LOCAL LIMIT THEOREMS AND EQUIDISTRIBUTION OF RANDOM WALKS ON THE HEISENBERG GROUP

E. BREUILLARD

To the memory of Martine Babillot

Abstract. We prove local limit theorems for products of independent random variables on the Heisenberg group which are identically distributed with respect to an arbitrary centered and compactly supported probability measure μ. We also provide uniform estimates for translates of a bounded set by comparing μ^n to the associated heat kernel. This, in turn, enables us to show the equidistribution of Heisenberg-unipotent random walks on finite volume homogeneous spaces G/Γ.

The goal of this paper is to show the local limit theorem for centered probability measures on the Heisenberg group along with some of its refinements and applications. The local limit problem on non-commutative Lie groups has been studied by many authors in the last thirty or forty years (Ito-Kawada, Arnold-Krylov, Kazhdan, Bougerol, Le Page, Guivarc'h, Varopoulos, etc.). In the classical commutative case or in the compact group case the local limit theorem is available under the weakest assumptions on the probability measure (see [Sto] and [ItK]). In [Bou], Bougerol solves the local limit problem for absolutely continuous probability measures on semisimple Lie groups. In a recent work, Alexopoulos [Ale] obtains a very precise local limit theorem and estimates à la Berry-Essen for probability measures with a continuous density of compact support on an arbitrary connected Lie group of polynomial growth.

However, in this problem, an assumption of absolute continuity of the probability measure with respect to the Haar measure is often made, while the case of a possibly singular (e.g. finitely supported) measure remains generally open. Such cases include finitely supported probability measures on the group of isometries of the Euclidean 3-space (see [Kaz], [Gui2]) or the case of semisimple groups [Bou]. Similarly the speed of convergence to equidistribution is not well understood and seems to depend on difficult arithmetic questions (cf. the spectral gap conjecture [Sar] about the equidistribution on the sphere). In this paper however, we treat the case of an arbitrary, possibly singular, measure. We will focus on the simplest nilpotent Lie group: the first Heisenberg group.

Comparing the convolution powers of the measure with the associated heat kernel, we also obtain a uniform version of the local limit theorem yielding a fairly precise estimate on the asymptotic behavior of centered random walks on the Heisenberg group. This generalizes a result of Stone [Sto] in the commutative case. This estimate allows to show further equidistribution results for random walks on homogeneous spaces. Following this strategy, and making use of Ratner's theorem on orbits of unipotent flows (see [Rat], [Sha], [Sta]), we show at the end of the paper that centered Heisenberg-unipotent random walks on homogeneous spaces G/Γ, where Γ is discrete of finite co-volume in a Lie group G, are equidistributed in the closure of the orbit on which they live.

In the non-centered case, an interesting phenomenon can occur: non-centered unipotent random walks may not converge to any probability measure on G/Γ. For instance if Γ is not co-compact, they may stay outside an arbitrary compact set with high probability at arbitrary large times. Hence the hypothesis that the walk should be centered is crucial (unless of course the Haar measure is uniquely ergodic for the unipotent subgroup).

Finally let us remark that the results of this paper should extend to the case of an arbitrary simply connected nilpotent Lie group, but the technical difficulty of the forthcoming proofs forced us to restrict our attention to the Heisenberg group.

1. Statement of the results

Let G be the group of 3×3 upper-triangular unipotent matrices and e the identity in G. Let us fix the Haar measure on G, $dg = dxdydz$ where

$$g = \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$$

is simply denoted by $g = (x, y, z)$. We also denote by $|A|$ the Haar measure of a Borel set A and we fix a homogeneous norm $\|g\| = \max\{|x|, |y|, |z|^{1/2}\}$ on G.

We consider a probability measure μ on G with the following properties:

- compactly supported.
- centered: $\int p(x)d\mu(x) = 0$ where $p : G \to G/[G,G]$ is the canonical map.
- aperiodic: for any proper closed subgroup $H \subsetneq G$ and any $x \in G$, $\mu(xH) < 1$.

In particular, we make no assumption of smoothness for μ, which can be for instance finitely supported.

The convolution product of measures is denoted by $\mu * \nu$ and convolution powers simply denoted by μ^n.

The central limit theorem for G is well known (see the work of Wehn [Weh], as well as Tutubalin [Tut] and Crépel-Raugi [Rau]). It states that if $(d_t)_t$ is the semigroup of dilations given by $d_t(x, y, z) = (tx, ty, t^2z)$ then the sequence $d_{\frac{1}{\sqrt{n}}}(\mu^n)$ converges to some gaussian measure ν on G (in the sense of probability measures, i.e. $\int f \circ d_{\frac{1}{\sqrt{n}}} d\mu^n \to \int f d\nu$ for every bounded continuous function on G). The measure ν lies inside a gaussian semigroup of probability measures $(\nu_t)_{t>0}$ defined by its generating distribution

$$\mathcal{A}f = \frac{d}{dt}_{t=0} \int f(x) d\nu_t(x) = \overline{z}\partial_z f(e) + \overline{xy}\partial^2_{xy} f(e) + \frac{1}{2}\overline{x^2}\partial^2_x f(e) + \frac{1}{2}\overline{y^2}\partial^2_y f(e) \tag{1}$$

where $\overline{z} = \int z d\mu(x,y,z)$ and $\overline{xy} = \int xy d\mu(x,y,z)$. Then $\nu = \nu_1$ and $\nu_t = p_t(g)dg$ where $p_t(g)$ is the heat kernel associated to the operator defined by (1), hence is a strictly positive fastly decreasing smooth function on G (see [VSC]). It is the density of the Brownian Motion corresponding to $\mathcal{A}$ on G.

Let $c(\mu) = p_1(e) > 0$. For general references about gaussian semigroups and the Levy-Khintchin-Hunt formula, see [Gre], [Neu] and the original article of Hunt [Hun], as well as the survey article [Breu2] where a self-contained proof of Wehn's central limit theorem can be found.

We say that μ satisfies *Cramér's condition* if

$$\sup_{t^2+s^2\geq 1} \left| \int e^{i(tx+sy)} d\mu(x,y,z) \right| < 1$$

We obtain the theorems below without this assumption, but at some point we get a slightly stronger result if this assumption is made (apparently for technical reasons, but we have no guess whether it is necessary).

Let us state the local limit theorem for G together with a uniform version for translates of a bounded set. These results generalize to the Heisenberg group the well known theorems of classical probability theory on $\mathbb{R}^n$ (see [Bre] and [Sto]).

The method of proof makes use of the unitary representations of G to obtain a crucial spectral bound (Proposition 3.2). The main point after that is the study of an auxiliary quadratic dynamical system (see equations (30)), in order to obtain a key estimate similar to the domination condition appearing in the classical proof of the local limit theorem on the real line (see [Bre]).

Theorem 1.1. *(Local limit theorem) Let μ be a compactly supported aperiodic centered probability measure on G. Let f be a compactly supported continuous function on G. Then the following convergence holds uniformly when z varies in compact subsets of G*

$$\lim_{n\to+\infty} n^2 \int_G f(gz) d\mu^n(g) = c(\mu) \int_G f(g) dg$$

Theorem 1.2. *(Uniform local limit theorem) Let μ be a compactly supported aperiodic centered probability measure on G and $(v_t)_{t>0}$ the corresponding limit gaussian semigroup. Then for any bounded Borel subset $B \subset G$ with $|\partial B| = 0$,*

$$\lim_{n\to+\infty} \sup_{x\in G} n^2 \left| \mu^n(xB) - \nu_n(xB) \right| = 0$$

If we assume additionally Cramér's condition, then

$$\lim_{n\to+\infty}\sup_{x,y\in G} n^2\left|\mu^n(xBy)-\nu_n(xBy)\right|=0$$

Let us remark that the choice of $(\nu_t)_{t>0}$ depends on the choice we made of a semi-group of dilations $(d_t)_{t>0}$. Any other choice $(d'_t)_{t>0}$ for the semi-group of dilations is of the form $\phi\circ d_t\circ\phi^{-1}$ for some automorphism ϕ of G. The associated gaussian semi-group $(\nu'_t)_{t>0}$ is obtained from $(\nu_t)_{t>0}$ by composing by some automorphism of G and Theorem 1.2 remains valid if we take $(\nu'_t)_{t>0}$ instead of $(\nu_t)_{t>0}$.

Theorem 1.3. *(Concentration function) Under the assumptions above for μ, for every bounded set $K\subset G$, there is a constant C_K such that for all integers n*

$$\sup_{x\in G}\mu^n(xK)\le\frac{C_K}{n^2}$$

If we suppose additionally Cramér's condition, then

$$\sup_{x,y\in G}\mu^n(xKy)\le\frac{C_K}{n^2}$$

This uniform version of the local limit theorem for translates enables to show the following corollary. The point of this result is that the function f is not assumed to tend to 0 at infinity, and in particular, can have a periodic type of behavior. For the classical commutative case see [Breu].

Corollary 1.4. *Let f be an arbitrary left-uniformly continuous function on G such that the following limit exists*

$$\lim_{T_x\to\pm\infty,T_y\to\pm\infty,T_z\to\pm\infty}\frac{1}{T_xT_yT_z}\int_0^{T_x}\int_0^{T_y}\int_0^{T_z}f(x,y,z)dxdydz=\ell \tag{2}$$

where $\ell\in\mathbb{C}$. Then

$$\lim_{n\to+\infty}\int_G f(g)d\mu^n(g)=\ell$$

This corollary enables to prove further equidistribution results on homogeneous spaces. In the last section we show how to derive the equidistribution of Heisenberg-unipotent random walks on homogeneous spaces, such as horospheric random walks on complex hyperbolic manifolds, namely,

Theorem 1.5. *Let G be a connected Lie group and Γ a lattice in G. Let H be a closed subgroup of G consisting of unipotent elements and isomorphic to the Heisenberg group. Let μ be a centered compactly supported aperiodic probability measure on H. Then for an arbitrary $x\in G/\Gamma$ and for any bounded and continuous function f on G/Γ,*

$$\lim_{n\to+\infty}\int_H f(hx)d\mu^n(h)=\int_G f(g)dm_x(g)$$

where m_x is the unique H-invariant ergodic probability measure on G/Γ whose support is the closure of the orbit $\overline{Hx}$.

The existence of the measure m_x is given by Ratner's theorem (see [Rat1 to 3], [Sta]). The corresponding deterministic result was proved for one parameter subgroups by Ratner [Rat] elaborating on a weaker qualitative recurrence result due to Margulis [Mar] and subsequently generalized by Dani [Dan], and for general unipotent groups by Shah [Sha].

Note that this also implies that the only μ-stationary probability measures on G/Γ (i.e. the measures ν such that $\mu * \nu = \nu$) are the H-invariant ones. This means, in the terminology of Furstenberg (cf. [Fus]), that the action of H on G/Γ is *stiff* with respect to μ. But this also follows from the fact that (H, μ) has the Choquet-Deny property (i.e. the absence of bounded μ-harmonic functions) as follows from the work of Guivarc'h [Gui].

Also note that according to a general random ergodic theorem of Oseledec (cf. [Ose]), which is proved in the case when μ is symmetric (i.e. $\mu(A) = \mu(A^{-1})$), the convergence of Theorem 1.5 above holds for m-almost every x in G/Γ for any H-ergodic probability measure m on G/Γ. We underline that, in Theorem 1.5, we capture the behavior of the random walk for every starting point x.

In the above theorem, the assumption that μ is centered cannot be removed. As will be shown in §10.2.1, a simple use of the central limit theorem shows that for certain (in fact almost all) lattices in $\mathbb{R}^2$, any non centered unipotent random walk starting at that point in the space of lattices $SL_2(\mathbb{R})/SL_2(\mathbb{Z})$ will diverge, i.e. may remain very far with high probability at some arbitrary large time.

When H is uniquely ergodic on G/Γ (e.g. the horocycle flow on a compact Riemann surface), then Theorem 1.5 follows easily from an equidistribution theorem due to Guivarc'h (Théorème V.5. in [Gui]) and holds even when μ is not centered.

This application was originally motivated by the work of Eskin and Margulis [EsM], where they studied the case of random walks on G/Γ obtained by a measure μ whose support is Zariski dense in a semisimple group. Their main result is that the sequence $\mu^n * \delta_x$ is relatively compact in the space of probability measures on G/Γ.

2. Notations and outline of the paper

We keep the notations and terminology introduced in the last section. Let ρ be the regular representation of G on the functions on G: $\rho(g)f(x) = f(g^{-1}x)$. We denote by $\langle , \rangle$ the scalar product on $\mathbb{L}^2(G)$.

G is identified with its Lie algebra $\mathfrak{g}$ by writing $g = (x, y, z)$ with the help of the diffeomorphism

$$\begin{array}{rcl} \mathfrak{g} & \to & G \\ xX + yY + zZ & \mapsto & e^{yY}e^{xX}e^{zZ} \end{array} \tag{3}$$

where X, Y and Z are the upper triangular elementary matrices, with $[X, Y] = Z$.

The product in G is given by

$$gg' = (x + x', y + y', z + z' + xy')$$

In the sequel, we will need to look at possible other parametrizations of G, in particular at those of the form

$$\begin{array}{rcl} \phi_\sigma & : & \mathbb{R}^3 \to G \\ (x, y, z) & \mapsto & (x, y, z + \sigma(x, y)) \end{array}$$

where $\sigma(x, y)$ is a quadratic form in x and y. The Fourier transform of a function on G can be defined in different ways depending on the choice of a parametrization. Given a function $f : G \to \mathbb{C}$, we shall denote by $F_\sigma(f) : (\mathbb{R}^3)^\vee \to \mathbb{C}$ the Fourier transform taken in the parametrization defined by ϕ_σ. When $\sigma = 0$, we simply write $F_0(f) = \widehat{f}$. The variable in the dual space $(\mathbb{R}^3)^\vee$ will be denoted by $\xi = (t, s, \lambda)$. The formula reads:

$$\begin{array}{rcl} F_\sigma(f) & : & (\mathbb{R}^3)^\vee \to \mathbb{C} \\ \xi & = & (t, s, \lambda) \mapsto \dfrac{1}{(2\pi)^{3/2}} \displaystyle\int f(\phi_\sigma(x, y, z)) e^{i(tx+sy+\lambda z)} dxdydz \end{array} \tag{4}$$

Let $C_c(G)$ be the space of continuous and compactly supported functions on G. Let μ be a probability measure satisfying the properties listed in the introduction and (X, Y, Z) a random variable on G distributed according to μ. We set once and for all

$$\alpha = \frac{\mathbb{E}(XY)}{2\mathbb{E}(Y^2)} \tag{5}$$

Theorem 2.1. *Let $\sigma(x, y) = \alpha y^2$ and F_σ the Fourier transform just defined. Let f and g be two functions on G with $f \in C_c(G)$, g integrable and $F_\sigma(g) \in C_c(G)$. Then*

$$\lim_{n \to +\infty} n^2 \langle \rho(\mu^n) f, g \rangle = c(\mu) \int_G f \int_G \overline{g}$$

and, if we suppose additionally that $F_\sigma(g)$ is absolutely continuous, then uniformly when z varies in compact subsets,

$$\lim_{n \to +\infty} n^2 \int g(z^{-1}x) d\mu^n(x) = c(\mu) \int_G g$$

Below, we derive Theorem 1.1 from Theorem 2.1. The strategy for proving Theorem 2.1 follows the general scheme provided by Stone's proof of the local limit theorem on $\mathbb{R}^d$ [Sto] and is as follows. Looking at the Fourier transform of the integral, we give an explicit decomposition of the regular representation of G into a continuous direct sum of primary representations and treat each part of the integral to show that only the part with small λ's and small s and t's gives a contribution. Then we gain control on this part by showing a domination condition on the integrand. This is achieved by performing a Taylor expansion in s, t, λ of the characteristic function of

μ^n. Lebesgue's dominated convergence theorem combined with the point-wise convergence granted by the central limit theorem on G completes the proof. The proof of Theorem (1.2) makes use of the estimates previously obtained and goes along similar lines.

3. Irreducible unitary representations of G

The irreducible unitary representations of G are well known. Apart from characters, there is a one-parameter family of irreducible unitary representations π_λ ($\lambda \in \mathbb{R}\backslash\{0\}$) modeled on $\mathbb{L}^2(\mathbb{R})$ by

$$\pi_\lambda(g)f(t) = e^{i(tx+\lambda z)}f(t+\lambda y)$$

The following two propositions will be crucial in the proof. The first is quite standard (see [Gre], and [Gui]):

Proposition 3.1. *Let μ be an aperiodic probability measure on G. Then for any closed interval $I \subset \mathbb{R}\backslash\{0\}$*

$$\sup_{\lambda\in I} \|\pi_\lambda(\mu)\| < 1 \tag{6}$$

Proof. Let $\lambda_n \to \lambda \in I$ be such that

$$\left\langle \pi_{\lambda_n}(\mu * \mu^{-1})f_n(t), f_n(t)\right\rangle \to 1$$

for some sequence of vectors $f_n \in \mathbb{L}^2(\mathbb{R})$ of norm 1. Then, up to taking a subsequence, for $\mu * \mu^{-1}$-almost every x,

$$\langle \pi_{\lambda_n}(x)f_n(t), f_n(t)\rangle \to 1$$

But $\Gamma = \{x \in G, \langle \pi_{\lambda_n}(x)f_n(t), f_n(t)\rangle \to 1\}$ is clearly a subgroup of G, and $\mu * \mu^{-1}(\Gamma) = 1$. Since μ is aperiodic, Γ is dense in G, hence $[\Gamma,\Gamma]$ is dense in the center of G. In particular, we can find $(0,0,z) \in \Gamma$, such that $\lambda z \notin 2\pi\mathbb{Z}$. Then

$$\left\| e^{i\lambda_n z} f_n - f_n \right\| \to 0$$

which implies $e^{i\lambda z} = 1$ and provides the desired contradiction. □

The next proposition gives an estimate of the norm of the operators $\pi_\lambda(\mu)$. When μ is taken to be the symmetric Dirac measure on $(1,0,0)$ and $(0,1,0)$, this operator can be viewed as acting on $\ell^2(\mathbb{Z})$ and then coincide with the well-known Harper operator (see [BVZ]) studied in mathematical physics.

Proposition 3.2. *Let μ be a probability measure on G whose support is not contained in a coset of an abelian subgroup of G. Then we have*

$$\liminf_{\lambda\to 0} \frac{1-\|\pi_\lambda(\mu)\|}{|\lambda|} > 0 \tag{7}$$

Proof. Note that if we define $\mu^{-1}(B)=\mu(B^{-1})$ for every Borel subset B of G, then $\pi_\lambda(\mu^{-1})$ is the adjoint of $\pi_\lambda(\mu)$ and $\pi_\lambda(\mu^{-1}*\mu)$ is self-adjoint and non-negative.

If (7) does not hold, then we can find unit vectors $f_\lambda \in \mathbb{L}^2(\mathbb{R})$ such that for arbitrarily small λ's

$$\left\langle \pi_\lambda(\mu^{-1}*\mu)f_\lambda, f_\lambda \right\rangle \geq 1-|o(\lambda)|$$

From the assumption made on μ, we can find two non-commuting elements x_0 and x_1 lying in the support of $\mu^{-1}*\mu$. For each λ we can then find x_0^λ close to x_0 and x_1^λ close to x_1 (i.e. $\left\|x_1^\lambda - x_1\right\|$ and $\left\|x_0^\lambda - x_0\right\| < \|x_1 - x_0\|/3$) such that for $i=0,1$

$$Re\left\langle \pi_\lambda(x_i^\lambda)f_\lambda, f_\lambda \right\rangle \geq 1-|o(\lambda)|$$

The commutator $(x_0^\lambda, x_1^\lambda)$ belongs to the center, hence is of the form $(0,0,c_\lambda)$. From the choice of x_i^λ it follows that $c<c_\lambda<1/c$ for some constant $c\in(0,1)$. From the Stone-Von Neumann theorem (cf. [CoG]) we can then find an isometry I_λ of $\mathbb{L}^2(\mathbb{R})$ such that, conjugating by I_λ, $\pi_\lambda(x_0^\lambda)$ is turned into the translation by $c_\lambda\lambda$ and $\pi_\lambda(x_1^\lambda)$ is turned into the multiplication by e^{it}. Hence we can assume that for arbitrarily small λ's

$$\begin{aligned} Re\left\langle f_\lambda(t+c_\lambda\lambda), f_\lambda(t)\right\rangle &\geq 1-|o(\lambda)| \\ Re\left\langle e^{it}f_\lambda(t), f_\lambda(t)\right\rangle &\geq 1-|o(\lambda)| \end{aligned}$$

Or equivalently

$$\|f_\lambda(t+c_\lambda\lambda)-f_\lambda(t)\| = o(\sqrt{|\lambda|}) \tag{8}$$

$$\left\|e^{it}f_\lambda(t)-f_\lambda(t)\right\| = o(\sqrt{|\lambda|}) \tag{9}$$

Let $A_\varepsilon=\{t\in\mathbb{R}, d(t,2\pi\mathbb{Z})<\varepsilon\}$. We deduce from (9) that

$$\int_{A_\varepsilon^c} |f_\lambda(t)|^2 dt = o(|\lambda|/\varepsilon^2) \tag{10}$$

and from (8) that for any positive integer n

$$\|f_\lambda(t+nc_\lambda\lambda)-f_\lambda(t)\| = o(n\sqrt{|\lambda|})$$

or

$$Re\left\langle f_\lambda(t+nc_\lambda\lambda), f_\lambda(t)\right\rangle \geq 1-o(n^2\lambda) \tag{11}$$

We now take $n=[\frac{1}{2\sqrt{|\lambda|}}]+1$ and $\varepsilon=c_\lambda\sqrt{|\lambda|}/12$. Then for small enough λ we have $1>|nc_\lambda\lambda|>3\varepsilon$. So for small λ and as soon as $\varepsilon<(\pi-1)/2$, making use of (10) and applying the Cauchy-Schwarz inequality we have

$$
\begin{aligned}
|\langle f_\lambda(t+nc_\lambda\lambda), f_\lambda(t)\rangle| &\leq \sqrt{\int_{A_\varepsilon} |f_\lambda(t+nc_\lambda\lambda)|^2 \int_{A_\varepsilon} |f_\lambda(t)|^2} + \\
&\quad \sqrt{\int_{A_\varepsilon^c} |f_\lambda(t+nc_\lambda\lambda)|^2 \int_{A_\varepsilon^c} |f_\lambda(t)|^2} \\
&\leq 2\sqrt{\int_{A_\varepsilon^c} |f_\lambda(t)|^2} = o(\sqrt{|\lambda|}/\varepsilon) = o(1)
\end{aligned}
$$

which yields the desired contradiction with (11). □

4. Reducing to small values of λ

Here we will begin the proof of Theorem (2.1). Recall that G is identified with its Lie algebra via the choice of the basis (X, Y, Z) and the coordinates (x, y, z) defined as in equation (3). The center of G is the one parameter subgroup $H = e^{\mathbb{R}Z}$. We are going to decompose the regular representation of G into a continuous direct sum of other unitary representations. Every character of H is determined by a number $\lambda \in \widehat{H} \cong \mathbb{R}^\vee$. If $f \in \mathbb{L}^1(G)$, we define for $\mathbf{x} \in G$

$$
f_\lambda(\mathbf{x}) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(\mathbf{x}e^{zZ}) e^{i\lambda z} dz
$$

We check that $f_\lambda(\mathbf{x}e^{zZ}) = e^{-i\lambda z} f_\lambda(\mathbf{x})$. By the Fourier isometry, if $z \mapsto f(\mathbf{x}e^{zZ})$ is in $\mathbb{L}^2(\mathbb{R})$, then $\lambda \mapsto f_\lambda(\mathbf{x})$ is in $\mathbb{L}^2(\mathbb{R}^\vee)$, and

$$
\int_{\mathbb{R}} |f(\mathbf{x}e^{zZ})|^2 dz = \int_{\mathbb{R}^\vee} |f_\lambda(\mathbf{x})|^2 d\lambda
$$

By Fubini's theorem, it follows that, if f also belongs to $\mathbb{L}^2(G)$ then $|f_\lambda(\mathbf{x})|$ is in $\mathbb{L}^2(G/H)$ for almost every λ. Let us write $\mathcal{H}_\lambda$ the Hilbert space of measurable functions F on G such that $F(\mathbf{x}e^{zZ}) = e^{-i\lambda z} F(\mathbf{x})$ and $|F(\mathbf{x})|$ is square integrable on G/H. Then $\mathcal{H}_\lambda$ is a realization of the induced representation $\rho_\lambda = Ind_H^G \lambda$, where G acts by left translations. The above Plancherel formula for f_λ shows that we have the continuous sum decomposition

$$
\rho = \int^\oplus \rho_\lambda d\lambda
$$

and if f, g belong to $\mathbb{L}^2(G)$

$$
\langle \rho(\mu^n) f, g\rangle = \int \langle \rho_\lambda(\mu^n) f_\lambda, g_\lambda\rangle_{\mathcal{H}_\lambda} d\lambda
$$

It is easy to see that ρ_λ is a primary unitary representation of G and that the representation π_λ defined above is the only irreducible representation of G contained in ρ_λ. Moreover, its multiplicity is infinite:

$$\rho_\lambda = \int^{\oplus} \pi_\lambda(s)ds \tag{12}$$

with $\pi_\lambda(s) \simeq \pi_\lambda$ for all s.

Now from Proposition (3.2) and from (12) we obtain that for some $\varepsilon \in (0,1)$ there exists some $c > 0$ such that for all $\lambda \in \mathbb{R}$ with $|\lambda| \leq \varepsilon$ and all $n \in \mathbb{N}$

$$\|\rho_\lambda(\mu^n)\| \leq \|\rho_\lambda(\mu)\|^n \leq e^{-c|\lambda|n} \tag{13}$$

therefore for any integer $k_0 \geq 3$, whenever $D > k_0/c$ we have

$$\begin{aligned}
\left| n^2 \int_{D\frac{\log n}{n} \leq |\lambda| \leq \varepsilon} \langle \rho_\lambda(\mu^n) f_\lambda, g_\lambda \rangle_{\mathcal{H}_\lambda} d\lambda \right| &\leq n^2 \int_{D\frac{\log n}{n} \leq |\lambda| \leq \varepsilon} e^{-k_0 \log n} \|f_\lambda\|_{\mathcal{H}_\lambda} \|g_\lambda\|_{\mathcal{H}_\lambda} d\lambda \\
&\leq \frac{1}{n^{k_0-2}} \int_{\lambda \in \mathbb{R}} \|f_\lambda\|_{\mathcal{H}_\lambda} \|g_\lambda\|_{\mathcal{H}_\lambda} d\lambda \\
&\leq \frac{1}{n^{k_0-2}} \|f\|_{L^2(G)} \|g\|_{L^2(G)} \to 0 \text{ as } n \to \infty
\end{aligned} \tag{14}$$

The last step follows from the Cauchy-Schwarz inequality. Hence this part of the integral tends to 0. Similarly if I denotes any of the intervals $[-A, -\varepsilon]$ or $[\varepsilon, A]$ where A is some positive number such that g_λ is identically zero outside $[-A, A]$, then it follows from Proposition (3.2) and from (12) that there is some constant $\beta \in (0,1)$ such that

$$\sup_{\lambda \in I} \|\rho_\lambda(\mu^n)\| \leq \beta^n$$

Hence

$$\begin{aligned}
\left| n^2 \int_{\lambda \in I} \langle \rho_\lambda(\mu^n) f_\lambda, g_\lambda \rangle d\lambda \right| &\leq n^2 \beta^n \int_{\lambda \in \mathbb{R}} \|f_\lambda\|_{\mathcal{H}_\lambda} \|g_\lambda\|_{\mathcal{H}_\lambda} d\lambda \\
&\leq n^2 \beta^n \|f\|_{L^2(G)} \|g\|_{L^2(G)} \\
&\leq \frac{C}{n^{k_0-2}} \|f\|_{L^2(G)} \|g\|_{L^2(G)}
\end{aligned} \tag{15}$$

for some constant $C > 0$ depending on β. The right hand side clearly tends to 0 as n tends to infinity.

Now observe that if $F_\sigma(g)$ (defined in Theorem 2.1) has compact support, then g_λ is identically zero outside some bounded set of values of λ. More precisely, if for some fixed λ, $F_\sigma(g)(t, s, \lambda) = 0$ for all t and s, then g_λ vanishes identically. The last argument allows then to reduce to small values of λ (i.e. less that $D\frac{\log n}{n}$ for some fixed $D > 0$).

We are now going to perform Fourier integration one step further, i.e. on G/H. We fix an arbitrary Borel section $\overline{\sigma} : G/H \to G$ (given in the above coordinates by $(x,y) \mapsto (x,y,\sigma(x,y))$ where σ is some measurable function on $\mathbb{R}^2$). Then $f_\lambda(\overline{\sigma}(\overline{\mathbf{x}}))$ is in $\mathbb{L}^2(G/H)$ for almost every λ.

In those coordinates, we can write down explicitely the decomposition of ρ into the continuous sum of the ρ_λ's. Indeed, let f and g be two functions in $\mathbb{L}^1(G) \cap \mathbb{L}^2(G)$ and $\mathbf{y}$ be given in G. We have on the one hand:

$$\langle \rho(\mathbf{y})f, g\rangle = \int_{\mathbb{R}^\vee} \langle \rho_\lambda(\mathbf{y}^{-1})f_\lambda, g_\lambda\rangle_{\mathcal{H}_\lambda} d\lambda$$

and on the other hand,

$$\begin{aligned}\langle \rho(\mathbf{y})f, g\rangle &= \int_{G/H}\int_H f(\mathbf{y}^{-1}\mathbf{x}e^{zZ})\overline{g}(\mathbf{x}e^{zZ})dzd\overline{\mathbf{x}} \\ &= \int_{G/H}\int_{\mathbb{R}^\vee} f_\lambda(\mathbf{y}^{-1}\mathbf{x})\overline{g}_\lambda(\mathbf{x})d\lambda d\overline{\mathbf{x}} \\ &= \int_{\mathbb{R}^\vee}\int_{G/H} f_\lambda(\mathbf{y}^{-1}\mathbf{x})\overline{g}_\lambda(\mathbf{x})d\overline{\mathbf{x}}d\lambda \\ &= \int_{\mathbb{R}^\vee}\int_{\mathbb{R}^2} f_\lambda(\mathbf{y}^{-1}\overline{\sigma}(x,y))\overline{g}_\lambda(\overline{\sigma}(x,y))dxdyd\lambda\end{aligned}$$

For notational convenience we set $\phi_{\mathbf{y},\sigma,\lambda}(x,y) := f_\lambda(\mathbf{y}^{-1}\overline{\sigma}(x,y))$ and $\psi_{\sigma,\lambda}(x,y) := \overline{g}_\lambda(\overline{\sigma}(x,y))$. For almost all $\lambda \in \mathbb{R}$, these functions belong to $\mathbb{L}^2(\mathbb{R}^2)$. Performing Fourier transform on $\mathbb{L}^2(\mathbb{R}^2)$ now we get:

$$\langle \rho(\mathbf{y})f, g\rangle = \int_{\mathbb{R}^\vee}\int_{\mathbb{R}^{2\vee}} \widehat{\phi_{\mathbf{y},\sigma,\lambda}}(t,s)\overline{\widehat{\psi_{\sigma,\lambda}}}(t,s)dtdsd\lambda$$

Now a straightforward computation yields that $\widehat{\psi_{\sigma,\lambda}}(t,s)$ is the Fourier transform at (t,s,λ) of g^σ defined by

$$g^\sigma := (x,y,z) \mapsto g(x,y,\sigma(x,y)+z) \tag{16}$$

$$\widehat{\psi_{\sigma,\lambda}}(t,s) = \widehat{g^\sigma}(t,s,\lambda)$$

Similarly if $\mathbf{y} = (\mathbf{y}_x, \mathbf{y}_y, \mathbf{y}_z)$ we compute,

$$\widehat{\phi_{\mathbf{y},\sigma,\lambda}}(t,s) = \frac{1}{(2\pi)^{3/2}}\int e^{i(\mathbf{y}_x(t+\lambda y)+\mathbf{y}_y s+\lambda(\mathbf{y}_z-\sigma(x+\mathbf{y}_x,y+\mathbf{y}_y)}e^{i(tx+sy+\lambda z)}f(x,y,z)dxdydz$$

Hence, if we suppose additionally that f has compact support on G and $F_\sigma(g) = \widehat{g^\sigma}$ has compact support on $\mathbb{R}^3$, we obtain

$$\langle \rho(\mathbf{y})f, g\rangle = \frac{1}{(2\pi)^{3/2}}\int d\mathbf{x}f(\mathbf{x})\int_{\mathbb{R}^\vee}\int_{\mathbb{R}^{2\vee}} e^{i[\mathbf{y}_x(t+\lambda y)+\mathbf{y}_y s+\lambda(\mathbf{y}_z-\sigma(x+\mathbf{y}_x,y+\mathbf{y}_y))]}e^{i(tx+sy+\lambda z)}\overline{\widehat{g^\sigma}}(t,s,\lambda)d\xi$$

in other words,

$$\langle \rho(\mathbf{y})f, g\rangle = \int_G \frac{d\mathbf{x}}{(2\pi)^{3/2}} f(\mathbf{x}) \left\langle e^{i\psi(\xi,\mathbf{y},\mathbf{x})}, \widehat{g^{\sigma}}(t-\lambda y, s, \lambda)\right\rangle_{\xi=(t,s,\lambda)\in\mathbb{R}^3}$$

where $\mathbf{x} = (x, y, z)$, $\mathbf{y} = (\mathbf{y}_x, \mathbf{y}_y, \mathbf{y}_z)$, $\xi = (t, s, \lambda)$ and $\psi(\xi, \mathbf{y}, \mathbf{x}) = t(\mathbf{y}_x + x) + s(\mathbf{y}_y + y) + \lambda(z - xy + \mathbf{y}_z - \sigma(x + \mathbf{y}_x, y + \mathbf{y}_y))$.

Let S is a random variable in G with distribution μ^n. The quantity we are interested in is

$$n^2\mathbb{E}\left(\langle \rho(S)f, g\rangle\right) = \int_G \frac{d\mathbf{x}}{(2\pi)^{3/2}} f(\mathbf{x}) \left\langle \mathbb{E}(e^{i\psi(\xi,S,\mathbf{x})}), \widehat{g^{\sigma}}(t-\lambda y, s, \lambda)\right\rangle_{\xi=(t,s,\lambda)\in\mathbb{R}^3}$$

Let $D > 0$. We split it in two parts and write $n^2\mathbb{E}\left(\langle \rho(S)f, g\rangle\right) = A_n + B_n$ where

$$\begin{aligned} A_n &= \int_G \frac{d\mathbf{x}}{(2\pi)^{3/2}} f(\mathbf{x}) J_n(\mathbf{x}) \\ B_n &= \int_G \frac{d\mathbf{x}}{(2\pi)^{3/2}} f(\mathbf{x}) I_n(\mathbf{x}) \end{aligned}$$

and

$$J_n(\mathbf{x}) = n^2 \int_{|\lambda|\geq D\frac{\log n}{n}} \left\langle \mathbb{E}(e^{i\psi(\xi,S,\mathbf{x})}), \widehat{g^{\sigma}}(t-\lambda y, s, \lambda)\right\rangle_{(t,s)\in\mathbb{R}^2} d\lambda$$

$$I_n(\mathbf{x}) = n^2 \int_{|\lambda|\leq D\frac{\log n}{n}} \left\langle \mathbb{E}(e^{i\psi(\xi,S,\mathbf{x})}), \widehat{g^{\sigma}}(t-\lambda y, s, \lambda)\right\rangle_{(t,s)\in\mathbb{R}^2} d\lambda$$

The part A_n has already been dealt with, because the above computations show that

$$A_n = n^2 \int_{|\lambda|\geq D\frac{\log n}{n}} \langle \rho_\lambda(\mu^n) f_\lambda, g_\lambda\rangle_{\mathcal{H}_\lambda} d\lambda$$

and (applying (14) and (15)) there is $C \geq 0$ (depending on D, μ and the size of the set $\{\lambda, \exists (t, s), F_g(g)(s, t, \lambda) \neq 0\}$) such that if $n \geq 1$

$$|A_n| \leq \frac{C}{n^{k_0-2}} \|f\|_{L^2(G)} \|g\|_{L^2(G)} \tag{17}$$

which tends to zero as soon as k_0 is taken such that $D \geq k_0/c \geq 3/c$ (where c was the constant defined in (13)).

Hence in the sequel, fixing $\mathbf{x} = (x, y, z) \in G$, we shall focus on the term $I_n(\mathbf{x})$. Before going further, we shall fix once and for all the section σ. We take it of the form proposed in Theorem 2.1, that is $\sigma(x, y) = \alpha y^2$ where α is defined in terms of the moments of μ in equation (5). In this case,

$$\begin{aligned}\mathbb{E}(e^{i\psi(\xi,S,\mathbf{x})}) &= \mathbb{E}(e^{i[t(S_x+x)+s(S_y+y)+\lambda(z+S_z-\sigma(x+S_x,y+S_y))]}) \\ &= e^{i(tx+sy+\lambda z)}\mathbb{E}(e^{i\left[tS_x+sS_y+\lambda S_z-\lambda\alpha(y+S_y)^2\right]}) \\ &= e^{i(tx+sy+\lambda(z-\alpha y^2))}\mathbb{E}(e^{i\left[tS_x+(s-2\alpha\lambda y)S_y+\lambda(S_z-\alpha S_y^2)\right]}\end{aligned}$$

Hence,

$$I_n(\mathbf{x}) = n^2 \int_{|\lambda|\le D\frac{\log n}{n}} \left\langle \mathbb{E}(\theta_n(t,s,\lambda)), e^{-i\phi}\widehat{g^\sigma}(t-\lambda y, s+2\alpha\lambda y,\lambda)\right\rangle_{(t,s)\in\mathbb{R}^2} d\lambda$$

where $\phi = tx + sy + \lambda(z - xy + \alpha y^2)$ and

$$\theta_n(t,s,\lambda) = e^{i\left[tS_x+sS_y+\lambda(S_z-\alpha S_y^2)\right]} \tag{18}$$

We shall next estimate $\mathbb{E}(\theta_n(t,s,\lambda))$ for small values of λ.

5. Evaluation of the integral for small λ

Let us fix $D > 0$. The remainder of this section is devoted to finding a suitable bound for the expectation $\mathbb{E}(\theta_n(t,s,\lambda))$ when $|\lambda| \le D\frac{\log n}{n}$ and s and t take values that are bounded away from 0 and infinity.

Let $G_n = (X_n, Y_n, Z_n)$ be a sequence of independent random variables identically distributed according to the probability measure μ on G. We write $S_n = G_n \cdot ... \cdot G_1$ the product of these variables. The law of S_n is μ^n. Bearing in mind the form of the product on G, we get $S_n = (S_{n,x}, S_{n,y}, S_{n,z})$ where

$$\begin{aligned}S_{n,x} &= X_1 + ... + X_n \\ S_{n,y} &= Y_1 + ... + Y_n \\ S_{n,z} &= Z_1 + ... + Z_n + X_2Y_1 + X_3(Y_1+Y_2) + ... + X_n(Y_1 + ... + Y_{n-1})\end{aligned}$$

Let θ_n be the random variable defined in (18) as follows

$$\begin{aligned}\theta_n &= \theta_n(t,s,\lambda) = e^{i(tS_{n,x}+sS_{n,y}+\lambda(S_{n,z}-\alpha S_{n,y}^2))} \\ &= e^{-i\lambda\alpha U_n^2}\prod_{1\le k\le n} e^{iX_k(t+\lambda U_{k-1})}e^{i(sY_k+\lambda Z_k)} \\ \text{and } \theta_0 &= 1\end{aligned} \tag{19}$$

where $U_k = Y_1 + ... + Y_k$ and $U_0 = 0$ as above. In the sequel we fix the value of α (as in equation (5)) to be

$$\alpha = \mathbb{E}(XY)/2\mathbb{E}(Y^2).$$

We will also use the following notation:

$$\beta(t,s) = \mathbb{E}(e^{i(X_1t+Y_1s)})$$

5.1. **Moderate deviations.** We shall need the following lemma about moderate deviations. Since we could not find a reference for this precise form of the estimate we want (about the maximum of the random walk up to time n), we include a proof. It follows the well known argument of Cramér via Laplace transforms.

Lemma 5.1. *Let $U_0 = 0$ and $U_n = Y_1 + ... + Y_n$ be a sum of independent identically distributed real random variables Y_n's, which are assumed of compact support and centered. Let $\mathcal{A}_n$ be the event $\{\max_{0\leq k\leq n}|U_k| \leq \sqrt{n}\log n\}$ and let $\mathcal{A}_n^c$ be the complementary event. Then for every non-negative integer p, there are constants $c_p > 0$ and $C_p > 0$ such that for all integers $n \in \mathbb{N}$*

$$\mathbb{E}(\max_{0\leq k\leq n}|U_k|^p \mathbf{1}_{\mathcal{A}_n^c}) \leq C_p e^{-c_p \log^2 n}$$

Proof. Let Y be some random variable distributed according to the common distribution of the Y_n's. Since Y has compact support, if $D > 0$ is a bound for the support, we obviously get

$$|U_n| \leq Dn$$

for all n. Hence for any fixed $p \geq 0$

$$\mathbb{E}(\max_{0\leq k\leq n}|U_k|^p \mathbf{1}_{\mathcal{A}_n^c}) \leq D^p n^p \mathbb{P}(\mathcal{A}_n^c)$$

Therefore it is enough to show the lemma when $p = 0$.

We can assume that Y is not identically 0. Define the Laplace transform $\Lambda(\lambda)$ of Y by $e^{\Lambda(\lambda)} = \mathbb{E}(e^{\lambda Y})$ for a positive real $\lambda > 0$. Then define the Fenchel transform $\Lambda^*(x) = \sup_{\lambda>0}(\lambda x - \Lambda(\lambda))$ for a given $x > 0$. Clearly, $\Lambda^*(x)$ is a non-decreasing function of $x > 0$. The function $\Lambda(\lambda)$ is strictly convex, since its second derivative $\Lambda''(\lambda) = e^{-2\Lambda(\lambda)}(\mathbb{E}(e^{\lambda Y})\mathbb{E}(Y^2 e^{\lambda Y}) - \mathbb{E}(Ye^{\lambda Y})^2)$ is > 0 from the Cauchy-Schwarz inequality. In particular the supremum $\Lambda^*(x)$ is attained for a unique value λ_x of λ given by the equation $\Lambda'(\lambda_x) = x$. And if $x > 0$ then $\lambda_x > 0$ because $\Lambda'(0) = 0$ since Y is centered. Differentiating the relation $\Lambda'(\lambda_x) = x$, we obtain

$$\frac{d\lambda_x}{dx} = \frac{1}{\Lambda''(\lambda_x)}$$

hence

$$\frac{d\Lambda^*}{dx} = \lambda_x$$

and

$$\frac{d^2\Lambda^*}{dx^2} = \frac{1}{\Lambda''(\lambda_x)} > 0$$

Therefore $\Lambda^*(x)$ is strictly convex for $x > 0$.

Since Y has compact support, it has moments of any order and in particular we have the following Taylor expansion

$$\Lambda(\lambda) = \frac{\lambda^2}{2}\mathbb{E}(Y^2) + O(\lambda^3)$$

and

$$\Lambda'(\lambda) = \lambda \mathbb{E}(Y^2) + O(\lambda^2) \tag{20}$$

In particular, as x tends to 0 the value λ_x tends to 0 too. We set $\Lambda^*(0) = 0$. By definition of Λ^*, we have for any $x > x_0 > 0$

$$\Lambda^*(x) \geq \lambda_{x_0} x - \Lambda(\lambda_{x_0}) \tag{21}$$

Now let us study the function Λ^* near 0. From (20), we get

$$\lambda_x = \frac{x}{\mathbb{E}(Y^2) + O(\lambda_x)}$$

and

$$\lim_{x \to 0^+} \frac{\lambda_x \mathbb{E}(Y^2)}{x} = 1$$

Hence

$$\begin{aligned} \Lambda^*(x) &= \lambda_x x - \Lambda(\lambda_x) \\ &= \lambda_x x - \lambda_x^2 \frac{\mathbb{E}(Y^2)}{2} + O(\lambda_x^3) \\ &= \frac{x^2}{\mathbb{E}(Y^2)}(u - \frac{u^2}{2}) + O(ux)^3 \end{aligned}$$

where $u = \lambda_x \mathbb{E}(Y^2)/x$. As x tends to 0, u tends to 1, therefore

$$\Lambda^*(x) \geq \frac{x^2}{4\mathbb{E}(Y^2)} \tag{22}$$

for any sufficiently small x.

Similarly

$$\begin{aligned} \frac{\Lambda(\lambda_x)}{x} &= \frac{\lambda_x^2}{2x}\mathbb{E}(Y^2) + O(\frac{\lambda_x^3}{x}) \\ &= \lambda_x \frac{u}{2} + O(\lambda_x^2 u) \end{aligned}$$

Hence

$$\frac{\Lambda(\lambda_x)}{x} \leq \frac{3}{4}\lambda_x \tag{23}$$

for all sufficiently small x. Fix $x_0 > 0$ such that both (22) and (23) hold for all $x \in (0, x_0]$. From (21) we obtain immediately for all $x \geq x_0$

$$\Lambda^*(x) \geq \lambda_{x_0}(x - \frac{3}{4}x_0) \tag{24}$$

Now let us now apply these estimates obtained above for $\Lambda^*(x)$ to the probability $\mathbb{P}(U_k > \sqrt{n}\log(n))$. We write for $x > 0$ and $\lambda > 0$

$$\mathbb{E}(e^{\lambda U_n}) \geq e^{\lambda x}\mathbb{P}(U_n > x)$$

or

$$\mathbb{P}(U_n > x) \leq e^{-n(\lambda \frac{x}{n} - \Lambda(\lambda))}$$

Hence taking the supremum of $\lambda > 0$

$$\mathbb{P}(U_n > x) \leq \exp(-n\Lambda^*(\frac{x}{n})) \tag{25}$$

Take two integers k and n with $k \leq n$. We have

$$\mathbb{P}(U_k > \sqrt{n}\log(n)) \leq \exp(-k\Lambda^*(\frac{\sqrt{n}\log n}{k}))$$

Suppose first that $\frac{\sqrt{n}\log n}{k} \geq x_0$. Then it follows from (24) that

$$\begin{aligned}\mathbb{P}(U_k &> \sqrt{n}\log(n)) \leq \exp(-\lambda_{x_0}\sqrt{n}\log n + \frac{3}{4}kx_0\lambda_{x_0}) \\ &\leq \exp(-\frac{1}{4}\lambda_{x_0}\sqrt{n}\log n) \leq e^{-c(\log n)^2}\end{aligned}$$

where we can take $c = \lambda_{x_0}/4$.

Now assume that $\frac{\sqrt{n}\log n}{k} \leq x_0$. Then it follows from (22) that

$$\mathbb{P}(U_k > \sqrt{n}\log(n)) \leq e^{-c(\log n)^2}$$

where c can be taken to be $1/4\mathbb{E}(Y^2)$.

By taking the lesser of the two c above we can now write

$$\mathbb{P}(U_k > \sqrt{n}\log(n)) \leq e^{-c(\log n)^2}$$

for all integers n and k with $n \geq k \geq 1$. Hence

$$\mathbb{P}(\mathcal{A}_n^c) \leq 2ne^{-c(\log n)^2}$$

Therefore there is a constant $c_0 > 0$ smaller that c such that when n is larger than say n_0 we have

$$\mathbb{P}(\mathcal{A}_n^c) \leq e^{-c_0(\log n)^2}$$

For C_0 we may take $e^{c_0(\log n_0)^2}$ and we have obtained the desired inequality. $\square$

Obviously, we may, and do, assume that $c_{p+1} < c_p$ for all p. This lemma will enable us to reduce to the case when U_k does not take very big values. Let us also remark that in the above proof, the constant c_p can be chosen to depend only on the size of the support of Y, i.e. on $M = \min\{t, P(|Y| < t) = 1\}$. In the sequel, we will use freely the result of Lemma 5.1, in particular the fact that $\mathbb{E}(|U_n|^p) = O_p(n^{p/2}\log^p(n))$ for any $p \geq 0$.

5.2. **Estimate for "large" values of s and t.**

Proposition 5.2. *Let ε, A, D be positive numbers with $A > \varepsilon > 0$. Let k be an arbitrary positive integer. Then the following estimate holds uniformly when $\lambda \in [-D\frac{\log n}{n}, D\frac{\log n}{n}]$ and $\varepsilon < s^2 + t^2 < A$:*

$$\mathbb{E}(\theta_n) = O(\frac{1}{n^k})$$

where the constant in O depends on ε, A, k, D and μ.

Proof. The proof will proceed by induction. Recall that $\mathcal{A}_n$ is the event $\{\max_{0\leq k\leq n}|U_k| \leq \sqrt{n}\log n\}$. In fact, we will take for induction hypothesis the following statement, which we denote by $H_{j,k}$ for $j \in \mathbb{N}$ and $k \in \mathbb{Z}$:

$$\mathbb{E}(\theta_n U_n^j) = O(\frac{\log^{2k+3j} n}{n^{k/2}})$$

holds uniformly when the parameters (s, t, λ) vary in the range defined in the statement of the proposition, and where O depends only on μ, ε, A, D, k and j.

Clearly $H_{0,k}$ for all $k \geq 0$ implies the proposition. Also note that $H_{j,k}$ implies $H_{j,k-1}$.

Let $k + j \geq 0$. Suppose that $D > 0$ is given and let $\lambda \in [-D\frac{\log n}{n}, D\frac{\log n}{n}]$. We do not make an assumption on (s, t) for the moment. Making use of Lemma 5.1 and of the independence of the random variables G_i's, we have the following Taylor expansion. For all positive integers $p \leq n$,

$$\begin{aligned}
\mathbb{E}(\theta_p U_p^j) &= \mathbb{E}(\theta_p U_p^j \mathbf{1}_{\mathcal{A}_p}) + O_{\mu,j}(e^{-c_j \log^2 p}) \\
&= \mathbb{E}(\theta_{p-1} U_p^j e^{-i\lambda\alpha(2U_{p-1}Y_p + Y_p^2)} e^{iX_p(t+\lambda U_{p-1})} e^{i(sY_p+\lambda Z_p)} \mathbf{1}_{\mathcal{A}_p}) + O_{\mu,j}(e^{-c_j \log^2 p}) \\
&= \mathbb{E}[\theta_{p-1} U_p^j e^{i(X_p t + Y_p s)} \mathbf{1}_{\mathcal{A}_p} \sum_{l=0}^{k+j} \frac{(i\lambda)^l}{l!} \\
&\sum_{q=0}^{l} C_l^q (Z_p - \alpha Y_p^2)^{l-q} ((X_p - 2\alpha Y_p) U_{p-1})^q] + \\
&+ \lambda^{k+j+1} O_{\mu,D,k,j}((\sqrt{p}\log p)^{k+2j+1}) + O_{\mu,j}(e^{-c_j \log^2 p})
\end{aligned}$$

Further expanding and using Lemma 5.1 again:

$$\begin{aligned}
\mathbb{E}(\theta_p U_p^j) &= \mathbb{E}[\theta_{p-1}\left(\sum_{r=0}^{j} C_j^r U_{p-1}^r Y_p^{j-r}\right) e^{i(X_p t+Y_p s)} \sum_{l=0}^{k+j} \frac{(i\lambda)^l}{l!}\cdot \\
&\sum_{q=0}^{l} C_l^q (Z_p-\alpha Y_p^2)^{l-q}((X_p-2\alpha Y_p)U_{p-1})^q] + O_{\mu,D,k,j}(\frac{\log^{2k+3j+2} p}{p^{(k+1)/2}}) \\
&= \sum_{l=0}^{k+j}\sum_{q=0}^{l}\sum_{r=0}^{j} C_j^r C_l^q \frac{(i\lambda)^l}{l!}\mathbb{E}(\theta_{p-1}U_{p-1}^{r+q}) \\
&\cdot \mathbb{E}(e^{i(X_1 t+Y_1 s)} Y_1^{j-r}(Z_1-\alpha Y_1^2)^{l-q}(X_1-2\alpha Y_1)^q)+ \\
&+ O_{\mu,D,k,j}(\frac{\log^{2(k+1)+3j} p}{p^{(k+1)/2}})
\end{aligned}$$

The last expression can be written as a sum of $\beta(t,s)\mathbb{E}(\theta_{p-1}U_{p-1}^j)$ (where $\beta(t,s) = \mathbb{E}(e^{i(X_1 t+Y_1 s)})$) and a linear combination with bounded coefficients of a bounded number of terms (the bounds depend only on μ, D, k and j) of the form $\lambda^l \mathbb{E}(\theta_{p-1}U_{p-1}^m)$ with $0 \le m \le j+l$ and $0 \le l \le k+j$ and $(m,l) \ne (j,0)$, plus a remainder term.

We now fix ε, A and D and consider (t,s,λ) in the range defined in the statement of the proposition. In the rest of the proof, when the Landau notion O is used, we implicitly mean that the corresponding constant depends only on μ, ε, A, D, k and j.

Let $j \ge 0$ and $k+j \ge 0$. Now suppose $H_{m,k+1-2l}$ holds for all $0 \le m \le j+l$ and $0 \le l \le k+j$ except when $(m,l)=(j,0)$. We are going to show that it implies $H_{j,k+1}$. Since $|\lambda| \le D\frac{\log n}{n} \le D\frac{\log p}{p}$, we have for all these values of m and l:

$$\lambda^l \mathbb{E}(\theta_{p-1}U_{p-1}^m) = O(\frac{\log^{l+2(k+1-2l)+3m} p}{p^{(k+1)/2}}) = O(\frac{\log^{2(k+1)+3j} p}{p^{(k+1)/2}})$$

Hence,

$$\mathbb{E}(\theta_p U_p^j) = \beta(t,s)\mathbb{E}(\theta_{p-1}U_{p-1}^j) + O(\frac{\log^{2(k+1)+3j} p}{p^{(k+1)/2}})$$

Therefore, recursively on p, we obtain for all n

$$\begin{aligned}
(26)\quad \mathbb{E}(\theta_n U_n^j) &= \sum_{p=1}^{n} \beta(s,t)^{n-p} O(\frac{\log^{2(k+1)+3j} p}{p^{(k+1)/2}}) \\
&= n\beta(s,t)^{n/2} O(n^{|k|}) + (1+\ldots+\beta(t,s)^{n/2+1}) O(\frac{\log^{2(k+1)+3j} n}{n^{(k+1)/2}})
\end{aligned}$$

As above, since μ is aperiodic, it follows that the law of (X, Y) on $\mathbb{R}^2$ is aperiodic. Hence

$$\sup_{\varepsilon < s^2+t^2 < A} |\beta(t,s)| < 1$$

therefore (26) yields

$$\mathbb{E}(\theta_n U_n^j) = O(\frac{\log^{2(k+1)+3j} n}{n^{(k+1)/2}})$$

Thus we obtain $H_{j,k+1}$.

Now, we note that from Lemma 5.1, $H_{j,-j}$ holds for all $j \geq 0$. This also guarantees $H_{j,k}$ when $k + j \leq 0$. Then we proceed by induction on $h = j + k$. From the considerations above, we obtain $H_{j,k}$ from the knowledge of other $H_{j',k'}$ with $j' + k' < j + k$. So we are done.

□

Remark 5.3. *If we make the following additional assumption on* μ

$$\sup_{t^2+s^2 \geq 1} |\beta(t,s)| < 1$$

then Proposition (5.2) *holds uniformly in* A *and*

$$\mathbb{E}(\theta_n) = O(\frac{1}{n^k})$$

holds uniformly in (t, s, λ) *when* $\lambda \in [-D\frac{\log n}{n}, D\frac{\log n}{n}]$ *and* $t^2 + s^2 \geq \varepsilon$ *and the constant in* O *depends only* ε, k, D *and* μ*.*

5.3. **Estimate for intermediate values of s and t.** Now we will treat the case when $|\lambda| \leq D\frac{\log n}{n}$ and $D\frac{\log^8 n}{\sqrt{n}} \leq |(s,t)| \leq \varepsilon$ for some fixed (large) $D > 0$.

Proposition 5.4. *With the notation above, there is some* $\varepsilon = \varepsilon(\mu) > 0$*, such that uniformly in* λ*,* $|\lambda| \leq D\frac{\log n}{n}$

$$\int_{D\frac{\log^8 n}{\sqrt{n}} \leq |(s,t)| \leq \varepsilon} |\mathbb{E}(\theta_n(t,s,\lambda))| dt ds = O(\frac{1}{n \log^2 n})$$

where the constant in O *depends on* μ *and* D*.*

Proof. Let us write $r := |(t,s)| = \sqrt{t^2+s^2}$ and now choose $\varepsilon = \varepsilon(\mu) > 0$ so that for $|(t,s)| \leq \varepsilon$, $|\partial_1 \beta(t,s)| = O_\mu(r)$ and $|\partial_2 \beta(t,s)| = O_\mu(r)$. Choosing ε smaller if necessary, we can assume that when $|(t,s)| \leq \varepsilon$

$$\exists C = C(\mu) > 0, |\beta(t,s)| \leq e^{-Cr^2}$$

or equivalently,

$$\frac{1}{1 - |\beta(t,s)|} = O_\mu(\frac{1}{r^2})$$

We now suppose that $|\lambda| \leq D\frac{\log n}{n}$. Until the end of the proof, when we use Landau's notation O we implicitly mean that the underlying constant depends only on μ and D.

From the computation in the subsection above (with $j = 2$ and $k = -1$), we obtain that $\mathbb{E}(\theta_p U_p^2)$ is a sum of $\beta(t, s)\mathbb{E}(\theta_{p-1}U_{p-1}^2)$ and a linear combination with bounded coefficients of the terms $\lambda\mathbb{E}(\theta_{p-1}U_{p-1}^m)$ with $0 \leq m \leq 2$, $\mathbb{E}(\theta_{p-1})$, $\mathbb{E}(\theta_{p-1}U_{p-1})\partial_2\beta(t, s)$, and $\lambda\mathbb{E}(\theta_{p-1}U_{p-1}^3)(\partial_1\beta(t, s)+\partial_2\beta(t, s))$, with a remainder of order $O_{\mu,D}(\log^6 n)$. Hence, making use of Lemma 5.1, for all n and p, $p \leq n$, and if $D\frac{\log^2 n}{\sqrt{n}} \leq r \leq \varepsilon$, we have

$$\mathbb{E}(\theta_p U_p^2) = \beta(t, s)\mathbb{E}(\theta_{p-1}U_{p-1}^2) + rO(\sqrt{n}\log^4 n)$$

where O is independent of p and n. For any given p, $\frac{n}{2} \leq p \leq n$, we iterate this equation $\frac{n}{4}$ times. We obtain:

$$\begin{aligned}
\mathbb{E}(\theta_p U_p^2) &= \beta^{n/4}(s, t)O(n\log^2 n) + (1 + ... + \beta(t, s)^{n/4})rO(\sqrt{n}\log^4 n) \\
&= \frac{1}{1 - |\beta(t, s)|}rO(\sqrt{n}\log^4 n) \\
&= \frac{1}{r}O(\sqrt{n}\log^4 n) \qquad (27)
\end{aligned}$$

since $\beta^{n/4}(s, t)O(n\log^2 n) = e^{-Cnr^2/4}O(n\log^2 n) = \frac{1}{r}O(\sqrt{n}\log^4 n)$, because $D\frac{\log n}{\sqrt{n}} \leq r \leq \varepsilon$.

Similarly we can express $\mathbb{E}(\theta_p)$ as above (taking $j = 0$ and $k = 1$ in the proof of prop. 5.2) as a sum of $\beta(t, s)\mathbb{E}(\theta_{p-1})$ and a linear combination with bounded coefficients of the terms $\lambda\mathbb{E}(\theta_{p-1})$ and $\lambda\mathbb{E}(\theta_{p-1}U_{p-1})(\partial_1\beta + \partial_2\beta)$ with a rest of order $O(\frac{\log^4 n}{n})$. Since $D\frac{\log^2 n}{\sqrt{n}} \leq r \leq \varepsilon$ we have

$$\mathbb{E}(\theta_p) = \beta(t, s)\mathbb{E}(\theta_{p-1}) + rO(\frac{\log^2 n}{\sqrt{n}})$$

Iterating as above, we obtain for $\frac{n}{2} \leq p \leq n$

$$\mathbb{E}(\theta_p) = \frac{1}{r}O(\frac{\log^2 n}{\sqrt{n}}) \qquad (28)$$

Now we look at $\mathbb{E}(\theta_p U_p)$. As above (take $j = 1$, $k = 0$) it is a sum of $\beta(t, s)\mathbb{E}(\theta_{p-1}U_{p-1})$ and a linear combination with bounded coefficients of the terms $\partial_2\beta(t, s)\mathbb{E}(\theta_{p-1})$, $\lambda\mathbb{E}(\theta_{p-1})$, $\lambda\mathbb{E}(\theta_{p-1}U_{p-1})$, and $\lambda(\partial_1\beta + \partial_2\beta)\mathbb{E}(\theta_{p-1}U_{p-1}^2)$, with a rest of order $O(\frac{\log^5 n}{\sqrt{n}})$. But thanks to (28) and (27) all these terms are $O(\frac{\log^5 n}{\sqrt{n}})$ when $\frac{n}{2} \leq p \leq n$. Hence,

$$\mathbb{E}(\theta_p U_p) = \beta(t, s)\mathbb{E}(\theta_{p-1}U_{p-1}) + O(\frac{\log^5 n}{\sqrt{n}})$$

And then, iterating this relation, for all p, $\frac{3n}{4} \leq p \leq n$, we get

$$\mathbb{E}(\theta_p U_p) = \frac{1}{r^2} O(\frac{\log^5 n}{\sqrt{n}}) \tag{29}$$

Finally we again decompose $\mathbb{E}(\theta_p)$ but pushing one step further the Taylor expansion and we see by the above calculation (take $j = 0$, $k = 2$) that it is a sum of $\beta(t,s)\mathbb{E}(\theta_{p-1})$ and a linear combination with bounded coefficients of the terms $\lambda\mathbb{E}(\theta_{p-1})$, $\lambda\mathbb{E}(\theta_{p-1}U_{p-1})\partial_1\beta(t,s)$, $\lambda^2\mathbb{E}(\theta_{p-1})$, $\lambda^2\mathbb{E}(\theta_{p-1}U_{p-1})$ and $\lambda^2\mathbb{E}(\theta_{p-1}U_{p-1}^2)$ with a rest of order $O(\frac{\log^6 n}{n\sqrt{n}})$. Thanks to (27), (28), (29) and Lemma 5.1, they are all of order at most $\frac{1}{r}O(\frac{\log^6 n}{n\sqrt{n}})$ when $\frac{3n}{4} \leq p \leq n$. Thus

$$\mathbb{E}(\theta_p) = \beta(t,s)\mathbb{E}(\theta_{p-1}) + \frac{1}{r} O(\frac{\log^6 n}{n\sqrt{n}})$$

and consequently for large n

$$\mathbb{E}(\theta_n) = \frac{1}{r^3} O(\frac{\log^6 n}{n\sqrt{n}})$$

Then, integrating over r when $D\frac{\log^8 n}{\sqrt{n}} \leq r \leq \varepsilon$ we obtain:

$$\begin{aligned}
\int_{D\frac{\log^8 n}{\sqrt{n}} \leq |(s,t)| \leq \varepsilon} |\mathbb{E}(\theta_n)| dt ds &\leq O(\frac{\log^6 n}{n\sqrt{n}}) \int_{D\frac{\log^8 n}{\sqrt{n}} \leq r \leq \varepsilon} \frac{1}{r^3} r dr \\
&\leq O(\frac{1}{n \log^2 n})
\end{aligned}$$

This concludes the proof of the proposition. ☐

6. Study of a dynamical system

In this section we study a quadratic dynamical system in the complex plane and give precise estimates that will be crucial in the proof of the domination condition in the next section (i.e. Proposition 7.2).

We first fix three real numbers x, y, z satisfying the following condition

$$\Delta := \det \begin{pmatrix} 1 & x & z \\ x & 1 & y \\ z & y & 1 \end{pmatrix} = 1 + 2xyz - x^2 - y^2 - z^2 \geq 0$$

And suppose additionally that $|x| \leq 1$, $|y| \leq 1$ and $|z| < 1$. Now define the following three sequences recursively:

$$\begin{aligned} a_{k+1} &= a_k + \frac{1}{2} - 2\lambda^2 c_k^2 + 2i\lambda c_k y \\ b_{k+1} &= b_k + \frac{1}{2} - 2\lambda^2 b_k^2 + 2i\lambda b_k z \\ c_{k+1} &= c_k + \frac{x}{2} - 2\lambda^2 b_k c_k + i\lambda b_k y + i\lambda c_k z \end{aligned} \tag{30}$$

where the initial values a_0, b_0 and c_0 are arbitrary and λ is a given real number. In the applications below, λ will be small (of order $\log(n)/n$) and the above dynamical system can be viewed as a perturbation of that given when $\lambda = 0$. In this section we will study the behavior of the above three sequences depending on initial values and also on the values of the parameters x, y, z and λ.

We first note that (b_k) is a quadratic dynamical system and is therefore conjugate to $P_\lambda : u \mapsto u^2 + c_\lambda$ for some complex number c_λ. A straightforward computation shows that if we set

$$x_k := \frac{1}{2} + i\lambda z - 2\lambda^2 b_k$$

then we have $x_{k+1} = P_\lambda(x_k)$ where $c_\lambda = \frac{1}{4} - \lambda^2(1 - z^2)$. In the limit when λ tends to 0, then c_λ tends to $\frac{1}{4}$, which is on the boundary of the Mandelbrot set. As long as λ is small enough and non-zero, then the Fatou set corresponding to P_λ has exactly one bounded connected component. Moreover, as soon as $|\lambda| < 1$, for every starting point x_0 lying inside this component, the resulting sequence of iterates (x_k) will converge to the attracting fixed point x_λ given by

$$x_\lambda = \frac{1}{2} - |\lambda|\sqrt{1 - z^2}$$

Thus for $\lambda \neq 0$ and $|\lambda| < 1$, the sequence (b_k) converges to

$$b_\lambda = \frac{iz + sgn(\lambda)\sqrt{1 - z^2}}{2\lambda}$$

Now let us define

$$v_k := c_k + \frac{yz - x}{1 - z^2} b_k$$

Then it is easy to check directly from the equations (30) that v_k satisfies

$$v_{k+1} - v = (x_k + \frac{1}{2})(v_k - v) \tag{31}$$

where

$$v = \frac{i}{2\lambda} \frac{y - xz}{1 - z^2} \tag{32}$$

Let us write $y_k = \frac{1}{2} + x_k$, then we obtain

$$v_k - v = (v_0 - v) y_0 \cdot \ldots \cdot y_{k-1} \tag{33}$$

Similarly, if we let

$$f_k := a_k + \frac{yz - x}{1 - z^2} c_k$$

then we find that

$$f_{k+1} - f_k = \frac{\Delta}{2(1-z^2)} + (v_k - v)(1 - (\frac{1}{2} + x_k))\frac{yz - x}{1 - z^2} - 2\lambda^2 (v_k - v)^2$$

Making use of (31), it follows that for $k \geq 1$

$$\begin{aligned} a_k &= f_0 + \left(\frac{yz-x}{1-z^2}\right)^2 b_k + \frac{yz-x}{1-z^2}(v_0 - 2v_k) + \\ &\quad k\frac{\Delta}{2(1-z^2)} - 2\lambda^2 (v_0 - v)^2 \sum_{p=0}^{k-1} (y_0 \cdot \ldots \cdot y_{p-1})^2 \end{aligned} \tag{34}$$

We are now going to study the dynamical system (30) in the particular case when the initial values are defined by

$$\begin{aligned} a_0 &= 0 \\ b_0 &= \frac{i}{2\lambda} z \\ c_0 &= \frac{i}{2\lambda} w \end{aligned} \tag{35}$$

where w is a fixed real number. Then $x_0 = \frac{1}{2}$ belongs to the filled Julia set of P_λ, and the sequence (x_k) (hence (y_k) too) stays on the real line and satisfies

$$\frac{1}{2} - |\lambda|\sqrt{1 - z^2} = x_\lambda \leq x_k \leq \frac{1}{2}$$

Additionally,

$$v_0 = \frac{i}{2\lambda}(w + z\frac{yz - x}{1 - z^2})$$

Together with (31) and (32) this shows that v_k belongs to $i\mathbb{R}$ for all k. With these initial values, (34) takes the form

$$\begin{aligned} a_k &= \frac{1}{2\lambda}\left(\frac{yz-x}{1-z^2}\right)^2 (iz + \frac{1}{\lambda}(1 - y_k)) + \frac{i}{2\lambda}\left(\frac{yz-x}{1-z^2}\right)(2w - \frac{2y}{1-z^2} + z\frac{yz-x}{1-z^2}) - \\ &\quad -\frac{i}{\lambda}\left(\frac{yz-x}{1-z^2}\right)(w - y) y_0 \cdot \ldots \cdot y_{k-1} + \\ &\quad + k\frac{\Delta}{2(1-z^2)} + \frac{1}{2}(w-y)^2 \sum_{p=0}^{k-1} (y_0 \cdot \ldots \cdot y_{p-1})^2 \end{aligned}$$

taking the real part we get

$$Re(a_k) = \frac{1}{2|\lambda|}\left(\frac{yz-x}{1-z^2}\right)^2 \frac{1-y_k}{|\lambda|} + k\frac{\Delta}{2(1-z^2)} + \frac{1}{2}(w-y)^2 \sum_{p=0}^{k-1} (y_0 \cdot \ldots \cdot y_{p-1})^2 \tag{36}$$

The following lemma summarizes the computations above and encloses the information that will be relevant to the sequel:

Lemma 6.1. *In the dynamical system defined by* (30)*, with initial values given by* (35)*, the following holds.*

Take $|\lambda| \leq 1/2$*. Then for all* $k \geq 0$

(i) $|\lambda k| \leq 1/\sqrt{1-z^2}$ *implies* $1 - y_k \geq k\frac{1}{2}\lambda^2(1-z^2)$*, and* $Re(b_k) \geq k\frac{1-z^2}{4}$*, and*

$$Re(a_k) \geq \frac{k}{4}[\frac{(yz-x)^2}{1-z^2} + 2e^{-4}(w-y)^2]$$

(ii) $|\lambda k| \geq 1/\sqrt{1-z^2}$ *implies* $|\lambda|\sqrt{1-z^2} \geq 1 - y_k \geq \frac{1}{2}|\lambda|\sqrt{1-z^2}$*, and* $Re(b_k) \geq \frac{1}{4|\lambda|}\sqrt{1-z^2}$ *and*

$$Re(a_k) \geq \frac{1}{|\lambda|}\frac{1}{4\sqrt{1-z^2}}[\frac{(yz-x)^2}{1-z^2} + \frac{1}{4}(w-y)^2]$$

(iii) $|\lambda b_k| \leq \frac{1}{2}$ *and* $Re(b_k) \geq 0$

(iv) $|\lambda v_k| \leq 2/(1-z^2) + |w| + 1$

(v) $|\lambda c_k| \leq 3/(1-z^2) + |w| + 1$

(vi) $Re(a_k)Re(b_k) \geq Re(c_k)^2$

(vii) $Re(a_k)$ *is a non-decreasing sequence.*

Proof. All these points are easy to check from what was done above. The proof of (i) follows by induction ; it is true when $k = 0$ and, assuming the inequality for k, we get

$$\begin{aligned} 1 - y_{k+1} &= (1-y_k)y_k + \lambda^2(1-z^2) \\ &\geq k\frac{1}{2}\lambda^2(1-z^2)(1-|\lambda|\sqrt{1-z^2}) + \lambda^2(1-z^2) \\ &\geq k\frac{1}{2}\lambda^2(1-z^2) + \frac{1}{2}\lambda^2(1-z^2)(2-|k\lambda|\sqrt{1-z^2}) \\ &\geq (k+1)\frac{1}{2}\lambda^2(1-z^2) \end{aligned} \tag{37}$$

Besides, for all $p \leq k$ we have $y_p \geq 1 - |\lambda|\sqrt{1-z^2} \geq e^{-2|\lambda|\sqrt{1-z^2}}$ since $1 - t \geq e^{-2t}$ if $t \in [0, \frac{1}{2}]$. Hence

$$(y_0 \cdot \ldots \cdot y_{p-1})^2 \geq e^{-4|\lambda p|\sqrt{1-z^2}} \geq e^{-4}$$

The inequality for $Re(a_k)$ in (i) now follows instantly from (36) and the fact that $\Delta \geq 0$.

Point (ii) follows from the fact (granted by (37)) that $1 - y_k \geq c|\lambda|$ implies $1 - y_{k+1} \geq c|\lambda|$ for any real number c with $\sqrt{1-z^2} \geq c \geq 0$.

We have $2\lambda b_k = iz + (1 - y_k)/\lambda$ and $0 \leq 1 - y_k \leq |\lambda|\sqrt{1-z^2}$ so $|2\lambda b_k| \leq \sqrt{z^2 + (1 - z^2)} \leq 1$ yields (iii). Similarly $v_k - v = (v_0 - v)y_0 \cdot \ldots \cdot y_{k-1}$, hence

$$2\lambda v_k = i(y - xz)/(1 - z^2) + i(w - y)y_0 \cdot \ldots \cdot y_{k-1}$$

Since $|y_i| \leq 1$ for all i, we get $|2\lambda v_k| \leq |y - xz|/|1 - z^2| + |w - y|$ so $|2\lambda v_k| \leq 2/|1-z^2|+|w|+1$, and we have (iv) and also (v) because of (iii) and $v_k := c_k + \frac{yz-x}{1-z^2} b_k$.

For (vi), we have $Re(a_k) \geq \frac{1-y_k}{2\lambda^2}\left(\frac{yz-x}{1-z^2}\right)^2 = Re(b_k)\left(\frac{yz-x}{1-z^2}\right)^2$, so (recall that the v_k belong to $i\mathbb{R}$)

$$Re(a_k)Re(b_k) \geq \left(Re(b_k)\frac{yz-x}{1-z^2}\right)^2 = Re(c_k)^2$$

Finally (vii) is easily checked from (36). $\square$

7. Proof of the domination condition for small values of the parameters s, t, λ

In this section, we give a domination estimate for a particular type of trigonometric sum that will arise in the proof of local limit theorem. We then apply these estimates to treat the part of the integral that yields a contribution to the limit, that is when s, t, λ are small.

7.1. **Estimating a trigonometric sum.** Let us consider throughout this section a sequence of independent and identically distributed random variables $(A_k, B_k, C_k, D_k)_{k\geq 1}$ in $\mathbb{R}^4$. We assume that the distribution has compact support in $\mathbb{R}^4$, and that $\mathbb{E}(A_k) = \mathbb{E}(B_k) = \mathbb{E}(C_k) = 0$. Let us use the shorthand $\overline{X}$ to denote the expectation $\mathbb{E}(X)$ of a random variable X. We fix the following notations for the correlations

$$x = \frac{\overline{A_1B_1}}{\sqrt{\overline{A_1^2}\,\overline{B_1^2}}},\ y = \frac{\overline{A_1C_1}}{\sqrt{\overline{A_1^2}\,\overline{C_1^2}}},\ z = \frac{\overline{B_1C_1}}{\sqrt{\overline{B_1^2}\,\overline{C_1^2}}}$$

We also assume that A_1 is not identically 0 and that the distribution of the marginal (B_k, C_k) is not degenerate, i.e. is not supported on a line. This is equivalent to the condition $|z| < 1$.

Fix $w \in \mathbb{R}$ and then consider the trigonometric product for r, λ in $\mathbb{R}$,

(38)
$$\theta_n = \left(\prod_{1\leq k\leq n} e^{ir(A_k/\sqrt{\overline{A_1^2}} - wC_k/\sqrt{\overline{C_1^2}})} e^{i\lambda D_k}\right)\left(\prod_{1\leq p<q\leq n} e^{i\lambda B_q C_p/\sqrt{\overline{B_1^2}\,\overline{C_1^2}}}\right) e^{-i\frac{\lambda}{2}z(C_1+\ldots+C_n)^2/\overline{C_1^2}}$$

The following proposition yields the desired estimate for the trigonometric sum $\mathbb{E}(\theta_n)$.

Proposition 7.1. *Let us fix $D > 0$ a positive number and $m \geq 1$ an integer. For any integer $n \geq 1$ and any distribution (A_1, B_1, C_1, D_1), of compact support and as described above, the following estimate holds uniformly when r varies in $[-D\frac{\log^{2m} n}{\sqrt{n}}, D\frac{\log^{2m} n}{\sqrt{n}}]$ and λ varies in $[-D\frac{\log n}{n}, D\frac{\log n}{n}]\backslash[-\frac{2}{n\sqrt{1-z^2}}, \frac{2}{n\sqrt{1-z^2}}]$,*

$$|\mathbb{E}(\theta_n)| \leq \exp(-\left[n|\lambda|\frac{\sqrt{1-z^2}}{8} + \frac{r^2}{|\lambda|}\frac{C}{16\sqrt{1-z^2}}\right]) + \left[\frac{1+|w|}{1-z^2}\right]^4 O(\frac{\log^{6m} n}{\sqrt{n}})$$

where $C = \left((w-y)^2 + \frac{(yz-x)^2}{1-z^2}\right)$.

Similarly, if r varies in $[-D\frac{\log^{2m} n}{\sqrt{n}}, D\frac{\log^{2m} n}{\sqrt{n}}]$ and λ varies in $[-\frac{2}{n\sqrt{1-z^2}}, \frac{2}{n\sqrt{1-z^2}}]$, we obtain

$$|\mathbb{E}(\theta_n)| \leq \exp[-nr^2\frac{Ce^{-4}}{4}] + \left[\frac{1+|w|}{1-z^2}\right]^4 O(\frac{\log^{6m} n}{\sqrt{n}})$$

The constant in $O(\cdot)$ depends only on D and on the size of the distribution

$$M = \max\{|A_1|/\sqrt{\overline{A_1^2}}, |B_1|/\sqrt{\overline{B_1^2}}, |C_1|/\sqrt{\overline{C_1^2}}, |D_1|\}$$

Proof. Let $n \in \mathbb{N}$ and $r, \lambda \in \mathbb{R}$ be as in the statement of the proposition. Let $U_0 = 0$ and for $k \geq 1$,

$$U_k = (C_1 + ... + C_k)/\sqrt{\overline{C^2}}$$

We first set a few notations. Let q_k be the quadratic form

$$q_k(u, v) = a_k u^2 + b_k v^2 + 2c_k uv$$

where the coefficients are defined right below. Then let $\pi_0 = 1$ and for $k \geq 1$

$$\pi_k = e^{ir(A_1+...+A_k)/\sqrt{\overline{A_1^2}}} \left(\prod_{1 \leq p < q \leq k} e^{i\lambda B_q C_p/\sqrt{\overline{B_1^2 C_1^2}}}\right) e^{i\lambda(D_1+...+D_k)}$$

And finally set

$$P_k(r, \lambda) = \mathbb{E}(\pi_{n-k} e^{-q_k(r, \lambda U_{n-k})})$$

We shall define the coefficients a_k, b_k, c_k recursively as follows. We set:

$$\begin{aligned} a_{k+1} &= a_k + \frac{1}{2} - 2\lambda^2 c_k^2 + 2i\lambda c_k y \\ b_{k+1} &= b_k + \frac{1}{2} - 2\lambda^2 b_k^2 + 2i\lambda b_k z \\ c_{k+1} &= c_k + \frac{x}{2} - 2\lambda^2 b_k c_k + i\lambda b_k y + i\lambda c_k z \end{aligned} \tag{39}$$

and the initial values are taken to be: $a_0 = 0$, $c_0 = \frac{i}{2\lambda} w$ (recall that $w \in \mathbb{R}$), and $b_0 = \frac{i}{2\lambda} z$.

These relations are precisely the recurrence relations (30) defined in the last section. As was discussed there, Lemma 6.1 applies and, as soon as $|\lambda| \leq D\frac{\log n}{n}$ and $|r| \leq$

$D\frac{\log^{2m} n}{\sqrt{n}}$, we see that λb_k and λc_k are uniformly bounded (by $const/(1-z^2)+|w|+1$) and that $Re(a_k)Re(b_k) \geq (Re(c_k))^2$. Also note that $P_0 = \mathbb{E}(\pi_n e^{-q_0(r,\lambda U_n)}) = \mathbb{E}(\pi_n e^{-\frac{i}{2}z\lambda U_n^2 - riwU_n})$, so that

$$P_0(r,\lambda) = \mathbb{E}(\theta_n).$$

And

$$P_n = \mathbb{E}(e^{-q_n(r,0)}) = e^{-a_n r^2}.$$

In order to obtain the desired estimate on $\mathbb{E}(\theta_n)$, that is on $P_0(r,\lambda)$, we are going to start from $P_0(r,\lambda)$ and proceed by induction on k. To achieve this goal, we need the following key induction step:

Claim: For all k and n with $k \leq n-1$, and if $|\lambda| \leq D\frac{\log n}{n}$ and $|r| \leq D\frac{\log^{2m} n}{\sqrt{n}}$, we have

$$P_k(t,s) = e^{-b_k\lambda^2 + i\overline{D}\lambda} P_{k+1}(t,s) + \left[\frac{1+|w|}{1-z^2}\right]^4 O(\frac{\log^{6m} n}{n\sqrt{n}}) \tag{40}$$

where the constant in O depends only on D and on the size M of the support of the distribution $(A_1, ..., D_1)$.

This crucial estimate follows from the computation made below. First, note that

$$\begin{aligned} q_k(r,\lambda U_{n-k}) &= q_k(r,\lambda U_{n-k-1} + \lambda C_{n-k}/\sqrt{\overline{C^2}}) \\ &= q_k(r,\lambda U_{n-k-1}) + b_k\lambda^2 C_{n-k}^2/\overline{C^2} + \\ &\quad 2b_k\lambda^2 U_{n-k-1}C_{n-k}/\sqrt{\overline{C^2}} + 2c_k\lambda r C_{n-k}/\sqrt{\overline{C^2}} \end{aligned}$$

We also recall that $\mathcal{A}_n$ is the event $\{\max_{0\leq k\leq n}|U_k| \leq \sqrt{n}\log n\}$.

We will use the expansion $e^t = 1 + t + t^2/2 + O(t^3)$ near $t=0$. Applying Lemma 5.1 and bearing in mind that μ is centered we can write

$$\begin{aligned} P_k &= \mathbb{E}\left(\pi_{n-k-1} e^{irA_{n-k}/\sqrt{\overline{A^2}}} e^{i\lambda B_{n-k}U_{n-k-1}/\sqrt{\overline{B^2}}} e^{i\lambda D_{n-k}} e^{-q_k(r,\lambda U_{n-k})}\right) \\ &= \mathbb{E}(\mathbf{1}_{\mathcal{A}_n}\pi_{n-k-1} e^{-q_k(r,\lambda U_{n-k-1})} e^{irA_{n-k}/\sqrt{\overline{A^2}}} e^{i\lambda B_{n-k}U_{n-k-1}/\sqrt{\overline{B^2}}} \\ &\quad e^{i\lambda D_{n-k}} e^{-b_k\lambda^2 C_{n-k}^2/\overline{C^2} - 2b_k\lambda^2 U_{n-k-1}C_{n-k}/\sqrt{\overline{C^2}} - 2c_k\lambda r C_{n-k}/\sqrt{\overline{C^2}}}) + e^{-c_0\log^2 n} \\ &= \mathbb{E}(\mathbf{1}_{\mathcal{A}_n}\pi_{n-k-1} e^{-q_k(r,\lambda U_{n-k-1})}(1 + irA_{n-k}/\sqrt{\overline{A^2}} + i\lambda B_{n-k}U_{n-k-1}/\sqrt{\overline{B^2}} + \\ &\quad i\lambda D_{n-k} - b_k\lambda^2 C_{n-k}^2/\overline{C^2} - 2b_k\lambda^2 U_{n-k-1}C_{n-k}/\sqrt{\overline{C^2}} - 2c_k\lambda r C_{n-k}/\sqrt{\overline{C^2}} + \\ &\quad \frac{1}{2}\left[\begin{array}{c} irA_{n-k}/\sqrt{\overline{A^2}} + i\lambda B_{n-k}U_{n-k-1}/\sqrt{\overline{B^2}} + i\lambda D_{n-k} \\ -b_k\lambda^2 C_{n-k}^2/\overline{C^2} - 2b_k\lambda^2 U_{n-k-1}C_{n-k}/\sqrt{\overline{C^2}} - 2c_k\lambda r C_{n-k}/\sqrt{\overline{C^2}} \end{array}\right]^2 \\ &\quad + O(\frac{\log^{6m} n}{n\sqrt{n}}))) + e^{-c_0\log^2 n} \end{aligned}$$

Expanding further, we get

$$\begin{aligned}
P_k &= \mathbb{E}(\pi_{n-k-1}e^{-q_k(r,\lambda U_{n-k-1})}[1-b_k\lambda^2+i\overline{D}\lambda-r^2(\frac{1}{2}-2\lambda^2c_k^2+2i\lambda c_k y)- \\
&\quad \lambda^2U_{n-k-1}^2(\frac{1}{2}-2\lambda^2b_k^2+2i\lambda b_k z)-2\lambda rU_{n-k-1}(\frac{x}{2}-2\lambda^2b_kc_k+i\lambda b_k y+i\lambda c_k z)] \\
&\quad +O(\frac{\log^{6m} n}{n\sqrt{n}})) \\
&= \mathbb{E}(\pi_{n-k-1}e^{-q_k(r,\lambda U_{n-k-1})}[exp\big(-b_k\lambda^2+i\overline{D}\lambda-a_{k+1}r^2-b_{k+1}\lambda^2U_{n-k-1}^2-2c_{k+1}\lambda rU_{n-k-1}\big) \\
&\quad +O(\frac{\log^{6m} n}{n\sqrt{n}})]+O(\frac{\log^{6m} n}{n\sqrt{n}})
\end{aligned}$$

where the coefficients a_{k+1}, b_{k+1} and c_{k+1} are the ones defined above. Hence we indeed obtain

$$P_k = e^{-b_k\lambda^2+i\overline{D}\lambda}P_{k+1}+O(\frac{\log^{6m} n}{n\sqrt{n}})$$

This computation makes sense as long as λb_k and λc_k remain uniformly bounded when n grows, and we checked that it is indeed the case. We also need to insure that $Re(q_k)\geq 0$ everywhere, i.e. $Re(a_k)Re(b_k)\geq (Re(c_k))^2$ and $Re(a_k)\geq 0$, but we also checked that above.

With the help of Lemma 5.1 and the remark following it, we verify that the constant involved in the O in the above calculations can be taken of the form $c\cdot D^4\cdot K^4$, where $c>0$ depends only on the size of the support of the distribution of (A_1,B_1,C_1,D_1) and where K is that number that bounds λc_k and λb_k. Hence the claim (40) is proved.

We can now iterate (40), going from P_0 to P_n. We deduce

$$\mathbb{E}(\theta_n)=e^{-\lambda^2\sum_{k=0}^{n-1}b_k}e^{i\lambda n\overline{D}}e^{-a_nr^2}+O(\frac{\log^{6m} n}{\sqrt{n}})$$

The constant involved here in the O is bounded by some $cD^4(1+|w|)^4/(1-z^2)^4$ where c is a constant depending only on $M=\max\{|A_1|/\sqrt{\overline{A_1^2}},|B_1|/\sqrt{\overline{B_1^2}},|C_1|/\sqrt{\overline{C_1^2}},|D_1|\}$.

$$|\mathbb{E}(\theta_n)|\leq e^{-\lambda^2\sum_{k=0}^{n-1}Reb_k}e^{-r^2Rea_n}+O(\frac{\log^{6m} n}{\sqrt{n}})$$

From Lemma 6.1, if $|\lambda|\geq\frac{2}{\sqrt{1-z^2}}\frac{1}{n}$ then $Reb_k\geq\frac{1}{4|\lambda|}\sqrt{1-z^2}$ for all $k\geq n/2$ and $Re(b_k)\geq 0$ for all k. So the first factor in the above equation leads to the bound $e^{-n|\lambda|\sqrt{1-z^2}/8}$. While $Re(a_n)\geq\frac{1}{|\lambda|}\frac{1}{4(1-z^2)^{1/2}}\left(\frac{1}{4}(w-y)^2+\frac{(yz-x)^2}{1-z^2}\right)$.

If $|\lambda|\leq\frac{2}{\sqrt{1-z^2}}\frac{1}{n}$, then $Re(a_n)\geq Re(a_{n/2})\geq\frac{n}{4}e^{-4}\left((w-y)^2+\frac{(yz-x)^2}{1-z^2}\right)$.

So we obtain the desired inequalities.

□

7.2. **Domination condition.** The purpose of this subsection is to give the precise estimate we wanted in the course of the proof of the local limit theorem (control of the part of the integral where all parameters s, t, λ are small). This is explained is the following proposition:

Proposition 7.2. *Let μ be a probability measure on the Heisenberg group with the properties described in the introduction. Let θ_n be the random variable defined in (19) at the beginning of section 5. There exist positive numbers $c_1 > 0$ and $c_2 > 0$ depending only on μ, such that, if $D > 0$ denotes some positive number, then the following estimates hold uniformly*

(i) when s and t vary in $[-D\frac{\log^8 n}{\sqrt{n}}, D\frac{\log^8 n}{\sqrt{n}}]$ and λ varies in $[-D\frac{\log n}{n}, D\frac{\log n}{n}]\backslash[-\frac{c_2}{n}, \frac{c_2}{n}]$,

$$|\mathbb{E}(\theta_n)| \leq \exp(-c_1 \left[n|\lambda| + \frac{t^2+s^2}{|\lambda|}\right]) + O(\frac{\log^{24} n}{\sqrt{n}})$$

(ii) when s and t vary in $[-D\frac{\log^8 n}{\sqrt{n}}, D\frac{\log^8 n}{\sqrt{n}}]$ and λ varies in $[-\frac{c_2}{n}, \frac{c_2}{n}]$,

$$|\mathbb{E}(\theta_n)| \leq \exp(-c_1 n(t^2+s^2)) + O(\frac{\log^{24} n}{\sqrt{n}})$$

where the constant in O depends only on μ and on D.

Proof. We are going to apply the results of the last section. In order to do so, we need to choose carefully the variables A_1, B_1, C_1, D_1 as well as the coefficient w. Take $D_1 = Z/\sqrt{\overline{X^2} \cdot \overline{Y^2}}$, $C_1 = Y/\sqrt{\overline{Y^2}}$, $B_1 = X/\sqrt{\overline{X^2}}$ and $A_1 = \cos(\theta)\frac{X}{\sqrt{\overline{X^2}}} + 2\sin(\theta)\frac{Y}{\sqrt{\overline{Y^2}}}$ and $w = \frac{\sin(\theta)}{\alpha(\theta)}$, where $\alpha(\theta)^2 = \overline{A_1^2} = 1 + 3\sin^2(\theta) + 4z\sin(\theta)\cos(\theta)$ and $\alpha(\theta) \geq 0$. An easy computation shows that $8 \geq \alpha(\theta)^2 \geq (1-z^2)/5$. Clearly the variables A_1, B_1 and C_1 are linearly dependent, hence $\Delta = 0$. But B_1 and C_1 are linearly independent since μ is aperiodic, hence $|z| < 1$.

Then the two expressions θ_n in (19) and (38) agree if we change λ into $\lambda/\sqrt{\overline{X^2 Y^2}}$ and let r and θ be determined by

$$\begin{aligned} t &= \frac{r\cos(\theta)}{\alpha(\theta)\sqrt{\overline{X^2}}} \\ s &= \frac{r\sin(\theta)}{\alpha(\theta)\sqrt{\overline{Y^2}}} \end{aligned}$$

From the above inequality on $\alpha(\theta)$, we conclude that if $M_{X,Y} = \max\{\frac{1}{\overline{X^2}}, \frac{1}{\overline{Y^2}}\}$ and $m_{X,Y} = \min\{\frac{1}{\overline{X^2}}, \frac{1}{\overline{Y^2}}\}$,

$$m_{X,Y}\frac{1-z^2}{5}(t^2+s^2) \leq \frac{r^2}{\overline{X^2 Y^2}} \leq 8M_{X,Y}(t^2+s^2)$$

To compute the constant C appearing in Proposition (7.1), let us first compute x, y and z.

$$\begin{aligned} z &= \overline{XY}/\sqrt{\overline{X^2}\cdot\overline{Y^2}} \\ y &= \frac{z\cos(\theta)+2\sin(\theta)}{\alpha(\theta)} \\ x &= \frac{\cos(\theta)+2z\sin(\theta)}{\alpha(\theta)} \end{aligned}$$

hence

$$\begin{aligned} C &= (w-y)^2 + \frac{(yz-x)^2}{1-z^2} \\ &= \left(\frac{\sin(\theta)+z\cos(\theta)}{\alpha(\theta)}\right)^2 + \frac{(1-z^2)\cos^2(\theta)}{\alpha(\theta)^2} \\ &= \frac{1}{\alpha(\theta)^2}\left[1-2z\cos(\theta)\sin(\theta)\right] \\ &\geq \frac{1}{\alpha(\theta)^2}(1-|z|) \geq \frac{1-z^2}{2\cdot 8} \end{aligned}$$

Also note that $|w| \leq 1/\alpha(\theta) \leq \sqrt{5/(1-z^2)}$. The first estimate in (7.1) now yields

$$\begin{aligned} |\mathbb{E}(\theta_n)| &\leq \exp(-\left[n|\lambda|\sqrt{\overline{X^2}\cdot\overline{Y^2}}\frac{\sqrt{1-z^2}}{8} + \frac{r^2}{|\lambda|}\frac{C}{4\sqrt{1-z^2}}\right]) + O(\frac{\log^6 n}{\sqrt{n}}) \\ &\leq \exp(-\sqrt{\overline{X^2}\cdot\overline{Y^2}}\sqrt{1-z^2}\left[\frac{n|\lambda|}{8} + \frac{r^2}{|\lambda|\overline{X^2}\cdot\overline{Y^2}}\frac{1}{64}\right]) + O(\frac{\log^6 n}{\sqrt{n}}) \\ &\leq \exp(-\sqrt{\overline{X^2}\cdot\overline{Y^2}-\overline{XY}^2}\left[\frac{n|\lambda|}{8} + \frac{M_{X,Y}(t^2+s^2)}{64\cdot|\lambda|}\right]) + O(\frac{\log^6 n}{\sqrt{n}}) \\ &\leq \exp(-c_1\left[n|\lambda| + \frac{(t^2+s^2)}{|\lambda|}\right]) + O(\frac{\log^6 n}{\sqrt{n}}) \end{aligned}$$

where we can take $c_1 \leq \frac{\sqrt{\overline{X^2}\cdot\overline{Y^2}-\overline{XY}^2}}{8}\min\{1, M_{X,Y}/8\}$, and the constant in O in the last line depends only on μ and D.

We thus have obtained (i), and (ii) follows similarly with $c_1 \leq \frac{e^{-4}}{80}(\overline{X^2}\cdot\overline{Y^2} - \overline{XY}^2)$. □

8. Proofs of the main theorems

We can now finish the proof of Theorem 2.1. Let f and g be as in the statement of Theorem 2.1. Recall that the number α was defined in (5). As was remarked at the end of section 3, we can write

$$n^2 \langle \rho(\mu^n) f, g\rangle = A_n + B_n$$

where A_n and B_n are defined as follows:

$$A_n = n^2 \int_{|\lambda| \geq D\frac{\log n}{n}} \langle \rho_\lambda(\mu^n) f_\lambda, g_\lambda\rangle_{\mathcal{H}_\lambda} \, d\lambda$$

and there is a constant C depending on D, μ and the size of the set $\{\lambda, \exists (t,s), F_g(g)(s,t,\lambda) \neq 0\}$ such that

$$|A_n| \leq \frac{C}{n^{k_0-2}} \|f\|_{L^2(G)} \|g\|_{L^2(G)} \tag{41}$$

as soon as the integer k_0 satisfies $D > k_0/c$ (where $c > 0$ is a constant depending on μ only defined in (13)). And

$$B_n = \int_G \frac{d\mathbf{x}}{(2\pi)^{3/2}} f(\mathbf{x}) I_n(\mathbf{x}) \tag{42}$$

with

$$I_n(\mathbf{x}) = n^2 \int_{|\lambda| \leq D\frac{\log n}{n}} \left\langle \mathbb{E}(\theta_n(t,s,\lambda)), e^{-i\phi} F_\sigma(g)(t-\lambda y, s + 2\alpha\lambda y, \lambda)\right\rangle_{\mathbb{L}^2(\mathbb{R}^2)} d\lambda$$

where, keeping the notations of Theorem 2.1 $F_\sigma(g)$ is the Fourier transform defined in (4), $\theta_n(t,s,\lambda)$ is defined in (19) and $\mathbf{x} = (x,y,z) \in G$.

Splitting the integral on $\mathbb{R}^2$ in the expression of $I_n(\mathbf{x})$ above into the parts when $|(t,s)| \leq D\frac{\log^8 n}{\sqrt{n}}$ on the one hand and $|(t,s)| \geq D\frac{\log^8 n}{\sqrt{n}}$ on the other hand, we can write:

$$I_n(\mathbf{x}) = I_n^S(\mathbf{x}) + I_n^L(\mathbf{x}) \tag{43}$$

and assert that if $\|\mathbf{x}\| = \max\{|x|,|y|,|z|\}$ and $2|\alpha|$ are less than say $K > 1$ then $F_\sigma(g)(t-\lambda y, s+2\alpha\lambda y, \lambda) \neq 0$ implies that $\max\{|t|,|s|,|\lambda|\} \leq A(1+K^2)$ where $A > 0$ is a number such that the support of $F_\sigma(g)$ lies inside $[-A,A]^3$. Then we may write:

$$\begin{aligned}
|I_n^L(\mathbf{x})| &\leq n^2 \|F_\sigma(g)\|_\infty \int_{|\lambda| \leq D\frac{\log n}{n}} \int_{D\frac{\log^8 n}{\sqrt{n}} \leq |(t,s)| \leq A(1+K^2)} |\mathbb{E}(\theta_n(t,s,\lambda))| dt ds d\lambda \\
&\leq n^2 \|F_\sigma(g)\|_\infty \int_{|\lambda| \leq D\frac{\log n}{n}} O(\frac{1}{n\log^2(n)}) d\lambda \qquad (44) \\
&\leq \|F_\sigma(g)\|_\infty O(\frac{1}{\log(n)})
\end{aligned}$$

where line 44 is granted by the two Propositions 5.2 and 5.4. The constant involved here in O depends only on μ, D and the size A of the support of $F_\sigma(g)$ and on the maximum of $\|\mathbf{x}\|$. In particular it is uniform when $\mathbf{x}$ varies in compact subsets of G.

We can now concentrate on the part $I_n^S(x)$ of the integral which actually gives a contribution to the limit. From Proposition 7.2, we deduce that if $|(t,s)| \leq D\log^8 n$ and $c_2 \leq |\lambda| \leq D\log n$

$$|\mathbb{E}(\theta_n(\frac{t}{\sqrt{n}}, \frac{s}{\sqrt{n}}, \frac{\lambda}{n}))| \leq e^{-c_1(|\lambda|+t^2/|\lambda|+s^2/|\lambda|)} + O(\frac{\log^{24} n}{\sqrt{n}})$$

and if $|(t,s)| \leq D\log^8 n$ and $|\lambda| \leq c_2$

$$|\mathbb{E}(\theta_n(\frac{t}{\sqrt{n}}, \frac{s}{\sqrt{n}}, \frac{\lambda}{n}))| \leq e^{-c_1(t^2+s^2)} + O(\frac{\log^{24} n}{\sqrt{n}})$$

where the constant in O depends only on μ and D. But one can check that the function

$$(t,s,\lambda) \mapsto e^{-c(|\lambda|+t^2/|\lambda|+s^2/|\lambda|)}$$

is integrable over $\mathbb{R}^3$. And

$$(t,s,\lambda) \mapsto e^{-c(t^2+s^2)}$$

is integrable in $(t,s,\lambda) \in \mathbb{R}^2 \times [-c_2, c_2]$. Moreover

$$\int_{|\lambda|\leq D\log n} \int_{|(t,s)|\leq D\log^8 n} O(\frac{\log^{24} n}{\sqrt{n}}) = O(\frac{\log^{41} n}{\sqrt{n}}) \to 0$$

And finally, from the central limit theorem (see [Tut] or [Rau]) the following limit holds point-wise in t, s, λ :

$$\mathbb{E}(\theta_n(\frac{t}{\sqrt{n}}, \frac{s}{\sqrt{n}}, \frac{\lambda}{n})) \to \mathbb{E}(e^{i(tX+sY+\lambda(Z-\alpha Y^2))})$$

where (X,Y,Z) is the limit random variable with gaussian distribution ν_1 as mentioned in the introduction. Hence if we use the shorthand $t_n := t/\sqrt{n}$, $s_n = s/\sqrt{n}$ and $\lambda_n := \lambda/n$ we have (see the definition of $\phi_n = \phi(\frac{t}{\sqrt{n}}, \frac{s}{\sqrt{n}}, \frac{\lambda}{n})$ in (18) above), uniformly when $\mathbf{x}$ varies in compact subsets of G

$$\lim_{n\to+\infty} \mathbb{E}(\theta_n(t_n, s_n, \lambda_n))e^{i\phi_n}\overline{F_\sigma(g)}(t_n - \lambda_n y, s_n + 2\alpha\lambda_n y, \lambda_n) = \mathbb{E}(e^{i(tX+sY+\lambda(Z-\alpha Y^2))})\overline{F_\sigma(g)}(0)$$

Since the heat kernel corresponding to μ is a fastly decreasing smooth function $p(x,y,z)$ on $\mathbb{R}^3$ (see [VSC]) it follows that $\mathbb{E}(e^{i(tX+sY+\lambda(Z-\alpha Y^2))})$ is integrable in $(t,s,\lambda) \in \mathbb{R}^3$ and we compute by the Fourier inversion formula

$$\int_{\mathbb{R}^3} \mathbb{E}(e^{i(tX+sY+\lambda(Z-\alpha Y^2))})dtdsd\lambda = (2\pi)^3 p(0,0,0) = (2\pi)^3 c(\mu)$$

Therefore, by Lebesgue's dominated convergence theorem, we get uniformly when $\mathbf{x}$ varies in compact subsets

$$\begin{aligned}\lim_{n\to+\infty} I_n(\mathbf{x}) &= \lim_{n\to+\infty} I_n^S(\mathbf{x})\\ &= n^2\int_{|\lambda|\le D\frac{\log n}{n}}\int_{|(t,s)|\le D\frac{\log^8 n}{\sqrt{n}}} \mathbb{E}(\theta_n(t,s,\lambda))e^{i\phi}\overline{F_\sigma(g)}(t-\lambda y, s+2\alpha\lambda y,\lambda)dtdsd\lambda\\ &= (2\pi)^3 c(\mu)\overline{F_\sigma(g)}(0)\end{aligned}$$

Integrating in $\mathbf{x}$ we finally obtain

$$\begin{aligned}\lim_{n\to\infty} B_n &= \lim_{n\to+\infty}\int_G \frac{d\mathbf{x}}{(2\pi)^{3/2}} f(\mathbf{x})I_n(\mathbf{x})\\ &= (2\pi)^{3/2}c(\mu)\overline{F_\sigma(g)}(0)\int_G f\\ &= c(\mu)\int_G f\int_G \overline{g} \qquad (45)\end{aligned}$$

Combining (45) and (41) we get

$$\lim_{n\to\infty} n^2\langle\rho(\mu^n)f,g\rangle = c(\mu)\int_G f\int_G\overline{g} \qquad (46)$$

as desired.

We also remark that any translation of g (to the left ${}_zg(x) = g(z^{-1}x)$ or to the right $g_z(x) = g(xz)$) has again a compactly supported Fourier transform $F_\sigma({}_zg)$ and $F_\sigma(g_z)$. If z remains in a compact subset, then the supports of $F_\sigma({}_zg)$ and $F_\sigma(g_z)$ also remain within a prescribed bounded set. Moreover $\|F_\sigma(g_z)\|_\infty = \|F_\sigma({}_zg)\|_\infty$ and $\|g_z\|_2 = \|{}_zg\|_2$ are independent of z. This follows from the computation of $F_\sigma({}_zg)$ which yields

$$F_\sigma({}_zg)(t,s,\lambda) = e^{i\tau}\cdot F_\sigma(g)(t, s-\lambda(2\alpha z_y - z_x),\lambda) \qquad (47)$$

where $z = (z_x, z_y, z_z)$ and $\tau = z_xt+z_ys+\lambda(z_z-\alpha z_y^2)$. Consequently all the calculations above, and (46) in particular, hold uniformly for translates of g on compact subsets.

Now take f_n a Dirac sequence of positive functions supported on a neighborhood of 0 of diameter of order $1/n^3$. Then $\|f_n\|_2^2 = O(n^9)$. Choosing D large enough (so that A_n in (41) remains negligeable) we see that we can replace f by f_n in the above calculations, hence

$$\lim_{n\to\infty} n^2\langle\rho(\mu^n)f_n,g\rangle = c(\mu)\int_G\overline{g} \qquad (48)$$

And the same holds uniformly when z varies in compact subsets and g is replaced by g_z or ${}_zg$.

Assume now that $F_\sigma(g)$ is also absolutely continuous (that is belongs to $W^{1,1}(\mathbb{R}^3)$, i.e. its derivative distribution is a function in $\mathbb{L}^1(\mathbb{R}^3)$), then it follows[1] that $g \circ \phi_\sigma$ is real analytic and that $\|x\| \, d_x(g \circ \phi_\sigma)$ is a bounded function on $\mathbb{R}^3$ (where $\|x\|$ is the max of the coordinates of $x \in \mathbb{R}^3$). For any compact K we can then find a constant $C = C(g, K)$ such that whenever y is small enough

$$\sup_{z \in K} \|_{y^{-1}z} g -_z g\|_\infty \leq C \|y\|$$

and

$$\sup_{z \in G} \|_{y^{-1}} g_z - g_z\|_\infty \leq C \|y\|$$

Hence uniformly for z in compact subsets, $\|_z g * f_n -_z g\|_\infty = O(1/n^3)$ and $\|g_z * f_n - g_z\|_\infty = O(1/n^3)$. From (48) it now follows that uniformly for z in compact subsets:

$$\lim_{n \to \infty} n^2 \int g(z^{-1}x) d\mu^n(x) = c(\mu) \int_G g$$

and

$$\lim_{n \to \infty} n^2 \int g(xz) d\mu^n(x) = c(\mu) \int_G g \tag{49}$$

We now deduce Theorem 1.1 from Theorem 2.1.

Proof. For all $\varepsilon > 0$ one can find[2] a strictly positive function h in $\mathbb{L}^1(\mathbb{R}^3)$ such that $\widehat{h}$ is C^1 and compactly supported, and such that there exists $C > 0$, with $h(\mathbf{x}) \|\mathbf{x}\|^{6+\varepsilon} \geq C$ for $\mathbf{x} \in \mathbb{R}^3$ large enough. Let $g = h \circ \phi_\sigma^{-1}$. Now for all $\xi \in \mathbb{R}^3$ and $z \in G$, $e^{i\xi \cdot \phi_\sigma^{-1}(x)} g_z(x)$ also has C^1 Fourier transform F_σ of compact support, hence (49) holds for it uniformly when z varies in compact subsets. By Lévy's criterion for weak convergence of finite measures, this shows that the sequence of finite measures $d\nu_n^z = n^2 g_z d\mu^n$ converges weakly (in the space of finite measures on G) to $c(\mu) g_z(x) dx$ uniformly when z varies in compact subsets. Now let f be a function on G as in the statement of the theorem. Then f/g is a bounded continuous function, hence

$$n^2 \int f_z d\mu^n = \int f_z / g_z d\nu_n^z \to c(\mu) \int_G f$$

uniformly when z varies in compact subsets. $\square$

---

[1]Note that if a function $h(x) \in \mathbb{L}^1(\mathbb{R})$ is such that $\widehat{h}(t) \in C_c(\mathbb{R})$ and $\frac{d\widehat{h}}{dt} \in \mathbb{L}^1(\mathbb{R})$ (i.e. $\widehat{h}$ absolutely continuous) then $x\frac{dh}{dx}$ is the Fourier transform of $\frac{d}{dt}(t\widehat{h})$, hence is bounded.

[2]It is enough to find a function $f \in \mathbb{L}^1(\mathbb{R})$ such that $\widehat{f}$ is C^1 of compact support and $f(x) \geq \frac{C}{1+|x|^{2+\varepsilon}}$ for some $C > 0$ (e.g. see [Breu] section 3.2). Then take $h(\mathbf{x}) = f(x)f(y)f(z)$ if $\mathbf{x} = (x, y, z)$.

9. Uniform local limit theorem for translates of a bounded set

We intend here to prove Theorem 1.2. The probability measure μ is assumed to be aperiodic, centered and compactly supported. The letter ν denotes the associated gaussian probability distribution. Its density function, the heat kernel, is denoted by $p_1(x, y, z)$. It is a fastly decreasing smooth function on G.

We start by fixing a non-negative function K on G such that $F_\sigma(K)$ is smooth and has compact support and $F_\sigma(K)(0) = 1$. Then we form a Dirac family $(K_a)_{a>0}$ by letting $K_a(x) = a^4 K(d_a(x))$. Then $F_\sigma(K_a)(\xi) = F_\sigma(K)(d_{\frac{1}{a}}(\xi))$. Let us also write $K_a^\vee(z) = K_a(z^{-1})$. They also form a Dirac family when $a \to +\infty$.

Lemma 9.1. *There are two sequences of positive numbers $(\varepsilon_n)_n$ and $(a_n)_n$, depending only on μ, with $\varepsilon_n \to 0$ and $a_n \to +\infty$, such that for all bounded Borel sets $B \subset G$ for which $\max\{|y_1|, |y_2|\} \leq n/\log n$ whenever $\mathbf{y} = (y_1, y_2, y_3) \in B$, and all $\mathbf{x} \in G$, if we set $P_n^B(\mathbf{x}) = \mu^n(\mathbf{x}B)$ and $Q_n^B(\mathbf{x}) = \nu^n(\mathbf{x}B)$, the following inequality holds for all positive integers n,*

$$n^2 \left| P_n^B * K_{a_n}^\vee(\mathbf{x}) - Q_n^B * K_{a_n}^\vee(\mathbf{x}) \right| \leq \varepsilon_n \max\{1, |B|\} \tag{50}$$

where $|B|$ denotes the Haar measure of B.

Proof. We may write

$$\begin{aligned} P_n^B * K_a^\vee(\mathbf{x}) &= \int_G \mu^n(\mathbf{x}z^{-1}B) K_a^\vee(z) dz \\ &= \int_G \mathbb{E}(\mathbf{1}_{S_n^{-1} z \in B^{-1}}) K_a(\mathbf{x}^{-1} z) dz \\ &= \langle \rho(\mu^n) f, {}_{\mathbf{x}} g \rangle \end{aligned}$$

where ${}_{\mathbf{x}}g$ is the translate ${}_{\mathbf{x}}g(z) := g(\mathbf{x}^{-1}z)$ and $f := \mathbf{1}_{B^{-1}}$ and $g(z) := K_a(z)$ (note that K_a is real). Then we can make use of the calculations performed in the previous sections to estimate this scalar product. We use the notations introduced in section 8.

As noted at the beginning of section 8, we can write

$$n^2 \langle \rho(S_n) f, {}_{\mathbf{x}} g \rangle = A_n(\mu) + B_n(\mu)$$

where A_n is controlled by the estimation (41) and B_n given by (42). Since $\|f\|_{L^2(G)} = \sqrt{|B|}$ and $\|{}_x g\|_{L^2(G)} = a^2 \|K\|_{L^2(G)}$ we obtain (take $k_0 = 3$ and $D = 3/c$)

$$|A_n(\mu)| \leq C(a) \cdot \frac{a^2}{n} \cdot \sqrt{|B|} \cdot \|K\|_{L^2(G)} \tag{51}$$

where $C(a)$ is a positive constant depending only on a. Integrating the decomposition $I_n(\mathbf{y}) = I_n^S(\mathbf{y}) + I_n^L(\mathbf{y})$ with respect to $f(\mathbf{y})d\mathbf{y}$ (see equations (42) and (43)), we obtain that $B_n(\mu)$ can be written as a sum $B_n^S(\mu) + B_n^L(\mu)$. Note that $\|F_\sigma({}_x K_a)\|_\infty =$

$\|F_\sigma(K)\|_\infty$ as follows from (47). Spliting $B_n^L(\mu)$ into part when $\varepsilon \geq |(t,s)| \geq D\frac{\log^8 n}{n}$ and the part when $|(t,s)| \geq \varepsilon$ we have from (42) and (44)

$$B_n^L(\mu) = B_n^{L_0}(\mu) + B_n^{L_1}(\mu)$$

$$|B_n^{L_0}(\mu)| \leq |B| \cdot \|F_\sigma(K)\|_\infty \cdot O_\mu(\frac{1}{\log n}) \tag{52}$$

as follows from Proposition 5.4, and

$$|B_n^{L_1}(\mu)| \leq \frac{|B| \cdot \|F_\sigma(K)\|_\infty}{(2\pi)^{3/2}} \cdot L \cdot n^2 \sup_{T \geq |(t,s)| \geq \varepsilon, |\lambda| \leq D\frac{\log n}{n}} |\mathbb{E}(\theta_n^\mu(t,s,\lambda))| \tag{53}$$

where T is the size of the support of the functions $F_\sigma({}_{\mathbf{x}}g)$ and L is the Lebesgue measure of that support:

$$(t,s) \mapsto F_\sigma(K)(\frac{t - \lambda y_2}{a}, \frac{s + 2\lambda\alpha(y_2 - x_2) + \lambda x_1}{a}, \frac{\lambda}{a^2})$$

where $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2, y_3)$. Note that L is a fixed multiple of a^2. Since $|\lambda| \leq D\frac{\log n}{n}$ in (44) and by assumption $\max\{|y_1|, |y_2|\} \leq n/\log n$, if we suppose additionally that $\max\{|x_1|, |x_2|\} \leq 2n/\log n$ then the constant T in the above equation (53) is bounded by some fixed function of a. In Proposition 5.2 we showed that there is a constant C_T such that

$$\sup_{T \geq |(t,s)| \geq \varepsilon, |\lambda| \leq D\frac{\log n}{n}} |\mathbb{E}(\theta_n^\mu(t,s,\lambda))| \leq \frac{C_T}{n^3}$$

Hence for some number $\beta(a) < +\infty$

$$|B_n^L(\mu)| \leq |B| \cdot \frac{\beta(a)}{\log n} \tag{54}$$

The estimations above can be carried out in a similar way for ν instead of μ. In particular

$$\begin{aligned} Q_n^B * K_{a_n}^\vee(\mathbf{x}) &= \langle \rho(\nu^n) f, {}_{\mathbf{x}} g \rangle \\ &= A_n(\nu) + B_n(\nu) \end{aligned}$$

The term $A_n(\nu)$ is dealt with in exaclty the same way since estimate (41) is also valid for ν and we have

$$|A_n(\nu)| \leq C(a) \cdot \frac{a^2}{n} \cdot \sqrt{|B|} \cdot \|K\|_{L^2(G)} \tag{55}$$

In order to control $B_n(\mu)$ we made use of the compact support assumption on μ. For ν we can use the following direct argument because $F_\sigma(p_1)$ is integrable since $p_1(x,y,z)$

decays rapidly when (x, y, z) is large. Recall that $(\nu_t)_t$ is a stable semi-group, i.e. $\nu_n = \nu_1^n = d_{\sqrt{n}}(\nu_1)$. From (44) we have

$$
\begin{aligned}
|B_n^L(\nu)| &\leq \frac{1}{(2\pi)^{3/2}} \cdot |B| \cdot \|F_\sigma(K)\|_\infty n^2 \int_{|(t,s)|\geq D\frac{\log^8 n}{\sqrt{n}}} |\mathbb{E}(\theta_n^\nu(t,s,\lambda))| dt ds d\lambda \\
&\leq \frac{1}{(2\pi)^{3/2}} \cdot |B| \cdot \|F_\sigma(K)\|_\infty \cdot \int_{|(t,s)|\geq D\log^8 n} |F_\sigma(p_1)(t,s,\lambda)| dt ds d\lambda \\
&\leq \frac{1}{(2\pi)^{3/2}} \cdot |B| \cdot \|F_\sigma(K)\|_\infty \cdot o(1) \qquad (56)
\end{aligned}
$$

Therefore it remains to treat $B_n^S(\mu) - B_n^S(\nu)$. From (42) and (43), we have

(57)

$$
\left|B_n^S(\mu) - B_n^S(\nu)\right| \leq A \cdot n^2 \int_{|\lambda|\leq D\frac{\log n}{n}} \int_{|(t,s)|\leq D\frac{\log^8 n}{\sqrt{n}}} |\mathbb{E}(\theta_n^\mu(t,s,\lambda)) - \mathbb{E}(\theta_n^\nu(t,s,\lambda))| dt ds d\lambda
$$

with

$$
A = \frac{1}{(2\pi)^{3/2}} \cdot |B| \cdot \|F_\sigma(K)\|_\infty
$$

Now the integral on the right hand side tends to 0 as it follows from Lebesgue's dominated converge theorem like we did in section 8. In section 7 we showed that there exists an integrable function $\psi(t, s, \lambda)$ such that for n sufficiently large and for all (t, s, λ) such that $|\lambda| \leq D \log n$ and $|(t, s)| \leq D \log^8 n$,

$$
\left|\mathbb{E}(\theta_n^\mu(\frac{t}{\sqrt{n}}, \frac{s}{\sqrt{n}}, \frac{\lambda}{n}))\right| \leq \psi(t,s,\lambda) + O(\frac{\log^6 n}{\sqrt{n}})
$$

On the other hand $F_\sigma(p_1)$ is integrable and

$$
\mathbb{E}(\theta_n^\nu(\frac{t}{\sqrt{n}}, \frac{s}{\sqrt{n}}, \frac{\lambda}{n})) = F_\sigma(p_1)(t,s,\lambda)
$$

And by the central limit theorem, we had point-wise

$$
\lim_{n\to\infty} \mathbb{E}(\theta_n^\mu(\frac{t}{\sqrt{n}}, \frac{s}{\sqrt{n}}, \frac{\lambda}{n})) - \mathbb{E}(\theta_n^\nu(\frac{t}{\sqrt{n}}, \frac{s}{\sqrt{n}}, \frac{\lambda}{n})) = 0
$$

Hence by Lebesgue's dominated convergence theorem,

$$
\lim_{n\to+\infty} n^2 \int_{|\lambda|\leq D\frac{\log n}{n}} \int_{|(t,s)|\leq D\frac{\log^8 n}{\sqrt{n}}} |\mathbb{E}(\theta_n^\mu(t,s,\lambda)) - \mathbb{E}(\theta_n^\nu(t,s,\lambda))| dt ds d\lambda = 0
$$

Therefore (57) reads:

$$
\left|B_n^S(\mu) - B_n^S(\nu)\right| \leq |B| \cdot o(1) \qquad (58)
$$

And finally combining (51), (55), (56), (54) and (58), we have that for some sequence $\varepsilon_n \to 0$

$$n^2 \left| P_n^B * K_a^\vee(\mathbf{x}) - Q_n^B * K_a^\vee(\mathbf{x}) \right| \leq \varepsilon_n |B| + \frac{\beta(a)}{\log n}|B| + C(a)\frac{a^2}{n} \cdot 2 \|K\|_2 \sqrt{|B|}$$

whenever $\mathbf{x}$ satisfies $\max\{|x_1|, |x_2|\} \leq 2n/\log n$. But we can choose a sequence $a_n \to +\infty$ such that $\frac{\beta(a_n)}{\log n}$ and $C(a_n)\frac{a_n^2}{n}$ tend to 0. Changing ε_n if necessary, we obtain for these values of $\mathbf{x}$

$$n^2 \left| P_n^B * K_{a_n}^\vee(\mathbf{x}) - Q_n^B * K_{a_n}^\vee(\mathbf{x}) \right| \leq \varepsilon_n \max\{|B|, \sqrt{|B|}\}$$

Now let us examine the case when $\mathbf{x}$ takes large values, i.e. when $\max\{|x_1|, |x_2|\} \geq 2n/\log n$. Then

$$P_n^B * K_{a_n}^\vee(\mathbf{x}) = \int_G \mu^n(xzB) K_{a_n}(z)dz \leq P(|T_n| \text{ or } |U_n| \geq n/2\log n) + \int_{|z| \geq \frac{n}{2\log n}} K_{a_n}(z)dz$$

where U_n is as before the sum $Y_1 + ... + Y_n$ and $T_n = X_1 + ... + X_n$. Now since $F_\sigma(K)$ is smooth, K decays rapidly and in particular

$$\int_{|z| \geq \frac{n}{2\log n}} K_{a_n}(z)dz = \int_{|z| \geq a_n \frac{n}{2\log n}} K(z)dz = O(\frac{1}{n^k})$$

for any $k \geq 0$. Hence we get from Lemma 5.1

$$P_n^B * K_{a_n}^\vee(\mathbf{x}) = o(1/n^2)$$

Similarly the same holds for $Q_n^B * K_{a_n}^\vee(\mathbf{x})$ when $\max\{|x_1|, |x_2|\} \geq 2n/\log n$. Changing $(\varepsilon_n)_n$ if necessary, we obtain the desired conclusion, i.e. inequality (50) for all $\mathbf{x}$. ☐

The restrictions on the size of B in the above lemma disappear if we make the additional assumption (Cramér's condition) that

$$\sup_{t^2+s^2 \geq 1} |\mathbb{E}(e^{i(tX+sY)})| < 1$$

Indeed in this case, we can control $\mathbb{E}(\theta_n^\mu(t,s,\lambda))$ uniformly for arbitrary large values of t and s (see remark (5.3)). Therefore we can take $T = +\infty$ in the estimation (53) above and the restriction on B is unecessary. We obtain

Lemma 9.2. *If we suppose additionally that μ satisfies Cramér's condition* (9)*, then there are two sequences of positive numbers $(\varepsilon_n)_n$ and $(a_n)_n$, depending only on μ, with $\varepsilon_n \to 0$ and $a_n \to +\infty$, such that for all bounded Borel sets $B \subset G$ and all $x \in G$, if we set $P_n^B(x) = \mu^n(xB)$ and $Q_n^B(x) = \nu^n(xB)$, the following inequality holds for all positive integers n,*

$$n^2 \left| P_n^B * K_{a_n}^\vee(x) - Q_n^B * K_{a_n}^\vee(x) \right| \leq \varepsilon_n \max\{1, |B|\}$$

where $|B|$ denotes the Haar measure of B.

Lemma 9.3. *Let $(a_n)_n$ be a sequence of positive numbers such that $a_n \to +\infty$. Then there exists another sequence $(\varepsilon_n)_n$ with $\varepsilon_n \to 0$ and a sequence of neighborhoods of identity (U_n) converging to identity, such that for all bounded Borel sets $B \subset G$, the following inequality holds for all positive integers n,*

$$n^2 \left|Q_n^B(x) - Q_n^B * K_{a_n}^\vee(x)\right| \leq p(|V_n B \backslash B| + \varepsilon_n |B|)$$

where $|B|$ denotes the Haar measure of B and where we have set $Q_n^B(x) = \nu^n(xB)$ and $p = max\big(|p_1(g)|, g \in G\big) > 0$.

Proof. The proof is straightforward. We first note that for any bounded Borel set B, $\nu^n(B) \leq p|B|/n^2$. Then we simply write:

$$\begin{aligned} |Q_n^B * K_{a_n}^\vee(x) - Q_n^B(x)| &\leq \int |Q_n^B(xz) - Q_n^B(x)| K_{a_n}(z)dz \\ &\leq \int_{z \in U_n^{-1}} \nu^n(x(zB\Delta B) K_{a_n}(z)dz \\ &\quad +2p\frac{|B|}{n^2} \int_{z \notin U_n^{-1}} K_{a_n}(z)dz \\ &\leq p(|U_n B \backslash B| + \varepsilon_n |B|)/n^2 \end{aligned}$$

where U_n is a sequence of neighborhoods of identity tending to identity such that $\int_{z \notin U_n^{-1}} K_{a_n}(z)dz$ tends to 0 at infinity. □

Now, let us complete the proof of Theorem (1.2). Let B be an arbitrary bounded Borel set satisfying the condition of Lemma (9.1), that is $\max\{|y_1|, |y_2|\} \leq n/\log n$ whenever $y = (y_1, y_2, y_3) \in B$ (resp. satisfying no additional condition if we assume Cramér's condition). In the sequel like above, the Landau notations o and O will correspond to functions depending only on μ. We keep notations of Lemma (9.1),

$$\begin{aligned} P_n^{U_n B} * K_{a_n}^\vee(x) &= \int \mu^n(xzU_n B) K_{a_n}(z)dz \\ &\geq \mu^n(xB) \int_{U_n} K_{a_n}(z)dz \end{aligned}$$

where U_n is as in Lemma 9.3. Now making use of Lemma 9.1 we get uniformly in $x \in G$,

$$\begin{aligned} P_n^B(x) &\leq (1 + o(1)) P_n^{U_n B} * K_{a_n}^\vee(x) \\ &\leq (1 + o(1))[Q_n^{U_n B} * K_{a_n}^\vee(x) + (1 + |U_n B|)o(1/n^2)] \end{aligned} \tag{59}$$

And by Lemma 9.3,

$$\begin{aligned} Q_n^{U_nB} * K_{a_n}^{\vee}(x) &\leq Q_n^{U_nB}(x) + \frac{p}{n^2}(|U_n^2B\backslash B| + \varepsilon_n|U_nB|) \\ &\leq Q_n^B(x) + \frac{p}{n^2}|U_nB\backslash B| + \frac{p}{n^2}(|U_n^2B\backslash B| + \varepsilon_n|U_nB|) \\ &\leq Q_n^B(x) + \frac{p}{n^2}(2|U_n^2B\backslash B| + \varepsilon_n|U_nB|) \end{aligned}$$

In particular we have

$$Q_n^{U_nB} * K_{a_n}^{\vee}(x) \leq \frac{4p}{n^2}|U_n^2B|$$

and, from (59),

$$P_n^B(x) \leq |U_n^2B|O(\frac{1}{n^2}) + o(\frac{1}{n^2}) \tag{60}$$

Additionally,

$$P_n^B(x) \leq Q_n^B(x) + \frac{2p}{n^2}|U_n^2B\backslash B| + (|U_n^2B| + 1)o(\frac{1}{n^2}) \tag{61}$$

Now let us turn to the other direction of the inequality. We have, making use of (60)

$$\begin{aligned} P_n^B * K_{a_n}^{\vee}(x) &= \int \mu^n(xzB)K_{a_n}(z)dz \\ &\leq \int_{U_n} \mu^n(xzB)K_{a_n}(z)dz + \int_{U_n^c} \mu^n(xzB)K_{a_n}(z)dz \\ &\leq \mu^n(xU_nB)\int_{U_n} K_{a_n}(z)dz + (|U_n^2B| + 1)o(\frac{1}{n^2}) \\ &\leq P_n^{U_nB}(x) + (|U_n^2B| + 1)o(\frac{1}{n^2}) \end{aligned}$$

But from (60),

$$\begin{aligned} P_n^{U_nB}(x) - P_n^B(x) &= \mu^n(x(U_nB\backslash B)) \\ &\leq |U_n^2(U_nB\backslash B)|O(\frac{1}{n^2}) + o(\frac{1}{n^2}) \end{aligned}$$

Hence

$$P_n^B * K_{a_n}^{\vee}(x) \leq P_n^B(x) + |U_n^2(U_nB\backslash B)|O(\frac{1}{n^2}) + (|U_n^2B| + 1)o(\frac{1}{n^2})$$

Now it follows from Lemma (9.1) that

$$Q_n^B * K_{a_n}^{\vee}(x) \leq P_n^B * K_{a_n}^{\vee}(x) + (1 + |U_nB|)o(1/n^2)$$

and from Lemma (9.3)

$$Q_n^B(x) \leq Q_n^B * K_{a_n}^{\vee}(x) + \frac{p}{n^2}(|U_nB\backslash B| + \varepsilon_n|B|)$$

Combining the last three inequalities, we get

$$Q_n^B(x) \leq P_n^B(x) + |U_n^2(U_nB\backslash B)|O(\frac{1}{n^2}) + (|U_n^2B|+1)o(\frac{1}{n^2}) \tag{62}$$

Equations (61) and (62) yield the desired result

$$|P_n^B(x) - Q_n^B(x)| \leq |U_n^2(U_nB\backslash B)|O(\frac{1}{n^2}) + (|U_n^2B|+1)o(\frac{1}{n^2}) \tag{63}$$

But clearly $\bigcap_{n\geq 0} U_n^2(U_nB\backslash B)$ is contained in $\overline{B}\backslash\overset{\circ}{B}$. Hence for every bounded measurable set B such that $|\partial B| = 0$ we have

$$\lim_{n\to 0}\sup_{x\in G} n^2|\mu^n(xB) - \nu^n(xB)| = 0 \tag{64}$$

And if μ satisfies Cramér's condition the estimate (63) holds without the above restriction on B. Hence (63) holds uniformly on y for all By. We conclude

$$\lim_{n\to 0}\sup_{x,y\in G} n^2|\mu^n(xBy) - \nu^n(xBy)| = 0 \tag{65}$$

Remark 9.4. *It is easy to see from* (63) *that the limits in* (64) *and* (65) *are uniform in B when B ranges over the set of balls for a given norm on G lying in a given compact subset of G.*

Finally note that Theorem 1.3 follows instantly from the inequality (60) above.

10. Applications

10.1. **An equidistribution result for bounded uniformly continuous functions.** In this section, we intend to give a proof of Corollary (1.4). The proof splits into two steps.

Lemma 10.1. *Suppose f is a continuous and bounded function on G satisfying condition* (2) *of Corollary* (1.4) *then*

$$\lim_{n\to\infty}\int_G f(g)d\nu^n(g) = \ell$$

where $\nu = \nu_1$ is an arbitrary gaussian measure on G.

Proof. Let $(\nu_t)_t$ be the one-parameter semigroup of gaussian measures in which ν is embedded. By scaling invariance, ν_t coincide with the image of ν_1 under the automorphism $d_{\sqrt{t}}$ of G, where $d_t(x,y,z) = (tx, ty, t^2z)$. Hence

$$\int_G f(g)d\nu^n(g) = \int_G f\circ d_{\sqrt{n}}(g)p(g)dg$$

where $p(g)$ is the density of ν. It is known that p as well as its derivatives are smooth, fastly decreasing functions on G. Let $L>0$ be a Lipschitz constant for p in the sense that $|p(g_1) - p(g_2)| \leq L\|g_1 - g_2\|$ for all $g_1, g_2 \in G$ where $\|g\| = \max\{|x|, |y|, |z|\} \geq C$ for $g = (x, y, z)$. Fix $\varepsilon > 0$ and let $C > 0$ be such that

$$\int_{\|g\| \geq C} p(g)dg \leq \varepsilon$$

We now have

$$\left| \int_{\|g\| \geq C} f(d_{\sqrt{n}}(g))p(g)dg \right| \leq \varepsilon \|f\|_\infty$$

Let us denote by $R(g, h)$ the rectangle $[x, x + h) \times [y, y + h) \times [z, z + h)$ where $g = (x, y, z) \in G$ and $h > 0$. We can find a subdivision of the hypercube $\{\|g\| \leq C\}$ by small cubes of the form $R(g_i, h)$. Hence

$$\int_{\|g\| \leq C} f(d_{\sqrt{n}}(g))p(g)dg = \sum \int_{R(g_i,h)} f(\sqrt{n}x, \sqrt{n}y, nz)p(g)dg$$

Also

$$\left| \sum \int_{R(g_i,h)} f(\sqrt{n}x, \sqrt{n}y, nz)p(g)dg - \sum p(g_i) \int_{R(g_i,h)} f(\sqrt{n}x, \sqrt{n}y, nz)dg \right| \leq Lh \|f\|_\infty (2C)^3$$

Now, note that, by viewing $R(g, h)$ as a difference of several rectangles in $\mathbb{R}^3$, the assumption (2) made on f easily implies that for all $g \in \mathbb{G}$ and $h > 0$

$$\lim_{T \to +\infty} \frac{1}{T^2} \int f(x)\chi_{R(g,h)}(d_{1/\sqrt{T}}(x))dx = \ell \cdot h^3 \tag{66}$$

since $\chi_{R(g,h)} \circ d_{1/\sqrt{T}}$ can be written as a finite sum of terms of the form $\pm\chi_{[0,T_1] \times \ldots \times [0,T_3]}$ for some positive or negative $T_1, \ldots, T_3$. Hence by (66)

$$\lim_{T \to +\infty} \sum p(g_i) \int_{R(g_i,h)} f(\sqrt{n}x, \sqrt{n}y, nz)dg = \ell \cdot \sum p(g_i)h^3$$

But

$$\begin{aligned} \left| \sum p(g_i)h^3 - 1 \right| &\leq \int_{\|g\| \geq C} p(g)dg + \sum \int_{R(g_i,h)} |p(g) - p(g_i)|\, dg \\ &\leq \varepsilon + Lh(2C)^3 \end{aligned}$$

Therefore, combining the above inequalities, for n large enough

$$\left| \int f(d_{\sqrt{n}}(g))p(g)dg - \ell \right| \leq \ell(\varepsilon + Lh(2C)^3) + Lh \|f\|_\infty (2C)^3 + \varepsilon \|f\| + \varepsilon$$

We finally obtain the desired result since h can be taken arbitrarily small. $\square$

The second step is about comparing the integrals with respect to the probability measure μ and its associated gaussian distribution ν. Here, we make use of the uniform version of the local limit theorem (Theorem 1.2). Namely,

Lemma 10.2. *Let f be a bounded and uniformly continuous function (with respect to either right or left uniform structure on G). Let μ be a compactly supported aperiodic and centered probability measure on G. And let ν be its associated gaussian distribution. Then we have*

$$\lim_{n\to+\infty}\left|\int f(g)d\mu^n(g)-\int f(g)d\nu^n(g)\right|=0$$

Proof. We may assume $\|f\|_\infty \leq 1$. Fix $\varepsilon > 0$ and let $\omega > 0$ be a modulus of continuity for f relatively to ε, i.e. $|f(ux)-f(x)| \leq \varepsilon$ if $\|u\| \leq \omega$ and $x \in G$, where $\|g\| = \max\{|x|,|y|,|z|\}$ for $g = (x,y,z) \in G$. As follows from the central limit theorem, we can find a number $C > 0$ such that if we let $A_n = \{g = (x,y,z) \in G, |x|,|y| \leq C\sqrt{n}$ and $|z| \leq Cn\}$, we have for large n

$$\mu^{*n}(A_n^c) \leq \varepsilon$$

and

$$\nu^{*n}(A_n^c) \leq \varepsilon$$

We can then find a cover B_n of the cube A_n by less than $O(n^2/\omega^4)$ disjoint translates $R_\omega h$ of a small cube of the form $R_\omega = d_\omega(R)$ where $h \in d_\omega(G(\mathbb{Z}))$, d_ω is the dilation on G with coefficient of contraction ω, and R is a fundamental domain for the co-compact lattice $G(\mathbb{Z})$ in G. Now we can write

$$\begin{aligned}
\left|\int fd\mu^{*n}-\int fd\nu^{*n}\right| &\leq \left|\int_{B_n} fd\mu^{*n}-\int_{B_n} fd\nu^{*n}\right|+2\varepsilon\\
&\leq \sum_i f(h_i)\left|\mu^{*n}(Rh_i)-\nu^{*n}(Rh_i)\right|+4\varepsilon\\
&\leq O(\frac{n^2}{\omega^4})\sup_{h\in G}\left|\mu^{*n}(Rh)-\nu^{*n}(Rh)\right|+4\varepsilon
\end{aligned}$$

Here we can apply the uniform local limit theorem (Theorem 1.2) and get

$$\lim_{n\to\infty} n^2\sup_{h\in G}\left|\mu^{*n}(Rh)-\nu^{*n}(Rh)\right|=0$$

Thus, we obtain the desired result. □

The proof of Corollary 1.4 now follows immediately from the combination of the last two lemmas.

10.2. **Unipotent random walks and equidistribution on homogeneous spaces.** Here we shall conclude this paper and give a proof of Theorem 1.5.

Let G be a connected real Lie group and Γ a lattice in G, that is, a discrete subgroup of G such that the homogeneous space G/Γ bears a finite Borel measure invariant by the left action of G. An element $u \in G$ is called *Ad-unipotent*, when the automorphism $Ad(u) \in GL(\mathfrak{g})$ of the Lie algebra $\mathfrak{g}$ of G is unipotent, i.e. every eigenvalue of $Ad(u)$ equals 1. A subgroup $U \subset G$ is called *Ad-unipotent* or simply *unipotent* is every element $u \in U$ is Ad-unipotent. The action of U on G/Γ is called a *unipotent flow.*

In the early nineties, in a series of papers (see [Rat1-3]), M. Ratner proved the validity in full generality of the Raghunathan-Dani conjectures for the action of connected Ad-unipotent subgroups on G/Γ. These results have had a number of far reaching applications (some of which were obtained earlier by proving special cases of the conjecture, like in Margulis' proof of the Oppenheim conjecture (1986)), especially to number theory and lattice points counting problems (see the recent survey [Bab]). The results, can be summarized as follows. First, if U is a connected Ad-unipotent subgroup of G, for every $x \in G/\Gamma$, the orbit Ux has a "nice algebraic" closure, that is, there exists a closed subgroup $H \subset G$ such that $\overline{Ux} = Hx$ is closed and bears a unique H-invariant probability measure m_x. Secondly, every U-ergodic probability measure on G/Γ is of the form m_x for some $x \in G/\Gamma$. We refer the reader to the surveys [Rat3] and [Sta] for a detailed exposition of these results and further references (see also [MaT] for an alternative proof).

One of the main steps in the proof of the latter conjecture is the following equidistribution theorem for the action of one-parameter unipotent flows:

Theorem 10.3 (M. Ratner)**.** *Suppose G is a Lie group and Γ a lattice in G. Let $U = \{u(t), t \in \mathbb{R}\}$ be a one-parameter Ad-unipotent subgroup of G. Then for any $x \in G/\Gamma$, there is a closed subgroup H of G, such that Hx is closed and bears an H-invariant probability m_x, and the orbit Ux is equidistributed in Hx with respect to m_x. In other words, for all continuous and bounded functions f on G/Γ, we have*

$$\lim_{T \to +\infty} \frac{1}{T} \int_0^T f(u(t)x)dt = \int_{Hx} f dm_x$$

Let us emphasize the fact that this equidistribution holds for every point $x \in G/\Gamma$ and not only almost everywhere with respect to some U-ergodic measure.

10.2.1. *A counter-example.* In this paragraph, we illustrate the phenomenon described in the introduction of a non-centered random walk that may diverge in G/Γ. More precisely, we have:

Proposition 10.4. *Let $G = SL_2(\mathbb{R})$, $\Gamma = SL_2(\mathbb{Z})$ and U be a one-parameter unipotent subgroup of G. Let μ be a non-centered probability measure on U with standard deviation $\sigma < +\infty$ and average $d \neq 0$. Then for any compact subset K in G/Γ and λ-almost all points $x \in G/\Gamma$ (λ is a G-invariant measure on G/Γ), we have*

$$\liminf_{n \to +\infty} \mu^n * \delta_x(K) = 0 \tag{67}$$

Let $S_n = X_1 + ... + X_n$ be a sum of centered i.i.d. variables on the line with standard deviation $\sigma > 0$. Then the random walk $(S_n + nd)_n$ is non-centered with drift d, and we can choose X_1 so that the probability law of $S_n + nd$ is precisely μ^n.

We identify G/Γ with the space of lattices in $\mathbb{R}^2$ and we let $\|\cdot\|$ be the standard 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 any lattice $x \in K$ and any non-zero vector $v \in x$. We set $x_0 = \mathbb{Z}^2$.

Let $(u(t))_t$ be a one-parameter unipotent subgroup in $SL_2(\mathbb{R})$. It stabilizes a line D in $\mathbb{R}^2$ and we let $\theta \in \mathbb{R}$ be the slope (that we assume finite) of this line. In the canonical coordinates of $\mathbb{R}^2$, the action of $u(t)$ reads:

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

for some $\alpha \in \mathbb{R}\backslash\{0\}$.

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

$$|x(y - \theta x)| < \varepsilon \text{ and } \sigma x(y - \theta x) > 0$$

For the Lebesgue measure on $\mathbb{R}$, almost all θ are well approximable on both sides.

Let us fix an arbitrary compact subset $K \subset G/\Gamma$ and let $\delta > 0$ be small enough so that $\|v\| > \delta$ for any lattice $x \in K$ and any non-zero vector $v \in x$. Let $\Omega = \{t, u(t)x_0 \notin K\}$. Then we claim:

Claim 10.5. *If $\theta \notin \mathbb{Q}$ is well approximable on both sides, then for any $C > 0$ and $n_0 > 0$ there is an integer $n > n_0$ such that*

$$S_n + nd \in \Omega$$

whenever $|S_n| < C\sqrt{n}$.

Proof of claim: For $(x, y) \in \mathbb{Z}^2\backslash\{(0,0)\}$ we set $A_{x,y} = \{t \in \mathbb{R}, |u(t)(x,y)| < \delta\}$ and $\kappa = |2\alpha| \max\{|\theta|, 1\}$ and $t_{x,y} = \frac{x}{\alpha(\theta x - y)}$, $I_{x,y} = \{t, |t| < \frac{\delta}{\kappa|\theta x - y|}\}$ so that

$$\bigcup_{x,y \in \mathbb{Z}^2\backslash\{(0,0)\}} A_{x,y} \subset \Omega$$

If δ is small enough, since $\theta \notin \mathbb{Q}$ is well approximable on both sides, one can find arbitrarily large integers x, y such that

$$|x(y - \theta x)| < \delta^4 \tag{68}$$

and $t_{x,y}$ has the sign we want. Note that if δ is small enough, then $t_{x,y} + I_{x,y} \subset A_{x,y}$. We can also assume $|I_{x,y}| > 2|d|$. Hence, for some integer n,

$$nd \in t_{x,y} + \frac{1}{2} I_{x,y} \tag{69}$$

Now if S_n is not too large, i.e. if $|S_n| < \frac{1}{4}|I_{x,y}|$, then $S_n + nd \in A_{x,y} \subset \Omega$. On the other hand, if n satisfies (69) then $nd \leq \frac{2|x|}{|\alpha||\theta x - y|}$. Thanks to (68), it follows that $|S_n| < \frac{1}{4}|I_{x,y}|$ whenever $|S_n| < C_\delta\sqrt{n}$, with $C_\delta = \frac{1}{\delta}\sqrt{|\alpha d|/2\kappa^2}$. Taking a smaller δ is necessary, we obtain the desired conclusion. $\square$

Now the proof of the proposition follows easily from the central limit theorem: if $\varepsilon > 0$ one can find $C > 0$ so that for all large enough n, $\mathbb{P}(|S_n| < C\sqrt{n}) \geq 1 - \varepsilon$. Hence,

$$\limsup \mathbb{P}(u(S_n + nd)x_0 \notin K) = 1 \tag{70}$$

We thus have established (67) for $x = x_0$ and as soon as the slope θ of the line D fixed by U satisfies the diophantine condition defined above. If E is the set of $g \in G$ such that (67) holds for $x = g^{-1}x_0$, then E contains the set of $g \in G$ such that the slope of gD is irrational and well approximable on both sides. The map from G to $\mathbb{R}$ which sends g to the slope of gD has no critical points (it identifies with left translation on G/P). Hence E is a set of full Haar measure, and this ends the proof of the proposition.

Remark 10.6. *The same idea shows that a non-centered random walk can stay very close to a closed orbit of U at arbitrary large time and with high probability. In any case, it prevents the walk to equidistribute in the closure of the orbit. A similar phenomenon arises as soon as the action of U on G/Γ is not uniquely ergodic, i.e. if U stabilizes a proper homogeneous subspace of G/Γ.*

10.2.2. *Proof of Theorem 1.5.* Making use of the equidistribution theorem above (Theorem 10.3), N. Shah (cf. [Sha]) subsequently extended this result to the action of an arbitrary simply connected *Ad*-unipotent subgroup U of G. Let us introduce N. Shah's result.

Let U be any simply connected nilpotent Lie group and $(v_1, ..., v_k)$ be a basis of the Lie algebra $\mathfrak{u}$ of U. This basis is called a *triangular basis* (or strong Malcev basis) if the subspaces spanned by $(v_i, ..., v_k)$ for any i are ideals of $\mathfrak{u}$, that is $[v_i, v_j] \in span(v_m, ..., v_k)$ where $m = \max\{i, j\} + 1$. Such a basis gives rise to polynomial coordinates on U, i.e. the map

$$\begin{array}{rcl} \phi & : & \mathbb{R}^k \to U \\ (t_1, ..., t_k) & \longmapsto & \exp(t_k v_k) \cdot ... \cdot \exp(t_1 v_1) \end{array}$$

is polynomial diffeomorphism. It also sends the Lebesgue measure on $\mathbb{R}^k$ to the Haar measure on U. With this terminology, Shah proved (cf. [Sha] Cor. 1.3.)

Theorem 10.7 (N. Shah)**.** *Suppose G is a Lie group and Γ a lattice in G. Let U be a simply connected Ad-unipotent subgroup of G. Let $(v_1, ..., v_k)$ be a triangular basis for U, and $x \in G/\Gamma$. Then for any continuous and bounded function f on G/Γ,*

$$\lim_{T_1 \to \infty, ..., T_k \to \infty} \frac{1}{T_1 ... T_k} \int_{[0,T_1] \times ... \times [0,T_k]} f(\phi(t_1, ..., t_k)x) dt_1 ... dt_k = \int_{Hx} f dm_x$$

where m_x is the H-invariant probability measure on $\overline{Ux} = Hx$.

This theorem is precisely what we need to apply Corollary 1.4 to the situation of Theorem 1.5. Keeping the notations of the statement of Theorem 1.5, let f be a compactly supported function on G/Γ, and suppose that U is isomorphic to the Heisenberg group with triangular basis given by (3) in section 2. Then the function $F(u) = f(ux) = f(\phi(u_x, u_y, u_z)x)$ is a bounded uniformly continuous function for the left uniform structure on U satisfying the condition of Corollary 1.4 with limit $\ell = \int_{Hx} f dm_x$. Therefore this is the end of the proof and of this paper.

Acknowledgments 10.8. *This work is part of the author's Yale University Ph.D. thesis. I sincerely thank my supervisor Gregory Margulis for suggesting this problem and offering his invaluable help and guidance throughout the last few years. I am also very grateful to Yves Guivarc'h for our long discussions and to Martine Babillot whose inspiring enthusiasm I now miss dearly.*

REFERENCES

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

[ArK] 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.

[Bab] M. Babillot, *Points entiers et groupes discrets, de l'analyse aux systèmes dynamiques*, in "Rigidité, groupe fondamental et dynamique", Panoramas et Synthèses, **13**, p. 1–119, Soc. Math. France, Paris, (2002).

[BVZ] C. Beguin, A. Valette, A. Zuk, *On the spectrum of a random walk on the discrete Heisenberg group and the norm of the Harper operator*, J. Geom. Phys. 21 (1997) p. 337-356.

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

[Bre] Breiman, *Probability*, Addison-Wesley (1968)

[Breu] E. Breuillard, *Distributions diophantiennes et théorème limite local sur* $\mathbb{R}^d$, Probab. Theory Rel. Fields, (2004).

[Breu2] E. Breuillard, *Random walks on Lie groups*, Chapter I in Yale Univ. dissertation, (2004).

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

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

[Dan] S.G. Dani, *On orbits of unipotent flows on homogeneous spaces*, Ergodic Theory Dynam. Systems **6** (1986), no. 2, p. 167–182.

[EsM] A. Eskin, G.A. Margulis, *Recurrence properties of random walks on finite volume homogeneous manifolds*, in Random walks and geometry, p. 431–444, Walter de Gruyter, Berlin, (2004).

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

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

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

[Gui] Y. Guivarc'h, *Croissance polynômiale et périodes des fonctions harmoniques*, Bull. Soc. math. France, **101**, (1973), 333-379.

[Gui2] Y. Guivarc'h, *Equirépartition dans les espaces homogènes*, (French) Théorie ergodique (Actes Journées Ergodiques, Rennes, 1973/1974), pp. 131–142. Lecture Notes in Math., Vol. 532, Springer, Berlin, (1976).

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

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

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

[Mar] G.A. Margulis, *On the action of unipotent subgroups in the space of lattices*, Mat. Sb. **86** (1971), 552-556.

[MaT] 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, 347–392.

[Neu] D. Neuenschwander, *Probabilities on the Heisenberg group*, LNM 1630, Springer-Verlag (1996).

[Ose] V.I. Oseledec, *Markov chains, skew-products, and ergodic thoerems for "general" dynamical systems*, Th. Prob. & App. (1965), **10**, 3, 551-557.

[Rat1] M. Ratner, *On Raghunathan's measure conjecture*, Ann. Math. **134** (1991), 545-607.

[Rat2] M. Ratner, *Raghunathan's topological conjecture and distributions of unipotent flows*, Duke Math. J. **63** (1991), 235-290.

[Rat3] M. Ratner, *Interaction between Ergodic Theory, Lie Groups, and Number Theory*, in Proc. ICM (Zurich, 1994) 157-182, Birkhaüser, 1995.

[Rau] A. Raugi, *Théorème de la limite centrale sur les groupes Nilpotents,* Z. Wahrsch. Verw. Gebiete **43** (1978), no. 2, 149–172.

[Sar] P. Sarnak, *Some applications of modular forms*, Cambridge Tracts in Mathematics, 99 CUP, Cambridge, (1990).

[Sha] N. Shah, *Limit distributions of polynomial trajectories on homogeneous spaces*, Duke Math. J. **75**, 3, (1994), 711-732.

[Sta] A.N. Starkov, *Dynamical systems on homogeneous spaces*, Transl. of Math. Monographs, **190**, AMS Prov. RI, (2000).

[Sto] Ch. Stone, *A local limit theorem for nonlattice multi-dimensional distribution functions*, Ann. Math. Stat, **36** (1965) p. 546-551.

[Tut] V.N. Tutubalin, *Compositions of measures on the simplest nilpotent group*, (Russian) Teor. Verojatnost. i Primenen **9**, (1964) p. 531–539.

[VSC] Varopoulos, Saloff-Coste, et Coulhon, *Analysis and geometry on groups*, Cambridge Tracts in Mathematics, 100, CUP, Cambridge, (1992).

[Weh] D. Wehn, *Probabilities on Lie groups*, Proc. Nat. Acad. Sci. U.S.A. 48 (1962) p. 791–795.

Emmanuel Breuillard, IHES, Le Bois-Marie, 35 route de Chartres, F-91440 Bures-sur-Yvette, FRANCE

E-mail address: emmanuel.breuillard@ihes.fr