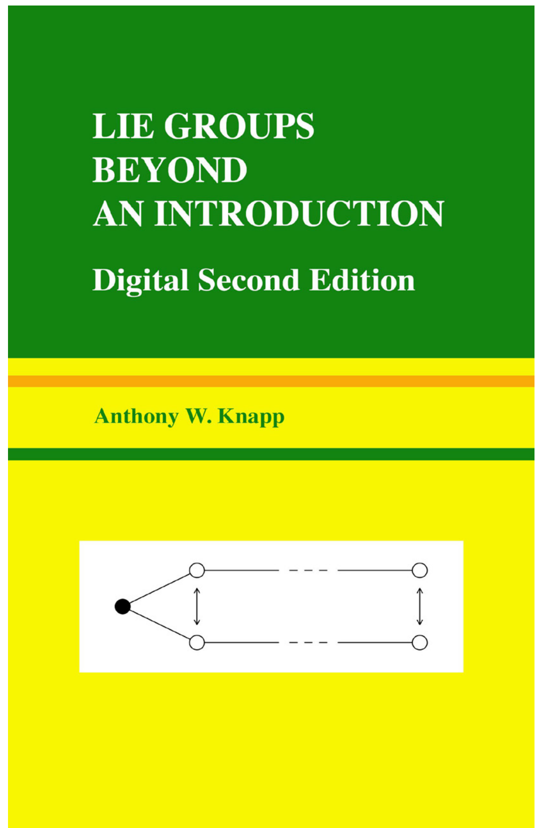
Appendix B. Lie's Third Theorem, 659-682

DOI: 10.3792/euclid/9798989504206-13

from

Lie Groups
Beyond an Introduction
Digital Second Edition, 2023

Anthony W. Knapp



Full Book DOI: 10.3792/euclid/9798989504206
ISBN: 979-8-9895042-0-6

Anthony W. Knapp
81 Upper Sheep Pasture Road
East Setauket, N.Y. 11733–1729
U.S.A.
Email to: aknapp@math.stonybrook.edu
Homepage: www.math.stonybrook.edu/~aknapp

Lie Groups Beyond an Introduction, Digital Second Edition

Pages vii–xviii and 1–812 are the same in the digital and printed second editions. A list of corrections as of June 2023 has been included as pages 813–820 of the digital second edition. The corrections have not been implemented in the text.

Cover: Vogan diagram of $\mathfrak{sl}(2n, \mathbb{R})$. See page 399.

AMS Subject Classifications: 17-01, 22-01

APPENDIX B

Lie's Third Theorem

Abstract. A finite-dimensional real Lie algebra is the semidirect product of a semisimple subalgebra and the solvable radical, according to the Levi decomposition. As a consequence of this theorem and the correspondence between semidirect products of Lie algebras and semidirect products of simply connected analytic groups, every finite-dimensional real Lie algebra is the Lie algebra of an analytic group. This is Lie's Third Theorem.

Ado's Theorem says that every finite-dimensional real Lie algebra admits a one-one finite-dimensional representation on a complex vector space. This result sharpens Lie's Third Theorem, saying that every real Lie algebra is the Lie algebra of an analytic group of matrices.

The Campbell–Baker–Hausdorff Formula expresses the multiplication rule near the identity in an analytic group in terms of the linear operations and bracket multiplication within the Lie algebra. Thus it tells constructively how to pass from a finite-dimensional real Lie algebra to the multiplication rule for the corresponding analytic group in a neighborhood of the identity.

1. Levi Decomposition

Chapter I omits several important theorems about general finite-dimensional Lie algebras over $\mathbb{R}$ related to the realization of Lie groups, and those results appear in this appendix. They were omitted from Chapter I partly because in this treatment they use a result about semisimple Lie algebras that was not proved until Chapter V. One of the results in this appendix uses also some material from Chapter III.

Lemma B.1. Let φ be an $\mathbb{R}$ linear representation of the real semisimple Lie algebra $\mathfrak{g}$ on a finite-dimensional real vector space V. Then V is completely reducible in the sense that there exist invariant subspaces $U_1, \dots, U_r$ of V such that $V = U_1 \oplus \cdots \oplus U_r$ and such that the restriction of the representation to each U_i is irreducible.

PROOF. It is enough to prove that any invariant subspace U of V has an invariant complement W. By Theorem 5.29, there exists an invariant complex subspace W' of $V^{\mathbb{C}}$ such that $V^{\mathbb{C}} = U^{\mathbb{C}} \oplus W'$. Let P be the $\mathbb{R}$ linear projection of $V^{\mathbb{C}}$ on V along iV, and put

$$W = P(W' \cap (V \oplus iU)).$$

Since P commutes with $\varphi(\mathfrak{g})$, we see that $\varphi(\mathfrak{g})(W) \subseteq W$. To complete the proof, we show that $V = U \oplus W$.

Let a be in $U \cap W$. Then $a + ib$ is in $W' \cap (V \oplus iU)$ for some $b \in V$. The element b must be in U, and we know that a is in U. Hence $a + ib$ is in $U^{\mathbb{C}}$. But then $a + ib$ is in $U^{\mathbb{C}} \cap W' = 0$, and $a = 0$. Hence $U \cap W = 0$.

Next let $v \in V$ be given. Since $V^{\mathbb{C}} = U^{\mathbb{C}} + W'$, we can write $v = (a + ib) + (x + iy)$ with $a \in U$, $b \in U$, and $x + iy \in W'$. Since v is in V, $y = -b$. Therefore $x + iy$ is in $V \oplus iU$, as well as W'. Since $P(x + iy) = x$, x is in W. Then $v = a + x$ with $a \in U$ and $x \in W$, and $V = U + W$.

Theorem B.2 (Levi decomposition). If $\mathfrak{g}$ is a finite-dimensional Lie algebra over $\mathbb{R}$, then there exists a semisimple subalgebra $\mathfrak{s}$ of $\mathfrak{g}$ such that $\mathfrak{g}$ is the semidirect product $\mathfrak{g} = \mathfrak{s} \oplus_{\pi} (\operatorname{rad} \mathfrak{g})$ for a suitable homomorphism $\pi : \mathfrak{s} \to \operatorname{Der}_{\mathbb{R}}(\operatorname{rad} \mathfrak{g})$.

PROOF. Let $\mathfrak{r} = \operatorname{rad} \mathfrak{g}$. We begin with two preliminary reductions. The first reduction will enable us to assume that there is no nonzero ideal $\mathfrak{a}$ of $\mathfrak{g}$ properly contained in $\mathfrak{r}$. In fact, an argument by induction on the dimension would handle such a situation: Proposition 1.11 shows that the radical of $\mathfrak{g}/\mathfrak{a}$ is $\mathfrak{r}/\mathfrak{a}$. Hence induction gives $\mathfrak{g}/\mathfrak{a} = \mathfrak{s}/\mathfrak{a} \oplus \mathfrak{r}/\mathfrak{a}$ with $\mathfrak{s}/\mathfrak{a}$ semisimple. Since $\mathfrak{s}/\mathfrak{a}$ is semisimple, $\mathfrak{a} = \operatorname{rad} \mathfrak{s}$. Then induction gives $\mathfrak{s} = \mathfrak{s}' \oplus \mathfrak{a}$ with $\mathfrak{s}'$ semisimple. Consequently $\mathfrak{g} = \mathfrak{s}' \oplus \mathfrak{r}$, and $\mathfrak{s}'$ is the required complementary subalgebra.

As a consequence, $\mathfrak{r}$ is abelian. In fact, otherwise Proposition 1.7 shows that $[\mathfrak{r}, \mathfrak{r}]$ is an ideal in $\mathfrak{g}$, necessarily nonzero and properly contained in $\mathfrak{r}$. So the first reduction eliminates this case.

The second reduction will enable us to assume that $[\mathfrak{g}, \mathfrak{r}] = \mathfrak{r}$. In fact, $[\mathfrak{g}, \mathfrak{r}]$ is an ideal of $\mathfrak{g}$ contained in $\mathfrak{r}$. The first reduction shows that we may assume it is 0 or $\mathfrak{r}$. If $[\mathfrak{g}, \mathfrak{r}] = 0$, then the real representation ad of $\mathfrak{g}$ on $\mathfrak{g}$ descends to a real representation of $\mathfrak{g}/\mathfrak{r}$ on $\mathfrak{g}$. Since $\mathfrak{g}/\mathfrak{r}$ is semisimple, Lemma B.1 shows that the action is completely reducible. Thus $\mathfrak{r}$, which is an invariant subspace in $\mathfrak{g}$, has an invariant complement, and we may take this complement as $\mathfrak{s}$.

As a consequence,

$$\mathfrak{r} \cap Z_{\mathfrak{g}} = 0. \tag{B.3}$$

In fact $\mathfrak{r} \cap Z_{\mathfrak{g}}$ is an ideal of $\mathfrak{g}$. It is properly contained in $\mathfrak{r}$ since $\mathfrak{r} \cap Z_{\mathfrak{g}} = \mathfrak{r}$ implies that $[\mathfrak{g}, \mathfrak{r}] = 0$, in contradiction with the second reduction. Therefore the first reduction implies (B.3).

With the reductions in place, we imitate some of the proof of Theorem 5.29. That is, we put

$$V = \{\gamma \in \operatorname{End}\mathfrak{g} \mid \gamma(\mathfrak{g}) \subseteq \mathfrak{r} \text{ and } \gamma|_{\mathfrak{r}} \text{ is scalar}\}$$

and define a representation σ of $\mathfrak{g}$ on $\operatorname{End}\mathfrak{g}$ by

$$\sigma(X)\gamma = (\operatorname{ad} X)\gamma - \gamma(\operatorname{ad} X) \qquad \text{for } \gamma \in \operatorname{End}\mathfrak{g} \text{ and } X \in \mathfrak{g}.$$

The subspace V is an invariant subspace under σ, and

$$U = \{\gamma \in V \mid \gamma = 0 \text{ on } \mathfrak{r}\}$$

is an invariant subspace of codimension 1 in V such that $\sigma(X)(V) \subseteq U$ for $X \in \mathfrak{g}$. Let

$$T = \{\operatorname{ad} Y \mid Y \in \mathfrak{r}\}.$$

This is a subspace of U since $\mathfrak{r}$ is an abelian Lie subalgebra. If X is in $\mathfrak{g}$ and $\gamma = \operatorname{ad} Y$ is in T, then $\sigma(X)\gamma = \operatorname{ad}[X, Y]$ with $[X, Y] \in \mathfrak{r}$. Hence T is an invariant subspace under σ.

From $V \supseteq U \supseteq T$, we can form the quotient representations V/T and V/U. The natural map of V/T onto V/U respects the $\mathfrak{g}$ actions, and the $\mathfrak{g}$ action of V/U is 0 since $\sigma(X)(V) \subseteq U$ for $X \in \mathfrak{g}$. If X is in $\mathfrak{r}$ and γ is in V, then

$$\sigma(X)\gamma = (\operatorname{ad} X)\gamma - \gamma(\operatorname{ad} X) = -\gamma(\operatorname{ad} X)$$

since image $\gamma \subseteq \mathfrak{r}$ and $\mathfrak{r}$ is abelian. Since γ is a scalar $\lambda(\gamma)$ on $\mathfrak{r}$, we can rewrite this formula as

$$\sigma(X)\gamma = \operatorname{ad}(-\lambda(\gamma)X). \tag{B.4}$$

Equation (B.4) exhibits $\sigma(X)\gamma$ as in T. Thus $\sigma|_{\mathfrak{r}}$ maps V into T, and σ descends to representations of $\mathfrak{g}/\mathfrak{r}$ on V/T and V/U. The natural map of V/T onto V/U respects these $\mathfrak{g}/\mathfrak{r}$ actions.

Since $\dim V/U = 1$, the kernel of $V/T \to V/U$ is a $\mathfrak{g}/\mathfrak{r}$ invariant subspace of V/T of codimension 1, necessarily of the form W/T with $W \subseteq V$. Since $\mathfrak{g}/\mathfrak{r}$ is semisimple, Lemma B.1 allows us to write

$$V/T = W/T \oplus (\mathbb{R}\gamma_0 + T)/T \tag{B.5}$$

for a 1-dimensional invariant subspace $(\mathbb{R}\gamma_0 + T)/T$. The directness of this sum means that γ_0 is not in U. So γ_0 is not 0 on $\mathfrak{r}$. Normalizing, we may assume that γ_0 acts by the scalar -1 on $\mathfrak{r}$. In view of (B.4), we have

$$\sigma(X)\gamma_0 = \operatorname{ad} X \qquad \text{for } X \in \mathfrak{r}. \tag{B.6}$$

Since $(\mathbb{R}\gamma_0 + T)/T$ is invariant in (B.5), we have $\sigma(X)\gamma_0 \in T$ for each $X \in \mathfrak{g}$. Thus we can write $\sigma(X)\gamma_0 = \operatorname{ad}\varphi(X)$ for some $\varphi(X) \in \mathfrak{r}$. The element $\varphi(X)$ is unique by (B.3), and therefore φ is a linear function $\varphi : \mathfrak{g} \to \mathfrak{r}$. By (B.6), φ is a projection. If we put $\mathfrak{s} = \ker\varphi$, then we have $\mathfrak{g} = \mathfrak{s} \oplus \mathfrak{r}$ as vector spaces, and we have only to show that $\mathfrak{s}$ is a Lie subalgebra. The subspace $\mathfrak{s} = \ker\varphi$ is the set of all X such that $\sigma(X)\gamma_0 = 0$. This is the set of all X such that $(\operatorname{ad} X)\gamma_0 = \gamma_0(\operatorname{ad} X)$. Actually if γ is any element of $\operatorname{End}\mathfrak{g}$, then the set of $X \in \mathfrak{g}$ such that $(\operatorname{ad} X)\gamma = \gamma(\operatorname{ad} X)$ is always a Lie subalgebra. Hence $\mathfrak{s}$ is a Lie subalgebra, and the proof is complete.

2. Lie's Third Theorem

Lie's Third Theorem, which Lie proved as a result about vector fields and local Lie groups, has come to refer to the following improved theorem due to Cartan.

Theorem B.7. Every finite-dimensional Lie algebra over $\mathbb{R}$ is isomorphic to the Lie algebra of an analytic group.

PROOF. Let $\mathfrak{g}$ be given, and write $\mathfrak{g} = \mathfrak{s} \oplus_\pi \mathfrak{r}$ as in Theorem B.2, with $\mathfrak{s}$ semisimple and $\mathfrak{r}$ solvable. Corollary 1.126 shows that there is a simply connected Lie group R with Lie algebra isomorphic to $\mathfrak{r}$. The group $\operatorname{Int}\mathfrak{s}$ is an analytic group with Lie algebra $\operatorname{ad}\mathfrak{s}$ isomorphic to $\mathfrak{s}$ since $\mathfrak{s}$ has center 0. Let S be the universal covering group of $\operatorname{Int}\mathfrak{s}$. By Theorem 1.125 there exists a unique action τ of S on R by automorphisms such that $d\bar{\tau} = \pi$, and $G = S \times_\tau R$ is a simply connected analytic group with Lie algebra isomorphic to $\mathfrak{g} = \mathfrak{s} \oplus_\pi \mathfrak{r}$.

3. Ado's Theorem

Roughly speaking, Ado's Theorem is the assertion that every Lie algebra over $\mathbb{R}$ has a one-one representation on some finite-dimensional complex vector space. This theorem can be regarded as sharpening Lie's Third Theorem: Each real Lie algebra is not merely the Lie algebra of an analytic group; it is the Lie algebra of an analytic group of complex matrices.

Throughout this section, $\mathfrak{g}$ will denote a finite-dimensional Lie algebra over $\mathbb{R}$, and $U(\mathfrak{g}^{\mathbb{C}})$ will be the universal enveloping algebra of its complexification.

Theorem B.8 (Ado's Theorem). Let $\mathfrak{g}$ be a finite-dimensional Lie algebra over $\mathbb{R}$, let $\operatorname{rad}\mathfrak{g}$ be its radical, and let $\mathfrak{n}$ be its unique largest nilpotent ideal given as in Corollary 1.41. Then there exists a one-one finite-dimensional representation φ of $\mathfrak{g}$ on a complex vector space such that $\varphi(Y)$ is nilpotent for every Y in $\mathfrak{n}$. If $\mathfrak{g}$ is complex, then φ can be taken to be complex linear.

The proof of the theorem will be preceded by two lemmas. The second lemma is the heart of the matter, using the left Noetherian property of universal enveloping algebras (Proposition 3.27) to prove that a certain natural representation is finite dimensional.

The last statement of the theorem is something that we shall dispose of now. Proving this extension of the theorem amounts to going over the entire argument to see that, in every case, real vector spaces and Lie algebras can be replaced by complex vector spaces and Lie algebras and that Lie algebras that get complexified when $\mathfrak{g}$ is real do not need to be complexified when $\mathfrak{g}$ is complex. In Theorem B.2 the representation ad is complex linear, and no new analog of Lemma B.1 is needed; Theorem 5.29 is enough by itself. In the proof of Theorem B.2 and the argument that is about to come, when $\mathfrak{g}$ is complex, so is $\operatorname{rad}\mathfrak{g}$ and so is the unique largest nilpotent ideal. In Lemmas B.9 and B.12, $U(\mathfrak{g}^{\mathbb{C}})$ and $T(\mathfrak{g}^{\mathbb{C}})$ are simply to be replaced by $U(\mathfrak{g})$ and $T(\mathfrak{g})$, and $\operatorname{Der}_{\mathbb{R}}\mathfrak{g}$ and $\operatorname{End}_{\mathbb{R}}\mathfrak{g}$ are to be replaced by $\operatorname{Der}_{\mathbb{C}}\mathfrak{g}$ and $\operatorname{End}_{\mathbb{C}}\mathfrak{g}$. The details are all routine, and we omit them.

As in Appendix A, a **derivation** $D : A \to A$ of an associative algebra A with identity is a linear mapping such that $D(uv) = (Du)v + u(Dv)$ for all u and v in A. A derivation automatically has $D(1) = 0$.

Lemma B.9. Any derivation d of a real Lie algebra $\mathfrak{g}$ extends uniquely to a derivation $\widetilde{d}$ of $U(\mathfrak{g}^{\mathbb{C}})$ to itself.

PROOF. Uniqueness is clear since monomials span $U(\mathfrak{g}^{\mathbb{C}})$ and since the assumptions determine $\widetilde{d}$ on monomials.

For existence we use Proposition A.16 to construct a derivation D of $T(\mathfrak{g}^{\mathbb{C}})$ extending d. To get D to descend to a derivation $\widetilde{d}$ of $U(\mathfrak{g}^{\mathbb{C}})$, we need to see that D carries

$$\ker(T(\mathfrak{g}^{\mathbb{C}}) \to U(\mathfrak{g}^{\mathbb{C}})) \tag{B.10}$$

to itself, i.e., that

$$D(u(X \otimes Y - Y \otimes X - [X, Y])v) \quad \text{is in (B.10)} \tag{B.11}$$

for all monomials u and v in $T(\mathfrak{g}^{\mathbb{C}})$ and for all X and Y in $\mathfrak{g}$. The derivation D acts on one factor of a product at a time. If it acts in a factor of u or v, then the factor $(X \otimes Y - Y \otimes X - [X, Y])$ is left alone by D, and the corresponding term of (B.11) is in (B.10). Next suppose it acts on the middle factor, leaving u and v alone. Since d is a derivation of $\mathfrak{g}$, we have

$$\begin{aligned} D(X \otimes Y - Y \otimes X - [X, Y]) & \\ &= (dX \otimes Y + X \otimes dY) - (dY \otimes X + Y \otimes dX) \\ &\quad - ([dX, Y] + [X, dY]) \\ &= (dX \otimes Y - Y \otimes dX - [dX, Y]) \\ &\quad + (X \otimes dY - dY \otimes X - [X, dY]). \end{aligned}$$

The right side is the sum of two members of (B.10), and thus the remaining terms of (B.11) are in (B.10). Thus D descends to give a definition of $\widetilde{d}$ on $U(\mathfrak{g}^{\mathbb{C}})$.

Lemma B.12. Let $\mathfrak{g}$ be a real solvable Lie subalgebra of $\mathfrak{gl}(N, \mathbb{C})$, let $\mathfrak{d}$ be the Lie subalgebra $\mathrm{Der}_{\mathbb{R}}\, \mathfrak{g}$ of $\mathrm{End}_{\mathbb{R}}\, \mathfrak{g}$, and let π be the natural action of $\mathfrak{d}$ on $\mathfrak{g}$. Suppose that all members of the largest nilpotent ideal $\mathfrak{n}$ of $\mathfrak{g}$ are nilpotent matrices. Then there exists a one-one representation φ of the semidirect product $\mathfrak{d} \oplus_\pi \mathfrak{g}$ such that $\varphi(d + Y)$ is nilpotent whenever Y is in $\mathfrak{n}$ and the member d of $\mathfrak{d}$ is nilpotent as a member of $\mathrm{End}_{\mathbb{R}}\, \mathfrak{g}$.

PROOF. Let $\mathcal{G}$ be the complex associative algebra of matrices generated by $\mathfrak{g}$ and 1. By Proposition 3.3 the inclusion of $\mathfrak{g}$ into $\mathcal{G}$ extends to an associative algebra homomorphism $\rho : U(\mathfrak{g}^{\mathbb{C}}) \to \mathcal{G}$ sending 1 into 1. Let I be the kernel of ρ. Since $\mathcal{G}$ is finite dimensional, I is a two-sided ideal of finite codimension in $U(\mathfrak{g}^{\mathbb{C}})$.

Using Lemma B.9, we extend each derivation d of $\mathfrak{g}$ to a derivation $\widetilde{d}$ of $U(\mathfrak{g}^{\mathbb{C}})$. Let $\mathcal{D}$ be the complex associative algebra of linear mappings of $U(\mathfrak{g}^{\mathbb{C}})$ into itself generated by 1 and all the extensions $\widetilde{d}$.

Let $I_0 \subseteq I$ be the subset of all $u \in I$ such that Du is in I for all $D \in \mathcal{D}$. We prove that I_0 is an ideal in $U(\mathfrak{g}^{\mathbb{C}})$. It is certainly a vector subspace. To see that it is a left ideal, let a be in $U(\mathfrak{g}^{\mathbb{C}})$, let u be in I_0, and let $D = \widetilde{d}_1 \cdots \widetilde{d}_k$ be a monomial in $\mathcal{D}$. When we apply D to au, we obtain a sum of 2^k terms; each term is of the form $(D_1 a)(D_2 u)$, with D_1 equal to the product of a subset of the $\widetilde{d}_j$ and D_2 equal to the product of the complementary subset. Since u is in I_0, each $D_2 u$ is in I, and hence $(D_1 a)(D_2 u)$ is in I. Consequently $D(au)$ is in I for all $D \in \mathcal{D}$, and u is in I_0. Thus I_0 is a left ideal, and a similar argument shows that it is a right ideal.

Recall that the members of $\mathfrak{g}$ are N-by-N matrices. We are going to obtain the space of the desired representation φ as $U(\mathfrak{g}^{\mathbb{C}})/I_0$. The finite dimensionality of this space will follow from Corollary 3.28 (a consequence of the left Noetherian property of $U(\mathfrak{g}^{\mathbb{C}})$) once we prove that

$$I^N \subseteq I_0 \subseteq I. \tag{B.13}$$

By Lie's Theorem (Corollary 1.29) we may regard the N-by-N matrices in $\mathfrak{g}$ as upper triangular. By assumption the matrices in $\mathfrak{n}$ are nilpotent. Since the latter matrices are simultaneously upper triangular and nilpotent, we see that $Y_1 \cdots Y_N$ is the 0 matrix for any $Y_1, \ldots, Y_N$ in $\mathfrak{n}$. Lifting this result back via ρ to a statement about $U(\mathfrak{g}^{\mathbb{C}})$, we conclude that

$$Y_1 \cdots Y_N \quad \text{is in} \quad I \tag{B.14}$$

whenever all Y_j lie in $\mathfrak{n} \subseteq U(\mathfrak{g}^{\mathbb{C}})$.

Let J be the two-sided ideal in $U(\mathfrak{g}^{\mathbb{C}})$ generated by the members of $\mathfrak{n}$. Toward proving (B.13), we first show that (B.14) implies

$$J^N \subseteq I. \tag{B.15}$$

Let us begin by showing that inductively on s that if Y is in $\mathfrak{n}$ and $X_1, \ldots, X_s$ are in $\mathfrak{g}$, then

$$X_1 \cdots X_s Y \quad \text{is in} \quad \mathfrak{n}U(\mathfrak{g}^{\mathbb{C}}). \tag{B.16}$$

This is trivial for $s = 0$. If s is ≥ 1 and if (B.16) holds for $s - 1$, then

$$X_1 \cdots X_s Y = X_1 \cdots X_{s-1} Y X_s + X_1 \cdots X_{s-1}[X_s, Y].$$

Since $[X_s, Y]$ is in $\mathfrak{n}$, the inductive hypothesis shows that both terms on the right side are in $\mathfrak{n}U(\mathfrak{g}^{\mathbb{C}})$. Thus (B.16) follows for s. Consequently we obtain

$$U(\mathfrak{g}^{\mathbb{C}})\mathfrak{n} \subseteq \mathfrak{n}U(\mathfrak{g}^{\mathbb{C}}). \tag{B.17}$$

From (B.17) it follows that $(u_1Y_1u'_1)(u_2Y_2u'_2)$ is a sum of terms of the form $u_1Y_1Y'_2u''_2$. Thus we can argue inductively on r that

$$(u_1Y_1u'_1)(u_2Y_2u'_2)\cdots(u_rY_ru'_r) \tag{B.18}$$

is a sum of terms of the form $u_1Y_1Y'_2\cdots Y'_ru''_r$. For $r = N$, the latter terms are in I by (B.14). Every member of J^N is a sum of terms (B.18) with $r = N$, and thus (B.15) follows.

From Proposition 1.40 we know that $d(X)$ is in $\mathfrak{n}$ for any $d \in \mathfrak{d}$ and $X \in \mathfrak{g}$. If $\widetilde{d}$ denotes the extension of d to $U(\mathfrak{g}^{\mathbb{C}})$, then it follows from the derivation property of $\widetilde{d}$ that

$$\widetilde{d}(U(\mathfrak{g}^{\mathbb{C}})) \subseteq J. \tag{B.19}$$

From another application of the derivation property, we obtain $\widetilde{d}(J^N) \subseteq J^N$. Taking products of such derivations and using (B.15), we see that $D(J^N) \subseteq J^N \subseteq I$ for all $D \in \mathcal{D}$. Therefore

$$J^N \subseteq I_0. \tag{B.20}$$

Now we can finish the proof of (B.13), showing that $I^N \subseteq I_0$. Certainly $I^N \subseteq I$. Let $u_1, \ldots, u_N$ be in I, and let D be a monomial in $\mathcal{D}$. By the derivation property, $D(u_1\cdots u_N)$ is a linear combination of terms $(D_1u_1)\cdots(D_Nu_N)$ with D_j a monomial in $\mathcal{D}$. If some D_j has degree 0, then D_ju_j is in I, and the corresponding term $(D_1u_1)\cdots(D_Nu_N)$ is in I since I is a two-sided ideal. If all D_j have degree > 0, then (B.19) shows that all D_ju_j are in J. The corresponding term $(D_1u_1)\cdots(D_Nu_N)$ is then in J^N and is in I by (B.15). Thus all terms of $D(u_1\cdots u_N)$ are in I, and $u_1\cdots u_N$ is in I_0. This proves (B.13).

As was mentioned earlier, it follows from Corollary 3.28 that $\mathcal{G}^* = U(\mathfrak{g}^{\mathbb{C}})/I_0$ is finite dimensional. Let $u \mapsto u^*$ be the quotient map. Then $\mathcal{G}^*$ is a unital $U(\mathfrak{g}^{\mathbb{C}})$ module, and we obtain a representation φ of $\mathfrak{g}$ on it by the definition

$$\varphi(X)(u^*) = (Xu)^*. \tag{B.21a}$$

Since I_0 is stable under $\mathcal{D}$, each d in $\mathfrak{d}$ induces a derivation $\varphi(d)$ of $\mathcal{G}^*$ given by

$$\varphi(d)u^* = (\widetilde{d}u)^*. \tag{B.21b}$$

Formula (B.21b) defines a representation of $\mathfrak{d} = \mathrm{Der}_{\mathbb{R}}\, \mathfrak{g}$ on $\mathcal{G}^*$ because the uniqueness in Lemma B.9 implies that

$$\widetilde{[d_1, d_2]} = \widetilde{d}_1 \widetilde{d}_2 - \widetilde{d}_2 \widetilde{d}_1.$$

Proposition 1.22 observes that $\mathfrak{d} \oplus_\pi \mathfrak{g}$ becomes a semidirect-product Lie algebra, and φ, as defined in (B.21), is a representation of $\mathfrak{d} \oplus_\pi \mathfrak{g}$ because

$$\begin{aligned}
[\varphi(d), \varphi(X)]u^* &= \varphi(d)\varphi(X)u^* - \varphi(X)\varphi(d)u^* \\
&= \varphi(d)(Xu)^* - \varphi(X)(\widetilde{d}u)^* \\
&= (\widetilde{d}(Xu))^* - (X\widetilde{d}u)^* \\
&= ((\widetilde{d}X)u + X\widetilde{d}u)^* - (X\widetilde{d}u)^* \\
&= \varphi(dX)u^* \\
&= \varphi([d, X])u^*.
\end{aligned}$$

Now let us show that φ is one-one as a representation of $\mathfrak{d} \oplus_\pi \mathfrak{g}$. If $\varphi(d + X) = 0$, then

$$0 = \varphi(d + X)1^* = (\widetilde{d}1)^* + (X1)^* = X^*.$$

Then X is in $I_0 \subseteq I$, and $X = 0$ as a member of $\mathfrak{g}$. So $\varphi(d) = 0$. Every X' in $\mathfrak{g}$ therefore has

$$0 = \varphi(d)(X')^* = (\widetilde{d}X')^* = (dX')^*.$$

Hence dX' is in $I_0 \subseteq I$, and $dX' = 0$ as a member of $\mathfrak{g}$. So d is the 0 derivation. We conclude that φ is one-one.

To complete the proof, we show that $\varphi(d + Y)$ is nilpotent whenever Y is in $\mathfrak{n}$ and d is nilpotent as a member of $\mathrm{End}_{\mathbb{R}}\, \mathfrak{g}$. To begin with, $\varphi(Y)$ is nilpotent because (B.14) gives

$$(\varphi(Y))^N u^* = (Y^N u)^* = 0$$

for every u. Next, let us see that $\varphi(d)$ is nilpotent. In fact, let $\mathcal{G}_n = U_n(\mathfrak{g}^{\mathbb{C}})$, so that $\mathcal{G}_n^*$ is the subspace $U_n(\mathfrak{g}^{\mathbb{C}}) + I_0$ of $\mathcal{G}^*$. If $d^p = 0$, we show by induction on $n \geq 1$ that $\widetilde{d}^{np}(\mathcal{G}_n) = 0$. It is enough to handle monomials in $\mathcal{G}_n$. For $n = 1$, $\mathcal{G}_1$ is just $\mathbb{C} + \mathfrak{g}^{\mathbb{C}}$, and we have $\widetilde{d}^p(1) = 0$ and $\widetilde{d}^p X = d^p X = 0$ for X in $\mathfrak{g}^{\mathbb{C}}$. For general n, suppose that $\widetilde{d}^{(n-1)p}(\mathcal{G}_{n-1}) = 0$. Any monomial of $\mathcal{G}_n$ is of the form Xu with $X \in \mathfrak{g}^{\mathbb{C}}$ and $u \in \mathcal{G}_{n-1}$. Powers of a derivation satisfy the Leibniz rule, and therefore $\widetilde{d}^{np}(Xu) = \sum_{k=0}^{np} \binom{np}{k} (\widetilde{d}^k X)(\widetilde{d}^{np-k}u)$. The factor $\widetilde{d}^k X$ is 0 for $k \geq p$, and the factor $\widetilde{d}^{np-k}u$ is 0 for $k \leq p$; thus $\widetilde{d}^{np}(Xu) = 0$, and we have proved that $\widetilde{d}^{np}(\mathcal{G}_n) = 0$. Then we have $\varphi(d)^{np}(\mathcal{G}_n^*) = (\widetilde{d}^{np}\mathcal{G}_n)^* = 0$. Since $\mathcal{G}^*$ is finite dimensional and $\bigcup \mathcal{G}_n^* = \mathcal{G}^*$, $\varphi(d)^{np}(\mathcal{G}^*) = 0$ for n large enough. Hence $\varphi(d)$ is nilpotent.

Now that we know $\varphi(d)$ and $\varphi(Y)$ to be nilpotent, let us form the solvable Lie subalgebra $\mathbb{R}d \oplus_\pi \mathfrak{n}$ of $\mathfrak{d} \oplus_\pi \mathfrak{g}$. It is a Lie subalgebra since $d(\mathfrak{g}) \subseteq \mathfrak{n}$, and it is solvable since $\mathbb{R}d$ is abelian. By Lie's Theorem (Corollary 1.29), we may choose a basis of $\mathcal{G}^*$ such that the matrix of every member of $\varphi(\mathbb{R}d + \mathfrak{n})$ is upper triangular. Since $\varphi(d)$ and $\varphi(Y)$ are nilpotent, their matrices are strictly upper triangular and hence the sum of the matrices is strictly upper triangular. Consequently $\varphi(d + Y)$ is nilpotent.

PROOF OF THEOREM B.8. We begin with the special case in which $\mathfrak{g}$ is solvable, so that $\mathfrak{g} = \operatorname{rad}\mathfrak{g} \supseteq \mathfrak{n}$. We proceed by induction on $\dim \mathfrak{g}$. If $\dim \mathfrak{g} = 1$, then $\mathfrak{g} \cong \mathbb{R}$, and $\varphi_1(t) = \begin{pmatrix} 0 & t \\ 0 & 0 \end{pmatrix}$ is the required representation.

Suppose that $\mathfrak{g}$ is solvable with $\dim \mathfrak{g} = n > 1$, that the theorem has been proved for solvable Lie algebras of dimension $< n$, and that $\mathfrak{n}$ is the largest nilpotent ideal in $\mathfrak{g}$. By Proposition 1.23, $\mathfrak{g}$ contains an elementary sequence—a sequence of subalgebras going from 0 to $\mathfrak{g}$ one dimension at a time such that each is an ideal in the next. Moreover, the last members of this sequence can be taken to be any subspaces between $[\mathfrak{g}, \mathfrak{g}]$ and $\mathfrak{g}$ that go up one dimension at a time. Proposition 1.39 shows that $[\mathfrak{g}, \mathfrak{g}] \subseteq \mathfrak{n}$, and we may thus take $\mathfrak{n}$ to be one of the members of the elementary sequence.

Let $\mathfrak{h}$ be the member of the elementary sequence of codimension 1 in $\mathfrak{g}$, let $\mathfrak{n}_\mathfrak{h}$ be its largest nilpotent ideal, and let X be a member of $\mathfrak{g}$ not in $\mathfrak{h}$. By inductive hypothesis we can find a one-one finite-dimensional representation φ_0 of $\mathfrak{h}$ such that $\varphi_0(Y)$ is nilpotent for all $Y \in \mathfrak{n}_\mathfrak{h}$. There are now two cases.

Case 1: $\operatorname{ad} X = 0$. Then $\operatorname{ad} X$ is nilpotent and X lies in $\mathfrak{n}$. Our construction forces $\mathfrak{n} = \mathfrak{g}$. Hence $\mathfrak{h}$ is nilpotent and $\mathfrak{g}$ must be the direct sum of $\mathbb{R}X$ and $\mathfrak{h}$. Let us write members of $\mathfrak{g}$ as pairs (t, Y) with $t \in \mathbb{R}$ and $Y \in \mathfrak{h}$. Then $\varphi(t, Y) = \varphi_1(t) \oplus \varphi_0(Y)$ is the required representation.

Case 2: $\operatorname{ad} X \neq 0$. We apply Lemma B.12 to the solvable Lie algebra $\varphi_0(\mathfrak{h})$. Let $\mathfrak{d} = \operatorname{Der}_\mathbb{R} \mathfrak{h}$. The lemma gives us a one-one finite-dimensional representation φ of the semidirect product $\mathfrak{d} \oplus \mathfrak{h}$ such that $\varphi(d + Y)$ is nilpotent for all $Y \in \mathfrak{n}_\mathfrak{h}$ and all nilpotent $d \in \mathfrak{d}$. We restrict this to the Lie subalgebra $\mathbb{R}(\operatorname{ad} X) \oplus \mathfrak{h}$, which is isomorphic with $\mathfrak{g}$. We consider separately the subcases that $\mathfrak{g}$ is nilpotent and $\mathfrak{g}$ is not nilpotent.

Subcase 2a: $\mathfrak{g}$ is nilpotent. Then the member $\operatorname{ad} X$ of $\mathfrak{d}$ is nilpotent by (1.31), and thus every member of $\varphi(\mathbb{R}(\operatorname{ad} X) \oplus \mathfrak{h})$ is nilpotent. So φ, interpreted as a representation of $\mathfrak{g}$, is the required representation.

Subcase 2b: $\mathfrak{g}$ is not nilpotent. Then $\mathfrak{n}$ is a nilpotent ideal of $\mathfrak{h}$ and we must have $\mathfrak{n} \subseteq \mathfrak{n}_{\mathfrak{h}}$. Again φ, interpreted as a representation of $\mathfrak{g}$, is the required representation: Since X is not in $\mathfrak{n}$, $\operatorname{ad} X$ is not nilpotent, and no nonzero derivation in $\mathbb{R}(\operatorname{ad} X)$ is nilpotent. We know that every member of $\varphi(\mathfrak{n}_{\mathfrak{h}})$ is nilpotent, and thus every member of $\varphi(\mathfrak{n})$ is nilpotent.

This completes the induction, and the theorem has now been proved for $\mathfrak{g}$ solvable.

Now we consider the general case in which $\mathfrak{g}$ does not need to be solvable. Let $\operatorname{rad} \mathfrak{g}$ be the largest solvable ideal of $\mathfrak{g}$, and let $\mathfrak{n}$ be the largest nilpotent ideal. By the special case we can find a one-one finite-dimensional representation ψ of $\operatorname{rad} \mathfrak{g}$ such that every member of $\psi(\mathfrak{n})$ is nilpotent. Let $\mathfrak{d} = \operatorname{Der}_{\mathbb{R}}(\operatorname{rad} \mathfrak{g})$. We apply Lemma B.12 to the solvable Lie algebra $\psi(\operatorname{rad} \mathfrak{g})$, obtaining a one-one finite-dimensional representation φ_1 of $\mathfrak{d} \oplus \psi(\operatorname{rad} \mathfrak{g})$ such that $\varphi_1(d + \psi(Y))$ is nilpotent whenever Y is in $\mathfrak{n}$ and d is a nilpotent member of $\mathfrak{d}$.

We apply the Levi decomposition of Theorem B.2 to write $\mathfrak{g}$ as a semidirect product $\mathfrak{s} \oplus \operatorname{rad} \mathfrak{g}$ with $\mathfrak{s}$ semisimple. For $S \in \mathfrak{s}$ and $X \in \operatorname{rad} \mathfrak{g}$, define $\varphi_2(S + X) = \operatorname{ad} S$ as a representation of $\mathfrak{g}$ on $\mathfrak{s}^{\mathbb{C}}$. Then we put

$$\varphi(S + X) = \varphi_1(\operatorname{ad} S + \psi(X)) \oplus \varphi_2(S + X)$$

as a representation of $\mathfrak{g}$ on the direct sum of the spaces for φ_1 and φ_2.

If $\varphi(S + X) = 0$, then $\varphi_2(S + X) = 0$ and $\operatorname{ad} S = 0$. Since $\mathfrak{s}$ is semisimple, $S = 0$. Therefore $\varphi(X) = 0$ and $\varphi_1(\psi(X)) = 0$. Since ψ if one-one on $\operatorname{rad} \mathfrak{g}$ and φ_1 is one-one on $\psi(\operatorname{rad} \mathfrak{g})$, we obtain $X = 0$. We conclude that φ is one-one.

Finally if Y is in $\mathfrak{n}$, then $\varphi_1(\psi(Y))$ is nilpotent by construction, and $\varphi_2(Y)$ is 0 since Y has no $\mathfrak{s}$ term. Therefore $\varphi(Y)$ is nilpotent for every Y in $\mathfrak{n}$.

4. Campbell–Baker–Hausdorff Formula

The theorem to be proved in this section is the following.

Theorem B.22 (Campbell–Baker–Hausdorff Formula). Let G be an analytic group with Lie algebra $\mathfrak{g}$. Then for all A and B sufficiently close to 0 in $\mathfrak{g}$, $\exp A \exp B = \exp C$, where

$$C = A + B + H_2 + \cdots + H_n + \cdots \tag{B.23}$$

is a convergent series in which $H_2 = \frac{1}{2}[A, B]$ and H_n is a finite linear combination of expressions $(\operatorname{ad} X_1) \cdots (\operatorname{ad} X_{n-1}) X_n$ with each X_j equal to either A or B. The particular linear combinations that occur may be taken to be independent of G, as well as of A and B.

A way of getting at the formula explicitly comes by thinking of G as $GL(N, \mathbb{C})$ and using the formula from complex-variable theory

$$z = \log e^z = \log(1 + (e^z - 1)) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k} \Big(\sum_{n=1}^{\infty} \frac{1}{n!} z^n \Big)^k,$$

valid for $|z| < \log 2$ since $|e^z - 1| \le e^{|z|} - 1$. Because the sum of a convergent power series determines its coefficients, an identity of this kind forces identities on the coefficients; for example, the sum of the contributions from the right side to the coefficient of z is 1, the sum of the contributions from the right side to the coefficient of z^2 is 0, etc. Hence the identity has to be correct in a ring of formal power series. Then we can substitute a matrix C, and we still have an identity if we have convergence. Thus we obtain

(B.24)

$$\begin{aligned} C &= \sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k} (e^C - 1)^k \\ &= \sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k} (e^A e^B - 1)^k \\ &= \sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k} \Big(\Big(\sum_{m=0}^{\infty} \frac{1}{m!} A^m \Big) \Big(\sum_{n=0}^{\infty} \frac{1}{n!} B^n \Big) - 1 \Big)^k \\ &= \sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k} \Big((A + B) + \frac{1}{2!}(A^2 + 2AB + B^2) \\ &\qquad\qquad + \frac{1}{3!}(A^3 + 3A^2 B + 3AB^2 + B^3) + \cdots \Big)^k, \end{aligned}$$

and H_n will have to be the sum of the terms on the right side that are homogeneous of degree n, rewritten in terms of brackets. For example, the quadratic term is

$$\tfrac{1}{2!}(A^2 + 2AB + B^2) - \tfrac{1}{2}(A + B)^2 = \tfrac{1}{2}(2AB - AB - BA) = \tfrac{1}{2}[A, B],$$

as stated in the theorem. Similar computation shows that

$$H_3 = \tfrac{1}{12}[A, [A, B]] + \tfrac{1}{12}[B, [B, A]] \quad \text{and} \quad H_4 = -\tfrac{1}{24}[A, [B, [A, B]]].$$

These formulas are valid as long as A and B are matrices that are not too large: The first line of (B.24) is valid if $\|C\| < \log 2$. The entire computation is valid if also $\|A\| + \|B\| < \log 2$ since

$$\begin{aligned}\|e^A e^B - 1\| &\le \|e^A - 1\|\|e^B\| + \|e^B - 1\| \\ &\le (e^{\|A\|} - 1)e^{\|B\|} + (e^{\|B\|} - 1) \\ &= e^{\|A\|+\|B\|} - 1.\end{aligned}$$

This calculation indicates two important difficulties in the proof of Theorem B.22. First, although the final formula (B.23) makes sense for any G, the intermediate formula (B.24) and its terms like A^2B do not make sense in general. We were able to use such expressions by using the matrix product operation within the associative algebra $\mathcal{A}_N$ of all N-by-N complex matrices. Thus (B.24) is a formula that may help with $GL(N, \mathbb{C})$, but it has no meaning for general G. To bypass this difficulty, we shall use Ado's Theorem, Theorem B.8. We formalize matters as in the first reduction below.

A second important difficulty is that it is not obvious even in $GL(N, \mathbb{C})$ that the homogeneous terms of (B.24) can be rewritten as linear combinations of iterated brackets. Handling this step requires a number of additional ideas, and we return to this matter shortly.

FIRST REDUCTION. In order to prove Theorem B.22, it is enough to prove, within the associative algebra of all N-by-N complex matrices, that the sum of the terms of

$$\sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k} \Big((A + B) + \frac{1}{2!}(A^2 + 2AB + B^2) \\ + \frac{1}{3!}(A^3 + 3A^2B + 3AB^2 + B^3) + \cdots \Big)^k$$

that are homogeneous of degree n, for $n \ge 2$, is a linear combination of expressions $(\operatorname{ad} X_1) \cdots (\operatorname{ad} X_{n-1}) X_n$ with each X_j equal to A or B, the particular combination being independent of N.

PROOF OF FIRST REDUCTION. The hypothesis is enough to imply the theorem for $GL(N, \mathbb{C})$. In fact, choose an open neighborhood U about $C = 0$ in $\mathfrak{gl}(N, \mathbb{C})$ where the exponential map is a diffeomorphism, then choose neighborhoods of $A = 0$ and $B = 0$ such that $e^A e^B$ lies in $\exp U$, and then cut down the neighborhoods of $A = 0$ and $B = 0$ further so that the computation (B.24) is valid. The hypothesis then allows us to rewrite the homogeneous terms of (B.24) as iterated brackets, and the theorem follows.

Let G be a general analytic group, and use Theorem B.8 to embed its Lie algebra $\mathfrak{g}$ in some $\mathfrak{gl}(N, \mathbb{C})$. Let G_1 be the analytic subgroup of $GL(N, \mathbb{C})$ with Lie algebra $\mathfrak{g}$, so that G and G_1 are locally isomorphic and it is enough to prove the theorem for G_1. Choose an open neighborhood U_1 about $C = 0$ in $\mathfrak{g}$ where $\exp : \mathfrak{g} \to G_1$ is a diffeomorphism, then choose open neighborhoods of $A = 0$ and $B = 0$ in $\mathfrak{g}$ such that $\exp A \exp B$ lies in $\exp U_1$, and then, by continuity of the inclusions $\mathfrak{g} \subseteq \mathfrak{gl}(N, \mathbb{C})$ and $G_1 \subseteq GL(N, \mathbb{C})$, cut down these neighborhoods so that they lie in the neighborhoods constructed for $GL(N, \mathbb{C})$ in the previous paragraph. The partial sums in (B.23) lie in $\mathfrak{g}$, and they converge in $\mathfrak{gl}(N, \mathbb{C})$. Thus they converge in $\mathfrak{g}$. Since the exponential maps for G_1 and $GL(N, \mathbb{C})$ are continuous and are consistent with each other, formula (B.23) in $GL(N, \mathbb{C})$ implies validity of (B.23) in G_1.

Let A and B denote distinct elements of some set, and define a 2-dimensional complex vector space by $V = \mathbb{C}A \oplus \mathbb{C}B$. Let $T(V)$ be the corresponding tensor algebra. We shall omit the tensor signs in writing out products in $T(V)$. For u in V and v in $T(V)$, define $(\operatorname{ad} u)v$ and $[u, v]$ to mean $uv - vu$. By Proposition A.14, the linear map ad of V into $\operatorname{End}_{\mathbb{C}} T(V)$ extends to an algebra homomorphism ad of $T(V)$ into $\operatorname{End}_{\mathbb{C}} T(V)$ sending 1 to 1. For this extension, $\operatorname{ad}(u_1 u_2)v$ is $(\operatorname{ad} u_1)(\operatorname{ad} u_2)v$, not $u_1 u_2 v - v u_1 u_2$.

SECOND REDUCTION. In order to prove Theorem B.22, it is enough to prove, within the tensor algebra $T(V)$, that the sum of the terms of the formal sum

$$\sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k} \Big((A + B) + \frac{1}{2!}(A^2 + 2AB + B^2) + \frac{1}{3!}(A^3 + 3A^2B + 3AB^2 + B^3) + \cdots \Big)^k \tag{B.25}$$

that are homogeneous of degree n, for $n \geq 2$, is a finite linear combination of expressions $(\operatorname{ad} X_1) \cdots (\operatorname{ad} X_{n-1}) X_n$ with each X_j equal to A or B.

PROOF OF SECOND REDUCTION. Let $\mathcal{A}_N$ be the associative algebra of all N-by-N complex matrices, and let A and B be given in $\mathcal{A}_N$. The linear mapping of V into $\mathcal{A}_N$ that sends the abstract elements A and B into the matrices with the same names extends to an associative algebra homomorphism of $T(V)$ into $\mathcal{A}_N$. If the asserted expansion in terms of brackets in $T(V)$ is valid, then it is valid in $\mathcal{A}_N$ as well, and the first reduction shows that Theorem B.22 follows.

Now we come to the proof that the expression in (B.25) may be written as asserted in the second reduction. We isolate three steps as lemmas and then proceed with the proof.

Lemma B.26. For any X in $T(V)$ and for $m \geq 1$,

$$\begin{aligned} XB^{m-1} + BXB^{m-2} + \cdots + B^{m-1}X \\ &= \binom{m}{1} XB^{m-1} + \binom{m}{2}((\operatorname{ad} B)X)B^{m-2} \\ &\quad + \binom{m}{3}((\operatorname{ad} B)^2 X)B^{m-3} + \cdots + \binom{m}{m}(\operatorname{ad} B)^{m-1}X. \end{aligned}$$

PROOF. If X is a polynomial $P(B)$ in B, then the identity reduces to $mP(B)B^{m-1} = mP(B)B^{m-1}$, and there is nothing to prove. Thus we may assume that X is not such a polynomial.

Write $L(B)$ and $R(B)$ for the operators on $T(V)$ of left and right multiplication by B. These commute, and $L(B) = R(B) + \operatorname{ad} B$ shows that $R(B)$ and $\operatorname{ad} B$ commute. Therefore the binomial theorem may be used to compute powers of $R(B) + \operatorname{ad} B$, and we obtain

$$\begin{aligned} &(\operatorname{ad} B)(L(B)^{m-1} + L(B)^{m-2}R(B) + \cdots + R(B)^{m-1}) \\ &= (L(B) - R(B))(L(B)^{m-1} + L(B)^{m-2}R(B) + \cdots + R(B)^{m-1}) \\ &= L(B)^m - R(B)^m \\ &= (R(B) + \operatorname{ad} B)^m - R(B)^m \\ &= \binom{m}{1}R(B)^{m-1}(\operatorname{ad} B) + \binom{m}{2}R(B)^{m-2}(\operatorname{ad} B)^2 + \cdots + \binom{m}{m}(\operatorname{ad} B)^m \\ &= (\operatorname{ad} B)\Big(\binom{m}{1}R(B)^{m-1} + \binom{m}{2}R(B)^{m-2}(\operatorname{ad} B) + \cdots + \binom{m}{m}(\operatorname{ad} B)^{m-1}\Big). \end{aligned}$$

We apply both sides of this identity to X. If H denotes the difference of the left and right sides in the statement of the lemma, what we have just showed is that $(\operatorname{ad} B)H = 0$. A look at H shows that H is of the form

$$c_0 B^{m-1}X + c_1 B^{m-2}XB + \cdots + c_{m-1}XB^{m-1},$$

and ad B of this is

$$c_0 B^m X + (c_1 - c_0)B^{m-1}XB + \cdots + (c_{m-1} - c_{m-2})BXB^{m-1} - c_{m-1}XB^m.$$

To obtain the conclusion $H = 0$, which proves the lemma, it is therefore enough to show that the elements $B^m X, B^{m-1}XB, \ldots, XB^m$ are linearly independent in $T(V)$.

Since X is not a polynomial in B, we can write $X = (c + PA + QB)B^k$ with $k \geq 0$, $c \in \mathbb{C}$, $P \in T(V)$, $Q \in T(V)$, and $P \neq 0$. Assume a linear relation among $B^m X, B^{m-1}XB, \ldots, XB^m$, and substitute for X in it. The resulting monomials with A as close as possible to the right end force all coefficients in the linear relation to be 0, and the linear independence follows. This proves the lemma.

It will be handy to express the above lemma in a slightly different language. For X in $T(V)$, let d_X be the linear map of V into $T(V)$ given by

$$d_X(aA + bB) = bX,$$

and extend d_X to a derivation D_X of $T(V)$ by means of Proposition A.16.

If $P(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_M z^M$ is any ordinary polynomial, we define $P(B) = a_0 + a_1 B + a_2 B^2 + \cdots + a_M B^M$. The derivatives $P'(z), P''(z), \ldots$ are polynomials as well, and thus it is meaningful to speak of $P'(B), P''(B), \ldots$.

Lemma B.27. If $P(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_M z^M$ is a polynomial of degree M, then

$$\begin{aligned} D_X(P(B)) = X\,\frac{P'(B)}{1!} &+ ((\operatorname{ad} B)X)\,\frac{P''(B)}{2!} \\ &+ ((\operatorname{ad} B)^2 X)\,\frac{P'''(B)}{3!} + \cdots + ((\operatorname{ad} B)^{M-1}X)\,\frac{P^{(M)}(B)}{M!}\,. \end{aligned}$$

PROOF. The special case of this result when $P(z) = z^m$ is exactly Lemma B.26. In fact, $D_X(P(B))$ is the left side of the expression in that lemma, and the right side here is the right side of the expression in that lemma. Thus Lemma B.27 follows by taking linear combinations.

Let $T^{\leq M}(V) = \bigoplus_{n=0}^{M} T^n(V)$ and $T^{>M}(V) = \bigoplus_{n=M+1}^{\infty} T^n(V)$. The space $T^{>M}(V)$ is a two-sided ideal in $T(V)$, but $T^{\leq M}(V)$ is just a subspace. We have $T(V) = T^{\leq M}(V) \oplus T^{>M}(V)$ as vector spaces. Because $T^{>M}(V)$ is an ideal, the projection π_M of $T(V)$ on $T^{\leq M}(V)$ along $T^{>M}(V)$ satisfies

$$\pi_M(uv) = \pi_M((\pi_M u)(\pi_M v)) = \pi_M(u(\pi_M v)) \tag{B.28}$$

for all u and v in $T(V)$.

Again let X be a member of $T(V)$. From now on, we assume X has no constant term. Since X has no constant term, the derivation D_X carries $T^n(V)$ to $T^{>n-1}(V)$ for all n. Then it follows that

$$\pi_M D_X = \pi_M D_X \pi_M. \tag{B.29}$$

Since $T^{\leq M}(V)$ is finite dimensional, the exponential of a member of $\mathrm{End}_{\mathbb{C}}(T^{\leq M}(V))$ is well defined. For z in $\mathbb{C}$, we apply this observation to $z\pi_M D_X \pi_M$. We shall work with

$$\pi_M \exp(z\pi_M D_X \pi_M) = \exp(z\pi_M D_X \pi_M)\pi_M. \tag{B.30}$$

Put

$$C(z) = C_M(z) = \pi_M \exp(z\pi_M D_X \pi_M)(B). \tag{B.31}$$

For each $z \in \mathbb{C}$, this is a member of $T(V)$ without constant term. For $z = 0$, we have $C(0) = B$ for all $M > 0$.

Lemma B.32. For any integer $k \geq 0$,

$$\pi_M(C(z)^k) = \pi_M \exp(z\pi_M D_X \pi_M)(B^k).$$

PROOF. Without loss of generality we may assume $k \geq 1$. Then

$$\begin{aligned}
\frac{dC(t)}{dt} &= \frac{d}{dt}\pi_M \exp(t\pi_M D_X \pi_M)(B) \\
&= \pi_M \frac{d}{dt} \exp(t\pi_M D_X \pi_M)(B) && \text{since } \pi_M \text{ is linear} \\
&= (\pi_M D_X \pi_M) \exp(t\pi_M D_X \pi_M)(B) && \text{by Proposition 0.11d} \\
&= (\pi_M D_X) C(t),
\end{aligned}$$

and hence

$$\begin{aligned}
\frac{d}{dt}\,\pi_M(C(t)^k) &= \pi_M\Big(\frac{d}{dt}\,C(t)^k\Big) \\
&= \pi_M\Big(\frac{dC(t)}{dt}\,C(t)^{k-1} + C(t)\,\frac{dC(t)}{dt}\,C(t)^{k-2} + \cdots + C(t)^{k-1}\,\frac{dC(t)}{dt}\Big) \\
&= \pi_M((\pi_M D_X C(t))C(t)^{k-1} + C(t)(\pi_M D_X C(t))C(t)^{k-2} \\
&\qquad + \cdots + C(t)^{k-1}(\pi_M D_X C(t))) \\
&= \pi_M((D_X C(t))C(t)^{k-1} + C(t)(D_X C(t))C(t)^{k-2} \\
&\qquad + \cdots + C(t)^{k-1}(D_X C(t))) \qquad \text{by (B.28) and (B.29)} \\
&= (\pi_M D_X)(C(t)^k).
\end{aligned}$$

Therefore, using (B.29), we find

$$\Big(\frac{d}{dt}\Big)^m \pi_M(C(t)^k) = (\pi_M D_X)^m(C(t)^k).$$

Since $z \mapsto \pi_M(C(z)^k)$ is analytic,

$$\begin{aligned}
\pi_M(C(z)^k) &= \pi_M \sum_{m=0}^{\infty} \frac{z^m}{m!}\,\Big(\frac{d}{dt}\Big)^m \pi_M(C(t)^k)\Big|_{t=0} \\
&= \pi_M \sum_{m=0}^{\infty} \frac{z^m}{m!}\,(\pi_M D_X)^m(C(0)^k) \\
&= \pi_M \sum_{m=0}^{\infty} \frac{z^m}{m!}\,(\pi_M D_X \pi_M)^m(C(0)^k) \quad \text{by (B.29)} \\
&= \pi_M \exp(z\pi_M D_X \pi_M)(C(0)^k),
\end{aligned}$$

and the lemma follows.

PROOF OF THEOREM B.22. According to the statement of the second reduction, what needs proof is that, in the formal expression (B.25), the sum of the terms homogeneous of each particular degree greater than 1 is a finite linear combination of iterated brackets involving A and B. Let M be an odd integer greater than the degree of homogeneity to be addressed. Let X be an element in $T(V)$ without constant term; X will be specified shortly.

Define $E_M(z) = 1 + z + z^2/2! + \cdots + z^M/M!$ to be the M^{th} partial sum of the power series for e^z. Then $E_M(B)$ is in $T^{\le M}(V)$. The derivatives of this particular polynomial have the property that $\pi_{M-k}(E_M^{(k)}(B)) = \pi_{M-k}(E_M(B))$ for $0 \le k \le M$. If Y_k is in $T^{>k-1}(V)$, then it follows that

$$\pi_M(Y_k E_M^{(k)}(B)) = \pi_M(Y_k E_M(B)).$$

Applying Lemma B.27 with $P = E_M$ and taking $Y_k = (\operatorname{ad} B)^{k-1} X$ for $1 \le k \le M$, we obtain

$$\begin{aligned}
&\pi_M(D_X(E_M(B))) \\
&= \pi_M\Big(\Big(X + \frac{(\operatorname{ad} B)X}{2!} + \frac{(\operatorname{ad} B)^2 X}{3!} + \cdots + \frac{(\operatorname{ad} B)^{M-1} X}{M!}\Big) E_M(B)\Big) \\
&= \pi_M\Big(\pi_M\Big(\Big(1 + \frac{(\operatorname{ad} B)}{2!} + \frac{(\operatorname{ad} B)^2}{3!} + \cdots + \frac{(\operatorname{ad} B)^{M-1}}{M!}\Big)(X)\Big)(E_M(B))\Big)
\end{aligned}$$

by (B.28).

From complex-variable theory we have

$$\frac{z}{e^z - 1} = \Big(1 + \frac{z}{2!} + \frac{z^2}{3!} + \cdots\Big)^{-1} = 1 - \frac{z}{2} + \frac{b_1}{2!} z^2 + \frac{b_2}{4!} z^4 + \cdots,$$

where $b_1 = \frac{1}{6}$, $b_2 = -\frac{1}{30}$, $\ldots$ are Bernoulli numbers apart from signs. Remembering that M is odd, we can finally define X:

$$X = \Big(1 - \frac{\operatorname{ad} B}{2} + \frac{b_1}{2!} (\operatorname{ad} B)^2 + \frac{b_2}{4!} (\operatorname{ad} B)^4 + \cdots + \frac{b_{(M-1)/2}}{(M-1)!} (\operatorname{ad} B)^{M-1}\Big) A.$$

The element X is in $T^{\le M}(V)$. Substituting for X in the expression

$$\pi_M\Big(\Big(1 + \frac{(\operatorname{ad} B)}{2!} + \frac{(\operatorname{ad} B)^2}{3!} + \cdots + \frac{(\operatorname{ad} B)^{M-1}}{M!}\Big)(X)\Big)$$

above, we find that

$$\pi_M(D_X(E_M(B))) = \pi_M(A E_M(B)). \tag{B.33}$$

We shall now prove by induction for $m \ge 1$ that

$$(\pi_M D_X \pi_M)^m (E_M(B)) = \pi_M(A^m E_M(B)). \tag{B.34}$$

The result for $m = 1$ is just (B.33). Assuming the result for $m - 1$, we use (B.28), (B.29), and (B.33) repeatedly to write

$$\begin{aligned}
(\pi_M D_X \pi_M)^m (E_M(B)) &= (\pi_M D_X \pi_M)(\pi_M D_X \pi_M)^{m-1}(E_M(B)) \\
&= (\pi_M D_X \pi_M)(\pi_M (A^{m-1} E_M(B))) \\
&= \pi_M D_X \pi_M (A^{m-1} E_M(B)) \\
&= \pi_M D_X (A^{m-1} E_M(B)) \\
&= \pi_M (A^{m-1} D_X(E_M(B))) \quad \text{since } D_X(A) = 0 \\
&= \pi_M (A^{m-1} \pi_M D_X(E_M(B))) \\
&= \pi_M (A^{m-1} \pi_M (A E_M(B))) \\
&= \pi_M (A^m E_M(B)).
\end{aligned}$$

This completes the induction and proves (B.34).

Next we shall prove that

$$\text{(B.35)} \qquad \pi_M D_X \pi_M \quad \text{is nilpotent on} \quad T^{\le M}(V).$$

To do so, we shall exhibit a basis of $T^{\le M}(V)$ with respect to which the matrix of $\pi_M D_X \pi_M$ is strictly lower triangular. The basis begins with

$$1,\ B,\ A,\ B^2,\ BA,\ AB,\ A^2,$$

and it continues with bases of $T^3(V)$, $T^4(V)$, and so on. The basis of $T^m(V)$ begins with B^m, then contains all monomials in A and B with 1 factor A and $m-1$ factors B, then contains all monomials in A and B with 2 factors A and $m-2$ factors B, and so on. Take a member of this basis, say a monomial in $T^m(V)$ with k factors of A and $m-k$ factors of B. When we apply $\pi_M D_X \pi_M$, the right-hand π_M changes nothing, and the D_X acts on the monomial as a derivation. Since $D_X A = 0$, we get $m - k$ terms, each obtained by replacing one instance of B by X. The definition of X shows that X is the sum of A and higher-order terms. When we substitute for X, the A gives us a monomial in $T^m(V)$ with one more A and one less B, and the higher-order terms give us members of $T^{>m}(V)$. Application of the final π_M merely throws away some of the terms. The surviving terms are linear combinations of members of the basis farther along than our initial monomial, and (B.35) follows.

Because of (B.35), we may assume that $(\pi_M D_X \pi_M)^{M'} = 0$, where M' is $\ge M$. Multiplying (B.34) by $1/m!$ and summing up to M', we obtain

$$\text{(B.36)} \qquad \pi_M \exp(\pi_M D_X \pi_M)(E_M(B)) = \pi_M(E_M(A) E_M(B)).$$

Meanwhile if we multiply the formula of Lemma B.32 by $1/k!$ and sum for $0 \le k \le M$, we have

$$\pi_M(E_M(C(z))) = \pi_M \exp(z\pi_M D_X \pi_M)(E_M(B)). \tag{B.37}$$

Put $C = C(1)$. For $z = 1$, equations (B.36) and (B.37) together give

$$\pi_M(E_M(C)) = \pi_M(E_M(A)E_M(B)). \tag{B.38}$$

We can recover C from this formula by using the power series for $\log(1-z)$ in the same way as in the first line of (B.24), and we see from (B.38) that C is the member of $T(V)$ whose expression in terms of brackets we seek.

To obtain a formula for C, we use (B.31) with $z = 1$ to write

$$\begin{aligned}
C &= \pi_M \exp(\pi_M D_X \pi_M)(B)\\
&= \pi_M\Big(1 + (\pi_M D_X \pi_M) + \frac{(\pi_M D_X \pi_M)^2}{2!} + \cdots + \frac{(\pi_M D_X \pi_M)^{M'}}{M'!}\Big)(B)\\
&= B + \Big(1 + \frac{(\pi_M D_X \pi_M)}{2!} + \cdots + \frac{(\pi_M D_X \pi_M)^{M'-1}}{M'!}\Big)(\pi_M D_X \pi_M)(B)\\
&= B + \Big(1 + \frac{(\pi_M D_X \pi_M)}{2!} + \cdots + \frac{(\pi_M D_X \pi_M)^{M'-1}}{M'!}\Big)(X)\\
&= B + \Big(1 + \frac{(\pi_M D_X \pi_M)}{2!} + \cdots + \frac{(\pi_M D_X \pi_M)^{M'-1}}{M'!}\Big)\\
&\qquad \times \Big(1 - \frac{(\mathrm{ad}\, B)}{2} + \frac{b_1}{2!}(\mathrm{ad}\, B)^2 + \cdots + \frac{b_{(M-1)/2}}{(M-1)!}(\mathrm{ad}\, B)^{M-1}\Big)(A)\\
&= A + B + \Big(1 + \frac{(\pi_M D_X \pi_M)}{2!} + \cdots + \frac{(\pi_M D_X \pi_M)^{M'-1}}{M'!}\Big)\\
&\qquad \times \Big(-\frac{(\mathrm{ad}\, B)}{2} + \frac{b_1}{2!}(\mathrm{ad}\, B)^2 + \cdots + \frac{b_{(M-1)/2}}{(M-1)!}(\mathrm{ad}\, B)^{M-1}\Big)(A),
\end{aligned}$$

the last step holding since $D_X(A) = 0$. The right side is the sum of $A + B$, a linear combination of various bracket terms $(\mathrm{ad}\, B)^m(A)$ with $m \ge 1$, and terms $(\pi_M D_X \pi_M)^k((\mathrm{ad}\, B)^m(A))$ with $k \ge 1$ and $m \ge 1$.

To complete the proof, we are to show that each of the terms

$$(\pi_M D_X \pi_M)^k((\mathrm{ad}\, B)^m(A)) \tag{B.39}$$

with $k \ge 1$ and $m \ge 1$ is a linear combination of iterated brackets. It is enough to prove that if

$$(\mathrm{ad}\, X_1)(\mathrm{ad}\, X_2)\cdots(\mathrm{ad}\, X_{n-1})X_n, \quad \text{with each } X_j \text{ equal to } A \text{ or } B, \tag{B.40}$$

is given, then $(\pi_M D_X \pi_M)$ of it is a linear combination of other terms of the same general form as (B.40) with various n's.

Let us prove inductively on k that $\operatorname{ad}((\operatorname{ad} B)^k A)$ is a linear combination of terms

$$(\operatorname{ad} B)^j (\operatorname{ad} A)(\operatorname{ad} B)^{k-j}, \quad 0 \le j \le k. \tag{B.41}$$

This is trivial for $k = 0$. If it is true for $k - 1$, then

$$\begin{aligned}
\operatorname{ad}((\operatorname{ad} B)^k A) &= \operatorname{ad}((\operatorname{ad} B)((\operatorname{ad} B)^{k-1} A)) \\
&= \operatorname{ad}(B((\operatorname{ad} B)^{k-1} A) - ((\operatorname{ad} B)^{k-1} A)B) \\
&= (\operatorname{ad} B)(\operatorname{ad}((\operatorname{ad} B)^{k-1} A)) - (\operatorname{ad}((\operatorname{ad} B)^{k-1} A))(\operatorname{ad} B),
\end{aligned}$$

and substitution of the result for $k - 1$ yields the result for k.

Since X is a linear combination of terms $(\operatorname{ad} B)^k A$, we see from the above conclusion that $\operatorname{ad} X$ is a linear combination of terms (B.41).

Next we observe the formula

$$D_X((\operatorname{ad} u)v) = (\operatorname{ad}(D_X u))v + (\operatorname{ad} u)(D_X v). \tag{B.42}$$

In fact,

$$\begin{aligned}
D_X((\operatorname{ad} u)v) &= D_X(uv - vu) \\
&= (D_X u)v + u(D_X v) - (D_X v)u - v(D_X u) \\
&= (\operatorname{ad}(D_X u))v + (\operatorname{ad} u)(D_X v).
\end{aligned}$$

Now suppose that (B.40) is given. In applying $\pi_M D_X \pi_M$, we may disregard the occurrences of π_M at the ends. Formula (B.42) allows us to compute the effect of D_X on (B.40). We get the sum of n terms. In the first $n-1$ terms the factor $(\operatorname{ad} X_j)$ gets replaced by $(\operatorname{ad} X)$ if $X_j = B$ or by 0 if $X_j = A$; we have seen that $(\operatorname{ad} X)$ is a linear combination of terms (B.41), and thus substitution in these $n - 1$ terms give terms of the same general form as (B.40). In the last term that we obtain by applying D_X to (B.40), the factor X_n gets replaced by X if $X_n = B$ or by 0 if $X_n = A$; since X is a linear combination of terms $(\operatorname{ad} B)^k A$, substitution yields terms of the same general form as (B.40). This proves that application of $(\pi_M D_X \pi_M)$ to (B.40) yields terms of the same general form. The theorem follows.

Using the same notation $V = \mathbb{C}A \oplus \mathbb{C}B$ as in the last part of the proof of Theorem B.22, we can derive an explicit formula for how (B.25) may be expressed as the sum of $A + B$ and explicit iterated brackets. Being an associative algebra, $T(V)$ is also a Lie algebra under the bracket operation $[u, v] = uv - vu$. Let $L(V)$ be the Lie subalgebra of $T(V)$ generated by the elements of V. This consists of linear combinations of iterated brackets of elements of V.

Proposition B.43. The unique linear map $p : T(V) \to T(V)$ such that $p(1) = 0$, $p(v) = v$ for v in V, and

$$p(v_1 \cdots v_n) = n^{-1}(\operatorname{ad} v_1) \cdots (\operatorname{ad} v_{n-1}) v_n$$

whenever $n > 1$ and $v_1, \ldots, v_n$ are all in V has the property of being a projection of $T(V)$ onto $L(V)$.

REMARKS. Since we know from Theorem B.22 that the sum of all terms in (B.25) with a given homogeneity is in $L(V)$, we can apply the map p to such a sum to get an expression in terms of iterated brackets. For example, consider the cubic terms. Many terms, like AB^2 and $(A+B)^3$, map to 0 under p. For the totality of cubic terms,

$$\begin{aligned}
&p(\tfrac{1}{6}(A^3 + 3A^2B + 3AB^2 + B^3) - \tfrac{1}{2}(\tfrac{1}{2}(A^2 + 2AB + B^2)(A+B)) \\
&\quad - \tfrac{1}{2}((A+B)\tfrac{1}{2}(A^2 + 2AB + B^2)) + \tfrac{1}{3}((A+B)^3)) \\
&= \tfrac{1}{3}\{\tfrac{1}{2}(\operatorname{ad} A)^2 B - \tfrac{1}{4}((\operatorname{ad} A)^2 B + 2(\operatorname{ad} A)(\operatorname{ad} B)A + (\operatorname{ad} B)^2 A) \\
&\quad - \tfrac{1}{4}(2(\operatorname{ad} A)^2 B + 2(\operatorname{ad} B)(\operatorname{ad} A)B)\} \\
&= \tfrac{1}{12}((\operatorname{ad} A)^2 B + (\operatorname{ad} B)^2 A).
\end{aligned}$$

PROOF. The map p is unique since the monomials in V generate $T(V)$. For existence, we readily define p on each $T^n(V)$ by means of the universal mapping property of n-fold tensor products. It is clear that p carries $T(V)$ into $L(V)$. To complete the proof, we show that p is the identity on $L(V)$.

Recall that ad has been extended from V to $T(V)$ as a homomorphism, so that $\operatorname{ad}(AB)A$, for example, is $(\operatorname{ad} A)(\operatorname{ad} B)A = 2ABA - A^2B - BA^2$, not $(AB)A - A(AB)$. However, we shall prove that

$$(\operatorname{ad} x)u = xu - ux \quad \text{for } x \in L(V). \tag{B.44}$$

It is enough, for each n, to consider elements x that are n-fold iterated brackets of members of V, and we proceed inductively on n. For degree $n = 1$, (B.44) is the definition. Assuming (B.44) for degree $< n$, we suppose that x and y are iterated brackets of members of V and that the sum of their degrees is n. Then

$$\begin{aligned}
(\operatorname{ad}[x,y])u &= \operatorname{ad}(xy - yx)u \\
&= (\operatorname{ad} x \operatorname{ad} y - \operatorname{ad} y \operatorname{ad} x)u \\
&= x(yu - uy) - (yu - uy)x - y(xu - ux) + (xu - ux)y \\
&= [x,y]u - u[x,y],
\end{aligned}$$

and the induction is complete.

To prove that p is the identity on $L(V)$, we introduce an auxiliary mapping $p^* : T(V) \to L(V)$ defined in the same way as p except that the coefficient n^{-1} is dropped in the definition on $v_1 \cdots v_n$. The map p^* has the property that

$$p^*(uv) = (\operatorname{ad} u)\, p^*(v) \tag{B.45}$$

for all u and v in $T(V)$ as long as v has no constant term. In fact, it is enough to consider the case of monomials, say $u = u_1 \cdots u_m$ and $v = v_1 \cdots v_n$ with $n \geq 1$. Then

$$\begin{aligned} p^*(uv) &= (\operatorname{ad} u_1) \cdots (\operatorname{ad} u_m)(\operatorname{ad} v_1) \cdots (\operatorname{ad} v_{n-1}) v_n \\ &= (\operatorname{ad} u)(\operatorname{ad} v_1) \cdots (\operatorname{ad} v_{n-1}) v_n \\ &= (\operatorname{ad} u)\, p^*(v), \end{aligned}$$

and (B.45) is proved.

Next let us see that

(B.46) $\quad p^*$ restricted to $L(V)$ is a derivation of $L(V)$.

In fact, if x and y are in $L(V)$, (B.44) and (B.45) yield

$$\begin{aligned} p^*[x, y] &= p^*(xy - yx) = (\operatorname{ad} x)\, p^*(y) - (\operatorname{ad} y)\, p^*(x) \\ &= [x, p^*(y)] - [y, p^*(x)] = [x, p^*(y)] + [p^*(x), y], \end{aligned}$$

and (B.46) is proved.

Using (B.46), we prove inductively on the degree of the bracket that if $x \in L(V)$ is an iterated bracket involving n elements of V, then $p^*(x) = nx$. This is true by definition of p^* for $n = 1$. Suppose it is true for all degrees less than n. Let x and y be members of $L(V)$ given as d-fold and $(n - d)$-fold iterated brackets of members of V. Then

$$p^*[x, y] = [x, p^* y] + [p^* x, y] = (n - d)[x, y] + d[x, y] = n[x, y],$$

and the induction goes through. Thus p^* acts on $L(V)$ as asserted, and p acts on $L(V)$ as the identity. Thus p is indeed a projection of $T(V)$ onto $L(V)$.