J. OPERATOR THEORY
56:1(2006), 89–110

ON MULTI-COMMUTATORS AND SUMS OF SQUARES OF GENERATORS OF ONE PARAMETER GROUPS

DAVIDE DI GIORGIO and A.F.M. TER ELST

Communicated by William B. Arveson

ABSTRACT. Let $\mathcal{X}$ be a Banach space and let $A_1, \ldots, A_d$ be the generators of strongly continuous groups. We prove that under suitable assumptions (roughly speaking if all the multi-commutators of $A_1, \ldots, A_d$ of order $r+1$ are zero) every linear combination of the multi-commutators is closable on its natural domain and the closure generates a strongly continuous group. Moreover the sum of the squares of $A_1, \ldots, A_d$ is closable and the closure of $-\sum\limits_{k=1}^{d} A_k^2$ generates a holomorphic semigroup. Finally, as an application of our theorem we obtain the Kolmogorov operator and the Grushin operator.

KEYWORDS: *Semigroups, commutators, integrability of Lie algebras.*

MSC (2000): 47D03, 17B15.

1. INTRODUCTION

Let A and B be generators of strongly continuous semigroups in a Banach space $\mathcal{X}$. If A and B commute, i.e., their resolvents commute, then it is well known (see p. 64 of [7]) that the sum $A + B \colon D(A) \cap D(B) \to \mathcal{X}$ is closable and the closure generates a strongly continuous semigroup. Other results regarding commuting operators which do not necessarily generate a strongly continuous semigroup are in [4] and [13].

If A and B are (possibly non commuting) generators of strongly continuous groups, then it is a natural problem to find sufficient conditions in order that the closure of the Laplacian $-(A^2 + B^2)$ with its natural domain generates an analytic semigroup. Similar problems can be posed for more than two operators.

These problems are substantially open in a Banach space setting but there are some results if $\mathcal{X}$ is finite dimensional or if the restrictions of the operators to some dense subspace span a finite dimensional Lie algebra [15], [12], [11], [1]. Alternatively, let G be a Lie group with Lie algebra $\mathfrak{g}$ and let $a, b \in \mathfrak{g}$ be close to zero in $\mathfrak{g}$. Then the Campbell–Baker–Hausdorff formula states that there is a unique $c \in \mathfrak{g}$ such that

$$\exp a \exp b = \exp c$$

where $c = a + b + \frac{1}{2}[a,b] + \frac{1}{12}[a,[a,b]] - \frac{1}{12}[b,[a,b]] + \cdots$ is a power series in multi-commutators in a and b. Hence

$$\exp(a+b) = \exp a \, \exp b \, \exp\big(-\tfrac{1}{2}[a,b]\big) \, \exp d$$

where d is a power series in (higher order) multi-commutators in a and b. It is interesting to understand whether these results can be extended, under suitable assumptions, to a Banach space setting when a and b are replaced by generators of strongly continuous groups. Recently, using as starting point the Campbell–Baker–Hausdorff formula, in [5] the following result for non commuting operators was proved.

THEOREM 1.1. *Let A, B be generators of strongly continuous groups e^{-tA} and e^{-tB} in a Banach space $\mathcal{X}$. Suppose that there exists a dense subspace D of $\mathcal{X}$ with $D \subset D(AB) \cap D(BA)$. Assume that for all $\lambda, \mu \in \mathbb{R}$ with $|\lambda|, |\mu|$ sufficiently large one has:*

(i) *$(\lambda I - A)^{-1} D \subset D$, $(\mu I - B)^{-1} D \subset D$, and*
(ii) *$[[A,B], (\lambda I - A)^{-1}]x = 0 = [[A,B], (\mu I - B)^{-1}]x$ for all $x \in D$.*

Then the operators $A + B \colon D(A) \cap D(B) \to \mathcal{X}$ and $[A,B] \colon D(AB) \cap D(BA) \to \mathcal{X}$ are closable and their closures generate strongly continuous groups given by

$$\begin{aligned} &\mathrm{e}^{t^2 \overline{[A,B]}} = \mathrm{e}^{tA}\mathrm{e}^{tB}\mathrm{e}^{-tA}\mathrm{e}^{-tB}, \quad \mathrm{e}^{-t^2 \overline{[A,B]}} = \mathrm{e}^{tB}\mathrm{e}^{tA}\mathrm{e}^{-tB}\mathrm{e}^{-tA}, \\ &\mathrm{e}^{t\overline{(A+B)}} = \mathrm{e}^{tA}\mathrm{e}^{tB}\mathrm{e}^{-(t^2/2)\overline{[A,B]}}, \end{aligned} \tag{1.1}$$

for all $t \in \mathbb{R}$.

Condition (ii) states roughly that the commutators of order 3 in A and B vanish. On a Lie group the Campbell–Baker–Hausdorff formula gives identities similarly to (1.1) if the commutators of order 3 vanish. The main aim of this paper is to generalize Theorem 1.1 to d generators of strongly continuous groups under the assumption that roughly all the commutators of order $r+1$ are zero, where $d, r \in \mathbb{N}$.

In order to state the main theorem of this paper we need some multi-index notation. Let $d \in \mathbb{N}$. Set

$$J^+(d) = \bigcup_{n=1}^{\infty} \{1, \ldots, d\}^n \quad \text{and} \quad J^+_N(d) = \bigcup_{n=1}^{N} \{1, \ldots, d\}^n$$

for all $N \in \mathbb{N}$. If $\alpha = (k_1, \ldots, k_n) \in J^+(d)$ then we set $|\alpha| = n$. Let $A_1, \ldots, A_d$ be (possibly unbounded) operators in a Banach space $\mathcal{X}$. Set $A^\alpha = A_{k_1} \cdots A_{k_n}$, with natural domain, if $\alpha = (k_1, \ldots, k_n)$. Moreover, if A and B are two operators in $\mathcal{X}$ we define the operator $[A,B]$ in $\mathcal{X}$ by

$$D([A,B]) = D(AB) \cap D(BA); \quad [A,B]x = ABx - BAx \quad \text{for all } x \in D([A,B]).$$

Then we define the operator $A_{[\alpha]}$ by

$$A_{[\alpha]} = [A_{k_1}, [A_{k_2}, \ldots, [A_{k_{n-1}}, A_{k_n}], \ldots]]$$

if $\alpha = (k_1, \ldots, k_n) \in J^+(d)$. Obviously $A_{[\alpha]}x = 0$ for all $x \in D(A_{[\alpha]})$ if $k_{n-1} = k_n$, but if A_{k_n} is unbounded then $D(A_{[\alpha]}) \neq \mathcal{X}$. In the main theorem this would give an unnecessary restriction on a domain D, which we would like to avoid. Therefore, for all $r \in \mathbb{N}$ define

$$Z_r = \{(k_1, \ldots, k_r) \in \{1, \ldots, d\}^r : k_{r-1} = k_r\}$$

and

$$\Delta_r = \begin{cases} \{(k_1, \ldots, k_r) \in \{1, \ldots, d\}^r : k_1 = \cdots = k_r\} & \text{if } r \geqslant 2, \\ \emptyset & \text{if } r = 1. \end{cases}$$

Then

$$\bigcap_{\beta \in J_r^+(d) \setminus \Delta_r} D(A^\beta) \subset D(A_{[\alpha]})$$

for all $\alpha \in J_r^+(d) \setminus Z_r$.

If A is the generator of a strongly continuous group S in $\mathcal{X}$ then the exponential growth bound of S is the infimum of all $\omega \in \mathbb{R}$ for which there is an $M \geqslant 1$ such that $\|S_t\| \leqslant M\, e^{\omega|t|}$ for all $t \in \mathbb{R}$. The main theorem of this paper is the following.

THEOREM 1.2. *Let $d, r \in \mathbb{N}$ and D a dense subspace of a Banach space $\mathcal{X}$. Moreover, let $A_1, \ldots, A_d$ be generators of strongly continuous groups in $\mathcal{X}$ with exponentially growth bounds $\omega_1, \ldots, \omega_d$. Suppose for all $k \in \{1, \ldots, d\}$ there exist $\lambda_{k,0}, \lambda_{k,1} \in \mathbb{C}$ with $\operatorname{Re} \lambda_{k,0} > \omega_k$ and $\operatorname{Re} \lambda_{k,1} < -\omega_k$ such that*

$$D \subset \bigcap_{\alpha \in J_r^+(d) \setminus \Delta_r} D(A^\alpha), \quad (\lambda I - A_k)^{-1} D \subset D \tag{1.2}$$

and

$$(\lambda I - A_k)^{-1} A_{[\alpha]} x = A_{[\alpha]} (\lambda I - A_k)^{-1} x \quad \text{for all } x \in D \text{ and } \alpha \in J_r^+(d) \setminus Z_r \text{ with } |\alpha| = r, \tag{1.3}$$

for all $\lambda \in \{\lambda_{k,0}, \lambda_{k,1}\}$.

Then for all $\alpha \in J_r^+(d) \setminus Z_r$ the operator $A_{[\alpha]}|_D$ is closable and the closure generates a strongly continuous group on $\mathcal{X}$. In fact, if $F \subset J_r^+(d) \setminus Z_r$ and if $c_\alpha \in \mathbb{R}$ for all $\alpha \in F$ then the operator $\sum_{\alpha \in F} c_\alpha A_{[\alpha]}$, with its natural domain, is closable, the closure generates a strongly continuous group on $\mathcal{X}$ and D is a core for its generator. Moreover, for all $n_1, \ldots, n_d \in \mathbb{N}$ the operator

$$\sum_{k=1}^{d} (-1)^{n_k} A_k^{2n_k} \tag{1.4}$$

with domain $\bigcap_{k=1}^{d} D(A_k^{2n_k})$ *is closable and the closure generates a strongly continuous holomorphic semigroup on* $\mathcal{X}$. *The space* $\bigcap_{\alpha \in J^+(d)} D(A^\alpha)$ *is a core. Finally, if* $d' \in \{1, \ldots, d\}$ *and* $C = (c_{kl})$ *is a real symmetric positive semi-definite matrix then the operator*

$$-\sum_{k,l=1}^{d'} c_{kl}\, A_k\, A_l + \sum_{k=d'+1}^{d} A_k \tag{1.5}$$

with domain $\bigcap_{k,l=1}^{d'} D(A_k\, A_l) \cap \bigcap_{k=d'+1}^{d} D(A_k)$ *is closable and the closure generates a strongly continuous semigroup on* $\mathcal{X}$. *The space* $\bigcap_{\alpha \in J^+(d)} D(A^\alpha)$ *is again a core.*

Note that the assumptions of this theorem are expressed in terms of resolvents. It is still an open problem whether it is possible to express the assumptions in terms of the groups generated by the A_k (see p. 195 of [12]).

The proof of this theorem uses the theory of Lie algebras and Lie groups. If the domain D was invariant under the operators $A_1, \ldots, A_d$ then we could define a representation T of a suitable Lie algebra $\mathfrak{g}$ in the Banach space $\mathcal{X}$ and show that it is integrable, i.e., it lifts to a representation of the corresponding Lie group G. Then the statements of Theorem 1.2 follow easily. Unfortunately it is unclear whether D has some dense subspace D_0 such that D_0 is invariant under the operators $A_1, \ldots, A_d$. Nevertheless, even if D is not invariant under the A_k then one can still define the map $T \colon \mathfrak{g} \to \mathcal{L}(\mathcal{X})$. Rusinek [18] has given sufficient conditions to ensure that T lifts to a representation of the Lie group G in the Banach space $\mathcal{X}$ and we will prove that these conditions are satisfied.

Note that Theorem 1.2 extends Theorem 1.1 even if $d = r = 2$, since in Theorem 1.2 the conditions (1.2) and (1.3) are required only for one large and for one small λ, for each A_k, whilst in Theorem 1.1 the conditions (i) and (ii) are required for all $\lambda, \mu \in \mathbb{R}$ with $|\lambda|, |\mu|$ large. The proof of Theorem 1.2 consists of several steps and one step is much easier if one assumes that the conditions (1.2) and (1.3) are valid for all $\lambda \in \mathbb{R}$ with $|\lambda|$ sufficiently large. In Section 2 we give a transparent proof of Theorem 1.2 under these additional assumptions. The general case requires the same ideas and steps, but in addition one more technical lemma. This will be proved in Section 3. It is then also possible to extend the Campbell–Baker–Hausdorff formula in our setting for unbounded operators in a Banach space. In the last section we write some explicit formulae for the groups generated by the commutators and in particular for the case when $r = 3$. Moreover we give applications of Theorem 1.2. The Kolmogorov operator and the Grushin operators are two examples where our theorem works well.

In this paper we follow the sign convention of the generators as in the book of Robinson. Thus H is the generator of the semigroup $t \mapsto e^{-tH}$ and if U is a representation of a Lie group G then $dU(a)$ is the generator of the group $t \mapsto U(\exp(-ta))$ for all $a \in \mathfrak{g}$, the Lie algebra of G.

2. INTEGRABILITY OF A LIE ALGEBRA REPRESENTATION

As the first step, in this section we shall prove Theorem 1.2 if in addition (1.2) and (1.3) are valid for all $\lambda \in \mathbb{R}$ with $|\lambda| > \omega + 1$ and $k \in \{1, \ldots, d\}$, where $\omega = \max\{\omega_1, \ldots, \omega_d\}$. In the third section, we shall give the full proof that, essentially, uses the main ideas of this section but, since the assumptions are weaker, is more technical.

The proof is inspired by the proof in the paper [17]. We first introduce the nilpotent Lie algebra in d generators which is free of step r. Let $a_1, \ldots, a_d$ be a basis of a real vector space V. Set

$$\mathcal{T}_\infty = \bigoplus_{n=1}^{\infty} V^{\otimes n} \quad \text{and} \quad \mathcal{T}_r = \bigoplus_{n=1}^{r} V^{\otimes n}.$$

If $\alpha = (k_1, \ldots, k_n) \in J^+(d)$ set $a^\alpha = a_{k_1} \otimes \cdots \otimes a_{k_n} \in \mathcal{T}_\infty$. Define the linear function $T \colon \mathcal{T}_r \to \mathrm{Lin}(D, \mathcal{X})$, the space of all linear operators from D into $\mathcal{X}$, such that

$$T(a^\alpha) = \begin{cases} A^\alpha|_D & \text{if } \alpha \in J_r^+(d) \backslash \Delta_r\,, \\ 0 & \text{if } \alpha \in \Delta_r\,. \end{cases}$$

The space $\mathcal{T}_\infty$ is a Lie algebra with Lie bracket

$$[t_1, t_2] = t_1 \otimes t_2 - t_2 \otimes t_1.$$

Let $\mathfrak{G}$ be the Lie subalgebra of $\mathcal{T}_\infty$ generated by $a_1, \ldots, a_d$ and let $\mathcal{I}$ be the ideal of $\mathfrak{G}$ spanned by $\{a_{[\alpha]} : \alpha \in J^+(d),\ |\alpha| \geqslant r+1\}$, where

$$a_{[\alpha]} = [a_{k_1}, [a_{k_2}, \ldots, [a_{k_{n-1}}, a_{k_n}], \ldots]]$$

if $\alpha = (k_1, \ldots, k_n) \in J^+(d)$. Define $\mathfrak{g} = \mathfrak{G}/\mathcal{I} = \mathrm{span}\{a_{[\alpha]} + \mathcal{I} : \alpha \in J_r^+(d)\}$. Then $\mathfrak{g}$ is the nilpotent Lie algebra in d generators which is free of step r. If $c_\alpha \in \mathbb{R}$ for all $\alpha \in J_r^+(d)$ and $\sum_{\alpha \in J_r^+(d)} c_\alpha\, a_{[\alpha]} \in \mathcal{I}$ then $\sum_{\alpha \in J_r^+(d)} c_\alpha\, a_{[\alpha]} = 0$ and $T\Big(\sum_{\alpha \in J_r^+(d)} c_\alpha\, a_{[\alpha]}\Big) = 0$. Hence there exists a unique linear map $\widetilde{T} \colon \mathfrak{g} \to \mathrm{Lin}(D, \mathcal{X})$ such that $\widetilde{T}(a_{[\alpha]} + \mathcal{I}) = T(a_{[\alpha]})$ for all $\alpha \in J_r^+(d)$.

From now on we write a_k for $a_k + \mathcal{I}$, T for $\widetilde{T}$ and $[\,\cdot\,,\,\cdot\,]$ for the Lie bracket on $\mathfrak{g}$. Then $T \colon \mathfrak{g} \to \mathrm{Lin}(D, \mathcal{X})$ is linear and $T(a_{[\alpha]}) = A_{[\alpha]}|_D$ for all $\alpha \in J_r^+(d) \backslash Z_r$. Moreover, $T(a_{[\alpha]}) = T(0) = 0$ if $\alpha \in Z_r$. Set $R_k(\lambda) = (\lambda I - A_k)^{-1}$ for all $k \in \{1, \ldots, d\}$ and $\lambda \in \rho(A_k)$.

LEMMA 2.1. *Let $k \in \{1, \ldots, d\}$ and $\lambda \in \mathbb{R}$ with $|\lambda| > \omega_k$ and suppose that* (1.2) *and* (1.3) *are valid. Then*

$$T(a)\, R_k(\lambda)x = R_k(\lambda)\, T(a)x - R_k(\lambda)\, T([a_k, a])\, R_k(\lambda)x \tag{2.1}$$

for all $a \in \mathfrak{g}$ *and* $x \in D$.

Proof. It suffices to prove the lemma for $a = a_{[\alpha]}$ for all $\alpha \in J_r^+(d) \backslash Z_r$. If $|\alpha| = r$ then $[a_k, a_{[\alpha]}] = 0$ by the rank of $\mathfrak{g}$ and (2.1) follows from the hypothesis (1.3).

Alternatively, if $\alpha \in J_{r-1}^+(d)$ then the statement is trivial if $\alpha \in Z_{r-1}$, or if $\alpha = (k)$, so we may assume that $\alpha \notin Z_{r-1}$ and $\alpha \neq (k)$. Then $D \subset D(A_k A_{[\alpha]}) \cap D(A_{[\alpha]} A_k)$. Since $R_k(\lambda) D \subset D$, one has $R_k(\lambda) x \in D(A_k A_{[\alpha]}) \cap D(A_{[\alpha]} A_k)$. Therefore

$$\begin{aligned}
T(a_{[\alpha]})\, R_k(\lambda)x - R_k(\lambda)\, T(a_{[\alpha]})x \\
&= A_{[\alpha]}\, R_k(\lambda)x - R_k(\lambda)\, A_{[\alpha]}x \\
&= R_k(\lambda)(\lambda I - A_k)A_{[\alpha]}\, R_k(\lambda)x - R_k(\lambda)\, A_{[\alpha]}(\lambda I - A_k)R_k(\lambda)x \\
&= -R_k(\lambda)\, [A_k, A_{[\alpha]}]\, R_k(\lambda)x = -R_k(\lambda)\, T([a_k, a_{[\alpha]}])\, R_k(\lambda)x
\end{aligned}$$

and the lemma follows. ∎

COROLLARY 2.2. *Let* $k \in \{1, \ldots, d\}$ *and* $\lambda \in \mathbb{R}$ *with* $|\lambda| > \omega_k$ *and suppose that* (1.2) *and* (1.3) *are valid. Then*

$$T(a)\, R_k(\lambda)^m x = R_k(\lambda)^m \sum_{i=0}^{r-1} (-1)^i \binom{m-1+i}{i} R_k(\lambda)^i\, T((\mathrm{ad} a_k)^i a)x \tag{2.2}$$

for all $a \in \mathfrak{g}$, $m \in \mathbb{N}$ *and* $x \in D$.

Proof. The proof is by induction on m. It follows by iteration of Lemma 2.1 that

$$\begin{aligned}
&T(a)\, R_k(\lambda)x \\
&\quad = R_k(\lambda) \sum_{i=0}^{r-1} (-1)^i R_k(\lambda)^i\, T((\mathrm{ad} a_k)^i a)x + (-1)^r R_k(\lambda)^r\, T((\mathrm{ad} a_k)^r a) R_k(\lambda)x.
\end{aligned}$$

But $(\mathrm{ad} a_k)^r a = 0$, so (2.2) is valid if $m = 1$. In fact, by nilpotency it follows that

$$T(a)\, R_k(\lambda)^m x = R_k(\lambda)^m \sum_{i=0}^{\infty} (-1)^i \binom{m-1+i}{i} R_k(\lambda)^i\, T((\mathrm{ad} a_k)^i a)x \tag{2.3}$$

for $m = 1$. Note that there is no convergence problem since there are only finitely many terms nonzero. But

$$\sum_{i=0}^{n} \binom{m-1+i}{i} = \binom{m+n}{n} \tag{2.4}$$

for all $m \in \mathbb{N}$ and $n \in \mathbb{N}_0$. The identity (2.4) follows easily by a double induction. Hence it follows by induction on m that (2.3) is valid for all $m \in \mathbb{N}$. ∎

For all $k \in \{1, \ldots, d\}$, $m \in \mathbb{N}$ and $t \in \mathbb{R} \backslash \{0\}$ with $|mt^{-1}| > \omega + 1$ define the operator $W_m^{(k)}(t) \in \mathcal{L}(\mathcal{X})$ by

$$W_m^{(k)}(t) = (mt^{-1})^m R_k(mt^{-1})^m.$$

Then $W_m^{(k)}(t) D \subset D$ if (1.2) and (1.3) are valid for all $\lambda \in \mathbb{R}$ with $|\lambda| > \omega + 1$ and $k \in \{1, \ldots, d\}$. If $U^{(k)}$ is the strongly continuous group generated by $-A_k$, i.e., $U^{(k)}(t) = \mathrm{e}^{tA_k}$, then $\lim_{m \to \infty} W_m^{(k)}(t)x = U^{(k)}(t)x$ for all $t \in \mathbb{R} \backslash \{0\}$ and $x \in \mathcal{X}$ by Corollary III.5.5 of [7]. It follows from Corollary 2.2 that

$$T(a)\, W_m^{(k)}(t)x = W_m^{(k)}(t) \sum_{i=0}^{r-1} (-1)^i \binom{m-1+i}{i} R_k(mt^{-1})^i\, T((\mathrm{ad} a_k)^i a)x \tag{2.5}$$

for all $a \in \mathfrak{g}$, $m \in \mathbb{N}$, $k \in \{1, \ldots, d\}$, $t \in \mathbb{R} \backslash \{0\}$ and $x \in D$ with $m > (\omega + 1)|t|$ if (1.2) and (1.3) are valid for all $\lambda \in \mathbb{R}$ with $|\lambda| > \omega + 1$ and $k \in \{1, \ldots, d\}$. We would like to take the limit $m \to \infty$, but in general D is not invariant under $U^{(k)}(t)$ and we do not yet know whether each $T(a)$ is closable for every $a \in \mathfrak{g}$. If $r > 2$ then it is even not clear whether the operator $[A_1, A_2]|_D$ is closable. We will circumvent this problem proving that the operators $T(a)$ are 'jointly' closable. The next lemma handles all problems.

LEMMA 2.3. *Suppose* (1.2) *and* (1.3) *are valid for all* $\lambda \in \mathbb{R}$ *with* $|\lambda| > \omega + 1$ *and* $k \in \{1, \ldots, d\}$. *Let* $x_1, x_2, \ldots \in D$ *be a sequence,* $x \in \mathcal{X}$ *and* $\Phi \colon \mathfrak{g} \to \mathcal{X}$ *a linear function. Suppose* $\lim_{n \to \infty} x_n = x$ *and* $\lim_{n \to \infty} T(a)x_n = \Phi(a)$ *for all* $a \in \mathfrak{g}$. *Then*

$$\lim_{m \to \infty} W_m^{(k)}(t)x_m = U^{(k)}(t)x \tag{2.6}$$

$$\lim_{m \to \infty} T(a)\, W_m^{(k)}(t)x_m = U^{(k)}(t)\, \Phi(\mathrm{e}^{-t\mathrm{ad} a_k} a) \tag{2.7}$$

for all $a \in \mathfrak{g}$, $k \in \{1, \ldots, d\}$ *and* $t \in \mathbb{R} \backslash \{0\}$.

Proof. Fix $k \in \{1, \ldots, d\}$ and $t \in \mathbb{R} \backslash \{0\}$. Then the Hille–Yosida theorem implies that there exists an $\omega' > 0$ such that $\sup_{m > \omega'} \|W_m^{(k)}(t)\| < \infty$. Since $\lim_{m \to \infty} W_m^{(k)}(t)y = U^{(k)}(t)y$ for all $y \in \mathcal{X}$ the limit (2.6) follows.

Next fix $i \in \{0, \ldots, r-1\}$. Then write

$$\begin{aligned}(-1)^i \binom{m-1+i}{i} & R_k(mt^{-1})^i\, T((\mathrm{ad} a_k)^i a)x \\ &= m^{-i} \binom{m-1+i}{i} (mt^{-1} R_k(mt^{-1}))^i\, T((-t\mathrm{ad} a_k)^i a)x\end{aligned}$$

and note that $\lim_{m\to\infty} m^{-i}\binom{m-1+i}{i} = (i!)^{-1}$. Since $\lim_{|\lambda|\to\infty} \lambda\, R_k(\lambda)y = y$ for all $y \in \mathcal{X}$ and $\sup_{|\lambda|>\omega'} \|\lambda\, R_k(\lambda)\| < \infty$ for some $\omega' > 0$ it follows that

$$\lim_{m\to\infty} (mt^{-1}R_k(mt^{-1}))^i\, T((-t\mathrm{ad}a_k)^i a)x_m = \Phi((-t\mathrm{ad}a_k)^i a).$$

Hence

$$\begin{aligned}\lim_{m\to\infty} \sum_{i=0}^{r-1} (-1)^i \binom{m-1+i}{i} R_k(mt^{-1})^i\, T((\mathrm{ad}a_k)^i a)x_m &= \sum_{i=0}^{r-1} (i!)^{-1} \Phi((-t\mathrm{ad}a_k)^i a) \\ &= \Phi(\mathrm{e}^{-t\mathrm{ad}a_k} a). \qquad (2.8)\end{aligned}$$

Arguing as in the proof of (2.6) the limit (2.7) follows from (2.8) and (2.5). ∎

The next lemma states that the operators $T(a)$, with $a \in \mathfrak{g}$, are 'jointly' closable. Recall that we still do not know that $T(a)$ is closable for all $a \in \mathfrak{g}$.

LEMMA 2.4. *Suppose* (1.2) *and* (1.3) *are valid for all* $\lambda \in \mathbb{R}$ *with* $|\lambda| > \omega + 1$ *and* $k \in \{1, \ldots, d\}$. *Let* $x_1, x_2, \ldots \in D$ *and for all* $a \in \mathfrak{g}$ *let* $y_a \in \mathcal{X}$. *Suppose that* $\lim_{n\to\infty} x_n = 0$ *and* $\lim_{n\to\infty} T(a)x_n = y_a$ *for all* $a \in \mathfrak{g}$. *Then* $y_a = 0$ *for all* $a \in \mathfrak{g}$.

Proof. Since T is linear we may assume that the map $a \mapsto y_a$ is linear.

For all $l \in \{1, \ldots, r\}$ let $P(l)$ be the following hypothesis:

If $x_1, x_2, \ldots \in D$ and $\Phi\colon \mathfrak{g} \to \mathcal{X}$ is linear such that $\lim_{n\to\infty} x_n = 0$ and $\lim_{n\to\infty} T(a)x_n = \Phi(a)$ for all $a \in \mathfrak{g}$, then $\Phi(a_{[\alpha]}) = 0$ for all $\alpha \in J_l^+(d)$.

Obviously the hypothesis $P(1)$ is valid since $T(a_k) = A_k|_D$ is closable for all $k \in \{1, \ldots, d\}$, because A_k is closed.

Let $l \in \{1, \ldots, r-1\}$ and suppose that $P(l)$ is valid. Let $x_1, x_2, \ldots \in D$, $\Phi\colon \mathfrak{g} \to \mathcal{X}$ linear and suppose that $\lim_{n\to\infty} x_n = 0$ and $\lim_{n\to\infty} T(a)x_n = \Phi(a)$ for all $a \in \mathfrak{g}$. Let $k \in \{1, \ldots, d\}$ and $\alpha \in J^+(d)$ with $|\alpha| = l$. We shall prove that $\Phi([a_k, a_{[\alpha]}]) = 0$. Let $t > 0$. Then $W_m^{(k)}(t)x_m \in D$ for all $m > (\omega+1)t$ and $\lim_{m\to\infty} W_m^{(k)}(t)x_m = 0$ by Lemma 2.3. Moreover, by the same lemma,

$$\lim_{m\to\infty} T(a)\, W_m^{(k)}(t)x_m = U^{(k)}(t)\, \Phi(\mathrm{e}^{-t\mathrm{ad}a_k} a)$$

for all $a \in \mathfrak{g}$ and the map $a \mapsto U^{(k)}(t)\, \Phi(\mathrm{e}^{-t\mathrm{ad}a_k} a)$ is linear. Hence by the induction hypothesis applied to the sequence $(W_m^{(k)}(t)x_m)_{m>(\omega+1)t}$, it follows that $U^{(k)}(t)\, \Phi(\mathrm{e}^{-t\mathrm{ad}a_k} a_{[\alpha]}) = 0$. Since $U^{(k)}(t)$ is invertible, one has $\Phi(\mathrm{e}^{-t\mathrm{ad}a_k} a_{[\alpha]}) = 0$ for all $t > 0$. So $\sum_{i=0}^{r-1} (-t)^i\, (i!)^{-1} \Phi((\mathrm{ad}a_k)^i a_{[\alpha]}) = 0$ for all $t > 0$. But the sum is a polynomial in t. Therefore the linear term has to be zero. Thus $\Phi([a_k, a_{[\alpha]}]) = 0$ and the hypothesis $P(l+1)$ is valid. ∎

Because of the previous lemma the following makes sense if (1.2) and (1.3) are valid for all $\lambda \in \mathbb{R}$ with $|\lambda| > \omega + 1$. Let E be the vector space of all $x \in \mathcal{X}$ such that there exist a sequence $x_1, x_2, \ldots \in D$ and for all $a \in \mathfrak{g}$ a $y_a \in \mathcal{X}$ such that $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} T(a)x_n = y_a$ for all $a \in \mathfrak{g}$. Since the y_a are unique by Lemma 2.4 one can define for all $a \in \mathfrak{g}$ the operator $\widetilde{T}(a) \colon E \to \mathcal{X}$ by $\widetilde{T}(a)x = y_a$. Clearly $E \subset D(A_k)$ for all $k \in \{1, \ldots, d\}$ since A_k is closed and $T(a_k) = A_k|_D$.

LEMMA 2.5. *Suppose* (1.2) *and* (1.3) *are valid for all* $\lambda \in \mathbb{R}$ *with* $|\lambda| > \omega + 1$ *and* $k \in \{1, \ldots, d\}$. *Then the following statements hold:*

(i) *The space* E *is dense in* $\mathcal{X}$.

(ii) *The map* $a \mapsto \widetilde{T}(a)$ *is linear.*

(iii) $U^{(k)}(t)E \subset E$ *for all* $k \in \{1, \ldots, d\}$ *and* $t \in \mathbb{R}$.

(iv) *If* $a \in \mathfrak{g}$, $k \in \{1, \ldots, d\}$ *and* $t \in \mathbb{R}$ *then*

$$U^{(k)}(t)\, \widetilde{T}(a)\, U^{(k)}(-t)x = \widetilde{T}(e^{t\,\mathrm{ad}\,a_k} a)x$$

for all $x \in E$.

Proof. Obviously $D \subset E$ and $\widetilde{T}(a)|_D = T(a)$ for all $a \in \mathfrak{g}$. In particular, E is dense in $\mathcal{X}$. Statement (ii) is easy. Let $x \in E$. Then there are $x_1, x_2, \ldots \in D$ such that $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} T(a)x_n = \widetilde{T}(a)x$ for all $a \in \mathfrak{g}$. Let $k \in \{1, \ldots, d\}$ and $t \in \mathbb{R}\backslash\{0\}$. Then $W_m^{(k)}(t)x_m \in D$ for all $m > (\omega + 1)|t|$. Moreover, it follows from Lemma 2.3 that $\lim_{m\to\infty} W_m^{(k)}(t)x_m = U^{(k)}(t)x$ and $\lim_{m\to\infty} T(a)\, W_m^{(k)}(t)x_m = U^{(k)}(t)\, \widetilde{T}(e^{-t\,\mathrm{ad}\,a_k} a)x$ for all $a \in \mathfrak{g}$. Hence $U^{(k)}(t)x \in E$ and $\widetilde{T}(a)\, U^{(k)}(t)x = U^{(k)}(t)\, \widetilde{T}(e^{-t\,\mathrm{ad}\,a_k} a)x$. This proves Statements (iii) and (iv). ∎

Proof of Theorem 1.2 *under restrictions.* Suppose (1.2) and (1.3) are valid for all $\lambda \in \mathbb{R}$ with $|\lambda| > \omega + 1$ and $k \in \{1, \ldots, d\}$.

Since E is dense in $\mathcal{X}$, invariant under $U^{(k)}$ and $E \subset D(A_k)$ it follows from Corollary 3.1.7 of [2], that E is a core for A_k for all $k \in \{1, \ldots, d\}$. Lemma 2.5 together with the one parameter groups $U^{(k)}$ generated by $\overline{\widetilde{T}(a_k)}$ and the fact that $a_1, \ldots, a_d$ generate $\mathfrak{g}$ are precisely the conditions of Theorem 2.1 in [18]. Hence if G is the connected simply connected Lie group with Lie algebra $\mathfrak{g}$, then there exists a strongly continuous representation $U \colon G \to \mathcal{L}(\mathcal{X})$ such that $E \subset D(dU(a))$ and $dU(a)x = \widetilde{T}(a)x$ for all $a \in \mathfrak{g}$ and $x \in E$, where $dU(a)$ is the infinitesimal generator of the group $t \mapsto U(\exp(-ta))$. In particular, $dU(a_{[\alpha]})|_D = A_{[\alpha]}|_D$ for all $\alpha \in J_r^+(d)\backslash Z_r$. We next show that E is invariant under U.

First let $k \in \{1, \ldots, d\}$. Since $E \subset D(dU(a_k))$ and $A_k|_D = dU(a_k)|_D$ it follows from the definition of E that $A_k|_E = dU(a_k)|_E$. So $dU(a_k)$ is a closed extension of $A_k|_E$. But E is a core for A_k. Hence $A_k = \overline{A_k|_E} \subset dU(a_k)$. Then $A_k = dU(a_k)$ because a semigroup generator has no strict semigroup generator extension.

Let $G_0 = \{g \in G : U(g)E \subset E\}$. Then G_0 is a Lie subgroup of G. Let $\mathfrak{g}_0$ be its Lie algebra. Then $\mathfrak{g}_0$ is a subalgebra of $\mathfrak{g}$. Moreover, $\exp ta_k \in G_0$ for all $t \in \mathbb{R}$ by Lemma 2.5 (iii). Therefore $a_k \in \mathfrak{g}_0$ for all $k \in \{1, \ldots, d\}$. But $a_1, \ldots, a_d$ generates $\mathfrak{g}$. So $\mathfrak{g}_0 = \mathfrak{g}$ and $G_0 = G$.

Now let $a \in \mathfrak{g}$. Then E is a dense subspace of $\mathcal{X}$, $E \subset D(dU(a))$ and E is invariant under the group $(U(\exp ta))_{t \in \mathbb{R}}$. Hence E is a core for $dU(a)$. In particular, $\widetilde{T}(a)$ is closable. By definition of E the space D is then a core for $\widetilde{T}(a)$ and D is a core for $dU(a)$. Thus $T(a)$ is closable and the closure generates a strongly continuous group for all $a \in \mathfrak{g}$.

Next let $F \subset J_r^+(d) \backslash Z_r$ and for all $\alpha \in F$ let $c_\alpha \in \mathbb{R}$. Set $b = \sum_{\alpha \in F} c_\alpha \, a_{[\alpha]}$ and $B = \sum_{\alpha \in F} c_\alpha \, A_{[\alpha]}$, with its natural domain. We first show that $B \subset dU(b)$. Expanding the commutators it follows that there is a subset $\widetilde{F} \subset J_r^+(d) \backslash \Delta_r$ and for all $\beta \in \widetilde{F}$ there is a $\tilde{c}_\beta \in \mathbb{R}$ such that

$$D(B) = \bigcap_{\beta \in \widetilde{F}} D(A^\beta),$$

where $Bx = \sum_{\beta \in \widetilde{F}} \tilde{c}_\beta \, A^\beta x$ for all $x \in D(B)$ and $b = \sum_{\beta \in \widetilde{F}} \tilde{c}_\beta \, a^\beta \in \mathcal{T}_r$. Let U_* be the dual representation of U in the dual space $\mathcal{X}^*$. Moreover, let $\mathcal{X}_\infty(U)$ and $\mathcal{X}^*_\infty(U_*)$ denote the space of all C^∞-vectors for U and U_*. Then $dU(b)x = \sum_{\beta \in \widetilde{F}} \tilde{c}_\beta \, A^\beta x$ for all $x \in \mathcal{X}_\infty(U)$. So

$$dU_*(-b)f = dU(b)^* f = \sum_{\beta \in \widetilde{F}} \tilde{c}_\beta \, (A^\beta)^* f = \sum_{\beta \in \widetilde{F}} (-1)^{|\beta|} \tilde{c}_\beta \, (dU_*(a))^{\beta_*} f$$

for all $f \in \mathcal{X}^*_\infty(U_*)$, where $\beta_* = (k_n, \ldots, k_1)$ if $\beta = (k_1, \ldots, k_n)$. Now let $x \in D(B)$. Then for all $f \in \mathcal{X}^*_\infty(U_*)$ one has

$$(dU_*(-b)f, x) = \sum_{\beta \in \widetilde{F}} (-1)^{|\beta|} \tilde{c}_\beta \, ((dU_*(a))^{\beta_*} f, x) = \sum_{\beta \in \widetilde{F}} \tilde{c}_\beta \, (f, A^\beta x) = (f, Bx).$$

Since $\mathcal{X}^*_\infty(U_*)$ is a core for $dU_*(-b)$ it follows that

$$x \in D((dU_*(-b))^*) = D(dU(b))$$

and $dU(b)x = Bx$. So $B \subset dU(b)$. But $\mathcal{X}_\infty(U)$ is a core for $dU(b)$. Therefore B is closable and $\overline{B} = dU(b)$.

Next, the statement of Theorem 1.2 concerning the operator $\sum_{k=1}^{d} (-1)^{n_k} A_k^{2n_k}$ follows from Theorem 1.1 and Example 4.4 of [9]. The space $\bigcap_{\alpha \in J^+(d)} D(A^\alpha)$ is a core by the argument given in the proof of Lemma 2.4 in [8], or the proof of Proposition II.8.1 in [6]. Finally, consider the operator $H = - \sum_{k,l=1}^{d'} c_{kl} \, A_k \, A_l +$

$\sum_{k=d'+1}^{d} A_k$. We may assume that there is a $d'' \in \{1,\ldots,d'\}$ such that $c_{kl} = 1$ if $k = l \leqslant d''$ and $c_{kl} = 0$ otherwise. Then $H = -\sum_{k=1}^{d''} A_k^2 + \sum_{k=d'+1}^{d} A_k$ and the statements follow from [10] or Theorem IV.4.5 of [16], applied to the operator $-\sum_{k=1}^{d''} dV(a_k)^2 + dV(a_0)$, where $a_0 = \sum_{k=d'+1}^{d} a_k$ and V is the restriction of U to G_1 with G_1 is the connected simply connected subgroup of G with Lie algebra $\mathfrak{g}_1$ and $\mathfrak{g}_1$ the Lie subalgebra of $\mathfrak{g}$ generated by $a_0, a_1, \ldots, a_{d''}$. The claims about the domain and core follow again from the argument given in the proof of Lemma 2.4 in [8]. ▮

REMARK 2.6. The above proof of (the restricted version of) Theorem 1.2 shows that the statements on operators of the form (1.4) follow from [9]. In fact, (1.4) is merely an example of the Lie group theory developed in [9]. In Theorem 1.1 and Example 4.4 of [9] there are many more examples of weighted subcoercive operators which like (1.4) all pre-generate a holomorphic semigroup. We refer the interested reader to [9].

In the course of the (restricted) proof of Theorem 1.2 we proved an integrability result.

THEOREM 2.7. *Assume the hypothesis of Theorem* 1.2. *Then there exists a unique strongly continuous representation U of G in $\mathcal{X}$ such that D is a core for $dU(a_k)$ and $A_k|_D = dU(a_k)|_D$ for all $k \in \{1,\ldots,d\}$, where $\mathfrak{g}$ is the nilpotent Lie algebra with generators $a_1,\ldots,a_d$ which is free of step r and G is the connected simply connected Lie group with Lie algebra $\mathfrak{g}$.*

It follows that D is a core for $dU(a)$ for all $a \in \mathfrak{g}$.

Proof. Everything has been proved, except the uniqueness. But if V is another representation then

$$dU(a_k) = \overline{dU(a_k)|_D} = \overline{dV(a_k)|_D} = dV(a_k)$$

for all $k \in \{1,\ldots,d\}$. Since $a_1,\ldots,a_d$ generate $\mathfrak{g}$ it follows that $U = V$. ▮

3. COMPLETE PROOF OF THEOREM 1.2

In this section we assume merely the hypothesis of Theorem 1.2 without any additional assumptions. The key element in the proof in the previous section is the existence of the operators $W_m^{(k)}(t)$ satisfying (2.6) and (2.7). By definition $W_m^{(k)}(t) = (mt^{-1})^m R_k(mt^{-1})^m$ whenever $|mt^{-1}| > \omega + 1$. The additional restrictions in Section 2 ensured immediately that $W_m^{(k)}(t)D \subset D$ and Corollary 2.2 gave the bounds needed to prove (2.6) and (2.7).

In fact, we need the following improvement of Lemma 2.3. As before, $U^{(k)}$ denotes the strongly continuous group generated by $-A_k$, so $U^{(k)}(t) = e^{tA_k}$.

LEMMA 3.1. *Assume the assumptions of Theorem* 1.2. *Then for all* $k \in \{1, \ldots, d\}$, $t \in \mathbb{R}\backslash\{0\}$ *and* $m \in \mathbb{N}$ *there exists a linear operator* $\mathcal{W}_m^{(k)}(t) \colon D \to D$ *such that*

$$\lim_{m\to\infty} \mathcal{W}_m^{(k)}(t)x_m = U^{(k)}(t)x \tag{3.1}$$

$$\lim_{m\to\infty} T(a)\,\mathcal{W}_m^{(k)}(t)x_m = U^{(k)}(t)\,\Phi(e^{-t\,\mathrm{ad}\,a_k}a) \tag{3.2}$$

for every sequence $x_1, x_2, \ldots \in D$, $x \in \mathcal{X}$, *linear* $\Phi \colon \mathfrak{g} \to \mathcal{X}$ *and* $a \in \mathfrak{g}$ *such that* $\lim\limits_{n\to\infty} x_n = x$ *and* $\lim\limits_{n\to\infty} T(b)x_n = \Phi(b)$ *for all* $b \in \mathfrak{g}$.

Proof of Theorem 1.2. If Lemma 3.1 has been proved, then Lemmas 2.4, 2.5 and the restricted proof of Theorem 1.2 in Section 2 extend line by line to the general case. ∎

Thus it remains to prove Lemma 3.1. The outline is as follows. If $t > 0$, $k \in \{1, \ldots, d\}$ and $m \in \mathbb{N}$ is large then $R_k(mt^{-1})$ in the definition of $W_m^{(k)}(t)$ is defined, but one has no control over it. It is even not clear whether it leaves D invariant. Only $R_k(\lambda_{k,0})$ behaves well enough by the assumptions (1.2) and (1.3). In particular each polynomial in $R_k(\lambda_{k,0})$ leaves D invariant. But $\lambda \mapsto R_k(\lambda)$ is analytic in a neighbourhood of $\lambda_{k,0}$ and for λ close to $\lambda_{k,0}$ one can approximate $R_k(\lambda)$ by a polynomial in $R_k(\lambda_{k,0})$. Similarly, for λ' close to λ one can approximate $R_k(\lambda')$ by a polynomial in $R_k(\lambda)$, so by a polynomial in $R_k(\lambda_{k,0})$. Continuing this way one can find an approximation of $R_k(\lambda)$ by a polynomial in $R_k(\lambda_{k,0})$ for all $\lambda \geqslant \operatorname{Re}\lambda_{k,0}$. This then provides an approximation for $W_m^{(k)}(t) = (mt^{-1})^m R_k(mt^{-1})^m$, which will be the new $\mathcal{W}_m^{(k)}(t)$.

We build the approximate resolvent only for $t > 0$ using the $\lambda_{k,0}$. First, in order to make the notation less heavy, we fix $k \in \{1, \ldots, d\}$. Set $\lambda_0 = \lambda_{k,0}$. We suppose that $\lambda_0 \in \mathbb{R}$, since the complex case follows similarly. Denote $R = R(\lambda_0, A_k)$ and $R(\lambda) = R(\lambda, A_k)$ for all $\lambda > \omega_k$. Let $\widetilde{\omega} = 2^{-1}(\omega_k + \lambda_0)$. There exists an $\widetilde{M} \geqslant 1$ such that $\|R(\lambda)\| \leqslant \widetilde{M}(\lambda - \widetilde{\omega})^{-1}$ for all $\lambda \geqslant \lambda_0$. Set $M = 3\widetilde{M}(\lambda_0 - \widetilde{\omega})^{-1}$. For $n \in \mathbb{N}$ set $\lambda_n = \lambda_0 + n\,M^{-1}$. Then

$$\|R(\lambda_n)\| \leqslant \widetilde{M}(\lambda_n - \omega)^{-1} \leqslant 3^{-1}M \tag{3.3}$$

for all $n \in \mathbb{N}_0$.

For all $N_1 \in \mathbb{N}$ set

$$\widetilde{R}_{N_1}(\lambda_1) = \sum_{i=0}^{N_1} (-1)^i\, M^{-i}\, R^{i+1}$$

and by induction, for all $n \in \mathbb{N}$ with $n \geqslant 2$ and all $N_1, \ldots, N_n \in \mathbb{N}$ set

$$\widetilde{R}_{N_1,\ldots,N_n}(\lambda_n) = \sum_{i=0}^{N_n} (-1)^i \, M^{-i} \, \widetilde{R}_{N_1,\ldots,N_{n-1}}(\lambda_{n-1})^{i+1}.$$

Then $\widetilde{R}_{N_1,\ldots,N_n}(\lambda_n)$ is a polynomial in R. So it leaves D invariant. If no confusion is possible then we write $\widetilde{R}(\lambda_n) = \widetilde{R}_{N_1,\ldots,N_n}(\lambda_n)$. Moreover, we set $\widetilde{R}(\lambda_0) = R$. So $\widetilde{R}(\lambda_0) = R(\lambda_0)$ and $\widetilde{R}(\lambda_n) = \sum_{i=0}^{N_n} (-1)^i \, M^{-i} \, \widetilde{R}(\lambda_{n-1})^{i+1}$ for all $n \in \mathbb{N}$.

The next lemma shows that $\widetilde{R}_{N_1,\ldots,N_n}(\lambda_n)$ is an approximation of $R(\lambda_n)$ if $N_1, \ldots, N_n$ are large enough.

LEMMA 3.2. *If $n \in \mathbb{N}$, $N_1, \ldots, N_n \in \mathbb{N}$ and $\sum_{j=1}^{n-1} 3^{-N_j+n-2-j} \leqslant 6^{-1}$ then*

$$\|R(\lambda_n) - \widetilde{R}_{N_1,\ldots,N_n}(\lambda_n)\| \leqslant \sum_{j=1}^{n} 3^{-N_j+n-1-j} M.$$

Hence if $\sum_{j=1}^{n} 3^{-N_j+n-1-j} \leqslant 6^{-1}$ then

$$\|\widetilde{R}_{N_1,\ldots,N_n}(\lambda_n)\| \leqslant 2^{-1} M.$$

Proof. The proof is by induction. Since $\lambda_n - \lambda_{n-1} = M^{-1} < \|R(\lambda_{n-1})\|^{-1}$ by (3.3), it follows that

$$\begin{aligned}\Big\| R(\lambda_n) - \sum_{i=0}^{N} (-1)^i \, M^{-i} \, R(\lambda_{n-1})^{i+1} \Big\| &\leqslant \sum_{i=N+1}^{\infty} M^{-i} \, \|R(\lambda_{n-1})\|^{i+1} \\ &\leqslant \sum_{i=N+1}^{\infty} M^{-i} \, (3^{-1} M)^{i+1} \leqslant 3^{-N-1} \, M\end{aligned}$$

for all $n, N \in \mathbb{N}$. Therefore the lemma follows for $n = 1$ from (3.3).

Next let $n \in \mathbb{N}$, $N_1, \ldots, N_n \in \mathbb{N}$ and suppose that $\sum_{j=1}^{n-1} 3^{-N_j+n-2-j} \leqslant 6^{-1}$ and $n \geqslant 2$. Then it follows from the induction hypothesis that

$$\begin{aligned}&\|R(\lambda_n) - \widetilde{R}_{N_1,\ldots,N_n}(\lambda_n)\| \\ &\quad \leqslant \Big\| R(\lambda_n) - \sum_{i=0}^{N_n} (-1)^i \, M^{-i} \, R(\lambda_{n-1})^{i+1} \Big\| \\ &\qquad + \sum_{i=0}^{N_n} M^{-i} \|R(\lambda_{n-1})^{i+1} - \widetilde{R}_{N_1,\ldots,N_{n-1}}(\lambda_{n-1})^{i+1}\|\end{aligned}$$

$$\leqslant 3^{-N_n-1}M + \sum_{i=0}^{N_n} M^{-i} \sum_{l=0}^{i} \|R(\lambda_{n-1})\|^l \|R(\lambda_{n-1}) - \widetilde{R}_{N_1,\ldots,N_{n-1}}(\lambda_{n-1})\|$$
$$\cdot \|\widetilde{R}_{N_1,\ldots,N_{n-1}}(\lambda_{n-1})\|^{i-l}$$
$$\leqslant 3^{-N_n-1}M + \sum_{i=0}^{N_n} M^{-i} \sum_{l=0}^{i} (3^{-1}M)^l \sum_{j=1}^{n-1} 3^{-N_j+n-2-j} M\,(2^{-1}M)^{i-l}$$
$$\leqslant 3^{-N_n-1}M + \sum_{j=1}^{n-1} 3^{-N_j+n-2-j} M \sum_{i=0}^{\infty}\sum_{l=0}^{i} 3^{-l} 2^{-(i-l)} = \sum_{j=1}^{n} 3^{-N_j+n-1-j}M,$$

from which the lemma easily follows. ∎

For all $n \in \mathbb{N}$, $N_1, \ldots, N_n \in \mathbb{N}$ and $j \in \mathbb{N}_0$ define

$$Z^{(j)}_{N_1,\ldots,N_n}(\lambda_n)$$
$$= \sum_{\substack{0 \leqslant n_1, n_2 \leqslant N_n \\ N_n < n_1+n_2}} (-1)^{n_1+n_2+j} M^{-(n_1+n_2)} \binom{n_2+j}{j} \widetilde{R}_{N_1,\ldots,N_{n-1}}(\lambda_{n-1})^{n_1+n_2+j+2}.$$

If no confusion is possible then we write $Z^{(j)}(\lambda_n) = Z^{(j)}_{N_1,\ldots,N_n}(\lambda_n)$. In addition we set $Z^{(j)}(\lambda_0) = 0$ for all $j \in \mathbb{N}_0$. Moreover, for all $n \in \mathbb{N}_0$, $m \in \mathbb{N}$ and $a \in \mathfrak{g}$ define the operator $P(\lambda_n, a, m) \colon D \to \mathcal{X}$ such that

$$T(a)\,\widetilde{R}(\lambda_n)^m x$$
$$= \widetilde{R}(\lambda_n)^m \sum_{i=0}^{r-1} (-1)^i \binom{m-1+i}{i} \widetilde{R}(\lambda_n)^i\, T((\mathrm{ad} a_k)^i a)x + P(\lambda_n, a, m)x \tag{3.4}$$

for all $x \in D$. Note that $P(\lambda_n, a, m)$ depends on $N_1, \ldots, N_n$. In addition one deduces that $P(\lambda_0, a, m) = 0$ by Corollary 2.2 applied to λ_0.

The next lemma is a version of Lemma 2.1 for $\widetilde{R}(\lambda_n)$ with a remainder term.

LEMMA 3.3. *If $n \in \mathbb{N}$, $N_1, \ldots, N_n \in \mathbb{N}$ and $a \in \mathfrak{g}$ then*

$$T(a)\,\widetilde{R}(\lambda_n)x = \widetilde{R}(\lambda_n)\,T(a)x - \widetilde{R}(\lambda_n)\,T([a_k, a])\,\widetilde{R}(\lambda_n)x + P(\lambda_n, a)x \tag{3.5}$$

for all $x \in D$, where

$$P(\lambda_n, a) = \sum_{j=0}^{r-2} Z^{(j)}(\lambda_n)\, T((\mathrm{ad} a_k)^{j+1} a)$$
$$+ \sum_{n_1=0}^{N_n}\sum_{n_2=0}^{N_n} (-1)^{n_1+n_2} M^{-(n_1+n_2)} \widetilde{R}(\lambda_{n-1})^{n_1+1} P(\lambda_{n-1}, [a_k, a], n_2+1)$$
$$+ \sum_{j=0}^{N_n} (-1)^j M^{-j} P(\lambda_{n-1}, a, j+1).$$

Proof. Using the definition of $\widetilde{R}(\lambda_n)$ and (3.4) one deduces that

$$\begin{aligned}
T(a)\,\widetilde{R}(\lambda_n)x &= \sum_{j=0}^{N_n}(-1)^j M^{-j}T(a)\,\widetilde{R}(\lambda_{n-1})^{j+1}x \\
&= \sum_{j=0}^{N_n}(-1)^j M^{-j}\widetilde{R}(\lambda_{n-1})^{j+1}\sum_{i=0}^{r-1}(-1)^i\binom{j+i}{i}\widetilde{R}(\lambda_{n-1})^i\,T((\mathrm{ad}a_k)^i a)x \\
&\qquad + \sum_{j=0}^{N_n}(-1)^j M^{-j}P(\lambda_{n-1},a,j+1).
\end{aligned}$$

But by (2.4) one establishes that

$$\begin{aligned}
&\sum_{j=0}^{N_n}(-1)^j M^{-j}\widetilde{R}(\lambda_{n-1})^{j+1}\sum_{i=0}^{r-1}(-1)^i\binom{j+i}{i}\widetilde{R}(\lambda_{n-1})^i\,T((\mathrm{ad}a_k)^i a)x - \widetilde{R}(\lambda_n)\,T(a)x \\
&= \sum_{j=0}^{N_n}\sum_{i=1}^{r-1}(-1)^j M^{-j}\widetilde{R}(\lambda_{n-1})^{j+1}(-1)^i\binom{j+i}{i}\widetilde{R}(\lambda_{n-1})^i T((\mathrm{ad}a_k)^i a)x \\
&= \sum_{j=0}^{N_n}\sum_{i=0}^{r-2}(-1)^{i+j+1}M^{-j}\binom{j+i+1}{i+1}\widetilde{R}(\lambda_{n-1})^{i+j+2}\,T((\mathrm{ad}a_k)^{i+1}a)x \\
&= \sum_{j=0}^{N_n}\sum_{i=0}^{r-2}(-1)^{i+j+1}M^{-j}\sum_{n_2=0}^{j}\binom{n_2+i}{i}\widetilde{R}(\lambda_{n-1})^{i+j+2}\,T((\mathrm{ad}a_k)^{i+1}a)x \\
&= \sum_{\substack{0\leqslant n_1,n_2\leqslant N_n \\ n_1+n_2\leqslant N_n}}\sum_{i=0}^{r-2}(-1)^{n_1+n_2+i+1}M^{-(n_1+n_2)}\binom{n_2+i}{i}\widetilde{R}(\lambda_{n-1})^{n_1+n_2+i+2}\,T((\mathrm{ad}a_k)^{i+1}a)x
\end{aligned}$$

for all $x \in D$. Alternatively, it follows again from (3.4) that

$$\begin{aligned}
&\widetilde{R}(\lambda_n)\,T([a_k,a])\,\widetilde{R}(\lambda_n)x \\
&\quad= \sum_{n_1=0}^{N_n}(-1)^{n_1}M^{-n_1}\widetilde{R}(\lambda_{n-1})^{n_1+1}\,T([a_k,a])\,\sum_{n_2=0}^{N_n}(-1)^{n_2}M^{-n_2}\widetilde{R}(\lambda_{n-1})^{n_2+1}x \\
&\quad= \sum_{n_1=0}^{N_n}\sum_{n_2=0}^{N_n}\sum_{i=0}^{r-2}(-1)^{n_1+n_2+i}M^{-(n_1+n_2)}\binom{n_2+i}{i}\widetilde{R}(\lambda_{n-1})^{n_1+n_2+i+2}T((\mathrm{ad}a_k)^{i+1}a)x \\
&\qquad+ \sum_{n_1=0}^{N_n}\sum_{n_2=0}^{N_n}(-1)^{n_1+n_2}M^{-(n_1+n_2)}\widetilde{R}(\lambda_{n-1})^{n_1+1}P(\lambda_{n-1},[a_k,a],n_2+1).
\end{aligned}$$

Then the lemma follows by addition. ∎

LEMMA 3.4. *For all* $n \in \mathbb{N}_0$, $N_1,\ldots,N_n \in \mathbb{N}$, $m \in \mathbb{N}$ *and* $i \in \{0,\ldots,r-1\}$ *there exists an operator* $p_{i,m,n}\colon \mathcal{X} \to \mathcal{X}$ *such that*

$$P(\lambda_n,a,m)x = \sum_{i=0}^{r-1} p_{i,m,n}\,T((\mathrm{ad}a_k)^i a)x$$

for all $x \in D$ and $a \in \mathfrak{g}$. Moreover, $p_{i,m,n}$ can be chosen such that it is a polynomial in the $\widetilde{R}(\lambda_l)$ and $Z^{(j)}(\lambda_l)$ with $j \in \mathbb{N}_0$ and $l \in \{0, \ldots, n\}$; the coefficients of $p_{i,m,n}$ are independent of $N_1, \ldots, N_n$ and each term contains at least one factor $Z^{(j)}(\lambda_l)$ for some $j \in \mathbb{N}_0$ and $l \in \{0, \ldots, n\}$.

Proof. For all $n \in \mathbb{N}_0$ and $m \in \mathbb{N}$ let $H(n, m)$ be the following hypothesis:

> For all $N_1, \ldots, N_n \in \mathbb{N}$ and $i \in \{0, \ldots, r-1\}$ there exists an operator $p_{i,m,n} \colon \mathcal{X} \to \mathcal{X}$ such that
>
> $$P(\lambda_n, a, m)x = \sum_{i=0}^{r-1} p_{i,m,n}\, T((\mathrm{ad} a_k)^i a)x$$
>
> for all $x \in D$ and $a \in \mathfrak{g}$. Moreover, $p_{i,m,n}$ can be chosen such that it is a polynomial in the $\widetilde{R}(\lambda_l)$ and $Z^{(j)}(\lambda_l)$ with $j \in \mathbb{N}_0$ and $l \in \{0, \ldots, n\}$; the coefficients of $p_{i,m,n}$ are independent of $N_1, \ldots, N_n$ and each term contains at least one factor $Z^{(j)}(\lambda_l)$ for some $j \in \mathbb{N}_0$ and $l \in \{0, \ldots, n\}$.

If $n = 0$ and $m \in \mathbb{N}$ then clearly $H(0, m)$ is valid since $P(\lambda_0, a, m) = 0$ by Corollary 2.2 applied to λ_0. Let $n \in \mathbb{N}_0$ and suppose that $H(n, m)$ is valid for all $m \in \mathbb{N}$. By iteration it follows from Lemma 3.3 that

$$\begin{aligned} T(a)\, \widetilde{R}(\lambda_{n+1}) = \widetilde{R}(\lambda_{n+1}) \sum_{i=0}^{r-1} (-1)^i \widetilde{R}(\lambda_{n+1})^i\, T((\mathrm{ad} a_k)^i a)x \\ + \sum_{i=0}^{r-1} (-1)^i \widetilde{R}(\lambda_{n+1})^i P(\lambda_{n+1}, (\mathrm{ad} a_k)^i a). \end{aligned} \tag{3.6}$$

Hence $H(n+1, 1)$ is valid by Lemma 3.3 and the induction hypothesis.

Now let $m \in \mathbb{N}$ and suppose that $H(n+1, m)$ is valid. Then it follows as in the proof of Corollary 2.2 that

$$\begin{aligned} P(\lambda_{n+1}, a, m+1)x &= P(\lambda_{n+1}, a, m)\, \widetilde{R}(\lambda_{n+1})x \\ &\quad + \widetilde{R}(\lambda_{n+1})^m \sum_{i=0}^{r-1} (-1)^i \binom{m-1+i}{i} \widetilde{R}(\lambda_{n+1})^i \\ &\quad \circ \sum_{j=0}^{r-1} (-1)^j\, \widetilde{R}(\lambda_{n+1})^j\, P(\lambda_{n+1}, (\mathrm{ad} a_k)^{i+j} a)x \\ &= P(\lambda_{n+1}, a, m)\, \widetilde{R}(\lambda_{n+1})x \\ &\quad + \sum_{i=0}^{r-1} (-1)^i \binom{m+i}{i} \widetilde{R}(\lambda_{n+1})^{m+i}\, P(\lambda_{n+1}, (\mathrm{ad} a_k)^i a)x. \end{aligned}$$

Using (3.4) one can rewrite the first term as

$$\begin{aligned}P(\lambda_{n+1},a,m)\,\widetilde{R}(\lambda_{n+1})x &= \sum_{i=0}^{r-1} p_{i,m,n+1}T((\mathrm{ad}a_k)^i a)\,\widetilde{R}(\lambda_{n+1})x \\ &= \sum_{i,j=0}^{r-1}(-1)^j p_{i,m,n+1}\,\widetilde{R}(\lambda_{n+1})^{j+1}T((\mathrm{ad}a_k)^{i+j}a)x \\ &\quad + \sum_{i=0}^{r-1} p_{i,m,n+1}P(\lambda_{n+1},(\mathrm{ad}a_k)^i a,1)x.\end{aligned}$$

Hence $H(n+1,m+1)$ is valid. By induction the lemma follows. ∎

LEMMA 3.5. *If $n\in\mathbb{N}$, $N_1,\ldots,N_{n-1}\in\mathbb{N}$ and $\sum_{j=1}^{n-1}3^{-N_j+n-2-j}<6^{-1}$ then*

$$\lim_{N_n\to\infty}\|Z^{(j)}_{N_1,\ldots,N_n}(\lambda_n)\|=0$$

for all $j\in\mathbb{N}_0$.

Proof. For all $N_n\in\mathbb{N}$ one has by Lemma 3.2

$$\begin{aligned}\|Z^{(j)}_{N_1,\ldots,N_n}(\lambda_n)\| &\leqslant \sum_{\substack{0\leqslant n_1,n_2\leqslant N_n\\ N_n<n_1+n_2}} M^{-(n_1+n_2)}\binom{n_2+j}{j}(2^{-1}M)^{n_1+n_2+j+2} \\ &\leqslant (2^{-1}M)^{j+2}\sum_{n=N_n+1}^{\infty}(n+1)\,(n+j)^j\,2^{-n}\end{aligned}$$

from which the lemma follows. ∎

From now on fix $t>0$. For all $m\in\mathbb{N}$ let

$$n_m=\min\{n\in\mathbb{N}:\lambda_n\geqslant mt^{-1}\}.$$

LEMMA 3.6. *For all $m\in\mathbb{N}$ there exist $N_1,\ldots,N_{n_m}\in\mathbb{N}$ such that:*

$$\sup_{0\leqslant j\leqslant m}\lambda_{n_m}^m\|\widetilde{R}_{N_1,\ldots,N_{n_m}}(\lambda_{n_m})^j-R(\lambda_{n_m})^j\|\leqslant m^{-1} \tag{3.7}$$

$$\sup_{0\leqslant i\leqslant r-1}\lambda_{n_m}^m\|p_{i,m,n_m}\|\leqslant m^{-1} \tag{3.8}$$

where $p_{i,m,n}$ is as in Lemma 3.4.

Proof. This is a direct consequence of Lemmas 3.2, 3.4 and 3.5. ∎

Before we can prove Lemma 3.1 we need one more theorem.

THEOREM 3.7. *Let $\mathcal{X}$ be a Banach space, $F\colon(0,\infty)\to\mathcal{L}(X)$ a function, D a dense subspace of $\mathcal{X}$, $A\colon D\to\mathcal{X}$ a linear map and $\omega\geqslant 0$. Suppose the following:*

(i) *There exists an $M\geqslant 1$ such that $\|F(\rho)^k\|\leqslant Me^{\omega k\rho}$ for all $\rho>0$ and $k\in\mathbb{N}$.*

(ii) *$Ax=\lim_{\rho\downarrow 0}\rho^{-1}(F(\rho)x-x)$ for all $x\in D$.*

(iii) *There exists a* $\lambda \in \mathbb{C}$ *with* $\operatorname{Re}\lambda > \omega$ *such that* $(\lambda I - A)(D)$ *is dense in X.*
Then A is closable, the closure of $-A$ *generates a strongly continuous semigroup T and for all* $t > 0$, $\rho_1, \rho_2, \ldots > 0$ *and* $k_1, k_2, \ldots \in \mathbb{N}$ *with* $\lim_n \rho_n = 0$, $k_n \uparrow \infty$ *and* $\lim_n k_n \rho_n = t$ *one has*

$$T_t x = \lim_{n\to\infty} F(\rho_n)^{k_n} x$$

for all $x \in \mathcal{X}$.

Proof. See Corollaries III.5.3 and III.5.4 of [7]. ∎

Proof of Lemma 3.1. Recall that $k \in \{1, \ldots, d\}$ and $t > 0$ were fixed. For all $m \in \mathbb{N}$ set

$$\mathcal{W}_m^{(k)}(t) = \lambda_{n_m}^m \widetilde{R}_{N_1,\ldots,N_{n_m}}(\lambda_{n_m})^m,$$

where $N_1, \ldots, N_{n_m} \in \mathbb{N}$ are as in Lemma 3.6. (The value of, for example, N_1 is allowed to depend on m.) Define $F\colon (0,\infty) \to \mathcal{L}(\mathcal{X})$ by $F(\rho) = \rho^{-1} R(\rho^{-1})$ if $\rho^{-1} > \omega_k$ and $F(\rho) = I$ if $\rho^{-1} \leqslant \omega_k$. Then it follows from Theorem 3.7 that

$$\lim_{m\to\infty} F(\lambda_{n_m}^{-1})^m y = U^{(k)}(t) y$$

for all $y \in \mathcal{X}$. In addition, $\sup_m \|F(\lambda_{n_m}^{-1})^m\| < \infty$. Therefore $\lim_{m\to\infty} F(\lambda_{n_m}^{-1})^m x_m = U^{(k)}(t)x$. Hence by (3.7) one deduces that $\lim_{m\to\infty} \mathcal{W}_m^{(k)}(t) x_m = U^{(k)}(t)x$.

Let $\alpha \in J_r^+(d)$. Then (3.2) follows for $a = a_{[\alpha]}$ similarly to the proof in Lemma 2.3. Although there is one more term in (3.4) it follows from (3.8) that the additional item gives no contribution. Since $\{a_{[\alpha]} : \alpha \in J_r^+(d)\}$ spans $\mathfrak{g}$ the limit (3.2) is valid for all $a \in \mathfrak{g}$. ∎

The proof of Theorem 1.2 is complete.

4. THE CAMPBELL–BAKER–HAUSDORFF FORMULA AND EXAMPLES

In this section we give a consequence of the existence of the strongly continuous representation U of the Lie group G in the Banach space $\mathcal{X}$ obtained in Section 2. By Theorem 1.2 the closure of every commutator of order at most r generates a strongly continuous group and thanks to the representation U it is possible to find a formula for this group in terms of the strongly continuous groups generated by $A_1, \ldots, A_d$. Since $e^{t\overline{T(a)}} = U(\exp(ta))$ for all $a \in \mathfrak{g}$ and $t \in \mathbb{R}$, it suffices to find a formula for $\exp(ta)$ expressed in the $\exp(ta_k)$, and then apply the representation U. We assume the assumptions of Theorem 1.2 and use the notation of Section 2. Since $\mathfrak{g}$ is nilpotent and G is connected and simply connected the exponential map $\exp\colon \mathfrak{g} \to G$ is a bijection by Theorem 1.2.1 of [3].

In the sequel we need an expression for multi-commutators. This expression is given in Lemma 2.21 of [14], which we state here for convenience.

LEMMA 4.1. *If $l \in \mathbb{N}$ and $b_1, \ldots, b_l \in \{a_1, \ldots, a_d\}$ then for all $\alpha \in J^+(d)$ with $|\alpha| > l$ there are $c_\alpha \in \mathbb{R}$ such that*

$$\exp\big(t^l [b_1, [b_2, \ldots, [b_{l-1}, b_l], \ldots]] + \sum_{l < |\alpha| \leqslant r} t^{|\alpha|} c_\alpha a_{[\alpha]}\big) = C_l(t, b_1, \ldots, b_l) \tag{4.1}$$

for all $t \in \mathbb{R}$, where $C_1(t, b_1) = \exp(tb_1)$ and by recursion

$$C_l(t, b_1, \ldots, b_l) = \exp(tb_1)\, C_{l-1}(t, b_2, \ldots, b_l)\, \exp(-tb_1)\, (C_{l-1}(t, b_2, \ldots, b_l))^{-1}.$$

Proof. See Lemma 2.21 of [14]. Note that $a_{[\alpha]} = 0$ if $|\alpha| > r$. ∎

By Lemma 4.1 we obtain the following result if $l = r$.

PROPOSITION 4.2. *If $b_1, \ldots, b_r \in \{a_1, \ldots, a_d\}$ then*

$$\exp(t^r [b_1, [b_2, \ldots, [b_{r-1}, b_r], \ldots]]) = C_r(t, b_1, \ldots, b_r)$$

for all $t \in \mathbb{R}$.

Now consider a multi-commutator $a = [b_1, [b_2, \ldots, [b_{r-2}, b_{r-1}], \ldots]]$ of order $r - 1$. Then by Lemma 4.1 we have

$$\exp(t^{r-1} [b_1, [b_2, \ldots, [b_{r-2}, b_{r-1}], \ldots]] + t^r R_r) = C_{r-1}(t, b_1, \ldots, b_{r-1})$$

for all $t \in \mathbb{R}$, where R_r is a linear combination of commutators of order r. Multiplying by $\exp(-t^r R_r)$ on the right and using that a and R_r commute one deduces from the Campbell–Baker–Hausdorff formula that

$$\exp(t^{r-1} [b_1, [b_2, \ldots, [b_{r-2}, b_{r-1}], \ldots]]) = C_{r-1}(t, b_1, \ldots, b_{r-1}) \exp(-t^r R_r).$$

This gives an explicit formula for $\exp(t^{r-1} a)$ expressed in the $\exp(ta_k)$ whenever a is a multi-commutator of order $r - 1$. Using the Campbell–Baker–Hausdorff formula once again one can find an explicit formula for $\exp(t^{r-1} a)$ whenever a is a linear combination of multi-commutators of order $r - 1$. Continuing this way, starting with (4.1) and using the Campbell–Baker–Hausdorff formula repeatedly, it is possible to find an explicit formula for $\exp\big(\sum_{|\alpha| \geqslant j} t^{|\alpha|} c_\alpha a_{[\alpha]}\big)$ expressed in the $\exp(ta_k)$ whenever the c_α are constant. This is by downward induction on j. Since $\mathfrak{g}$ is nilpotent one needs to apply the Campbell–Baker–Hausdorff formula only a finite number of times. Nevertheless, the final explicit formulae get very long if r is large. We write down an explicit formula only for the case $r = 3$, since the formulae for $r \in \{1, 2\}$ are well known.

PROPOSITION 4.3. *Let $d \in \mathbb{N}$, $t \in \mathbb{R}$, $r = 3$ and $b_1, b_2, b_3 \in \{a_1, \ldots, a_d\}$. Then*

$$\begin{aligned}
\exp(t^3[b_1,[b_2,b_3]]) &= \exp(tb_1)\,\exp(tb_2)\,\exp(tb_3)\,\exp(-tb_2)\,\exp(-tb_3) \\
&\quad\cdot \exp(-tb_1)\,\exp(tb_3)\,\exp(tb_2)\,\exp(-tb_3)\,\exp(-tb_2), \\
\exp(t^2[b_1,b_2]) &= \exp(tb_1)\,\exp(tb_2)\,\exp(-tb_1)\,\exp(-tb_2) \\
&\quad\cdot \exp\big(-\tfrac{t^3}{2}[b_1+b_2,[b_1,b_2]]\big) \text{ and} \\
\exp(t(b_1+b_2)) &= \exp(tb_1)\,\exp(tb_2)\,\exp\big(-\tfrac{t^3}{12}([b_1,[b_1,b_2]]-[b_2,[b_1,b_2]])\big) \\
&\quad\cdot \exp\big(-\tfrac{t^2}{2}[b_1,b_2]\big)\,\exp\big(\tfrac{t^3}{4}[b_1+b_2,[b_1,b_2]]\big).
\end{aligned}$$

Proof. We only prove the last identity. By the Campbell–Baker–Hausdorff formula and the assumption $r = 3$ one has

$$\exp(tb_1)\,\exp(tb_2) = \exp\big(t(b_1+b_2) + \tfrac{t^2}{2}[b_1,b_2] + \tfrac{t^3}{12}([b_1,[b_1,b_2]]-[b_2,[b_1,b_2]])\big).$$

Multiplying by $\exp\big(-\frac{t^3}{12}([b_1,[b_1,b_2]]-[b_2,[b_1,b_2]])\big)$ on the right one deduces that

$$\begin{aligned}
&\exp\big(t(b_1+b_2) + \tfrac{t^2}{2}[b_1,b_2]\big) \\
&\qquad = \exp(tb_1)\,\exp(tb_2)\,\exp\big(-\tfrac{t^3}{12}([b_1,[b_1,b_2]]-[b_2,[b_1,b_2]])\big).
\end{aligned}$$

Now multiply by $\exp\big(-\frac{t^2}{2}[b_1,b_2]\big)$ on the right. Since it does not commute with $\exp(t(b_1+b_2))$, one obtains

$$\begin{aligned}
&\exp\big(t(b_1+b_2) - \tfrac{t^3}{4}[b_1+b_2,[b_1,b_2]]\big) \\
&\quad = \exp(tb_1)\exp(tb_2)\exp\big(-\tfrac{t^3}{12}([b_1,[b_1,b_2]]-[b_2,[b_1,b_2]])\big)\exp\big(-\tfrac{t^2}{2}[b_1,b_2]\big).
\end{aligned}$$

Finally multiplying by $\exp\big(\frac{t^3}{4}[b_1+b_2,[b_1,b_2]]\big)$ we find the required formula for $\exp(t(b_1+b_2))$. ∎

EXAMPLE 4.4. Let $m \in \mathbb{N}$. Let $p \in (1,\infty)$ and set $\mathcal{X} = L_p(\mathbb{R}^{2m})$ or set $\mathcal{X} = C_0(\mathbb{R}^{2m})$. Let $N \in \{1,\ldots,m\}$ and $n_1,\ldots,n_N \in \mathbb{N}_0$. Set $D = \mathcal{S}(\mathbb{R}^{2m})$, the Schwartz space. For all $k \in \{1,\ldots,N\}$ and $\varphi \in D$ define $\mathcal{A}_k, \mathcal{A}_{N+k} \colon D \to D$ by

$$\mathcal{A}_k\varphi = \partial_k\varphi \quad\text{and}\quad (\mathcal{A}_{N+k}\varphi)(x) = x_k^{n_k}\,(\partial_{m+k}\varphi)(x).$$

Then $\mathcal{A}_k$ and $\mathcal{A}_{N+k}$ are closable and the closures $A_k = \overline{\mathcal{A}_k}$ and $A_{N+k} = \overline{\mathcal{A}_{N+k}}$ generate strongly continuous groups on $\mathcal{X}$. It is not hard to verify that the assumptions of Theorem 1.2 are satisfied with $d = 2N$ and $r = \max\{n_1,\ldots,n_N\}$. Hence an arbitrary linear combination of (multi-)commutators of arbitrary order is closable and the closure generates a strongly continuous group. Moreover, by (1.4), the operator

$$-\sum_{k=1}^{N}(\partial_k^2 + x_k^{2n_k}\,\partial_{N+k}^2)$$

is closable on its natural domain and the closure generates an analytic semigroup. If one chooses $d' = N$ and $c_{kl} = \delta_{kl}$ then one deduces from (1.5) that the operator

$$-\sum_{k=1}^{N} \partial_k^2 + \sum_{k=1}^{N} x_k^{n_k} \partial_{N+k}$$

is closable on its natural domain and the closure generates a strongly continuous semigroup. In particular, choosing $N = m$ and $n_1 = \cdots = n_N = 1$ one obtains the so called Kolmogorov operator in $\mathbb{R}^{2m}$. Alternatively, if one chooses $N = m = 1$ and $n_1 = n$ then one obtain the operator

$$-\partial_1^2 - x_1^{2n}\, \partial_2^2,$$

which is the Grushin operator.

Finally, choose $N = m$, $n_1 = \cdots = n_N = 0$, and define $\mathcal{B}\colon D \to D$ by $(\mathcal{B}\varphi)(x) = \mathrm{i}p(x)\, \varphi(x)$, where p is a polynomial. Then $\mathcal{B}$ is closable and the closure $B = \overline{\mathcal{B}}$ generates a strongly continuous group on $\mathcal{X}$. Moreover, the operators $B, A_1, \ldots, A_{2m}$ satisfy the assumptions of Theorem 1.2 with $d = 2m + 1$ and r equal to the degree of the polynomial p. Hence all the conclusion of Theorem 1.2 are valid and in particular by (1.4) the closure of the operator

$$-\sum_{k=1}^{2m} \partial_k^2 + p^2 I$$

is the generator of an analytic semigroup.

Acknowledgements. The first author would like to thank Alessandra Lunardi for her helpful suggestions and encouragement on this subject and Wolfgang Arendt for his hospitality in Ulm and his interesting discussions. The main part of this paper was written whilst the second author was visiting the University of Ulm. He wishes to thank Wolfgang Arendt for his hospitality, the many discussions and financial support.

REFERENCES

[1] O. BRATTELI, F.M. GOODMAN, P.E.T. JØRGENSEN, D.W. ROBINSON, The heat semigroup and integrability of Lie algebras, *J. Funct. Anal.* **79**(1988), 351–397.

[2] O. BRATTELI, D.W. ROBINSON, *Operator Algebras and Quantum Statistical Mechanics*, vol. 1, Second edition, Springer-Verlag, New York 1987.

[3] L. CORWIN, F.P. GREENLEAF, *Representations of Nilpotent Lie Groups and their Applications. Part 1: Basic Theory and Examples*, Cambridge Stud. Adv. Math., vol. 18, Cambridge Univ. Press, Cambridge 1990.

[4] G. DA PRATO, P. GRISVARD, Sommes d'opérateurs linéaires et équations différentielles opérationnelles, *J. Math. Pures Appl.* **54**(1975), 305–387.

[5] D. DI GIORGIO, Sums and commutators of generators of noncommuting strongly continuous groups, *J. Math. Anal. Appl.* **275**(2002), 165–187.

[6] N. DUNGEY, A.F.M. TER ELST, D.W. ROBINSON, *Analysis on Lie Groups with Polynomial Growth*, Progr. Math., vol. 214, Birkhäuser Boston Inc., Boston 2003.

[7] K.-J. ENGEL, R. NAGEL, *One-Parameter Semigroups for Linear Evolution Equations*, Graduate Texts in Math., vol. 194, Springer-Verlag, New York 2000.

[8] A.F.M. TER ELST, D.W. ROBINSON, Subcoercive and subelliptic operators on Lie groups: variable coefficients, *Publ. Res. Inst. Math. Sci.* **29**(1993), 745–801.

[9] A.F.M. TER ELST, D.W. ROBINSON, Weighted subcoercive operators on Lie groups, *J. Funct. Anal.* **157**(1998), 88–163.

[10] P.E.T. JØRGENSEN, Representations of differential operators on a Lie group, *J. Funct. Anal.* **20**(1975), 105–135.

[11] P.E.T. JØRGENSEN, *Operators and Representation Theory*, North-Holland, Amsterdam 1988.

[12] P.E.T. JØRGENSEN, R.T. MOORE, *Operator Commutation Relations*, D. Reidel Publishing Co., Dordrecht 1984.

[13] R. LABBAS, Some results on the sum of linear operators with nondense domains, *Ann. Mat. Pura Appl.* **154**(1989), 91–97.

[14] A. NAGEL, E.M. STEIN, S. WAINGER, Balls and metrics defined by vector fields. I: basic properties, *Acta Math.* **155**(1985), 103–147.

[15] E. NELSON, Analytic vectors, *Ann. Math.* **70**(1959), 572–615.

[16] D.W. ROBINSON, *Elliptic Operators and Lie Groups*, Oxford Math. Monographs, Oxford Univ. Press, Oxford 1991.

[17] J. RUSINEK, A theorem on the integration of a Lie algebra representation, *Bull. Acad. Polon. Sci. Sér. Sci. Math.* **28**(1980), 427–431.

[18] J. RUSINEK, Analytic vectors and integrability of Lie algebra representations, *J. Funct. Anal.* **74**(1987), 10–23.

DAVIDE DI GIORGIO, DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DI PISA, VIA BUONARROTI 2, 56127 PISA, ITALY

E-mail address: davedg@virgilio.it

A.F.M. TER ELST, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF AUCKLAND, P.B. 92019, AUCKLAND, NEW ZEALAND

E-mail address: terelst@math.auckland.ac.nz

Received September 22, 2004; revised November 18, 2005.