Probability on Groups: Random Walks and Invariant Diffusions

Laurent Saloff-Coste

What do card shuffling, volume growth, and Harnack inequalities have to do with each other? They all arise in the study of random walks on groups. Probability on groups is concerned with probability measures and random processes whose properties are dictated in part by an underlying group structure. It is a diverse area where one finds both sophisticated theories and the analysis of concrete problems. Although there are many other fascinating examples, we will focus on random walks and invariant diffusions, which both are processes with independent stationary increments. Random walks proceed by jumps, whereas diffusions have continuous paths. The two share important properties but differ in some aspects, including the nature of the typical underlying group: finitely generated for random walks, connected for diffusions. We will focus on very basic properties of these processes and leave out many developments, some of which can be found in [V+], [W].

Our aim is to present the theory of random walks and invariant diffusions on general groups, with an emphasis on the relations with algebra, analysis, and geometry. By studying these processes we hope to learn something about the underlying group and related objects. For instance, certain properties of the solutions of the Laplace and heat diffusion equations on the universal cover of a compact manifold relate to the behavior of random walks on the fundamental group. From this viewpoint, understanding the basic properties of random walks on large classes of groups is more important than the detailed study of specific examples. The accumulated knowledge concerning the example of d-dimensional Brownian motion (including the beautiful recent advances by Lawler, Schramm, and Werner) serves as a remote, perhaps unreachable, ideal. It provides natural questions, ideas, and insights, but one may have to settle for much less on general groups.

This article has two parts: one treating random walks, the other diffusions. The two parts are related in many ways—at the level of ideas as well as on firmer mathematical ground—and more so than this article can possibly convey. Behind the difference in settings, ranging from the symmetric group S_n to the Lie group $SL_n(\mathbb{R})$ to the infinite dimensional torus $\mathbb{T}^\infty$, there is unity in the problems that are discussed, and any substantial progress in one particular context sheds light on the entire subject.

Random Walks

Let G be a group generated by a finite symmetric set S. That is, $s \in S$ implies $s^{-1} \in S$, and $G = \cup_{n=0}^\infty S^n$. The *Cayley graph* (G, S) has vertex set G and an edge from x to y if and only if $y = xs$ for some $s \in S$. To capture the basic idea of *random walk*, imagine a walker whose position is a vertex of this graph. At each stage the walker takes a step along one of the adjacent edges, choosing uniformly at random from the possibilities. Where will the walker be after n steps?

Laurent Saloff-Coste is professor of mathematics at Cornell University, currently on leave from the Centre National de la Recherche Scientifique (Toulouse, France). His e-mail address is `lsc@math.cornell.edu`.

More generally, given a probability measure p on G, the associated random walk $(X_n)_{n \geq 0}$ proceeds at each step by picking s in G with probability $p(s)$ and moving to $X_{n+1} = X_n s$. The distribution after n steps is the convolution power $p^{(n)}$, where $p * q(x) = \sum_y p(y) q(y^{-1} x)$.

Shuffling Cards

Why would anyone want to study random walks on groups? Maybe simply because everyone uses random walks, just as Molière's Monsieur Jourdain uses prose without realizing it. Indeed, most card-shuffling methods can be modeled as random walks on the symmetric group S_n ($n = 52$), where the shuffling mechanism is interpreted as choosing at random among a certain set of permutations. A single question obviously takes center stage: how many shuffles are needed to mix up the cards? Bayer and Diaconis made the *New York Times* for proving that seven riffle shuffles are necessary and sufficient. Not only is the question broadly appealing, but the mathematics of riffle shuffles is surprisingly rich and beautiful. That such a precise answer can be given is in itself an interesting fact which has been studied and publicized by Diaconis under the name of cut-off phenomenon.

Card shuffling was discussed much earlier in mathematics: for instance, by Poincaré, Borel, and others (see [Ho]). However, the first quantitative theorem is the following result due to Diaconis and Shahshahani concerning random transposition. To describe this process, we imagine the cards laid out neatly in a row on a table. Two cards are picked independently and uniformly at random, and the cards are switched. For random transposition, a sudden convergence to the uniform distribution occurs after about $\frac{1}{2} n \log n$ repetitions, an example of the cut-off phenomenon. For a standard deck of 52 cards this means that about 100 random transpositions are appropriate to mix up the deck. To state a precise result, we consider the *total variation distance* between two probability measures p and q given by $\|p - q\|_{TV} = \sup |p(A) - q(A)|$, where the supremum is over all subsets A of G.

Theorem 1. *For random transposition on the symmetric group S_n, let $p^{(k)}$ be the law after k steps. Let $k(n, c) = \frac{1}{2} n(c + \log n)$. Then there exists a constant A such that for all n and for all $c > 0$,*

$$\|p^{(k(n,c))} - u_n\|_{TV} \leq A e^{-c},$$

where u_n denotes the uniform probability measure on S_n. Moreover, there exist a constant B and a positive function f satisfying $\lim_{c \to 0} f(c) = 0$ such that for all n and for all $c < 0$,

$$\|p^{(k(n,c))} - u_n\|_{TV} \geq 1 - f(c) - B n^{-1} \log n.$$

Adjacent transposition (adjacent cards are transposed) and random insertion (a card is picked at random and inserted at an independent random position) are two other simple examples of shuffling mechanisms that have been studied. To mix up the cards uniformly takes order $n^3 \log n$ shuffles for adjacent transposition and order $n \log n$ shuffles for random insertion. In both cases the exact multiplicative constant is not known, and the existence of a precise cut-off time is an open question. Educated guesses are that it takes about 30,000 adjacent transpositions and a few hundred random insertions to mix 52 cards.

For large n, about three of every four pairs of permutations generate the symmetric group, but one has no clue how many shuffles are typically needed to mix up the cards using such a pair of permutations. Varied techniques have been used in the last twenty years by Aldous, Diaconis, and their many collaborators and followers to understand random walks on the symmetric groups and other finite groups. We will now describe two very different approaches in some detail. For more, see [D].

Fix a given shuffling mechanism. In the probabilistic method known as "coupling", two dependent copies (X_n, Y_n) of the process—the first stationary, the second started from a fixed arbitrary state—are constructed with the property that they agree with higher and higher probability as time evolves. Let T be the random time equal to the smallest n at which X_n and Y_n coincide. This T is called the coupling time, and the total variation distance between the law $p^{(n)}$ of Y_n and the stationary measure u (i.e., the law of X_n) can be bounded by

$$\|p^{(n)} - u\|_{TV} \leq \text{Prob}(T > n).$$

Thus the problem becomes that of constructing a good coupling for which $\text{Prob}(T > n)$ can be estimated. This method has the advantage of not being restricted to random walks on finite groups and is used widely in other contexts.

Representation theory (e.g., of the symmetric group) offers great possibilities when the walk possesses extra symmetries. Studying a random walk on a large finite group can be viewed as the manipulation of a large matrix, namely, the transition probability matrix of the walk. Representation theory helps reduce the size of the problem by providing a partial diagonalization of the matrix into blocks. But the blocks can still have large dimension. For instance, for the symmetric group S_n, the starting matrix has size $n! \times n!$, and after the breakup according to irreducible representations, the largest blocks are still of order $\sqrt{n!} \times \sqrt{n!}$. However, if the walk is invariant under inner automorphisms (i.e., $x \mapsto axa^{-1}$, a typical example of the extra symmetries alluded to above), then each block is a scalar matrix, and refined results can be obtained, as in the case of random transposition. Another very useful approach

involves comparisons of different random walks and elementary combinatoric considerations, including the geometry of paths in the corresponding finite Cayley graphs. For instance, Diaconis and the author have used comparison with random transposition to bound efficiently the number of shuffles needed for adjacent transposition, random insertion, and many other examples.

Nonetheless, results such as Theorem 1 exist only for a small number of specific examples. Although there are satisfactory weaker results for a few larger classes of random walks on finite groups, there is no real global understanding of the behavior of random walks on finite groups, especially for walks based on small sets of generators.

Thus, there are many challenging questions and open problems. One is as follows. In any given graph, the *boundary* ∂A of a set A is the set of all edges connecting A to its complement A^c. A *family of (k, c)-expanders* is an infinite collection of finite graphs for which any vertex has at most k neighbors, and for any subset A,

$$\min\{\#A, \#A^c\} \leq c\#\partial A.$$

These graphs have very good connectivity properties. They are of practical interest as models for communication networks. Random walks on expanders have few local moves but converge rapidly to equilibrium. The first examples of expanders were produced by Margulis using the representation theory of the infinite group $SL_n(\mathbb{Z})$ in the form of Kazhdan's property (T) (see [L]). Whether or not the symmetric groups can yield a family of (k, c)-expanders is an open question.

Before we leap to infinite groups, let us emphasize that random walks on finite and finitely generated groups are related in many ways. Results concerning specific infinite groups (e.g., Kazhdan's property (T)) can lead to interesting results concerning finite quotients; conversely, many infinite groups can be approximated by finite groups. After all, a short-sighted random walker walking on a large finite cyclic group $\mathbb{Z}/N\mathbb{Z}$ will not immediately realize that the group is not $\mathbb{Z}$. A recent success story which illustrates this point is the computation by Grigorchuk and Żuk of the spectral measure (i.e., the measure μ on $[-1, 1]$ such that $p^{(n)}(e) = \int_{-1}^{1} \lambda^n d\mu(\lambda)$) of a random walk on the wreath product $\mathbb{Z}_2 \wr \mathbb{Z}$ (this group is described below in the section on solvable groups). They proceed by approximation by random walks on finite groups. With Linnel and Schick they show that this computation provides a negative answer to a question of Atiyah concerning divisibility properties of the L^2-Betti numbers of coverings of compact manifolds.

The Birth of Random Walks on Groups

For any random walk on a finitely generated group, let $\phi(n)$ denote the probability of being back at the starting point after $2n$ steps. Thus, if p denotes the probability measure driving the walk, we have $\phi(n) = p^{(2n)}(e)$. We assume throughout that the support of p generates the group and that p is finitely supported and symmetric, i.e., satisfies $p(x) = p(x^{-1})$. Because of the symmetry assumption, $\phi(n) = p^{(2n)}(e)$ is a decreasing function of n (the behavior of $p^{(2n+1)}(e)$ is less interesting: for the simple random walk on the integers, $p^{(2n+1)}(e) = 0$ for all n).

A random walk is *recurrent* if, with probability 1, it comes back to its starting point infinitely often. Around 1920 Pólya proved that the simple random walk on the square lattice is recurrent in dimensions 1 and 2 and not in dimensions 3 and higher. Indeed, elementary results from probability theory show that recurrence is equivalent to $\sum_n \phi(n) = +\infty$, and for the d-dimensional square lattice, $\phi(n) \sim c(d)n^{-d/2}$. Understanding the behavior of $\phi(n)$ is the most basic question in random walk theory from our viewpoint.

In 1958, in his Ph.D. thesis, Kesten—guided by a question of Kac concerning the product of random 2×2 matrices—created the subject of random walks on finitely generated groups. In a sequel to his thesis he proved that $\phi(n)$ decays exponentially fast with n if and only if the group is *nonamenable*. A topological group G is *amenable* if there exists a continuous linear functional ν defined on the space of all Borel measurable bounded functions and such that $\nu(f) \geq 0$ when $f \geq 0$, $\nu(\mathbf{1}) = 1$, and $\nu(f_x) = \nu(f)$ where $f_x(y) = f(xy)$. Such a linear functional is called a left-invariant mean. Although amenability relates to the algebraic structure of the group, there is no satisfactory algebraic description of the dichotomy between amenable and nonamenable groups. Just before Kesten's work, Følner characterized amenability in terms of isoperimetry and proved that a group is nonamenable if and only if there is a constant C such that $\#A \leq C\#\partial A$ for any finite set $A \subset G$. These early results well illustrate how random walk theory relates to algebraic and geometric notions.

All Abelian groups and, more generally, all solvable groups are amenable. See Figure 3 below. The free group $\mathbb{F}_k$ on $k \geq 2$ generators and the fundamental group of a two-dimensional orientable surface of genus $g \geq 2$ are nonamenable. Surprising examples of nonamenable groups include some groups all of whose elements have the same finite order (these deep examples are due to Adyan, and the proof uses the co-growth criterion of Grigorchuk). The natural simple random walk on the free group $\mathbb{F}_k$ ($k \geq 2$) has

$$\phi(n) \sim c(k)n^{-3/2}(2\sqrt{k}/(k + 1))^{2n} \text{ as } n \to \infty,$$

but for most random walks on nonamenable groups the exact rate of exponential decay of ϕ, i.e., the spectral radius $\rho = \lim_{n \to \infty} \phi(n)^{1/2n}$, is hard to compute and is not known.
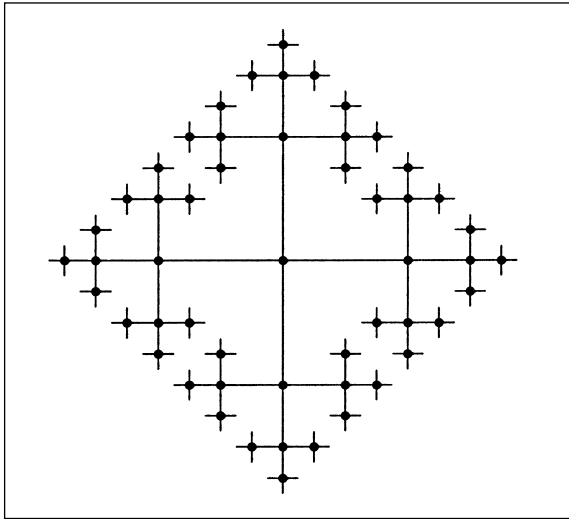
Figure 1. The ball of radius 4 in the free group $\mathbb{F}_2$ on two generators.

For twenty years after Kesten's thesis little progress was made concerning the basic behavior of random walks on finitely generated groups. The conjecture that the only infinite groups that carry recurrent random walks are the finite extensions of $\mathbb{Z}$ and $\mathbb{Z}^2$ became known as Kesten's conjecture. As we shall see, it was solved positively by Varopoulos in the mid 1980s [V+], [W]. An analogous conjecture for connected Lie groups was solved in 1977 by Baldi, Lohoué, and Peyrière using work of Guivarc'h, Keane, and Roynette, but because of the structure theory of Lie groups, this is a rather different story.

Quasi-isometric Invariants

In the 1980s Gromov popularized the notion of *quasi-isometry* between metric spaces and the idea of looking at Cayley graphs of groups as basic geometric objects in their own right. Quasi-isometries are maps that do not distort large distances too much while imposing no restriction on small distances and local topology. For instance, the universal cover of a compact Riemannian manifold and its fundamental group are quasi-isometric objects. Two Cayley graphs corresponding to two different finite generating sets of the same group G are quasi-isometric. A finitely generated group is quasi-isometric to any of its finite index subgroups.

Given a Cayley graph (G, S), the *volume growth function* $V(n)$ is the number of elements in the ball of radius n around the identity element e, that is, the number of elements of the group that can be written as a product of at most n generators. The *isoperimetric profile* is the function

$$I(n) = \inf\{\#\partial A : A \subset G, \ \#A \geq n\}.$$

The behavior at infinity of the volume growth function V and of the isoperimetric profile I are quasi-isometric invariants. A much less obvious example of a quasi-isometric invariant is the behavior of the random walk function ϕ (see [W]). Looking at random walks from this viewpoint turns

out to be very fruitful. A natural question that arises is whether or not these three invariants, V, I, and ϕ, all carry the same information about the group G.

It is plain that the volume growth function does not determine either I or ϕ: all nonamenable groups have exponential volume growth, but there are also many amenable groups with such volume growth. We shall see that, even among amenable groups, the function V does not determine the behavior of I or ϕ and that the relation between the isoperimetric profile I and the probability of return function ϕ is not completely understood. The difficulty in attacking this kind of question comes from the diversity and complexity of the algebraic structures of arbitrary finitely generated groups. This is why Gromov's celebrated theorem asserting that any group whose volume growth is bounded above by a polynomial contains a nilpotent subgroup of finite index is remarkable. For random walks the breakthrough came from the following theorem of Varopoulos. See [V+].

Theorem 2. *Assume that there exists a positive constant c such that $V(n) \geq cn^d$ for all n. Then there are positive constants C_1 and c_1 such that $\phi(n) \leq C_1 n^{-d/2}$ and $I(n) \geq c_1 n^{1-1/d}$ for all n.*

Observe that the hypothesis of this theorem puts very little constraint on the group. On the one hand, this means that one cannot use sophisticated tools to prove such a result. On the other hand, the group structure is essential since there are regular graphs—not arising as Cayley graphs—that have exponential growth from any base point even though the simple random walk is recurrent. The key to both Varopoulos's original proof and the argument outlined below is a simple calculus-type inequality. On any Cayley graph (G, S), for any $y \in G$ and any finitely supported function f,

$$\sum_{x \in G} |f(xy) - f(x)| \leq |y| \sum_{x \in G} |df(x)|,$$

where $|y|$ is the word length of y (the smallest k such that $y = s_1 \cdots s_k$ with $s_i \in S$), and $|df(x)| = \sum_{s \in S} |f(xs) - f(x)|$ is the discrete analog of the gradient. This inequality can be used to prove the following functional inequality involving the inverse function $w(t) = \inf\{n : V(n) > t\}$ of V. Setting $\Psi(t) = Cw^2(8t)$, we have

$$\text{(N)} \qquad \|f\|_2^2 \leq \Psi(\|f\|_1^2 / \|f\|_2^2) \ \|df\|_2^2$$

for any finitely supported function f on G, where ℓ^p-norms are with respect to the counting measure. If $V(n) \geq cn^d$, we find that $\Psi(t) \leq Ct^{2/d}$, and inequality (N) is then analogous to an inequality on $\mathbb{R}^d$ introduced by Nash in his celebrated paper concerning the Hölder continuity of solutions of parabolic equations. Nash used his inequality to control the behavior of certain heat diffusion
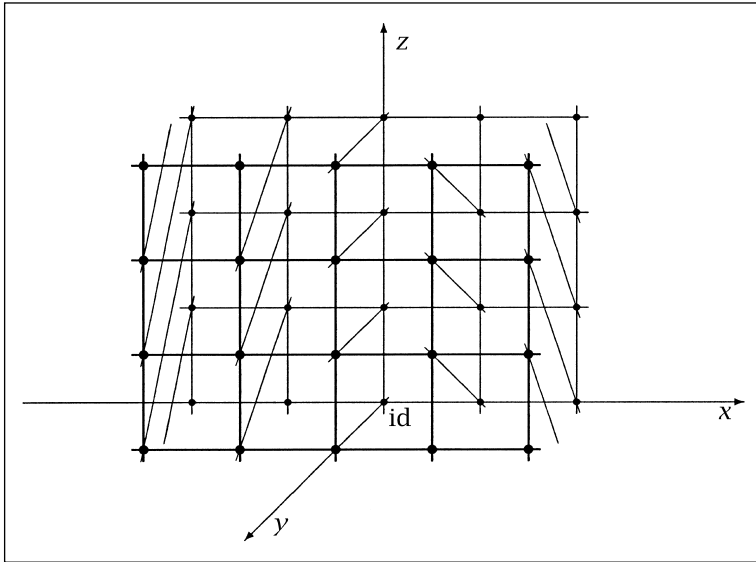
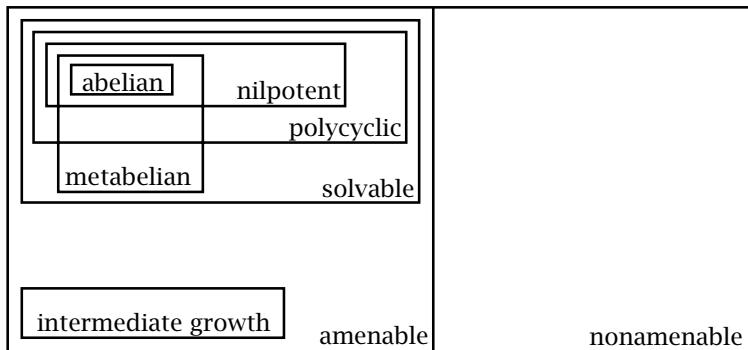**Figure 2. A piece of the Cayley graph of the Heisenberg group.**

Figure 3. A schematic diagram of the inclusion relations between various classes of finitely generated groups.

semigroups. In the present context, (N) leads to the conclusion of Theorem 2 concerning ϕ.

The Polynomial Realm

With a little work, Gromov's polynomial growth theorem and Theorem 2 give a positive solution of Kesten's conjecture: the only finitely generated groups that admit a recurrent random walk are the finite extensions of $\{0\}$, $\mathbb{Z}$, and $\mathbb{Z}^2$. The works of Gromov and Varopoulos are also the main ingredients for the more precise results described below, where for two positive functions, $f(n) \approx g(n)$ means that there are constants c and C such that $0 < c \leq f(n)/g(n) \leq C < +\infty$.

Theorem 3. *For a finitely generated group G, the following are equivalent properties: (1) $V(n) \approx n^d$; (2) $I(n) \approx n^{1-1/d}$; (3) $\phi(n) \approx n^{-d/2}$; (4) G contains a nilpotent subgroup N of finite index, and $d = \sum_i i r_i$, where r_i is the torsion-free rank of the abelian group N_i/N_{i+1}, and (N_i) is the lower central series of N defined by $N_1 = N$, $N_{i+1} = [N, N_i]$.*

Thus, in the polynomial/nilpotent realm, V, I, and ϕ contain essentially the same information. The simplest non-abelian group with polynomial growth is the countable Heisenberg group

$$\mathbb{H} = \left\{ \begin{pmatrix} 1 & y & z \\ 0 & 1 & x \\ 0 & 0 & 1 \end{pmatrix} : x, y, z \in \mathbb{Z} \right\}.$$

It is generated by the four matrices obtained by setting $x = \pm 1$ with $y = z = 0$ and $y = \pm 1$ with $x = z = 0$. The corresponding Cayley graph is shown in Figure 2. It has $V(n) \approx n^4$, $I(n) \approx n^{3/4}$, $\phi(n) \approx n^{-2}$.

Superpolynomial Growth

There are many groups whose volume grows faster than any polynomial. In fact, most groups have this property, even among amenable groups. Hence the following result is a useful complement to Theorem 2.

Theorem 4. *Fix $\alpha \in [0, 1]$. Assume that there exists a positive constant c such that $\log V(n) \geq cn^\alpha$ for all n. Then there are positive constants c_1 and c_2 such that $\log \phi(n) \leq -c_1 n^{\alpha/(\alpha+2)}$ and $I(n) \geq c_2 n/[\log n]^{1/\alpha}$ for all n.*

The bound on ϕ is due to Varopoulos; the isoperimetric bound, to Coulhon and the author. This theorem says that any group with exponential growth has $\log \phi(n) \leq -c_1 n^{1/3}$ and $I(n) \geq c_2 n/\log n$. We shall see below that these bounds are sharp for some groups but not for all. Theorem 4 is also useful for groups of *intermediate growth* whose volume grows faster than any polynomial but slower than any exponential. The existence of such groups was discovered by Grigorchuk in the mid 1980s. Little is known about random walks on these groups, but there is a growing body of work on the structure of a large class of examples (see [BGS]).

Solvable Groups

By a result of Milnor and Wolf, solvable groups have either polynomial or exponential growth. See Figure 3. We start with a very satisfactory result concerning *polycyclic* groups. By a deep structure theorem, polycyclic groups are, up to finite extension, the amenable discrete subgroups of connected Lie groups. Here discrete refers to the topology inherited by the subgroup from the ambient group. Because of their specific algebraic structure, polycyclic groups can be understood quite well, and this yields the following satisfactory result.

Theorem 5. *Let G be an amenable discrete subgroup of a connected Lie group. Then G is finitely generated, and either there exists an integer d such that $V(n) \approx n^d$ or $V(n)$ grows exponentially. In the latter case, $I(n) \approx n/\log n$ and $\log \phi(n) \approx -n^{1/3}$.*

The lower bound on $\log \phi$ is due to Alexopoulos and the upper bound on I to Pittet. The other bounds follow from Theorem 4. One of the simplest examples of a polycyclic group with exponential growth is the semidirect product $\mathbb{Z} \ltimes \mathbb{Z}^2$, whose group operation is defined for $(x, u), (y, v) \in \mathbb{Z} \times \mathbb{Z}^2$ by

$$(x, u) \cdot (y, v) = (x + y, u + A^x v) \text{ with } A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

It is a discrete version of the Lie group Sol, which gives one of the eight geometries used to describe 3-manifolds in Thurston's geometrization program.

Things are different for general solvable groups of exponential volume growth. Among solvable groups the simplest are the *metabelian* groups, the ones whose commutator group is abelian. Even in this class of groups, the behavior of the functions I and ϕ can vary widely. For $\lambda > 1$, let $\mathbb{A}_\lambda$ be the subgroup of the affine group $ax + b$ generated by $u_\pm(x) = x \pm 1$ and $v_\pm^\lambda(x) = \lambda^{\pm 1} x$. These groups are metabelian and have exponential volume growth. They are not discrete in $ax + b$, and most are not polycyclic. When λ is an integer, $\mathbb{A}_\lambda$ can be presented as $\langle a, b : aba^{-1} = b^\lambda \rangle$ with $a = v_+$ and $b = u_+$ (these are also known as Baumslag-Solitar groups).

Figures 4 and 5 describe the Cayley graph of $\mathbb{A}_2$. When λ is algebraic, $\mathbb{A}_\lambda$ has $I(n) \approx n/\log n$ and $\log \phi(n) \approx -n^{1/3}$. When λ is transcendental, $\mathbb{A}_\lambda$ is isomorphic to the wreath product $\mathbb{Z} \wr \mathbb{Z}$ and has $\log \phi(n) \approx -n^{1/3}(\log n)^{2/3}$.

In the study of metabelian groups the wreath products $\mathbb{Z}_2 \wr \mathbb{Z}$ play an important role. These are also known as "lamplighter groups". Imagine $\mathbb{Z}^d$ as the map of a (multidimensional and infinite) American city. At each street crossing there is a lamp which can be on or off (only finitely many lamps can be on). In addition, a lamplighter is wandering around the city turning lamps on or off. An element of $\mathbb{Z}_2 \wr \mathbb{Z}^d$ can be thought of as a "scenery" formed by the lamps and the lamplighter, standing somewhere. See Figure 6. Note that this picturesque description fails to capture how two elements are multiplied. Nevertheless, the basic moves of a natural random walk on $\mathbb{Z}_2 \wr \mathbb{Z}^d$ can be described as the $2d$ possible steps of the lamplighter to adjacent nodes together with the action of turning on or off the lamp at the current node. A theorem of Donsker and Varadhan asserts that N_n, the number of points visited by the simple random walk on $\mathbb{Z}^d$ up to time n, satisfies $\log \mathbb{E}\left(e^{-\lambda N_n}\right) \sim -c(\lambda, d)n^{d/(d+2)}$ as $n \to \infty$, where $\mathbb{E}$ denotes the expectation. It turns out that this is just what one needs to prove that $\log \phi(n) \approx -n^{d/(d+2)}$ on the lamplighter group $\mathbb{Z}_2 \wr \mathbb{Z}^d$ [PSa].

The examples above show that for finitely generated (in fact, with additional arguments, for finitely presented) metabelian groups of
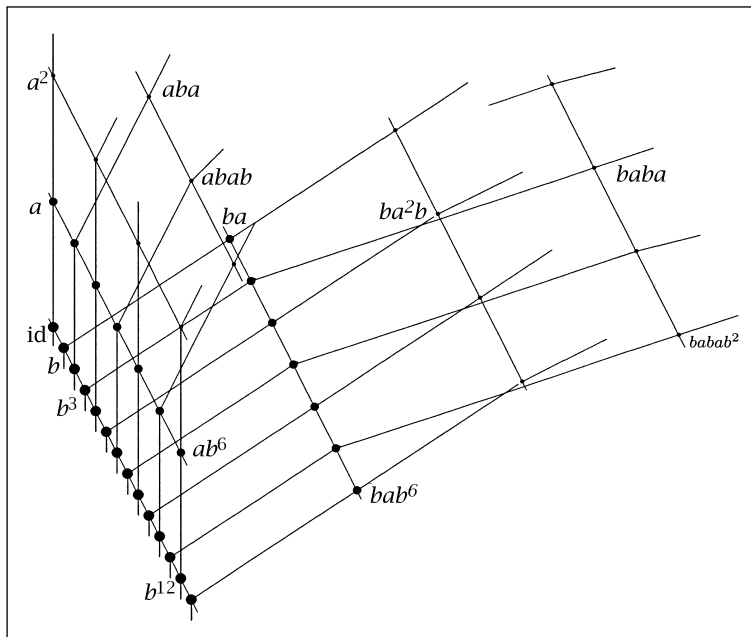

Figure 4. A piece of the Cayley graph of $\mathbb{A}_2 = \langle a, b : aba^{-1} = b^2 \rangle$. **This Cayley graph has** $\log \phi(n) \approx -n^{1/3}, I(n) \approx n/\log n$.
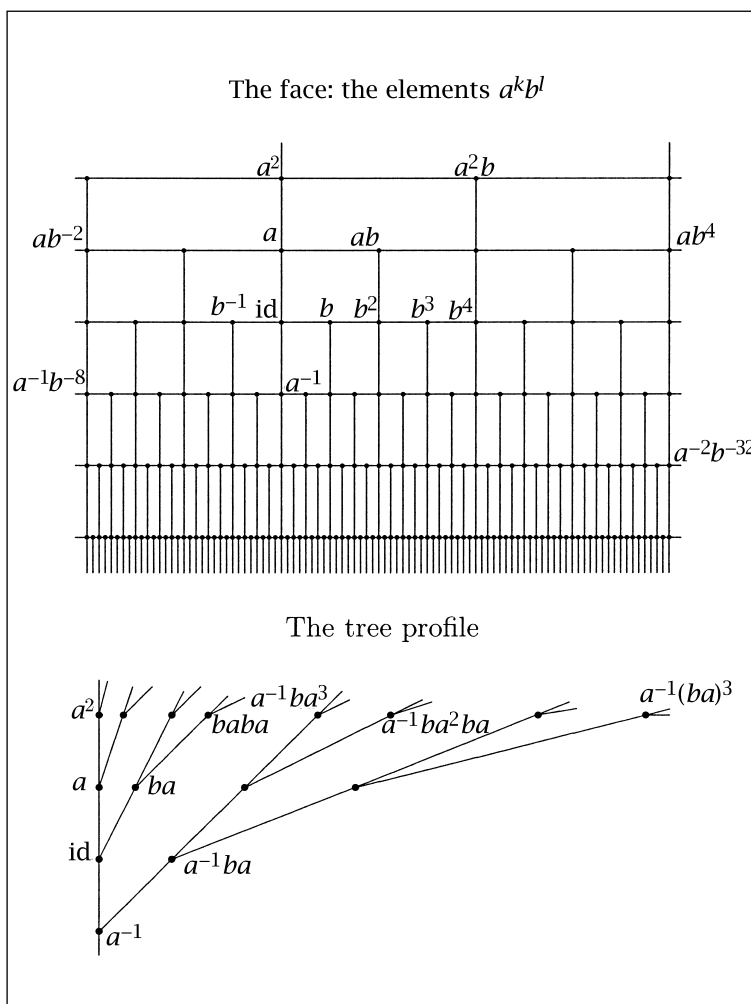

Figure 5. Face and profile of $\mathbb{A}_2 = \langle a, b : aba^{-1} = b^2 \rangle$.
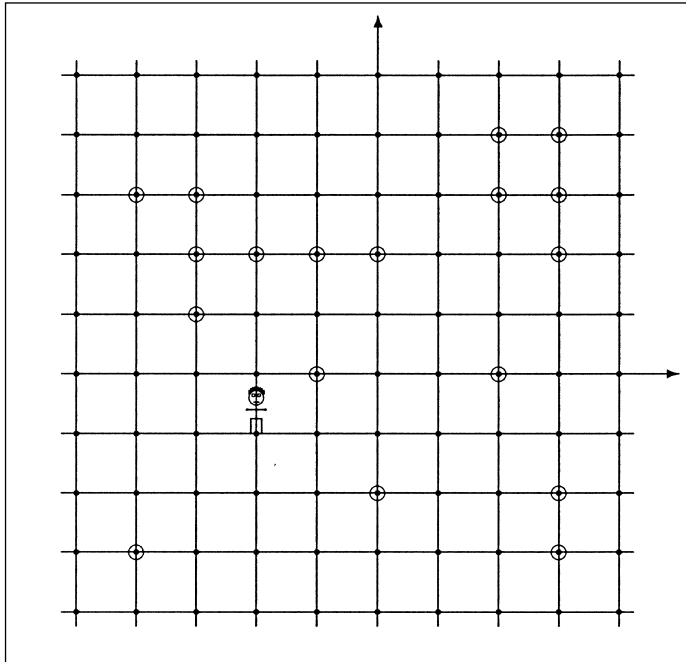
Figure 6. An element of the 2**-dimensional lamplighter group** $\mathbb{Z}_2 \wr \mathbb{Z}^2$**.**

exponential growth, there are infinitely many different behaviors for ϕ. The same is true for the isoperimetric profile I, and Erschler(Dyubina) is developing new precise isoperimetric bounds for wreath products in a promising work in progress. Two related challenging open problems concerning finitely generated metabelian groups are: (a) classify all the possible behaviors of ϕ and I; (b) does the behavior of ϕ determine the behavior of I and vice versa? Thus, despite some remarkable achievements, a complete understanding of the behavior of random walks and isoperimetry on finitely generated groups is still very much out of reach, even for solvable groups. This contrasts with the results we are about to describe concerning invariant diffusions on connected Lie groups, where thanks to the existence of a simpler and more satisfactory structure theory, a complete picture has emerged.

Invariant Diffusion Processes

Let us now change the setting and consider left-invariant diffusion processes on locally compact connected groups. Brownian motion on $\mathbb{R}^d$ is the classical and most studied example. These processes can be characterized in different ways, but they have the crucial properties of having independent stationary increments and continuous paths. Equivalently, by a theorem of Hunt their infinitesimal generators are second-order differential operators that can be written in the form

$$L = \sum_i X_i^2 + X_0,$$

where the X_i's are left-invariant vector fields and can thus be viewed as elements of the Lie algebra.

This makes sense even if G is not a Lie group, because locally compact connected groups are projective limits of Lie groups (see, e.g., [H]). The parallel with random walks is striking, the role of the generators being played now by the X_i's.

A family of probability measures $(\mu_t)_{t>0}$ forms a *convolution semigroup* if $\mu_t * \mu_s = \mu_{t+s}$ and $\mu_t \to \delta_e$ as $t \to 0$. Such a semigroup is *Gaussian* if $t^{-1}\mu_t(G \setminus U) \to 0$ as $t \to 0$ for each neighborhood U of e. This last property is equivalent to the continuity of the sample paths of the associated stochastic process. For a left-invariant diffusion process $Z = (Z_t)$ on a group G, the laws μ_t of Z_t, $t > 0$, form a Gaussian convolution semigroup such that the function $u(t,x) = \int_G f(x,y)d\mu_t(y)$ solves the heat diffusion equation $(\partial_t - L)u = 0$, $u(0,x) = f(x)$.

To complete this picture with a geometric perspective, one introduces a natural "distance function" $d(x,y)$ (allowing the value ∞) called the *intrinsic distance* or *Carnot-Carathéodory distance* and defined by

$$d(x,y) = \sup \{f(x) - f(y) :$$
$$f \in C^\infty(G), \Gamma(f,f) \le 1\},$$

where $\Gamma(f,f) = \frac{1}{2}(Lf^2 - 2fLf) = \sum_i |X_i f|^2$ is the "carré du champs". This definition is more general but essentially equivalent to others based on suitable notions of length of paths. In particular, if G is a Lie group and L is the Laplace-Beltrami operator of a left-invariant Riemannian structure, then d equals the Riemannian distance. The corresponding volume growth function $V(t)$ is defined as the volume of any metric ball of radius t with respect to the left-invariant Haar measure on G.

The main questions concerning these diffusions are: Does μ_t have a smooth density with respect to Haar measure? And if it does, what is the behavior of this density? How does this behavior relate to properties of the distance function d, to the volume growth function, to the family of vector fields (X_i)? How does this relate to the algebraic structure of the group G? Assuming that $(\mu_t)_{t>0}$ admits a continuous density $x \mapsto \mu_t(x)$, the value $\mu_t(e)$ at the origin is the exact analog of the probability of return $\phi(n)$ in the first part of the article, and the most basic question concerns the behavior of $\mu_t(e)$ as t tends either to zero or to infinity.

For brevity we will concentrate on the case where $L = \sum X_i^2$, i.e., $X_0 = 0$. The case where $X_0 \ne 0$ is interesting and requires both additional arguments and some different ideas, but this is more of a technical matter.

Local Theory

Let G be a connected Lie group of dimension n. The natural hypothesis in this context is that the family (X_i) generates the Lie algebra of G. This means that the X_i's, together with their commutators of

all orders, linearly span the Lie algebra. We always make this hypothesis. For the second-order differential operator L, it corresponds to the celebrated subellipticity condition of Hörmander. The local theory that we are about to describe can (and should) be viewed as a model for the deeper and more difficult study of general subelliptic second-order differential operators. The geometry of the distance d is an area of research in its own right under the name of sub-Riemannian geometry and is closely related to control theory. In some sense, the group structure is irrelevant here, although it leads to significant simplifications (see [V+]).

Under Hörmander's condition, μ_t has a smooth positive density for all $t > 0$, the distance d is Hölder continuous with respect to any fixed locally Euclidean distance, and the operator L is hypoelliptic. There exists an integer $m \in [n, 1 + \binom{n}{2}]$, depending on the family (X_i), such that $\mu_t(e) \sim ct^{-m/2}$ as $t \to 0$. This m is also characterized by the fact that $V(t) \sim bt^m$ as $t \to 0$. Much like for harmonic functions in Euclidean space, there exists a constant $C = C_L$ such that, for any $r \in (0, 1)$ and for any positive solution of $Lu = 0$ in the d-ball $B(x, r)$, we have

(GH) $$\sup_{B(x,r/2)} u \le C \inf_{B(x,r/2)} u.$$

The geometric nature of this *Harnack inequality* makes it a powerful tool and illustrates the role played by the distance d.

Thus, under Hörmander's condition, symmetric Gaussian semigroups on Lie groups are very well behaved. Before discussing their large-scale and long-time behavior, we briefly consider what happens locally when G is not a Lie group. This case illustrates in a highly nontrivial way the general theory of analysis and geometry on Dirichlet spaces. Simple-minded but already interesting examples are the product of countably many circle groups and the product of countably many orthogonal groups in different dimensions. In such cases, can $(\mu_t)_{t>0}$ have a nice continuous density for all $t > 0$? Although the theory of such Gaussian semigroups is developed in [H], this question is not answered there. It is natural to focus (at least at first) on bi-invariant, i.e., central, Gaussian semigroups on compact groups. In fact, there are many interesting and challenging open questions already in the case of the infinite-dimensional torus $\mathbb{T}^\infty$, where the infinitesimal generator can simply be written $L = \sum_{i,j} a_{i,j} \partial_i \partial_j$, and this infinite sum is easily interpreted as acting on functions depending only on finitely many coordinates.

A recent result of Bendikov and the author is that any compact, connected, locally connected, metrizable group G carries a host of central Gaussian semigroups having a smooth continuous positive density with respect to Haar measure. The quantity $\mu_t(e)$ can explode in many different ways as t tends to zero, including behaviors such as $e^{[\log 1/t]^{1+\lambda}}$, $e^{t^{-\lambda}}$, $e^{e^{t^{-\lambda}}}$, etc., with $\lambda > 0$.

A sufficient but far-from-necessary condition for $(\mu_t)_{t>0}$ to have a continuous density is that the associated intrinsic distance be continuous. This condition also implies an elliptic Harnack inequality. Namely, if d is continuous, then for any domain Ω and any compact set K in Ω, there exists a constant $C(\Omega, K)$ such that any positive continuous solution of $Lu = 0$ in Ω satisfies

(H) $$\sup_K u \le C(\Omega, K) \inf_K u.$$

Observe that the geometric nature of the inequality has been lost here in the sense that one does not know how to make the constant $C(\Omega, K)$ scale invariant by choosing the pair (Ω, K) to be suitable concentric balls as in (GH). The surprising fact that such a Harnack inequality can hold in infinite dimensions was discovered in the mid-1970s by Bendikov and Berg independently in their Ph.D. theses. Remarkably, inequality (H) can be characterized in terms of the behavior of $\mu_t(e)$.

Theorem 6. *Let L be the infinitesimal generator of a central symmetric Gaussian semigroup $(\mu_t)_{t>0}$ on a compact connected group G. Then L satisfies the elliptic Harnack inequality (H) if and only if $\log \mu_t(e) = o(1/t)$ as $t \to 0$ [BS].*

One of the crucial ingredients in the proof of Theorem 6 is a study of bi-invariant diffusions on compact simple *Lie* groups that brings out the role played by the dimension and that involves small-, medium-, and large-time behaviors.

Long-Time Behavior on Lie Groups

We now turn to noncompact connected Lie groups and discuss the long-time behavior of $\mu_t(e)$ as t tends to infinity, under the standing condition that the family (X_i) generates the Lie algebra. This long-time behavior is really the heart of the matter, since it is where the group structure plays the most significant part.

To start, the behavior of the volume growth function V at infinity is independent of the choice of the family (X_i). Guivarc'h proved in the early 1970s that the volume growth at infinity is either exponential or comparable to a power function whose exponent D is an integer depending only on the underlying group. The groups for which V has a polynomial behavior are called groups of type (R), for rigid. The *adjoint representation* of G is obtained by lifting the action of the inner automorphisms $x \mapsto axa^{-1}$ to the Lie algebra. Groups of type (R) can be characterized algebraically in terms of the adjoint representation, whose eigenvalues must be purely imaginary. The group $U(m)$ of all $m \times m$ unipotent upper-triangular real matrices is of type (R) and has $V(t) \approx t^D$ with $D = \frac{1}{6}(m-1)m(m+1)$.

Groups of type (R) are amenable and unimodular, that is, have bi-invariant Haar measures, but there are many amenable unimodular Lie groups of exponential growth (hence, not of type (R)). The simplest such group is the group Sol mentioned above in the section on random walks on solvable groups. Sol can be described as the semidirect product of $\mathbb{R}^2$ by $\mathbb{R}$ with the action given by multiplication by $\begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$.

The behavior of $\mu_t(e)$ on amenable unimodular Lie groups is described by the following theorem.

Theorem 7. *For any amenable unimodular connected Lie group, if $V(t) \approx t^D$ as $t \to \infty$, then $\mu_t(e) \approx t^{-D/2}$ as $t \to \infty$. If V is exponential, then $\log \mu_t(e) \approx -t^{1/3}$ as $t \to \infty$.*

The two-sided bound under polynomial growth is due to Varopoulos. In the exponential growth case the lower bound is due to Alexopoulos and the upper bound to Varopoulos, with independent distinct proofs of the latter by Hebisch and Robinson. See [V+]. Theorem 7 is analogous to Theorem 5 and can be complemented by a statement concerning the isoperimetric profile, as in Theorem 5 (see [P], [V+]). The two results, for random walks and for diffusions, emerged simultaneously and can be proved by similar methods. A recent work of Alexopoulos [A] complements Theorem 7 with long-time asymptotics on groups with polynomial volume growth, i.e., groups of type (R). Alexopoulos's approach, which is tightly connected to the algebraic structure of type (R) groups, is adapted from techniques and ideas of the area of PDE known as homogenization theory, which deals with the large-scale behavior of differential operators having periodic coefficients in $\mathbb{R}^n$.

Of course, for nonamenable groups, $\mu_t(e)$ decays exponentially fast at a rate described by the *spectral gap* λ of $L = \sum X_i^2$, which is defined as the infimum of the Raleigh quotient

$$\frac{\int_G \sum |X_i f|^2 d\nu}{\int_G |f|^2 d\nu}, \ \ f \neq 0, \ f \in L^2(G, \nu),$$

where ν is a right-invariant Haar measure on G. The spectral gap λ vanishes if and only if G is amenable. It should be noted that, in sharp contrast with the case of finitely generated groups, there exists a satisfactory structure theory that distinguishes between amenable and nonamenable groups in the class of connected locally compact groups [Pa]. Typical nonamenable connected Lie groups are all the semisimple noncompact Lie groups such as $SL_n(\mathbb{R})$, and the connected component of the identity in $SO(p,q)$. One of the early results concerning diffusions on Lie groups is the local central limit theorem of Bougerol, which gives, for semisimple Lie groups, a precise asymptotic result of the form $\mu_t(e) \sim c t^{-a/2} e^{-\lambda t}$ as $t \to +\infty$ for some integer $a \geq 3$ and $\lambda > 0$ as above. In general, such

a precise result is hard to obtain. In either the commutative or the semisimple case, representation theory is the tool of choice for this purpose, but for other groups, including nilpotent and solvable groups, representation theory fails to a large extent to provide useful information about the behavior of $\mu_t(e)$.

For many years precisely understanding the long-time behavior of $\mu_t(e)$ in full generality seemed hopeless, despite the structure theory of Lie groups. However, in the last ten years Varopoulos has developed a theory that describes what happens for any symmetric Gaussian semigroup on any connected real Lie group, amenable or not, unimodular or not. The form of the main result is similar to Theorem 7, but the proofs are quite different. The proof of Theorem 7 is mostly analytic in nature, whereas the proof of Theorem 8 below also involves probability, algebra, and geometry.

Varopoulos [Vb] separates real connected Lie groups into two classes, (B) and (NB). This algebraic classification is too involved to describe here precisely. All (noncompact) semisimple groups, e.g., $SL_n(\mathbb{R})$, are in (NB). In the case of amenable groups, this classification reduces to a simpler one, (C) versus (NC), which can be understood in terms of the adjoint representation and the geometry of (generalized) roots [Va]. The class (R) of rigid groups coincides exactly with the class of those (NC) groups that are unimodular. Further examples of (NC) groups are the groups AN coming from the KAN Iwasawa decomposition of semisimple groups, for instance the group $ax + b$. To describe the simplest family of examples where both (C) and (NC) groups arise, let $\mathbb{S}_\ell = \mathbb{R} \ltimes_\ell \mathbb{R}^2$, where $\ell = (\ell_1, \ell_2) \in \mathbb{R}^2$ and the product is given by

$$(x, u) \cdot (y, v) = (x + y, u + A_\ell^x v)$$

$$\text{where } A_\ell = \begin{pmatrix} e^{\ell_1} & 0 \\ 0 & e^{\ell_2} \end{pmatrix}.$$

Then $\mathbb{S}_\ell$ is of type (C) if $\ell_1 \ell_2 < 0$ and of type (NC) if $\ell_1 \ell_2 > 0$.

Varopoulos's main result describes the classes (B) and (NB) (hence also the classes (C) and (NC)) in terms of the long-time behavior of $\mu_t(e)$ and classifies all the possible behaviors.

Theorem 8. *(1) For groups of type (NB), for each $L = \sum X_i^2$ there exists a nonnegative real number a (which may depend on L) such that $\mu_t(e) \approx t^{-a} e^{-t\lambda}$ as $t \to \infty$. (2) For groups of type (B), $\log(e^{t\lambda} \mu_t(e)) \approx -t^{1/3}$ as $t \to \infty$. Here λ denotes the spectral gap of the corresponding operator L.*

The factors t^{-a} and $e^{-t^{1/3}}$ appearing respectively in the (NB) and (B) cases can be interpreted in terms of the probability that a certain Euclidean Brownian motion stays in a certain convex region up to time t. The exact nature of the Brownian motion (i.e., its covariance matrix) is determined

by the algebraic structure of the group and by L. The convex region is determined by the geometry of the roots. It is compact or not depending on whether the group is (B) or (NB), and this accounts for the $e^{-t^{1/3}}$ versus polynomial behavior. A precise knowledge of the covariance matrix of the Brownian motion and of the convex region above are necessary to determine the constant a in the (NB) case. In fact, typically, the exact value of a is very hard to compute and can vary continuously with L.

These results extend to give a description of the behavior of the convolution powers of any continuous compactly supported symmetric nonnegative function f. This behavior precisely mimics the behavior of symmetric Gaussian convolution semigroups depending on whether the group is (B) or (NB). When expressed in terms of convolution powers, the result can be formulated in a straightforward way in the setting of locally compact connected groups. The restriction that G be connected is essential, as shown by the finitely generated groups $\mathbb{Z}_2 \wr \mathbb{Z}^d$ discussed in the section on solvable groups.

To conclude, there is a geometric description of the classes (B) and (NB) which adds a final touch to this remarkable classification [Vc]. It involves *filling invariants*. These have been considered in various contexts, in particular by Gromov. The 2-dimensional filling invariant $\psi_2(t)$ of a simply connected Riemannian manifold is defined as follows. For any given loop of length at most t, consider all immersed disks having this loop as their boundary, and find the infimum of the areas of all such disks. Then $\psi_2(t)$ is the supremum of these infimal areas over all such loops. For each dimension $k = 2, \ldots, n-1$, where n is the topological dimension of the manifold, there is a k-filling invariant. In particular, $\psi_{n-1}(t)$ gives the largest possible volume of a compact set with smooth boundary of $(n-1)$-volume at most t and is closely related to the isoperimetric profile. Essentially, a group is (NB) if and only if all its filling invariants are bounded above polynomially, whereas a group is a (B) group if and only if at least one of its filling invariants is growing faster than any polynomial. Thus, for connected real Lie groups, one has three equivalent classifications: the analytic/probabilistic classification according to the long-time behavior of symmetric Gaussian semigroups, the geometric classification in terms of filling invariants, and the (B) versus (NB) algebraic classification. There is no doubt that these fundamental results will lead to further progress concerning invariant diffusions, harmonic analysis, and geometry on Lie groups.

References

[A] G. ALEXOPOULOS, Centered sub-Laplacians on Lie groups of polynomial volume growth, *Mem. Amer. Math. Soc.*, to appear.

[BGS] L. BARTHOLDI, R. GRIGORCHUK, and Z. ŠUNIĆ, *Branch Groups,* Handbook of Algebra, vol. 3 (M. Hazewinkel, ed.), North-Holland, Amsterdam, 2001.

[BS] A. BENDIKOV and L. SALOFF-COSTE, Central Gaussian semigroups of measures with continuous density, *J. Funct. Anal.*, to appear.

[D] P. DIACONIS, *Group Representations in Probability and Statistics*, IMS Lecture Notes Monograph Ser., Hayward, CA, 1986.

[H] H. HEYER, *Probability Measures on Locally Compact Groups*, Springer-Verlag, Berlin-Heidelberg-New York, 1977.

[Ho] M. HOSTINSKY, *Méthodes Générales du Calcul des Probabilités*, Gauthier-Villars, Paris, 1931.

[L] A. LUBOTZKY, *Discrete Groups, Expanding Graphs and Invariant Measures*, Birkhäuser, Basel, 1994.

[Pa] A. PATERSON, *Amenability,* Math. Surveys and Monographs, vol. 29, Amer. Math. Soc., Providence, RI, 1988.

[P] CH. PITTET, The isoperimetric profile of homogeneous Riemannian manifolds, *J. Differential Geom.* **54** (2000), 255–302.

[PSa] CH. PITTET and L. SALOFF-COSTE, Amenable groups, isoperimetric profiles and random walks, *Geometric Group Theory Down Under, Canberra 1996* (J. Cossey et al., eds.), de Gruyter, Berlin, 1999, pp. 293–316.

[Va] N. VAROPOULOS, Diffusion on Lie groups I, II, III, *Canad. J. Math.* **46** (1994), 438–448, 1073–1093; **48** (1996), 641–672.

[Vb] ____ , Analysis on Lie groups, *Rev. Mat. Iberoamericana* **12** (1996), 791–917.

[Vc] ____ , A geometric classification of Lie groups*, Rev. Mat. Iberoamericana* **16** (2000), 49–136.

[V+] N. VAROPOULOS, L. SALOFF-COSTE, and T. COULHON, *Analysis and Geometry on Groups,* Cambridge Tracts in Math., vol. 100, Cambridge University Press, Cambridge, 1992.

[W] W. WOESS, *Random Walks on Infinite Graphs and Groups*, Cambridge Tracts in Math., vol. 138, Cambridge University Press, Cambridge, 2000.