HEAT KERNEL AND SEMIGROUP ESTIMATES FOR
SUBLAPLACIANS WITH DRIFT ON LIE GROUPS

NICK DUNGEY

Abstract

Let $G$ be a Lie group. The main new result of this paper is an estimate in $L^2(G)$ for the Davies perturbation of the semigroup generated by a centered sublaplacian $H$ on $G$. When $G$ is amenable, such estimates hold only for sublaplacians which are centered. Our semigroup estimate enables us to give new proofs of Gaussian heat kernel estimates established by Varopoulos on amenable Lie groups and by Alexopoulos on Lie groups of polynomial growth.

1. Introduction

In this paper we establish an estimate for the Davies perturbation of the semigroup generated by a centered sublaplacian on a Lie group. This result constitutes a new and useful form of $L^2$ off-diagonal estimate for such semigroups. From our result, we derive a new proof of a Gaussian estimate of Varopoulos [18] for the heat kernel of a centered sublaplacian, and also recover a more precise Gaussian estimate of Alexopoulos [2], [1] for the case where the Lie group has polynomial volume growth.

The analyses of [18] and [2] are based on probabilistic methods to study the diffusion associated with a sublaplacian. The approach of the present paper, on the other hand, is functional-analytic rather than probabilistic. Note that while [18] and [2] make essential use of the detailed structure theory and geometry for amenable or polynomial growth Lie groups, we largely avoid this and for this reason we feel our approach is often technically simpler.

Our methods can be extended to study convolution powers of centered probability densities, but this extension is not trivial and will be described elsewhere [7]. See [3], [18] for centered densities on Lie groups.

2000 Mathematics Subject Classification. 22E30, 35B40, 58J35.
Key words. Lie group, heat kernel, Gaussian estimates, sublaplacian.
To state our results precisely we fix some notation. Let $G$ be a connected Lie group. Denote by $dg$ the left Haar measure on $G$ and let $d\bar{g} = d(g^{-1}) = \Delta(g^{-1}) \, dg$ be the corresponding right Haar measure where $\Delta: G \to (0, \infty)$ is the modular function. Let $\mathfrak{g}$ be the Lie algebra of $G$, consisting of all right invariant vector fields on $G$. Suppose that $A_0, A_1, \ldots, A_d$ are elements of $\mathfrak{g}$ such that $A_1, \ldots, A_d$ algebraically generate the Lie algebra $\mathfrak{g}$, and consider the subelliptic sublaplacian

$$H = - \sum_{i=1}^{d'} A_i^2 + A_0$$

with drift term $A_0$. It is well known (see for example [14, Section IV.4]) that $H$ generates a contraction semigroup $S_t = e^{-tH}$ in the spaces $L^p := L^p(G; dg)$, $1 \leq p \leq \infty$, and we denote by $K_t: G \to (0, \infty)$ the corresponding heat kernel which satisfies

$$(1) \quad (S_t f)(g) = (K_t * f)(g) = \int_G dh \, K_t(h) f(h^{-1}g) = \int_G dh \, K_t(gh^{-1}) f(h)$$

for all $t > 0$, $g \in G$ and $f \in C_c^\infty(G)$.

The notion of centeredness is defined as follows (cf. [2], [18]). Let $G_0$ be the closure of the group $[G, G]$ in $G$, and consider the canonical homomorphism $\pi_0: G \to G/G_0$. Note that $G/G_0$ is a connected abelian Lie group and it can therefore be written as a direct product $\mathbb{R}^{n_1} \times \mathbb{T}^{n_2}$, where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. Set $G_1 = \pi_0^{-1}(\{0\} \times \mathbb{T}^{n_2}) \subseteq G$ and let $\mathfrak{g}_1$ be the Lie algebra of $G_1$. Then $\mathfrak{g}_1$ is an ideal of $\mathfrak{g}$ and $[\mathfrak{g}, \mathfrak{g}] \subseteq \mathfrak{g}_1$. (If $G$ is simply connected, one just has $\mathfrak{g}_1 = [\mathfrak{g}, \mathfrak{g}]$.) Observe that $G/G_1 \cong \mathbb{R}^{n_1}$. One says that $H = - \sum_{i=1}^{d'} A_i^2 + A_0$ is centered if $A_0 \in \mathfrak{g}_1$.

It is not difficult to show that $H$ is centered if and only if $H\eta = 0$ for every continuous homomorphism $\eta: G \to \mathbb{R}$ (one observes that any such homomorphism vanishes on $G_1$, so that $A_0\eta = 0$ when $A_0 \in \mathfrak{g}_1$).

By a well known lemma (see [2, Section 1]), every sublaplacian $H$ on $G$ is conjugate via a character with a centered sublaplacian. Explicitly, one can find a multiplicative character $\Phi: G \to (0, \infty)$, so $\Phi(gh) = \Phi(g)\Phi(h)$, and a constant $\beta \geq 0$ such that $H = \Phi^{-1}(H_c + \beta)\Phi$ where $H_c$ is a centered sublaplacian.

In this sense, the study of general sublaplacians essentially reduces to the centered case.

Let us state our basic result for the Davies perturbation of the semigroup $S_t$. Let $\mathcal{R}^+(G)$ denote the algebra of all real, right invariant differential operators on $G$ without constant term, so $\mathcal{R}^+(G)$ is linearly spanned by all monomials $X_1 \ldots X_j$ where $j \geq 1$ and $X_1, \ldots, X_j \in \mathfrak{g}$. Introduce the set $\mathcal{E}$ consisting of all $C^\infty$-smooth functions $\psi: G \to \mathbb{R}$
such that $P\psi \in L^\infty$ for all $P \in \mathcal{R}^\dagger(G)$; note that $\psi$ itself may be unbounded. We write $\| \cdot \|_{p\rightarrow q}$ for the norm of a bounded linear operator from $L^p = L^p(G; dg)$ to $L^q$.

**Theorem 1.1.** Let $H$ be centered, let $\psi \in \mathcal{E}$ and set $H_\lambda = e^{\lambda\psi}He^{-\lambda\psi}$. Then there exists $k > 0$ such that

$$\text{Re}(H_\lambda f, f) \geq -k\lambda^2 \| f \|_2^2$$

for all $\lambda \in \mathbb{R}$ and $f \in C_c^\infty(G)$. Moreover, $H_\lambda$ extends to the generator of a semigroup $e^{-tH_\lambda} = e^{\lambda\psi}Se^{-\lambda\psi}$ in $L^2$ satisfying

$$\|e^{-tH_\lambda}\|_{2\rightarrow 2} \leq e^{k\lambda^2 t}$$

for all $t > 0$ and $\lambda \in \mathbb{R}$.

Davies perturbation estimates are well known for various classes of elliptic or subelliptic operators on manifolds: see, for example, [5], [14] and [19]. In our situation, however, the drift term $A_0$ creates new and non-trivial difficulties in the proof. These can be overcome by carefully utilizing the properties of the class $\mathcal{E}$ defined above.

When $G$ is amenable, the estimates of Theorem 1.1 fail for non-centered sublaplacians. We shall give a precise version of this statement in Section 2 below. The main fact about amenability needed there is that $G$ is amenable if and only if

$$\mu_0 := \inf_{0 \neq f \in C_c^\infty(G)} \frac{\sum_{i=1}^d \|A_i f\|_2^2}{\|f\|_2^2}$$

is the bottom of the $L^2$ spectrum of the symmetric sublaplacian $-\sum_{i=1}^d A_i^2$.

If $G$ is non-amenable, the statements of Theorem 1.1 hold for all sublaplacians, centered or not, as an easy consequence of $\mu_0 > 0$ (see remarks of Section 2 below). See also [18] for discussions of the role of amenability.

We define a distance on $G$ in the following standard way. Fix a compact neighborhood $U$ of the identity $e$ of $G$ which is symmetric ($U = U^{-1}$) and define $\rho: G \rightarrow \mathbb{N} = \{1, 2, 3, \ldots\}$ by

$$\rho(g) = \inf \{ n \in \mathbb{N}: g \in U^n \}, \quad g \in G,$$

where $U^n := \{ g_1g_2 \cdots g_n : g_j \in U \}$.

Although $\rho \notin \mathcal{E}$ in general, we shall construct in Section 3 a function $\psi_0 \in \mathcal{E}$ with the same global growth as $\rho$, that is, with $c^{-1}\rho \leq \psi_0 \leq c\rho$ for some $c > 1$. Together with Theorem 1.1, this leads to a new proof of the following Gaussian estimate of Varopoulos [18].
Theorem 1.2 ([18]). Let $H$ be centered. Then there are $c, b > 0$ such that

$$K_t(g) \leq c \Delta(g)^{-1/2} e^{-b \rho(g)^2/t}$$

for all $t \geq 1$, $g \in G$.

Recall (cf. [10]) that $G$ is said to have polynomial growth of order $D$ if an estimate $c^{-1} n^D \leq d g(U^n) \leq cn^D$ holds for some $c > 1$ and all $n \in \mathbb{N}$; alternatively, if $d g(U^n) \geq a e^{an}$ for some constant $a > 0$ and all $n \in \mathbb{N}$, then we say $G$ has exponential volume growth.

We shall give a new proof of the following theorem which was proved by Alexopoulos [2] with a quite difficult proof. (In the special case where $G$ is nilpotent, an easier proof is given by Melzi [13]. The possibility of using Davies perturbation estimates to prove Alexopoulos’ theorem is mentioned in [18, p. 435] without giving details.)

Theorem 1.3 ([2]). Suppose $G$ has polynomial growth of order $D$, and let $H$ be centered. Then there are $c, b > 0$ such that

$$K_t(g) \leq c t^{-D/2} e^{-b \rho(g)^2/t}$$

for all $t \geq 1$ and $g \in G$.

Our proof of Theorem 1.3 is based on Theorem 1.1 and on ideas of [6] and involves certain weighted Nash inequalities for convolution operators.

The semigroup $S_t = e^{-tH}$ generated by a sublaplacian with drift is not bounded analytic in general (it is, however, analytic if $A_0$ is in the linear span of the fields $\{A_i, [A_i, A_j]: i, j \in \{1, \ldots, d\}\}$: see [9], [15], [16] for small time estimates in this case). Thus the estimate $\|HS_t\|_{2\to2} = \|\partial_t S_t\|_{2\to2} \leq ct^{-1}$, $t > 0$, fails in general. Nevertheless, when $H$ is centered we have the following interesting large time regularity result for $S_t$.

Theorem 1.4. Let $H$ be centered and let $p \in (1, \infty)$. Then there is $c = c(p) > 0$ such that $\|HS_t\|_{p\to p} \leq ct^{-1}$ for all $t \geq 1$.

From Theorem 1.4 and since $\sum_{i=1}^{d'} \|A_i S_t f\|_2^2 = \text{Re}(HS_t f, S_t f)$ for all $f \in L^2$, one deduces the estimate of spatial derivatives

$$\|A_i S_t\|_{2\to2} \leq ct^{-1/2}$$

for all $i \in \{1, \ldots, d\}$, $t \geq 1$, when $H$ is centered. This estimate is apparently new for centered sublaplacians on general Lie groups.

In the particular case where $G$ has polynomial growth, Theorem 1.4 is contained in results of [2] (which also imply the analogous estimate
for \( p = 1 \) and \( p = \infty \). Our proof of Theorem 1.4 for general \( G \) will be given elsewhere, since the proof seems best approached via a more general study of regularity of the convolution powers \( K^{(n)} = K \ast K \ast \cdots \ast K \) of a fixed probability density \( K \) on \( G \), rather than by analyzing directly the sublaplacian \( H \).

2. Proof of Theorem 1.1

In this section, we prove Theorem 1.1, and then give a converse result showing that similar estimates fail for non-centered sublaplacians.

In general, \( c, c' \) and so on denote positive constants whose value may change from line to line when convenient. Also, sums over the variable \( i \) are always taken over the range \( i \in \{1, \ldots, d'\} \) unless otherwise indicated.

A key idea in the proof of Theorem 1.1 is to write the vector field \( A_0 \in \mathfrak{g}_1 \) in terms of second or higher order derivatives (or differences). To do this we require the following algebraic result.

**Proposition 2.1.** There exists a compact, connected, abelian, possibly trivial subgroup \( C \) of \( G \) such that \( g_1 = c + [g, g] \), where \( c \) is the Lie algebra of \( C \) and + denotes a sum of vector spaces which is not necessarily direct.

Some related descriptions of the algebra \( g_1 \) are given in [18, Appendix] in the case where \( G \) is amenable. For completeness, we give a proof of Proposition 2.1 in the Appendix of this paper.

We remark that compactness of \( C \), but not the fact that \( C \) is abelian, will be essential in the arguments to follow.

To begin the proof of Theorem 1.1, let us fix \( \psi \in \mathcal{E} \). The norm estimate \( \|e^{-tH}\|_{2 \to 2} \leq e^{k \lambda^2 t} \) follows in a routine manner from the first estimate of the theorem. Indeed, using the first estimate one obtains the differential inequality

\[
\frac{d}{dt}(\|e^{-tH}\|_{2 \to 2}^2) = -2 \Re(H_\lambda e^{-tH_\lambda} f, e^{-tH_\lambda} f) \leq 2k \lambda^2 \|e^{-tH_\lambda} f\|_2^2
\]

which implies the desired norm estimate (compare, for example, [5, Section 4]). Thus we concentrate on proving the first estimate of the theorem.

Observing the basic identity \( e^{\lambda \psi} X (e^{-\lambda \psi} f) = X f - \lambda (X \psi) f \) for \( \lambda \in \mathbb{R} \), \( X \in \mathfrak{g} \), \( f \in C_0^\infty (G) \), since \( H = - \sum_i A_i^2 + A_0 \) we find that

\[
H_\lambda f = - \sum_i A_i^2 f + A_0 f + \lambda \sum_i (A_i \psi) A_i f + \lambda \sum_i A_i ((A_i \psi) f)
- \lambda^2 \sum_i (A_i \psi)^2 f - \lambda (A_0 \psi) f.
\]
Then
\[(H_\lambda f, f) = \sum_i \|A_if||_2^2 + (A_0f, f) + \lambda \sum_i (A_if, (A_\psi)f) + \lambda \sum_i ((A_\psi)f, A_if) - \lambda^2 \sum_i \|(A_\psi)f||_2^2 - \lambda((A_0\psi)f, f).\]

Because the terms \((A_0f, f)\) and \((A_if, (A_\psi)f) - ((A_\psi)f, A_if)\) are purely imaginary, then
\[\text{Re}(H_\lambda f, f) = \sum_i \|A_if||_2^2 - \lambda^2 \sum_i \|(A_\psi)f||_2^2 - \lambda((A_0\psi)f, f)\]
\[\geq \sum_i \|A_if||_2^2 - c\lambda^2\|f\|_2^2 - \lambda((A_0\psi)f, f),\]
where we used the finiteness of the norms \(\|A_if\|_\infty, i \in \{1, \ldots, d\}\). Our main task will be to prove an estimate
\[|(A_0\psi)f, f)| \leq c' \sum_i \|A_if\|_2 \|f\|_2\]
for all \(f \in C_c^\infty(G)\). Note the elementary estimate that
\[\|A_if\|_2 \|f\|_2 \leq 2^{-1} \varepsilon \|A_if\|_2^2 + 2^{-1} \varepsilon^{-1} \lambda^2 \|f\|_2^2\]
for all \(\varepsilon > 0, \lambda \in \mathbb{R}\) and \(i \in \{1, \ldots, d\}\). Then by choosing \(\varepsilon > 0\) sufficiently small and using (3) and (4), we obtain for some \(k > 0\) an estimate of the desired form \(\text{Re}(H_\lambda f, f) \geq -k\lambda^2\|f\|_2^2, \lambda \in \mathbb{R}\).

Thus, it remains to establish (4). Apply Proposition 2.1 to decompose \(A_0 = A_0' + A_0''\) with \(A_0' \in [\mathfrak{g}, \mathfrak{g}]\) and \(A_0'' \in c\). We will estimate the terms \((A_0\psi)f, f\) and \((A_0'\psi)f, f\) separately.

Note that \([\mathfrak{g}, \mathfrak{g}]\) is linearly spanned by all commutators of the form \([A_{i_1}, \ldots, A_{i_k}]\) with \(k \geq 2\) and \(i_1, \ldots, i_k \in \{1, \ldots, d\}\), since \(A_1, \ldots, A_{d'}\) generate the Lie algebra \(\mathfrak{g}\). Therefore \(A_0' \in [\mathfrak{g}, \mathfrak{g}]\) is expressible as a sum of terms each of the form \(A_if\) for some \(i \in \{1, \ldots, d\}\) and \(P \in \mathcal{R}^+(G)\). Note that
\[\|(A_i P\psi)f, f\) = (A_i((P\psi)f), f) - ((P\psi)(A_if), f)\]
\[= -((P\psi)f, A_if) - (A_if, (P\psi)f)\]
and hence \(|(A_i P\psi)f, f)| \leq 2\|P\psi\|_\infty \|f\|_2 \|A_if\|_2\). Since \(\|P\psi\|_\infty\) is finite for any \(P \in \mathcal{R}^+(G)\), this argument shows that
\[|(A_0'\psi)f, f)| \leq c \sum_i \|A_if\|_2 \|f\|_2\]
for all \(f \in C_c^\infty(G)\).
To handle the term \( ((A_0'' \psi) f, f) \), define an operator \( P \) acting on \( C^\infty \) functions \( \varphi: G \to \mathbb{R} \) by

\[
(P \varphi)(g) = \int_C ds \varphi(sg) = \left( \int_C ds L(s) \varphi \right) (g), \quad g \in G,
\]

where \( ds \) denotes Haar measure on the compact group \( C = \exp(c) \) normalized so that \( ds(C) = 1 \). Here, \( L \) is the left regular representation of \( G \) which acts by the formula \( (L(g) \varphi)(h) = \varphi(g^{-1} h) \), \( g, h \in G \).

Since \( A_0'' \in \mathfrak{c} \) it is easy to see that \( PA_0'' \varphi = 0 \). Defining the difference operators \( \partial_s := I - L(g) \) for \( g \in G \), we may therefore write the operator \( A_0'' \) in the form

\[
A_0'' = (I - P) A_0'' = \int_C ds \partial_s A_0''.
\]

Observing the general identity \( (\partial_s A_0'' \psi) f, f) = (\partial_s (A_0'' \psi) f, f) - ((L(s) A_0'' \psi) \partial_s f, f) \)

\[
= (A_0'' \psi f, \partial_s -1 f) - ((L(s) A_0'' \psi) \partial_s f, f)
\]

for all \( s \in C \), since \( \partial_s -1 \) is adjoint to \( \partial_s \). Since \( \|L(s) A_0'' \psi\|_{\infty} = \|A_0'' \psi\|_{\infty} \) is finite and \( \|\partial_s f\|_2 = \|\partial_s -1 f\|_2 \), we see that

\[
\|((\partial_s A_0'' \psi) f, f)\|_2 \leq 2 \|A_0'' \psi\|_{\infty} \|f\|_2 \|\partial_s f\|_2
\]

for all \( s \in C \). Recall the standard inequality (cf. [14, pp. 267–268])

\[
(5) \quad \|\partial_s f\|_p \leq c \rho(g) \sum_i \|A_i f\|_p
\]

which is valid for all \( g \in G \) and \( p \in [1, \infty] \). From the above observations, and the compactness of \( C \) which implies that \( \sup\{\rho(s): s \in C\} \) is finite, we deduce that

\[
\|((A_0'' \psi) f, f)\|_2 \leq \int_C ds \|((\partial_s A_0'' \psi) f, f)\|_2 \leq c \sum_i \|A_i f\|_2 \|f\|_2
\]

for all \( f \in C^\infty_c(G) \). This completes the proof of (4) and of Theorem 1.1. \( \square \)

**Remarks.**

- In general, the constant \( k \) in Theorem 1.1 depends on the choice of \( \psi \in \mathcal{E} \). However, if \( \mathcal{E}' \) is some subset of \( \mathcal{E} \) which is uniformly bounded, in the sense that

\[
\sup_{\psi \in \mathcal{E}'} \|P \psi\|_{\infty} < \infty
\]
for each \( P \in \mathcal{F}^+(G) \), then it easily follows from the above proof that one can choose the same constant \( k \) uniformly for all \( \psi \in \mathcal{E}' \).

- The proof of the theorem, and (3), yield the following stronger inequality for \( (H_\lambda f, f) \): for each \( \varepsilon \in (0, 1) \) there is a \( c(\varepsilon) > 0 \) such that

\[
\Re(H_\lambda f, f) \geq (1 - \varepsilon) \sum_i \|A_i f\|^2_2 - c(\varepsilon) \lambda^2 \|f\|^2_2
\]

for all \( f \in C^\infty_c \) and \( \lambda \in \mathbb{R} \). This inequality will be needed in the proof of Theorem 1.3. For a symmetric sublaplacian (that is, \( A_0 = 0 \)) one can choose \( \varepsilon = 0 \) in (6), as follows from (3). We do not know if one can take \( \varepsilon = 0 \) for a general centered sublaplacian.

- For an arbitrary sublaplacian (possibly non-centered), inequality (3) implies the crude estimate

\[
\Re(H_\lambda f, f) \geq (\mu_0 - c\lambda^2 - c'|\lambda|)\|f\|^2_2
\]

for \( \lambda \in \mathbb{R}, f \in C^\infty_c(G) \), where \( \mu_0 \) is as in (2) and \( c' = \|A_0 \psi\|_{\infty} \).

If \( G \) is non-amenable, then since \( \mu_0 > 0 \) we see that the estimates of Theorem 1.1 hold even for non-centered sublaplacians.

We next prove a result converse to Theorem 1.1.

**Theorem 2.2.** Suppose \( G \) is amenable and let \( H \) be a sublaplacian which is not centered. Then the estimate of Theorem 1.1 fails when \( \lambda \) is close to zero. More precisely, there exist a homomorphism \( \Phi: G \to \mathbb{R} \) (so \( \Phi(gh) = \Phi(g) + \Phi(h), g, h \in G \) with \( \Phi \in \mathcal{E} \), and constants \( \alpha, \beta > 0 \) such that \( H_\lambda := e^{\lambda \Phi} H e^{-\lambda \Phi} \) satisfies

\[
\inf_{0 \neq f \in C^\infty_c} \frac{\Re(H_\lambda f, f)}{\|f\|^2_2} = -\beta \lambda^2 - \alpha \lambda
\]

and

\[
\|e^{-tH_\lambda}\|_{2 \to 2} = e^{(\beta \lambda^2 + \alpha \lambda)t}
\]

for all \( \lambda \in \mathbb{R} \) and \( t > 0 \). It follows that \( \sup \{e^{-k\lambda^2 t}\|e^{-tH_\lambda}\|_{2 \to 2}: t \geq 1, 0 < \lambda \leq t^{-1/2}\} = \infty \) for any number \( k > 0 \).

**Proof:** Consider the homomorphism \( \pi_1: G \to G/G_1 \cong \mathbb{R}^{n_1} \) and identify \( G/G_1 \) with \( \mathbb{R}^{n_1} \). The vector fields \( A'_i := d\pi_1(A_i), i \in \{0, 1, \ldots, d'\} \), are constant coefficient fields on \( \mathbb{R}^{n_1} \), and \( A'_0 \neq 0 \) because \( H \) is not centered. Therefore one can find \( b \in \mathbb{R}^{n_1} \) such that the function \( F(x) := \langle b, x \rangle \) satisfies \( A'_0 F = 1 \), where \( \langle \cdot, \cdot \rangle \) is the usual inner product in \( \mathbb{R}^{n_1} \). Because \( A_1, \ldots, A_{d'} \) generate \( \mathfrak{g} \), the fields \( A'_1, \ldots, A'_{d'} \) must linearly span
the tangent space of \( \mathbb{R}^{n_1} \), so at least one of the constants \( \lambda_i := A_i' F_i \), \( i \in \{1, \ldots, d'\} \), is non-zero.

Then \( \Phi := F \circ \pi_1 : G \to \mathbb{R} \) is a homomorphism satisfying \( A_i \Phi = \lambda_i \), \( i \in \{1, \ldots, d'\} \), and \( A_0 \Phi = 1 \). Clearly \( P \Phi \) is constant for any \( P \in \mathcal{R}^+(G) \), so \( \Phi \in \mathcal{E} \). Note that \( \beta := \sum_{i=1}^{d'} \lambda_i^2 > 0 \), and by calculating as in the proof of Theorem 1.1,

\[
H_\lambda = e^{\lambda \Phi} H e^{-\lambda \Phi} = -\sum_i A_i^2 + A_0 + 2 \sum_i (\lambda \lambda_i) A_i - (\beta \lambda^2 + \lambda).
\]

Then \( \text{Re}(H_\lambda f, f)/(\|f\|_2^2) = \sum_i \|A_i f\|_2^2/(\|f\|_2^2) - (\beta \lambda^2 + \lambda) \) and taking infimums over \( f \in C_c^\infty(G) \) yields the first statement of the theorem, since amenability means that \( \mu_0 = 0 \) in (2).

That \( \|e^{-tH_\lambda}\|_{2 \to 2} \leq e^{(\beta \lambda^2 + \lambda)t} \) follows from the first statement of the theorem. To prove the reverse inequality, consider \( H'_\lambda := d\pi_1(H_\lambda) \) which is an elliptic operator on \( \mathbb{R}^{n_1} \) with constant coefficients and with constant term \(-\beta \lambda^2 - \lambda\). It is easy to see, via the Fourier theory of \( L^2(\mathbb{R}^{n_1}) \), that \( \|e^{-tH'_\lambda}\|_{2 \to 2} = e^{(\beta \lambda^2 + \lambda)t} \). Since \( G \) is amenable, a well known transference theorem [4, Theorem 2.4] gives \( \|e^{-tH_\lambda}\|_{2 \to 2} \leq \|e^{-tH'_\lambda}\|_{2 \to 2} \), and (7) is proved.

The final statement of the theorem follows directly from (7) upon choosing \( \lambda \sim t^{-1/2} \).

\[\square\]

3. A smooth distance

The aim of this section is to prove the following lemma providing a smooth distance function \( \psi_0 \in \mathcal{E} \) on any connected Lie group. This is required for the applications of Theorem 1.1. Let \( \rho : G \to \mathbb{N} \) be defined as in Section 1.

**Lemma 3.1.** There exists a \( \psi_0 \in \mathcal{E} \) satisfying \( \psi_0(g) \geq 1 \), \( \psi_0(g) = \psi(g^{-1}) \), \( \psi_0(gh) \leq \psi_0(g) + \psi_0(h) \) and \( \psi_0(g) \leq c \rho(g) \) for all \( g, h \in G \) and some constant \( c > 1 \).

As an aside, the function of Lemma 3.1 automatically satisfies \( P \psi_0 \in L^\infty \) for left invariant differential operators \( P \) without constant term on \( G \). (This follows easily from \( \psi_0(g) = \psi_0(g^{-1}) \) since the inversion map \( g \to g^{-1} \) intertwines left invariant with right invariant operators.)

**Proof of Lemma 3.1:** Consider the modulus \( \rho_B : G \to [0, \infty) \) associated with a fixed vector space basis \( B_1, \ldots, B_N \) of the Lie algebra \( \mathfrak{g} \). Then \( \rho_B(g) \) is the distance from \( g \) to \( e \), with respect to the right invariant Riemannian metric on \( G \) such that \( B_1, \ldots, B_N \) are orthonormal. The functions \( \rho \) and \( \rho_B \) are equivalent at infinity (cf. [19, p. 41]) in the sense
that $c^{-1} \rho \leq \rho_B + 1 \leq c \rho$ for some $c > 1$. Note that $\|B_j \rho_B\|_\infty \leq 1$ for all $j$ and that $\rho_B(g) = \rho_B(g^{-1})$, $\rho_B(gh) \leq \rho_B(g) + \rho_B(h)$ for all $g, h \in G$.

Fix a function $\tau \in C_\infty_c(G)$ with $\tau \geq 0$, $\int_G \tau(g) = 1$, $\tau(g) = \tau(g^{-1})$, $g \in G$, and with support contained in the ball $U_B := \{g \in G : \rho_B(g) < 1\}$. We define

$$\psi_0 := \tau * \rho_B * \tau, \quad \psi_0 := \psi_0 + 6.$$ 

Then $\psi_0 : G \to \mathbb{R}$ is a $C^\infty$-smooth function and $\psi_0 \geq 6$. Because $\rho_B(g) = \rho_B(g^{-1})$ and $\tau(g) = \tau(g^{-1})$, it is clear that $\psi_0(g) = \psi_0(g^{-1})$.

Next, since $|\rho_B(hg) - \rho_B(g)| \leq 1$ and $|\rho_B(gh) - \rho_B(g)| \leq 1$ whenever $h \in U_B$, it is straightforward to verify that

$$\|\psi_0 - \rho_B\|_\infty \leq \|\rho_B - (\tau * \rho_B)\|_\infty + \|(\tau * \rho_B) - (\tau * \rho_B * \tau)\|_\infty \leq 1 + 1 = 2.$$ 

Then $\|\psi_0 - \rho_B\|_\infty \leq 8$, and because $\rho_B$ is equivalent to $\rho$, it easily follows that $c^{-1} \rho \leq \psi_0 \leq c \rho$ for some $c > 1$.

The estimates $\|\psi_0 - \rho_B\|_\infty \leq 2$ and $\rho_B(gh) \leq \rho_B(g) + \rho_B(h)$ imply that $\psi_0(gh) \leq \psi_0(g) + \psi_0(h) + 6$, so that $\psi_0(gh) \leq \psi_0(g) + \psi_0(h)$.

It remains to show that $\psi_0 \in \mathcal{E}$. Let $\langle \cdot, \cdot \rangle$ be the inner product on $\mathfrak{g}$ such that $B_1, \ldots, B_N$ are orthonormal, and define $\chi_{ij} \in C_\infty(G)$ by $\chi_{ij}(g) := \langle \text{Ad}(g^{-1})B_i, B_j \rangle$ where $\text{Ad}$ is the adjoint representation of $G$. For any $\varphi \in C_\infty(G)$, the general identity (cf. [8, p. 200])

$$B_i L(g)\varphi = L(g)(\text{Ad}(g^{-1})B_i)\varphi = \sum_{j=1}^N L(g)\chi_{ij}(g)B_j \varphi$$

implies that

$$B_i (\tau * \varphi) = \sum_{j=1}^N (\chi_{ij} \tau) * (B_j \varphi).$$

Let $Q$ be any right invariant differential operator on $G$ (possibly of order zero). Then $QB_i(\tau * \varphi) = \sum_{j=1}^N (Q(\chi_{ij} \tau)) * (B_j \varphi)$, and setting $\varphi = \rho_B * \tau$ yields

$$QB_i \psi_0 = QB_i \psi_0 = \sum_{j=1}^N (Q(\chi_{ij} \tau)) * (B_j \rho_B) * \tau.$$

Since $B_j \rho_B \in L^\infty$ and $\tau, Q(\chi_{ij} \tau) \in C_\infty_c(G)$, we see that $QB_i\psi_0 \in L^\infty$. Because $Q$ was arbitrary, this shows that $\psi_0 \in \mathcal{E}$. \hfill \Box

4. Proof of Theorem 1.2

Given Theorem 1.1 and Lemma 3.1, we can obtain Theorem 1.2 by adapting essentially known arguments (compare, for example, [19, pp. 126–127]).
First, we need the following fact. For each $t_0 > 0$ there exist $c, b > 0$, depending on $t_0$, such that

$$K_{t_0}(g) \leq ce^{-b\psi_0(g)}$$

for all $g \in G$. This Gaussian estimate for a fixed time is known: see [17, Appendix A.4] and [11]. Actually, (8) could be deduced from the semigroup estimate of Theorem 1.1 by applying a local Harnack inequality: see [19, pp. 126–127] for details which could be adapted to our situation.

Next, fix a $\psi_0 \in E$ with properties as in Lemma 3.1, and for $t \in \mathbb{R}$ write

$$S_t^\lambda := e^{\lambda \psi_0} S_t e^{-\lambda \psi_0}, \quad \widetilde{S}_t^\lambda := \Delta^{1/2} S_t^\lambda \Delta^{-1/2}.$$

One sees from (1) that the integral kernel of the operator $\widetilde{S}_t^\lambda$ with respect to right Haar measure $d\bar{g} = \Delta(g^{-1})dg$ is given by

$$\widetilde{K}_t^\lambda(g, h) := e^{\lambda \psi_0(g)} \Delta^{1/2}(g)K_t(gh^{-1})\Delta^{-1/2}(h)e^{-\lambda \psi_0(h)}$$

for $g, h \in G$. Choosing $h = 1$, one derives an estimate of type

$$e^{\lambda \psi_0(g) - \psi_0(e)} \Delta^{1/2}(g)K_t(g) \leq \|\widetilde{S}_t^\lambda\|_{1 \rightarrow \infty}$$

$$\leq \|\widetilde{S}_t^{1/3}\|_{2 \rightarrow \infty} \|\widetilde{S}_t^{2/3}\|_{2 \rightarrow 2} \|\widetilde{S}_t^{1/3}\|_{1 \rightarrow 2}$$

$$\leq ce^{\omega t}$$

for all $t \geq 1$, $g \in G$ and $\lambda \in \mathbb{R}$, where $\| \cdot \|_{p \rightarrow q}$ denotes the norm from $L^p(G; d\bar{g})$ to $L^q(G; d\bar{g})$. Here we applied an estimate $\|\widetilde{S}_t^\lambda\|_{2 \rightarrow 2} \leq e^{\omega t}$, which follows from Theorem 1.1 since $\Delta^{1/2}: L^2(G; d\bar{g}) \rightarrow L^2(G; d\bar{g})$ is a unitary isomorphism, and estimates of type $\|\widetilde{S}_t^{1/3}\|_{2 \rightarrow \infty} + \|\widetilde{S}_t^{1/3}\|_{1 \rightarrow 2} \leq ce^{\omega t}$ which follow by a standard integration of (8) with $t_0 = 1/3$. Theorem 1.2 follows by optimizing over $\lambda$. □

5. Proof of Theorem 1.3

The proof of Theorem 1.3 in this section is inspired by ideas of [6], and is based on certain Nash type inequalities.

Let $G$ have polynomial growth of order $D \geq 1$. It is well known (cf. [19, p. 124]) that $G$ is unimodular.
We shall fix \( \psi_0 \in \mathcal{E} \) as in Lemma 3.1 and denote by \( U_\lambda \) the operator of pointwise multiplication by \( e^{\lambda \psi_0} \), for \( \lambda \in \mathbb{R} \). To prove Theorem 1.3, it suffices to establish for some \( \omega > 0 \) that
\[
\|U_\lambda K_t\|_2 \leq c t^{-D/2} e^{\omega \lambda^2 t}
\]
for all \( t \geq 1 \) and \( \lambda \geq 0 \). Indeed, since \( K_{2t} = K_t * K_t \) and applying (9) one gets
\[
e^{\lambda \psi_0(g)} K_{2t}(g) \leq \int_G dh \ e^{\lambda \psi_0(h)} K_t(h) e^{\lambda \psi_0(h^{-1} g)} K_t(h^{-1} g)
\leq \|U_\lambda K_t\|_2^2 \leq c t^{-D/2} e^{2\omega \lambda^2 t}
\]
for all \( g \in G, \lambda \geq 0, t \geq 1 \), and then Theorem 1.3 follows by optimizing over \( \lambda \).

We will derive (9) as a consequence of the following proposition, whose proof is temporarily deferred. For a locally integrable function \( f: G \to \mathbb{C} \), denote by \( L(f) \) the convolution operator \( L(f) f_1 := f * f_1 \), which is well defined at least on bounded, compactly supported functions \( f_1 \). Note that \( S_t = L(K_t) \).

**Proposition 5.1.** There exist constants \( k' > 1, r_0 \geq 1 \), such that
\[
\|U_\lambda K_1\|_2 \leq c e^{k' \lambda^2},
\]
\[
\text{Re}(U_\lambda H f, U_\lambda f) + k' \lambda^2 \|U_\lambda f\|_2^2 \geq \lambda^2 \|U_\lambda f\|_2^2,
\]
\[
\|U_\lambda S_t U_\lambda\|_{2 \to 2} \leq e^{k' \lambda^2 t}
\]
for all \( \lambda \geq 0, f \in C_c^\infty(G), t > 0 \), and
\[
\|U_\lambda f\|_2^2 \leq cr^2 \left\{ \text{Re}(U_\lambda H f, U_\lambda f) + k' \lambda^2 \|U_\lambda f\|_2^2 \right\}
+ cr^{-D} (\|U_\lambda L(f) U_\lambda\|_{2 \to 2})^2
\]\nfor all \( r \geq r_0, \lambda \geq 0 \) and \( f \in C_c^\infty(G) \).

We call inequality (10) a weighted convolution Nash inequality (cf. [6]). It differs essentially from standard Nash type inequalities (see for example [14]) in replacing the \( L^1 \) norm of \( f \) with the \( L^2 \) operator norm of the convolution operator \( L(f) \). This replacement allows one to avoid semigroup estimates in \( L^1 \) which occur in the use of standard Nash inequalities.

Let us show how to obtain Theorem 1.3 from Proposition 5.1. Fix \( k' \) as in the proposition, and define
\[
J_\lambda(t) = e^{-2k' \lambda^2 t} \|U_\lambda K_t\|_2^2
\]
for $t \geq 1$ and $\lambda \geq 0$. (That $J_{\lambda}(t)$ is finite follows from (8) or Theorem 1.2.) Because $(d/dt)K_t = -HK_t$, differentiation with respect to $t$ gives

$$J'_{\lambda}(t) = -2e^{-2k^2t}((U_{\lambda}HK_t, U_{\lambda}K_t) + k^2||U_{\lambda}K_t||_2^2) \leq 0$$

for all $t \geq 1$, $\lambda \geq 0$. The first estimate of Proposition 5.1 implies that sup\{$J_{\lambda}(1): \lambda \geq 0$\} is finite, and since $J'_{\lambda}(t) \leq 0$ it follows that

$$(11) \quad c_0 := \sup\{J_{\lambda}(t): t \geq 1, \lambda \geq 0\} < \infty.$$ 

In inequality (10) let us set $f = e^{-k^2t^2}K_t$ for $t \geq 1$, $\lambda \geq 0$, and observe that $\|U_{\lambda}L(f)U_{\lambda}\|_{2\to 2} \leq 1$ by the third estimate of Proposition 5.1. One gets

$$(12) \quad J_{\lambda}(t) \leq cr^2(-2^{-1}J'_{\lambda}(t)) + cr^{-D}$$

for all $t \geq 1$, $r \geq r_0$ and $\lambda \geq 0$. Now for each $t \geq 1$ and $\lambda \geq 0$ set $r = (\varepsilon J_{\lambda}(t))^{-1/D}$, where $\varepsilon > 0$ is a fixed constant. By choosing $\varepsilon$ sufficiently small, we may arrange that $r \geq r_0$ because of (11), and that $cr^{-D} \leq 2^{-1}J_{\lambda}(t)$ in the right side of (12). Then subtracting $2^{-1}J_{\lambda}(t)$ from both sides of (12) and rearranging yields, for some $c' > 0$,

$$J_{\lambda}(t)^{1+(2/D)} \leq -c'J'_{\lambda}(t)$$

for all $t \geq 1$ and $\lambda \geq 0$. Thus

$$(d/dt)[J_{\lambda}(t)^{-2/D}] \geq c'' > 0,$$

which implies, recalling (11), that

$$J_{\lambda}(t)^{-2/D} \geq J_{\lambda}(1)^{-2/D} + \int_1^t c'' \geq c_0^{-2/D} + c''(t-1) \geq ct$$

for all $t \geq 1$, $\lambda \geq 0$. That is, $J_{\lambda}(t) \leq c't^{-D/2}$ for all $t \geq 1$, $\lambda \geq 0$, which proves (9) and Theorem 1.3 follows.

It remains to prove Proposition 5.1. The first estimate of the proposition follows from Theorem 1.2, or just from (8), while the second and third estimates follow for sufficiently large $k^2$ from Theorem 1.1.

It is left to prove (10). In fact, it is enough to obtain inequality (10) in the case where $\lambda r \leq 1$. For given this case, we can assume (increasing $c$ if necessary) that $c > 1$; then by the second estimate of Proposition 5.1, we see that (10) holds automatically in the case where $\lambda r \geq 1$. 
Set \( V_0(r) = dq(\{g \in G : \psi_0(g) \leq r\}) \) for \( r > 0 \). Let \( w : G \to (0, \infty) \) be any continuous function satisfying \( w(g) = w(g^{-1}) \), \( g \in G \), and denote also by \( w \) the multiplication operator \( f \mapsto wf \). Setting \( \|w\|_{\infty,r} = \sup\{w(g) : g \in G, \psi_0(g) \leq r\} \), one has the “convolution Nash inequality”

\[
\|wf\|_2 \leq \sup_{g \in G, \psi_0(g) \leq r} \|w\partial_g f\|_2 + \|w\|_{\infty,r} V_0(r)^{-1/2} \|w^{-1}L(f)w\|_2 \to_2
\]

for all \( r > 0 \) and \( f \in C_c(G) \). See [6, Lemma 2.2] for the straightforward proof of this estimate, which is valid in a general setting of unimodular locally compact groups.

Since \( G \) has polynomial growth of order \( D \), there exists \( r_0 \geq 1 \) such that \( V_0(r) \geq c_r^D \) for all \( r \geq r_0 \). Setting \( w = e^{\lambda \psi_0} \) and squaring the last displayed inequality, we therefore obtain

\[
\|U_\lambda f\|_2^2 \leq 2 \sup_{\psi_0(g) \leq r} \|U_\lambda \partial_g f\|_2^2 + cr^{2\lambda - D}(\|U_\lambda L(f)U_\lambda\|_2 \to_2)^2
\]

for all \( r \geq r_0 \) and \( \lambda \geq 0 \). We will prove the \( L^2 \) estimate that

\[
(13) \quad \|U_\lambda \partial_g f\|_2^2 \leq c\psi_0(g)^2 Re(U_\lambda Hf, U_\lambda f) + c\psi_0(g)^2 2\lambda c^{\lambda \psi_0(g) \to 2} \|U_\lambda f\|_2^2
\]

uniformly for all \( g \in G, \lambda \geq 0 \) and \( f \in C_c^\infty(G) \). Inserting (13) in the preceding displayed inequality yields (10) for all \( r \geq r_0 \) and \( \lambda \geq 0 \) with \( \lambda r \leq 1 \), and Proposition 5.1 then follows.

To verify (13), first note the identity

\[
U_\lambda \partial_g f = \partial_g U_\lambda f + [1 - e^{\lambda \psi_0}] L(g) U_\lambda f.
\]

Applying (5), (6), and the fact that \( \rho \leq c\psi_0 \), we find that

\[
\|\partial_g U_\lambda f\|_2^2 \leq c\rho(g)^2 \sum_i \|A_i U_\lambda f\|_2^2
\]

\[
\leq c\psi_0(g)^2 \{Re(U_\lambda Hf, U_\lambda f) + c_2 \lambda^2 \|U_\lambda f\|_2^2\}
\]

\[
= c\psi_0(g)^2 \{Re(U_\lambda Hf, U_\lambda f) + c_2 \lambda^2 \|U_\lambda f\|_2^2\}.
\]

From (5) and \( \rho \leq c\psi_0 \), we have \( \|\partial_g \psi_0\|_\infty \leq c \rho(g) \sum_i \|A_i \psi_0\|_\infty \leq c\psi_0(g) \), and because \( |1 - e^s| \leq |s| e^{|s|}, s \in \mathbb{R}, \) then

\[
\|1 - e^{\lambda \partial_g \psi_0}\|_2 \leq 1 - e^{\lambda \partial_g \psi_0} \|U_\lambda f\|_2 \leq c\psi_0(g) \lambda c^{\lambda \psi_0(g) \to 2} \|U_\lambda f\|_2
\]

for all \( g \in G \) and \( \lambda \geq 0 \). Combination of these estimates establishes (13), and the proof of Proposition 5.1 and Theorem 1.3 is complete. \( \square \)
6. Appendix

In this appendix we give a proof of Proposition 2.1 by appealing to Lie structure theory (see [12]). Some similar results about \( g_1 \) are described in [18, Appendix].

Let \( G \) be any connected Lie group, and recall from Section 1 the notations \( G_0 = \left[ G, G \right], \pi_0: G \to G/G_0 \cong \mathbb{R}^{n_1} \times T^{n_2} \) and \( G_1 = \pi_0^{-1}(\{0\} \times T^{n_2}) \).

Let \( M \) be a maximal compact subgroup of \( G \); such subgroups exist (see [12, p. 186]). Because \( \pi_0(M) \) is a compact subgroup of \( G/G_0 \), it is contained in \( \{0\} \times T^{n_2} \), and hence \( M \subseteq G_1 \). We claim that

\[
G_1 = M[G, G].
\]

The proof of (14) requires several observations. Note that \( G_0 \) is a closed, connected normal subgroup of \( G_1 \), and \( G_1/\mathbb{G}_0 \cong T^{n_2} \) is compact and connected; since \( M \subseteq G_1 \) is maximal compact, then \( G_1 = MG_0 \) by applying [12, p. 186, Theorem 3.7]. Next, if \( K \) is an arbitrary compact subgroup of \( G \), the result of [12, pp. 180–181, Theorem 3.1] shows that \( gKg^{-1} \subseteq M \) for some \( g \in G \). Thus \( K \subseteq g^{-1}MG \subseteq M[G, G] \), proving that any compact subgroup of \( G \) is contained in the subgroup \( M[G, G] \). It follows that \( M[G, G] \) is closed in \( G \), for otherwise (see [12, pp. 191–192]) there would exist a one-parameter subgroup \( \{\gamma(t): t \in \mathbb{R}\} \) of \( M[G, G] \) whose closure is compact and not contained in \( M[G, G] \), a contradiction. Closedness of \( M[G, G] \) implies that \( G_0 = [G, G] \subseteq M[G, G] \), and because \( G_1 = MG_0 \), our claim (14) follows.

Passing to Lie algebras in (14) yields \( g_1 = m + [g, g] \) where \( m \) is the Lie algebra of \( M \) and the sum need not be direct. By a standard decomposition theorem for compact Lie groups (cf. [12, p. 144]), \( m = m_1 \oplus c \) where \( m_1, c \) are respectively semisimple and abelian ideals of \( m \) and \( C = \exp(c) \) is compact, connected and abelian. Since \( m_1 = [m_1, m_1] \subseteq [g, g] \), we see that \( g = c + [g, g] \) and the proof of Proposition 2.1 is complete.

Acknowledgements

This work was carried out with financial support from the Australian Research Council (ARC) Centre of Excellence for Mathematics and Statistics of Complex Systems (MASCOS).

The referee for an earlier version of this paper pointed out Varopoulos’ article [18] of which I was previously unaware.
References


School of Mathematics
The University of New South Wales
Sydney 2052
Australia
E-mail address: dungey@maths.unsw.edu.au

Primera versió rebuda el 15 de desembre de 2004,
daarrera versió rebuda el 22 d’abril de 2005.