RANDOM WALKS ON LIE GROUPS

E. BREUILLARD

The goal of these notes is to give an introduction to random walks and limit theorems on Lie groups, mostly amenable Lie groups, with an emphasis on equidistribution problems. We also state a number of open problems.

In Section 1 we define the basic notions regarding probability measures on groups, convergence in distribution, recurrence, equidistribution, Brownian motions, limit theorems, etc. Most notably we state and prove the Central Limit Theorem on Lie groups. Oddly enough, the proof of this fundamental theorem cannot be found in the literature in a simple and complete form, although it was established almost 45 years ago by Wehn. The exposition here will I hope make up for that.

In Section 2 we discuss the equidistribution properties of random walks in Lie groups and survey the existing ratio limit theorems and local limit theorems in this context. We give several examples including the Itô-Kawada equidistribution theorem for compact groups and the local limit theorem for random walks by isometries on the plane, of which we give a full proof thus generalizing an old theorem of Kazhdan.

In Section 3 we study in more detail the case of nilpotent Lie groups, rephrase the central limit theorem there, discuss the absence of harmonic functions and Guivarc'h's theorem on the L^1-vanishing of convolution powers, and describe how harmonic analysis and Kirillov theory come into play to show equidistribution and local limit theorems on nilpotent Lie groups. We also survey some of the recent work of Alexopoulos on finitely generated nilpotent groups and convolution powers of continuous densities on nilpotent Lie groups.

In Section 4 we go back to the classical theory of sums of independent and identically distributed random variables on R^d and discuss the speed of convergence in the classical local limit theorem. Although such random walks have been very thoroughly studied in the past, in particular in the context of renewal theory, we put forward here a new simple characterization, introduced in [22], of the probability measures for which the speed of convergence to equidistribution is optimal. We call such measures *diophantine* because they are characterized by the property that they cannot be well approximated by measures supported on arithmetic progressions.

Finally in Section 5 we presents some of the results of [21] and [20] concerning equidistribution of unipotent random walks on homogeneous spaces. Making use of precise uniform versions of the existing local limit theorems on commutative or nilpotent Lie groups, as well as Ratner's classification of invariant measures, we show that centered unipotent random walks equidistribute to the homogeneous measure supported on the closed orbit on which they live, while non-centered random walks may wander back and forth and no equidistribution can be expected in this case.

For another point of view and an introduction to random walks on Lie groups in a broader context, we refer the reader to the surveys [51] and [50] and the references therein.

Date: March 2004.

1. Definitions and basic notions

Let G be a separable locally compact group generated by a compact neighborhood of the identity $e \in G$. We consider G as a measurable space for the σ-algebra of Borel subsets. We endow G with a left invariant Haar measure dg. Also we will denote by μ a probability measure on G, that is a non-negative σ-additive measure defined on Borel subsets such that $\mu(G) = 1$. The support of μ is by definition the smallest closed subset of G of μ-measure 1. Let us start with a few definitions:

Definition 1.1. *A probability measure μ on G is said to be* ***non-degenerate*** *if its support generates a dense subgroup of G.*

Definition 1.2. *A probability measure μ on G is said to be* ***aperiodic*** *if its support is not contained in a (left or right) coset of a proper closed subgroup of G.*

Let U be a generating compact neighborhood of the identity. We set:

$$\delta_U(g) = \inf\{n \geq 0, g \in U^n\}$$

The function δ_U is sub-additive and gives a way to measure the distance from the identity e in G.

Definition 1.3. *We say that μ has a* ***moment*** *of order $\alpha > 0$ if*

$$\int \delta_U(g)^\alpha d\mu(g) < +\infty$$

It can be easily checked that this definition does not depend on the choice of U. In the case of connected Lie groups, we can equivalently replace $\delta_U(g)$ by $d(e, g)$ where d is any left invariant metric on G, since one can easily see that for any such d there is a constant $c > 0$ such that $c^{-1}d(e, g) \leq \delta_U(g) \leq cd(e, g)$ when$g$ lies outside U.

The existence of a moment of order 1 implies that the continuous homomorphisms $G \to (\mathbb{R}, +)$ belong to $\mathbb{L}^1(\mu)$.

Definition 1.4. *We say that μ is* ***centered*** *if μ has a moment of order 1 and if*

$$\int \chi(g)d\mu(g) = 0$$

for any continuous homomorphism $\chi : G \to (\mathbb{R}, +)$.

Definition 1.5. *We say that μ is* ***symmetric*** *if $\mu(E) = \mu(E^{-1})$ for any Borel set E in G.*

Let us note that symmetric measures with a moment of order 1 are centered and that if μ symmetric, then μ is aperiodic if and only if μ^2 is non-degenerate.

We say that a random variable X with values in G is distributed according to μ, or equivalently that μ is the **law** of the random variable X if the probability $P(X \in A)$ that X belongs to A equals $\mu(A)$ for every Borel subset A in G.

1.1. Random walks and recurrence.

In analogy with the theory of sums of real random variables, we define a random walk on a group G, to be a stochastic process $S_n = X_1 \cdot ... \cdot X_n$ given by the product of n independent and identically distributed random variables with values in G. The random variables X_i's are independent and distributed according to a single probability measure μ on G. Similarly, we speak about a random walk on a space X, on which G operates, if we apply independently at each step a transformation from G chosen randomly according to a given probability measure on G.

We usually assume the measure μ to be *non-degenerate*. This is no serious assumption, since we could always restrict G to a smaller closed subgroup. The probability law of S_n is simply the n-th convolution power of μ, that is μ^{*n}. The convolution of two measures is defined by the formula:

$$\int fd(\mu * \nu) = \int f(gh)d\mu(h)d\nu(h)$$

where f is an arbitrary continuous and bounded function on G. Let X be a random variable with values in G and whose probability law is μ. If Y is another random variable with law ν this time, then the law of the product XY is precisely the measure $\mu * \nu$.

One of the first simple observation that can be made on the behavior of $(S_n)_n$ is the following dichotomy (see [48] for a proof):

• either the walk S_n diverges to infinity almost surely (meaning that for every compact subset K, $S_n \notin K$ after some large time), we then say that the random walk is *transient.*

• or, almost surely, the walk spends an infinite amount of time in each open subset of G, we then say that the random walk is *recurrent.*

We also observe that the walk is transient if and only if the potential $\sum_{n\geq 0} \mu^{*n}$ is a locally finite measure on G (i.e. a measure that gives finite mass to compact subsets). In the opposite case, the potential of any open set is infinite.

1.2. **Convergence in distribution and unitary representations.** The classical theorem of Lévy, which asserts the equivalence between the convergence in distribution of a sequence of probability measures and the point-wise convergence of the Fourier transforms, or characteristic functions, extends naturally to non-commutative harmonic analysis, with essentially the same statement, as we will see below.

If G is a locally compact group, the set of irreducible unitary representations of G is called the *unitary dual* of G and is denoted by $\widehat{G}$. Recall that a unitary representation $\pi \in \widehat{G}$ is a continuous action of G on a Hilbert space $\mathcal{H}$ by automorphisms preserving the scalar product. It is said to be irreducible if it has no closed invariant subspace.

If μ is a probability measure defined on Borel subsets of G, and π is a unitary representation of G, we can associate naturally to them an operator $\pi(\mu)$ defined for $\xi \in \mathcal{H}$ by

$$\pi(\mu)\xi = \int_G \pi(g)\xi\mu(dg)$$

For example, if $G = \mathbb{R}^d$, the unitary representations are of dimension 1 and are given by the characters $\pi_t(x) = e^{i(t,x)}$, where $t \in \widehat{G} \cong \mathbb{R}^d$ and $x \in G = \mathbb{R}^d$. In this case, $\pi_t(\mu)$ is simply a complex number of modulus less or equal to one: this is the classical Fourier transform, or characteristic function $\widehat{\mu}(t)$ of μ. Moreover, every unitary representation π of G gives rise to a Banach algebra homomorphism between the space of (bounded) complex measures on G and the bounded operators on the Hilbert $\mathcal{H}$ of the representation π. In other words,

$$\pi(\mu * \nu) = \pi(\mu)\pi(\nu)$$

In order to understand the behavior of the random walk S_n, we are led to study the powers of the operators $\pi(\mu)$ for $\pi \in \widehat{G}$.

With these notations, the desired generalization of Lévy's convergence criterion can be stated as follows:

Theorem 1.1. *Let G be a locally compact group and $\widehat{G}$ unitary dual.*

• *If μ and ν are two probability measures on G such that $\pi(\mu) = \pi(\nu)$ for all $\pi \in \widehat{G}$, then $\mu = \nu$ (uniqueness of the Fourier transform).*

• *If $(\mu_n)_n$ is a sequence of probability measures on G which converges in distribution toward another probability measure ν (that is $\int f d\mu_n \to \int f d\nu$ for all continuous and bounded functions f on G), then $\pi(\mu_n)$ converges strongly to $\pi(\nu)$ (that is $\pi(\mu_n)\xi \to \pi(\nu)\xi$ for all $\xi \in \mathcal{H}$).*

• *If $\pi(\mu_n)$ converges weakly to $\pi(\nu)$ (that is $(\pi(\mu_n)\xi, \eta) \to (\pi(\nu)\xi, \eta)$ for all $\xi, \eta \in \mathcal{H}$), then $(\mu_n)_n$ converges in distribution to ν.*

The proof of this statement (see [43]) is a natural generalization of the classical proof of Lévy's theorem by approximation of continuous functions on compact subsets by trigonometric polynomials (i.e. by linear combinations of matrix elements $(\pi(g)\xi, \eta)$).

1.3. **Equidistribution, local limit theorem.**

Definition 1.6. *We will say that the random walk associated to a probability measure μ on G is* ***equidistributed*** *if there exists a locally finite Borel measure m on G such that*

$$\lim_{n\to+\infty} \frac{\int f d\mu^n}{\int g d\mu^n} = \frac{\int f dm}{\int g dm} \tag{1}$$

for all compactly supported continuous functions f and g on G with $g \geq 0$ non identically zero.

If the random walk is equidistributed, we also say, equivalently, that the measure μ satisfies a *ratio limit theorem.* When we can find an *explicit sequence* $(a_n)_n$ of positive real numbers such that

$$\lim_{n\to+\infty} a_n \int f d\mu^n = \int f dm$$

for all compactly supported continuous functions f on G, we say that μ satisfies a *local limit theorem.*

In later sections, we will mainly focus on this property of random walks on groups and study it in several special cases.

1.4. **Diffusion processes and the Lévy-Khinchin-Hunt formula.**

In his 1956 paper [62], Hunt gave a general formula characterizing the continuous semi-groups of probability measures on connected Lie groups. As in the classical case where $G = \mathbb{R}^d$ these semi-groups correspond to stochastic processes with independent and stationary increments, also called Lévy processes. When $G = \mathbb{R}^d$ the well-known Lévy-Khinchin theorem characterizes their probability law by giving an explicit formula for the characteristic function (see [19] and [40]). On a connected Lie group the natural generalization of this idea is to write down explicitly the action on C^2 functions on G of the infinitesimal generator associated to the semi-group of measures.

1.4.1. *Continuous semi-groups of probability measures.* Let G be a connected Lie group. We will work in the Banach space $B := C_0(G)$, the space of continuous real valued functions on G tending to zero at infinity, endowed with the supremum norm $\|f\| = \|f\|_\infty$. By a **continuous semi–group of probability measures** on G, we mean a family $(\mu_t)_{t>0}$ of probability measures on G such that:

(i) $\mu_t * \mu_s = \mu_{t+s}$ for all $s, t > 0$.

(ii) $\mu_t \Rightarrow \delta_e$ when $t \to 0$ (i.e. $\int f d\mu_t \to f(e)$ for all functions $f \in C_0(G)$).

We let $T_t f(g) = \int f(gh) d\mu_t(h)$. Then the $(T_t)_{t>0}$ form a semi-group of operators on the space $B = C_0(G)$. Each T_t is a contraction, i.e. $\|T_t f\| \leq \|f\|$ for any $f \in B$. When $t \to 0$,

$T_t f$ converges uniformly toward f for all $f \in B$. Moreover, we define the infinitesimal generator L of T_t by the formula

$$Lf(g) = \lim_{t \to 0} \frac{1}{t}(T_t f(g) - f(g))$$

L is an unbounded operator in B with domain $Dom(L)$ defined as the set of elements $f \in B$ such that the above limit holds in B, that is uniformly in $g \in G$. The domain $Dom(L)$ is a dense subspace in B and actually the theorem below shows that it contains $C_c^2(G)$ (i.e. the space of C^2 functions with compact support on G). It is easy to see that L determines the semi-group $(T_t)_{t>0}$. Also $Dom(L)$ is invariant under T_t. For background and basic properties on semi-groups of operators on Banach spaces, see Davies [31].

A stochastic process with independent and stationary increments, or **Lévy process**, on G is a stochastic process $(X_t, \mathbb{P}, \Omega)$ such that :

(i) for all s, t $0 < s < t < +\infty$ the law of $X_s^{-1} X_t$ depends only on $t - s$.

(ii) for all $t_1, ..., t_k$ such that $0 < t_1 < ... < t_k$ the random variables $X_{t_i}^{-1} X_{t_{i+1}}$ are independent.

(iii) when $t \to 0$, X_t converges in law to e.

The Lévy processes are in bijective correspondence with the continuous semi-groups of probability measures on G. More precisely, let $(X_t, \mathbb{P}, \Omega)$ be a Lévy process, and let μ_t be the probability law of X_t, then the $(\mu_t)_{t>0}$ form a continuous semi-group of probability measures on G. And vice versa, if $(\mu_t)_t$ is any such semi-group, then one can construct a probability space $(\mathbb{P}, \Omega)$ and a Lévy process (X_t) defined on $(\mathbb{P}, \Omega)$ such that μ_t is the probability law of X_t. Moreover, one can always find a càdlàg version of this process (i.e. such that sample paths are almost surely right continuous with limits on the left at every $t > 0$).

We fix a basis $X_1, ..., X_d$ of the vector space of the Lie algebra $\mathfrak{g}$ of G and a relatively compact neighborhood U_0 of the identity in G on which the logarithm is a well defined diffeomorphism. This allows to parametrize elements $g \in U_0$ by their coordinates $(x_i(g))_{i=1,...,d}$ determined by the equation

$$g = \exp(\sum x_i(g) X_i)$$

We extend each coordinate function $x_i(g)$ smoothly to the whole of G with the requirement that $x_i(g) = 1$ if g lies outside some compact set containing U_0. We define the function ϕ on G to be $\phi(g) = \sum x_i(g)^2$. The elements of the Lie algebra $\mathfrak{g}$ of G can be seen as elements of the universal enveloping algebra of $\mathfrak{g}$, which is identified with the algebra of left-invariant differential operators on G. In particular, for a differentiable function f on G and for $X \in \mathfrak{g}$, we set

$$Xf(g) = \frac{d}{dt}_{|t=0} f(ge^{Xt})$$

Then we have:

Theorem 1.2. *(Hunt [62]) Let $(\mu_t)_{t>0}$ be a continuous semi-group of probability measures on G and (T_t) be the associated semi-group of operators on $B = C_0(G)$. Then the domain of the infinitesimal generator L of (T_t) contains C_c^2 (compactly supported C^2 functions on G) where it admits the following form:*

(2)

$$Lf(g) = \sum_i b_i X_i f(g) + \frac{1}{2} \sum_{i,j} a_{ij} X_i X_j f(g) + \int_G \left\{ f(gh) - f(g) - \sum_i X_i f(g) x_i(h) \right\} \frac{d\eta(h)}{\phi(h)}$$

where (b_i) and (a_{ij}) are real numbers such that the matrix (a_{ij}) is positive semi-definite and $d\eta$ is a finite positive measure on G such that $\eta(\{e\}) = 0$. Moreover, the second order differential operator $\sum_{i,j} a_{ij} X_i X_j$ and the measure $\frac{d\eta(g)}{\phi(g)}$ are independent of the choices of the basis (X_i) and the function $\phi(g)$. Finally, the semi-group $(\mu_t)_t$ is determined in a unique way by the operator L restricted to $C_c^2(G)$.

Conversely, every operator L defined on $C_c^2(G)$ by the formula (2) is the infinitesimal generator of a continuous semi-group of probability measures on G.

Let us remark that the operator $Id - L$ maps $Dom(L)$ *onto* $B = C_0(G)$ and that the inverse map, defined on B, is a contraction. To see these two facts, set for $f \in B$

$$Rf = \int_0^\infty e^{-t} T_t f dt$$

Since the T_t's are contractions, the integral gives a well defined element of B, and R is a linear operator on B which is also a contraction, i.e. $\|Rf\| \leq \|f\|$ for any $f \in B$. Now Rf belongs to $Dom(L)$ since one checks directly that

$$\lim_{s \to 0} \frac{1}{s}(T_s Rf - Rf) = Rf - f$$

where the limit holds in B. It follows that $(Id - L)R = Id$ on B. Similarly, since $\frac{d}{dt} T_t f = T_t L f$ we obtain $R(Id - L) = Id$ on $Dom(L)$. In particular $Im(Id - L) = B$ and $\mathrm{Im} R = Dom(L)$ (which is dense in B). Hence R is the inverse map of $Id - L$. The operator $R = (Id - L)^{-1}$ is called the **resolvent** of L.

We end by the following important proposition which will turn useful in the next section (see Hirsch [60] Théorème 13, and also [55] Satz 4.1)

Proposition 1.1. *The space $C_c^\infty(G)$ of compactly supported smooth functions on G is a **core** for the infinitesimal generator L of the semi-group $(T_t)_{t>0}$, that is for any $f \in Dom(L)$ one can find a sequence $(f_n)_n$ in $C_c^\infty(G)$ such that $f_n \to f$ and $Lf_n \to Lf$ where the convergence holds in B (i.e. uniformly on G). Equivalently $(Id - L)C_c^\infty(G)$ is dense in B.*

Proof. To see that the two assertions are equivalent is easy. The first assertion follows from the second since if $g_n \in C_c^\infty(G)$ is such that $g_n - Lg_n$ converges to $f - Lf$ in B then, applying the contraction $R = (Id - L)^{-1}$, we deduce that g_n converges to f in B. Hence also Lg_n tends to Lf, and $C_c^\infty(G)$ is a core for L. The converse follows from the fact that $(Id - L)Dom(L) = Im(Id - L) = B$ as noted above.

To show that $(Id - L)C_c^\infty(G)$ is dense we can equivalently show that no non-zero bounded measure ν on G is orthogonal to $(Id - L)C_c^\infty(G)$. Indeed the dual Banach space to B is the space of bounded Borel measures on G. First note that if $f \in Dom(L)$ then $\nu * f \in Dom(L)$ because $T_t(\nu * f) = \nu * T_t f$ for all $f \in B$ and $t > 0$, and thus $\nu * Lf = L(\nu * f)$.

Let ν be a bounded Borel measure on G which is orthogonal to $(Id - L)C_c^\infty(G)$, i.e. such that $\int (f - Lf)(g) d\nu(g) = 0$ for all $f \in C_c^\infty(G)$. Since $\nu * f$ belongs to $B = C_0(G)$, up to changing f into $-f$ if necessary, we can find an element $g_0 \in G$ such that $\nu * f(g_0) = \|\nu * f\|$. Now since $f * \delta_{g_0} \in C_c^\infty(G)$ we have

$$\int (f * \delta_{g_0} - L(f * \delta_{g_0}))(g) d\nu(g) = 0$$

which can also be written as

$$\begin{aligned}\nu^{-1} * f * \delta_{g_0}(e) &= \nu^{-1} * L(f * \delta_{g_0})(e)\\ &= L(\nu^{-1} * f * \delta_{g_0})(e)\end{aligned}$$

Let $h = \nu^{-1} * f * \delta_{g_0}$. On the one hand we have $h(e) = L(h)(e)$ and on the other hand we have $h(e) = \|h\|$. But from the very definition of L

$$Lh(e) = \lim_{t\to 0} \frac{1}{t} \int (h(x) - h(e)) d\mu_t(x)$$

it follows that $Lh(e) \le 0$. Hence $\|h\| = h(e) \le 0$ and $h = 0$. In particular $h(g_0^{-1}) = \int f d\nu = 0$. Since f was arbitrary, we conclude that $\nu = 0$. □

□

1.4.2. *Gaussian semi-groups and Brownian motions.* When the measure $\frac{\eta}{\phi}$ (called the *Lévy measure*) is identically zero, then we say that the semi-group $(\mu_t)_t$ is **gaussian** (equivalently, $(\mu_t)_t$ is gaussian if and only if for any neighborhood U of the identity $\mathbb{P}(X_t \notin U) = o(t)$), and that the associated stochastic process $(X_t)_t$ is a (left-invariant) **Brownian motion** on G. In this case and this case only, the stochastic process (X_t) has continuous sample paths almost surely. Moreover, the gaussian semi-group (μ_t) is symmetric (i.e. $X_t \overset{d}{=} X_t^{-1}$) if and only if the b_i's are all zero.

The second order differential operator $\sum_i b_i X_i + \frac{1}{2}\sum_{i=1}^d a_{ij} X_i X_j$ can be put in the form $E_0 + \sum_{i=1}^p E_i^2$, $1 \le p \le d$, for certain vectors E_i in $\mathfrak{g}$. Let $\mathfrak{h}$ be the Lie algebra generated by the vectors $E_1, ..., E_p$ together with all brackets of all orders between different E_i's, $0 \le i \le p$, in which E_0 appears at least once. We say that the operator L_μ is a sub-laplacian, if the differential operator $\frac{\partial}{\partial t} - L_\mu$ on $\mathbb{R}_+^* \times G$ is hypoelliptic, which amounts to say that $\mathfrak{h} = \mathfrak{g}$ thanks to Hörmander's theorem. In this case and this case only, the gaussian semi-group $(\mu_t)_{t>0}$ has a density $p_t(x)dx$ with respect to the Haar measure on G. This density is the *heat kernel* associated to L_μ (see for instance [116] in the case when μ_t is symmetric, and [101], [102] for the general case). The functions $(p_t(x))_{t>0}$ are C^∞ and constitute a Dirac family. They also satisfy the heat equation $(\frac{\partial}{\partial t} - L_\mu)p_t = 0$. It can be shown that, on a unimodular group, $p_t(x)$ decays faster, at fixed t, than $e^{-cd(e,x)^2}$ for a certain constant $c = c(t) > 0$ (see [113] for a detailed study of the heat kernel p_t in terms of the geometry of the group G). If the matrix (a_{ij}) is positive definite (i.e. $p = d$), then the densities $p_t(x)$ are analytic functions on G (see [58] Theorem 6.3.1).

Obviously, the support of μ_t is contained in the analytic subgroup G_0 of G whose Lie algebra $\mathfrak{g}_0$ is generated by the vectors $E_0, ..., E_p$. The subalgebra $\mathfrak{h}$ is a ideal in g_0 and $\mathfrak{h} = \mathfrak{g}_0$ if and only if the μ_t's are absolutely continuous with respect to the Haar measure of G_0. If $\mathfrak{h}$ is a proper ideal in $\mathfrak{g}_0$ then the support of μ_t is contained in the coset He^{tE_0} where H is the subgroup of G corresponding to the Lie subalgebra $\mathfrak{h}$. On the other hand, the support of μ_t contains the coset Me^{tE_0} where M is the subgroup of G corresponding to the Lie algebra $\mathfrak{m}$ generated by the vectors $E_1, ..., E_p$ ([101] Theorem 4). In particular, if $E_1, ..., E_p$ generate all of $\mathfrak{g}$ then $supp(\mu_t) = G$. When G is a simply connected nilpotent Lie group, then $\mathfrak{m}$ is proper is and only if $\mathfrak{h}$ is proper. Hence in this case, $supp(\mu_t) = G$ if and only if μ_t has a density with respect to the Haar measure. However, for an arbitrary Lie group G, it may happen that the support of μ_t is proper, although μ_t is absolutely continuous (see [101] example 3.4b, in particular $p_t(x)$ is not analytic in general).

1.5. **Law of large numbers, Central limit theorem.** There are different ways to generalize the law of large numbers and the central limit theorem to random walks on

Lie groups. One approach consists in studying the random variables $d(e,S_n)$ where d is a left-invariant Riemannian metric on G in order to obtain a law of large numbers for these variables, in other words, a convergence of the type

$$\frac{d(e,S_n)}{n} \to \gamma$$

where γ is a non-negative real number. In [49], Guivarc'h shows that such a convergence holds almost surely for every locally compact group G and that one can be more precise about the limit γ in many cases. For instance, when G is not amenable then $\gamma > 0$. More precisely, one can ask whether a central limit theorem holds, that is whether the following convergence holds

$$\frac{d(e,S_n)-\gamma n}{\sqrt{n}} \to X$$

where X is a non-degenerate centered gaussian random variable. We will not consider this question here (cf. [49] and [52] and the book [18]).

Another approach is to consider the products $P_n = X_{1,n}\cdot\ldots\cdot X_{n,n}$ of independent random variables such that the typical size of $X_{k,n}$ is very small of the order of $1/\sqrt{n}$. When G is $\mathbb{R}^d$ or a homogeneous nilpotent Lie group (see below 3.1.2), then G has a one-parameter semi-group of dilations (δ_t) which are automorphisms of G with the property that $\delta_t(K) \to e$ when $t \to 0$ for every compact set K in G. Then one can set $X_{k,n} = \delta_{\frac{1}{\sqrt{n}}}(X_k)$ and study $P_n = \delta_{\frac{1}{\sqrt{n}}}(S_n)$. When $G = \mathbb{R}^d$, the classical central limit theorem says that P_n converges in distribution to a gaussian variable, under the assumption that the law of X_k has a finite moment of order 2. In the Lie group case, this infinitesimal approach yields the fundamental limit theorem stated below.

We keep the notations from the last subsection. In particular G is a connected Lie group and the coordinate functions $x_i(g)$ are as defined above.

We are given a sequence $(\mu_n)_{n\geq 1}$ of probability measures on G together with independent random variables $(X_{k,n})_{1\leq k\leq n}$ with law μ_n. The random variable $S_n = X_{1,n}\cdot\ldots\cdot X_{n,n}$ is distributed according to μ_n^{*n}. Then the following theorem holds:

Theorem 1.3. ***(CLT on Lie groups**, Wehn [114] [115] [43]) We make the following assumption. There are real numbers b_i's and a_{ij}'s such that for every $i,j = 1,...,d$ we have*

(i) $\int x_i(g)d\mu_n(g) = \frac{b_i}{n} + o(\frac{1}{n})$

(ii) $\int x_i(g)x_j(g)d\mu_n(g) = \frac{a_{ij}}{n} + o(\frac{1}{n})$

(iii) $\mu_n(U^c) = o(\frac{1}{n})$ *for every neighborhood U of the identity in G.*

Then the sequence of measures $(\mu_n^{*n})_n$ converges to the measure ν_1, which belongs to the gaussian semi-group of probability measures $(\nu_t)_{t>0}$ on G, whose infinitesimal generator is given by

$$L_\nu = \sum_{i=1}^d b_i X_i + \frac{1}{2}\sum_{i=1}^d a_{ij}X_iX_j$$

In the above formula, the elements of the Lie algebra X_i's are considered as differential operators acting on $C_c^2(G)$. The matrix (a_{ij}) is by definition positive semi-definite and the semi-group $(\nu_t)_{t>0}$ associated to L_ν by the Lévy-Khinchin-Hunt formula is gaussian.

When $G = \mathbb{R}^d$, Wehn's theorem implies the classical central limit theorem: the existence of a finite second moment is equivalent to condition (ii) above and implies condition (iii). Similarly Wehn's theorem implies the now well-known central limit theorem on graded nilpotent groups (see below theorem 3.11), and it can be checked that conditions

(ii) and (iii) are automatically satisfied when there is a finite second moment. However the conditions of Theorem 1.3 are weaker in this case. We give below an example of a measure on the Heisenberg group with no finite second moment, but for which the conditions of the theorem are still fulfilled (see Example 3.1).

Wehn's original statement involved commutative triangular arrays of probability measures and some additional assumptions (see Wehn [114] [115] and Grenander [43]). His original proof is only sketched in Grenander's book [43] while the full argument remains virtually impossible to find in the available literature[1]. His proof was based on the theory of semi-groups and certain convergence theorems for sequences of semi-groups established by Hunt. For the sake of completeness, I give below a complete proof of Wehn's theorem using essentially his original method.

Let me mention that Stroock and Varadhan [110] have generalized this result for non-commutative triangular systems by giving a similar necessary and sufficient condition for convergence in law to $(\nu_t)_{t>0}$, for the process obtained by gluing piecewise the $X_{k,n}$'s into a single $X_t(n)$ defined by $X_{1,n} \cdot ... \cdot X_{[nt],n}$ on the interval $[\frac{[nt]}{n}, \frac{[nt]+1}{n}[$. Their proof is very different and purely probabilistic but, as Wehn's proof, does not require any result from the structure theory of Lie groups.

Proof of Theorem 1.3. The proof uses ideas and techniques from the beautiful theory of continuous semi-groups of operators in Banach spaces as exposed in Davies [31] for example. In particular, Lemma 1.2 below is part of the well-known Trotter-Kato theorems on perturbations of semi-groups. However, the exposition we give here is entirely self-contained and does not require anything but very basic functional analysis.

Let L_n be the Laplacian associated to μ_n rescaled by a factor n, that is the bounded operator defined for $f \in B = C_0(G)$ by

$$L_n f = n(f * \mu_n - f)$$

We define the one-parameter group $(e^{tL_n})_{t\in\mathbb{R}}$ by the formula

$$e^{tL_n} = \sum_{k\geq 0} \frac{t^k}{k!} L_n^k$$

Since L_n is a bounded operator (in fact $\|L_n\| \leq 2n$), this sum is well defined. Moreover $e^{tL_n} = e^{-nt}e^{nt\mu_n}$ and hence $\left\|e^{tL_n}\right\| \leq 1$. We denote by R_n the resolvent of L_n, defined as was done above for $(T_t)_t$ and equal to $(Id - L_n)^{-1}$. Similarly we have $\|R_n\| \leq 1$. Let $(T_t)_t$ be the gaussian semi-group of probability measures defined in the statement of the theorem and let L be its infinitesimal generator. The main lemma is the following:

Lemma 1.1. *Let $f \in C_c^\infty(G)$ then $L_n f \to Lf$ in B as n tends to infinity.*

From this lemma and the result of Proposition 1.1, the conclusion of the theorem follows by a direct application of the Trotter-Kato theory of perturbations of operator semi-groups. Yet, for the sake of completeness, we indicate the full argument below.

Lemma 1.2. *We have successively:*
(i) $R_n f \to Rf$ for any $f \in B$.
(ii) $e^{tL_n} f \to T_t f$ for any $f \in B$ and $t > 0$.
(iii) $\left\|\mu_n^{*n} f - e^{L_n} f\right\| \to 0$ for any $f \in B$.

[1]His Yale dissertation, made under the supervision of S. Kakutani in 1959, is only accessible through the archives department of the Sterling Memorial Library at Yale university and it bears a special notice prohibiting its photocopying.

It is clear that the conclusion of the theorem follows from the combination of (ii) and (iii) in this lemma.

Proof of Lemma 1.1. The lemma follows by writing down the Taylor expansion of f and using the three conditions assumed in the theorem. Let $f \in C_c^\infty(G)$ and U be a neighborhood of the identity in G where the exponential map is a diffeomorphism. Write $y = e^Y$, $Y \in Lie(G)$ and $Y = \sum y_i X_i$ for $y \in U$. Then by the classical Taylor-Lagrange theorem applied to $t \mapsto f(xe^{tY})$

$$f(xy) - f(x) = Yf(x) + \frac{1}{2}Y^2 f(xe^{\theta Y})$$

for some number $\theta \in [0,1]$ depending on x and y. We can write

$$L_n f(x) = n \int_U (f(xy) - f(x))\, d\mu_n(y) + \varepsilon_n(x)$$

where $\varepsilon_n \in B$ and $\|\varepsilon_n\| \leq 2n\,\|f\|\,\mu_n(U^c) \to 0$ from assumption (iii). Moreover

$$n \int_U Yf(x) d\mu_n(y) = \sum_{1\leq i\leq d} n \left(\int_U y_i d\mu_n(y)\right) X_i f(x)$$

converges uniformly in $x \in G$ to

$$\sum_{1\leq i\leq d} b_i X_i f(x)$$

as follows from the combination of assumptions (i) and (iii). We proceed similarly for the quadratic term. More precisely we first have from (ii) and (iii)

$$n \int_U Y^2 f(x) d\mu_n(y) \to \sum_{1\leq i,j\leq d} a_{ij} X_i X_j f(x)$$

uniformly in $x \in G$. It remains to show that

$$n \int_U \left(Y^2 f(xe^{\theta Y}) - Y^2 f(x)\right) d\mu_n(y)$$

can be made arbitrarily small uniformly in $x \in G$ when n is large enough and U is small enough. Since $f \in C_c^\infty(G)$, $Y^3 f$ also belongs to $C_c^\infty(G)$ and by Taylor's inequality, for all $x \in G$ and $y \in U$ we have

$$\left|Y^2 f(xe^{\theta Y}) - Y^2 f(x)\right| \leq \left\|Y^3 f\right\|$$

Moreover

$$\left\|Y^3 f\right\| \leq \sum_{i,j,k} |y_i y_j y_k|\, \|X_i X_j X_k f\|$$

Hence it is enough to show that each integral

$$n \int_U |y_i y_j y_k|\, d\mu_n(y)$$

can be made arbitrarily small for n large enough and U small enough. Let $\varepsilon > 0$ and take U so small that $|y_i| < \varepsilon$ for all $i = 1, ..., d$ and $y \in U$. Then

$$\begin{aligned} n \int_U |y_i y_j y_k|\, d\mu_n(y) &\leq \varepsilon n \int_U |y_i y_j|\, d\mu_n(y) \\ &\leq \varepsilon \left(n \int_U y_i^2 d\mu_n(y)\right)^{1/2} \left(n \int_U y_j^2 d\mu_n(y)\right)^{1/2} \end{aligned}$$

Since both factors in the right hand side converge to $a_{ii}^{1/2}$ and $a_{jj}^{1/2}$ respectively by assumption (ii), we are done.□

Proof of Lemma 1.2. Let us begin by (i). Since R_n and R are contractions, it is enough to show the convergence for all f belonging to a dense subspace of B, for example the subspace $(Id - L)C_c^\infty(G)$, which is dense by Proposition 1.1. Let $f \in C_c^\infty(G)$. From Lemma 1.1 we have that $(Id - L_n)f$ converges to $(Id - L)f$ in B. Applying R_n on both sides, we get that $R_n(Id - L)f$ converges to f (because R_n is a contraction). In other words $R_n g$ converges to Rg for any g of the form $g = (Id - L)f$, $f \in C_c^\infty(G)$. We are done.

For (ii) note first that for $t \geq 0$ and for every $f \in Dom(L)$ we have

$$e^{tL_n}f - T_t f = e^{tL_n}(f - e^{-tL_n}T_t f) = e^{tL_n}\int_0^t e^{-sL_n}(L_n - L)T_s f ds$$

Changing f into Rf and applying R_n on both sides, we now get for an arbitrary $f \in B$

$$R_n(e^{tL_n} - T_t)Rf = e^{tL_n}\int_0^t e^{-sL_n}(R_n - R)T_s f ds$$

The quantity under the integral sign is bounded in norm by $2\|f\|$ while it converges for every $s \in (0,t)$ to 0 owing to part (i) of the lemma and the fact that the e^{-sL_n} are contractions. By dominated convergence we obtain that

$$R_n(e^{tL_n} - T_t)Rf \to 0$$

for every $f \in B$. However, since $ImR = Dom(L)$ is dense in B and R_n and the semigroups e^{tL_n} and T_t are contractions, we obtain

$$R_n(e^{tL_n} - T_t)f \to 0$$

for every $f \in B$. But we can write

$$R_n(e^{tL_n} - T_t)f = e^{tL_n}(R_n - R)f - (R_n - R)T_t f + e^{tL_n}Rf - T_t Rf$$

Since both $e^{tL_n}(R_n - R)f$ and $(R_n - R)T_t f$ tends to 0 for any $f \in B$ by (i), we obtain

$$e^{tL_n}Rf - T_t Rf \to 0$$

Again, since ImR is dense, we conclude that $e^{tL_n}f \to T_t f$ for all $f \in B$.

Finally we prove (iii). Note that μ_n and e^{tL_n} commute for all $t > 0$. In particular

$$\mu_n^{*n} - e^{L_n} = (\mu_n - e^{L_n/n})\sum_{k=0}^{n-1}\mu_n^{n-1-k}e^{kL_n/n}$$

and taking the norms, for any $f \in B$

$$\left\|\mu_n^{*n}f - e^{L_n}f\right\| \leq n\left\|\mu_n f - e^{L_n/n}f\right\|$$

On the other hand $\mu_n = 1 + \frac{L_n}{n}$ and therefore

$$\mu_n f - e^{L_n/n}f = \sum_{k\geq 0}\frac{1}{(k+2)!}(\frac{L_n}{n})^k \cdot (\frac{L_n}{n})^2 f$$

Since $\|L_n\| \leq 2n$ we obtain

$$\left\|\mu_n f - e^{L_n/n}f\right\| \leq e^2\frac{\left\|L_n^2 f\right\|}{n}$$

However $\left\|L_n^2 f\right\|$ might not be bounded for any $f \in B$ or even $C_c^\infty(G)$. To go around this problem we change f into $R_n^2 f$. Then $\left\|L_n^2 R_n^2\right\| \leq \|L_n R_n\|^2 \leq 4$ because $L_n R_n = R_n - Id$ and R_n is a contraction. Hence

$$\left\|\mu_n R_n^2 f - e^{L_n/n} R_n^2 f\right\| \to 0$$

Moreover it follows from (i), and the fact that R_n is a contraction, that $R_n^2 f \to R^2 f$ for every $f \in B$. Since μ_n and $e^{L_n/n}$ are contractions, we obtain that

$$\left\|\mu_n R^2 f - e^{L_n/n} R^2 f\right\| \to 0$$

for every $f \in B$. But again the image of R is dense and R is a bounded operator, hence the image of R^2 is also dense and this implies that the convergence $\left\|\mu_n f - e^{L_n/n} f\right\| \to 0$ holds for every $f \in B$. $\square$

2. Local limit theorems on compact or abelian groups and their extensions

2.1. Local limit theorems on abelian groups.

On $\mathbb{R}^d$ one has the following theorem, classically known as *the local limit theorem,* and whose proof relies on real analysis and the use of the Fourier transform.

Theorem 2.1. *(see* [19] *Theorem 10.17) Let μ be a centered and aperiodic probability measure on $\mathbb{R}^d$ with a finite moment of order 2 and let K be its covariance matrix. Then for any continuous function f of compact support on $\mathbb{R}^d$ we have*

$$\lim_{n\to+\infty} n^{d/2} \int f d\mu^n = \frac{1}{\sqrt{(2\pi)^d \det K}} \int_{\mathbb{R}^d} f(x)dx$$

Let us remark that the aperiodicity condition is necessary to obtain the convergence above. It is therefore not surprising that we find it in all other equidistribution statements. In a way, the case of $\mathbb{R}^d$ is ideal, and most studied, and the general goal is to try to obtain similar theorems for other Lie groups. In these notes, we will consider almost exclusively amenable groups, or groups with polynomial growth. For non-amenable groups, the behavior of μ^n is very different, but the question of equidistribution can be addressed in a similar fashion (see §6).

On more general abelian groups, we always have the following ratio limit theorem:

Theorem 2.2. *(Stone* [107]*) Let G be an abelian locally compact second countable group generated by a compact neighborhood of the identity. Let μ be an aperiodic and centered probability measure on G. Then we have*

$$\lim_{n\to+\infty} \frac{\mu^n(A)}{\mu^n(B)} = \frac{|A|}{|B|}$$

for all relatively compact Borel subsets A and B in G with positive Haar measure and with negligible boundary.

We stress that in this theorem, we assume the existence of a finite moment of order 1 only.

2.2. Compact groups.

If the group is compact, there always is equidistribution. This is an old theorem of Itô and Kawada (cf. [65]).

Theorem 2.3. *(Itô-Kawada) Let μ be an aperiodic probability measure on G. Then the sequence μ^n converges in distribution to the normalized Haar measure on G.*

The proof is based on Lévy's criterion for convergence of measures (i.e. Theorem 1.1). Indeed, irreducible unitary representations of G are finite dimensional and the operators $\pi(\mu)$ are just matrices. It is enough to check that if π is not the trivial representation, then the powers $\pi(\mu)^n$ converge to 0. To see this is suffices to prove that the operator norm of $\pi(\mu)$ is strictly less than 1, or equivalently that all eigenvalues of $\pi(\mu)$ are of modulus < 1. But the existence of a modulus 1 eigenvalue contradicts easily the aperiodicity assumption. This method is essentially the method used already by Poincaré in his 1912 treatise on Probability Theory [86], in order to show this theorem in the particular case when G is the finite group of the permutations of n elements.

This theorem implies the equidistribution of every random walk on a compact group. For example, the equidistribution problem of Arnol'd and Krylov (see [5]). Every random walk on S^2, where each move is a random isometry, is equidistributed, as long as the group generated by the support of the probability is dense in $SO(3)$.

The problem of determining the speed of convergence in the above theorem is a delicate issue. Is f is a C^∞ function on G, then one can try to estimate the speed of convergence of $\int f d\mu^n$ toward $\int_G f(g)dg$. By decomposing f into a sum of harmonics corresponding to the decomposition of the regular representation of G as a direct sum of irreducible representations $(\pi_n)_n$, one can reduce partly this question to the study of the operator norms $\|\pi_n(\mu)\|$. We will not deal with this difficult question here. Let us just make a few remarks. If μ is not singular with respect to the Haar measure, then there is a "spectral gap", that is a number $\alpha < 1$ such that $\|\pi_n(\mu)\| \leq \alpha$ for all $n \neq 0$. However, if μ is singular, and in particular if μ is atomic, then the problem is widely open when G is not commutative. When G is a semisimple Lie group, Dolgopyat [32] has recently obtained a polynomial upper bound of the form $\|\pi_n(\mu)\| \leq 1 - \frac{c}{n^\beta}$ for some constants $c, \beta > 0$. This yields a polynomial speed of convergence for C^∞ functions in the local limit theorem. However it has been conjectured, for instance for the group $G = SO(3)$, that there always is a spectral gap (see the book of Sarnak [96]). For this group, only very special examples of measures with a spectral gap have been found (see [96] [39] et [73]), all of which come from some arithmetic considerations related to the solution to the well-known Ruziewicz problem (see [73]).

2.3. **Equidistribution in the plane.** In this paragraph, we will consider the case of a random walk evolving in the Euclidean space $\mathbb{R}^d$, seen as a homogeneous space for the group of motions of $\mathbb{R}^d$. So we let $G = SO(d) \cdot \mathbb{R}^d$, $X = \mathbb{R}^d$ and $S_n \cdot x$ be a random walk starting at x. From x, we apply at each step a random motion according to a given probability law μ on G. An element of G can be written as $g = (\tau, \rho)$ where $\tau \in \mathbb{R}^d$ is a translation and $\rho \in SO(d)$ is a rotation. The first important result concerning this random walk is the following central limit theorem:

Theorem 2.4. ***(CLT for** $SO(d) \cdot \mathbb{R}^d$**)*** *Let μ be a non-degenerate probability measure on $G = SO(d) \cdot \mathbb{R}^d$ such that the image of μ into $\mathbb{R}^d$ under the projection $G \to \mathbb{R}^d$, $g \mapsto \tau$ has a finite moment of order 2. Let $S_n = (\tau_n, \rho_n)$ be the product of n independent random variables distributed according to the same law μ. Then the random variable $\frac{1}{\sqrt{n}}\tau_n$ converges in distribution towards a gaussian law $\mathcal{N}(0, \sigma^2)$, which is centered and of diagonal covariance $\sigma^2 Id$.*

It follows that $\frac{1}{\sqrt{n}} S_n \cdot x$ converges also in distribution to the same limit. This theorem is a direct consequence of the well-known Ibragimov-Billingsley for martingale differences (cf. [63] and [13] Theorem 35.12). Such stochastic processes always satisfy a central limit theorem under the Lindeberg condition. The sequence S_n can be written in the form $X_1 \cdot ... \cdot X_n$ (in this order) and, when projected to $\mathbb{R}^d$, is a martingale as soon as $\mathbb{E}(T_i) = 0$. Indeed, it can be written as $T_1 + \rho_1(T_2) + ... + \rho_{n-1}(T_n)$ where T_i is the projection of X_i on $\mathbb{R}^d$. See also [48] for another proof by Roynette and also Gorostiza's paper [42].

Let us now focus on the question of equidistribution. In 1965, Kazhdan (cf. [66]) proved the following theorem (the proof was later corrected and completed by Guivarc'h in [47]).

Theorem 2.5. *([66] et [47]). Let $G = SO(2) \cdot \mathbb{R}^2$ be the group of Euclidean motions of the plane. Let μ be a non-degenerate measure on G which we assume symmetric and of finite support. Then for every $x \in \mathbb{R}^2$, and for all positive continuous functions with compact support ϕ and ψ on $\mathbb{R}^2$, we have the following ratio limit theorem:*

$$\lim_{n \to +\infty} \frac{\int_G \phi(g \cdot x) d\mu^{2n}(g)}{\int_G \psi(g \cdot x) d\mu^{2n}(g)} = \frac{\int_{\mathbb{R}^2} \phi(y) dy}{\int_{\mathbb{R}^2} \psi(y) dy}$$

In [72], Le Page gave another proof (which also allows to treat the case of any symmetric compactly supported measure) based on the absence of positive μ-harmonic functions on $\mathbb{R}^2$.

Below, we generalize this result by proving the corresponding local limit theorem in complete generality. If μ is aperiodic on G, then the map $G \to \mathbb{R}^2, g \mapsto \int_G g \cdot x d\mu$ has a unique fixed point $x_0 \in \mathbb{R}^2$. We set $\sigma^2 = \int_G |g \cdot x_0|^2 d\mu(g)$. Then we have

Theorem 2.6. ***(LLT in the plane)*** *Let μ be an aperiodic probability measure on $G = SO(2) \cdot \mathbb{R}^2$ with a finite moment of order 2. Then for every $x \in \mathbb{R}^2$, and for every continuous function f of compact support on $\mathbb{R}^2$, we have*

$$\lim_{n \to +\infty} n \int_G f(g \cdot x) d\mu^n(g) = \frac{1}{2\pi\sigma^2} \int_{\mathbb{R}^2} f(y) dy \tag{3}$$

Proof. The idea is to generalize the classical proof of the local limit theorem on $\mathbb{R}$ by using the Fourier transform. We first make the following classical reduction (see [19]):

Claim 2.1. It is enough to show the convergence (3) for functions f on $\mathbb{R}^2$ whose Fourier transform is continuous and of compact support (these functions are integrable but no longer compactly supported).

Proof of claim: Let h be a continuous function which we take integrable and strictly positive on $\mathbb{R}^2$ such that its Fourier transform $\widehat{h}$ is compactly supported (such a function does exist, see [19] 10.2). From Lévy's criterion for convergence of measures, the sequence of finite measures $\nu_n = nh(y)\mu^n * \delta_x(dy)$ converges to the measure $\frac{1}{2\pi\sigma^2} h(y)dy$, since characteristic functions $\widehat{\nu_n}(t)$ do converge point-wise toward the characteristic function of the limit measure. Indeed, $\widehat{\nu_n}(t)$ is nothing else but $n \int_G h_t(g \cdot x) d\mu^n(g)$ where $h_t(y) = e^{it \cdot y} h(y)$ and $\widehat{h_t}$ is the translate of $\widehat{h}$ by t (hence compactly supported). Now since f is continuous with compact support on $\mathbb{R}^2$ then f/h is also continuous with compact support and the expression $\int f/h d\nu_n$, which also equals $n \int f d(\mu^n * \delta_x)$, converges to the desired limit.

□

Let us come back to the proof of the theorem. Clearly, it is enough to prove the equidistribution when $x=0$. Then we have

$$\begin{aligned}\int_G f(\tau)d\mu^n(g) &= \int_G\int_{\mathbb{R}^2}\widehat{f}(x)e^{i\tau\cdot x}dxd\mu^n(g)\\ &= \int_0^{+\infty}\int_G\int_0^{2\pi}\widehat{f}(r\theta)e^{ir\tau\cdot\theta}d\theta d\mu^n(g)rdr\\ &= \int_0^{+\infty}\left\langle\pi_r(\mu)^n\widehat{f}_r,1\right\rangle_{\mathbb{L}^2(S^1)}rdr\end{aligned}\tag{4}$$

where $\widehat{f}_r$ is the function defined on the circle S^1 by $\widehat{f}(r\theta)$ and π_r is the irreducible unitary representation of G defined on the space $\mathbb{L}^2(S^1)$ by

$$\pi_r(g)\phi(\theta)=e^{ir\tau\cdot\theta}\phi(\rho^{-1}\theta)\tag{5}$$

where $r>0$ and $g=(\tau,\rho)\in G$. Then we do the following substitution $r\to r\sqrt{n}$ in (4) and we observe that by the central limit theorem 2.4, for any $r>0$ and $\theta\in S^1$, we have

$$\begin{aligned}\lim_{n\to+\infty}\pi_{r/\sqrt{n}}(\mu)^n\widehat{f}_{r/\sqrt{n}}(\theta) &= \lim_{n\to+\infty}\mathbb{E}\left(e^{ir\frac{S_n}{\sqrt{n}}\cdot\theta}\widehat{f}(\frac{r}{\sqrt{n}}R_n^{-1}\theta)\right)\\ &= \mathbb{E}(e^{irN\cdot\theta})\widehat{f}(0)\\ &= e^{-r^2\sigma^2/2}\widehat{f}(0)\end{aligned}$$

where N is the gaussian law $\mathcal{N}(0,\sigma^2)$ and R_n is the rotation part of S_n. Consequently, $\left\langle\pi_{r/\sqrt{n}}(\mu)^n\widehat{f}_{r/\sqrt{n}},1\right\rangle$ converges for each $r>0$ to $2\pi\widehat{f}(0)e^{-r^2\sigma^2/2}$. If we can interchange the integral sign and the limit, then we will have the desired conclusion, that is

$$\lim_{n\to+\infty}n\int_G f(\tau)d\mu^n(g)=\widehat{f}(0)\int_{\mathbb{R}}2\pi e^{-r^2\sigma^2/2}rdr=\frac{1}{2\pi\sigma^2}\int_{\mathbb{R}^2}f(y)dy$$

since $\widehat{f}(0)=\frac{1}{(2\pi)^2}\int f$.

Since $\widehat{f}$ has compact support, we can restrict the domain of integration in (4) to a compact set $[0,M]$. Then we get the following lemma:

Lemma 2.1. *There is a constant $c>0$ such that $\|\pi_r(\mu)\|\le 1-cr^2$ when r lies in a neighborhood of 0, i.e. $[0,\varepsilon]$, and moreover $s=\sup_{r\in[\varepsilon,M]}\|\pi_r(\mu)\|<1$.*

Before we proceed to the proof of the lemma, let us explain how to deduce the theorem from it. From the lemma, we obtain:

$$\left|\left\langle\pi_{r/\sqrt{n}}(\mu)^n\widehat{f}_{r/\sqrt{n}},1\right\rangle\right|\le s^n\left\|\widehat{f}\right\|_\infty\longrightarrow 0$$

as soon as $r\in[\varepsilon\sqrt{n},M\sqrt{n}]$, hence this term is negligible. And

$$\left|\left\langle\pi_{r/\sqrt{n}}(\mu)^n\widehat{f}_{r/\sqrt{n}},1\right\rangle\right|\le(1-\frac{cr^2}{n})^n\left\|\widehat{f}\right\|_\infty\le e^{-cr^2}\left\|\widehat{f}\right\|_\infty$$

for the interval $r\in[0,\varepsilon\sqrt{n}]$. This allows to apply Lebesgue's dominated convergence theorem to justify the convergence. □

Proof of the lemma. The proof of these inequalities relies on the essential fact that G is a solvable Lie group because $d=2$. Let us first observe that since $\|\pi_r(\mu)\|^2=\left\|\pi_r(\mu*\mu^{-1})\right\|$, we can assume that μ is symmetric and non-degenerate. Let us show the first inequality. We fix two non-commuting elements x_0 and y_0 belonging to the support of μ. The commutator (x_0,y_0) is a non-trivial pure translation. Similarly we fix two other non-commuting elements w_0 and z_0 in the support of μ such that the pure translation (w_0, z_0) is not co-linear to (x_0, y_0). This is possible since μ is non-degenerate. We then fix a neighborhood U of the identity in G such that the norm of the translation vector (x, y), where $x \in x_0U$ and $y \in y_0U$, and that of (w, z), $w \in w_0U$ and $z \in z_0U$, are bounded below by some fixed positive real number, say $\alpha > 0$. We can also assume that the angle between any two of these translation vectors in bounded below by α. We have:

$$\mu = \int \nu_{x,y,w,z}\mu(dx)\mu(dy)\mu(dw)\mu(dz)$$

where

$$\nu_{x,y,w,z} = \frac{1}{8}\left(\delta_x + \delta_{x^{-1}} + \delta_y + \delta_{y^{-1}} + \delta_w + \delta_{w^{-1}} + \delta_z + \delta_{z^{-1}}\right)$$

Let μ_{x_0U} and μ_{y_0U} be the normalized restrictions of μ to x_0U and y_0U. Then there exists $c_0 > 0$ and two probability measures ν_1 and ν_2 such that

$$\begin{aligned} \mu &= c_0\nu_1 + (1 - c_0)\nu_2 \\ \nu_1 &= \int \nu_{x,y,w,z}\mu_{x_0U}(dx)\mu_{y_0U}(dy)\mu_{w_0U}(dw)\mu_{z_0U}(dz) \end{aligned}$$

Therefore

$$\begin{aligned} 1 - \|\pi_r(\mu)\| &\geq c_0(1 - \|\pi_r(\nu_1)\|) \\ &\geq c_0(1 - \sup_{x \in x_0U, \ldots, z \in z_0U} \|\pi_r(\nu_{x,y,w,z})\|) \end{aligned}$$

It remains to show the desired inequality for $\nu_{x,y,w,z}$ uniformly when $x \in x_0U, ..., z \in z_0U$. But it is now easy to obtain such a bound since it is enough to get it for $\nu_{x,y,w,z}^4$ in place of $\nu_{x,y,w,z}$ and $\nu_{x,y,w,z}^4$ is a finitely supported symmetric probability measure whose support contains the pure translations (x, y), $(x, y)^{-1}$, (w, z) and $(w, z)^{-1}$. The length of these translations and the angle between them are bounded below by $\alpha > 0$, hence we deduce immediately that for any sufficiently small $r > 0$ and uniformly in $x, ..., z$

$$1 - \|\pi_r(\nu_{x,y,w,z})\| \geq C \cdot r^2$$

for some constant $C > 0$ depending only on μ. This follows from the fact that since the angle between $t_1 = (x, y)$ and $t_2 = (w, z)$ is bounded below, the quantity $(t_1 \cdot \theta)^2 + (t_2 \cdot \theta)^2$ is bounded below by some positive constant, uniformly when θ varies in S^1. This ends the proof of the first estimate.

For the second estimate, things are easier and we refer the reader to Proposition 3.1 below, or to [21] where a similar argument (in fact essentially present already in [46]) will be found. □

Note that the theorem is stated only for dimension 2. Its validity in higher dimensions is an open question as of now. Even the extension of Kazhdan's ratio limit theorem to higher dimensions is unknown.

This above proof illustrates the fact that a crucial step in proving equidistribution is to obtain an estimate of the norm $\|\pi(\mu)\|$ for a given unitary representation π and in particular the proof of a "spectral gap" $\|\pi(\mu)\| < 1$.

3. The case of nilpotent groups

In this section N will denote a simply connected nilpotent Lie group. We will also denote by $(C^p(N))_{p=1,...,r}$ the descending central series corresponding to N, where $C^1(N) =$

N and r is the largest index for which $C^r(N)$ is non trivial. Moreover, we write

$$d(N) = \sum_{p \geq 1} p \cdot \dim(C^p(N)/C^{p+1}(N)). \tag{6}$$

a call this number the *exponent of polynomial growth* of the group N, or equivalently the *homogeneous dimension* of N.

3.1. **Polynomials and dilations.** The main source for this paragraph is the book by Goodman [41].

3.1.1. *Polynomial maps.* Let $\mathcal{N}$ be the Lie algebra of N. The exponential map gives rise to a diffeomorphism $exp : \mathcal{N} \to N$. A polynomial map on N is a map $P : N \to \mathbb{R}$ such that $P(\log(x))$ is a polynomial map on the real vector space $\mathcal{N}$.

Let $(e_i)_{i=1,...,n}$ be a basis of $\mathcal{N}$, with $n = \dim_{\mathbb{R}} \mathcal{N}$. We assume further that this basis is *triangular* in the sense that $[e_i, e_j]$ belongs to *span* $(e_k, k \geq \max\{i,j\})$ for all i, j. For every index $i = 1, ..., n$ we set d_i to be the largest positive integer r such that $e_i \in C^r(\mathcal{N})$.

This allows to define a notion **degree** of a polynomial map on N in the following way. Let $(x_i)_{i=1,..,n}$ be the coordinate maps (i.e. $x_i(y) = y_i$ if $y = y_1e_1 + ... + y_ne_n$), then any polynomial map on N is a linear combination of monomials of the form $x_1^{\alpha_1} \cdot ... \cdot x_n^{\alpha_n}$ which we simply denote by x^α, where $\alpha = (\alpha_1, ..., \alpha_n)$ is a multi-index and α_i's are non-negative integers. Then we *define* the degree of x_i to be the number d_i defined above. The degree of the monomial x^α will be equal to $d(\alpha) := \alpha_1 d_1 + ... + \alpha_n d_n$, and the degree of an arbitrary polynomial map is equal to the maximum of degree of each of its monomials. For example, let N be the Heisenberg group of upper triangular three by three matrices, i.e. the step-2 nilpotent group whose Lie algebra is defined a basis (X, Y, Z) and the relation $[X, Y] = Z$. The coordinate maps x and y are polynomials of degree 1 although the coordinate map z has degree 2. This definition is easily seen to be independent of the choice of the triangular basis used to define it.

The coordinates of the product of two elements in the basis $(e_i)_{i=1,..,n}$ are obtained from the Campbell-Hausdorff formula as follows (see Goodman [41] p. 14):

$$(xy)_i = x_i + y_i + P_i(x_1, ..., x_{i-1}, y_1, ..., y_{i-1}) \tag{7}$$

where P_i is a polynomial map on $N \times N$ with the following properties. Its total degree, as a polynomial on $N \times N$, is less or equal to d_i. It depends only the first $i-1$ coordinates of x and y. As a polynomial of x (resp. y) it has degree less or equal to $d_i - 1$ and has valuation greater or equal to 1.

3.1.2. *Associated graded algebra, dilations.* We identify N and $\mathcal{N}$ via the exponential map. We will write $gr(\mathcal{N})$ for the graded Lie algebra canonically associated to $\mathcal{N}$. By definition $gr(\mathcal{N}) = \bigoplus_{p \geq 1} C^p(\mathcal{N})/C^{p+1}(\mathcal{N})$ is endowed by the Lie bracket induced from that of $\mathcal{N}$. Let $D_k = \bigoplus_{p \geq k} C^p(\mathcal{N})/C^{p+1}(\mathcal{N})$. The D_k's form a filtration of $gr(\mathcal{N})$, that is

$$gr(\mathcal{N}) = D_1 \supsetneq D_2 \supsetneq ... \supsetneq D_{r+1} = \{0\} \text{ et } [D_k, D_l] \subset D_{k+l}$$

in the same way that the $C^p(\mathcal{N})$'s formed a filtration of $\mathcal{N}$.

From now on, we will fix a collection of sub-vector spaces $(m_p)_{p \geq 1}$ of $\mathcal{N}$ such that for each i,

$$C^p(\mathcal{N}) = C^p(\mathcal{N}) \oplus m_p.$$

Then $\mathcal{N} = \oplus_{p\geq 1} m_p$ and in this decomposition, any element x in $\mathcal{N}$ (or N by abuse of notation) will be written in the form

$$\mathbf{x} = \sum_{p\geq 1} \mathbf{x}_p$$

This allows to define on $\mathcal{N}$ another Lie algebra structure. Let $\mathcal{N}'$ be the same vector space $\mathcal{N}$ endowed with this new Lie bracket defined by $[\mathbf{x}_p, \mathbf{y}_q]' = [\mathbf{x}_p, \mathbf{y}_q]_{p+q}$.

An equivalent way to define this new Lie algebra structure is as follows. We can define a linear map

$$\phi = (\phi_1, ..., \phi_r) : \mathcal{N} \to gr(\mathcal{N})$$

by the property that if $\mathbf{x} \in m_p$, then $\phi(\mathbf{x}) = \phi_i(\mathbf{x}) = \mathbf{x}\ mod(C^{p+1}(\mathcal{N}))$. In this way, ϕ is an isomorphism of vector spaces which preserves the respective filtrations, i.e. $\phi(C^p(\mathcal{N})) \subset D_p$. Moreover, ϕ induces on m_p an isomorphism with $C^p(\mathcal{N})/C^{p+1}(\mathcal{N})$ which coincide with the canonical quotient map. Then ϕ establishes an isomorphism of Lie algebras between $\mathcal{N}'$ and $gr(\mathcal{N})$.

A choice of supplementary subspaces $(m_p)_{p\geq 1}$ allows to define a semi-group $(\delta_t)_{t\geq 0}$ of linear transformations of $\mathcal{N}$ called **dilations** as follows:

$$\delta_t(\mathbf{x}_p) = t^p \mathbf{x}_p$$

Conversely, the semi-group $(\delta_t)_{t\geq 0}$ determines the $(m_p)_{p\geq 1}$ since they appear as eigenspaces of each δ_t. The dilations δ_t's preserve the new Lie algebra structure $\mathcal{N}'$ but do not preserve *a priori* the original Lie algebra structure on $\mathcal{N}$. Moreover, we see that $\mathcal{N}$ and $\mathcal{N}'$ (or $gr(\mathcal{N})$) are isomorphic as Lie algebras if and only if the δ_t's are automorphisms of $\mathcal{N}$.

We will say that N (or $\mathcal{N}$) is **homogeneous** if one can find a sequence of supplementary subspaces $(m_p)_{p\geq 1}$ such that the corresponding δ_t's are automorphisms (equivalently $[m_p, m_q] \subset m_{p+q}$).

3.2. **Homogeneous norms.** For this paragraph, we refer the reader to the following two articles of Y. Guivarc'h [49] and [46].

Let U be a compact neighborhood of the identity in N. We define as above the function δ_U as follows

$$\delta_U(g) = \inf\{n \geq 0, g \in U^n\}$$

The function δ_U is sub-additive (i.e. $\delta_U(gh) \leq \delta_U(g) + \delta_U(h)$) and any other choice of U would lead to a *coarsely equivalent* function (i.e. there are positive constants A and B such that $\delta_U(g) \leq A\delta_V(g) + B$ and vice-versa). This allows to define a left-invariant distance on the Lie group by letting $\delta_U(x, y) = \delta_U(x^{-1}y)$. This distance is usually called the *word metric* associated to the generating set U. The following fact is straightforward, and valid on any connected Lie group:

Proposition 3.1. *Every left-invariant Riemannian metric is coarsely equivalent to δ_U.*

On a simply connected nilpotent Lie group N endowed with a semi-group of dilations $(\delta_t)_t$ and their eigenspaces $(m_p)_p$ (we do not assume that the δ_t's are automorphisms), it is convenient to define the following notion.

Definition 3.1. *A continuous function $|\cdot| : N \to \mathbb{R}_+$ is called on homogeneous norm on N for the dilations $(\delta_t)_t$ if it satisfies the following properties*

(i) $|x| = 0 \Leftrightarrow x = 0$

(ii) $|\delta_t(x)| = t|x|$ *for all* $t > 0$

Clearly a homogeneous norm is determined by its sphere of radius 1 and two homogeneous norms are always equivalent in the sense that $\frac{1}{c}|\cdot|_1 \leq |\cdot|_2 \leq c|\cdot|_1$ for some constant $c>0$. Also changing $|x|$ into $(|x|+|x^{-1}|)/2$ gives rise to a symmetric homogeneous norm. More generally, the homogeneous norms are easily seen to satisfy the following properties. There is a constant $C>0$ such that

(*i*) $|x^{-1}| \leq C \cdot |x|$.

(*ii*) $|x_i| \leq C \cdot |x|^{d_i}$ if $x = x_1 e_1 + ... + x_n e_n$.

(*iii*) $|xy| \leq C(|x|+|y|)$

It can be a problem that the constant C in (*iii*) may not be 1. However the following proposition is often a good enough remedy to this situation. Let $\|\cdot\|_p$ be an arbitrary norm on the vector space m_p.

Proposition 3.2. *([46] lemme II.1) Up to rescaling each $\|\cdot\|_p$ into a proportional norm $\lambda_p \|\cdot\|_p$ ($\lambda_p > 0$) if necessary, the function $|x|_N := \max_p \|\mathbf{x}_p\|_p^{1/p}$ (where $x = \mathbf{x}_1 + ... + \mathbf{x}_r$, $\mathbf{x}_p \in m_p$) becomes a homogeneous norm on N which coarsely equivalent to δ_U and is almost sub-additive, i.e. there exists a constant $c>0$ such that $|xy| \leq |x|+|y|+c$ for all $x, y \in N$.*

The proof is based on the Campbell-Hausdorff formula (7). As we show below this proposition has a very nice immediate application, namely the computation of the exponent of growth of nilpotent groups.

The following definition is useful.

Definition 3.2. *Let $(X_1, ..., X_n)$ be a basis of $\mathcal{N}$ adapted to the choice of supplementary subspaces $(m_p)_p$. A subset of N is called a* ***rectangle*** *in N is it consists of all elements $x = t_1 X_1 + ... + t_n X_n$ such that $t_i \in [a_i, b_i]$ for some intervals $[a_i, b_i]$.*

Remark 3.1. *It is also possible to define an almost sub-additive function on N that is coarsely equivalent to δ_U by considering exponential coordinates of second kind to parametrize N instead of taking the coordinates from the Lie algebra. Let $(X_1, ..., X_n)$ be an adapted basis of $\mathcal{N} = \oplus_{p \geq 1} m_p$ such that $span(X_k, ..., X_n)$ is an ideal of $\mathcal{N}$ for each k, and consider the map*

$$\begin{array}{rcl} \phi & : & \mathcal{N} \to N \\ \sum t_i X_i & \mapsto & \prod \exp(t_i X_i) \end{array}$$

The ϕ is a polynomial diffeomorphism together with its inverse (see [27]). Moreover set $\delta(x) = \max_i |t_i|^{1/d_i}$. Then as in proposition 3.2, up to rescaling the basis X_i's if necessary, we obtain an almost sub-additive function on N which is coarsely equivalent to δ_U.

3.3. **The growth of nilpotent groups.** In this section $|\cdot|$ will denote the homogeneous norm introduced in Proposition 3.2. The existence of such a homogeneous norm, which is coarsely equivalent to $\delta_U(g) = \inf\{n \geq 0, g \in U^n\}$ has the following direct corollary.

Theorem 3.1. *(Guivarc'h) Let U be a compact neighborhood of the identity in the simply connected nilpotent Lie group N. Then there are positive constants C_1 and C_2 such that for any positive integer n*

$$C_1 \cdot n^{d(N)} \leq vol(U^n) \leq C_2 \cdot n^{d(N)} \tag{8}$$

where $vol(X)$ denotes the Haar measure of X and $d(N)$ is the exponent of growth defined in (6).

The proof follows immediately from the fact that, since $|\cdot|$ is coarsely equivalent to δ_U, there is a constant $c > 0$ such that for all positive integers n,

$$B_{n/c} \subset U^n \subset B_{cn}$$

where $B_t = \{x \in N, |x| \leq t\}$. Moreover $|\cdot|$ is a homogeneous norm and hence $B_t = \delta_t(B_1)$, so $vol(B_t) = t^{d(N)} \cdot vol(B_1)$ where $d(N)$ is given by formula (6).

Now if Γ is a finitely generated nilpotent group, say torsion-free, then according to a theorem of Malcev (see [87] Theorem 2.18), there exists a simply connected nilpotent Lie group $\widetilde{\Gamma}$ such that Γ embeds as a co-compact discrete subgroup of $\widetilde{\Gamma}$. We immediately deduce

Theorem 3.2. *(Guivarc'h) Let Γ be a finitely generated nilpotent group and let S be a symmetric generating set for Γ, then there are two positive constants C_1 and C_2 such that for all positive integers n*

$$C_1 \cdot n^{d(\Gamma)} \leq \#B_S(n) \leq C_2 \cdot n^{d(\Gamma)}$$

where $d(\Gamma) = d(\widetilde{\Gamma})$ is the exponent $\sum i \cdot rk(C^i(\Gamma)/C^{i+1}(\Gamma))$ and $B_S(n)$ is the ball of radius n in the word metric defined by S.

This theorem has also been established independently by H. Bass by a purely combinatorial method (see [53] for a short discussion of this result and older references, see also [64]).

3.4. **Irreducible unitary representations of nilpotent Lie groups.**

3.4.1. *Harmonic functions.* Given a probability measure μ on a group G, a function f on G is called μ-harmonic if it satisfies the convolution equation $\mu * f = f$. Harmonic functions are important in the study of equidistribution and ratio limit theorems (see most notably the article of Le Page [72]). A recent result of A. Raugi [93] shows that bounded μ-harmonic functions on locally compact second countable nilpotent groups are constant whenever μ is non-degenerate. Raugi's proof is based on the martingale convergence theorem and puts a definite end to a question that was pending for some time. The result had been established previously in many special cases. For $G = \mathbb{Z}^d$ this was first treated by Blackwell, then subsequently by Choquet and Deny for an arbitrary locally compact second countable abelian group (see [26]). For nilpotent groups, assuming G is finitely generated and the support of μ generates G as a semi-group, the result was proved by Dynkin and Maliutov in [33]. In [46], Guivarc'h treats the case of an arbitrary locally compact second countable nilpotent group, but his argument requires that the measure μ has a finite moment of positive order. Guivarc'h's method belongs to harmonic analysis and yields a more precise and quite remarkable result which is a sort of weak local limit theorem on G.

Theorem 3.3. *(Guivarc'h) Let G be a locally compact second countable nilpotent group and μ an aperiodic probability measure on G. Then for every integrable function f on G such that $\int_G f = 0$, we have*

$$\lim_{n \to +\infty} \|\mu^n * f\|_{\mathbb{L}^1(G)} = 0$$

This theorem is applied in [20] in order to prove certain equidistribution results on homogeneous spaces. The absence of non-constant μ-harmonic functions can be derived as follows: if ϕ is a bounded and continuous μ-harmonic function, then $\phi * \mu^n = \phi$ for every positive integer n. Thus

$$\|\phi * f\|_\infty = \|\phi * \mu^n * f\|_\infty \leq \|\phi\|_\infty \|\mu^n * f\|_1$$

and by the above theorem $\phi * f$ must be identically zero for every integrable function f whose integral on G vanishes. This means precisely that ϕ is constant.

The proof of Guivarc'h relies on representation theory and on the following spectral gap property for operators associated to an aperiodic probability measure:

Proposition 3.1. *Let G be a locally compact second countable nilpotent group and μ and probability measure on G. Then the following conditions are equivalent:*
(i) μ is aperiodic
(ii) for any non-trivial irreducible unitary representation π of G we have $\|\pi(\mu)\| < 1$.

See [46] (Prop. V.4). Similar arguments can already be found in [43].

Proof. Suppose that μ is aperiodic and π is a non-trivial irreducible unitary representation of G. Let us first show $(i) \Rightarrow (ii)$. When G is abelian then π is a non trivial character χ of G and if $\|\pi(\mu)\| = |\chi(\mu)| = 1$, then there must be a complex number z of modulus 1 such that for μ-almost all $g \in G$ we have $\chi(g) = z$. This contradicts the aperiodicity of μ because $\ker \chi$ is a proper closed subgroup of G.

Now assume that G is arbitrary and that $\|\pi(\mu)\| = 1$. From Shur's lemma, the restriction of π to $C^r(G)$ (which is a subgroup of the center of G) coincide with a character of G, that is a homomorphism $\chi : C^r(G) \to \mathbb{C}^\times$. Let (ξ_n) be such that $\|\pi(\mu)\xi_n\| \to 1$. Equivalently, this means that

$$\int_G \langle \pi(g)\xi_n, \xi_n \rangle \, \mu * \mu^{-1}(dg) \to 1$$

Now passing to a subsequence if necessary, we can assume that $\mu * \mu^{-1}$-almost everywhere $\langle \pi(g)\xi_n, \xi_n \rangle \to 1$, or in other words $\pi(g)\xi_n - \xi_n \to 0$. But $\Gamma = \{g \in G, \pi(g)\xi_n - \xi_n \to 0\}$ is a subgroup of G such that $\mu * \mu^{-1}(\Gamma) = 1$. The aperiodicity of μ implies that Γ is dense in G. It follows that $C^r(\Gamma)$ is dense in $C^r(G)$. Moreover, if $\gamma \in C^r(\Gamma)$, $(\chi(\gamma) - 1)\xi_n \to 0$ and this implies that $\chi(\gamma) = 1$. Consequently, $\chi(g) = 1$ for all $g \in C^r(G)$ and χ is the trivial character. It follows that the representation π induces a representation $\overline{\pi}$ of the quotient $G/C^r(G)$ and $\pi(\mu) = \overline{\pi}(\overline{\mu})$ where $\overline{\mu}$ is the image of μ in the canonical projection onto $G/C^r(G)$. Moreover $\overline{\pi}$ is a non trivial irreducible unitary representation of $G/C^r(G)$. But since μ is aperiodic, $\overline{\mu}$ is also aperiodic and it follows by induction on the degree of the central descending series of G that $\|\overline{\pi}(\overline{\mu})\| < 1$. A contradiction.

To prove the converse, note that if μ were not aperiodic there would exist a proper closed subgroup H in G and $g \in G$ such that μ is supported on gH. To conclude it suffices to apply the following fact[2]:

Lemma 3.1. *([45]) Let H be a proper closed subgroup of a locally compact nilpotent group G. Then H lies in the kernel of a non-trivial continuous character of G.*

Indeed we then obtain some $\theta \in \mathbb{R}$ such that $\chi(x) = e^{i\theta}$ for μ-almost all x in G. Hence $|\chi(\mu)| = 1$.

To prove the lemma, first note that if H is normal, then we can take any non trivial character of the quotient G/H. Otherwise, we proceed by induction on the largest integer p such that H contains $C^{p+1}(G)$. If $p = 1$, then the image of H under the canonical

---

[2] In particular n elements in N generate a dense subgroup in N if and only if their projection in the maximal abelian quotient $N/[N, N]$ generate a dense subgroup.

projection on $G/\overline{[G,G]}$ is a proper closed subgroup of the abelian group $G/\overline{[G,G]}$, so it is normal. If $p>1$ and H is not normal, its normalizer L is a proper closed subgroup containing $C^p(G)$. By induction L (hence also H) must be contained in some $\ker\chi$ for some non trivial character χ of G.$\square$

3.4.2. *Quantitative estimates of the spectral gap.* The classical proof of the local limit theorem on $\mathbb{R}$ is based on an estimation of the modulus of the characteristic function $\widehat{\mu}(t)$ of the measure μ when t lies in some neighborhood of zero. In the above proof of the local limit theorem on the group of motions of the plane (lemma 2.1), a similar estimate is needed. In [21], we prove the local limit theorem for centered measures on the Heisenberg group by following a similar strategy. We have to estimate the norm $\|\pi(\mu)\|$ when π is a unitary representation of G lying in a neighborhood of the trivial representation for the Fell topology on the unitary dual of G. For the definition of the Fell topology, we refer the reader to [69] and [54].

For a nilpotent Lie group, the unitary dual is fairly well understood thanks to the Dixmier-Kirillov theory (see [70]). For every linear form ℓ on the Lie algebra $Lie(N)$ of N, one can find a subalgebra $\mathfrak{m}$ such that $[\mathfrak{m},\mathfrak{m}]\subset\ker\ell$ which is maximal for this property. This allows to define first a character of the group $M=\exp(\mathfrak{m})$ by setting $\chi_\ell(\exp(v))=e^{i\ell(v)}$, and then a representation $\pi_{\ell,\mathfrak{m}}$ of N by inducing this character to the whole of N, $\pi_{\ell,\mathfrak{m}}=Ind_M^N\chi_\ell$. The we have (see [70]),

Theorem 3.4. *(Kirillov) The representation $\pi_{\ell,\mathfrak{m}}$ is irreducible and any two choices for the maximal subalgebra $\mathfrak{m}$ lead to equivalent representations. Two linear forms ℓ and ℓ' are conjugate under the action of N if and only if they lead to equivalent representations. Moreover any irreducible unitary representation of N is equivalent to a representation of this form.*

It follows that the unitary dual of N is naturally identified to the space of orbits for the action of N on the space of linear forms on $Lie(N)$, i.e. the dual vector space of $Lie(N)$ (co-adjoint action). It turns out that this identification is also a homeomorphism between the Fell topology on the one hand and the quotient topology on the other hand (Brown [23]).

The proof of Theorem 3.4 begins with the preliminary study of an important special case: the Heisenberg group. For this group the theorem is a consequence of the following theorem of Stone and Von Neumann.

Theorem 3.5. *(Stone-Von Neumann) Let $\mathcal{H}$ be a separable Hilbert space, and A and B be two self-adjoint operators in $\mathcal{H}$ such that $AB-BA=id$. Then $\mathcal{H}$ decomposes as a completed direct sum $\mathcal{H}=\widehat{\oplus}_{i\geq 0}\mathcal{H}_i$ of countably many invariant subspaces $\mathcal{H}_i$ each isomorphic to $\mathbb{L}^2(\mathbb{R})$ and on which A and B are simultaneously equivalent (i.e. conjugate via a single unitary transformation) to the operators $T_1=i\frac{d}{dt}$ and $T_2=t$, the multiplication by t in $\mathbb{L}^2(\mathbb{R})$.*

For one-parameter subgroups of unitary operators, this theorem admits the following equivalent formulation: if ρ_1 and ρ_2 are two unitary representations of the additive group $(\mathbb{R},+)$ in $\mathcal{H}$ such that $\rho_1(u)\rho_2(v)\rho_1^{-1}(u)\rho_2^{-1}(v)=e^{icuv}$ (for some $c>0$) then $\mathcal{H}$ decomposes as $\mathcal{H}=\widehat{\oplus}_{i\geq 0}\mathcal{H}_i$ where $\mathcal{H}_i\simeq\mathbb{L}^2(\mathbb{R})$ and ρ_1 and ρ_2 are conjugate to multiplication by e^{it} and translation by c.

This theorem allows to classify the irreducible unitary representations of the Heisenberg group as follows[3]. Recall that the Heisenberg group is by definition the group of upper triangular matrices in $GL_3(\mathbb{R})$ with 1's on the diagonal. In the standard matrix coordinates, the multiplication has the form:

$$(x, y, z) \cdot (x', y', z') = (x + x', y + y', z + z' + xy')$$

Irreducible unitary representations are of two types: there are characters ($\chi(x, y, z) = e^{i(ax+by)}$ for some real numbers a and b) and there is a one-parameter family of representations of infinite dimension $(\pi_\lambda)_\lambda$ parametrized by the non-zero characters λ of the center of H. One possible model for π_λ is given on $\mathbb{L}^2(\mathbb{R})$ by

$$\pi_\lambda(x, y, z) \cdot f(t) = e^{i\lambda(z+yt)} f(t + x)$$

In [21], we make use of the Stone-Von Neumann theorem to obtain the following estimate, which turns out to be crucial in the proof of the local limit theorem on H:

Proposition 3.2. *Let μ be a probability measure on the Heisenberg group H whose support is not contained in a coset of a proper abelian subgroup of H. Then there exists a constant $c > 0$ such that for every $\lambda \neq 0$ sufficiently small :*

$$\|\pi_\lambda(\mu)\| < 1 - c|\lambda| \tag{9}$$

In particular, this result applies to the symmetric measure $\mu_0 = \frac{1}{4}(\delta_a + \delta_{a^{-1}} + \delta_b + \delta_{b^{-1}})$ where $a = (1, 0, 0)$ and $b = (0, 1, 0)$. The operator $\pi_\lambda(\mu_0)$ can be seen as acting on $\ell^2(\mathbb{Z})$. Then it coincide with the well-known discrete Schrodinger operator called the Harper operator and whose spectrum has been extensively studied in mathematical physics: see most notably Figure 1 in [10] and Theorem 2.1. in [14] which justify the estimate (9).

3.5. **Limit theorems for discrete nilpotent groups.** In this paragraph, we will quote some of the numerous known results on the asymptotic behavior of convolution powers of a probability measure on a finitely generated group. We will restrict mostly to groups of polynomial growth, that is virtually nilpotent groups according to Gromov's theorem. If the group G is discrete, then any probability measure on G is absolutely continuous with respect to the Haar measure of G, i.e. the counting measure on G. In this case, the analytic methods initiated by N. Varopoulos and developed subsequently by many other mathematicians turn out to be remarkably efficient (see the books [113] and [117]). However these methods do not allow, as far as I know, to treat the case of singular probability measures (for example atomic probability measures) on Lie groups, a case which involves some arithmetic aspects related to how dense subgroups equidistribute.

Le us begin with a very general result due to Avez [6] :

Theorem 3.6. *(Avez) Let Γ be a finitely generated amenable group and μ a symmetric probability measure with finite support on Γ. Then for any x and y in Γ,*

$$\lim_{n\to+\infty} \frac{\mu^{2n}(x)}{\mu^{2n}(y)} = 1$$

In [113], Varopoulos obtains a coarse local limit theorem for finitely supported symmetric measures on a finitely generated group of polynomial growth. More precisely:

[3] Instead of the Stone-Von Neumann theorem, one can also use a theorem of Mackey on induced representations: see for instance [118] Theorem 7.3.1 and Example 7.3.2.

Theorem 3.7. *(Varopoulos) Let Γ finitely generated group of polynomial growth and μ a non-degenerate symmetric probability measure with finite support on Γ. Then there exists a constant $c > 1$ such that for all $n \geq 1$*

$$\frac{1}{c}\frac{1}{n^{d(\Gamma)/2}} \leq \mu^{2n}(e) \leq c\frac{1}{n^{d(\Gamma)/2}}$$

The exponent $d(\Gamma)$ is the exponent of growth of the group Γ. According to Gromov's theorem, the polynomial growth group Γ must contain a nilpotent subgroup Γ_n of finite index. Then $d(\Gamma)$ is simply given by the formula (6) with Γ_n in place of N.

Since μ is symmetric, $\mu^{2n}(x) \leq \mu^{2n}(e)$ for all $x \in \Gamma$ (as follows immediately from the Cauchy-Schwarz inequality: see [6]). More generally, if we drop the hypothesis of symmetry, we still get the following upper bound:

Theorem 3.8. *(Varopoulos,* [112]*) Let Γ finitely generated group of polynomial growth and μ a non-degenerate probability measure with finite support on Γ. Then there is a constant $C > 0$ such that for all $n \geq 1$*

$$\sup_{x \in \Gamma} \mu^n(x) \leq \frac{C}{n^{d(\Gamma)/2}}$$

These results have been improved by Hebisch and Saloff-Coste in [56] (symmetric case) and by Alexopoulos [2] (centered case) who obtain a gaussian estimate valid on a large portion of the support of μ^n.

Theorem 3.9. *(*[56] *et* [2]*) Assume that Γ has polynomial growth and μ is a centered probability measure on Γ whose support is finite and contains a symmetric system of generators S ($S \ni e$) of Γ. Then there exist constants $c > 1$ and $\theta \in (0, 1)$ such that*

(i) for all integers $n \geq 1$ and all $x \in \Gamma$

$$\mu^n(x) \leq c\frac{1}{n^{d(\Gamma)/2}} \exp(-\frac{d(e,x)^2}{cn})$$

(ii) for all integers $n \geq 1$ and all $x \in \Gamma$ such that $d(e, x) \leq \theta n$

$$\mu^n(x) \geq \frac{1}{cn^{d(\Gamma)/2}} \exp(-c\frac{d(e,x)^2}{n})$$

where d is the left-invariant distance induced by the system of generators S (word metric).

In [2], Alexopoulos improves these gaussian estimates and obtains the local limit theorem with a sharp and uniform control on the speed of convergence as follows. We assume as before that Γ has polynomial growth and that μ is a centered probability measure on Γ whose support is finite and contains a symmetric set of generators $S \ni e$ of Γ. From Gromov's theorem, Γ contains a subgroup of finite index which is nilpotent. It is easy to see that, up to passing to smaller finite index subgroup, this finite index nilpotent subgroup can be taken to be normal and torsion-free (see [87] lemma 4.6). According to Malcev's theorem (see [87]), there exists a simply connected nilpotent Lie group N, unique up to isomorphism, such that Γ_N is isomorphic to a discrete co-compact subgroup of N.

Let $g_1, ..., g_k$ be representatives of the cosets of Γ_N in Γ. By Malcev's rigidity theorem ([87] 2.11), the automorphisms $h \mapsto g_i h g_i^{-1}$ extend to the whole of N. This allows to define the Lie group $G = \{hg_i, h \in N, 1 \leq i \leq k\}$, in which Γ injects as a discrete co-compact subgroup.

Alexopoulos then defines a left-invariant sub-laplacian L_μ on N associated to the measure μ. When Γ is itself nilpotent the sub-laplacian has the simple form given by the formula (10) appearing in the central limit theorem. The associated heat kernel $\widetilde{p}_t(x, y)$

$(x, y \in N)$ is smooth (see above §1.4) and its coarse asymptotic behavior is well-known (see [113]). We extend it to the whole of G by setting:

$$p_t(g_i x, g_j y) = \frac{1}{k}\widetilde{p}_t(x, y).$$

Theorem 3.10. *(Alexopoulos [2]) We have*

(*i*) *There is a real number $C(\mu) > 0$ such that for all $x \in \Gamma$*

$$\lim_{n\to\infty} n^{d(\Gamma)/2} \cdot \mu^n(x) = C(\mu)$$

(*ii*) *There is a constant $c > 0$ such that for any integer $n \geq 1$ and any $x \in \Gamma$, we have*

$$|\mu^n(x) - p_n(e, x)| \leq \frac{c}{n^{(d(\Gamma)+1)/2}} \exp\left(-\frac{d(e, x)^2}{cn}\right)$$

To a large extent, this theorem is optimal (obviously, it generalizes the classical theorems obtained in [40] in the abelian case by means of Fourier techniques) and it allows to answer many questions on the behavior of centered random walks on a finitely generated group of polynomial growth (recurrence, equidistribution, harmonic functions, etc). The proofs are based on a certain parabolic Harnack inequality for solutions of the heat equation associated to μ. Let us note that the speed of convergence here is of order $1/\sqrt{n}$. In [22] we observe that in the non-discrete case, and already in the case of $\mathbb{R}$, the speed of convergence in the local limit theorem can be arbitrarily slow for an arbitrary measure μ and in fact depends on the diophantine properties of μ.

3.6. **Limit theorems for nilpotent Lie groups.**

3.6.1. *The central limit theorem.* We have already proved the central limit theorem on Lie groups (see above Wehn's theorem §2.4). In the case of simply connected nilpotent Lie groups, Wehn's theorem admits the simple corollary stated below as Theorem 3.11. Let us fix a sequence of supplementary subspaces $(m_p)_p$ (cf. §3.1.2) and let δ_t be the semi-group of dilations associated to this choice:

$$C^p(\mathcal{N}) = m_p \oplus C^{p+1}(\mathcal{N})$$

and $\delta_t(x) = t^p x$ whenever $x \in m_p$. Let us denote by N' the corresponding graded Lie group and $\mathcal{N}'$ its Lie algebra. As vector spaces $\mathcal{N} = \mathcal{N}'$ but the Lie algebra structure may differ. Let $X_1, ..., X_d$ be left-invariant vector fields on N such that $X_1(e), ..., X_d(e)$ forms a basis of $\mathcal{N}$ which is adapted to the direct sum decomposition $\mathcal{N} = \oplus_{p\geq 1} m_p$ (i.e. the $X_i(e)$'s belonging to m_p form a basis of m_p). If $g \in N$ we write $x_i(g)$ to be the coordinate of $\log(g)$ in this basis, i.e. $\log(g) = \sum x_i(g) X_i(e)$. We also introduce the following notation: we let $n_i = \dim(m_1 \oplus ... \oplus m_i)$ and X'_i be the left-invariant vector fields on N' such that $X'_i(e) = X_i(e)$.

Theorem 3.11. *(**CLT centered case**) Let μ be a centered probability measure on N admitting a finite moment of order 2. Let S_n be the random walk associated to μ starting at the identity in N. Then we have the following convergence in distribution (i.e. convergence of the probability measures):*

$$\delta_{\frac{1}{\sqrt{n}}}(S_n) \xrightarrow{d} X$$

where $X = X_1$ is the value at time 1 of the gaussian process $(X_t)_{t\geq 0}$ on $\mathcal{N}'$ whose infinitesimal generator (see §2.4) is the N'-left-invariant sub-laplacian given on $C^2_c(N)$ by

$$L_\mu = \sum_{i=n_1+1}^{n_2} b_i X'_i + \frac{1}{2} \sum_{1\leq i,j\leq n_1} a_{ij} X'_i X'_j \tag{10}$$

where

$$a_{ij} = \int x_i(x)x_j(x)d\mu(x) \text{ if } 1 \leq i,j \leq n_1 \tag{11}$$
$$b_i = \int x_i(x)d\mu(x) \text{ if } n_1 < i \leq n_2$$

The limit law is non-degenerate if and only if the n_1 by n_1 matrix $(a_{ij})_{1\leq i,j\leq n_1}$ is positive definite or equivalently if the support of the measure μ is not contained in a closed proper connected subgroup of N. In this case the limit law has a smooth density with respect to the Haar measure on N (see above 1.4).

The stochastic process $(X_t)_t$ is *stable* for (δ_t). This means that for each t the random variable X_t has the same distribution as the random variable $\delta_{\sqrt{t}}(X_1)$ (denoted $X_t \stackrel{d}{=} \delta_{\sqrt{t}}(X_1)$).

Note that in the expression of L_μ, the elements of $\mathcal{N}$ are considered as differential operators acting on $C^2(N)$. Also note that the limit (X et L_μ) depends on the choice of the semi-group of dilations, i.e. on the choice of the supplementary subspaces $(m_p)_p$. If we consider two possible choices, say (a) and (b) such that $\mathcal{N} = \oplus_{i\geq 1} m_i^{(a)} = \oplus_{i\geq 1} m_i^{(b)}$ then one can check that the limit diffusion processes $X_t^{(a)}(\mu)$ and $X_t^{(b)}(\mu)$ satisfy the following identity in law

$$\phi_{ab}(X_t^{(a)}(\mu)) \stackrel{d}{=} X_t^{(b)}(\mu) \tag{12}$$

where ϕ_{ab} is the endomorphism of the vector space $\mathcal{N}$ which maps every element of $m_p^{(a)}$ to its projection on $m_p^{(b)}$. The map ϕ_{ab} establishes an isomorphism between the induced Lie algebras structures $\mathcal{N}^{(a)}$ and $\mathcal{N}^{(b)}$ defined on $\mathcal{N}$ by the choice (a) or (b) a supplementary subspaces (cf. §3.1.2). Moreover $\phi_{ab} = \phi_{ba}^{-1}$. The relation between the corresponding semi-groups of dilations is given by

$$\phi_{ab} \circ \delta_t^{(a)} = \delta_t^{(b)} \circ \phi_{ab}$$

The relation (12) follows easily from the theorem and from the following observation: we have the following identity in law $X_t^{(a)}(\phi_{ba}(\mu)) \stackrel{d}{=} X_t^{(a)}(\mu)$ because $a_{ij}(\phi_{ba}(\mu)) = a_{ij}(\mu)$ if $1 \leq i,j \leq n_1$ and $b_i(\phi_{ba}(\mu)) = b_i(\mu)$ if $n_1 < i \leq n_2$. Indeed we check that $x_i^{(a)} \circ \phi_{ba} = x_i^{(a)}$ if $1 \leq i \leq n_1$ and $x_i^{(a)} \circ \phi_{ba} = x_i^{(a)} + \ell_i(x_1^{(a)}, ..., x_{n_1}^{(a)})$ where ℓ_i is some linear form if $n_1 < i \leq n_2$. Then taking the averages with respect to μ, and making use of the fact that μ is centered, in obtain the desired identities.

When N is homogeneous (see the definition above in §3.1.2), the theorem is a particular case of Wehn's theorem, since the condition of existence of a finite second moment for μ is easily seen to imply conditions (ii) and (iii) of Theorem 1.3 for $\mu_n := \delta_{\frac{1}{\sqrt{n}}}(\mu)$. For another short proof of this theorem, see [52]. The general case reduces to the case when N is homogeneous thanks to a lemma due to Crépel et Raugi ([30] Lemme 3.5) asserting that the difference $\delta_{\frac{1}{\sqrt{n}}}(S_n) - \delta_{\frac{1}{\sqrt{n}}}(S'_n)$ converges to zero in L^2, where S'_n is the random walk associated to μ on the graded nilpotent group N' corresponding to the dilations δ_t's.

It is also possible to give an Edgeworth expansion for $\delta_{\frac{1}{\sqrt{n}}}(S_n)$. This has been done by Bentkus and Pap in [12] where they apply Lindeberg's method to give yet another proof of Theorem 3.11 that yields (as it does in the classical $\mathbb{R}^d$ case) an estimate of the speed of convergence (in $1/\sqrt{n}$).

In the non-centered case, the situation is quite different, but a central limit theorem does exist (see [92] and §6 below).

The following example shows a probability measure on the Heisenberg group H which has no finite moment of order 2 and yet satisfies the central limit theorem above. This constitutes a major difference with the case of $\mathbb{R}^d$. A similar example already appears in [97].

Example 3.1. *Let (X, Y, Z) be distributed according to a probability measure μ on H. Suppose that X and Y have a finite second moment, i.e. $\mathbb{E}(X^2) < \infty$ and $\mathbb{E}(Y^2) < \infty$. Suppose further that Z has a density $f(z)$ such that*

$$f(z) = \frac{1}{z^2 \log |z|}$$

whenever $|z| \geq 3$. Then $\mathbb{E}(|Z|) = +\infty$, hence μ has infinite second moment on H. However it is straightforward to check that conditions (ii) and (iii) of Theorem 1.3 are satisfied when $\mu_n = \delta_{\frac{1}{\sqrt{n}}}(\mu)$ is the probability distribution of $(\frac{X}{\sqrt{n}}, \frac{Y}{\sqrt{n}}, \frac{Z}{n})$. Hence μ_n^{*n} converges toward a gaussian measure on H.

3.6.2. *Local limit theorems and equidistribution.* Let us now pass to equidistribution properties of random walks on nilpotent groups. The first theorem of this type for nilpotent groups is due to Le Page and deals with finitely supported symmetric measures μ.

Theorem 3.12. *(Le Page [72]) Let G be a locally compact second countable nilpotent group and μ be a finitely supported symmetric probability measure on G such that $\mu(e) > 0$. Then for all continuous functions ϕ and ψ assumed positive and with compact support, we have*

$$\lim_{n \to +\infty} \frac{\int \phi d\mu^n}{\int \psi d\mu^n} = \frac{\int \phi dg}{\int \psi dg}$$

The proof of this theorem relies on the fact that $(\int \phi d\mu^n)^{1/n}$ converges to 1 because G is amenable (see Kesten [68] and Guivarc'h [49]) and on the absence of harmonic positive functions proved in this case by Margulis in [78].

Another old result of this type deals with the equidistribution in homogeneous spaces and is a corollary of Theorem 3.3,

Theorem 3.13. *(Guivarc'h [46] [47]) Let μ be an aperiodic probability measure on N. Let X be a compact metric space endowed with of probability measure m. Suppose that N acts continuously on X by homeomorphisms preserving m and so that m is the only invariant measure under the action of N (unique ergodicity). Then for every $x \in X$ and every continuous function f on X, we have*

$$\lim_{n \to +\infty} \int f(g \cdot x) d\mu^n(g) = \int_X f(y) dm(y) \tag{13}$$

For example, X can be a compact homogeneous space of N or simply the tangent unit sphere bundle over a compact Riemann surface with $N = \mathbb{R}$ acting via the horocyclic flow (its unique ergodicity was first proved by Furstenberg in [37]). Note that we do not assume μ to be centered in this theorem. This additional assumption can be waived here thanks to the unique ergodicity. Below, we give an example of a non-compact Riemann surface X of finite volume for which the convergence (4.1) (for the horocyclic flow) does not hold for every starting point x (although it holds for almost all x) as soon as μ is not centered.

When the measure μ has a compactly supported continuous density with respect to the Haar measure, the recent results of Alexopoulos allow to obtain the analogue of Theorem

3.10. The following result is obtained by the same methods as 3.10. Let $d(x,y)$ be a left-invariant metric on N.

Theorem 3.14. *(Alexopoulos [3]) Let μ be a centered probability measure on N which admits a continuous density of compact support, i.e $\mu = \phi(g)dg$. Suppose additionally that $\phi(e) > 0$. Let L_μ be the left-invariant sub-laplacian defined on N by*

$$L_\mu = \sum_{i=n_1+1}^{n_2} b_i X_i + \frac{1}{2} \sum_{1\leq i,j\leq n_1} a_{ij} X_i X_j$$

where the coefficients a_{ij} and b_i are defined by the relations (11). Let $p_t(x)$ be the heat kernel associated to L_μ on N (See §1.4). Then there exists a constant $c > 0$ such that for all $x \in N$

$$|\phi^{*n}(x) - p_n(x)| \leq \frac{c}{n^{(d(N)+1)/2}} \exp\left(-\frac{|x|^2}{cn}\right) \tag{14}$$

Moreover, there exists a constant $c(\phi) > 0$ such that, uniformly when x remains in a given compact of N

$$\lim_{n\to+\infty} n^{d(N)/2}\phi^{*n}(x) = c(\phi) \tag{15}$$

Alexopoulos also obtains a similar estimate in the general case of connected Lie groups of polynomial growth, but in this case, the definition of L_μ is more delicate.

When μ does not admit a density with respect to the Haar measure, the analogous statements, (15) or (14), that is the local limit theorem and its uniform version, remain open problems. In [21] however, we show that these results holds for the Heisenberg group. The proof remains valid for all step-2 nilpotent groups. Applying the same method to the general case seems quite intricate. We have:

Theorem 3.15. *Let N be the Heisenberg group of upper triangular 3 by 3 matrices. Let μ be an aperiodic, centered probability measure with compact support on N. Let $(\nu_t)_t$ be the gaussian semi-group associated to μ and determined by its infinitesimal generator L_μ defined in (10). Then for every continuous function f of compact support on N, we have*

$$\lim_{n\to+\infty} n^2 \int f(x)d\mu^n(x) = c(\mu)\int_N f(x)dx$$

where $c(\mu) > 0$ is the value at e of the density p_1 corresponding to ν_1 (i.e. the heat kernel for the sub-laplacian L_μ). Moreover for every bounded Borel subset B of N whose boundary is negligible with respect to Lebesgue measure (i.e. $|\partial B| = 0$) we have

$$\lim_{n\to+\infty} n^2 \sup_{x\in N} |\mu^n(xB) - \nu_n(xB)| = 0 \tag{16}$$

In this theorem several choices for the gaussian semi-group $(\nu_t)_t$ are possible. The result remains of course true for all choices. Note that N is a graded nilpotent group and a choice of a $(\nu_t)_t$ corresponds to a choice of a semi-group of dilations $(\delta_t)_t$ or equivalently to a choice of a vector subspace of the Lie algebra which is in direct sum with the center of $\mathcal{N}$. The semi-group (ν_t) is stable with respect to the corresponding semi-group of dilations (δ_t), i.e. $\nu_t = \delta_{\sqrt{t}}(\nu_1)$. If X, Y, Z is a basis of $\mathcal{N}$ such that $[X,Y] = Z$, then the group of automorphisms of $\mathcal{N}$ is

$$\left\{ \begin{pmatrix} A & 0 \\ C & b \end{pmatrix},\ A \in GL_2(\mathbb{R}),\ \det A = b \right\}$$

Let δ_t be the automorphism defined by $A = tId$, $b = t^2$, $C = 0$ and let (ν_t) be the corresponding gaussian semi-group. Then the other possible semi-groups of dilations are all of the form $\phi \circ \delta_t \circ \phi^{-1}$ for some automorphism ϕ such that $A = Id$ and $b = 1$. The gaussian semi-group corresponding to $(\phi \circ \delta_t \circ \phi^{-1})_t$ is $(\phi(\nu_t))_t$. The corresponding densities have the same value at the identity e and hence $c(\mu)$ is indeed independent of the choice of the gaussian semi-group $(\nu_t)_t$.

Note that in Theorem 3.14, the estimate (14) shows that we have a speed of convergence in at least $\frac{1}{\sqrt{n}}$ in the local limit theorem (15) (because $n^{d/2}p_n(x)$ converges also with a speed of order $1/\sqrt{n}$, by [3] 1.9.2). As we already pointed out in the commutative case (i.e. $N = \mathbb{R}^d$), such a speed depends strongly on the regularity of the measure μ and is related to the diophantine properties of μ (see §4.2). In the case when $N = \mathbb{R}^d$ and μ has a density like in the statement of 3.14, we have in fact a speed of convergence in $1/n$ as follows for example from [36] ch. XVI, Theorem 1. If however the support of μ is finite and very well approximable by elements from a discrete subgroup of N, then we cannot hope for any good speed of convergence in (16).

In [20], we obtain coarser estimates than those of Theorem 3.15 but which are valid for an arbitrary nilpotent Lie group. The method is very different and yields the following:

Theorem 3.16. *Let N be a simply connected nilpotent Lie group and μ non-degenerate finitely supported symmetric probability measure on N. Then there exists a constant $C > 0$ such that for every Borel set B with positive Haar measure and negligible boundary, we have*

$$\frac{1}{C}\frac{|B|}{n^{d(N)/2}} \leq \mu^n(B) \leq C\frac{|B|}{n^{d(N)/2}}$$

as soon as $n \geq n_1 = n_1(B)$. Moreover the upper bound in this estimate is uniform when B is changed into a translate xB, $x \in N$.

The method used here takes advantage of the known results for random walks on discrete groups, in particular Theorem 3.9 due to Hebisch and Saloff-Coste in the case of a symmetric measure, and combines them with a generalization of Weyl's equidistribution theorem for finitely generated dense subgroups of nilpotent Lie groups.
Other limit theorems are known on nilpotent Lie groups (one can have a look at [84]), most notably a large deviation principle due to Baldi and Caramelino [9]. We will not need them in the sequel.

4. Equidistribution and refinements of the local limit theorem

4.1. **Equidistribution along an orbit.** Let μ be an aperiodic and centered probability measure on $\mathbb{Z}$. Let D be a subset of $\mathbb{Z}$ admitting a density $d(D)$, i.e. the following limit exists

$$\lim_{n\to+\infty} \frac{\#\{D \cap [-n, n]\}}{2n+1} = d(D) \geq 0$$

A natural question arises: do we have

$$\lim_{n\to+\infty} \mu^n(D) = d(D) \ ? \tag{17}$$

As we will show below, the answer is yes as soon as the measure μ admits a finite moment σ^2 of order 2. This type of question arises when trying to prove the equidistribution of a random walk. Assume that (X, T, m) is an ergodic invertible and measure preserving dynamical system on a topological space X and assume further that the full orbit $\{T^n x\}_{n\in\mathbb{Z}}$

of the point x is equidistributed, that is

$$\lim_{N\to+\infty}\frac{1}{2N+1}\sum_{n=-N}^{N}f(T^nx)=\int_X fdm$$

for every continuous function f of compact support on X (from Birkhoff's ergodic theorem this holds for m-almost every $x\in X$). Then the random walk associated to μ starting at x and walking along the orbit is also equidistributed, that is

$$\lim_{n\to+\infty}\int f(T^kx)d\mu^n(k)=\int_X fdm \tag{18}$$

Indeed, it is enough to show this convergence for indicator functions $f=\mathbf{1}_B$ where B is, say, an open subset of X such that $m(\partial X)=0$. Then we take $D=\{k\in\mathbb{Z},T^kx\in B\}$ and we reduced the problem to the question above. It is worth noting that the convergence (18) holds for *the same* starting point x. If we seek (18) only for m-almost all points, then the conclusion is easier and follows from a random ergodic theorem (for instance [85]).

The proof of (17) is simple and requires the uniform version of the local limit theorem on $\mathbb{Z}$, that is

$$\lim_{n\to+\infty}\sqrt{n}\sup_{x\in\mathbb{Z}}\left|\mu^n(x)-p(x/\sqrt{n})\sqrt{n}\right|=0 \tag{19}$$

where p is the centered gaussian with variance σ^2. This theorem is classical for $\mathbb{Z}$ and proved for example in [40] by means of Fourier analysis. If μ has finite support, this is a special case of Theorem 3.10 (ii). From the central limit theorem, we also have : for every $\varepsilon>0$ there exists $C>0$ such that

$$\sum_{|x|\geq C\sqrt{n}}\mu^n(x)\leq\varepsilon \tag{20}$$

Combining (19) and (20) we see that it is enough to show the convergence for $p(x/\sqrt{n})/\sqrt{n}$, that is

$$\lim_{n\to+\infty}\sum_{x\in D}p(x/\sqrt{n})/\sqrt{n}=d(D)$$

Now approximating p by a piecewise constant function, we reduce to prove that for every set $I=\{x\in\mathbb{R},|x|\in[a,b]\}$, $b>a\geq 0$ we have

$$\lim_{n\to+\infty}\frac{1}{\sqrt{n}}\#\{I\sqrt{n}\cap D\}=d(D)(b-a)$$

But this follows immediately from the assumption on D that it has a density $d(D)$.

In a similar way, we can show that if μ is no longer assumed centered then there still is a Cesaro type of convergence.More precisely, if the average of μ is >0 and if

$$\lim_{n\to+\infty}\frac{\#\{D\cap[1,n]\}}{n}=d(D)\geq 0$$

then

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{k=0}^{n-1}\mu^k(D)=d(D)$$

All the above discussion was made on $\mathbb{Z}$, but easily extends to $\mathbb{R}$ and $\mathbb{R}^d$. The main tool, the uniform local limit theorem (19), admits the following form on $\mathbb{R}$. Its proof, rather delicate, is due to Stone [106].

Theorem 4.1. *(Stone) Let μ be an aperiodic and centered probability measure on $\mathbb{R}$ admitting a finite moment σ^2 of order 2 and let ν be the centered gaussian law with variance σ^2. Then there exists a decreasing sequence ε_n tending to zero and depending only on μ such that for every closed interval I of $\mathbb{R}$ we have*

$$\lim_{n\to+\infty} \sqrt{n} \sup_{x\in\mathbb{R}} |\mu^n(I+x) - \nu^n(I+x)| \leq \varepsilon_n(1+|I|)$$

where $|I|$ is the Lebesgue measure of I.

As above, the theorem has an analogous corollary. The details of the proof are given in [22].

Theorem 4.2. *Let μ be an aperiodic and centered probability measure on $\mathbb{R}$ admitting a finite moment of order 2 a Suppose that f is a uniformly continuous function which is bounded on $\mathbb{R}$ and such that the following limit exists*

$$\lim \frac{1}{T}\int_0^T f(t)dt = \ell \tag{21}$$

as $|T| \to +\infty$. Then

$$\lim_{n\to+\infty} \int f d\mu^n = \ell$$

Remark 4.1. *Theorem 4.2 is stated for centered random walks and the same conclusion is evidently wrong for non-centered random walks. However, in this case, we still have a convergence in the sense of Cesaro, as was mentioned above for random walks on $\mathbb{Z}$. More precisely, if μ is aperiodic and non-centered and if f satisfies the assumptions of the theorem then*

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} \int f d\mu^k = \ell$$

The proof of this assertion is in all respects similar to that of the theorem.

Analogously, Theorem 4.2 allows to get probabilistic equidistribution as soon as there is deterministic equidistribution. More precisely, let X be a locally compact space with a Borel measure m and $(\phi_t)_t$ be a flow acting continuously on X and by measure preserving transformations. The following corollary shows that if the orbit of a point $x \in X$ is equidistributed by the flow, then any centered random walk along this orbit is also equidistributed. We fix a probability measure μ, which we assume centered and aperiodic with a finite moment of order 2, and we let S_n be the associated random walk.

Corollary 4.1. *Assume that the orbit $(\phi_t)_{t\in\mathbb{R}} \cdot x$ is equidistributed with respect to m, that is*

$$\lim_{|T|\to+\infty} \frac{1}{T}\int_0^T f(\phi_t \cdot x)dt = \int_X f(y)dm(y)$$

for every continuous function with compact support f on X. Then we also have

$$\lim_{n\to+\infty} \mathbb{E}(f(\phi_{S_n} \cdot x)) = \int_X f(y)dm(y)$$

The proof follows immediately from Theorem 4.2 applied to the function $t \mapsto f(\phi_t \cdot x)$ which is bounded and uniformly continuous on $\mathbb{R}$ because f has compact support. In this statement, the fact that μ is centered is essential. If μ is not centered, we still have a convergence in the sense of Cesaro.

4.2. **Local limit theorem and speed of convergence.** In [22] we study the problem of the speed of convergence in the local limit theorem on $\mathbb{R}$ (i.e. Theorem 4.1). This question is related to diophantine properties of the probability measure μ. It is convenient to introduce the following definition. We say that μ is diophantine if it cannot be well approximated by a measure supported on an arithmetic progression of $\mathbb{R}$. More precisely,

Definition 4.1. *Let $l \geq 0$. A probability measure μ on $\mathbb{R}$ is said to be **l-diophantine** if one of the following equivalent assertions is satisfied:*

(i) $\exists C > 0$ such that if $|x|$ is large enough

$$\inf_{y \in \mathbb{R}} \int \{xa + y\}^2 d\mu(a) \geq \frac{C}{|x|^l}$$

where $\{t\}$ denotes the fractional part of t.

(ii) $\exists C > 0$ such that if $|x|$ is large enough

$$|\widehat{\mu}(x)| \leq 1 - \frac{C}{|x|^l}$$

where $\widehat{\mu}$ is the characteristic function of μ.

We then obtain the following results about the speed of convergence in the local limit theorem (see [22])

Theorem 4.3. *Let μ be a centered probability measure on $\mathbb{R}$ admitting a finite moment of order 4.*

(i) Assume that μ is l-diophantine, and let $k > 3l/2 + 1$. Then for every C^k function f with compact support there exists a constant $C = C(f) > 0$ such that

$$\left| \int f d\mu^n - \frac{1}{\sqrt{2\pi\sigma^2 n}} \int f(x)dx \right| \leq \frac{C(f)}{n\sqrt{n}}$$

(ii) If μ is symmetric of the form $\mu = \nu * \nu^{-1}$ and if μ is not l-diophantine for any $l \geq 0$ then for any $\varepsilon > 0$ and any integer $p > 0$ there exists a compactly supported C^p function f such that

$$\overline{\lim_{n \to +\infty}} n^{\frac{1}{2}+\varepsilon} \left| \int f d\mu^n - \frac{1}{\sqrt{2\pi\sigma^2 n}} \int f(x)dx \right| = +\infty$$

(iii) The measure μ is l-diophantine for some $l \geq 0$ if and only if there exists $k \geq 0$ such that for every compactly supported C^k function on $\mathbb{R}$ we have

$$\sup_{t \in \mathbb{R}} \left| \int f(t + \cdot)d\mu^n - \int f(t + \cdot)d\nu^n \right| = O(\frac{1}{n})$$

where ν is a gaussian law is the same variance as μ.

4.3. **Local limit theorems and large deviations.** Another phenomenon related to the two last paragraphs arises when trying to determine the asymptotic behavior of the expression

$$\int f d\mu^n$$

when f is a continuous and bounded function on $\mathbb{R}$ which is not assumed to have a limit at infinity. A preliminary result in this direction is Theorem 4.2 quoted above. But what happens if f no longer satisfies the assumption (21) ? Then it is natural to try to compare the behavior of $\int f d\mu^n$ to that of $\int f d\nu^n$ where ν is the gaussian law associated to μ. Indeed the quantity $\int f d\nu^n$ is easier to understand since ν is known explicitly. The next theorem shows that, under certain regularity hypothesis, the asymptotic behavior of $\int f d\nu^n$ is the same for all probability measures μ with the same variance. More precisely,

Theorem 4.4. *Let μ be a centered probability measure on $\mathbb{R}$ admitting a finite moment of order 4 and let ν be the associated gaussian law. We assume that μ is l-diophantine. The for every $k > 3l/2 + 1$ we have*

$$\lim_{n\to+\infty} \frac{\int f d\mu^n}{\int f d\nu^n} = 1 \tag{22}$$

for every non-zero bounded C^k function $f \geq 0$ on $\mathbb{R}$ whose derivatives up to order k are bounded.

In [22], we give for each $k \geq 0$ an example of a C^k function f_k whose derivatives up to order k tend to zero at infinity and a diophantine measure μ_k which is centered and finitely supported for which the limit (22) above does not hold.

The proof of Theorem 4.2 required a uniform control of the quantity $\mu^n(x + I)$ on a large interval of values of x (here I is a fixed closed interval). Such information is given by Stone's uniform local limit theorem (Theorem 4.1). Let us remark however that this theorem gives a significant information only when x is not too large, namely when $|x| \leq C\sqrt{n}$ where C is a positive constant. As an immediate consequence of Theorem 4.1, we have

$$\lim_{n\to+\infty} \frac{\mu^n(I+x)}{\nu^n(I+x)} = 1$$

uniformly in x when $|x|/\sqrt{n}$ stays bounded. If we try to estimate $\mu^n(I + x)$ for larger values of x, then we face a large deviations type of question and an additional moment assumption on μ is necessary. For moderate deviations, we obtain the following local limit theorem:

Theorem 4.5. *Let μ be a centered and aperiodic probability measure on $\mathbb{R}$ with variance σ^2 and admitting a finite moment of order $r > 2$ and let ν be the associated centered gaussian law (i.e. with variance σ^2). Let $I = [a, b]$ be a closed interval in $\mathbb{R}$ and c a real number $\in]0, r - 2[$. Then we have*

$$\lim_{n\to+\infty} \frac{\mu^n(I+x)}{\nu^n(I+x)} = 1$$

uniformly when $|x| \leq \sqrt{c\sigma^2 n \log n}$.

This theorem extends Stone's local limit theorem to moderate deviations. Let me remark that in the particular case when μ has a density with respect to the Lebesgue measure on $\mathbb{R}$ then the analogous result is due to Amosova (see [4]).

5. Unipotent random walks and Ratner's theorem

5.1. **Equidistribution of unipotent orbits in G/Γ.** In this paragraph, we briefly present Ratner's theorem and certain related results. Let G be a connected Lie group and Γ a discrete subgroup of finite co-volume in G. We say that an element $g \in G$ is *unipotent* if the automorphism $Ad(g)$ of the Lie algebra of G is unipotent, that is all its eigenvalues are equal to 1. A subgroup U of G is said to be unipotent if all its elements are unipotent. Quite surprisingly the behavior of the orbits of points in G/Γ under the action of a unipotent subgroup, or more generally a subgroup generated by unipotent elements is quite tame. This situation contrasts greatly with the case of diagonal actions for example.

This behavior has been conjectured by Raghunathan, Dani and Margulis during the seventies and eighties and subsequently partially proved by many authors (Dani, Margulis, Shah, etc.) before M. Ratner completed the proof in full generality in the early nineties (cf. [105], [88]).

Theorem 5.1. *(Ratner) Let H be a closed connected subgroup of G which is generated by unipotent elements. Let $x \in G/\Gamma$. Then there exists a closed subgroup $F \supset H$ of G such that $\overline{H \cdot x} = F \cdot x$ and such that the closed orbit $F \cdot x$ admits an F-invariant probability measure denoted by m_x. Moreover the measure m_x is H-ergodic. Finally every H-ergodic probability measure on G/Γ is of this form.*

An essential ingredient of the proof is the following recurrence theorem due to Dani.

Theorem 5.2. *(Dani) Let $U = (u_t)_t$ be a one-parameter unipotent subgroup of G. Fix $x \in G/\Gamma$ and $\varepsilon > 0$. Then there exists a compact subset K in G/Γ such that*

$$\overline{\lim_{T \to +\infty}} \frac{1}{T} \left| \{ t \in [0, T], u_t \cdot x \notin K \} \right| \leq \varepsilon$$

This theorem implies in particular the following non-trivial fact: no unipotent orbit escapes to infinity in G/Γ. This phenomenon was first studied by Margulis in [75] for the case of $SL_n(\mathbb{R})/SL_n(\mathbb{Z})$ and was also a crucial step in the Margulis' first proof of the arithmeticity of non co-compact lattices in higher rank. The proof of Theorem 5.2 relies on the polynomial nature of unipotent flows and uses the ideas introduced in [75].

In order to prove Theorem 5.1, Ratner generalized Dani's theorem and showed that the orbits of one-parameter unipotent subgroups are in fact equidistributed in their closure.

Theorem 5.3. *(Ratner) If $U = \{u_t, t \in \mathbb{R}\}$ is a one-parameter unipotent subgroup of G, the for all $x \in G/\Gamma$, there exists a closed subgroup F of G such that the orbit $F \cdot x$ is closed and bears an F-invariant probability measure m_x such that*

$$\lim_{T \to +\infty} \frac{1}{T} \int_0^T f(u_t \cdot x) dt = \int_{G/\Gamma} f dm_x$$

for all bounded and continuous function f on G/Γ.

Later, answering a question of Ratner, N. Shah has generalized this result to the case of an arbitrary simply connected unipotent subgroup (see [99]). Let U be a simply connected unipotent subgroup of G and $X_1, ..., X_n$ a triangular basis for the Lie algebra $Lie(U)$ (i.e. for every $k \in [1, n]$ the vectors $X_1, ..., X_k$ generate an ideal in $Lie(U)$. We introduce the following subsets of U (rectangles, see above §3.2),

$$S(s_1, ..., s_n) = \left\{ \prod \exp(t_i X_i) \in U, 0 \leq t_i \leq s_i \right\}$$

Theorem 5.4. *(Shah) Let U be a simply connected unipotent subgroup of G and $x \in G/\Gamma$. Then we have*

$$\lim \frac{1}{\lambda(S(s_1, ..., s_n))} \int_U f(u \cdot x) \lambda(du) = \int_{G/\Gamma} f dm_x$$

where λ denote a Haar measure on U and f is a bounded continuous function on G/Γ and all s_i's tend to $+\infty$.

In [100], Shah generalizes Ratner's theorems for the action of a subgroup H such that $Ad(H)$ is contained in the Zariski closure of the subgroup generated by the unipotent elements of $Ad(H)$. In particular H is no longer assumed connected. This allows to treat the case of discrete subgroups generated by unipotent elements, for example non co-compact lattices in semisimple Lie groups. He obtains

Theorem 5.5. *(Shah) For all $x \in G/\Gamma$ there exists a closed subgroup $F \supset H$ of G such that $\overline{H \cdot x} = F \cdot x$ and such that the closed orbit $F \cdot x$ bears an F-invariant locally finite measure denoted by m_x. Moreover the measure m_x is H-ergodic. Finally every H-ergodic locally finite measure on G/Γ is of this form.*

We observe that the measure m_x here is locally finite (i.e. finite on compact subsets). Shah conjectures in [100] that the m_x are in fact finite and show that this conjecture reduces to the case when G is semisimple of higher rank and Γ is an irreducible lattice. Typically the fact that m_x is finite implies that if H is discrete then every discrete orbit of H is in fact finite. In the same article he conjectures that locally finite H-ergodic measures are finite. Recently, Eskin and Margulis have answered this question positively and they show in [34] that if H is Zariski dense in G semisimple, then any locally finite H-invariant measure on G/Γ is finite. Their proof is quite remarkable and consists in showing a recurrence property of random walks living on an orbit of H in G/Γ. We describe their result in the next paragraph.

5.2. **A probabilistic version of Ratner's equidistribution theorem.** Let G be a connected Lie group and G a discrete subgroup of finite co-volume in G. We keep the notations of the last paragraph. In particular $m_{\mathbf{x}}$ denotes the invariant measure associated to $\mathbf{x}$ in Ratner's theorem (Theorem 5.1).

Eskin and Margulis obtain in [34] the following theorem:

Theorem 5.6. *(Eskin-Margulis) Assume G is semisimple with finite center and Γ is an irreducible lattice in G. Let μ be a probability measure on G whose support is Zariski-dense in G. Then for every $\varepsilon > 0$ there exists a compact subset K of G/Γ such that for every $\mathbf{x} \in G/\Gamma$ we have*

$$\mu^n * \delta_{\mathbf{x}}(K) \geq 1 - \varepsilon$$

as soon as $n \geq n_0 = n_0(\mathbf{x})$. Moreover $\mathbf{x} \mapsto n_0(\mathbf{x})$ is locally bounded.

In particular the sequence of measures $(\mu^n * \delta_{\mathbf{x}})_n$ is relatively compact in the space of probability measures on G/Γ. The proof uses properties of the convolution operator given by μ on functions on G and most notably the positivity of the first Lyapunov exponent due to Furstenberg (see [52] for a quick proof of this theorem) and also makes use of the techniques introduced in [35] to show the quantitative version of the Oppenheim conjecture (see the survey [7]). As an immediate corollary, they obtain that every locally finite μ-stationary measure on G/Γ is in fact finite, thus answering N. Shah's question quoted above.

In this theorem, the support of μ is Zariski dense in G and G is assumed semisimple. One can ask what if the support of μ is contained in, say, a unipotent subgroup of G. This case is in a sense simpler because we already know much about unipotents orbits thanks to Ratner's theorem, in particular the orbits are equidistributed (theorems 5.3 et 5.4). Below, in [20] and in [21], we study this problem and show that centered random walks along a unipotent orbit are equidistributed. It is possible to deduce the equidistribution of the random walk from the equidistribution of the orbit by making use of a refined version of the local limit theorem on the corresponding unipotent group (see corollary 4.1).

When such a theorem is available, we immediately obtain as a corollary the equidistribution of the corresponding random walk on the homogeneous space G/Γ. For one-parameter unipotent subgroups, or more generally for commutative unipotent subgroups, we have:

Theorem 5.7. *Let $U = \{u_t, t \in \mathbb{R}\}$ be a one-parameter unipotent subgroup of G and μ be an aperiodic and centered probability measure on U which admits a finite moment of order 2. If $\mathbf{x} \in G/\Gamma$, then*

$$\lim_{n \to +\infty} \mu^n * \delta_{\mathbf{x}} = m_{\mathbf{x}} \tag{23}$$

(weak convergence of probability measures).

This theorem is a direct consequence of the combination of Theorem 5.3 and Corollary 4.1, which we have deduced from Stone's local limit theorem in the last section.

When U is not commutative, the local limit theorem is still an open problem. In [21], we show the local limit theorem and its uniform version (i.e. the analogous statement to Stone's Theorem 4.1) for the Heisenberg group (see Theorem 3.15 below) and we derive, as a corollary, the equidistribution of centered random walks on Heisenberg-unipotent orbits in G/Γ. This is the case for instance for centered random walks on horospheres of complex hyperbolic manifolds of finite volume.

Note that if the measure μ admits a continuous density of compact support with respect to the Haar measure, then Alexopoulos' theorem yields in the same way the convergence (23).

In [20], we show a coarse local limit theorem for finitely supported symmetric measures (see Theorem 3.16 above) on an arbitrary simply connected nilpotent Lie group. Combining this result with Guivarc'h's Theorem 3.3, we succeed in proving the equidistribution of the corresponding random walk on G/Γ. We obtain:

Theorem 5.8. *Let U be a simply connected unipotent subgroup of G and μ a finitely supported symmetric and aperiodic probability measure on U, then for all $\mathbf{x} \in G/\Gamma$*

$$\lim_{n \to +\infty} \mu^n * \delta_{\mathbf{x}} = m_{\mathbf{x}}$$

5.3. **Cesaro convergence.** In this paragraph, we show that there always is convergence in the sense of Cesaro even if the random walk is not centered.

Proposition 5.1. *Let U be a one-parameter unipotent subgroup of G and let μ an aperiodic probability measure on U with a finite moment of order 2. Then for all $\mathbf{x} \in G/\Gamma$, we have the following convergence of probability measures*

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu^k * \delta_{\mathbf{x}} = m_{\mathbf{x}} \tag{24}$$

In order to show this proposition we could do as in the centered case and use Remark 4.1. However we will use a slightly different method which has the advantage of not demanding such detailed information about the random walk as is provided by the uniform version of the local limit theorem. In [20], we apply this idea to the case of nilpotent groups.

We are going to show that the sequence of Cesaro averages is relatively compact in the space of probability measures on G/Γ and then that any limit measure is absolutely continuous with respect to $m_{\mathbf{x}}$. This will follow from the lemma below. However, every limit measure of the sequence of Cesaro average is μ-stationary. According to the Choquet-Deny theorem [26], every μ-stationary measure is necessarily invariant under the whole of U. Since $m_{\mathbf{x}}$ is U-ergodic, it follows that every limit measure is equal to $m_{\mathbf{x}}$, thus establishing the convergence. We have to show the following result:

Lemma 5.1. *Let μ be an aperiodic probability measure on $\mathbb{R}$ with finite variance $\sigma^2 < +\infty$ and mean $d = 1$. We denote by $S_n + n$ the associated random walk (S_n is the corresponding* *centered random walk). Then for all* $\varepsilon > 0$ *there exists* $\delta > 0$ *such that if* f *is a bounded function* $0 \leq f \leq 1$ *which is uniformly continuous on* $\mathbb{R}$, *and such that*

$$\overline{\lim_{T\to+\infty}} \frac{1}{T}\int_0^T f(t)dt \leq \delta \tag{25}$$

then

$$\overline{\lim_{N\to+\infty}} \frac{1}{N}\sum_{n=0}^{N-1} \mathbb{E}(f(S_n+n)) \leq \varepsilon \tag{26}$$

Let ϕ be a continuous function on G/Γ taking its values in $[0,1]$ and such that $\phi \equiv 0$ on a compact subset K and $\phi \equiv 1$ outside a bigger compact subset. Taking $f(t) = \phi(u_t \cdot \mathbf{x})$ in the above lemma, we obtain that the sequence of measures $(\frac{1}{n}\sum_{k=0}^{n-1}\mu^k * \delta_{\mathbf{x}})_n$ is relatively compact. Note that (25) holds if K is large enough according to Dani's recurrence theorem.

Now if ϕ is continuous and compactly supported on G/Γ and satisfies $\int \phi(y)dm_{\mathbf{x}}(y) \leq \delta$ then $f(t) = \phi(u_t \cdot \mathbf{x})$ satisfies (25) according to Ratner's equidistribution theorem. Hence, applying the lemma, for every limit measure of the sequence of Cesaro averages, say ν, we have $\int \phi(y)d\nu(y) \leq \varepsilon$. This shows that ν is absolutely continuous with respect to $m_{\mathbf{x}}$.

Proof of the Lemma: First note that it is enough to show that

$$\overline{\lim_{N\to+\infty}} \frac{1}{N}\sum_{N/2\leq n\leq N} \mathbb{E}(f(S_n+n)) \leq \delta \tag{27}$$

instead of (26) as can been by splitting the sum in (26) into summands going from $N/2^{i+1}$ to $N/2^i$. Moreover, since f is uniformly continuous, it is enough to show (27) for a steps functions g of the form

$$g(t) = \sum_{i\in\mathbb{Z}} f_i \mathbf{1}_{I_i}$$

where I_i is the interval $[ih,(i+1)h]$ and $h > 0$ is fixed and $f_i \in [0,1]$ (to obtain (27) for a general f we will take $h < \omega(\delta/2)$ where ω is a modulus of uniform continuity for f). The only probabilistic information used here is the following upper bound, which is a consequence of Stone's local limit theorem: there is a constant $C > 0$ such that for every closed interval I in $\mathbb{R}$ we have, for n larger than some integer n_0 depending only on I

$$\sup_{x\in\mathbb{R}} \mathbb{P}(S_n \in I + x) \leq \frac{C\cdot|I|}{\sqrt{n}} \tag{28}$$

This being given, we fix another large constant $D > 0$ and we write

$$\frac{1}{N}\sum_{N/2\leq n\leq N} \mathbb{E}(g(S_n+n)) \leq \frac{1}{N}\sum_{N/2\leq n\leq N} \mathbb{E}(g(S_n+n)1_{|S_n|\leq D\sqrt{N}}) + \max_{N/2\leq n\leq N} \mathbb{P}(|S_n| \geq D\sqrt{N})$$

We can choose D large enough so that the remainder in the term on the right hand side be smaller than $\varepsilon/2$ as soon as N is large enough. For the first term, we write, for large enough N, using (28):

$$\begin{aligned}
\frac{1}{N}\sum_{N/2\leq n\leq N}\mathbb{E}(g(S_n+n)1_{|S_n|\leq D\sqrt{N}}) &\leq \frac{1}{N}\sum_{N/2\leq n\leq N}\sum_{|ih-n|\leq D\sqrt{N}} f_i\mathbb{P}(S_n\in I_i-n)\\
&\leq \frac{1}{N}\sum_{N/2\leq n\leq N}\frac{C}{\sqrt{N/2}}\int_{n-D\sqrt{N}}^{n+D\sqrt{N}} g(t)dt\\
&\leq \frac{1}{N}\sum_{N/2\leq n\leq N}\frac{C}{\sqrt{N/2}}\int_{-D\sqrt{N}}^{D\sqrt{N}} g(t+n)dt\\
&\leq \frac{1}{N}\sum_{N/2\leq n\leq N}\frac{C}{\sqrt{N/2}}\sum_{|k|\leq D\sqrt{N}}\int_0^1 g(t+n+k)dt\\
&\leq \frac{C}{\sqrt{N/2}}\sum_{|k|\leq D\sqrt{N}}\frac{1}{N}\int_{N/2+k}^{N+k} g(t)dt
\end{aligned}$$

but thanks to the assumption on g, that is (25), each term in the above sum is say $\leq 2\delta$ for N large enough. Hence we get

$$\overline{\lim_{N\to+\infty}}\frac{1}{N}\sum_{N/2\leq n\leq N}\mathbb{E}(g(S_n+n)1_{|S_n|\leq D\sqrt{N}})\leq 8CD\delta$$

It suffices to choose δ so that $8CD\delta\leq\varepsilon/2$ and this ends the proof..□

Remark 5.1. *Note that Proposition 5.1 is not a corollary of Kakutani's random ergodic theorem. According to this theorem, for an arbitrary probability measure m on G/Γ, if m is U-ergodic (and here U could be any subgroup of G) then the convergence in 24 holds for m-almost all starting point $\mathbf{x}$ in G/Γ (under the assumption that μ is non-degenerate in U). Here on the other hand, we try to capture the behavior of the random walk for every orbit of U on which it may live.*

5.4. **A counter-example in the non-centered case.** We give here an example that shows that if the flow U is not uniquely ergodic, then one cannot remove the assumption that μ be centered in Theorem 5.7.

Proposition 5.1. *Let $G = SL_2(\mathbb{R})$ and $\Gamma = SL_2(Z)$ and let U be a one-parameter unipotent subgroup of G. Let μ be a non-centered probability measure on U with variance $\sigma^2 < +\infty$ and mean $d\neq 0$. Then for every compact subset K of G/Γ and for almost every point $\mathbf{x}\in G/\Gamma$, we have*

$$\lim\inf_{n\to+\infty}\mu^n*\delta_{\mathbf{x}}(K)=0 \tag{29}$$

We will say that a real number θ is *well approximable from both sides* if for every $\varepsilon>0$ and $\sigma\in\{-1,1\}$ one can find two integers x and y in $\mathbb{Z}^2\backslash\{(0,0)\}$ such that

$$\begin{aligned}
|x(y-\theta x)| &< \varepsilon\\
\sigma x(y-\theta x) &> 0
\end{aligned}$$

It is easy to check that θ is well approximable from both sides if and only if its continuous fraction expansion $[a_0,a_1,...,a_n,...]$ is such that both subsequences $(a_{2n})_n$ and $(a_{2n+1})_n$ are unbounded. Moreover almost every real number θ (for the Lebesgue measure) is well approximable from both sides.

We identify $SL_2(\mathbb{R})/SL_2(\mathbb{Z})$ with the space of unimodular lattices in $\mathbb{R}^2$ and we set as usual $SL_2(\mathbb{R}) = G$ and $SL_2(\mathbb{Z}) = \Gamma$. Let $\|\cdot\|$ be the canonical Euclidean norm on $\mathbb{R}^2$. Recall that according to *Mahler's criterion*, a subset $K \subset G/\Gamma$ is relatively compact if and only if there exists $\delta > 0$ such that $\|v\| > \delta$ for all lattices $\mathbf{x} \in K$ and all non-zero vectors $v \in \mathbf{x}$. We let also $\mathbf{x}_0 = \mathbb{Z}^2$.

Let $(u_t)_t$ be a unipotent subgroup of $SL_2(\mathbb{R})$ and D a line in $\mathbb{R}^2$ which is invariant under $(u_t)_t$. Let $\theta \in \mathbb{R}$ be the slope (which we assume finite) of this line. Then in the canonical basis of $\mathbb{R}^2$ the action of $(u_t)_t$ can be written as

$$u_t \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + \alpha t(y - \theta x) \\ y + \alpha\theta t(y - \theta x) \end{pmatrix}$$

for some $\alpha \in \mathbb{R}\backslash\{0\}$. We fix an arbitrary compact subset K of G/Γ and we fix a number $\delta > 0$ which correspond to it by Mahler's criterion (one can choose δ small enough so that $2|\theta|\delta^3 < 1$). We denote by Ω the set of all real numbers t such that $u_t \cdot \mathbf{x}_0 \notin K$. For $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{Z}^2\backslash\{(0,0)\}$ we set $A_{x,y} = \left\{ t \in \mathbb{R}, \left| u_t \begin{pmatrix} x \\ y \end{pmatrix} \right| < \delta \right\}$. Then

$$\Omega = \{t, u_t \cdot \mathbf{x}_0 \notin K\} = \bigcup_{x,y \in \mathbb{Z}^2\backslash\{(0,0)\}} A_{x,y}$$

Now we assume that $\theta \notin \mathbb{Q}$ and that θ is well approximable from both sides. Let us fix $\kappa = |2\alpha| \max\{|\theta|, 1\}$ and denote $t_{x,y} = \frac{x}{\alpha(\theta x - y)}$ and $I_{x,y} = \{t, |t| < \frac{\delta}{\kappa|\theta x - y|}\}$. We can then find arbitrarily large integers x and y such that

$$|x(y - \theta x)| < \delta^4 \tag{30}$$

and such that $t_{x,y}$ has whichever sign we desire. We check that, since δ was chosen small enough

$$t_{x,y} + I_{x,y} \subset A_{x,y}$$

Now let $S_n = X_1 + ... + X_n$ be a centered random walk with finite variance $\sigma^2 > 0$ and $d \in \mathbb{R}\backslash\{0\}$. Then $S_n + nd$ is a non-centered random walk with a drift term equal to d, so S_n is distributed according to μ^n. Since x can be taken as large as we want, we can assume that $|I_{x,y}| > 2|d|$. Hence we can find a positive integer n such that

$$nd \in t_{x,y} + \frac{1}{2} I_{x,y} \tag{31}$$

Then, if S_n is not too large, i.e. if $|S_n| < \frac{1}{4}|I_{x,y}|$, we will have $S_n + nd \in \Omega$. But if n satisfies (31) then

$$n \leq \frac{2|x|}{|\alpha||\theta x - y|}$$

Thanks to (30), it follows that if $|S_n| < C_\delta \sqrt{n}$ where $C_\delta = \frac{1}{\delta}\sqrt{|\alpha d|/2\kappa^2}$ then $|S_n| < \frac{1}{4}|I_{x,y}|$.

So we have found an arbitrarily large positive integer n (because $t_{x,y}$ itself is arbitrarily large) such that

$$S_n + nd \in \Omega$$

as soon as $|S_n| < C_\delta \sqrt{n}$. But since δ is as small as we want, C_δ is arbitrarily large and for all $\varepsilon > 0$ we have

$$\mathbb{P}(|S_n| < C_\delta \sqrt{n}) \geq 1 - \varepsilon$$

as it follows from the classical central limit theorem. Hence for every compact subset $K \subset G/\Gamma$,

$$\limsup \mathbb{P}(u_{S_n + nd} \cdot \mathbf{x}_0 \notin K) = 1 \tag{32}$$

We have established (29) for $\mathbf{x} = \mathbf{x}_0$ and whenever the slope θ of the invariant line fixed by U satisfies the diophantine condition stated at the beginning of the paragraph. Let us come back to the situation of the proposition. Let E be the set of all $g \in G$ such that (29) holds for all $\mathbf{x} = g^{-1} \cdot \mathbf{x}_0$. Then E is precisely the set of elements $g \in G$ such that (32) holds for every subset K when (u_t) is changed into its conjugate $(gu_tg^{-1})_t$ whose fixed line is $g^{-1}D$. Therefore E contains the set of all $g \in G$ such that the slope of gD is irrational and well approximable from both sides. The map from G to $\mathbb{R}$ sending g to the slope of gD is differentiable and with no critical points (it identifies with the left translation in G/P where P is the stabilizer of D). It follows that, for the Haar measure on G, almost every g belongs to E. End of the proof.

Remark 5.2. *In this example, we have shown that a non-centered random walk can stay in the cusp with high probability at arbitrarily large times. But the same idea shows that the walk can also stay very close to a closed orbit of U with high probability at arbitrarily large times. The example described above is for $SL_2(\mathbb{R})/SL_2(\mathbb{Z})$, but a similar phenomenon arises as soon as U is not uniquely ergodic on G/Γ, i.e. if there are proper homogeneous subspaces invariant under U.*

6. Questions and further remarks

(i) Let us come back to Theorem 3.15 and observe that the uniform estimate (16) implies in particular
$$\lim_{n\to+\infty} \frac{\mu^n(xB)}{\nu_n(xB)} = 1$$
uniformly when $|x|/\sqrt{n}$ stays bounded. Indeed it follows from the gaussian lower bound satisfied by the heat kernel $p_t(x)$ associated to μ (i.e. the density of ν_t). However, this tells nothing about the behavior of this ratio on a larger set of values of x. It would be interesting to obtain a result similar to Theorem 4.5 in the case of nilpotent groups under some additional moment assumption. The question can be asked already when μ has a density, i.e. in Theorem 3.14.

(ii) We did not mention non-centered random walks on non-commutative nilpotent Lie groups. In this case we have Raugi's central limit theorem [92] which asserts that under a certain renormalization $d_{1/\sqrt{n}}$ of $\mathcal{N}$ by linear dilations (which are not assumed to be automorphisms) determined by the law μ, the sequence of probability measures $d_{1/\sqrt{n}}(\mu^n \circ \delta_{e^{-n}X})$ (where $X \in \mathcal{N}$ represents the mean of μ in $\mathcal{N}/[\mathcal{N},\mathcal{N}]$) converges to a non-degenerate probability measure, which is the law at time 1 of a certain diffusion process on the vector space $\mathcal{N}$ whose infinitesimal generator is given explicitly. The non-centered case differs remarkably from the centered case since the renormalization $d_{1/\sqrt{n}}$ can be strictly stronger in the non-centered case. More precisely, let $D(\mu)$ be the exponent of dilation of the d_t's, i.e. the integer such that $vol(d_t(B)) = t^{D(\mu)}vol(B)$ for every compact subset B. Then we always have $d(N) \leq D(\mu) \leq 2d(N) - 1$. And $D(\mu) = d(N)$ if and only if the restriction of $ad(X)$ to each $\mathcal{N}^i/\mathcal{N}^{i+1}$ is zero. Hence in general, we can have $D(\mu) > d(N)$; this phenomenon is explained by the fact that inner automorphisms spread out the random walk on a larger domain. It is natural to ask how this phenomenon translates at the level of local limit theorems. In particular in the case of a finitely generated nilpotent group Γ which we embed as a co-compact lattice in a nilpotent Lie group, if μ is not centered, do we have (see Varopoulos' Theorem 3.8)
$$\sup_{x\in\Gamma} \mu^n(x) \leq \frac{C}{n^{D(\mu)/2}} \ ?$$

(*iii*) Let us now come back to the equidistribution in the Euclidean space (§ 2.3). The proof of the local limit theorem that we gave above for the group of motions of the plane makes use of an estimate of the norm of some operators associated to μ by the irreducible unitary representations of G (Lemma 2.1). When $d = 2$, this estimate is made possible mainly because G is then a solvable Lie group. If $d \geq 3$, G has a non-trivial Levi factor $SO(d)$ which is a semisimple compact Lie group and the proof of Lemma 2.1 fails. However it is natural to ask whether we still have a spectral gap in this case, that is

$$\|\pi_r(\mu)\| < 1$$

as soon as $r > 0$. The unitary representation π_r of G is irreducible if $r > 0$ and defined on $\mathbb{L}^2(S^{d-1})$ by the equation (5). When $r = 0$ we find the regular representation of $SO(d)$. The existence of a spectral gap when $r = 0$ is precisely the question asked by Sarnak in [96] and seems very delicate.

(*iv*) Let us consider Theorem 5.6. It implies in particular that the sequences of averages $(\frac{1}{n}\sum_{k=0}^{n-1}\mu^k * \delta_{\mathbf{x}})_n$ is relatively compact in the space of probability measures on G/Γ. One can ask whether this sequence converges to the invariant probability measure m_x from Theorem 5.5. Clearly any limit measure is μ-stationary. So the question boils down to showing that μ-stationary measures are in fact invariant under the elements of the support of μ, that is, in some sense, a Choquet-Deny property. In [38], Furstenberg calls this property *stiffness* and shows that it holds if μ is not singular with respect to the Haar measure on G. He also asks whether it holds if we only assume that the support of μ generate of dense or only Zariski dense subgroup in G.

(*v*) We did not mention here the problem of equidistribution in non-amenable groups. On semisimple Lie groups, we have Bougerol's local limit theorem [17] for probability measures μ such that some power of it is not singular with respect to the Haar measure. However the problem remains open for singular measures. In particular, let G be a semisimple group and a and b be two elements chosen randomly (for the Haar measure) in a small neighborhood of the identity. As can be shown, the group generated by a and b is free and dense in G, almost surely. Let then μ be the probability measure giving a weight of $1/4$ to a and b and their inverses. Do we have a ratio limit theorem for μ? Similarly, one can ask the analogous question when averages are made on balls of the free group instead.

References

[1] G. Alexopoulos, *Centered sub-laplacians with drift on Lie groups of polynomial volume growth*, Mem. Amer. Math. Soc. **155**, n. 739, (2002)

[2] G. Alexopoulos, *Random walks on discrete groups of polynomial volume growth*, Annals of Proba., vol **30**, n. 2, (2002) p. 723-801.

[3] G. Alexopoulos, *Centered densities on Lie groups of polynomial volume growth*, Probab. Theory Relat. Fields **124**, (2002) p. 112-150.

[4] N. Amosova, *Local limit theorems for probabilities of moderate deviations*, Lit. Mat. Sb. **14**, no3. (1974), p. 401-409.

[5] V.I. Arnol'd et A.L. Krylov, *Uniform distribution of points on a sphere and certain ergodic properties of solutions of linear ordinary differential equations in a complex domain*, Dokl. Akad. Nauk SSSR **148**, (1963), p. 9-12.

[6] A. Avez, *Limite de quotients pour des marches aléatoires sur des groupes,* C. R. Acad. Sci. Paris Sér. A-B **276** (1973), A317–A320.

[7] M. Babillot, *Points entiers et groupes discrets, de l'analyse aux systèmes dynamiques*, Panoramas et synthèses, (2002).

[8] P. Baldi, *Caractérisation des groupes de Lie connexes récurrents,* Ann. Inst. H. Poincaré Sect. B (N.S.) 17 (1981), no. **3**, 281–308.

[9] P. Baldi, L. Caramelino, *Large and moderate deviations for random walks on nilpotent groups*, J. Theoret. Probab. **12** (1999), no. 3, p. 779–809.

[10] C. Béguin, A. Valette, A. Zuk, *On the spectrum of a random walk on the discrete Heisenberg group and the norm of Harper's operator*, J. Geom. and Phys. (1997), vol **21**, n. 4, p. 337-356.

[11] J. Bellissard, B. Simon, *Cantor spectrum for the almost Mathieu Equation*, J. Funct. Anal. (1982), vol **48**, n. 3, p. 408-419.

[12] V. Bentkus, G. Pap, *The accuracy of Gaussian approximations in nilpotent Lie groups*, J. Theoret. Probab. **9** (1996), no. 4, p. 995–1017.

[13] P. Billingsley, *Measure and Probability*, 2ème édition. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, 1999

[14] F. Boca, A. Zaharescu, *Norm estimates of almost Mathieu operators*, preprint ArXiv 2003.

[15] A. Borel, *Introduction aux groupes arithmétiques*, Hermann, Paris (1969).

[16] A. Borel, *Linear Algebraic Groups*, 2nd edition, GTM **126**, Springer-Verlag, (1991) 288 pp.

[17] P. Bougerol, *Théorème central limite local sur certains groupes de Lie*, Ann. Sci. École Norm. Sup. (4) 14 (1981), no. **4**, 403–432 (1982).

[18] P. Bougerol, J. Lacroix, *Products of random matrices with applications to Schrödinger operators*, Birkhäuser, Prog. in Probab. & Stat. (1985).

[19] L. Breiman, *Probability*, Addison-Wesley (1968).

[20] E. Breuillard *Equidistribution of unipotent random walks on homogeneous spaces*, Chapter 2 in Thèse Université Paris Sud, Dec. 2003.

[21] E. Breuillard *Local limit theorems and equidistribution of random walks on the Heisenberg group*, Chapter 3 in Thèse Université Paris Sud, Dec. 2003.

[22] E. Breuillard *Diophantine measures and the speed of convergence in the local limit theorem on R^d*, [In French], Chapter 4 in Thèse Université Paris Sud, Dec. 2003.

[23] I. D. Brown, *Dual topology of a nilpotent Lie group*, Ann. Sci. École Norm. Sup. (4) **6** (1973), p. 407–411.

[24] H. Carlsson, *Error estimates in d-dimensional renewal theory*, Compositio Math. **46** (1982), no. 2, 227–253.

[25] H. Carlsson, *Remainder term estimates of the renewal function* Ann. Probab. **11** (1983), no. 1, p. 143-157.

[26] G. Choquet, J. Deny, *Sur l'équation de convolution $\mu = \mu * \sigma$*, C. R. Acad. Sc. Paris, t. **250** (1960) p. 799-801.

[27] L. Corwin et F.P. Greenleaf, *Representations of nilpotent Lie groups and their applications*, Cambridge studies in advanced mathematics 18, CUP (1990).

[28] H. Cramér, *Les sommes et les fonctions de variables aléatoires*, Paris Hermann, (1938).

[29] H. Cramér, *Elements of probability theory and some of its applications*, Cambridge (1962).

[30] P. Crépel et A. Raugi, *Théorème central limite sur les groupes nilpotents*, Ann. Inst. H. Poincaré sect. B, Prob. & Stat., vol XIV, 2, (1978) 145-164.

[31] E. B. Davies, *One-parameter semigroups*, London Mathematical Society Monographs, **15**, Acad. Press, (1980), 230 pp.

[32] D. Dolgopyat, *On mixing properties of compact group extensions of hyperbolic systems*, Israel J. Math. **130** (2002), p. 157–205.

[33] E. Dynkin, M. Malioutov, *Random walk on groups with a finite number of generators*, Dokl. Akad. Nauk SSSR **137** (1961) p.1042–1045.

[34] A. Eskin, G.A. Margulis, Recurrence properties of random walks on homogeneous spaces, (preprint E. Schrodinger Institute, Vienna 2001).

[35] A. Eskin, G.A. Margulis, S. Mozes, *Upper bounds and asymptotics in a quantitative version of the Oppenheim conjecture*, Ann. of Math. (2) **147** (1998), no. 1, p. 93–141.

[36] W. Feller, *An introduction to probability theory and its applications*, Vol. II. 2nd ed. John Wiley Sons, (1971).

[37] H. Furstenberg, *The unique ergodicity of the horocycle flow*, Recent Advances in Topological Dynamics (A. Besk ed.) Springer Verlag, 1972, p. 95-115.

[38] H. Furstenberg, *Stiffness of group actions*, in Lie groups and ergodic theory (Mumbai, 1996), 105–117, Tata Inst. Fund. Res. Stud. Math., 14, 1998.

[39] A. Gamburd, D. Jakobson, P. Sarnak, *Spectra of elements in the group ring of $SU(2)$*, J. Eur. Math. Soc. (JEMS) **1** (1999), no. 1, 51–85.

[40] B. Gnedenko, A. Kolmogorov, *Limit distributions for sums of independent random variables*, Revised edition Addison-Wesley, (1968) 293 pp.

[41] R. W. Goodman, *Nilpotent Lie groups: structure and applications to analysis*, LNM 562, Springer-Verlag (1976).

[42] L. Gorostiza, *The central limit theorem for random motions of d-dimensional Euclidean space*, Ann. of Proba. (1973).

[43] U. Grenander, *Probabilities on algebraic structures,* John Wiley & Sons ; Almqvist & Wiksell, Stockholm-Göteborg-Uppsala (1963), 218 pp.

[44] M. Gromov, *Groups of polynomial growth and expanding maps*, Publ. Math. IHES **53** (1981).

[45] Y. Guivarc'h, *Générateurs des groupes résolubles*, Publications des Sminaires de Mathmatiques de l'Universit de Rennes, 1967–1968, Exp. No. 1, 17 pp., Rennes, (1968).

[46] Y. Guivarc'h, *Croissance polynomiale des groupes de Lie et périodes des fonctions harmoniques*, Bull. Soc. Math. France **101**, (1973), p. 333-379.

[47] Y. Guivarc'h, *Equirépartition dans les espaces homogènes,* Théorie ergodique (Actes Journées Ergodiques, Rennes, 1973/1974), pp. 131–142. Lecture Notes in Math., Vol. 532, Springer, Berlin, (1976).

[48] Y. Guivarc'h, M. Keane, P. Roynette, *Marches aléatoires sur les groupes de Lie,* LNM Vol. 624, Springer-Verlag, (1977).

[49] Y. Guivarc'h, *Sur la loi des grands nombres et le rayon spectral des marches aléatoires*, Conf. on Random Walks (Kleebach, 1979), 3, Astérisque, **74**, (1980) p. 47–98.

[50] Y. Guivarc'h, *Marches aléatoires sur les groupes et problèmes connexes*, Fascicule de probabilits, 39 pp., Publ. Inst. Rech. Math. Rennes, Univ. Rennes I, Rennes, (1993).

[51] Y. Guivarc'h, *Marches aléatoires sur les groupes*, Development of mathematics 1950–2000, 577–608, Birkhuser, Basel, (2000).

[52] Y. Guivarc'h, *Limit theorems for random walks and products of random matrices*, Lectures held at the Tata Institute, Summer School in Probab. Theory, Sept. 2002.

[53] P. de la Harpe,*Topics in Geometric Group Theory*, Chicago Univerity Press, (2001).

[54] P. de la Harpe, A. Valette, *La proprit (T) de Kazhdan pour les groupes localement compacts, (avec un appendice de Marc Burger)*, Astrisque No. **175** (1989), 158 pp.

[55] W. Hazod, *Stetige Faltungshalbgruppen von Wahrscheinlichkeitsmassen und erzeugende Distributionen*, Lecture Notes in Mathematics, Vol. 595. Springer-Verlag, (1977), 157 pp.

[56] W. Hebisch, L. Saloff-Coste, *Gaussian estimates for Markov chains and random walks on groups*, Annals of Proba. **21** (1993) p. 673-709.

[57] W. Hebisch, A. Sikora, *A smooth subadditive homogeneous norm on a homogeneous group*, Studia Math. **96** (1990), no. 3, p. 231–236.

[58] H. Heyer, *Probability measures on locally compact groups*, Ergeb. **94** Spinger Verlag (1977).

[59] K. Hinderer, *Remarks on Directly Riemann Integrable Functions*, Math. Nachr. **130** (1987) p. 225-230.

[60] F. Hirsch and J.P. Roth, *Opérateurs dissipatifs et codissipatifs invariants sur un espace homogène*, Théorie du potentiel et analyse harmonique, Strasbourg, 1973, pp. 229–245. Lecture Notes in Math., Vol. 404, Springer, (1974).

[61] T. Höglund, *A multi-dimensional renewal theorem*, Bull. Sc. Math, 2ème série, **112** (1988) p. 111-138.

[62] G.A. Hunt, *Semi-groups of measures on Lie groups*, Trans. Amer. Math. Soc. **81** (1956), p. 264–293.

[63] I. Ibragimov, *A central limit theorem for a class of dependent random variables*, Teor. Verojatnost. i Primenen, **8**, (1963), p. 89–94.

[64] J. W. Jenkins, *Growth of locally compact groups*, J. Functional Analysis **12** (1973), p. 113-127.

[65] Y. Kawada, K. Itô, *On the probability distribution on a compact group,* I. Proc. Phys.-Math. Soc. Japan (3) **22**, (1940) p. 977-998.

[66] D.A. Kazhdan, *Uniform distribution on a plane*, Trudy Moskov. Mat. Ob. **14**, (1965) , p. 299-305.

[67] R. Keener, *Asymptotic expansions in multivariate renewal theory*, Stochastic Process. Appl. 34 (1990), no. 1, p. 137-153.

[68] H. Kesten, *Symmetric random walks on groups*, Trans. Amer. Math. Soc. **92**, (1959), p. 336-354.

[69] A. Kirillov, *Unitary representations of nilpotent Lie groups*, Uspehi Mat. Nauk **17** (1962) no. 4 (106), 57–110.

[70] A. Kirillov, *Éléments de la théorie des représentations*, Éditions Mir, Moscow, (1974) 347 pp.

[71] E. Le Page, *Théorèmes quotients pour certaines marches aléatoires*, Comptes Rendus Acad. Sc., **279**, série A, n. 2, (1974).

[72] E. Le Page, *Un théorème local sur le groupe de Heisenberg*, preprint IRMAR, Univ. Rennes 1.

[73] A. Lubotzky, *Discrete groups, expanding graphs and invariant measures*, with an appendix by J. D. Rogawski. Progress in Math. **125**, Birkhuser, 1994, 195 pp.

[74] G.A. Margulis, *Positive harmonic functions on nilpotent groups,* Dokl. Akad. Nauk SSSR **166** (1966) p. 1054–1057
[75] G.A. Margulis, *On the action of unipotent subgroups in the space of lattices*, Mat. Sb. 86 (1971), 552-556.
[76] G.A. Margulis, *Oppenheim conjecture*, Fields Medallists' lectures, 272–327, World Sci. Ser. 20th Century Math., 5, World Sci. Publishing, River Edge, NJ, 1997
[77] G.A. Margulis, Random walks and Borel Harish-Chandra theorem, preprint.
[78] G. A. Margulis, *Discrete subgroups of semi-simple Lie groups*, Ergeb. Math. Grenz. (3) **17** Springer Verlag, (1991), 388 pp.
[79] G.A. Margulis, G.M. Tomanov, *Invariant measures for actions of unipotent groups over local fields on homogeneous spaces*, Invent. Math. **116** (1994), no. 1-3, p. 347-392.
[80] R. Michel, *Results on probabilities of moderate deviations*, Ann. Prob., no. 2 (1974), p. 349-353.
[81] A. V. Nagaev, *Renewal theorems in* $\mathbb{R}^d$, Teor. Veroyatnost. i Primenen. **24** (1979), no. 3, p. 565-573.
[82] A. V. Nagaev, *Integral limit theorems with regards to large deviations when Cramér's condition is not satisfied*, Prob. Th. Appl. 14 (1969), p.51-63
[83] A. V. Nagaev, *Large deviations of sums of independant random variables*, Ann. of Prob. **7**, no 5., (1979), p. 745-789.
[84] D. Neuenschwander, *Probabilities on the Heisenberg group*, LNM 1630, Springer-Verlag (1996).
[85] V.I. Oseledec, *Markov chains, skew-products, and ergodic thoerems for "general" dynamical systems*, Th. Prob. & App. (1965), **10**, 3, 551-557.
[86] H. Poincaré, Calcul des Probabilités, 2ème éd., Gauthier Villars, (1912)
[87] M. S. Raghunathan, *Discrete Subgroups of Discrete Groups*, Springer Verlag (1972).
[88] M. Ratner, *Interaction between Ergodic Theory, Lie Groups, and Number Theory*, in Proc. ICM (Zurich, 1994) 157-182, Birkhaüser, 1995.
[89] M. Ratner, *On Raghunathan's measure conjecture*, Ann. Math. **134** (1991), 545-607.
[90] M. Ratner, *Raghunathan's topological conjecture and distributions of unipotent flows*, Duke Math. J. **63** (1991), 235-290.
[91] M. Ratner, *Interaction between Ergodic Theory, Lie Groups, and Number Theory*, in Proc. ICM (Zurich, 1994) 157-182, Birkhaüser, 1995.
[92] A. Raugi, *Théorème de la limite centrale sur les groupes Nilpotents,* Z. Wahrsch. Verw. Gebiete **43** (1978), no. 2, 149–172.
[93] A. Raugi, *A general Choquet-Deny theorem for a large class of locally compact second countable groups*, preprint IRMAR, Univ. Rennes 1, (2002).
[94] H. Rubin, J. Sethuraman, *Probabilities of moderate deviations*, Sankhyā Ser. A, **27** (1965), p. 325-346.
[95] W. Rudin, *Real and Complex Analysis*, Addison-Wesley.
[96] P. Sarnak, *Some applications of modular forms*, Cambridge Tracts in Mathematics, **99** CUP, Cambridge, (1990)
[97] H. P. Scheffler, *Domains of attraction of stable measures on the Heisenberg group*, Probab. Math. Statist. 14 (1993), no. 2, 327–345 (1994).
[98] J.P. Serre, *Lie algebras and Lie groups*, Springer LNM 1500.
[99] N. Shah, *Limit distributions of polynomial trajectories on homogeneous spaces*, Duke Math. J. **75**, 3, (1994), 711-732.
[100] N. Shah, *Invariant measures and orbit closures on homogeneous spaces for actions of subgroups generated by unipotent elements.* Lie groups and ergodic theory (Mumbai, 1996), 229–271, Tata Inst. Fund. Res. Stud. Math., **14**, Bombay, 1998
[101] E. Siebert, *Absolute continuity, singularity, and supports of Gauss semigroups on a Lie group*, Monat. für Math. **93**, (1982), p. 239-253.
[102] E. Siebert, *Densities and differentiability properties on Gauss semigroups on a Lie group*, Proc. Amer. Math. Soc. vol **91**, n. 2, (1984), p.298-305.
[103] A. D. Slastnikov, *Limit theorems for moderate deviations probabilities*, Probab. Theo. and Appl. **23** (1978), no2., p. 325-340.
[104] A. J. Stam, *Renewal theory in r dimensions I.* Compositio Math. **21** (1969) p. 383-399 and **23** (1971) p. 1-13.
[105] A. N. Starkov, *Dynamical systems on homogeneous spaces*, Transl. of Math. Monographs, **190**, AMS Prov. RI, (2000).
[106] Ch. Stone, *A local limit theorem for nonlattice multi-dimensional distribution functions*, Ann. Math. Stat, **36** (1965) 546-551.

[107] Ch. Stone, *Ratio limit theorems for random walks on groups*, Trans. Amer. Math. Soc. **125** (1966) p. 86–100.

[108] Ch. Stone, *Ratio local and ratio limit theorems*, 5th Berkeley Symp. on Math. Stat. and Prob., Berkeley and Los Angeles, UCP (1966), Vol II, 2, p. 217-224

[109] Ch. Stone, *Application of unsmoothing and Fourier Analysis to random walks*, Markov Processes and Potential Theory, Madison Wis, (1967), p. 165-192.

[110] D. Stroock, S.R.S Varadhan, *Limit theorems for random walks on Lie groups*, Indian J. of Stat. Sankhyā, Ser. A, **35**, (1973) p. 277-294.

[111] V.N. Tutubalin, *Compositions of measures on the simplest nilpotent group*, Teor. Verojatnost. i Primenen **9**, 1964, p. 531-539.

[112] N. Varopoulos, *Wiener-Hopf theory and nonunimodular groups*, J. Funct. Anal., **120,** (1994) p.467-483.

[113] N. Varopoulos, L. Saloff-Coste, et Th. Coulhon, *Analysis and geometry on groups*, Cambridge Tracts in Mathematics, **100**, CUP, Cambridge, 1992.

[114] D. Wehn, *Probabilities on Lie groups*, Proc. Nat. Acad. Sci. USA, vol **48**, (1962) p.791-795

[115] D. Wehn, *Limit distributions on Lie groups*, Yale thesis, 1960.

[116] D. Wehn, *Some remarks on Gaussian Distributions on a Lie Group*, Z. Wahrsch. Geb. **30**, p. 255-263.

[117] W. Woess, *Random walks on infinite graphs and groups*, Cambridge Tracts in Mathematics, **138**, CUP, Cambridge, (2000)

[118] R. Zimmer, *Ergodic theory and semi-simple groups*, Monographs in Math., **81**, Birkhäuser Verlag, (1984), 209 pp.