The Annals of Probability
1993, Vol. 21, No. 2, 673–709

GAUSSIAN ESTIMATES FOR MARKOV CHAINS AND RANDOM WALKS ON GROUPS

BY W. HEBISCH AND L. SALOFF-COSTE

Wrocław University and University of Chicago, and CNRS, Université Paris VI

A Gaussian upper bound for the iterated kernels of Markov chains is obtained under some natural conditions. This result applies in particular to simple random walks on any locally compact unimodular group G which is compactly generated. Moreover, if G has polynomial volume growth, the Gaussian upper bound can be complemented with a similar lower bound. Various applications are presented. In the process, we offer a new proof of Varopoulos' results relating the uniform decay of convolution powers to the volume growth of G.

1. Introduction. The first result proved in this paper is a fairly general Gaussian upper bound for Markov chains. This bound applies in particular to simple random walks on locally compact compactly generated unimodular groups. When the group has polynomial volume growth, the iterated convolution kernel governing the random walk is shown to satisfy a two sided Gaussian estimate. Various applications of this estimate are discussed.

Our Gaussian upper bound for Markov chains is as follows. Consider a symmetric Markov kernel k defined on a measure space X, and assume that there is a distance function ρ on X such that $k(x,y) = 0$ whenever x, y satisfy $\rho(x,y) \geq r_0$. Also assume that the iterated kernels k_n satisfy the uniform estimate

$$\sup_{x,y}\{k_n(x,y)\} \leq C_0 n^{-D/2}, \qquad n = 1,2,\dots$$

for some $D \geq 0$. Then, we prove that

$$k_n(x,y) \leq C' n^{-D/2} \exp(-\rho^2(x,y)/Cn), \qquad x,y \in X, n = 1,2\dots.$$

Taken in this general setting, this result is similar to an estimate obtained by Varopoulos in [26] (see also [6] for a very nice proof of the Varopoulos estimate). Indeed, under the above hypotheses and if X is countable, Varopoulos' estimate yields

$$k_n(x,y) \leq C'_\varepsilon n^{-D/2+\varepsilon} \exp(-\rho^2(x,y)/C_\varepsilon n), \qquad x,y \in X, n = 1,2,\dots,$$

for any $\varepsilon > 0$. Such a result is also implicitly contained in [5]. However, the

Received November 1991.
AMS 1991 *subject classifications*. 60J15, 60B15.
Key words and phrases. Markov chain, random walk, convolution, groups, Gaussian estimates.

strength of our estimate comes from the fact that it is sharp in certain situations. This will be crucial for most of the applications presented in this paper. There are various ways of discussing the sharpness of such an estimate. One of them is to obtain a corresponding lower bound and we will be able to do so in some cases. A more trivial way to test a Gaussian estimate is to try to integrate it. When the volume of the set $\{y/\rho(x,y) \le n\}$ is uniformly bounded by Cn^D, the integral of our Gaussian bound over X is uniformly bounded, and this is a good sign. We also want to stress the fact that the different approaches used in [26, 6, 5] are adapted from methods that give sharp results when applied to the study of various (continuous time) heat kernels. Hence, it is rather surprising that these methods do not yield optimal results in the discrete time setting. Somehow, the discrete time case is more resistant.

From what has been said above, it becomes apparent that Gaussian estimates are especially interesting when one can first link the uniform decay of the iterated kernels k_n to the volume growth of balls. Thanks to the work of N. Varopoulos (see [25, 28, 31, 32, 20, 9, 21, 33]), one knows that such a link does exist in the setting of Markov chains on groups. In this paper, we present a new and simple proof of Varopoulos' uniform decay estimates of convolution powers (i.e., translation invariant Markov kernels); see Theorem 4.1 and Section 4. Our approach is adapted from the work [15] of the first of us where heat flow semigroups on Lie groups are studied. It yields sharp results in both the polynomial and superpolynomial volume growth cases. Let us emphasize here that these results do belong to harmonic analysis. In the absence of a group structure, there is no general link between uniform decay of Markov kernels and volume growth of balls; see [31]. This is one natural reason why this paper is mainly concerned with random walks on groups.

In the main part of this work, we apply the above results to the case of simple random walks on compactly generated groups having polynomial volume growth. In this case, the Gaussian upper bound can be complemented with a similar lower bound. Let us describe these results in more details in a special but typical case. Let Γ be a finitely generated group with neutral element e, and fix a finite set $\{\gamma_1, \dots, \gamma_s\}$ of generators. Define $\rho(x)$ to be the smallest integer such that $x = x_1 \cdots x_n$, where $x_i \in \{e, \gamma_1^{\pm 1}, \dots, \gamma_s^{\pm 1}\}$. The group Γ is said to have polynomial volume growth of order D if the number of elements x such that $\rho(x) < n$ is comparable to n^D. For instance, if $\Gamma = \mathbb{Z}^D$, then ρ is comparable to the Euclidean norm, and $\mathbb{Z}^D$ has polynomial volume growth of order D. Let p be the probability density corresponding to the uniform distribution on the set $\{e, \gamma_1^{\pm 1}, \dots, \gamma_s^{\pm 1}\}$ (what is important here is that p is symmetric compactly supported and charges a set of generators). Consider the random walk on Γ governed by p. It is a translation invariant Markov chain, and the iterated kernels are given by the convolution powers $p^{(n)}$. When Γ has polynomial volume growth of order D, our general Gaussian upper bound, together with the uniform decay estimate, yields

$$p^{(n)}(x) \le C' n^{-D/2} \exp\left(-\rho(x)^2/Cn\right), \qquad x \in \Gamma,\ n = 1, 2, \dots .$$

But we also prove the "gradient" estimate

$$\left|p^{(n)}(x) - p^{(n)}(x\gamma_i^{\pm 1})\right| \le C' n^{-(D+1)/2} \exp\left(-\rho(x)^2/Cn\right),$$

$$x \in \Gamma, n = 1, 2, \dots,$$

and the Gaussian lower bound [note that $p^{(n)}(x) = 0$ if $\rho(x) > n$]

$$p^{(n)}(x) \ge (C'n)^{-D/2} \exp\left(-C\rho(x)^2/n\right) \quad \text{for } \rho(x) \le n, n = 1, 2, \dots .$$

Such estimates are powerful tools. For instance, they imply that positive or sublinear p-harmonic functions (i.e., $u = u * p$) are constant. They also yield Sobolev and isoperimetric inequalities, as well as partial results concerning operators that are analogous to the classical Riesz transforms. In Section 9, the rate of escape of the random walk governed by p is studied. The main result in this direction is a generalization of a theorem due to Dvoretzky and Erdös when $\Gamma = \mathbb{Z}^D$. Another important application of the above Gaussian estimates is that they yield similar estimates for the kernels of associated Markov chains on homogeneous spaces. This is developed in the last section of this paper. It is worth emphasizing that most of these applications depend on having both an upper and lower Gaussian bound. A remarkable feature of the approach used in this work is that it does not depend on any result describing the structure of the underlying group.

2. Gaussian upper bounds. Let X be a measurable space endowed with a positive σ-finite measure dx. Let ρ be a (measurable) distance function on X, and denote by $B(x, r)$, $x \in X$, $r > 0$, the ball of center x and radius r. Let $k(x, y), (x, y) \in X^2$, be a bounded symmetric Markov kernel such that

$$(1) \qquad \{y \in X/k(x, y) \ne 0\} \subset B(x, r_0), \qquad x \in X$$

for some fixed $r_0 > 0$. The iterated kernel k_n is defined by $k_n(x, y) = \int k(x, z) k_{n-1}(z, y)\, dz$. Here is the result which is the powerhouse of this paper.

THEOREM 2.1. *Let k be a symmetric bounded Markov kernel which satisfies (1), and assume that*

$$(2) \qquad \sup_{x,y}\{k_n(x, y)\} \le C_0 n^{-D/2}, \qquad n = 1, 2, \dots .$$

Then, there exist two constants C, C' such that

$$k_n(x, y) \le C' C_0 n^{-D/2} \exp\left(-\rho^2(x, y)/Cn\right)$$

for all $x, y \in X$, and all $n = 1, 2, \dots$. Here, the constant C depends only on r_0 whereas C' depends on D and r_0.

This section is devoted to the proof of this theorem. First, we introduce some notation. The Dirichlet form $\mathscr{D}$, associated with k, is defined by

$$\mathscr{D}(f,g) = \langle (I-K)f, g\rangle = \tfrac{1}{2}\int (f(x)-f(y))(g(x)-g(y))k(x,y)\,dx\,dy,$$

$$f, g \in L^2,$$

where K is the symmetric Markov operator $Kf(x) = \int k(x,y)f(y)\,dy$. Consider the weight functions w_s, $s \in \mathbb{R}$, given by $w_s(x) = \exp(s\rho(x_0, x))$ where x_0 is any fixed point in X, and define the operator K_s by $K_s f(x) = w_{-s}K(w_s f)(x)$. Remark that the hypothesis (1) easily implies $\|K_s\|_{p\to p} \le \exp(r_0|s|)$ and

$$\|K_s^n\|_{p\to p} \le \exp(r_0|s|n). \tag{3}$$

Also, (1) and (2) imply

$$\|K_s\|_{p\to\infty} \le C_0^{1/p}\exp(r_0|s|).$$

Unfortunately, this is not quite enough to show that k_n admits a Gaussian upper bound. What is needed to obtain a Gaussian upper bound is indicated in the following technical lemma.

LEMMA 2.2. *Let k be a symmetric bounded Markov kernel which satisfies the uniform estimate* (2), *and such that*

$$\|K_s\|_{p\to\infty} \le C_0^{1/p}\exp(C_1(s^2+1))$$

for all $s \in \mathbb{R}$ and $p \ge 2$. Assume also that

$$\|K_s^n\|_{2\to 2} \le \exp(C_1(s^2 n + 1)), \qquad s \in \mathbb{R},\ n = 1, 2, \dots. \tag{4}$$

Then, there exist two constants C, C_1' such that

$$k_n(x,y) \le C C_0 n^{-D/2}\exp(-\rho^2(x,y)/C_1' n)$$

for all $x, y \in X$, and all $n = 1, 2 \dots$. The constant C_1' depends only on C_1 whereas C depends only on D and C_1.

PROOF. Using the fact that K is a contraction on each L^p space, and classical interpolation techniques, we deduce from (2) that

$$\|K^n\|_{p\to+\infty} \le C_0^{1/p} n^{-D/2p}$$

for all $p \in [1, +\infty[$. Similarly, we deduce from (4) that $\|K_{\theta s}^n\|_{p\to p} \le \exp(C_1\theta(s^2 n + 1))$ where $1/p = \theta/2 + (1-\theta)/\infty$, with $p \ge 2$. After replacing θs by s in this last inequality, it reads

$$\|K_s^n\|_{p\to p} \le \exp(C_1((p/2)s^2 n + 2/p)).$$

We interpolate between these two inequalities and get

$$\|K_{\theta s}^n\|_{p\to q} \le \left(C_0^{1/p} n^{-D/2p}\right)^{1-\theta}\exp(C_1\theta((p/2)s^2 n + 2/p)),$$

where $1/q = \theta/p + (1-\theta)/\infty$, $q \geq p \geq 2$. Replacing again θs by s, we obtain

$$\|K_s^n\|_{p\to q} \leq \left(C_0 n^{-D/2}\right)^{1/p-1/q} \exp\left(C_1((q/2)s^2 n + 2/q)\right).$$

Set $p_i = 2i^2$, $\eta_i = ci^{-5}$ with $c = (\Sigma_1^{+\infty} i^{-5})^{-1}$. For any fixed integer n, define $N = N(n)$ to be the largest integer such that $n\eta_N \geq 1$. Also, for $i = 2, \dots, N-1$, define $n_i = n_i(n)$ to be the largest integer less or equal to $n\eta_i$, and set $n_1 = n - \Sigma_2^{N-1} n_i$. Remark that we have $n_1 \geq c_n$, $\Sigma_1^\infty \eta_i p_{i+1} < +\infty$, and $\sup_n\{n^{D/p_N}\} < +\infty$. Armed with this notation and the last estimate on $\|K_s^n\|_{p\to q}$, we obtain

$$\begin{aligned}
\|K_s^{n+1}\|_{2\to\infty} &\leq \|K_s\|_{p_N\to\infty} \prod_1^{N-1} \|K_s^{n_i}\|_{p_i\to p_{i+1}} \\
&\leq C_0^{1/p_N} \prod_1^{N-1} \left(C_0 n_i^{-D/2}\right)^{(1/p_i - 1/p_{i+1})} \\
&\quad \times \exp\left(C_1\left((p_{i+1}/2)n_i s^2 + 2p_{i+1}^{-1} + s^2 + 1\right)\right) \\
&\leq C_0^{1/2} n^{-D(1/4-1/2p_N)} \left(\prod \eta_i^{-D/p_i}\right) \\
&\quad \times \exp\left(C_1\left(\left(\sum \eta_i p_{i+1}\right)s^2 n + 2\sum p_{i+1}^{-1} + s^2 + 1\right)\right) \\
&\leq C(D) C_0^{1/2} n^{-D/4} \exp\left(C_1'(s^2 n + 1)\right).
\end{aligned}$$

Here, $C(D)$ depends only on D whereas $C_1' = CC_1$ for some numerical constant C (the exact value of $C(D)$ and C_1' may change from line to line in the estimates below). Notice that, since K_{-s} is the adjoint of K_s, we also have

$$\|K_s^n\|_{1\to 2} \leq C(D) C_0 n^{-D/4} \exp\left(C_1'(s^2 n + 1)\right).$$

Hence, we get

$$\|K_s^n\|_{1\to\infty} \leq C(D) C_0 n^{-D/2} \exp\left(C_1'(s^2 n + 1)\right),$$

or, equivalently,

$$k_n(x,y) \leq C(D) C_0 n^{-D/2} \exp\left(C_1'(s^2 n + 1) + s\left(\rho(x_0, x) - \rho(x_0, y)\right)\right).$$

Choosing $x_0 = x$, and $s = \rho(x,y)/2C_1' n$, we obtain

$$k_n(x,y) \leq C(D) C_0 n^{-D/2} \exp\left(-\rho^2(x,y)/4C_1' n + C_1'\right),$$

which ends the proof of Lemma 2.2. □

Lemma 2.2 reduces the proof of Theorem 2.1 to checking that (4) holds under the hypothesis (1). With this goal in mind, we first state a lemma whose usefulness will be apparent later on. We could have extracted this lemma from [5]. However, the machinery introduced in [5] is not needed here and we include a proof for the sake of completeness.

LEMMA 2.3. *Let k be a symmetric bounded Markov kernel which satisfies (1). There exists a constant C depending only on r_0 such that*

$$\mathscr{D}(w_s f, w_{-s} f) \geq -Cs^2\|f\|_2^2$$

for all positive $f \in L^2$, and all s such that $|s| \leq 1$.

PROOF. Fix $s \in [-1, 1]$, and set $w = w_s$. Note that

(5) $\quad |w(x) - w(y)| \leq r_0|s|(w(x) + w(y)), \qquad x \in X, y \in B(x, r_0).$

Replacing f by wf, we see that Lemma 2.3 can be reduced to the claim that

$$\mathscr{D}(w^2 f, f) \geq -Cs^2\|wf\|_2^2.$$

To prove this claim, we write

$$\begin{aligned} 4\mathscr{D}(f, w^2 f) &= 2\int (f(x) - f(y))(w^2 f(x) - w^2 f(y))k(x,y)\,dx\,dy \\ &= \int (f(x) - f(y))(f(x) - f(y)) \\ &\qquad \times (w^2(x) + w^2(y))k(x,y)\,dx\,dy \\ &\qquad + \int (f(x) - f(y))(f(x) + f(y)) \\ &\qquad\quad \times (w(x) - w(y))(w(x) + w(y))k(x,y)\,dx\,dy \\ &= E_1 + E_2. \end{aligned}$$

The first term E_1 is nonnegative. Using the Cauchy–Schwarz inequality and (5), the second term can be estimated by

$$\begin{aligned} |E_2| &\leq \left(\int |f(x) - f(y)|^2 (w(x) + w(y))^2 k(x,y)\,dx\,dy\right)^{1/2} \\ &\qquad \times \left(\int (f(x) + f(y))^2 (w(x) - w(y))^2 k(x,y)\,dx\,dy\right)^{1/2} \\ &\leq C|s|E_1^{1/2}\left(\int (f(x) + f(y))^2 (w(x) + w(y))^2 k(x,y)\,dx\,dy\right)^{1/2} \\ &\leq C'|s|E_1^{1/2}\|wf\|_2 \leq E_1 + C''s^2\|wf\|_2^2. \end{aligned}$$

Here the constants depend only on r_0 (note that we used the fact that $|s| \leq 1$ to obtain the third inequality). Hence, we obtain $E_1 + E_2 \geq -C''s^2\|wf\|_2^2$, which proves the claim and Lemma 2.3. □

LEMMA 2.4. *Let k be a symmetric bounded Markov kernel which satisfies (1). There exists a constant C, depending only on r_0, such that*

$$\|K_s^n\|_{2\to 2} \leq \exp(C(s^2 n + 1)), \qquad s \in \mathbb{R}, n = 1, 2, \ldots.$$

PROOF. The proof has two steps. First, consider the symmetric Markov semigroup

$$\Gamma_t = e^{-t(I-K)} = e^{-t}\sum_{n=0}^{+\infty}\frac{t^n}{n!}K^n,$$

and define the perturbed semigroups $\Gamma_{s,t}$ by

$$\Gamma_{s,t}f(x) = w_{-s}(x)\Gamma_t(w_s f)(x), \qquad s \in \mathbb{R}.$$

We have $\partial_t\|\Gamma_{s,t}f\|_2^2 = -2\langle w_{-s}(I-K)(w_s\Gamma_{s,t}f), (\Gamma_{s,t}f)\rangle$ and, using Lemma 2.3, we obtain $\partial_t\|\Gamma_{s,t}f\|_2^2 \le Cs^2\|\Gamma_{s,t}f\|_2^2$ for $|s| \le 1$. From this we deduce immediately that

$$\|\Gamma_{s,t}f\|_2 \le \exp(Cs^2t)\|f\|_2, \qquad f \in L^2, t > 0, |s| \le 1.$$

The second step consists in passing information from $\Gamma_{s,t}$ to K_s^n. Denote by $\mathscr{E}(n)$ the set of the even integers i satisfying $n - \sqrt{n} \le i \le n$. For any nonnegative function f in L^2, $t > 0$, $|s| \le 1$, $n = 1, 2, \dots,$ we have

$$\left\| e^{-t}\sum_{\mathscr{E}(n)}\frac{t^i}{i!}K_s^i f\right\|_2^2 \le \|\Gamma_{s,t}f\|_2^2 \le \exp(Cs^2t)\|f\|_2^2.$$

Note also that

$$K_{-s}^j f \le \exp(C|s|j)K_s^j f, \qquad s \in \mathbb{R}, j = 1, 2, \dots.$$

From this remark and (3), it follows that

$$\begin{aligned}\left\| e^{-t}\sum_{i\in\mathscr{E}(n)}\frac{t^i}{i!}K_s^i f\right\|_2^2 &= e^{-2t}\sum_{i\in\mathscr{E}(n)}\sum_{j\in\mathscr{E}(n)}\frac{t^{i+j}}{i!j!}\langle K_s^i f, K_s^j f\rangle \\ &\ge e^{-(2t+C|s|\sqrt{n})}\sum_{i\in\mathscr{E}(n)}\sum_{j\in\mathscr{E}(n)}\frac{t^{i+j}}{i!j!}\|K_s^{(i+j)/2}f\|_2^2 \\ &\ge e^{-C'|s|\sqrt{n}}\|K_s^n f\|_2^2\left(e^{-t}\sum_{i\in\mathscr{E}(n)}\frac{t^i}{i!}\right)^2.\end{aligned}$$

Stirling's formula yields $e^{-n}\sum_{i\in\mathscr{E}(n)}n^i/i! \ge c$ for some positive constant c independent of n. Hence, choosing $t = n$, we deduce from the above that

$$\|K_s^n\|_{2\to 2} \le \exp\big(C(s^2n + |s|\sqrt{n} + 1)\big), \qquad n = 1, 2, \dots,$$

for $|s| \le 1$. But, using (3), we see that this estimate also holds for $|s| > 1$. This ends the proof of Lemma 2.4. Using Lemma 2.2 and Lemma 2.4 we also have a proof of Theorem 2.1. □

Theorem 2.1 yields estimates on the corresponding Green kernel. More precisely, denote by Θ the modified Green potential operator $\Theta = \sum_1^{+\infty}K^n$ (the usual Green potential is $I + \Theta$) and let $\Theta(x,y) = \Sigma_1^{+\infty} k_n(x,y)$ be its kernel. From Theorem 2.1, we easily deduce the following:

THEOREM 2.5. *Let k be a symmetric bounded Markov kernel which satisfies* (1) *and* (2) *for some $D > 2$. Then there exists a constant C such that*

$$\Theta(x,y) \le C\rho(x,y)^{-D+2}$$

for all $x, y \in X$.

REMARK 1. Inspection of the above proof, together with few simple adaptations, shows that the hypothesis (1) can be relaxed a little and replaced by the condition that $p(x) \le A\exp(-a\rho(x)^2)$ for some fixed positive constants a, A.

REMARK 2. Here is a question left open by the above, and which corresponds to the case of time dependent coefficients in the classical continuous time setting. Let h_i, $i = 1, 2, \ldots$ be a sequence of bounded symmetric Markov kernels. Assume that the h_i's satisfy uniformly the condition that $h_i(x,y) = 0$ whenever $\rho(x,y) \ge r_0$, for some fixed $r_0 > 0$. Denote by H_i the corresponding operators, and let $k_{i,j}$ be the kernel of $K_{i,j} = H_j H_{j-1} \cdots H_{i+1}$ for $1 \le i < j$. Finally, assume that the $k_{i,j}$'s satisfy the uniform estimate

$$\sup_{x,y}\{k_{i,j}(x,y)\} \le C(j-i)^{D/2}, \quad \text{for all } j > i \ge 1$$

(see [33] for examples where these hypotheses are satisfied). It is natural to conjecture that, under these circumstances, $k_{i,j}$ also satisfies the Gaussian estimate

$$k_{i,j}(x,y) \le C(j-i)^{D/2}\exp\left(-\rho(x,y)^2/C(j-i)\right).$$

The above argument does not yield this result. Note however that Lemma 2.2 can easily be generalized to this setting.

3. Preliminary considerations on groups. In this section, we introduce some notation which will be used throughout the rest of this paper. Apart from the general results obtained in Section 2, most of the results in this paper deal with the case where the underlying space X is a group and the Markov kernel k is invariant under left translation. More precisely, consider a locally compact unimodular group G, and assume that G is compactly generated. Let μ be a probability measure on G. After n steps, the distribution of the random walk on G governed by μ is given by the convolution power $\mu^{(n)}$. Assume that μ has a bounded symmetric density p with respect to the Haar measure of G, and denote by P the Markov operator defined by

$$Pf(x) = f * p(x) = \int f(y)p(y^{-1}x)\,dy.$$

Of course, this situation corresponds to the case when the Markov kernel k

satisfies $k(x, y) = k(zx, zy)$, $x, y, z \in G$. In this case, setting $p(x) = k(x, e)$, where e denote the neutral element of G, we have $k_n(x, y) = p^{(n)}(y^{-1}x)$.

Fix a symmetric open neighborhood Ω of e which is relatively compact and generates G (i.e., $G = \bigcup_0^{+\infty} \Omega^n$). The volume growth function V is defined by

$$V(n) = |\Omega^n|, \qquad n = 1, 2, \dots,$$

where $|A|$ is the Haar measure of the set $A \subset G$. There is also a left invariant distance function ρ associated with Ω. Namely, for $x \in G$, we set

$$\rho(x) = \rho(e, x) = \inf\{n | x \in \Omega^n\}$$

and $\rho(x, y) = \rho(x^{-1}y) = \rho(y^{-1}x)$. The sets $x\Omega^n$ are balls for the distance ρ. In general, we set $B(x, r) = \{y \in G,\ \rho(x, y) \le r\}$. If Ω_1, Ω_2 are two neighborhoods of e as above, it is not very difficult to check that there exists $C > 0$ such that $C^{-1} \le \rho_2/\rho_1 \le C$ and that the corresponding growth functions satisfy $V_1 \simeq V_2$, by which we mean that $V_1(C^{-1}n) \le V_2(n) \le V_1(Cn)$, $n = 1, 2, \dots$. Hence, in some sense, V and ρ are invariants attached to G, and it is possible to define the type of growth of the group G as a notion independent of Ω. It is easy to see that V is at most of exponential type. Several deep theorems relate the volume growth to the structure of G. The approach used in this work does not depend on any of these theorems. However, they certainly help understanding the real meaning of the results obtained in this paper, and we briefly recall them now.

1. (Y. Guivarc'h [14]) If G is a connected Lie group, or if G is solvable, then either there exists $D = 0, 1, \dots,$ such that $V(n) \simeq n^D$, or $V(n) \simeq \exp(n)$.
2. (M. Gromov; see [12, 27]) If G is finitely generated, then either G is almost nilpotent and $V(n) \simeq n^D$ for some $D = 0, 1, \dots,$ or V is of superpolynomial growth.
3. (R. Grigorchuk [13]) There exist $0 < \beta \le \alpha < 1$ and a finitely generated group such that

$$\exp(C^{-1}n^{\beta}) \le V(n) \le \exp(Cn^{\alpha}), \qquad n = 1, 2, \dots.$$

We say that G has polynomial volume growth of order D when $V(n) \simeq n^D$. Nilpotent groups are examples of groups having polynomial volume growth; see [14, 3], and the Appendices to [12] by Tits. Suppose, for instance, that G is a nilpotent simply connected Lie group with Lie algebra $\mathscr{G}$ and set $\mathscr{G}_1 = \mathscr{G}$, $\mathscr{G}_i = [\mathscr{G}, \mathscr{G}_{i-1}]$ for $i = 2, \dots$. Since G is nilpotent, there exists an integer m such that $\mathscr{G}_{m+1} = \{0\}$. The order D of the polynomial volume growth of G is given by $D = \sum_1^m i \dim(\mathscr{G}_i/\mathscr{G}_{i+1})$. Similarly, if G is a finitely generated nilpotent group with lower central series $\{0\} \subset G_m \subset \cdots \subset G_2 \subset G_1 = G$, then G has polynomial volume growth of order $D = \sum_1^m i \operatorname{rk}(G_i/G_{i+1})$, where $\operatorname{rk}(H)$ is the (torsion-free) rank of the abelian finitely generated group H. Note however that certain solvable (but not nilpotent) Lie groups also have polynomial volume growth; see [14, 33].

The following notation will turn out to be useful. The "gradient" ∇f of a function f is defined by

$$\nabla f(x) = \sup_{y \in \Omega} \{ | f(x) - f(xy) | \}.$$

Also, define $\nabla_2 f(x) = (\int_{\Omega^2} |f(x) - f(xy)|^2 \, dy)^{1/2}$ and $\nabla_1 f(x) = \int_\Omega |f(x) - f(xy)| \, dy$. All these notions are essentially equivalent. More precisely, we have:

LEMMA 3.1. *For all $x \in G$ and $f \in L^1 \cap L^\infty$, the gradients $\nabla, \nabla_1, \nabla_2$ satisfy*

$$\nabla_1 f(x) \le |\Omega|^{1/2} \nabla_2 f(x) \le 2 |\Omega^2| \sup_{z \in \Omega} \{ \nabla f(xz) \},$$

$$\nabla f(x) \le 2 |\Omega|^{-1/2} \sup_{z \in K} \{ \nabla_2 f(xz) \}.$$

PROOF. The first string of inequalities is clear. To prove the last inequality stated in the lemma note that, for $y \in \Omega$, we have

$$\begin{aligned} | f(x) - f(xy) | &\le \left| f(x) - |\Omega|^{-1} \int_\Omega f(xz) \, dz \right| + \left| f(xy) - |\Omega|^{-1} \int_\Omega f(sz) \, dz \right| \\ &\le |\Omega|^{-1} \left(\int_\Omega | f(x) - f(xz) | \, dz + \int_\Omega | f(xy) - f(xz) | \, dz \right) \\ &\le |\Omega|^{-1/2} \left(\int_\Omega | f(x) - f(xz) |^2 \, dz \right)^{1/2} \\ &\quad + \left(\int_{\Omega^2} | f(xy) - f(xyz) |^2 \, dz \right)^{1/2}, \end{aligned}$$

from which the desired conclusion follows. □

The next simple lemma (which is taken from [9]) will play an important role.

LEMMA 3.2. *Let K be a symmetric Markov operator, and fix an integer l. There exists C_l such that we have*

$$\left\| (I - K^{2l})^{1/2} K^n f \right\|_2 \le C_l n^{-1/2} \| f \|_2, \qquad f \in L^2, \, n = 1, 2, \dots .$$

PROOF. Let $K^2 = \int_0^1 \lambda \, dE_\lambda$ be a spectral decomposition of the symmetric positive operator K^2. Clearly, it is enough to prove the lemma when $n = 2s$ is even. In this case, we have

$$\left\| (I - K^{2l})^{1/2} K^{2s} f \right\|_2^2 = \int_0^1 (\lambda^{2s} - \lambda^{l+2s}) \, dE_\lambda(f, f)$$

and the lemma follows since $\sup\{\lambda^{2s} - \lambda^{l+2s}, \lambda \in [0, 1]\} \le l/4s$. Note that this lemma has nothing to do with the group structure. □

4. Uniform decay of convolution powers. In this section, we give a new proof of results which are essentially due to Varopoulos, and which relate the uniform decay of convolution powers (i.e., the decay of $p^{(n)}(e)$) to the volume growth of the group G. Here is the result that we prove:

THEOREM 4.1. *Let p be a symmetric bounded probability density on G and assume that there exists an open generating neighborhood U of e such that*

$$\inf\{p(x), x \in U\} > 0. \tag{6}$$

(i) *If $V(n) \geq Cn^D$ for some $D \geq 0$, then $p^{(n)}(e) = O(n^{-D/2})$ as n tends to $+\infty$.*

(ii) *If $V(n) \geq \exp(Cn^\alpha)$ for some $0 < \alpha \leq 1$, then there exists $\kappa > 0$ such that*

$$p^{(n)}(e) = O\big(\exp(-\kappa n^{\alpha/(\alpha+2)})\big), \quad \text{as } n \text{ tends to } +\infty.$$

Other type of growth can also be considered; see the last paragraph at the end of this section. Apart from technical details, the above theorem is due to Varopoulos. Part (i) is proved in [27, 28]. A simpler proof is also given in [20]. Part (ii) is proved in [31, 32] in the case of finitely generated groups. There is a common feature to all these proofs. They depend on the equivalence between the decay properties of $p^{(n)}(e)$ and certain functional inequalities (e.g., the Sobolev inequality); see [9, 20, 21, 27, 28, 31, 32]. The proof presented below differs completely from the above ones. It does not use any kind of functional inequality. The arguments are adapted from an approach discovered by the first of us in the context of the heat flow semigroup on Lie groups; see [15]. The fact that the Markov chains which we consider are translation invariant (i.e., have a convolution kernel) plays a crucial role in these arguments. The following result is crucial for our proof of Theorem 4.1. It can be interpreted as a sort of weak Harnack inequality.

THEOREM 4.2. *Let p be a symmetric bounded probability density satisfying (6). There exists C_0 such that, for all $n, m = 1, 2, \ldots,$ we have*

$$\left|p^{(2n+m)}(x) - p^{(2n+m)}(e)\right| \leq C_0 \rho(x) m^{-1/2} p^{(2n)}(e).$$

PROOF. First, note that

$$\left|p^{(2n+m)}(x) - p^{(2n+m)}(e)\right| \leq \rho(x) \left\|\nabla p^{(2n+m)}\right\|_\infty \leq C\rho(x) \left\|\nabla_2 p^{(2n+m)}\right\|_\infty,$$

where the last inequality follows from Lemma 3.1. This reduces the proof of Theorem 4.2 to the claim that

$$\left\|\nabla_2 p^{(2n+m)}\right\|_\infty \leq Cm^{-1/2} p^{(2n)}(e). \tag{7}$$

Writing $p^{(2n+m)} = p^{(n)} * p^{(n+m)}$, we obtain

$$|\nabla_2 p^{(2n+m)}(x)| = \left(\int_{\Omega^2}\left|\int_G (p^{(n+m)}(y^{-1}x) - p^{(n+m)}(y^{-1}xz))p^{(n)}(y)\,dy\right|^2 dz\right)^{1/2}$$

$$\le \int_G \left\{\int_{\Omega^2}|p^{(n+m)}(y^{-1}x) - p^{(n+m)}(y^{-1}xz)|^2\,dz\right\}^{1/2} p^{(n)}(y)\,dy.$$

Using the Cauchy–Schwarz inequality and the invariance of the Haar measure (this is the only place where the group structure and the translation invariance of the Markov chain are really used), we get

$$\|\nabla_2 p^{(2n+m)}\|_\infty \le \|p^{(n)}\|_2 \|\nabla_2 p^{(n+m)}\|_2. \tag{8}$$

The hypothesis (6) implies that there exists some integer n_0 such that (see for instance [9], page 433, Proposition 5)

$$\inf\{p^{(2n_0)}(x),\, x \in \Omega^2\} > 0. \tag{9}$$

Hence, setting $p_0 = p^{(2n_0)}$, we have $\nabla_2 f(x) \le C(\int|f(x) - f(xy)|^2 p_0(y)\,dy)^{1/2}$, which yields

$$\|\nabla_2 p^{(n+m)}\|_2 = \left(\int_G\int_{\Omega^2}|p^{(n+m)}(x) - p^{(n+m)}(xy)|^2\,dy\,dx\right)^{1/2}$$

$$\le C\left(\int_G\int_G|p^{(n+m)}(x) - p^{(n+m)}(xy)|^2 p_0(y)\,dy\,dx\right)^{1/2}$$

$$= C'\|(I - P^{2n_0})^{1/2}P^m p^{(n)}\|_2.$$

By (8) and Lemma 3.2, we get

$$\|\nabla_2 p^{(2n+m)}\|_\infty \le Cm^{-1/2}\|p^{(n)}\|_2^2 = Cm^{-1/2}p^{(2n)}(e),$$

which proves the claim (7) and thus Theorem 4.2.

PROOF OF THEOREM 4.1. Given n, $m = 1, 2, \dots,$ set

$$r(n,m) = m^{1/2}p^{(2n+m)}(e)/2C_0 p^{(2n)}(e),$$

where C_0 is the constant appearing in Theorem 4.2. With this notation, Theorem 4.2 yields

$$p^{(2n+m)}(x) \ge \tfrac{1}{2}p^{(2n+m)}(e), \qquad \rho(x) \le r(n,m).$$

Integrating over the ball $\{\rho(x) \le r(n,m)\}$ and setting $A(n) = p^{(n)}(e)$, we obtain

$$A(2n+m) \le 2V(r(n,m))^{-1}. \tag{10}$$

Armed with this fact, we now give a proof of Theorem 4.1.

PROOF OF PART (i) (Polynomial case). Assume that $V(n) \geq Cn^D$, $n = 1, 2, \ldots$. In this case, (10) becomes

$$A(2n+m) \leq C\big(m^{1/2}A(2n+m)/A(2n)\big)^{-D}.$$

Choosing $m = 2n$ and setting $\theta = D/(1+D)$, we get

$$A(4n) \leq \big(Cn^{-1/2}A(2n)\big)^{\theta}.$$

For $n \geq 3$, define $\sigma(n)$ to be the smallest integer such that $2^{-\sigma(n)-1}n \leq 1$. We have $n > 2^{\sigma(n)}$ and

$$A(n) \leq A(2^{\sigma(n)}) \leq \prod_{1}^{\sigma(n)-1} \{C^{\theta^i}2^{\theta^i(i-\sigma(n))/2}\}A(2)^{\theta^{\sigma(n)-1}}$$

$$\leq C'2^{-D\sigma(n)/2} \leq C'(n/2)^{-D/2},$$

which proves the first part of Theorem 4.1.

PROOF OF PART (ii) (Superpolynomial case). Let $n, m = 1, 2, \ldots$ be such that $m \leq n$. On the one hand, if for some integer $i \in [1, n/m]$, we have

$$A(2n+2im)/A(2n+2(i-1)m) > 1/2,$$

then, by (10) and the fact that A is nonincreasing, we obtain

$$A(4n) \leq 2\{V(m^{1/2}/C)\}^{-1}.$$

On the other hand, if

$$A(2n+2im)/A(2n+2(i-1)m) \leq 1/2$$

for all integers $i \in [1, n/m]$, then we get

$$A(4n) \leq 2^{2-n/m}A(2n) \leq 2^{2-n/m}A(2).$$

Hence, we always have

$$A(4n) \leq \max\{2V(m^{1/2}/C)^{-1}, 2^{2-n/m}A(2)\}, \qquad m \leq n. \tag{11}$$

If we assume that $V(n) \geq \exp(Cn^{\alpha})$, $n = 1, 2, \ldots$, for some $0 < \alpha \leq 1$, choosing $m \sim n^{2/(2+\alpha)}$ in (11) yields

$$A(4n) = O\big(\exp(-\kappa n^{\alpha/(\alpha+2)})\big),$$

for some $\kappa > 0$, which ends the proof of Theorem 4.1. □

We now discuss the sharpness of Theorem 4.1. When doing so, we have to restrict the class of functions p under consideration: If p is too spread out, $p^{(n)}(e)$ might have a faster decay than the one given by Theorem 4.1. Hence, we assume in the following discussion that p satisfies the hypotheses of Theorem 4.1 and has compact support. Concerning the polynomial case, it is known that, if G is a finitely generated nilpotent group, or a nilpotent Lie group, with $V(n) \simeq n^D$, then

$$C^{-1} \leq p^{(2n)}(e) n^{D/2} \leq C \qquad n = 1, 2, \dots;$$

see [27] and the references given there. In the next section, we show that this holds on any group having polynomial growth of order D. When G has exponential growth, it may happen that $p^{(n)}(e)$ has exponential decay: This is the case if and only if G is nonamenable (note that any nonamenable group has exponential volume growth). However, there are examples of groups G having exponential volume growth, and for which the type of decay given by Theorem 4.1, part (ii) is correct. Namely, let G be a finitely generated group which is polycyclic but not almost nilpotent. By Theorem 4.1 and a result of Alexopoulos [1] we have, in this case,

$$\exp(-Cn^{1/3}) \leq p^{(2n)}(e) \leq \exp(-n^{1/3}/C), \qquad n = 1, 2, \dots$$

(note that polycyclic groups are solvable hence also amenable, and that a finitely generated solvable group that is not almost nilpotent has exponential growth). Hence, part (ii) of Theorem 4.1 is rather sharp in this case.

As a closing remark, it may be of interest to note that other types of growth can be considered. For instance, suppose that G has a slow superpolynomial volume growth in the sense that

$$V(n) \geq Cn^{\lambda} \exp\bigl(\gamma(\log(n))^{\beta}\bigr), \qquad n = 1, 2, \dots,$$

for some $\lambda \in \mathbb{R}$, $\gamma \geq 0$, $\beta > 1$. Then, choosing $m = \kappa n/(\log(n))^{\beta} \leq n$ in (11), we obtain

$$p^{(4n)}(e) \leq \max\Bigl\{C_1 e^{(-\gamma\{\log(\kappa^{1/2} n^{1/2}/C_2(\log(n))^{\beta/2})\}^{\beta})}, e^{(-(\log(n))^{\beta}/C_3\kappa)}\Bigr\}$$

and, choosing κ small enough,

$$p^{(n)}(e) \leq n^{\lambda/2} e^{(-\gamma((1/2)\log(n))^{\beta} + O(C(\log(n))^{\beta-1}(\log(\log(n)))))}.$$

Unfortunately, we do not know if there exist groups having this type of growth.

5. Gaussian bounds for convolution powers. We now apply the results of Sections 2 and 4 to obtain a Gaussian upper bound for convolution powers. By taking advantage of the fact that we are dealing with convolution kernels, we are able to complement the Gaussian upper bound with a gradient estimate. For the same reason, we easily obtain a Gaussian lower bound. We want to emphasize that a number of key points of the argument presented below are specific to the case of translation invariant Markov chains (i.e., random walks). A further discussion of this aspect is given at the end of this section. The main estimates proved below are gathered in the following theorem.

THEOREM 5.1. *Assume that G has polynomial volume growth of order D. Let p be a symmetric bounded probability density. Assume that there exists an open generating neighborhood U of e such that* $\inf\{p(x),\ x \in U\} > 0$ [*i.e.*, p

satisfies (6)]. *Also assume that there exists* r_0 *such that*

$$\text{(12)} \qquad p \text{ has support in } B(e, r_0).$$

Then, there exist three positive constants C, C', C'' *such that, for all* $x \in G$ *and all integers* n, *we have*

$$\text{(13)} \qquad p^{(n)}(x) \leq Cn^{-D/2}\exp\bigl(-\rho(x)^2/C'n\bigr),$$

$$\text{(14)} \qquad \nabla p^{(n)}(x) \leq Cn^{-(D+1)/2}\exp\bigl(-\rho(x)^2/C'n\bigr),$$

$$\text{(15)} \qquad p^{(n)}(x) \geq (Cn)^{-D/2}\exp\bigl(-C'\rho(x)^2/n\bigr) \quad \text{if } x \in B(e, n/C'').$$

Moreover, if $D > 2$, *the modified Green kernel* $\Theta = \sum_1^\infty p^{(n)}$ *satisfies*

$$(C\rho(x))^{-D+2} \leq \Theta(x) \leq C\rho(x)^{-D+2}, \qquad x \in G$$

and

$$\nabla\Theta(x) \leq C\rho(x)^{-D+1}, \qquad x \in G.$$

Note that the statement concerning the Green kernel follows easily from (13), (14) and (15). The inequality (13) is an immediate application of Theorem 2.1 and Theorem 4.1. The rest of this section is devoted to the proofs of (14) and (15). We start with an easy consequence of (13).

LEMMA 5.2. *Set* $w_s(x) = \exp(s\rho(x))$, $x \in G$. *Assume that* G *has polynomial volume growth of order* D *and let* p *be a symmetric bounded density of probability satisfying* (12) *and* (6). *Then, for* $1 \leq q \leq +\infty$, *we have*

$$\|p^{(n)}w_s\|_q \leq Cn^{-D(1-1/q)/2}\exp(C's^2n), \qquad s > 0, n = 1, 2, \dots.$$

This result follows from the elementary fact that

$$\exp(s\rho - \rho^2/Cn) \leq \exp(Cs^2n - \rho^2/2Cn), \qquad s > 0, \rho > 0, n = 1, 2, \dots,$$

together with (13) since we have

$$\int e^{-\rho(x)^2/Cn}\,dx \leq \int_{\rho(s)^2 \leq n} e^{-\rho(x)^2/Cn}\,dx + \sum_0^{+\infty} \int_{2^i n < \rho(x)^2 \leq 2^{i+1}n} e^{-\rho(x)^2/Cn}\,dx$$
$$\leq C'n^{-D/2}\sum_0^{+\infty} 2^{iD/2}e^{-2^i/C} \leq C''n^{-D/2}.$$

PROOF OF (14). We adapt a method used in [22] in a continuous time setting. Fix $s > 0$, $\nu = n + m$, and note that $e^{s\rho(x)} \leq e^{s\rho(y^{-1}x)}e^{s\rho(y)}$ and $p^{(\nu)} = p^{(n)} * p^{(m)}$. This implies

$$\text{(16)} \qquad \begin{aligned} e^{s\rho(x)}\nabla_2 p^{(\nu)}(x) &\leq \int e^{s\rho(y^{-1}x)}\nabla_2 p^{(n)}(y^{-1}x)e^{s\rho(y)}p^{(m)}(y)\,dy \\ &\leq \|w_s\nabla_2 p^{(n)}\|2\|w_s p^{(m)}\|_2. \end{aligned}$$

Lemma 5.2 yields a good bound for $\| w_s p^{(m)} \|_2$. Using (9) and the notation introduced right after it, the other factor can be estimated by

$$\begin{aligned}\| w_s \nabla_2 p^{(n)} \|_2^2 &\le \int w_{2s}(x) \int \left| p^{(n)}(x) - p^{(n)}(xy) \right|^2 p_0(y)\, dy\, dx \\ &= \int w_{2s}(x) \left\{ \left| p^{(n)}(x) \right|^2 - 2p^{(n)}(x) p^{(n+2n_0)}(x) \right. \\ &\qquad \left. + \int \left| p^{(n)}(xy) \right|^2 p_0(y)\, dy \right\} dx \\ &= 2 \int w_{2s}(x) p^{(n)}(x) (I - P^{2n_0}) p^{(n)}(x)\, dx \\ &\qquad + \int (w_{2s}(x) - w_{2s}(xy)) \left| p^{(n)}(xy) \right|^2 p_0(y)\, dy\, dx \\ &= E_1 + E_2.\end{aligned}$$

Using the Cauchy–Schwarz inequality, Lemma 3.2 and Lemma 5.2, we get

$$E_1 \le \| w_{2s} p^{(n)} \|_2 \| (I - P^{2n_0}) p^{(n)} \|_2 \le C n^{-1-d/2} \exp(C' s^2 n). \tag{17}$$

To estimate E_2, first note that the invariance of the Haar measure and the symmetry of p imply

$$\begin{aligned} E_2 &= \int (w_{2s}(xy) - w_{2s}(s)) \left| p^{(n)}(x) \right|^2 p_0(y)\, dy\, dx \\ &= \tfrac{1}{2} \int (w_{2s}(xy) - w_{2s}(x)) \left(\left| p^{(n)}(x) \right|^2 - \left| p^{(n)}(xy) \right|^2 \right) p_0(y)\, dy\, dx. \end{aligned}$$

Using (5) and the Cauchy–Schwarz inequality, this yields

$$E_2 \le Cs \left\| \left((I - P^{2n_0})^{1/2} p^{(n)} \right) \right\|_2 \| w_{2s} p^{(n)} \|_2.$$

Finally, by Lemma 3.2 and Lemma 5.2, we get

$$E_2 \le Csn^{-1/2-D/2} \exp(C' s^2 n). \tag{18}$$

Hence, choosing $n = m$ or $n = m + 1$ depending on whether ν is even or odd, and using (16), (17), (18) and Lemma 5.2, we obtain

$$\exp(s\rho(x)) \nabla_2 p^{(\nu)}(x) \le C \left(1 + s\sqrt{\nu}\right)^{1/2} \nu^{-D/2-1/2} \exp(C' s^2 \nu).$$

Choosing $s = \rho/2C'\nu$ in this last inequality yields the estimate

$$\nabla_2 p^{(\nu)}(x) \le C\nu^{-1/2-D/2} \exp\left(-\rho(x)^2 / C'\nu\right),$$

and by Lemma 3.1, this ends the proof of (14). □

PROOF OF (15). We start by proving the following much weaker version of (15):

$$p^{(n)}(e) \ge (Cn)^{-D/2}, \quad n = 1, 2, \ldots . \tag{19}$$

Note that it is enough to prove the above when n is even since p satisfies (6). Using the Gaussian upper bound (13), we see that, for some fixed A large enough,

$$\int_{\rho(x)^2 \geq An} p^{(2n)}(x)\,dx \leq C'A^{D/2} \sum 2^{iD/2} \exp(-2^i A/C) \leq 1/2,$$

uniformly for $n = 1, 2, \dots$. Hence, we obtain

$$C(An)^{D/2} p^{(2n)}(e) \geq \int_{\rho(x)^2 \leq An} p^{(2n)}(x)\,dx \geq 1/2,$$

which yields the desired estimate. In a second step towards our Gaussian lower bound, we improve (19) by using Theorem 4.2 [or the gradient estimate (14)]. Indeed, we have

$$\left|p^{(n)}(x) - p^{(n)}(e)\right| \leq Cn^{-D/2-1/2}\rho(x)$$

which, together with (19), shows that there exist two positive constants C_0, C_1 such that

$$p^{(n)}(x) \geq (C_0 n)^{-D/2} \quad \text{for all } x, n \text{ such that } \rho(x) \leq \sqrt{n}/C_1. \tag{20}$$

From here, (15) is obtained by adapting a classical chaining argument. Fix $x \in G$ and an integer n. If $\rho(x) \leq \sqrt{n}/C_1$, (20) yields the desired estimate. Hence, assume that $\rho(x) \geq \sqrt{n}/C_1$. Write $n = n_1 + n_2 + \cdots + n_j$ where $n_i \sim n/j$ and $j \leq n$ is an integer to be fixed later. Also, fix a sequence $x_1 = e$, $x_2, \dots, x_{j+1} = x$ of elements of G such that $\rho(x_i^{-1}x_{i+1}) \leq \rho(x)/j$. For n large enough, choose j to be the smallest integer such that $\sqrt{j} \geq 10C_1\rho(x)/\sqrt{n}$. This is compatible with the condition $j \leq n$ when $10C_1\rho(x) \leq n$. Assuming that x, n satisfy this additional condition, (20) yields

$$\inf\left\{p^{(n_i)}\left(y_i^{-1}y_{i+1}\right), y_i \in B_i, y_{i+1} \in B_{i+1}\right\} \geq (C_0 n_i)^{-D/2},$$

where $B_i = B(x_i, \sqrt{n_i}/10C_1)$ and $1 \leq i \leq j$. Hence, we obtain

$$\begin{aligned} p^{(n)}(x) &\geq \int \cdots \int p^{(n_1)}(y_2) p^{(n_2)}\left(y_2^{-1}y_3\right) \cdots p^{(n_j)}\left(y_j^{-1}x\right) dy_2 \cdots dy_j \\ &\geq \int_{B_2} \cdots \int_{B_j} p^{(n_1)}(y_2) \cdots p^{(n_j)}\left(y_j^{-1}x\right) dy_2 \cdots dy_j \\ &\geq (n/j)^{-D/2}(1/C)^j \geq (C'n)^{-D/2} \exp\left(-C''\rho(x)^2/n\right), \end{aligned}$$

where the last inequality comes from the fact that $j \sim \rho(x)^2/n$. This ends the proof of the Gaussian lower bound (15). □

As we mentioned before, the gradient estimate obtained above for convolution powers is specific to invariant Markov chains. Hence, the proof of the Gaussian lower bound given above is also specific to this case. However, we have the following general Gaussian upper bound.

THEOREM 5.3. *Assume that G satisfies $V(n) \geq Cn^D$, $n = 1, 2, \dots$. Let k be a symmetric bounded Markov kernel on G satisfying* (1), *and assume that there exists an open symmetric generating neighborhood U of e such that*

$$\inf\{k(x,y),\, x \in X,\, y \in xU\} > 0. \tag{21}$$

Then there exist two constants C', C'' such that

$$k_n(x,y) \leq C'' n^{-D/2} \exp\bigl(-\rho(x,y)^2/C'n\bigr), \qquad x, y \in X,\, n = 1, 2, \dots.$$

Also, if $D > 2$, the modified Green kernel $\Theta = \sum_1^{+\infty} k_n$ satisfies

$$\Theta(x,y) \leq C' \rho(x,y)^{-D+2}, \qquad x, y \in G.$$

The proof follows from Theorem 2.1, Theorem 2.5 and the fact that under the above hypotheses we have $k_n(x,y) \leq Cn^{-D/2}$, $n = 1, 2, \dots$. Indeed, this uniform estimate is a (by now well known) consequence of the volume growth hypothesis and (21). It follows, for instance, from Theorem 4.1 and Proposition 4, page 430, of [9].

REMARK 1. Note that the volume growth hypothesis in Theorem 5.3 is weaker than the one in Theorem 5.1. If, for instance, we know that G has exponential volume growth, then we can combine Theorem 4.1 and the preceding result to conclude that any bounded symmetric probability density p satisfying (6) and (12) also satisfies

$$p^{(n)}(x) \leq C_1 \exp\left(-\left(\frac{n^{1/3}}{C_2} + \frac{\rho(x)^2}{C_3 n}\right)\right), \qquad x \in G,\, n = 1, 2, \dots.$$

We do not elaborate on this because such an estimate (whatever the constant C_2 is) is rather poor. For instance, the integral of the right-hand side is not uniformly bounded as n tend to infinity. Much work seems still to be needed in order to obtain better estimates in the nonpolynomial case.

REMARK 2. Another aspect of Theorem 5.3 is that it offers an estimate from above for Markov kernels instead of convolution powers. We have no doubt that, if G has polynomial volume growth, a corresponding Gaussian lower bound [similar to (15)] holds for Markov kernels as well. However, we have not been able to prove this result. We hope to come back to this question in the future. Note that the gradient estimate similar to (14) does not hold in general for Markov chains. We expect a Hölder continuity estimate to hold instead. Such a Hölder continuity estimate would follow from a Gaussian lower bound similar to (15).

6. Harmonic functions. In this section we present some applications of Theorem 5.1 to the potential theory associated with a symmetric bounded probability density p on a group G where p and G are as in Theorem 5.1 (in particular G has polynomial growth of order D). More precisely, we show that the p-harmonic functions which are either nonnegative or of sublinear growth are necessarily constants. First, recall that a function u is said to be p-harmonic if it satisfies the equation $Pu = u$. Also, a function f is of sublinear growth if

$$Mf(r) = \sup\{|f(x)|, x \in B(e,r)\}$$

is such that $r^{-1}Mf(r)$ tends to zero as r tends to infinity. The estimate (14) in Theorem 5.1 yields easily the triviality of p-harmonic function having sublinear growth. Indeed, if u is p-harmonic, we have for any integer n,

$$\begin{aligned}
\nabla u(e) &\le \int \nabla p^{(n)}(z)|u(z)|\,dz \\
&\le C_1 n^{-D/2-1/2} \int \exp(-\rho(z)^2/Cn)|u(z)|\,dz \\
&\le C_2 n^{-D/2-1/2}\left\{ Mu(\sqrt{n})\int_{\rho^2\le n} dx + \sum_i Mu(2^{i/2}\sqrt{n})\int_{\rho^2\sim 2^i n} e^{-\rho(x)^2/Cn}\,dx\right\} \\
&\le C_3 n^{-1/2}\sum_i 2^{iD/2}\exp(-2^i/C)Mu(2^{i/2}\sqrt{n}).
\end{aligned}$$

If u has sublinear growth, given $\varepsilon > 0$, we have $Mu(\sqrt{2^i n}) \le \varepsilon\sqrt{2^i n}$, $i = 0, 1, \dots$ for n large enough. Hence, we get $\nabla u(e) \le C\varepsilon$, thus $\nabla u(e) = 0$ since $\varepsilon > 0$ is arbitrary. Applying this to the p-harmonic function $u_x(z) = u(xz)$ we obtain that $\nabla u(x) = 0$ for all $x \in G$. Hence, we have proved the following:

THEOREM 6.1. *Assume that G has polynomial volume growth. Let p be a symmetric bounded density of probability satisfying* (12) *and* (6). *Then, any p-harmonic function which has sublinear growth is constant on G.*

In [18], Margulis proved that nonnegative p-harmonic functions are constant when G is nilpotent (hence extending to a noncommutative setting some of the results obtain in [7] by Choquet and Deny). Since nilpotent groups have polynomial volume growth, the following theorem extends Margulis' result. (Note that our arguments are completely different from those of [18]. Also note that [18] offers a description of nonnegative p-harmonic functions even when p is not symmetric nor has compact support).

THEOREM 6.2. *Assume that G has polynomial volume growth. Let p be a symmetric bounded density of probability satisfying* (12) *and* (6). *Then, any p-harmonic function which is bounded from below is constant on G.*

PROOF. Note that the Gaussian estimates (13) and (15) given by Theorem 5.1 imply that there exist an integer a and a constant $C > 0$ such that, for any integer n and any $x, y \in G$ such that $\rho(x^{-1}y) \le \sqrt{n}$, we have $p^{(n)}(x) \le Cp^{(an)}(y)$. Hence, any nonnegative p-harmonic function u satisfies $u(x) = P^n u(x) \le CP^{an}u(y) = Cu(y)$ for $\rho(x^{-1}y) \le \sqrt{n}$. Since n is arbitrary, u is bounded and by Theorem 6.1, u is constant. Of course, the case where u is bounded from below is easily reduced to the above. It is worth noting that the above argument can be applied to prove a Harnack inequality. Namely, we obtain the following result.

THEOREM 6.3. *Assume that G has polynomial volume growth. Let p be a symmetric bounded density of probability satisfying* (12) *and* (6). *There exist an integer a and a constant C such that, for all $x \in G$, all integers n, and any sequence of nonnegative functions u_i, $i = 1, 2, \dots,$ satisfying $u_{i+1} = Pu_i$, we have*

$$\sup_{y \in B(x, \sqrt{n})} \{u_n(y)\} \le C \inf_{y \in B(x, \sqrt{n})} \{u_{an}(y)\}$$

and

$$\sup_{y \in B(x, \sqrt{n})} \{\nabla u_n(y)\} \le C n^{-1/2} \inf_{y \in B(x, \sqrt{n})} \{u_{an}(y)\}.$$

7. Sobolev and isoperimetric inequalities. Recall that, in the Euclidean space, there are two kinds of Sobolev inequalities. The first kind involves the operator $\Delta^{1/2}$ (where Δ is the Laplace operator), and reads

$$\| f \|_{Dp/(D-p)} \le C(D, p) \| \Delta^{1/2} f \|_p, \qquad f \in \mathscr{C}_0^\infty(\mathbb{R}^D),\ 1 < p < D.$$

The second kind involves the gradient instead, as in

$$\| f \|_{Dp/(D-p)} \le C(D, p) \| \nabla f \|_p, \qquad f \in \mathscr{C}_0^\infty(\mathbb{R}^D),\ 1 \le p < D.$$

Of course, in this classical setting, and for $1 < p < +\infty$, the fact that the Riesz transform $\nabla \Delta^{-1/2}$ is bounded on L^p establishes a direct bridge between these two families of inequalities (note that when $p = 2$ the two inequalities are identical). However, the most powerful of the inequalities of the second kind, namely

$$\| f \|_{D/(D-1)} \le C \| \nabla f \|_1, \qquad f \in \mathscr{C}_0^\infty(\mathbb{R}^D),$$

has no full counterpart in the first family. This inequality is of special interest since it is equivalent to the isoperimetric inequality

$$\mathrm{Vol}_n(U)^{(D-1)/D} \le C \,\mathrm{Vol}_{n-1}(\partial U),$$

where U is an open bounded set with smooth boundary ∂U. Because of this, the second family of Sobolev inequalities bears a more geometric meaning.

In the setting of Markov chains, inequalities analogous to those of the first family have been obtained in [25, 9]. Indeed, it is shown in [9] that if k is a bounded symmetric Markov kernel satisfying

$$\sup_{x, y} \{k_n(x, y)\} \le C n^{-D/2}, \quad n = 1, 2, \dots,$$

it follows that

$$\| Kf \|_{Dp/(D-p)} \le C(p) \left\| (I - K)^{1/2} f \right\|_p, \quad f \in L^p,\ 1 < p < D.$$

In this section, we make use of Gaussian estimates obtained in Section 5 to prove some geometric Sobolev inequalities on groups (i.e., inequalities analogous to those of the second family).

Let G be a group having polynomial growth of order D. Denote by Q the operator of convolution associated with the probability density $\xi = |\Omega|^{-1}1_\Omega$, where 1_Ω is the indicator function of Ω which is the fixed generating neighborhood of e used to define the distance ρ and the volume growth function V; see Section 3. For $y \in \Omega$, define ∂_y setting $\partial_y f(x) = f(x) - f(xy) = (I - \delta_{y^{-1}})f(x)$, $x \in G$, where δ_z is the operator of convolution associated with the Dirac mass at z, and consider the operators

$$S_y = Q(I - Q)^{-1}\partial_{y^{-1}} = \sum_1^{+\infty} Q^n \partial_{y^{-1}}$$

and $S_{y,n} f = Q^n \partial_{y^{-1}} f$. Using the symmetry of ξ, we see that $S_{y,n}$ is given by

$$S_{y,n} f(x) = \int \sigma_{y,n}(x^{-1}z) f(z)\, dz,$$

where $\sigma_{y,n}(z) = \partial_y \xi(z)$. Set

$$\sigma_n(z) = \sup_{y\in\Omega}\{|\sigma_{y,n}(z)|\}, \qquad S_n f(x) = \int \sigma_n(x^{-1}z) f(z)\, dz.$$

Note that σ_n satisfies $\sigma_n(x) \le Cn^{-D/2-1/2}\exp(-\rho(x)^2/Cn)$. It follows that $\|\sigma_n\|_\infty \le Cn^{-D/2-1/2}$. Reasoning as in the proof of Lemma 5.2, it also implies $\|\sigma_n\|_1 \le Cn^{-1/2}$. From this we deduce that

$$\|S_n\|_{r-s} \le Cn^{-1/2-D(1/r-1/s)/2}, \qquad n = 1,2,\dots, \tag{22}$$

for $1 \le r \le s \le +\infty$. Set $S = \Sigma_1^{+\infty} S_n$. The operator S is a kind of inverse of the gradient ∇_1. Indeed, the symmetry of Ω implies

$$\frac{1}{2|\Omega|}\int_\Omega S_y \partial_y\, dy = Q(I - Q)^{-1}\left(\frac{1}{2|\Omega|}\int_\Omega (2\delta_e - \delta_y - \delta_{y^{-1}})\, dy\right)$$
$$= Q(I - Q)^{-1}(I - Q) = Q,$$

which yields

$$|Qf| \le \tfrac{1}{2} S \nabla_1 f. \tag{23}$$

LEMMA 7.1. *Fix $1 \le p < D$ and $Dp/(D - p) \le q < +\infty$. We have*

$$|\{x \in G \mid |Sf(x)| > \lambda\}| \le \left(C_{p,q}\lambda^{-1}\|f\|_p\right)^q, \qquad \lambda > 0,\ f \in L^p.$$

PROOF. Assume that $\|f\|_p = 1$ and $q = Dp/(D - p)$. For an integer m to be chosen later, set

$$F_1 = \sum_1^{m-1} S_n f, \qquad F_\infty = \sum_m^{+\infty} S_n f.$$

Using (22) with $r = p$, $s = +\infty$, we see that

$$\|F_\infty\|_\infty \le \sum_{m}^{+\infty} n^{-1/2-D/2p} \le C_1 m^{-D/2q}.$$

Hence, writing $\{|Sf| > \lambda\} \subset \{|F_\infty| > \lambda/2\} \cup \{|F_1 > \lambda/2\}$ and choosing m to be the smallest integer such that $C_1 m^{-D/2q} \le \lambda/2$, we obtain

$$|\{|Sf| > \lambda\}| \le |\{|F_1| > \lambda/2\}| \le \left(C\lambda^{-1}\|F_1\|_p\right)^p.$$

But (22) with $p = r = s$ yields $\|F_1\|_p \le C\sum_1^{m-1} n^{-1/2} \le C'(m-1)^{1/2}$ and

$$|\{|Sf| > \lambda\}| \le \left(C\lambda^{-1}(m-1)^{1/2}\right)^p.$$

Since $m = 1$ for large λ, and $m^{-D/2q} \simeq \lambda$ otherwise, we get $|\{|Sf| > \lambda\}| \le (C\lambda)^{-q}$ for all $\lambda > 0$, which ends the proof of the lemma when $q = Dp/(D-p)$. Replacing D by a smaller number in the above yields the rest of the lemma. □

The Marcinkiewicz interpolation theorem, (23) and Lemma 7.1 yield the following result.

THEOREM 7.2. *Assume that G has polynomial volume growth of order D, and let Q be as above. We have*

$$\|Qf\|_q \le C(D,p,q)\|\nabla_1 f\|_p, \qquad f \in L^p,$$

for all $1 < p < D$ and $Dp/(D-p) \le q < +\infty$. Moreover, if $D/(D-1) < q < +\infty$, we have

$$\|Qf\|_q \le C(D,q)\|\nabla_1 f\|_1, \qquad f \in L^1.$$

An isoperimetric inequality can be obtained as follows. Given an open relatively compact set $A \subset G$, define the "boundary" of A to be the set $\partial A = \{x \in A \mid \text{there exists } y \in \Omega \text{ such that } xy \notin A\}$.

THEOREM 7.3. *Assume that G has polynomial growth of order D, and fix a neighbourhood W of e. There exists a constant C such that*

$$|A|^{(D-1)/D} \le C|\partial A|$$

for any open relatively compact set A that contains zW for some $z \in G$.

PROOF. Fix $\eta \in]0,1[$ and a neighbourhood W' of e such that $Q1_W(y) > \eta$ for $y \in W'$ [such η and W' exist because $Q1_W$ is continuous and $Q1_W(e) > 0$; they can be taken to depend only on Ω and W, which are fixed]. Consider the sets

$$A_{\text{fat}} = \{x \in A, Q1_A(x) > \eta\}, \qquad A_{\text{thin}} = \{x \in A, Q1_A(x) \le \eta\}.$$

Using Lemma 7.1 and (23), we get $|A_{\text{fat}}|^{(D-1)/D} \le C_\eta \|\nabla_1 1_A\|_1$ and it is easy to check that $\|\nabla_1 1_A\|_1 \le C|\partial A|$. We also have $A_{\text{thin}} \subset \partial A$, $A_{\text{fat}} \cup A_{\text{thin}} = A$ and, since A contains zW for some $z \in G$, $|A_{\text{fat}}| \geq C_{\eta, W'}$. Hence, we can conclude that $|A|^{(D-1)/D} \leq C|\partial A|$, which proves the theorem. □

It is worthwhile specializing these results to the case where $G = \Gamma$ is finitely generated. In this case, the operator S_y introduced above can be replaced by $S'_y = (I - Q)^{-1}\partial_{y^{-1}}$, and the result becomes (here $|A|$ is the cardinality of the set $A \subset \Gamma$):

THEOREM 7.4. *Assume that Γ is a finitely generated group of polynomial growth of order D. For all finite sets $A \subset \Gamma$, we have*

$$|A|^{(D-1)/D} \leq C(D)|\partial A|.$$

Moreover, we also have

$$\|f\|_{Dp/(D-p)} \leq C(D,p)\|\nabla_1 f\|_p, \qquad f \in L^p, 1 \leq p < D.$$

To obtain the second statement with $p = 1$, we use the isoperimetric inequality and the proposition of Section 4 in [25] (i.e., we use a kind of co-area formula adapted to the situation).

REMARK. The results stated in Theorem 7.4 are known and due to Varopoulos (see [27], [33]): By a theorem of Gromov (see Section 3), we can restrict ourselves to the case when Γ is nilpotent. Then, we can assume that Γ has no torsion. Thanks to a theorem of Malcev, such a Γ can be embedded as a cocompact lattice in a nilpotent Lie group G, and the isoperimetric inequality on Γ can be deduced from the isoperimetric inequality on G (once this last one has been obtained one way or another; see [33] for more details). The proof given in this paper is more direct, and does not use any structure theorem. Also, see [9] for a simple elementary proof when $G = \mathbb{Z}^D$.

8. Riesz transforms. In the Euclidean space $\mathbb{R}^D$, the Riesz transforms are the operators $(\partial/\partial x_i)\Delta^{-1/2}$, $i = 1, \ldots, D$. They are bounded on L^p for $1 < p < +\infty$. Generalizations of this result in the setting of Lie groups have been studied by several authors. A recent and difficult result of Alexopoulos [2], is that on a Lie group having polynomial growth, the Riesz transforms associated to a family of left invariant vector fields satisfying the Hörmander condition are bounded on L^p, $1 < p < +\infty$. Prior to Alexopoulos' theorem, partial results were obtained in [22] in the same setting. Clearly, analogous questions can be asked in the context of the present paper. Since the results we are able to obtain are not complete, we will be brief and sketchy. Also, for the sake of simplicity, we restrict ourselves to the case of a finitely generated group Γ. In this case, the set $\Omega = \{e, \omega_1, \ldots, \omega_m\}$ is a finite symmetric set of generators. Setting

$$\partial_i f(x) = \partial_{\omega_i} f(x) = f(x) - f(x\omega_i),$$

and $\xi = (1/|\Omega|)1_\Omega$, $Qf = f * \xi$, the Riesz transforms can be defined as the operators $\partial_i(I - Q)^{-1/2}$ and $(I - Q)^{-1/2}\partial_i$, for $i = 1, \dots, m$. Note that $(I - Q)^{-1/2}\partial_\omega$ is the adjoint of $\partial_{\omega^{-1}}(I - Q)^{-1/2}$. It is natural to ask whether or not these operators are bounded on L^p, $1 < p < +\infty$. Equivalently, one can ask whether or not

$$\left\|(I - Q)^{1/2} f\right\|_p \simeq \|\nabla f\|_p, \qquad 1 < p < +\infty,$$

where, according to our previous notation, $\nabla f = \sup_i\{|\partial_i f|\}$ (note that, in this discrete setting, all the different gradients $\nabla, \nabla_2, \nabla_1$ are comparable). Note that the Riesz transforms are obviously bounded on L^2 since we have $\|(I - Q)^{1/2} f\|_2 \simeq \|\nabla f\|_2$. Denote by R_i the operator $\partial_i(I - Q)^{-1/2}$ and by r_i its convolution kernel. We have $r_i = \Sigma_0^{+\infty} a_n \partial_i \xi^{(n)}$, where the a_n's are such that $(1 - x)^{-1/2} = \Sigma a_n x^n$. Hence, we have $|a_n| \le Cn^{-1/2}$, and using the Gaussian estimate (14), we obtain

$$|r_i(x)| \le C\rho(x)^{-D}, \qquad x \in \Gamma.$$

Set $\tilde{\partial}_y f(x) = f(x) - f(yx)$. Using the fact that the operators ∂_y and $\tilde{\partial}_z$ commute, and the same method of proof that we used to obtain (14), we obtain (see [22] for more details in a continuous time setting) $|\partial_i \tilde{\partial}_j \xi^{(n)}(x)| \le C'n^{-1-D/2}\exp(-\rho(x)^2/Cn)$, $x \in \Gamma$, $n = 1, \dots, i$, $j = 1, \dots, m$. From this it is not hard to deduce that

$$\left|\tilde{\partial}_y r_i(x)\right| \le C\rho(y)\rho(x)^{-D-1} \text{ for all } x, y \in \Gamma \text{ such that } 2\rho(y) \le \rho(x).$$

Because of the above estimates on r_i and the fact that R_i is bounded on L^2, we can use the general Calderon–Zygmund theory on spaces of "homogeneous type" developed by Coifman and Weiss to obtain that the R_i's are bounded from L^1 to weak-L^1. Hence, (by interpolation) the R_i's are also bounded on L^p for $1 < p \le 2$ (see [8], Theorem (2.4), page 74). Of course, the adjoints $(I - Q)^{-1/2}\partial_i$ are thus bounded on L^p for $2 \le p < +\infty$. The above approach fails to yield a complete result because the methods used to estimate $|\partial_i \tilde{\partial}_j \xi^{(n)}|$ do not work for $|\partial_i \partial_j \xi^{(n)}|$. For a better understanding of this fact we refer the interested reader to [22] and Alexopoulos' paper [2]. However, there is a case where the above analysis yields complete results. Indeed, if Γ is abelian, the above argument yields estimates on $|\partial_{i_1} \cdots \partial_{i_s} \xi^{(n)}|$, and we obtain the following result.

THEOREM 8.1. *Assume that Γ is a finitely generated abelian group. For each $1 < p < +\infty$ there exist two constants C, C' such that*

$$C^{-1}\|\nabla f\|_p \le \left\|(I - Q)^{1/2} f\right\|_p \le C'\|\nabla f\|_p, \qquad f \in L^p.$$

9. Rate of escape. One of the motivating problems in the study of the uniform decay of convolution powers is to establish whether simple random walks on a given group are recurrent or not. Indeed, the random walk governed by p is recurrent if and only if the series $\Sigma p^{(n)}(e)$ diverges. Clearly, Theorem 4.1 contains more than enough information to settle this question.

For instance, the only finitely generated groups that are recurrent are the finite extensions of $\{0\}$, $\mathbb{Z}$, $\mathbb{Z}^2$ (Varopoulos [27]). Let x_n be the random variable representing the position, after n steps, of the random walk governed by p and started at e. Assume that $V(n) \geq Cn^D$ for some $D > 2$, and that the symmetric bounded probability density p satisfies (6) and (12). Setting $r_n = \rho(x_n)$, the transience of the random walk x_n amounts to the fact that, for any $r > 0$, $\liminf(r_n) \geq r$ almost surely. Our main interest in this section is to generalize a result of Dvoretzky and Erdös. In [10], they proved that if x_n is the simple random walk on $\mathbb{Z}^D$ with $D > 2$ and ψ a decreasing function, $\mathscr{P}(\{\|x_n\| < \psi(n)\sqrt{n}\ \text{i.o.}\})$ equals 0 or 1 according as the series $\Sigma\psi(2^n)^{D-2}$ converges or diverges. In other words, $\liminf(\|x_n\|/\psi(n)\sqrt{n})$ equals 0 or $+\infty$ according to the above test (here, $\|x\|$ is the Euclidean norm of x). For instance, $\lim(\|x_n\|n^{-1/2}(\log(n))^{\varepsilon+1/(D-2)}) = +\infty$ a.s. for any $\varepsilon > 0$, but $\liminf(\|x_n\|\, n^{-1/2}(\log(n))^{1/(D-2)}) = 0$ a.s. Also note that, in this classical setting, the law of the iterated logarithm asserts that

$$\limsup\left(\|x_n\|/\sigma(2n \log\log(n))^{1/2}\right) = 1 \quad \text{a.s.},$$

where σ^2 is the variance of p. In what follows, we prove a weak version of the law of the iterated logarithm and generalize the Dvoretzky–Erdös result in the case of simple random walks on groups of polynomial growth. The proofs are adapted from the classical case, but depend on the estimates given by Theorem 5.1. Some details are given for the sake of completeness.

In order to describe the random walk governed by p, let W be the compact set containing the support of p and set $\mathscr{W} = G \times \mathscr{W}^{\mathbb{N}}$. On $\mathscr{W}$, consider the family of the probability measures $\mathscr{P}_x = \delta_x \otimes \mu^{\otimes\mathbb{N}}$, where $x \in G$ and μ is the measure of density p. When $x = e$, we set $\mathscr{P} = \mathscr{P}_e$. For $w = (x, w_1 \dots) \in \mathscr{W}$, set $x_n = xw_1 \cdots w_n$, $r_n = \rho(x_n)$. The random variable x_n represents the position, after n steps, of the random walk governed by p and started at x. Our main interest here is in the real random variables r_n.

Before considering the case where G has polynomial growth, it is of interest to note that, by the subadditive ergodic theorem, there always exists a real $\alpha \geq 0$ such that $\lim(r_n/n) = \alpha$ $\mathscr{P}$ a.s. Moreover, as explained in [26], $\alpha = 0$ if and only if any bounded p-harmonic function is constant. Of course, if G has polynomial volume growth, $\alpha = 0$, but a stronger result holds.

THEOREM 9.1. *Assume that G has polynomial volume growth. Let p be a symmetric bounded density of probability satisfying* (12) *and* (6). *There exists a positive constant C such that*

$$C^{-1} \leq \limsup(r_n/\lambda_n) \leq C \qquad \mathscr{P}\ a.s.,$$

where $\lambda_n = (n \log\log(n))^{1/2}$.

PROOF. The upper bound follows easily from the Borel–Cantelli lemma and the claim that

$$\mathscr{P}\left(\left\{\sup_{1\leq\nu\leq n}\{r_\nu\} \geq m\right\}\right) \leq C\exp(-m^2/Cn)$$

for all integers n, m. But the Gaussian upper bound (13) shows that

$$\mathscr{P}(\{r_n \geq m\}) \leq C\exp(-m^2/Cn)$$

and this is enough to prove the claim; see [23], Lemma 3, for a proof which can easily be adapted to the present setting. In order to prove the lower bound, write $x_n = x(n)$, $r_n = r(n)$, $\lambda_n = \lambda(n)$, and consider the events

$$A_n = \{\rho(x^{-1}(a^n)x(a^{n+1})) > C^{-1}\lambda(a^{n+1} - a^n)\},$$

where a is a large positive constant. The A_n's are independent, and when C is large enough, the Gaussian lower bound (15) shows that $\Sigma\mathscr{P}(A_n) = +\infty$. Hence, $\mathscr{P}$ almost surely, we have

$$r(a^{n+1}) \geq \rho(x^{-1}(a^n)x(a^{n+1})) - r(a^n) > C^{-1}\lambda(a^{n+1} - a^n) - r(a^n)$$

for infinitely many n, when C is large enough. The desired conclusion follows by using the upper bound and choosing a large enough. Note that the proof of the estimate from above in Theorem 9.1 depends only on the Gaussian upper bound and the Markov property. Hence, this part of Theorem 9.1 also holds for Markov chains governed by a kernel k satisfying (1) and (21) (see Theorem 5.3).

We now pass to the generalization of the Dvoretzky–Erdös result.

THEOREM 9.2. *Assume that G has polynomial volume growth of order $D > 2$. Let p be a symmetric bounded density of probability satisfying* (12) *and* (6). *For any decreasing function $\psi \geq 0$ we have*

$$\mathscr{P}\left(\left\{r_n < \psi(n)\sqrt{n} \text{ for infinitely many } n\right\}\right) = 0 \text{ or } 1$$

according as the series $\Sigma\psi(2^n)^{D-2}$ converges or diverges. In other words, we have

$$\lim\left(r_n/\psi(n)\sqrt{n}\right) = +\infty \quad \mathscr{P}\ a.s. \quad \text{or} \quad \liminf\left(r_n/\psi(n)\sqrt{n}\right) = 0 \quad \mathscr{P}\ a.s.$$

according as the above series converges or diverges.

The proof depends on two lemmas which are applications of Theorem 5.1 and are of some independent interest. In what follows, G and p are as in Theorem 9.2.

LEMMA 9.3. *Fix $x \in G$ and set $N = \rho(x)$. For $n = 1, \dots$ denote by $\mathscr{E}(x, n)$ the probability that the random walk started at e ever enters $x\Omega^n$ (i.e. the ball of center x and radius n). There exists a constant $C > 0$ such that*

$$\min\left\{1, C^{-1}(n/N)^{D-2}\right\} \leq \mathscr{E}(x,n) \leq \min\left\{1, C(n/N)^{D-2}\right\}.$$

PROOF. Consider the function $u(z) = u_{x,n}(z)$ which is equal to the probability that the random walk started at $z \in G$ ever enters the set $x\Omega^n$. It is not hard to check that u is a positive bounded p-superharmonic function (i.e., $u * p \leq u$). Also, $u = 1$ on $x\Omega^n$, and u is p-harmonic outside $x\Omega^n$. Elementary potential theory shows that $u = (I + \Theta)v + h$ where $v = (I - P)u$, $h = \lim_{i \to +\infty} P^i u$ is a bounded p-harmonic function, and $I + \Theta = \Sigma_0^{+\infty} P^i$ is the Green potential operator associated to P. By Theorem 6.1, we know that h is constant. This constant has to be 0 (to see this, one can show that $\lim_{x \to \infty} u(x) = 0$ by using the Gaussian upper bound, or alternatively, one can interpret h has the probability of visiting $x\Omega^n$ infinitely often, and conclude that $h = 0$ because of the transience of the random walk). Moreover, v has support in $x\Omega_n \setminus x\Omega^{n-1}$. Hence, we get

$$u(z) = \Theta v(z) = \int_{x\Omega^n \setminus x\Omega^{n-1}} \Theta(y^{-1}z) v(y)\, dy.$$

This inequality, the fact that $u(x) = 1$, and the two sided estimate on the kernel Θ given by Theorem 5.1 yield

$$\int_{x\Omega^n \setminus x\Omega^{n-1}} v \simeq n^{D-2}.$$

From the same estimate on Θ, we also get when $\rho(x) = N > n$,

$$\mathscr{E}(x,n) = u(e) \simeq N^{-D+2} \int_{x\Omega^n \setminus x\Omega^{n-1}} v \simeq (n/N)^{D-2}.$$

This ends the proof of Lemma 9.3. Note that the proof of each of the bounds stated in Lemma 9.3 depends on a two-sided estimate of the Green kernel. □

Consider now the probabilities

$$\mathscr{F}(x,n,N) = \mathscr{P}_x(\{r_\nu \le n \text{ for some } \nu \ge N\})$$

and

$$\mathscr{F}(x,n,N,M) = \mathscr{P}_x(\{r_\nu \le n \text{ for some } N \le \nu < M\}).$$

LEMMA 9.4. *There exists a constant $C > 0$ such that, for all integers n, N, M and any $x, y \in G$, we have*

$$\mathscr{F}(e,n,N) \ge C^{-1}\big(n/\sqrt{N}\big)^{D-2} \exp(-Cn^2/N),$$

$$\mathscr{F}(x,n,N) \le C\big(n/\sqrt{N}\big)^{D-2},$$

$$|\mathscr{F}(x,n,N,M) - \mathscr{F}(y,n,N,M)| \le C\big(\rho(x,y)/\sqrt{N}\big)\big(n/\sqrt{N}\big)^{D-2}.$$

Hence, there exist a constant $C' > 0$ and an integer M_0 such that, for all integers n, M, M satisfying $n \le \sqrt{N}$ and $M \ge M_0 N$, we have

$$\mathscr{F}(e,n,N,M) \ge C'^{-1}\big(n/\sqrt{N}\big)^{D-2}$$

and

$$|\mathscr{F}(x,n,N,M) - \mathscr{F}(y,n,N,M)| \le C'\big(\rho(x,y)/\sqrt{N}\big)\mathscr{F}(e,n,2N,M).$$

PROOF. The conditional probability that $r_\nu \le n$ for some $\nu \ge N$, knowing that $x_N = z$, is given by the function $u_{e,n}(z)$ introduced in the proof of Lemma

9.3, and is comparable to $\min\{1, (n/\rho(z))^{D-2}\}$. Moreover, we have

$$\mathscr{F}(x, n, N) = \int u_{e,n}(z) p^{(N)}(z^{-1}x)\, dz.$$

Hence, the first part of Lemma 9.4 follows from the Gaussian estimates (13) and (15) of Theorem 5.1 by arguments similar to those of the proof of Lemma 5.2. The Hölder continuity of $\mathscr{F}(x, n, N, M)$ follows from the Gaussian estimate (15) and the fact that

$$\mathscr{F}(x, n, N, M) = \int \mathscr{F}(z, n, N-N', M-N') p^{(N')}(z^{-1}x)\, dz$$

for all $N' < N$ (here we take $N' \sim N/2$). The rest of Lemma 9.4 follows clearly from what we just proved since $\mathscr{F}(x, n, N, M) \geq \mathscr{F}(x, n, N) - \mathscr{F}(x, n, M)$. □

PROOF OF THEOREM 9.2. First, consider the case when the series $\Sigma\psi(2^n)^{D-2}$ converges. Set $A_n = \{r_\nu \leq \psi(\nu)\sqrt{\nu}$ for some $2^n \leq \nu < 2^{n+1}\}$. Since ψ is decreasing, we have $\mathscr{P}(A_n) \leq \mathscr{F}(e, \psi(2^n)2^{(n+1)/2}, 2^n) \leq C\psi(2^n)^{D-2}$. Hence, by the Borel–Cantelli lemma, $\mathscr{P}(\{r_n \leq \psi(n)\sqrt{n}$ for infinitely many $n\}) = 0$.

Suppose now that the series $\Sigma\psi(2^n)^{D-2}$ diverges. Then, we can find a sequence of integers m_i such that

$$\sum_{i=1}^{\infty} \psi(2^{m_i})^{D-2} = +\infty, \qquad \lim_{i\to\infty} (m_{i+1} - m_i) = +\infty. \tag{24}$$

Set $A_n = \{r_\nu \leq \psi(\nu)\sqrt{\nu}$ for some $2^{m_n} \leq \nu < M_0 2^{m_n}\}$, where M_0 is the integer given by Lemma 9.4. Using Lemma 9.4 and (24), we find that $\Sigma_i \mathscr{P}(A_i) = +\infty$. Hence, by a well-known extension of the Borel–Cantelli lemma (see [4], Theorem 6.4), the proof of Theorem 9.2 reduces to the claim that

$$\lim_{k,n\to\infty} \frac{\mathscr{P}(A_k)\mathscr{P}(A_n)}{\mathscr{P}(A_k \cap A_n)} = 1.$$

To prove the claim, consider two integers k, n large enough and such that $k < n$. Set $j = M_0 2^{m_k}$, and

$$l = \inf\{\nu : 2^{m_k} \leq \nu \text{ and } r_\nu \leq \psi(\nu)\sqrt{\nu}\}.$$

Also, define two measures λ, λ_0 by setting, for any measurable set $E \in G$, $\lambda_0(E) = \mathscr{P}(x_j \in E | A_k)$, $\lambda(E) = p(x_j \in E)$. By the Markov property, we have

$$\int \rho(x)\, d\lambda_0(x) = (\mathscr{P}(A_k))^{-1} \sum_{i<j} \int_{\{l=i\}} \int_G p^{(j-i)}(x_i, x)\rho(x)\, dx\, d\mathscr{P}$$

$$\leq \sup\left\{ \int_G p^{(j-i)}(z, x)\rho(x)\, dx,\ i < j,\ \rho(z) \leq Cj^{1/2} \right\} \leq C'j^{1/2}.$$

We also have $\int \rho(x)\, d\lambda(x) \le C' j^{1/2}$. The Markov property and the above yield

$$\begin{aligned}
&|\mathscr{P}(A_n|A_k) - \mathscr{F}(e, n, 2^{m_n} - j, M_0 2^{m_n} - j)| \\
&\quad \le \int |\mathscr{F}(x, n, 2^{m_n} - j, M_0 2^{m_n} - j) \\
&\qquad - \mathscr{F}(e, n, 2^{m_n} - j, M_0 2^{m_n} - j)|\, d\lambda_0(x) \\
&\quad \le C 2^{-m_n/2} \mathscr{F}(e, n, 2^{m_n} - j, M_0 2^{m_n} - j) \int \rho(x)\, d\lambda_0(x) \\
&\quad \le C 2^{-(m_n - m_k)/2} \mathscr{F}(e, n, 2^{m_n} - j, M_0 2^{m_n} - j)
\end{aligned}$$

and similarly,

$$\begin{aligned}
&|\mathscr{P}(A_n) - \mathscr{F}(e, n, 2^{m_n} - j, M_0 2^{m_n} - j)| \\
&\qquad \le C 2^{-(m_n - m_k)/2} \mathscr{F}(e, n, 2^{m_n} - j, M_0 2^{m_n} - j).
\end{aligned}$$

This implies that

$$\lim_{k, n \to \infty} \frac{\mathscr{P}(A_n|A_k)}{\mathscr{F}(e, n, 2^{m_n} - M_0 2^{m_k}, M_0(2^{m_n} - 2^{m_k}))} = 1$$

and

$$\lim_{k, n \to \infty} \frac{\mathscr{P}(A_n)}{\mathscr{F}(e, n, 2^{m_n} - M_0 2^{m_k}, M_0(2^{m_n} - 2^{m_k}))} = 1.$$

Hence,

$$\lim_{k, n \to \infty} \frac{\mathscr{P}(A_k)\mathscr{P}(A_n)}{\mathscr{P}(A_k \cap A_n)} = \lim_{k, n \to \infty} \frac{\mathscr{P}(A_n)}{\mathscr{P}(A_n|A_k)} = 1.$$

This proves our claim, and ends the proof of Theorem 9.2. □

REMARK 1. We want to emphasize the fact that the above proof of Theorem 9.2 depends crucially on the two-sided Gaussian estimate obtained in Theorem 5.1.

REMARK 2. It is worth noting that Theorems 9.1 and 9.2 also hold in the following related context. Let G be a connected Lie group having polynomial growth. Fix a set $L = \{L_1, \dots, L_s\}$ of left invariant vector fields on G, and assume that $L_1, \dots, L_s$ together with their Lie brackets of all order generate the Lie algebra of G (Hörmander condition). Set $\Delta = -\sum_1^s L_i^2$. Consider the heat semigroup $e^{-t\Delta}$, its convolution kernel h_t, and the corresponding Markov Process (Brownian motion) with (continuous) trajectories X_t. There is a natural left invaraint distance function ρ associated with L (see [29], for instance). Set $R_t = \rho(X_t)$. Then, the statements analogous to Theorems 9.1 and 9.2 hold true in this setting. The proofs also are analogous. They rest on Gaussian estimates satisfied by the kernel h_t. The relevant Gaussian upper bound is proved in [29]. The corresponding gradient estimate and lower bound are proved in [22]. In this continuous time setting the Gaussian upper bound is more precise and, in the statement corresponding to Theorem 9.1, we obtain

$$\limsup_{t\to+\infty}\left(\frac{R_t}{(t\log\log(t))^{1/2}}\right)\leq 2.$$

Moreover, when G is a nilpotent Lie group, the lower Gaussian estimate obtained by Varopoulos in [30] can be used to show that

$$\limsup_{t\to+\infty}\left(\frac{R_t}{(t\log\log(t))^{1/2}}\right)= 2.$$

10. Markov chains on homogeneous spaces. The aim of this last section is to generalize some of the preceding results in the context of Markov chains on homogeneous spaces of groups having polynomial volume growth. A nice account of what is known on this subject is given in [24]. A simple example is presented in [16] where a description of all recurrent homogeneous spaces is given in the context of nilpotent finitely generated groups. The work [11] contains a similar description in the case of some Lie groups having polynomial volume growth. Hence, the question of recurrence or transience is well understood in a number of important cases where it has been shown that the answer can be read in the volume growth of the homogeneous space; see [24, 16, 11]. Note however, that the proofs of these results given in [24, 16, 11] are quite intricate. Moreover, none of the above works contains an estimate of the iterated kernels of the Markov chain on the homogeneous space in terms of the volume growth. Using the full strength of our two-sided Gaussian estimate for convolution powers on groups, we are able to establish similar estimates for the iterated kernel of the induced Markov chains on homogeneous spaces. These estimates easily imply the known results on recurrence. They also prove a conjecture made in [24], page 571. Namely, it follows from the estimates obtained below that if G is a connected Lie group having polynomial growth (i.e., of rigid type), and H is a closed subgroup of G, then the homogeneous space $H\setminus G$ is recurrent if and only if it has polynomial volume growth of degree less or equal to 2; see [24, 11].

We now introduce some notation and our basic hypotheses. Given a group G and a subgroup H of G, denote by $M = H\setminus G = \{\bar{g} = Hg,\ g\in G\}$ the set of the right cosets of H in G. In what follows, we assume that G is locally compact, compactly generated, and has polynomial volume growth (hence, G is unimodular). Also, we assume that H is a closed subgroup of G. Note that H is unimodular. Indeed, H is a sum of its open compactly generated subgroups, each of which is of polynomial growth, hence unimodular. This easily implies that H is unimodular. Consequently (see [19] for instance), $M = H\setminus G$ admits an invariant measure. Under these hypotheses, the Haar measures on G and H and the invariant measure on $M = H\setminus G$ can be chosen so that

$$\int_G f(g)\,dg = \int_{H\setminus G}\left(\int_H f(hg)\,dh\right)d\bar{g} \tag{25}$$

for any continuous function f with compact support: see [19].

Now, let p be a bounded probability density on G with associated Markov operator $Pf(x) = \int p(y^{-1}x) f(y)\,dy = \int p(y) f(x^{-1}y)\,dy$. Using (25), we see that p induces a Markov operator and a Markov kernel k on M defined by

$$Kf(\bar{x}) = \int_M k(\bar{x}, \bar{y}) f(\bar{y})\,d\bar{y} = \int_G p(y^{-1}x) f(\bar{y})\,dy,$$

where

$$k(\bar{x}, \bar{y}) = p(x^{-1}Hy) = \int_H p(x^{-1}hy)\,dh.$$

In another language, the projection of the random walk on G associated with p is the Markov chain on M associated with k. If p is symmetric, then k is symmetric as well. In what follows, we assume that p is a bounded symmetric probability density on G which satisfies the hypotheses (12) and (6) introduced in Section 5. Namely, we assume that there is $r_0 > 0$ such that p is supported in $B(e, r_0)$, and that there exists a generating open neighborhood U of e such that $\inf_U\{p\} > 0$.

Recall that G is compactly generated and let ρ be the distance function on G introduced in Section 3 and associated with a fixed open symmetric relatively compact neighbourhood Ω of $e \in G$ which generates G. The induced distance on M is also denoted by ρ and is given by

$$\rho(\bar{x}, \bar{y}) = \inf\{\rho(x^{-1}hy),\ h \in H\}.$$

For $\xi \in M$, define the distance ball $B(\xi, r) \subset M$ by

$$B(\xi, r) = \{\xi \in M,\ \rho(\xi, \zeta) \le r\},$$

and set $V_\xi(r) = |B(\xi, r)|$. It follows from Lemma I.1 in [14] that the volume function satisfies the doubling property

$$V_\xi(2r) \le CV_\xi(r)$$

for all $\xi \in M$ and $r > 1$. Given a bounded function f on M, define the gradient ∇f by

$$\nabla f(\xi) = \sup\{|f(\xi) - f(\zeta)|,\ \rho(\xi, \zeta) \le 1\}.$$

When taking the gradient of a function of two variables ξ, ζ, we use the notation ∇^ξ and ∇^ζ.

In general, it is difficult to transfer information from p to k. However there are certain properties that pass easily to the quotient. Here is a powerful example.

THEOREM 10.1. *Let G, H, M and p, k be as above. There exists an integer a and a constant $C > 0$ such that, for any $\xi \in M$, any integer j and for any sequence of nonnegative functions u_i, $k = 1, 2, \ldots$ on M satisfying $u_{i+1} = Ku_i$, we have*

$$\sup_{B(\xi, \sqrt{j})} \{u_j\} \le C \inf_{B(\xi, \sqrt{j})} \{u_{aj}\}$$

and

$$\sup_{B(\xi,\sqrt{j})}\{\nabla u_j\} \le Cj^{-1/2} \inf_{B(\xi,\sqrt{j})}\{u_{aj}\}.$$

PROOF. Given u_i as in the theorem, set $\tilde{u}_i(x) = u_i(\bar{x})$. Then, $\tilde{u}_i$ is a sequence of nonnegative functions on G and, using (25), we check that $\tilde{u}_{i+1} = P\tilde{u}_i$. Hence, the above theorem follows from Theorem 6.3 and the (easy) fact that the canonical projection of the ball $B(x,r) \subset G$ is equal to the ball $B(\bar{x},r) \subset M$. As a corollary to Theorem 10.1, we obtain:

THEOREM 10.2. *There exist an integer a and two positive constants C, C' such that for any $\xi, \zeta \in M$ and any integer n the iterated kernel k_n satisfies*:

(i) $\sup_{E(\xi,\zeta,n)}\{k_n(\xi',\zeta')\} \le C \inf_{E(\xi,\zeta,n)}\{k_{an}(\xi',\zeta')\}$
where $E(\xi,\zeta,n) = B(\xi,\sqrt{n}) \times B(\zeta,\sqrt{n})$.

(ii) $k_n(\xi,\zeta) \le C \min\{V_\xi(\sqrt{n})^{-1}, V_\zeta(\sqrt{n})^{-1}\}$.

(iii) $\max\{k_n(\xi,\xi), k_n(\zeta,\zeta)\} \le Ck_{an}(\xi,\zeta)\exp(C\rho(\xi,\zeta)^2/n)$,
for $\rho(\xi,\zeta) \le n/C'$.

(iv) $\nabla^\zeta k_n(\xi,\zeta) \le Cn^{-1/2}k_{an}(\xi,\zeta)$.

PROOF. The assertions (i) and (iv) follow easily from Theorem 10.1 applied to $u_i(\zeta) = k_i(\xi,\zeta)$, together with the symmetry of k_i. The assertion (ii) follows from (i) and the fact that $\int_M k_i(\xi,\zeta)\,d\zeta = 1$. Assertion (iii) also follows from (i) and a well-known chaining argument. Note that (ii) implies in particular that

$$k_n(\xi,\xi) \le CV_\xi(\sqrt{n})^{-1}, \qquad \xi \in M, n = 1,2,\dots.$$

In order to obtain a converse inequality, we first claim that there exists a constant $A > 0$ such that

$$\int_{\rho(\xi,\zeta)>\sqrt{An}} k_n(\xi,\zeta)\,d\zeta \le 1/2 \tag{26}$$

for all $\xi \in M$ and all integers n. Indeed, setting $\xi = \bar{x}$, $\zeta = \bar{y}$, we have $k_n(\xi,\zeta) = \int_H p^{(n)}(x^{-1}hy)\,dh$ and

$$\begin{aligned}\int_{\rho(\xi,\zeta)>\sqrt{An}} k_n(\xi,\zeta)\,d\zeta &= \int_M 1_{\{\rho(\bar{x},\bar{y})>\sqrt{An}\}}\int_H p^{(n)}(x^{-1}hy)\,dh\,d\bar{y}\\ &\le \int_M\int_H 1_{\{\rho(x^{-1}hy)>\sqrt{An}\}}p^{(n)}(x^{-1}hy)\,dh\,d\bar{y}\\ &= \int_{\rho(x,y)>\sqrt{An}} p^{(n)}(x,y)\,dy.\end{aligned}$$

Hence, the claim follows from the similar statement on G which has been proved in Section 5. From (26), it follows that

$$\int_{\rho(\xi,\zeta)\le\sqrt{An}} k_n(\xi,\zeta)\,d\zeta \ge 1/2,$$

which together with Theorem 10.2 and the doubling property of the volume, implies

$$\left(CV_\xi(\sqrt{n})\right)^{-1} \le k_n(\xi,\xi) \tag{27}$$

for some constant $C > 0$. □

We now state the main result of this section, which generalizes Theorem 5.1.

THEOREM 10.3. *Let G, H, M and p, k be as above. Then, there exist three positive constants C, C', C'' such that, for any $\xi, \zeta \in M$ and any integer n, we have*

$$k_n(\xi,\zeta) \le CV_\xi(\sqrt{n})^{-1}\exp\left(-\rho(\xi,\zeta)^2/C'n\right), \tag{28}$$

$$\nabla^\zeta k_n(\xi,\zeta) \le Cn^{-1/2}V_\xi(\sqrt{n})^{-1}\exp\left(-\rho(\xi,\zeta)^2/C'n\right), \tag{29}$$

$$k_n(\xi,\zeta) \ge C^{-1}V_\xi(\sqrt{n})^{-1}\exp\left(-C'\rho(\xi,\zeta)^2/n\right) \quad \text{if } \rho(\xi,\zeta) \le n/C''. \tag{30}$$

In the above estimates, $V_\xi(\sqrt{n})$ can be replaced by $V_\zeta(\sqrt{n})$ or by $(V_\xi(\sqrt{n})V_\zeta(\sqrt{n}))^{1/2}$.

PROOF. First, we note that (30) follows from Theorem 10.2 and (27). Also, (29) follows from (28) and Theorem 10.2. Hence, we are left with the task of proving (28). Note that the method of Section 2 does not apply directly here. This is because the behavior of the function $n \to V_\xi(n)$ is not uniform in $\xi \in M$. However, there are at least two ways of proving (28). We choose to present in some detail an approach which combines a technique introduced by Carne in [6] with assertion (i) of Theorem 10.2 (i.e., a Harnack inequality). Carne's idea is to obtain a "Hadamard's transmutation formula" for Markov chains. Namely, let X_n be the symmetric random walk on $\mathbb{Z}$ which starts from $X_0 = 0$ [i.e., $\mathscr{P}(X_{n+1} - X_n = \pm 1) = 1/2$]. Also, let

$$\mathscr{Q}_n(z) = \tfrac{1}{2}\left(\left(z + (z^2-1)^{1/2}\right)^n + \left(z - (z^2-1)^{1/2}\right)^n\right)$$

be the Chebyshev polynomials with $n = 0, 1, \dots$. Then, we have $K^n = \sum_{i=0}^\infty \mathscr{P}(|X_n| = i)\mathscr{Q}_i(K)$; see [6], Theorem 2, page 400. Fix $\xi, \zeta \in M$, two integers n, ν, and set $r = \rho(\xi,\zeta)$. Also, set $B = B(\xi,\nu)$, $B' = B(\zeta,\nu)$. Remark that the function $K^i 1_{B'}$ is supported in $B(\zeta, ir_0 + \nu)$, which implies that $\mathscr{Q}_i(K)1_{B'}$ is also supported in $B(\zeta, ir_0 + \nu)$. Hence, with this notation and for $2\nu \le r$, we can write

$$\langle K^n 1_{B'}, 1_B \rangle = \sum_{i=0}^{\infty} \mathscr{P}(|X_n| = i) \langle \mathscr{Q}_i(K) 1_{B'}, 1_B \rangle$$
$$\leq \left(V_\xi(\nu) V_\zeta(\nu)\right)^{1/2} \sum_{i \geq (r-2\nu)/r_0} \mathscr{P}(|X_n| = i).$$

By a classical estimate (see [6]), we have $\mathscr{P}(X_n \geq s) = \sum_{i \geq s} \mathscr{P}(X_n = i) \leq e^{-s^2/2n}$. Hence, we obtain

$$\langle K^n 1_{B'}, 1_B \rangle \leq 2\left(V_\xi(\nu) V_\zeta(\nu)\right)^{1/2} \exp\left(-(r-2\nu)^2/2r_0^2 n\right). \tag{31}$$

Assume now that $r \geq 4\sqrt{n}$ [otherwise (28) follows directly from Theorem 10.2]. Choose $\nu = \sqrt{n}$ and apply the above with n replaced by an where a is as in Theorem 10.2. Using assertion (i) of Theorem 10.2 to estimate the left-hand side of (31) from below, we obtain

$$k_n(\xi, \zeta) \leq C\left(V_\xi(\sqrt{n}) V_\zeta(\sqrt{n})\right)^{-1/2} \exp(-r^2/C'n),$$

where $C' = 2r_0 a$. This ends the proof of Theorem 10.3. □

An alternative proof of (28) can be outlined as follows: Use Lemma 2.4 to estimate $\langle K_s^n 1_{B'}, 1_B \rangle$, where K_s^n is as in Section 2 and B, B' as above. Then, use assertion (i) of Theorem 10.2 to obtain a pointwise estimate, and choose s as at the end of Section 2.

Recall that the (modified) Green kernel Θ is defined by $\Theta(\xi, \zeta) = \sum_{i=1}^{\infty} k_i(\xi, \zeta)$.

THEOREM 10.4. *Let G, H, M and p, k be as above. There exists a positive constant C such that Green kernel Θ satisfies*

$$C^{-1} \sum_{i \geq \rho(\xi,\zeta)^2} V_\xi(\sqrt{i})^{-1} \leq \Theta(\xi, \zeta) \leq C \sum_{i \geq \rho(\xi,\zeta)^2} V_\xi(\sqrt{i})^{-1},$$

$$\nabla^\xi \Theta(\xi, \zeta) \leq C \sum_{i \geq \rho(\xi,\zeta)^2} i^{-1/2} V_\xi(\sqrt{i})^{-1}.$$

Moreover, $V_\xi(\sqrt{i})$ can be replaced by $V_\zeta(\sqrt{i})$ or by $(V_\xi(\sqrt{i}) V_\zeta(\sqrt{i}))^{1/2}$.

PROOF. The lower bound follows immediately from the lower bound in Theorem 10.3. The upper estimate of Θ follows from (28) and the claim that

$$\sum_{1 \leq i \leq r^2} k_i(\xi, \zeta) \leq C \sum_{i \geq r^2} V_\xi(\sqrt{i})^{-1}$$

when $r = \rho(\xi, \zeta)$. To prove this claim, note that a repeated use of the doubling property of the volume yields

$$V_\xi(j^{1/2})^{-1} \leq CV_\xi\left((r^2+j)^{1/2}\right)^{-1} \exp\left(\sqrt{\frac{r^2+j}{j}}\right),$$

(in fact $\exp(\sqrt{(r^2+j)}/j)$ can even be replaced by a power of $\sqrt{(r^2+j)}/j$). Hence, by Theorem 10.3, we have

$$\sum_{1 \leq i \leq r^2} k_i(\xi, \zeta) \leq C \sup_{j=1,\dots,r^2} \left\{\exp\left(\frac{r}{\sqrt{j}} - \frac{r^2}{C'j}\right)\right\} \sum_{i \geq r^2} V_\xi(\sqrt{i})^{-1}$$
$$\leq C'' \sum_{i \geq r^2} V_\xi(\sqrt{i})^{-1}.$$

The bound on the gradient is obtained similarly. This ends the proof of Theorem 10.4. □

REMARK 1. Assume that k is a left H-invariant kernel on G, that is $k(hx, hy) = k(x, y)$ for all $h \in H$, and that k satisfies assumptions of Theorem 5.3. Then, consider the kernel $\tilde{k}(Hx, Hy) = \int_H k(x, hy)\,dh$. By Theorem 5.3 and Theorem 5.1, we can find a convolution kernel p on G satisfying the hypotheses of Theorem 5.1 and such that $k_n(x, hy) \leq Cp^{an}(x^{-1}hy)$ on G for some constants a, C. Integrating this inequality over H, we deduce from Theorem 10.3 a Gaussian upper bound for the kernel $\tilde{k}$ on M.

REMARK 2. The study of harmonic functions on M follows easily from the results on G. The same is true for the study of Riesz transforms. Theorem 10.3 and the technique of Section 7 yield some Sobolev and isoperimetric inequalities if we assume that $V_\xi(n) \geq cn^D$ for some positive constants c and D. Also, Theorem 10.3 can be used to prove the statement analogous to the upper bound in Theorem 9.1 (law of the iterated logarithm). However, it is not clear what the statement analogous to Theorem 9.2 should be in the present setting.

REMARK 3. There are various interesting questions concerning Markov chains on homogeneous spaces which are still open. For instance, let Γ be a finitely generated solvable group (even polycyclic) which is not almost nilpotent. What can be said about the Markov chains on the homogeneous spaces of Γ? The above is no help since Γ has exponential volume growth. Even the basic question of the recurrence or transience of such a homogeneous space has yet to be understood.

REMARK 4. Some of the techniques used above are similar to the one used by Maheux in [17] where he studies subelliptic heat kernels on nilmanifolds.

REFERENCES

[1] Alexopoulos, G. (1987). Functions harmoniques bornées sur les groupes résolubles. *C. R. Acad. Sci. Paris Sér. I* **305** 777–779.

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

[3] Bass, H. (1972). The degree of polynomial growth of finitely generated nilpotent groups. *Proc. London Math. Soc.* **25** 603–614.

[4] Billingsley, P. (1979). *Probability and Measure*. Wiley, New York.

[5] Carlen, E., Kusuoka, S. and Stroock, D. (1987). Upper bounds for symmetric Markov transition functions. *Ann. Inst. H. Poincaré* **23** 245–287.

[6] Carne, T. K. (1985). A transmutation formula for Markov chains. *Bull. Sci. Math.* (2) **109** 399–405.

[7] Choquet, G. and Deny, J. (1960). Sur l'équation de convolution $\mu = \mu * \sigma$. *C. R. Acad. Sci. Paris* **250** 799–801.

[8] Coifman, R. and Weiss, G. (1971). *Analyse Harmonique Non-Commutative sur Certains Espaces Homogènes*. *Lecture Notes in Math.* **242**. Springer, Berlin.

[9] Coulhon, Th. and Saloff-Coste, L. (1990). Puissances d'un opérateur régularisant. *Ann. Inst. H. Poincaré* **26** 419–436.

[10] Dvoretzky, A. and Erdös, P. (1951). Some problems on random walk in space. *Proc. Second Berkeley Symp. Math. Statist. Probab.* 353–367. Univ. California Press, Berkeley.

[11] Gallardo, L. and Schott, R. (1980). Marches aléatoires sur les espaces homogènes de certains groupes de type rigide. *Journées sur les Marhes Aléatoires. Astérisque* **74** 149–170.

[12] Gromov, M. (1981). Groups of polynomial growth and expanding maps. *Publ. Ihes* **53** 53–78.

[13] Grigorchuk, R. I. (1985). Degrees of growth of finitely generated groups and the theory of invariant means. *Math. USSR-Izv.* **25** 259–300.

[14] Guivarc'h, Y. (1973). Croissance polynômiale et périodes des functions harmoniques. *Bull. Soc. Math. France* **101** 333–379.

[15] Hebisch, W. (1992). On the heat kernel on Lie groups. *Math. Zeit.* **210** 593–606.

[16] Hennion, H. (1976). Marches aléatoires sur les espaces homogènes des groupes nilpotents à génération finie. *Z. Wahrsch. Verw. Gebiete* **34** 245–267.

[17] Maheux, P. (1992). Analyse et géomètrie sur les espaces homogènes. *C. R. Acad. Sci. Paris Sér. I* **314** 441–444.

[18] Margulis, G. A. (1966). Positive harmonic functions on nilpotent groups. *Soviet Math. Dokl.* **7** 241–244.

[19] Raghunatham, M. S. (1972). *Discrete Subgroups of Lie Groups*. Springer, Berlin.

[20] Saloff-Coste, L. (1989). Sur la décroissance des puissances de convolution sur les groupes. *Bull. Sci. Math.* (2) **113** 3–21.

[21] Saloff-Coste, L. (1991). Sobolev inequalities and polynomial decay of convolution powers and random walks. *Proceeding of the 1989 Lisbon Conference Stochastic Analysis and Applications*. 176–189. Birkhäuser, Boston.

[22] Saloff-Coste, L. (1990). Analyse sur les groupe de Lie à croissance polynômiale. *Ark. Mat.* **28** 315–331.

[23] Saloff-Coste, L. and Stroock, D. (1991). Opérateurs uniformèment sous-elliptiques sur les groupes de Lie. *J. Funct. Anal.* **98** 97–121.

[24] Schott, R. (1983). Random walks on homogeneous spaces. *Probability Measures on Groups VII. Lecture Notes in Math.* **1064** 564–575. Springer, Berlin.

[25] Varopoulos, N. Th. (1985). Isoperimetric inequalities and Markov chains. *J. Funct. Anal.* **63** 215–239.

[26] Varopoulos, N. Th. (1985). Long range estimates for Markov chains. *Bull. Sci. Math.* (2) **109** 225–252.

[27] Varopoulos, N. Th. (1986). Théorie du potentiel sur des groupes et des variétés. *C. R. Acad. Sci. Paris Ser. I* **302** 203–205.

[28] VAROPOULOS, N. TH. (1987). Convolution powers on locally compact groups. *Bull. Sci. Math.* (*2*) **111** 333–342.

[29] VAROPOULOS, N. TH. (1988). Analysis on Lie Groups. *J. Funct. Anal.* **76** 346–410.

[30] VAROPOULOS, N. TH. (1990). Small time Gaussian estimates of heat diffusion kernels. II. The theory of large deviations. *J. Funct. Anal.* **93** 1–33.

[31] VAROPOULOS, N. TH. (1991). Analysis and geometry on groups. *Proc. I.C.M., Kyoto, 1990.* **2** 951–957.

[32] VAROPOULOS, N. TH. (1991). Groups of super polynomial growth. *Proceed. I.C.M. Satellite Conf. on Harmonic Anal.* 194–200. Springer, Berlin.

[33] VAROPOULOS, N., SALOFF-COSTE, L. and COULHON, TH. (1993). *Analysis and Geometry on Groups*. Cambridge Univ. Press. To appear.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CHICAGO
CHICAGO, ILLINOIS 60637

LABORATOIRE ANALYSE COMPLEXE ET GÉOMÉTRIE,
TOUR 46-0, 5ÈME ÉTAGE
UNIVERSITÉ PARIS VI
4, PLACE JUSSIEU
75252 PARIS CEDEX 05
FRANCE