Revista Matemática Iberoamericana
Vol. 12, N.° 1, 1996

The heat kernel on Lie groups

N. Th. Varopoulos

0. Introduction.

0.1. Background: Lie groups.

In this paper G will denote a real connected amenable Lie group; $N \subset Q \subset G$ will denote the radical and the nilradical. Amenability on G is equivalent to the fact that the (semisimple) group $G/Q = S$ is compact. The group S can then be identified locally to some closed subgroup of G (a Levi subgroup).

If G is simply connected then all the groups N, Q, S are simply connected and $G = Q \lambda S$ is a semidirect product. We shall suppose throughout until the final section that G is simply connected. The choice of S is not unique but $G/N = Q/N \times S$ where the product is now direct (and not just semidirect *cf.* [1]) and $Q/N = V (\cong \mathbb{R}^a)$ furthermore S identified to a subgroup of G/N is then unique. Indeed if $\pi : G/N \longrightarrow Q/N \cong \mathbb{R}^a$ is the projection that we obtain from such an identification then $\pi(S_1) = \{0\}$ for any other subgroup S_1 that is either compact or semisimple. Let now $\Delta = -\sum X_j^2$ be some subelliptic left invariant "Laplacian" on G ,*i.e.*, $X_1, \dots, X_k$ are left invariant vector fields (*i.e.*, $(Xf)_g = Xf_g$; $f_g(x) = f(gx)$) that satisfy the Hörmander condition and generate together with all their brackets the tangent space on G (*cf.* [2]). Using the natural (uniquely defined!) projections

$$G \longrightarrow V \times S \longrightarrow V \tag{0.1}$$

we can define $\dot{\Delta}$ an elliptic operator V by projecting Δ on V. There exists therefore one and only one (up to orthogonal transformation) choice of coordinates $V \cong \mathbb{R}^a$ (in other words the corresponding scalar product on V is uniquely defined!) for which

$$\dot{\Delta} = -\sum_{i=1}^{a} \frac{\partial^2}{\partial^2 x_i}\,. \tag{0.2}$$

I shall denote by $\langle\cdot\,,\cdot\rangle_\Delta = \langle\cdot\,,\cdot\rangle$ the above well determined scalar product on V.

If we make the additional assumption that the Lie algebra $\mathfrak{g}$ of G is algebraic (or equivalently that G coincides locally with an algebraic group) then the structure of G simplifies further since we can then find $\mathbb{R}^a \cong V \subset Q$ such that $Q = N\lambda V$, and S an appropriate Levi subgroup, such that V and S commute (*cf.* [3]). We have thus a representation

$$G = N\lambda(V \times S)\,. \tag{0.3}$$

This is the model that the reader should keep in mind in what follows. We shall explain the correct substitute of (0.3) for the general (*i.e.* not necessarily algebraic) groups.

0.2. Background: The heat diffusion semigroup and kernel.

The notations are as in Section 0.1. I shall denote by $T_t = e^{-t\Delta}$ the heat diffusion semigroup generated by Δ and I shall also denote by

$$d^l g = dg\,, \quad d^r g = dg^{-1}\,, \quad m(g) = d^r g/dg$$

the left and right Haar measure on G and the modular function. I shall also denote by

$$\tilde{T}_t = m^{1/2}\, T_t\, m^{-1/2} = \exp(-t\tilde{\Delta}) = \exp(-t\, m^{1/2}\, \Delta\, m^{-1/2})\,, \quad t > 0\,.$$

The semigroup T_t (respectively $\tilde{T}_t$) is symmetric with respect to $d^r g$ (respectively dg). I shall denote by $\phi_t(g)$ (respectively $\psi_t(g)$) the convolution kernel of T_t (respectively $\tilde{T}_t$) with respect to dg and by $\mu_t \in P(G)$ the corresponding convolution measures:

$$T_t f(x) = \int \phi_t(y^{-1}x)\, f(y)\, dy = f * \mu_t(x) = \int f(xy^{-1})\, d\mu_t(y)\,,$$

$$\tilde{T}_t f(x) = \int \psi_t(y^{-1}x)\, f(y)\, dy \qquad x \in G\,,\ t > 0\,,\ f \in C_0^\infty(G)\,,$$

where

$$\psi_t(g) = m^{1/2}(g)\,\phi_t(g) = \psi_t(g^{-1})$$

and

$$\begin{aligned} d\mu_t(g) &= d\mu_t(g^{-1}) = \phi_t(g)\,d^r g = m(g)\,\phi_t(g)\,dg \\ &= m^{1/2}(g)\,\psi_t(g)\,dg = m^{-1/2}(g)\,\psi_t(g)\,d^r g\,. \end{aligned}$$

The semigroup T_t defines a diffusion on G: $\Omega = \{z(t) \in G : \ t > 0\}$ and for the corresponding probabilities on the path space Ω we denote as usual

$$\mathbf{P}_x(z(t) \in dy) = P_t(x, dy) = \mathbf{P}(z(t) \in dy : \ z(0) = x)$$

and we have

$$T_t f(x) = \int f(y)\,P_t(x, dy)\,.$$

We therefore have

$$d\mu_t(g) = d\mu_t(g^{-1}) = P_t(e, dg)\,.$$

0.3. The disintegration of the Haar measure and the L^p-norms.

If we use the projection $G \longrightarrow G/N = A = V \times S$ we can disintegrate the Haar measure

$$\int f(g)\,d^r g = \int_A \Big(\int_N f(na)\,dn \Big)\,da$$

this we shall summarise by saying

$$g = n\,a\,, \quad d^r g = dn\,da\,, \quad n \in N\,, \quad a \in A\,,$$

similarly

$$\int f(g)\,dg = \int_A \Big(\int_N f(an)\,dn \Big)\,da\,,$$

where $g = a\,n$, $dg = dn\,da$, $n \in N$, and $a \in A$.

The notations in the above integrals are of course clear enough but somewhat abusive. Indeed A cannot necessarily be identified to a subgroup of G but there is always some section $A \longrightarrow G$ (of $G \longrightarrow A$) (*cf.* [4]) and this identifies A as a subset of G. (In fact the above iterated integrals make obvious sense even without the existence of the above section (*cf.* [5])). We shall now fix $A \subset G$ such a section and express the diffusion

$$\Omega = \{z(t) = n(t)\, a(t) : \ t > 0\}\,, \quad n(t) \in N\,,\ a(t) \in A\,,\ t > 0\,.$$
$$\Omega_A = \{a(t) \in A : \ t > 0\}\,.$$

Ω_A corresponds to the path space of the diffusion on $A = V \times S$ generated by Δ_A the image of Δ by $G \longrightarrow G/N = A$.

Let us denote by $G_t^A(x)$ ($x \in A$) the convolution kernel of $e^{-t\Delta_A}$ on A. Here the notation G_t^A has been deliberately chosen to invoque the Gaussian functions:

$$G_t^a(x) = \frac{1}{(4\pi t)^{a/2}} \exp\big(-\frac{|x|^2}{4t}\big)\,, \qquad x \in \mathbb{R}^a\,,\ t > 0\,.$$

because when $S = \{e\}$ reduces to the identity we have $G_t^A = G_t^a$. The case $S = \{e\}$ is the model that the reader should keep in mind.

The reader should also keep in mind that we always have

$$\int_S G_t^A(x,s)\, ds = G_t^a(x)\,, \qquad x \in V \cong \mathbb{R}^a\,,$$

this by the standard local Harnack estimate gives

$$C^{-1} G_{t-1}^a(x) \le G_t^A(x,s) \le C\, G_{t+1}^a(x)\,, \tag{0.4}$$

for all $t \ge 10$, and $(x,s) \in V \times S$. With the above notations we have

$$\phi_t(na)\, dn = \mathbf{P}_e(n(t) \in dn : \ a(t) = a)\, G_t^A(a)\,. \tag{0.5}$$

The above disintegration allows us now to write the following formulaes:

$$\begin{aligned}
\|\tilde{T}_t\|_{1\to p}^p = \|\psi_t\|_p^p &= \int \psi_t^p(g)\, dg = \int \psi_t^p(g)\, d^r g \\
&= \int \psi_t^p(g)\, m(g)\, dg = \int \psi_t^p(g)\, m^{-1}(g)\, d^r g \\
&= \int_A dx \int_N \psi_t^p(xn)\, dn = \int_A dx \int_N \psi^p(nx)\, dn \\
&= \int_A dx\, m(x) \int_N \psi_t^p(xn)\, dn = \int_A dx\, m^{-1}(x) \int_N \psi^p(nx)\, dn\,,
\end{aligned}$$

$$\int_N \phi_t(nx)\,dn = m(x)\int \phi_t(xn)\,dn = G_t^A(x)\,,$$

$$\int \psi_t(nx)\,dn = m^{1/2}(x)\,G_t^A(x)\,, \qquad \int \psi_t(xn)\,dn = m^{-1/2}(x)\,G_t^A(x)\,.$$

here and throughout we denote by $\|\cdot\|_p$ the norm in $L^p(G;dg)$ and for any operator we denote by $\|\cdot\|_{p\to q}$ the norm in $L^p(G;dg) \longrightarrow L^q(G;dg)$.

0.4. The roots and the modular function.

Let $\mathfrak{n} \subset \mathfrak{q} \subset \mathfrak{g}$ be the Lie algebra of G, its radical and its nilradical. We have of course $\mathfrak{g}/\mathfrak{n} = (\mathfrak{q}/\mathfrak{n}) \times \mathfrak{s}$ where $\mathfrak{s}$ is the Lie algebra of S. The *ad* action on $\mathfrak{g}$ induces

$$ad_{\mathfrak{n}_{\mathbb{C}}} q = \begin{pmatrix} \rho_1(q) & & * \\ & \ddots & \\ 0 & & \rho_n(q) \end{pmatrix}, \qquad q \in \mathfrak{q}\,.$$

linear endomorphisms on $\mathfrak{n}_{\mathbb{C}} = \mathfrak{n} \otimes \mathbb{C}$ the complexified of $\mathfrak{n}$ which for an appropriate basis of $\mathfrak{n}_{\mathbb{C}}$ can be simultaneously triangulated for every $q \in \mathfrak{q}$ (Lie's theorem, *cf.* [1], [6], *cf.* also [7] where the above roots are systematically examined from a point of view that is adapted to our needs). We can identify $\rho_j \in \mathrm{Hom}_{\mathbb{R}}(\mathfrak{q};\mathbb{C})$ and we shall denote $L_j = \mathrm{Re}\,\rho_j \in \mathfrak{q}^* = \mathrm{Hom}_{\mathbb{R}}(\mathfrak{q};\mathbb{R})$, $(j = 1,\dots,n)$. It is clear that $\rho_j|_{\mathfrak{n}} \equiv 0$ and therefore $\rho_j \in \mathrm{Hom}_{\mathbb{R}}(\mathfrak{q}/\mathfrak{n};\mathbb{C})$.

The space $\mathfrak{q}/\mathfrak{n}$ can be canonically identified with $V = Q/N$ and therefore we can identify the ρ_j's with functions on $V \times S$ and therefore also on G. (What is more natural is to identify the e^{ρ_j}'s with functions on G).

The definition of *ad* that we have adopted above is the differential at the identity of $Ad : G \longrightarrow GL(\mathfrak{g})$. And

$$Ad(g) = d_h I_g(h)|_{h=e}\,, \qquad I_g : G \longrightarrow G\,, \quad I_g(h) = h^g = ghg^{-1},$$

in terms of the roots we have then

$$Ad(q)|_{\mathfrak{n}_{\mathbb{C}}} = \begin{pmatrix} e^{\rho_1(q)} & & * \\ & \ddots & \\ 0 & & e^{\rho_n(q)} \end{pmatrix}, \qquad q \in Q\,.$$

Indeed this clearly holds when $q \in \operatorname{Exp}(\mathfrak{q})$ and therefore also for all q.

An elementary verificati on that has to be carried out by the reader shows therefore that with the above identifications we have:

$$\exp\left(\sum \rho_j(g)\right) = \exp\left(\sum L_j(g)\right) = m(g)\,, \qquad g \in G\,.$$

We shall finally recall the following.

Definition. *We shall say that the group G is* NC-*(respectively* WNC) *if there exists $x \in V$ such that $L_j(x) > 0$ for all $L_j \neq 0$, $j = 1, \dots, n$ (respectively, $L_j(x) \geq 0$, $j = 1, \dots, n$, $\sum L_j(x) > 0$). When G is semi-simple or $\{0\}$ and the roots are not defined we say that G is* NC *but not* WNC.

The C-condition was first introduced in [8] (C stands there for "condition". I guess I should have called it the V condition, now is too late to do anything about it). NC-stands for Non-C, WNC stands for weak NC which is slightly abusive. Observe that if we suppose that the L_i's span V^*, the negation of WNC is that for every $0 \neq x \in V$ there exist $1 \leq j, k \leq n$ such that $L_j(x) < 0$, $L_k(x) > 0$. A unimodular group is NC if and only if $L_j = 0$ $(j = 1, 2, \dots, n)$ and such a group is not WNC.

0.5. The main estimate.

The upper estimate. All the notations are as before. We have for the supremum on $n \in N$ over each fiber:

$$\begin{aligned} \sup_{n\in N} \phi_t(nx) &= \sup_n \phi_t(xn) \\ &\leq C_\varepsilon \exp\left(-\sum L_j^+(x)\right) G^A_{(1+\varepsilon)t}(x)\,, \end{aligned} \tag{0.6}$$

for all

$$t \geq t_\varepsilon\,, \qquad x \in G/N = A = V \times S\,,$$

and where we denote as usual $r^+ = \sup\{r, 0\}$, $r^- = \inf\{r, 0\}$, $(r \in \mathbb{R})$ and where $0 < \varepsilon < 1$ is arbitrary and C_ε, $t_\varepsilon > 0$ depends on ε.

Equivalently the above estimates can be formulated in terms of ψ_t:

$$\begin{aligned} \sup_n \psi_t(nx) &= \sup_n \psi_t(xn) \\ &\leq C_\varepsilon \exp\left(-\frac{1}{2}\sum |L_j(x)|\right) G^A_{(1+\varepsilon)t}(x)\,. \end{aligned} \tag{0.7}$$

This estimate together with (0.5) implies at once

$$\begin{aligned}(0.8)\qquad &\Big(\int \phi_t^p(nx)\,dn\Big)^{1/p}\\ &\qquad \le C_\varepsilon \exp\Big(\Big(1-\frac{1}{p}\Big)\Big(-\sum L_j^+(x)\Big)\Big) G^A_{(1+\varepsilon)t}(x)\end{aligned}$$

and

$$\begin{aligned}(0.9)\qquad \int \psi_t^p(nx)\,dn \le C_\varepsilon \exp\Big(\Big(1-\frac{p}{2}\Big)&\sum L_j^+(x)\\ &+\frac{p}{2}\sum L_j^-(x)\Big)\,(G^A_{(1+\varepsilon)t}(x))^p\,,\end{aligned}$$

$$\begin{aligned}(0.10)\qquad \int \psi_t^p(xn)\,dn \le C_\varepsilon \exp\Big(-\frac{p}{2}&\sum L_j^+(x)\\ &+\Big(\frac{p}{2}-1\Big)\sum L_j^-(x)\Big)\,(G^A_{(1+\varepsilon)t}(x))^p\,,\end{aligned}$$

for all

$$t \ge t_\varepsilon\,,\qquad x\in A = V\times S\,,\qquad 1\le p\le +\infty\,.$$

If we integrate the above on A (with respect to dx) we obtain at once upper estimates of $\|\psi_t\|_p$. Observe also that because of (0.4) we can replace $G_t^A(\,\cdot\,)$ by the standard Gaussian on V.

The lower estimate. The most convenient way to express the lower estimate is through probabilistic language. We shall show that for all $x\in V$ and all $t\ge 10$ there exists $P\subset N$ some subset (that depends on x and t) such that:

$$\begin{aligned}(0.11)\qquad \text{Vol.} &= \text{Haar-mes}_N(P)\\ &\le C\exp\Big(c\,t^{1/3} + C\,\frac{|x|}{\sqrt{t}} + \sum L_j^+(x)\Big)\,,\end{aligned}$$

$$(0.12)\qquad \text{Pr.} = \mathbf{P}(z(t)\in P\times B_x) \ge C\exp(-c\,t^{1/3})\,G_t^a(x)\,,$$

where $B_x = \{y\in V:\ |x-y|\le 1\}\times S\subset A$ and where A has been identified to some fixed section of $G\longrightarrow G/N$ as in Section 0.3.

The above estimate gives of course information about the rapidity with which the mass goes to infinity under the diffusion $\{z(t) : t > 0\}$. An obvious application of Hölder's estimate gives at once

$$\text{Pr.} \leq \Big(\int_{N \times B_x} \phi_t^p(nx)\, dn\, dx \Big)^{1/p} (\text{Vol.})^{1-1/p}, \quad 1 \leq p \leq +\infty .$$

This together with the local Harnack estimate gives

$$\begin{aligned} &\Big(\int_N \phi_t^p(ny)\, dn \Big)^{1/p} \\ &\qquad \geq C_\varepsilon \exp\Big(- c\, t^{1/3} - \Big(1 - \frac{1}{p}\Big) \sum L_j^+(x) \Big) G^a_{(1-\varepsilon)t}(x), \end{aligned} \tag{0.13}$$

for all $1 \leq p \leq +\infty$, $t \geq 10$, $0 < \varepsilon < 1$, and $y = (x, \sigma) \in V \times S$.

It follows in particular that up to a factor $e^{-ct^{1/3}}$ (which in view of [4], [9] is perfectly natural) the upper and lower estimates that we have given are sharp.

NC-lower estimates. If we make the additional assumption that the group G is NC we can make a substantial improvement in the above lower estimates! In that case we can replace the factors $\exp(\pm c\, t^{1/3})$ in the right hand side of (0.11) and (0.12) by the polynomial factors $t^{\pm c}$ respectively. This allows us to improve the estimate (0.13) and obtain

$$\Big(\int_N \phi^p(ny)\, dn \Big)^{1/p} \geq C_\varepsilon\, t^{-c} \exp\Big(- \Big(1 - \frac{1}{p}\Big) \sum L_j^+(x) \Big) G^a_{(1-\varepsilon)t}$$

with the same notations as in (0.13).

An analogous refinement that refers to the upper estimate holds for C-groups (*cf.* end of Section 1.2).

0.6. The Hardy-Littlewood theory.

The first thing to observe is that

$$\liminf_{t \to \infty} \frac{1}{t} \log \|\tilde{T}_t\|_{p \to q} \geq 0, \qquad 1 \leq p, q < +\infty .$$

Let us consider the spectral decomposition of $\tilde{T}_t$ in $L^2(G; dg)$:

$$\tilde{T}_t = \int_0^\infty e^{-\lambda t}\, dE_\lambda .$$

By the amenability of G it follows that for all $\lambda > 0$ there exists $\varphi \in C_0^\infty$ such that $\langle E_\lambda \varphi, \varphi\rangle \geq c > 0$ and therefore $\langle \tilde{T}_t\varphi, \varphi\rangle > c\, e^{-\lambda t}$ $(t > 0)$.

We have, on the other hand,

$$\|\tilde{T}_t\|_{p\to q} = O(1)\,, \qquad 1 \leq p \leq 2 \leq q \leq +\infty\,.$$

This is easy and was pointed out in [10] (*cf.* (4.1) below). Let us consider

$$\ell(q) = \limsup_{t\to\infty} \frac{1}{t} \log \|\tilde{T}_t\|_{1\to q} = \limsup_{t\to\infty} \frac{1}{t} \log \|\psi_t\|_q$$

which is a convex function of $1/q \in [0,1]$ and identically 0 for $q \in [2,+\infty]$. It follows that $\ell(q)$ is continuous for $q \in (1,+\infty]$ and in fact it is also continuous for $q = 1$. This last point is of course a consequence of the explicit formulaes given below but can also be seen in a more "abstract" way. Indeed by the "general" Gaussian estimates for ψ_t we see that $\ell(q) = \limsup (1/t) \log \|\psi_t\|_q$ can be defined and is finite and convex in q for all $q \in (0,+\infty]$. The continuity follows. At any rate the estimates of Section 0.5 allows us to obtain the following more precise information.

Theorem 1. *Let $L_1,\dots,L_n \in V^* = (\mathfrak{q}/\mathfrak{n})^*$ be the real parts of the roots as in Section* 0.4 *and let $\langle\,\cdot\,,\,\cdot\,\rangle$ the scalar product on $V = \mathfrak{q}/\mathfrak{n}$ defined in Section* 0.4. *We have then*

$$\begin{aligned}\ell(p) &= \lim_{t\to\infty} \frac{1}{t} \log \|\psi_t\|_p \\ &= \lim_{t\to\infty} \frac{1}{t} \log \int_V \exp\Big(\Big(\frac{1}{p} - \frac{1}{2}\Big) \sum L_j^+(x) \\ &\qquad\qquad + \frac{1}{2} \sum L_j^-(x) - \frac{|x|^2}{4t}\Big)\, dx\,.\end{aligned} \tag{0.14}$$

In (0.14) $|\cdot|^2 = \langle\,\cdot\,,\,\cdot\,\rangle$ denotes of course the corresponding euclidean norm on V. We choose to set $V = \mathfrak{q}/\mathfrak{n}$ rather than Q/N because formulated like this the theorem makes sense even for non simply connected groups. The point is that the above theorem holds as stated for general, not necessarily simply connected, groups. An explicit formula for the limit (0.14) can also be given by an elementary computation.

What is significant is the first point where $\ell(q)$ vanishes:

$$L = \inf\{q \geq 1:\ \ell(q) = 0\} \leq 2\,.$$

We have $L = 1$ if and only if G is unimodular. To see this we can either use our theorem or we can use the value of $\|\tilde{T}_t\|_{1\to 1}$ (*cf.* Theorem 2.iv) and [10]) and the continuity of $\ell(q)$.

An important consequence of the theorem is the following:

Theorem 2. *Let $G, \tilde{T}_t$ and L be as above. Then*

i) *The constant $1 \le L \le 2$ depends only on G and is in particular independent of the particular choice of Δ.*

ii) *We have $L = 1$ if and only if G is unimodular. We have $L = 2$ if and only if G is* WNC.

iii) *If G is not* WNC, *for every $1 \le p < 2$ there exists $p < q < 2$ such that*

$$\|\tilde{T}_t\|_{p\to q} = O(1).$$

iv) *For any group and any $1 \le p \le +\infty$ we have*

$$\|\tilde{T}_t\|_{p\to p} = \exp\left(t\left(\frac{1}{2} - \frac{1}{p}\right)^2 \rho^2\right),$$

where ρ^2 is defined by $\Delta m = -\rho^2 m$. If G is WNC *we have*

$$\lim \frac{1}{t} \log \|\tilde{T}_t\|_{p\to q} = \left(\frac{1}{2} - \frac{1}{q}\right)^2 \rho^2, \qquad 1 \le p \le q \le 2.$$

v) *Conversely if G is non unimodular and if*

$$\limsup \frac{1}{t} \log \|\tilde{T}_t\|_{p\to q} \ge \left(\frac{1}{2} - \frac{1}{q}\right)^2 \rho^2 \tag{0.15}$$

for some $1 \le p < q < 2$, then G is a WNC *group.*

Once more the above theorem holds for general, not necessarily simply connected groups.

The main problem of Hardy-Littlewood theory in the above context is to find all the combinations $\alpha \in \mathbb{C}$, $\operatorname{Re}\alpha \ge 0$ for which the following mapping

$$\tilde{\Delta}^{-\alpha/2} : L^p(G; dg) \longrightarrow L^q(G; dg),$$

$$\tilde{\Delta}^{-\alpha/2} = c_\alpha \int_0^\infty t^{\alpha/2-1}\, \tilde{T}_t\, dt \tag{0.16}$$

is bounded. From our previous results on $\tilde{T}_t$ we can obtain some partial but significant information in that direction. First of all, the operator

$$\tilde{\Delta}^{-\alpha/2} = c_\alpha \int_0^\infty t^{\alpha/2-1}\, \tilde{T}_t\, dt \geq 0\,, \qquad \alpha > 0\,, \tag{0.17}$$

is positive and therefore $\tilde{\Delta}^{-\alpha/2} f$ is defined (possibly $+\infty$) for every $f \geq 0$. We have then

Corollary 1. *If G is a* WNC*-group then $\|\tilde{\Delta}^{-\alpha/2}\|_{p\to q} = +\infty$ for every $1 \leq p < q < 2$, $\alpha > 0$.*

The notation $\|\cdot\|_{p\to q} = +\infty$ is abusive but in view of (0.17) it has an obvious meaning. This corollary says that the mapping (0.16) cannot be bounded unless we cross the L^2- level. What happens for $1 \leq p \leq 2 \leq q \leq +\infty$ is in general an open question. Partial results can be found in [10], [11], [12], [13].

The above corollary essentially characterises WNC groups. Recall that quite generally there exists $\delta = 1, 2, \ldots$ such that $\psi_t(e) \sim t^{-\delta/2}$ $(t \to 0)$. We have then

Corollary 2. *Assume that G is neither a* NC *nor a* WNC *group, and let $1 \leq p < 2$. Then there exists $p \leq L(p) < 2$ such that for all $q \in]L(p), +\infty]$ we have*

$$\|\tilde{T}_t\|_{p\to q} = o(e^{-ct^{1/3}})\,, \qquad as\ \ t \to \infty\,,$$

for some $c > 0$. Furthermore if $1 < p \leq L(p) < q < +\infty$ the mapping (0.16) is bounded if and only if $1/p - 1/q \leq \operatorname{Re}\alpha/\delta$.

1. The upper estimate.

1.1. The general set up.

Let G be a simply connected Lie group, let $N \subset Q \subset G$ be the radical and nilradical, let $H = Z(N) \subset N$ be the center of N, let also $\alpha : G \longrightarrow G/H$ be the canonical projection. Let us also recall that $G/N \cong V \times S$ where $V \cong \mathbb{R}^m$ and S is semisimple (and compact if G is amenable). In this section I shall rely very heavily on the ideas the methods and even the notations introduced in [8] (*e.g.*, I shall set

$V \cong \mathbb{R}^m$ rather than $\mathbb{R}^a$) and there is no way at all that the reader can read this section without constantly refering back to [8].

The crucial thing is that (*cf.* [8, Section 1] where the notations are the same as above) G/H acts by inner automorphisms on H because H is abelian. Furthermore the estimates of [8, Section 1] hold. The extra "twist" for us refers to [8, Section 2]: H being central in N it follows that the above action of $G/H \longrightarrow GL(H)$ is trivial on N/H so that we can factorise this action by

$$G/H \xrightarrow{\pi} G/N = V \times S \longrightarrow GL(H).$$

Furthermore the estimates on the determinants given in [8, Section 2] apply. We then move on to [8, Section 3] to observe that everything applies *verbatim* here except that $G/H \not\cong V \times S$ and therefore in the last few lines [8, Section 3] we have to define

$$A_n(L_i) = \inf_{1 \le j \le n} \exp\left(c\,|g_j|^2_{G/H} - L_i(b_j)\right),$$
$$(b_j, \sigma_j) = \pi(s_j) \in V \times S, \qquad i = 1, 2, \ldots, p.$$

To simplify notations I have droped here and in what follows the cofactor d_i (the dimension of the corresponding root space). I shall denote by $L_1, \ldots, L_p$ the real parts of the roots of the action of G/H and identify them with a subset of the real parts of the roots of G.

With all the other notations being as in [8, Section 3] the estimate (3.3) of [8] therefore holds in our setting and if we follow closely [8, Section 3] we see that for every $r \in G/H$ we have

$$\begin{aligned}(1.1)\qquad \Big(\sup_{g \in \alpha^{-1}(r)} \phi_n(g)\Big)\, d^r r &\le C\,\mathbb{E}\big(A_n(L_1)\cdots A_n(L_p);\ z_{G/H}(n) \in dr\big) \\ &= C\,\mathbb{E}\big(A_n(L_1)\cdots A_n(L_p)\, I\{z_{G/H}(n) \in dr\}\big),\end{aligned}$$

where $n \ge 1$. This is an inequality between two measures on G/H where $\{z_{G/H}(t) \in G/H :\ t > 0\}$ is the diffusion naturally induced on G/H by our laplacian Δ. The next step is to project $\pi(r) = (x, \sigma) \in V \times K$ and to distinguish two cases:

Case i): The $x \in G/H$ (*cf.* (0.6)) is such that $L_i(x) \le 0$.

In the infimum that defines $A_n(L_i)$ we then set $j = 1$ and obtain

$$A_n(L_i) \le C\, \exp(c\,|g_1|^2 + c\,|g_1|) \le C\, \exp(2\,c\,|g_1|^2).$$

Case ii): $L_i(x) > 0$.

In the infimum that defines $A_n(L_i)$ we then set $j = n$ and obtain

$$A_n(L_i) \le C \exp(c\,|g_n|^2 - L_i^+(x)), \qquad b_n = x .$$

For $b_n = x$ we can therefore estimate the right hand side of (1.1) by

$$(1.2) \qquad \exp\Big(-\sum L_i^+(x)\Big) \cdot \mathbb{E}\big(\exp(2\,cp\,(|g_1|^2_{G/H} + |g_n|^2_{G/H});\ z_{G/H}(n) \in dr\big),$$

for $i = 1, 2, \dots, p$.

In the case when $H = N$ (which is the case considered in [8]) the above estimate simplifies since, S being compact, we can replace

$$|g_j|_{G/H} \approx |X_j|_V , \qquad (X_j, \tilde{\sigma}_j) = \pi(g_j) \in V \times S .$$

In general however this is not possible.

When $H = N$ the above estimate allows us to conclude very easily. Indeed we have then $dr = dx\, d\sigma$ and if we use the local Harnack estimate on the left hand side of (1.1) we see that since S is compact, by replacing n by $n-1$, and with $\alpha(g) = r = (x, \sigma) \in V \times S$ $(g \in G)$, we have

$$\phi_{n-1}(g)\, dx \le C \exp\Big(-\sum L_i^+(x)\Big) \cdot \mathbb{E}\big(\exp 2\,cp\,(|X_1|^2 + |X_n|^2);\ b(n) \in dx\big),$$

for $n > 1$. (The notations are of course as in [8, Section 4] and we use interchangeably $b_j = b(j)$ $(j = 1, 2, \dots)$ which is Brownian motion at times $t = 1, 2, \dots)$. The coefficient of dx in the above expectation is equal to $(m = \dim V)$

$$(1.3) \quad n^{-m/2} \iint \exp\Big(\Big(2\,p\,c - \frac{1}{4}\Big)(|\xi|^2 + |\varsigma|^2) - \frac{|x - (\xi + \varsigma)|^2}{4\,(n-2)}\Big)\, d\xi\, d\varsigma .$$

A direct computation on the Gaussian shows that the double integral is comparable to

$$\exp\Big(-\frac{|x|^2}{4}\Big(\frac{1}{n} + \frac{a}{n^2}\Big)\Big).$$

Alternatively Peter-Paul gives:

$$\begin{aligned}|x-(\xi+\varsigma)^2| &\geq |x|^2+|\xi+\varsigma|^2-\varepsilon_0\,|x|^2-\varepsilon_0^{-1}\,|\xi+\varsigma|^2\\ &=(1-\varepsilon_0)\,|x|^2+(1-\varepsilon_0^{-1})\,|\xi+\varsigma|^2\,, \qquad 0<\varepsilon_0\ll 1\,,\end{aligned}$$

which means that we can estimate (1.3) by

$$c_\varepsilon\, n^{-m/2}\exp\Big(-\frac{|x|^2}{(4+\varepsilon)\,n}\Big)\,, \qquad n\geq n_\varepsilon\,.$$

Provided of course that the c appearing in the exponential (1.2) can be taken small enough. As explained in [8] this is always possible, but highly *non* trivial to see. For that reason an alternative method towards estimating expressions as above was given in [7] and no assumption as to the smallness of $c>0$ was needed in that alternative method. The same thing applies here, but to avoid diverting the argument, we shall not give the details. With the help of [7] (or/and Section 1.4 further down) the reader can work this out for himself if he so wishes.

1.2. The inductive step.

Let G, N be as before (simply connected) and let:

$$H=Z(N)=N_k\subset N_{k-1}\subset\cdots\subset N_0=N$$

a central series of N so that H is the center of N. The upper estimates will be proved by induction on k. The case $k=0$ was dealt with in Section 1.1. What we shall do here is to prove the inductive step and show that if the estimate (0.6) holds for the group G/H, where the k is one unit lower, it also hold for G. The issue is clearly to estimate the expectation $\mathbb{E}(\cdots)$ that appears in (1.2). We shall change slightly the notations and write this expectation in the form

$$\mathbb{E}\big(\exp(c\,(|g_1|^2+|g_n|^2)) :\ g_1g_2\cdots g_n\in dr\big)\,, \tag{1.4}$$

where now $g_1,\ldots,g_j,\cdots\in G/H$ are independent, equidistributed, G/H valued, random variables with distribution $\mu_1^{G/H}(\,\cdot\,)$ the heat diffusion convolution measure on G/H at time $=1$ (*cf.* Section 0.2) (the $g_1g_2\cdots g_n$ in the expectation is of course a group product) using this formulation and the fact that $\mu_t(g) = \mu_t(g^{-1})$ it is clear that (1.4) is equal to

$$\iint_{\zeta,\xi\in G/H} \mathbb{E}\big(\exp(c(|\xi|^2+|\zeta|^2)) : g_2\cdots g_{n-1}\in \xi\, dr\, \zeta\big) \cdot \mathbf{P}(g_1\in d\xi)\,\mathbf{P}(g_n\in d\zeta)\,. \tag{1.5}$$

If we use the inductive hypothesis which ensures that the estimate (0.6) holds for the group G/H we see that we can estimate the expectation inside the integral as follows

$$\mathbb{E}(\cdots) \le C_\varepsilon \exp\Big(c(|\xi|^2_{G/H}+|\zeta|^2_{G/H}) - \sum_j \tilde{L}_j^+(\xi\, r\,\zeta) - \frac{|\pi(\xi\, r\,\zeta)|_V^2}{(4+\varepsilon)\,n}\Big)\, m(\xi)\, d^r r\,, \tag{1.6}$$

where $\pi : G/H \longrightarrow V$ is the canonical projection (as in Section 1.1) and $\tilde{L}_j$ are the real parts of the roots of G/H. What is important is to note that each $\tilde{L}_j$ factors through π and is additive. We can thus absorb the $\tilde{L}_j^+(\xi)$, $\tilde{L}_j^+(\zeta)$ with the $c(|\xi|^2+|\zeta|^2)$ and therefore, if we bear in mind that $|g|_{G/H}\ge C\,|\pi(g)|_V$ we can estimate the coefficient of $d^r r$ in (1.6) by

$$\exp\Big(-\sum \tilde{L}_j^+(x) + c(|\xi|^2+|\zeta|^2) - \frac{|x-\pi(\xi)-\pi(\zeta)|_V^2}{(4+\varepsilon)\,n}\Big)\,,$$

where $x=\pi(r)$. Peter-Paul is used as in (1.3) and it yields the estimate

$$\exp\Big(-\frac{|x|_V^2}{(4+\varepsilon)\,n} - \sum \tilde{L}_j^+(x)\Big)\,\exp\Big(c(|\xi|^2_{G/H}+|\zeta|^2_{G/H})\Big)\,,$$

provided that n is large enough (depending on ε). The c in (1.6) can again be chosen in advance and as small as we like. We shall insert this estimate in (1.5) and use the Gaussian decay of $\mathbf{P}(g_i\in d\xi)$ on G/H (and not just on V), *cf.* [11]. If we use the estimate in (1.2) and bear in mind that the real parts of the roots $L_1,\dots,L_p$ of the action of G/H on H (*cf.* Section 4.1) together with the root $\tilde{L}_1,\tilde{L}_2,\dots$ make up for all the roots of G we see that the inductive step follows. A slightly more subtle computation gives as before the estimate $\exp(-|x|^2(1/n-c/n^2)/4)$ for the Gaussian contribution (instead of $\exp(-|x|^2/(4+\varepsilon)n)$).

1.3. An alternative approach.

In this section I shall explain how we can simplify considerably the proof of the upper estimate if we are prepared to loose a little at the end result.

Let $X \subset G$ be some compact subset and let

$$Y = \alpha(X) \subset G/H, \qquad Z = \beta(X) \subset V,$$

where β is the composed mapping $\beta : G \longrightarrow G/N = V \times K \longrightarrow V$. Then (*cf.* (1.2)) there exists $C = C_X$ such that

$$\mathbf{P}(z(n) \in X) \le C \sup_{x \in Z} \exp(-\sum L_i^+(x))\, T \tag{1.7}$$

$$T = T(c) = \mathbb{E}\big(\exp(c\,|g_1|_G^2 + c|g_n|_G^2);\ \alpha(z(n)) \in Y\big), \tag{1.8}$$

where the expectation in (1.8) refers to diffusion on G and where we have replaced $|\alpha(g)|_{G/H}$ by $|g|_G$ which is larger. The c's appearing in the exponential of (1.8) can again be assumed as small as we like. We can now use Hölder to estimate $T(c)$. Indeed for a given $p > 1$ if c is sufficiently small we have $\mathbb{E}(\exp(p\,c\,(|g_1|_G^2 + |g_n|_G^2)) < +\infty$ by the Gaussian estimate (on G) of the g_i's (*cf.* [7]). It follows that we can replace T in (1.7) by

$$T^* = (\mathbf{P}(\alpha(z(n) \in Y))^{1/q},$$

where $1/p + 1/q = 1$ are conjugate indices.

The estimate (1.7) becomes thus an inductive step that reduced the estimate of $\mathbf{P}_G(z(n) \in X)$ to the estimate of $\mathbf{P}_{G/H}(\alpha(z(n)) \in \alpha(X))$ with an arbitrary small loss at the exponent: $1/q$. We shall then set X = "dg" = "an appropiate small element", as in Section 1.1 and we shall examine the dependence of C_X on X. The details of this computation are not trivial but they will not be given here; they can be found in [17].

The end result of the above induction is our upper estimate in Section 0.5 with the L_i's replaced by $(1-\varepsilon)\,L_i$'s for an arbitrary $\varepsilon > 0$. With this approach however we obtain the estimate for $t \ge 1$ and *not* only for $t \ge t_0(\varepsilon)$. For all practical purposes this estimate is as good as the one we have in Section 0.5.

1.4. A general overview of upper estimates.

First of all by the definition (*cf.* Section 1.1) we have

$$A_n(L_i) \leq B_n \cdot D_n \cdot C_n(L_i), \qquad i = 1, 2, \ldots, p. \tag{1.9}$$

where

$$C_n(L_i) = \inf_{1 \leq j \leq n} \exp(-L_i(b_j)),$$
$$B_n = \exp\big(\sup_{1 \leq r \leq n} c\big(\inf_{kr \leq j < (k+1)r} |g_j|_G^2 \big)\big),$$
$$D_n = \exp\big(C\, r \sup_{1 \leq j \leq n} |g_j|_G\big).$$

A few comments are in order: the c that appears in the definition of B_n is the same as the one in the definition of $A_n(L_i)$ but in the estimates that follow I shall not assume that it can be chosen as small as we like. The r in the definition of B is a new parameter and will be chosen later. To see (1.9) one simply samples the j in the infimum over $1 \leq j \leq n$ on the succesive blocks $[kr, (k+1)r)$ so as to pick up the inf $|g_j|_G^2$ in that block. Then one makes the appropriate correction bearing in mind that $|b_j - b_{j-1}|_V \leq C\,|g_j|_G$.

Let X, Y, Z and the other notations be as in Section 1.3 and let us condition with respect to the projected path on V

$$\underline{\beta} = \{\beta(z(j)) = b_j \in V : \ 1 \leq j \leq n\} \subset V^n.$$

For the corresponding conditional expectations, going back to [8, Section 3] and by the same reasoning as in sections 1.1 and 1.3, we obtain

$$\begin{aligned} \mathbf{P}(z_G(n) \in X \,|\, \underline{\beta}) \leq C_X\, \mathbb{E}\big(B_n^C D_n^C;\, z_{G/H}(n) \in Y \,|\, \underline{\beta}\big) \\ \cdot \prod_i C_n(L_i)\, I(b(n) \in Z), \end{aligned} \tag{1.10}$$

where the product $\prod_i$ in the right hand side is taken as before over the roots of the action of G/H on H (as in Section 1.1). This product factors out since it only depends on $\underline{\beta}$. $I(\cdots)$ is the indicator function.

Using Hölder we can estimate the conditional expectation $\mathbb{E}(\,\cdot\,|\,\underline{\beta})$ by

$$\mathbb{E}(B_n^{Cp} \,|\, \underline{\beta})^{1/p}\, \mathbb{E}(D_n^{Cp} \,|\, \beta)^{1/p}\, (\mathbf{P}(z_{G/H}(n) \in Y \,|\, \beta))^{1/q},$$

where $2/p+1/q=1$ and where clearly the q can be made as close to 1 as we like provided that p is large enough. It follows that (1.10) can be used as a recurrence formula that allows us to estimate $\mathrm{P}(z_G(n)\in X\,|\,\underline{\beta})$ in terms of $\mathrm{P}(z_{G/H}(n)\in\alpha(X)\,|\,\underline{\beta})$ with a small error $(1/q<1, 1/q\sim 1)$ in the exponent.

If we repeat this procedure we finally see that we can estimate the left hand side of (1.10) by

$$I(b(n)\in Z)\prod_i C_n((1-\varepsilon_i)\,L_i)\Big(\prod_\alpha \mathbb{E}(B_n^{Cp_\alpha}\,|\,\underline{\beta})^{1/p_\alpha}\,\mathbb{E}(D_n^{Cp_\alpha}\,|\,\underline{\beta})^{1/p_\alpha}\Big),$$

where the $0\le\varepsilon_i\ll 1$ can be made arbitrarily small and the p_α's are appropriately large and where all the roots of G are now involved in the first product. We shall integrate the above estimate over the path space and use Hölder once more.

It follows that we can estimate

$$\text{(1.11)}\quad \mathrm{P}(z_G(n)\in X)\le C_X\,\theta\,\mathbb{E}\prod_i\big(C_n((1-\eta_i)\,L_i)\,I(b(n)\in Z)\big)^{1-\delta_i},$$

where $0\le\eta_i,\delta_i\ll 1$ can be made arbitrarily small and where the cofactor θ is some product of $(\mathbb{E}(B_n^\alpha))^\beta$ and $(\mathbb{E}(D_n^\gamma))^\delta$ for various values of α, β, γ, $\delta>0$. The cofactor θ admits a polynomial bound. Indeed we have:

$$\text{(1.12)}\qquad \mathbb{E}(B_n^C),\ \mathbb{E}(D_n^C)=0(n^A).$$

This is clear for D_n by the (non trivial *cf.* [11]) Gaussian estimate that we have for the g_j's on G

$$\text{(1.13)}\qquad \mathrm{P}(|g_j|_G>\lambda)\le C\exp(-c_0\,\lambda^2)$$

for some fixed but positive $c_0>0$. This implies

$$\mathrm{P}(\sup_{1\le j\le n}|g_j|>\lambda)\le Cn\exp(-C_0\,\lambda^2).$$

The estimate (1.13) gives also

$$\mathrm{P}\big(\inf_{kr\le j<(k+1)r}|g_j|_G^2=\zeta_k>\lambda\big)\le C^r\exp(-c_0\,r\,\lambda),\quad r=1,2,\dots,$$

and therefore:

$$\mathbf{P}\big(\sup_{1\le k\le n}\xi_k \ge \lambda\big) \le n\,C^r \exp(-c_0\, r\,\lambda)\,.$$

If r is large enough we can clearly "absorb" any exponent M in $E(B_n^M)$ and (1.12) follows for B_n.

The final "move" in this general approach is to examine the principal term $\prod_i(\cdots)I[\cdots]$ in the right hand side of (1.11), which is a Brownian functional and which can therefore be estimated by purely probabilistic (Brownian or random-walks) methods. Two cases have been examined in details

1) X is some neighbourhood of 0 in G and the linear functionals satisfy the C-condition. This was done in [8] and the estimate obtained there is $O(e^{-cn^{1/3}})$

2) Z is some neighbourhood of $x \in V$, some point "far out" on V. This is what was done in this paper.

It is of course possible to incorporate the C-condition in the case 2) and obtain an extra factor $e^{-ct^{1/3}}$ in front of the upper estimates of Section 0. The details will be left to the interested reader.

2. Dilation structure on a Lie algebra.

2.1. Algebraic considerations.

Let $\Lambda^t \in GL(V)$ $(-\infty < t < +\infty)$ be a one parameter group of automorphisms of the real vector space $V(\cong \mathbb{R}^a,\ a \ge 1)$. Let $V_{\mathbb{C}} = V \otimes_{\mathbb{R}} \mathbb{C}$ be the complexified space and let

$$V_{\mathbb{C}} = V_1^{\mathbb{C}} \oplus \cdots \oplus V_k^{\mathbb{C}}$$

the corresponding root space decomposition of $V_{\mathbb{C}}$ under Λ^t. In other words if $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ are the corresponding roots we have $x \in V_j$ $(1 \le j \le k)$ if and only if

$$(\Lambda^t - e^{\lambda_j t})^N x = 0\,, \qquad t \in \mathbb{R},$$

for some $N \ge 1$ large enough (independently of j).

We shall combine together the subspaces $V_{i_s}^{\mathbb{C}}$, $1 \le s \le r$, for which the corresponding roots have the same real part $\operatorname{Re}\lambda_{j_s} = \gamma$, $1 \le s \le r$, and write

$$(2.1) \qquad V = V_1 \oplus \cdots \oplus V_p\,, \qquad ; V_j \otimes \mathbb{C} = V_{i_1}^{\mathbb{C}} \oplus \cdots \oplus V_{i_r}^{\mathbb{C}}\,.$$

The operator norms of Λ^t restricted to V_s satisfy

$$\|\Lambda^t|_{V_s}\| = O(e^{\gamma t}|t|^A)\,, \qquad \text{as } t \to \pm\infty\,.$$

This is simply because

$$\Lambda^t|_{V_s \otimes \mathbb{C}} = e^{\gamma t}\, T(t)\,,$$

where $T(t)$ is an upper triangular complex matrix with unimodular diagonal coefficients and therefore satisfies (*cf.* [7])

$$T(t) = T^{-1}(-t)\,, \qquad |T(t)| = O(|t|^A)\,.$$

Let us now suppose that $V = \mathfrak{g}$ is some real Lie algebra and that Λ^t are algebra automorphisms. We have then $\Lambda^t = \exp(tD)$ where $D \in \partial(\mathfrak{g})$ is some derivation of $\mathfrak{g}$. Let

$$\begin{gathered}\mathfrak{g} = \mathfrak{g}_{-\beta_1} \oplus \cdots \oplus \mathfrak{g}_{-\beta_s} \oplus \mathfrak{g}_0 \oplus \mathfrak{g}_{\alpha_1} \oplus \cdots \oplus \mathfrak{g}_{\alpha_r}\,,\\ -\beta_1 < \cdots < -\beta_s < 0 < \alpha_1 < \cdots < \alpha_r\,,\end{gathered}$$

be the corresponding decomposition as in (2.1) with

$$\gamma = \operatorname{Re}\lambda_{i_s} = -\beta_1, \ldots, 0, \alpha_1, \ldots$$

It is easy to see (*cf.* [6]) that we have $[\mathfrak{g}_{\gamma_1}, \mathfrak{g}_{\gamma_2}] \subset \mathfrak{g}_{\gamma_1+\gamma_2}$ if $\gamma_1 + \gamma_2$ is a real part of a root and $[\mathfrak{g}_{\gamma_1}, \mathfrak{g}_{\gamma_2}] = 0$ if not. It follows in particular that

$$\begin{aligned}\mathfrak{g}_+ &= \mathfrak{g}_{\alpha_1} \oplus \cdots \oplus \mathfrak{g}_{\alpha_r}\,, & \mathfrak{g}_+^0 &= \mathfrak{g}_+ \oplus \mathfrak{g}_0\,,\\ \mathfrak{g}_- &= \mathfrak{g}_{-\beta_1} \oplus \cdots \oplus \mathfrak{g}_{-\beta_s}\,, & \mathfrak{g}_-^0 &= \mathfrak{g}_- \oplus \mathfrak{g}_0\,,\end{aligned}$$

are subalgebras. Let $|\cdot|$ be some norm on $\mathfrak{g}$, it is then clear from the above that we have

$$(2.2) \qquad \begin{aligned}|\Lambda^t x| &= O(|t|^A)\,, \quad \text{as } t \to +\infty \quad &\Longleftrightarrow \quad x \in \mathfrak{g}_-^0\,,\\ |\Lambda^t x| &= O(|t|^A)\,, \quad \text{as } t \to -\infty \quad &\Longleftrightarrow \quad x \in \mathfrak{g}_+^0\,.\end{aligned}$$

From this it follows in particular that if $\mathfrak{h} \subset \mathfrak{g}$ is a subalgebra, or even a subspace, stable by Λ (*i.e.*, $\Lambda\mathfrak{h} \subset \mathfrak{h}$) then

$$\mathfrak{g}_{+}^{0} \cap \mathfrak{h} \subset \mathfrak{h}_{+}^{0}, \qquad \mathfrak{g}_{-}^{0} \cap \mathfrak{h} \subset \mathfrak{h}_{-}^{0}. \tag{2.3}$$

We shall now describe a rather technical construction that will be essential in what follows.

Proposition. *Let us assume that* $\mathfrak{g}$ *is nilpotent and let* $\mathfrak{h} \subset \mathfrak{g}$ *be a subalgebra that satisfies*

$$\mathfrak{g}_{+} \subset \mathfrak{h} \neq \mathfrak{g}, \ (\textit{respectively}, \ \mathfrak{g}_{+}^{0} \subset \mathfrak{h} \neq \mathfrak{g}), \qquad \Lambda^{t}\mathfrak{h} \subset \mathfrak{h}.$$

Then we can find $y, z \in \mathfrak{g}_{-}^{0}$ *(respectively,* $\mathfrak{g}_{-}$*) such that the subalgebra*

$$\mathfrak{h}_1 = \operatorname{Alg}\{\mathfrak{h}, y, z\}$$

is stable by the action of Λ^{t} *(i.e.,* $\Lambda^{t}\mathfrak{h}_1 \subset \mathfrak{h}_1$*) and furthermore*

$$\mathfrak{h} \neq \mathfrak{h}_1, \qquad [\mathfrak{h}, \mathfrak{h}_1] \subset \mathfrak{h}.$$

PROOF. Let us denote

$$\tilde{\mathfrak{h}} = (\mathfrak{h} \otimes \mathbb{C}) \land \mathbb{C}D \subset \tilde{\mathfrak{g}} = (\mathfrak{g} \otimes \mathbb{C}) \land \mathbb{C}D,$$

where the skew product is defined by the action of the derivation $D \in \partial(\mathfrak{g})$ on $\mathfrak{g}$. The algebra $\tilde{\mathfrak{g}}$ is a complex soluble algebra and let $V = \tilde{\mathfrak{g}}/\tilde{\mathfrak{h}}$ which is a complex vector space on which $\tilde{\mathfrak{h}}$ acts by adjoint action. By Lie's theorem there exists therefore

$$0 \neq X = y + iz + \tilde{\mathfrak{h}} \in V, \qquad y, z \in \mathfrak{g} \ \ (y + iz \notin \tilde{\mathfrak{h}}),$$

such that

$$\xi(X) = \lambda(\xi) \cdot X, \qquad \xi \in \tilde{\mathfrak{h}}, \qquad \lambda(\,\cdot\,) \in \operatorname{Hom}_{\mathbb{C}}(\tilde{\mathfrak{h}}; \mathbb{C}). \tag{2.4}$$

The fact that $\mathfrak{g}$ is nilpotent implies that every $\zeta \in \mathfrak{h}$ gives rise to a nilpotent transformation on V and therefore that

$$\lambda(\zeta) = 0, \qquad \zeta \in \mathfrak{h}. \tag{2.5}$$

The fact that $\mathfrak{h} \supset \mathfrak{g}_+$ (respectively, $\supset \mathfrak{g}_+^0$) implies furthermore that we can assume that $y, z \in \mathfrak{g}_-^0$ (respectively, $\mathfrak{g}_-$). Now (2.5) implies that

$$[X, \mathfrak{h}] = [y, \mathfrak{h}] + i\,[z, \mathfrak{h}] \subset \tilde{\mathfrak{h}}\,.$$

From this it follows that

$$[y, \mathfrak{h}]\,, [z, \mathfrak{h}] \subset \mathfrak{h}\,.$$

What we have shown is that the subalgebra $\mathfrak{h}_1$ generated by $\{\mathfrak{h}, y, z\}$ is strictly larger than $\mathfrak{h}$ and normalises $\mathfrak{h}$. By (2.4) it follows that

$$[D, y] + i\,[D, z] \in \lambda(D)(y + iz) + \tilde{\mathfrak{h}}\,.$$

Let $\lambda(D) = \alpha + i\beta$ $(\alpha, \beta \in \mathbb{R})$ we then have

$$[D, y] \in \alpha\, y - \beta\, z + \mathfrak{h}\,, \qquad [D, z] \in \beta\, y + \alpha\, z + \mathfrak{h}\,.$$

This shows that $[D, \mathfrak{h}_1] \subset \mathfrak{h}_1$ and therefore that $\Lambda^t \mathfrak{h}_1 \subset \mathfrak{h}_1$. The above algebra $\mathfrak{h}_1$ satisfies therefore all the conditions of our proposition.

From the above proposition it follows that we can construct

$$\mathfrak{g}_+^0 = \mathfrak{h}_0 \subset \mathfrak{h}_1 \subset \cdots \subset \mathfrak{h}_k = \mathfrak{g}\,, \tag{2.6}$$

a finite chain of subalgebras such that

$$\mathfrak{h}_{j+1} = \mathrm{Alg}\,\{\mathfrak{h}_j, y_j, z_j\}\,, \qquad j = 0, 1, \ldots, k-1\,,$$

with $y_j, z_j \in \mathfrak{g}_-$ and so that $\mathfrak{h}_{j+1}$ normalises $\mathfrak{h}_j$ $(j = 0, 1, \ldots, k-1)$ and

$$\Lambda^t \mathfrak{h}_j \subset \mathfrak{h}_j\,, \qquad j = 0, 1, 2, \ldots, k\,.$$

We shall end up this section by recalling an important lemma due to Hörmander (*cf.* [2], [11]).

Lemma (L. Hörmander). *Let $\mathfrak{g}$ be a nilpotent Lie algebra and let G be some Lie group that corresponds to $\mathfrak{g}$. Then there exists $N \geq 0$, $c_\alpha \in \mathbb{R}$ $(0 \leq \alpha \leq N)$, $k_\alpha = 0, 1, \ldots$ $(1 \leq j \leq N)$ such that*

$$\mathrm{Exp}\,(x + y) = \mathrm{Exp}\,(c_0\, y) \prod_{\alpha=1}^{N} \mathrm{Exp}\,(c_\alpha\, x_\alpha) \in G\,, \qquad x, y \in \mathfrak{g}\,,$$

where

$$x_\alpha = (ad(y))^{k_\alpha} x = [y, [y[\cdots[y, x]\cdots]\,, \qquad 1 \leq \alpha \leq N\,.$$

2.2. Positive and negative roots.

All the notations of the previous section will be preserved. $\mathfrak{g}$ will be assumed throughout to be nilpotent and we shall denote by $|\cdot|$ some norm on $\mathfrak{g}$ (the exact value of $|\cdot|$ will be irrelevant here). We shall denote by

$$A(a) = \{x \in \mathfrak{g} : \ |x| \leq a\}.$$

We shall say that $\mathfrak{g}$ is a positive (respectively, negative) algebra if $\mathfrak{g} = \mathfrak{g}^0_+$ (respectively, $\mathfrak{g} = \mathfrak{g}^0_-$).

We shall also consider $\mathrm{Exp} : \mathfrak{g} \longrightarrow G$ the exponential mapping where G is the simply connected nilpotent group associated to $\mathfrak{g}$. The Haar measure of G will be denoted by m_G (not to be confused with the previous notation for the modular function that here is identically 1) which is the image by Exp of Lebesgue measure on $\mathfrak{g}$. The Jacobian of $\Lambda^t : \mathfrak{g} \longrightarrow \mathfrak{g}$ is

$$\mathrm{Jacb}\,(\Lambda^t) = \exp\left(\left(\sum \alpha_j - \sum \beta_j\right) t\right),$$

where here and in what follows we count the α_j's and the β_j's with multiplicity (*i.e.*, they are tacitly multiplied by the dimensionality of the corresponding root spaces $\mathfrak{g}_{\alpha_j}, \mathfrak{g}_{-\beta_j}$).

The group Λ^t induces a one parameter group, also denoted by $\Lambda^t : G \longrightarrow G$, of group automorphisms and Exp interwines Λ^t. The norm $|\cdot|$ on $\mathfrak{g}$ induces a left (or right) invariant distance on G (*cf.* [11]), where for $g \in G$ we set

$$d(e,g) = \inf\left\{ \sum |t_j| : \ g = \mathrm{Exp}\,(t_1 x_1)\,\mathrm{Exp}\,(t_2 x_2)\cdots, \quad x_j \in \mathfrak{g},\ |x_j| \leq 1 \right\}$$

and clearly

$$d(\Lambda x, \Lambda y) \leq |\Lambda|\, d(x,y), \qquad x,y \in G,$$

where $|\Lambda|$ denotes the operator norm of $\Lambda : \mathfrak{g} \longrightarrow \mathfrak{g}$.

Observe finally that when $\Lambda = Ad(g)$, $g \in G$ on $\mathfrak{g}$, then $\Lambda = I_g : h \mapsto ghg^{-1} = h^g$ on G. It follows therefore that for any normal subgroup $H \subset G$ we have

$$d_H(x^g, y^g) \leq |Ad(g)|\, d_H(x,y), \qquad g \in G,\ x,y \in H,$$

where $d_H(\cdot\,,\cdot)$ denotes the "intrinsic " distance on H (and not the distance induced by the ambient G).

Observe also that since the algebra $\mathfrak{g}$ is nilpotent the operator norm of $Ad(g)$ satisfies

$$|Ad(g)| \leq C\,(|g|+1)^C .$$

I shall denote

$$\begin{aligned} B(a) &= \{g \in G : \ d(e,g) \leq a\} \\ D(a) &= D_0(a) = \mathrm{Exp}\,(A(a)) , \qquad a > 0 . \end{aligned} \tag{2.8}$$

We have then

$$D(a) \subset B(a) , \qquad B(a) \subset D(a^C) , \qquad a > 1 , \tag{2.9}$$

for some $C > 0$ where the second inclusion followes from the Baker-Campbell-Hausdorff formula (*cf.* [1, Section 2.15]). I shall denote

$$D_t(a) = \mathrm{Exp}\,(\Lambda^t A(a)) = \Lambda^t \mathrm{Exp}\,(A(a)) = \Lambda^t D_0(a) , \tag{2.10}$$

for $t \in \mathbb{R}$, $a > 0$. From the above we have

$$m_G(D_t(a)) = C \exp\Big(\Big(\sum \alpha_j - \sum \beta_j\Big)t\Big)\, a^{\dim \mathfrak{g}} . \tag{2.11}$$

Let us first suppose that $\mathfrak{g}$ is a *negative* algebra. We have then

$$D_{t_1}(a_1)\cdots D_{t_n}(a_n) \subset D_0\left(C\Big(\sum a_j + \sum t_j + n\Big)^C\right) , \tag{2.12}$$

for $a_j \geq 0$ and $t_j \geq 0$ and where $C > 0$ only depends on Λ^t (the left hand side of (2.12) is, of course, a group product in G). Two special cases of the inclusion (2.12) are easy to prove. The case $t_1 = t_2 = \cdots = t_n = 0$ is a consequence of the Baker-Campbell-Hausdorff formula. The case $n = 1$ is a consequence of (2.2). The inclusion (2.12) in general follows immediately from the above two special cases.

Let us now suppose that $\mathfrak{g}$ is a *positive* algebra, let $t_1, \ldots, t_n \in \mathbb{R}$ and let $t = \max_{1 \leq j \leq n} t_j$. Then

$$\begin{aligned} D_{t_1}(a_1)\cdots D_{t_n}(a_n) &= \Lambda^t\big(D_{t_1-t}(a_1)\cdots D_{t_n-t}(a_n)\big) \\ &\subset D_t\left(C\Big(\sum a_j + \sum |t_j| + n\Big)^C\right) , \end{aligned} \tag{2.13}$$

where the second inclusion follows from (2.12) and the fact that $t_j - t \leq 0$ $(1 \leq j \leq n)$. Indeed $\Lambda^{t_j - t} = \tilde{\Lambda}^{t - t_j}$ and $\mathfrak{g}$ is a negative algebra for the dilation Λ^{-t}.

Together with the above observations we shall need the following

Lemma. *Let us assume that* $\mathfrak{g}$ *is a positive algebra, then for each* $t, s > 0$ *there exist*

$$z_\nu, \tilde{z}_\nu \in G\,, \qquad 1 \leq \nu \leq M \leq C\,(s+t+1)^C \exp\Big(t \sum \alpha_j\Big)$$

that satisfy

$$D_t(s) \subset \Big(\bigcup_\nu z_\nu\, D_0(1)\Big) \bigcap \Big(\bigcup_\nu D_0(1)\, \tilde{z}_\nu\Big) = R(M)\,. \tag{2.14}$$

This lemma will be combined with (2.13) to yield

$$\begin{aligned} D_{t_1}(a_1) \cdots D_{t_n}(a_n) \subset R\Big(C\Big(\sum a_j + \sum |t_j| + n\Big)^C\Big) \\ \cdot \exp\Big(t \sum \alpha_j\Big)\,. \end{aligned} \tag{2.15}$$

where we suppose that $t = \max(t_j) \geq 0$.

PROOF. Let us fix $d(\,\cdot\,,\,\cdot\,)$ some left invariant Riemannian distance on G and let z_ν, $1 \leq \nu \leq M$, be some maximal ε-net in $D_t(s)$ (for $\varepsilon > 0$ appropriately small but fixed), *i.e.*, we choose

$$z_\nu \in D_t(s)\,,\ (1 \leq \nu \leq M)\,, \qquad d(z_\nu, z_\mu) \geq \varepsilon\,,\ \nu \neq \mu\,,$$

and the set $z_1, \ldots, z_M$ is maximal under the above two conditions. It is clear that (2.14) is then verified. What remains is to give a bound of M. To achieve this it suffices to observe that

$$\bigcup_\nu z_\nu\, D_0(\varepsilon_1) \subset D_t(s) \cdot D_0(\varepsilon_1) \subset D_t\big(C\,(s+t+1)^C\big)\,, \tag{2.16}$$

where the union in (2.16) is disjoined if $\varepsilon_1 > 0$ is appropriately small. The volume estimate (2.11) of the right hand side of (2.16) (we have $\beta_j = 0$ now) gives then the required bound for M.

2.3. General nilpotent algebras.

All our previous notations will be preserved but here $\mathfrak{g}$ will not be assumed to be either a positive or a negative algebra. What will be proved in this section is that for all $s > 0$, $N = 1, 2, \dots$, we can find

$$(2.17) \qquad z_\nu \in G\,, \quad \nu = 1, 2, \dots M \leq C(s + N + 1)^C \exp\left(N \sum \alpha_j\right)$$

such that

$$(2.18) \qquad D_0(s) \cdots D_N(s) \subset \bigcup_\nu D_0(1)\, z_\nu\,.$$

The consequence that we shall draw from this is that

$$(2.19) \qquad m_G(D_0(s) \cdots D_N(s)) \leq C(s + N + 1)^C \; exp\left(N \sum \alpha_j\right).$$

At this point the reader is *strongly* encouraged to give for himself a proof of (2.19) when $G = \mathbb{R}^d$. A direct proof of (2.19) is highly not trivial even for the Heisenberg group with the dilation structure $X \longrightarrow \lambda X$, $Y \longrightarrow \lambda^{-1} Y$, $Z \longrightarrow Z$. ($X, Y, Z = [X, Y]$ is here the standard basis of the corresponding algebra).

We shall consider the chain of subalgebras (2.6) and we shall prove (2.18) by induction on the k of (2.6). For $k = 0$ (2.18) is contained in the Lemma of Section 2.2.

We shall use of the notations of sections 2.1 and 2.2. The norm in $\mathfrak{g}$ will be chosen so that

$$A_j(s) = \{x \in \mathfrak{h}_j : \; |x| \leq s\}$$

satisfy

$$(2.20) \qquad A_{j+1}(s) = A_j(s) + \theta\, y_j + \theta\, z_j\,.$$

The above notation is abusive: the two θ's are *not* identical. They are θ_1 and θ_2 and the right hand side of (2.20) is to be understood as the union over all the $|\theta_i| \leq s$, $i = 1, 2$. This kind of abusive but convenient notation will be adopted throughout the rest of the proof. We shall also denote by H_j the subgroup that corresponds to $\mathfrak{h}_j$. The inductive hypothesis is that for some $j \geq 1$ in the algebra $\mathfrak{h}_j$ we have

$$(L_j) \qquad \mathrm{Exp}\,(A_j(s))\, \mathrm{Exp}\,(\Lambda A_j(s)) \cdots \mathrm{Exp}\,(\Lambda^N A_j(s)) \subset \bigcup_\nu D_0(1)\, z_\nu\,.$$

$D_0(1) \subset H_j$ of course here corresponds to $\mathfrak{h}_j$. Using (L_j) we shall procede to prove (L_{j+1}) in the algebra $\mathfrak{h}_{j+1}$.

The first step is to apply Hörmander's lemma (*cf.* Section 2.1) on each factor

$$\begin{aligned}\operatorname{Exp}(\Lambda^m A_{j+1}(s)) &= \operatorname{Exp}(\Lambda^m A_j(s) + \theta\,\Lambda^m y_j + \theta\,\Lambda^m z_j) \\ &= \operatorname{Exp}(e_m) \prod_\alpha \operatorname{Exp}(\Lambda^m Z_\alpha) \\ &= \operatorname{Exp}(e_m)\, M_m = E_m\, M_m\,,\end{aligned}$$

where

$$e_m = c\,\theta\,\Lambda^m y_j + c\,\theta\,\Lambda^m z_j\,, \qquad Z_\alpha = \left(c_\alpha\,\theta\,ad(y_j) + c_\alpha\,\theta\,ad(z_j)\right)^{k_\alpha} A_j(s)\,.$$

Here we make the same abuse of notation over the θ's, and to simplify notations we have droped the j's. The left hand side of (L_{j+1}) is therefore

$$\prod_{m=1}^{N} E_m\, M_m\,. \tag{2.21}$$

The next step is to "commute backwards" all the E_k's through the M_p's that precede it ($i.e., p < k$) so as to put (2.21) in the form

$$E_1\, E_2 \cdots E_N\, \tilde{M}_1 \cdots \tilde{M}_N\,,$$

where

$$\begin{aligned}\tilde{M}_p &\subset Ad(E_N) \cdots Ad(E_{p+1})\,\Lambda^p B_j(c(s+1)) \\ &= \Lambda^p(Ad(\Lambda^{-p}E_N\,\Lambda^{-p}E_{N-1} \cdots \Lambda^{-p}E_{p+1})B_j(c(s+1)))\,,\end{aligned}$$

$$B_j(s) = \{h \in H_j :\ d_j(e,h) = |h|_j \le s\}\,,$$

and where d_j is the distance on H_j.

The fact that $y_j, z_j \in \mathfrak{g}_-$ implies (*cf.* (2.2), (2.3)) that $y_j, z_j \in (\mathfrak{h}_{j+1})^0_-$ and that $|e_m|_{j+1} \le C\,(s+N)^C$ and, more generally, that

$$|\Lambda^{-p} e_k|_{j+1} \le C\,(s+N)^C\,, \qquad k \ge p\,,$$

$$|\Lambda^{-p}E_N \cdots \Lambda^{-p}E_{p+1}|_{j+1} \le C\,(s+N)^C\,, \qquad p \ge 0\,.$$

We obtain therefore that

$$E_1 \cdot E_2 \cdots E_N \subset \operatorname{Exp}(A_{j+1}(C\,(s+N)^C)),$$

$$\tilde{M}_p \subset \Lambda^p B_j(C\,(s+N)^C), \tag{2.22}$$

$$\tilde{M}_1 \cdot \tilde{M}_2 \cdots \tilde{M}_N \subset \operatorname{Exp}(A_j(s')) \cdots \operatorname{Exp}(\Lambda^N A_j(s')),$$

with $s' = C\,(s+N)^C$. If we use (L_j) on right hand side of (2.22) we conclude therefore that the right hand side of of (L_{j+1}) is contained in

$$\bigcup_\nu \operatorname{Exp}(A_{j+1}(C\,(s+N)^C))\, D_0(1)\, z_\nu\,. \tag{2.23}$$

An obvious use of the Baker-Campbell-Hausdorff formula gives that (2.23) is contained in

$$\bigcup_\nu \operatorname{Exp}(A_{j+1}(C\,(s+N)^C))\, z_\nu\,.$$

To complete the inductive step it suffices therefore to observe that for obvious reasons (*cf.*, proof of the Lemma in Section 2.2) we have

$$\operatorname{Exp}(A_{j+1}(C\,(s+N)^C)) \subset \bigcup_{\mu=1}^{C(s+N)^C} D_0(1)\, u_\mu$$

for an appropiate choice of $u_\mu \in H_{j+1}$.

A simple use of the involution $g \longrightarrow g^{-1}$ in G shows that from (2.18) we have the symmetric result

$$D_N(s)\, D_{N-1}(s) \cdots D_0(s) \subset \bigcup_\nu z_\nu^{-1}\, D_0(1)\,. \tag{2.24}$$

It is worth observing also that the left hand side of (2.24) is

$$\Lambda^N(\tilde{D}_0(s) \cdots \tilde{D}_N(s)),$$

where $\tilde{D}_t(s) = \operatorname{Exp}(\tilde{\Lambda}^t A(s))$ with $\tilde{\Lambda}^t = \Lambda^{-t}$. The effect of replacing Λ by $\tilde{\Lambda}$ is to swap the positive roots with the negative ones. Using the above observations one can obtain several variants of (2.18) that are relevant in different contexts.

3. The lower estimate.

3.1. Algebraic groups.

In this section I shall follow very closely [4], including the notations. The proof in [4] simplifies considerably if I make the additional assumption that G (*cf.* Section 0) has the form

$$G = Q\lambda M = N\lambda(V \times M)\,, \qquad Q = N\lambda V\,, \tag{3.1}$$

where $V \cong \mathbb{R}^a$. This assumption is in particular verified for all algebraic groups (or more generally when $\mathfrak{g}$ the Lie algebra is algebraic). Indeed in that case I shall take $\Sigma = V$ for the section constructed in [4] and many of the geometric and algebraic difficulties disappear in one stroke. For the convenience of the reader I shall here first give the proof under the additional assumption (3.1) and then proceed to consider the general case.

All the other notations of [4, Section 3] are preserved:

$$Z_n = \gamma_1\,\gamma_2 \cdots \gamma_n = \dot{Z}_n\,\Lambda_n\,, \qquad \dot{Z}_n \in S\,,\ \Lambda_n \in N\,,$$

is the random walk on G controlled by $d\mu_1 = \phi_1\, d^r g$ (where however in [4] the above product $Z_s = \dot{Z}_s\,\Lambda_s$ was written in the other way round). We shall now fix $x \in V$, $s = 1, 2, \ldots$ and we shall find $A \subset \Omega$ a subset of the path space of this random walk on which the following conditions are verified

$$|\dot{Z}_s - x| \leq 1 \tag{3.2}$$

$$L_k\Big(\dot{Z}_j - \frac{j}{s}\,x\Big) \leq D\,, \quad j = 1, 2, \ldots s\,,\ k = 1, 2, \ldots, n\,, \tag{3.3}$$

$$|\gamma_j| \leq \delta\Big(c\,D + \frac{|x|^2}{4s}\Big)^{1/2} = \delta\,\lambda\,, \quad j = 1, 2, \ldots, s\,. \tag{3.4}$$

The $D \sim s^{1/3}$ and the δ will be chosen later. I will show that we can choose A so that

$$\mathbf{P}(A) \geq c\,\exp\Big(-c\,s^{1/3} - \frac{|x|^2}{4s}\Big)\,. \tag{3.5}$$

Indeed let $\Omega_D \subset \Omega$ be the subset of the path space determined by (3.2) and (3.3) then since M is compact, for $D = c\,s^{1/3}$, we have (*cf.* Appendix):

$$\mathbf{P}(\Omega_D) \geq \exp\Big(-c\,s^{1/3} - \frac{|x|^2}{4s}\Big) = \exp(-\lambda^2)\,.$$

Let $\Omega^\lambda \subset \Omega$ be the subset determined by (3.4) then

$$P(\Omega^\lambda) \geq 1 - s\, e^{-C\delta^2\lambda^2}$$

because the variables $\gamma_j \in G$ satisfy a Gaussian estimate on G (*cf.* [11]). We shall fix $D \sim s^{1/3}$ and δ large enough. We have then

$$P(\Omega_D \cap \Omega^\lambda) \geq C \exp(-\lambda^2)$$

and our assertion follows.

Let us now go back to [4] where we shall assume that $G = N\lambda(V \times M)$ and that $\Sigma = V$, $S = V \times M$. For any subset $E \subset N$ we shall denote $E^g = gEg^{-1}$ $(g \in G)$, we shall further denote B_α the ball in N of radius $\exp(C\alpha)$ centered at the identity. Analysing closely the argument in [4] it follows that on our subset $A \subset \Omega$ (that satisfies (3.2), (3.3), (3.4) and (3.5)) we have

$$Z_s \in (B_{\delta\lambda}^{\dot{Z}_1} B_{\delta\lambda}^{\dot{Z}_2} \cdots B_{\delta\lambda}^{\dot{Z}_s}) \dot{Z}_s\,. \tag{3.6}$$

On the other hand, from (3.3) it is clear that

$$B_{\delta\lambda}^{\dot{Z}_j} \subset (B_{\delta\lambda+D})^{jx/s}\,.$$

It follows therefore that

$$Z_s \in (B_{\delta\lambda+D}^{x/s} \cdots B_{\delta\lambda+D}^{jx/s} \cdots)\, \dot{Z}_s\,, \tag{3.7}$$

where because of (2.19) the m_N-measure of $(\cdots)$ is bounded above by

$$s^c \exp\Big(C(\delta\lambda + D) + \sum L_j^+(x)\Big)\,.$$

The estimates (0.11), (0.12) follow.

The only modification needed to obtain the improvement under the (NC)-condition is that we set $D \sim c \log s$ instead (*cf.* Appendix). The rest of the proof is identical.

3.2. The general case.

A thorough understanding of the geometric construction in [4, Section 1] is essential for this section. The notations here are those of the previous section and of [4] and we shall start the proof exactly as in the previous section. The point where we run into trouble is (3.6), (3.7). Indeed the section Σ is in general not a group and the elements of Σ do not commute between themselves.

The first step towards resolving these difficulties is to choose the generators $e_1, e_2, \ldots, e_m$ of the nilpotent algebra $\mathfrak{a}$ (*cf.* proof [4, Lemma 1.2]) so that our preasigned point $x \in V$ (on which we want to prove the lower estimates) is $x = \pi(\exp(|x|e_1))$, *i.e.*, x lies in the image by $\pi : Q \longrightarrow V$ of the first one parameter subgroup $x_1(t)$ of Σ. With the obvious identification of Σ and V and with the basis $d\pi(e_j)$, $j = 1, 2, \ldots, m$, on V we have then $x = (|x|, 0, 0, \ldots, 0)$.

Let us assume for simplicity that $G = Q$ is soluble, *i.e.*, that $M = \{e\}$. We have then as in [4, (3.5)]

$$\gamma_j = \dot\gamma_j n_j\,, \qquad \dot\gamma_j = \sigma(t_j) = x_1(t_j^{(1)}) \cdots x_m(t_j^{(m)}) \in \Sigma\,,$$

$$t_j = (t_j^{(1)}, \ldots, t_j^{(m)}) \in V\,, \qquad j = 1, 2, \ldots, s\,,$$

and because of (3.2), (3.3), (3.4) we have

$$|t_1 + t_2 + \cdots + t_s - x| \le 1\,,$$

$$L_k(t_1 + t_2 + \cdots + t_j - jx/s) \le D\,,$$

$$|t_j^{(i)}| \le \delta\Big(D + \frac{|x|^2}{4s}\Big) = \delta\,\lambda\,,$$

$$j = 1, 2, \ldots, s\,, \qquad k = 1, 2, \ldots, n\,, \qquad i = 1, 2, \ldots, m\,.$$

The critical step is to prove that

$$\begin{aligned} Z_s &= \dot\gamma_1 n_1 \dot\gamma_2 n_2 \cdots \dot\gamma_s n_s \\ &\in x_1(t_1^{(1)})\,B\,x_1(t_2^{(1)})\,B \cdots \\ &\quad \cdot x_1(t_s^{(1)})\,B\,x_2(t_1^{(2)} + \cdots + t_s^{(2)}) \cdots x_m(t_1^{(m)} + \cdots + t_s^{(m)})\,, \end{aligned} \tag{3.8}$$

where $B \subset N$ is the ball of radius

$$\exp(C\,\delta\,\lambda + c\,D)\,s^C\,|x|^C \le s^C\Big(1 + \frac{|x|^2}{4s}\Big)^C \exp(C\,\delta\,\lambda + C\,D)\,. \tag{3.9}$$

To see this we must shift each $x_2(t_j^{(2)})\cdots x_m(t_j^{(m)})$ (*i.e.*, the co-factors of $x_1(t_j^{(1)})$ in $\dot{\gamma}_j$) through to product. Every time we cross a ball $|n_k|_N \leq \exp(C\,\delta\,\lambda)$ by

$$x_2\big(t_1^{(2)}+\cdots+t_k^{(2)}\big)\,x_3\big(t_1^{(3)}+\cdots+t_k^{(3)}\big)\cdots=\zeta_k\,. \tag{3.10}$$

We multiply the radius of the ball by $\exp(CD)$ hence the exponential term in the left hand side of (3.9). The extra factor $s^C\,|x|^C$ in (3.9) is the bound of the number of commutators of terms of the form $x_j(t)$ ($t\in\mathbb{R}$, $|t|\leq 1$, $j=1,2,\ldots,n$) (*cf.*, Nil-Gp Lemma in [4]). Each of these commutators lie in some fixed ball in N (for the $|\cdot|_N$ distance). The above commutators are placed between the $x_1(t_j^{(1)})$'s ($j=1,2,\ldots,s$) and the B's and they arise because we have to put the factors $x_i(t_j^{(i)})$ in the "right order" to make a term ζ_k as in (3.10).

We are now in a position to finish the proof. Indeed one more set of commutations, with the $x_1(t_j^{(1)})$'s this time, brings the right hand side of (3.8) in the form

$$\big(B^{x/s}_{\delta\lambda+D}\cdots B^{jx/s}_{\delta\lambda+D}\cdots\big)\,\sigma(t_1+\cdots+t_s)\,.$$

This expression is essentially the same as (3.7) and the proof finishes as before.

Finally since the elements of Σ commute with M if we do not assume that $M=\{e\}$ nothing changes in the above argument (*cf.* [4, Section 1]).

4. The Hardy-Littlewood theory.

Theorem 1 is an obvious integration of the estimates (0.9), (0.13) on V which also proves the first estimate (4.1) below. The reader should observe that the change of variable $x\mapsto x^{-1}$ in G shifts left to the right measure and therefore the "left" and "right" $\|\cdot\|$-normes of ψ_t are the same. What is important in Theorem 1 is to understand the exact role played by Δ on G and by the induced scalar product $\langle\,\cdot\,,\,\cdot\,\rangle_\Delta$ on V (*cf.* Section 0.1).

The expression of $\ell(q)$ given by theorem does of course depend on that scalar product and therefore on the particular choice of the Laplacian Δ on G. One of the consequences of Theorem 2 is that L is a *linear invariant*, *i.e.*, that it is independent of the particular scalar product used in (0.2). This is very easy to see. Indeed let

$$L(p;x)=\Big(\frac{1}{p}-\frac{1}{2}\Big)\sum L_j^+(x)+\frac{1}{2}\sum L_j^-(x)\,,\quad 1\le p\le+\infty\,,\ x\in V\,.$$

We clearly have

$$L(1;x)=\frac{1}{2}\sum L_j(x)\,,\qquad L(2;x)=\frac{1}{2}\sum L_j^-(x)\,.$$

We also have

$$L=\inf\{1\le p\le+\infty:\ L(p;x)\le 0 \text{ for all } x\in V\}\,.$$

Indeed assume that there exists $x_0\in V$ such that $\lambda_0=L(p,x_0)>0$; by the homogeneity of degree 1 of $L(p,\cdot\,)$ it follows that

$$L(p,x)\ge C\,|x|\,\lambda_0\ ;\qquad x\in V\,,\ |x-|x|\,x_0|\le 1\,,$$

provided that $|x|$ is large enough. And if we integrate the "shifted" Gaussian inside the tube $\{x:\ |x-|x|\,x_0|\le 1\}$ we see that $\ell(p)>0$. Conversely of course $\ell(p)=0$ if $L(p;x)\le 0$ $(x\in V)$.

Parts i) and ii) of Theorem 2 follow at once from this, we also have

$$\begin{aligned}&\|\tilde T_t\|_{1\to p}=\|\psi_t\|_p=O(1)\,,\quad \text{as } t\to+\infty\,,\ L\le p\le+\infty\,.\\ &\|\tilde T_t\|_{r\to 2}=O(1)\,,\quad \text{as } t\to+\infty\,,\ 1\le r\le 2\,.\end{aligned}\tag{4.1}$$

For $r=2$ the second estimate is the definition of amenability; for $r=1$ it is very easy and has been verified in [10]; we interpolate to obtain the other values of r.

If we assume that $L<2$ and interpolate between $\|\cdot\|_{r\to 2}$ and $\|\cdot\|_{1-L}$ we obtain iii).

Part iv) of theorem is a trifle more subtle and it relies on the a priori knowledge of

$$\|\tilde T_t\|_{p\to p}=\exp(\rho^2\,(1/2-1/p)^2\,t)\,,\qquad 1\le p\le+\infty\,,\tag{4.2}$$

where ρ^2 is defined by $\Delta m=-\rho^2 m$. (4.2) is an immediate consequence of amenability and of the formula (*cf.* [10])

$$\tilde T_t f=(f\,m^{-1/p}*\mu_t\,m^{1/2-1/p})\,m^{1/p}\,.$$

On the other hand we clearly have by the semigroup property:

$$(4.3)\quad \|\tilde{T}_{t+2}\|_{1\to p} \le C_2\,\|\tilde{T}_{t+1}\|_{r\to p} \le C_1\,\|\tilde{T}_t\|_{p\to p}\,, \quad 1\le r\le p\le +\infty\,.$$

The only thing that has to be done to complete the proof of iv) is to show that under the WNC-condition, up to negligible error, we have

$$\exp(\rho^2\,(1/2-1/p)^2\,t) \le \|\psi_t\|_p\,.$$

To see this let us choose $x_0 \in V$ so that

$$L_j(x)\ge 0\,; \qquad j=1,2,\dots,n\,,\ \ \langle x,x_0\rangle \ge 0\,,$$

which is possible by the WNC-condition. We clearly have

$$L(p,x) = \Big(\frac{1}{p}-\frac{1}{2}\Big)\sum L_j(x)\,, \qquad \langle x,x_0\rangle \ge 0\,,$$
$$L(p,x) = \frac{1}{2}\sum L_j(x)\,, \qquad \langle x,x_0\rangle \le 0\,.$$

It follows therefore that up to negligible error

$$(4.4)\quad \begin{aligned}&\int \exp\Big(L(p,x)-\frac{|x|^2}{4t}\Big)\,dx\\ &\qquad = \int \exp\Big(\Big(\frac{1}{p}-\frac{1}{2}\Big)\sum L_j(x)-\frac{|x|^2}{4t}\Big)\,dx\\ &\qquad\quad + \text{negligible error.}\end{aligned}$$

A moment reflexion gives on the other hand

$$\Big|\sum L_j\Big|_{V^*} = \rho$$

for the dual norm in V^*. The right hand side of (4.4) can therefore be explicitely computed, this together with our theorem proves our assertion and it gives iv).

To prove v) observe that for every fixed q the function

$$\ell(p,q) = \limsup_{t\to+\infty}\frac{1}{t}\,\log\|\tilde{T}_t\|_{p\to q}$$

is, as a function of p, non decreasing (*cf.* (4.3)) and convex in $1/p$ (by Riesz-Thorin). Under the condition (0.15) of v) it follows therefore that

$$\ell(q) = \ell(1,q) = \left(\frac{1}{2} - \frac{1}{q}\right)^2 \rho^2$$

for some $1 < q < 2$. To finish the proof of v) we shall insert in (0.14) the above value of $\ell(q)$ and prove that this implies the WNC-condition on the L_i's. Towards that observe that we can assume that the L_i's span the space V^* for otherwise we can quotient out $\cap Ker L_i$ and this reduces the integral in (0.14) to a lower dimensional Gaussian. Then (*cf.* the remark that follows the definition of Section 0.4) since

$$L(p,x) = \left(\frac{1}{p} - \frac{1}{2}\right) \sum L_j(x) + \left(1 - \frac{1}{p}\right) \sum L_j^-(x)\,,$$

if G is not WNC, we must have

$$L(p,x) \leq \left(\frac{1}{p} - \frac{1}{2}\right) \sum L_j(x) - \varepsilon\,, \qquad |x| = 1\,,$$

for some $\varepsilon > 0$. This will immediatly give us a contradiction and will complete the proof of v).

The corollaries 1 and 2 are easy to prove. Indeed

$$\left\| \int_t^{t+1} \tilde{T}_s \psi_1\, ds \right\|_p \geq C\, \|\psi_t\|_p\,, \qquad t > 1\,,$$

by the local Harnack principle. Corollary 1 follows.

If G is as in Corollary 2 we have (*cf.* [8], [10])

$$\|\tilde{T}_t\|_{1\to+\infty}\,,\ \|\tilde{T}_t\|_{1\to 2} = O(e^{-ct^{1/3}})\,, \qquad \|\tilde{T}_t\|_{2\to 2} = O(1)\,,$$

as $t \to +\infty$ for some $c > 0$. By interpolation it follows that

$$\|\tilde{T}_t\|_{p\to q} = O(e^{-ct^{1/3}})\,, \qquad 1 \leq p \leq 2 \leq q \leq +\infty\,,\ p < q\,.$$

If we interpolate this estimate with (4.1) we see that for all $1 \leq p < 2$ there exists $L(p) < 2$ such that

$$\|\tilde{T}_t\|_{p\to q} = O(e^{-ct^{1/3}})\,, \qquad \text{as } t \to +\infty\,,\ L(p) < q \leq +\infty\,.$$

This estimate takes care of the convergence of

$$\int_1^{+\infty} t^{\alpha/2-1}\,\tilde{T}_t\,dt$$

the convergence of $\int_0^1$ is taken care off by standard methods (*cf.* [11]). Corollary 2 follows.

5. Non simply connected groups.

The key to the analysis of non simply connected groups is the following observation. Let G be some connected Lie group and let $K \subset G$ be some compact normal subgroup. We can then use the projected sublaplacian $d\pi(\Delta) = \dot{\Delta}$ by the projection $\pi : G \longrightarrow G/K$ to define T_t, $\tilde{T}_t$, ϕ_t, ψ_t both on G and on G/K. Let $\dot{\psi}_t$ be the corresponding kernel on G/K. It is then clear (*cf.* [8]) that

$$\dot{\psi}_t(x) = \int_K \psi_t(xk)\,dk = \int_K \psi_t(kx)\,dk\,.$$

By an easy application of the local Harnack estimates it follows therefore that

$$\|\dot{\psi}_t\|_{L^p(G/K)} \approx \|\psi_t\|_{L^p(G)}\,, \qquad t \geq 1\,,\ 1 \leq p \leq +\infty\,.$$

This means that, at least the Hardy-Littlewood estimates and the results of Section 0.6 pass through a quotient by some compact subgroup.

To go further we shall have to introduce some notations. We shall say that some soluble connected Lie group Q is *admissible* soluble if there exists some simply connected group $\tilde{Q}$ and some covering map $\theta : \tilde{Q} \longrightarrow Q$ such that $\operatorname{Ker}\theta \cap \tilde{N} = \{e\}$ where $\tilde{N} \subset \tilde{Q}$ is the nilradical of $\tilde{Q}$. The fact that $\theta(\tilde{N}) = N \subset Q$ is the nilradical of Q (*i.e.*, a closed subgroup) implies that

$$\operatorname{dist}(\tilde{N}, \operatorname{Ker}\theta \setminus \{e\}) > 0\,,$$

$$Q/N \cong V \times T = \tilde{Q}/N \cdot \operatorname{Ker}\theta\,,$$

$$V \approx \mathbb{R}^a\,, \qquad T \approx \mathbb{T}^b = (\mathbb{R} \bmod 2\pi)^b\,.$$

We shall say that G some connected Lie group is *admissible* if $G \approx Q\lambda M$ where Q is admissible soluble and M is compact.

Let G now be some amenable connected Lie group. By standard global structure theorems (*cf.* [1], [7], [14]) we see that we can "cover" G by $\theta : Q\lambda M \longrightarrow G$ with $\operatorname{Ker}\theta$ finite (*i.e.*, an "isogeny"). Inside the center of the nilradical of Q on the other hand we can find some compact subgroup $K(\approx \mathbb{T}^b)$ such that G/K is admissible.

The consequence of the above structure theorems and of our previous considerations is that if we can extend our Hardy-Littlewood theory of Section 0.6 from simply connected groups to admissible groups then we automatically have it for all amenable groups. For these theorems it should be observed that the non zero real parts of the roots of G and of the corresponding admissible group (as constructed above) are up to obvious identification the same.

When G is an admissible group then its nilradical $N \subset G$ is simply connected and $G/N \cong V \times K$ where $K = T \times M$ so that the only difference with the simply connected case is the presence of the torus T as a cofactor of M.

By going through the proofs of both the upper and the lower estimates of Section 0 we see therefore that strictly nothing changes neither in the proofs nor in the statements of the results.

The conclusion is in particular that theorems 1 and 2 and their corollaries are valid as stated, for general, not necessarily simply connected, groups.

APPENDIX: The brownian bridge.

Let us recall here some facts on the Brownian bridge (*cf.* [15], [16]) *i.e.*, brownian motion $b(t) \in \mathbb{R}^a = V$ conditioned by

$$b(0) = x\,, \qquad b(t) = y\,.$$

We shall denote this process by $b^t_{x,y}(s)$, $0 \leq s \leq t$, and recall the following well known facts ($\simeq$ denotes equidistributed processes)

i) $b^t_{x,y}(s) \simeq x + s(y-x)/t + b^t_{0,0}(s)$,

ii) $b^t_{0,0}(s) \simeq b(s) - s\,b(t)/t$,

iii) $b^{\lambda t}_{\sqrt{\lambda}x,\sqrt{\lambda}y}(\lambda s) \simeq \sqrt{\lambda}\, b^t_{x,y}(s)$,

iv) $b^t_{0,0}(s) \simeq (1 - s/t)\, b(st/(t-s))$,

v) $b^t_{x,y}(s) \simeq b^t_{y,x}(t-s)$, (*cf.* [15], [16] together with the scalling properties of standard brownian motion).

vi) The standard Markov property implies that with respect to the conditional probability $\mathbf{P}[\cdot \mid b^t_{x,y}(t/2)]$ the two "halfs" of the brownian bridge:

$$\{b^t_{x,y}(s) : \ 0 < s < t/2] , \qquad \{b^t_{x,y}(s) : \ t/2 < s < t\} ,$$

are independents.

We shall now denote by

$$P(D) = \mathbf{P}(L_j(b^t_{0,0}(s)) \leq D : \ 0 < s < t , \ j = 1, 2, \ldots, n) ,$$

where $L_1, \ldots, L_n \in V^*$ are linear functionals of V. We shall say that the L_j's satisfy the NC-condition if there exists $x \in V$ such that $L_j(x) > 0$, $j = 1, 2, \ldots, n$ (*cf.* Section 0.4). In fact in what follows the L_js will be the real parts of the roots as defined in Section 0. We shall prove the following

Lemma 1. *There exist constants $C, c > 0$ independent of t and D such that:*

$$P(D) \geq C\, e^{-ct/D^2} , \qquad t, \ D > 0 . \tag{A.1}$$

If we suppose in addition that the $L_j \in V^*$, $j = 1, 2, \ldots,$ satisfy the NC-condition then

$$P(D) \geq C \left(\frac{t}{D^2}\right)^{-c} . \tag{A.2}$$

The estimate (A.1) is a consequence of the following

$$\mathbf{P}(|b^t_{0,0}(s)| \leq D ; \ 0 < s < t) \geq C\, e^{-ct/D^2} , \tag{A.3}$$

but (A.3) immediately reduces because of ii) or iv) to the corresponding estimate for brownian motion which is well known (see *e.g.* [7]).

The estimate (A.2) is more subtle to prove. We use v) and vi) to see that

$$P(D) = \int \left(\mathbf{P}(L_j(b^t_{0,0}(s)) \leq D ; \ 1 \leq j \leq n , \ 0 < s < t/2 \mid b^t_{0,0}(t/2) = z)\right)^2 \cdot \mathbf{P}(b^t_{0,0}(t/2) \in dz)$$

$$\geq \Big(\int (\cdots)\, \mathbf{P}(b^t_{0,0}(t/2) \in dz) \Big)^2$$
$$= \big(\mathbf{P}(L_j(b^t_{0,0}(s)) \leq D\,;\ 1 \leq j \leq n\,;\ 0 \leq s \leq t/2) \big)^2 .$$

But because of iv) the right hand sideof the above can be replaced by the coresponding expression with $b^t_{0,0}(\,\cdot\,)$ replaced by $b(\,\cdot\,)$. The corresponding lower estimate is easy to obtain (*cf.* [7]).

References.

[1] Varadarajan, V. S., *Lie groups, Lie algebras and their representations.* Prentice-Hall, 1984.

[2] Hörmander, L., Hypoelliptic second order equations. *Acta Math.* **119** (1967), 147-171.

[3] Chevalley, C., *Théorie des groupes de Lie*, tomes II & III. Hermann 1955.

[4] Varopoulos, N. Th., Diffusion on Lie groups (II). *Can. J. Math.* **46**, (1994), 1073-1093.

[5] Jacobson, *Lie algebras.* Interscience, 1962.

[6] Bourbaki, N., *Eléments de mathématiques, Livre* VI, *Intégration.* Hermann, 1963.

[7] Varopoulos, N. Th., Analysis on Lie group, preprint.

[8] Varopoulos, N. Th., Diffusion on Lie groups (I). *Can. J. Math.* **46** (1994), 438-448.

[9] Alexopoulos, G., An application of homogenization theory to harmonic analysis: Harnack inequalities and Riesz transforms on Lie groups of polynomial growth. *Can. J. Math.* **44** (1992), 691-727.

[10] Varopoulos, N. Th., Thèorie de Hardy-Littlewood sur les groupes de Lie. *C. R. Acad. Sci. Paris, Serie I* **316** (1993), 999-1003.

[11] Varopoulos, N. Th., Saloff-Coste, L., Coulhon, Th., *Analysis and geometry on groups.* Cambridge Univ. Press, 1992.

[12] Lohoué, N., Estimées de type Hardy pour l' opérateur $\Lambda + \lambda$ d' un espace symétrique non compact.*C. R. Acad. Sci. Paris.* **308** (1989), 11-14.

[13] Lohoué, N., Inégalités de Sobolev pour les sous Laplaciens de certains groups unimodulaires. *Geom. Funct. Anal.* **4** (1992), 394-420.

[14] Varopoulos, N. Th., Hardy-Littlewood theory on unimodular groups, to appear in *Ann. Inst. H. Poincaré.*

[15] Knight, F. B., Essentials of Brownian motion and diffusion. *Math. Survey* **18**. Amer. Math. Soc., 1981.

[16] Ikeda, N., Watanabe, S., *Stochastic differential equations and diffusion processes.* North-Holland **24**, 1989.
[17] Varopoulos, N. Th., Mustapha, S., Lecture notes Université Paris VI.

Recibido: 15 de febrero de 1.995

N. Th. Varopoulos
I.U.F. and
Departement de Mathématiques
Université de Paris VI
75005 Paris, FRANCE