The Annals of Probability
2002, Vol. 30, No. 2, 723–801

RANDOM WALKS ON DISCRETE GROUPS OF POLYNOMIAL VOLUME GROWTH

BY GEORGIOS K. ALEXOPOULOS

Université de Paris-Sud

Let μ be a probability measure with finite support on a discrete group Γ of polynomial volume growth. The main purpose of this paper is to study the asymptotic behavior of the convolution powers μ^{*n} of μ. If μ is centered, then we prove upper and lower Gaussian estimates. We prove a central limit theorem and we give a generalization of the Berry–Esseen theorem. These results also extend to noncentered probability measures. We study the associated Riesz transform operators. The main tool is a parabolic Harnack inequality for centered probability measures which is proved by using ideas from homogenization theory and by adapting the method of Krylov and Safonov. This inequality implies that the positive μ-harmonic functions are constant. Finally we give a characterization of the μ-harmonic functions which grow polynomially.

1. Introduction and statement of the results. Let Γ be a finitely generated discrete group of polynomial volume growth, let μ be a probability measure with finite support on Γ and let $\mu^{*n} = \mu * \mu * \cdots * \mu$ be the nth convolution power of μ.

The main purpose of this paper is to study the asymptotic behavior of μ^{*n}. We obtain generalizations of certain results concerning the lattice valued distributions in $\mathbb{R}^n$ (cf. [20, 34]). We shall also extend certain results of [2, 15, 23] to nonsymmetric probability measures.

The measure μ can be either centered or not centered. It turns out that if μ is not centered, then we can conjugate μ by a multiplicative function and obtain another centered measure μ'. So it is enough to consider only centered measures.

According to a famous theorem of Gromov [22], Γ is a finite extension of a nilpotent subgroup $\Gamma_N \triangleleft \Gamma$. By considering a subgroup of Γ_N if necessary we can assume that Γ_N can be embedded as a lattice in a simply connected nilpotent Lie group N. We can associate with μ a centered left invariant sub-Laplacian on N denoted by $L_{H\mu}$. $L_{H\mu}$ is defined by a formula similar to the one we have in classical homogenization theory (cf. [13, 26]).

Let $p_t^{H\mu}(x, y)$ be the heat kernel $L_{H\mu}$ [i.e., of the fundametal solution of the associated heat equation $(\frac{\partial}{\partial t} + L)u = 0$]. Comparing μ^{*n} with $p_t^{H\mu}(x, y)$, we can obtain information on the distribution of the mass of μ^{*n}. Using this information, together with a result of Varopoulos [49] which gives a uniform upper bound

Received September 1999; revised August 2001.
AMS 2000 subject classifications. 43A80, 60J15, 60B15, 20F65, 22E25, 22E30.
Key words and phrases. Random walk, group, convolution, Harnack inequality, heat kernel.

on μ^{*n}, it is possible to adapt the method of Krylov and Safonov (cf. [29, 30, 41]) and obtain a parabolic Harnack inequality.

Applying this inequality we can obtain upper and lower Gaussian bounds for μ^{*n}. We can also prove, adapting some ideas of Bergström (cf. [14, 41]), that as $n \to \infty$, the values of μ^{*n} approach the values of $p_t^{H\mu}(x, y)$ with uniform speed $1/n^{\gamma/2}$, for some $\gamma \in (0, 1]$. Of course by the classical Berry–Esseen theorem (cf. [20, 34]), the optimal rate of convergence is $1/\sqrt{n}$. This is proved with the same method a posteriori, once we have the appropriate estimate for the space differences of μ^{*n}.

The Berry–Esseen estimate implies that, on large balls, the μ-harmonic functions look like $L_{H\mu}$-harmonic functions. Using this observation, we can adapt some ideas of Avellaneda and Lin [3, 4, 8–10] and prove a Taylor formula for the μ-harmonic functions. This formula gives Harnack inequalities for the time and space differences of μ^{*n}. It can also be used to obtain a caracterization of the μ-harmonic functions which grow polynomially.

Finally, we prove Berry–Esseen estimates for the time and space differences of μ^{*n}. We apply these estimates to study the associated Riesz transform operators.

1.1. *Centered probability measures.* The group $\Gamma/[\Gamma, \Gamma]$ is finitely generated and Abelian and hence it can be written as a direct product $\mathbb{Z}^k \times A$, where A is finite and Abelian. Let π be the canonical projection $\pi : G \to \mathbb{Z}^k$ and let $H = \mathrm{Ker}(\pi)$.

Let μ be a probability measure on Γ whose support is finite and generates Γ. We say that μ is centered if the first order moments of its projection $\pi(\mu)$ on $\mathbb{Z}^k$ vanish, that is, if $\sum_{x\in\Gamma} \pi(x)_i \mu(x) = 0$, $1 \le i \le k$, where y_i is the ith coordinate of the element $y = (y_1, \dots, y_k) \in \mathbb{Z}^k$, $1 \le i \le k$.

1.2. *The passage from a noncentered to a centered probability measure.* We say that $\chi : \Gamma \to \mathbb{R}^+$ is multiplicative if $\chi(xy) = \chi(x)\chi(y)$, $x, y \in \Gamma$. Note that then χ can be written as

$$\chi = \phi \circ \pi \qquad \text{with } \phi(x) = e^{\langle b, x\rangle},$$

where π is the quotient map $\pi : G \to \mathbb{Z}^k \cong G/H$ and where $\langle b, x\rangle = b_1x_1 + \cdots + b_kx_k$ for $b = (b_1, \dots, b_k)$, $x = (x_1, \dots, x_k) \in \mathbb{R}^k$.

Let μ be a probability measure on Γ whose support is finite and generates Γ. We have the following well-known lemma.

LEMMA 1.1. *If μ is not centered, there are a multiplicative function χ, a constant $\beta_\mu > 0$ and another centered probability measure μ' on Γ such that*

$$\mu(x) = e^{-\beta_\mu} \mu'(x) \chi(x), \qquad x \in \Gamma. \tag{1.1}$$

Note that (1.1) implies that

$$\mu^{*n}(x) = e^{-\beta_\mu n}\mu'^{*n}(x)\chi(x), \qquad x \in \Gamma. \tag{1.2}$$

PROOF. Let $\pi(\mu)$ be the image of μ under the quotient map $\pi:\Gamma \to \mathbb{Z}^k \cong G/H$ and let us consider the function

$$F(a) = \sum_{x\in\mathbb{Z}^k} \pi(\mu)(x)e^{\langle a,x\rangle},$$

where $\langle a, x\rangle = a_1x_1 + \cdots + a_kx_k$ for $a = (a_1,\dots,a_k)$, $x = (x_1,\dots,x_k) \in \mathbb{R}^k$.

We observe that F is a positive smooth function on $\mathbb{R}^k$ and that $F(a) \to \infty$ as $|a| \to \infty$. So F attains its minimum $b_\varphi = \min\{F(a) : a \in \mathbb{R}^k\}$ at some point $a_0 \in \mathbb{R}^k$. Also $F(0) = 1$ and, since φ is not centered, $\nabla F(0) \neq 0$. Hence $b_\varphi = F(a_0) < 1$ and $a_0 \neq 0$. The lemma follows by taking

$$\beta_\mu = -\log b_\mu, \qquad \chi(x) = e^{-\langle a_0,\pi(x)\rangle}$$

and

$$\mu'(x) = \frac{1}{b_\mu}\mu(x)e^{\langle a_0,\pi(x)\rangle}, \qquad x \in \Gamma.$$

The fact that $\nabla F(a_0) = 0$ implies that μ' is indeed a centered probability measure. □

1.3. *The geometry of* Γ *and the sub-Laplacian* $L_{H\mu}$. Let us fix a subset U of Γ such that the following hold:

1. U is finite and generates Γ;
2. $e \in U$ (e is the identity element of Γ);
3. U is symmetric; that is, $x \in U$ if and only if $x^{-1} \in U$.

Let $U^n = \{x_1x_2\cdots x_n : x_i \in U, 1 \le i \le n\}$ and set

$$|x|_\Gamma = \min\{n : x \in U^n\}.$$

Also, let $|A|$ denote the number of elements of $A \subseteq \Gamma$.

In this article we assume that Γ has polynomial volume growth, that is, that there are constants $c > 0$ and $A \in \mathbb{N}$ such that $|U^n| \le cn^A$, for all $n \in \mathbb{N}$. By a theorem of Gromov [22], this assumption implies that there is a nilpotent subgroup $\Gamma_N \triangleleft \Gamma$ such that $|\Gamma/\Gamma_N| < \infty$. Hence, by a theorem of Bass [11], there is an integer $D \ge 0$ such that

$$\frac{1}{c}n^D \le |U^n| \le cn^D, \qquad n \in \mathbb{N}. \tag{1.3}$$

We call D the homogeneous dimension of Γ. Note that D does not depend on the choice of U.

Let π the quotient map $\pi : \Gamma \to \Gamma/\Gamma_N$ and let us choose elements $g_0 = e$, $g_1, \ldots, g_k \in \Gamma$ such that

$$\Gamma/\Gamma_N = \{\pi(g_0), \pi(g_1), \ldots, \pi(g_k)\}.$$

Every element $g \in \Gamma$ can be written uniquely as $g = yg_j$, with $y \in \Gamma_N$, $0 \le j \le k$. We set

$$\overline{g} = g_j \quad \text{and} \quad g_N = y.$$

Γ_N has a torsion-free subgroup $\Gamma_N^1 \lhd \Gamma_N$ of finite index, that is, such that $|\Gamma_N/\Gamma_N^1| < \infty$ (cf. [35]). Let $\Gamma_N^2 = \bigcap_{0 \le i \le k} g_i \Gamma_N^1 g_i^{-1}$. Then Γ_N^2 is still nilpotent and torsion free. Furthermore, $\Gamma_N^2 \lhd \Gamma$ and $|\Gamma/\Gamma_N^2| < \infty$. So by replacing Γ_N with Γ_N^2, if necessary, we assume that Γ_N has the following properties:

1. $\Gamma_N \lhd \Gamma$;
2. $|\Gamma/\Gamma_N| < \infty$;
3. Γ_N is finitely generated, nilpotent and torsion free.

Let $U_{\Gamma_N} \subseteq \Gamma_N$ be a finite and symmetric subset which generates Γ_N and set

$$|x|_{\Gamma_N} = \min\{n \in \mathbb{N} : x \in U_{\Gamma_N}^n\}, \qquad x \in \Gamma_N. \tag{1.4}$$

Then there is a $c \ge 1$ such that

$$\frac{1}{c}|x|_{\Gamma_N} \le |xg_i|_\Gamma \le c|x|_{\Gamma_N} \tag{1.5}$$

for all $x \in \Gamma_N$ and $0 \le i \le k$, or more generally

$$\frac{1}{c}|x^{-1}y|_{\Gamma_N} \le |g^{-1}h|_\Gamma \le c|x^{-1}y|_{\Gamma_N} \tag{1.6}$$

for all $g = xg_i, h = yg_j \in \Gamma,\ x, y \in \Gamma_N,\ 0 \le i, j \le k$.

Property 3 above implies that Γ_N is isomorphic to (and hence can be identified with) a uniform lattice in a simply connected nilpotent Lie group N (cf. [35]). Note that N/Γ_N is a compact neighborhood. Let us fix a fundamental domain Ω for Γ_N and let dg be the Haar measure on N which satisfies dg-measure$(\Omega) = \mathrm{vol}(N/\Gamma_N) = 1$.

Let V be a compact neighborhood of the identity element e of N and set

$$|x|_N = \min\{n \in \mathbb{N} : x \in V^n,\ x \in N\}. \tag{1.7}$$

Then there is a $c \ge 1$ such that

$$\frac{1}{c}|x|_{\Gamma_N} \le |x|_N \le c|x|_{\Gamma_N}, \qquad x \in \Gamma_N. \tag{1.8}$$

The isomorphisms $y \to g_i y g_i^{-1}$, $0 \le k \le k$, can be extended to isomorphisms of N (cf. [35]). So we can consider the group

$$G = \{yg_i,\ y \in N,\ 0 \le i \le k\}$$

with multiplication law defined by

$$xg_i yg_j = xg_i yg_i^{-1}(g_i g_j)_N \overline{g_i g_j}, \qquad x, y \in N,\ 0 \le i, j \le k.$$

If Γ is nilpotent, then a better way to proceed is to consider the torsion subgroup $\tau(\Gamma)$ of Γ (cf. [12]). $\tau(\Gamma)$ is the set of elements of finite order in Γ, it is a normal subgroup of Γ and $\Gamma/\tau(\Gamma)$ is torsion free. So, we can set $\Gamma_N = \Gamma/\tau(\Gamma)$.

Let $\mathfrak{n}$ be the Lie algebra of N. We identify $\mathfrak{n}$ with the left invariant vector fields on N.

By a left invariant sub-Laplacian on N, we mean an operator

$$L = -(E_1^2 + \cdots + E_p^2) + E_0,$$

where $E_0, E_1, \ldots, E_p$ are left invariant vector fields on N and where the vector fields $E_1, \ldots, E_p$ satisfy Hörmander's condition; that is, they generate together with their successive Lie brackets $[E_{i_1}, [E_{i_2}, [\ldots, E_{i_k}]\ldots]]]$, $1 \le i_j \le p$, $1 \le j \le k$, the Lie algebra $\mathfrak{n}$ of N.

We shall say that L is centered if $E_0 \in [\mathfrak{n}, \mathfrak{n}]$.

Let us fix a discrete probability measure μ on Γ, let $\operatorname{supp}(\mu) = \{g \in G : \mu(g) > 0\}$ and let us assume that the following hold:

1. $|\operatorname{supp}(\mu)| < \infty$;
2. $U \subseteq \operatorname{supp}(\mu)$;
3. μ is centered.

Our goal is to associate with μ a centered left invariant sub-Laplacian $L_{H\mu}$ on N, in such a way that the asymptotic behavior of the convolution powers μ^{*n} can be compared to the large-time behavior of the heat kernel $p_t^{H\mu}(x, y)$ of $L_{H\mu}$. If $\Gamma = \Gamma_N$ or G is nilpotent, then the definition of $L_{H\mu}$ is rather straightforward (cf. [19]). In this case we use the notation L_μ and $p_t^{\mu}(x, y)$ instead of $L_{H\mu}$ and $p_t^{H\mu}(x, y)$, respectively. If Γ is not nilpotent, then $L_{H\mu}$ is defined by a method inspired by the theory of the homogenization (cf. [13, 26]). We call $L_{H\mu}$ the homogenized sub-Laplacian (associated with μ).

1.4. *Notation.* Given another measure ν we define the convolution $\mu * \nu$ by $\mu * \nu(x) = \sum_{y \in \Gamma} \mu(y)\nu(y^{-1}x)\, dy,\ x \in \Gamma$.

Given a kernel $K(x, y)$ we set

$$K(x, A) = \sum_{y \in A} K(x, y) \quad \text{and} \quad Kf(x) = \sum_{y \in \Gamma} K(x, y) f(y).$$

If $S(x, y)$ is another kernel, then we denote by KS the kernel

$$KS(x, y) = \sum_{z \in \Gamma} K(x, z) S(z, y).$$

We also set

$$\|K\|_1 = \sup\{\|K(x,.)\|_1, \|K(.,y)\|_1;\ x, y \in \Gamma\},$$
$$\|K\|_\infty = \sup\{|K(x,y)| : x, y \in \Gamma\}.$$

To simplify the notation, we set $\mu^{*n} = \mu^n$, $n \in \mathbb{N}$, and $\mu^0 = \delta_e$, where δ_x is the Dirac mass at x. We also denote by μ^n the kernel

$$\mu^n(x,y) = \mu^{*n}(x^{-1}y), \qquad x, y \in \Gamma.$$

For $n = 1$, we just write $\mu(x,y)$ instead of $\mu^1(x,y)$.

Note that $\mu^{n+1} = \mu\mu^n$ and that

$$\mu^n f(x) = \sum_{y\in\Gamma} \mu^n(x,y) f(y) = \sum_{y\in\Gamma} \mu^{*n}(y) f(xy).$$

We say that a function u is μ-harmonic in $A \subseteq \Gamma$ if $\mu u(x) = u(x)$, $x \in A$.

We say that a function u is a space–time μ-harmonic function in $A \subseteq \mathbb{Z} \times \Gamma$ if $(\mu u(n,\cdot))(x) = u(n+1,x)$, $(n,x) \in A$.

We denote by ∂_1 and ∂_z, $z \in \Gamma$, respectively the difference operators

$$\partial_1 u(n,x) = u(n+1,x) - u(n,x) \quad \text{and} \quad \partial_z u(n,x) = u(n,xz) - u(n,x).$$

Note that u is a space–time μ-harmonic function if and only if $(\partial_1 + (I-\mu))u = 0$.

We also set

$$\partial_z \mu^{*n}(x) = \mu^{*n}(xz) - \mu^{*n}(x) \quad \text{and} \quad \partial_1 \mu^{*n}(x) = \mu^{*(n+1)}(x) - \mu^{*n}(x).$$

If $A \subseteq \Gamma$, then we set

$$\nabla_A u(n,x) = \sup\{|\partial_z u(n,x)|;\ z \in U\}.$$

We say that a function f is of type P if $f(xg) = f(g)$, $x \in \Gamma_N$, $g \in \Gamma$.

If f is such a function, then we denote by $\langle f\rangle$ its mean value

$$\langle f\rangle = \frac{1}{k+1} \sum_{0\le i\le k} f(g_i).$$

Note that if f is a function of type P, then μf is also a function of type P. If we also have $\langle f\rangle = 0$, then the function

$$u = \sum_{n\ge 0} \mu^n f$$

is well defined and satisfies

$$(I-\mu)u = f.$$

If Γ is nilpotent and we set $\Gamma_N = \Gamma/\tau(\Gamma)$, then the type P functions will just be the constant functions.

If $K(x, y)$ is a kernel initially defined on N, then we use the same notation $K(x, y)$ to denote its restriction to Γ_N and its extension to Γ. The extension of $K(x, y)$ to Γ is defined by

$$K(xg_i, yg_j) = \frac{1}{k+1} K(x, y), \qquad x, y \in N,\ 0 \le i, j \le k.$$

If Γ is nilpotent and we set $\Gamma_N = \Gamma/\tau(\Gamma)$, then we extend $K(x, y)$ to Γ by setting

$$K(zx, wy) = \frac{1}{|\tau(\Gamma)|} K(\dot{x}, \dot{y}),$$

where $z, w \in \tau(\Gamma)$ and $\dot{x} = x\tau(\Gamma)$, $\dot{y} = y\tau(\Gamma)$, $x, y \in \Gamma$.

In particular, we use this notation for the heat kernels $p_t^{H\mu}(x, y)$, $p_t^{\mu}(x, y)$ and their derivatives $X_1 \cdots X_n p_t^{H\mu}(x, y)$, $X_1 \cdots X_n p_t^{\mu}(x, y)$, $X_1, \dots, X_n \in \mathfrak{n}$.

A function f on Γ_N will be extended to Γ by setting

$$f(xg_i) = f(x), \qquad x \in \Gamma_N,\ 0 \le i \le k.$$

If Γ is nilpotent and we set $\Gamma_N = \Gamma/\tau(\Gamma)$, then we extend f to Γ by setting $f(zx) = f(\dot{x})$, for $z \in \tau(\Gamma)$ and $\dot{x} = x\tau(\Gamma)$, $x \in \Gamma$.

We do this, in particular, when f is a harmonic function or a polynomial.

We denote by $[\![a, b]\!]$ the interval $[a, b] \cap \mathbb{Z}$.

Given a nonempty subset A of Γ we set $A^0 = \{e\}$ and $A^r = A^{[r]}$, $r > 0$.

The different constants are always denoted by the same letter c. When their dependence or independence is significant, it is clearly stated.

1.5. *A parabolic Harnack inequality.* The following Harnack inequality plays a central role in this article.

THEOREM 1.2. *For all $a, b \ge 1$ there are $\beta > \alpha > 1$, $c > 1$ and $\lambda > 0$ such that, for all $r \ge 1$ and all $u \ge 0$ satisfying*

$$\big(\partial_1 + (I - \mu)\big)u = 0 \qquad \text{in } [\![0, (\beta + b^2)r^2]\!] \times U^{cr},$$

we have

$$\sup\{u; [\![\alpha r^2, (\alpha + a^2)r^2]\!] \times U^{ar}\} \le \lambda \inf\{u; [\![\beta r^2, (\beta + b^2)r^2]\!] \times U^{br}\}. \tag{1.9}$$

For the case when μ is symmetric [i.e., $\mu(x^{-1}) = \mu(x)$, $x \in \Gamma$], the above inequality was proved in [23] by a different method.

We prove this inequality by adapting the method of Krylov and Safonov [29]. This method uses certain information on the growth of the positive space–time μ-harmonic functions. To obtain this information we use the following three results.

The first two results concern the distribution of the mass of the convolution powers μ^{*n} as $n \to \infty$:

PROPOSITION 1.3. *For all $a > 1$ there are $r_0 \geq 1$ and $\partial > 0$ such that*

$$\sum_{y \in U^r} \mu^n(x, y) > \partial, \tag{1.10}$$

for all $(n, x) \in [\![a^{-2}r^2, a^2r^2]\!] \times U^{ar}$, $r \geq r_0$.

PROPOSITION 1.4. *For all $\partial > 0$ there is an $a > 1$ such that*

$$\sum_{y \notin U^{ar}} \mu^n(e, y) < \partial, \tag{1.11}$$

for all $k \in [\![1, r^2]\!]$.

The third result is a theorem of Varopoulos [49] which asserts that the convolution powers μ^{*n} decay with a certain uniform speed as $n \to \infty$:

THEOREM 1.5 [49]. *Let μ be a (not necessarily centered) probability measure whose support generates Γ. Then there is a constant $c > 0$ such that*

$$\|\mu^{*n}\|_\infty \leq c\, n^{-D/2}, \qquad n \in \mathbb{N}. \tag{1.12}$$

If u is a function defined on $B \subseteq \mathbb{Z} \times \Gamma$, then let us set

$$\operatorname{Osc}(u, B) = \sup\{|u(k, x) - u(m, y)| : (k, x), (m, y) \in B\}.$$

To prove Theorem 1.2, we proceed as follows. Using the above three results, we prove an analogue of the first growth lemma of [29]. From this and arguing in the same way as in [29] we obtain a second growth lemma.

A direct consequence of the second growth lemma is the following:

PROPOSITION 1.6. *There are $c \geq 1$ and $\gamma \in (0, 1]$ such that, for all $t \in \mathbb{R}$, $r \geq 1$ and all functions u satisfying $(\partial_1 + (I - \mu))u = 0$ in $[\![t - c^2r^2, t]\!] \times U^{cr}$,*

$$\operatorname{Osc}(u, [\![t - r^2, t]\!] \times U^r) \leq \gamma \operatorname{Osc}(u, [\![t - c^2r^2, t]\!] \times U^{cr}). \tag{1.13}$$

Theorem 1.2 follows from the second growth lemma and the above proposition by a standard argument (see, e.g., [7, 30, 40, 44]). It can also be proved by arguing in the same way as in [29] (but this is less obvious).

An immediate consequence of (1.9) and (1.13) is the following:

COROLLARY 1.7. *Every positive μ-harmonic function is constant.*

1.6. *Gaussian estimates.* Making use of the Harnack inequality (1.9) we prove the following upper Gaussian estimate:

THEOREM 1.8. *There is a constant $c > 0$ such that*

$$\mu^{*n}(x) \leq c\, n^{-D/2} \exp\left(-\frac{|x|_\Gamma^2}{cn}\right), \qquad x \in \Gamma,\ n \in \mathbb{N}. \tag{1.14}$$

Once we have the upper Gaussian estimate then, again by using the Harnack inequality (1.9), we can obtain a lower Gaussian estimate (cf. [23] and [50], pages 47–50):

COROLLARY 1.9. *There is a constant $c > 0$ such that*

$$\mu^{*n}(x) \geq \frac{1}{c}\, n^{-D/2} \exp\left(-c\frac{|x|_\Gamma^2}{n}\right) \tag{1.15}$$

for all $n \in \mathbb{N}$ and $x \in \Gamma$ satisfying $|x|_\Gamma \leq n/c$.

Combining (1.2) and (1.14) we have the following:

COROLLARY 1.10. *Let us assume that μ is not centered and let β_μ and χ be as in Lemma* 1.1. *Then there is a $c > 0$ such that*

$$\mu^{*n}(x) \leq cn^{-D/2} \exp(-\beta_\mu n)\chi(x) \exp\left(-\frac{|x|_\Gamma^2}{cn}\right) \tag{1.16}$$

for all $x \in \Gamma$ and $n \in \mathbb{N}$.

1.7. *A Taylor formula for the space–time μ-harmonic functions.* Using the exponential coordinates we shall identify N, as a differential manifold, with $\mathbb{R}^q$. So a monomial $P(x)$ on N will be just a monomial on $\mathbb{R}^q$. A monomial $P(x)$ on Γ_N will be just the restriction to Γ_N of a monomial $P(x)$ on N. We extend the monomials $P(x)$ to Γ by setting $P(xg_i) = P(x),\ x \in \Gamma_N, 1 \leq i \leq k$.

In the rest of this article, we do not make any distinction between the restriction of a monomial $P(x)$ to Γ_N and its extension to Γ.

For every monomial $P(x)$, there are an integer $d \geq 0$ and a constant $c > 0$ such that

$$\frac{1}{c}n^d \leq \sup\{|P(x)|,\ x \in U^n\} \leq cn^d, \qquad n \in \mathbb{N}. \tag{1.17}$$

We say then that $P(x)$ has homogeneous degree $\deg_H P = d$.

We say that $P(t,x)$ is a monomial on $\mathbb{Z} \times \Gamma$ (resp. $\mathbb{R} \times G$) if $P(t,x) = t^m Q(x)$, with $Q(x)$ a monomial on Γ (resp. G). We define the homogeneous degree $\deg_H P(t,x)$ of $P(t,x)$ by

$$\deg_H P(t,x) = 2m + \deg_H Q(x).$$

By polynomials we of course mean linear combinations of monomials. The homogeneous degree of a polynomial is therefore the maximum of the homogeneous degrees of its monomials.

To fix the notation, we use

$$P_0(t,x),\, P_2(t,x), \ldots, P_{\nu_d}(t,x)$$

from now on to denote the monomials with homogeneous degree less than or equal to d. With every such monomial $P_i(t,x)$ we associate another, more convenient "corrected" monomial $Q^{\psi}_{P_i}(t,x)$ written as

$$Q^{\psi}_{P_i}(t,x) = P_i(t,x) + \sum_{0\le j\le \nu_{k-1}} \psi^i_j(x) P_j(t,x),$$

where $k = \deg_H P_i$ and where the functions ψ^i_j are of type P.

Note that when Γ is nilpotent the ψ^i_j will just be constant functions.

The following result gives a Taylor formula for the space–time μ-harmonic functions.

THEOREM 1.11. *For all $n \in \mathbb{N}$ there is a constant $c_n > 0$ such that, for all $R \ge r \ge 1$ and all functions u satisfying*

$$\big(\partial_1 + (I-\mu)\big)u = 0 \qquad \text{in } [\![-R^2, R^2]\!] \times U^R,$$

we have

$$\sup\left\{\left|u - \sum_{0\le i\le \nu_n} A_i R^{-\deg_H P_i} Q^{\psi}_{P_i}\right|; [\![r^2, r^2]\!] \times U^r\right\} \le c_n \left(\frac{R}{r}\right)^{-(n+1)} \|u\|_\infty, \tag{1.18}$$

where the constants A_i satisfy

$$|A_i| \le c_n \|u\|_\infty,$$

for all $0 \le i \le \nu_n$, and

$$\big(\partial_1 + (I-\mu)\big)\left(\sum_{\nu_{d-1} < i \le \nu_d} A_i Q^{\psi}_{P_i}\right) = 0,$$

for all $1 \le d \le n$.

The proof of the above result is based on ideas of Avellaneda and Lin (cf. [9, 10]). These ideas have already been used in the context of Lie groups in [3, 4, 8]. The interest of the method lies in the fact that we do not make use of any a priori control on the differences.

1.8. *μ-harmonic functions of polynomial growth.* We say that a function u on Γ grows polynomially if there is a $c > 0$ such that

$$\sup\{|u|; U^n\} \le cn^c, \qquad n \in \mathbb{N}. \tag{1.19}$$

The following result is a consequence of Theorem 1.11:

THEOREM 1.12. *Every μ-harmonic function u which grows polynomially is equal to a linear combination of the monomials $Q^{\psi}_{P_i}$.*

A result of this type was first proved by Avellaneda and Lin [10] in the case of differential operators with periodic coefficients in $\mathbb{R}^n$. It was generalized in [8] in the context of connected Lie groups of polynomial volume growth, where it was used to prove a Sobolev inequality. We state below the discrete analogue of that inequality. The proof is similar and is omitted.

COROLLARY 1.13. *Let f be a function, not necessarily with compact support, such that $\nabla_U f \in L^p$, $1 < p < D$. Then there are a universal constant $c > 0$ and a constant C_f depending on f such that $f - C_f \in L^{Dp/(D-p)}$, and*

$$\|f - C_f\|_{Dp/(D-p)} \le c\|\nabla_U f\|_p.$$

1.9. *Harnack inequalities for the differences.* A consequence of Theorem 1.11 is the following result:

THEOREM 1.14. *For all $a, b \ge 1$ and all $k \in \mathbb{N}$ there are $\beta > \alpha > 1$, $c > 1$ and $\lambda > 0$ such that, for all $z \in U$, $r \ge 1$ and all $u \ge 0$ satisfying*

$$\big(\partial_1 + (I - \mu)\big)u = 0 \qquad \textit{in } [\![0, cr^2]\!] \times U^{cr},$$

we have

$$\begin{aligned} &\sup\big\{|\partial_1^k \partial_z u|; [\![\alpha r^2, (\alpha + a^2)r^2]\!] \times U^{ar}\big\} \\ &\qquad \le \lambda r^{-2k-1} \inf\big\{u; [\![\beta r^2, (\beta + b^2)r^2]\!] \times U^{br}\big\}. \end{aligned} \tag{1.20}$$

If Γ is nilpotent, then we can also control higher order spacial differences:

THEOREM 1.15. *If $\Gamma = \Gamma_N$, then for all $a, b \ge 1$ and all $k, m \in \mathbb{N}$ there are $\beta > \alpha > 1$, $c > 1$ and $\lambda > 0$ such that, for all $z_1, \dots, z_m \in U$, $r \ge 1$ and all $u \ge 0$ satisfying*

$$\big(\partial_1 + (I - \mu)\big)u = 0 \qquad \textit{in } [\![0, cr^2]\!] \times U^{cr},$$

we have

$$\begin{aligned} &\sup\big\{|\partial_1^k \partial_{z_1} \cdots \partial_{z_m} u|; [\![\alpha r^2, (\alpha + a^2)r^2]\!] \times U^{ar}\big\} \\ &\qquad \le \lambda r^{-2k-m} \inf\big\{u; [\![\beta r^2, (\beta + b^2)r^2]\!] \times U^{br}\big\}. \end{aligned} \tag{1.21}$$

Note that if Γ is not nilpotent, then (1.21) is not necessarily true for $m \geq 2$. This is due to the existence of the functions ψ_j^i in the definition of the monomials $Q_{P_i}^{\psi}$ [see Sections 1.7 and 19 as well as (1.30) below].

1.10. *Berry–Esseen estimates.* Let $p_t^{H\mu}(x, y)$ be the heat kernel of the homogenized sub-Laplacian $L_{H\mu}$ associated with μ. As mentioned in Section 1.4, we extend $p_t^{H\mu}(x, y)$ to G, by setting

$$(1.22) \qquad p_t^{H\mu}(xg_i, yg_j) = \frac{1}{k+1} p_t^{H\mu}(x, y), \qquad x, y \in N,\ 0 \leq i, j \leq k.$$

We have the following analogue of the Berry–Esseen theorem (cf. [20, 26, 28, 34, 51]):

THEOREM 1.16. *There is a* $c > 0$ *such that*

$$(1.23) \qquad |\mu^n(x, y) - p_n^{H\mu}(x, y)| \leq cn^{-(D+1)/2}$$

for all $x, y \in \Gamma$ *and* $n \in \mathbb{N}$.

The reader can observe that by a straightforward adaptation of the proof of the above result we can also obtain a similar L^1 estimate, that is, that there is a $c > 0$ such that

$$\|\mu^n - p_n^{H\mu}\|_1 \leq c/\sqrt{n}, \qquad n \in \mathbb{N}.$$

It was proved in [2, 7] that there are constants $c, C_L \geq 0$ such that

$$(1.24) \qquad |p_t^{H\mu}(e, e) - C_{L_{H\mu}} t^{-D/2}| \leq ct^{-(D+1)/2}, \qquad t \geq 1.$$

Combining (1.23) and (1.24) we have the following:

COROLLARY 1.17. *There are constants* $c, C_\mu \geq 0$ *such that*

$$(1.25) \qquad |\mu^{*n}(e) - C_\mu n^{-D/2}| \leq cn^{-(D+1)/2}, \qquad n \in \mathbb{N}.$$

Combining (1.2) and (1.25) we have the following:

COROLLARY 1.18. *Let us assume that* μ *is not centered and let* β_μ *be as in Lemma* 1.1. *Then there are constants* $c, C_\mu \geq 0$ *such that*

$$(1.26) \qquad |\mu^{*n}(e) - C_\mu n^{-D/2} e^{-\beta_\mu n}| \leq cn^{-(D+1)/2} e^{-\beta_\mu n}, \qquad n \in \mathbb{N}.$$

By interpolating (1.14) and (1.22) we can have the following:

COROLLARY 1.19. *For all* $\varepsilon \in (0,1)$ *there is a* $c > 0$ *such that*

$$|\mu^n(x,y) - p_n^{H\mu}(x,y)| \le cn^{-(D+\varepsilon)/2} \exp\left(-\frac{|x^{-1}y|_\Gamma^2}{cn}\right) \tag{1.27}$$

for all $x, y \in \Gamma$ *and* $n \in \mathbb{N}$.

Concerning the space and time differences we have the following results:

THEOREM 1.20. *There is a constant* $c > 0$ *such that*

$$|\partial_1 \mu^n(x,y) - \partial_1 p_n^{H\mu}(x,y)| \le cn^{-(D+3)/2} \tag{1.28}$$

for all $x, y \in \Gamma$ *and* $n \in \mathbb{N}$.

THEOREM 1.21. *If* Γ *is nilpotent, then there is a constant* $c > 0$ *such that, for all* $z \in U$ *and all* $x, y \in \Gamma$ *and* $n \in \mathbb{N}$,

$$|\partial_z \mu^n(x,y) - \partial_z p_n^{\mu}(x,y)| \le cn^{-(D+2)/2}. \tag{1.29}$$

If Γ is not nilpotent, then the situation is quite different. More precisely, if $X_1, \dots, X_q$ is a convenient basis of $\mathfrak{n}$ and if $\psi^1, \dots, \psi^{n_1}$ are the associated first order correctors (see Section 14 for the exact definitions), then we have the following result:

THEOREM 1.22. *There is a* $c > 0$ *such that, for all* $z \in U$ *and all* $x, y \in \Gamma$ *and* $n \in \mathbb{N}$,

$$\begin{aligned} &\left| \partial_z \mu^n(x,y) - \partial_z p_n^{H\mu}(x,y) - \sum_{1 \le j \le n_1} \big(\partial_z \psi^j(x)\big) X_j p_n^{H\mu}(x,y) \right| \\ &\qquad \le cn^{-(D+2)/2}. \end{aligned} \tag{1.30}$$

Combining (1.14) and (1.30) we have the following:

COROLLARY 1.23. *For all* $\varepsilon \in (0,1)$ *there is a constant* $c > 0$ *such that, for all* $z \in U$ *and all* $x, y \in \Gamma$ *and* $n \in \mathbb{N}$,

$$\begin{aligned} &\left| \partial_z \mu^n(x,y) - \partial_z p_n^{H\mu}(x,y) - \sum_{1 \le j \le n_1} \big(\partial_z \psi^j(x)\big) X_j p_n^{H\mu}(x,y) \right| \\ &\qquad \le cn^{-(D+1+\varepsilon)/2} \exp\left(-\frac{|x^{-1}y|_\Gamma^2}{cn}\right). \end{aligned} \tag{1.31}$$

Inequalities (1.27) and (1.31) above actually hold with $\varepsilon = 1$. This can be proved by arguing in a similar way as in the proofs of Theorems 1.16 and 1.22. However, the proofs become much more technical, while (1.27) and (1.31) are sufficient for the application that we have in mind, namely the proof of Theorem 1.24.

1.11. *Riesz transforms.* Let us denote by $(I-\mu)^{-1/2}$ the operator defined by $(I-\mu)^{-1/2}=\sum_{n\geq 0} a_n\mu^n$, where the a_n's are as in $(1-t)^{-1/2}=\sum_{n\geq 0} a_n t^n$.

THEOREM 1.24. *For all $z\in\Gamma$ the Riesz transform operators $R_z=\partial_z(I-\mu)^{-1/2}$ and $R_z^*=(I-\mu)^{-1/2}\partial_z$ are bounded on L^p, for $1<p<\infty$ and from L^1 to weak-L^1.*

If Γ is nilpotent, then we can also consider higher order Riesz transforms.

THEOREM 1.25. *If Γ is nilpotent, then for all $z_1,\dots,z_k\in\Gamma$ the Riesz transform operators $R_k=\partial_{z_1}\cdots\partial_{z_k}(I-\mu)^{-k/2}$ and $R_k^*=(I-\mu)^{-k/2}\partial_{z_k}\cdots\partial_{z_1}$, are bounded on L^p, for $1<p<\infty$ and from L^1 to weak-L^1 .*

If Γ is not nilpotent then, as we can see from (1.31), the second order Riesz transforms $R_2=\partial_{z_1}\partial_{z_2}(I-\mu)^{-1}$ and $R_2^*=(I-\mu)^{-1}\partial_{z_2}\partial_{z_1}$ may be unbounded even on L^2 (cf. [3]).

2. Organization of the article. We have tried to give the proof of the results in the simplest possible context. The proof of the parabolic Harnack inequality (1.9) from Varopoulos's theorem (Theorem 1.5) and by assuming Propositions 1.3 and 1.4, does not use any particular result from the structure of Γ and so it is given already in Section 3. The construction of the operator $L_{H\mu}$ is much simpler when $\Gamma=\Gamma_N$. So those proofs that are essentially the same, whether $\Gamma=\Gamma_N$ or not, are only given in the case $\Gamma=\Gamma_N$. This is the case for the Gaussian estimate (1.14), the Taylor formula (1.18) and the main part of the proof of Propositions 1.3 and 1.4.

The proof of the Berry–Esseen estimate (1.23) is much more complicated when $\Gamma\neq\Gamma_N$. So, to illustrate the ideas better, we also give the proof in the case $\Gamma=\Gamma_N$.

3. The proof of the Harnack inequality from Varopoulos's theorem and Propositions 1.3 and 1.4. In this section we give the proof of Theorem 1.2 from Varopoulos's theorem (Theorem 1.5) and by assuming Propositions 1.3 and 1.4. This has already been done in [7] in the context of left invariant sub-Laplacians on connected Lie groups of polynomial volume growth. We give below an adaptation of that proof in the context of discrete groups.

We first prove an analogue of the first growth lemma of [29] by using (1.10), (1.11) and (1.12).

Next, we prove an analogue of the second growth lemma of [29]. To do this, we follow closely [29] and we adapt in our context their covering lemmas.

The proof of Proposition 1.6 and of Theorem 1.2 from the second growth lemma is standard in the literature (cf. [7, 30, 40, 44]; since it is also long, it will be omitted. We point out again that the argument given in [29] can also be used.

If $A \subseteq \mathbb{Z} \times \Gamma$, then we denote by $|A|$ the number of its elements.

If $A \subseteq \mathbb{R} \times \Gamma$, then we set

$$|A| = \sum_{x \in \Gamma} |A_x|,$$

where $|A_x|$ is the Lebesgue measure of $A_x = A \cap \mathbb{R} \times \{x\}$.

3.1. *The first growth lemma.* If $B \subseteq \mathbb{Z} \times \Gamma$, $A \subseteq B$ and $(t, x) \in B$ then, adopting the notation of [29], we set

$$\Psi\big((t,x), A, B\big) = \inf\big\{u(t,x) : u \geq 0,\ u(s,y) \geq 1 \text{ for } (s,y) \in A \\ \text{and } \big(\partial_1 + (I - \mu)\big)u = 0 \text{ in } B\big\}.$$

If $A' \subseteq B$, then we set

$$\Psi(A', A, B) = \inf\big\{\Psi\big((t,x), A, B\big) : (t,x) \in A'\big\}.$$

Note that if $v \geq 0$, $u = \mu v$ and $a = \min\{\mu(x), x \in U\}$, then

$$u(x) \geq a v(xy), \qquad y \in U. \tag{3.1}$$

LEMMA 3.1 (First growth lemma). *For all $a > 1$, there are $r_0 \geq 1$, $c > a$ and $\partial, \xi \in (0,1)$ such that*

$$\Psi\big([\![a^{-2}r^2, a^2r^2]\!] \times U^{ar}, A, [\![0, a^2r^2]\!] \times U^{cr}\big) > \delta \tag{3.2}$$

for all $r \geq r_0$ and every $A \subseteq [\![0, r^2]\!] \times U^r$ satisfying

$$|A| > \xi |[\![0, r^2]\!] \times U^r|.$$

An immediate consequence of (3.1) and (3.2) is the following:

COROLLARY 3.2. *For all $a > 1$ there are $c > a$, $r_0 \geq 1$, $m \in \mathbb{N}$ and $\partial > 0$ such that for all $r \geq r_0$ and all $u \geq 0$ satisfying*

$$\big(\partial_1 + (I - \Phi)\big)u = 0 \qquad \textit{in } [\![0, a^2r^2]\!] \times U^{cr}$$

we have

$$\inf\big\{u; [\![a^{-2}r^2, a^2r^2]\!] \times U^{ar}\big\} \geq \delta u(1, e) r^{-m}. \tag{3.3}$$

Moreover, if for some $1 \leq R \leq r$,

$$\inf\big\{u; [\![0, R^2]\!] \times U^R\big\} \geq 1,$$

then

$$\inf\big\{u; [\![a^{-2}r^2, a^2r^2]\!] \times U^{ar}\big\} \geq \delta \left(\frac{R}{r}\right)^m. \tag{3.4}$$

3.2. *Proof of Lemma* 3.1. The following lemma is an immediate consequence of Theorem 1.5 and Proposition 1.3.

LEMMA 3.3. *For all $a > 1$ there are $r_0 \geq 1$, $\partial > 0$ and $\xi \in (0, 1)$ such that*

$$\sum_{y \in A} \mu^n(x, y) > \delta \tag{3.5}$$

for all $r \geq r_0$, $(n, x) \in [\![a^{-2}r^2, a^2r^2]\!] \times U^{ar}$ and $A \subseteq U^r$ satisfying

$$|A| > \xi |U^r|.$$

Let Z_n be the random walk with transition probabilities

$$P[Z_{n+1} = y | Z_n = x] = \mu(x, y).$$

Let us also denote by P_x, $x \in G$, the probability measures satisfying

$$P_x[Z_0 = x] = 1 \quad \text{and} \quad P_x[Z_n = y] = \mu^n(x, y).$$

If $r > 0$ and $x \in \Gamma$, then we denote by τ_r^x the first exit time

$$\tau_r^x = \min\{n \in \mathbb{N} : Z_n \notin xU^r\}.$$

LEMMA 3.4. *For all $\varepsilon > 0$ there is a constant $c = c(\varepsilon) > 0$ such that, for all $r \geq 1$,*

$$P_x[\tau_{cr}^x \leq r^2] \leq \varepsilon. \tag{3.6}$$

PROOF. By Proposition 1.3, there is a $\delta > 0$ such that, for all $r \geq 1$,

$$\sum_{y \in U^r} \mu^n(e, y) \geq \partial, \qquad 1 \leq n \leq r^2.$$

Let us fix $\varepsilon > 0$. Then, by Proposition 1.4, there is a $c \geq 1$ such that, for all $r \geq 1$,

$$\sum_{y \notin U^{cr}} \mu^{r^2}(e, y) \leq \varepsilon\delta.$$

By choosing a larger constant c if necessary we can also assume that $xU^r \cap U^{cr} = \varnothing$ for $x \notin U^{2cr}$, $r \geq 1$.

We have

$$\begin{aligned}
\varepsilon\delta &\geq \sum_{y \notin xU^{cr}} \mu^{r^2}(x, y) \\
&= P_x[Z_{r^2} \notin xU^{cr}] \\
&\geq E^{P_x}\big[\mu^{r^2 - \tau_{2cr}^x}(Z_{\tau_{2cr}^x}, G \setminus U^{cr}); \tau_{2cr}^x \leq r^2\big]
\end{aligned}$$

$$
\begin{aligned}
&\geq E^{P_x}\big[\mu^{r^2-\tau^x_{2cr}}(Z_{\tau^x_{2cr}}, Z_{\tau^x_{2cr}}U^r); \tau^x_{2cr} \leq r^2\big]\\
&= E^{P_x}\big[\mu^{r^2-\tau^x_{2cr}}(e, U^r); \tau^x_{2cr} \leq r^2\big]\\
&\geq \delta P_x[\tau^x_{2cr} \leq r^2]
\end{aligned}
$$

and hence

$$P_x[\tau^x_{2cr} \leq r^2] \leq \varepsilon,$$

which proves the lemma. □

PROOF OF LEMMA 3.1. Let $t_0 \in [\![0, \frac{1}{2}a^{-2}r^2]\!]$ and let

$$A_{t_0} = A \cap \{t_0\} \times U^r.$$

Let $c > a$ and let $u \geq 0$ satisfy $u(s, y) \geq 1$, for $(s, y) \in A$ and $(\partial_1 + (I - \mu))u = 0$ in $[\![0, a^2r^2]\!] \times U^{2cr}$. Then, for all $(t, x) \in [\![a^{-2}r^2, a^2r^2]\!] \times U^{ar}$,

$$
\begin{aligned}
u(t, x) &\geq E^{P_x}[u(t_0, Z_{t-t_0}); \tau^x_{cr} > t - t_0]\\
&\geq E^{P_x}[\mathbb{1}_{A_{t_0}}(Z_{t-t_0}); \tau^x_{cr} > t - t_0] \qquad (3.7)\\
&= \mu^{t-t_0}(x, A_{t_0}) - P_x[\tau^x_{cr} \leq t - t_0].
\end{aligned}
$$

Now, by Lemma 3.3, there are $\partial > 0$, $r_0 \geq 1$ and $\xi_0 \in (0, 1)$ such that

$$\mu^{t-t_0}(x, A_{t_0}) > 2\delta$$

for all $(t, x) \in [\![a^{-2}r^2, a^2r^2]\!] \times U^{ar}$, if $r \geq r_0$ and A_{t_0} satisfies

$$|A_{t_0}| > \xi_0|U^r|. \qquad (3.8)$$

If we assume that $|A| > \xi|[\![0, r^2]\!] \times U^r|$, with $\xi \in [\xi_0, 1)$ close enough to 1, then A will always have a section A_{t_0} with $t_0 \in [\![0, \frac{1}{2}a^{-2}r^2]\!]$ and satisfying (3.8).

Also, by Lemma 3.4, if the constant c is large enough, then there is a $\delta > 0$ such that

$$P_x[\tau^x_{cr} \leq t - t_0] < \delta, \qquad t \in [\![a^{-2}r^2, a^2r^2]\!],$$

and hence, for all $(t, x) \in [\![a^{-2}r^2, a^2r^2]\!] \times U^{ar}$,

$$
\begin{aligned}
u(t, x) &\geq \mu^{t-t_0}(x, A_{t_0}) - P_x[\tau^x_{cr} \leq t - t_0]\\
&\geq 2\delta - \delta = \delta,
\end{aligned}
$$

which proves the lemma. □

3.3. *The second growth lemma.*

LEMMA 3.5 (Second growth lemma). *For all $b > 1$, there are $\beta > 1$, $c > 1$, $\delta > 0$ and $m \in \mathbb{N}$ such that*

$$\Psi\big([\![\beta r^2, (\beta + b^2)r^2]\!] \times U^{br}, A, [\![0, (\beta + b^2)r^2]\!] \times U^{cr}\big) > \delta\left(\frac{|A|}{|[\![1, r^2]\!] \times U^r|}\right)^m \tag{3.9}$$

for all $r \geq 1$ and every $A \subseteq [\![1, r^2]\!] \times U^r$.

The above lemma will actually be a consequence of the following:

LEMMA 3.6. *For all $b \geq 1$, there are $c, \beta > 0$, $\delta > 0$, $\delta' > 0$, $\theta > 0$ and $m \in \mathbb{N}$ such that for all $r \geq 1$ and every $A \subseteq [\![1, r^2]\!] \times U^r$ either*

$$\Psi\big([\![\beta r^2, (\beta + b^2)r^2]\!] \times U^{br}, A, [\![0, (\beta + b^2)r^2]\!] \times U^{cr}\big) > \partial\left(\frac{|A|}{|[\![1, r^2]\!] \times U^r|}\right)^m \tag{3.10}$$

or there is an $A_0 \subseteq [\![1, r^2]\!] \times U^r$ such that

$$|A_0| > (1 + \theta)|A| \tag{3.11}$$

and

$$\Psi\big(A_0, A, [\![0, r^2]\!] \times U^{cr}\big) > \delta'. \tag{3.12}$$

3.4. *Proof of Lemma* 3.6. We use the notation

$$Q(s, t, x) = \big[\!\big[t - \tfrac{1}{2}s^2, t + \tfrac{1}{2}s^2\big]\!\big] \times xU^s.$$

By Lemma 3.1, there are $s_0 \geq 1$, $\delta_1, \xi \in (0, 1)$ and $c_1 > 4$ such that

$$\Psi\big(\big[\!\big[\tfrac{1}{16}s^2, 16s^2\big]\!\big] \times U^{4s}, V, [\![0, 16s^2]\!] \times U^{cs}\big) > \delta_1 \tag{3.13}$$

for all $s \geq s_0$ and every measurable subset $V \subseteq [\![0, s^2]\!] \times U^s$ satisfying

$$|V| \geq \xi |[\![0, s^2]\!] \times U^s|.$$

Let us fix $r_0 > 3s_0$. Then by (3.1) there are $k > r_0^2$, $c_2 \geq 1$ and $\delta_2 > 0$ such that

$$\Psi\big(Q(r_0, 2k, x), \{(k, x)\}, \big[\!\big[0, 2k + \tfrac{1}{2}r_0^2\big]\!\big] \times xU^{c_2 r_0}\big) > \delta_2. \tag{3.14}$$

Let us assume that $r^2 \geq 6k$, $c > c_1 + c_2$ and consider the sets

$$A_1 = A \cap [\![3k, r^2 - 3k]\!] \times U^r \quad \text{and} \quad A_2 = A \setminus A_1.$$

Note that for $r^2 < 6k$ the lemma follows from (3.1).
Let $\eta \in (0, 1)$ be determined later.

Case I ($|A_1| \le \eta |A|$). Then $|A_2| > (1-\eta)|A|$ and hence

$$\begin{aligned}\frac{|A|}{|[\![1,r^2]\!]\times U^r|} &\le \frac{1}{1-\eta}\frac{|A_2|}{|[\![1,r^2]\!]\times U^r|}\\ &\le \frac{1}{1-\eta}\frac{|[\![1,3k]\!]\times U^r|+|[\![r^2-3k,r^2]\!]\times U^r|}{|[\![1,r^2]\!]\times U^r|}\\ &\le \frac{6k+2}{(1-\eta)r^2}.\end{aligned}$$

If $A \neq \varnothing$, then there is a $(t,x) \in [\![1,r^2]\!] \times U^r$ such that $u(t,x) \ge 1$ and so (3.10) follows from (3.3).

Case II ($|A_1| > \eta |A|$). Let

$$A_k = \{(t,x) : (t-k,x) \in A_1\}.$$

Then

$$A_k \subseteq [\![4k, r^2-2k]\!] \times U^r \quad \text{and} \quad |A_k| = |A_1| > \eta |A|.$$

We set

$$A_{\delta_2} = \bigcup_{(t,x)\in A_k} Q(r_0,t,x) \cap [\![1,r^2]\!] \times U^r.$$

Then of course $A_k \subseteq A_{\delta_2}$ and, by (3.14),

$$\Psi\big(A_{\delta_2}, A, [\![0,r^2]\!] \times U^{cr}\big) > \delta_2. \tag{3.15}$$

We consider the set of balls

$$\begin{aligned}\mathcal{Q} = \big\{&[\![t,t+s^2]\!] \times xU^s \subseteq [\![1,r^2]\!] \times U^r : s \ge s_0,\\ &|x|_\Gamma + |y|_\Gamma \le r,\ y \in U^s \text{ and } |Q \cap A_{\delta_2}| \ge \xi |Q|\big\}.\end{aligned}$$

With every ball $Q = [\![t,t+s^2]\!] \times xU^s \in \mathcal{Q}$ we associate a ball Q^0 as follows:

If $s + |x|_\Gamma < r$, then we set $Q' = [\![t,t+(s+1)^2]\!] \times xU^{s+1}$. If $s + |x|_\Gamma \ge r$, then we consider $y_1, \dots, y_{|x|_\Gamma} \in U$ such that $x = y_1 \cdots y_{|x|_\Gamma}$ and we set $x' = y_1 \cdots y_{|x|_\Gamma - 1}$ and $Q' = [\![t,t+(s+1)^2]\!] \times x'U^{s+1}$. If $Q' = |Q' \cap A_{\delta_2}| < \xi |Q'|$, then we take $Q^0 = Q'$. If not, then we repeat the same proceedure.

We set

$$\mathcal{Q}^0 = \{Q^0 : Q \in \mathcal{Q}\} \quad \text{and} \quad W^0 = \bigcup_{Q^0 \in \mathcal{Q}^0} Q^0.$$

Note that $A_k \subseteq W^0$.

LEMMA 3.7. *There is a* $\theta_1 = \theta_1(\xi) > 0$ *such that*

$$|W^0| > (1+\theta_1)|A_k| \tag{3.16}$$

and hence

$$|W^0| > (1+\theta_1)\eta|A|. \tag{3.17}$$

The proof of the above lemma is given later.

By using (3.13) repeatedly, we can see that there are $\delta_3 > 0$, $\delta_3 \leq \min(\delta_1, \delta_2)$ and $m \in \mathbb{N}$ such that, for all $\sigma \geq 1$ and $s \geq s_0$,

$$\begin{aligned}\Psi(&[\![s^2, 16\sigma^2 s^2]\!] \times U^{4s}, [\![s^2, 16s^2]\!] \times U^{4s}, [\![s^2, 16\sigma^2 s^2]\!] \times U^{c\sigma s}) \\ &\geq \delta_3 \sigma^{-m}.\end{aligned} \tag{3.18}$$

Let us fix $\sigma > 2$ such that

$$(1+\theta_1)\frac{16(\sigma-1)^2 - 1}{16(\sigma-1)^2} > 1 + \frac{\theta_1}{2}. \tag{3.19}$$

If $Q^0 = [\![t, t+s^2]\!] \times xU^s \in \mathcal{Q}^0$, then we set

$$Q^1 = [\![t, t+16\sigma^2 s^2]\!] \times xU^s, \qquad \mathcal{Q}^1 = \{Q^1, Q^0 \in \mathcal{Q}^0\}$$

and

$$W^1 = \bigcup_{Q^1 \in \mathcal{Q}^1} Q^1.$$

We also set

$$Q^1_{\mathbb{R}} = \big(t, t+16(\sigma-1)^2 s^2\big) \times xU^s$$

and define $\mathcal{Q}^1_{\mathbb{R}}$ and $W^1_{\mathbb{R}}$ similarly.

If $Q^1 = [\![t, t+16\sigma^2 s^2]\!] \times xU^s \in \mathcal{Q}^1$, then we set

$$Q^2 = [\![t+s^2, t+16\sigma^2 s^2]\!] \times xU^s, \qquad \mathcal{Q}^2 = \{Q^2, Q^1 \in \mathcal{Q}^1\}$$

and

$$W^2 = \bigcup_{Q^2 \in \mathcal{Q}^2} Q^2.$$

We also set

$$Q^2_{\mathbb{R}} = \big(t+s^2, t+16(\sigma-1)^2 s^2\big) \times xU^s$$

and define $\mathcal{Q}^2_{\mathbb{R}}$ and $W^2_{\mathbb{R}}$ similarly.

It follows from (3.15) and (3.18) that if $\delta_4 = \delta_1\delta_3\sigma^{-m}$, then

$$\Psi\big(W^2, A, [\![0, c^2 r^2]\!] \times U^{cr}\big) > \delta_4. \tag{3.20}$$

Let

$$\gamma = \frac{|A|}{|[\![1, r^2]\!] \times U^r|}$$

and let $\omega \in (0,1)$ be determined later.

Case IIa ($|W^2 \setminus [\![1, r^2]\!] \times U^r| \geq \omega|A|$). This assumption implies that

$$|W^2 \setminus [\![1, r^2]\!] \times U^r| \geq \omega\gamma|[\![1, r^2]\!] \times U^r|.$$

So there is a ball

$$Q^2 = [\![t + s^2, t + 16\sigma^2 s^2]\!] \times xU^s \in \mathcal{Q}^2$$

such that

$$16\sigma^2 s^2 \geq \omega\gamma r^2.$$

Now, by Corollary 3.2, for all $a_1 > 0$, there are $c \geq a_1$, $\delta_5 > 0$ and $m \in \mathbb{N}$ such that if s_0 is chosen large enough and $R \geq 4\sigma s$,

$$\begin{aligned}
&\Psi\big([\![t + R^2, t + (1 + a_1^2)R^2]\!] \times xU^{a_1 R}, Q^2, \ [\![t, t + (1 + a_1^2)R^2]\!] \times xU^{cR}\big) \\
&\quad \geq \delta_5\left(\frac{4\sigma s}{R}\right)^m \\
&\quad \geq \delta_5\omega^{m/2}\left(\frac{r}{R}\right)^m \gamma^{m/2}.
\end{aligned} \tag{3.21}$$

The lemma follows from (3.21) above, by taking a_1 large enough and by replacing R by an appropriate multiple of r.

Case IIb ($|W^2 \setminus [\![1, 1 + r^2]\!] \times U^r| < \omega|A|$). Let us first observe that

$$|W^0| \leq |W^1_{\mathbb{R}}| \tag{3.22}$$

and that

$$|W^2_{\mathbb{R}}| \leq |W^2|. \tag{3.23}$$

The following lemma is the analogue of Lemma 2.3 in [29], page 158.

LEMMA 3.8.

$$|W^1_{\mathbb{R}}| \leq \frac{16(\sigma - 1)^2}{16(\sigma - 1)^2 - 1}|W^2_{\mathbb{R}}|. \tag{3.24}$$

The proof of the above lemma is given in Section 3.6.

Combining (3.19), (3.22), (3.23) and (3.24) we have

$$\begin{aligned}
|W^2| &\geq \frac{16(\sigma - 1)^2 - 1}{16(\sigma - 1)^2}|W^0| \\
&\geq \frac{16(\sigma - 1)^2 - 1}{16(\sigma - 1)^2}(1 + \theta_1)\eta|A| \\
&\geq \left(1 + \frac{\theta_1}{2}\right)\eta|A|.
\end{aligned}$$

We set

$$A_0 = W^2 \cap [\![1, r^2]\!] \times U^r.$$

Then

$$\begin{aligned} |A_0| &= |W^2 \cap [\![1, r^2]\!] \times U^r| \\ &= |W^2| - |W^2 \setminus [\![1, r^2]\!] \times U^r| \\ &\geq |W^2| - \omega|A| \\ &\geq \left(1 + \frac{\theta_1}{2}\right)\eta|A| - \omega|A| \\ &\geq \left[\left(1 + \frac{\theta_1}{2}\right)\eta - \omega\right]|A|. \end{aligned}$$

It follows that if we chose $\eta \in (0, 1)$ so that

$$\left(1 + \frac{\theta_1}{2}\right)\eta > 1 + \frac{\theta_1}{4}$$

and

$$\omega \in \left(0, \frac{\theta_1}{8}\right),$$

then we would have

$$|A_0| > \left(1 + \frac{\theta_1}{8}\right)|A|,$$

which proves (3.11).

3.5. *Proof of Lemma* 3.7. If $Q = Q(s, t, x) = [\![t - \frac{1}{2}s^2, t + \frac{1}{2}s^2]\!] \times xU^s$, then we denote by Q^* the ball

$$Q^*(s, t, x) = \left[\!\left[t - \tfrac{25}{2}s^2, t + \tfrac{25}{2}s^2\right]\!\right] \times xU^{5s}.$$

Using a standard Vitalli type of argument (cf., e.g., [7, 27, 39]) we can prove that there is a finite sequence of balls $Q_1^0, Q_2^0, Q_3^0, \ldots, Q_n^0 \in \mathcal{Q}^0$ such that the following hold:

1. $Q_i^0 \in \mathcal{Q}^0$, $1 \leq i \leq n$;
2. $Q_i^0 \cap Q_j^0 = \varnothing$, $i \neq j$, $1 \leq i, j \leq n$;
3. $W^0 \subseteq \bigcup_{i=1}^n Q_i^{0*}$.

By (1.3) there is a constant $c \geq 1$ such that

$$\frac{1}{c} \leq \frac{|Q^*(s, t, x)|}{|Q(s, t, x)|} \leq c$$

for all $s \geq 1$.

So

$$
\begin{aligned}
\frac{|W^0|}{|A_k|} &= \frac{|A_k| + |W^0 \setminus A_k|}{|A_k|} = 1 + \frac{|W^0 \setminus A_k|}{|A_k|} \\
&\geq 1 + \frac{|W^0 \setminus A_{\delta_1}|}{|W^0|} \geq 1 + \frac{|W^0 \setminus A_{\delta_1}|}{|\bigcup_{i=1}^n Q_i^{0*}|} \\
&\geq 1 + \frac{|W^0 \setminus A_{\delta_1}|}{\sum_{i=1}^n |Q_i^{0*}|} \geq 1 + \frac{|W^0 \setminus A_{\delta_1}|}{c \sum_{i=1}^n |Q_i^0|} \\
&\geq 1 + \frac{|\bigcup_{i=1}^n Q_i^0 \setminus A_{\delta_1}|}{c \sum_{i=1}^n |Q_i^0|} = 1 + \frac{\sum_{i=1}^n |Q_i^0 \setminus A_{\delta_1}|}{c \sum_{i=1}^n |Q_i^0|}.
\end{aligned}
\tag{3.25}
$$

Since $|Q^i \cap A_{\delta_1}| < \xi |Q_i^0|$, we have

$$
\begin{aligned}
|Q_i^0 \setminus A_{\delta_1}| &= |Q_i^0| - |Q_i^0 \cap A_{\delta_1}| \\
&\geq |Q_i^0| - \xi |Q_i^0| = (1 - \xi)|Q_i^0|.
\end{aligned}
\tag{3.26}
$$

Combining (3.25) and (3.26) we have that

$$
\begin{aligned}
\frac{|W^0|}{|A_k|} &\geq 1 + \frac{\sum_{i=1}^n (1-\xi)|Q_i^0|}{c \sum_{i=1}^n |Q_i^0|} \\
&= 1 + \frac{1-\xi}{c},
\end{aligned}
$$

which proves the lemma.

3.6. *Proof of Lemma* 3.8. If $x \in U^r$, then we set

$$
W_{\mathbb{R}x}^1 = W_{\mathbb{R}}^1 \cap \mathbb{R} \times \{x\} \quad \text{and} \quad W_{\mathbb{R}x}^2 = W_{\mathbb{R}}^2 \cap \mathbb{R} \times \{x\}
$$

for $x \in U^r$.

It is enough to prove that

$$
|W_{\mathbb{R}x}^1| \leq \frac{(\sigma - 1)^2 \eta^2}{(\sigma - 1)^2 \eta^2 - 1} |W_{\mathbb{R}x}^2|.
\tag{3.27}
$$

This follows from Lemma 2.2 in [29], page 157, by taking

$$
\kappa = \frac{16(\sigma - 1)^2 \eta^2}{16(\sigma - 1)^2 \eta^2 - 1}
$$

and by setting

$$
g\bigl((t_1, t_2)\bigr) = \bigl(t_2 - \kappa(t_2 - t_1), t_2\bigr).
$$

4. A first difference estimate. Repeated use of Proposition 1.6 yields the following:

THEOREM 4.1. *There are* $\gamma \in (0,1]$ *and* $c > 0$ *such that, for all* $r \geq 1$, $z, x \in U$ *and every function* u *satisfying* $(\partial_1 + (I - \mu))u = 0$ *in* $[\![-r^2, 0]\!] \times U^r$,

$$|\partial_z u(0,x)| \leq c r^{-\gamma} \|u\|_\infty. \tag{4.1}$$

Combining this result with Varopoulos's theorem (Theorem 1.5) we have the following:

COROLLARY 4.2. *There are* $\gamma \in (0,1]$ *and* $c > 0$ *such that, for all* $n \in \mathbb{N}$, $z \in U$,

$$\|\partial_z \mu^{*n}\|_\infty \leq c n^{-(D+\gamma)/2}. \tag{4.2}$$

5. Results on the algebraic structure of N. In this section we recall certain well-known results on the algebraic structure of N (cf. [19, 21, 33, 45, 46, 50]).

5.1. *The filtration of the Lie algebra.* Let $\mathfrak{n}$ be the Lie algebra of N, which we identify with the left invariant vector fields on N.

We set $\mathfrak{n}_1 = \mathfrak{n}$ and $\mathfrak{n}_{i+1} = [\mathfrak{n}_1, \mathfrak{n}_i]$, $i \geq 1$. Since $\mathfrak{n}$ is nilpotent, we have the filtration

$$\mathfrak{n} = \mathfrak{n}_1 \supseteq \mathfrak{n}_2 \supseteq \cdots \supseteq \mathfrak{n}_m \supseteq \mathfrak{n}_{m+1} = \{0\}, \qquad \mathfrak{n}_m \neq \{0\}.$$

We consider linear subspaces $\mathfrak{a}_1, \ldots, \mathfrak{a}_m$ of $\mathfrak{n}$ such that

$$\mathfrak{n}_i = \mathfrak{a}_i \oplus \cdots \oplus \mathfrak{a}_m, \qquad 1 \leq i \leq m.$$

We set

$$\begin{aligned} n_0 &= 0, \qquad n_i = \dim(\mathfrak{a}_1 \oplus \cdots \oplus \mathfrak{a}_i), \qquad 1 \leq i \leq m, \\ \sigma(j) &= i, \qquad \text{for } n_{i-1} < j \leq n_i, \\ q &= n_m = \dim(\mathfrak{n}). \end{aligned}$$

Notice that the homogeneous dimension D of N is given by

$$D = \sigma(1) + \cdots + \sigma(q).$$

We consider a basis $\{X_1, \ldots, X_q\}$ of $\mathfrak{n}$ such that $\{X_{n_{i-1}+1}, \ldots, X_{n_i}\}$ is a basis of $\mathfrak{a}_i$, $1 \leq i \leq m$.

On the linear space $\mathfrak{n}$, we define the Lie bracket $[\cdot,\cdot]_0$ by setting

$$[X_i, X_j]_0 = \mathrm{pr}_{\mathfrak{a}_{\sigma(i)+\sigma(j)}}[X_i, X_j].$$

We denote by $\mathfrak{n}_0$ the Lie algebra $\mathfrak{n}_0 = (\mathfrak{n}, [\cdot,\cdot]_0)$. Note that $\mathfrak{n}_0$ is nilpotent.

5.2. *Dilations and the exponential coordinates.* Using the exponential coordinates of the second kind (or Malcev coordinates)

$$\phi:\mathbb{R}^q \to N, \qquad \phi: x = (x_q, \dots, x_1) \to \exp x_q X_q \cdots \exp x_1 X_1$$

we identify N, as a differential manifold, with $\mathbb{R}^q$.

Let τ_ε, $\varepsilon > 0$, be the family of dilations of N defined by

$$\tau_\varepsilon : (x_q, \dots, x_1) \to \big(\varepsilon^{\sigma(q)} x_q, \dots, \varepsilon^{\sigma(1)} x_1\big).$$

Also, let $*_\varepsilon$, $\varepsilon > 0$, be the family of group products defined by

$$x *_\varepsilon y = \tau_\varepsilon[(\tau_{\varepsilon^{-1}} x)(\tau_{\varepsilon^{-1}} y)]$$

and let

$$x *_0 y = \lim_{\varepsilon \to 0} x *_\varepsilon y.$$

Then $N_0 = (N, *_0)$ is a stratified nilpotent Lie group whose Lie algebra is isomorphic to $\mathfrak{n}_0$. We identify $\mathfrak{n}_0$ with the $*_0$-left invariant vector fields.

If $X \in \mathfrak{n}$ is a left invariant vector field on N, then we denote by X_0 the $*_0$-left invariant vector field satisfying $X_0(e) = X(e)$.

In particular we denote by X_{0i} the $*_0$-left invariant vector fields satisfying $X_{0i}(0) = X_i(0)$, $1 \le i \le q$.

Note that

$$X_{0i} = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon^{\sigma(i)}}\, d\tau_\varepsilon(X_i), \qquad 1 \le i \le q, \tag{5.1}$$

and that

$$X_{0i} = \frac{1}{\varepsilon^{\sigma(i)}}\, d\tau_\varepsilon(X_{0i}), \qquad 1 \le i \le q. \tag{5.2}$$

We now give an expression of the left invariant vectors fields of N as vector fields on $\mathbb{R}^q$.

If $X = a_1 X_1 + \cdots + a_q X_q$, then we set $\mathrm{pr}_i(X) = a_i$, $i = 1, \dots, q$.

We also denote by $\overline{\mathrm{ad}} X_i$ the linear transformations of $\mathfrak{n}$ defined by

$$\overline{\mathrm{ad}}(X_i) X_j = \begin{cases} 0, & \text{for } i \ge j, \\ \mathrm{ad}(X_i) X_j, & \text{for } i < j. \end{cases}$$

LEMMA 5.1 (cf. [4]). *Let X be a left invariant vector field on N. Then $X(x) = a_q(x)\frac{\partial}{\partial x_q} + \cdots + a_1(x)\frac{\partial}{\partial x_1}$ with*

$$\begin{aligned} a_i(x) &= \mathrm{pr}_i\big[e^{x_{i-1}\overline{\mathrm{ad}} X_{i-1}} \cdots e^{x_1 \overline{\mathrm{ad}} X_1}(X)\big] \\ &= \mathrm{pr}_i\Bigg[\sum_{\lambda_1\sigma(1)+\cdots+\lambda_{i-1}\sigma(i-1) \le \sigma(i)-1} \frac{1}{\lambda_1!}\cdots\frac{1}{\lambda_{i-1}!} x_1^{\lambda_1}\cdots x_{i-1}^{\lambda_{i-1}} \\ &\qquad \times (\overline{\mathrm{ad}} X_{i-1})^{\lambda_{i-1}} \cdots (\overline{\mathrm{ad}} X_1)^{\lambda_1}(X)\Bigg]. \end{aligned} \tag{5.3}$$

Note that if X_0 is the associated $*_0$-left invariant vector field satisfying $X_0(e) = X(e)$, then $X_0(x) = a_{0q}(x)\frac{\partial}{\partial x_q} + \cdots + a_{01}(x)\frac{\partial}{\partial x_1}$ with

$$a_{0i}(x) = \mathrm{pr}_i \Bigg[\sum_{\lambda_1\sigma(1)+\cdots+\lambda_{i-1}\sigma(i-1)=\sigma(i)-1} \frac{1}{\lambda_1!}\cdots\frac{1}{\lambda_{i-1}!} x_1^{\lambda_1}\cdots x_{i-1}^{\lambda_{i-1}} \times (\overline{\mathrm{ad}}X_{i-1})^{\lambda_{i-1}}\cdots(\overline{\mathrm{ad}}X_1)^{\lambda_1}(X)\Bigg]. \tag{5.4}$$

Let us set, for $f \in C^\infty$ and $\ell \in \mathbb{N}$,

$$\nabla^\ell f(x) = \sum_{\substack{a\le\ell \\ \sigma(i_1)+\cdots+\sigma(i_a)\ge\ell}} \left|\frac{\partial}{\partial x_{i_1}}\cdots\frac{\partial}{\partial x_{i_a}} f(x)\right|,$$

$$\nabla_X^\ell f(x) = \sum_{\substack{a\le\ell \\ \sigma(i_1)+\cdots+\sigma(i_a)\ge\ell}} |X_{i_1}\cdots X_{i_a} f(x)|.$$

Then it follows from (5.3) that there are $c > 0$ and $k \in \mathbb{N}$ such that, for all $x \in N$ and $f \in C^\infty$,

$$\frac{1}{c(1+|x|)^k}\nabla^\ell f(x) \le \nabla_X^\ell f(x) \le c(1+|x|)^k \nabla^\ell f(x). \tag{5.5}$$

5.3. *Taylor expansions.*

LEMMA 5.2. *Let $f \in C^\infty$. Then*

$$\frac{\partial}{\partial x_i} f(0) = X_i f(0), \qquad 1 \le i \le q, \tag{5.6}$$

$$\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} f(0) = X_i X_j f(0), \qquad 1 \le j \le i \le n_1 \tag{5.7}$$

and

$$\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} f(0) = X_i X_j f(0) - \sum_{n_1<\lambda\le q} (\mathrm{pr}_\lambda[X_i, X_j]) X_\lambda f(0), \qquad 1 \le i < j \le n_1. \tag{5.8}$$

PROOF. Equation (5.6) follows immediately from (5.3). Equation (5.7) follows also from (5.3), since

$$X_i X_j f(0) = \frac{\partial}{\partial y_i} X_j f(0) = \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} f(0), \qquad 1 \le j \le i \le n_1.$$

Finally, to prove (5.8) we observe that

$$\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} f(0) = \frac{\partial}{\partial x_j}\frac{\partial}{\partial x_i} f(0)$$

and hence, by (5.6) and (5.7),

$$\begin{aligned}\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} f(0) &= X_j X_i f(0)\\ &= \big(X_i X_j - [X_i, X_j]\big) f(0)\\ &= X_i X_j f(0) - \sum_{n_1<\lambda\le q} (\mathrm{pr}_\lambda [X_i, X_j]) X_\lambda f(0). \qquad \square\end{aligned}$$

Let us now assume that V is a compact neighborhood of the identity element e of N which, viewed as subset of $\mathbb{R}^q$, is also convex. Also, let us denote by P_i the monomial $p_i(x) = x_i$, $x = (x_1, \dots, x_q) \in \mathbb{R}^q$, $1 \le i \le q$.

LEMMA 5.3. *Let V be as above. Then there is a constant $c > 0$ such that, for all $f \in C^\infty(N)$ and all $x \in N$, $y \in V$,*

$$f(xy) = f(x) + F_x(y) \qquad \textit{with } |F_x(y)| \le c\|\nabla_X f\|_{L^\infty(xV)}, \tag{5.9}$$

$$f(xy) = f(x) + \sum_{1\le i\le n_1} P_i(y) X_i f(x) + F_x(y) \tag{5.10}$$

with $|F_x(y)| \le c\|\nabla_X^2 f\|_{L^\infty(xV)}$ and

$$\begin{aligned} f(xy) = f(x) &+ \sum_{1\le i\le n_1} P_i(y) X_i f(x) + \tfrac{1}{2}\sum_{1\le i,j\le n_1} P_i(y)P_j(y) X_i X_j f(x)\\ &+ \sum_{n_1<i\le n_2}\Bigg(P_i(y) - \tfrac{1}{2}\sum_{1\le\lambda<\mu\le n_1} P_\lambda(y)P_\mu(y)\,\mathrm{pr}_i[X_\lambda, X_\mu]\Bigg) X_i f(x)\\ &+ F_x(y)\end{aligned} \tag{5.11}$$

with $|F_x(y)| \le c\|\nabla_X^3 f\|_{L^\infty(xV)}$.

PROOF. We only give the proof of (5.11). The proofs of (5.9) and (5.10) are similar. Let $f'(y) = f(xy)$. If $y \in V$, then by the Taylor formula (in $\mathbb{R}^q$)

$$\begin{aligned} f'(y) = f'(0) &+ \sum_{1\le i\le q} P_i(y)\frac{\partial}{\partial y_i} f'(0)\\ &+ \frac{1}{2}\sum_{1\le i,j\le q} P_i(y)P_j(y)\frac{\partial}{\partial y_i}\frac{\partial}{\partial y_j} f'(0) + F'(y)\end{aligned}$$

$$= f'(0) + \sum_{1\le i\le n_1} P_i(y)\frac{\partial}{\partial y_i} f'(0) + \frac{1}{2}\sum_{1\le i,j\le n_1} P_i(y)P_j(y)\frac{\partial}{\partial y_i}\frac{\partial}{\partial y_j} f'(0)$$
$$+ \sum_{n_1<i\le n_2} P_i(y)\frac{\partial}{\partial y_i} f'(0) + F''(y),$$

where

$$|F'(y)| \le c\|\nabla^3 f'\|_{L^\infty(V)}$$

and hence also

$$|F''(y)| \le c\|\nabla^3 f'\|_{L^\infty(V)}.$$

So, by Lemma 5.2,

$$f'(y) = f'(0) + \sum_{1\le i\le n_1} P_i(y)X_i f'(0) + \tfrac{1}{2}\sum_{1\le i,j\le n_1} P_i(y)P_j(y)X_iX_j f'(0)$$
$$+ \sum_{n_1<i\le q}\left(P_i(y) - \tfrac{1}{2}\sum_{1\le\lambda<\mu\le n_1} P_\lambda(y)P_\mu(y)\,\mathrm{pr}_i[X_\lambda, X_\mu]\right)X_i f'(0)$$
$$+F''(y).$$

If we set

$$F(y) = F''(y) + \sum_{n_2<i\le q}\left(P_i(y) - \tfrac{1}{2}\sum_{1\le\lambda<\mu\le n_1} P_\lambda(y)P_\mu(y)\,\mathrm{pr}_i[X_\lambda, X_\mu]\right)X_i f'(0)$$

then, by (5.5),

$$|F(y)| \le c\|\nabla_X^3 f'\|_{L^\infty(V)}.$$

Also,

$$f'(y) = f'(0) + \sum_{1\le i\le n_1} P_i(y)X_i f'(0) + \tfrac{1}{2}\sum_{1\le i,j\le n_1} P_i(y)P_j(y)X_iX_j f(0)$$
$$+ \sum_{n_1<i\le n_2}\left(P_i(y) - \tfrac{1}{2}\sum_{1\le\lambda<\mu\le n_1} P_\lambda(y)P_\mu(y)\,\mathrm{pr}_i[X_\lambda, X_\mu]\right)X_i f'(0)$$
$$+F(y).$$

Given the left invariance of the vector fields X_i, this implies (5.11). □

COROLLARY 5.4. *There is a constant $c > 0$ such that, for all $f \in C^\infty(N)$ and all $x \in N$, $y \in V$, the following hold*:

(i) *if* $1 \leq \nu \leq n_1$, *then*

$$X_\nu f(xy) = X_\nu f(x) + \sum_{1 \leq i \leq n_1} P_i(y) X_i X_\nu f(x) + F_x(y) \tag{5.12}$$

with $|F_x(y)| \leq c \|\nabla_X^3 f\|_{L^\infty(xV)}$;

(ii) *if* $n_1 < \nu \leq n_2$,

$$X_\nu f(xy) = X_\nu f(x) + F_x(y) \qquad \textit{with } |F_x(y)| \leq c \|\nabla_X^3 f\|_{L^\infty(xV)}. \tag{5.13}$$

6. Centered sub-Laplacians on N. Let $L = -(E_1^2 + \cdots + E_p^2) + E_0$ be a left invariant sub-Laplacian on N and let us assume that it is centered, that is, that $E_0 \in [\mathfrak{n}, \mathfrak{n}]$.

Let $\{X_1, \ldots, X_q\}$ be the basis of $\mathfrak{n}$ introduced in Section 5. Since the vector fields E_i are linear combinations of the vector fields X_i, the sub-Laplacian L can also be written as

$$L = -\sum_{1 \leq i, j \leq q} a_{ij} X_i X_j - \sum_{n_1 < i \leq q} a_i X_i. \tag{6.1}$$

Note that $a_{ij} = a_{ji}$, $1 \leq i, j \leq q$. Also the assumption that the vector fields $E_1, \ldots, E_p$ satisfy Hörmander's condition implies that the $(n_1 \times n_1)$ matrix $B = (b_{ij})$ with entries $b_{ij} = a_{ij}$, $1 \leq i, j \leq n_1$, is positive definite.

We associate with L the limit (at ∞) sub-Laplacian

$$L_0 = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon^2} d\tau_\varepsilon(L) = -\sum_{1 \leq i, j \leq n_1} a_{ij} X_{0i} X_{0j} - \sum_{n_1 < i \leq n_2} a_i X_{0i}.$$

Note that L_0 is dilation invariant; that is, it satisfies

$$L_0 = \frac{1}{\varepsilon^2} d\tau_\varepsilon(L_0), \qquad \varepsilon > 0. \tag{6.2}$$

6.1. *Polynomials.* Since we have identified N, as a differential manifold, with $\mathbb{R}^q$ using the exponential coordinates of the second kind, the monomials on $\mathbb{R} \times N$ will just be monomials on $\mathbb{R} \times \mathbb{R}^q$.

Let $P(t, x) = t^{i_0} x_1^{i_1} \cdots x_q^{i_q}$ be such a monomial. Then the homogeneous degree $\deg_H P$ of P is also given by

$$\deg_H P = 2i_0 + i_1 \sigma(1) + \cdots + i_q \sigma(q).$$

By (5.3),

$$\deg_H \left(\frac{\partial}{\partial t} + L \right) P(x) \leq \deg_H P(t, x) - 2. \tag{6.3}$$

Also, by (5.4),

$$\left(\frac{\partial}{\partial t} + L \right) P(t, x) = \left(\frac{\partial}{\partial t} + L_0 \right) P(t, x) + Q(t, x), \tag{6.4}$$

where $Q(t,x)$ is a polynomial satisfying

$$\deg_H Q \leq \deg_H P - 3.$$

Using induction on the dimension q of the Lie algebra $\mathfrak{n}$ of N and the homogeneous degree $\deg_H P$ of P we can prove that there is a polynomial $Q(t,x)$ satisfying

$$\left(\frac{\partial}{\partial t} + L_0\right) Q(t,x) = P(t,x), \tag{6.5}$$
$$\deg_H Q(t,x) = \deg_H P(t,x) + 2.$$

Combining (6.3), (6.4) and (6.5) we have the following:

LEMMA 6.1. *With every monomial $P(t,x)$ as above we can associate a polynomial*

$$Q_P(t,x) = P(t,x) + W(t,x)$$

satisfying

$$\deg_H W \leq \deg_H P - 1, \tag{6.6}$$
$$\left(\frac{\partial}{\partial t} + L_0\right) P(t,x) = \left(\frac{\partial}{\partial t} + L\right) Q_P(t,x).$$

Note that the polynomial $Q_P(t,x)$ in the above lemma, is not necessarily unique.

From now on, for all $d \in \mathbb{N}$, we denote by

$$P_0(t,x),\, P_2(t,x),\, \ldots,\, P_{\nu_d}(t,x)$$

the monomials with homogeneous degree less than or equal to d. Given a centered left invariant sub-Laplacian L on N (in this article, this is either L_μ or $L_{H\mu}$), we associate with those monomials, polynomials

$$Q_{P_0}(t,x),\, Q_{P_2}(t,x),\, \ldots,\, Q_{P_{\nu_d}}(t,x)$$

satisfying (6.6).

Note that, for $0 \leq i \leq \nu_2$, we can take $Q_{P_i} = P_i$. Note also that $\nu_0 = 0$, $\nu_1 = n_1$ and $\nu_1 < n_2 < \nu_2$. So, we assume that $P_0(t,x) = Q_{P_0}(t,x) = 1$, that $P_i(t,x) = Q_{P_i}(t,x) = x_i$, for $1 \leq i \leq n_2$, and that $P_i = Q_{P_i}$, for $\nu_1 < i \leq \nu_2$.

6.2. *A Taylor formula for the heat functions.* The following Taylor formula for the heat functions is proved in [7].

THEOREM 6.2. *For all $n \in \mathbb{N}$ there is a $c_n > 0$ such that for all $r, R \in \mathbb{N}$, $R \geq r \geq 1$, and all functions u satisfying*

$$\left(\frac{\partial}{\partial t} + L\right)u = 0 \qquad \text{in } (-R^2, R^2) \times V^R$$

we have

$$\sup\left\{\left|u - \sum_{0 \leq i \leq \nu_n} A_i R^{-\deg_H P_i} Q_{P_i}\right|; (r^2, r^2) \times V^r\right\} < c_n \left(\frac{R}{r}\right)^{-(n+1)} \|u\|_\infty, \tag{6.7}$$

where the constants A_i satisfy

$$|A_i| \leq c_n \|u\|_\infty$$

for all $0 \leq i \leq \nu_n$ and

$$\left(\frac{\partial}{\partial t} + L\right)\left(\sum_{\nu_{d-1} < i \leq \nu_d} A_i Q_{P_i}\right) = 0$$

for all $1 < k \leq n$.

6.3. *A Harnack inequality.*

THEOREM 6.3 [7]. *Let V be a compact connected neighborhood of the identity element e of N, let $\alpha, a, \beta, b \in \mathbb{N}$, $1 \leq \alpha < \alpha + a < \beta$, and let $k, \ell \in \mathbb{N}$. Then there is a $c \in \mathbb{N}$ such that for all $r \in \mathbb{N}$, $r \geq 1$, and all $u \geq 0$ satisfying*

$$\left(\frac{\partial}{\partial t} + L\right)u = 0 \qquad \text{in } \big(0, (\beta^2 + b^2)r^2\big) \times V^{cr},$$

we have

$$\sup\left\{\left|\frac{\partial^k}{\partial t^k} X_{i_1} \cdots X_{i_\ell} u\right|; \big(\alpha r^2, (\alpha + a^2)r^2\big) \times V^{ar}\right\} \leq c r^{-2k-\sigma(i_1)-\cdots-\sigma(i_\ell)} \inf\{u; \big(\beta r^2, (\beta + b^2)r^2\big) \times V^{br}\}. \tag{6.8}$$

6.4. *Estimates for the heat kernel.* The heat kernel $p_t(x, y)$ of L satisfies the following Gaussian estimate (cf. [7]).

THEOREM 6.4. *There is a constant $c > 0$ such that, for all $x, y \in N$ and $t \geq 1$,*

$$\frac{1}{c} t^{-D/2} \exp\left(-c\frac{|x^{-1}y|_N^2}{t}\right) \leq p_t(x, y) \leq c t^{-D/2} \exp\left(-\frac{|x^{-1}y|_N^2}{ct}\right). \tag{6.9}$$

Combining (6.9) with (6.8) we have the following:

COROLLARY 6.5. *For all $k, \ell \in \mathbb{N}$ there is a constant $c > 0$ such that*

$$\left|\frac{\delta^k}{\partial t^k} X_{i_1} \cdots X_{i_n} p_t(x, y)\right| \leq c t^{-(D+2k+\sigma(i_1)+\cdots+\sigma(i_\ell))/2} \exp\left(-\frac{|x^{-1}y|_N^2}{ct}\right), \tag{6.10}$$

for all $t \geq 1$, $x, y \in N$ and for all $1 \leq i_j \leq q$, $1 \leq j \leq n$.

Let $p_t^0(x, y)$ be the heat kernel of L_0. Then we have the following analogue of the classical Berry–Esseen estimate (cf. [7, 20]):

THEOREM 6.6. *There is a constant $c > 0$ such that, for all $t \geq 1$ and $x \in N$,*

$$|p_t(x, e) - p_t^0(x, e)| \leq c t^{-(D+1)/2}. \tag{6.11}$$

It follows from (6.2) that

$$p_t^0(x, y) = \varepsilon^D p_{\varepsilon^2 t}^0(\tau_\varepsilon x, \tau_\varepsilon y), \qquad \varepsilon > 0,$$

and hence there is a $C_{L_0} > 0$ such that

$$p_t^0(e, e) = C_{L_0} t^{-D/2}. \tag{6.12}$$

Combining (6.11) with (6.12) we have the following:

COROLLARY 6.7. *There are constants $C_L > 0$ and $c > 0$ such that*

$$|p_t(x, e) - C_L t^{-D/2}| \leq c t^{-(D+1)/2}, \qquad t \geq 1. \tag{6.13}$$

7. A smooth substitute for $|x|_N$. The following proposition furnishes a positive smooth function $\rho(x)$ on a simply connected nilpotent Lie group N, which will replace $|x|_N$ in the proof of the Gaussian estimate (1.14). This function will be a convenient power of the Green function of a symmetric sub-Laplacian L on N.

We use the notation of the previous section.

PROPOSITION 7.1. *There is a function $\rho(x) \in C^\infty(N)$ with the following properties: For all $n \in \mathbb{N}$ there is a constant $c \geq 1$ such that, for all $x \in N$ and all $1 \leq i_j \leq q$, $1 \leq j \leq n$,*

$$\begin{gathered} \rho(x) \geq 0, \qquad x \in N, \\ \frac{1}{c}\,|x|_N \leq \rho(x) \leq \ c\,|x|_N \qquad \text{for } |x| \geq 2, \\ |X_{i_1} \cdots X_{i_n} \rho(x)| \leq \frac{c}{|x|_N^{\sigma(i_1)+\cdots+\sigma(i_n)-1}} \qquad \text{for } |x| \geq 2. \end{gathered} \tag{7.1}$$

PROOF. If the homogeneous dimension D of N is $D \leq 2$, then N is isomorphic either to $\mathbb{R}$ or to $\mathbb{R}^2$ and then we can take as ρ the Euclidean norm. So let us assume that $D > 2$. Let L be a symmetric left invariant sub-Laplacian on N and let $p_t(x, y)$ be its heat kernel. The Green function of L is given by

$$G_L(x, y) = \int_0^\infty p_t(x, y)\, dt.$$

Let $G_L(x) = G_L(x, e)$. Then it follows from (6.9) that there is a $c > 0$ such that

$$\frac{1}{c}\frac{1}{|x|_N^{D-2}} \leq G_L(x) \leq c\frac{1}{|x|_N^{D-2}}, \qquad |x|_N \geq 2. \tag{7.2}$$

Since $LG_L(x) = 0$, in $N \setminus \{e\}$ it follows from (6.10) that for all $n \in \mathbb{N}$ there is a $c > 0$ such that

$$|X_{i_1} X_{i_2} \cdots X_{i_n} G_L(x)| \leq \frac{c}{|x|_N^{D-2+\sigma(i_1)+\cdots+\sigma(i_n)}}, \qquad |x|_N \geq 2,\ 0 \leq \varepsilon \leq 1. \tag{7.3}$$

The function

$$\rho(x) = \bigl(G_L(x)\bigr)^{-1/(D-2)}$$

satisfies (7.1). $\square$

8. Construction of the sub-Laplacian L_μ when Γ is nilpotent. In this section we give the definition of the operator $L_{H\mu}$ when Γ is nilpotent. If Γ is not nilpotent, then the action of Γ/Γ_N on Γ_N gives rise to phenomena of homogenization and this makes the definition of $L_{H\mu}$ more complicated. To make this distinction, in the nilpotent case, we use the notation L_μ instead of $L_{H\mu}$.

We assume that $\Gamma = \Gamma_N$. If Γ is nilpotent and the torsion subgroup $\tau(\Gamma)$ is not trivial, then we define L_μ to be the same as the operator $L_{\pi(\mu)}$ associated with the image $\pi(\mu)$ of μ under the quotient map $\pi : \Gamma \to \Gamma_N = \Gamma/\tau(\Gamma)$.

We use the notation of Sections 5 and 6.

The operator L_μ will be a centered sub-Laplcian which can be written as

$$L_\mu = -\sum_{1 \leq i, j \leq n_1} a_{ij} X_i X_j - \sum_{n_1 < i \leq n_2} a_i X_i.$$

The coefficients a_{ij} and a_i are defined as follows.

The coefficients a_{ij} are given by

$$a_{ij} = \tfrac{1}{2} \sum_{x \in \Gamma} P_i(x) P_j(x) \mu(x), \qquad 1 \leq i, j \leq n_1.$$

Let

$$b_i = \sum_{x \in \Gamma} P_i(x) \mu(x), \qquad 1 \leq i \leq n_2.$$

Note that since μ is centered $b_i = 0$, $1 \le i \le n_1$.

The coefficients a_i of L_μ are given by

$$a_i = b_i - \frac{1}{2} \sum_{1 \le \lambda < \mu \le n_1} a_{\lambda\mu} \operatorname{pr}_i [X_\lambda, X_\mu], \qquad n_1 < i \le n_2.$$

Let V be as in in Section 5.3 and let us assume that V is large enough that $\operatorname{supp} \mu \subseteq V$.

The following lemma explains the relation between μ and L_μ. It is an immediate consequence of Lemma 5.3.

LEMMA 8.1. *There is a constant $c > 0$ such that, for all functions $f \in C^\infty(N)$ and all $x \in N$,*

$$|(I - \mu) f(x)| \le c \|\nabla_X^2 f\|_{L^\infty(xV)} \tag{8.1}$$

and

$$(I - \mu) f(x) = L_\mu f(x) + F(x), \tag{8.2}$$

with

$$|F(x)| \le c \|\nabla_X^3 f\|_{L^\infty(gV)}.$$

COROLLARY 8.2. *There is a constant $c > 0$ such that, for all functions $u(t, x) \in C^\infty(\mathbb{R} \times N)$ and all $x \in N$,*

$$\begin{aligned} u(t+1, x) - \mu u(t, x) &= \big(\partial_1 + (I - \mu)\big) u(t, x) \\ &= \left(\frac{\partial}{\partial t} + L_\mu\right) u(t, x) + F(t, x), \end{aligned} \tag{8.3}$$

where $F(t, x)$ satisfies

$$|F(t, x)| \le c \left\| \left| \frac{\partial^2}{\partial s^2} u \right| + |\nabla_X^3 u| \right\|_{L^\infty([t, t+1] \times xV)}.$$

The following lemma asserts that L_μ is indeed a sub-Laplacian.

LEMMA 8.3. *The $n_1 \times n_1$ matrix (a_{ij}) is positive definite.*

PROOF. It is enough to prove that

$$\sum_{1 \le i, j \le n_1} a_{ij} \xi_i \xi_j > 0 \tag{8.4}$$

for all $\xi = (\xi_1, \ldots, \xi_{n_1}) \in \mathbb{R}^{n_1}$, $\xi \ne 0$.

To this end, let us fix $\xi = (\xi_1, \dots, \xi_{n_1}) \neq 0$ and consider the function

$$u(x) = \big(\xi_1 P_1(x) + \cdots + \xi_{n_1} P_{n_1}(x)\big)^2 = \sum_{1 \le i, j \le n_1} \xi_i \xi_j P_i(x) P_j(x).$$

By (8.2),

$$(I - \mu)\big(P_i(x) P_j(x)\big) = L_\mu\big(P_i(x) P_j(x)\big) = -(a_{ij} + a_{ji})$$

and hence

$$(I - \mu) u(x) = 2 \sum_{1 \le i, j \le n_1} a_{ij} \xi_i \xi_j.$$

If we had

$$\sum_{1 \le i, j \le n_1} a_{ij} \xi_i \xi_j = 0,$$

then we would have $(I - \mu) u = 0$, that is, $\sum_{y \in \Gamma} u(xy) \mu(y) = u(x)$, $x \in N$. Since $u(x) = 0$, this would imply that $u(x) = 0$ for all $x \in \Gamma_N$, which is false. Hence, (8.4) holds and the lemma follows. $\square$

9. Proof of Propositions 1.3 and 1.4 when Γ is nilpotent. The goal of this section is to prove Propositions 1.3 and 1.4 when Γ is nilpotent. Note that there is no loss of generality if we assume that the torsion subgroup $\tau(\Gamma)$ is trivial and hence that $\Gamma = \Gamma_N$.

We use the same notation $p_t^\mu(x, y)$ to denote both the heat kernel of the sub-Laplacian L_μ and its restriction to Γ_N.

The proofs are based on the following lemma.

LEMMA 9.1. *There is a constant $c > 0$ such that, for all $n \in \mathbb{N}$, $T \ge 1$,*

$$\| p^\mu_{n+T} - \mu^n p^\mu_T \|_\infty \le c\, T^{-(D+1)/2}. \tag{9.1}$$

PROOF. We have

$$\begin{aligned} p^\mu_{n+T} - \mu^n p^\mu_T &= p^\mu_{n+T} - \mu^{n-1} p^\mu_{1+T} + \mu^{n-1} p^\mu_{1+T} - \mu^n p^\mu_T \\ &= \sum_{0 \le i \le n-1} \mu^i \big(p^\mu_{n-i+T} - \mu p^\mu_{n-i-1+T}\big). \end{aligned} \tag{9.2}$$

On the other hand, it follows from (8.3) that

$$p^\mu_{t+1} - \mu p^\mu_t = \left(\frac{\partial}{\partial t} + L_\mu\right) p^\mu_t + V_t = V_t$$

with V_t satisfying

$$|V_t(x,y)| \le c_1 \left\| \left| \frac{\partial^2}{\partial s^2} p_s^\mu(\cdot, y) \right| + |\nabla_X^3 p_s^\mu(\cdot, y)| \right\|_{L^\infty([t,t+1]\times x V)}.$$

So, by (6.10) there is a $c > 0$ such that

$$\|p_{t+1}^\mu - \mu p_t^\mu\|_\infty \le c t^{-(D+3)/2}, \qquad t \ge 1. \tag{9.3}$$

Combining (9.2) and (9.3) we have

$$\begin{aligned} \|p_{n+T}^\mu - \mu^n p_T^\mu\|_\infty &\le \sum_{0\le i\le n-1} \|\mu^i\|_1 \|p_{n-i+T}^\mu(x,y) - \mu p_{n-i-1+T}^\mu\|_\infty \\ &\le c \sum_{0\le i\le n-1} (n-i-1+T)^{-(D+3)/2} \\ &\le c\, T^{-(D+1)/2}, \end{aligned}$$

which proves the lemma. □

9.1. *Proof of Proposition* 1.3. We have

$$\begin{aligned} \sum_{y\in U^r} \mu^n(x,y) &\ge \frac{1}{\|p_T^\mu\|_\infty} \sum_{y\in U^r} \mu^n(x,y) p_T^\mu(y,e) \\ &= \frac{1}{\|p_T^\mu\|_\infty} \left(\mu^n p_T^\mu(x,e) - \sum_{\{y\notin U^r\}} \mu^n(x,y) p_T^\mu(y,e) \right) \\ &\ge \frac{1}{\|p_T^\mu\|_\infty} \biggl(p_{n+T}^\mu(x,e) - \|\mu^n p_T^\mu - p_{n+T}^\mu\|_\infty \\ &\qquad\qquad - \sum_{\{y\notin U^r\}} \mu^n(x,y) p_T^\mu(y,e) \biggr). \end{aligned}$$

Let us fix $a > 1$. Then, by (6.9), there is a $c > 1$ such that, for all $r, T \ge 1$,

$$\inf\{p_{t+T}^\mu(x,e) : x \in U^{ar}, a^{-2}r^2 \le t \le a^2 r^2\} \ge \frac{1}{c}(a^2r^2 + T)^{-D/2}$$

and

$$\sup\{p_T^\mu(x,e) : x \notin U^r\} \le cT^{-D/2} \exp\left(-\frac{r^2}{cT}\right).$$

Also, by (9.1) there is a constant $c' > 0$ such that, for all $t, T \ge 1$,

$$\|\mu^n p_T^\mu - p_{n+T}^\mu\|_\infty \le c'\, T^{-(D+1)/2}.$$

It follows that, for all $(n, x) \in [\![a^{-2}r^2, a^2r^2]\!] \times U^{ar}$,

$$\sum_{y\in U^r} \mu^n(x,y)$$
$$\geq \frac{1}{c}T^{D/2}\left(\frac{1}{c}(a^2r^2+T)^{-D/2} - c'T^{-(D+1)/2} - cT^{-D/2}\exp\left(-\frac{r^2}{cT}\right)\right).$$

If $T = \varepsilon r^2$, for some $\varepsilon \in (0,1)$, then we have

$$\begin{aligned}\sum_{y\in U^r} \mu^n(x,y) &\geq \frac{1}{c}\varepsilon^{D/2}\left(\frac{1}{c}(a^2+\varepsilon)^{-D/2} - c\varepsilon^{-(D+1)/2}r^{-1} - c\varepsilon^{-D/2}\exp\left(-\frac{1}{c\varepsilon}\right)\right)\\ &\geq \frac{1}{c}\varepsilon^{D/2}\bigg(\frac{1}{c}(a^2+1)^{-D/2}\\ &\qquad - c\varepsilon^{-(D+1)/2}r^{-1} - c\varepsilon^{-D/2}\exp\left(-\frac{1}{c\varepsilon}\right)\bigg). \qquad \square\end{aligned}$$

The proposition follows by choosing first ε small enough and next r large enough.

9.2. *Proof of Proposition* 1.4. Define V as in Section 5.3, let $a_1 > a_2 \geq 1$ and set

$$M = \sup\left\{\int_{V^{a_2r}} p^{\mu}_{r^2}(y,z)\,dz : y \notin U^{a_1r}\right\}.$$

We have, for $r \in \mathbb{N}$,

$$\begin{aligned}\sum_{y\in U^{a_1r}} \mu^{r^2}(e,y) &\geq \sum_{y\in U^{a_1r}} \int_{V^{a_2r}} \mu^{r^2}(e,y)p^{\mu}_{r^2}(y,z)\,dz\\ &= \int_{V^{a_2r}} \mu^{r^2}p^{\mu}_{r^2}(e,z)\,dz - \sum_{\{y\notin U^{a_1r}\}}\int_{V^{a_2r}} \mu^{r^2}(e,y)p^{\mu}_{r^2}(y,z)\,dz\\ &\geq \int_{V^{a_2r}} \mu^{r^2}p^{\mu}_{r^2}(e,z)\,dz - M \qquad (9.4)\\ &= \int_{V^{a_2r}} p^{\mu}_{2r^2}(e,z)\,dz + \int_{V^{a_2r}} (\mu^{r^2}p^{\mu}_{r^2} - p^{\mu}_{2r^2})(e,z)\,dz - M\\ &\geq \int_{V^{a_2r}} p^{\mu}_{2r^2}(e,z)\,dz - \|\mu^{r^2}p^{\mu}_{r^2} - p^{\mu}_{2r^2}\|_\infty |V^{a_2r}| - M.\end{aligned}$$

Now, by (9.1) there is a $c_1 > 0$ such that, for all $r \in \mathbb{N}$,

$$\|\mu^{r^2}p^{\mu}_{r^2} - p^{\mu}_{2r^2}\|_\infty \leq c_1\, r^{-D-1}. \tag{9.5}$$

Also, there is a $c_2 > 0$ such that

$$|V^r| \leq c_2 r^D, \qquad r \in \mathbb{N}. \tag{9.6}$$

Combining (9.4), (9.5) and (9.6) we have

$$\sum_{y\in U^{a_1 r}} \mu^{r^2}(e,y) \geq \int_{V^{a_2 r}} p^{\mu}_{2r^2}(e,z)\,dz - c_1 c_2 a_2^D r^{-1} - M. \tag{9.7}$$

Now let us fix $\delta > 0$ and let us chose $a_2 > 1$ such that, for all $r \geq 1$,

$$\int_{V^{a_2 r}} p^{\mu}_{2r^2}(e,z)\,dz \geq 1 - \frac{\delta}{3}.$$

Let us choose $a_1 > a_2$ so that $M \leq \partial/3$ and $r_0 \geq 1$ so that

$$c_1 c^2 a_2^D r_0^{-1} \leq \frac{\delta}{3}.$$

Then it follows from (9.7) that

$$\sum_{y\in U^{a_1 r}} \mu^{r^2}(e,y) \geq 1 - \delta$$

and the proposition follows. □

10. The proof of the Gaussian estimate when Γ is nilpotent. In this section we give the proof of the Gaussian estimate (1.14) when the group Γ is nilpotent. Note that there is no loss of generality if we assume that the torsion subgroup $\tau(\Gamma)$ is trivial and hence that $\Gamma = \Gamma_N$.

10.1. *The functions* $\rho_k,\ k \geq 1$. Let $\rho(x)$ be as in Section 7 and let the family of dilations τ_ε, $\varepsilon > 0$, be as in Section 5.2.

Let V be a compact neighborhood of the identity element e of N, as in Section 5.3, and let $|\cdot|_N$ be defined as in (1.7). Then there is a $C > 1$ such that

$$\tau_{1/Ck} V^k \subseteq V, \qquad k \in \mathbb{N}. \tag{10.1}$$

Let $0 \leq \phi \in C^\infty(N)$, such that

$$\phi(x) = \begin{cases} 0, & \text{for } |x|_N \leq 1, \\ 1, & \text{for } |x|_N \geq 4, \end{cases}$$

and set

$$\phi_k(x) = \phi(\tau_{1/C\sqrt{k}} x), \qquad k \geq 1.$$

Then

$$\phi_k(x) = 0, \qquad |x|_N \leq \sqrt{k}, \tag{10.2}$$

and there is a constant $\zeta > 0$ such that

$$\phi_k(x) = 1, \qquad |x|_N \geq \zeta\sqrt{k}. \tag{10.3}$$

Also, if $\{X_1, \dots, X_q\}$ is the basis of $\mathfrak{n}$ introduced in Section 5.1, then for every $n \in \mathbb{N}$ there is a constant $c > 0$ such that

$$|X_{i_1} X_{i_2} \cdots X_{i_n} \phi_k(x)| \le c k^{-(\sigma(i_1)+\cdots+\sigma(i_n))/2} \tag{10.4}$$

for all $x \in N$ and $1 \le i_j \le q$, $1 \le j \le n$.

We set

$$\rho_k(x) = \phi_k(x)\rho(x).$$

In the next lemma we gather the properties of the functions $\rho_k(x)$ that we need in the proof of Gaussian estimate (1.14). These properties are immediate consequences of (7.1).

LEMMA 10.1. *For all $n \in \mathbb{N}$ there is a constant $c \ge 1$ such that, for all $k \ge 1$ and all $1 \le i_j \le q$, $1 \le j \le n$,*

$$\begin{gathered} \rho_k(x) \ge 0, \qquad x \in N, \\ \rho_k(x) = 0 \qquad \textit{for } |x|_N \le \sqrt{k}, \\ \frac{1}{c}\,|x|_N \le \rho_k(x) \le c\,|x|_N \qquad \textit{for } |x| \ge \zeta\sqrt{k}, \\ |X_{i_1} \cdots X_{i_n} \rho_k(x)| \le \frac{c}{|x|_N^{\sigma(i_1)+\cdots+\sigma(i_n)-1}}, \qquad x \in N. \end{gathered} \tag{10.5}$$

10.2. *The functions H_k, $k \ge 1$.* For fixed constants $A > 0$ and $B > 0$ we consider the family of functions H_k, $k \ge 1$, defined by

$$H_k(t,x) = \exp\biggl(-\frac{(\rho_k(x) + B\sqrt{k})^2}{A(k+t)}\biggr), \qquad t \ge 0,\ x \in N.$$

LEMMA 10.2. *There are constants $A > 0$ and $B > 0$ such that*

$$H_k(t+1,x) > \mu H_k(t,x) \tag{10.6}$$

for all $(t,x) \in [0,k] \times N$ and $k \ge 1$.

PROOF. We observe that

$$H_k(t+1,x) - \mu H_k(t,x) = H_k(t+1,x) - H_k(t,x) + (I-\mu)H_k(t,x)$$

and that

$$H_k(t+1,x) - H_k(t,x) \ge \inf_{t \le s \le t+1} \frac{\partial}{\partial s} H_k(s,x). \tag{10.7}$$

Let V be a compact neighborhood of e, as in Section 5.3, and let us assume that $\operatorname{supp}\mu \subseteq V$. Then, by Lemma 8.1,

$$|(I-\mu)H_k(t,x)| \le c \sup_{y \in xV} |\nabla_X^2 H_k(t,y)|. \tag{10.8}$$

We have

$$\frac{\partial}{\partial t}H_k(t,x)=\frac{1}{A}\frac{1}{k+t}\frac{(\rho_k(x)+B\sqrt{k})^2}{k+t}H_k(t,x). \tag{10.9}$$

Also, for all $X, Y \in \mathfrak{n}$,

$$\begin{aligned} YH_k(t,x) &= -\frac{1}{A}\frac{1}{k+t}2\big(\rho_k(x)+B\sqrt{k}\big)Y\rho_k(x)\ H_k(t,x),\\ XYH_k(t,x) &= -\frac{1}{A}\frac{1}{k+t}2X\rho_k(x)Y\rho_k(x)\ H_k(t,x)\\ &\quad -\frac{1}{A}\frac{1}{k+t}2\big(\rho_k(x)+B\sqrt{k}\big)XY\rho_k(x)\ H_k(t,x)\\ &\quad +\frac{1}{A^2}\frac{1}{(k+t)^2}4\big(\rho_k(x)+B\sqrt{k}\big)^2X\rho_k(x)Y\rho(x)\ H_k(t,x). \end{aligned} \tag{10.10}$$

Case I ($|x|_N \le \sqrt{k}-1$ and $t\in[0,k]$). By construction, for all $|x|_N\le\sqrt{k}$,

$$H_k(t,x)=\exp\Big(-\frac{B^2k}{A(k+t)}\Big).$$

Hence, for all $|x|_N\le\sqrt{k}-1$,

$$H_k(t+1,x)>H_k(t,x)=\big(\mu H_{k+1}(t,\cdot)\big)(x).$$

Case II ($\sqrt{k}-1\le|x|_N\le\zeta\sqrt{k}+1$ and $0\le t\le k$). By (10.7) and (10.9) there is a $c_1>0$ such that

$$\begin{aligned} H_k(t+1,x)-H_k(t,x) &\ge \frac{1}{A}\frac{1}{k+t+1}\frac{B^2k}{k+t+1}H_k(t,x)\\ &\ge \frac{1}{A}\frac{1}{k+t}\,c_1\frac{B^2}{2}H_k(t,x). \end{aligned}$$

Also, by (10.8) and (10.10) there is a constant $c_2>0$ such that

$$\begin{aligned} |(I-\mu)H_k(t,x)| &\le \frac{1}{A}\frac{1}{k+t}\Big[(c_2\sqrt{k}+B\sqrt{k})c_2\frac{1}{\sqrt{k}}+c_2\\ &\qquad +(c_2\sqrt{k}+B\sqrt{k})\frac{1}{\sqrt{k}}\\ &\qquad +\frac{1}{A}\frac{1}{k}(c_2\sqrt{k}+B\sqrt{k})^2\Big]\ H_k(t,x)\\ &\le \frac{1}{A}\frac{1}{k+t}\Big[c_2(c_2+B)+c_2+(c_2+B)\\ &\qquad +\frac{1}{A}(c_2+B)^2\Big]\ H_k(t,x). \end{aligned}$$

Hence

$$
\begin{aligned}
&H_k(t+1,x) - \mu H_k(t,x) \\
&\qquad \geq \frac{1}{A}\frac{1}{k+t}\left[c_1\frac{B^2}{2} - c_2(c_2+B) - c_2 - (c_2+B) - \frac{1}{A}(c_2+B)^2\right] H_k(t,x).
\end{aligned}
$$

So, by choosing B large enough that

$$
c_1\frac{B^2}{4} > c_2 + c_2(c_2+B) + c_2(c_2+B)
$$

and A large enough that

$$
c_1\frac{B^2}{4} > \frac{1}{A}(c_2+B)^2
$$

we have

$$
H_k(t+1,x) > \mu H_k(t,x).
$$

Case III ($|x|_N > \zeta\sqrt{k}+1$ and $0 \leq t \leq k$). By (10.7) and (10.9), there is a $c_1 > 0$ such that

$$
H_k(t+1,x) - H_k(t,x) \geq \frac{1}{A}\frac{1}{k+t}\frac{(c_1|x|_N + B\sqrt{k})^2}{2k} H_k(t,x).
$$

Also, by (10.8) and (10.10) there is a $c_2 > 0$ such that

$$
\begin{aligned}
&|(I-\mu)H_k(t,x)| \\
&\quad \leq \frac{1}{A}\frac{1}{k+t}\left[c_2(c_2|x|_N + B\sqrt{k})\frac{1}{|x|_N} + c_2 + c_2(c_2|x|_N + B\sqrt{k})\frac{1}{|x|_N}\right. \\
&\qquad\qquad \left. + c_2\frac{1}{A}\frac{1}{k}(c_2|x|_N + B\sqrt{k})^2\right] H_k(t,x) \\
&\quad = \frac{1}{A}\frac{1}{k+t}\left[c_2^2 + c_2B\frac{\sqrt{k}}{|x|_N}c_2 + c_2^2 + c_2B\frac{\sqrt{k}}{|x|_N}\right. \\
&\qquad\qquad \left. + c_2\frac{1}{A}\frac{(c_2|x|_N + B\sqrt{k})^2}{k}\right] H_k(t,x) \\
&\quad \leq \frac{1}{A}\frac{1}{k+t}\left[c_2^2 + c_2^2B + c_2^2 + c_2B + c_2\frac{1}{A}\frac{(c_2|x|_N + B\sqrt{k})^2}{k}\right] H_k(t,x).
\end{aligned}
$$

Hence

$$
\begin{aligned}
&H_k(t+1,x) - \mu H_k(t,x) \\
&\qquad \geq \frac{1}{A}\frac{1}{k+t}\left[\frac{(c_1|x|_N + B\sqrt{k})^2}{2k} - c_2^2 - c_2^2B - c_2^2 - c_2B\right. \\
&\qquad\qquad \left. - c_2\frac{1}{A}\frac{(c_2|x|_N + B\sqrt{k})^2}{k}\right] H_k(t,x).
\end{aligned}
$$

So, by choosing B large enough that

$$\frac{(c_1|x|_N + B\sqrt{k})^2}{4k} > c_2 + c_2^2 + c_2 + Bc_2^2 + c_2 B$$

and A large enough that

$$\frac{1}{4}(c_1|x|_N + B\sqrt{k})^2 > c_2\frac{1}{A}(c_2|x|_N + B\sqrt{k})^2$$

we have

$$H_k(t+1, x) > \mu H_k(t, x). \qquad \square$$

10.3. *Proof of Theorem* 1.8. It is enough to prove that there is a $c > 0$ such that

$$\mu^n(x, e) \leq cn^{-D/2}\exp\left(-\frac{|x|_N^2}{cn}\right), \qquad n \geq 1,\ x \in N. \tag{10.11}$$

Let us fix constants $A > 0$ and $B > 0$ such that the family of functions

$$H_k(t, x) = \exp\left(-\frac{(\rho_k(x) + B\sqrt{k})^2}{A(k+t)}\right), \qquad k \geq 1,$$

satisfy (10.6).

Let us consider the function

$$u(n, x) = \sum_{y \in U^{\sqrt{k}}} \mu^n(x, y), \qquad x \in \Gamma_N,\ n \in \mathbb{N}.$$

Let us also fix a constant $C > 0$ such that

$$CH_k(0, x) > 1, \qquad |x|_N \leq 3\sqrt{k},$$

and consider the function

$$F(n, x) = CH_k(n, x) - u(n, x).$$

Then $F(n, x)$ satisfies

$$\begin{aligned} F(n+1, x) &> \mu F(n, x), \qquad x \in \Gamma_N,\ n \in [0, k] \\ F(0, x) &> 0, \qquad x \in \Gamma_N, \end{aligned}$$

and hence

$$F(t, x) > 0, \qquad t \in [0, k],\ x \in \Gamma_N. \tag{10.12}$$

It follows that, for all $x \in \Gamma_N$ and $k \geq 1$,

$$\sum_{y \in U^{\sqrt{k}}} \mu^k(x, y) \leq C \exp\left(-\frac{(\rho_k(x) + B\sqrt{k})^2}{2Ak}\right).$$

On the other hand, it follows from (1.9) that there are $\beta \in \mathbb{N}$ and $\lambda > 0$ such that, for all $x \in \Gamma_N$ and $n \in \mathbb{N}$,

$$\mu^n(x, e) \leq \lambda \ \inf\{\mu^{\beta n}(x, y), y \in U^{\sqrt{n}}\}.$$

Since $|U^n| \leq cn^D$, $n \in \mathbb{N}$, we have that

$$\begin{aligned}\mu^n(x, e) &\leq \lambda \frac{1}{|U^{\sqrt{n}}|} \sum_{y \in U^{\sqrt{n}}} \mu^{\beta n}(x, y) \\ &\leq \lambda c n^{-D/2} C \exp\left(-\frac{(\rho_{\beta n}(x) + B\sqrt{\beta n})^2}{2A\beta n}\right)\end{aligned}$$

for all $x \in \Gamma_N$ and $n \geq 1$. This proves (10.11) and the theorem follows. $\square$

11. The proof of the Berry–Esseen estimate when Γ is nilpotent. In this section we assume that $\Gamma = \Gamma_N$. If Γ is nilpotent and the torsion subgroup $\tau(\Gamma)$ is not trivial, then we can just extend the different kernels from $\Gamma_N = \Gamma/\tau(\Gamma)$ to Γ, as explained in Section 1.4, and then the proofs remain exactly the same.

Let L_μ be the centered left invariant sub-Laplacian associated with μ and let $p_t^\mu(x, y)$ be its heat kernel L_μ.

By (4.2) there are $\gamma \in (0, 1]$ and $c > 0$ such that

$$\|\nabla_U \mu^n\|_\infty \leq cn^{-(D+\gamma)/2}, \qquad n \in \mathbb{N}. \tag{11.1}$$

In this section we prove the following Berry–Esseen estimate (cf. [20, 34]):

THEOREM 11.1. *There is a $c > 1$ such that, for all $x, y \in \Gamma_N$ and $n \in \mathbb{N}$,*

$$|\mu^n(x, y) - p_n^\mu(x, y)| \leq cn^{-(D+\gamma)/2}. \tag{11.2}$$

Once we have proved Theorem 1.14, then (11.1) and hence (11.2) will hold with $\gamma = 1$.

For the case when μ is symmetric, the above result was proved for $\gamma = 1$ in [2]. We give below an adaptation of that proof.

Let Ω be a fundametal domain for Γ_N (see Section 1.3) and let

$$S_t(x, y) = \int_\Omega p_t^\mu(xh, y)\, dh, \qquad x, y \in \Gamma_N.$$

The proof of (11.2) is based on the following two lemmas, which are inspired by [14] (see also [7, 41]).

LEMMA 11.2. *There are constants $a, b \geq 1$ such that, for all $T \geq 1$ and $n \in \mathbb{N}$,*

$$\|\mu^n - p_n^\mu\|_\infty \leq a\|(\mu^n - p_n^\mu)S_T\|_\infty + b\sqrt{T}n^{-(D+\gamma)/2}. \tag{11.3}$$

LEMMA 11.3. *There is a constant $c \geq 1$ such that if, for some $n \in \mathbb{N}$,*

$$\|\mu^k - p_k^\mu\|_\infty \leq Ak^{-(D+\gamma)/2} \qquad \text{for all } 1 \leq k \leq n-1, \tag{11.4}$$

then

$$\|(\mu^n - p_n^\mu)S_T\|_\infty \leq c\left(1 + \frac{A}{\sqrt{T}}\right)n^{-(D+\gamma)/2}. \tag{11.5}$$

PROOF OF THEOREM 11.1. If

$$\|\mu^k - p_k^\mu\|_\infty \leq Ak^{-(D+\gamma)/2}, \qquad 1 \leq k \leq n-1,$$

then by (11.3) and (11.5),

$$\begin{aligned}\|\mu^n - p_n^\mu\|_\infty &\leq \alpha c\left(1 + \frac{A}{\sqrt{T}}\right)n^{-(D+\gamma)/2} + b\sqrt{T}n^{-(D+\gamma)/2}\\ &\leq \left(ac + ac\frac{A}{\sqrt{T}} + b\sqrt{T}\right)n^{-(D+\gamma)/2}.\end{aligned}$$

So (11.2) can be proved by induction provided that for all A large enough there is a $T \geq 1$ such that

$$ac + ac\frac{A}{\sqrt{T}} + b\sqrt{T} \leq A. \tag{11.6}$$

To this end, let us consider the function

$$\varphi(x) = ac + acA\frac{1}{x} + bx$$

(note that $a, b, c \geq 1$) and take

$$A \geq 9a^2b^2c^2 \quad \text{and} \quad T = \frac{acA}{b}.$$

Then we have

$$\varphi(\sqrt{T}) \leq A.$$

This proves (11.6) and the theorem follows. $\square$

11.1. *Proof of Lemma* 11.2. Let us set

$$H_n(x,y) = \mu^n(x,y) - p_n^\mu(x,y)$$

and assume that

$$-\|H_n\|_\infty = \min\{H_n(x,y),\ x,y \in \Gamma_N\}$$

(the case $\|H_n\|_\infty = \max\{H_n(x,y),\ x,y \in \Gamma_N\}$ can be treated in the same way). There are $x_0, y_0 \in \Gamma_N$ such that

$$H_n(x_0,y_0) = -\|H_n\|_\infty.$$

Then

$$\begin{aligned}
-\|H_n S_T\|_\infty &\leq \sum_{z\in\Gamma_N} H_n(x_0,z)S_T(z,y_0)\\
&= H_n(x_0,y_0) \sum_{|y_0^{-1}z|_{\Gamma_N}\leq c\sqrt{T}} S_T(z,y_0)\\
&\quad + \sum_{|y_0^{-1}z|_{\Gamma_N}\leq c\sqrt{T}} [H_n(x_0,z) - H_n(x_0,y_0)]S_T(z,y_0)\\
&\quad + \sum_{|y_0^{-1}z|_{\Gamma_N}\geq c\sqrt{T}} H_n(x_0,z)S_T(z,y_0)\\
&\leq -\|H_n\|_\infty \sum_{|y_0^{-1}z|_{\Gamma_N}\leq c\sqrt{T}} S_T(z,y_0)\\
&\quad + c\sqrt{T}\|\nabla_U H_n(x_0,\cdot)\|_\infty \sum_{|y_0^{-1}z|_{\Gamma_N}\leq c\sqrt{T}} S_T(z,y_0)\\
&\quad + \|H_n\|_\infty \sum_{|y_0^{-1}z|_{\Gamma_N}\geq c\sqrt{T}} S_T(z,y_0).
\end{aligned}$$

Hence, if

$$\lambda = \sum_{|y_0^{-1}z|_{\Gamma_N}\leq c\sqrt{T}} S_T(z,y_0),$$

then

$$-\|H_n S_T\|_\infty \leq -\|H_n\|_\infty\lambda + c\sqrt{T}\lambda n^{-(D+\gamma)/2} + \|H_n\|_\infty(1-\lambda),$$

or

$$(2\lambda-1)\|H_n\|_\infty \leq \|H_n S_T\|_\infty + c\lambda\sqrt{T}\lambda n^{-(D+\gamma)/2}.$$

By choosing c large enough, so that $\lambda > 1/2$, we get

$$\|H_n\|_\infty \leq \frac{1}{2\lambda-1}\|H_n S_T\|_\infty + \frac{c\lambda}{2\lambda-1}\sqrt{T}n^{-(D+\gamma)/2},$$

which proves the lemma. $\square$

11.2. *Proof of Lemma* 11.3. It follows from (6.10) that there is a $c > 0$ such that, for all $x, y \in \Gamma_N$ and $t \geq 1$,

$$|p^{\mu}_{t+T}(x, y) - p^{\mu}_t S_T(x, y)| \leq ct^{-(D+1)/2}.$$

So, it is enough to prove that

$$\|\mu^n S_T - p^{\mu}_{n+T}\|_\infty \leq c\left(1 + \frac{A}{\sqrt{T}}\right)n^{-(D+\gamma)/2}. \tag{11.7}$$

We have

$$\begin{aligned}
p^{\mu}_{n+T} - \mu^n S_T &= p^{\mu}_{n+T} - \mu^{n-1} p^{\mu}_{1+T} + \mu^{n-1} p^{\mu}_{1+T} - \mu^n S_T \\
&= \sum_{0 \leq i \leq n-2} \mu^i (p^{\mu}_{n-i+T} - \mu p^{\mu}_{n-i-1+T}) + \mu^{n-1}(p^{\mu}_{1+T} - \mu S_T) \\
&= \sum_{0 \leq i \leq n/2} \mu^i (p^{\mu}_{n-i+T} - \mu p^{\mu}_{n-i-1+T}) \\
&\quad + \sum_{n/2 < i \leq n-2} (\mu^i - p^{\mu}_i)(p^{\mu}_{n-i+T} - \mu p^{\mu}_{n-i-1+T}) \\
&\quad + (\mu^{n-1} - p^{\mu}_{n-1})(p^{\mu}_{1+T} - \mu S_T) \\
&\quad + \sum_{n/2 < i \leq n-2} p^{\mu}_i (p^{\mu}_{n-i+T} - \mu p^{\mu}_{n-i-1+T}) \\
&\quad + p^{\mu}_{n-1}(p^{\mu}_{1+T} - \mu S_T).
\end{aligned} \tag{11.8}$$

Now, by (6.10) and (8.3),

$$\begin{aligned}
&\sum_{0 \leq i \leq n/2} \|\mu^i\|_1 \|p^{\mu}_{n-i+T} - \mu p^{\mu}_{n-i-1+T}\|_\infty \\
&\qquad \leq c \sum_{0 \leq i \leq n/2} (n - i - 1 + T)^{-(D+3)/2} \\
&\qquad \leq c\left(\frac{n}{2} + T\right)^{-(D+1)/2} \\
&\qquad \leq cn^{-(D+1)/2}.
\end{aligned} \tag{11.9}$$

By the induction hypothesis (11.4),

$$\begin{aligned}
&\sum_{n/2 < i \leq n-2} \|\mu^i - p^{\mu}_i\|_\infty \|p^{\mu}_{n-i+T} - \mu p^{\mu}_{n-i-1+T}\|_1 \\
&\quad \leq \sum_{n/2 < i \leq n-2} A i^{-(D+\gamma/2}(n - i - 1 + T)^{-3/2} \\
&\quad \leq cA \frac{1}{\sqrt{T}} n^{-(D+\gamma)/2}.
\end{aligned} \tag{11.10}$$

To estimate the term $p^{\mu}_{1+T} - \mu S_T$, let us consider the function $\varphi = \sum_{g\in\Gamma_N} \mu(g) \times \mathbb{1}_{g\Omega}$. Then

$$\mu S_T(x, y) = \int \varphi(h) p^{\mu}_T(xh, y)\, dh, \qquad x, y \in \Gamma_N.$$

Since $\|\varphi\|_1 = 1$, it follows from (6.10) that

$$\|\mu S_T - p^{\mu}_T\|_1 \le \frac{c}{\sqrt{T}}, \qquad T \ge 1.$$

Since we also have

$$\|p^{\mu}_{T+1} - p^{\mu}_T\|_1 \le \frac{c}{T}, \qquad T \ge 1,$$

we conclude that

$$\begin{aligned} \|(\mu^{n-1} - p^{\mu}_{n-1})(p^{\mu}_{1+T} - \mu S_T)\|_\infty &\le \|\mu^{n-1} - p^{\mu}_{n-1}\|_\infty \|p^{\mu}_{1+T} - \mu S_T\|_1 \\ &\le cA\frac{1}{\sqrt{T}} n^{-(D+\gamma)/2}. \end{aligned} \tag{11.11}$$

To estimate the remaining term in (11.8) we observe that

$$\begin{aligned} &\sum_{n/2<i\le n-2} p^{\mu}_i (p^{\mu}_{n-i+T} - \mu p^{\mu}_{n-i-1+T}) + p^{\mu}_{n-1}(p^{\mu}_{1+T} - \mu S_T) \\ &\quad = p^{\mu}_{[n/2]+1} p^{\mu}_{n-[n/2]-1+T} - p^{\mu}_{n-1}\mu S_T \\ &\qquad + \sum_{n/2<i\le n-2} (p^{\mu}_i \mu - p^{\mu}_{i+1}) p^{\mu}_{n-i-1+T}. \end{aligned} \tag{11.12}$$

By (6.10) and (8.3),

$$\begin{aligned} \sum_{n/2<i\le n-2} \|p^{\mu}_i \mu - p^{\mu}_{i+1}\|_\infty \|p^{\mu}_{n-i-1+T}\|_1 &\le \sum_{n/2<i\le n-2} c i^{-(D+3)/2} \\ &\le cn^{-(D+1)/2}. \end{aligned} \tag{11.13}$$

Also, by (6.10), for all $x, y \in \Gamma_N$,

$$\left| p^{\mu}_{n+T}(x, y) - \sum_{z\in\Gamma_N} p^{\mu}_{[n/2]+1}(x, z) p^{\mu}_{n-[n/2]-1+T}(z, y) \right| \le cn^{-(D+1)/2}$$

and

$$\left| p^{\mu}_{n+T}(x, y) - \sum_{z\in\Gamma_N} p^{\mu}_{n-1}(x, z)(\mu S_T)(z, y) \right| \le cn^{-(D+1)/2}.$$

Hence

$$\|p^{\mu}_{[n/2]+1} p^{\mu}_{n-[n/2]-1+T} - p^{\mu}_{n-1}\mu S_T\|_\infty \le cn^{-(D+1)/2}. \tag{11.14}$$

The lemma follows by summing (11.9)–(11.14). $\square$

12. Proof of the Taylor formula. In this section we give the proof of the Taylor formula (1.18) under the assumption that the Berry–Esseen estimate (11.2) holds.

We assume that $\Gamma_N = \Gamma$. The proof in the general case under a similar assumption (cf. Section 17) is exactly the same.

If Γ is nilpotent and the torsion subgroup $\tau(\Gamma)$ is not trivial, then we can just extend the different kernels and functions from $\Gamma_N = \Gamma/\tau(\Gamma)$ to Γ, as explained in Section 1.4, and then the proofs remain exactly the same.

12.1. *Polynomials on* Γ_N. We use the notation of Sections 5 and 6.

Let us fix a monomial $P(t,x) = t^{i_0} x_1^{i_1} \cdots x_q^{i_q}$. Then, by Corollary 8.3,

$$\big(\partial_1 + (I-\mu)\big) P(t,x) = \left(\frac{\partial}{\partial t} + L_\mu\right) P(t,x) + Q(t,x),$$

where $Q(t,x)$ is a polynomial satisfying

$$\deg_H Q \le \deg_H P - 3. \tag{12.1}$$

On the other hand, it follows from (6.5) and (6.6) that, given any monomial $P(t,x)$ as above, there is a polynomial $W(t,x)$ such that

$$\left(\frac{\partial}{\partial t} + L_\mu\right) W(t,x) = P(t,x)$$

and

$$\deg_H W(t,x) = \deg_H P(t,x) + 2. \tag{12.2}$$

The following proposition is a consequence of the above observations.

PROPOSITION 12.1. *Given any monomial* $P(t,x)$ *on* N, *there is a polynomial* $P^\mu(t,x)$ *satisfying*

$$\begin{gathered} \left(\frac{\partial}{\partial t} + L_\mu\right) P = \big(\partial_1 + (I-\mu)\big) P^\mu, \\ P^\mu = P + W, \\ \deg_H W \le \deg_H P - 1. \end{gathered} \tag{12.3}$$

Let the polynomials P_i, $i = 0,1,2,\dots$, be as in Section 6.1 and let us associate with those monomials and the sub-Laplacian L_μ, polynomials Q_{P_i}, $i = 0,1,2,\dots$, satisfying (6.6). We associate with the polynomials Q_{P_i}, $i = 0,1,2,\dots$, and fix, polynomials $Q^\mu_{P_i}$, $i = 0,1,2,\dots$, satisfying (12.3) above.

Note that for $0 \le i \le \nu_2$ we can take $Q^\mu_{P_i} = P_i$. Note also that $\nu_0 = 0$, $\nu_1 = n_1$ and $n_1 \le n_2 < \nu_2$. So we assume that $P_0(t,x) = Q^\mu_{P_0}(t,x) = 1$, that $P_i(t,x) = Q^\mu_{P_i}(t,x) = x_i$, for $1 \le i \le \nu_2$, and that $P_i = Q^\mu_{P_i}$, for $\nu_1 < i \le \nu_2$.

12.2. *A uniform approximation of a space–time μ-harmonic function by an L_μ-heat function.* In this section we use the Berry–Esseen estimate (11.2) to prove the following lemma:

LEMMA 12.2. *Let V be as in Section 5.3 and let us assume that $\operatorname{supp}\mu \subseteq V$. Then there is a constant $c > 0$ and $\beta \in (0,1)$ such that for all $\Theta_1 \geq 4\Theta_2$, $\Theta_2 \geq 2$, $r \geq 1$ and all functions u satisfying*

$$(\partial_1 + (I - \mu))u = 0 \qquad \text{in } [\![-\Theta_1^2 r^2, \Theta_1^2 r^2]\!] \times U^{\Theta_1 r}$$

we can associate a function u^μ satisfying

$$\left(\frac{\partial}{\partial t} + L_\mu\right)u^\mu = 0 \qquad \text{in } (-\Theta_2^2 r^2, \Theta_2^2 r^2) \times V^{\Theta_1 r}$$

as well as $\|u^\mu\|_\infty \leq \|u\|_\infty$ and

$$\sup\{|u - u^\mu|; [\![-r^2, r^2]\!] \times U^r\} \leq c\Theta_2^{-\beta} r^{-\beta} + ce^{-\Theta_1^2/c\Theta_2^2}. \tag{12.4}$$

As in Section 3.2, we use Z_n to denote the right random walk with transition kernel $\mu(x,y)$ and by τ_r^x the stopping time

$$\tau_r^x = \inf\{n : Z_n \notin xU^r\}.$$

Using the Gaussian estimate (1.14), we can obtain the following improvement of Lemma 3.4.

LEMMA 12.3. *There is a constant $c > 0$ such that, for all $r, n \in \mathbb{N}$,*

$$P_x[\tau_r^x \leq n] \leq c\exp\left(-\frac{r^2}{cn}\right). \tag{12.5}$$

PROOF. The proof follows the same lines as the proof of Lemma 3.4.
Let $a \geq 1$ such that $xU^r \cap yU^r = \varnothing$ when $y \notin xU^{ar}$. We have

$$\begin{aligned}
\sum_{y \notin xU^r} \mu^n(x,y) &= P_x[Z_n \notin xU^r] \\
&\geq E^{P_x}\big[\mu^{n-\tau_{ar}^x}(Z_{\tau_{ar}^x}, \Gamma_N \setminus xU^r); \tau_{ar}^x \leq n\big] \\
&\geq E^{P_x}\big[\mu^{n-\tau_{ar}^x}(Z_{\tau_{ar}^x}, Z_{\tau_{ar}^x} U^r); \tau_{ar}^x \leq n\big] \\
&\geq E^{P_x}\big[\mu^{n-\tau_{ar}^x}(e, U^r); \tau_{ar}^x \leq n\big].
\end{aligned}$$

Now, we observe that (12.5) is interesting only for $r^2 \geq n$ and that in that case, by (1.14), there is a $\delta > 0$ such that

$$\mu^{n-\tau_{ar}^x}(e, U^r) \geq \delta.$$

So

$$\sum_{y \notin xU^r} \mu^n(x, y) \geq \delta P_x[\tau_{ar}^x \leq r^2].$$

Since, by (1.14), there is a constant $c > 0$ such that, for all $r, n \in \mathbb{N}$,

$$\sum_{y \notin xU^r} \mu^n(x, y) \leq c \exp\left(-\frac{r^2}{cn}\right),$$

we conclude that

$$P_x[\tau_{ar}^x \leq n] \leq \frac{1}{\delta} c \exp\left(-\frac{r^2}{cn}\right),$$

which proves the lemma. □

LEMMA 12.4. *Let $r, n \in \mathbb{N}$ and let u be a function satisfying*

$$(\partial_1 + (I - \mu))u = 0 \qquad \text{in } [\![-2n, 2n]\!] \times xU^{2r}.$$

Then

$$\left| u(n, x) - \sum_{y \in xU^r} \mu^n(x, y) u(0, y) \right| \leq 2\|u\|_\infty P_x[\tau_r^x \leq n]. \tag{12.6}$$

PROOF. We have

$$\begin{aligned} u(n, x) &= E^{P_x}[u(0, Z_n); \tau_r^x > n] + E^{P_x}[u(\tau_r^x, Z_{\tau_r^x}); \tau_r^x \leq n] \\ &= \sum_{y \in xU^r} \left(\mu^n(x, y) - E^{P_x}\left[\mu^{n-\tau_r^x}(Z_{\tau_r^x}, y); \tau_r^x \leq n\right]\right) u(0, y) \\ &\quad + E^{P_x}[u(\tau_r^x, Z_{\tau_r^x}); \tau_r^x \leq n]. \end{aligned}$$

Hence

$$\begin{aligned} &\left| u(n, x) - \sum_{y \in xU^r} \mu^n(x, y) u(0, y) \right| \\ &\qquad \leq \|u\|_\infty \left(E^{P_x}\left[\mu^{n-\tau_r^x}(Z_{\tau_r^x}, xU^r); \tau_r^x \leq n\right] + P_x[\tau_r^x \leq n]\right) \\ &\qquad \leq 2\|u\|_\infty P_x[\tau_r^x \leq n] \end{aligned}$$

and the lemma follows. □

PROOF OF LEMMA 12.2. Let u satisfy

$$(\partial_1 + (I - \mu))u = 0 \qquad \text{in } [\![-\Theta_1^2 r^2, \Theta_1^2 r^2]\!] \times xU^{\Theta_1 r}$$

and let us define, for $n > -\Theta_2^2 r^2$ and $x \in \Gamma_N$,

$$u_1(n,x) = \sum_{y \in U^{\Theta_1 r/2}} \mu^{n+\Theta_2^2 r^2}(x,y) u(-\Theta_2^2 r^2, y)$$

and for $t > -\Theta_2^2 r^2$, $x \in N$,

$$u^{\mu}(t,x) = \sum_{y \in U^{\Theta_1 r/2}} p^{\varphi}_{t+\Theta_2^2 r^2}(x,y) u(-\Theta_2^2 r^2, y).$$

Now by (12.5) and (12.6),

$$\begin{aligned}\sup\{|u - u_1|; [\![-\Theta_2^2 r^2, \Theta_2^2 r^2]\!] \times U^{\Theta_2 r}\} &\le 2\|u\|_\infty P_x[\tau^x_{\Theta_1 r/2} \le \Theta_2^2 r^2] \\ &\le c\|u\|_\infty e^{-\Theta_1^2/c\Theta_2^2}.\end{aligned} \tag{12.7}$$

Also, by interpolating the Berry–Esseen estimate (11.2) and the Gaussian estimates (1.14) and (6.9), we have that there is a $\beta \in (0, 1/2)$ such that

$$\|\mu^n - p_n^{\mu}\|_1 \le c n^{-\beta}, \qquad n \in \mathbb{N}.$$

It follows that

$$\sup\{|u_1 - u^{\mu}|; [\![-r^2, r^2]\!] \times U^r\} \le c\|u\|_\infty \Theta_2^{-\beta} r^{-\beta}. \tag{12.8}$$

The lemma follows by summing (12.7) and (12.8). $\square$

12.3. *The iteration argument.* The following lemmas are inspired by Avellaneda and Lin [9, 10].

LEMMA 12.5. *For all $n \in \mathbb{N}$ and $\eta \in (0,1)$ there are $r_0 > 1$, $\Theta > 1$ and $c_n > 0$ such that, for all $r \ge r_0$ and all functions u satisfying*

$$(\partial_1 + (I - \mu))u = 0 \quad \textit{in } [\![-\Theta^2 r^2, \Theta^2 r^2]\!] \times U^{\Theta r}, \tag{12.9}$$

we have

$$\sup\left\{\left| u - \sum_{0 \le i \le \nu_n} A_i (\Theta r)^{-\deg_H P_i} Q^{\mu}_{P_i} \right|; [\![r^2, r^2]\!] \times U^r \right\} < \Theta^{-(n+\eta)} \|u\|_\infty \tag{12.10}$$

where the constants A_i satisfy

$$|A_i| \le c_n (\log \Theta)^{\deg_H P_i} \|u\|_\infty,$$

for all $0 \le i \le \nu_n$, and

$$(\partial_1 + (I - \mu))\left(\sum_{\nu_{d-1} \le i \le \nu_d} A_i Q^{\mu}_{P_i}\right) = 0,$$

for all $1 \le d \le n$.

PROOF. Let us fix $n \in \mathbb{N}$, $\mu \in (0,1)$, $\Theta > 16$ and a function u satisfying (12.9).

Then, by Lemma 12.2 and by taking $\Theta_1 = \Theta$ and $\Theta_2 = \Theta/\log\Theta$, there is a function u^μ satisfying

$$\left(\frac{\partial}{\partial t} + L_\mu\right)u^\mu = 0 \qquad \text{in } \left[-\left(\frac{\Theta}{\log\Theta}\right)^2 r^2, \left(\frac{\Theta}{\log\Theta}\right)^2 r^2\right] \times V^{\Theta r/\log\Theta},$$

as well as

$$\|u^\mu\|_\infty \le \|u\|_\infty$$

and

$$\begin{aligned} \sup\{|u - u^\mu|; [\![-r^2, r^2]\!] \times U^r\} \\ \le c\|u\|_\infty\big(\Theta^{-\beta}(\log\Theta)^\beta r^{-\beta} + e^{-(\log\Theta)^2/c}\big). \end{aligned} \tag{12.11}$$

Also, by Theorem 6.2,

$$\begin{aligned} &\sup\left\{\left|u^\mu - \sum_{0\le i\le \nu_n} B_i\left(\frac{\Theta}{\log\Theta}r\right)^{-\deg_H P_i} Q_{P_i}\right|; [-r^2, r^2] \times V^r\right\} \\ &\quad < c_n\left(\frac{\Theta}{\log\Theta}\right)^{-(n+1)}\|u^\mu\|_\infty, \end{aligned} \tag{12.12}$$

where the constants B_i satisfy

$$|B_i| \le c_n\|u^\mu\|_\infty,$$

for all $0 \le i \le \nu_n$, and

$$\left(\frac{\partial}{\partial t} + L_\mu\right)\left(\sum_{\nu_{d-1}\le i\le \nu_d} B_i Q_{P_i}(t,x)\right) = 0,$$

for all $1 \le d \le n$.

Now let us observe that there is a constant $c > 0$ such that

$$\sup\{|Q_{P_i} - Q^\mu_{P_i}|; [-r^2, r^2] \times V^r\} \le c r^{\deg_H P_i - 1}$$

and hence

$$\begin{aligned} &\left(\frac{\Theta}{\log\Theta}r\right)^{-\deg_H P_i}\sup\{|Q_{P_i} - Q^\mu_{P_i}|; [-r^2, r^2] \times V^r\} \\ &\quad \le c\left(\frac{\Theta}{\log\Theta}\right)^{-\deg_H P_i} r^{-1}. \end{aligned} \tag{12.13}$$

Let us take

$$A_i = B_i(\log\Theta)^{\deg_H P_i}, \qquad 0 \le i \le \nu_n.$$

Then, combining (12.11), (12.12) and (12.13), we have that

$$
\begin{aligned}
&\sup\left\{\left|u-\sum_{0\le i\le \nu_n} A_i(\Theta r)^{-\deg_H P_i} Q^{\mu}_{P_i}\right|; [\![-r^2,r^2]\!]\times U^r\right\}\\
&\quad\le \sup\left\{\left|u^{\mu}-\sum_{0\le i\le \nu_n} B_i(\log\Theta)^{\deg_H P_i}(\Theta r)^{-\deg_H P_i} Q_{P_i}\right|; [\![-r^2,r^2]\!]\times U^r\right\}\\
&\qquad+\sup\{|u-u^{\mu}|; [\![-r^2,r^2]\!]\times U^r\}\\
&\qquad+\sum_{0\le i\le \nu_n}|B_i|(\log\Theta)^{\deg_H P_i}(\Theta r)^{-\deg_H P_i}\sup\{|Q_{P_i}-Q^{\mu}_{P_i}|; [r^2,r^2]\times U^r\}\\
&\quad\le c_n\Theta^{-(n+1)}(\log\Theta)^{n+1}\|u^{\mu}\|_\infty + c\|u\|_\infty\big((\log\Theta)^{\beta}\Theta^{-\beta}r^{-\beta}+e^{-(\log\Theta)^2/c}\big)\\
&\qquad+c(\log\Theta)^n\Theta^{-1}r^{-1}\sum_{0\le i\le\nu_n}|B_i|\\
&\quad\le c_n\Theta^{-(n+1)}(\log\Theta)^{n+1}\|u\|_\infty + c\|u\|_\infty\big((\log\Theta)^{\beta}\Theta^{-\beta}r^{-\beta}+e^{-(\log\Theta)^2/c}\big)\\
&\qquad+c(\log\Theta)^n\Theta^{-1}r^{-1}c_n\nu_n\|u\|_\infty.
\end{aligned}
$$

The lemma follows by taking Θ and r_0 large enough. □

LEMMA 12.6. *Let μ,Θ and r_0 be as in the previous lemma. Then there is a constant $c_n>0$ such that, for all $m\in\mathbb{N}$, all $r\ge\Theta^{m-1}r_0$ and all functions u satisfying*

$$\big(\partial_1+(I-\mu)\big)u=0 \qquad \text{in } [\![-\Theta^{2m}r^2,\Theta^{2m}r^2]\!]\times U^{\Theta^m r}, \tag{12.14}$$

we have

$$
\begin{aligned}
&\sup\left\{\left|u-\sum_{0\le i\le\nu_n}A_i^m\big(\Theta^m r\big)^{-\deg_H P_i}Q^{\mu}_{P_i}\right|; [\![-r^2,r^2]\!]\times U^r\right\}\\
&\qquad<\Theta^{-m(n+\eta)}\|u\|_\infty,
\end{aligned}
\tag{12.15}
$$

where the constants A_i^m satisfy

$$|A_i^m|\le c_n(\log\Theta)^{\deg_H P_i}\|u\|_\infty,$$

for all $0\le i\le\nu_n$, and

$$\big(\partial_1+(I-\mu)\big)\left(\sum_{\nu_{d-1}<i\le\nu_d}A_i^m Q^{\mu}_{P_i}\right)=0,$$

for all $1\le d\le n$.

PROOF. We prove the lemma by induction on m. For $m=1$ we are in the case of the previous lemma. So let us assume that (12.15) is true for some $m \in \mathbb{N}$. To prove that it is also true for $m+1$, let us assume, for simplicity, that $\|u\|_\infty \le 1$. By the induction hypothesis

$$\sup\left\{\left|u - \sum_{0\le i\le \nu_n} A_i^m (\Theta^{m+1} r)^{-\deg_H P_i} Q_{P_i}^{\mu}\right|; [\![-\Theta^2 r^2, \Theta^2 r^2]\!] \times U^{\Theta r}\right\} \le \Theta^{-m(n+\eta)}.$$

We consider the function

$$w = \Theta^{m(n+\eta)} = \left(u - \sum_{0\le i\le \nu_n} A_i^m (\Theta^{m+1} r)^{-\deg_H P_i} Q_{P_i}^{\mu}\right).$$

Then

$$\big(\partial_1 + (I-\mu)\big) w = 0 \qquad \text{in } [\![-\Theta^2 r^2, \Theta^2 r^2]\!] \times U^{\Theta r}$$

and

$$\sup\{|w|;\ [\![-\Theta^2 r^2, \Theta^2 r^2]\!] \times U^{\Theta r}\} \le 1.$$

So, by Lemma 12.5, we have that

$$\sup\left\{\left|w - \sum_{0\le i\le \nu_n} B_i (\Theta r)^{-\deg_H P_i} Q_{P_i}^{\mu}\right|; [\![-r^2, r^2]\!] \times U^{r}\right\} < \Theta^{-(n+\eta)},$$

where the constants B_i satisfy

$$|B_i| \le c_n (\log \Theta)^{-\deg_H P_i},$$

for all $0 \le i \le \nu_n$, and

$$\big(\partial_1 + (I-\mu)\big)\left(\sum_{\nu_{d-1} < i \le \nu_d} B_i Q_{P_i}^{\mu}\right) = 0,$$

for all $1 \le d \le n$.

So, if we set

$$A_i^{m+1} = A_i^m + \Theta^{-m(n+\mu-\deg_H P_i)} B_i,$$

then we have

$$\sup\left\{\left|u - \sum_{0\le i\le \nu_n} A_i^{m+1} (\Theta^{m+1} r)^{-\deg_H P_i} Q_{P_i}^{\mu}\right|; [\![-r^2, r^2]\!] \times U^{r}\right\} < \Theta^{-(m+1)(n+\eta)},$$

which proves the inductive step and the lemma follows. $\square$

12.4. *The Taylor formula.* The following result is the analogue of Theorem 1.11 when $\Gamma = \Gamma_N$.

THEOREM 12.7. *For all $n \in \mathbb{N}$ there is a constant $c_n > 0$ such that, for all $R \geq r \geq 1$ and all functions u satisfying*

$$(\partial_1 + (I - \mu))u = 0 \qquad \text{in } [\![-R^2, R^2]\!] \times U^R,$$

we have

$$\sup\left\{\left|u - \sum_{0 \leq i \leq \nu_n} A_i R^{-\deg_H P_i} Q^{\mu}_{P_i}\right|; [\![-r^2, r^2]\!] \times U^r\right\} \leq c_n \left(\frac{R}{r}\right)^{-(n+1)} \|u\|_\infty, \tag{12.16}$$

where the constants A_i satisfy

$$|A_i| \leq c_n \|u\|_\infty,$$

for all $0 \leq i \leq \nu_n$, and

$$(\partial_1 + (I - \mu))\left(\sum_{\nu_{d-1} < i \leq \nu_d} A_i Q^{\mu}_{P_i}\right) = 0,$$

for all $1 \leq d \leq n$.

PROOF. If $R \geq r \geq r_0$, then $\Theta^{m-1} r \leq R < \Theta^m r$ for some $m \in \mathbb{N}$ and hence (12.16) follows from Lemma 12.6.

If $R \geq r_0 > r \geq 1$, then $R/r \leq r_0 R/r_0$ and hence (12.16) follows in the same way from Lemma 12.6.

If $r_0 \geq R \geq r \geq 1$, then (12.16) is trivial. □

13. Harmonic functions of polynomial growth. In this section we give the proof of Theorem 1.12. We assume that $\Gamma_N = \Gamma$. The proof in the general case is exactly the same. If Γ is nilpotent and we have set $\Gamma_N = \Gamma/\tau(\Gamma)$, then we can just extend the different polynomials from Γ_N to Γ (see Section 1.4) and then the proof below also works as is.

PROOF OF THEOREM 1.12. Let u be a μ-harmonic function on G which grows polynomially; that is, there are $c > 0$ and $n \in \mathbb{N}$ such that

$$\sup\{|u|; U^r\} \leq c r^n, \qquad r \geq 1. \tag{13.1}$$

Let the polynomials $Q^{\mu}_{P_i}(t, x)$ be as in section 12.1 and let us denote by $Q^{\mu}_{P_i}(x)$ their restrictions to N, that is, $Q^{\mu}_{P_i}(x) = Q^{\mu}_{P_i}(0, x)$, $x \in N$.

By (12.16), there is a $c > 0$ such that, for all $r \geq k$

$$\sup\left\{\left|u - \sum_{0\leq i\leq \nu_n} A_i^r r^{-\deg_H P_i} Q_{P_i}^{\mu}\right|; U^k\right\} < c\left(\frac{r}{k}\right)^{-(n+1)} \|u\|_{L^\infty(U^r)} \leq ck^{n+1}r^{-1}, \tag{13.2}$$

where the constants A_i^r are such that

$$(I-\mu)\left(\sum_{0\leq i\leq \nu_n} A_i^r r^{-\deg_H P_i} Q_{P_i}^{\varphi}\right) = 0.$$

For each $k \in \mathbb{N}$, let us choose $r_k \in \mathbb{N}$ such that

$$ck^{n+1}r_k^{-1} \leq \frac{1}{k}.$$

We set $C_{k,i} = 0$, for $Q_{P_i}^{\mu} = 0$ and

$$C_{k,i} = A_i^{r_k} r_k^{-\deg_H P_i}$$

otherwise. Then (13.1) and (13.2) imply that there is a $c > 0$ such that, for all $k \in \mathbb{N}$,

$$\sup\left\{\left|u - \sum_{0\leq i\leq \nu_n} C_{k,i} Q_{P_i}^{\mu}\right|; U^k\right\} < \frac{1}{k}, \tag{13.3}$$

with

$$(I-\mu)\left(\sum_{0\leq i\leq \nu_n} C_{k,i} Q_{P_i}^{\mu}\right) = 0. \tag{13.4}$$

Now, there are a subsequence $C_{k_j,i}$ and constants C_i such that

$$C_{k_j,i} \to C_i \qquad \text{as } j \to \infty \tag{13.5}$$

for all $0 \leq i \leq \nu_n$.

To see this, let us observe that if this were not the case, then we would have that

$$M_k = \max\{|C_{k,i}|,\ 0 \leq i \leq \nu_n\} \to \infty \qquad \text{as } k \to \infty.$$

Since $|C^{k,i}|/M_k \leq 1$, there are a subsequence $C^{k_\ell,i}$ and constants B_i such that

$$\frac{C^{k_\ell,i}}{M_{k_\ell}} \to B_i \qquad \text{as } k \to \infty$$

for all $0 \leq i \leq \nu_n$. Note that the subsequence $C^{k_\ell,i}$ can be chosen is such a way that some of the constants B_i are equal to 1 (and hence not all of them vanish). Let

$$R(x) = \sum_{0\leq i\leq \nu_n} B_i Q_{P_i}^{\varphi}(x).$$

By (13.3)

$$\frac{1}{M_{k_\ell}}\left(u(x) - \sum_{0\le i\le \nu_n} C_{k_\ell,i}\, Q^{\mu}_{P_i}(x)\right) \to 0 \qquad \text{as } k \to \infty$$

for all $x \in \Gamma_N$. So, $R(x) = 0$ for all $x \in \Gamma_N$.

Since

$$\sup\{|R(x)|;\, U^r\} \sim \sup\left\{\left|\sum_{\nu_{n-1}<i\le\nu_n} B_i P_i(x)\right|;\, U^r\right\} \qquad \text{as } r \to \infty,$$

we have that

$$\sum_{\nu_{n-1}<i\le\nu_n} B_i P_i(x) = 0, \qquad x \in \Gamma_N,$$

and hence $B_i = 0$, $\nu_{n-1} < i \le \nu_n$.

Arguing in the same way, we can prove successively that, for all $k = n - 1, \dots, 1$, $B_i = 0, \nu_{k-1} < i \le \nu_k$, and that $B_0 = 0$. This is absurd because, by construction, not all of the coefficients B_i vanish. We conclude therefore that (13.5) holds.

By letting $j \to \infty$, it follows from (13.3) and (13.4) that

$$u(x) = \sum_{0\le i\le\nu_n} C_i\, Q^{\mu}_{P_i}(x)$$

for all $x \in \Gamma_N$ and the theorem follows. □

14. The homogenized sub-Laplacian $L_{H\mu}$. The goal of this section is to define the homogenized operator $L_{H\mu}$ associated with the centered probability measure μ on Γ, when $\Gamma \ne \Gamma_N$. $L_{H\mu}$ is a centered left invariant sub-Laplacian on N which, with the notation of Section 6, we write as

$$L_{H\mu} = -\sum_{1\le i,j\le n_1} q_{ij} X_i X_j - \sum_{n_1<i\le n_2} q_i X_i.$$

The coefficients q_{ij} and q_i is are constant [and the $n_1 \times n_1$ matrix (q_{ij}) is symmetric and positive definite].

The way $L_{H\mu}$ and μ are related is illustrated by (14.9) below.

14.1. *The coefficients a_{ij} and a_i.* Let $f \in C^\infty(N)$ and let us extend f to Γ by setting $f(xg_i) = f(x), x \in \Gamma_N, 0 \le i \le k$.

Let the monomials P_i be as in Section 6.1. We extend these monomials to Γ by setting $P_i(t, xg_i) = P_i(t, x),\ x \in N, 0 \le i \le k$. We set $P_i(x) = P_i(0, x)$.

Let

$$
\begin{aligned}
b_i(g_\ell) &= \sum_{h\in\Gamma} P_i(g_\ell h)\mu(h), \qquad 1 \le i \le n_2,\\
a_{ij}(g_\ell) &= \sum_{h\in\Gamma} P_i(g_\ell h)P_j(g_\ell h)\mu(h), \qquad 1 \le i,\, j \le n_1,\\
a_i(g_\ell) &= b_i(g_\ell), \qquad 1 \le i \le n_1,\\
a_i(g_\ell) &= b_i(g_\ell) - \tfrac{1}{2} \sum_{1\le\lambda<\mu\le n_1} a_{\lambda\mu}(g_\ell)\,\mathrm{pr}_i[X_\lambda, X_\mu]_N, \qquad n_1 < i \le n_2.
\end{aligned}
\tag{14.1}
$$

Note that, by setting $a_i(xg_\ell) = a_i(g_\ell)$ and $a_{ij}(xg_\ell) = a_{ij}(g_\ell)$, $x \in \Gamma_N$, these coefficients become functions of type P (cf. Section 1.4).

LEMMA 14.1. *We have*

$$
\sum_{0\le\ell\le k} a_i(g_\ell) = 0, \qquad 1 \le i \le n_1. \tag{14.2}
$$

PROOF. If $g = xg_\ell$ with $x \in \Gamma_N$ and $0 \le \ell \le k$, then, using the notation of Section 1.3, we set $g_N = x$ and $\overline{g} = g_\ell$. Let

$$
\zeta_i(g) = \sum_{\overline{w}} P_i(\overline{w}g), \qquad 1 \le i \le n_1.
$$

Clearly, to prove (14.2) it is enough to prove that

$$
\sum_{g\in\Gamma} \zeta_i(g)\mu(g) = 0, \qquad 1 \le i \le n_1. \tag{14.3}
$$

We have

$$
\begin{aligned}
\zeta_i(gh) &= \sum_{\overline{w}} P_i(\overline{w}gh) = \sum_{\overline{w}} P_i\big((\overline{w}g)_N\overline{\overline{w}g}h\big)\\
&= \sum_{\overline{w}} P_i\big((\overline{w}g)_N\big) + \sum_{\overline{w}} P_i(\overline{\overline{w}g}h) = \sum_{\overline{w}} P_i(\overline{w}g) + \sum_{\overline{w}} P_i(\overline{w}h)\\
&= \zeta_i(g) + \zeta_i(h).
\end{aligned}
$$

This shows that the functions ζ_i, $1 \le i \le n_1$, are additive, and hence (14.3) follows from the definition of a centered probability measure. $\square$

Let V be as in Section 6.5 and let us also assume that $(g_\ell h)_N \in V$ for all $h \in \operatorname{supp}\mu$ and $0 \le i \le k$.

The following lemma is a consequence of Lemma 5.3 and (14.1) and (14.2) above.

LEMMA 14.2. *Let f be as above. Also, let ψ be a function of type P. Then there is a $c > 0$ independent of f such that, for all $x \in \Gamma_N$ and $0 \le \ell \le k$, the following hold*:

(i)

$$(I-\mu)f(xg_\ell) = -\sum_{1\le i\le n_1} a_i(g_\ell)X_i f(x) - \frac{1}{2}\sum_{1\le i,j\le n_1} a_{ij}(g_\ell)X_iX_j f(x) - \sum_{n_1<i\le n_2} a_i(g_\ell)X_i f(x) + F(xg_\ell) \tag{14.4}$$

with

$$|F(xg_\ell)| \le c\|\nabla_X^3 f\|_{L^\infty(xV)};$$

(ii) *for all* $1\le \nu \le n_1$,

$$(I-\mu)(\varphi X_\nu f)(xg_\ell) = \big((I-\mu)\psi\big)(g_\ell)X_\nu f(x) - \sum_{1\le i\le n_1} \mu(\psi P_i)(g_\ell)X_iX_n f(x) + F(xg_\ell) \tag{14.5}$$

with

$$|F(xg_\ell)| \le c\|\nabla_X^3 f\|_{L^\infty(xV)}.$$

14.2. *The correctors and the homogenized operator $L_{H\varphi}$.* The definition of the correctors ψ^j, $j=1,\dots,n_1$, is motivated by (14.4).

DEFINITION. *We define the* (*first order*) *correctors* ψ^j, $1\le j\le n_1$ (cf. [13, 26]), *as functions of type P satisfying*

$$(I-\mu)\psi^j = a_j \quad \textit{and} \quad \langle \psi^j\rangle = 0. \tag{14.6}$$

Note that the correctors ψ^j are well defined and they are given by $\psi^j = \sum_{n\ge 0}\mu^n a_j$, $1\le j\le n_1$ (cf. Section 1.4).

Let

$$b_{ij}(g_\ell) = \sum_{h\in\Gamma}\psi^j(g_\ell h)P_i(g_\ell h)\mu(h), \qquad 1\le i,j\le n_1.$$

If f is as in Lemma 14.2, then combining (14.4), (14.5) and (14.6) we have that, for all $x\in\Gamma_N$ and $0\le \ell\le k$,

$$(I-\mu)\Big(f+\sum_{1\le j\le n_1}\psi^j X_j f\Big)(xg_\ell) = -\sum_{1\le i,j\le n_1}\big(\tfrac{1}{2}a_{ij}(g_\ell)+b_{ij}(g_\ell)\big)X_iX_j f(x) - \sum_{n_1<i\le n_2} a_i(g_\ell)X_i f(x) + F(xg_\ell) \tag{14.7}$$

with

$$|F(xg_\ell)| \le c\|\nabla_X^3 f\|_{L^\infty(xV)}.$$

The following definitions are motivated by the expression (14.7) above.

DEFINITION. *The homogenized sub-Laplacian $L_{H\mu}$ associated with μ is defined with be the operator*

$$L_{H\mu} = -\sum_{1\le i,j\le n_1} q_{ij} X_i X_j - \sum_{n_1 < i \le n_2} q_i X_i$$

with coefficients defined by

$$q_{ij} = \langle \tfrac{1}{2} a_{ij} + b_{ij} \rangle, \qquad 1 \le i, j \le n_1,$$
$$q_i = \langle b_i \rangle, \qquad n_1 < i \le n_2.$$

DEFINITION. *We define the* (*second order*) *correctors ψ^{ij}*, $1 \le i, j \le n_1$ (cf. [13, 26]), *as functions of type P satisfying*

$$(I-\mu)\psi^{ij} = \tfrac{1}{2} a_{ij} + b_{ij} - q_{ij}, \qquad \langle \psi^{ij} \rangle = 0.$$

We also define the (*second order*) *correctors ψ^j, $n_1 < j \le n_2$, as continuous functions on M satisfying*

$$(I-\mu)\psi^j = a_j - q_j, \qquad \langle \psi^j \rangle = 0.$$

The following lemma is a direct consequence of (14.7) and the above definitions.

LEMMA 14.3. *There is a $c > 0$ such that, for all functions f as in Lemma* 14.2 *and all $x \in \Gamma_N$ and $0 \le \ell \le n_1$,*

$$\left|(I-\mu)\left(f(xg_\ell) + \sum_{1\le j\le n_2} \psi^j(g_\ell) X_j f(x)\right)\right| \le c\|\nabla_X^2 f\|_{L^\infty(xV)}, \tag{14.8}$$

$$(I-\mu)\left(f + \sum_{1\le j\le n_2} \psi^j X_j f + \sum_{1\le i,j\le n_1} \psi^{ij} X_j f\right) = L_{H\varphi} f + F \tag{14.9}$$

with the function F satisfying

$$|F(xg_\ell)| \le c\|\nabla_X^3 f(x)\|_{L^\infty(xV)}.$$

COROLLARY 14.4. *Let $u \in C^\infty(\mathbb{R} \times N)$ and let us extend u to $\mathbb{R} \times \Gamma$ by setting $u(t, xg_\ell) = u(t, x)$, $x \in \Gamma_N$, $0 \le \ell \le n_1$. Also, let*

$$U(t, xg_\ell) = u(t, xg_\ell) + \sum_{1 \le j \le n_2} \psi^j(g_\ell) X_j u(t, xg_\ell) + \sum_{1 \le i, j \le n_1} \psi^{ij}(g_\ell) X_i X_j u(t, xg_\ell).$$

Then there is a constant $c > 0$ independent of u such that, for all $t \in \mathbb{R}$, $x \in N$ and $0 \le \ell \le k$,

$$U(t+1, xg_\ell) - \mu U(t, xg_\ell) = \left(\frac{\partial}{\partial t} + L_{H\mu}\right) U(t, x) + V(t, xg_\ell) \tag{14.10}$$

with

$$|V(t, xg_\ell)| \le c_1 \left\| \left| \frac{\partial^2}{\partial s^2} u(s, x) \right| + \left| \nabla_X \frac{\partial}{\partial s} u(s, x) \right| + |\nabla_X^3 u(t, x)| \right\|_{L^\infty([t, t+1] \times xV)}.$$

14.3. *$L_{H\mu}$ is a centered sub-Laplacian on N.* The following lemma asserts that $L_{H\mu}$ is indeed a sub-Laplacian on N.

LEMMA 14.5. *For all $\xi = (\xi_1, \dots, \xi_{n_1}) \in \mathbb{R}^{n_1}$, $\xi \ne 0$,*

$$\sum_{1 \le i, j \le n_1} q_{ij} \xi_i \xi_j > 0. \tag{14.11}$$

PROOF. Let us fix $\xi = (\xi_1, \dots, \xi_{n_1}) \ne 0$ and consider the function

$$u = \sum_{1 \le i \le n_1} \xi_i (P_i + \psi^i).$$

Since by construction $(I - \mu)(P_i + \psi^i) = 0$, $1 \le i \le n_1$, we have that $(I - \mu) \times u = 0$; that is, u is μ-harmonic. Since the function $f(t) = t^2$ is convex, we have that $(I - \mu) u^2 \le 0$.

We have

$$\begin{aligned} u^2 &= \sum_{1 \le i, j \le n_1} \xi_i \xi_j (P_i + \psi^i)(P_j + \psi^j) \\ &= \sum_{1 \le i, j \le n_1} \xi_i \xi_j (P_i P_j + P_i \psi^j + P_j \psi^i + \psi^i \psi^j). \end{aligned}$$

By Lemma 14.2, for $1 \le i, j \le n_1$,

$$\begin{aligned} (I - \mu)(P_i P_j) &= -P_i a_j - P_j a_i - a_{ij}, \\ (I - \mu)(P_i \psi^j) &= P_i a_j - b_{ij}, \\ (I - \mu)(P_j \psi^i) &= P_j a_i - b_{ji}. \end{aligned}$$

Hence

$$(I-\mu)[(P_i+\psi^i)(P_j+\psi^j)] = -a_{ij} - b_{ij} - b_{ji} + [(I-\mu)(\psi^i\psi^j)].$$

It follows that $(I-\mu)u^2$ is a function of type P. Since $\langle (I-\mu)(\psi^i\psi^j)\rangle = 0$, we have

$$\big\langle (I-\mu)\big((P_i+\psi^i)(P_j+\psi^j)\big)\big\rangle = -2q_{ij}.$$

Hence

$$\langle (I-\mu)u^2\rangle = -2\sum_{1\le i,j\le n_1} q_{ij}\xi^i\xi^j.$$

Now, if we had $\sum_{1\le i,j\le n_1} q_{ij}\xi_i\xi_j = 0$, then we would have $(I-\mu)u^2 = 0$. Since the function $f(t) = t^2$ is strictly convex, this would imply that u =const, which is absurd. $\square$

15. Proof of Propositions 1.3 and 1.4 in the general case. The proof of Propositions 1.3 and 1.4 in the general case is similar to the proof in the case $\Gamma = \Gamma_N$ (see Section 9). The only difference is that instead of Lemma 9.1 we must use the following generalization:

LEMMA 15.1. *There is a constant $c > 0$ such that, for all $n \in \mathbb{N}$, $T \ge 1$,*

$$\| p^{H\mu}_{n+T} - \mu^n p^{H\varphi}_T \|_\infty \le c\, T^{-(D+1)/2}. \tag{15.1}$$

PROOF. Let

$$U_t = p^{H\mu}_t - \sum_{1\le j\le n_2} \psi^j X_j p^{H\mu}_t - \sum_{1\le i,j\le n_1} \psi^{ij} X_i X_j p^{H\mu}_t.$$

By (6.10), there is a $c > 0$ such that, for all $T \ge 1$ and $t \ge 0$,

$$\| X_j p^{H\mu}_{t+T}\|_\infty \le c\,(t+T)^{-(D+1)/2}, \qquad \| X_i X_j p^{H\mu}_{t+T}\|_\infty \le c\,(t+T)^{-(D+2)/2}.$$

So to prove (15.1) it is enough to prove that there is a $c > 0$ such that, for all $T \ge 1$ and $t \ge 0$,

$$\| U_{n+T} - \mu^n U_T\|_\infty \le c\, T^{-(D+1)/2}. \tag{15.2}$$

By (14.10) and (6.10) there is a $c > 0$ such that, for all $t \ge 1$,

$$\| U_{t+1} - \mu U_t\|_\infty \le c t^{-(D+3)/2}. \tag{15.3}$$

We have

$$\begin{aligned} U_{n+T} - \mu^n U_T &= U_{n+T} - \mu^{n-1}U_{1+T} + \mu^{n-1}U_{1+T} - \mu^n U_T \\ &= \sum_{0\le i\le n-1} \mu^i\big(U_{n-i+T} - \mu U_{n-i-1+T}\big). \end{aligned}$$

So, by (15.3),

$$\begin{aligned}\|U_{n+T}-\mu^n U_T\|_\infty &\le \sum_{0\le i\le n-1}\|\mu^i\|_1\|U_{n-i+T}-\mu U_{n-i-1+T}\|_\infty\\ &\le c\sum_{0\le i\le n-1}(n-i-1+T)^{-(D+3)/2}\\ &\le c\,T^{-(D+1)/2}\end{aligned}$$

and the lemma follows. □

16. Proof of the Gaussian estimate in the general case. The of proof of (1.14) follows the same lines as in the case $\Gamma_N=\Gamma$ (see Section 10). The only change is that we must replace the functions H_k by their modifications U_k, $k\ge 1$.

16.1. *The functions U_k.* Let the family of functions ρ_k be as in Section 10.1, let us fix $A>0$ and $B>0$ and let

$$H_k(t,x)=\exp\biggl(-\frac{(\rho_k(x)+B\sqrt{k})^2}{A(k+t)}\biggr),\qquad t\ge 0,\ x\in N.$$

We extend the functions ρ_k and H_k to Γ by setting $\rho_k(xg_\ell)=\rho_k(x)$ and $H_k(t,xg_\ell)=H_k(t,x)$, $x\in\Gamma_N$, $0\le\ell\le k$.

We do the same for the derivatives $XY\cdots ZH_k$, $X,Y,\dots,Z\in\mathfrak{n}$.

We consider the functions

$$U_k=H_k(t,x)+\sum_{1\le j\le n_1}\psi^j X_j H_k.$$

LEMMA 16.1. *There are constants $A>0$ and $B>0$ such that, for all $k\ge 1$, $t\in[0,k]$ and all $|x|_\Gamma\le ak$,*

$$0<\tfrac12 H_k(t,x)<U_k(t,x)<2H_k(t,x)\tag{16.1}$$

and

$$U_k(t+1,x)>\mu U_k(t,x).\tag{16.2}$$

PROOF. For all $X\in\mathfrak{n}$, we have

$$XH_k(t,x)=-\frac{1}{A}\frac{1}{k+t}2\bigl(\rho_k(x)+B\sqrt{k}\bigr)X\rho_k(x)\ H_k(t,x).\tag{16.3}$$

It follows that there is a $c>0$ such that, for all $|x|_\Gamma\le ak$,

$$|XH_k(t,x)|\le c\frac{1}{A}\biggl(a+\frac{B}{\sqrt{k}}\biggr)H_k(t,x),$$

which proves (16.1).

The proof of (16.2) is similar to the proof of (10.6) in the case $\Gamma = \Gamma_N$. We observe again that

$$U_k(t+1,x) - \mu U_k(t,x) = U_k(t+1,x) - U_k(t,x) + (I-\mu)U_k(t,x).$$

Since the correctors ψ^j are functions defined on Γ/Γ_N, which is finite, there is a $c > 0$ such that

$$\begin{aligned} U_k(t+1,x) - U_k(t,x) \geq \inf&\left\{\frac{\partial}{\partial s}H_k(s,x); s \in [t,t+1]\right\} \\ &- c\sup\left\{\left|\nabla_X \frac{\partial}{\partial s}H_k(s,y)\right|; s \in [t,t+1], y \in xV\right\}. \end{aligned}$$

Also, by (14.8),

$$|(I-\mu)U_k(t,x)| \leq c\sup\{\nabla_X^2 H_k(t,y); y \in xV\}.$$

It follows that

$$\begin{aligned} &U_k(t+1,x) - \mu U_k(t,x) \\ &\quad \geq \inf\left\{\frac{\partial}{\partial s}H_k(s,x); s \in [t,t+1]\right\} \\ &\quad\quad - c\sup\left\{\left|\nabla_X \frac{\partial}{\partial s}H_k(s,y)\right|; s \in [t,t+1], y \in xV\right\} \\ &\quad\quad - c\sup\{\nabla_X^2 H_k(t,y); y \in xV\}. \end{aligned} \tag{16.4}$$

As in the proof of (10.6) we observe that

$$\frac{\partial}{\partial t}H_k(t,x) = \frac{1}{A}\frac{1}{k+t}\frac{(\rho_k(x) + B\sqrt{k})^2}{k+t}H_k(t,x) \tag{16.5}$$

and that, for all $X, Y \in \mathfrak{n}$,

$$\begin{aligned} \frac{\partial}{\partial t}XH_k(t,x) &= \frac{1}{A}\frac{1}{k+t}\Big[2\frac{\rho_k(x) + B\sqrt{k}}{k+t}X\rho_k(x) \\ &\qquad - \frac{(\rho_k(x) + B\sqrt{k})^2}{k+t}\frac{1}{A}\frac{1}{k+t}2\big(\rho_k(x) \\ &\qquad + B\sqrt{k}\big)X\rho_k(x)\Big]H_k(t,x) \\ &= \frac{1}{A}\frac{1}{k+t}\Big[2\frac{\rho_k(x) + B\sqrt{k}}{k+t}X\rho_k(x) \\ &\qquad - 2\frac{1}{A}\frac{1}{(k+t)^2}\big(\rho_k(x) + B\sqrt{k}\big)^3 X\rho_k(x)\Big]H_k(t,x) \end{aligned} \tag{16.6}$$

and

$$
\begin{aligned}
&XYH_k(t,x)\\
&\quad= -\frac{1}{A}\frac{1}{k+t}\Big[2X\rho_k(x)Y\rho_k(x) - 2\big(\rho_k(x)+B\sqrt{k}\big)XY\rho_k(x)\\
&\qquad\qquad + \frac{1}{A}\frac{1}{(k+t)}4\big(\rho_k(x)+B\sqrt{k}\big)^2 X\rho_k(x)Y\rho_k(x)\Big]H_k(t,x).
\end{aligned}
\tag{16.7}
$$

Case I ($|x|_\Gamma \le \sqrt{k}-1$ and $t \in [0,k]$). By construction, for all $|x|_\Gamma \le \sqrt{k}$,

$$
U_k(t,x) = H_k(t,x) = \exp\Big(-\frac{B^2k}{A(k+t)}\Big).
$$

Hence, for all $|x|_\Gamma \le \sqrt{k}-1$,

$$
U_k(t+1,x) > U_k(t,x) = \big(\mu U_{k+1}(t,\cdot)\big)(x).
$$

Case II ($\sqrt{k}-1 \le |x|_\Gamma \le \zeta\sqrt{k}+1$ and $0 \le t \le k$). Then by (16.5)–(16.7) there is a $c > 0$ such that

$$
\begin{aligned}
&U_k(t+1,x) - \mu U_k(t,x)\\
&\quad\ge \frac{1}{A}\frac{1}{k+t+1}\Big[\frac{B^2k}{k+k+1} - c\frac{c\sqrt{k}+B\sqrt{k}}{k}\\
&\qquad\qquad - c\frac{1}{A}\frac{1}{k^2}(c\sqrt{k}+B\sqrt{k})^3 - c - c(c\sqrt{k}+B\sqrt{k})\frac{1}{\sqrt{k}}\\
&\qquad\qquad\quad - c\frac{1}{A}\frac{1}{k}(c\sqrt{k}+B\sqrt{k})^2 - c(c\sqrt{k}+B\sqrt{k})\frac{1}{\sqrt{k}}\Big]H_k(t,x)\\
&\quad\ge \frac{1}{A}\frac{1}{k+t+1}\Big[\frac{B^2}{3} - c\frac{c+B}{\sqrt{k}} - c\,\frac{1}{A}\frac{1}{\sqrt{k}}(c+B)^3\\
&\qquad\qquad - c - c(c+B) - c\,\frac{1}{A}(c+B)^2 - c(c+B)\Big]H_k(t,x).
\end{aligned}
$$

So, by choosing B large enough that

$$
\frac{B^2}{6} > +c(c+B) + c + c(c+B) + c(c+B)
$$

and then A large enough that

$$
\frac{B^2}{6} > c\frac{1}{A}(c+B)^3 + c\frac{1}{A}(c+B)^2,
$$

we have

$$
U_k(t+1,x) > \mu U_k(t,x).
$$

Case III ($\zeta\sqrt{k}+1 < |x|_\Gamma < ak$ and $0 \le t \le k$). Then by (16.5)–(16.7) there is a $c > 0$ such that

$$
\begin{aligned}
&U_k(t+1,x) - \mu U_k(t,x)\\
&\quad\ge \frac{1}{A}\frac{1}{k+t+1}\biggl[\frac{(|x|_\Gamma + B\sqrt{k})^2}{k+t+1} - c\frac{c|x|_\Gamma + B\sqrt{k}}{k+t}\\
&\qquad\qquad - c\frac{1}{A}\frac{1}{k^2}(|x|_\Gamma + B\sqrt{k})^3 - c - c(|x|_\Gamma + B\sqrt{k})\frac{1}{|x|_\Gamma}\\
&\qquad\qquad - c\frac{1}{A}\frac{1}{k}(|x|_\Gamma + B\sqrt{k})^2 - c(|x|_\Gamma + B\sqrt{k})\frac{1}{|x|_\Gamma}\biggr]H_k(t,x)\\
&\quad\ge \frac{1}{A}\frac{1}{k+t+1}\biggl[(|x|_\Gamma + B\sqrt{k})\frac{|x|_\Gamma + B\sqrt{k}}{3k} - c\frac{|x|_\Gamma + B\sqrt{k}}{k}\\
&\qquad\qquad - c\frac{1}{A}\biggl(\frac{|x|_\Gamma + B\sqrt{k}}{k}\biggr)^2(|x|_\Gamma + B\sqrt{k}) - c - c\biggl(1 + \frac{B\sqrt{k}}{|x|_\Gamma}\biggr)\\
&\qquad\qquad - c\frac{1}{A}\frac{|x|_\Gamma + B\sqrt{k}}{k}(|x|_\Gamma + B\sqrt{k}) - c\biggl(1 + \frac{B\sqrt{k}}{|x|_\Gamma}\biggr)\biggr]H_k(t,x).
\end{aligned}
$$

So, by choosing B large enough that

$$
(|x|_\Gamma + B\sqrt{k})\frac{|x|_\Gamma + B\sqrt{k}}{6k} > c\frac{|x|_\Gamma + B\sqrt{k}}{k} + c + c\biggl(1 + \frac{B\sqrt{k}}{|x|_\Gamma}\biggr) + c\biggl(1 + \frac{B\sqrt{k}}{|x|_\Gamma}\biggr)
$$

and A large enough that

$$
\begin{aligned}
(|x|_\Gamma + B\sqrt{k})\frac{|x|_\Gamma + B\sqrt{k}}{6k} &> \frac{1}{A}\biggl(\frac{|x|_\Gamma + B\sqrt{k}}{k}\biggr)^2(|x|_\Gamma + B\sqrt{k})\\
&\quad + c\frac{1}{A}\frac{|x|_\Gamma + B\sqrt{k}}{k}(|x|_\Gamma + B\sqrt{k}),
\end{aligned}
$$

we have

$$
U_k(t+1,x) > \mu U_k(t,x),
$$

which ends the proof of the lemma. □

17. The proof of the Berry–Esseen estimate in the general case. In this section, we give the proof of Theorem 11.1 in the case when $\Gamma \ne \Gamma_N$. The general strategy is the same.

Let $L_{H\mu}$ be the homogenized sub-Laplacian associated with μ and let $p_t^{H\mu}(x,y)$ be its heat kernel. We extend $p_t^{H\mu}(x,y)$ to Γ by setting

$$
p_t^{H\varphi}(xg_i, yg_j) = \frac{1}{k+1}p_t^{H\varphi}(x,y), \qquad x, y \in N,\ 0 \le i, j \le k.
$$

Let us also recall that by (4.2) there are $\gamma \in (0,1]$ and $c > 0$ such that

$$\|\nabla_U \mu^n\|_\infty \le cn^{-(D+\gamma)/2}, \qquad n \in \mathbb{N}. \tag{17.1}$$

The following result is a generalization of Theorem 11.1.

THEOREM 17.1. *There is a $c > 1$ such that, for all $x, y \in \Gamma$ and $n \in \mathbb{N}$,*

$$|\mu^n(x,y) - p_n^{H\mu}(x,y)| \le cn^{-(D+\gamma)/2}. \tag{17.2}$$

We point out again that, once we have proved Theorem 1.14, the above inequality will hold with $\gamma = 1$.

Let the kernel $S_t(x,y)$ be as in Section 11 and let us extend it to Γ by setting

$$S_t(xg_i, yg_j) = \frac{1}{k+1} S_t(x,y), \qquad x, y \in N,\ 0 \le i, j \le k.$$

Theorem 17.1 is proved in exactly the same way as Theorem (11.1), once we have the following analogues of Lemmas 11.2 and 11.3.

LEMMA 17.2. *There are constants $a, b \ge 1$ such that, for all $T \ge 1$ and $n \in \mathbb{N}$,*

$$\|\mu^n - p_n^{H\mu}\|_\infty \le a\|(\mu^n - p_n^{H\mu})S_T\|_\infty + b\sqrt{T} n^{-(D+\gamma)/2}. \tag{17.3}$$

The proof of the above lemma is exactly the same as the proof of Lemma 11.2.

LEMMA 17.3. *There is a constant $c \ge 1$ such that if, for some $n \in \mathbb{N}$,*

$$\|\mu^k - p_k^{H\mu}\|_\infty \le Ak^{-(D+\gamma)/2}, \qquad 1 \le k \le n-1, \tag{17.4}$$

then

$$\|(\mu^n - p_n^{H\mu})S_T\|_\infty \le c\left(1 + \frac{A}{\sqrt{T}}\right) n^{-(D+\gamma)/2}. \tag{17.5}$$

The proof of the above lemma, although similar in spirit to the proof of Lemma 11.3, is technically more complicated. For the case when μ is symmetric, a proof of the above lemma is given in [2]. We give below an adaptation of that proof.

17.1. *Proof of Lemma* 17.3. Let $\mu^{\vee n}(x,y) = \mu^n(y,x)$, $x, y \in \Gamma$, and let $L_{H\mu^\vee}$ be the homogenized sub-Laplacian associated with $\mu^\vee$. Note that $L_{H\mu^\vee}$ is just the formal adjoint of $L_{H\mu}$ and that its heat kernel $p_t^{\vee H\mu}(x,y)$ satisfies $p_t^{\vee H\mu}(x,y) = p_t^{H\mu}(y,x)$, $x, y \in N$.

Finally, let $\psi^{\vee j}$, $1 \le j \le n_2$, and $\psi^{\vee ij}$, $1 \le i, j \le n_1$, be respectively the first and second order correctors associated with $\mu^\vee$.

We set

$$W_t(x,y) = \sum_{1\le j\le n_2} \psi^j(x) X_j^x p_t^{H\mu}(x,y) + \sum_{1\le i,j\le n_1} \psi^{ij}(x) X_i^x X_j^x p_t^{H\mu}(x,y)$$

and

$$W_t^\vee(x,y) = \sum_{1\le j\le n_2} \psi^{\vee j}(y) X_j^y p_t^{H\mu}(x,y) + \sum_{1\le j\le n_1} \psi^{\vee ij}(y) X_i^y X_j^y p_t^{H\mu}(x,y),$$

where the superscripts x and y denote differentiation with respect to the x and y variables respectively.

Let

$$U_t = p_t^{H\mu} + W_t,$$
$$U_t^\vee = p_t^{H\mu} + W_t^\vee.$$

Let us also fix a $T \ge 1$. Then, to prove (17.5), it is enough to prove that

$$\|\mu^n S_T - U_{n+T}\|_\infty \le c\left(1 + \frac{A}{\sqrt{T}}\right) n^{-(D+\gamma)/2}. \tag{17.6}$$

We have

$$\begin{aligned}
U_{n+T} - \mu^n S_T &= U_{n+T} - \mu^{n-1}U_{1+T} + \mu^{n-1}U_{1+T} - \mu^n S_T \\
&= \sum_{0\le i\le n-2} (\mu^i U_{n-i+T} - \mu^{i+1} U_{n-i-1+T}) + \mu^{n-1}U_{1+T} - \mu^n S_T \\
&= \sum_{0\le i\le n-2} \mu^i (U_{n-i+T} - \mu U_{n-i-1+T}) + \mu^{n-1}U_{1+T} - \mu^n S_T \\
&= \sum_{0\le i\le n/2} \mu^i (U_{n-i+T} - \mu U_{n-i-1+T}) \\
&\quad + \sum_{n/2<i\le n-2} \mu^i (U_{n-i+T} - \mu U_{n-i-1+T}) \\
&\quad + \mu^{n-1}(U_{1+T} - \mu S_T) \qquad (17.7)\\
&= \sum_{0\le i\le n/2} \mu^i (U_{n-i+T} - \mu U_{n-i-1+T}) \\
&\quad + \sum_{n/2<i\le n-2} (\mu^i - U_i^\vee)(U_{n-i+T} - \mu U_{n-i-1+T}) \\
&\quad + \sum_{n/2<i\le n-2} U_i^\vee (U_{n-i} - \mu U_{n-i-1+T}) \\
&\quad + (\mu^{n-1} - U_{n-1}^\vee)(U_{1+T} - \mu S_T) \\
&\quad + U_{n-1}^\vee (U_{1+T} - \mu S_T).
\end{aligned}$$

In the rest of the proof, we make repeated use of (14.10) and (6.10). We have

$$
\begin{aligned}
&\sum_{0\le i\le n/2} \|\mu^i(U_{n-i+T}-\mu U_{n-i-1+T})\|_\infty \\
&\quad\le \sum_{0\le i\le n/2} \|\mu^i\|_1\|(U_{n-i}-\mu U_{n-i-1})\|_\infty \\
&\quad\le \sum_{0\le i\le n/2} c(n-i-1+T)^{-(D+3)/2} \\
&\quad\le cn^{-(D+1)/2}.
\end{aligned}
\tag{17.8}
$$

By the inductive hypothesis (17.4),

$$
\begin{aligned}
&\sum_{n/2<i\le n-2} \|(\mu^i-U_i^\vee)(U_{n-i+T}-\mu U_{n-i-1+T})\|_\infty \\
&\quad\le \sum_{n/2<i\le n-2} \|(\mu^i-U_i^\vee)\|_\infty\|(U_{n-i+T}-\mu U_{n-i-1+T})\|_1 \\
&\quad\le \sum_{n/2<i\le n-2} Ai^{-(D+\gamma)/2}c(n-i-1+T)^{-3/2} \\
&\quad\le cA\frac{1}{\sqrt{T}}n^{-(D+\gamma)/2}.
\end{aligned}
\tag{17.9}
$$

We have (arguing as in the proof of Lemma 11.3) that

$$
\|\mu S_T - p_T^{H\mu}\|_1 \le \frac{c}{\sqrt{T}}, \qquad T\ge 1.
$$

Also, by (6.10),

$$
\|p_{T+1}^{H\mu} - p_T^{H\mu}\|_1 \le \frac{c}{T}, \qquad T\ge 1.
$$

Hence

$$
\begin{aligned}
&\|(\mu^{n-1}-U_{n-1}^\vee)(U_{1+T}-\mu S_T)\|_\infty \\
&\quad\le \|\mu^{n-1}-U_{n-1}^\vee\|_\infty\|U_{1+T}-\mu S_T\|_1 \le cA\frac{1}{\sqrt{T}}n^{-(D+\gamma)/2}.
\end{aligned}
\tag{17.10}
$$

To estimate the remaining term in (17.7), we observe that

$$
\begin{aligned}
&\sum_{n/2<i\le n-2} U_i^\vee(U_{n-i+T}-\mu U_{n-i-1+T}) + U_{n-1}^\vee(U_{1+T}-\mu S_T) \\
&\quad= U_{[n/2]+1}^\vee U_{n-[n/2]-1+T} - U_{n-1}^\vee\mu S_T \\
&\qquad+\sum_{n/2<i\le n-2}(U_{i+1}^\vee - U_i^\vee\mu)U_{n-i-1+T}.
\end{aligned}
$$

Now

$$\begin{aligned} &\sum_{n/2<i\le n-2} \|(U_{i+1}^{\vee} - U_i^{\vee}\mu)U_{n-i-1+T}\|_\infty \\ &\qquad \le \sum_{n/2<i\le n-2} \|U_{i+1}^{\vee} - U_i^{\vee}\mu\|_\infty \|U_{n-i-1+T}\|_1 \\ &\qquad \le \sum_{n/2<i\le n-1} ci^{-(D+3)/2} \\ &\qquad \le cn^{-(D+1)/2}. \end{aligned} \tag{17.11}$$

Also

$$\begin{aligned} &U_{[n/2]+1}^{\vee} U_{n-[n/2]-1+T} - U_{n-1}^{\vee}\mu S_T \\ &\quad = \big(p_{[n/2]+1}^{H\mu} + W_{[n/2]+1}^{\vee}\big)\big(p_{n-[n/2]-1+T}^{H\mu} + W_{n-[n/2]-1+T}\big) \\ &\qquad - \big(p_{n-1}^{H\mu} + W_{n-1}^{\vee}\big)\mu S_T \\ &\quad = U_{[n/2]+1}^{\vee} W_{n-[n/2]-1+T} + W_{[n/2]+1}^{\vee} U_{n-[n/2]-1+T} \\ &\qquad + W_{n-1}^{\vee}\mu S_T + p_{[n/2]+1}^{H\mu} p_{n-[n/2]-1+T}^{H\mu} - p_{n-1}^{H\mu}\mu S_T. \end{aligned}$$

By (6.10), for all $x, y \in \Gamma$,

$$\left| p_{n+T}^{H\mu}(x,y) - \sum_{z\in\Gamma} p_{[n/2]+1}^{H\mu}(x,z) p_{n-[n/2]-1+T}^{H\mu}(z,y) \right| \le cn^{-(D+1)/2}$$

and

$$\left| p_{n+T}^{H\mu}(x,y) - \sum_{z\in\Gamma} p_{n-1}^{H\mu}(x,z)(mS_T)(z,y) \right| \le cn^{-(D+1)/2}.$$

Hence

$$\| p_{[n/2]+1}^{H\mu} p_{n-[n/2]-1+T}^{H\mu} - p_{n-1}^{H\mu}\mu S_T \|_\infty \le cn^{-(D+1)/2}.$$

It follows that

$$\begin{aligned} &\|U_{[n/2]+1}^{\vee} U_{n-[n/2]-1+T} - U_{n-1}^{\vee}\mu S_T\|_\infty \\ &\quad \le \|U_{[n/2]+1}^{\vee}\|_1 \|W_{n-[n/2]-1+T}\|_\infty \\ &\qquad + \|W_{[n/2]+1}^{\vee}\|_\infty \|U_{n-[n/2]-1+T}\|_1 + \|W_{n-1}^{\vee}\|_\infty \|\mu S_T\|_1 \\ &\qquad + \|p_{[n/2]+1}^{H\mu} p_{n-[n/2]-1+T}^{H\mu} - p_{n-1}^{H\mu}\mu S_T\|_\infty \\ &\quad \le c(n-[n/2]-1+T)^{-(D+1)/2} + c([n/2]+1)^{-(D+1)/2} \\ &\qquad + c(n-1)^{-(D+1)/2} + cn^{-(D+1)/2} \\ &\quad \le cn^{-(D+1)/2}. \end{aligned} \tag{17.12}$$

Summing (17.8)–(17.12), we obtain (17.6) and the lemma follows.

18. The corrected monomials $Q_{P_i}^{\psi}$. The goal of this section is to construct the corrected monomials $Q_{P_i}^{\psi}$ appearing in Theorem 1.11.

Let the monomials P_i, $i = 0, 1, 2, \ldots$, be as in Section 6.1 and let us associate with these monomials and the sub-Laplacian $L_{H\mu}$ polynomials Q_{P_i} satisfying (6.6).

Note that, by (6.5) and (6.6), we can associate with every polynomial $P(t, x)$ on $\mathbb{R} \times N$ another polynomial $Q(t, x)$ satisfying

$$\left(\frac{\partial}{\partial t} + L_{H\mu}\right) Q(t, x) = P(t, x), \qquad \deg_H Q(t, x) = \deg_H P(t, x) + 2. \tag{18.1}$$

Then the corrected monomials $Q_{P_i}^{\psi}$ will be furnished by the following:

PROPOSITION 18.1. *With every monomial $Q_{P_i}(t, g)$ with $\deg_H Q_{P_i} = d$, as above, we can associate a polynomial*

$$Q_{P_i}^{\psi}(t, x) = P_i(t, x) + \sum_{0 \le j \le \nu_{d-1}} \psi_j^i(x) P_j(t, x) \tag{18.2}$$

satisfying

$$\left(\frac{\partial}{\partial t} + L_{H\mu}\right) Q_{P_i} = \big(\partial_1 + (I - \mu)\big) Q_{P_i}^{\psi} \tag{18.3}$$

and where the functions ψ_j^i are of type P.

Before we continue with the proof of Proposition 18.1, let us observe that by (14.10) we can take (1) for $1 < i \le \nu_1$ (note that $\nu_1 = n_1$),

$$Q_{P_i}^{\psi} = P_i + \psi^i \tag{18.4}$$

and (2) for $\nu_1 < i \le \nu_2$,

$$Q_{P_i}^{\psi} = P_i + \sum_{1 \le j \le n_2} \psi^j X_j P_i + \sum_{1 \le \ell, j \le n_1} \psi^{\ell j}(z) X_\ell X_j P_i. \tag{18.5}$$

PROOF OF PROPOSITION 18.1. By (18.4) and (18.5) we can assume that $k \ge 3$. Then as a first approximation to $Q_{P_i}^{\psi}$ we take

$$Q_{P_i}^{\psi,1} = Q_{P_i} + \sum_{1 \le j \le n_2} \psi^j X_j Q_{P_i} + \sum_{1 \le \ell, j \le n_1} \psi^{\ell j} X_\ell X_j Q_{P_i}.$$

By (14.10),

$$(\partial_1 + (I-\mu))Q_{P_i}^{\psi,1} = \left(\frac{\partial}{\partial t} + L_{H\mu}\right)Q_{P_i} + \sum_{0\le j\le \nu_{d-3}} f_j^{i,1} P_j, \tag{18.6}$$

where the functions $f_j^{i,1}$ are of type P.

Making use of (18.1), for every $\nu_{d-4} < \ell \le \nu_{d-3}$, we consider a polynomial R_ℓ satisfying

$$\begin{gathered}\left(\frac{\partial}{\partial t} + L_{H\mu}\right)R_\ell = -P_\ell,\\ \deg_H Q_\ell = \deg_H P_\ell + 2 = d-1.\end{gathered} \tag{18.7}$$

Arguing in the same way as for the definition of the correctors, we consider functions $\phi_\mu^{i,1}$ which are of type P and which satisfy

$$(I-\mu)\phi_\ell^{i,1} = -f_\ell^{i,1} + \langle f_\ell^{i,1}\rangle \quad \text{and} \quad \langle \phi_\ell^{i,1}\rangle = 0. \tag{18.8}$$

Let

$$R_\ell^f = \langle f_\ell^{i,1}\rangle\left(R_\ell + \sum_{1\le j\le n_2} \psi^j X_j R_\ell + \sum_{1\le \lambda, j\le n_1} \psi^{ij} X_\lambda X_j R_\ell\right) + \phi_\ell^{i,1} P_\ell.$$

As a second approximation to $Q_{P_i}^\psi$ we consider the corrected polynomial

$$Q_{P_i}^{\psi,2} = Q_{P_i}^{\psi,1} + \sum_{\nu_{d-4}<\ell\le\nu_{d-3}} R_\ell^f.$$

This polynomial satisfies

$$(\partial_1 + (I-\mu))Q_{P_i}^{\psi,2} = \left(\frac{\partial}{\partial t} + L_{H\mu}\right)Q_{P_i} + \sum_{0\le j\le n_{d-4}} f_j^{i,2} P_j,$$

where the functions $f_j^{i,2}$ are again of type P.

We repeat the same procedure another $d-2$ times. The polynomial $Q_{P_i}^{\psi,k}$ that we obtain in the end will satisfy (18.2) and (18.3). □

19. Harnack inequalities for higher order spatial differences. If Γ is not nilpotent, then the analogue of (1.21) for higher order spatial differences is in general false. To see this, let us assume for simplicity that there is a finite subgroup $M \le \Gamma$ such that $\Gamma = \Gamma_N M$ and $\Gamma \cap M = \{e\}$ (i.e., Γ is isomorphic to the semidirect product $\Gamma_N \rightthreetimes M$) and let us consider the function

$$u = P_j + \psi^j,$$

where $k_1 < j \le n_1$. This function grows linearly; that is, there is a $c > 0$ such that

$$\sup\{|u|; U^r\} \le cr, \qquad r \ge 1,$$

and it satisfies $(I-\mu)u=0$. Also, $\partial_z\partial_w u=\partial_z\partial_w\psi^j$ for all $z,w\in M$. So, if $\partial_z\partial_w\psi^j\neq 0$, then the inequality

$$\sup\{|\partial_z\partial_w u|; M\}\leq cr^{-2}\sup\{|u|; U^r)\},\qquad r\geq 1,$$

is false.

20. Berry–Esseen estimates for the differences. The goal of this section is to prove the Berry–Esseen estimates (1.28) and (1.30).

We use the notation of Section 17. We set

$$W_t(x,y)=\sum_{1\leq j\leq n_2}\psi^j(x)X_j^x p_t^{H\mu}(x,y)+\sum_{1\leq i,j\leq n_1}\psi^{ij}(x)X_i^x X_j^x p_t^{H\mu}(x,y)$$

and

$$U_t(x,y)=p_t^{H\mu}(x,y)+W_t(x,y).$$

Then, by (1.23),

$$\|\mu^n-U_n\|_\infty\leq cn^{-(D+1)/2}. \tag{20.1}$$

20.1. *Proof of Theorem* 1.21. It is enough to prove that there is a constant $c>0$ such that, for all $z\in U$,

$$\|\partial_z\mu^n-\partial_z U_n\|_\infty\leq cn^{-(D+2)/2}. \tag{20.2}$$

We have

$$\begin{aligned}U_n-\mu^n&=\sum_{0\leq i<[n/2]}\mu^i U_{n-i}-\mu^{i+1}U_{n-i-1}+\mu^{[n/2]}U_{n-[n/2]}-\mu^n\\&=\sum_{0\leq i<[n/2]}\mu^i(U_{n-i}-\mu U_{n-i-1})+\mu^{[n/2]}\big(U_{n-[n/2]}-\mu^{n-[n/2]}\big)\end{aligned}$$

Hence, by (14.10) and (6.10),

$$\begin{aligned}\|\partial_z\mu^n-\partial_z U_n\|_\infty&\leq\sum_{0\leq i<[n/2]}\|\partial_z\mu^i\|_1\|U_{n-i}-\mu U_{n-i-1}\|_\infty\\&\qquad+\|\partial_z\mu^{[n/2]}\|_1\|U_{n-[n/2]}-\mu^{n-[n/2]}\|_\infty\\&\leq\sum_{0<i<[n/2]}ci^{-1/2}(n-i-1)^{-(D+3)/2}+cn^{-1/2}n^{-(D+1)/2}\\&\leq cn^{-(D+2)2},\end{aligned}$$

which proves (20.2) and the theorem follows. $\square$

20.2. *Proof of Theorem* 1.22. It is enough to prove that there is a constant $c > 0$ such that

$$\|\partial_1 \mu^n - \partial_1 U_n\|_\infty \le c n^{-(D+3)/2}, \qquad n \in \mathbb{N}. \tag{20.3}$$

We have

$$\begin{aligned}
\partial_1 U_n - \partial_1 \mu^n &= \sum_{0 \le i < [n/2]} \mu^i \partial_1 U_{n-i} - \mu^{i+1} \partial_1 U_{n-i-1} \\
&\quad + \mu^{[n/2]} \partial_1 U_{n-[n/2]} - \partial_1 \mu^{[n/2]} \mu^{n-[n/2]} \\
&= \sum_{0 \le i < [n/2]} \mu^i (\partial_1 U_{n-i} - \mu \partial_1 U_{n-i-1}) \\
&\quad + \mu^{[n/2]} \big(\partial_1 + (I-\mu)\big) U_{n-[n/2]} - \mu^{[n/2]} (I-\mu) U_{n-[n/2]} \\
&\quad - \big(\partial_1 \mu^{[n/2]}\big) \mu^{n-[n/2]} \\
&= \sum_{0 \le i < [n/2]} \mu^i \big(\partial_1 + (I-\mu)\big) \partial_1 U_{n-i-1} \\
&\quad + \mu^{[n/2]} \big(\partial_1 + (I-\mu)\big) U_{n-[n/2]} + \partial_1 \mu^{[n/2]} \big(U_{n-[n/2]} - \mu^{n-[n/2]}\big).
\end{aligned}$$

Hence, by (14.10) and (6.10),

$$\begin{aligned}
\|\partial_1 \mu^n - \partial_1 U_n\|_\infty &\le \sum_{0 \le i < [n/2]} \|\mu^i\|_1 \big\|\big(\partial_1 + (I-\mu)\big)\partial_1 U_{n-i-1}\big\|_\infty \\
&\quad + \|\mu^{[n/2]}\|_1 \big\|\big(\partial_1 + (I-\mu)\big) U_{n-[n/2]}\big\|_\infty \\
&\quad + \|\partial_1 \mu^{[n/2]}\|_1 \big\|U_{n-[n/2]} - \mu^{n-[n/2]}\big\|_\infty \\
&\le \sum_{0 \le i < [n/2]} c(n-i-1)^{-(D+5)2} \\
&\quad + c n^{-(D+3)/2} + c n^{-1} n^{-(D+1)/2} \\
&\le c n^{-(D+3)/2},
\end{aligned}$$

which proves the theorem. □

21. Riesz transforms.

21.1. *Proof of Theorem* 1.25. The kernel K_k of the operator R_k is given by

$$K_k(x, y) = \sum_{n \ge 0} a_n \, \partial_{z_1} \cdots \partial_{z_k} \mu^n(x, y),$$

where the a_n's are as in the series $(1-t)^{-1/2} = \sum_{n \ge 0} a_n t^n$.

By (1.14) and (1.21), $K_k(x, y)$ satisfies the standard estimates

$$|K_k(x, y)| \le \frac{c}{|y^{-1}x|_{\Gamma_N}^D} \tag{21.1}$$

and

$$\nabla_U^x K_k(x, y) + \nabla_U^y K_k(x, y) \le \frac{c}{|y^{-1}x|_{\Gamma_N}^{D+1}}, \tag{21.2}$$

where the superindices x and y denote differences with respect to the variables x and y respectively.

So by the Calderon–Zygmund theory (cf. [17, 42]), to prove Theorem 1.25, it is enough to prove that the operator R_k is bounded on L^2. This can be done by an almost orthogonality argument (cf. [43], Chapter 7).

Let us denote by T_j, $j \in \mathbb{N}$, the operators with kernel K_j given by

$$K_j(x, y) = \sum_{2^{j-1} \le n < 2^j} a_n \, \partial_{z_1} \cdots \partial_{z_k} \mu^n.$$

Then $R_k = \partial_{z_1} \cdots \partial_{z_k} + \sum_{j \ge 1} T_j$. Also the kernels $K_j(x, y)$ satisfy

$$\sum_{y \in \Gamma_N} K_j(x, y) = \sum_{x \in \Gamma_N} K_j(x, y) = 0. \tag{21.3}$$

Furthermore, there is a $c > 0$ such that, for all $j \in \mathbb{N}$,

$$\|K_j\|_1 \le \sum_{2^{j-1} \le n < 2^j} |a_n| \|\partial_{z_1} \cdots \partial_{z_k} \mu^n\|_1 \le c$$

Hence,

$$\sup_{j \in \mathbb{N}} \|T_j\|_{L^2 \to L^2} < \infty. \tag{21.4}$$

Finally, by a straightforward calculation we can see that there is a $c > 0$ such that, for all $j \in \mathbb{N}$ and $x \in \Gamma_N$,

$$\begin{gathered}
\sum_{x \in \Gamma_N} |x|_{\Gamma_N} |K_j(x, y)| \le c 2^{j/2}, \\
\sum_{y \in \Gamma_N} |y|_{\Gamma_N} |K_j(x, y)| \le c 2^{j/2}, \\
\sum_{y \in \Gamma_N} |K_j(x, y) - K_j(e, y)| \le c 2^{-j/2} |x|_{\Gamma_N}, \\
\sum_{x \in \Gamma_N} |K_j(x, e) - K_j(x, y)| \le c 2^{-j/2} |y|_{\Gamma_N}.
\end{gathered} \tag{21.5}$$

It follows from (21.3), (21.4) and (21.5) that there is a $c > 0$ such that

$$\|T_i T_j^*\|_{L^2 \to L^2} + \|T_i^* T_j\|_{L^2 \to L^2} \le c 2^{-|i-j|/2},$$

and from this we conclude that R_k is bounded on L^2 (for details we refer the reader to [43], pages 623–625).

The same arguments also apply to the operator R_k^*. $\square$

21.2. *Proof of Theorem* 1.24. If Γ is not nilpotent, then the kernel K_z of the Riesz transform R_z does not necessarily satisfy the estimate (21.2). So to prove theorem 1.24 we use (1.22) and Theorem 1.25.

More precisely, let us consider the kernels

$$K_z^{H\mu}(x, y) = \sum_{n \geq 0} a_n \, \partial_z p_n^{H\mu}(x, y), \qquad x, y \in \Gamma,$$

$$K_j^{H\mu}(x, y) = \sum_{n \geq 0} a_n X_j p_n^{H\mu}(x, y), \qquad 1 \leq j \leq n_1, \ x, y \in \Gamma,$$

and let us denote by $R_z^{H\varphi}$ and $R_j^{H\mu}$ respectively the associated operators.

Arguing as in the proof of Theorem 1.25 in the previous section, we can prove that the operators $R_j^{H\mu}$ are bounded on $L^p(\Gamma)$, $1 < p < \infty$, and from $L^1(\Gamma)$ to weak-$L^1(\Gamma)$.

Also, if $x, y, h \in \Gamma_N$, $0 \leq i, j, \ell \leq k$ and $z = hg_\ell$, then

$$\begin{aligned} \partial_z p_t^{H\mu}(xg_i, yg_j) &= p_t^{H\mu}(xg_i hg_\ell, yg_j) - p_t^{H\mu}(xg_i, yg_j) \\ &= p_t^{H\mu}(x(g_i h)_N, y) - p_t^{H\mu}(x, y) \\ &= \partial_{(g_i h)_N} p_t^{H\mu}(x, y). \end{aligned}$$

So, if $w \in N$, then

$$\partial_w \partial_z p_t^{H\mu}(xg_i, yg_j) = \partial_w \partial_{(g_i h)_N} p_t^{H\mu}(x, y)$$

and hence there is a $c > 0$ such that, for all $w \in V$ and $t \geq 1$,

$$|\partial_w p_t^{H\mu}(xg_i, yg_j)| \leq ct^{-(D+2)/2} \exp\left(-\frac{|x^{-1}y|_\Gamma^2}{ct}\right).$$

It follows that

$$|\partial_w K_z^{H\mu}(xg_i, yg_j)| \leq \frac{c}{|x^{-1}y|_\Gamma^{D+1}}.$$

So the operator $R_z^{H\mu}$ is also bounded on $L^p(\Gamma)$, $1 < p < \infty$, and from $L^1(\Gamma)$ to weak-$L^1(\Gamma)$.

Let us consider the kernel $K_z^{H\mu}(x, y)$ that satisfies the estimate

$$S_z(x, y) = K_z(x, y) - K_z^{H\mu}(x, y) - \sum_{1 \leq j \leq n_1} (\partial_z \psi^j(x)) K_j^{H\mu}(x, y), \qquad x, y \in \Gamma.$$

Then it follows from (1.31) that for all $\varepsilon \in (0, 1)$ there is a $c > 0$ such that

$$|S(x, y)| \leq \frac{c}{|x^{-1}y|_\Gamma^{D+\varepsilon}}, \qquad x, y \in G;$$

that is, the kernel $S(x, y)$ is integrable and hence the operator

$$S = R_z - R_z^{H\varphi} - \sum_{1 \le j \le n_1} (\partial_z \psi^j)\ R_j^{H\mu}$$

is bounded on $L^p(\Gamma)$, $1 \le p \le \infty$.

Hence R_z is bounded on L^p, $1 < p < \infty$, and from $L^1(\Gamma)$ to weak-$L^1(\Gamma)$.

The same arguments also apply to the operator R_z^*. $\square$

REFERENCES

[1] ALEXOPOULOS, G. K. (1989). Quelques propriétés des fonctions harmoniques sur les groupes de Lie à croissance polynômiale. *C. R. Acad. Sci. Paris Sér. I Math.* **308** 337–338.

[2] ALEXOPOULOS, G. K. (1990). Convolution powers on discrete groups of polynomial volume growth. In *Harmonic Analysis and Number Theory: Papers in Honor of Carl S. Herz* **21** (S. W. Drury and M. Ram Murty, eds.) 31–57. Amer. Math. Soc., Providence, RI.

[3] ALEXOPOULOS, G. K. (1992). An application of homogenisation theory to harmonic analysis: Harnack inequalities and Riesz transforms on Lie groups of polynomial growth. *Canad. J. Math.* **44** 1–37.

[4] ALEXOPOULOS, G. K. (1993). An application of Homogenization theory to Harmonic analysis on solvable Lie groups of polynomial growth. *Pacific J. Math.* **159** 19–45.

[5] ALEXOPOULOS, G. K. (1997). Puissances de convolution sur les groupes à croissance polynômiale du volume. *C. R. Acad. Sci. Paris Sér. I Math.* **324** 771–776.

[6] ALEXOPOULOS, G. K. (1998). Sous-laplaciens et densités centrées sur les groupes de Lie à croissance polynômiale du volume. *C. R. Acad. Sci. Paris Sér. I Math.* **326** 539–542.

[7] ALEXOPOULOS, G. K. (2002). Centered Sub-Laplacians with drift on Lie groups of polynomial volume growth. *Mem. Amer. Math. Soc.* **155** Number 739.

[8] ALEXOPOULOS, G. K. and LOHOUÉ, N. (1993). Sobolev inequalities and harmonic functions of polynomial growth. *J. London Math. Soc.* **48** 452–464.

[9] AVELLANEDA, M. and LIN, F. H. (1987). Compactness methods in the theory of Homogenization. *Comm. Pure Appl. Math.* **40** 803–847.

[10] AVELLANEDA, M. and LIN, F. H. (1989). Un théorème de Liouville pour des equations elliptiques avec coefficients périodiques. *C. R. Acad. Sci. Paris Sér. I Math.* **609** 245–250.

[11] BASS, H. (1972). The degree of polynomial growth of finitely generated nilpotent groups. *Proc. London Math. Soc.* **25** 603–614.

[12] BAUMSLAG, G. (1971). *Lecture Notes on Nilpotent Groups*. Amer. Math. Soc., Providence, RI.

[13] BENSOUSSAN, A., LIONS, J. L. and PAPANICOLAOU, G. (1978). *Asymptotic Analysis of Periodic Structures*. North-Holland, Amsterdam.

[14] BERGSTRÖM, H. (1969). On the Central Limit Theorem in $\mathbb{R}^k$. *Z. Wahrsch. Verw. Gebiete* **14** 113–126.

[15] CHRIST, M. (1995). Temporal regularity of random walks on discrete nilpotent groups. In *Proceedings of the Conference in Honor of J.-P. Kahane* 141–151. CRC Press, Boca Raton, FL.

[16] CHRIST, M. and GELLER, D. (1984). Singular integral characterizations of Hardy spaces on homogeneous groups. *Duke Math. J.* **51** 547–598.

[17] COIFMAN, R. and WEISS, G. (1971). *Analyse non-commutative sur certains espaces homogènes*. *Lecture Notes in Math.* **242**. Springer, New York.

[18] COLDING, T. H. and MINICOZZI, W. P., II (1997). Harmonic functions with polynomial growth. *J. Differential Geom.* **46** 1–77.

[19] CRÉPEL, P. and RAUGI, A. (1978). Théorème central limite sur les groupes nilpotents. *Ann. Inst. H. Poincaré Probab. Statist.* **14** 145–164.

[20] FELLER, W. (1971). *An Introduction to the Theory of Probability and Its Applications* **2**. Wiley, New York.

[21] FOLLAND, G. B. and STEIN, E. (1982). *Hardy Spaces on Homogeneous Groups*. Princeton Univ. Press.

[22] GROMOV, M. (1981). Groups of polynomial growth and expanding maps. *Inst. Hautes Études Sci. Publ. Math.* **53** 53–78.

[23] HEBISCH, W. and SALOFF-COSTE, L. (1993). Gaussian estimates for Markov chains and random walks on groups. *Ann. Probab.* **21** 673–709.

[24] HÖRMANDER, L. H. (1967). Hypoelliptic second order differential operators. *Acta. Math.* **16** 147–171.

[25] JACOBSON, N. (1962). *Lie Algebras*. Wiley, New York.

[26] JIKOV, V. V., KOZLOV, S. M. and OLEINIK, O. A. (1994). *Homogenization of Differential Operators and Integral Functionals*. Springer, New York.

[27] KENIG, C. E. (1994). *Harmonic Analysis Techniques for Second Order Elliptic Boundary Value Problems*. Amer. Math. Soc., Providence, RI.

[28] KOZLOV, S. M. (1980). Asymptotics of fundamental solutions for second order differential equations. *Mat. Sb.* **113** 302–323. [English translation (1982) *Math. USSR.-Sb.* **41** 249–267.]

[29] KRYLOV, N. V. and SAFONOV, M. V. (1981). A certain property of solutions of parabolic equations with measurable coefficients. *Math. USSR-Izv.* **16** 151–164.

[30] LAWLER, G. F. (1991). Estimates for differences and Harnack inequality for difference operators coming from random walks with symmetric, spacially inhomogeneous, increments. *Proc. London Math. Soc.* **63** 552–568.

[31] LOHOUÉ, N. and VAROPOULOS, N. TH. (1985). Remarques sur les transformés de Riesz sur les groupes de Lie nilpotents. *C. R. Acad. Sci. Paris Sér. I Math.* **301** 11.

[32] MOUSTAPHA, S. and VAROPOULOS, N. TH. Cambridge Univer. Press. To appear.

[33] NAGEL, A., RICCI, F and STEIN, E. (1990). Harmonic Analysis and fundamental solutions on nilpotent Lie groups. In *Analysis and Partial Differential Equations*: *A Collection of Papers Dedicated to M. Cotlar* (C. Sadosky ed.) 249–275. Dekker, New York.

[34] PETROV, V. V. (1975). *Sums of Independent Random Variables*. Springer, New York.

[35] RAGHUNATHAN, M. S. (1972). *Discrete Subgroups of Lie Groups*. Springer, New York.

[36] RAUGI, A. (1978). Théorème de la limite centrale sur les groupes de Lie nilpotents. *Z. Wahrsch. Verw. Gebiete* **43** 149–172.

[37] RAUGI, A. (1978). Théorème de la limite centrale pour un produit semidirect d'un groupe de Lie résoluble simplement connexe de type rigide par un groupe compact. *Probability Measures on Groups. Lecture Notes in Math.* **706** 257–324. Springer, New York.

[38] RIESZ, F. and VON SZ. NAGY, B. (1968). *Leçons d'analyse fonctionelle*, 5th ed. Gauthier-Villars, Paris.

[39] SAFONOV, M. V. (1983). Harnack's inequality for elliptic equations and the Hölder property of their solutions. *J. Soviet Math.* **21** 851–863.

[40] SALOFF-COSTE, L. (1990). Analyse sur les groupes de Lie à croissance polynômiale. *Ark. Mat.* **28** 315–331.

[41] SAZONOV, V. V. (1981). *Normal Approximation—Some Recent Advances. Lecture Notes in Math.* **879**. Springer, New York.

[42] STEIN, E. (1970). *Singular Integrals and Differentiability Properties of Functions*. Princeton Univ. Press.

[43] STEIN, E. (1993). *Harmonic Analysis: Real-Variable Methods, Orthogonality and Oscillatory Integrals*. Princeton Univ. Press.

[44] STROOCK, D. W. (1988). Diffusion semigroups corresponding to uniformly elliptic divergence form operators. *Séminaire de probabilités XXII. Lecture Notes in Math.* **1321** 316–347. Springer, New York.
[45] VARADARAJAN, V. S. (1984). *Lie Groups, Lie Algebras and Their Representations*. Springer, New York.
[46] VAROPOULOS, N. TH. (1986). Analysis on nilpotent groups. *J. Funct. Anal.* **66** 406–431.
[47] VAROPOULOS, N. TH. (1987). Fonctions harmoniques positives sur les groupes de Lie. *C. R. Acad. Sci. Paris Sér. I Math.* **304** 519–521.
[48] VAROPOULOS, N. TH. (1988). Analysis on Lie groups. *J. Funct. Anal.* **76** 346–410.
[49] VAROPOULOS, N. TH. (1994). Wiener–Hopf theory and nonunimodular groups. *J. Funct. Anal.* **120** 467–483.
[50] VAROPOULOS, N. TH., SALOFF-COSTE, L. and COULHON, TH. (1993). *Analysis and Geometry on Groups*. Cambridge Univ. Press.
[51] ZHIKOV, V. V. (1989). Spectral approach to asymptotic diffusion problems. *Differensial'nye Uravneniya* **25** 44–50. [English translation (1989) *Differential Equations* **25** 33–39.]

MATHÉMATIQUES
UNIVERSITÉ DE PARIS-SUD
BÂT. 425
91405 ORSAY CEDEX
FRANCE