Forum Math. Galley Proof, 13 pages
DOI 10.1515/forum-2012-0085
Forum Mathematicum

Faithful representations of infinite-dimensional nilpotent Lie algebras

Ingrid Beltiţă and Daniel Beltiţă

Communicated by Karl-Hermann Neeb

Abstract. For locally convex, nilpotent Lie algebras we construct faithful representations by nilpotent operators on a suitable locally convex space. In the special case of nilpotent Banach–Lie algebras we get norm continuous representations by bounded operators on Banach spaces.

Keywords. Nilpotent Banach–Lie group, polynomial, faithful representation.

2010 Mathematics Subject Classification. Primary 22E66; secondary 22E65, 22E25.

1 Introduction

The aim of this short note is to provide an affirmative answer to a question raised in the seminal paper by K.-H. Neeb [13, page 440, Problems VII.2 and VIII.6 (b)], regarding suitable versions of Birkhoff's embedding theorem in infinite dimensions. In particular, we prove that every nilpotent Banach–Lie algebra has a bounded faithful representation by nilpotent operators on a suitable Banach space; see Corollary 2.14 below.

In order to put this result in a proper perspective, we recall a few classical facts:

- Every finite-dimensional nilpotent Lie algebra can be realized as a Lie algebra of nilpotent matrices, by Birkhoff's embedding theorem; see [4].
- In a more general situation, every finite-dimensional Lie algebra can be faithfully represented by matrices, by Ado's theorem; see [1].
- The above statement of Ado's theorem cannot be directly extended to the Banach category, by replacing the finite-dimensional vector spaces by Banach spaces. In fact, it has been long known that there exist Banach–Lie algebras that cannot be faithfully represented by Banach space operators; see [8] and also [13, page 438, Remark VIII.7] for a broader discussion.

This research has been partially supported by the Grant of the Romanian National Authority for Scientific Research, CNCS-UEFISCDI, project number PN-II-ID-PCE-2011-3-0131.

Besides this old open problem concerning the extension of Ado's theorem to infinite dimensions, there also exists a current interest in questions regarding the finite-dimensional situation. Thus, one has been interested in finding good estimates for the size of the matrices that have to be used in order to realize a given nilpotent Lie algebra; see [6, 7, 10], and the references therein. Among the various methods developed in [6], the second one is in some sense related to the setting of the present paper, inasmuch as the dual of the universal enveloping algebra of any Lie algebra is closely related to the space of polynomial functions on that Lie algebra. Our point here is to provide an analytic version of that method which is eventually applicable to some (infinite-dimensional) topological Lie algebras, namely to locally convex, nilpotent Lie algebras ([11, 13]).

The main result of our paper is Theorem 2.13 and provides faithful representations for algebras of this type by nilpotent operators, and for the corresponding locally convex Lie groups by unipotent operators. In the special case of Banach–Lie algebras, we get in Corollary 2.14 bounded representations.

We conclude the introduction by pointing out that the aforementioned corollary provides a partial solution to a more general problem which was raised in [15] and is still open in its full generality: A quasinilpotent Banach–Lie algebra is a Banach–Lie algebra $\mathfrak{g}$ with the property that for every $x \in \mathfrak{g}$ the bounded linear operator $\mathrm{ad}_{\mathfrak{g}} x \colon \mathfrak{g} \to \mathfrak{g}$ is quasinilpotent, in the sense that its spectrum is equal to $\{0\}$ or equivalently $\lim_{n\to\infty} \|(\mathrm{ad}_{\mathfrak{g}} x)^n\|^{1/n} = 0$, where the operator norm is computed with respect to any norm which defines the topology of $\mathfrak{g}$. With this terminology, it was asked in [15, Question 2] whether every quasinilpotent Banach–Lie algebra admits a faithful continuous representation on some Banach space. It is clear that if a Banach–Lie algebra is nilpotent then it is also quasinilpotent, hence Corollary 2.14 in the present paper solves in the affirmative the above question in the special case of the nilpotent Banach–Lie algebras.

2 Main results

Notation 2.1. We will work in the following setting:

- Unless otherwise stated, $\mathfrak{g}$ is a Hausdorff, locally convex, nilpotent Lie algebra over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ with the nilpotency index denoted by $N+1$, where $N \in \mathbb{N}$. Thus, if we define $\mathfrak{g}^{(1)} = \mathfrak{g}$ and $\mathfrak{g}^{(j)} := [\mathfrak{g}, \mathfrak{g}^{(j-1)}]$ for $j \geq 1$, then we have $\mathfrak{g}^{(N)} \neq \{0\} = \mathfrak{g}^{(N+1)}$.
- $G = (\mathfrak{g}, *)$ is the corresponding simply connected, locally convex, nilpotent Lie group, whose group operation $*$ is defined by the Baker–Campbell–Hausdorff formula on $\mathfrak{g}$ (see [13, Example IV 1.6.6]). Note that in this case there exists a smooth exponential mapping, which is in fact the identity. Therefore

Note 1: Red parts indicate major changes. Please check them carefully.

we use the same letters to denote elements from $\mathfrak{g}$ and G, and use the notation G only when it is useful to emphasize the group structure.

- We denote by $\mathcal{C}(\mathfrak{g},\mathbb{K})$ the space of $\mathbb{K}$-valued continuous functions on $\mathfrak{g}$ endowed with the topology of uniform convergence on the bounded sets, and by

$$\lambda\colon G \to \mathrm{End}\,(\mathcal{C}(\mathfrak{g},\mathbb{K})), \quad (\lambda(x)\phi)(y) = \phi((-x) * y)$$

the corresponding left-regular representation of G. If $\phi \in \mathcal{C}(\mathfrak{g},\mathbb{K})$, $x, y \in \mathfrak{g}$, then we define

$$(\mathrm{d}\lambda(x)\phi)(y) := \lim_{t\to 0} \frac{\phi((-tx) * y) - \phi(y)}{t} \tag{2.1}$$

whenever the above limit exists. (Compare [3, Definition 2.1].)

- For every locally convex space $\mathcal{Y}$ over $\mathbb{C}$ and every integer $m \geq 0$ we denote by $\mathcal{P}_m(\mathfrak{g},\mathcal{Y})$ the linear space of $\mathcal{Y}$-valued, continuous polynomial functions of degree $\leq m$ on $\mathfrak{g}$ and by $\mathcal{P}^m(\mathfrak{g},\mathcal{Y})$ its subspace consisting of the homogeneous polynomial functions of degree m (see [5]). We endow $\mathcal{P}_m(\mathfrak{g},\mathcal{Y})$ with the topology of uniform convergence on the bounded sets. We have a topological direct sum decomposition

$$\mathcal{P}_m(\mathfrak{g},\mathcal{Y}) = \mathcal{P}^0(\mathfrak{g},\mathcal{Y}) \dotplus \mathcal{P}^1(\mathfrak{g},\mathcal{Y}) \dotplus \cdots \dotplus \mathcal{P}^m(\mathfrak{g},\mathcal{Y}), \tag{2.2}$$

and then it easily follows by [5, Theorem A, Theorem 2] that $\mathcal{P}_m(\mathfrak{g},\mathcal{Y})$ is a closed subspace of $\mathcal{C}(\mathfrak{g},\mathbb{K})$.

Remark 2.2. Assume here that $\mathfrak{g}$ is a Banach–Lie algebra and let $m \in \mathbb{N}$. The space $\mathcal{P}^m(\mathfrak{g},\mathbb{K})$ is a Banach space and one norm defining its topology can be described as follows. For $\phi \in \mathcal{P}^m(\mathfrak{g},\mathbb{K})$ we denote by $\tilde{\phi}\colon \mathfrak{g}\times\cdots\times\mathfrak{g} \to \mathbb{K}$ the symmetric m-linear bounded functional satisfying $\phi(x) = \tilde{\phi}(x,\ldots,x)$ for every $x \in \mathfrak{g}$. Then we set $\|\phi\|_{\mathcal{P}^m(\mathfrak{g},\mathbb{K})} := \|\tilde{\phi}\|$. (See [5, Proposition 1].) Here we use the norm

$$\|\tilde{\phi}\| = \sup\{|\tilde{\phi}(x_1,\ldots,x_m)| \mid x_1,\ldots,x_m \in \mathfrak{g},\ \|x_j\| \leq 1,\ j = 1,\ldots,m\}.$$

The norm $\|\cdot\|_{\mathcal{P}^m(\mathfrak{g},\mathbb{K})}$ is equivalent (see [5, Proposition 1]) to the norm

$$\phi \mapsto \sup_{\|x\|\leq 1} |\phi(x)|.$$

We have the following formula for the directional derivatives:

$$(\forall x, y \in \mathfrak{g}) \quad \phi'_y(x) = m\widetilde{\phi}(x, y, \ldots, y)$$

and this implies that the estimate

$$|\phi_y'(x)| \le m\|x\| \cdot \|y\|^{m-1}\|\phi\|_{\mathcal{P}^m(\mathfrak{g},\mathbb{K})}$$

for every $x, y \in \mathfrak{g}$.

Remark 2.3. It is not clear in general that if $\phi \in \mathcal{C}(\mathfrak{g},\mathbb{K})$, then the mapping $G \to \mathcal{C}(\mathfrak{g},\mathbb{K})$, $x \mapsto \lambda(x)\phi$ is continuous. However, it is easily checked that this is the case if $\mathfrak{g}$ is finite-dimensional or if it is a Banach–Lie algebra and

$$\phi \in \bigcup_{m\ge 0} \mathcal{P}_m(\mathfrak{g},\mathbb{K}).$$

It then follows by [14, Proposition 5.1] that if $\mathcal{V}$ is a closed subspace of $\mathcal{P}_m(\mathfrak{g},\mathbb{K})$ for some $m \ge 0$ (hence $\mathcal{V}$ is a Banach space) such that $\lambda(G)\mathcal{V} \subseteq \mathcal{V}$, then the mapping

$$G \times \mathcal{V} \to \mathcal{V}, \quad (x,\phi) \mapsto \lambda(x)\phi$$

is continuous. As we will see by Corollary 2.14 below, the representation of G on $\mathcal{V}$ defined by λ is actually norm-continuous.

Lemma 2.4. *If $m \in \mathbb{N}$, $\Phi \subseteq \mathcal{P}_m(\mathfrak{g},\mathbb{K})$, and we denote*

$$\mathcal{V}_\Phi := \overline{\mathrm{span}}\,(\lambda(G)\Phi)$$

then $\mathcal{V}_\Phi$ is an invariant subspace for the regular representation λ and moreover we have $\mathcal{V}_\Phi \subseteq \mathcal{P}_{mN}(\mathfrak{g},\mathbb{K})$.

Proof. Since λ is a group representation, it follows at once that $\mathcal{V}_\phi$ is an invariant subspace. Furthermore, let $x \in \mathfrak{g}$ arbitrary and define

$$L_x\colon \mathfrak{g} \to \mathfrak{g}, \quad L_x(y) = (-x) * y.$$

Since $\mathfrak{g}$ is an $(N+1)$-step nilpotent Lie algebra, it follows that $L_x \in \mathcal{P}_N(\mathfrak{g},\mathfrak{g})$, and then for $\phi \in \mathcal{P}_m(\mathfrak{g},\mathbb{K})$ we get $\lambda(x)\phi = \phi \circ L_x \in \mathcal{P}_{mN}(\mathfrak{g},\mathbb{K})$ by [5, Proposition 2, Corollary 1, page 63]. Thus $\lambda(G)\Phi \subseteq \mathcal{P}_{mN}(\mathfrak{g},\mathbb{K})$. Since $\mathcal{P}_{mN}(\mathfrak{g},\mathbb{K})$ is a closed subspace of $\mathcal{C}(\mathfrak{g},\mathbb{K})$, we get $\mathcal{V}_\Phi \subseteq \mathcal{P}_{mN}(\mathfrak{g},\mathbb{K})$, which concludes the proof. □

Remark 2.5. If $\dim \mathfrak{g} = n < \infty$, then it follows by Lemma 2.4 that every polynomial function on $\mathfrak{g}$ is a finite vector for the regular representation λ, in the sense that it generates a finite-dimensional cyclic subspace. Even more precisely, if $\phi \in \mathcal{P}_m(\mathfrak{g},\mathbb{K})$ and $\Phi := \{\phi\}$, then

$$\dim \mathcal{V}_\Phi \le \dim \mathcal{P}_{mN}(\mathfrak{g},\mathbb{K}) = \frac{(mN+n)!}{n!(mN)!}.$$

We recall how the latter formula is obtained: The linear space of real polynomial functions in n variables of degree at most mN is naturally isomorphic (by using homogeneous coordinates) to the space of real homogeneous polynomial functions in $n+1$ variables of degree precisely mN, and the dimension of the latter space is $\frac{(mN+n)!}{n!(mN)!}$; see for instance [12, Theorem 1.7.5].

From the point of view of the question we address here, the drawback of the above estimate on $\dim \mathcal{V}_\Phi$ is that it depends on the dimension of $\mathfrak{g}$. Lemmas 2.6, 2.7, 2.10 and 2.11 below are aimed to fix this problem.

The following observation is well known, and the simple proof is included for the sake of completeness. We will need this result below in Lemma 2.7, in order to get a uniform estimate for the dimension of the subalgebra generated by a finite subset of a nilpotent Lie algebra.

Lemma 2.6. *Let $\mathfrak{m}$ be an arbitrary Lie algebra over $\mathbb{K}$ and denote by $\mathfrak{m}_S$ the Lie subalgebra generated by some subset $S \subseteq \mathfrak{m}$. Then*

$$\mathfrak{m}_S = \mathrm{span}_{\mathbb{K}}(S \cup \{(\mathrm{ad}_{\mathfrak{m}} v_r) \cdots (\mathrm{ad}_{\mathfrak{m}} v_1) w \mid v_1, \dots, v_r, w \in S,\ r = 1, 2, \dots\}).$$

Proof. Let us denote the right-hand side by $\mathfrak{h}$. We have $S \subseteq \mathfrak{h}$ and it is clear that if some subalgebra of $\mathfrak{m}$ contains S, then it also contains $\mathfrak{h}$.

Therefore, the proof will be completed as soon as we have proved that $\mathfrak{h}$ is a subalgebra of $\mathfrak{m}$. Since $\mathfrak{h}$ is by definition a linear subspace of $\mathfrak{m}$, we are left with checking that for every $x \in \mathfrak{h}$ we have $[x, \mathfrak{h}] \subseteq \mathfrak{h}$. To this end let us consider the normalizer of $\mathfrak{h}$ in $\mathfrak{m}$, that is, $\mathcal{N}_{\mathfrak{m}}(\mathfrak{h}) := \{x \in \mathfrak{m} \mid [x, \mathfrak{h}] \subseteq \mathfrak{m}\}$. By using the Jacobi identity under the form

$$(\forall x, y \in \mathfrak{m}) \quad \mathrm{ad}_{\mathfrak{m}}([x, y]) = [\mathrm{ad}_{\mathfrak{m}} x, \mathrm{ad}_{\mathfrak{m}} y] = (\mathrm{ad}_{\mathfrak{m}} x)(\mathrm{ad}_{\mathfrak{m}} y) - (\mathrm{ad}_{\mathfrak{m}} y)(\mathrm{ad}_{\mathfrak{m}} x) \tag{2.3}$$

and the fact that $\mathfrak{h}$ is a linear subspace, it follows that if $x, y \in \mathcal{N}_{\mathfrak{m}}(\mathfrak{h})$, then

$$\mathrm{ad}_{\mathfrak{m}}([x, y])\mathfrak{h} \subseteq \mathfrak{h},$$

hence $[x, y] \in \mathcal{N}_{\mathfrak{m}}(\mathfrak{h})$. Therefore $\mathcal{N}_{\mathfrak{m}}(\mathfrak{h})$ is a subalgebra of $\mathfrak{m}$. On the other hand, it follows by the definition of $\mathfrak{h}$ that for every $x \in S$ we have $[x, \mathfrak{h}] \subseteq \mathfrak{h}$, hence the subalgebra $\mathcal{N}_{\mathfrak{m}}(\mathfrak{h})$ contains S. Then $\mathcal{N}_{\mathfrak{m}}(\mathfrak{h}) \supseteq \mathfrak{h}$ by the remark at the very beginning of the proof. By the definition of $\mathcal{N}_{\mathfrak{m}}(\mathfrak{h})$ we then get $[\mathfrak{h}, \mathfrak{h}] \subseteq \mathfrak{h}$, hence $\mathfrak{h}$ is a subalgebra of $\mathfrak{m}$, and this completes the proof. $\square$

Lemma 2.7. *If $q \geq 1$, then the subalgebra $\mathfrak{g}_{x_1,\dots,x_q}$ generated by any elements $x_1, \dots, x_q \in \mathfrak{g}$ is finite dimensional and we have an estimate depending only on*

N and q, namely

$$\dim_{\mathbb{K}}(\mathfrak{g}_{x_1,\dots,x_q}) \le \sum_{r=0}^{N-1} q^{r+1}.$$

Proof. We may assume that we have distinct elements $x_1,\dots,x_q \in \mathfrak{g}$, so that we have a subset with q elements $S = \{x_1,\dots,x_q\} \subseteq \mathfrak{g}$. By using the fact that $\mathfrak{g}^{(N+1)} = \{0\}$ we get by Lemma 2.6

$$\mathfrak{g}_{x_1,\dots,x_q} = \mathrm{span}_{\mathbb{K}}(S \cup \{(\mathrm{ad}_{\mathfrak{m}} v_r)\cdots(\mathrm{ad}_{\mathfrak{m}} v_1)w \mid v_1,\dots,v_r,w \in S,\\ 1 \le r \le N-1\}).$$

Thus $\mathfrak{g}_{x_1,\dots,x_q}$ is linearly generated by a set containing at most $\sum_{r=0}^{N-1} q^{r+1}$ elements, hence we get the required estimate. □

Remark 2.8. We record the following fact for later use. If $x_0,x_1,\dots,x_q \in \mathfrak{g}$, $\alpha_1,\dots,\alpha_q \in \mathbb{N}$, and $\psi \in \mathcal{C}^\infty(\mathfrak{g},\mathbb{K})$, the value $\big((\mathrm{d}\lambda(x_1))^{\alpha_1}\cdots(\mathrm{d}\lambda(x_q))^{\alpha_q}\psi\big)(x_0)$ depends only on the values of ψ on the finite-dimensional subalgebra $\mathfrak{g}_{x_0,x_1,\dots,x_q}$ (see Lemma 2.7). This is a direct consequence of the definition (2.1).

Notation 2.9. We denote by $\mathbb{N}^{(\mathbb{N})}$ the set of all sequences $\alpha = (\alpha_0,\alpha_1,\dots)$ consisting of nonnegative integers such that there exists $k_\alpha \in \mathbb{N}$ with the property $\alpha_j = 0$ if $j > k_\alpha$. If $\alpha \in \mathbb{N}^{(\mathbb{N})}$, then for any sequence $x_0,x_1,\dots \in \mathfrak{g}$ and any $y \in \mathfrak{g}$ we denote

$$(\mathrm{ad}_{\mathfrak{g}} x_0)^{\alpha_0}(\mathrm{ad}_{\mathfrak{g}} x_1)^{\alpha_1}\cdots y := (\mathrm{ad}_{\mathfrak{g}} x_0)^{\alpha_0}(\mathrm{ad}_{\mathfrak{g}} x_1)^{\alpha_1}\cdots(\mathrm{ad}_{\mathfrak{g}} x_k)^{\alpha_k} y$$

for any $k > k_\alpha$, where $(\mathrm{ad}_{\mathfrak{g}} x)^0 := \mathrm{id}_{\mathfrak{g}}$ for every $x \in \mathfrak{g}$, and $k_\alpha \in \mathbb{N}$ is as above.

Moreover, for every $\alpha = (\alpha_0,\alpha_1,\dots) \in \mathbb{N}^{(\mathbb{N})}$ we denote $|\alpha| = \sum_{k\ge 0}\alpha_k$.

Lemma 2.10. *Let $\phi \in \mathfrak{g}^*$ and $x_0 \in \mathfrak{g}$. For all $\alpha,\beta \in \mathbb{N}^{(\mathbb{N})}$ define*

$$p_{\alpha\beta}\colon \mathfrak{g} \to \mathbb{K}, \quad p_{\alpha\beta}(y) = \phi((\mathrm{ad}_{\mathfrak{g}} x_0)^{\alpha_0}(\mathrm{ad}_{\mathfrak{g}} y)^{\beta_0}(\mathrm{ad}_{\mathfrak{g}} x_0)^{\alpha_1}(\mathrm{ad}_{\mathfrak{g}} y)^{\beta_1}\cdots y).$$

If we denote $\mathcal{V} = \mathrm{span}_{\mathbb{K}}(\{p_{\alpha\beta} \mid \alpha,\beta \in \mathbb{N}^{(\mathbb{N})}\} \cup \{\mathbf{1}\})$, then

(i) $\dim_{\mathbb{K}} \mathcal{V} \le 2^{N-1}+1$,

(ii) *for all $\alpha,\beta \in \mathbb{N}^{(\mathbb{N})}$ we have*

$$\mathrm{d}\lambda(x_0)(p_{\alpha\beta}) \in \mathrm{span}_{\mathbb{K}}(\{p_{\gamma\delta} \mid (\gamma,\delta) \in I_{\alpha\beta}\} \cup \{\mathbf{1}\}),$$

where $I_{\alpha\beta}$ denotes the set of all pairs $(\gamma,\delta) \in \mathbb{N}^{(\mathbb{N})} \times \mathbb{N}^{(\mathbb{N})}$ satisfying either $|\gamma|+|\delta| > |\alpha|+|\beta|$, or $|\gamma|+|\delta| = |\alpha|+|\beta|$ and $|\gamma| > |\alpha|$,

(iii) $\mathrm{d}\lambda(x_0)\mathcal{V} \subseteq \mathcal{V}$ *and* $(\mathrm{d}\lambda(x_0))^{2^{N-1}+1} = 0$ *on* $\mathcal{V}$.

Proof. (i) It suffices to show that for at most 2^{N-1} pairs $(\alpha, \beta) \in \mathbb{N}^{(\mathbb{N})} \times \mathbb{N}^{(\mathbb{N})}$ we have $p_{\alpha\beta} \neq 0$. For any such pair we must have

$$|\alpha| + |\beta| \leq N - 1,$$

since $\mathfrak{g}^{(N+1)} = \{0\}$. Moreover, if we define $k := \max\{j \in \mathbb{N} \mid \beta_j \geq 1\}$, then $\alpha_{k+1} \geq 1$. For every $r \in \{0, \dots, N-1\}$, it is then easily seen that the set

$$\{(\alpha, \beta) \in \mathbb{N}^{(\mathbb{N})} \times \mathbb{N}^{(\mathbb{N})} \mid |\alpha| + |\beta| = r \quad \text{and} \quad p_{\alpha\beta} \neq 0\}$$

has at most $\max\{1, 2^{r-1}\}$ elements. In fact, if (α, β) belongs to the above set, then $p_{\alpha\beta}(y) = \phi((T_1(y) \cdots T_r(y))y)$ where each of $T_1(y), \dots, T_r(y)$ is either $\mathrm{ad}_{\mathfrak{g}} x_0$ or $\mathrm{ad}_{\mathfrak{g}} y$, and moreover $T_r(y) \neq \mathrm{ad}_{\mathfrak{g}} y$, and there are $\max\{1, 2^{r-1}\}$ such possible choices. Consequently there exist at most $1 + 1 + 2 + \cdots + 2^{N-2} = 2^{N-1}$ pairs $(\alpha, \beta) \in \mathbb{N}^{(\mathbb{N})} \times \mathbb{N}^{(\mathbb{N})}$ with $p_{\alpha\beta} \neq 0$.

(ii) For every $y \in \mathfrak{g}$ we define

$$R_y \colon \mathfrak{g} \to \mathfrak{g}, \quad z \mapsto z * y.$$

Then for every $\psi \in \mathcal{C}^\infty(\mathfrak{g}, \mathbb{K})$ and $y \in \mathfrak{g}$ we have

$$(\mathrm{d}\lambda(x_0)\psi)(y) = (\psi \circ R_y)'_0(-x_0) = -\psi'_y((R_y)'_0 x_0).$$

It then follows by the Baker–Campbell–Hausdorff formula that there exist some universal constants $c_0, \dots, c_N \in \mathbb{Q}$ such that

$$(\forall \psi \in \mathcal{C}^\infty(\mathfrak{g}, \mathbb{K}))(\forall y \in \mathfrak{g}) \quad (\mathrm{d}\lambda(x_0)\psi)(y) = \sum_{j=0}^{N} c_j \psi'_y((\mathrm{ad}_{\mathfrak{g}} y)^j x_0). \tag{2.4}$$

If $\alpha = \beta = 0$, that is, for $p_{\alpha\beta} = \phi$, it follows by the above formula that for all $y \in \mathfrak{g}$ we have

$$\begin{aligned}(\mathrm{d}\lambda(x_0)\phi)(y) &= \sum_{j=0}^{N} c_j \phi((\mathrm{ad}_{\mathfrak{g}} y)^j x_0) \\ &= c_0 \phi(x_0) - \sum_{j=1}^{N} c_j \phi((\mathrm{ad}_{\mathfrak{g}} y)^{j-1} (\mathrm{ad}_{\mathfrak{g}} x_0) y)\end{aligned}$$

hence $\mathrm{d}\lambda(x_0)\phi$ is a linear combination of a constant function and of elements in the set $\{p_{\gamma\delta} \mid |\gamma| = 1\}$. Thus the assertion holds true for $\alpha = \beta = 0$.

Now let $\alpha, \beta \in \mathbb{N}^{(\mathbb{N})}$ with $|\alpha| + |\beta| \geq 1$ and $p_{\alpha\beta} \neq 0$. Then we may assume that there exist $\varepsilon \in \{\pm 1\}$ and $k \in \mathbb{N}$ such that $\beta_0, \alpha_1, \beta_1, \dots, \alpha_k, \beta_k \geq 1$, $\alpha_i = \beta_i = 0$ for $i \geq k+1$, and

$$(\forall y \in \mathfrak{g}) \quad p_{\alpha\beta}(y) = \varepsilon \phi((\mathrm{ad}_{\mathfrak{g}} x_0)^{\alpha_0} (\mathrm{ad}_{\mathfrak{g}} y)^{\beta_0} \cdots (\mathrm{ad}_{\mathfrak{g}} x_0)^{\alpha_k} (\mathrm{ad}_{\mathfrak{g}} y)^{\beta_k} x_0)$$

(since we have $(\mathrm{ad}_{\mathfrak{g}} x_0)^q y = -(\mathrm{ad}_{\mathfrak{g}} x_0)^{q-1}(\mathrm{ad}_{\mathfrak{g}} y)x_0$ if $q \ge 1$, which was already used above). Then for all $y, z \in \mathfrak{g}$ we have

$$(p_{\alpha\beta})'_y(z) = \sum_{i=0}^{k}\sum_{l=0}^{\beta_i-1} \varepsilon\phi((\mathrm{ad}_{\mathfrak{g}} x_0)^{\alpha_0}(\mathrm{ad}_{\mathfrak{g}} y)^{\beta_0}\cdots(\mathrm{ad}_{\mathfrak{g}} y)^{l}(\mathrm{ad}_{\mathfrak{g}} z)(\mathrm{ad}_{\mathfrak{g}} y)^{\beta_i-l-1} \times\cdots\times(\mathrm{ad}_{\mathfrak{g}} x_0)^{\alpha_k}(\mathrm{ad}_{\mathfrak{g}} y)^{\beta_k}x_0).$$

Therefore, by using (2.4), we get

$$\begin{aligned}(\mathrm{d}\lambda(x_0)(p_{\alpha\beta}))(y) = \sum_{j=0}^{N}\sum_{i=0}^{k}\sum_{l=0}^{\beta_i-1} c_j\varepsilon\phi((\mathrm{ad}_{\mathfrak{g}} x_0)^{\alpha_0}(\mathrm{ad}_{\mathfrak{g}} y)^{\beta_0}\\ \times\cdots\times(\mathrm{ad}_{\mathfrak{g}} y)^{l}(\mathrm{ad}_{\mathfrak{g}}((\mathrm{ad}_{\mathfrak{g}} y)^{j}x_0))(\mathrm{ad}_{\mathfrak{g}} y)^{\beta_i-l-1}\\ \times\cdots\times(\mathrm{ad}_{\mathfrak{g}} x_0)^{\alpha_k}(\mathrm{ad}_{\mathfrak{g}} y)^{\beta_k}x_0).\end{aligned}$$

On the other hand, by using the Jacobi identity (2.3), we see that for every $j \in \mathbb{N}$ and $s \in \{0,\dots,j\}$ there exists $a_{js} \in \mathbb{Z}$ such that

$$\begin{aligned}\mathrm{ad}_{\mathfrak{g}}((\mathrm{ad}_{\mathfrak{g}} y)^{j}x_0) &= \underbrace{[\mathrm{ad}_{\mathfrak{g}} y,\dots,[\mathrm{ad}_{\mathfrak{g}} y}_{j\text{ times}},\mathrm{ad}_{\mathfrak{g}} x_0]\cdots]\\ &= \sum_{s=0}^{j} a_{js}(\mathrm{ad}_{\mathfrak{g}} y)^{s}(\mathrm{ad}_{\mathfrak{g}} x_0)(\mathrm{ad}_{\mathfrak{g}} y)^{j-s}.\end{aligned}$$

This equality along with the above formula for $\mathrm{d}\lambda(x_0)(p_{\alpha\beta})$ show that this function is a linear combination of the following polynomial functions of $y \in \mathfrak{g}$, for $0 \le j \le N$, $0 \le i \le k$, $0 \le l \le \beta_i - 1$, and $0 \le s \le j$,

$$\phi((\mathrm{ad}_{\mathfrak{g}} x_0)^{\alpha_0}(\mathrm{ad}_{\mathfrak{g}} y)^{\beta_0}\cdots(\mathrm{ad}_{\mathfrak{g}} x_0)^{\alpha_i}(\mathrm{ad}_{\mathfrak{g}} y)^{l+s}(\mathrm{ad}_{\mathfrak{g}} x_0)(\mathrm{ad}_{\mathfrak{g}} y)^{j-s+\beta_i-l-1} \times\cdots\times(\mathrm{ad}_{\mathfrak{g}} x_0)^{\alpha_k}(\mathrm{ad}_{\mathfrak{g}} y)^{\beta_k}x_0)$$

that is,

$$-\phi((\mathrm{ad}_{\mathfrak{g}} x_0)^{\alpha_0}(\mathrm{ad}_{\mathfrak{g}} y)^{\beta_0}\cdots(\mathrm{ad}_{\mathfrak{g}} x_0)^{\alpha_i}(\mathrm{ad}_{\mathfrak{g}} y)^{l+s}(\mathrm{ad}_{\mathfrak{g}} x_0)(\mathrm{ad}_{\mathfrak{g}} y)^{j-s+\beta_i-l-1} \times\cdots\times(\mathrm{ad}_{\mathfrak{g}} x_0)^{\alpha_k}(\mathrm{ad}_{\mathfrak{g}} y)^{\beta_k-1}(\mathrm{ad}_{\mathfrak{g}} x_0)y).$$

The above polynomial function is equal to $-p_{\gamma\delta}$, where $|\gamma| = |\alpha| + 2$ and $|\gamma| + |\delta| = |\alpha| + |\beta| + j$. Hence the assertion follows.

(iii) We have already noted in the above proof of assertion (i) that

$$\sup\{|\gamma| + |\delta| \mid \gamma,\delta \in \mathbb{N}^{(\mathbb{N})} \text{ and } p_{\gamma\delta} \ne 0\}(\le N-1) < \infty$$

hence also

$$\sup\{|\gamma| \mid \gamma \in \mathbb{N}^{(\mathbb{N})} \text{ and there exists } \delta \in \mathbb{N}^{(\mathbb{N})} \text{ with } p_{\gamma\delta} \ne 0\}(\le N-1) < \infty.$$

As a direct consequence of assertion (ii), we then see that for every $\alpha, \beta \in \mathbb{N}^{(\mathbb{N})}$ we have, besides $\mathrm{d}\lambda(x_0)(p_{\alpha\beta}) \in \mathcal{V}$, also $(\mathrm{d}\lambda(x_0))^m(p_{\alpha\beta}) = 0$ for sufficiently large $m \geq 1$. Then $\mathrm{d}\lambda(x_0)\mathcal{V} \subseteq \mathcal{V}$ and, since $\dim_{\mathbb{K}} \mathcal{V} \leq 2^{N-1} + 1$ by assertion (i), we also get $(\mathrm{d}\lambda(x_0))^{2^{N-1}+1} = 0$ on $\mathcal{V}$. This completes the proof. □

Lemma 2.11. *If $x_0 \in \mathfrak{g}$, then for every $m \geq 0$ and $\phi \in \mathcal{P}_m(\mathfrak{g}, \mathbb{K})$ we have*

$$(\mathrm{d}\lambda(x_0))^{2^{N-1}m+1}\phi = 0.$$

Proof. We have to prove that for every $y \in \mathfrak{g}$ we have

$$((\mathrm{d}\lambda(x_0))^{2^{N-1}m+1}\phi)(y) = 0.$$

By using Remark 2.8, we easily see that it suffices to obtain the conclusion under the additional assumption $\dim \mathfrak{g} < \infty$. Moreover, we may assume that ϕ is a homogeneous polynomial (of degree m). Then for every $\alpha \in \mathbb{N}^{(\mathbb{N})}$ with $|\alpha| = m$ there exist $\phi_{\alpha,1}, \dots, \phi_{\alpha,m} \in \mathfrak{g}^*$ such that

$$\phi = \sum_{\substack{\alpha \in \mathbb{N}^{(\mathbb{N})} \\ |\alpha| = m}} \phi_{\alpha,1} \cdots \phi_{\alpha,m}.$$

Now denote $Q := 2^{N-1}m+1$. Since $\mathrm{d}\lambda(x_0)\colon \mathcal{C}^\infty(\mathfrak{g}, \mathbb{K}) \to \mathcal{C}^\infty(\mathfrak{g}, \mathbb{K})$ is a derivation, we get

$$(\mathrm{d}\lambda(x_0))^Q\phi = \sum_{\substack{\alpha \in \mathbb{N}^{(\mathbb{N})} \\ |\alpha| = m}} \sum_{\substack{q_1,\dots,q_m \in \mathbb{N} \\ q_1+\dots+q_m = Q}} \frac{Q!}{q_1! \cdots q_m!} \times (\mathrm{d}\lambda(x_0))^{q_1}\phi_{\alpha,1} \cdots (\mathrm{d}\lambda(x_0))^{q_m}\phi_{\alpha,m}.$$

If $q_1, \dots, q_m \in \mathbb{N}$ and $q_1 + \dots + q_m = Q > 2^{N-1}m$, then there exists an index $i \in \{1, \dots, m\}$ such that $q_i > 2^{N-1}$, hence $(\mathrm{d}\lambda(x_0))^{q_i}\phi_{\alpha,i} = 0$ by Lemma 2.10. Then $(\mathrm{d}\lambda(x_0))^Q\phi = 0$, and this completes the proof. □

Lemma 2.12. *If $m \in \mathbb{N}$ and $\phi \in \mathcal{P}_m(\mathfrak{g}, \mathbb{K})$, then the mapping*

$$\mathfrak{g} \to \mathcal{P}_{m-1}(\mathfrak{g}, \mathbb{K}), \quad x \mapsto \mathrm{d}\lambda(x)\phi,$$

is linear and for $x_1, x_2 \in \mathfrak{g}$ we have

$$\mathrm{d}\lambda([x_1, x_2])\phi = \mathrm{d}\lambda(x_1)\mathrm{d}\lambda(x_2)\phi - \mathrm{d}\lambda(x_2)\mathrm{d}\lambda(x_1)\phi.$$

Proof. By using Remark 2.8, it suffices to obtain the conclusion under the additional assumption $\dim \mathfrak{g} < \infty$. In this case we see that $\mathcal{V}_\phi := \mathrm{span}(\lambda(G)\phi)$ is a finite-dimensional invariant subspace for the regular representation λ, by using Remark 2.5. Let $\mathrm{d}\lambda_0\colon \mathfrak{g} \to \mathrm{End}\,(\mathcal{V}_\phi)$, $x \mapsto \mathrm{d}\lambda(x)|_{\mathcal{V}_\phi}$, denote the corresponding finite-dimensional representation. Note that $\lambda_0\colon G \to \mathrm{End}\,(\mathcal{V}_\phi)$, $x \mapsto \lambda(x)|_{\mathcal{V}_\phi}$, is a finite-dimensional representation of the Lie group G, whose derived representation is just $\mathrm{d}\lambda_0\colon \mathfrak{g} \to \mathrm{End}\,(\mathcal{V}_\phi)$. Therefore, $\mathrm{d}\lambda_0$ is a homomorphism of Lie algebras, and we are done. □

In the next statement we use a function space which is invariant to translations, and that appeared in [2] in the case of finite-dimensional nilpotent Lie groups. See also the proof of Birkhoff's embedding theorem provided in [9].

Theorem 2.13. *If we define*

$$\mathcal{F}_G = \overline{\mathrm{span}}\,(\lambda(G)\mathfrak{g}^*)$$

then the following assertions hold:

(i) *The function space $\mathcal{F}_G$ is a closed linear subspace of $\mathcal{P}_N(\mathfrak{g},\mathbb{K})$ which is invariant under the left regular representation and contains the constant functions.*

(ii) *The mapping $\lambda_G\colon G \to \mathrm{End}\,(\mathcal{F}_G)$, $x \mapsto \lambda(x)|_{\mathcal{F}_G}$, is a faithful smooth representation of the Lie group G. Moreover, for every $x \in G$ we have*

$$(\lambda_G(x) - \mathbf{1})^{2^{N-1}N+1} = 0.$$

(iii) *The mapping $\mathrm{d}\lambda_G\colon \mathfrak{g} \to \mathrm{End}\,(\mathcal{F}_G)$, $x \mapsto \mathrm{d}\lambda(x)|_{\mathcal{F}_G}$, is a faithful representation of the Lie algebra $\mathfrak{g}$ and for every $x \in \mathfrak{g}$ we have*

$$\mathrm{d}\lambda_G(x)^{2^{N-1}N+1} = 0.$$

Proof. For assertion (i) note that $\mathfrak{g}^* \subseteq \mathcal{P}_1(\mathfrak{g},\mathbb{K})$, hence it follows by Lemma 2.4 that $\mathcal{F}_G \subseteq \mathcal{P}_N(\mathfrak{g},\mathbb{K})$ and $\mathcal{F}_G$ is invariant under the representation λ. Moreover, if $\mathfrak{z}$ denotes the center of $\mathfrak{g}$ and we pick $v \in \mathfrak{z}$ and $\xi \in \mathfrak{g}^*$, then $\lambda(v)\xi = -\xi(v)\mathbf{1} + \xi$. We thus see that the constant functions belong to $\mathcal{F}_G$.

We now prove assertions (ii)–(iii). Note that $\mathrm{d}\lambda_G$ is a homomorphism of Lie algebras by Lemma 2.12. The representation λ_G is smooth (that is, its representation space consists only of smooth vectors) since every continuous polynomial function on $\mathfrak{g}$ has arbitrarily high Gâteaux derivatives (see for instance [5, Section 3]). Moreover, Lemma 2.11 shows that for every $x \in \mathfrak{g}$ we have

$$(\mathrm{d}\lambda_G(x))^{2^{N-1}N+1} = 0.$$

By considering the Taylor expansion of the function $\mathbb{K} \mapsto \mathcal{F}_G$, $t \mapsto \lambda_G(tx)\phi$, for arbitrary $\phi \in \mathcal{F}_G$, it then easily follows that

$$\lambda_G(x) = \mathrm{e}^{\mathrm{d}\lambda_G(x)} \in \mathrm{End}\,(\mathcal{F}_G). \tag{2.5}$$

We then get $(\lambda_G(x) - \mathbf{1})^{2^{N-1}N+1} = 0$.

In order to prove that λ_G is a faithful representation, it suffices to check that if $x_0 \in G$ and $\lambda_G(x_0) = \mathbf{1} \in \mathrm{End}\,(\mathcal{F}_G)$, then $x_0 = 0 \in \mathfrak{g}$. In fact, since $\mathfrak{g}^* \subseteq \mathcal{F}_G$, it follows that for every $\xi \in \mathfrak{g}^*$ we have $\lambda_G(x_0)\xi = \xi$. By evaluating both sides of this equation at $0 \in \mathfrak{g}$ we get $\xi(x_0) = 0$ for every $\xi \in \mathfrak{g}^*$. Since the locally convex space underlying $\mathfrak{g}$ is a Hausdorff space, it then follows by the Hahn–Banach Theorem that $x_0 = 0 \in \mathfrak{g}$.

Finally, since the smooth representation λ_G is faithful and the Lie group G has a smooth exponential map, it follows that also the derived representation $\mathrm{d}\lambda_G$ is faithful. In fact, if $x_0 \in \mathfrak{g}$ and $\mathrm{d}\lambda_G(x_0) = 0$, then by (2.5) we get

$$\lambda_G(x_0) = \mathrm{e}^{\mathrm{d}\lambda_G(x_0)} = \mathbf{1} \in \mathrm{End}\,(\mathcal{F}_G).$$

Therefore $x_0 = 0 \in \mathfrak{g}$. □

Corollary 2.14. *Each connected, simply connected, $(N+1)$-step nilpotent Banach–Lie group has a faithful, unipotent, norm continuous representation on a suitable Banach space, with the unipotence index at most $2^{N-1}N+1$. Every $(N+1)$-step nilpotent Banach–Lie algebra has a faithful bounded representation by nilpotent operators on some Banach space, with the nilpotence index at most $2^{N-1}N+1$.*

Proof. Using Theorem 2.13 (iii) in the special case when $\mathfrak{g}$ is a nilpotent Banach–Lie algebra, we are only left with proving that λ_G is a norm continuous representation and $\mathrm{d}\lambda_G$ is a bounded linear map. It follows by Remark 2.2 that the space of polynomials $\mathcal{P}_N(\mathfrak{g},\mathbb{K})$ is a Banach space, hence so is its closed subspace $\mathcal{F}_G$. By using (2.4) and then Remark 2.2 we get for suitable constants C_1, C_2 and C_3

$$\begin{aligned}
\|\dot{\lambda}(x_0)\psi\|_{\mathcal{P}^{m-1}(\mathfrak{g},\mathbb{K})} &\le C_1 \sup_{\|y\|\le 1} |(\dot{\lambda}(x_0)\psi)(y)| \\
&\le C_2 \sup_{\|y\|\le 1,\, 0\le j\le N} |\psi_y'((\mathrm{ad}_{\mathfrak{g}} y)^j x_0)| \\
&\le C_2 \sup_{\|y\|\le 1,\, 0\le j\le N} \|\psi\|_{\mathcal{P}^m(\mathfrak{g},\mathbb{K})} \cdot \|(\mathrm{ad}_{\mathfrak{g}} y)^j x_0\| \cdot \|y\|^{m-1} \\
&\le C_3 \|\psi\|_{\mathcal{P}^m(\mathfrak{g},\mathbb{K})} \|x_0\|
\end{aligned}$$

for every $m = 1, \dots, N$. By using the direct sum decomposition in (2.2) we easily see that $\dot{\lambda}(x_0)$ is a bounded linear operator on $\mathcal{F}_G$ and the representation

$d\lambda_G\colon \mathfrak{g} \to \mathrm{End}\,(\mathcal{F}_G)$ is norm-continuous. It then integrates to a norm continuous representation of G, which coincides with λ_G since G is connected. $\square$

Remark 2.15. The second part of Corollary 2.14 is implied by the first part by using a more general argument. In fact, it is easily seen that if a homomorphism of locally convex Lie groups with smooth exponential maps is injective, then its derivative is also injective. Moreover, a bounded linear operator on a Banach space is nilpotent of index $\le q$ if and only if it is the infinitesimal generator of a one-parameter group of unipotent operators with unipotence index $\le q$.

Acknowledgments. We wish to thank the referee for several useful remarks and suggestions that helped us improve the exposition.

Bibliography

[1] I. Ado, On the representation of finite continuous groups by means of linear substitutions (in Russian), *Izv. Fiz.-Mat. Obsch. Kazan* **7** (1934/1935), 3–43.

[2] I. Beltiţă and D. Beltiţă, Continuity of magnetic Weyl calculus, *J. Funct. Anal.* **260** (2011), no. 7, 1944–1968.

[3] I. Beltiţă and D. Beltiţă, On differentiability of vectors in Lie group representations, *J. Lie Theory* **21** (2011), no. 4, 771–785.

[4] G. Birkhoff, Representability of Lie algebras and Lie groups by matrices, *Ann. of Math. (2)* **38** (1937), no. 2, 526–532.

[5] J. Bochnak and J. Siciak, Polynomials and multilinear mappings in topological vector spaces, *Studia Math.* **39** (1971), 59–76.

[6] D. Burde, B. Eick and W. de Graaf, Computing faithful representations for nilpotent Lie algebras, *J. Algebra* **322** (2009), no. 3, 602–612.

[7] D. Burde and W. A. Moens, Faithful Lie algebra modules and quotients of the universal enveloping algebra, *J. Algebra* **325** (2011), 440–460.

[8] W. T. van Est and S. Świerczkowski, The path functor and faithful representability of Banach Lie algebras, *J. Aust. Math. Soc.* **16** (1973), 54–69.

[9] R. W. Goodman, *Nilpotent Lie Groups: Structure and Applications to Analysis*, Lecture Notes in Mathematics 562, Springer-Verlag, Berlin, 1976.

[10] W. A. de Graaf, Constructing faithful matrix representations of Lie algebras, in: *Proceedings of the 1997 International Symposium on Symbolic and Algebraic Computation* (Kihei 1997), ACM Press, New York (1997), 54–59.

[11] N. Limić, Nilpotent locally convex Lie algebras and Lie field structures, *Comm. Math. Phys.* **14** (1969), 89–107.

[12] R. Merris, *Combinatorics*, Second edition, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2003.

[13] K.-H. Neeb, Towards a Lie theory of locally convex groups, *Jpn. J. Math.* **1** (2006), no. 2, 291–468.

[14] K.-H. Neeb, On differentiable vectors for representations of infinite dimensional Lie groups, *J. Funct. Anal.* **259** (2010), no. 11, 2814–2855.

[15] W. Wojtyński, Quasinilpotent Banach–Lie algebras are Baker–Campbell–Hausdorff, *J. Funct. Anal.* **153** (1998), no. 2, 405–413.

Received June 19, 2012; revised July 24, 2012.

Author information

Ingrid Beltiţă, Institute of Mathematics "Simion Stoilow" of the Romanian Academy, Research Unit 1, P.O. Box 1-764, Bucharest, Romania.
E-mail: `Ingrid.Beltita@imar.ro`

Daniel Beltiţă, Institute of Mathematics "Simion Stoilow" of the Romanian Academy, Research Unit 1, P.O. Box 1-764, Bucharest, Romania.
E-mail: `Daniel.Beltita@imar.ro`