The accuracy of Gaussian approximations in nilpotent Lie groups

by Vidmantas Bentkus and Gyula Pap

We consider the Central Limit Theorem with Gaussian limit distributions in stratified nilpotent Lie groups. We obtain estimates of the rate of convergence and Edgeworth expansions for expectations of smooth functionals.

1. Introduction

The role of nilpotent Lie groups in investigations of the Central Limit Theorem (CLT) in compact topological groups may be explained by an important result due to Hazod and Siebert [24], [8]. They showed that if a probability measure on a locally compact topological group has a nonempty domain of attraction then its support lies in a subgroup isomorphic to a simply connected nilpotent Lie group whose Lie algebra admits a positive graduation.

In this paper we find refinements of the following version of the CLT due to Raugi [19] (see also the survey paper Pap [16]). Let $|\cdot|$ denote a homogeneous norm on G, let $\delta_t \in \mathrm{Aut}(G)$, $t > 0$ denote a semigroup of standard dilations, and let μ be a centered probability measure on G. For $t > 0$ define the measure $\delta_{1/t}\mu(A) = \mu(\delta_t A)$, where $A \subset G$. Denote by μ^n the n-fold convolution of μ. Then the condition $\int |x|^2 \mu(dx) < \infty$ implies that the measures $\mu_n = \delta_{1/\sqrt{n}}\mu^n$, $n = 1, 2, 3, \ldots$ converge weakly to a Gaussian measure, say ν.

Under the second moment assumption in Section 4 a Berry-Esseen type bound is given for $\int f(x)(\mu_n - \nu)(dx)$ provided $f : G \to \mathbb{R}$ is a sufficiently smooth function. In particularly, this bound implies Raugi's version of the CLT.

In Section 5 we obtain the Edgeworth expansion

$$\int f(x)(\mu_n - \nu)(dx) = \alpha n^{-1/2} + O(n^{-1}),$$

where the coefficient $\alpha = \alpha(G, f, \mu)$ depends on G, several lower order moments of μ and derivatives of f only.

Sections 2 and 3 contain some preliminaries about Lie groups, Gaussian measures and differential calculus on these groups (see the book of Heyer [11] for the related information). The stratified Taylor expansion plays an important role. Notice that the proofs of the results

AMS 1991 *subject classifications.* Primary 60B15; secondary 60F05.

Key words and phrases. Central Limit Theorem, Berry-Esseen bound, Edgeworth expansion, Gaussian semigroups of measures on Lie groups, stratified Lie groups, stratified Taylor expansion.

rely heavily on differential calculus, although they are formally independent of a concrete version of it. Therefore we choose and adopt a version of this calculus most convenient for our purposes.

Our main aim was to develop methods and technique applicable in the non-commutative case. The approach here is related to the Lindeberg method as it is described in Bentkus et al. [1]. Due to the non-commutativity a direct application of this method is not possible. In this paper we restrict ourselves to the class of stratified nilpotent Lie groups of step s. This class seems to be sufficiently rich to reflect the structure of general nilpotent Lie groups. We restrict ourselves as well to investigation of $\int f(x)(\mu_n - \nu)(dx)$ with smooth f only since already in this case algebraic diffuculties are present in the full extent. As usually the investigation of the Berry-Esseen bounds and Edgeworth expansions for probabilities requires additionally ingenious analytic technique, and we are going to publish related results elsewhere. Remark that the only result known on nilpotent Lie groups is a Berry-Esseen type bound for probabilities of homogeneous balls due to Pap [14]. This bound is obtained under relatively stringent moment assumptions whereas the assumptions in this paper seem to be optimal. Edgeworth expansions on Lie groups are considered here for the first time. Remark that the term α in the Edgeworth expansion has certain new features specific only for the non-commutative case. The corresponding classical results for the group $(\mathbb{R}^k, +)$ follow from our theorems substituting $s = 1$.

The rates of convergence due to Khokhlov [12] and Rachev and Yukich [18] are not comparable with our ones since they are obtained for the Euclidean motion group, which is the semidirect product of $(\mathbb{R}^k, +)$ and of the compact group of orthogonal transformations of $\mathbb{R}^k$.

2. Preliminaries on nilpotent Lie groups

The letters c, C (respectively $c(\cdot)$, $C(\cdot)$) with or without indices will denote positive generic constants (respectively positive constants depending only on the quantities in parentheses); the same symbol may stand for different constants.

Let $\mathcal{G}$ be a m-dimensional s-step nilpotent Lie algebra with descending central series

$$\mathcal{G}^{(1)} = \mathcal{G}, \quad \mathcal{G}^{(2)} = [\mathcal{G}, \mathcal{G}^{(1)}], \ \ldots, \ \mathcal{G}^{(k+1)} = [\mathcal{G}, \mathcal{G}^{(k)}], \ \ldots, \ \mathcal{G}^{(s+1)} = \{0\}.$$

Let V_k be the vector subspace complementary to $\mathcal{G}^{(k+1)}$ in $\mathcal{G}^{(k)}$. Then $\mathcal{G} = \oplus_{k=1}^{s} V_k$ is a vector space decomposition. $\mathcal{G}$ is called *stratified* if V_1 generates $\mathcal{G}$ as an algebra. Let $\{X_1, \ldots, X_m\}$ be a basis in $\mathcal{G}$ adapted to the above decomposition, i. e. there exists a mapping

$$d : \{1, 2, \ldots m\} \to \{1, 2, \ldots, s\}$$

such that $\{X_i : d(i) = k\}$ is a basis of V_k. We shall write $d_i = d(i)$. Let G be a simply connected nilpotent Lie group whose Lie algebra $\mathcal{G}$ is stratified. The neutral element of G is denoted by e. It is known that the exponential mapping $\exp : \mathcal{G} \to G$ is an analytic diffeomorphism. An element $X \in \mathcal{G}$ can be regarded as a left-invariant differential operator on G: for $f \in \mathcal{C}^1(G)$ we put

$$Xf(x) = \lim_{t\to 0} \frac{f(x \exp tX) - f(x)}{t}.$$

Similarly, the right-invariant differential operator $\widetilde{X}$ is defined by

$$\widetilde{X}f(x) = \lim_{t\to 0} \frac{f((\exp tX)x) - f(x)}{t}.$$

For $x \in G$ the global canonical coordinates of the first kind $\{x_1, \ldots, x_m\}$ are defined by

$$x = \exp\left(\sum_{i=1}^m x_i X_i\right).$$

We equip $\mathcal{G}$ and G with the *natural dilations* by extending

$$\mathring{\delta}_t(X_i) = t^{d(i)} X_i, \qquad t > 0, \quad i = 1, \ldots, m$$

by linearity to $\mathcal{G}$ and transferring it to G by

$$\delta_t = \exp \circ \mathring{\delta}_t \circ \exp^{-1}, \qquad t > 0.$$

The family $(\delta_t)_{t>0}$ is a continuous one-parameter group in $\mathrm{Aut}(G)$ and plays the role of multiplication by scalars $t > 0$. A function $f : G \setminus \{e\} \to \mathbb{R}$ is called *homogeneous of degree* α $(\alpha \in \mathbb{R})$ if $f(\delta_r x) = r^\alpha f(x)$ for $x \in G$, $r > 0$. Similarly, a function $f : G\times G \setminus \{(e,e)\} \to \mathbb{R}$ is called homogeneous of degree α $(\alpha \in \mathbb{R})$ if $f(\delta_r x, \delta_r y) = r^\alpha f(x,y)$ for $x, y \in G$, $r > 0$.

We shall use the *homogeneous norm*

$$|x| := \sum_{i=1}^m |x_i|^{1/d(i)}, \qquad x \in G,$$

which is a homogeneous function of degree 1, $|\delta_t x| = t|x|$. An analogue of the triangular inequality holds, $|xy| \le c(G)(|x| + |y|)$. Moreover, $|x_i| \le |x|^{d_i}$. In fact, an arbitrary homogeneous norm can be fixed, because they are equivalent as norms, see Pap [14].

We adopt the following multiindex notation. For a multiindex $I = (i_1, \ldots, i_m) \in \mathbb{Z}_+^m$, where $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$ we set $x^I := x_1^{i_1} \ldots x_m^{i_m}$, $x \in G$. Further, $|I| := i_1 + \ldots + i_m$ and $d(I) := d_1 i_1 + \ldots + d_m i_m$. For $I = (i_1, \ldots, i_m)$ and $J = (j_1, \ldots, j_m)$ we put $I + J := (i_1 + j_1, \ldots, i_m + j_m)$. We shall denote by $[j]$ the multiindex with 1 in the j^{th} place and zeros elsewhere.

A function $P : G \to \mathbb{R}$ is called a *polynomial* if $P(\exp X)$, $X \in \mathcal{G}$ is a polynomial on $\mathcal{G}$. Every polynomial on G can be written uniquely as

$$P(x) = \sum_I a_I x^I,$$

where all but finitely many of the coefficients a_I vanish. The *homogeneous degree* of the polynomial P is $\max\{d(I) : a_I \neq 0\}$. A polynomial P is called *homogeneous* if it is homogeneous as a function. If P is a homogeneous polynomial of degree d then it can be written uniquely as

$$P(x) = \sum_{d(I)=d} a_I x^I,$$

and there is a constant $c_P > 0$ such that

$$|P(x)| \leq c_P |x|^d.$$

Similarly, a function $P : G \times G \to \mathbb{R}$ is called a polynomial if $P(\exp X, \exp Y)$, $X, Y \in \mathcal{G}$ is a polynomial on $\mathcal{G}$. It can be written uniquely as

$$P(x,y) = \sum_{I,J} a_{IJ} x^I y^J, \qquad x, y \in G,$$

where all but finitely many of the coefficients a_{IJ} vanish. Its homogeneous degree is $\max\{d(I) + d(J) : a_{IJ} \neq 0\}$. If it is a homogeneous function of degree d then it can be written uniquely as

$$P(x,y) = \sum_{d(I)+d(J)=d} a_{IJ} x^I y^J, \qquad x, y \in G,$$

and there is a constant $c_P > 0$ such that

$$|P(x,y)| \leq c_P (1 + |x|^d + |y|^d), \qquad x \in G.$$

The multiplication of G can be reconstructed by the help of the Campbell-Hausdorff series

$$(\exp\ X)(\exp\ Y) = \exp\left(X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] - \frac{1}{12}[Y,[X,Y]] + \ldots \right) \tag{1}$$

where the series terminates after finitely many terms and defines a polynomial map from $\mathcal{G} \times \mathcal{G}$ to $\mathcal{G}$. It follows that

$$(xy)_k = x_k + y_k + \sum_{I,J \neq 0,\, d(I)+d(J)=d_k} c^k_{IJ} x^I y^J, \qquad x, y \in G, \tag{2}$$

where $c_k^{IJ} \in \mathbb{R}$. Therefore

$$X_j f(x) = \left(\partial_j + \sum_{d_k \geq d_j + 1,\, d(I) = d_k - d_j} c^k_{I[j]} x^I \partial_k \right) f(x),$$

cf. Folland and Stein [6], Proposition 1.26. Thus we have

$$X_j f(x) = \left(\partial_j + \sum\nolimits_{d_k \geq d_j+1} P_{jk}(x)\partial_k\right) f(x),$$

where P_{jk} are homogeneous polynomials of degree $d_k - d_j$. Similarly

$$(xy)^K = x^K + y^K + \sum\nolimits_{I,J\neq 0,\, d(I)+d(J)=d(K)} c^K_{IJ} x^I y^J \tag{3}$$

and

$$X^J f(x) = \sum\nolimits_{d(K)\geq d(J),\, |K|\leq |J|} P_{JK}(x)\partial^K f(x),$$

where P_{JK} are homogeneous polynomials of degree $d(K) - d(J)$. Similar formulas are valid for right-invariant differential operators. Combining these formulas we obtain

$$\widetilde{X}^J f(x) = \sum\nolimits_{d(K)\geq d(J),\, |K|\leq |J|} Q_{JK}(x) X^K f(x), \tag{4}$$

where Q_{JK} are homogeneous polynomials of degree $d(K) - d(J)$ (cf. Folland and Stein [6], Proposition 1.29).

For a multiindex $I \in \mathbb{Z}^m_+$ we shall use the left-invariant differential operator

$$S^I := \frac{1}{|I|!} \sum\nolimits_{[j_1]+\ldots+[j_{|I|}]=I} X_{j_1} \ldots X_{j_{|I|}}.$$

It can be considered as the symmetrization of X^I. The right-invariant differential operator $\widetilde{S}^I$ is defined similarly. We remark that by the Poincaré–Birkhoff–Witt theorem (cf. Bourbaki [2], I.2.7), the operators X^I form a basis for the algebra of left-invariant differential operators on G. In particularly,

$$S^I = \sum\nolimits_{d(J)=d(I),\, |J|\leq |I|} C^I_J X^J \tag{5}$$

and

$$X^I X^J = \sum\nolimits_{d(K)=d(I)+d(J),\, |K|\leq |I|+|J|} C^{IJ}_K X^K. \tag{6}$$

Combining formulas (4) and (6) we get

$$\widetilde{X}^J X^I f(x) = \sum\nolimits_{d(K)\geq d(I)+d(J),\, |K|\leq |I|+|J|} P^{IJ}_K(x) X^K f(x), \tag{7}$$

where P^{IJ}_K are homogeneous polynomials of degree $d(K) - d(I) - d(J)$.

For $k \in \mathbb{N}$ and $f \in \mathcal{C}^k(G)$ let

$$\|D^{(k)} f\|_\infty := \max\nolimits_{d(I)\leq k} \sup\nolimits_{x\in G} |X^I f(x)|.$$

Lemma 1. *Let* $I \in \mathbb{Z}^m_+$, $f \in \mathcal{C}^{|I|}(G)$, $x, z, u, w \in G$. *Consider the functions* $g_{x,z} : G \to \mathbb{R}$ *and* $h_{x,u,w} : G \to \mathbb{R}$ *defined by* $g_{x,z}(y) := f(xyz)$, $y \in G$ *and* $h_{x,u,w}(v) := \widetilde{S}^I g_{x,uvw}(e)$, $v \in G$. *Then*

$$\widetilde{S}^I g_{x,z}(y) = \sum\nolimits_{d(J)\geq d(I),\, |J|\leq |I|} P^I_J(x) X^J f(xyz), \tag{8}$$

where P_J^I *are homogeneous polynomials of degree* $d(J)-d(I)$, *and*

$$|\widetilde{X}^I g_{x,z}(y)| \leq c(G)\left(1+|x|^{(s-1)d(I)}+|y|^{(s-1)d(I)}\right)\|D^{(sd(I))}f\|_\infty,$$

$$|\widetilde{X}^J h_{x,u,w}(v)| \leq c(G)\left(1+|xuv|^{(s-1)(d(I)+d(J))}\right)\|D^{(s(d(I)+d(J)))}f\|_\infty.$$

Proof. Introducing the function $h_z : G \to \mathbb{R}$ defined by $h_z(u) := f(uz)$ and putting $u = xy$ we have

$$\begin{aligned}\widetilde{X}^I g_{x,z}(y) &= \widetilde{X}^I h_z(u) = \sum\nolimits_{d(J)\geq d(I),\,|J|\leq|I|} P_{IJ}(u)X^J h_z(u)\\ &= \sum\nolimits_{d(J)\geq d(I),\,|J|\leq|I|} P_{IJ}(xy)X^J f(xyz),\end{aligned}$$

where P_{IJ} are homogeneous polynomials of degree $d(J)-d(I)$. Combining this formula with the right-invariant analogue of (5) we obtain (8). Clearly $|J|\leq|I|$ implies $d(J)\leq s|J|\leq s|I|\leq sd(I)$, thus

$$\begin{aligned}|\widetilde{X}^I g_{x,z}(y)| &\leq c(G)\sum\nolimits_{d(J)\geq d(I),\,|J|\leq|I|} |xy|^{d(J)-d(I)}\|D^{(d(J))}f\|_\infty\\ &\leq c(G)\left(1+|xy|^{(s-1)d(I)}\right)\|D^{(sd(I))}f\|_\infty\\ &\leq c(G)\left(1+|x|^{(s-1)d(I)}+|y|^{(s-1)d(I)}\right)\|D^{(sd(I))}f\|_\infty.\end{aligned}$$

The last estimate can be proved similarly. □

Let $k\in\mathbb{N}\cup\{0\}$, $f\in\mathcal{C}^k(G)$ and $x\in G$. The *left Taylor polynomial of* f *at* x *of homogeneous degree* k is the unique polynomial $P_x^{(k)}$ of homogeneous degree k such that $X^I P_x^{(k)}(e) = X^I f(x)$ for $I\in\mathbb{Z}_+^m$ with $d(I)\leq k$. It is known (see Dieudonné [4] 19.5.8 and 19.9.5) that

$$P_x^{(k)}(y) = \sum\nolimits_{d(I)\leq k} S^I f(x) y^I.$$

We shall use the following *left stratified Taylor expansion* valid up to the k^{th} order (see Folland and Stein [6], 1.44 Corollary)

$$f(xy) = P_x^{(k)}(y) + R_{k+1}^f(x,y), \tag{9}$$

where

$$|R_{k+1}^f(x,y)| \leq c_k|y|^{k+1}\sup\left\{|X^I f(xz)| : d(I)=k+1,\ |z|\leq b^{k+1}|y|\right\},$$

and $b\geq 1$ is a constant (depending only on G). The right Taylor polynomials can be defined similarly, and a right stratified Taylor expansion is valid as well.

3. Moments of Gaussian measures

A family $(\nu_t)_{t\geq 0}$ of probability measures on G is said to be a *(continuous) convolution semigroup* if we have $\nu_s * \nu_t = \nu_{s+t}$ for all $s,t\geq 0$, and $\lim_{t\downarrow 0}\nu_t = \nu_0 = \varepsilon_e$, where ε_e is the Dirac measure in $e \in G$. Its *infinitesimal generator* $(N, \mathcal{N})$ is defined by

$$\mathcal{N} := \left\{ f \in \mathcal{C}^b(G) \,\Big|\, (Nf)(x) := \lim_{t \downarrow 0} t^{-1} \int (f(xy) - f(x)) \nu_t(dy) \quad \text{exists} \quad \right\},$$

where $\mathcal{C}^b(G)$ denotes the space of bounded continuous complex-valued functions on G equipped with the supremum norm $\|\cdot\|_\infty$. We have $\mathcal{D}(G) \subseteq \mathcal{N}$, where $\mathcal{D}(G)$ is the space of infinitely differentiable complex-valued functions with compact support on G.

A convolution semigroup $(\nu_t)_{t \geq 0}$ of non-degenerate probability measures is called a *Gaussian semigroup* if $\lim_{t \downarrow 0} t^{-1} \nu_t(G \setminus U) = 0$ for any neighborhood U of e. A non-degenerate probability measure ν is called a *Gaussian measure* if there exists a Gaussian semigroup $(\nu_t)_{t \geq 0}$ such that $\nu_1 = \nu$. If $(\mu_t)_{t \geq 0}$ and $(\nu_t)_{t \geq 0}$ are Gaussian semigroups with $\mu_1 = \nu_1$ then $\mu_t = \nu_t$ for all $t \geq 0$, see Pap [17]. If $(\nu_t)_{t \geq 0}$ is a Gaussian semigroup then its infinitesimal generator N has the form

$$(Nf)(x) = \sum\nolimits_{i=1}^m a_i (X_i f)(x) + \sum\nolimits_{i,j=1}^m a_{ij} (X_i X_j f)(x), \tag{10}$$

where $a_1, \ldots, a_m$ are real numbers, and $(a_{ij})_{1 \leq i,j \leq m}$ is a real symmetric positive semidefinite matrix. Furthermore, if $a_1, \ldots, a_m$ are real numbers, and $(a_{ij})_{1 \leq i,j \leq m}$ is a real symmetric positive semidefinite matrix then there exists exactly one convolution semigroup $(\nu_t)_{t \geq 0}$ with infinitesimal generator (10) (see Heyer [11]).

A probability measure μ on G is called *centered* if $\int x_i \mu(dx) = 0$ for $i \in \{1, \ldots, m\}$ with $d_i = 1$. A Gaussian semigroup $(\nu_t)_{t \geq 0}$ on G is centered and stable with respect to the one-parameter group $(\delta_{\sqrt{t}})_{t>0}$ of automorphisms in the sense of Hazod [7] (i.e. $\nu_t = \delta_{\sqrt{t}} \nu_1$ for $t > 0$) if and only if its infinitesimal generator has the form

$$\sum\nolimits_{d_i=2} a_i X_i + \frac{1}{2} \sum\nolimits_{d_i=d_j=1} a_{ij} X_i X_j \tag{11}$$

(see Pap [14]). In particular, if $(\nu_t)_{t \geq 0}$ is a Gaussian semigroup on G with infinitesimal generator (11) then $\nu_1 = \delta_{1/\sqrt{n}} \nu_1^n$ for $n \in \mathbb{N}$. It is also known that $\int \exp\{\gamma |x|^2\} \nu(dx) < \infty$ for $\gamma \leq C(G, \nu)$, where $C(G, \nu) > 0$ (see in Hebisch [10] the exponential bound for the density of ν with respect to the Haar measure). Hence all moments of ν are finite.

We shall need the following statement about the moments of a Gaussian measure.

Lemma 2. *Let ν be a Gaussian measure on G with infinitesimal generator (11). Then*

$$\int x^I \nu(dx) = \begin{cases} 0 & \text{if } d(I) \text{ is odd} \\ a_I & \text{if } d(I) = 2, \end{cases}$$

where

$$a_I = \begin{cases} a_i & \text{if } I = [i] \\ a_{ij} & \text{if } I = [i] + [j]. \end{cases}$$

Proof. Applying (3) and using $\nu = \delta_{1/\sqrt{2}}\nu^2 = (\delta_{1/\sqrt{2}}\nu)^2$ we obtain

$$\begin{aligned}\int x^I \nu(dx) &= \iint (xy)^I \delta_{1/\sqrt{2}}\nu(dx)\delta_{1/\sqrt{2}}\nu(dy) = 2^{-d(I)/2}\iint (xy)^I \nu(dx)\nu(dy)\\ &= 2^{-d(I)/2}\left(2\int x^I\nu(dx) + \sum\nolimits_{J,K\neq 0,\, d(J)+d(K)=d(I)} c^I_{JK}\int x^J\nu(dx)\int x^K\nu(dy)\right).\end{aligned}$$

For $d(I) = 1$ we get

$$\int x^I\nu(dx) = 2^{1/2}\int x^I\nu(dx),$$

hence $\int x^I\nu(dx) = 0$. If $d(I)$ is odd and $d(J)+d(K) = d(I)$ with $J, K \neq 0$ then $d(J)$ or $d(K)$ is odd and less then $d(I)$, hence by induction we obtain $\int x^I\nu(dx) = 0$.

For the case $d(I) = 2$ we calculate directly the infinitesimal generator of the Gaussian semigroup $(\nu_t)_{t\geq 0}$ with $\nu_1 = \nu$. Using the stability of $(\nu_t)_{t\geq 0}$ we have for $f \in \mathcal{D}(G)$,

$$\int (f(xy) - f(x))\nu_t(dy) = \int (f(xy) - f(x))\delta_{\sqrt{t}}\nu(dy) = \int (f(x\delta_{\sqrt{t}}(y)) - f(x))\nu(dy).$$

The left stratified Taylor expansion up to the second order for the function f at $x \in G$ yields

$$f(x\delta_{\sqrt{t}}(y)) = \sum_{d(I)\leq 2} S^I f(x) t^{d(I)/2} y^I + R^f(x,y)$$

where

$$|R^f(x,y)| \leq ct^{3/2}|y|^3\|D^{(3)}f\|_\infty .$$

Taking into account that the measure ν is centered we obtain

$$\begin{aligned}(Nf)(x) &= \lim_{t\downarrow 0} t^{-1}\int (f(xy)-f(x))\nu_t(dy) = \sum_{d(I)=2} S^I f(x)\int y^I\nu(dy)\\ &= \sum_{d_i=2} X_i f(x)\int y_i\nu(dy) + \frac{1}{2}\sum_{d_i=d_j=1} X_iX_j f(x)\int y_iy_j\nu(dy).\end{aligned}$$

The proof is complete. □

We introduce the following moments and truncated moments.

$$\begin{aligned}m_\beta(\mu) &= \int |x|^\beta \mu(dx)\\ \Lambda_\beta(\mu) &= n^{(2-\beta)/2}\int_{|x|\geq\sqrt{n}} |x|^\beta\mu(dx)\\ L_\beta(\mu) &= n^{(2-\beta)/2}\int_{|x|<\sqrt{n}} |x|^\beta\mu(dx).\end{aligned}$$

Let μ be a centered probability measure on G. If $m_2(\mu) < \infty$ then by Lemma 2 there exists a centered Gaussian measure ν such that the moments of μ and ν are equal up to the second homogeneous order (i.e. $\int x^I\mu(dx) = \int x^I\nu(dx)$ for $d(I) = 1, 2$) and the infinitesimal generator of the corresponding Gauss semigroup is defined by (11). We shall say that ν *is the Gaussian measure associated with* μ and write $\nu = \mathrm{Ass}(\mu)$. Furthermore, the measures $\delta_{1/\sqrt{n}}\mu^n$ converge weakly to ν, i.e. μ satisfies the CLT (see Raugi [19]). We remark that $m_2(\mu) < \infty$ is not necessary for the CLT in general (see Scheffler [21]).

Lemma 3. *Let ν be a Gaussian measure on G with infinitesimal generator (11). Let $\beta \geq 2$. Then there exists a constant $C(G,\nu,\beta) > 0$ such that*

$$m_\beta(\nu) \leq C(G,\nu,\beta) m_\beta(\mu)$$

for any centered probability measure μ on G with $m_2(\mu) < \infty$ and $\mathrm{Ass}(\mu) = \nu$.

Proof. If $\exists i \in \{1,\ldots,m\}$ with $d_i = 1$ and $a_{ii} > 0$ then

$$\begin{aligned} m_\beta(\mu) &= \int \left(\sum_{i=1}^m |x_i|^{1/d(i)} \right)^\beta \mu(dx) \geq c(G,\beta) \left(\sum_{d_i=1} \int x_i^2 \mu(dx) \right)^{\beta/2} \\ &= c(G,\beta) \left(\sum_{d_i=1} a_{ii}^2 \right)^{\beta/2} = m_\beta(\nu)/C(G,\nu,\beta) \end{aligned}$$

where $C(G,\nu,\beta) = m_\beta(\nu)(c(G,\beta))^{-1} \left(\sum_{d_i=1} a_{ii}^2 \right)^{-\beta/2}$.

If $\exists i \in \{1,\ldots,m\}$ with $d_i = 2$ and $a_i \neq 0$ then

$$\begin{aligned} m_\beta(\mu) &\geq c(G,\beta) \left(\sum_{d_i=2} \int |x_i| \mu(dx) \right)^{\beta/2} \geq c(G,\beta) \left(\sum_{d_i=2} \left| \int x_i \mu(dx) \right| \right)^{\beta/2} \\ &= c(G,\beta) \left(\sum_{d_i=2} |a_i| \right)^{\beta/2} = m_\beta(\nu)/C(G,\nu,\beta) \end{aligned}$$

where $C(G,\nu,\beta) = m_\beta(\nu)(c(G,\beta))^{-1} \left(\sum_{d_i=2} |a_i| \right)^{-\beta/2}$. □

Lemma 4. *Let ν be a Gaussian measure on G with infinitesimal generator (11). Let $\gamma \geq \beta \geq 2$. Then there exists a constant $C(G,\nu,\beta,\gamma) > 0$ such that*

$$\Lambda_\beta(\nu) + L_\gamma(\nu) \leq C(G,\nu,\beta,\gamma)\left(\Lambda_\beta(\mu) + L_\gamma(\mu)\right)$$

for any centered probability measure μ on G with $m_2(\mu) < \infty$ and $\nu = \mathrm{Ass}(\mu)$.

Proof. First we remark that

$$\Lambda_\beta(\nu) + L_\gamma(\nu) \leq 2 m_\gamma(\nu) n^{(2-\gamma)/2}$$

since

$$\Lambda_\beta(\nu) \leq n^{(2-\gamma)/2} \int_{|x| \geq \sqrt{n}} |x|^\gamma \nu(dx) \leq m_\gamma(\nu) n^{(2-\gamma)/2}$$

and

$$L_\gamma(\nu) \leq m_\gamma(\nu) n^{(2-\gamma)/2}.$$

In the case if $\Lambda_\beta(\mu) \geq n^{(2-\beta)/2} m_\beta(\mu)/2$ we have

$$\Lambda_\beta(\mu) + L_\gamma(\mu) \geq \Lambda_\beta(\mu) \geq n^{(2-\beta)/2} m_\beta(\mu)/2 \geq n^{(2-\gamma)/2} c(G,\nu,\beta) \geq \frac{\Lambda_\beta(\nu) + L_\gamma(\nu)}{C(G,\nu,\beta,\gamma)}$$

where $C(G,\nu,\beta,\gamma) = 2m_\gamma(\nu)/c(G,\nu,\beta)$.

In the case if $\Lambda_\beta(\mu) \leq n^{(2-\beta)/2} m_\beta(\mu)/2$ we have $L_\beta(\mu) \geq n^{(2-\beta)/2} m_\beta(\mu)$ hence

$$\int_{|x|<\sqrt{n}} |x|^\beta \mu(dx) \geq m_\beta(\mu)/2 \geq c(G,\nu,\beta) > 0.$$

By Hölder inequality

$$\int_{|x|<\sqrt{n}} |x|^\beta \mu(dx) \leq \left(\int_{|x|<\sqrt{n}} |x|^\gamma \mu(dx) \right)^{\beta/\gamma}$$

hence

$$\int_{|x|<\sqrt{n}} |x|^\gamma \mu(dx) \geq (c(G,\nu,\beta))^{\gamma/\beta}.$$

Consequently,

$$\Lambda_\beta(\mu) + L_\gamma(\mu) \geq L_\gamma(\mu) \geq (c(G,\nu,\beta))^{\gamma/\beta} n^{(2-\gamma)/2} \geq \frac{\Lambda_\beta(\nu) + L_\gamma(\nu)}{C(G,\nu,\beta,\gamma)}$$

where $C(G,\nu,\beta,\gamma) = 2m_\gamma(\nu)(c(G,\nu,\beta))^{-\gamma/\beta}$. □

4. Rate of convergence

Let μ be a centered probability measure on G with $m_2(\mu) < \infty$. Let $\nu = \mathrm{Ass}(\mu)$ be the Gaussian measure associated with μ (see Section 3). For the sake of simplicity we shall use the notation

$$\tau = n^{-1/2}, \qquad \text{for} \quad n = 1, 2, \ldots$$

where it does not cause misunderstanding. Put

$$\Delta_n(f) = \int f(x)(\delta_\tau \mu^n - \nu)(dx).$$

Theorem 1. *If* $m_2(\mu) < \infty$ *and* $f \in \mathcal{C}^{3s}(G)$ *then*

$$|\Delta_n(f)| \leq r_1 + r_2 + r_3$$

where

$$\begin{aligned}
r_1 &= \|f\|_\infty \Lambda_0(\mu + \nu) \\
r_2 &= C(G)\|D^{(2s)} f\|_\infty (1 + m_{2(s-1)}(\nu)) \Lambda_2(\mu + \nu) \\
r_3 &= C(G)\|D^{(3s)} f\|_\infty (1 + m_{3(s-1)}(\nu)) L_3(\mu + \nu).
\end{aligned}$$

Corollary 1. *If* $m_2(\mu) < \infty$ *and* $f \in \mathcal{C}^{3s}(G)$ *then*

$$|\Delta_n(f)| \le C(G,\nu)\|D^{(3s)}f\|_\infty(\Lambda_2(\mu) + L_3(\mu)).$$

In particular,

$$|\Delta_n(f)| \le C(G,\nu)\|D^{(3s)}f\|_\infty m_\beta(\mu) n^{(2-\beta)/2}$$

for $2 \le \beta \le 3$. *If* $m_\beta(\mu) < \infty$ *then* $\Delta_n(f) = o\left(n^{(2-\beta)/2}\right)$ *as* $n \to \infty$ *provided* $2 \le \beta < 3$. *The condition* $m_2(\mu) < \infty$ *implies the CLT for* μ.

Proof. Using the decomposition

$$\begin{aligned}
\delta_\tau\mu^n - \nu &= (\delta_\tau\mu)^n - (\delta_\tau\nu)^n = \sum_{k=1}^{n}(\delta_\tau\nu)^{k-1} * (\delta_\tau\mu - \delta_\tau\nu) * (\delta_\tau\mu)^{n-k} \\
&= \sum_{k=1}^{n} \delta_\tau\nu^{k-1} * \delta_\tau(\mu - \nu) * \delta_\tau\mu^{n-k}
\end{aligned}$$

we obtain

$$\int f(x)(\delta_\tau\mu^n - \nu)(dx) = \sum_{k=1}^{n} \iiint f(xyz)\delta_\tau\nu^{k-1}(dx)\delta_\tau(\mu-\nu)(dy)\delta_\tau\mu^{n-k}(dz). \qquad (12)$$

Splitting the integrals into to parts we get

$$\int f(x)(\delta_\tau\mu^n - \nu)(dx) = I_1 + I_2 \qquad (13)$$

where

$$\begin{aligned}
I_1 &= \sum_{k=1}^{n} \iiint_{|y|\ge 1} f(xyz)\delta_\tau\nu^{k-1}(dx)\delta_\tau(\mu-\nu)(dy)\delta_\tau\mu^{n-k}(dz) \\
I_2 &= \sum_{k=1}^{n} \iiint_{|y|< 1} f(xyz)\delta_\tau\nu^{k-1}(dx)\delta_\tau(\mu-\nu)(dy)\delta_\tau\mu^{n-k}(dz).
\end{aligned}$$

We have

$$|I_1| \le \|f\|_\infty \Lambda_0(\mu + \nu). \qquad (14)$$

To estimate I_2 we use the right stratified Taylor expansion up to the second order for the function $g_{x,z}(y) = f(xyz)$ at $e \in G$ for fixed $x, z \in G$.

$$g_{x,z}(y) = \sum\nolimits_{d(I)\le 2} \widetilde{S}^I g_{x,z}(e) y^I + R^{g_{x,z}}(y),$$

where

$$|R^{g_{x,z}}(y)| \le c|y|^3 \sup\left\{|\widetilde{X}^I g_{x,z}(u)| : d(I) = 3,\ |u| \le b^3|y|\right\}.$$

For $d(I) \le 2$ we have

$$\int y^I \nu(dy) = \int y^I \mu(dy) \qquad (15)$$

hence

$$\left|\int_{|y|<1} y^I \delta_\tau(\mu-\nu)(dy)\right| = \left|\int_{|y|\ge 1} y^I \delta_\tau(\mu-\nu)(dy)\right| \le \int_{|y|\ge 1} |y|^2 \delta_\tau(\mu+\nu)(dy) = n^{-1}\Lambda_2(\mu+\nu)$$

and Lemma 1 gives the estimate

$$|\widetilde{S}^I g_{x,z}(e)| \le c(G)\left(1+|x|^{d(I)(s-1)}\right)\|D^{(sd(I))}f\|_\infty \le c(G)\left(1+|x|^{2(s-1)}\right)\|D^{(2s)}f\|_\infty$$

thus

$$\left|\int_{|y|<1} \widetilde{S}^I g_{x,z}(e) y^I \delta_\tau(\mu-\nu)(dy)\right| \le n^{-1} c(G)\|D^{(2s)}f\|_\infty \left(1+|x|^{2(s-1)}\right)\Lambda_2(\mu+\nu). \tag{16}$$

Again by Lemma 1

$$|R^{g_{x,z}}(y)| \le c|y|^3\left(1+|x|^{3(s-1)}+|y|^{3(s-1)}\right)\|D^{(3s)}f\|_\infty$$

hence

$$\left|\int_{|y|<1} R^{g_{x,z}}(y)\delta_\tau(\mu-\nu)(dy)\right| \le n^{-1}c(G)\|D^{(3s)}f\|_\infty\left(1+|x|^{3(s-1)}\right)L_3(\mu+\nu).$$

Collecting the estimates we obtain $|I_2| \le r_2 + r_3$. □

5. Edgeworth expansion

We continue to use notations of Section 4.

Theorem 2. *If* $m_2(\mu)<\infty$ *and* $f \in \mathcal{C}^{6s}(G)$ *then*

$$\int f(x)\delta_{1/\sqrt{n}}\mu^n(dx) = \int f(x)\nu(dx) + \alpha_n n^{-1/2} + R_n \tag{17}$$

where $\alpha_n = \alpha_n(G,f,\mu)$ *and* $R_n = R_n(G,f,\mu)$ *depend only on* G, f, μ *and* n. *Furthermore,*

$$\alpha_n = \sum_{d(I)=3} \iiint_0^1 \widetilde{S}^I g_{x,z}(e)\nu_t(dx)\nu_{1-t}(dz)dt \int_{|y|<\sqrt{n}} y^I(\mu-\nu)(dy)$$

and $|R_n| \le \sum_{i=1}^7 r_i$, *where*

$$\begin{aligned}
r_1 &= \|f\|_\infty \Lambda_0(\mu+\nu)\\
r_2 &= C(G)\|D^{(2s)}f\|_\infty(1+m_{2(s-1)}(\nu))\Lambda_2(\mu+\nu)\\
r_3 &= C(G)\|D^{(3s)}f\|_\infty(1+m_{3(s-1)}(\nu))\Lambda_0(\mu+\nu)L_3(\mu+\nu)\\
r_4 &= C(G)\|D^{(4s)}f\|_\infty(1+m_{4(s-1)}(\nu))L_4(\mu+\nu)\\
r_5 &= C(G)\|D^{(5s)}f\|_\infty(1+m_{5(s-1)}(\nu))\Lambda_2(\mu+\nu)L_3(\mu+\nu)\\
r_6 &= C(G)\|D^{(6s)}f\|_\infty(1+m_{6(s-1)}(\nu))(L_3(\mu+\nu))^2\\
r_7 &= C(G)\|D^{(4s)}f\|_\infty(m_1(\nu)+m_{4s-3}(\nu))L_3(\mu+\nu)n^{-1}.
\end{aligned}$$

Corollary 2. *The coefficient* α_n *in (17) satisfies*

$$|\alpha_n| \le C(G)\|D^{(3s)}f\|_\infty(1+m_{3(s-1)}(\nu))L_3(\mu+\nu)n^{1/2}$$

and

$$|\alpha_n-\alpha| \le C(G)\|D^{(3s)}f\|_\infty(1+m_{3(s-1)}(\nu))\Lambda_3(\mu+\nu)n^{1/2}$$

where $\alpha=\alpha(G,f,\mu)$ *depends only on* G, f, μ *and*

$$\alpha = \sum_{d(I)=3}\iiint_0^1 \tilde S^I g_{x,z}(e)\nu_t(dx)\nu_{1-t}(dz)dt\int y^I(\mu-\nu)(dy).$$

If $m_\beta(\mu)<\infty$ *with* $3\le\beta\le4$ *then*

$$\int f(x)\delta_{1/\sqrt{n}}\mu^n(dx) = \int f(x)\nu(dx)+\alpha n^{-1/2}+\tilde R_n \tag{18}$$

where $\tilde R_n=\tilde R_n(G,f,\mu)$ *depends only on* G, f, μ *and* n. *Moreover,*

$$|\tilde R_n| \le C(G,\nu)\|D^{(6s)}f\|_\infty\left(m_\beta(\mu)+n^{(\beta-4)/2}(m_3(\mu))^2+n^{(\beta-5)/2}m_3(\mu)\right)n^{(2-\beta)/2}.$$

If $m_\beta(\mu)<\infty$ *with* $3\le\beta<4$ *then* $\Delta_n(f)-\alpha n^{-1/2}=o\left(n^{(2-\beta)/2}\right)$ *as* $n\to\infty$.

Remark 1. It should be mentioned that

$$\begin{aligned}\tilde S^I g_{x,z}(e) &= S^I g_{x,z}(e)\\ &= \frac{1}{|I|!}\sum\nolimits_{[j_1]+\ldots+[j_{|I|}]=I}\frac{d^{|I|}}{dt_1\ldots dt_{|I|}}f\left(x(\exp t_1X_{j_1})\ldots(\exp t_{|I|}X_{j_{|I|}})z\right)\Big|_{t_1=0,\ldots,t_{|I|}=0}\end{aligned}$$

hence the terms α_n and α in the Edgeworth expansions (17) and (18) are independent of the concrete choice of differential calculus on the group G. Using Lemma 1 and the representation of the differential operators X^J in terms ∂^J we obtain

$$\tilde S^I g_{x,z}(e) = \sum\nolimits_{d(J)\ge d(I),\,|J|\le|I|}P^I_J(x,z)\partial^J f(xz),$$

where P^I_J are homogeneous polynomials of degree $d(J)-d(I)$.

Remark 2. Let us formulate the result in the language of random elements. Let $\xi,\xi_1,\xi_2,\ldots$ be i. i. d. random elements taking values in G with distribution μ. Let $(\eta_t)_{t\ge0}$ be a Wiener process in G with infinitesimal generator (11). Let $\eta:=\eta_1$. We restrict ourselves to the case $\mathbb{E}|\xi|^4<\infty$. Then for $f\in\mathcal{C}^{6s}(G)$,

$$\mathbb{E}f(\delta_{1/\sqrt{n}}(\xi_1\cdots\xi_n)) = \mathbb{E}f(\eta)+\alpha n^{-1/2}+\tilde R_n$$

where

$$\alpha = \sum\nolimits_{d(I)=3}\sum\nolimits_{d(J)\ge3,\,|J|\le|I|}\mathbb{E}P^I_J(\eta_U,\eta_U^{-1}\eta_1)\partial^J f(\eta_1)\mathbb{E}\xi^I,$$

U is a random variable with uniform distribution on $[0,1]$ independent of the Wiener process $(\eta_t)_{t\geq 0}$, and

$$|\widetilde{R}_n| \leq C(G,\nu)\|D^{(6s)}f\|_\infty \left(\mathbb{E}|\xi|^4 + (\mathbb{E}|\xi|^3)^2 + n^{-1/2}\mathbb{E}|\xi|^3\right) n^{-1}.$$

Proof. We use again the decomposition (12) and the splitting (13). We have the bound (14).

For estimation of I_2 we use now the right stratified Taylor expansion up to the third order for the function $g_{x,z}(y) = f(xyz)$ at $e \in G$ for fixed $x, z \in G$.

$$g_{x,z}(y) = \sum\nolimits_{d(I)\leq 3} \widetilde{S}^I g_{x,z}(e) y^I + R^{g_{x,z}}(y),$$

where

$$|R^{g_{x,z}}(y)| \leq c|y|^4 \sup\left\{|\widetilde{X}^I g_{x,z}(u)| : d(I) = 4,\ |u| \leq b^4|y|\right\}.$$

For $d(I) \leq 2$ we have again (16). Again by Lemma 1

$$\left|\int_{|y|<1} R^{g_{x,z}}(y)\delta_\tau(\mu-\nu)(dy)\right| \leq n^{-1}c(G)\|D^{(4s)}f\|_\infty \left(1+|x|^{4(s-1)}\right) L_4(\mu+\nu).$$

Hence

$$I_2 = A + r$$

where

$$A = \sum_{k=1}^{n}\sum_{d(I)=3} \iint \widetilde{S}^I g_{x,z}(e)\delta_\tau\nu^{k-1}(dx)\delta_\tau\mu^{n-k}(dz)\int_{|y|<1} y^I\delta_\tau(\mu-\nu)(dy)$$

and

$$|r| \leq r_2 + r_4.$$

Using the decomposition

$$\delta_\tau\mu^{n-k} = \delta_\tau\nu^{n-k} + \sum_{l=1}^{n-k}\delta_\tau\nu^{l-1} * \delta_\tau(\mu-\nu) * \delta_\tau\mu^{n-k-l}$$

we obtain

$$A = (A_1 + A_2)\int_{|y|<1} y^I\delta_\tau(\mu-\nu)(dy),$$

where

$$A_1 = \sum_{d(I)=3}\sum_{k=1}^{n}\iint \widetilde{S}^I g_{x,z}(e)\delta_\tau\nu^{k-1}(dx)\delta_\tau\nu^{n-k}(dz)$$

and

$$A_2 = \sum_{d(I)=3}\sum_{k=1}^{n}\sum_{l=1}^{n-k}\iiiint \widetilde{S}^I g_{x,uvw}(e)\delta_\tau\nu^{k-1}(dx)\delta_\tau\nu^{l-1}(du)\delta_\tau(\mu-\nu)(dv)\delta_\tau\mu^{n-k-l}(dw).$$

First we analyse A_2. Splitting again the integrals into two parts we obtain

$$A_2 = A_{2,1} + A_{2,2}$$

where $A_{2,1}$ contains the integrals on the set $\{|v| \geq 1\}$ and $A_{2,2}$ contains the integrals on the set $\{|v| < 1\}$. Using Lemma 1 we get

$$|A_{2,1}| \leq nc(G)\|D^{(3s)}f\|_\infty \left(1 + m_{3(s-1)}(\nu)\right) \Lambda_0(\mu+\nu).$$

To estimate $A_{2,2}$ we use now the right stratified Taylor expansion up to the second order for the function $h_{x,u,w}(v) = \widetilde{S}^I g_{x,uvw}(e)$ at $e \in G$ for fixed $x, u, w \in G$.

$$h_{x,u,w}(v) = \sum\nolimits_{d(J)\leq 2} \widetilde{S}^J h_{x,u,w}(e) v^J + R^{h_{x,u,w}}(v),$$

where

$$|R^{h_{x,u,w}}(v)| \leq c|v|^3 \sup\left\{|\widetilde{X}^J h_{x,u,w}(z)| : d(J) = 3,\ |z| \leq b^3|v|\right\}.$$

For $d(J) \leq 2$ Lemma 1 gives the estimate

$$\left|\int_{|v|<1} \widetilde{S}^J h_{x,u,w}(e) v^J \delta_\tau(\mu-\nu)(dv)\right| \leq n^{-1}c(G)\|D^{(5s)}f\|_\infty \left(1 + |x|^{5(s-1)}\right) \Lambda_2(\mu+\nu)$$

(see the proof of (16)). Again by Lemma 1,

$$\left|\int_{|v|<1} R^{h_{x,u,w}}(v)\delta_\tau(\mu-\nu)(dv)\right| \leq n^{-1}c(G)\|D^{(6s)}f\|_\infty \left(1 + |x|^{6(s-1)}\right) L_3(\mu+\nu).$$

Taking into account the estimate

$$\left|\int_{|y|<1} y^I \delta_\tau(\mu-\nu)(dy)\right| \leq n^{-1} L_3(\mu+\nu)$$

we obtain

$$\left|A_2 \int_{|y|<1} y^I \delta_\tau(\mu-\nu)(dy)\right| \leq r_3 + r_5 + r_6.$$

Now we analyse A_1. Remarking the stability of the convolution semigroup $(\nu_t)_{t\geq 0}$ we obtain

$$\begin{aligned} n^{-1} &\sum_{k=1}^{n} \iint \widetilde{S}^I g_{x,z}(e) \delta_\tau \nu^{k-1}(dx) \delta_\tau \nu^{n-k}(dz) \\ &= n^{-1} \sum_{k=1}^{n} \iint \widetilde{S}^I g_{x,z}(e) \delta_{\sqrt{(k-1)/n}} \nu(dx) \delta_{\sqrt{(n-k)/n}} \nu(dz) \\ &= \iiint_0^1 \widetilde{S}^I g_{x,z}(e) \delta_{\sqrt{[nt]/n}} \nu(dx) \delta_{\sqrt{[n(1-t)]/n}} \nu(dz) dt. \end{aligned}$$

We shall use the notation

$$\kappa = \sqrt{[nt]/n}, \qquad \lambda = \sqrt{[n(1-t)]/n}.$$

Using equation (8) we have

$$\iint \widetilde{S}^I g_{x,z}(e)\delta_\kappa\nu(dx)\delta_\lambda\nu(dz) = \iint \widetilde{S}^I g_{x,z}(e)\delta_{\sqrt{t}}\nu(dx)\delta_{\sqrt{1-t}}\nu(dz) + B_1 + B_2,$$

where

$$B_1 = \sum \iint P^I_J(\delta_\kappa(x)\delta_\lambda(z))(X^J f(\delta_\kappa(x)\delta_\lambda(z)) - X^J f(\delta_{\sqrt{t}} x\delta_{\sqrt{1-t}} z))\nu(dx)\nu(dz)$$

$$B_2 = \sum \iint X^J f(\delta_{\sqrt{t}}(x)\delta_{\sqrt{1-t}}(z)))(P^I_J(\delta_\kappa(x)\delta_\lambda(z)) - P^I_J(\delta_{\sqrt{t}}(x)\delta_{\sqrt{1-t}}(z)))\nu(dx)\nu(dz),$$

the sums are taken over $J \in \mathbb{Z}^m_+$ with $d(J) \geq 3$, $|J| \leq |I|$, and P^I_J are homogeneous polynomials of degree $d(J) - d(I)$.

Clearly

$$\delta_\kappa(x) = \delta_{\sqrt{t-\kappa}}(x^{-1})\delta_{\sqrt{t}}(x),$$

hence applying the right stratified Taylor expansion up to the second order and formula (7) we get

$$\begin{aligned}
&\left|X^J f(\delta_\kappa(x)\delta_\lambda(z)) - X^J f(\delta_{\sqrt{t}}(x)\delta_\lambda(z))\right| \\
&\quad\leq c|\delta_{\sqrt{t-\kappa}}(x)| \sup\left\{\left|\widetilde{X}^K X^J f(u\delta_\kappa(x)\delta_\lambda(z))\right| : d(K) = 1,\ |u| \leq b|\delta_{\sqrt{t-\kappa}}(x)|\right\} \\
&\quad\leq n^{-1}c(G)t^{-1/2}|x|(1 + |x|^{4(s-1)} + |z|^{4(s-1)})\|D^{(4s)}f\|_\infty.
\end{aligned}$$

Applying the left stratified Taylor expansion up to the second order and the formula (6) we obtain

$$\begin{aligned}
&\left|X^J f(\delta_{\sqrt{t}}(x)\delta_\lambda(z)) - X^J f(\delta_{\sqrt{t}}(x)\delta_{\sqrt{1-t}}(z))\right| \\
&\quad\leq c|\delta_{\sqrt{1-t-\lambda}}(z)| \sup\left\{\left|X^K X^J f(\delta_{\sqrt{t}}(x)\delta_\lambda(z)u)\right| : d(K) = 1,\ |u| \leq b|\delta_{\sqrt{1-t-\lambda}}(z)|\right\} \\
&\quad\leq n^{-1}c(G)(1-t)^{-1/2}|z|(1 + |x|^{3(s-1)} + |z|^{3(s-1)})\|D^{(3s+1)}f\|_\infty.
\end{aligned}$$

Thus using the Hölder inequality we conclude

$$|B_1| \leq n^{-1}c(G)(t^{-1/2} + (1-t)^{-1/2})(m_1(\nu) + m_{4s-3}(\nu))\|D^{(4s)}f\|_\infty.$$

For estimating B_2 first we remark that for a homogeneous polynomial $P(x, y)$ of homogeneous degree d we have

$$|P(x, y) - P(z, u)| \leq c_P \textstyle\sum_{d(I)+d(J)=d,\, I\neq 0} \left(|x^I - z^I||y^J| + |y^I - u^I||z^J|\right).$$

Applying formula (3) we get

$$\left|(\delta_\kappa(x))^I - (\delta_{\sqrt{t}}(x))^I\right| \leq n^{-1}c(G)t^{-1/2}|x|^{d(I)}.$$

Using the Hölder inequality we conclude

$$|B_2| \leq n^{-1}c(G)(t^{-1/2} + (1-t)^{-1/2})(m_1(\nu) + m_{3(s-1)}(\nu))\|D^{(3s)}f\|_\infty.$$

Collecting the estimates for $|B_1|$ and $|B_2|$ we conclude

$$A_1 \int_{|y|<1} y^I \delta_\tau(\mu - \nu)(dy) = \alpha_n n^{-1/2} + r,$$

where $|r| \leq r_7$. Hence the assertion. □

6. Example: the Heisenberg group

Let us consider the simplest non-commutative nilpotent Lie group, the Heisenberg group H, which can be obtained furnishing $\mathbb{R}^3$ with its natural topology and with the product

$$(x_1, x_2, x_3)(y_1, y_2, y_3) = (x_1 + y_1, x_2 + y_2, x_3 + y_3 + \frac{1}{2}(x_1 y_2 - x_2 y_1)).$$

The corresponding Lie algebra $\mathcal{H}$ of the Lie group H can be realized as the vector space $\mathbb{R}^3$ furnished with the multiplication

$$[(x_1, x_2, x_3), (y_1, y_2, y_3)] = (0, 0, x_1 y_2 - x_2 y_1).$$

The natural basis of $\mathcal{H}$ is denoted by $\{X_1, X_2, X_3\}$.

It is well known that the left invariant Brownian motion on H with infinitesimal generator

$$a_3 X_3 + \frac{1}{2} \sum_{i=1}^{2} \sum_{j=1}^{2} a_{i,j} X_i X_j$$

(where $A = (a_{i,j})_{i,j=1,2}$ is a symmetric positive semidefinite matrix) can be obtained as

$$(Z_1(t),\, Z_2(t),\, L(t))\,, \qquad t \geq 0,$$

where $(Z_1(t), Z_2(t))_{t \geq 0}$ is a Brownian motions on $\mathbb{R}^2$ with zero mean and covariance A, and

$$L(t) := a_3 t + \frac{1}{2} \int_0^t (Z_1(t)\, dZ_2(t) - Z_2(t)\, dZ_1(t))$$

is the Lévy's stochastic area process (see, for example, Feinsilver and Schott [5] and Roynette [20]).

Let us formulate our results in terms of random elements.

Let $(\xi, \eta), (\xi_1, \eta_1), (\xi_2, \eta_2), \ldots \in \mathbb{R}^2$ and $\zeta, \zeta_1, \zeta_2, \ldots \in \mathbb{R}^1$ denote i. i. d. sequences. Assume that $\mathbb{E}\xi = \mathbb{E}\eta = 0$, $\mathbb{E}\zeta = a_3$, $\mathbb{E}\xi^2 = a_{1,1}$, $\mathbb{E}\eta^2 = a_{2,2}$, $\mathbb{E}\xi\eta = a_{1,2}$. Consider the statistic

$$T_n := \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i,\ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \eta_i,\ \frac{1}{n} \sum\nolimits_{i=1}^{n} \zeta_i + \frac{1}{2n} \sum\nolimits_{1 \leq i < j \leq n} (\xi_i \eta_j - \xi_j \eta_i) \right).$$

Then for $f \in \mathcal{C}^6(H)$,

$$|\mathbb{E}f(T_n) - \mathbb{E}f(Z_1(1), Z_2(1), L(1))| \leq C(H) \|D^{(6)} f\|_\infty (\mathbb{E}|\xi|^3 + \mathbb{E}|\eta|^3 + \mathbb{E}|\zeta|^{3/2}) n^{-1/2}.$$

If $\mathbb{E}|\xi|^4 < \infty$, $\mathbb{E}|\eta|^4 < \infty$ and $\mathbb{E}|\zeta|^2 < \infty$ then for $f \in \mathcal{C}^{12}(H)$,

$$\mathbb{E}f(T_n) = \mathbb{E}f(Z_1(1), Z_2(1), L(1)) + \alpha n^{-1/2} + \widetilde{R}_n,$$

where the remainder term $|\widetilde{R}_n| \leq r_1 + r_2$, and where

$$\begin{aligned} r_1 &= C(H)\|D^{(12)}f\|_\infty \left(\mathbb{E}|\xi|^4 + \mathbb{E}|\eta|^4 + \mathbb{E}|\zeta|^2 + (\mathbb{E}|\xi|^3)^2 + (\mathbb{E}|\eta|^3)^2 + (\mathbb{E}|\zeta|^{3/2})^2\right) n^{-1} \\ r_2 &= C(H)\|D^{(12)}f\|_\infty \left(\mathbb{E}|\xi|^3 + \mathbb{E}|\eta|^3 + \mathbb{E}|\zeta|^{3/2})\right) n^{-3/2}. \end{aligned}$$

The coefficient α is given by

$$\begin{aligned} \alpha = \quad & \frac{1}{6}\sum_{k=1}^{3}\binom{3}{k}\mathbb{E}D_1^k D_2^{3-k} f(W_1(1), W_2(1), L(1))\,\mathbb{E}\xi^k\eta^{3-k} \\ & +\frac{1}{3}\mathbb{E}D_1 D_3 f(W_1(1), W_2(1), L(1))\,\mathbb{E}\xi\zeta \\ & +\frac{1}{3}\mathbb{E}D_2 D_3 f(W_1(1), W_2(1), L(1))\,\mathbb{E}\eta\zeta, \end{aligned}$$

where random differential operators D_1, D_2 and D_3 are defined by

$$\begin{aligned} D_1 &= \partial_1 + \frac{W_2(1) - 2W_2(U)}{2}\partial_3 \\ D_2 &= \partial_2 + \frac{2W_1(U) - W_1(1)}{2}\partial_3 \\ D_3 &= \partial_3, \end{aligned}$$

and where U is a random variable with uniform distribution on $[0,1]$, independent of all other random variables.

References

[1] Bentkus, V., Götze, F., Paulauskas, V. and Račkauskas, A. (1990). The accuracy of Gaussian approximation in Banach spaces. Preprint No 90 – 100 Sonderforschungsbereich 343, Diskrete Strukturen in der Mathematik, Bielefeld.

[2] Bourbaki, N. (1960, 1972). *Éléments de Mathématique, Groupes et Algèbres de Lie, Chapitres 1,2,3., Actual. Scient. Ind. 1285, 1349.* Hermann, Paris.

[3] Crépel, P. and Raugi, A. (1978). Théorème central limite sur les groupes nilpotents. *Ann. Inst. H. Poincaré Probab. Statist.* **14** 145–164.

[4] Dieudonné, N. (1971). *Éléments d'Analyse, Tome IV, Chapitres XVIII à XX.* Gauthier–Villars Éditeur, Paris.

[5] Feinsilver, Ph. and Schott, R. (1989). An operator approach to processes on Lie groups. In: Probability Theory on Vector Spaces IV. Proceedings, Lancut, 1987. (Lecture Notes in Math., vol. 1391, pp. 59–65) Springer, Berlin Heidelberg New York Tokyo.

[6] Folland, G. B. and Stein, E. M. (1982). *Hardy Spaces on Homogeneous Groups.* Princeton University Press, New Jersey.

[7] Hazod, W. (1984). Stable and semistable probabilities on groups and on vector spaces. In: Szynal, D., Weron, A. (eds.) Probability Theory on Vector Spaces III. Proceedings, Lublin 1983. (Lecture Notes in Math., vol. 1080, pp. 69–89) Springer, Berlin Heidelberg New York Tokyo.

[8] Hazod, W. and Siebert, E. (1986). Continuous automorphism groups on a locally compact group contracting modulo a compact subgroup and applications to stable convolution semigroups. *Semigroup Forum* **33** 111–143.

[9] Hazod, W. and Scheffler, H. P. (1992). The domains of partial attraction of probabilities on groups and on vector spaces. To appear in *J. Theor. Prob.*

[10] Hebisch, W. (1989). Sharp pointwise estimate for kernels of the semigroup generated by sums of even powers of vector fields on homogeneous groups. *Studia Math.* **95** 93–106.

[11] Heyer, H. (1977). *Probability Measures on Locally Compact Groups.* Springer, Berlin Heidelberg New York.

[12] Khokhlov, Yu. S. (1987). The estimation of the rate of convergence in the integral limit theorem in the Euclidean motion group. In: Stability problems for stochastic models. Proceedings, Varna 1985. (Lecture Notes in Math., vol. 1233, pp. 1–10) Springer, Berlin New York.

[13] Nobel, S. (1991). Limit theorems for probability measures on simply connected nilpotent Lie groups. *J. Theoret. Probab.* **4** 261–284.

[14] Pap, G. (1991). Rate of convergence in CLT on stratified groups. *J. Multivariate Anal.* **38** 333–365.

[15] Pap, G. (1991). A new proof of the central limit theorem on stratified Lie groups. In: Heyer, H. (ed.) Probability Measures on Groups X. Proceedings, Oberwolfach 1990, pp. 329–336. Plenum Press, New York.

[16] Pap, G. (1994). Central limit theorems on nilpotent Lie groups. *Probability and Mathematical Statistics* **14** vol. 2.

[17] Pap, G. (1994). Uniqueness of embedding into a Gaussian semigroup on a nilpotent Lie group. *Archiv der Math.* **49** ?.

[18] Rachev, S. T. and Yukich, J. E. (1991). Rates of convergence of α-stable random motions. *J. Theor. Probab.* **4** 333–352.

[19] RAUGI, A. (1978). Théorème de la limite centrale sur les groupes nilpotents. *Z. Wahrsch. Verw. Gebiete* **43** 149–172.

[20] ROYNETTE, B. (1975). Croissance et mouvements browniens d'un groupe de Lie nilpotent et simplement connexe. Z. Wahrsch. verw. Gebiete **32** 133–138.

[21] SCHEFFLER, H. P. (1993). Domains of attraction of stable measures on the Heisenberg group. To appear.

[22] SIEBERT, E. (1981). Fourier analysis and limit theorems for convolution semigroups on a locally compact group. *Adv. Math.* **39** 111–154.

[23] SIEBERT, E. (1982). Absolute continuity, singularity, and supports of Gauss semigroups on a Lie group. *Monatsh. Math.* **93** 239–253.

[24] SIEBERT, E. (1986). Contractive automorphisms on locally compact groups. *Math. Zeitschrift* **191** 73–90.

[25] TAYLOR, M. E. (1986). *Noncommutative Harmonic Analysis.* (Mathematical Surveys and Monographs, no. 22) American Mathematical Society, Providence, Rhode Island.

[26] VIRTSER, A. D. (1974). Limit theorems for compositions of distribution on certain nilpotent Lie groups. *Theory of Probab. and Appl.* **19** 86–105.

Vidmantas Bentkus
Department of Mathematics
University of Bielefeld
Postfach 100131
33501 Bielefeld
Germany

bentkus@mathematik.uni-bielefeld.de

Gyula Pap
Institute of Mathematics
Lajos Kossuth University
Pf.12
H–4010 Debrecen
Hungary

papgy@math.klte.hu