



HAL
open science

DIFFUSION BY OPTIMAL TRANSPORT IN HEISENBERG GROUPS

Nicolas Juillet

► **To cite this version:**

Nicolas Juillet. DIFFUSION BY OPTIMAL TRANSPORT IN HEISENBERG GROUPS. 2009. hal-00372216v1

HAL Id: hal-00372216

<https://hal.science/hal-00372216v1>

Preprint submitted on 31 Mar 2009 (v1), last revised 13 Apr 2018 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DIFFUSION BY OPTIMAL TRANSPORT IN THE HEISENBERG GROUP

NICOLAS JUILLET

ABSTRACT. We prove that the hypoelliptic diffusion of the Heisenberg group \mathbb{H}_n describes in the space of probability measures over \mathbb{H}_n a gradient flow of the Boltzmann entropy Ent_∞ in the sense of optimal transport. We prove that conversely any gradient flow of Ent_∞ that satisfy a further derivability assumption satisfy the hypoelliptic heat equation.

INTRODUCTION

The Heisenberg group \mathbb{H}_n is apparently one of the meeting point of various mathematical and scientific domains. This can be partially explained by the fact that the first Heisenberg group appears as the less elaborated Lie group after the Euclidean spaces. It includes only a few non-commutativity and its Lie algebra presents a unique non-trivial relation (that can be interpreted in quantic physic as the uncertainty principle). That is a reason why it is occurring so much in different domains. The subRiemannian distance, called for \mathbb{H}_n Carnot-Carathéodory distance d_c provides an exotic structure that still allows elementary computations. It makes of \mathbb{H}_n one of the spaces aimed by some theory on metric spaces (see e.g. [12] for geometric measure theory, [19] for conformal geometry, [6] for embedding problems). This paper is devoted to a metric approach of the hypoelliptic diffusion in the Heisenberg group. We will consider in the main theorems (Theorem 0.1 and Theorem 0.2) some curves of $\mathcal{P}_2(\mathbb{H}_n)$ — a space of probability measures called Wasserstein space —that are defined as the gradient flow of the Boltzmann entropy with respect to the Wasserstein metric structure on $\mathcal{P}(\mathbb{H}_n)$. We prove that these curves correspond to the solutions of the hypoelliptic heat equation provided by $\Delta_{\mathbb{H}}$, the subRiemannian “sum of square” operator, the so-called Kohn operator. The hypoellipticity of this operator is the consequence of a famous theorem by Hörmander [14] about “sum of square” operators. Roughly speaking we give a metric way to characterize the classical diffusion of the Heisenberg group by using a gradient flow in the Wasserstein space $\mathcal{P}_2(\mathbb{H}_n)$.

We say now a little more about the origin of this approach. The breakthrough on this topic are the seminal papers of Otto ([15, 24] the first one with Jordan and Kinderlehrer) where the Wasserstein space $\mathcal{P}_2(\mathbb{R}^n)$ is considered for the first time formally as an infinite dimensional Riemannian manifold. This approach is sometimes called “Otto calculus” as in [28]. Otto and his coauthors realized that the solutions of the heat equation are densities of measures describing a special curve on $\mathcal{P}_2(\mathbb{R}^n)$ and justify it at the formal level. The Boltzmann entropy (with

respect to \mathcal{L}), defined by

$$(1) \quad \text{Ent}_\infty(\mu) = \begin{cases} \int \rho \ln \rho d\mathcal{L} & \text{if } \mu = \rho\mathcal{L} \\ +\infty & \text{if } \mu \text{ is not absolutely continuous,} \end{cases}$$

can be regarded as a function on the formal manifold $\mathcal{P}_2(\mathbb{R}^n)$ and the diffusion curve moves with a speed and direction determined by the gradient of this function (the vector field $\frac{\nabla \rho}{\rho}$). This discovery initiated the study of the gradient flow of different functionals in the Wasserstein spaces starting with the relative Rényi entropy recovering the porous medium equation [24]. People continued this approach in various spaces (X, d) , following various definitions of the gradient flow on $\mathcal{P}_2(X)$, sometimes with numerical aspirations [1, 7, 9, 18, 23, 28, 26]. Nowadays, the most documented book on this subject is probably the book by Ambrosio, Gigli and Savaré [1] where this theory is developed in fine analysis for Hilbert spaces. An interesting topic is to study the gradient flows of the entropy in $\mathcal{P}_2(X)$ (relatively to a reference measure \mathcal{L}_X) considering it as the definition for the heat equation in non-smooth settings. Our results show that this point of view is valuable in (\mathbb{H}_n, d_c) , a metric space where it is quite unequivocal what the “heat diffusion” shall be.

With the “Otto calculus” on $\mathcal{P}_2(\mathbb{H}_n)$ our results are formally true. Actually the gradient of $\text{Ent}_\infty(\rho\mathcal{L})$ is formally the vector field $\nabla_{\mathbb{H}}\rho_s/\rho_s$ and this observation leads to a weak form of the hypoelliptic heat equation. The difficulty lies in the execution : for any $K \in \mathbb{R}$ the Boltzmann entropy is not K -convex along the geodesics of $\mathcal{P}_2(\mathbb{H}_n)$ [16]. This property on $\mathcal{P}_2(M)$ is equivalent for Riemannian manifolds M to having a Ricci curvature bounded below by K . In this case studied by Erbar [7], Savaré [26] and Villani [28], a gradient flow of Ent_∞ is solution of the Riemannian heat equation $\Delta_M \rho_s = \frac{\partial}{\partial s} \rho_s$ or conversely solutions of the heat equation are gradient flows. For manifolds without bounds, nothing is known yet because the convexity of Ent_∞ is necessary in the today’s proofs. Our approach of the gradient flow is based on the approximation of our space by the Wasserstein spaces, $\mathcal{P}_2(\mathbb{H}_n^\varepsilon)$, of approximating Riemannian manifolds \mathbb{H}_n^ε . The manifolds \mathbb{H}_n^ε have bounds on the Ricci curvature but these bounds tends to $-\infty$ when ε goes to 0. Hence it is not possible to apply directly a stability theorem about the existence of a gradient flow for a limit of metric spaces with gradient flows (see [26]).

We will present the results as follow. In the first part we will introduce the Heisenberg group, the optimal transport theory and will state the definition we adopt for a gradient flow. In the second part, we will review what we need about the variations of the entropy in the Wasserstein space $\mathcal{P}_2(\mathbb{H}_n^\varepsilon)$ of the approximating manifolds. This part states for Riemannian the results that we hope to have for \mathbb{H}_n and that seem to be very difficult to obtain directly, that is without any reference to the approximating manifolds. In part three and four we study the speed of the curves in $\mathcal{P}_2(\mathbb{H}_n)$ as well as the slope of Ent_∞ in this space. In the fifth part we prove the theorems: (the notations and terminology are given in this paper)

Theorem 0.1. *Let $(\rho_s)_{s \in]0, +\infty[}$ be a solution of the hypoelliptic heat equation*

$$\begin{cases} \Delta_{\mathbb{H}} \rho_s = \partial_s \rho_s \\ \rho_0 d\mathcal{L} = \mu_0 \end{cases}$$

in \mathbb{H}_n where μ_0 has a compact support. The curve $(\mu_s)_{s \geq 0}$ of measures $\rho_s d\mathcal{L} = \mu_s$ is a gradient flow of the entropy Ent_∞ .

Theorem 0.2. *Let $(\mu_s)_{s \in I}$ be a gradient flow of Ent_∞ in $\mathcal{P}_2(\mathbb{H}_n)$. Assume that for almost every $s \in I$, there exists a weak derivative $\mathbf{T}\rho_s$ and a function $w_s^{\mathbf{T}} \in L^2(\mu_s)$ such that $\mathbf{T}\rho_s = w_s^{\mathbf{T}} \rho_s$. Then the density $(\rho_s)_{s \in I}$ satisfies the ‘‘hypoelliptic heat equation’’*

$$\partial_s \rho_s = \Delta_{\mathbb{H}} \rho_s.$$

1. DEFINITIONS

1.1. Heisenberg groups. Let n be a positive integer. The Heisenberg group \mathbb{H}_n is a Lie group that can be represented by $\mathbb{R}^{2n+1} = \mathbb{C}^n \times \mathbb{R}$ with the multiplicative structure

$$(z, t) \cdot (z', t') = (z + z', t + t' - \frac{1}{2} \sum_{k=1}^n \Im(z_k \bar{z}'_k))$$

where z writes $(z_1, \dots, z_n) = (x_1 + \mathbf{i}y_1, \dots, x_n + \mathbf{i}y_n)$ and \Im is the imaginary part of a complex number. A basis for the Lie algebra of left-invariant vector field is given by $(\mathbf{X}_1, \mathbf{Y}_1, \dots, \mathbf{X}_n, \mathbf{Y}_n, \mathbf{T})$ where

$$\mathbf{X}_k = \partial_{x_k} - \frac{y_k}{2} \partial_t \quad \text{and} \quad \mathbf{Y}_k = \partial_{y_k} + \frac{x_k}{2} \partial_t \quad \text{for } k \in \{1, \dots, n\}$$

and $\mathbf{T} = \partial_t$. It is possible to define two metric structures on \mathbb{H}_n . For any $\varepsilon > 0$, we note $(\mathbb{H}_n^\varepsilon, d_\varepsilon)$ the Riemannian manifold defined from $(\mathbf{X}_1, \mathbf{Y}_1, \dots, \mathbf{X}_n, \mathbf{Y}_n, \varepsilon \mathbf{T})$ as orthonormal basis in any point. The metric space (\mathbb{H}_n, d_c) will denote the Heisenberg group with the subRiemannian metric obtained from $(\mathbf{X}_1, \mathbf{Y}_1, \dots, \mathbf{X}_n, \mathbf{Y}_n)$. The functions $\|\cdot\|_\varepsilon$ and $\|\cdot\|_{\mathbb{H}}$ are then defined by

$$\begin{aligned} \left\| \left(\sum_{k=1}^n a_k \mathbf{X}_k + b_k \mathbf{Y}_k \right) + c \varepsilon \mathbf{T} \right\|_\varepsilon &= \sqrt{\left(\sum_{k=1}^n a_k^2 + b_k^2 \right) + c^2} \\ \left\| \sum_{k=1}^n a_k \mathbf{X}_k + b_k \mathbf{Y}_k + c \mathbf{T} \right\|_{\mathbb{H}} &= \begin{cases} \sqrt{\sum_{k=1}^n a_k^2 + b_k^2} & \text{if } c = 0, \\ +\infty & \text{otherwise.} \end{cases} \end{aligned}$$

The Lebesgue measure \mathcal{L}^{2n+1} is up to a constant the Haar measure of the group. Up to a constant it is for \mathbb{H}_n^ε the Riemannian volume and the $2n + 2$ -dimensional Hausdorff measure for \mathbb{H}_n .

We precise that the Carnot-Carathéodory distance d_c , the subRiemannian distance of \mathbb{H}_n is obtained by minimizing the subRiemannian length between two points. This length $l(\gamma)$ is defined for absolutely continuous curve γ of \mathbb{H}_n as the limit of $l_\varepsilon(\gamma)$, the length computed in \mathbb{H}_n^ε when ε tends to 0. Precisely if for almost every s , $\dot{\gamma}(s) = \sum_{k=1}^n (a_k(s) \mathbf{X}_k(\gamma(s)) + b_k(s) \mathbf{Y}_k(\gamma(s)) + c(s) \mathbf{T}(\gamma(s)))$, then

$$(2) \quad l(\gamma) = \int \|\dot{\gamma}(s)\|_{\mathbb{H}} ds = \begin{cases} \int \sqrt{\sum_{k=1}^n (a_k^2 + b_k^2)(s)} ds & \text{if } c = 0 \text{ for almost every } s \text{ (*)}, \\ +\infty & \text{otherwise.} \end{cases}$$

The curves satisfying condition (*) are called horizontal. They are contained in the horizontal subspace, that is $T\mathbb{H}_n := \text{Vect}(\mathbf{X}_1, \mathbf{Y}_1, \dots, \mathbf{X}_n, \mathbf{Y}_n)$. The distance d_c is finite because of the so-called Hörmander condition [14] (see also [22]): the horizontal subspace and the vector fields that it generates using the Lie bracket

span the whole tangent space. Here we have simply $[\mathbf{X}_k, \mathbf{Y}_k] = \mathbf{T}$. From there any two points of \mathbb{H}_n can be connected by an horizontal curve. There are actually geodesics, i.e. curve minimizing the length between two points. As in Riemannian geometry one can also consider an exponential map $\exp^{\mathbb{H}}$ (a variant of the ones in [2, 16]). For $p \in \mathbb{H}_n$, $\vartheta \in \mathbb{R}$ and V an horizontal vector of $T_p\mathbb{H}_n$, the point $\exp_p^{\mathbb{H}}(V, \vartheta)$ is the end of a local geodesic of length $\|V\|_{\mathbb{H}}$ starting in p tangentially to V . The only difference with the exponential map of Riemannian geometry is that there are several such geodesics ; they are parametrized by ϑ .

An other important consequence of the Hörmander condition is the hypoellipticity of the operators $\Delta_{\mathbb{H}} := \sum_{k=1}^n (\mathbf{X}_k^2 + \mathbf{Y}_k^2)$ and $\Delta_{\mathbb{H}} - \partial_s$ which in particular means that weak solutions of the heat equation on $\mathbb{H}_n \times]0, +\infty[$

$$(3) \quad \partial_s \rho_s = \Delta_{\mathbb{H}} \rho_s$$

are also solutions in the classical meaning. In the precise case of the group \mathbb{H}_n , more is known about the solutions of this equation. For an initial condition ρ_0 , Gaveau [13] proved that the solution is given by the convolution with a fundamental solution \mathfrak{h}_s (depending on the dimension n). This \mathfrak{h}_s is non negative, smooth in time s and space p and is given by the formula

$$(4) \quad \mathfrak{h}_s(z, t) = \frac{2}{(4\pi s)^{n+1}} \int_{\mathbb{R}} \exp\left(\frac{\lambda}{s}\left(it - \frac{|z|^2}{4} \coth \lambda\right)\right) \left(\frac{\lambda}{\sinh \lambda}\right)^n d\lambda.$$

Then

$$(5) \quad \rho_s(p) = (\rho_0 *_{\mathbb{H}} \mathfrak{h}_s)(p) = \int_{\mathbb{H}_n} \mathfrak{h}_s^n(q^{-1} \cdot p) \rho_0(q) d\mathcal{L}(q).$$

There is a stochastic equation attached to this diffusion such that it is not difficult to see that if ρ_0 is the density of a probability measures, $\int \rho_s(p) d\mathcal{L}^{2n+1}(p)$ is 1 too. Then $(\rho_s d\mathcal{L})_{s \geq 0}$ is a curve in the space of measure. In this paper we consider the solutions of (3) as a curve in the Wasserstein space, the space of probability measures adapted to the distance of optimal transport.

1.2. Wasserstein spaces. Let (X, d) be a Polish metric space. Then the Wasserstein distance of two Borel probability measures μ_0 and μ_1 is given by

$$(6) \quad W^X(\mu_0, \mu_1) = \inf_{\pi} \sqrt{\int_{X \times X} d^2(p, q) d\pi(p, q)}$$

where the infimum is taken over all couplings π of μ_0 and μ_1 , i.e. over all probability measures π over $X \times X$ such that $p_{\#}\pi = \mu_0$ and $q_{\#}\pi = \mu_1$. Here p (resp. q) stand for the first (resp. second) coordinate map and $\#$ is the push-forward symbol. The infimum in (6) is not always finite. The Wasserstein space $\mathcal{P}_2(X)$ is the space of the measures with finite Wasserstein distance to the Dirac measures. On this space, W^X is finite thanks to the triangle inequality. Also because of this inequality, $\mathcal{P}_2(X)$ can be defined as

$$\mathcal{P}_2(X) = \{\mu \in \mathcal{P}(X) \mid \int d(p, p_0)^2 d\mu(p) < +\infty\}$$

for any $p_0 \in X$.

For any Polish space X , there exists minimizing couplings π in (6). In the case of $\mathcal{P}_2(\mathbb{H}_n^c)$ and $\mathcal{P}_2(\mathbb{H}_n)$, it has been proved in [21] and [2] that if μ_0 is absolutely continuous, (6) is achieved by an unique π and this coupling can be expressed

by $\pi = (\text{Id} \otimes T)_{\#} \mu_0$ for a μ_0 -almost everywhere uniquely defined map T . Moreover these authors prove that in these two geodesic spaces a geodesic between p and $T(p)$ is μ_0 -almost surely unique. We denote it by $(T_s(p))_{s \in [0,1]}$ such that in particular $T_0 = \text{Id}$ and $T_1 = T$. From there usual theorems in optimal transport [1, 28] ensure that there exists a unique geodesic $(\mu_s)_{s \in [0,1]}$ between μ_0 and μ_1 , defined by $\mu_s = (T_s)_{\#} \mu_0$. Still under the assumption that μ_0 is absolutely continuous, the measures μ_s for $s < 1$ are absolutely continuous as well [10, 11].

A related result [28, Corollary 10.44 and Theorem 13.5] states that on a manifold if μ_0 is absolutely continuous, for U_s in a form that we will precise, $\mu_s = (U_s)_{\#} \mu_0$ is absolutely continuous too for s small enough. It is actually sufficient that U_s writes $p \rightarrow \exp_p(s \nabla \varphi(p))$ where \exp_p and ∇ are the usual exponential map and gradient of Riemannian geometry and φ is a smooth function with compact support. This proposition of course apply to \mathbb{H}_n^ε for any ε but there is also a adapted version for \mathbb{H}_n [17, Section 2.2.3] with a quite similar proof. We state this result here ($\nabla_{\mathbb{H}}$ is defined in (8)):

Proposition 1.1. *Let $\psi, \vartheta \in C_c^\infty(\mathbb{R}^{2n+1})$ and μ_0 an absolutely continuous measure on \mathbb{H}_n . Define $U_s(p) = \exp_p^{\mathbb{H}}(s \nabla_{\mathbb{H}} \psi(p), s \vartheta(p))$ and $\mu_s = (U_s)_{\#} \mu_0$. There is a $s_0 > 0$ such that for any $s < s_0$, U_s is injective, smooth and $\text{Jac}(U_s) \neq 0$. It follows that μ_s is absolutely continuous. These results is also true (for an other s_0) if we replace $\nabla_{\mathbb{H}} \psi$ by a smooth horizontal field V with compact support.*

1.3. Absolutely continuous curves. A curve $(\gamma_s)_{s \in I}$ in a metric space (X, d) is said to be absolutely continuous on I if there exists a $m \in L^1(\mathbb{R})$ such that for any $a < b$ in I ,

$$d(\gamma(a), \gamma(b)) \leq \int_a^b m(s) ds.$$

It is proved in [4] that if γ is absolutely continuous, for almost every $s \in I$ the metric derivative

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

exists and

$$(7) \quad l(\gamma) = \int_a^b |\dot{\gamma}_s| ds$$

equals the classical metric length of the curve γ between a and b . In a geodesic space $l(\gamma) \geq d(a, b)$. If $l(\gamma) = d(a, b)$, some reparametrizations of γ are (constant-speed) geodesics. We denote the space of absolutely continuous curves by $AC(X)$. Let $AC_2(X) \subset AC(X)$ be the subspace of absolutely continuous curves such that $|\dot{\gamma}_s|^2$ is locally integrable.

1.4. Gradient flow. In this paper, curves will take values in the Wassertein spaces $\mathcal{P}_2(\mathbb{H}_n)$ and $\mathcal{P}_2(\mathbb{H}_n^\varepsilon)$. We will study the slope of the functional entropy Ent_∞ introduced in (1). Note that $\mathcal{P}_2(\mathbb{H}_n) = \mathcal{P}_2(\mathbb{H}_n^\varepsilon)$ because $d_\varepsilon \leq d_c \leq d_\varepsilon + \pi\varepsilon$ as is proved in [17, Subsection 3.5.2]. The slope of Ent_∞ is defined as

$$\text{Slope}^X(\text{Ent}_\infty)(\mu) = \max \left(0, \limsup_{\nu \rightarrow \mu} \frac{\text{Ent}_\infty(\mu) - \text{Ent}_\infty(\nu)}{W^X(\mu, \nu)} \right).$$

This quantity is positive and quantify how much the entropy can locally decrease. In $\mathcal{P}_2(\mathbb{H}_n)$ we will denote the slope and the Wasserstein distance by Slope and W . In $\mathcal{P}_2(\mathbb{H}_n^\varepsilon)$ we use the notations Slope^ε and W^ε . As $d_c \geq d_\varepsilon$ we have for the

Wasserstein spaces $W \geq W^\varepsilon$ where W^ε . So $\text{Slope}^\varepsilon(\text{Ent}_\infty)(\mu) \geq \text{Slope}(\text{Ent}_\infty)(\mu)$ in every μ of finite entropy.

We will use in this paper a very metric definition for the gradient flow of the entropy. It refers to v) in Theorem 5.3 of [3] or to ii) of Proposition 23.2 of [28]. A curve of the Wasserstein space $(\mu_s)_{s \in I}$ is said to be a gradient flow of Ent_∞ if

- it is an absolutely continuous curve of $\mathcal{P}_2(X)$ and for almost every $s \in I$,

$$|\dot{\mu}_s| = \text{Slope}(\text{Ent}_\infty)(\mu_s),$$

- the function $E(s) = \text{Ent}_\infty(\mu_s)$ is absolutely continuous and for almost every $s \in I$,

$$\dot{E}(s) = -\text{Slope}(\text{Ent}_\infty)(\mu_s) \cdot |\dot{\mu}_s|.$$

This definition has to be understood as the decomposition of a formal “ $\dot{\mu}_s = -\nabla \text{Ent}_\infty(\mu_s)$ ” into the norm equality and the indication of the direction.

Remark 1.2. Of course then $\dot{E}(s) = -|\dot{\mu}_s|^2$ for almost every s , but in the first formulation it is easier to recognize a chain rule derivation where it appears that the curve falls off in the direction of the slope. Nevertheless it follows from this remark that $(\mu_s)_{s \in I}$ is in $AC^2(X)$.

1.5. Functional spaces, tangent spaces. The vector spaces we will introduce now are defined using weak formulations that require test functions. In the sequel the space of test functions will be $\mathcal{C}_c^\infty(\mathbb{R}^{2n+1})$, the space of smooth functions of \mathbb{R}^{2n+1} with a compact support. For any $k \in \{1, \dots, n\}$ and f a function of $L^1_{\text{loc}}(\mathbb{H}_n)$ (a \mathcal{L} -locally integrable function on \mathbb{H}_n), we will note $\mathbf{X}_k f$, $\mathbf{Y}_k f$ and $\mathbf{T}f$ the distributional derivatives of f if these distributions are functions of $L^1_{\text{loc}}(\mathbb{H}_n)$. We explicitate the defining relation for \mathbf{T} :

$$\forall \psi \in \mathcal{C}_c^\infty(\mathbb{R}^{2n+1}), \int (\mathbf{T}f)\psi d\mathcal{L} + \int f(\mathbf{T}\psi) d\mathcal{L} = 0$$

Note that these notations are coherent with the the usual ones when f is smooth. Actually in this case, $f\psi$ is a test function. Then (this time for \mathbf{X}_k)

$$\begin{aligned} \int \mathbf{X}_k(f\psi) d\mathcal{L} &= \int (\partial_{x_k} - \frac{1}{2}y\partial_t)(f\psi) d\mathcal{L} \\ &= \int \partial_{x_k}(f\psi) d\mathcal{L} - \frac{1}{2} \int \left(\int (y_k \partial_t)(f\psi) dt \right) dx dy = 0 \end{aligned}$$

But the first integral is also $\int (\mathbf{X}_k f)\psi d\mathcal{L} + \int f(\mathbf{X}_k \psi) d\mathcal{L}$ as we want.

We say that f has a gradient $\nabla_{\mathbb{H}} f$ if it has a weak derivative Uf for any $U \in \{\mathbf{X}_1, \mathbf{Y}_1, \dots, \mathbf{X}_n, \mathbf{Y}_n\}$. Then we can define the \mathbb{H} -gradient $\nabla_{\mathbb{H}} f$ by

$$(8) \quad \nabla_{\mathbb{H}} f := \sum_{k=1}^n (\mathbf{X}_k f \mathbf{X}_k + \mathbf{Y}_k f \mathbf{Y}_k).$$

If moreover f has a weak \mathbf{T} -derivative, we can define the ε -gradient $\nabla_\varepsilon f$,

$$\nabla_\varepsilon f := \nabla_{\mathbb{H}} f + (\varepsilon \mathbf{T}) f (\varepsilon \mathbf{T}).$$

We note $W_{\text{loc}}^{1,1}(\mathbb{H}_n)$ the space of locally integrable functions f with a \mathbb{H} -gradient $\nabla_{\mathbb{H}} f$. Let $W_{\text{loc}}^{1,1}(\mathbb{H}_n^\varepsilon) \subset W_{\text{loc}}^{1,1}(\mathbb{H}_n)$ be the subspace of the functions f with a weak derivative $\mathbf{T}f$.

For a given absolutely continuous measure μ , we define $\text{Tan}_{\mathbb{H}}(\mu)$, the tangent space at $\mu \in \mathcal{P}_2(\mathbb{H}_n)$ as the smallest Hilbert space that contains the vector fields $\nabla_{\mathbb{H}}\psi$ for a test function ψ . Precisely

$$\text{Tan}_{\mathbb{H}}(\mu) = \overline{\{\nabla_{\mathbb{H}}\psi \mid \psi \in C_c^\infty(\mathbb{R}^{2n+1})\}}^{L_{\mathbb{H}}^2(\mu)}$$

with $L_{\mathbb{H}}^2(\mu) = \{\xi \in \Gamma(T\mathbb{H}_n) \mid \int_{\mathbb{H}_n} \|\xi\|_{\mathbb{H}}^2 < +\infty\}$. Similarly at $\mu \in \mathcal{P}_2(\mathbb{H}_n^\varepsilon)$, the tangent space $\text{Tan}_\varepsilon(\mu)$ is defined by

$$\text{Tan}_\varepsilon(\mu) = \overline{\{\nabla_\varepsilon\psi \mid \psi \in C_c^\infty(\mathbb{R}^{2n+1})\}}^{L_\varepsilon^2(\mu)}$$

where $L_\varepsilon^2(\mu) = \{\xi \in \Gamma(T\mathbb{H}_n^\varepsilon) \mid \int_{\mathbb{H}_n^\varepsilon} \|\xi\|_\varepsilon^2 < +\infty\}$.

2. SOME RESULTS CONCERNING THE APPROXIMATING MANIFOLDS AND THEIR WASSERSTEIN SPACES

We state here, for the approximating manifolds \mathbb{H}_n^ε a proposition that Erbar [8, Proposition 4.3.] proved for Riemannian manifolds. It is the translation of the same statement for Euclidean spaces by Ambrosio and Savaré [3, Theorem 4.16]. The proof of Erbar also strongly relies on Theorem 23.14 of [28] whose proof is long and difficult.

Proposition 2.1. *Let $\varepsilon > 0$ and $d\mu = \rho d\mathcal{L}$ a probability measure. Then the following statements are equivalent:*

- (i) $\text{Slope}^\varepsilon(\text{Ent}_\infty)(\mu) < +\infty$
- (ii) $\rho \in W_{loc}^{1,1}(\mathbb{H}_n^\varepsilon)$, $\nabla_\varepsilon \rho = \rho w^\varepsilon$ \mathcal{L} -almost surely for some $w^\varepsilon \in L_\varepsilon^2(\mu)$.

In this case $w^\varepsilon \in \text{Tan}_\varepsilon(\mu)$ and $\text{Slope}(\text{Ent}_\infty)(\mu) = \|w^\varepsilon\|_{L_\varepsilon^2(\mu)}$.

Remark 2.2. The vector field w^ε is sometime simply written “ $\nabla_\varepsilon \rho / \rho$ ”. Actually as $\mu = \rho \mathcal{L}$, the function ρ is μ -almost surely non-zero such that $\nabla_\varepsilon \rho / \rho$ makes sense in $L_\varepsilon^2(\mu)$. It shows that there is a unique possible $w^\varepsilon \in L_\varepsilon^2(\mu)$. However, in order to avoid bad interpretations of this quotient we will rather use w^ε .

Remark 2.3. If statement (ii) is true for some ε , it also holds for other $\varepsilon' > 0$. It follows that if the slope is finite in $\mathcal{P}_2(\mathbb{H}_n^\varepsilon)$, it is also finite in the other Wasserstein spaces $\mathcal{P}_2(\mathbb{H}_n^{\varepsilon'})$ even for $\varepsilon' > \varepsilon$.

We state now a mixing of a propositions by Villani [28, Theorem 23.14] and by Erbar [8]. In fact the second point (10) is not in [28, Theorem 23.14]. It is obtained from it, approximating v_s^ε , by the gradient of c -convex functions and using [8, Lemma 2.7]. The first point is a mixing of Proposition 2.5 of [8] and [28, Theorem 13.8]. It is proved exactly in the same way as we will prove Proposition 3.1. Finally (11) corresponds to [28, Remark 23.16]. It seems to be very hard to prove Theorem 23.14 of [28] for spaces without assumption on the Ricci curvature. Nevertheless in this paper we need this type of result for \mathbb{H}_n . Our strategy is to state weaker relations for \mathbb{H}_n^ε in Proposition 2.4 and obtain the corresponding relations for \mathbb{H}_n using the approximation.

Proposition 2.4. *Let $\varepsilon > 0$ and $(\mu_s)_{s \in I}$ an AC^2 -curve of $\mathcal{P}_2(\mathbb{H}_n^\varepsilon)$. Assume that for almost every $s \in I$, the slope $\text{Slope}^\varepsilon(\text{Ent}_\infty)(\mu_s)$ is finite. Let $w_s^\varepsilon \in L_\varepsilon^2(\mu_s)$ be the corresponding vector with respect to Proposition (2.1). Then there is a subset $I'_\varepsilon \subset I$ of full-measure such that for any $s' \in I'_\varepsilon$ there is a vector field $v_{s'}^\varepsilon \in \text{Tan}_\varepsilon(\mu_{s'})$ satisfying the two following statements*

- For every test function $\psi \in \mathcal{C}_c^\infty$,

$$(9) \quad \frac{\partial}{\partial s} \mu_s(\psi)(s') = \int \langle v_{s'}^\varepsilon \mid \nabla_\varepsilon \psi \rangle_\varepsilon d\mu_{s'},$$

- the entropy evolves in a way such that

$$(10) \quad \text{Ent}_\infty(\mu_s) \geq \text{Ent}_\infty(\mu_{s'}) + \int \langle (s - s') v_{s'}^\varepsilon \mid w_{s'}^\varepsilon \rangle_\varepsilon d\mu_{s'} + o(|s - s'|),$$

when s goes to s' .

Moreover for any μ and ν in $\mathcal{P}_2(\mathbb{H}_n^\varepsilon)$ such that $\text{Slope}^\varepsilon(\text{Ent}_\infty)(\mu) < +\infty$,

$$(11) \quad \text{Ent}_\infty(\nu) \geq \text{Ent}_\infty(\mu) - W^\varepsilon(\mu, \nu) \sqrt{\int \|w^\varepsilon\|_\varepsilon^2 d\mu} - \frac{1}{2 \cdot 2\varepsilon^2} W^\varepsilon(\mu, \nu)^2.$$

Remark 2.5. Here $d\mu_s = \rho_s d\mathcal{L}$. In the proposition it is possible to change \mathcal{L} in vol_ε , the Riemannian volume of \mathbb{H}_1^ε and ρ_s in ρ_s^ε , the density with respect to this volume. Then $v_{s'}^\varepsilon$ must be multiplied by the same constant. This form is closer to the standard statement for the Riemannian manifolds.

Remark 2.6. For Riemannian manifolds M for which K is a lower bound of $\text{Ric}(p)$ for any $p \in M$, the bound $\frac{-1}{2 \cdot 2\varepsilon^2}$ in (11) should be replaced by $\frac{1}{2}K$.

In Proposition 3.1 we will prove a similar statement to (9) for the “true” Heisenberg group \mathbb{H}_n . In Section 4 we will let ε go to zero in (11) and get a result on the slope of the entropy in $\mathcal{P}_2(\mathbb{H}_n)$, the Wasserstein space of the “true” Heisenberg group. Inequality (10) will also be interpreted in the context of $\mathcal{P}_2(\mathbb{H}_n)$.

3. SPEED AND VELOCITY

Equality (9) in Proposition 2.4 shows that for the AC^2 -curves $(\mu_s)_{s \in I}$ of the Wasserstein spaces $\mathcal{P}_2(\mathbb{H}_n^\varepsilon)$ (where it is known that there is a metric speed $|\dot{\mu}_s|$) it is possible to define a velocity (speed and direction) thanks to a vector field v_s^ε . We show that there is a similar velocity for $\mathcal{P}_2(\mathbb{H}_n)$, the Wasserstein space of the Heisenberg group. The next proposition will appear in the proof of Theorem 0.2 where we will apply it to gradient flows of Ent_∞ .

Proposition 3.1. *Let $(\mu_s)_{s \in I}$ be an AC^2 -curve of $(\mathcal{P}_2(\mathbb{H}_n), W)$. Then there is a subset $I' \subset I$ of full measure such that in any $s' \in I'$ there is a vector field $v_{s'} \in \text{Tan}_{\mathbb{H}}(\mu_{s'})$ so that*

$$(12) \quad \partial_{s'} \mu_{s'} + \text{div}_{\mathbb{H}}(v_{s'} \mu_{s'}) = 0$$

in a the weak sense. It means that for every function $\psi \in \mathcal{C}_c^\infty(\mathbb{R}^{2n+1})$,

$$\left(\frac{\partial}{\partial s} \int \psi d\mu_s \right) |_{s'} = \int \langle \nabla_{\mathbb{H}} \psi \mid v_{s'} \rangle_{\mathbb{H}} d\mu_{s'}.$$

Moreover, $\|v_{s'}\|_{L_{\mathbb{H}}^2(\mu_{s'})} \leq |\dot{\mu}_{s'}|$ for any $s' \in I'$.

Proof. Let $\psi \in \mathcal{C}_c^\infty(\mathbb{R}^{2n+1})$. We assume also that ψ it is a 1-Lipschitz function of (\mathbb{H}_n, d_c) . Then for $s < t$ in I , by using an optimal transport plan with respect to the 1-Wasserstein cost (the L^1 version of (6) raising the distance W_1), we get

$$\left| \int_{\mathbb{H}} \psi d\mu_s - \int_{\mathbb{H}} \psi d\mu_t \right| \leq W_1(\mu_s, \mu_t) \leq W(\mu_s, \mu_t).$$

Then $\zeta^\psi(s) := \int_{\mathbb{H}} \psi d\mu_s$ is an absolutely continuous function. It is derivable for almost every $s \in I$. Let $\pi_{s,t}$ be an optimal transport plan between μ_s and μ_t . We define now Ψ by

$$\Psi(x, y) = \begin{cases} \frac{|\psi(p) - \psi(q)|}{d_c(p, q)} & \text{if } p \neq q \\ \|\nabla_{\mathbb{H}} \psi(p)\|_{\mathbb{H}} & \text{else.} \end{cases}$$

Then Ψ is bounded above by 1 and it is upper semi-continuous. Let $s' \in I$ such that ζ^ψ is differentiable and the metric derivative $|\dot{\mu}_{s'}|$ exists. We get

$$\begin{aligned} \left| \frac{d}{ds} \int \psi d\mu_s \right| (s') &\leq \liminf_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \left| \int \psi d\mu_{s'} - \int \psi d\mu_{s'+\varepsilon} \right| \\ &\leq \liminf_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int |\psi(q) - \psi(p)| d\pi_{s', s'+\varepsilon}(p, q) \\ &\leq \left(\liminf_{\varepsilon \downarrow 0} \sqrt{\int \Psi(p, q)^2 d\pi_{s', s'+\varepsilon}(p, q)} \frac{\sqrt{\int d(p, q)^2 d\pi_{s', s'+\varepsilon}}}{\varepsilon} \right) \\ &\leq \left(\liminf_{\varepsilon \downarrow 0} \sqrt{\int \Psi(p, q)^2 d\pi_{s', s'+\varepsilon}(x, y)} \frac{W(\mu_{s'}, \mu_{s'+\varepsilon})}{\varepsilon} \right) \\ &\leq \left(|\dot{\mu}_{s'}| \liminf_{\varepsilon \downarrow 0} \sqrt{\int \Psi(p, q)^2 d\pi_{s', s'+\varepsilon}(p, q)} \right). \end{aligned}$$

Since Ψ is upper semi-continuous and $\pi_{s', s'+\varepsilon}$ weakly converges to $(\text{Id} \otimes \text{Id})_{\#} \mu_{s'}$ (see [28, Theorem 5.19]) when $\varepsilon \downarrow 0$, we get

$$\begin{aligned} \left| \frac{d}{ds} \int \psi d\mu_s \right| &\leq |\dot{\mu}_s| \sqrt{\int |\Psi(x, x)|^2 d\mu_s(x)} \\ (13) \qquad \qquad \qquad &= |\dot{\mu}_s| \sqrt{\int \|\nabla_{\mathbb{H}} \psi(x)\|_{\mathbb{H}}^2 d\mu_s(x)}. \end{aligned}$$

This is the key estimate of the proof. We have assumed that ψ is 1-Lipschitz. The estimate also holds without this assumption.

We already know that for any $\psi \in \mathcal{C}_c^\infty(\mathbb{R}^{2n+1})$, the function ζ^ψ is differentiable at almost every $s \in I$. We will now prove that for almost every $s \in I$, every function ζ^ψ is differentiable. We use the fact that there is sequence $(\psi_k)_{k \in \mathbb{N}}$ of test functions such that $\nabla_{\mathbb{H}} \psi_k$ is dense in $\text{Tan}_{\mathbb{H}}(\mu_s)$. Moreover, one assume that the sequence $(\nabla_{\mathbb{H}} \psi_k)_{k \in \mathbb{N}}$ is dense in $\{\nabla_{\mathbb{H}} \psi \mid \psi \in \mathcal{C}_c^\infty(\mathbb{R}^{2n+1})\}$ for the norm of $L^\infty(T\mathbb{H}_1)$. The functions $\zeta^{\psi_k} = \int \psi_k d\mu_s$ are countable and derivable in almost every $s \in I$. Thus there is a set $I'_1 \subset I$ with full measure in I such that in each $s' \in I'_1$, the metric derivative $|\dot{\mu}_{s'}|$ exists and the ζ^{ψ_k} are differentiable.

For $s' \in I'_1$ thanks to (13) we define the bounded operator $T_{s'}$ on $\{\nabla_{\mathbb{H}} \psi_k \mid k \in \mathbb{N}\}$ by $T_{s'}(\nabla_{\mathbb{H}} \psi_k) = \frac{d}{ds} \zeta^{\psi_k}(s')$. This set is dense in $\text{Tan}_{\mathbb{H}}(\mu_{s'})$ so that we can extend $T_{s'}$ on $\text{Tan}_{\mathbb{H}}(\mu_{s'})$ and represent it by a vector field $v_{s'} \in \text{Tan}_{\mathbb{H}}(\mu_{s'})$:

$$T_{s'}(w) = \int \langle w | v_{s'} \rangle_{\mathbb{H}} d\mu_{s'}.$$

We show now that for any $\psi \in \mathcal{C}_c^\infty(\mathbb{R}^{2n+1})$,

$$T_{s'}(\nabla_{\mathbb{H}}\psi) = \lim_{\varepsilon \rightarrow 0} \frac{\zeta^\psi(s' + \varepsilon) - \zeta^\psi(s')}{|\varepsilon|}.$$

For every $k \in \mathbb{N}$

$$\begin{aligned} \limsup_{\varepsilon \rightarrow 0} \left\| T_{s'}(\nabla_{\mathbb{H}}\psi) - \frac{\zeta^\psi_{s'+\varepsilon} - \zeta^\psi_{s'}}{\varepsilon} \right\| &\leq \|T_{s'}(\nabla_{\mathbb{H}}\psi) - T_{s'}(\nabla_{\mathbb{H}}\psi_k)\| \\ &+ \left\| T_{s'}(\nabla_{\mathbb{H}}\psi_k) - \frac{\zeta^{\psi_k}_{s'+\varepsilon} - \zeta^{\psi_k}_{s'}}{\varepsilon} \right\| + \left\| \frac{\zeta^{(\psi_k - \psi)}_{s'+\varepsilon} - \zeta^{(\psi_k - \psi)}_{s'}}{\varepsilon} \right\|. \end{aligned}$$

As the curves $\zeta_s^{(\psi_k - \psi)}$ are absolutely continuous, by using the estimate (13) its differentiability set we can estimate the previous expression by

$$\begin{aligned} &\|T_{s'}(\nabla_{\mathbb{H}}\psi) - T_{s'}(\nabla_{\mathbb{H}}\psi_k)\| + \left\| T_{s'}(\nabla_{\mathbb{H}}\psi_k) - \frac{\zeta^{\psi_k}_{s'+\varepsilon} - \zeta^{\psi_k}_{s'}}{\varepsilon} \right\| \\ &+ \frac{1}{\varepsilon} \sqrt{\int_{s'}^{s'+\varepsilon} |\dot{\mu}_s|^2 ds} \sqrt{\int_{s'}^{s'+\varepsilon} \|\nabla_{\mathbb{H}}(\psi_k - \psi)\|_{L^2(\mu_s)}^2 ds}. \end{aligned}$$

For a given k , we let ε go to 0 and we obtain

$$\limsup_{\varepsilon \rightarrow 0} \left\| T_{s'}(\nabla_{\mathbb{H}}\psi) - \frac{\zeta_{s'+\varepsilon} - \zeta_{s'}}{\varepsilon} \right\| \leq 2|\dot{\mu}_{s'}| \cdot \|\nabla_{\mathbb{H}}(\psi_k - \psi)\|_\infty$$

on $\{s' \in I'_1 \mid s' \text{ is a Lebesgue points of } s \rightarrow |\dot{\mu}_s|^2\}$. Using the density of $\{\nabla_{\mathbb{H}}\psi_k \mid k \in \mathbb{N}\}$ in $L^\infty(T\mathbb{H}_1)$, we have the first part of the proposition.

We will now prove $\|v_{s'}\|_{L^2(\mu_{s'})} \leq |\dot{\mu}_{s'}|$. For $s' \in I'$, we consider $T_{s'}(\nabla_{\mathbb{H}}\psi_k)$ where $(\nabla_{\mathbb{H}}\psi'_k)_{k \in \mathbb{N}}$ is a sequence tending to $v_{s'}$ in $\text{Tan}_{\mathbb{H}}(\mu_{s'})$. On the one hand this sequence tends to

$$\lim_{k \rightarrow \infty} \int \langle v_{s'} \mid \nabla_{\mathbb{H}}\psi'_k \rangle_{\mathbb{H}} d\mu_{s'} = \int \|v_{s'}\|_{\mathbb{H}}^2 d\mu_{s'}.$$

On the other hand, from estimate (13) it is smaller than

$$|\dot{\mu}_{s'}| \lim_{k \rightarrow \infty} \sqrt{\int \|\nabla_{\mathbb{H}}\psi'_k\|_{\mathbb{H}}^2 d\mu_{s'}} = |\dot{\mu}_{s'}| \sqrt{\int \|v_{s'}\|_{\mathbb{H}}^2 d\mu_{s'}}.$$

Then $\|v_{s'}\|_{L^2(\mu_{s'})}$ is smaller than $|\dot{\mu}_{s'}|$. □

Remark 3.2. If one carefully read the proof, the set I' in Proposition 3.1 can be chosen as the intersection of the following sets: the set where $|\dot{\mu}_s|$ exists, the differentiation set of the functions ζ^{ψ_k} and the Lebesgue points of $s \rightarrow |\dot{\mu}_s|^2$.

4. SLOPE

After the previous Section we can represent the velocity of $(\mu_s)_{s \in I}$ by a vector field of $\text{Tan}_{\mathbb{H}}(\mu_{s'})$. Proposition 4.1 makes the picture more precise and permits to identify “the gradient of the entropy” as a vector field. For this proof we will not only assume that the slope of the entropy in $\mathcal{P}_2(\mathbb{H}_n)$ is finite but also that the slope of Ent_∞ is finite in $\mathcal{P}_2(\mathbb{H}_n^\varepsilon)$ for some ε . Proposition 4.1 has to be read in relation with Proposition 2.1.

Proposition 4.1. *Let μ be an absolutely continuous probability measure of density ρ and $\varepsilon > 0$. Assume that $\text{Slope}(\text{Ent}_\infty)(\mu) < +\infty$. Then there is an horizontal vector field $w^\mathbb{H} \in L^2_\mathbb{H}(\mu)$ such that \mathcal{L} -almost everywhere $\nabla_\mathbb{H}\rho = \rho w^\mathbb{H}$. Moreover, the two following statements are equivalent*

- (i) $\text{Slope}^\varepsilon(\text{Ent}_\infty)(\mu) < +\infty$,
- (ii) *The weak gradient $\mathbf{T}\rho$ exists and there is $w^\mathbf{T} \in L^2(\mu)$ such that $\mathbf{T}\rho = \rho w^\mathbf{T}$ (\mathcal{L} -almost everywhere).*

If these assumptions hold, there is actually as in Proposition 2.1 a vector field $w^\varepsilon \in \text{Tan}_\varepsilon(\mu)$ such that $\nabla_\varepsilon\rho = \rho w^\varepsilon$ and $\text{Slope}^\varepsilon(\text{Ent}_\infty)(\mu) = \|w^\varepsilon\|_{L^2_\varepsilon(\mu)}$. Moreover, $w^\mathbb{H} \in \text{Tan}_\mathbb{H}(\mu)$, $w^\mathbb{H} + \varepsilon w^\mathbf{T}(\varepsilon\mathbf{T}) = w^\varepsilon$ and

$$\text{Slope}(\text{Ent}_\infty)(\mu) = \|w^\mathbb{H}\|_{L^2_\mathbb{H}(\mu)}.$$

Remark 4.2. As in Remark 2.2, $w^\mathbb{H}$ could be written “ $\nabla_\mathbb{H}\rho/\rho$ ” and $w^\mathbf{T}$ should be understood as “ $\mathbf{T}\rho/\rho$ ”.

Remark 4.3. Contrarily to Proposition 2.1, we are not able to state that the existence of $w^\mathbb{H}$ implies that the slope is finite. Namely in Proposition 2.1 the implication *ii*) \Rightarrow *i*) is a consequence of Theorem 23.14 of [28].

Proof. Let V be a smooth horizontal vector field with compact support and $\vartheta \in C_c^\infty(\mathbb{R}^{2n+1})$. We consider then $U_s(p) = \exp_p^\mathbb{H}(sV(p), s\vartheta(p))$. Because of Proposition 1.1 for s small enough the map U_s is smooth, one-to-one and $\text{Jac}(U_s)$ does not vanish. Note that $V(p)$ is the speed vector of $s \rightarrow U_s(p)$ at time $s = 0$ and that this curve is a geodesic of (constant) speed $\|V(p)\|_\mathbb{H}$. Then the metric speed of $((U_s)_\# \mu)_{s>0}$ is smaller than $\|V\|_{L^2_\mathbb{H}(\mu)}$. The derivate in time of $\text{Ent}_\infty(\mu_s)$ can be classically computed using the fact that $(p, s) \rightarrow U_s(p)$ is the flow of a smooth vector field corresponding to V in $s = 0$. Hence we have

$$\frac{d \text{Ent}_\infty(\mu_s)}{ds} \Big|_{s=0} = - \int \rho \text{div}_\mathbb{H}(V) d\mathcal{L}.$$

Note that if the entropy grows, which we want to avoid, we can replace V by $-V$. Therefore the following inequality with slope and speed on the right hand-side holds:

$$\left| \int \rho \text{div}_\mathbb{H}(V) d\mathcal{L} \right| \leq \text{Slope}(\text{Ent}_\infty)(\mu) \cdot \|V\|_{L^2_\mathbb{H}(\mu)}.$$

Because $\text{Slope}(\text{Ent}_\infty)(\mu)$ is finite, the Riesz representation theorem provides an horizontal vector field $w^\mathbb{H} \in L^2_\mathbb{H}(\mu)$ with $\|w^\mathbb{H}\|_{L^2_\mathbb{H}(\mu)} \leq \text{Slope}(\text{Ent}_\infty)(\mu)$ such that for any $V \in \Gamma(T\mathbb{H}_n)$ with compact support,

$$- \int \rho \text{div}_\mathbb{H}(V) d\mathcal{L} = \int \langle w^\mathbb{H} | V \rangle_\mathbb{H} d\mu.$$

From there, one observes that $\rho \in W_{\text{loc}}^{1,1}(\mathbb{H}_1)$ with $\nabla_\mathbb{H}\rho = \rho w^\mathbb{H}$.

For the second part of the proposition it is clear from Proposition 2.1 that *ii*) is a consequence of *i*). For the inverse implication, we assume that the weak gradient $\mathbf{T}\rho$ exists and that there is $w^\mathbf{T} \in L^2(\mu)$ such that $\mathbf{T}\rho = \rho w^\mathbf{T}$. With the first part of the proposition, $\rho \in W_{\text{loc}}^{1,1}(\mathbb{H}^\varepsilon)$ and $w^\varepsilon := w^\mathbb{H} + \varepsilon w^\mathbf{T}(\varepsilon\mathbf{T}) \in L^2_\varepsilon(\mu)$ satisfy

$$\rho w^\varepsilon = \rho w^\mathbb{H} + \rho \varepsilon w^\mathbf{T}(\varepsilon\mathbf{T}) = \nabla_\mathbb{H}\rho + \varepsilon \mathbf{T}\rho(\varepsilon\mathbf{T}) = \nabla_\varepsilon\rho.$$

Then Proposition 2.1 states that $\text{Slope}^\varepsilon(\text{Ent}_\infty)(\mu) < +\infty$ and $w^\varepsilon \in \text{Tan}_\varepsilon(\mathbb{H}_1)$ such that *ii*) implies *i*).

Therefore if these statements are true it is possible to approach w^ε in $L^2(\mathbb{H}_1^\varepsilon)$ by a sequence $(\nabla_\varepsilon \psi_k)_{k \in \mathbb{N}}$ where every $\psi_k \in C_c^\infty(\mathbb{R}^{2n+1})$. It follows that $(\nabla_{\mathbb{H}} \psi_k)_{k \in \mathbb{N}}$ tends to the horizontal part of w^ε . Hence $w^{\mathbb{H}}$ is in the tangent space $\text{Tan}_{\mathbb{H}}(\mu)$.

We already know the inequality $\|w^{\mathbb{H}}\|_{L_{\mathbb{H}}^2(\mu)} \leq \text{Slope}(\text{Ent}_\infty)(\mu)$, we will prove the opposite inequality thanks to inequality (11) in Proposition 2.4. In this inequality, we first replace every $W^\varepsilon(\mu, \nu)$ by $W(\mu, \nu)$. It is allowed because the second is greater. Then we state ε as a function of ν with $\varepsilon = W(\mu, \nu)^{1/3}$. But

$$\|w^{\mathbb{H}}\|_{L_{\mathbb{H}}^2} \leq \|w^\varepsilon\|_{L_{\mathbb{H}}^2} \leq \sqrt{\|w^\varepsilon\|_{L_{\mathbb{H}}^2}^2 + (\varepsilon \|w^{\mathbf{T}}\|_{L^2(\mu)})^2} \leq \|w^{\mathbb{H}}\|_{L_{\mathbb{H}}^2} + \frac{1}{2} \frac{(\varepsilon \|w^{\mathbf{T}}\|_{L^2(\mu)})^2}{\|w^{\mathbb{H}}\|_{L_{\mathbb{H}}^2}}$$

because the graph of $\sqrt{\cdot}$ is under the tangent line in $\|w^{\mathbb{H}}\|_{L_{\mathbb{H}}^2}^2$. It follows

$$\begin{aligned} \text{Ent}_\infty(\nu) &\geq \text{Ent}_\infty(\mu) - W(\mu, \nu) \left[\|w^{\mathbb{H}}\|_{L_{\mathbb{H}}^2} + \frac{W(\mu, \nu)^{2/3}}{2\|w\|_{L_{\mathbb{H}}^2}} \|w^{\mathbf{T}}\|_{L^2(\mu)}^2 \right] \\ &\quad - W(\mu, \nu)^{4/3} \\ &\geq \text{Ent}_\infty(\mu) - W(\mu, \nu) \|w^{\mathbb{H}}\|_{L_{\mathbb{H}}^2} - O(W(\mu, \nu)^{4/3}). \end{aligned}$$

when ν tends to μ . Thus $\text{Slope}(\text{Ent}_\infty)(\mu) \leq \|w^{\mathbb{H}}\|_{L_{\mathbb{H}}^2}$ and the equality follows. \square

5. HEAT EQUATIONS ON THE HEISENBERG GROUP

In this section we prove two theorems that justify the belief that the gradient flows of the entropy in the Heisenberg group are exactly the solutions of the hypoelliptic diffusion, equation (3). Before that we say more about these solutions. For this purpose we present asymptotic estimates of \mathfrak{h}_s and its derivative that we need in the proof of Theorem 0.1. From the relation $\mathfrak{h}_s(z, t) = \frac{1}{s^{n+1}} \mathfrak{h}_1(\frac{z}{\sqrt{s}}, \frac{t}{s})$, we state for \mathfrak{h}_s two theorems by Li [20] initially stated for \mathfrak{h}_1 .

Theorem 5.1. *Let $k \in \mathbb{N}$ and $(\alpha_1, \dots, \alpha_k) \in \mathbb{N}^k$ a multi-index. For $j \in \{1, \dots, k\}$ let $U_j \in \{\mathbf{X}_1, \mathbf{Y}_1, \dots, \mathbf{X}_n, \mathbf{Y}_n\}$, then there is a $C > 0$ only depending on n and $|\alpha| = \sum_{j=1}^k \alpha_j$ such that for any $s > 0$,*

$$|U_1^{\alpha_1} \dots U_k^{\alpha_k} \mathfrak{h}_s(p)| \leq \frac{C}{\sqrt{s}^{|\alpha|}} \left(1 + \frac{d_c(p)}{\sqrt{s}}\right)^{|\alpha|} \mathfrak{h}_s(p).$$

Theorem 5.2. *There is constant $A \geq 1$ depending on n such that for any $s > 0$,*

$$A^{-1} P_s(n, |z|, d_c(z, t)) \leq \mathfrak{h}_s(z, t) \leq A P_s(n, |z|, d_c(z, t))$$

where $d_c(z, t) = d_c(0, (z, t))$ and

$$P_s(n, r, d) = \frac{e^{-d^2/(4s)}}{s^{n+1} \sqrt{1+rd/s}} \left(\frac{1+d^2/s}{1+rd/s} \right)^{n-1}.$$

A further estimate related to these theorems is the fact that

$$(14) \quad \mathbf{T} \mathfrak{h}_s \leq A s^{-1} P_s(n, |z|, d_c(z, t)) \leq A^2 s^{-1} \mathfrak{h}_s$$

where we can consider that A is the same as before. This can be easily proved using the same remarks as in [20, Theorem 2] that principally relies on [5].

If we now assume that the support of ρ_0 is compact and included in a ball of center $0_{\mathbb{H}}$ and radius d_0 we get new rough estimates for the decays of ρ_s , $|\mathbf{T} \rho_s|$, $|\nabla_{\mathbb{H}} \rho_s|$ and $|\Delta_{\mathbb{H}} \rho_s|$ by simply considering that ρ_s is given by a convolution (5) and

using Theorem 5.1, Theorem 5.2, and (14) for the decay of \mathfrak{h}_s and its derivatives. In the next estimates, we chose to reduce some factors in the exponential factor by using the well-known fact that $|z| \leq d_c(z, t)$. For $s > 0$ and $\eta \in]0, 1[$ there is a constant C_η such that

$$\begin{aligned} C_\eta^{-1} \exp\left(-\frac{(d_c(z, t) + d_0)^2}{4s}(1 + \eta)\right) s^{-(n+1)} \\ \leq |\rho_s| \\ \leq C_\eta \exp\left(-\frac{(d_c(z, t) - d_0)^2}{4s}(1 - \eta)\right) s^{-(n+1)}, \end{aligned} \quad (15)$$

$$\begin{aligned} |\mathbf{T}\rho_s| = |\rho_0 *_{\mathbb{H}}(\mathbf{T}\mathfrak{h}_s)| \leq \rho_0 *_{\mathbb{H}} |\mathbf{T}\mathfrak{h}_s| \\ \leq C \exp\left(-\frac{(d_c(z, t) - d_0)^2}{4s}(1 - \eta)\right) s^{-(n+2)}, \end{aligned} \quad (16)$$

$$\begin{aligned} |\nabla_{\mathbb{H}}\rho_s| = |\rho_0 *_{\mathbb{H}}(\nabla_{\mathbb{H}}\mathfrak{h}_s)| \leq \rho_0 *_{\mathbb{H}} |\nabla_{\mathbb{H}}\mathfrak{h}_s| \\ \leq C_\eta \exp\left(-\frac{(d_c(z, t) - d_0)^2}{4s}(1 - \eta)\right) s^{-(n+(3/2))}, \end{aligned} \quad (17)$$

$$\begin{aligned} |\Delta_{\mathbb{H}}\rho_s| = |\rho_0 *_{\mathbb{H}}(\Delta_{\mathbb{H}}\mathfrak{h}_s)| \leq \rho_0 *_{\mathbb{H}} |\Delta_{\mathbb{H}}\mathfrak{h}_s| \\ \leq C_\eta \exp\left(-\frac{(d_c(z, t) - d_0)^2}{4s}(1 - \eta)\right) s^{-(n+2)}. \end{aligned} \quad (18)$$

5.1. Proof of Theorem 0.1. Although our definition of gradient flow in Subsection 1.4 is just metric and does not even tell what is the “gradient”, $\nabla_{\mathbb{H}}\rho_s/\rho_s$ is morally the gradient of Ent_∞ in $\rho_s d\mathcal{L}$. We will show that $-\nabla_{\mathbb{H}}\rho_s/\rho_s$ satisfy (12) in the place of v_s . From there it will be easier to compute first the speed of $(\mu_s)_{s \in I}$, then $\text{Slope}(\text{Ent}_\infty)(\mu_s)$ and finally the derivative of $E(s) = \text{Ent}_\infty(\mu_s)$.

Let $\psi \in \mathcal{C}_c^\infty(\mathbb{R}^{2n+1})$. We recall that ρ_s is smooth in space and time for $s > 0$. Moreover, $\psi\rho_s$ is in $\mathcal{C}_c^\infty(\mathbb{R}^{2n+1})$ and its support is in the support of ψ . Then

$$\begin{aligned} \frac{d}{ds} \int_{\mathbb{H}} \psi \rho_s d\mathcal{L} &= \int_{\mathbb{H}} \psi \partial_s \rho_s d\mathcal{L} = \int_{\mathbb{H}} \psi \Delta_{\mathbb{H}} \rho_s d\mathcal{L} \\ &= \int_{\mathbb{H}} \langle \nabla_{\mathbb{H}} \psi \mid -\nabla_{\mathbb{H}} \rho_s \rangle_{\mathbb{H}} d\mathcal{L} = \int_{\mathbb{H}} \langle \nabla_{\mathbb{H}} \psi \mid -\frac{\nabla_{\mathbb{H}} \rho_s}{\rho_s} \rangle_{\mathbb{H}} d\mu_s. \end{aligned} \quad (19)$$

Hence

$$\partial_s(\mu_s) + \text{div}_{\mathbb{H}}\left(-\frac{\nabla_{\mathbb{H}}\rho_s}{\rho_s}\mu_s\right) = 0 \quad (20)$$

holds for every s . We prove now that $-\frac{\nabla_{\mathbb{H}}\rho_s}{\rho_s}$ is in $L^2_{\mathbb{H}}(\mu_s)$. Indeed the integral $\int \frac{\|\nabla_{\mathbb{H}}\rho_s\|_{\mathbb{H}}^2}{\rho_s} d\mu_s$ is finite because we can estimate the numerator from above with (17) and the denominator from below with (15). Moreover, the domination is such that $\frac{\nabla_{\mathbb{H}}\rho_s}{\rho_s}$ is continuous on I .

Starting from (20) because of the Mass Conservation Formula [28, Chapter 1] there is a probability measure Π on $\mathcal{C}(I, \mathbb{H}_n)$, the space of curves over \mathbb{H}_n with the Borel sigma-field, satisfying two conditions:

- the curve γ is $\Pi(\gamma)$ -almost certainly is an integral curve of the smooth vector field $-\nabla_{\mathbb{H}}\rho_s/\rho_s$,

$$\dot{\gamma}_s = -\frac{\nabla_{\mathbb{H}}\rho_s}{\rho_s}$$

- the law of the point $\gamma(s)$ with respect to Π is μ_s .

Then

$$\begin{aligned} W^2(\mu_s, \mu_t) &\leq \int_{\mathcal{C}(I, \mathbb{H}_n)} d_c^2(\gamma_s, \gamma_t) d\Pi(\gamma) \\ &\leq \int_{\mathcal{C}(I, \mathbb{H}_n)} \left((t-s) \int_s^t \left\| \frac{\nabla_{\mathbb{H}}\rho_\tau}{\rho_\tau}(\gamma_\tau) \right\|_{\mathbb{H}}^2 d\tau \right) d\Pi(\gamma) \\ &\leq (t-s) \int_s^t \left(\int_{\mathbb{H}_n} \left\| \frac{\nabla_{\mathbb{H}}\rho_\tau}{\rho_\tau} \right\|_{\mathbb{H}}^2 d\mu_\tau \right) d\tau \\ &\leq (t-s)^2 \sup_{\{\tau \in [s, t]\}} \|\nabla_{\mathbb{H}}\rho_\tau/\rho_\tau\|_{L_{\mathbb{H}}^2(\mu_\tau)}^2. \end{aligned}$$

Hence $(\mu_s)_{s \in I}$ is locally Lipschitz and therefore this curve is absolutely continuous. The norm $\|\nabla_{\mathbb{H}}\rho_s/\rho_s\|_{L_{\mathbb{H}}^2(\mu_s)}$ is continuous such that letting t go to s we get

$$(21) \quad |\dot{\mu}_s| \leq \|\nabla_{\mathbb{H}}\rho_s/\rho_s\|_{L_{\mathbb{H}}^2(\mu_s)}.$$

It is not enough to obtain $|\dot{\mu}_s| = \|\nabla_{\mathbb{H}}\rho_s/\rho_s\|_{L_{\mathbb{H}}^2(\mu_s)}$ but (21) will be sufficient for the proof.

We now compute the Slope of Ent_∞ . Actually by using (15) and (16) one get that

$$\|\nabla_\varepsilon \rho_s/\rho_s\|_{L_\varepsilon^2(\mu_s)} = \sqrt{\|\nabla_{\mathbb{H}}\rho_s/\rho_s\|_{L_{\mathbb{H}}^2(\mu_s)}^2 + \varepsilon^2 \int \left(\frac{\mathbf{T}\rho_s}{\rho_s} \right)^2 \rho_s d\mathcal{L}}$$

is finite. But this quantity is $\text{Slope}^\varepsilon(\text{Ent}_\infty)(\mu_s)$ such that condition (ii) of Proposition 4.1 is satisfied and as ρ_s does not vanish the vector field $w_s^{\mathbb{H}}$ of Proposition 4.1 is $\nabla_{\mathbb{H}}\rho_s/\rho_s$. From there in the Wasserstein space of the “true” Heisenberg group we have $\text{Slope}(\text{Ent}_\infty)(\mu_s) = \|\nabla_{\mathbb{H}}\rho_s/\rho_s\|_{L_{\mathbb{H}}^2(\mu_s)}$.

Finally we compute $\dot{E}(s)$ where $E(s) = \text{Ent}_\infty(\mu_s)$ as in Subsection 1.4. Firstly

$$(22) \quad |\partial_s(\rho_s \ln(\rho_s))| = |(1 + \ln(\rho_s))\partial_s \rho_s| = |(1 + \ln(\rho_s))\Delta_{\mathbb{H}}\rho_s|.$$

Using (15) and (18) we see that about every time $s_0 > 0$, (22) is dominated independently of s by a function of $L^1(\mathbb{H}_n)$ (use for instance $|1 + \ln(X)| < X + X^{-1/2}$). Therefore it is allowed to derivate in s under the integral sign and we obtain

$$\dot{E}(s) = \int (1 + \ln(\rho_s))\Delta_{\mathbb{H}}\rho_s.$$

We want now to justify the partial integration

$$(23) \quad \dot{E}(s) = - \int \langle \nabla_{\mathbb{H}}(1 + \ln(\rho_s)) \mid \nabla_{\mathbb{H}}\rho_s \rangle$$

which is less obvious as the one in (19) because the supports of $1 + \ln(\rho_s)$, $\mathbf{X}\rho_s$ and $\mathbf{Y}\rho_s$ are not compact. However, if for a smooth function f we integrate $\mathbf{X}f =$

$\partial_x f - \frac{y}{2} \partial_t f$ on $[-R, R]^3$, we get

$$\begin{aligned} \left| \int_{[-R, R]^3} \mathbf{X}f d\mathcal{L} \right| &= \left| \int_{\{x=R\} \times [-R, R]^2} f - \int_{\{x=-R\} \times [-R, R]^2} f \right. \\ &\quad \left. - \frac{1}{2} \int_{[-R, R]^2 \times \{t=R\}} yf + \frac{1}{2} \int_{[-R, R]^2 \times \{t=-R\}} yf \right| \\ &\leq R \int_{\partial[-R, R]^3} |f| \end{aligned}$$

where $\partial[-R, R]^3$ is the border of $[-R, R]^3$. We are interested in $f = (1 + \ln(\rho_s)) \mathbf{X} \rho_s$. As we have (15) and because d_c is Lipschitz-equivalent to the Korányi-Reimann distance (see [17, Subsection 1.1.3]), the previous integral tends to zero when R goes to infinity. A similar computation holds for $\left| \int_{[-R, R]^3} \mathbf{Y}((1 + \ln(\rho_s)) \mathbf{Y} \rho_s) \right|$. It also tends to zero when R goes to infinity. From there the partial integration in (23) is justified. It follows

$$\dot{E}(s) = - \int \left\langle \frac{\nabla_{\mathbb{H}} \rho_s}{\rho_s} \mid \frac{\nabla_{\mathbb{H}} \rho_s}{\rho_s} \right\rangle_{\mathbb{H}} \rho_s d\mathcal{L} = - \int \left\| \frac{\nabla_{\mathbb{H}} \rho_s}{\rho_s} \right\|^2 d\mu_s = - \left\| \frac{\nabla_{\mathbb{H}} \rho_s}{\rho_s} \right\|_{L^2_{\mathbb{H}}(\mu_s)}^2.$$

Because of the differentiability properties of $E(s)$, we have

$$(24) \quad \dot{E}(s) \leq \text{Slope}(\text{Ent}_{\infty})(\mu_s) \cdot |\dot{\mu}_s|$$

in almost every $s \in I$. However, in this proof we have shown that at almost every $s \in I$

$$\left\{ \begin{array}{l} |\dot{\mu}_s| \leq \left\| \frac{\nabla_{\mathbb{H}} \rho_s}{\rho_s} \right\|_{L^2_{\mathbb{H}}(\mu_s)} \\ \text{Slope}(\text{Ent}_{\infty})(\mu_s) = \left\| \frac{\nabla_{\mathbb{H}} \rho_s}{\rho_s} \right\|_{L^2_{\mathbb{H}}(\mu_s)} \\ \dot{E}(s) = - \left\| \frac{\nabla_{\mathbb{H}} \rho_s}{\rho_s} \right\|_{L^2_{\mathbb{H}}(\mu_s)}^2 \end{array} \right.$$

such that inequality (24) is an equality for almost every $s \in I$. Thus $(\mu_s)_{s \in I}$ is a gradient flow of Ent_{∞} .

5.2. Proof of Theorem 0.2. Let $(\mu_s)_{s \in I}$ be a gradient flow of the entropy in $\mathcal{P}_2(\mathbb{H}_n)$. This curve is absolutely continuous (even in AC^2). Hence $\text{Slope}(\text{Ent}_{\infty})(\mu_s) = |\dot{\mu}_s|$ is finite in almost every $s \in I$. With the assumption about $w^{\mathbf{T}}$ we get that Proposition 4.1 applies, in particular for any $\varepsilon > 0$ we have $\rho_s \in W_{\text{loc}}^{1,1}(\mathbb{H}_n^{\varepsilon})$ in almost every s . Therefore Proposition 2.1 and Proposition 2.4 apply too. Indeed $(\mu_s)_{s \in I}$ is a gradient flow of Ent_{∞} such that with Remark 1.2 it is a AC^2 -curve of $\mathcal{P}_2(\mathbb{H}_n)$. Consequently it is a AC^2 -curve of $\mathcal{P}_2(\mathbb{H}_n^{\varepsilon})$ too.

From Proposition 4.1, we know that $\text{Slope}(\text{Ent}_{\infty})(\mu_s) = \|w_s^{\mathbb{H}}\|_{L^2_{\mathbb{H}}}$ in almost every s where $w_s^{\mathbb{H}} \in \text{Tan}_{\mathbb{H}}(\mu_s)$ is the horizontal part of $w_s^{\varepsilon} \in \text{Tan}_{\varepsilon}(\mu_s)$ with $\text{Slope}^{\varepsilon}(\text{Ent}_{\infty})(\mu_s) = \|w_s^{\varepsilon}\|_{L^2_{\varepsilon}(\mu_s)}$. We fix now some $s' \in I'_{\varepsilon} \cap I'$ with the notation of Proposition 2.4 and Proposition 3.1. Then we would like to interpret the scalar product $\int \langle v_{s'}^{\varepsilon} \mid w_{s'}^{\varepsilon} \rangle_{\varepsilon} d\mu_{s'}$ appearing in (10) of Proposition 2.4 in terms of the Wasserstein space $\mathcal{P}_2(\mathbb{H}_n)$. First of all in the tangent space $\text{Tan}_{\varepsilon}(\mu_s)$ we approach the vector field w_s^{ε} by a sequence $(\nabla_{\varepsilon} \psi_k)_{k \in \mathbb{N}}$ where $\psi_k \in \mathcal{C}_c^{\infty}$. Then because of

Proposition 2.4, we see that $\int \langle w_{s'}^\varepsilon | v_{s'}^\varepsilon \rangle_\varepsilon d\mu_{s'}$ is the limit of $\frac{\partial}{\partial s} \mu_s(\psi_k)(s')$. But we know from Proposition 3.1 that the previous derivative is also $\int \langle \nabla_{\mathbb{H}} \psi_k | v_{s'} \rangle_{\mathbb{H}} d\mu_{s'}$. From there the limit is simply $\int \langle w_{s'}^{\mathbb{H}} | v_{s'} \rangle_{\mathbb{H}} d\mu_{s'}$. Indeed $(\nabla_{\mathbb{H}} \psi_k)_{k \in \mathbb{N}}$ tends to the horizontal part w_s^ε .

Then in almost every $s' \in I$, (10) writes

$$\text{Ent}_\infty(\mu_s) - \text{Ent}_\infty(\mu_{s'}) \geq (s - s') \int \langle w_{s'}^{\mathbb{H}} | v_{s'} \rangle_{\mathbb{H}} d\mu_{s'} + o(|s - s'|).$$

But $(\mu_s)_{s \in I}$ is a gradient flow. It satisfies $\dot{E}_{s'} = -(\text{Slope}(\text{Ent}_\infty)(\mu_{s'})) \cdot |\dot{\mu}_{s'}|$ where $|\dot{\mu}_{s'}| \geq \|v_{s'}\|_{L^2_{\mathbb{H}}(\mu_{s'})}$ and $\text{Slope}(\text{Ent}_\infty)(\mu_{s'}) = \|w_{s'}^{\mathbb{H}}\|_{L^2_{\mathbb{H}}(\mu_s)}$. Then because of the Cauchy-Schwarz inequality, the only possibility is that $v_{s'}$ and $w_{s'}^{\mathbb{H}}$ are negatively collinear in $\text{Tan}_{\mathbb{H}}(\mu_{s'})$. Moreover, as the gradient flows satisfy $\text{Slope}(\text{Ent}_\infty)(\mu_{s'}) = |\dot{\mu}_{s'}|$, we have simply $v_{s'} = -w_{s'}^{\mathbb{H}}$ in almost every $s' \in I$. Replacing $v_{s'}$, one can therefore rewrite the continuity equation of Proposition 3.1:

$$\partial_{s'} \mu_{s'} + \text{div}_{\mathbb{H}}(-w_{s'}^{\mathbb{H}} \mu_{s'}).$$

Because of relation $\nabla_{\mathbb{H}} \rho_{s'} = \rho_{s'} w_{s'}^{\mathbb{H}}$ it is also

$$(25) \quad \partial_s \mu_s + \text{div}_{\mathbb{H}}(-\nabla_{\mathbb{H}} \rho_{s'} d\mathcal{L})$$

where $\nabla_{\mathbb{H}} \rho_{s'}$ is the weak gradient of $\rho_{s'}$. Remind that (25) means that for any $\psi \in C_c^\infty(\mathbb{R}^{2n+1})$,

$$(26) \quad \left(\frac{\partial}{\partial s} \int \psi d\mu_s \right) |_{s'} = \int \langle \nabla_{\mathbb{H}} \psi | -\nabla_{\mathbb{H}} \rho_{s'} \rangle_{\mathbb{H}} d\mathcal{L}.$$

We know from the proof of Proposition 3.1 that the integral on the left-hand side, namely ζ^ψ is absolutely continuous in s . It follows that we can integrate (26) on an interval $[\sigma, \tau]$ and obtain

$$\int \psi \rho_\tau d\mathcal{L} - \int \psi \rho_\sigma d\mathcal{L} = - \int_\sigma^\tau \int \langle \nabla_{\mathbb{H}} \psi | \nabla_{\mathbb{H}} \rho_{s'} \rangle_{\mathbb{H}} d\mathcal{L} ds'$$

for any $\psi \in C_c^\infty(\mathbb{R}^{2n+1})$. We recognize a weak formulation of the ‘‘hypoelliptic heat equation’’. By using classical references about hypoelliptic operator as [25] or [27] and the references therein, this concludes the proof.

5.3. Remarks.

Remark 5.3. It is also possible to make similar proofs in the Albanese torus \mathbb{T}_n , a compact quotient of the Heisenberg group approximating it by \mathbb{T}_n^ε (see [17, Subsection 1.2] where it is presented for $n = 1$). The fact that \mathbb{T}_n is compact change small elements in the proofs. The partial integration is still right in this case and one does not require rapid decay. Furthermore the assumption on the support of μ_0 in Theorem 0.1 is necessary satisfied. It follows that one can mix Theorem 0.1 and Theorem 0.2 in the following way

Theorem 5.4. *Let $(\mu_s)_{s \in I}$ be a curve of $\mathcal{P}_2(\mathbb{T}_n)$ and ρ_s the density curve of μ_s with respect to $\mathcal{L}_{\mathbb{T}}$, that is the quotient of the Lebesgue measure. Both statements are equivalent*

- *The density curve $(\rho_s)_{s \in]0, +\infty[}$ is a solution of the hypoelliptic heat equation of \mathbb{T}_n ,*

$$\Delta_{\mathbb{T}} \rho_s = \partial_s \rho_s,$$

where $\Delta_{\mathbb{T}} = \sum_{k=1}^n \mathbf{X}_{k, \mathbb{T}}^2 + \mathbf{Y}_{k, \mathbb{T}}^2$.

- The curve $(\mu_s)_{s \in I}$ of measures with density $\rho_s = d\mu_s/d\mathcal{L}$ is a gradient flow of the entropy Ent_∞ and the weak derivative $\mathbf{T}\rho_s$ exists with $w^{\mathbf{T}}$ in $L^2_{\mathbb{T}}$ defined by $w_s^{\mathbf{T}}\rho_s = \mathbf{T}\rho_s$.

Remark 5.5. In [18, Section 6], Khesin and Lee prove a similar result to Theorem 0.1. Their paper takes place in the wide class of bracket-generating distribution τ on a connected and compact manifold M (it includes the Albanese torus). They also approximate their metric space by Riemannian manifolds completing the horizontal tangent space by the other directions. However, their result does not include Theorem 0.1 because their result does not apply to the non-compact \mathbb{H}_n . The proof of Khesin and Lee is more algebraic than the proofs of this paper and the Wasserstein space they are considering is a “smooth” Wasserstein space. It is restricted to smooth measures and the distance is possibly different. Actually the authors begin to give a tangential structure to $\mathcal{P}_2(M_\tau)$ defining the length of curves on the Wasserstein space and then defining the distance. From there the definition of the gradient flow is different: the solution of the hypoelliptic equation goes in the “smooth” direction with the greatest slope but a rough $\text{Slope}(\text{Ent}_\infty)$ is not defined. In [18], there is no analogous result to Theorem 0.2.

ACKNOWLEDGEMENT

I wish to thank my advisers Hervé Pajot and Karl-Theodor Sturm for introducing me in the study of optimal transport in the Heisenberg group. I thank the professors Ambrosio and Villani for helpful discussions as well as my colleagues Matthias Erbar and Miguel Rodrigues.

REFERENCES

- [1] 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.
- [2] L. Ambrosio and S. Rigot. Optimal mass transportation in the Heisenberg group. *J. Funct. Anal.*, 208(2):261–301, 2004.
- [3] L. Ambrosio and G. Savaré. Gradient Flows of Probability Measures. In *Handbook of differential equations, evolutionary equations*, volume 3, pages 1–136. 2006.
- [4] L. Ambrosio and P. Tilli. *Topics on analysis in metric spaces*, volume 25 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2004.
- [5] R. Beals, B. Gaveau, and P.C. Greiner. Hamilton-Jacobi theory and the heat kernel on Heisenberg groups. *J. Math. Pures Appl. (9)*, 79(7):633–689, 2000.
- [6] J. Cheeger and B. Kleiner. Generalized differential and bi-Lipschitz nonembedding in L^1 . *C. R. Math. Acad. Sci. Paris*, 343(5):297–301, 2006.
- [7] M. Erbar. Die Wärmeleitungsgleichung auf Mannigfaltigkeiten als Gradientenfluss im Wassersteinraum. Master’s thesis, Bonn Universität, 2008.
- [8] M. Erbar. The heat equation on manifolds as a gradient flow in the Wasserstein space. *to appear in Annales de l’Institut Henri Poincaré (B) Probabilités et Statistiques*, 2008.
- [9] S. Fang, J. Shao, and K.-T. Sturm. Wasserstein space over the Wiener space. *preprint 383, SFB 611 (Bonn)*, 2008.
- [10] A. Figalli and N. Juillet. Absolute continuity of Wasserstein geodesics in the Heisenberg group. *J. Funct. Anal.*, 255(1):133–141, 2008.
- [11] A. Figalli and L. Rifford. Mass Transportation on sub-Riemannian Manifolds. *preprint*, 2008.
- [12] B. Franchi, R. Serapioni, and F. Serra Cassano. Rectifiability and perimeter in the Heisenberg group. *Math. Ann.*, 321(3):479–531, 2001.
- [13] 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.

- [14] L. Hörmander. Hypoelliptic second order differential equations. *Acta Math.*, 119:147–171, 1967.
- [15] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17 (electronic), 1998.
- [16] N. Juillet. Ricci curvature bounds and geometric inequalities in the Heisenberg group. *preprint SFB611 Bonn*, 2006.
- [17] N. Juillet. *Optimal transport and geometric analysis in Heisenberg groups*. PhD thesis, Université Grenoble 1, Bonn Universität, <http://tel.archives-ouvertes.fr/tel-00345301>, 2008.
- [18] B. Khesin and Lee P. A Nonholonomic Moser theorem and Optimal mass transport. *preprint*, 2008.
- [19] A. Korányi and H. M. Reimann. Quasiconformal mappings on the Heisenberg group. *Invent. Math.*, 80(2):309–338, 1985.
- [20] H.-Q. Li. Estimations asymptotiques du noyau de la chaleur sur les groupes de Heisenberg. *C. R. Math. Acad. Sci. Paris*, 344(8):497–502, 2007.
- [21] R. J. McCann. Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.*, 11(3):589–608, 2001.
- [22] R. Montgomery. *A Tour of Subriemannian Geometries, Their Geodesics and Applications*, volume 91 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2002.
- [23] S.-I. Ohta. Gradient flows on Wasserstein spaces over compact Alexandrov spaces. *to appear in Am. J. Math.*, 2006.
- [24] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.
- [25] L. P. Rothschild and E. M. Stein. Hypoelliptic differential operators and nilpotent groups. *Acta Math.*, 137(3-4):247–320, 1976.
- [26] G. Savaré. Gradient flows and diffusion semigroups in metric spaces under lower curvature bounds. *C. R. Math. Acad. Sci. Paris*, 345(3):151–154, 2007.
- [27] E. M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, *Monographs in Harmonic Analysis*, III.
- [28] C. Villani. *Optimal transport, old and new*, volume 383. Grundlehren der mathematischen Wissenschaften (Springer), 2008.

INSTITUT FOURIER BP 74, UMR 5582, UNIVERSITÉ GRENOBLE 1, 38402 SAINT-MARTIN-D'HÈRES CEDEX, FRANCE

E-mail address: nicolas.juillet@ujf-grenoble.fr

INSTITUT FÜR ANGEWANDTE MATHEMATIK, UNIVERSITÄT BONN, POPPELSDORFER ALLEE 82, 53115 BONN, GERMANY

E-mail address: juillet@wiener.iam.uni-bonn.de