Properties of random walks on discrete groups: Time regularity and off-diagonal estimates

Nick Dungey

Department of Mathematics, Macquarie University, NSW 2109 Australia

Received 14 March 2007
Available online 14 April 2007

Abstract

In this paper we study some properties of the convolution powers $K^{(n)} = K*K*\cdots*K$ of a probability density $K$ on a discrete group $G$, where $K$ is not assumed to be symmetric. If $K$ is centered, we show that the Markov operator $T$ associated with $K$ is analytic in $L^p(G)$ for $1 < p < \infty$, and prove Davies–Gaffney estimates in $L^2$ for the iterated operators $T^n$. This enables us to obtain Gaussian upper bounds for the convolution powers $K^{(n)}$. In case the group $G$ is amenable, we discover that the analyticity and Davies–Gaffney estimates hold if and only if $K$ is centered. We also estimate time and space differences, and use these to obtain a new proof of the Gaussian estimates with precise time decay in case $G$ has polynomial volume growth.

© 2007 Elsevier Masson SAS. All rights reserved.

MSC: 60B15; 60G50; 47D06

Résumé

Dans ce travail on étudie le comportement des puissances de convolution $K^{(n)} = K*K*\cdots*K$ d’une probabilité $K$ sur un groupe discret $G$, où $K$ n’est pas nécessairement symétrique. Si $K$ est centrée, on montre que l’opérateur de Markov $T$ associé à $K$ est analytique dans $L^p(G)$, $1 < p < \infty$, et l’on démontre des estimations de Davies–Gaffney dans $L^2$ pour les opérateurs $T^n$. Ceci permet d’obtenir des estimations gaussiennes des puissances $K^{(n)}$. Dans le cas où le groupe $G$ est moyennable, $T$ est analytique, ou bien les estimations de Davies–Gaffney ont lieu, si et seulement si $K$ est centrée. On étudie aussi les dérivées en temps et en espace des puissances $T^n$, et l’on obtient alors une nouvelle preuve des estimations gaussiennes précisées si $G$ est à croissance polynomiale.

© 2007 Elsevier Masson SAS. All rights reserved.

MSC: 60B15; 60G50; 47D06
1. Introduction and statement of results

The aim of this paper is to study some properties of non-symmetric random walks on a finitely generated discrete group $G$. We consider a probability density $K$ on $G$ and the random walk on $G$ governed by $K$. (For general background, see for example [20,22,23].) Under certain assumptions on $K$, we prove time regularity estimates and Gaussian off-diagonal estimates for this walk. While various authors have obtained similar estimates for symmetric random walks and Markov chains, our results make no symmetry assumption. Also, it is remarkable that our results apply to arbitrary discrete groups without any special assumptions on the group structure (though, as it turns out, the results are most interesting for discrete groups which are amenable).

To be more precise, recall (cf. [2,8]) that a linear operator $S \in L(X)$, where $X$ is a complex Banach space, is said to be analytic if there exists a $c > 0$ such that
\[
\|(I - S)S^n\| \leq cn^{-1}
\]
for all $n \in \mathbb{N}$. This notion is a discrete analogue of the usual notion of analyticity for a continuous time semigroup $(e^{-tA})_{t \geq 0}$ which corresponds to an estimate $\|Ae^{-tA}\| \leq ct^{-1}$, $t > 0$ (see, for example, [9, Section 2.5]). Our first main theorem is the fundamental result that the Markov operator $T$, associated with the density $K$ on $G$, is analytic in $L^2 = L^2(G)$ whenever $K$ is centered (the definition of "centered" is given below). In particular, this theorem provides a large and interesting class of examples of analytic operators which are not self-adjoint.

Our second main theorem gives Davies–Gaffney estimates, that is, $L^2$ off-diagonal estimates, for the iterated operators $T^n$ when $K$ is centered. This enables us to deduce Gaussian estimates for the $n$th convolution powers $K^{(n)}$ of $K$. Similar Gaussian estimates have been established for symmetric random walks in, for example, [4,15,21], but are apparently new for non-symmetric walks on general discrete groups.

We also discover that when the group $G$ is amenable, then the above properties, that is, analyticity of $T$ or the Davies–Gaffney estimates, hold if and only if the density $K$ is centered.

In the case where the discrete group $G$ has polynomial volume growth, Alexopoulos [1] derived comprehensive results for non-symmetric, centered random walks, including a parabolic Harnack inequality and precise upper and lower Gaussian estimates on $K^{(n)}$. His results imply analyticity and $L^p$ off-diagonal estimates on this special class of groups. However, his arguments are quite specific to groups of polynomial growth: for example, he uses the famous theorem of Gromov [14] that any such group contains a nilpotent subgroup of finite index, to show that analysis on $G$ is approximated by analysis on a nilpotent group. Our approach is necessarily more general in order to deal with groups which are not of polynomial growth. Moreover, in the polynomial growth case, our approach enables a new proof of the precise Gaussian estimates for $K^{(n)}$ which does not require Gromov’s theorem.

To state our results precisely, we fix some ideas and notation. Let $G$ be a finitely generated discrete group and let $dg$ be the counting measure on $G$. Let $K$ be a probability density on $G$, that is, $K : G \to [0, 1]$ and $\int_G dg \ K(g) = 1$. We will generally assume that:

1. The support $V := \{g \in G : K(g) > 0\}$ of $K$ is finite, and $G = \bigcup_{n=1}^{\infty} V^n$ where $V^n := \{g_1 \ldots g_n : g_1, \ldots, g_n \in V\}$.
2. $K(e) > 0$, or in other words $e \in V$, where $e$ is the identity of $G$. 

One says that $K$ is symmetric if $K(g) = K(g^{-1})$ for all $g \in G$, but we shall not assume this property.

The density $K$ determines a random walk on $G$, whose distribution after $n$ steps is given by the $n$th convolution power $K^{(n)} := K * K \cdots * K$. Here, in general the convolution of two functions $f_1, f_2$ on $G$ is defined by

$$(f_1 * f_2)(g) = \int_G dh f_1(h) f_2(h^{-1}g) = \int_G dh f_1(g h^{-1}) f_2(h)$$

for all $g \in G$. The Markov operator $T$ associated with $K$ is defined by

$$Tf = K * f$$

for any $f \in L^p := L^p(G; dg), 1 \leq p \leq \infty$. Then $\|T\|_{p \to q} \leq 1$ for all $p \in [1, \infty]$, where $\|\cdot\|_{p \to q}$ denotes the norm of a bounded linear operator from $L^p$ to $L^q$. Note that $T^n f = K^{(n)} * f$ for all $n \in \mathbb{N} = \{1, 2, 3, \ldots\}$.

The concept of centeredness is defined as follows (cf. [1, Section 1]). Let $G_0 := [G, G]$ be the commutator subgroup of $G$, and consider the canonical homomorphism $\pi_0 : G \to G/G_0$. Note that $G/G_0$ is a finitely generated abelian group, and therefore it can be written in the form $G/G_0 \cong \mathbb{Z}^q \times A$ where $q \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}$ and $A$ is a finite abelian group. Define $G_1 := \pi_0^{-1}([0] \times A)$. Then $G_1$ is a normal subgroup of $G$ with $[G, G] \subseteq G_1$, and $G/G_1$ is isomorphic to $\mathbb{Z}^q$. Consider the homomorphism $\pi_1 : G \to G/G_1 \cong \mathbb{Z}^q$ and let $\pi_1^{(j)} : G \to \mathbb{Z}$ be the $j$th component of $\pi_1$ for $j \in \{1, \ldots, q\}$. We say that $K$ is centered if the first order moments of the projection of $K$ onto $\mathbb{Z}^q$ vanish, that is, if

$$\int_G dg K(g) \pi_1^{(j)}(g) = 0$$

for all $j \in \{1, \ldots, q\}$. We quote from [1, Section 1] the following basic lemma which shows that a general density on $G$ is conjugate via a multiplicative function to a centered density.

**Lemma 1.1.** Let $K$ be a density on $G$ satisfying assumption (i) above. Suppose $K$ is not centered. Then there exist a centered density $K'$, a multiplicative function $\chi : G \to (0, \infty)$ (that is, $\chi(gh) = \chi(g)\chi(h)$ for all $g, h \in G$) and a constant $\delta \in (0, 1)$ such that

$$K(g) = \delta \chi(g) K'(g)$$

for all $g \in G$. Then, setting $T'f = K' * f$, we have the relations

$$K^{(n)}(g) = \delta^n \chi(g) K'^{(n)}(g), \quad T^n f = \delta^n \chi T^n (\chi^{-1} f),$$

for all $n \in \mathbb{N}$, $g \in G$, and $f \in L^p$, $1 \leq p \leq \infty$.

Because of Lemma 1.1, in the sequel we concentrate mainly on the study of centered densities. Note that in the rest of this paper, we shall always assume that the density $K : G \to [0, 1]$ satisfies assumptions (i) and (ii) above, except where otherwise stated.

Let $L = L_G$ be the left regular representation of $G$ which acts on functions $f : G \to \mathbb{C}$ by $(L(g)f)(h) = f(g^{-1}h)$, $g, h \in G$. Define the difference operators $\partial_g = L(g) - I$ for $g \in G$. By definition, the “discrete Laplacian” corresponding to $K$ is the operator

$$I - T = - \int_G dg K(g) \partial_g.$$  

(2)
It is not difficult to derive the following criterion: \( K \) is centered if and only if \( (I - T)\eta = 0 \) for every homomorphism \( \eta : G \to \mathbb{Z} \).

We now state our first main result.

**Theorem 1.2.** Let \( K \) be centered. Then \( T \) is analytic in \( L^2 \), that is, there exists \( c > 0 \) with \( \| (I - T)T^n \|_{2 \to 2} \leq cn^{-1} \) for all \( n \in \mathbb{N} \).

From Theorem 1.2 and some interpolation methods we deduce the following corollary (see Section 3 below for details).

**Corollary 1.3.** Let \( K \) be centered. Then \( T \) is analytic in \( L^p \) for each \( p \in (1, \infty) \).

Next we consider \( L^2 \) off-diagonal estimates for the operators \( T^n \). We need to define a distance on \( G \), as follows. Fix a finite set \( U \subseteq G \) which is symmetric \( (U = U^{-1}) \), and is such that \( e \in U \) and \( U \) generates \( G \). Thus \( G = \bigcup_{n=1}^{\infty} U^n \), and we define \( \rho : G \to \mathbb{N} \) by

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

We then define the “distance” between any subsets \( E \) and \( F \) of \( G \) as

\[
d(E,F) = \inf \{ \rho(gh^{-1}) : g \in E, h \in F \} \in \mathbb{N}.
\]

This distance is right invariant in the sense that \( d(Eg,Fg) = d(E,F), \ g \in G \), where \( Eg := \{ hg : h \in E \} \). Observe that \( \rho(g) = d(\{g\}, \{e\}) = \rho(g^{-1}) \) for all \( g \in G \).

Let \( \mathcal{E} \) be the set of all functions \( \psi : G \to \mathbb{R} \) such that \( \partial g\psi \in L^\infty \) for each \( g \in G \).

The operators \( \{ T_\lambda \}_{\lambda \in \mathbb{R}} \) defined by \( T_\lambda f = e^{\lambda \psi}T e^{-\lambda \psi}f \) may be called the “Davies perturbations” of the Markov operator \( T \), in analogy with the Davies perturbation of semigroups generated by differential operators (for the latter, see for example [10]).

**Theorem 1.4.** Let \( K \) be centered. Then for each \( \psi \in \mathcal{E} \), setting \( T_\lambda := e^{\lambda \psi}T e^{-\lambda \psi} \) for \( \lambda \in \mathbb{R} \), there exists \( \omega > 0 \) such that

\[
\| T_\lambda^n \|_{2 \to 2} \leq e^{\omega \lambda^2 n}
\]

for all \( n \in \mathbb{N} \) and \( \lambda \in \mathbb{R} \).

By applying Theorem 1.4 we will obtain the following estimates. Let \( \chi_E \) denote the characteristic function of a subset \( E \subseteq G \) (thus \( \chi_E(g) = 1 \) for \( g \in E \), \( \chi_E(g) = 0 \) for \( g \notin E \)), and also denote by \( \chi_E \) the operator of pointwise multiplication \( f \mapsto \chi_E f \).

**Theorem 1.5.** Let \( K \) be centered. Then there exist \( c, b > 0 \) such that

\[
\| \chi_E T^n \chi_F \|_{2 \to 2} \leq ce^{-bd(E,F)^2/n}
\]

for all \( n \in \mathbb{N} \) and all non-empty subsets \( E \) and \( F \) of \( G \), and

\[
\int_E dg \int_F dh K^{(n)}(gh^{-1}) \leq c|E|^{1/2}|F|^{1/2}e^{-bd(E,F)^2/n}
\]

for all \( n \in \mathbb{N} \) and all non-empty finite subsets \( E, F \) of \( G \), where \( |E| = dg(E) \) denotes the number of elements of \( E \).
We refer to the estimates of Theorems 1.4 and 1.5 as $L^2$ off-diagonal estimates or Davies–Gaffney estimates. Analogues of these estimates are well known for heat semigroups on manifolds (see for example [10,11,13]). An analogue of Theorem 1.4 for symmetric Markov chains occurs for example in [15], while a proof of versions of (3) and (4) for walks on graphs is given in [7]. These previous proofs do not work for non-symmetric walks and our approach to proving Theorems 1.4 and 1.5 is rather different.

By choosing $E = \{g\}$, $F = \{e\}$ in (4), we obtain the following general Gaussian upper bound.

**Corollary 1.6.** Let $K$ be centered. There exist $c, b > 0$ such that

$$K^n(g) \leq ce^{-bd(E,F)^2/n}$$

for all $n \in \mathbb{N}$ and $g \in G$.

Estimates analogous to (5) were given for symmetric walks in, for example, [4,7,21].

Let us also remark the following estimate in $L^p$, which follows immediately by interpolation between (3) and the obvious estimates $\|\chi_ET^n\chi_F\|_{1\to 1} \leq \|T^n\|_{1\to 1} \leq 1$, $\|\chi_ET^n\chi_F\|_{\infty\to \infty} \leq 1$.

**Corollary 1.7.** Let $K$ be centered and let $p \in (1, \infty)$. There exist $c = c(p)$, $b = b(p) > 0$ such that

$$\|\chi_ET^n\chi_F\|_{p\to p} \leq ce^{-bd(E,F)^2/n}$$

for all $n \in \mathbb{N}$ and non-empty $E, F \subseteq G$.

Theorems 1.2 and 1.4 can be regarded as the fundamental results of this paper. Let us explain an algebraic result which is crucial for their proofs and which is of independent interest. For $j \in \mathbb{N}$, let $D_j$ denote the linear span of all operators in $\mathcal{L}(L^2)$ of the form $\partial g_1 \ldots \partial g_k$ with $k \in \mathbb{N}$, $k \geq j$ and $g_1, \ldots, g_k \in G$. Then $D_j$ is an algebra of operators contained in $\mathcal{L}(L^2)$, and $D_1 \supseteq D_2 \supseteq D_3 \supseteq \cdots$. It is clear from (2) that the discrete Laplacian $I - T$ is always an element of $D_1$.

We have the following characterization of centered densities.

**Proposition 1.8.** The density $K$ is centered if and only if $I - T \in D_2$.

In the special case where $K$ is symmetric ($K(g) = K(g^{-1})$), the well known identity $I - T = 2^{-1} \int_G dgK(g)\partial g^{-1} \partial g$ immediately implies that $I - T \in D_2$. The proof that $I - T \in D_2$ for a general centered density is not so trivial.

When the group $G$ is amenable, the estimates of Theorems 1.2 and 1.4 hold if and only if $K$ is centered. More precisely, in Section 6 we will prove the following theorem which gives a number of other analytic conditions equivalent to centeredness. As one might expect, the properties of the operators $T^n$ are closely related to properties of the continuous time Markov semigroup $(e^{-t(I-T)})_{t \geq 0}$ generated by $I - T$. Since $\|T\|_{p\to p} \leq 1$, the semigroup $(e^{-t(I-T)})_{t \geq 0}$ is contractive in $L^p$ for all $p \in [1, \infty]$.

Recall that we always assume $K$ satisfies assumptions (i) and (ii).

**Theorem 1.9.** Suppose that $G$ is amenable. Then the following conditions (I) to (IX) are equivalent.

(I) $K$ is centered.
(II) \( I - T \) is an element of \( D_2 \).

(III) There exists \( c > 0 \) such that
\[
\left| \left( (I - T) f, f \right) \right| \leq c \sum_{u \in U} \| \partial_u f \|_2^2
\]
for all \( f \in L^2 \).

(IV) \( T \) is analytic in \( L^2 \).

(V) The operators \( (e^{-t(I-T)})_{t \geq 0} \) form a bounded analytic semigroup in \( L^2 \); that is, there exists \( c > 0 \) such that
\[
\left\| (I - T)e^{-t(I-T)} \right\|_{2 \rightarrow 2} \leq ct^{-1}
\]
for all \( t > 0 \).

(VI) For each \( \psi \in E \), setting \( T_\lambda = e^{\lambda \psi} Te^{-\lambda \psi} \), there exist \( c, \omega > 0 \) such that
\[
\| T_\lambda^n \|_{2 \rightarrow 2} \leq ce^{\omega \lambda^2 n}
\]
for all \( n \in \mathbb{N} \) and \( \lambda \in [-1, 1] \).

(VII) For each \( \psi \in E \), setting \( T_\lambda = e^{\lambda \psi} Te^{-\lambda \psi} \), there exist \( c, \omega > 0 \) such that
\[
\| e^{-t(I-T_\lambda)} \|_{2 \rightarrow 2} \leq ce^{\omega \lambda^2 t}
\]
for all \( t > 0 \) and \( \lambda \in [-1, 1] \).

(VIII) For each \( \psi \in E \), setting \( T_\lambda = e^{\lambda \psi} Te^{-\lambda \psi} \), there exists \( \omega > 0 \) such that
\[
\Re((I - T_\lambda) f, f) \geq -\omega \lambda^2 \| f \|_2^2
\]
for all \( f \in L^2 \) and \( \lambda \in [-1, 1] \).

(IX) For each \( \psi \in E \), setting \( T_\lambda = e^{\lambda \psi} Te^{-\lambda \psi} \), one has
\[
\| T_\lambda \|_{2 \rightarrow 2} = 1 + O(\lambda^2)
\]
for all \( \lambda \in [-1, 1] \).

In condition (IX) above, we used the standard notation \( O(\lambda^k) \), \( \lambda \in J \), to denote a function \( s = s(\lambda) \) which satisfies an estimate \( |s(\lambda)| \leq c|\lambda|^k \) for all \( \lambda \) in the interval \( J \subseteq \mathbb{R} \).

Analyticity of \( T \) means that the operators \( T^n \) have a certain regularity in time. The next result (which will be deduced from Theorem 1.2) is a form of spatial regularity.

**Theorem 1.10.** Let \( K \) be centered. There is a \( c > 0 \) such that
\[
\| \partial_g T^n \|_{2 \rightarrow 2} \leq c\rho(g)n^{-1/2}
\]
for all \( n \in \mathbb{N} \) and \( g \in G \).

In fact, one also has temporal and spatial regularity in \( L^2 \) for the perturbed operators \( T_\lambda \), as follows.

**Theorem 1.11.** Let \( K \) be centered, let \( \psi \in E \) and set \( T_\lambda = e^{\lambda \psi} Te^{-\lambda \psi} \). Then there exist \( c, \omega > 0 \) such that
\[
\left\| (I - T_\lambda) T^n \right\|_{2 \rightarrow 2} = \left\| e^{\lambda \psi}(I - T) T^n e^{-\lambda \psi} \right\|_{2 \rightarrow 2} \leq cn^{-1} e^{\omega \lambda^2 n}
\]
and
\[ \| e^{\lambda \psi} \partial_g T^n e^{-\lambda \psi} \|_{2 \to 2} \leq c \left( \rho(g)n^{-1/2} \right)e^{\omega \lambda^2 n} \]
for all \( n \in \mathbb{N} \), \( \lambda \in \mathbb{R} \) and \( g \in G \) satisfying \( \rho(g) \leq n^{1/2} \).

We will derive the first estimate of Theorem 1.11 from Theorem 1.2, by applying a theorem of Blunck [2] on perturbations of analytic operators.

As an important application of the estimates of Theorem 1.11, we will obtain the following theorem which improves on Corollary 1.6.

**Theorem 1.12.** Let \( K \) be centered, and suppose that there exist \( a, D > 0 \) with \( dg(U^n) \geq an^D \) for all \( n \in \mathbb{N} \). Then there are \( c, b > 0 \) such that
\[
K^{(n)}(g) \leq cn^{-D/2}e^{-b\rho(g)^2/n},
\]
\[
|\partial_h K^{(n)}(g)| \leq c \left( \rho(h)n^{-1/2} \right)n^{-D/2}e^{-b\rho(g)^2/n},
\]
\[
|K^{(n)}(g) - K^{(n+1)}(g)| \leq cn^{-1}n^{-D/2}e^{-b\rho(g)^2/n}
\]
for all \( n \in \mathbb{N} \), \( g \in G \), and \( h \in G \) satisfying \( \rho(h) \leq n^{1/2} \).

In the case where \( G \) has polynomial volume growth of order precisely \( D \) (that is, \( c^{-1}n^D \leq dg(U^n) \leq cn^D \) for all \( n \in \mathbb{N} \)), the estimates of Theorem 1.12 were established by Alexopoulos [1]. By appealing to Nash inequality methods of [12], we obtain a proof of Theorem 1.12 which is completely different and perhaps easier than the proofs of [1].

Let us make some further remarks about the above results.

- The basic results and ideas of this paper can be adapted for centered densities on other classes of locally compact groups such as Lie groups. The technical details are somewhat different and will be presented elsewhere. (In fact, since writing the present paper, the author has a preprint ‘Properties of centered random walks on locally compact groups and Lie groups’ (accepted for Rev. Mat. Iberoamericana); and a paper ‘Time regularity for random walks on locally compact groups’, Probab. Theory Relat. Fields 137 (2007) 429–442.)

- The question of analyticity in \( L^1 \), that is, whether Corollary 1.3 holds when \( p = 1 \), is not answered in general in this paper. Similarly, we do not know if the estimate of Corollary 1.7 holds when \( p = 1 \) (or equivalently, by duality, when \( p = \infty \)). In the particular case where \( G \) has polynomial growth, a standard integration of the Gaussian estimates of Theorem 1.12 shows that the above results do hold for \( p \in \{1, \infty\} \).

- For a density \( K \) on \( G \), it is well known that \( \| T \|_{2 \to 2} < 1 \) if and only if \( G \) is non-amenabe. For non-amenabe groups, our main results in \( L^2 \) (for example, analyticity of \( T \)) follow rather trivially from \( \| T \|_{2 \to 2} < 1 \). (The estimate of Theorem 1.4 follows trivially using Lemma 4.1.) Thus our \( L^2 \) results are of greatest value for amenable groups.

- One sometimes wants to consider random walks where assumption (ii) is not satisfied. An example is the simplest symmetric random walk on \( \mathbb{Z} \), which corresponds to the density
\[
K : \mathbb{Z} \to [0, 1], \quad K(1) = K(-1) = 1/2.
\]
In this example, because \( K^{(2n+1)}(0) = 0 \) for all \( n \in \mathbb{N} \), one may show that \( |K^{(2n)}(0) - K^{(2n+1)}(0)| = K^{(2n)}(0) \geq cn^{-1/2} \) for large \( n \), and from this it is not hard to deduce that the operator \( T \), with \( Tf := K * f \), is not analytic in \( L^2(\mathbb{Z}) \).
This example shows that assumption (ii) cannot be omitted from the hypotheses of Theorem 1.2. We shall see that assumption (ii) implies a condition on the spectrum of $T$ which is necessary for analyticity (see Lemma 3.2 below).

In cases of interest where $K(e) = 0$, one can usually find $n_0 \geq 2$ such that the density $K^{(n_0)}$ satisfies assumptions (i) and (ii), and then apply our results to $K^{(n_0)}$. In the special case where $K$ is symmetric, $T$ is a self-adjoint contraction in $L^2$ so that $T^2$ is self-adjoint with spectrum contained in $[0, 1]$, and the spectral theorem easily shows that $T^2$ is analytic (see also Theorem 3.1 below).

– The problem of obtaining uniform upper bounds of the form

$$\|K^{(n)}\|_{\infty} \leq M(n)$$

for all $n \in \mathbb{N}$ and with some function $M : \mathbb{N} \to (0, \infty)$, has been considered in many works. For general results and references to the literature see, for example, [19,20,22,23].

For example, it is known (see [22, Chapter VII]) that if the volume growth of $G$ satisfies $dg(U^n) \geq a e^{an\gamma}$ for some $a > 0$, $\gamma \in (0, 1]$ and all $n \in \mathbb{N}$, then (6) holds with $M(n) = c_1 e^{-c_2 n^\beta}$ where $c_i$ are positive constants and $\beta = \gamma/(2 + \gamma)$. This result is not optimal for all $G$: for example, there are solvable groups with exponential growth $dg(U^n) \geq a e^{an}$ for which $\|K^{(n)}\|_{\infty} \leq c_1 e^{-c_2 n^\beta}$ where $\beta > 1/3$ (see [19,20] and their references).

Note that while many authors study only the case of symmetric $K$, one can usually deduce uniform bounds for non-symmetric walks from the symmetric case: see Appendix A of this paper for details.

– If an estimate (6) holds and $K$ is centered, then interpolation with the bound of Corollary 1.6 yields that

$$K^{(n)}(g) \leq M(n)^{1-\varepsilon} e^{-b\rho(g)^2/n}$$

for any $\varepsilon \in (0, 1)$, $g \in G$ and $n \in \mathbb{N}$. For example, if (6) holds with $M(n) = c_1 e^{-c_2 n^\beta}$ where $\beta \in (0, 1]$, then we obtain $K^{(n)}(g) \leq c_3 e^{-c_4 n^\beta} e^{-c_4 \rho(g)^2/n}$ for some constants $c_3, c_4 > 0$. In particular, combining with the previous remark we obtain the following result. If $G$ has exponential growth $dg(U^n) \geq a e^{an}$, $n \in \mathbb{N}$, and $K$ is centered, then an estimate

$$K^{(n)}(g) \leq c e^{-b n^{1/3}} e^{-b \rho(g)^2/n}$$

holds for all $n \in \mathbb{N}$ and $g \in G$.

The paper is organized as follows. In Section 2 we prove Proposition 1.8, which is a key step in the proofs of Theorems 1.2 and 1.4 given in Sections 3 and 4. Sections 5 and 6 give the proofs of Theorems 1.5 and 1.9 respectively. Theorems 1.10, 1.11 and 1.12 are proved in Section 7. Finally, in Appendix A we give a general result on uniform bounds for non-symmetric random walks.

In general, $c, c', b$ and so on, will denote positive constants whose value may change from line to line when convenient.

2. Proof of Proposition 1.8

In this section we continue the notations of Section 1 and give the proof of Proposition 1.8. The proof requires some lemmas which are of independent interest. In particular, the first lemma below provides a simple characterization of the subgroup $G_1$ of $G$. 
Lemma 2.1. One has
\[
G_1 = \{ g \in G : \text{there exists } n \in \mathbb{N} \text{ with } g^n \in [G, G] \}.
\]

Proof. Let \( g \in G \) and suppose \( g^n \in [G, G] \) for some \( n \in \mathbb{N} \). Then \( (\pi_0(g))^n \) is the identity element of \( G/[G, G] \cong \mathbb{Z}^q \times A \), where \( A \) is a finite abelian group. It follows that \( \pi_0(g) \in \{0\} \times A \), in other words, \( g \in \pi_0^{-1}(\{0\} \times A) = G_1 \). Conversely, if \( g \in G_1 \) then we can reverse each step in the above argument to obtain that \( g^n \in [G, G] \) for some \( n \in \mathbb{N} \).

Lemma 2.2. For any \( n \in \mathbb{N} \) with \( n \geq 2 \) and \( g_1, \ldots, g_n \in G \), one has
\[
\partial g_1 \cdots g_n = \partial g_1 + \cdots + \partial g_n + \sum_{k=2}^{n} \sum_{i_1, \ldots, i_k} \partial g_{i_1} \cdots \partial g_{i_k},
\]
where the inner sum is over all \( i_1, \ldots, i_k \in \{1, \ldots, n\} \) with \( i_1 < i_2 < \cdots < i_k \). In particular, when \( n = 2 \),
\[
\partial g_1 g_2 = \partial g_1 + \partial g_2 + \partial g_1 \partial g_2.
\]
Moreover, \( \partial g^{-1} = -\partial g - \partial g \partial g^{-1} \) for all \( g \in G \). Finally, one has
\[
\partial g^k - k \partial g \in D_2
\]
for all \( g \in G \) and \( k \in \mathbb{Z} \).

Proof. The first identity of the lemma may be proved by induction on \( n \), beginning with the case \( n = 2 \). The identity for \( \partial g^{-1} \) follows from the case \( n = 2 \) upon setting \( g_1 = g, g_2 = g^{-1} \). The final statement of the lemma is easily deduced from the preceding identities.

For the proof of the next lemma, remark that the spaces \( D_j, j \in \mathbb{N} \), which were defined before Proposition 1.8, have the characterization
\[
D_j = \text{span}\{ L(g_0) \partial g_1 \cdots \partial g_j : g_0, g_1, \ldots, g_j \in G \}.
\]
Indeed, it is obvious that each element of \( D_j \) is expressible as a linear combination of terms \( L(g_0) \partial g_1 \cdots \partial g_j \). Since also
\[
L(g_0) \partial g_1 \cdots \partial g_j = \partial g_0 \partial g_1 \cdots \partial g_j + \partial g_1 \cdots \partial g_j \in D_j,
\]
our remark follows.

Lemma 2.3. For any \( g \in G_1 \), one has \( \partial g \in D_2 \).

Proof. Let us write \([g_1, g_2] = g_1 g_2 g_1^{-1} g_2^{-1}\) for \( g_1, g_2 \in G \). Direct computation yields the identity
\[
\partial [g_1, g_2] = -L(g_1 g_2) (\partial_{g_1^{-1} g_2^{-1}} - \partial_{g_1^{-1} g_2^{-1}}) \in D_2
\]
(actually, that \( \partial [g_1, g_2] \in D_2 \) can also be deduced from identities of Lemma 2.2). Since \([G, G]\) consists of all finite products of elements of form \([g_1, g_2]\), using Lemma 2.2 we see that \( \partial g \in D_2 \) for all \( g \in [G, G] \).

Next, suppose that \( g \in G_1 \). Then \( g^n \in [G, G] \) for some \( n \in \mathbb{N} \) by Lemma 2.1, so that \( \partial g^n \in D_2 \). Because \( \partial g^n - n \partial g \in D_2 \) by Lemma 2.2, we deduce that \( \partial g \in D_2 \).
Proof of Proposition 1.8. First suppose that \( K \) is centered. Recall the homomorphism \( \pi_1 : G \to G/G_1 \cong \mathbb{Z}^q \) and identify \( G/G_1 \) with \( \mathbb{Z}^q \). Let \( e_j = (\delta_{jk})_k^{q} \), \( j \in \{1, \ldots, q\} \), denote the standard basis vectors in \( \mathbb{Z}^q \), and fix elements \( x_1, \ldots, x_q \in G \) such that \( \pi_1(x_j) = e_j \) for all \( j \in \{1, \ldots, q\} \).

For each \( g \in G \) we can write
\[
g = x_1^{p_1} \cdots x_q^{p_q} g' \tag{7}
\]
where \( p_j := \pi_1^{(j)}(g) \in \mathbb{Z} \), \( j \in \{1, \ldots, q\} \), and \( g' \in G \). Observe that \( g' \in G_1 \), because \( G_1 \) is the kernel of the homomorphism \( \pi_1 \) and \( \pi_1(x_1^{p_1} \cdots x_q^{p_q}) = (p_1, \ldots, p_q) = \pi_1(g) \). By applying the identities of Lemma 2.2 to (7), and since \( \partial g' \in \mathcal{D}_2 \), it follows that
\[
\partial g = \pi_1^{(1)}(g) \partial x_1 + \cdots + \pi_1^{(q)}(g) \partial x_q + W_g
\]
where \( W_g \) is an element of \( \mathcal{D}_2 \) depending on \( g \). Since \( K \) is centered, from (1) and (2) we obtain
\[
I - T = - \int_G dg K(g) \partial g = - \int_G dg K(g) W_g \in \mathcal{D}_2,
\]
as required.

Conversely, let us suppose that \( I - T \in \mathcal{D}_2 \) and show that \( K \) must be centered. Let \( \eta : G \to \mathbb{Z} \) be any homomorphism. We have \( (\partial g_1, \eta)(h) = \eta(g_1^{-1} h) - \eta(h) = -\eta(g_1) \) for all \( h \in G \), so that
\[
\partial g_1, \eta = -\eta(g_1), \quad \partial g_2, \eta = 0,
\]
for all \( g_1, g_2 \in G \). This shows that \( W \eta = 0 \) for any \( W \in \mathcal{D}_2 \). Setting \( W = I - T \), we have
\[
((I - T)\eta)(e) = 0 \quad \text{so that}
\]
\[
0 = \eta(e) = (T\eta)(e) = \int_G dg K(g) \eta(g^{-1}) = - \int_G dg K(g) \eta(g).
\]

With \( \eta = \pi_1^{(j)} \), \( j \in \{1, \ldots, q\} \), we conclude that (1) holds and \( K \) is centered. This completes the proof of Proposition 1.8. \( \Box \)

3. Proof of Theorem 1.2

To prove Theorem 1.2 we will use the following general characterization of analytic operators due to Nevanlinna (see [17, Theorem 4.5.4], [18, Theorem 2.1], and [2]). Let \( \mathbb{D} := \{ z \in \mathbb{C}: |z| < 1 \} \) be the open unit disk.

**Theorem 3.1.** Let \( X \) be a complex Banach space and let \( S \in \mathcal{L}(X) \). The following two conditions are equivalent.

(I) \( S \) is power-bounded (that is, \( \sup \{ \| S^n \| : n \in \mathbb{N} \} < \infty \)) and analytic.

(II) \( (e^{-t(I-S)})_{t \geq 0} \) is a bounded analytic semigroup in \( X \), and the spectrum of \( S \) is contained in \( \mathbb{D} \cup \{1\} \).

We will verify condition (II) of Theorem 3.1 when \( S = T \) and \( X = L^2 \). The following lemma gives the desired condition on the spectrum of \( T \). Denote by \( \sigma_p(S) \) the spectrum of an operator \( S \in \mathcal{L}(L^p) \).
Lemma 3.2. Let $P: G \to [0, 1]$ be any probability density on $G$, and define the operator $S$ by $Sf = P \ast f$, $f \in L^p$. If $P(e) > 0$, then $\sigma_p(S)$ is contained in $\mathbb{D} \cup \{1\}$ for all $p \in [1, \infty]$.

Proof. Let $B = \{ g \in G : g \neq e \}$. We may write

$$S = P(e)I + \int_B dg P(g)L(g) = P(e)I + S',$$

where the operator $S'$ satisfies

$$\|S'\|_{p \to p} \leq \int_B dg P(g) = 1 - P(e).$$

Then $\sigma_p(S') \subseteq \{ z \in \mathbb{C} : |z| \leq 1 - P(e) \}$ and

$$\sigma_p(S) = P(e) + \sigma_p(S') \subseteq P(e) + \{ z \in \mathbb{C} : |z| \leq 1 - P(e) \} \subseteq \mathbb{D} \cup \{1\},$$

where the last inclusion used the hypothesis $P(e) > 0$. □

By Theorem 3.1 and Lemma 3.2, the proof of Theorem 1.2 is now reduced to showing that $e^{-t(I-T)}$ is a bounded analytic semigroup in $L^2$ when $K$ is centered. To show this, it suffices to prove an estimate

$$|((I-T)f, f)| \leq c \operatorname{Re}((I-T)f, f)$$

for all $f \in L^2$. Indeed, one has the following standard semigroup result (see for example [16, Theorem IX.1.24]), whose proof we sketch for the sake of completeness and for later use. For $\theta > 0$ define the open sector $\Lambda_\theta = \{ z \in \mathbb{C} : z \neq 0, \ |\arg z| < \theta \}$.

Lemma 3.3. Let $\mathcal{H}$ be a complex Hilbert space and $S \in \mathcal{L}(\mathcal{H})$. Suppose there exists $c_0 \geq 0$ such that

$$|\operatorname{Im}(Sf, f)| \leq c_0 \operatorname{Re}(Sf, f)$$

for all $f \in \mathcal{H}$. Then $e^{-tS}$ is a bounded analytic semigroup; more precisely, there exist constants $\theta_1 \in (0, \pi/2]$, $c_1 > 0$, depending only on $c_0$, such that

$$\|e^{-zS}\| \leq 1$$

for all $z \in \Lambda_{\theta_1}$ and $\|Se^{-tS}\| \leq c_1 t^{-1}$ for all $t > 0$.

Proof. Define $\theta_1 \in (0, \pi/2]$ by requiring that $\tan \theta_1 = c_0^{-1}$. Let $f \in \mathcal{H}$ and $\theta \in (-\theta_1, \theta_1)$, and set $\psi_t := e^{-t\theta}Sf$ for all $t > 0$. Differentiation with respect to $t$ yields

$$\frac{d}{dt}(\|\psi_t\|^2) = -2 \cos \theta \operatorname{Re}(S\psi_t, \psi_t) + 2 \sin \theta \operatorname{Im}(S\psi_t, \psi_t)$$

which is non-positive because $|\sin \theta| \leq (\cos \theta)c_0^{-1}$. Therefore $\|\psi_t\|^2 \leq \|f\|^2$ for all $t > 0$, which implies the desired estimate $\|e^{-zS}\| \leq 1$, $z \in \Lambda_{\theta_1}$. The bound on $Se^{-tS}$ follows by applying the Cauchy integral formula to the analytic function $z \mapsto e^{-zH}$ (compare for example [9, Lemma 2.38]). □
We turn to the verification of (8). It is convenient to define the “gradient” of a function $f : G \to \mathbb{C}$ by setting

$$(\nabla f)(g) := \left( \sum_{u \in U} (\partial uf)(g) \right)^{1/2}, \quad g \in G,$$

where $U = U^{-1} \subseteq G$ is the fixed finite generating set from Section 1. Observe that

$$\|\nabla f\|^2_2 = \sum_{u \in U} \|\partial uf\|^2_2$$

and

$$\|\partial g f\|_2 \leq \rho(g) \sup_{u \in U} \|\partial uf\|_2 \leq \rho(g) \|\nabla f\|^2_2$$

for all $g \in G$, $f \in L^2$. Inequality (9) is easily derived by writing $g = u_1 u_2 \ldots u_n$ with $u_j \in U$, $n = \rho(g)$, and observing that $\|\partial g f\|_2 \leq \sum_{j=1}^n \|\partial u_j f\|_2$.

The first step in the proof of (8) is the following observation.

**Lemma 3.4.** There exists a $c > 1$ such that $c^{-1}\|\nabla f\|^2_2 \leq \text{Re}((I - T)f, f) \leq c\|\nabla f\|^2_2$ for all $f \in L^2$.

**Proof.** Define $\hat{T} := 2^{-1}(T + T^*)$ where $T^*$ is the $L^2$-adjoint of $T$. Then $\hat{T}$ is self-adjoint and

$$\text{Re}((I - T)f, f) = 2^{-1}((I - T)f, f) + 2^{-1}(f, (I - T)f) = ((I - \hat{T})f, f)$$

for all $f \in L^2$. Observe that $\hat{T}f = \hat{K} \ast f$, $f \in L^2$, where $\hat{K}$ is the probability density defined by

$$\hat{K}(g) = 2^{-1}(K(g) + K(g^{-1})), \quad g \in G.$$

The support $V'$ of $\hat{K}$ contains the support $V$ of $K$, so that $V'$ generates $G$. The symmetry $\hat{K}(g) = \hat{K}(g^{-1})$ of $\hat{K}$ implies that

$$I - \hat{T} = - \sum_{g \in V'} \hat{K}(g) \partial g = 2^{-1} \sum_{g \in V'} \hat{K}(g) \partial g^{-1} \partial g.$$

Then

$$\text{Re}((I - T)f, f) = ((I - \hat{T})f, f) = 2^{-1} \sum_{g \in V'} \hat{K}(g) \|\partial g f\|^2_2,$$

and since $V'$ generates $G$ the lemma follows easily. □

**Remark.** Lemma 3.4 does not require $K$ to be centered; in fact the lemma holds for any density which satisfies assumption (i).

Let us continue the proof of (8). Let $K$ be centered. By Proposition 1.8, $((I - T)f, f)$ is expressible as a finite linear combination of terms each of the form

$$(\partial_{g_1} \ldots \partial_{g_k} f, f) = (\partial_{g_2} \ldots \partial_{g_k} f, \partial_{g_1^{-1}} f)$$

with $k \geq 2$ and $g_1, \ldots, g_k \in G$. By (9), each such term satisfies an estimate
\[ |(\partial_{g_2} \ldots \partial_{g_k} f, \partial_{g_1}^{-1} f)| \leq \|\partial_{g_2} \ldots \partial_{g_k} f\|_2 \|\partial_{g_1}^{-1} f\|_2 \]
\[ \leq 2^{k-2} \|\partial_{g_k} f\|_2 \|\partial_{g_1}^{-1} f\|_2 \leq c \|\nabla f\|_2^2, \]
and we conclude that
\[ |((I - T) f, f)| \leq c \|\nabla f\|_2^2 \tag{10} \]
for all \( f \in L^2 \). Together with Lemma 3.4, this yields (8). The proof of Theorem 1.2 is complete. \( \square \)

**Proof of Corollary 1.3.** The following interpolation theorem for analytic operators was proved by Blunck [3, Theorem 1.1].

**Theorem 3.5.** Let \( p, q \in [1, \infty] \) and let \( S \in \mathcal{L}(L^p) \) be power-bounded and analytic in \( L^p \). If \( S \) is power-bounded in \( L^q \) then \( S \) is power-bounded and analytic in \( L^r \) for all \( r \) strictly between \( p \) and \( q \).

Corollary 1.3 follows immediately from Theorems 1.2 and 3.5, since \( T \) is power-bounded in \( L^1 \) and in \( L^\infty \).

Alternatively, one may obtain Corollary 1.3 without invoking Theorem 3.5, as follows. The semigroup \((e^{-t(I-T)})_{t \geq 0}\) is bounded analytic in \( L^2 \) as we showed above, and it is also bounded in \( L^1 \) and in \( L^\infty \). By a standard application of the Stein interpolation theorem (see for example [10, Theorem 1.4.2]), it follows that \((e^{-t(I-T)})\) is bounded analytic in \( L^p \) for each \( p \in (1, \infty) \). Applying Theorem 3.1 and Lemma 3.2, we deduce that \( T \) is analytic in \( L^p \) for such \( p \). \( \square \)

**4. Proof of Theorem 1.4**

In this section we prove Theorem 1.4, and in the process derive other useful estimates involving the operators \( T_\lambda \).

Let us fix \( \psi \in \mathcal{E} \) and set \( T_\lambda = e^{\lambda \psi} T e^{-\lambda \psi} \) for \( \lambda \in \mathbb{R} \). We begin with a simple estimate of the difference \( T_\lambda - T \) which is valid whether \( K \) is centered or not.

**Lemma 4.1.** There exist \( c, \omega > 0 \) such that
\[ \|T_\lambda - T\|_{p \to p} \leq c|\lambda| e^{\omega |\lambda|} \]
for all \( \lambda \in \mathbb{R} \) and \( p \in [1, \infty] \).

**Proof.** Note the identities
\[ e^{\lambda \psi} L(g)(e^{-\lambda \psi} f) = e^{-\lambda \partial g \psi} L(g)f, \]
\[ e^{\lambda \psi} \partial g (e^{-\lambda \psi} f) = \partial g f + [e^{-\lambda \partial g \psi} - 1] L(g)f \tag{11} \]
for a function \( f : G \to \mathbb{C} \) and \( g \in G \). Since \( T = \int_G dg K(g)L(g) \), we find that
\[ (T_\lambda - T)f = \int_G dg K(g)[e^{-\lambda \partial g \psi} - 1] L(g)f. \]
By applying the inequality $|e^s - 1| \leq |s|e^{|s|}$, $s \in \mathbb{R}$, and noting an estimate $\|\partial_g \psi\|_\infty \leq c$ for all $g \in V$, where $V$ is the finite support of $K$, we have

$$\|e^{-\lambda \partial_g \psi} - 1\|_\infty \leq c|\lambda|e^{c|\lambda|}$$

for all $g \in V$ and $\lambda \in \mathbb{R}$. Then

$$\| (T_\lambda - T) f \|_p \leq c|\lambda|e^{c|\lambda|} \int_V dg K(g) \| L(g) f \|_p = c|\lambda|e^{c|\lambda|} \| f \|_p$$

for all $f \in L^p$, which proves the lemma. □

Since $\|T\|_{2 \rightarrow 2} \leq 1$, Lemma 4.1 implies that $\|T_\lambda\|_{2 \rightarrow 2} = 1 + O(|\lambda|)$ when $|\lambda|$ is small. To obtain Theorem 1.4, our main task will be to prove the improved estimate

$$\|T_\lambda\|_{2 \rightarrow 2} = 1 + O(\lambda^2) \quad (12)$$

for all $|\lambda| \leq 1$, when $K$ is centered. Indeed, if one combines (12) with Lemma 4.1, one obtains for some $\omega' > 0$ that $\|T_\lambda\|_{2 \rightarrow 2} \leq e^{\omega'/\lambda^2}$ for all $\lambda \in \mathbb{R}$. It then follows that $\|T_\lambda^n\|_{2 \rightarrow 2} \leq e^{\omega'/\lambda^2 n}$ for all $n \in \mathbb{N}$, which is the estimate of Theorem 1.4.

The main step in proving (12) is the following perturbation estimate. Define quadratic forms $Q$ and $Q_\lambda$, $\lambda \in \mathbb{R}$, by

$$Q(f) = \|f\|^2 - \|Tf\|^2, \quad Q_\lambda(f) = \|f\|^2 - \|T_\lambda f\|^2,$$

for all $f \in L^2$, so that $Q = Q_0$.

**Proposition 4.2.** Let $K$ be centered.

(I) There exists $c > 0$ such that

$$\left| (T_\lambda - T)(f_1, f_2) \right| \leq \varepsilon (\|\nabla f_1\|^2 + \|\nabla f_2\|^2) + c(1 + \varepsilon^{-1})\lambda^2 (\|f_1\|^2 + \|f_2\|^2)$$

for all $\varepsilon > 0$, $f_1, f_2 \in L^2$ and $|\lambda| \leq 1$.

(II) There exists $c' > 0$ such that

$$\left| Q(f) - Q_\lambda(f) \right| \leq \varepsilon \|\nabla f\|^2 + c'(1 + \varepsilon^{-1})\lambda^2 \|f\|^2$$

for all $\varepsilon > 0$, $f \in L^2$ and $|\lambda| \leq 1$.

**Proof.** To prove part (I), we use the fact that $I - T \in D_2$ by Proposition 1.8. It follows that

$$((T_\lambda - T)(f_1, f_2) = ((I - T)(f_1, f_2) - ((I - T_\lambda)(f_1, f_2)$$

is a finite linear combination of terms $t$ each of the form

$$t = (\partial_{g_1} \ldots \partial_{g_k} f_1, f_2) - (e^{\lambda \psi} \partial_{g_1} \ldots \partial_{g_k} e^{-\lambda \psi} f_1, f_2)$$

where $k \geq 2$ and $g_1, \ldots, g_k \in G$. By (11), each such term $t$ is a finite linear combination of terms $t'$ of the form

$$t' = (B_1(\lambda) \ldots B_k(\lambda)f_1, f_2),$$

where $k \geq 2$ and where for each $i$ either $B_i(\lambda) = \partial_{g_i}$ or $B_i(\lambda) = [e^{-\lambda \partial_{g_i}} - 1]L(g_i)$, with the second case occurring for at least one $i \in \{1, \ldots, k\}$. 


Let us consider three cases in (13). First, in case $B_1(\lambda) = [e^{-\lambda \partial g_i} \psi - 1] L(g_i)$ for at least two values of $i$, then since $|e^{-\lambda \partial g_i} \psi - 1| \to c|\lambda|$ for $|\lambda| \leq 1$, we easily obtain that

$$|t'| \leq c'\lambda^2 \|f_1\|_2 \|f_2\|_2 \leq 2^{-1}c'\lambda^2 \left(\|f_1\|_2^2 + \|f_2\|_2^2\right)$$

for all $|\lambda| \leq 1$. Secondly, if $B_1(\lambda) = [e^{-\lambda \partial g_i} \psi - 1] L(g_i)$ and $B_i = \partial g_i$ for all $i \in \{2, \ldots, k\}$, then using (9) gives

$$|t'| \leq c|\lambda| \|\nabla f_1\|_2 \|f_2\|_2 \leq 2^{-1}c(\varepsilon \|\nabla f_1\|_2^2 + \varepsilon^{-1}\lambda^2 \|f_2\|_2^2)$$

for all $|\lambda| \leq 1$ and $\varepsilon > 0$. Thirdly, in the remaining case one has $B_1 = \partial g_i$ and $B_i(\lambda) = [e^{-\lambda \partial g_i} \psi - 1] L(g_i)$ for some $i \in \{2, \ldots, k\}$, and by writing $t' = (B_2(\lambda) \ldots B_k(\lambda) f_1, \partial g_1 f_2)$ we get

$$|t'| \leq c|\lambda| \|f_1\|_2 \|\nabla f_2\|_2 \leq 2^{-1}c(\varepsilon \|\nabla f_2\|_2^2 + \varepsilon^{-1}\lambda^2 \|f_1\|_2^2)$$

for all $|\lambda| \leq 1$ and $\varepsilon > 0$. Part (I) follows by combining these three cases.

To prove part (II), write

$$Q(f) - Q_\lambda(f) = \|T_\lambda f\|_2^2 - \|T f\|_2^2 = (T f, (T_\lambda - T) f) + ((T_\lambda - T) f, T f) = 2 \text{Re}((T_\lambda - T) f, T f) + \|T_\lambda - T\| f_2^2.$$ 

Note that $\|T_\lambda - T\| f_2^2 \leq c\lambda^2 \|f\|_2^2$ for all $|\lambda| \leq 1$, thanks to Lemma 4.1. Also, part (I) gives

$$\|(T_\lambda - T) f, T f)\| \leq \varepsilon \left(\|\nabla f\|_2^2 + \|\nabla T f\|_2^2\right) + c(1 + \varepsilon^{-1})\lambda^2 \left(\|f\|_2^2 + \|T f\|_2^2\right)$$

$$\leq c_1 \varepsilon \|\nabla f\|_2^2 + 2c(1 + \varepsilon^{-1})\lambda^2 \|f\|_2^2$$

for all $|\lambda| \leq 1$ and $\varepsilon > 0$, since it is easy to see using (9) that $\|\nabla T f\|_2 \leq c\|\nabla f\|_2$ for all $f \in L^2$. By combining these inequalities we obtain part (II). □

Let us complete the proof of Theorem 1.4. First observe that one has an estimate

$$c^{-1}\|\nabla f\|_2^2 \leq Q(f) = \|f\|_2^2 - \|T f\|_2^2 \leq c\|\nabla f\|_2^2$$

(14)

for all $f \in L^2$. Indeed, $Q(f) = ((I - T^*T) f, f) = T^*T f = \tilde{K} f$, where $\tilde{K} := K^* K$ and $K^*(g) := K(g^{-1})$, $g \in G$. Note that $\tilde{K}$ is a symmetric probability density. Moreover, since $\tilde{K} \geq K^* K$ (where $K^* K$ is the convolution of $K$ with itself), so $\tilde{K}$ contains the support of $K$, so that $\tilde{V}$ generates $G$. Thus (14) follows from Lemma 3.4 applied to $T^* T$ in place of $T$.

Let $K$ be centered. From (14) and Proposition 4.2, part (II), there exists $c'' > 0$ such that

$$|Q(f) - Q_\lambda(f)| \leq c'' (1 + \varepsilon^{-1})\lambda^2 \|f\|_2^2$$

(15)

for all $f \in L^2$, $\varepsilon > 0$ and $|\lambda| \leq 1$. By choosing $\varepsilon < 1$ in (15), we see that there exists $\omega > 0$ such that

$$Q_\lambda(f) = \|f\|_2^2 - \|T_\lambda f\|_2^2 \leq -\omega \lambda^2 \|f\|_2^2,$$

or in other words $\|T_\lambda f\|_2^2 \leq (1 + \omega \lambda^2) \|f\|_2^2$ for all $f \in L^2$ and $|\lambda| \leq 1$. Thus $\|T_\lambda\|_{L^2 \to L^2} \leq (1 + \omega \lambda^2)^{1/2} \leq 1 + \omega \lambda^2$ for $|\lambda| \leq 1$. This proves (12), and Theorem 1.4 follows. □

We end the section with a remark which will be useful later.
Remark 4.3. In general, the constants in the estimates of this section depend on the choice of the function $\psi \in \mathcal{E}$. However, if $\mathcal{E}'$ is a subset of $\mathcal{E}$ which is uniformly bounded, in the sense that
$$\sup_{\psi \in \mathcal{E}'} \|\partial_g \psi\|_{\infty} < \infty$$
for each fixed $g \in G$, then the estimates hold uniformly for all $\psi \in \mathcal{E}'$. This remark follows easily from the above proofs. In particular, the bounds of Lemma 4.1, Proposition 4.2 and the estimate $\|e^{\lambda \psi} T^n e^{-\lambda \psi}\|_{2\rightarrow 2} \leq e^{\omega_2 n}$ of Theorem 1.4 hold uniformly for all $\psi \in \mathcal{E}'$.

5. Proof of Theorem 1.5

Let us prove the first estimate (3) of Theorem 1.5. We will assume that $d(E, F) \geq 2$, for otherwise the desired estimate follows trivially from the bound $\|\chi_E T^n \chi_F\|_{2\rightarrow 2} \leq \|T^n\|_{2\rightarrow 2} \leq 1$. For each non-empty set $F \subseteq G$, define $\psi_F : G \rightarrow \mathbb{N}$ by $\psi_F (g) = d(\{g\}, F) = \inf \{\rho(gh - 1) : h \in F\}$. The bound $|d(\{gh\}, F) - d(\{h\}, F)| \leq \rho(g) = \rho(g^{-1}), \quad g, h \in G,$
shows that $\|\partial_g \psi_F\|_{\infty} \leq \rho(g)$ for all $g \in G$. Therefore Theorem 1.4 and Remark 4.3 give an estimate
$$\|e^{\lambda \psi_F} T^n e^{-\lambda \psi_F}\|_{2\rightarrow 2} \leq e^{\omega_2 n}$$
for all $n \in \mathbb{N}$, $\lambda \in \mathbb{R}$, where $\omega > 0$ is a constant independent of $F$. Observe that $\psi_F (g) \geq d(E, F)$ for all $g \in E$, while $\psi_F (g) = 1$ for all $g \in F$. We find that
$$\|\chi_E T^n \chi_F\|_{2\rightarrow 2} = \|\chi_E e^{-\lambda \psi_F} \|_{2\rightarrow 2} \|e^{\lambda \psi_F} T^n e^{-\lambda \psi_F} \chi_F e^{\lambda \psi_F} \|_{2\rightarrow 2} \leq \exp(-\lambda d(E, F) + \lambda + \omega \lambda^2 n)$$
for all $n \in \mathbb{N}$, $\lambda \geq 0$ and all non-empty $E, F \subseteq G$ with $d(E, F) \geq 2$. Choosing $\lambda$ to be a small constant multiple of $d(E, F)/n$ then yields estimate (3).

Next, if $f_1, f_2 \in L^2$ are such that $f_1$ is supported in $F$ and $f_2$ is supported in $E$, then (3) implies the estimate
$$\left| \int_G dg \int_G dh K^{(n)}(gh^{-1}) f_1(h) \overline{f_2(g)} \right| = \left| (T^n f_1, f_2) \right|$$
$$\leq c \\|f_1\|_2 \\|f_2\|_2 e^{-bd(E, F)\sqrt{n}}$$
for all $n \in \mathbb{N}$. By taking $f_1 = \chi_F$, $f_2 = \chi_E$, we obtain (4), and the proof of Theorem 1.5 is complete.

6. Proof of Theorem 1.9

To prove the equivalence of conditions (I) to (IX) in Theorem 1.9, we will establish the chains of implications (I)$\Rightarrow$(II)$\Rightarrow$(III)$\Rightarrow$(IV)$\Rightarrow$(V)$\Rightarrow$(I), (I)$\Rightarrow$(IX)$\Rightarrow$(VI)$\Rightarrow$(VII)$\Rightarrow$(I), and (IX)$\Rightarrow$(VII)$\Rightarrow$(VIII)$\Rightarrow$(VII).
The implication (I)⇒(II) is just Proposition 1.8. Since condition (III) is precisely the inequality (10), the implication (II)⇒(III) is contained in the proof of Theorem 1.2. The implication (III)⇒(IV) also follows from the proof of Theorem 1.2. That (IV) implies (V) is just a special case of Theorem 3.1.

The implications (I)⇒(IX) and (IX)⇒(VI) are contained in the proof of Theorem 1.4. The implication (VI)⇒(VII) is derived by writing \( e^{-t(I - T)} = e^{-t} \sum_{n=0}^\infty (n!)^{-1} t^n T_\lambda^n \), so that

\[
\| e^{-t(I - T)} \|_{2 \to 2} \leq c e^{-t} \sum_{n=0}^\infty (n!)^{-1} t^n e^{\omega t^2} \\
= c \exp(-t + t e^{\omega t^2}) \\
\leq c \exp(e^{\omega^2 t})
\]

for all \( t > 0 \) and \( |\lambda| \leq 1 \). In the last step we used an estimate \( e^{\omega t^2} - 1 \leq c' \omega^2 \) for all \( |\lambda| \leq 1 \).

If (IX) holds, then for some \( \omega > 0 \) one has \( \text{Re}(T_\lambda f, f) \leq (1 + \omega \lambda^2) \| f \|_2^2 \) for all \( f \in L^2 \) and \( |\lambda| \leq 1 \), and this inequality rearranges to give (VIII). That (VIII) implies (VII) is standard semigroup theory: given (VIII), then

\[
(d/dt)(\| e^{-t(I - T)} \|_2^2) = -2 \text{Re}((I - T)e^{-t(I - T)} f, e^{-t(I - T)} f) \\
\leq 2 \omega \lambda^2 \| e^{-t(I - T)} f \|_2^2
\]

for all \( t > 0 \) and \( f \in L^2 \). This differential inequality implies that \( \| e^{-t(I - T)} f \|_2^2 \leq e^{2 \omega^2 t^2} \| f \|_2^2 \), which yields (VII).

It only remains to prove the implications (V)⇒(I) and (VII)⇒(I), and these are immediate consequences of the following theorem.

**Theorem 6.1.** Let \( G \) be amenable and suppose that \( K \) is not centered. Then for some \( c > 0 \),

\[
\| (I - T)e^{-t(I - T)} \|_{2 \to 2} \geq ct^{-1/2}
\]

for all \( t \geq 1 \). Thus the semigroup \( e^{-t(I - T)} \) is not bounded analytic. Moreover, there exists a homomorphism \( \Phi : G \to \mathbb{Z} \) (that is, \( \Phi(gh) = \Phi(g) + \Phi(h) \)), \( g, h \in G \) and constants \( \alpha > 0 \), \( \varepsilon > 0 \) such that \( T_\lambda := e^{\alpha \Phi T} e^{-\lambda \Phi} \) satisfies

\[
\| e^{-t(I - T_\lambda)} \|_{2 \to 2} \geq e^{\alpha t}
\]

for all \( t > 0 \) and \( \lambda \in [0, \varepsilon] \).

To prove Theorem 6.1 we first analyze, using Fourier theory, the special case where \( G = \mathbb{Z} \).

**Lemma 6.2.** Let \( P : \mathbb{Z} \to [0, 1] \) be a finitely supported probability density on \( \mathbb{Z} \) which is not centered, that is, \( a := \sum_{n \in \mathbb{Z}} n P(n) \neq 0 \). Define \( Sf := P * f \) for all \( f \in L^2 = L^2(\mathbb{Z}) \). Then there is \( c > 0 \) such that

\[
\| (I - S)e^{-t(I - S)} \|_{2 \to 2} \geq ct^{-1/2}
\]

for all \( t \geq 1 \). Moreover, there exists a homomorphism \( \psi : \mathbb{Z} \to \mathbb{Z} \) and constants \( \alpha > 0 \), \( \varepsilon \in (0, 1) \) such that \( S_\lambda := e^{\alpha \psi} S e^{-\lambda \psi} \) satisfies

\[
\| e^{-t(I - S_\lambda)} \|_{2 \to 2} \geq e^{\alpha t}
\]

for all \( t > 0 \) and \( \lambda \in [0, \varepsilon] \).
Proof of Lemma 6.2. For $f \in L^2 = L^2(\mathbb{Z})$ consider the $\mathbb{Z}$-Fourier transform

$$(\mathcal{F}f)(x) := \sum_{n \in \mathbb{Z}} f(n) e^{inx}, \quad x \in [-\pi, \pi],$$

which defines a unitary isomorphism $\mathcal{F} : f \mapsto \mathcal{F}f$ of $L^2(\mathbb{Z})$ onto the space $L^2([-\pi, \pi]; dx/(2\pi))$. Set $F := \mathcal{F}P$ and observe that $(\mathcal{F}Sf)(x) = F(x)(\mathcal{F}f)(x), x \in [-\pi, \pi]$, that is, the operator $S$ corresponds under $\mathcal{F}$ to multiplication by $F$. It follows that

$$\| (I - S)e^{-t(I - S)} \|_{2 \to 2} = \sup_{x \in [-\pi, \pi]} \left| (1 - F(x))e^{-t(1 - F(x))} \right| = \sup_{x \in [-\pi, \pi]} \left| 1 - F(x) \right| \exp(-t\Re(1 - F(x)))$$

for each $t > 0$. Note that $F(0) = 1, F'(0) = i\sum_n n P(n) = ia$ and that $F''(0) = -\sum_n n^2 P(n) = -b$ where $b > 0$. Thus the Taylor expansion of $F$ about 0 has the form $F(x) = 1 + iax - 2^{-1}bx^2 + O(x^3)$. Since $0 \neq a \in \mathbb{R}$, we deduce that there are constants $\delta \in (0, \pi)$ and $c_1, c_2 > 0$ such that

$$\left| 1 - F(x) \right| \geq c_1|x|, \quad \Re(1 - F(x)) \leq c_2x^2,$$

for all $x \in [-\delta, \delta]$. Then

$$\| (I - S)e^{-t(I - S)} \|_{2 \to 2} \geq \sup_{x \in [-\delta, \delta]} c_1|x| \exp(-tc_2x^2) \geq c_3t^{-1/2}$$

for all $t \geq 1$, where the last inequality follows by taking $x = \delta t^{-1/2}$.

To prove the second estimate of the lemma, let us assume that $a > 0$ and set $\psi(n) = n$ for all $n \in \mathbb{Z}$ (the proof in the case $a < 0$ is similar and is left to the reader). A simple calculation yields that

$$S_\lambda f = e^{\lambda \psi} S e^{-\lambda \psi} f = P_\lambda \ast f$$

for all $f \in L^2, \lambda \in \mathbb{R}$, where $P_\lambda(n) := e^{\lambda n} P(n), n \in \mathbb{Z}$. Thus $S_\lambda$ corresponds under $\mathcal{F}$ to multiplication by the function $F_\lambda(x) := \sum_n P(n)e^{\lambda n}$. Differentiating $F_\lambda(0) = \sum_n P(n)e^{\lambda n}$ with respect to $\lambda$ gives the expansion

$$F_\lambda(0) = 1 + a\lambda + 2^{-1}b\lambda^2 + O(|\lambda|^3)$$

for small $|\lambda|$, where $a$ and $b$ are as above. Hence, there is a $\varepsilon \in (0, \pi)$ such that $\Re(1 - F_\lambda(0)) = 1 - F_\lambda(0) \leq -a\lambda$ for all $\lambda \in [0, \varepsilon]$. We obtain that

$$\| e^{-t(I - S_\lambda)} \|_{2 \to 2} = \sup_{x \in [-\pi, \pi]} e^{-t\Re(1 - F_\lambda(x))} \geq e^{-t(1 - F_\lambda(0))} \geq e^{\alpha \lambda t}$$

for all $t > 0$ and $\lambda \in [0, \varepsilon]$. Lemma 6.2 is proved. \qed

Proof of Theorem 6.1. We will deduce Theorem 6.1 from Lemma 6.2 by a transference method. Suppose $K$ is not centered, and fix $j \in \{1, \ldots, q\}$ such that $\int_G dg K(g)\pi_1^{(j)}(g) \neq 0$. Let $P$ be the image under $\pi_1^{(j)}$ of $K$, that is, $P : \mathbb{Z} \to [0, 1]$ is the probability density which satisfies

$$\int_{\mathbb{Z}} \sum_{n \in \mathbb{Z}} P(n)f(n) = \int_G dg K(g) f(\pi_1^{(j)}(g))$$
for all \( f \in L^\infty(\mathbb{Z}) \). Then \( \sum_{n \in \mathbb{Z}} n P(n) \neq 0 \). Set \( Sf = P \ast f \) for \( f \in L^2(\mathbb{Z}) \). Since \( G \) is amenable, a standard transference theorem (see [5, Theorem 2.4]) yields that
\[
\| (I - S)e^{-t(I - S)} \|_{2 \rightarrow 2} \leq \| (I - T)e^{-t(I - T)} \|_{2 \rightarrow 2}
\]
for each \( t > 0 \). Thus the first statement of Theorem 6.1 follows by Lemma 6.2.

Next, let the homomorphism \( \psi : \mathbb{Z} \to \mathbb{Z} \) be as in Lemma 6.2. We define a homomorphism \( \Phi := \psi \circ \pi^{(1)} : G \to \mathbb{Z} \). Since \( T_\lambda := e^{\lambda \Phi} Te^{-\lambda \Phi} \) is a convolution operator on \( G \) (indeed, \( T_\lambda f = (e^{\lambda \Phi} K) * f \) for \( f \in L^2(G) \)), we may again apply the transference theorem of [5] to obtain
\[
\| e^{-t(I - S_\lambda)} \|_{2 \rightarrow 2} \leq \| e^{-t(I - T_\lambda)} \|_{2 \rightarrow 2}
\]
for all \( t > 0 \) and \( \lambda \in \mathbb{R} \). Hence the second part of Theorem 6.1 also follows from Lemma 6.2. This ends the proof of Theorems 6.1 and 1.9. \( \Box \)

7. Further estimates

In this section we prove Theorems 1.10, 1.11 and 1.12. We assume throughout the section that \( K \) is centered.

Proof of Theorem 1.10. Lemma 3.4 and Theorem 1.2 yield a bound
\[
\| \nabla T^n f \|^2 \leq c \text{Re}((I - T)T^n f, T^n f) \leq c' n^{-1} \| f \|^2
\]
for all \( f \in L^2 \) and \( n \in \mathbb{N} \). Since \( \| \partial_g T^n f \|_2 \leq \rho(g)\| \nabla T^n f \|_2 \) by (9), the estimate of Theorem 1.10 follows. \( \Box \)

Proof of Theorem 1.11. We will deduce the first estimate of the theorem from the analyticity of \( T \) by using the following theorem of Blunck [2] on perturbation of analytic operators. Denote by \( \sigma(S) \) the spectrum of a linear operator \( S \), and recall from Section 3 the notation \( \Lambda_\delta = \{ z \in \mathbb{C} : z \neq 0, |\arg z| < \delta \} \) for \( \delta > 0 \).

Theorem 7.1. (See [2].) Let \( X \) be a complex Banach space and let \( S \in \mathcal{L}(X) \) be power-bounded and analytic. Fix \( \alpha_0 \in \mathbb{C} \setminus \sigma(S) \), \( C > 0 \), and \( \delta \in (\pi/2, \pi) \). Then there exist \( \beta_0, r, M > 0 \) such that for all \( W \in \mathcal{L}(X) \) with \( \| W \| \leq C \) and all \( \beta \in [0, \beta_0] \), if one has
\[
\| \alpha(I - W + \beta I + \alpha I)^{-1} \| \leq C
\]
for all \( \alpha \in \Lambda_\delta \) and
\[
(\alpha_0 I - S)^{-1} - (\alpha_0 I - W)^{-1} \leq r,
\]
then \( \| (I - W)W^n \| \leq M(\beta + n^{-1})e^{\beta n} \) for all \( n \in \mathbb{N} \).

According to standard semigroup theory (see for example [9, Section 2.5]), estimate (16) means that the semigroup \( (e^{-t(I - W + \beta I)})_{t \geq 0} = (e^{-\beta t} e^{-t(I - W)})_{t \geq 0} \) is bounded analytic. Thus the essential content of Theorem 7.1 is that the analyticity of the semigroup \( e^{-t(I - W)} \) controls the analyticity of the operator \( W \) when \( W \) is a small perturbation of \( S \).

We will verify the hypotheses of Theorem 7.1 when \( X = L^2(G) \), \( S = T \) and \( W = T_\lambda \) for sufficiently small \( \lambda \). Applying Lemma 3.4 and Proposition 4.2, part (I), we obtain estimates of the form
\[
\| \nabla f \|^2 \leq c' \text{Re}((I - T)f, f) \leq c \text{Re}((I - T_\lambda)f, f) + c \lambda^2 \| f \|^2
\]
and, using (10),
$$
\left| (I - T\lambda)f, f \right| \leq c \left| (I - T)f, f \right| + c\lambda^2 \| f \|_2^2 \leq c' \| \nabla f \|_2^2 + c\lambda^2 \| f \|_2^2
$$
for all \( f \in L^2 \) and \( \lambda \in \mathbb{R} \) with \( |\lambda| \leq 1 \). Therefore, one can find constants \( c, \omega_1 > 0 \) large enough so that
$$
\left| (I - T\lambda + \omega_1\lambda^2 I)f, f \right| \leq c \text{Re}(\langle (I - T\lambda + \omega_1\lambda^2 I)f, f \rangle)
$$
for all \( f \in L^2 \) and \( |\lambda| \leq 1 \). By Theorem 1.4, there exists \( \omega > 0 \).

**Proof of Theorem 1.12.**

Theorem 1.12 follows from the \( L^2 \) off-diagonal estimates of Theorems 1.4 and 1.11, by applying a general theorem of the author [12] on Gaussian estimates for convolution powers on groups. The details are as follows.

Setting \( \psi = \rho \in \mathcal{E} \) in Theorems 1.4 and 1.11 gives estimates of the form
$$
\| e^{\lambda T} e^{-\lambda\rho} \|_{2 \to 2} \leq e^{2\alpha^2\lambda^2 n}, \quad \| e^{\lambda\rho} \partial_g T^n e^{-\lambda\rho} \|_{2 \to 2} \leq c\rho(g)n^{-1/2} e^{\alpha^2\lambda^2 n}
$$
for all \( n \in \mathbb{N}, \lambda \in \mathbb{R} \), and \( g \in G \) with \( \rho(g) \leq n^{1/2} \). Also, it follows easily from the bound \( K(g) \leq ce^{-b\rho(g)^2}, g \in G \), or just from the fact that \( K \) has finite support, that there exist \( c', \omega' > 0 \) with
$$
\| e^{\lambda T} K \|_2 \leq c' e^{\omega'\lambda^2}
$$
for all \( \lambda \geq 0 \).
By (18), (19) and the hypothesis that $dg(U^n) \geq an^D$ for all $n \in \mathbb{N}$, we may apply Theorem 2.3 of [12] to obtain immediately the first two estimates of Theorem 1.12. Note that the proof in [12] also yields, for some $c, \omega > 0$, the bound
\[ \|e^{\lambda \rho} K^{(n)}\|_2 \leq cn^{-D/4} e^{o\omega^2n} \] (20)
for all $n \in \mathbb{N}$ and $\lambda \geq 0$.

To complete the proof of Theorem 1.12 it remains to establish the Gaussian estimate on $K(n) - K(n+1)$. From the identities $K(n+m) = T^n(K(m))$, $K(n+m) - K(n+m+1) = (I - T) T^n(K(m))$, we have
\[ \|e^{\lambda \rho}(K^{(n+m)} - K^{(n+m+1)})\|_2 \leq \|e^{-\lambda \rho} (I - T) T^n e^{-\lambda \rho}\|_{2 \to 2} \|e^{\lambda \rho} K^{(m)}\|_2 \]
for all $n, m \in \mathbb{N}$. Then choosing $n = m$ or $n = m + 1$, and applying (20) and Theorem 1.11, we get for some $c, \omega > 0$ that
\[ \|e^{\lambda \rho}(K^{(n)} - K^{(n+1)})\|_2 \leq cn^{-1}n^{-D/4} e^{o\omega^2n} \] (21)
for all $n \in \mathbb{N}$ with $n \geq 2$ and $\lambda \geq 0$. The convolution identity $K^{(n+m)} - K^{(n+m+1)} = (K^{(n)} - K^{(n+1)}) * K^{(m)}$ implies that
\[ e^{\lambda \rho}(g) \left| K^{(n+m)}(g) - K^{(n+m+1)}(g) \right| \]
\[ \leq \int_G dh \ e^{\lambda \rho(h)} \left| K^{(n)}(h) - K^{(n+1)}(h) \right| e^{\lambda \rho(h^{-1}g)} K^{(m)}(h^{-1}g) \]
\[ \leq \left\| e^{\lambda \rho}(K^{(n)} - K^{(n+1)})\right\|_2 \left\| e^{\lambda \rho} K^{(m)}\right\|_2 \]
for all $n, m \in \mathbb{N}$, $\lambda \geq 0$ and $g \in G$. Hence, using (20) and (21) we deduce a bound of the form
\[ e^{\lambda \rho}(g) \left| K^{(n)}(g) - K^{(n+1)}(g) \right| \leq c'n^{-1}n^{-D/2} e^{o\omega^2n} \]
for all $n \geq 3$, $\lambda \geq 0$ and $g \in G$. Choosing $\lambda$ to be a small constant multiple of $\rho(g)/n$ gives the desired Gaussian estimate on $K^{(n)} - K^{(n+1)}$. \qed

**Acknowledgements**

During the writing of this paper in 2004–2005, the author was located at the University of New South Wales and was financially supported by the Australian Research Council (ARC) Centre of Excellence for Mathematics and Statistics of Complex Systems (MASCOS).

**Appendix A**

In this appendix, we observe rather generally that uniform bounds for non-symmetric random walks on $G$ can be deduced from uniform bounds for a symmetric random walk. This is essentially a consequence of Nash inequality methods; here we will follow Coulhon [6] who formulates these methods for discrete time semigroups on measure spaces.

A linear operator $T$ acting in the spaces $L^p(X; d\mu)$, where $(X, d\mu)$ is a $\sigma$-finite measure space, is said to be regularizing if $\|T\|_{p \to q} < \infty$ for all $1 \leq p \leq q \leq \infty$; this terminology is taken from [6,8]. If $d\mu$ is counting measure on a countable set $X$, then $L^p \subseteq L^q$ for $p \leq q$ so that $T$ is regularizing whenever $T$ is bounded in $L^p$ for all $p$. 
A function $M : (0, \infty) \to (0, \infty)$ is said to be of class $(\tilde{D})$ if $M$ is one-to-one, decreasing, $C^1$, one has $\inf \{M'(n+1)/M'(n): n \in \mathbb{N} \} > 0$, and the function $m := \log M$ is convex and satisfies $-m'(u) \geq -\alpha m'(t)$ for some $\alpha > 0$, all $t > 0$ and all $u \in [t, 2t]$. For example, the functions $t \mapsto c_1 t^{d/2}$ and $t \mapsto c_1 e^{-c_2 t^\beta}$ are of class $(\tilde{D})$, for any constants $c_1, c_2, d > 0$ and $\beta \in (0, 1)$.

From [6] one extracts:

**Theorem A.1.** Let $(X, d\mu)$ be a $\sigma$-finite measure space and $L^p := L^p(X ; d\mu)$. Let $T_1, T_2$ be regularizing operators satisfying $\|T_i\|_{p \to p} \leq 1$ for all $p \in [1, \infty]$, $i = 1, 2$. Suppose $T_1$ is self-adjoint in $L^2$ and that $M$ is a function of class $(\tilde{D})$ with $\|T_n^i\|_1 \to \infty \leq M(n)$ for all $n \in \mathbb{N}$. If there exists $c > 0$ with

$$\|f\|_2^2 - \|T_1 f\|_2^2 \leq c (\|f\|_2^2 - \|T_2 f\|_2^2)$$

for all $f \in L^2$, then there are $c_1, c_2 > 0$ such that

$$\|T_2 n\|_{1 \to \infty} \leq c_1 M(c_2 n)$$

for all $n \in \mathbb{N}$.

**Proof.** It follows by combining Propositions IV.1, IV.2 and the proof of Theorem IV.3 of [6]. See also the proof of Proposition IV.4 of [6].

As a consequence we obtain:

**Theorem A.2.** Let $K_1, K_2 : G \to [0, 1]$ be probability densities on the finitely generated discrete group $G$, whose supports $V_i := \{g \in G: K_i(g) > 0\}$ are finite. Assume that $K_1$ is symmetric, that $G = \bigcup_{n=1}^{\infty} V_2^n$ and that $K_2(e) > 0$. Let $M$ be a function of class $(\tilde{D})$ such that

$$\|K_1^{(n)}\|_{\infty} \leq M(n)$$

for all $n \in \mathbb{N}$. Then there are $c_1, c_2 > 0$ such that

$$\|K_2^{(n)}\|_{\infty} \leq c_1 M(c_2 n)$$

for all $n \in \mathbb{N}$.

**Proof.** Define $T_i f = K_i \ast f$ for $f \in L^p = L^p(G)$ and $i = 1, 2$. The symmetry of $K_1$ means that $T_1$ is self-adjoint, and it follows as in the proof of Lemma 3.4 that

$$\|f\|_2^2 - \|T_1 f\|_2^2 = ((I - T_1^2) f, f) \leq c \|\nabla f\|_2^2$$

for all $f \in L^2$. Also, using the assumptions on $K_2$, the proof of (14) shows that

$$\|\nabla f\|_2^2 \leq c (\|f\|_2^2 - \|T_2 f\|_2^2)$$

for $f \in L^2$. Since $\|K_1^{(n)}\|_{\infty} = \|T_1^n\|_{1 \to \infty}$, the theorem now follows from Theorem A.1.
Remark. In Theorem A.2, the assumptions $K_2(e) > 0$, $G = \bigcup_n V_2^n$ can be replaced by the weaker assumption that the set

$$V_2^{-1}V_2 := \{g^{-1}h: g, h \in V_2\}$$

generates $G$. To see this, note that $V_2^{-1}V_2$ equals the support of $\tilde{K} := K_2^* \ast K_2$ where $K_2^*(g) := K_2(g^{-1})$, $g \in G$. The proof of (14) shows that if $V_2^{-1}V_2$ generates $G$ then (22) holds, and the proof of Theorem A.2 remains valid.

References