A SIMPLE PROOF OF LOGARITHMIC SOBOLEV INEQUALITY ON THE HEISENBERG GROUPS

A. S. ÜSTÜNEL

Abstract: In this note we give a simple, dimension independent, proof of the logarithmic Sobolev inequality on the Heisenberg groups \( H_n = \mathbb{R}^{2n+1} \) using the measure preserving transformations of the Brownian motion.

2020 Mathematics Subject Classification. Primary 60H; Secondary 22E30, 22E66, 35A23, 35K08, 35R03, 60J65.

Keywords: Entropy, degenerate diffusions, relative entropy, Poincaré inequality, logarithmic Sobolev inequality, intertwining, Heisenberg group, left and right invariant vector fields, opérateur carré du champs.

1. Introduction

In this work we give a very simple and short proof of the logarithmic Sobolev inequality for Heisenberg group \( H_n \), for any \( n \geq 1 \). This inequality is very interesting since this is the simplest non-linear (i.e., non-gaussian) case with degenerate diffusion coefficients, hence outside the reach of the standard methods and it has been suggested to us by Leonard Gross. For \( n = 1 \), the proof is done by using the expression of the density of the law of \( x_1 \) (\( \mathbb{R} \)) which is strongly dimension dependent (cf., also \( \mathbb{R} \)). Here we give a totally different proof which is dimension independent with the best constant since it is the same as in the flat Gaussian case, cf. \( \mathbb{R} \). This is done with the help of the invariance properties of the Gauss law under Euclidean symmetries. The key point in our approach is the fact that the difficulties due to the nonlinearity of the problem can be avoided by using spontaneously the left Markov processes generated by the left invariant vector fields generating the Lie algebra of \( H_n \) and its mirror image. This approach reduces the problem to a linear problem, cf., \( \mathbb{R} \) and avoids all the difficulties due to the nonlinearity of the problem.

Last but not the least we are grateful to Leonard Gross who has indicated the problem as well as the related references.

2. Poincaré and Log-Sobolev inequalities on the Heisenberg Group

We denote \( \mathbb{R}^{2n} \times \mathbb{R} \) by \( H_n \) which is equipped by the non-commutative product defined as

\[
(v, z) \star (v', z') = (v + v', z + z' + \frac{1}{2} \omega(v, v')),
\]

\( v = (x, y), v' = (x', y') \) and \( x, y, x', y' \in \mathbb{R}^n \). \( \omega \) is an antisymmetric bilinear form defined as

\[
\omega(v, v') = \sum_{i=1}^{n}(x_i y'_i - x'_i y_i)
\]
Under the ⋆ operation defined above \( H_n \) becomes a Lie group with its Lie algebra, \( \mathcal{H} \), generated by the left invariant vector fields
\[
X_i = \partial_{x_i} - \frac{y_i}{2} \partial_z, \quad Y_i = \partial_{y_i} + \frac{x_i}{2} \partial_z, \quad Z = \partial_z,
\]
i = 1, \ldots, n. We define also the right invariant vector fields
\[
\hat{X}_i = \partial_{x_i} + \frac{y_i}{2} \partial_z, \\
\hat{Y}_i = \partial_{y_i} - \frac{x_i}{2} \partial_z, \\
\hat{Z} = Z = \partial_z.
\]

Let \( b \) and \( w \) be two \( \mathbb{R}^n \)-valued, independent Brownian motions, the diffusion process with which we are concerned is the unique solution of the following SDE which is written in terms of differential geometric language:
\[
f(x_t) = \int_0^t X_i f(x_s) \cdot db^j_s + \int_0^t Y_i f(x_s) \cdot dw^j_s,
\]
for any smooth function \( f \) on \( H_n \) and the integrals are taken in Stratonovitch sense\(^1\). The solution of this equation is given explicitly as \( x_t = (b_t, w_t, \frac{1}{2} z_t(b, w)) \), where \( z_t \) is the Lévy area process (with Itô integrals):
\[
z_t = \sum_{i=1}^n \int_0^t (b^i_s dw^i_s - w^i_s db^i_s).
\]

We shall denote by \( (Q_t, t \geq 0) \) the (left) semi-group corresponding to the Markov process defined by \( Q_t f(p) = E[f(p \ast x_t)] \), \( p \in H_n \). Note the hypoelliptic infinitesimal generator \( \mathbb{L} \) of \( (Q_t) \) is given by
\[
\mathbb{L} = \frac{1}{2} \sum_{i=1}^n (X_i^2 + Y_i^2).
\]

Let us note that \([X_i, Y_i] = Z, [X_i, Z] = [Y_i, Z] = 0 \) for \( n = 1, \ldots, n \), consequently the Lie algebra generated by \( \{X_i, Y_i, Z; i = 1, \ldots, n\} \) fulfills the Hörmander condition, which assures the hypoellipticity of the operator \( \mathbb{L} \).

It follows from direct calculation that \( \hat{X}_i \) and \( \hat{Y}_i \) commute with \( \mathbb{L} \) hence they commute also with the semigroup \( (Q_t) \) generated by \( \mathbb{L} \), which is constructed through the stochastic differential equation\(^2\)
\[
dx_t = X_i(x_t) db^i_t + Y_i(x_t) dw^i_t,
\]
written with Itô integrals. Let now \( A \) be the isometry defined by
\[
-A = \begin{bmatrix}
I_{\mathbb{R}^n} & 0 & 0 \\
0 & I_{\mathbb{R}^n} & 0 \\
0 & 0 & -1
\end{bmatrix}
\]
and let us denote by \((y_t, t \geq 0)\) the process defined by \( y_t = (-b_t, -w_t, \frac{1}{2} z_t(b, w)) = Ax_t \).

The following elementary property of \( A \) is important for the sequel:

**Lemma 1.** For any \( p, q \in H_n \), we have
\[
A(p \ast q) = Ap \ast Aq.
\]

\(^1\)denoted with a “dot”
\(^2\)we use the summation convention
Proof: Let $p = (x, y, z)$, $q = (x', y', z')$, then

$$A(p \star q) = A \left( (x + x'), (y + y'), z + z' + \frac{1}{2} \sum_i (x_i y'_i - x'_i y_i) \right)$$

$$= \left( -(x + x'), -(y + y'), z + z' + \frac{1}{2} \sum_i (x_i y'_i - x'_i y_i) \right).$$

On the other hand

$$A p \star A q = \left( -x - x', -y - y', z + z' + \frac{1}{2} \sum_i ((-x_i)(-y'_i) - (-x'_i)(-y_i)) \right)$$

$$= \left( -(x + x'), -(y + y'), z + z' + \frac{1}{2} \sum_i (x_i y'_i - x'_i y_i) \right)$$

$$= A(p \star q).$$

\[\square\]

**Lemma 2.** We have $y_t = Ax_t$ and the law of $y$ is equal to the law of $x$.

**Proof:** We have explicitly $x_t = (b_t, w_t, \frac{1}{2} z_t(b, w))$ and $y_t = (-b_t, -w_t, \frac{1}{2} z_t(b, w))$. From the Paul Lévy theorem, law$(b, w) = \text{law}(-b, -w)$, hence also

$$\text{law} \left( b, w, \frac{1}{2} z(b, w) \right) = \text{law} \left( -b, -w, \frac{1}{2} z(-b, -w) \right),$$

but $z(-b, -w) = z(b, w)$, consequently the laws of $x$ and $y$ on the path space are identical. \[\square\]

**Proposition 1.** We have the following commutation results:

\begin{align}
(X_t f)(A y_t) &= -\hat{X}_i(f \circ A)(y_t) \\
(Y_t f)(A y_t) &= -\hat{Y}_i(f \circ A)(y_t)
\end{align}

or equivalently

\begin{align}
(X_t f)(x_t) &= -\hat{X}_i(f \circ A)(Ax_t) = -\hat{X}_i(f \circ A)(y_t) \\
(Y_t f)(x_t) &= -\hat{Y}_i(f \circ A)(Ax_t) = -\hat{Y}_i(f \circ A)(y_t)
\end{align}

and in particular

$$\hat{I} f(A y_t) = \hat{I} (f \circ A)(y_t)$$

almost surely for any smooth function $f$.

**Proof:** These results follow from the Itô formula and the unique decomposition of a continuous semimartingale as a sum of a local martingale and a finite variation process, cf. [2]. In fact, for any smooth function $f$ on $\mathbb{R}^{2n+1}$, we can write

$$f(x_t) = f(0) + \int_0^t X_t f(x_s) \cdot db_s^i + \int_0^t Y_t(x_s) \cdot dw_s^i$$

$$= f(A y_t) = (f \circ A)(y_t)$$

$$= f(0) - \int_0^t \hat{X}_i(f \circ A)(y_s) \cdot db_s^i - \int_0^t \hat{Y}_i(f \circ A)(y_s) \cdot dw_s^i.$$
Hence the integrands in front of \( b \)'s of the first and third lines should be equal, the same reasoning holds for the integrands in front of \( w \)'s in the first and third lines and the proof of the relations (2.1) (2.2) follows. The remaining part of the proof follows by changing just the notations.

**Lemma 3.** For any \( t \geq 0 \), \( f \in C_0^\infty (H_n) \), we have

\[
E_1 \left[ f(y_T) \right] = Q_{T-t} f(y_t)
\]

Proof: Let \( \theta(b, w) \) be any bounded and \( \mathcal{F}_t(x) \)-measurable function. As \( \mathcal{F}_t(x) = \sigma(b_s, w_s, s \leq t) = \sigma(-b_s, -w_s, s \leq t) \) and as the law of \((b, w)\) is equal to the law of \((-b, -w)\), we get

\[
E_1 \left[ f(y_T(b, w)) \right] = E_1 \left[ f(y_T(-b, -w)) \right] = E \left[ f(x_T(b, w)) \theta(-b, -w) \right] = E \left[ Q_{T-t} f(x_t(b, w)) \theta(-b, -w) \right] = E \left[ Q_{T-t} f(y_t(b, w)) \theta(b, w) \right],
\]

by the invariance of the law of the Brownian motion under the isometries, i.e., by the relation

\[
\text{law}(b, w) = \text{law}(-b, -w)
\]

thanks to the Paul Lévy theorem.

**Lemma 4.** For any positive function \( f \) on \( H_n \), we have

\[
(Q_t f)(Ap) = Q_t (f \circ A)(p)
\]

for any \( p \in H_n \), \( t \geq 0 \).

Proof: We have

\[
(Q_t f)(Ap) = E \left[ f(A(p) \star x_t) \right] = E \left[ f(A(p) \star A(x_t)) \right] = E \left[ f(A(p) \star A(x_t)) \right] = Q_t (f \circ A)(p),
\]

where the equality (2.8) follows from the equality of the laws of \( x \) and \( Ax \) and equality (2.9) follows from Lemma 3.

**Proposition 2.** For any smooth \( f \), we have

\[
X_i Q_{1-t} f(x_t) = - Q_{1-t} (\hat{X}_i (f \circ A))(y_t)
\]
\[
Y_i Q_{1-t} f(x_t) = - Q_{1-t} (\hat{Y}_i (f \circ A))(y_t),
\]

\[
(X_t Q_{1-t} f(x_t) = - Q_{1-t} (\hat{X}_t (f \circ A))(y_t)
\]
\[
(Y_t Q_{1-t} f(x_t) = - Q_{1-t} (\hat{Y}_t (f \circ A))(y_t),
\]
for almost surely. In particular, the processes \((X_i Q_{1-t} f(x_i), t \in [0,1])\) and \((Y_i Q_{1-t} f(x_i), t \in [0,1])\) are martingales for \(i = 1, \ldots, n\).

**Proof:** Let \(f\) be a smooth (hence bounded) function on \(\mathbb{R}^{2n+1}\). We have, for \(i = 1, \ldots, n\), the following equalities:

\begin{align}
(2.11) \quad X_i Q_{1-t} f(x_i) & = -\hat{X}_i((Q_{1-t} f) \circ A)(y_i) \\
(2.12) & = -\hat{X}_i((Q_{1-t} (f \circ A))(y_i) \\
(2.13) & = -Q_{1-t} (\hat{X}_i(f \circ A))(y_i) \\
(2.14) & = -E[\hat{X}_i(f \circ A)(y_1)|\mathcal{F}_t],
\end{align}

where the equality \((2.11)\) follows from Proposition \([1]\) the relation \((2.12)\) is proven in Lemma \([4]\) the equality \((2.14)\) is a consequence of the commutativity of \(\mathbb{I} L\) and \(\hat{X}_i\) and the last one follows from Lemma \([3]\) The proof with the vector fields \((Y_i, i = 1, \ldots, n)\) is the same and omitted.

We have all we need to prove the Poincaré and log-Sobolev inequalities:

**Theorem 1.** For any smooth \(f\) on \(H_n\), we have

\[E[|f(x_1)| - E[f(x_1)]|^2] \leq \sum_{i=1}^{n} E[|X_i f(x_1)|^2 + |Y_i f(x_1)|^2].\]

**Proof:** From the Itô formula from the commutation results of Proposition \([2]\)

\[Q_{1-t} f(x_i) - Q_1 f(0) = \int_0^t X_i Q_{1-s} f(x_s) db_s^i + \int_0^t Y_i f(x_s) dw_s^i \]

\[(2.15) \quad = -\int_0^t Q_{1-s} \hat{X}_i(f \circ A)(y_s) db_s^i + \int_0^t Q_{1-s} \hat{Y}_i(f \circ A)(y_s) dw_s^i.\]

Again from Proposition \([2]\) the integrands in the equality \((2.15)\) are square integrable martingales provided that \(f\) has bounded first and second order derivatives. Hence all these terms as well as \((Q_{1-t} f(x_i), t \in [0,1])\) converge in \(L^2\) as \(t \to 1\). Consequently

\[E[|f(x_1)| - E[f(x_1)]|^2] = \lim_{t \to 1} \sum_{i=1}^{n} E[\int_0^1 [|Q_{1-s} \hat{X}_i(f \circ A)(y_s)|^2 + |Q_{1-s} \hat{Y}_i(f \circ A)(y_s)|^2] ds \]

\[\leq \sum_{i=1}^{n} E[|\hat{X}_i(f \circ A)(y_1)|^2 + |\hat{Y}_i(f \circ A)(y_1)|^2] \]

\[(2.16) \quad = \sum_{i=1}^{n} E[|X_i f(x_1)|^2 + |Y_i f(x_1)|^2].\]

where the line \((2.16)\) follows again from Proposition \([2]\) 

The last theorem will be the justification of this manuscript:

**Theorem 2.** For any smooth \(f\) on \(H_n\) with bounded derivatives up to the second order, the following inequality holds true:

\[E \left[ f^2(x_1) \log \frac{f^2(x_1)}{E[f^2(x_1)]} \right] \leq 2 \sum_{i=1}^{n} E[|X_i f(x_1)|^2 + |Y_i f(x_1)|^2].\]
Proof: We can assume that $f$ is positive and $E[f(x_1)] = 1$. In this case it suffices to prove, by replacing $f$ with its square root, that

$$E[f(x_1) \log f(x_1)] \leq \frac{1}{2} \sum_i E[\frac{1}{f(x_1)}(|X_i \log f(x_1)|^2 + |Y_i \log f(x_1)|^2)].$$

Thanks to our choice of $f$, we may look at $f(x_1)$ as a probability density and we write $d\nu = f(x_1)dP$ for the new probability. From the Itô formula and from the commutation relations given by Proposition $\text{[2]}$ employed as in the proof of Theorem $\text{[1]}$ we have

$$Q_{1-t}f(x_t) = 1 - \int_0^t Q_{1-s}X_i(f \circ A)(y_s)db^i_s - \int_0^t Q_{1-s}Y_i(f \circ A)(y_s)dw^i_s.$$

By dividing and multiplying each integrand above with $Q_{1-s}(f \circ A)(y_s)$, we can represent $Q_{1-t}f(x_t) = l_t$ as an exponential martingale: in fact let

$$u^i_t = \frac{1}{l_t} \left(Q_{1-t}X_i(f \circ A)(y_t), Q_{1-t}Y_i(f \circ A)(y_t)\right), i = 1, \ldots, n$$

and let

$$W^i_t = (b^i_t, w^i_t), i = 1, \ldots, n.$$

Then we can represent $l_t$ as

$$(2.17) \quad l_t = Q_{1-t}(f \circ A)(y_t) = \exp \left(-\int_0^t (u_s, dW_s)_{\mathbb{R}^2} - \frac{1}{2} \int_0^t |u_s|^2_{\mathbb{R}^2} ds \right).$$

From the Girsanov theorem, $(W_t + \int_0^t u_s ds, t \in [0, 1])$ is a $\nu$-Brownian motion, consequently

$$E[l_t \log l_t] = E_\nu[\log l_t] = \frac{1}{2} E_\nu \int_0^t |u_s|^2_{\mathbb{R}^2} ds$$

$$= \frac{1}{2} \sum_i E_\nu \int_0^t \frac{1}{P_s}(|Q_{1-s}X_i(f \circ A)(y_s)|^2 + |Q_{1-s}Y_i(f \circ A)(y_s)|^2) ds$$

$$= \frac{1}{2} \sum_i \left(E_\nu \int_0^t |E_\nu[X_i(f \circ A)(y_1)|F_s|^2 ds \right.$$

$$+ E_\nu \int_0^t |E_\nu[Y_i(f \circ A)(y_1)|F_s|^2 ds$$

$$\leq \frac{1}{2} \sum_i E_\nu[|X_i \log(f \circ A)(y_1)|^2 + |Y_i \log(f \circ A)(y_1)|^2]$$

$$= \frac{1}{2} \sum_i E \left[ \frac{1}{f(x_1)}(|\tilde{X}_i(f \circ A)(y_1)|^2 + |\tilde{Y}_i(f \circ A)(y_1)|^2) \right]$$

$$= \frac{1}{2} \sum_i E \left[ \frac{1}{f(x_1)}(|X_i f(x_1)|^2 + |Y_i f(x_1)|^2) \right],$$

where in the fourth and fifth lines we have used the fact that the map $t \to E[M_t^2]$ is increasing for a square integrable martingale $M = (M_t, t \in [0, 1])$ and the equality between the two last lines follow from the commutation relations proved in Proposition $\text{[2]}$. This inequality holds for any $t \in [0, 1]$, in particular it is also true when we replace the left hand side with $E[l_t \log l_t]$. \qed
References

[1] D. Bakry, F. Baudoin, M. Bonnefont, D. Chafai: “On gradient bounds for the heat kernel on the Heisenberg group”. Jour. Functional Analysis, 255, 1905-1938 (2008).

[2] C. Dellacherie et P.-A. Meyer: “Probabilités et Potentiel, Processus de Markov, Chapitres 17 à 24”. Collection Enseignement des Sciences, Hermann, 1992.

[3] B.K. Driver, T. Melcher: “Hypoelliptic heat kernel inequalities on the Heisenberg group”. Jour. Funct. Anal. 221 (2) 340–365 (2005).

[4] L. Gross: “Logarithmic Sobolev inequalities”. Amer. Jour. Math.97, p.1061-1083 (1975).

[5] H.-Q. Li: “Estimation optimale du gradient du semi-groupe de la chaleur sur le groupe de Heisenberg”. Jour. Funct. Anal. 236 (2) 369–394 (2006).

A. S. Üstünel, Bilkent University, Math. Dept., Ankara, Turkey
ustunel@fen.bilkent.edu.tr