathrm{d}s \; ,$$

where $|\dot{\mu}^t_{\gamma_s}|$ denotes the metric derivative with respect to W. Another equivalent formulation is

$$d_t(x,y) = \inf \sup \sum_{i=0}^{N-1} \widetilde{d}_t(\gamma_{s_i},\gamma_{s_{i+1}}) = \inf \sup \sum_{i=0}^{N-1} W(\mu^t_{\gamma_{s_i}},\mu^t_{\gamma_{s_{i+1}}}), \tag{2.4}$$

the supremum being taken over all partitions $0 = s_0 < s_1 < \cdots < s_N = 1$ and the infimum over all continuous curves $(\gamma_s)_{s\in[0,1]}$ connecting x to y.
In this general setup we have the following continuity properties.

Proposition 2.4. *For all $x,y \in X$, the curve $[0,\infty) \ni t \mapsto \widetilde{d}_t(x,y)$ is continuous and the curve $t \mapsto (X,\widetilde{d}_t)$ is continuous with respect to the pointed*

Gromov–Hausdorff convergence. Moreover, assume in addition to Assumption 2.1 that bounded sets in $(X, \widetilde{d}_t)$ are bounded in (X, d). Then the distances $\widetilde{d}_t$ and d_t induce the same topology as the original distance d.

Proof. We first prove the convergence statement. Let $(t_n)_n$ converge to t. As an immediate consequence of Assumption 2.1 we have that $\widetilde{d}_{t_n}(x, y) \to \widetilde{d}_t(x, y)$ for fixed $x, y \in X$. Moreover, by (2.1), for each compact set K in (X, d) the functions $\widetilde{d}_{t_n}(\cdot, \cdot)$ are equicontinuous on $K \times K$. Thus, they converge uniformly to $\widetilde{d}_t(\cdot, \cdot)$. This readily yields the convergence of $(X, \widetilde{d}_{t_n})$ to $(X, \widetilde{d}_t)$ in the pointed Gromov–Hausdorff sense. Now, we turn to the second statement. First, recall from (2.3) that $\widetilde{d}_t \leq d_t \leq C_t d$. Thus, it suffices to show that for any sequence $(x_n)_n$ in X with $\widetilde{d}_t(x, x_n) \to 0$ as $n \to \infty$ for some $x \in X$ we also have that $d(x_n, x) \to 0$. By assumption, the sequence x_n is bounded in (X, d). Thus, up to taking a subsequence we can assume that $d(x_n, x') \to 0$ for some $x' \in X$. Hence, also $\widetilde{d}_t(x_n, x') \to 0$ and we infer that $x' = x$. This being independent of the subsequence chosen, we conclude that the full sequence x_n converges to x in (X, d). □

Remark 2.5. While continuity of the map $t \mapsto \widetilde{d}_t(x, y)$ holds by assumption, the continuity of $t \mapsto d_t(x, y)$ may fail. This happens for the Heisenberg group at $t = 0$ as we will see in Section 4. This is in contrast to [17, Thm. 5.18] where right-continuity of this map is shown. Note however, that the Heisenberg group does not satisfy the RCD condition and our construction is slightly different in this case, see Remark 2.2.

2.3. **Riemannian manifolds and RCD spaces.** In [17] the preceding construction has been introduced and studied in the case where (X, d) is a Riemannian manifold or more generally a metric measure spaces satisfying the *Riemannian curvature-dimension condition* for some curvature parameter $K \in \mathbb{R}$, denoted by $\mathrm{RCD}(K, \infty)$. For short we call such spaces RCD spaces. In both cases the embedding ι_t is constructed using the heat kernel. Let us briefly recall the main results in [17].

Let (X, g) be a smooth compact and connected Riemannian manifold with metric tensor g and let d and vol be the associated Riemannian distance and volume measure. One can define a map $\iota_t : X \to \mathscr{P}_2(X)$ be setting $\iota_t(x) = \nu_x^t$, where $\nu_x^t(\mathrm{d}y) = p_t(x, y)\,\mathrm{vol}(\mathrm{d}y)$ is the heat kernel measure, i.e. $p_t(\cdot, \cdot)$ is the fundamental solution to the heat equation on X. It can be verified that Assumption 2.1 and Proposition 2.4 hold in this case.
Gigli and Mantegazza prove that the distances d_t are induced by a family of smooth metric tensors $(g_t)_{t \geq 0}$ and that this flow of tensors is initially tangent to the Ricci flow [17, Prop. 3.5,Thm. 4.6]. More precisely, for every geodesic $(\gamma_s)_{s \in [0,1]}$ with respect to $g = g_0$:

$$\frac{\mathrm{d}}{\mathrm{d}t} g_t(\dot{\gamma}_s, \dot{\gamma}_s)\big|_{t=0} \;=\; \mathrm{Ric}(\dot{\gamma}_s, \dot{\gamma}_s) \quad \text{for almost every } s \in (0, 1)\,,$$

where Ric denotes the Ricci tensor of g. Gigli and Mantegazza then generalize the construction for the initial data being a metric measure space satisfying the $\mathrm{RCD}(K, \infty)$ condition. Since we do not work in this general setting, we will describe it only briefly. For more details on RCD spaces we refer to [3, 4].

Roughly speaking, RCD spaces form a natural class of metric measure spaces that can be equipped with a canonical notion of Laplace operator and a well behaved associated heat kernel. The RCD(K,∞) condition is a reinforcement of the curvature-dimension condition CD(K,∞) introduced by Lott–Villani and Sturm [23, 30] as a synthetic definition of a lower bound K on the Ricci curvature for a metric measure space (X,d,m). The condition CD(K,∞) asks for the relative entropy

$$\mathrm{Ent}(\mu) = \int \rho \log \rho \,\mathrm{d}m\ , \quad \text{for } \mu = \rho m \in \mathscr{P}_2(X)$$

to be K-convex along Wasserstein geodesics, i.e.

$$\mathrm{Ent}(\mu_s) \leq (1-s)\,\mathrm{Ent}(\mu_0) + s\,\mathrm{Ent}(\mu_1) - \frac{K}{2}s(1-s)W(\mu_0,\mu_1)^2\ .$$

The RCD(K,∞) condition requires in addition that the 'heat flow' obtained as the Wasserstein gradient flow of the entropy in the spirit of Otto [28] is linear. This excludes e.g. Finslerian geometries. It is a deep insight that the two requirements can be encoded simultaneously in the following property (which we take as a definition of RCD spaces for the purpose of this paper).

Theorem 2.6 (Characterization via EVI [4, Thm. 5.1])**.** *Let K be a real number. The metric measure space (X,d,m) satisfies the Riemannian curvature-dimension condition RCD(K,∞) if and only if for every $\mu \in \mathscr{P}_2(X)$ there exists an absolutely continuous curve $(\mu_t)_{t\geq 0}$ in $(\mathscr{P}_2(X),W)$ starting from μ in the sense that $W_2(\mu,\mu_t)\to 0$ as $t\to 0$ and solving the* Evolution Variational Inequality *(in short EVI) of parameter K, i.e. for all $\chi \in \mathscr{P}_2(X)$ such that* $\mathrm{Ent}(\chi|m)<\infty$ *and a.e. $t>0$:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}W(\mu_t,\chi)^2 + \frac{K}{2}W(\mu_t,\chi)^2 \leq \mathrm{Ent}(\chi) - \mathrm{Ent}(\mu_t)\ .$$

In fact, the solution μ_t to the EVI is unique and, putting $H_t\mu = \mu_t$, one obtains a linear semigroup on $\mathscr{P}_2(X)$ which is called the heat flow (acting on measures) in X. The construction in [17] then proceeds as presented in Section 2 by choosing the map $\iota_t : X \to \mathscr{P}_2(X)$ to be $\iota_t(x) = H_t\delta_x$. A natural example of RCD spaces are Euclidean cones, see [21].

2.4. **Normed spaces.** For an example that can be studied rapidly and is rather different let us consider the flow for $\mathbb{R}^n$ equipped with a norm $\|\cdot\|$. Indeed, the metric measure space $(\mathbb{R}^n,\|\cdot\|,\mathrm{Leb})$ satisfies the condition CD$(0,\infty)$ but does not satisfy RCD$(0,\infty)$ unless $\|\cdot\|$ is induced by an inner product. It is possible to consider in this setting a non-linear heat equation, driven by a non-linear Laplace operator, see [26] for the a study in the much more general setting of Finsler manifolds. However, for a non-Hilbert norm there is no canonical choice of a heat kernel, i.e. a solution starting from a Dirac mass since contraction of the heat flow fails [27]. Note however, that a particular solution is given by the appealing formula [27, Example 4.3]

$$f_t(x) = \frac{C}{4\pi t}\exp\left(-\frac{\|x\|^2}{4t}\right)\ ,$$

where C is a normalization constant. Hence a choice satisfying Assumption 2.1 is $\iota_t(x) = f_t(\cdot - x)\,\mathrm{Leb}$. Any other reasonable choice should be translation invariant. Let us show that in this case the distance d_t coincides with the original one, i.e. $d_t(x,y) = \|x-y\|$. Indeed, consider $\iota_t : x \mapsto (\tau_x)_\# \nu_t$ where $\nu_t \in \mathscr{P}_2(\mathbb{R}^d, \|\cdot\|)$ is a measure and τ_x the translation by x. It is easily checked using Jensen's inequality on the convex function $(u,v) \mapsto \|u-v\|^2$ that $W_{(\mathbb{R}^n,\|\cdot\|)}\big(\iota_t(x), \iota_t(y)\big) = \|x-y\|$. The translation τ_{y-x} is an optimal map, in other words $(\tau_x,\tau_y)_\#\nu_t$ is an optimal coupling. Since the original distance was already a length distance we find $d_t(x,y) = \widetilde{d}_t(x,y) = \|x-y\|$. Hence the flow leaves the space invariant and does not regularize it to a Riemannian manifold.

Remark 2.7. We mention that the approximation of some normed spaces by Riemannian manifolds is possible by using periodic Riemannian metrics with a period diameter going to zero. Consider for instance the sequence $(\mathbb{R}^n, k^{-1}d_g)_{k\geq 1}$ where d_g is a fixed periodic Riemannian distance. It converges to $\mathbb{R}^n$ equipped with its "stable norm" as defined for instance in [9, section 8.5.2]. It is not clear whether any norm may be attained in this way and this question is related to the notorious open problem of characterizing the stable norms [10]. Finally, note that it is impossible to approximate a non-Hilbertian normed space in Gromov Hausdorff topology by Riemannian manifolds with non-negative Ricci curvature. This is because any such limit metric measure space that contains a line has to split as a product of $\mathbb{R}$ and another metric measure space by the splitting theorem for Ricci limit spaces established by Cheeger and Colding [12], see also [33, Conclusions and open problems]. This argument also applies to the Heisenberg group. Moreover it is proven in [19] that $(\mathbb{H}, d_{cc})$ also cannot be approximated by a sequence of Riemannian manifolds with *any* uniform lower bound on the Ricci curvature.

3. Gigli–Mantegazza flow starting from a cone

In this section we will analyse the construction of Gigli and Mantegazza in the case where the initial datum is an Euclidean cone. More precisely, we will consider the cones of angle π and $\pi/2$. We will show that for all times t the resulting metrics d_t retain a warped product form in both cases. In the first case, it has a conic singularity of angle $\sqrt{2}\pi$ at the apex for all t. In the second case, the asymptotic angle at the apex is zero for all t. Thus in these natural examples, the flow does not smoothen out the singularity.
In Sections 3.1 to 3.3 we will present the case of the cone of angle π in detail. For the cone of angle $\pi/2$ we will state the main results in Section 3.4 and omit some parts of the proofs, since the arguments are very similar.

3.1. **Preliminaries.**

We will first recall basic properties of Euclidean cones and give an explicit representation of the heat kernel on the cone of angle π in the sense of RCD spaces. Moreover, we will exhibit a convenient way to calculate Wasserstein distances in the cone, via a lifting procedure from the cone to $\mathbb{R}^2$.

3.1.1. *Euclidean cones and optimal transport.* The Euclidean cone $C(\theta)$ with angle $\theta \in [0, 2\pi]$ is defined as the quotient

$$C(\theta) \;=\; \Big([0,\infty) \times [0,\theta]\Big)\Big/ \sim \,,$$

where we write $(r,\alpha) \sim (s,\beta)$ if and only if $r = s = 0$ or $r = s$ and $|\alpha - \beta| \in \{0, \theta\}$. The cone distance d is given by

$$d\big((r,\alpha),(s,\beta)\big) \;=\; \sqrt{r^2 + s^2 - 2rs\cos\big(\min\big(|\alpha-\beta|, \theta - |\alpha-\beta|\big)\big)}\,,$$

which is well defined on the quotient. Note that the cone without the apex, i.e $C(\theta)\backslash\{o\}$, where o is the equivalence class of $(0,0)$, is an open Riemannian manifold with the metric tensor $(\mathrm{d}r)^2 + r^2(\mathrm{d}\alpha)^2$. Its geometry is locally Euclidean. The associated Riemannian distance is the cone distance and the distance on the full cone $C(\theta)$ is its metric completion.
We will be concerned in particular with the cone of angle π. In this case we have the alternative characterization as the quotient

$$C(\pi) = \mathbb{R}^2/\sigma\,,$$

where the map $\sigma : \mathbb{R}^2 \to \mathbb{R}^2$ is the reflection at the origin, i.e. $\sigma(x) = -x$. Let us denote by $P : \mathbb{R}^2 \to C(\pi)$ the canonical projection. Then the cone distance between $p, q \in C(\pi)$ can be written as

$$d(p,q) = \min\big(|x-y|, |x+y|\big)\,,$$

where $x, y \in \mathbb{R}^2$ are such that $P(x) = p, P(y) = q$. The Hausdorff measure on $C(\pi)$ is given as $m = \frac{1}{2}P_\# \operatorname{Leb}$, where Leb denotes the Lebesgue measure on $\mathbb{R}^2$.
Now, we show how to calculate efficiently Wasserstein distance in the cone $C(\pi)$. We will denote by $W_{\mathbb{R}^2}$ and $W_{C(\pi)}$ the L^2 transport distances on $\mathbb{R}^2$ and $C(\pi)$ built from the Euclidean distance and the cone distance d respectively. If no confusion can arise we shall simply write W.
Let us introduce the set of measures on $\mathbb{R}^2$ with finite second moment, that are symmetric with respect to the origin. We set

$$\mathscr{P}_2^{\mathrm{sym}}(\mathbb{R}^2) = \{\mu \in \mathscr{P}_2(\mathbb{R}^2) \;:\; \sigma_\#\mu = \mu\}\,.$$

Note that given a measure $\nu \in \mathscr{P}_2(C(\pi))$ there exists a unique measure $L(\nu) \in \mathscr{P}_2^{\mathrm{sym}}(\mathbb{R}^2)$ such that $P_\# L(\nu) = \nu$. We call $L(\nu)$ the *symmetric lift* of ν.
We have the following useful fact.

Lemma 3.1. *For any two measures $\mu, \nu \in \mathscr{P}_2(C(\pi))$ it holds*

$$W_{C(\pi)}(\mu,\nu) = W_{\mathbb{R}^2}(L(\mu), L(\nu))\,.$$

In other words, the mapping $\mathscr{P}_2^{\mathrm{sym}}(\mathbb{R}^2) \to \mathscr{P}_2(C(\pi))$, $\mu \mapsto P_\#\mu$ is an isometry. Moreover, for any two measures $\mu, \nu \in \mathscr{P}_2(\mathbb{R}^2)$ we have

$$W_{C(\pi)}(P_\#\mu, P_\#\nu) \le W_{\mathbb{R}^2}(\mu,\nu)\,.$$

Proof. Let us first prove the second statement. Let $\mu, \nu \in \mathscr{P}(\mathbb{R}^2)$ and π a transport plan between μ and ν. Define a transport plan $\bar{\pi}$ between $P_\#\mu$

and $P_{\#}\nu$ by setting $\bar\pi = (P \otimes P)_{\#}\pi$. Therefore,

$$\int |y-x|^2 \, \mathrm{d}\pi(x,y) \geq \int d(P(y),P(x))^2 \, \mathrm{d}\pi(x,y) = \int d^2 \, \mathrm{d}\bar\pi \; . \tag{3.1}$$

Taking the infimum over π, we get the second statement. We turn now to the first statement. Let $\mu, \nu \in \mathscr{P}_2^{\mathrm{sym}}(\mathbb{R}^2)$ and let $\bar\pi$ be a transport plan between $P_{\#}\mu$ and $P_{\#}\nu$. We can find a measurable map $Q : C(\pi) \times C(\pi) \to \mathbb{R}^2 \times \mathbb{R}^2$ such that $(P \otimes P) \circ Q = \mathrm{Id}$ and $|x-y| = d(p,q)$ for $Q(p,q) = (x,y)$. These properties also hold for $Q^- := -Q$. The marginals of the transport plan $\pi = \frac{1}{2}(Q_{\#}\bar\pi + Q^-_{\#}\bar\pi)$ are symmetric, hence they coincide with μ and ν. Moreover π is concentrated on the set $\{(x,y) \in \mathbb{R}^2 \times \mathbb{R}^2,\ d(P(x),P(y)) = |y-x|\}$ so that we have equality in (3.1). Taking the infimum over $\bar\pi$ and taking into account the second statement, we obtain the first statement. $\square$

3.1.2. *RCD structure and the heat kernel.* Here we verify that the cone $C(\pi)$ fits into the framework of RCD spaces considered in [17] and we give an explicit description of the heat kernel in this case.

Indeed, the metric measure space $(C(\pi), d, m)$ satisfies the condition $\mathrm{RCD}(0,\infty)$ as proven for instance in [21, Thm. 1.1]. In order to identify the heat semigroup H_t acting on measures and the heat kernel $H_t\delta_x$ in this example, it is sufficient to exhibit an explicit solution to the Evolution Variational Inequality using [4, Thm. 5.1], see Section 2.3. This will be done again via the lifting to $\mathbb{R}^2$.

We denote by γ_x^t the Gaussian measure with variance $2t$ centered at $x \in \mathbb{R}^2$:

$$\gamma_x^t(\mathrm{d}y) = \frac{1}{4\pi t} \exp\left(-\frac{|y-x|^2}{4t}\right) \, \mathrm{d}y \; .$$

The heat semigroup in $\mathbb{R}^2$ acting on measures is denoted by $H_t^{\mathbb{R}^2}$. More precisely, for any $\mu \in \mathscr{P}_2(\mathbb{R}^2)$ we set $H_t^{\mathbb{R}^2}\mu(\mathrm{d}x) = \int \gamma_y^t(\mathrm{d}x) \, \mathrm{d}\mu(y)$.

Now, put $\nu_p^t = P_{\#}(\gamma_x^t)$ where x is such that $P(x) = p$. We define a semigroup $H_t^{C(\pi)}$ acting on $\mathscr{P}(C(\pi))$ via

$$H_t^{C(\pi)}\mu(\mathrm{d}q) = \int \nu_p^t(\mathrm{d}q) \, \mathrm{d}\mu(p) \; .$$

Note that we have $H_t^{C(\pi)} = P_{\#} \circ H_t^{\mathbb{R}^2} \circ L$.

Lemma 3.2 (Evolution Variational Inequality)**.** *For every $\mu, \chi \in \mathscr{P}_2(C(\pi))$ such that $\mathrm{Ent}(\chi) < \infty$ and every $t \geq 0$ we have*

$$\frac{1}{2} W^2_{C(\pi)}(H_t^{C(\pi)}\mu, \chi) - \frac{1}{2} W^2_{C(\pi)}(\mu, \chi) \leq t\big[\, \mathrm{Ent}(\chi) - \mathrm{Ent}(H_t^{C(\pi)}\mu)\big] \; .$$

Proof. Let $L(\mu), L(\chi) \in \mathscr{P}_2^{\mathrm{sym}}(\mathbb{R}^2)$ be the lifts of μ, χ. Note that $H_t^{\mathbb{R}^2}L(\mu)$ is the symmetric lift of $H_t^{C(\pi)}\mu$. Since $H_t^{\mathbb{R}^2}$ satisfies the Evolution Variational Inequality, see e.g. [2, Thm. 11.2.5], we find

$$\begin{aligned} \frac{1}{2} W^2_{\mathbb{R}^2}(H_t^{\mathbb{R}^2}L(\mu), L(\chi)) &- \frac{1}{2} W^2_{\mathbb{R}^2}(L(\mu), L(\chi)) \\ &\leq t\big[\, \mathrm{Ent}(L(\chi)) - \mathrm{Ent}(H_t^{\mathbb{R}^2}L(\mu))\big] \; . \end{aligned}$$

Observing that $\mathrm{Ent}(L(\mu)) = \mathrm{Ent}(\mu)$ for any $\mu \in \mathscr{P}_2(C(\pi))$ and its symmetric lift $L(\mu)$ and using Lemma 3.1, this immediately yields the claim. $\square$

In view of [4, Thm. 5.1], this shows again that $(C(\pi), d, m)$ satisfies $\mathrm{RCD}(0,\infty)$ and that $H_t^{C(\pi)}$ is the associated heat semigroup. In particular, $\nu_p^t = H_t^{C(\pi)}\delta_p$ is the heat kernel at time t centered at p.
We finish this section by noting the following contraction property of the heat flow:

$$W_{C(\pi)}(\nu_p^t, \nu_q^t) \le d(p,q), \qquad \forall p, q \in C(\pi), t \ge 0 \; . \tag{3.2}$$

Indeed, choosing x, y with $P(x) = p, P(y) = q$ and $d(p,q) = |x-y|$, by Lemma 3.1, convexity of the squared Wasserstein distance together with the identity $W(\gamma_x^t, \gamma_y^t) = W(\gamma_{-x}^t, \gamma_{-y}^t)$ we have that

$$\begin{aligned} W_{C(\pi)}(\nu_p^t, \nu_q^t) &= W_{\mathbb{R}^2}\left(\frac{1}{2}(\gamma_x^t + \gamma_{-x}^t), \frac{1}{2}(\gamma_y^t + \gamma_{-y}^t)\right) \\ &\le W_{\mathbb{R}^2}(\gamma_x^t, \gamma_y^t) = |x-y| = d(p,q) \; . \end{aligned}$$

3.1.3. *Absolutely continuous curves and the continuity equation.* We recall the characterization of absolutely continuous curves in the Wasserstein space over Euclidean spaces via solutions to the continuity equation. Moreover, we formulate a convenient estimate on the driving vector field in the continuity equation.

Proposition 3.3 ([2, Thm. 8.3.1])**.** *A weakly continuous curve $(\mu_s)_{s\in[0,T]}$ in $\mathscr{P}_2(\mathbb{R}^n)$ is 2-absolutely continuous with respect to W if and only if there exists a Borel family of vector fields V_s with $\int_0^T \|V_s\|^2_{L^2(\mu_s;\mathbb{R}^n)}\,\mathrm{d}s < \infty$ such that the continuity equation*

$$\partial_s \mu + \mathrm{div}(\mu_s V_s) = 0$$

holds in distribution sense. In this case we have $|\dot\mu_s| \le \|V_s\|_{L^2(\mu_s;\mathbb{R}^n)}$ for a.e. s. Moreover, V_s is uniquely determined for a.e. s if we require

$$V_s \in T_{\mu_s}\mathscr{P}_2(\mathbb{R}^n) := \overline{\{\nabla\psi \mid \psi \in C_c^\infty(\mathbb{R}^n)\}}^{L^2(\mu_s;\mathbb{R}^n)}$$

and in this case it holds $|\dot\mu_s| = \|V_s\|_{L^2(\mu_s;\mathbb{R}^n)}$.

The next lemma states a simple condition for existence and uniqueness of solutions to the continuity equation.

Lemma 3.4. *Let $\mu \in \mathscr{P}_2(\mathbb{R}^n)$ with strictly positive Lebesgue density ρ and assume that μ satisfies the Poincaré inequality*

$$\int |f|^2 \,\mathrm{d}\mu \le C \int |\nabla f|^2 \,\mathrm{d}\mu \; ,$$

for all $f \in C_c^\infty(\mathbb{R}^n)$ with $\int f \,\mathrm{d}\mu = 0$. Let $s \in L^1(\mathbb{R}^n, \mathrm{Leb})$ be such that $\int s = 0$ and

$$\|s/\sqrt{\rho}\|^2_{L^2} \;=\; \int \frac{s^2(x)}{\rho(x)}\,\mathrm{d}x \;<\; \infty \; .$$

Then there exists a unique vector field $V \in T_\mu\mathscr{P}_2(\mathbb{R}^n)$ such that the equation

$$s + \mathrm{div}(\mu V) = 0$$

holds in distribution sense. Moreover, we have

$$\|V\|^2_{L^2(\mu;\mathbb{R}^n)} \;=\; \int |V|^2 \,\mathrm{d}\mu \;\le\; C\|s/\sqrt{\rho}\|^2_{L^2} \,. \tag{3.3}$$

Proof. For any $f \in C_c^\infty(\mathbb{R}^n)$ with $\int f \,\mathrm{d}\mu = 0$ we deduce from the Cauchy–Schwarz and Poincaré inequalities that the bilinear map $B : f \mapsto \int sf$ satisfies

$$B(f) \le \left(\int \frac{s^2}{\rho}\right)^{\frac{1}{2}} \left(\int f^2\rho\right)^{\frac{1}{2}} \le \|s/\sqrt{\rho}\|_{L^2}\sqrt{C}\left(\int |\nabla f|^2 \,\mathrm{d}\mu\right)^{\frac{1}{2}} \,.$$

Thus, identifying f with its gradient, the map B can be extended to a bounded linear functional on the Hilbert space $T_\mu := T_\mu\mathscr{P}_2(\mathbb{R}^n)$ equipped with the scalar product

$$\langle U, W\rangle_{L^2(\mu;\mathbb{R}^n)} = \int U\cdot W \,\mathrm{d}\mu \,.$$

Moreover, the norm of B is bounded by $\sqrt{C}\|s/\sqrt{\rho}\|_{L^2}$. Thus, by the Riesz representation theorem there exists a unique vector field $V \in T_\mu$ such that $B(W) = \langle V, W\rangle_{L^2(\mu;\mathbb{R}^n)}$ and $\|V\|_{L^2(\mu;\mathbb{R}^n)} \le \sqrt{C}\|s/\sqrt{\rho}\|_{L^2}$. In particular, for any f as above we have

$$\int sf \;=\; B(f) \;=\; \int V\cdot\nabla f \,\mathrm{d}\mu \,.$$

Thus V is the unique distributional solution to $s + \mathrm{div}(\mu V) = 0$ in T_μ. □

3.2. **Warped structure of the convoluted cone.** Having identified the heat kernel in Lemma 3.2, we can now analyse in detail the construction of [17] in the case of $C(\pi)$. Let us define $\iota_t : C(\pi) \to \mathscr{P}_2(C(\pi))$ via $\iota_t(p) = \nu_p^t$. This map is obviously injective and by (3.2) satisfies Assumption 2.1. Thus, as outlined in Section 2 we introduce

$$\widetilde{d}_t(p,q) = W_{C(\pi)}(\nu_p^t, \nu_q^t)$$

and define d_t to be the associated length distance as in (2.2). The rotational symmetry of $C(\pi)$ is preserved by this transformation so that the space $(C(\pi), d_t)$ will retain a warped structure.
We first give a partial converse to the Lipschitz estimate (2.1).

Lemma 3.5. *For any $t \ge 0$ and $r > 0$ there exists a constant $C(t,r)$ such that for all $p, q \in C(\pi)\setminus B_r$:*

$$C(t,r)d(p,q) \le \widetilde{d}_t(p,q) \,, \tag{3.4}$$

where $B_r = \{p \in C(\pi) : d(o,p) \le r\}$.

In particular, in view of Proposition 2.4 this shows that $\widetilde{d}_t$ and d_t induce the same topology as the cone distance on $C(\pi)$.

Proof. Pick $x, y \in \mathbb{R}^2$ with $d(p,q) = |x-y|$. Without restriction we can assume that $|x| \le |y|$ and that $x = (x_1, 0)$, $y = (y_1, y_2)$ with $y_1, y_2 \ge 0$. Let A be the line passing through the origin at angle $3\pi/8$ with the first coordinate axis and let pr_A denote the orthogonal projection onto L. Then, setting $\mu_x^t = \frac{1}{2}(\gamma_x^t + \gamma_{-x}^t)$, we have

$$\widetilde{d}_t(p,q) = W_{\mathbb{R}^2}(\mu_x^t, \mu_y^t) \geq W_{\mathbb{R}^1}\big((\mathrm{pr}_A)_\# \mu_x^t, (\mathrm{pr}_A)_\# \mu_y^t\big)$$

Note that $(\mathrm{pr}_A)_\# \mu_x^t$ is the mixture of two one-dimensional Gaussians with variance $2t$ and centers $\pm \mathrm{pr}_A x$. Note further that $|\mathrm{pr}_A(x) - \mathrm{pr}_A(y)| \geq \cos(3\pi/8)|x-y|$ since the angle of A with $y-x$ is less than $3\pi/8$. Thus it suffices to establish the following claim: For any $t,r > 0$ there exists a constant $C(t,r)$ such that for all $x, y \geq r$:

$$C(t,r)|x-y| \leq W_{\mathbb{R}}\big(\frac{1}{2}(\gamma_x^t + \gamma_{-x}^t), \frac{1}{2}(\gamma_y^t + \gamma_{-y}^t)\big) \; , \tag{3.5}$$

where by abuse of notation γ_x^t denotes also the one-dimensional Gaussian measure with variance $2t$ and center x. By scaling it suffices to consider $t = 1$. In dimension 1, the optimal transport plan is known to be the monotone rearrangement. The two measures in (3.5) are symmetric so that the mass on $\mathbb{R}^+$ is mapped on $\mathbb{R}^+$. Observe that the measure $\gamma_x^1 + \gamma_{-x}^1$ restricted to $\mathbb{R}^+$ is distributed as $\omega_\# \gamma_x^1$ where $\omega : x \mapsto |x|$. Hence the right hand side of (3.5) is $W_{\mathbb{R}}(\omega_\# \gamma_x^1, \omega_\# \gamma_y^1)$. Applying Jensen's inequality in the definition of $W_{\mathbb{R}}$ we see that the distance between the means of these measures is a lower bound. But the mean of $\omega_\# \gamma_x^1$ is $\int |s| \,\mathrm{d}\gamma_x^1(s) = \int |s-x| \,\mathrm{d}\gamma_0^1(s)$. As a function of $x \in \mathbb{R}^+$, this is a strictly convex function with derivative zero at zero on the right and tangent to the first bisector at $+\infty$. In fact the second derivative in distribution sense is $2\gamma_0^1$. The estimate (3.5) follows from these remarks together, provided $x, y \geq r$ for some $r > 0$. □

Next, we show that in the definition of d_t we can restrict the infimum to Lipschitz curves with respect to the cone distance d.

Lemma 3.6. *For any $t \geq 0$ and $p, q \in C(\pi)$ we have*

$$d_t(p,q) \;=\; \inf\left\{ \int_0^T |\dot{p}_s|_t \,\mathrm{d}s \right\} \; , \tag{3.6}$$

where the infimum is taken over all d-Lipschitz curves $(p_s)_{s\in[0,T]}$ such that $p_0 = p, p_T = q$ and $|\dot{p}_s|_t$ denotes the metric derivative with respect to $\widetilde{d}_t$. If $p \neq o$ and $q \neq o$, one can restrict to curves supported in $C(\pi) \setminus \{o\}$.

Thus the construction of d_t given here is consistent with the general construction in RCD spaces given in [17] (see Remark 2.2).

Proof. The inequality "$\leq$" follows immediately from the fact that any d-Lipschitz curve is also $\widetilde{d}_t$-Lipschitz by (3.2). To see the reverse inequality, first recall that by Remark 2.3 we can restrict the infimum in (2.2) to $\widetilde{d}_t$-Lipschitz curves. Then the statement follows from Lemma 3.5. Indeed, given $\varepsilon > 0$, let $(p_s)_{s\in[0,T]}$ be a $\widetilde{d}_t$-Lipschitz curve such that

$$\int_0^T |\dot{p}_s|_t \,\mathrm{d}s \leq d_t(p,q) + \varepsilon \; .$$

Recall that (p_s) is d-continuous. If it avoids the origin (and thus also a neighborhood around it) then (p_s) is also d-Lipschitz by (3.4). If the curve hits the origin, put for sufficiently small $r > 0$:

$$s_1 := \inf\{s \in [0,T] : p_s \in B_r\} , \qquad s_2 := \sup\{s \in [0,T] : p_s \in B_r\} .$$

We can construct a d-Lipschitz curve $(\widetilde{p}_s)$ by replacing the part $(p_s)_{s\in[s_1,s_2]}$ with a piece of a circle connecting p_{s_1} to p_{s_2}. From (3.2) we see that the $\widetilde{d}_t$-length of $(\widetilde{p}_s)$ is bounded by $d_t(p,q) + \varepsilon + \pi r$. Choosing r sufficiently small and using the arbitrariness of ε we obtain the inequality "$\geq$" in (3.6). In the case $p = o$ or $q = o$, the ray from or to the apex is a minimizing curve. Note that it is a d-Lipschitz curve. $\square$

Let us further observe the particular behavior of the distance under scaling of space and time.

Lemma 3.7. *For any $t > 0$ and $p, q \in C(\pi)$ we have*

$$d_t(p,q) = \sqrt{t} \cdot d_1\big(\sqrt{t}^{-1} p, \sqrt{t}^{-1} q\big) . \tag{3.7}$$

Here, for $\lambda \geq 0$ and $p = (r,\alpha) \in C(\pi)$ we set $\lambda p = (\lambda r, \alpha)$.

Proof. It suffices to establish the identity (3.7) with d_t replaced by $\widetilde{d}_t$. It then passes easily to the associated length distance. Recall that $\widetilde{d}_t(p,q) = W_{2,\mathbb{R}^2}(\frac{1}{2}(\gamma_x^t + \gamma_{-x}^t), \frac{1}{2}(\gamma_y^t + \gamma_{-y}^t)$ for x, y such that $P(x) = p, P(y) = q$. Introduce the dilation $s_\lambda : x \mapsto \lambda x$ and note that $\gamma_x^t = (s_{\sqrt{t}})_\# \gamma^1_{\sqrt{t}^{-1}x}$. Now the claim is immediate. $\square$

We have the following result on the metric structure of the convoluted cone.

Proposition 3.8. *The distance d_t is induced by a metric tensor g^t on the open manifold $C(\pi) \setminus \{o\}$ which is of warped product form*

$$g^t_{(r,\alpha)}(\cdot,\cdot) = R(r/\sqrt{t})\mathrm{d}r^2 + r^2 A(r/\sqrt{t})\mathrm{d}\alpha^2 , \tag{3.8}$$

where $R, A : (0,\infty) \to (0,1]$ are bounded functions. Moreover, the distance d_t on the full cone is obtained for $p_0, p_1 \in C(\pi)$ by

$$d_t\big(p_0,p_1\big) = \inf \int_0^1 \sqrt{R(r_s/\sqrt{t})|\dot{r}_s|^2 + r_s^2 A(r_s/\sqrt{t})|\dot{\alpha}_s|^2}\,\mathrm{d}s , \tag{3.9}$$

where the infimum is taken over all Lipschitz curves $(p_s)_{s\in[0,1]}$ of $(C(\pi), d)$ connecting p_0, p_1 and $|\dot{r}_s|, |\dot{\alpha}_s|$ denote the metric derivatives of the polar coordinates of p_s.

In (3.8) R and A stand for radial and angular.

Proof. Recall from Section 2 that

$$d_t(p,q) = \inf_{(p_s)} \int_0^T |\dot{\nu}^t_{p_s}|\,\mathrm{d}s ,$$

where $\nu^t_p = \iota_t(p)$ and $|\dot{\nu}^t_{p_s}|$ denotes the metric derivative with respect to $W_{C(\pi)}$. We will first use the lifting to $\mathbb{R}^2$ and the characterization of the Wasserstein metric derivative in terms of solutions to the continuity equation to relate d_t to a smooth metric tensor on $\mathbb{R}^2$ and then we will push this tensor to the cone to obtain the desired warped structure.

From Lemma 3.7 we immediately infer that it is sufficient to consider $t = 1$. For brevity let us set $\mu_x = \frac{1}{2}\gamma_x + \frac{1}{2}\gamma_{-x}$ and let f_x be its density, i.e. $f_x(y) =$

$\frac{1}{2}\eta(y-x)+\frac{1}{2}\eta(y+x)$, where $\eta(y)=\frac{1}{4\pi}\exp(-|y|^2/4)$ is the density of the 2-dimensional Gaussian at time 1.

We define a metric tensor $\widetilde{g}$ on $\mathbb{R}^2$ by setting for $x,w\in\mathbb{R}^2$:

$$\widetilde{g}_x(v,w) \;=\; \int\langle V_x^v, V_x^w\rangle\,\mathrm{d}\mu_x\;,$$

where V_x^w is the unique vector field in $T_{\mu_x}\mathscr{P}(\mathbb{R}^2)$ solving

$$\frac{\mathrm{d}}{\mathrm{d}h}\Big|_{h=0}f_{x+hw} \;=\; \frac{1}{2}w\cdot\big(\nabla\eta(x+y)-\nabla\eta(x-y)\big) \;=\; -\operatorname{div}(f_xV_x^w) \quad (3.10)$$

given by Lemma 3.4 (applied to $s=\frac{\mathrm{d}}{\mathrm{d}h}f_{x+hw}$ and $\mu=\mu_x$). Indeed, by uniqueness, V_x^w depends linearly on w, hence $\widetilde{g}_x(v,w)$ is a bilinear form.

Now, define a metric tensor g on the open manifold $C(\pi)\setminus\{o\}$ by setting for $p=(r,\alpha)\in C(\pi)\setminus\{o\}$ and $v,\theta\in\mathbb{R}$:

$$g_p\big((v,\theta),(v,\theta)\big)=\widetilde{g}_x(w,w)\;,$$

where

$$x=r\begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix}\;,\quad w=v\begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix}+\theta r\begin{pmatrix}-\sin\alpha\\ \cos\alpha\end{pmatrix}\;,$$

and extend via polarization. That g takes the form (3.8) is a consequence of the fact that $\widetilde{g}_x(w,w)$ is invariant under reflecting w at the line passing through the origin and x, implying that $g_p\big((v,\theta),(v,\theta)\big)=g_p\big((v,-\theta),(v,-\theta)\big)$, and the invariance of $\widetilde{g}_x(w,w)$ under simultaneous rotation of x,w. Explicitly, we have

$$R(r)=\widetilde{g}_{r(1,0)}\big((1,0),(1,0)\big)\;,\qquad A(r)=\widetilde{g}_{r(1,0)}\big((0,1),(0,1)\big)\;. \qquad (3.11)$$

Let us now prove (3.9), i.e. that d_1 is induced by the tensor g. By Lemma 3.6 we have

$$d_1(p_0,p_1)=\inf\int_0^1|\dot{p}_s|_1\,\mathrm{d}s\;,$$

where $|\dot{p}_s|_1$ is the metric derivative of p_s with respect to the distance $\widetilde{d}_1$ and the infimum is over Lipschitz curves in $(C(\pi),d)$. Let us consider a Lipschitz curve $(p_s)_{s\in[0,1]}$ in $(C(\pi),d)$ with polar coordinates $(r_s)_s$ and $(\alpha_s)_s$. Let $(x_s)_{s\in[0,1]}$ be a continuous a curve such that $P(x_s)=p_s$. By (3.2) the curves $\nu^1_{p_s}$ and μ_{x_s} are Lipschitz with respect to $W_{C(\pi)}$ and $W_{\mathbb{R}^2}$ respectively and by definition of $\widetilde{d}_1$ and Lemma 3.1 we have

$$|\dot{p}_s|_1=|\dot{\nu}^1_{p_s}|=|\dot{\mu}_{x_s}|\;,$$

where the latter two metric derivatives are calculated with respect to $W_{C(\pi)}$ and $W_{\mathbb{R}^2}$ respectively. By the characterization of absolutely continuous curves, there exists for a.e. s a vector field $V_s\in T_{\mu_{x_s}}\mathscr{P}(\mathbb{R}^2)$ such that the continuity equation $\partial_s\mu_{x_s}=-\operatorname{div}(\mu_{x_s}V_s)$ holds in distribution sense. But for a.e. s the left hand side is given by $\frac{1}{2}w_s\cdot\big(\nabla\eta(\cdot+x_s)-\nabla\eta(\cdot-x_s)\big)$ with

$$w_s=\dot{r}_s\begin{pmatrix}\cos\alpha_s\\ \sin\alpha_s\end{pmatrix}+\dot{\alpha}_sr_s\begin{pmatrix}-\sin\alpha_s\\ \cos\alpha_s\end{pmatrix}\;.$$

Hence, the uniqueness statement in Lemma 3.4 implies that for a.e. s we have $V_s = V_{x_s}^{w_s}$ and thus

$$|\dot{p}_s|_1^2 = \widetilde{g}_{x_s}(w_s, w_s) = g_{p_s}\big((\dot{r}_s, \dot{\alpha}_s), (\dot{r}_s, \dot{\alpha}_s)\big) = |\dot{r}_s|^2 R(r_s) + r_s^2 |\dot{\alpha}_s|^2 A(r_s) .$$

This yields that $d_t(p_0, p_1)$ is given by the right hand side in (3.9).
Finally, we turn to the boundedness of R and A. For w a vector of $\mathbb{R}^2$ the vector field $V : y \mapsto \lambda_x(y) w + (1 - \lambda_x(y))(-w)$ with

$$\lambda_x(y) = \frac{1}{2}\frac{\eta(y-x)}{f_x(y)} = \frac{\eta(y-x)}{\eta(y-x) + \eta(y+x)}\ , 1 - \lambda_x(y) = \frac{1}{2}\frac{\eta(y+x)}{f_x(y)}$$

satisfies (3.10) in place of V_x^w. Indeed, we have

$$\begin{aligned} -\operatorname{div}(f_x V) &= -\operatorname{div}\Big(w \cdot \frac{1}{2}\big(\eta(\cdot - x) - \eta(\cdot + x)\big)\Big) \\ &= -\frac{1}{2} w \cdot \big(\nabla\eta(\cdot - x) - \nabla\eta(\cdot + x))\big) = \frac{\mathrm{d}}{\mathrm{dh}}\Big|_{h=0} f_{x+hw}\ , \end{aligned}$$

V is an element of $L^2(\mu_x; \mathbb{R}^n)$ with norm smaller than or equal to $|w|$. The orthogonal projection on $T_{\mu_x}\mathscr{P}(\mathbb{R}^2)$ contracts the norm and provides another solution to (3.10). According to the uniqueness statement in Lemma 3.4 it is V_x^w. Hence we have proved $\widetilde{g}_x(w, w) \le |w|^2$. It follows that the functions A and R defined in (3.11) are bounded from above by 1.

□

Remark 3.9. We believe that the metric tensor g^t which induces d_t in the proposition above is smooth on $C(\pi) \setminus \{o\}$. From the explicit expression (3.16) for R given below in the proof of Theorem 3.11, it is readily checked that R is smooth. However, proving smoothness for A by general arguments seems non-trivial due to the non-compactness of the cone. The same situation arises for $C(\pi/2) \setminus \{o\}$ with the corresponding distance below. In Remark 3.19 we explain how the conjectured smoothness would allow to conlude the presence of unbounded negative curvature near the apex in this case.

Proposition 3.10. *As t goes to zero, the metric space $(C(\pi), d_t)$ tends to $(C(\pi), d)$ pointwise and in the pointed Gromov–Hausdorff topology.*

Proof. By construction of d_t and by the contractivity (3.2) we have the chain of inequalities

$$\widetilde{d_t} \le d_t \le d\ .$$

From Proposition 2.4 we already know that $\widetilde{d_t}$ converges pointwise to d as $t \to 0$ whence the convergence of d_t follows. The convergence in pointed Gromov–Hausdorff topology follows as in the proof of Proposition 2.4. □

As an immediate consequence of Proposition 3.8, we deduce that a d_t-minimizing curve connecting the apex $o = (0,0)$ to the point $(r, 0) \in C(\pi)$ is given by the curve $(sr, 0)_{s\in[0,1]}$. Hence the distance of $(r, 0)$ from the apex o is

$$d_t\big((r,0), o\big) = \int_0^1 r\sqrt{R\big(sr/\sqrt{t}\big)}\ \mathrm{d}s = \int_0^r \sqrt{R\left(s/\sqrt{t}\right)}\ \mathrm{d}s\ . \tag{3.12}$$

3.3. **Persistence of the conic singularity for $C(\pi)$.** We will show that the new distance d_t has a conic singularity at the origin of angle $\sqrt{2}\pi$ independent of t. In order to do so we will compare for small r the distance of a point $p_r = (r, 0)$ from the origin to the length of a circle around the origin passing through p_r.
More precisely, for $r > 0$ set $\rho_t(r) = d_t\big((0, r), o\big)$ and define

$$l_t(r) = \int_0^\pi |\dot{p}^r_s|_t \, \mathrm{d}s \; .$$

where the curve $p^r : [0, \pi] \to C(\pi)$ is given by $p^r_s = (r, s)$.

Theorem 3.11. *For each $t > 0$ we have*

$$\lim_{r\to 0} \frac{l_t(r)}{\rho_t(r)} = \sqrt{2}\pi \; .$$

In other words, the angle at the apex o is $\sqrt{2}\pi$. In particular, a singularity persists at o. With the notation of Proposition 3.8 we have more precisely $R(r) \sim r^2/2$ and $A(r) \sim r^2/4$ as $r \to 0$.
Moreover, $C(\sqrt{2}\pi)$ is both, the tangent space of $(C(\pi), d_t)$ at o, and the limit in the pointed Gromov–Hausdorff topology as t goes to infinity.

Remark 3.12. The discontinuity at $t = 0$ of the asymptotic angle at o might seem intriguing at first in view of the convergence of d_t to the original distance d given by Proposition 3.10. Note however, that the asymptotic angle is in a certain sense a first order quantity, while the convergence of distances is zero order. Intuitively, the discontinuity can be understood from the scaling property (3.7). After zooming in at scale r, the heat kernel measure at a very small time t looks like the heat kernel measure at the larger time $t/\sqrt{r}$ at the original scale.

Proof. We will calculate ρ_t and l_t asymptotically as $r \to 0$. From Proposition 3.8 and (3.12) we have

$$\rho_t(r) = \int_0^r \sqrt{R\big(s/\sqrt{t}\big)} \, \mathrm{d}s \tag{3.13}$$

$$l_t(r) = \pi r \sqrt{A\big(r/\sqrt{t}\big)} \; . \tag{3.14}$$

Thus, it remains to calculate R and A. Denote by f_x the density of $\mu^1_x = \frac{1}{2}\gamma^1_x + \frac{1}{2}\gamma^1_{-x}$. We set $x_r = (r, 0) \in \mathbb{R}^2$ and recall from the proof of Proposition 3.8 that

$$R(r) = ||V^{(1,0)}_{x_r}||^2_{L^2(\mu^1_{x_r};\mathbb{R}^2)} \; ,$$

$$A(r) = ||V^{(0,1)}_{x_r}||^2_{L^2(\mu^1_{x_r};\mathbb{R}^2)} \; ,$$

where the vector field $V^w_x \in T_{\mu^1_x}\mathscr{P}(\mathbb{R}^2)$ is defined uniquely by the continuity equation

$$\frac{\mathrm{d}}{\mathrm{d}h}\Big|_{h=0} f_{x+hw} \;=\; -\operatorname{div}(f_x V^w_x) \; . \tag{3.15}$$

Note that we have

$$f_x(y) = 1/2\big(\eta(y_1 - x_1)\eta(y_2 - x_2) + \eta(y_1 + x_1)\eta(y_2 + x_2)\big) \; ,$$

where we write $x = (x_1, x_2)$ and $y = (y_1, y_2)$ and by abuse of notation $\eta(s) = (4\pi)^{-1/2}\exp(-s^2/4)$ now also denotes the 1-dimensional Gaussian density. Let us first concentrate on R. Here, we have to solve

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}h}\Big|_{h=0} f_{x_r+h(1,0)}(y) &= \eta(y_2)\frac{1}{2}\big(-\eta'(y_1-r)+\eta'(y_1+r)\big)\\ &= -\mathrm{div}\left(f_{x_r}V_{x_r}^{(1,0)}\right)(y)\ .\end{aligned}$$

It is easily checked that the solution is given by

$$V_{x_r}^{(1,0)}(y) = \begin{pmatrix}1\\0\end{pmatrix}\frac{\eta(y_1-r)-\eta(y_1+r)}{\eta(y_1-r)+\eta(y_1+r)} = \begin{pmatrix}1\\0\end{pmatrix}\frac{1}{2}\eta(y_2)\frac{\eta(y_1-r)-\eta(y_1+r)}{f_{x_r}(y_1,y_2)}\ .$$

Indeed, we have

$$\begin{aligned}-\,\mathrm{div}(f_{x_r}V_{x_r}^{(1,0)})(y) &= -\,\mathrm{div}\left(\begin{pmatrix}1\\0\end{pmatrix}\frac{1}{2}\eta(y_2)\big(\eta(y_1-r)-\eta(y_1+r)\big)\right)\\ &= -\frac{1}{2}\eta(y_2)\big(\eta'(y_1-r)-\eta'(y_1+r)\big)\ ,\end{aligned}$$

as claimed. Obviously, $V_{x_r}^{(1,0)}$ belongs to $T_{\mu^1_{x_r}}\mathscr{P}(\mathbb{R}^2)$. Thus, we have

$$\begin{aligned}R(r) &= \int |V_{x_r}^{(1,0)}|^2\,\mathrm{d}\mu^1_{x_r} = \frac{1}{2}\int\int \eta(y_2)\frac{|\eta(y_1-r)-\eta(y_1+r)|^2}{\eta(y_1-r)+\eta(y_1+r)}\,\mathrm{d}y_1\,\mathrm{d}y_2\\ &= \frac{1}{2}\int\frac{|\eta(y_1-r)-\eta(y_1+r)|^2}{\eta(y_1-r)+\eta(y_1+r)}\,\mathrm{d}y_1\ . \end{aligned} \tag{3.16}$$

To determine the asymptotic behavior as $r\to 0$, we first note that

$$R(r) = \frac{1}{2}r^2\int\frac{|2\eta'(y_1)|^2}{2\eta(y_1)}\,\mathrm{d}x + o(r^2) = r^2\int\frac{y_1^2}{4}\eta(y_1)\,\mathrm{d}y_1 + o(r^2) = \frac{r^2}{2}+o(r^2)\ . \tag{3.17}$$

Let us turn to calculating A. First we note that the left hand side of the continuity equation (3.15) is

$$\begin{aligned}\beta^r(y) := \frac{\mathrm{d}}{\mathrm{d}h}\Big|_{h=0} f_{x_r+h(0,1)}(y) &= \eta'(y_2)\frac{1}{2}\big[-\eta(y_1-r)+\eta(y_1+r)\big]\\ &= r\eta'(y_1)\eta'(y_2)+o(r)\ .\end{aligned}$$

Unfortunately, we can not explicitly solve equation (3.15) in $T_{\mu^1_{x_r}}\mathcal{P}(\mathbb{R}^2)$ but we can approximate the solution. To this end introduce the function $\psi^r : \mathbb{R}^2\to\mathbb{R}$ given by

$$\psi^r(y) = r\frac{1}{4}y_1y_2\ .$$

We calculate

$$\begin{aligned}
\delta^r(y) :=&\ -\operatorname{div}\big(f_{x_r}\nabla\psi^r\big)(y)\\
=&-\frac{r}{4}\Big[y_2\eta(y_2)\frac{1}{2}\big(\eta'(y_1+r)+\eta'(y_1-r)\big)\\
&\qquad+\eta'(y_2)y_1\frac{1}{2}\big(\eta(y_1+r)+\eta(y_1-r)\big)\Big]\\
=&\ \frac{r}{4}\Big[2\eta'(y_2)\big(\eta'(y_1+r)+\eta'(y_1-r)\big)\\
&\qquad+r\eta'(y_2)\frac{1}{2}\big(\eta(y_1+r)-\eta(y_1-r)\big)\Big]\\
=&\ \frac{r}{2}\eta'(y_2)\big(\eta'(y_1+r)+\eta'(y_1-r)\big)\\
&\qquad+\frac{r^2}{8}\eta'(y_2)\big(\eta(y_1+r)-\eta(y_1-r)\big)\ ,
\end{aligned}$$

where in the third equality we used several times the identity $\eta'(t) = -(t/2)\eta(t)$. Considering an expansion in r at $r = 0$ we check that as $r \to 0$ we have

$$\frac{1}{r^2}\int\frac{|\beta^r-\delta^r|^2}{f_{x_r}}\to 0\ .$$

Since the measures $\mu^1_{x_r}$ satisfy the Poincaré inequality with constant independent of r, we deduce by Lemma 3.4 (applied to $\nabla\psi^r - V^{(0,1)}_{x_r}$ and $\mu^1_{x_r}$) that

$$\frac{1}{r}\|\nabla\psi^r - V^{(0,1)}_{x_r}\|_{L^2(\mu^1_{x_r};\mathbb{R}^2)}\to 0\ .$$

It thus suffices to calculate

$$\begin{aligned}
&||\nabla\psi^r||_{L^2(\mu^1_{x_r};\mathbb{R}^2)}\\
&=\ \frac{r}{4}\left(\int(y_1^2+y_2^2)\eta(y_2)\frac{1}{2}\big(\eta(y_1+r)+\eta(y_1-r)\big)\,\mathrm{d}y_1\,\mathrm{d}y_2\right)^{\frac{1}{2}}\\
&=\ \frac{r}{4}\sqrt{4+r^2}\ =\ \frac{r}{2}+o(r^2)\ .
\end{aligned}$$

Thus $\sqrt{A(r)} = \frac{r}{2} + o(r)$. This together with (3.13), (3.14) yields

$$\begin{aligned}
\rho_t(r) &= \frac{1}{2\sqrt{2t}}r^2+o(r^2)\ ,\\
l_t(r) &= \pi\frac{1}{2\sqrt{t}}r^2+o(r^2)\ .
\end{aligned}$$

This gives the claim on the limit ratio.

For the last part of the statement let us consider the reparametrization $\mathcal{T} : (r, \theta) \in C(\pi) \mapsto (\rho(r), \theta) \in C(\pi)$ where ρ stands for ρ_t at time 1. We note $(\bar r, \theta)$ the new coordinates. The function ρ is continuously differentiable of positive derivative so that $\mathcal{T}$ is a diffeomorphism outside the apex. A curve $(\gamma_s)_{s\in[0,T]}$ with support on $C(\pi)\setminus\{o\}$ is Lipschitz if and only if $(\mathcal{T}\circ\gamma_s)_s$ is Lipschitz. Moreover, a change of variable shows how to compute the length on the second curve with the tensor defined by

$$\bar R(\bar r) = 1 \quad\text{and}\quad \bar A(\bar r)\bar r^2 = A(\rho^{-1}(\bar r))\times\rho^{-1}(\bar r)^2$$

in place of R and $A(r)r^2$. We proved in Lemma 3.5 that in the minimisation problem (3.9) it is possible to use Lipschitz curves outside the apex, or Lipschitz rays from or to the apex. Both classes of curves are preserved by $\mathcal{T}$ and $\mathcal{T}^{-1}$. Finally similarly as in Lemma 3.5 the infimum of the length in the new coordinates remains the same if it is allowed to test with Lipschitz curves going through o. We obtain $\bar{A}^{1/2} \sim_{\bar{r}\to 0} \sqrt{2}$. Note that the equation $\bar{A} = c$ would corresponds to the metric of $C(c\pi)$. With the new coordinates we easily recognize that the tangent space at zero is $C(\sqrt{2}\pi)$. Together with the time-space scaling of Lemma 3.7 we obtain the same limit space when t goes to infinity. $\square$

Remark 3.13. The intuition for finding a good candidate $\nabla\varphi$ for the solution to

$$\beta^r = -\operatorname{div}(f_{x_r}\nabla\varphi)$$

is as follows. Since we are interested only in the limit $r \to 0$ we expand the continuity equation in r. Note that $f_{x_r}(y_1, y_2) = \eta(y_1)\eta(y_2) + o(r)$. We expand $\nabla\varphi = \nabla\varphi_{(0)} + r\nabla\varphi_{(1)} + o(r)$. Since $\beta^r = r\eta'(y_1)\eta'(y_2) + o(r)$ we conclude that $\nabla\varphi_0 = 0$ and that we must have

$$\begin{aligned}\eta'(y_1)\eta'(y_2) &= -\operatorname{div}(\eta\otimes\eta\nabla\varphi_{(1)})(y_1, y_2)\\ &= -\eta'(y_1)\eta(y_2)\partial_{y_1}\varphi_{(1)} - \eta(y_1)\eta'(y_2)\partial_{y_2}\varphi_{(1)}\\ &\quad - \eta(y_1)\eta(y_2)\Delta\varphi_{(1)}(y_1, y_2)\end{aligned}$$

or equivalently, since $\eta'(u) = -(u/2)\eta(u)$,

$$y_1y_2/4 - (y_1/2)\partial_{y_1}\varphi_{(1)}(y) - (y_2/2)\partial_{y_2}\varphi_{(1)}(y) + \Delta\varphi_{(1)}(y) = 0\ .$$

A solution to this is given by $\varphi_{(1)}(y) = \frac{1}{4}y_1y_2$.

Remark 3.14. Unsurprisingly, the asymptotic angle of $\big(C(\pi), d_t\big)$ at infinity remains π independently of t. More precisely, for any $t \geq 0$:

$$\lim_{r\to\infty}\frac{l_t(r)}{\rho_t(r)} \;=\; \pi\ . \tag{3.18}$$

Indeed, in view of (3.13), (3.14) we find after a change of variables in the integral that

$$\frac{l_t(r)}{\rho_t(r)} \;=\; \frac{\pi\sqrt{A(r/\sqrt{t})}\times(r/\sqrt{t})}{\int_0^{r/\sqrt{t}}\sqrt{R(s)}\,\mathrm{d}s}\ .$$

Thus we have

$$\lim_{r\to\infty}\frac{l_t(r)}{\rho_t(r)} \;=\; \lim_{t\to 0}\frac{l_t(r)}{\rho_t(r)} \;=\; \pi\ .$$

In the last equality we used Proposition 3.10 for $\rho_t(r) \to r$ and also a representation of the length of type (2.4) together with $\widetilde{d_t} \leq d_t \leq d$ for $l_r(t) \to \pi r$.

3.4. **The cone of angle $\pi/2$.** Like $C(\pi)$, the cone $C(\pi/2)$ admits an alternative characterization as a quotient of $\mathbb{R}^2$, however, it will be convenient to phrase this in terms of complex numbers. We have

$$C(\pi/2) = \mathbb{C}/\sigma \ ,$$

where the map $\sigma : \mathbb{C} \to \mathbb{C}$ is the rotation by $\pi/2$, i.e. $\sigma(z) = iz$. Let us denote by $P : \mathbb{C} \to C(\pi/2)$ the canonical projection. Then the cone distance between $p, q \in C(\pi/2)$ can be written as

$$d(p,q) = \min\left\{ |z - e^{ik\pi/2} z'| : k = 0,1,2,3 \right\} ,$$

where $z, z' \in \mathbb{C}$ are such that $P(z) = p$, $P(z') = q$. The Hausdorff measure on $C(\pi/2)$ is given as $m = \frac{1}{4} P_\# \operatorname{Leb}$, where Leb denotes the Lebesgue measure on $\mathbb{C}$.

As in section 3.1 we can calculate Wasserstein distances in the cone via lifting. Given $\nu \in \mathscr{P}_2(C(\pi/2))$ we denote by $L(\nu)$ the *symmetric lift* of ν, i.e. the unique measure in

$$\mathscr{P}_2^{\mathrm{sym}}(\mathbb{C}) := \{\mu \in \mathscr{P}_2(\mathbb{C}) \ : \ \sigma_\# \mu = \mu\} \ .$$

such that $P_\# L(\nu) = \nu$. Then, in analogy to Lemma 3.1, for any two measures $\mu, \nu \in \mathscr{P}_2(C(\pi/2))$ we obtain

$$W_{C(\pi/2)}(\mu,\nu) = W_{\mathbb{C}}(L(\mu), L(\nu)) \ .$$

Recall that γ_z^t denotes the two-dimensional Gaussian measure with variance $2t$ centered at $z \in \mathbb{C}$. We set $\nu_p^t = P_\#(\gamma_z^t)$, where $p = P(z)$. By the obvious analogue of Lemma 3.2, ν_p^t is the heat kernel measure on $C(\pi/2)$ in the sense of RCD spaces. Note that its lift is given by

$$L(\nu_p^t) = \frac{1}{4}\Big[\gamma_z^t + \gamma_{iz}^t + \gamma_{-z}^t + \gamma_{-iz}^t\Big] \ .$$

Let d_t be the length distance associated to $\widetilde{d}_t(p,q) = W_{C(\pi/2)}(\nu_p^t, \nu_q^t)$. It is found again to satisfy the scaling relation (3.7). Arguing exactly as in Proposition 3.8 and Proposition 3.10 we obtain

Proposition 3.15. *The distance d_t is induced by a metric tensor g^t on the open manifold $C(\pi/2) \setminus \{o\}$ which is of warped product form*

$$g^t_{(r,\alpha)}(\cdot,\cdot) = R(r/\sqrt{t})\mathrm{d}r^2 + r^2 A(r/\sqrt{t})\mathrm{d}\alpha^2 \ , \tag{3.19}$$

where $R, A : (0,\infty) \to (0,1]$ are bounded functions.

Of course, the precise form of the functions R and A is different for $C(\pi/2)$ and $C(\pi)$.

Proposition 3.16. *As t goes to zero, the metric space $(C(\pi/2), d_t)$ tends to $(C(\pi/2), d)$ pointwise and in the pointed Gromov–Hausdorff topology.*

In order to calculate the angle at the apex, for $r > 0$ let us set again $\rho_t(r) = d_t\big(o, (0,r)\big)$, where o denotes the apex, as well as

$$l_t(r) = \int_0^{\pi/2} |\dot{p}_s^r|_t \ \mathrm{d}s \ .$$

where the curve $p^r : [0,\pi/2] \to C(\pi/2) = \big([0,\infty) \times [0,\pi/2]\big)/\sim$ is given by $p_s^r = (r,s)$.

Theorem 3.17. *For each $t > 0$ we have*

$$\lim_{r\to 0} \frac{l_t(r)}{\rho_t(r)} = 0 \ .$$

In other words, the angle at the apex o is zero. We have more precisely $R(r) \sim r^2/4$ and $A(r) \in O(r^6)$ as $r \to 0$.
Moreover, $\mathbb{R}^+$ is both, the tangent space of $(C(\pi/2), d_t)$ at o, and the limit in the pointed Gromov–Hausdorff topology as t goes to infinity.

Proof. We will follow the same reasoning as in the proof of Theorem 3.11. Let us highlight the main steps. By scaling, we can again assume that $t = 1$. Let us denote by $\mu^1_z = \frac{1}{4}\big[\gamma^1_z + \gamma^1_{iz} + \gamma^1_{-z} + \gamma^1_{-iz}\big]$ the lift of $\nu^1_{P(z)}$ and denote by f_z its density with respect to the Lebesgue measure. Recalling the expressions (3.13), (3.14) (the latter with π replaced by $\pi/2$) for ρ and l, it is sufficient to calculate $R(r)$ and $A(r)$ asymptotically as $r \to 0$. We calculate R exactly in a similar way as in Theorem 3.11. We set $z_r = re^{i\pi/4}$ and recall that $R(r) = \|V_r\|^2_{L^2(\mu^1_{z_r};\mathbb{R}^2)}$, where V_r is the unique vector field in $T_{\mu^1_{z_r}}\mathscr{P}(\mathbb{R}^2)$ solving the continuity equation

$$\frac{\mathrm{d}}{\mathrm{d}h}\Big|_{h=0} f_{z_r+he^{i\pi/4}} + \operatorname{div}(f_{z_r} V_r) = 0 \ . \tag{3.20}$$

Using the explicit expression

$$f_{z_r}(z) = \frac{1}{4}\Big[\eta(x_1 - r/\sqrt{2}) + \eta(x_1 + r/\sqrt{2})\Big]\Big[\eta(x_2 - r/\sqrt{2}) + \eta(x_2 + r/\sqrt{2})\Big] \ ,$$

we readily check that

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}h}\Big|_{h=0} f_{z_r+he^{i\pi/4}} = \\
&\frac{1}{4\sqrt{2}}\Big[-\eta'(x_1 - r/\sqrt{2}) + \eta'(x_1 + r/\sqrt{2})\Big]\Big[\eta(x_2 - r/\sqrt{2}) + \eta(x_2 + r/\sqrt{2})\Big] \\
&+\frac{1}{4\sqrt{2}}\Big[\eta(x_1 - r/\sqrt{2}) + \eta(x_1 + r/\sqrt{2})\Big]\Big[-\eta'(x_2 - r/\sqrt{2}) + \eta'(x_2 + r/\sqrt{2})\Big] \ ,
\end{aligned}$$

and that the solution to (3.20) is given by

$$V_r(z) = \begin{pmatrix}\varphi_r(x_1)\\ \varphi_r(x_2)\end{pmatrix} \ , \qquad \varphi_r(x) = \frac{1}{\sqrt{2}}\frac{\eta(x - r/\sqrt{2}) - \eta(x + r/\sqrt{2})}{\eta(x + r/\sqrt{2}) + \eta(x - r/\sqrt{2})} \ .$$

Thus, we find

$$\begin{aligned}
R(r) = \int_{\mathbb{R}^2} |V_r|^2 f_{z_r} &= \frac{1}{2}\int \frac{|\eta(x + r/\sqrt{2}) - \eta(x - r/\sqrt{2})|^2}{\eta(x + r/\sqrt{2}) + \eta(x - r/\sqrt{2})}\,\mathrm{d}x \\
&= \frac{r^2}{2}\int \frac{|\eta'(x)|^2}{\eta(x)}\,\mathrm{d}x + o(r^2) = \frac{r^2}{4} + o(r^2) \ .
\end{aligned}$$

Let us turn to calculating A. Here, it is convenient to set $z_r = (r, 0)$ and recall that $A(r) = \|V_r\|^2_{L^2(\mu^1_{z_r};\mathbb{R}^2)}$, where V_r is the unique solution in $T_{\mu^1_{z_r}}\mathscr{P}(\mathbb{R}^2)$ to the continuity equation

$$\frac{\mathrm{d}}{\mathrm{d}h}\Big|_{h=0} f_{z_r+(0,h)} + \operatorname{div}(f_{z_0} V_r) = 0 \ . \tag{3.21}$$

We will again approximate the solution. First note that

$$\begin{aligned}\beta_r(x) :=& \frac{\mathrm{d}}{\mathrm{d}h}\Big|_{h=0} f_{z_r+(0,h)}(x)\\ =& \frac{1}{4}\Big[\eta'(x_1)\Big(\eta(x_2+r)-\eta(x_2-r)\Big)-\eta'(x_2)\Big(\eta(x_1+r)-\eta(x_1-r)\Big)\Big]\\ =& \frac{r^3}{12}\Big[\eta'(x_1)\eta'''(x_2)-\eta'''(x_1)\eta'(x_2)\Big]+o(r^3)\\ =& \frac{r^3}{192}(x_1^3x_2-x_1x_2^3)\eta(x_1)\eta(x_2)+o(r^3)\ ,\end{aligned}$$

where in the last equality we have used that $\eta'(u)=-(u/2)\eta(u)$ and $\eta'''(u)=(-\frac{1}{8}u^3+\frac{3}{4}u)\eta(u)$. Note that

$$f_{z_r}(x)=\frac{1}{4}\big[\eta(x_1)\big(\eta(x_2+r)+\eta(x_2-r)\big)+\eta(x_2)\big(\eta(x_1+r)+\eta(x_1-r)\big)\big]\ .$$

Now, set $\psi^r(x)=\frac{r^3}{384}[x_1^3x_2-x_1x_2^3]$ and calculate

$$\begin{aligned}\delta_r(x) :=& -\operatorname{div}(f_{z_r}\nabla\psi^r)(x)=-\partial_1 f_{z_r}\partial_1\psi^r-\partial_2 f_{z_r}\partial_2\psi^r\\ =& -\frac{r^3}{384}(3x_1^2x_2-x_2^3)\frac{1}{4}\Big[\eta'(x_1)\big(\eta(x_2-r)+\eta(x_2+r)\big)\\ &\qquad\qquad+\eta(x_2)\big(\eta'(x_1-r)+\eta'(x_1+r)\big)\Big]\\ &+\frac{r^3}{384}(3x_2^2x_1-x_1^3)\frac{1}{4}\Big[\eta(x_1)\big(\eta'(x_2-r)+\eta'(x_2+r)\big)\\ &\qquad\qquad+\eta'(x_2)\big(\eta(x_1-r)+\eta(x_1+r)\big)\Big]\\ =& -\frac{r^3}{384}\Big[(3x_1^2x_2-x_2^3)\eta'(x_1)\eta(x_2)-(3x_2^2x_1-x_1^3)\eta(x_1)\eta'(x_2)\Big]+o(r^3)\\ =&\frac{r^3}{192}(x_1^3x_2-x_1x_2^3)\eta(x_1)\eta(x_2)+o(r^3)\ .\end{aligned}$$

Hence, as $r\to 0$ we obtain that

$$\frac{1}{r^6}\int\frac{|\beta^r-\delta^r|^2}{f_{z_r}}\to 0\ .$$

Since the measures $\mu^1_{z_r}$ satisfy the Poincaré inequality with constant independent of r, we deduce by Lemma 3.4 that

$$\frac{1}{r^3}\|\nabla\psi^r-V_r\|_{L^2(\mu^1_{z_r};\mathbb{R}^2)}\to 0\ .$$

This yields that $\sqrt{A(r)}=\|\nabla\psi^r\|_{L^2(\mu^1_{z_r})}+o(r^3)=Cr^3+o(r^3)$ for a suitable constant C. Using finally (3.13), (3.14) we find that $\rho_1(r)$ is of order r^2, while $l_1(r)$ is of order r^4. This yields the claim on the ratio.
In analogy with the end of the proof of Theorem 3.11 concerning the transformation $\mathcal{T}$, and with the notation adapted from it we find $\bar{R}=1$ and $\bar{A}=o(1)$ when $\bar{r}$ goes to zero. One recognizes that the tangent cone is $\mathbb{R}^+$ and, using the space-time scaling similarly to Lemma 3.7 one sees that $\mathbb{R}^+$ is also the pointed Gromov–Hausdorff limit when t goes to infinity. □

Remark 3.18. A consequence of the previous theorem is that the space $\big(C(\pi/2),d_t\big)$ is not an Alexandrov space of curvature bounded below by $-K$

for any $K > 0$. To see this, recall that in such a space the angle between to geodesics γ, η connecting a point a to points b and c respectively is always larger than the angle at $\bar{a}$ in a comparison triangle $(\bar{a}, \bar{b}, \bar{c})$ in the hyperbolic space of curvature $-K$ defined by having the same sidelenghts as the original triangle (a, b, c). Let now $a = o$, the apex, and $b = (r, 0), c = (r, \pi/4)$ in $\big(C(\pi/2), d_t\big)$. Proposition 3.15 and the expansion $R(r) = r^2/4 + o(r^2)$, $A(r) = O(r^6)$ from Theorem 3.17 show that the angle between the geodesics connecting a to b and c is 0. But the comparison triangle in hyperbolic space clearly has no vanishing angle.

Remark 3.19. Let us assume that the metric tensor (3.19) inducing d_t is smooth. As in the case of $C(\pi)$, see Remark 3.9 above, R is easily checked to be smooth from the explicit formula given in the proof above. We strongly believe A to be smooth but proving this seems intricate due to the non-compactness of the cone. The previous remark then implies that the metric measure space $\big(C(\pi/2), d_t, m_t\big)$, where m_t is the Hausdorff measure induced by d_t, is not a $\mathrm{CD}(-K, N)$ space for any $K > 0$ and $N \geq 1$. Indeed, if it were a $\mathrm{CD}(-K, N)$ space, the two-dimensional open smooth Riemannian manifold $\big(C(\pi/2) \setminus \{o\}, g_t\big)$ would have curvature bounded below by $-K$. Then $\big(C(\pi/2) \setminus \{o\}, d_t\big)$ would be a (non-complete) Alexandrov space of curvature bounded below by $-K$ (note that the expansion of R, A guarantees that no minimizing geodesic passes through the apex). But then also the completion $\big(C(\pi/2), d_t\big)$ would be an Alexandrov space (for instance because the midpoint comparison condition for triangles passes to the limit if one point approaches the apex), which is not true as we saw above.

4. Smoothing the Heisenberg group

4.1. **Heisenberg group.** Most of the considerations in this section can be generalized to the higher-dimensional Heisenberg groups, but for simplicity we consider only the first Heisenberg group $\mathbb{H}$. This Lie group can be represented by $\mathbb{H} = \mathbb{C} \times \mathbb{R}$ with the multiplicative structure

$$(z, u) \cdot (z', u') = \left(z + z', u + u' - \frac{1}{2}\mathrm{Im}(z\bar{z}') \right) ,$$

where Im is the imaginary part of a complex number.
A basis for the Lie algebra is given by the left-invariant vector fields

$$\mathbf{X} = \partial_x - \frac{y}{2}\partial_u \ , \quad \mathbf{Y} = \partial_y + \frac{x}{2}\partial_u \ , \quad \mathbf{U} = \partial_u \ ,$$

and the relation $[\mathbf{X}, \mathbf{Y}] = \mathbf{U}$. We will also consider the right-invariant vector fields

$$\hat{\mathbf{X}} = \partial_x + \frac{y}{2}\partial_u \ , \quad \hat{\mathbf{Y}} = \partial_y - \frac{x}{2}\partial_u \ , \quad \hat{\mathbf{U}} = \mathbf{U} \ .$$

The Haar measure associated with the group structure is up to a constant multiple the 3-dimensional Lebesgue measure, denoted by $\mathcal{L}$, it is both left- and right-invariant.

4.2. **Riemannian and sub-Riemannian distances.** The Heisenberg group carries a sub-Riemannian structure given by the pseudo-norm

$$\|a\mathbf{X}+b\mathbf{Y}+c\mathbf{U}\|_{\mathrm{cc}}^2=\begin{cases}a^2+b^2 & \text{if } c=0,\\ +\infty & \text{otherwise.}\end{cases}$$

The Carnot–Carathéodory distance d_{cc} is obtained by minimizing the sub-Riemannian length of curves connecting two points. More precisely, given $p,q\in\mathbb{H}$ we have

$$d_{\mathrm{cc}}(p,q)=\inf\int_0^T\|\dot\gamma_s\|_{\mathrm{cc}}\,\mathrm{d}s\ ,$$

where the infimum is taken e.g. over all absolutely continuous curves $(\gamma_s)_{s\in[0,T]}$ with respect to the Euclidean distance such that $\gamma_0=p,\gamma_T=q$. Note that the sub-Riemannian length of γ is only finite if γ is horizontal, i.e. for a.e. s the tangent vector $\dot\gamma_s$ is contained in the horizontal sub-bundle

$$\mathrm{T}\mathbb{H}=\mathrm{Vect}(\mathbf{X},\mathbf{Y})\ .$$

As a consequence of the so-called Hörmander condition, namely that the horizontal vector fields generate the full tangent space, the distance d_{cc} is finite: any two points of $\mathbb{H}_n$ can be connected by a horizontal curve of finite length and even a minimizing curve can be found. Note that a curve is absolutely continuous with respect to the Carnot–Carathéodory distance if and only if it is absolutely continuous with respect to the Euclidean distance, its tangent vector is horizontal at almost every point and its sub-Riemannian length is finite.

The 3-dimensional Lebesgue measure $\mathcal{L}$ coincides with the 4-dimension Hausdorff measure of the metric space $(\mathbb{H},d_{\mathrm{cc}})$. It has been shown in [19] that the metric measure space $(\mathbb{H},d_{\mathrm{cc}},\mathcal{L})$ does not satisfy the curvature-dimension condition $\mathrm{CD}(K,N)$ for any K,N.

However, the sub-Riemannian pseudo-norm is naturally approximated by a family of Riemannian metrics indexed by $\varepsilon>0$ and defined via

$$\|a\mathbf{X}+b\mathbf{Y}+c\mathbf{U}\|_{\mathrm{Riem}(\varepsilon)}^2=a^2+b^2+(c/\varepsilon)^2\ .$$

We denote the associated Riemannian distance by $d_{\mathrm{Riem}(\varepsilon)}$. The associated Riemannian volume coincides with $\mathcal{L}$ up to a constant. One can check that the best lower bound on the Ricci curvature of $\|\cdot\|_{\mathrm{Riem}(\varepsilon)}$ is $-\frac{1}{2}\varepsilon^{-2}$, see e.g. [5]. We have the following comparison of d_{cc} and $d_{\mathrm{Riem}}:=d_{\mathrm{Riem}(1)}$.

Proposition 4.1 ([20, Lemma1.1])**.** *We have*

$$d_{\mathrm{Riem}}\le d_{\mathrm{cc}}\le d_{\mathrm{Riem}}+4\pi\ .$$

Moreover, there are positive constants c and C such that for any point $p=(z,u)\in\mathbb{H}=\mathbb{C}\times\mathbb{R}$:

$$\max(|z|,c(|z|+|u|^{1/2}))\le d_{\mathrm{cc}}(0_{\mathbb{H}},p)\le C(|z|+|u|^{1/2})\ .$$

4.3. **Isometries.** For every $p \in \mathbb{H}$, we denote by $\tau_p : \mathbb{H} \to \mathbb{H}$ and $\theta_p : \mathbb{H} \to \mathbb{H}$ the left and right translations respectively, i.e.

$$\tau_p(q) = p \cdot q = \theta_q(p) \ .$$

By definition a vector field V is left-invariant if and only if $D\tau_p(V) = V$ for every $p \in \mathbb{H}$. Hence, the left translation τ_p is an isometry for both distances d_{cc} and d_{Riem}. This is false for θ_q unless $q = 0_{\mathbb{H}}$.
Other isometries are

- the rotations $\rho_\alpha : \mathbb{H} \ni (z, u) \mapsto (\mathrm{e}^{i\alpha} z, u)$ defined for $\alpha \in \mathbb{R}$,
- the reflection $\xi : \mathbb{H} \ni (z, u) \mapsto (\bar{z}, -u)$,

and up to the multiplicative constant λ,

- the dilations $\delta_\lambda : \mathbb{H} \ni (z, u) \mapsto (\lambda z, \lambda^2 u)$ where $\lambda > 0$.

One has $D\delta_\lambda(V) = \lambda V$ if and only if V is horizontal. In general one has

$$D\delta_\lambda(a\mathbf{X} + b\mathbf{Y} + c\mathbf{U}) = \lambda(a\mathbf{X} + b\mathbf{Y}) + \lambda^2 c\mathbf{U} \ .$$

Therefore δ_λ is an isometry between $(\mathbb{H}, d_{\mathrm{cc}})$ and $(\mathbb{H}, \lambda^{-1} d_{\mathrm{cc}})$ as well as between $(\mathbb{H}, d_{\mathrm{Riem}(\varepsilon)})$ and $(\mathbb{H}, \lambda^{-1} d_{\mathrm{Riem}(\varepsilon\lambda)})$. Hence, all the Riemannian manifolds $(\mathbb{H}, d_{\mathrm{Riem}(\varepsilon)})_{\varepsilon>0}$ are isometric up to a multiplicative constant, which justifies that we mainly consider $(\mathbb{H}, d_{\mathrm{Riem}})$ corresponding to $\varepsilon = 1$.

4.4. **Wasserstein space over the Heisenberg group and absolutely continuous curves of measures.** Denote by $W_{\mathbb{H}}$ the L^2-Wasserstein distance build from the Carnot–Carathéodory distance d_{cc}. We will recall here the characterization of 2-absolutely continuous curves in $\big(\mathscr{P}_2(\mathbb{H}, d_{\mathrm{cc}}), W_{\mathbb{H}}\big)$ via solutions to the continuity equation.
Denote by $\operatorname{div} \mathbf{V}$ the divergence of a vector field $\mathbf{V}$ on $\mathbb{R}^3$ with respect to the Lebesgue measure. Note that the basis vector fields $\mathbf{X}, \mathbf{Y}, \mathbf{U}$ all have divergence zero and moreover, we have $\operatorname{div}(f\mathbf{X} + g\mathbf{Y} + h\mathbf{V}) = \mathbf{X}f + \mathbf{Y}g + \mathbf{U}h$ for every smooth functions f, g, h. We denote by

$$\nabla_{\mathbb{H}} f = (\mathbf{X}f)\mathbf{X} + (\mathbf{Y}f)\mathbf{Y} \ ,$$

the horizontal gradient of a function f. Then, for any smooth, compactly supported function f and vector field $\mathbf{V}$ we have the integration by parts formula

$$\int_{\mathbb{H}} f \operatorname{div} \mathbf{V} \, \mathrm{d}\mathcal{L} = -\int_{\mathbb{H}} \langle \nabla_{\mathbb{H}} f, \mathbf{V} \rangle_{\mathrm{cc}} \, \mathrm{d}\mathcal{L} \ .$$

Further let us denote by $L^2_{\mathrm{cc}}(\mu)$ the Hilbert space of Borel vector fields $\mathbf{V}$ equipped with the norm

$$\|\mathbf{V}\|^2_{L^2_{\mathrm{cc}}} = \int \|\mathbf{V}\|^2_{\mathrm{cc}} \, \mathrm{d}\mu \ .$$

Note that any $\mathbf{V} \in L^2_{\mathrm{cc}}(\mu)$ must be horizontal μ-a.e. Now, we have the following characterization of absolutely continuous curves.

Proposition 4.2 ([20, Proposition 3.1])**.** *A weakly continuous curve $(\mu_s)_{s\in[0,T]}$ in $\mathscr{P}_2(\mathbb{H})$ is 2-absolutely continuous with respect to $W_{\mathbb{H}}$ if and only if there* *exists a Borel family of vector fields* $\mathbf{V}_s$ *with* $\int_0^T \|\mathbf{V}_s\|^2_{L^2_{\mathrm{cc}}(\mu_s)}\,\mathrm{d}s < \infty$ *such that the continuity equation*

$$\partial_s \mu + \mathrm{div}(\mu_s \mathbf{V}_s) = 0$$

holds in distribution sense. In this case we have $|\dot{\mu}_s| \le \|\mathbf{V}_s\|_{L^2_{\mathrm{cc}}(\mu_s)}$ *for a.e.* s*. Moreover,* $\mathbf{V}_s$ *is uniquely determined for a.e.* s *if we require*

$$\mathbf{V}_s \in T_{\mu_s}\mathscr{P}_2(\mathbb{H}) := \overline{\{\nabla_{\mathbb{H}}\psi \mid \psi \in C_c^\infty(\mathbb{R}^3)\}}^{L^2_{\mathrm{cc}}(\mu_s)}$$

and there holds $|\dot{\mu}_s| = \|\mathbf{V}_s\|_{L^2_{\mathrm{cc}}(\mu_s)}$.

Following verbatim the argument of Lemma 3.4 we obtain a similar statement in the Heisenberg group.

Lemma 4.3. *Let* $\mu = \rho\mathcal{L} \in \mathscr{P}_2(\mathbb{H})$ *with strictly positive density* ρ*. Assume that* μ *satisfies the Poincaré type inequality*

$$\int |f|^2 \,\mathrm{d}\mu \le C \int \|\nabla_{\mathbb{H}} f\|^2_{cc}\,\mathrm{d}\mu\ , \tag{4.1}$$

for all $f \in C_c^\infty(\mathbb{H})$ *with* $\int f\,\mathrm{d}\mu = 0$*. Let* $s \in L^1(\mathbb{H},\mathcal{L})$ *be such that* $\int s\,\mathrm{d}\mathcal{L} = 0$ *and*

$$\|s/\sqrt{\rho}\|^2_{L^2} \ = \ \int \frac{s^2}{\rho}\,\mathrm{d}\mathcal{L} \ < \ \infty\ .$$

Then there exists a unique horizontal vector field $V \in T_\mu\mathscr{P}_2(\mathbb{H})$ *such that the equation*

$$s + \mathrm{div}(\mu V) = 0$$

holds in distribution sense. Moreover, we have

$$\|V\|^2_{L^2_{\mathrm{cc}}(\mu)} \ \le \ C\|s/\sqrt{\rho}\|^2_{L^2}\ . \tag{4.2}$$

4.5. **Heat kernel.** Another important consequence of the Hörmander condition is the hypoellipticity of the operators $\Delta_{\mathrm{cc}} = \mathbf{X}^2 + \mathbf{Y}^2$ and $\Delta_{\mathrm{cc}} - \partial_t$, which in particular means that distributional solutions $\rho : (0,\infty)\times\mathbb{H} \to \mathbb{R}$ of the heat equation

$$\partial_t \rho = \Delta_{cc}\rho,$$

are smooth. Note that the heat equation is left-invariant. As shown by Gaveau [16], the unique distributional solution $\mu_t = \rho_t\mathcal{L}$ with initial condition $\mu_0 \in \mathscr{P}_2(\mathbb{H})$ is given via convolution with a fundamental solution $\mathfrak{h}_t$:

$$\rho_t(p) = \int \mathfrak{h}_t(q^{-1}p)\,\mathrm{d}\mu_0(q)\ ,$$

where $\mathfrak{h}_t$ is given explicitly by

$$\mathfrak{h}_t(z,u) = \frac{2}{(4\pi t)^2}\int_{\mathbb{R}} \exp\left(\frac{\lambda}{t}\left(iu - \frac{|z|^2}{4}\coth\lambda\right)\right)\frac{\lambda}{\sinh\lambda}\ \mathrm{d}\lambda.$$

In fact $\mathfrak{h}_t$ is the density of $X_t = (B_{2t}, L_{2t})$ where the process $(B_t)_{t\ge 0}$ is a planar Brownian motion $B = B^1 + iB^2$ and $L_t = \frac{1}{2}\int_0^t (B_s^1\,\mathrm{d}B_s^2 - B_s^2\,\mathrm{d}B_s^1)$ is the Lévy area. Hence $\mathfrak{h}_t$ is a strictly positive probability density with respect to $\mathcal{L}$ for all t. Moreover, $\mathfrak{h}_t\mathcal{L} \in \mathscr{P}_2(\mathbb{H})$.

We will need the following estimates [5, (14) and proof of Thm. 3.1].

$$\int \big((\mathbf{X}\log\mathfrak{h}_t)^2+(\mathbf{Y}\log\mathfrak{h}_t)^2\big)\mathfrak{h}_t\,\mathrm{d}\mathcal{L}=\frac{2}{t}\;,\qquad \int(\mathbf{U}\log\mathfrak{h}_t)^2\mathfrak{h}_t\,\mathrm{d}\mathcal{L}<\infty\;. \tag{4.3}$$

The same estimates hold for the right-invariant vector fields $\hat{\mathbf{X}},\hat{\mathbf{Y}},\hat{\mathbf{U}}$. Note also the scaling relation

$$\mathfrak{h}_t(z,u)=\frac{1}{t^2}\mathfrak{h}_1(z/\sqrt{t},u/t)\;. \tag{4.4}$$

Given $t\geq 0$ and $q\in\mathbb{H}$ we define the measure $\nu_q^t\in\mathscr{P}_2(\mathbb{H})$ via

$$\nu_q^t=\begin{cases}\delta_q & \text{if } t=0\;,\\ (\tau_q)_\#(\mathfrak{h}_t\mathcal{L})=\mathfrak{h}_t(q^{-1}p)\mathcal{L}(\mathrm{d}p) & \text{otherwise,}\end{cases}$$

and call it the heat kernel measure centered at q.

Lemma 4.4. *The map $\iota_t:(\mathbb{H},d_{\mathrm{cc}})\ni q\mapsto\nu_q^t\in(\mathscr{P}_2(\mathbb{H}),W_{\mathbb{H}})$ is injective and Lipschitz. Moreover, $W_{\mathbb{H}}(\nu_p^t,\nu_q^t)$ tends to infinity as $d_{\mathrm{cc}}(p,q)$ goes to infinity.*

Before we go to the proof, let us stress that the isometries of $(\mathbb{H},d_{\mathrm{cc}})$ introduced in paragraph 4.3 give rise to isometries of $(\mathscr{P}_2(\mathbb{H}),W_{\mathbb{H}})$ via pushforward. In particular, translations of measures $(\tau_p)_\#$ are isometries.

Proof. We first check injectivity. From the probabilistic interpretation of $\mathfrak{h}_t$ one sees that the projection $P^{\mathbb{C}}:\mathbb{H}\to\mathbb{C}$ transports $\nu_{(z,u)}^t$ to the 2-dimensional Gaussian measure centered at z with covariance $2t\cdot\mathrm{Id}$. Therefore, we have $\nu_{(z,u)}^t\neq\nu_{(z',u')}^t$ provided $z\neq z'$. Further, note that $\nu_{(z,u)}^t=(\tau_{(0,u'-u)})_\#\nu_{(z,u')}^t$. Hence, the two measures are distinct provided $u\neq u'$ since $(\tau_{(0,u'-u)})_\#$ is an isometry. This proves injectivity of ι_t.
Lipschitz continuity of ι_t follows from Kuwada's duality between L^q-Wasserstein contraction estimates and L^p-gradient estimates on the heat kernel of the Heisenberg group, established e.g. in [5]. See [22, Prop. 4.1] and the discussion thereafter. The distance $W_{\mathbb{H}}(\delta_p,\nu_p^t)=W_{\mathbb{H}}(\delta_{0_{\mathbb{H}}},\nu_{0_{\mathbb{H}}}^t)$ is independent from p. Hence

$$W_{\mathbb{H}}(\nu_p^t,\nu_q^t)\geq W_{\mathbb{H}}(\delta_p,\delta_q)-2W_{\mathbb{H}}(\delta_{0_{\mathbb{H}}},\nu_{0_{\mathbb{H}}}^t)$$

tends to infinity as $d_{\mathrm{cc}}(p,q)=W_{\mathbb{H}}(\delta_p,\delta_q)$ tends to infinity. □

Finally, we note that for any $t>0$ the measures ν_q^t satisfy the Poincaré type inequality (4.1) with a constant $C=c\cdot t$ for some $c>0$, see [15, Thm. 1.7].

4.6. **Smoothing effect of the transformation.** In view of Lemma 4.4 the map $\iota_t:\mathbb{H}\ni q\mapsto\nu_q^t\in\mathscr{P}_2(\mathbb{H})$ satisfies Assumption 2.1 and the hypotheses of Proposition 2.4 are also satisfied. Thus as outlined in Section 2 we can introduce the new distance

$$\widetilde{d}_t(p,q)=W_{\mathbb{H}}(\nu_p^t,\nu_q^t)\;,$$

as well as the associated length distance

$$d_t(q,p)\;=\;\inf\int_0^T|\dot{p}_s|_t\,\mathrm{d}s=\inf\int_0^T|\dot{\nu}_{p_s}^t|\,\mathrm{d}s\;,$$

where the infimum is taken over absolutely continuous (or equivalently Lipschitz) curves $(p_s)_{s\in[0,T]}$ in $(\mathbb{H},\widetilde{d}_t)$ such that $p_0 = p, p_T = q$ and $|\dot{p}_s|_t$ and $|\dot{\nu}^t_{p_s}|$ denote the metric derivatives with respect to $\widetilde{d}_t$ and $W_\mathbb{H}$ respectively. We have the following scaling relation.

Proposition 4.5. *For $t > 0$ and every $p, q \in \mathbb{H}$ we have*

$$d_t(\delta_{\sqrt{t}}p, \delta_{\sqrt{t}}q) = \sqrt{t}\cdot d_1(p,q)\ . \tag{4.5}$$

In other words, the dilation $\delta_{\sqrt{t}}$ is an isometry from $(\mathbb{H},\sqrt{t}d_1)$ to $(\mathbb{H},d_t)$.

Proof. The measure dilation $(\delta_\lambda)_\#$ dilates the Wasserstein distance $W_\mathbb{H}$ by the factor λ. As a consequence of the scaling relation (4.4) we find that $(\delta_\lambda)_\#\nu^t_q = \nu^{\lambda^2 t}_{\delta_\lambda(q)}$. Thus, we obtain (4.5) with d replaced by $\widetilde{d}$, which then easily passes to the induced length distance. □

We can now state our main theorem about the smoothing effect of the transformation of the distance.

Theorem 4.6. *The distance d_t is induced by a left-invariant Riemannian metric tensor g_t. More precisely, we have $d_t = K\cdot d_{\mathrm{Riem}(\kappa\sqrt{t})}$, where the constant K,κ satisfy $K\geq 2$ and $K/\kappa < \sqrt{2}$.*

The numerical estimates on κ and K will be given in Remarks 4.7 and 4.8. The first remark explains the reason why the convolution procedure allows to recover the forbidden non-horizontal direction. The second remark relates K to the optimal constant in Wasserstein contraction estimates for the heat flow. Together with the convergence results in Proposition 4.9, this result proves Theorem 1.3 claimed in the introduction.

Proof of Theorem 4.6. By Proposition 4.5 it suffices to consider $t = 1$ and we will do so for the moment. We suppress the index $t = 1$ in the notation, setting $d = d_1, \widetilde{d} = \widetilde{d}_1$ and $\nu_q = \nu^1_q$.

Definition of g: For $q \in \mathbb{H}$ and $a, b, c \in \mathbb{R}$ we define:

$$g_q\big((a\mathbf{X}+b\mathbf{Y}+c\mathbf{U})(q)\big) = \|\mathbf{V}^{a,b,c}_q\|^2_{L^2_{\mathrm{cc}}(\nu_q)}\ , \tag{4.6}$$

where $\mathbf{V}^{a,b,c}_q$ is the unique vector field in $T_{\nu_q}\mathscr{P}_2(\mathbb{H})$ solving the continuity equation

$$\partial_s\big|_{s=0}\nu_{q_s} + \mathrm{div}(\nu_q\mathbf{V}^{a,b,c}_q) = 0 \tag{4.7}$$

with a curve $(q_s)_s$ such that $q_0 = q$ and $\dot{q}_s = (a\mathbf{X}+b\mathbf{Y}+c\mathbf{U})(q_s)$. Existence and uniqueness of $\mathbf{V}^{a,b,c}_q$ are ensured by Lemma 4.3. We will show below that g is indeed a metric tensor after polarizing it to a bilinear form. First we have to check that the assumptions of Lemma 4.3 are fulfilled.
Note that the continuity equation can be rewritten as

$$\partial_s\big|_{s=0}\rho_s + \mathrm{div}(\rho_0\mathbf{V}^{a,b,c}_q) = 0\ , \tag{4.8}$$

where $\rho_s(p) = \mathfrak{h}_1(q_s^{-1}p)$ is the density of ν_{q_s}. The derivation of $s\mapsto q_s^{-1}.q_s$ yields

$$\frac{\mathrm{d}}{\mathrm{d}s}q_s^{-1} = -q_s^{-1}\dot{q}_s q_s^{-1} = -(a\hat{\mathbf{X}}+b\hat{\mathbf{Y}}+c\hat{\mathbf{U}})(q_s^{-1})$$

where we use the equalities between left- and right-invariant vector fields at $0_{\mathbb{H}} = q_s^{-1}.q_s$. We obtain

$$\partial_s\big|_{s=0}\rho_s(p) = -\big((a\hat{\mathbf{X}} + b\hat{\mathbf{Y}} + c\hat{\mathbf{U}})\mathfrak{h}_1\big)(q^{-1}p) \ .$$

Thus, by the left invariance of $\mathcal{L}$ and (4.3), we have that

$$\int \frac{|\partial_s|_{s=0}\rho_s|^2}{\rho_0} \, \mathrm{d}\mathcal{L} = \int \frac{|(a\hat{\mathbf{X}} + b\hat{\mathbf{Y}} + c\hat{\mathbf{U}})\mathfrak{h}_1|^2}{\mathfrak{h}_1} \, \mathrm{d}\mathcal{L} < \infty \ .$$

Hence, Lemma 4.3 is indeed applicable.

The metric g is Riemannian: By linearity of (4.7) and uniqueness of the solution, $\mathbf{V}_q^{a,b,c}$ depends linearly on a, b, c. Thus $g_q(\cdot)$ is quadratic and indeed gives rise to a metric tensor after polarization. Note moreover, that $\mathbf{V}_q^{a,b,c} = D\tau_q(\mathbf{V}_{0_{\mathbb{H}}}^{a,b,c})$. This implies that g is left-invariant, i.e.

$$g_q\big((a\mathbf{X} + b\mathbf{Y} + c\mathbf{U})(q)\big) = g_{0_{\mathbb{H}}}\big((a\mathbf{X} + b\mathbf{Y} + c\mathbf{U})(0_{\mathbb{H}})\big) \ .$$

In particular, g_q depends smoothly on q.

Characterization of the Riemannian metrics obtained by convolution: It is readily checked that g is also invariant under rotations ρ_α. Left-invariant and rotation-invariant Riemannian metrics g on $\mathbb{H}$ form a two parameter family indexed by $K, \kappa > 0$ defined by $K = g(\mathbf{X})^{1/2} = g(\mathbf{Y})^{1/2}$ and $K/\kappa = g(\mathbf{U})^{1/2}$. Thus, we must have that $g = K^2\|\cdot\|^2_{\mathrm{Riem}(\kappa)}$.

The distance d coincides with the Riemannian distance: Let us denote by d_g the Riemannian distance induced by g and recall that

$$d_g(p,q) = \inf \int_0^T \sqrt{g_{q_s}(\dot{q}_s)} \, \mathrm{d}s \ ,$$

where the infimum is taken e.g. over all curves $(q_s)_{s\in[0,T]}$ connecting p to q that are Lipschitz with respect to Euclidean distance. To see that d coincides with d_g, it is sufficient to check that a curve $(q_s)_s$ is $\widetilde{d}$-Lipschitz if and only if it is Lipschitz in Riemannian sense and for any such curve we have

$$|\dot{q}_s|^2 = g_{q_s}(\dot{q}_s) \ , \tag{4.9}$$

where the left hand side is the metric derivative with respect to $\widetilde{d}$.
So let $(q_s)_{s\in[0,T]}$ be a curve that is Lipschitz for the Riemannian distance $d_{\mathrm{Riem}(1)}$ and such that $\dot{q}_s = (a_s\mathbf{X} + b_s\mathbf{Y} + c_s\mathbf{U})(q_s)$. Following the reasoning in the first part of the proof, we see that the continuity equation

$$\partial_s\nu_{q_s} + \mathrm{div}(\nu_{q_s}\mathbf{V}_{q_s}^{a_s,b_s,c_s}) = 0$$

holds with $\|\mathbf{V}_{q_s}^{a_s,b_s,c_s}\|^2_{L^2_{\mathrm{cc}}(\nu_{q_s})} = K^2(a_s^2 + b_s^2 + c_s^2/\kappa^2)$. Thus, by the characterization of absolutely continuous curves in $(\mathscr{P}_2(\mathbb{H}), W_{\mathbb{H}})$, Proposition 4.2, and the definition of $\widetilde{d}$, the curve $(q_s)_s$ is $\widetilde{d}$-absolutely continuous with metric derivative $K\sqrt{a_s^2 + b_s^2 + c_s^2/\kappa^2}$, and also $\widetilde{d}$-Lipschitz. Moreover, (4.9) holds by definition of g. Conversely, to see that any $\widetilde{d}$-Lipschitz curve is also Lipschitz for $d_{\mathrm{Riem}(1)}$, we restrict $\widetilde{d}$ to lines $L_p = \{\lambda p \in \mathbb{H} | \, \lambda \in \mathbb{R}\}$ where the Euclidean norm of p is one. On these sets $\widetilde{d}$ is a translation invariant and continuous distance. Therefore $\widetilde{d}(0_{\mathbb{H}}, \lambda p) \geq |\lambda|\widetilde{d}(0_{\mathbb{H}}, p)$ can be proved for every $\lambda \in [-1,1]$, using the triangle inequality and starting with $\lambda \in \mathbb{Q}$.

But on the Euclidean sphere $p \mapsto \widetilde{d}(0_{\mathbb{H}}, p)$ is bounded from below be a positive number. Hence $\widetilde{d}(0_{\mathbb{H}}, \cdot)$ is up to a constant greater than the Euclidean norm on the Euclidean ball. This proves that $\widetilde{d}$-Lipschitz curves are locally $d_{\mathrm{Riem}(1)}$-Lipschitz and hence $d_{\mathrm{Riem}(1)}$-Lipschitz because we are considering curves on segments. □

Remark 4.7 (Estimate on κ). The crucial feature of the regularized distance d_t as opposed to d_{cc} is that also non-horizontal curves can have finite length. This is due to the effect that even when the length of a curve $(q_s)_s$ with respect to d_{cc} is infinite, the length of $(\nu^t_{q_s})_s$ with respect to the Wasserstein distance build from d_{cc} may be finite. Let us make this more explicit for the special curve $q_s = (0, 0, s)$. This curve is not horizontal and has infinite length, actually $d_{\mathrm{cc}}(q_s, q_r) = c \cdot \sqrt{|s - r|}$, where $c = d_{\mathrm{cc}}(0_{\mathbb{H}}, q_1)$, which follows from the behavior of d_{cc} under translations and dilations. However, the curve $\nu^t_{q_s} = \rho^t_s \mathcal{L}$ with $\rho_s(p) = \mathfrak{h}_t(q_s^{-1} p)$ satisfies the continuity equation

$$\partial_s \rho_s = -\mathbf{U} \rho_s = -\left[\mathbf{X}, \mathbf{Y}\right] \rho_s = -\mathbf{X}(\mathbf{Y} \rho_s) + \mathbf{Y}(\mathbf{X} \rho_s) = -\operatorname{div}(\rho_s \mathbf{V}_s)$$

with the horizontal, but not of gradient type vector field $\mathbf{V}_s = (\mathbf{Y} \log \rho_s)\mathbf{X} - (\mathbf{X} \log \rho_s)\mathbf{Y}$. Hence, we have

$$|\dot{\nu}_{q_s}| < \|\mathbf{V}_s\|_{L^2_{\mathrm{cc}}} = \sqrt{\int \frac{(\mathbf{X}\mathfrak{h}_t)^2 + (\mathbf{Y}\mathfrak{h}_t)^2}{\mathfrak{h}_t} \, \mathrm{d}\mathcal{L}} = \sqrt{\frac{2}{t}}$$

by (4.3). Therefore the curve (ν_{q_s}) has indeed finite $W_{\mathbb{H}}$-length and $K/\kappa = g_1(\mathbf{U}) \leq \sqrt{2}$.

Remark 4.8 (Estimate on K). It has been proved by Kuwada [22] that the ratio $\widetilde{d}_t/d_{cc}$ is related to a gradient estimate established by Driver and Melcher [15]. In fact the constant C_2 in this estimate can be dually defined by

$$C_2 = \sup_{p \neq q} d_{cc}(p, q)^{-1} W(\nu^t_p, \nu^t_q),$$

so that it is in particular independent from t. From [15] it is known $C_2 \geq 2$ and a conjecture is $C_2 = 2$, see [5, Remark 3.2]. Let us show $K = C_2$, which gives a new understanding of this constant. We have

$$\widetilde{d}_t/d_{cc} = (d_t/d_{cc}) \times (\widetilde{d}_t/d_t)$$

with $d_t/d_{cc} \leq K$ and $\widetilde{d}_t/d_t \leq 1$. But for $q_s = (s, 0, 0)$ we see that that the quotient of the distances between 0 and q_s tend to K and 1 respectively as s goes to zero. Therefore $g_1(\mathbf{X}) = K = C_2 \geq 2$.

We conclude this section with an observation on the limiting behavior of the convoluted distances as $t \to 0$.

Proposition 4.9. *As $t \to 0$, for all $p, q \in \mathbb{H}$ we have:*

$$\begin{aligned} \widetilde{d}_t(p, q) &\to d_{\mathrm{cc}}(p, q) \ , \\ d_t(p, q) &\to K \cdot d_{\mathrm{cc}}(p, q) \ . \end{aligned}$$

Moreover, the metric spaces $(\mathbb{H}, d_t)$ converge to $(\mathbb{H}, d_{cc})$ in the pointed Gromov–Hausdorff sense.

Proof. The pointwise convergence of $\widetilde{d}_t$ to d_{cc} follows from Proposition 2.4. The pointwise convergence of d_t follows immediately from the explicit formula $d_t = K d_{\mathrm{Riem}(\kappa\sqrt{t})}$ in Theorem 4.6. As the usual approximation of the subRiemannian Heisenberg group holds in the pointed Gromov–Hausdorff sense, the space $(\mathbb{H}, d_t)$ tends to $(\mathbb{H}, K d_{cc})$. But as explained in paragraph 4.3, $(\mathbb{H}, K \cdot d_{\mathrm{cc}})$ is isometric to $(\mathbb{H}, d_{\mathrm{cc}})$ via the dilation δ_K, which implies the last statement. $\square$

References

[1] S. Alexander, V. Kapovitch, and A. Petrunin. An optimal lower curvature bound for convex hypersurfaces in Riemannian manifolds. *Illinois J. Math.*, 52(3):1031–1033, 2008.

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

[3] L. Ambrosio, N. Gigli, and G. Savaré. Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below. *Invent. Math.*, 195(2):289–391, 2014.

[4] L. Ambrosio, N. Gigli, and G. Savaré. Metric measure spaces with Riemannian Ricci curvature bounded from below. *Duke Math. J.*, 163(7):1405–1490, 2014.

[5] D. Bakry, F. Baudoin, M. Bonnefont, and D. Chafai. On gradient bounds for the heat kernel on the Heisenberg group. *Journal of Functional Analysis*, 255:1905–1938, 2008.

[6] L. Bandara, S. Lakzian, and M. Munn. Geometric singularities and a flow tangent to the Ricci flow. *to appear in Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, May 2015.

[7] A. Bellaïche. The tangent space in sub-Riemannian geometry. *J. Math. Sci. (New York)*, 83(4):461–476, 1997. Dynamical systems, 3.

[8] S. Bujalo. Shortest arcs on convex hypersurfaces of a Riemannian space. *Zap. Nauchn. Semin. Leningr. Otd. Mat. Inst. Steklova*, 66:114–132, 1976.

[9] D. Burago, Y. Burago, and S. Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.

[10] D. Burago, S. Ivanov, and B. Kleiner. On the structure of the stable norm of periodic metrics. *Math. Res. Lett.*, 4(6):791–808, 1997.

[11] X. Cao and L. Saloff-Coste. Backward Ricci flow on locally homogeneous 3-manifolds. *Comm. Anal. Geom.*, 17(2):305–325, 2009.

[12] J. Cheeger and T. H. Colding. Lower bounds on Ricci curvature and the almost rigidity of warped products. *Ann. of Math. (2)*, 144(1):189–237, 1996.

[13] B. Chow, P. Lu, and L. Ni. *Hamilton's Ricci flow*, volume 77 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI; Science Press, New York, 2006.

[14] A. Deruelle. Smoothing out positively curved metric cones by Ricci expanders. *ArXiv e-prints*, Feb. 2015.

[15] B. K. Driver and T. Melcher. Hypoelliptic heat kernel inequalities on the Heisenberg group. *Journal of Functional Analysis*, 221(2):340 – 365, 2005.

[16] B. Gaveau. Principe de moindre action, propagation de la chaleur et estimées sous elliptiques sur certains groupes nilpotents. *Acta Math.*, 139(1-2):95–153, 1977.

[17] N. Gigli and C. Mantegazza. A flow tangent to the Ricci flow via heat kernels and mass transport. *Adv. Math.*, 250:74–104, 2014.

[18] M. Gromov. *Metric structures for Riemannian and non-Riemannian spaces*. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, english edition, 2007. Based on the 1981 French original, With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates.

[19] N. Juillet. Geometric inequalities and generalized Ricci bounds in the Heisenberg group. *Int. Math. Res. Not. IMRN*, 2009(13):2347–2373, 2009.

[20] N. Juillet. Diffusion by optimal transport in Heisenberg groups. *Calc. Var. Partial Differential Equations*, 50(3-4):693–721, 2014.
[21] C. Ketterer. Cones over metric measure spaces and the maximal diameter theorem. *J. Math. Pures Appl. (9)*, 103(5):1228–1275, 2015.
[22] K. Kuwada. Duality on gradient estimates and Wasserstein controls. *J. Funct. Anal.*, 258(11):3758–3774, 2010.
[23] J. Lott and C. Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. Math. (2)*, 169(3):903–991, 2009.
[24] R. J. McCann and P. M. Topping. Ricci flow, entropy and optimal transportation. *Amer. J. Math.*, 132(3):711–730, 2010.
[25] A. D. Milka. Shortest lines on convex surfaces. *Dokl. Akad. Nauk SSSR*, 248(1):34–36, 1979.
[26] S.-I. Ohta and K.-T. Sturm. Heat flow on Finsler manifolds. *Comm. Pure Appl. Math.*, 62(10):1386–1433, 2009.
[27] S.-i. Ohta and K.-T. Sturm. Non-contraction of heat flow on Minkowski spaces. *Arch. Ration. Mech. Anal.*, 204(3):917–944, 2012.
[28] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.
[29] F. Schulze and M. Simon. Expanding solitons with non-negative curvature operator coming out of cones. *Math. Z.*, 275(1-2):625–639, 2013.
[30] K.-T. Sturm. On the geometry of metric measure spaces. I and II. *Acta Math.*, 196(1):65–177, 2006.
[31] K.-T. Sturm. Super-Ricci flows for metric measure spaces I. *arXiv:1603.02193*, 2016.
[32] A. Takatsu. Wasserstein geometry of Gaussian measures. *Osaka J. Math.*, 48(4):1005–1026, 2011.
[33] C. Villani. *Optimal transport, Old and new*, volume 338 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 2009.

University of Bonn, Institute for Applied Mathematics, Endenicher Allee 60, 53115 Bonn, Germany
E-mail address: erbar@iam.uni-bonn.de

Institut de Recherche Mathématique Avancée, UMR 7501, Université de Strasbourg et CNRS, 7 rue René Descartes, 67 000 Strasbourg, France
E-mail address: nicolas.juillet@math.unistra.fr