J. DIFFERENTIAL GEOMETRY
10 (1975) 511-521

# HIGHER ORDER ANALOGUES OF CLASSICAL GROUPS

R. H. BOWMAN & R. G. POND

## 1. Introduction

In [2] and [3] one of the present writers introduced a notion of canonical tangential resolution $({}^kM)$ $(k = 0, 1, 2, \cdots)$ for an arbitrary real $C^\infty$ finite-dimensional manifold M. Subsequently, various aspects of the higher-order terms of such a sequence have been investigated (see [4], [5], and [6]). While the local origins of the theory are to be found in the formalism of extensor analysis (see [7] as a general reference), the categorical context is co-equalization in the general theory of cotriples, the basic cotriple being the zero-section and the tangent functor in the category of $C^\infty$ manifolds (see [9]).

The present paper concerns the resolution $({}^kG)$ of a Lie group $G$ and the resolution $({}^k\phi)$ of a differentiable action $\phi$ of $G$ on a manifold. The principal results are the theorems in § 2 establishing matrix realizations for each ${}^kG$ and its associated Lie algebra $\mathscr{L}({}^kG)$ and interpreting the relevant exponential map in the case where $G$ is a Lie subgroup of some general linear group. The information developed here yields the foundation for a general theory of differentiable fiber bundle resolution and its interpretation, a systematic treatment of which will be given in later papers. The remainder of the introduction explains the notational conventions and special identifications used in the sequel. All manifolds are modeled on real Banach spaces and are at least of class $C^\infty$. The notation is intended to conform as closely as possible with that currently employed in such a context (see [1], [8], and [11]).

Let $M$ be a manifold modeled on the Banach space $B$. An element of the tangent bundle $T(M)$ will be viewed as an equivalence class $[\theta, b]_x$, where $b \in B$, $x \in M$, and $\theta$ is a local coordinate map about $x$. Thus, if $\phi: M \to N$ is a differentiable map, its associated tangent map $T(\phi): T(M) \to T(N)$ is described locally by

$$T(\phi)([\theta, b]_x) = [\psi, D(\psi \circ \phi \circ \theta^{-1})(\theta(x))b]_y \,, \tag{1}$$

where $y = \phi(x)$, $\psi$ is a local coordinate map about $y$, and $D(\psi \circ \phi \circ \theta^{-1})(\theta(x))b$ is the total differential of $\psi \circ \phi \circ \theta^{-1}$ at the point $\theta(x)$ evaluated at the vector $b$. When $V$ is an open set in a Banach space $C$, $T(V)$ will be viewed as the direct product $V \times C$ with $(v; c)$ denoting a tangent vector $c \in C$ located at the point

Communicated by K. Yano, February 4, 1974.

$v \in V$. In particular, given $\theta$ as above, $T(\theta)([\theta, b]_x) = (\theta(x); b)$, i.e., $T(\theta)$ is the usual lift of $\theta$ to a local coordinate map on $T(M)$. If $M$ is a direct product $K \times L$, $T(M)$ will be treated as $T(K) \times T(L)$. In particular, if $\phi: M \to N$ is a differentiable map, let ${}_u\phi: L \to N$ and $\phi_v: K \to N$ be given by

$$ {}_u\phi(v) = \phi(u, v) = \phi_v(u) \qquad \text{for each } (u, v) \in M . $$

In this case formula (1) becomes

$$ \begin{aligned} T(\phi)([\tau, a]_u, [\eta, d]_v) &= T(\phi_v)([\tau, a]_u) + T({}_u\phi)([\eta, d]_v) \\ &= [\psi, D(\psi \circ \phi_v \circ \tau^{-1})(\tau(u))a + D(\psi \circ {}_u\phi \circ \eta^{-1})(\eta(v))d]_y , \end{aligned} \tag{2} $$

where $[\tau, a]_u \in T(K)$ and $[\eta, d]_v \in T(L)$.

Following [2], the tangential resolution $({}^kM)$ for a manifold $M$ can be specified inductively. Let ${}^0M = M$, ${}^1M = T(M)$, ${}_0I =$ the identity map ${}^1M \to T({}^0M)$. For $k \geq 0$, suppose ${}^kM$ and ${}^{k+1}M$ have been defined together with an embedding ${}_kI: {}^{k+1}M \to T({}^kM)$. Let $\pi^k: T({}^kM) \to {}^kM$ and $\pi^{k+1}: T({}^{k+1}M) \to {}^{k+1}M$ denote the usual projections and let ${}_k\pi = \pi^k \circ {}_kI$. Then ${}^{k+2}M$ is defined as the set of all points in $T({}^{k+1}M)$ where ${}_kI \circ \pi^{k+1}$ and $T({}_k\pi)$ agree and ${}_{k+1}I: {}^{k+2}M \to T({}^{k+1}M)$ denotes the inclusion map. With tangential resolutions of manifolds thus defined, let $\phi: M \to N$ be any differentiable map. The tangential resolution $({}^k\phi)$ for $\phi$ can be specified as follows: let ${}^0\phi = \phi$ and, assuming ${}^k\phi: {}^kM \to {}^kN$ has been defined, let $\phi^{k+1}: {}^{k+1}M \to {}^{k+1}N$ be ${}_kJ^{-1} \circ T({}^k\phi) \circ {}_kI$ where ${}_kI$ and ${}_kJ$ are the embeddings of ${}^{k+1}M$ and ${}^{k+1}N$ in $T({}^kM)$ and $T({}^kN)$, respectively. It is readily seen that the resolution process for manifolds is functorial at each level $k$.

When $V$ is an open set in a Banach space $C$, ${}^{k+1}V$ will be treated as the direct product $V \times C^{k+1}$ with the embedding map ${}^{k+1}V \to T({}^kV) = (V \times C^k) \times C^{k+1}$ given by sending $(v_0, \cdots, v_{k+1})$ to $(v_0, \cdots, v_k; v_1, \cdots, v_{k+1})$. Local coordinatization of the point-set ${}^{k+1}M$ is achieved through functions of the form ${}^{k+1}\theta$ where $\theta$ is a local coordinate map for the manifold $M$. If $M$ is a direct product $K \times L$, ${}^{k+1}M$ will be treated as ${}^{k+1}K \times {}^{k+1}L$. In particular, if $\phi: M \to N$ is any differentiable map, ${}^{k+1}\phi: {}^{k+1}K \times {}^{k+1}L \to {}^{k+1}N$ is given locally by

$$ {}^{k+1}\phi(x, z) = {}_kJ^{-1}([{}^k\psi, D({}^k\psi \circ {}^k\phi_v \circ {}^k\tau^{-1})(s)t + D({}^k\psi \circ {}_u{}^k\phi \circ {}^k\eta^{-1})(q)r]_y) , \tag{3} $$

where $s = (s_0, \cdots, s_k)$, $t = (s_1, \cdots, s_{k+1})$, $x = {}^{k+1}\tau^{-1}(s_0, \cdots, s_{k+1})$, $u = {}^k\tau^{-1}(s)$, $q = (q_0, \cdots, q_k)$, $r = (q_1, \cdots, q_{k+1})$, $z = {}^{k+1}\eta^{-1}(q_0, \cdots, q_{k+1})$, $v = {}^k\eta^{-1}(q)$, and $y = {}^k\phi(x, z)$.

Before passing to Lie considerations, an alternate description of the sequence $({}^kM)$ can be given which suggests the rationale for its consideration. Let $T^0(M) = M$, and for each integer $k \geq 0$ let $\pi_k = T^0(\pi_k): T^{k+1}(M) \to T^k(M)$ denote (inductively) the tangent bundle over $T^k(M)$. For each $k \geq 1$ and each $m$ from 1 through $k$, let $T^m(\pi_{k-m}): T^{k+1}(M) \to T^k(M)$ denote (inductively) the tangent map associated with $T^{m-1}(\pi_{k-m})$. For each $k \geq 0$ the various $T^m(\pi_{k-m})$ $(0 \leq m \leq k)$ are all distinct, and ${}^{k+1}M$ can be viewed as precisely that subset of $T^{k+1}(M)$ on which all the $T^m(\pi_{k-m})$ coincide.

Now suppose $\sigma = \sigma^{(0)} : I \to M$ is any differentiable curve in $M$. Let $\sigma^{(k+1)} : I \to T^{k+1}(M)$ denote (inductively) the standard lift of $\sigma^{(k)}$ over $\pi_k$, $k \geq 0$. One readily checks that $\sigma^{(k+1)}$ actually has its range in the (losed embedded) submanifold ${}^{k+1}M$. Thus higher order ordinary differential equations over $M$ are properly formulated as given data relative to the ${}^kM$ rather than the ambient $T^k(M)$. Recognition of this phenomenon is tacit in the standard treatment of sprays, for instance, at the second tangential level (see [11] or [13]). A similar situation occurs in other higher order differential contexts as well.

Let $G$ be a (Banach modeled) Lie group with multiplication $\mu$ and inversion $\iota$. With the conventions on products in mind, one checks that each ${}^kG$ is again a Lie group with multiplication ${}^k\mu$ and inversion ${}^k\iota$. Moreover all the global maps involved in the construction of ${}^kG$ as a manifold are Lie group homomorphisms. If $\phi : G \times N \to N$ is a differentiable (left) action of $G$ on $N$, then each ${}^k\phi : {}^kG \times {}^kN \to {}^kN$ is a differentiable action of ${}^kG$ on ${}^kN$, and the sequence $({}^k\phi)$ is called the tangential resolution of $\phi$. In particular, if $N$ is a Banach space and $\phi$ preserves the linear structure on $N$, then ${}^k\phi$ preserves the linear structure on ${}^kN$ (viewed as $N^{k+1}$).

We close this introductory section with a decomposition theorem for ${}^kG$. Let $K_0$ be the trivial subgroup of $G$ and let $G_0 = G$. For $k \geq 1$ let $K_k = \ker ({}_0\pi \circ \cdots \circ {}_{k-1}\pi)$ and let $G_k$ be the subgroup of ${}^kG$ consisting of those $x$ for which ${}_{k-1}I(x)$ is the zero tangent vector at ${}_{k-1}\pi(x)$.

**Theorem 1.** *For each $k \geq 0$ the restriction of ${}_k\pi$ to $G_{k+1}$ is a Lie isomorphism onto $G_k$ and ${}_k\pi(K_{k+1}) = K_k$. In particular, each ${}^kG$ is the internal semidirect product $K_kG_k$.*

*Proof.* Letting $O_k$ denote the restriction of the zero section of $\pi^k : T({}^kG) \to {}^kG$ to $G_k$, one checks (inductively) that ${}_kI^{-1} \circ O_k$ is the inverse of the restriction of ${}_k\pi$ to $G_{k+1}$. Clearly ${}_k\pi(K_{k+1}) \subseteq K_k$. Thus (inductively again) each $K_k \cap G_k$ is trivial. To see that $K_kG_k$ is all of ${}^kG$ for $k \geq 1$, let $x \in {}^kG$ be arbitrary and let $g^{-1} = {}_0\pi \circ \cdots \circ {}_{k-1}\pi(x) \in G$. Letting $y$ be the element of $G_k$ for which ${}_0\pi \circ \cdots \circ {}_{k-1}\pi(y) = g^{-1}$ and letting $\mu$ denote multiplication in $G$, one has $x = {}^k\mu({}^k(\mu_g)(x), y)$ with ${}^k(\mu_g)(x)$ necessarily in $K_k$.

**Corollary.** *Let $\mathscr{L}({}^kG)$ denote the Lie algebra of ${}^kG$. Then $\mathscr{L}({}^kG)$ is the internal semidirect product $\mathscr{L}(K_k) \oplus \mathscr{L}(G_k)$ of the ideal $\mathscr{L}(K_k)$ with the subalgebra $\mathscr{L}(G_k)$.*

## 2. A matrix realization of ${}^kG$

Let $B$ be a Banach space, $GL(B)$ the Lie group of all continuous automorphisms of $B$, and $gl(B)$ the Banach algebra (and Lie algebra) of all continuous linear endomorphisms of $B$. For $k \geq 0$ each $S \in gl(B^{k+1})$ will be treated as a

$(k+1) \times (k+1)$ matrix $[S_i{}^j]$ $(0 \leq i, j \leq k)$ with lower index = row index. (That is, given $S$, the various $S_i{}^j \in gl(B)$ are obtained by requiring $S(v) = (\Sigma_j S_0{}^j(v_j), \cdots, \Sigma_j S_k{}^j(v_j))$ for all $v = (v_0, \cdots, v_k) \in B^{k+1}$.) For the remainder of the paper $G$ will be a Lie subgroup of $GL(B)$, and $\phi\colon G \times B \to B$ will denote the usual left action of $G$ on $B$. Thus the map $x \mapsto {}_x{}^k\phi$ amounts to a realization of ${}^kG$ as a Lie subgroup of $GL(B^{k+1})$. Our purpose in this section is to characterize the matrices $[{}_x{}^k\phi_i{}^j]$ which arise in this realization. In the process we obtain a characterization of the matrices which arise in the corresponding realization of $\mathscr{L}({}^kG)$ in $gl(B^{k+1})$ as well as an interpretation of ${}^k\exp\colon {}^k\mathscr{L}(G) \to {}^kG$ as an ordinary exponential map, where $\exp\colon \mathscr{L}(G) \to G$ is the restriction to $\mathscr{L}(G) \subseteq gl(B)$ of the usual exponential map.

**Theorem 2.** *Let $C(i,j) = i!/[(i-j)!j!]$ when $0 \leq j \leq i \leq k$. Then ${}_x{}^k\phi_i{}^j = C(i,j)_x{}^k\phi_{i-j}{}^0$ when $0 \leq j \leq i \leq k$, and ${}_x{}^k\phi_i{}^j = 0$ when $0 \leq i \leq j \leq k$.*

*Proof.* It is sufficient to consider the case $G = GL(B)$. We argue by induction on $k$, the case $k = 0$ being trivial. With $K = G$, $L = N = B$, $\tau =$ the standard injection $G \to gl(B)$, $\psi = \eta =$ the identity on $B$, and the convention that ${}_kJ(b_0, \cdots, b_{k+1}) = (b_0, \cdots, b_k; b_1, \cdots, b_{k+1})$ and in mind, formula (3) amounts to

$$(4) \qquad {}^{k+1}\phi(x,z) = {}_kJ^{-1}({}^k\phi(u,v); D({}^k\phi_v \circ {}^k\tau^{-1})(s)t + D({}_u{}^k\phi)(q)r) .$$

Now $D({}_u{}^k\phi)(q)r = {}_u{}^k\phi(r)$, since ${}_u{}^k\phi$ is a continuous linear map. Thus, by the inductive hypothesis for $k$, one has

$$(5) \qquad \begin{aligned} D({}_u{}^k\phi)(q)r = (C(0,0)s_0(q_1), C(1,0)s_1(q_1) + C(1,1)s_0(q_2), \\ \cdots, \Sigma_j C(k,j)s_{k-j}(q_{j+1})) . \end{aligned}$$

Define a continuous linear map $F\colon gl(B)^{k+1} \to B^{k+1}$ by

$$\begin{aligned} F(E_0, \cdots, E_k) = (C(0,0)E_0(q_0), C(1,0)E_1(q_1) + C(1,1)E_0(q_1), \\ \cdots, \Sigma_j C(k,j)E_{k-j}(q_j)) . \end{aligned}$$

By the induction assumption, the restriction of $F$ to $G \times gl(B)^k$ is precisely ${}^k\phi_v \circ {}^k\tau^{-1}$. Thus one has

$$(6) \qquad D({}^k\phi_v \circ {}^k\tau^{-1})(s)t = F(s_1, \cdots, s_{k+1}) .$$

Since $C(i,j) + C(i,j+1) = C(i+1,j+1)$, formulas (5) and (6) yield

$$(7) \qquad \begin{aligned} ({}^k\phi(u,v); D({}^k\phi_v \circ {}^k\tau^{-1})(s)t + D({}_u{}^k\phi)(q)r) = (F(s_0, \cdots, s_k); C(1,0)s_1(q_0) \\ + C(1,1)s_0(q_1), \cdots, \Sigma_j C(k+1,j)s_{k+1-j}(q_j)) . \end{aligned}$$

Apply ${}_kJ^{-1}$ to both sides of (7) to obtain (from (4))

$$ {}^{k+1}\phi(x, z) = (C(0,0)s_0(q_0), C(1,0)s_1(q_0) + C(1,1)s_0(q_1), \cdots, \Sigma_j C(k+1, j)s_{k+1-j}(q_j)) \ , \tag{8} $$

which completes the induction.

**Theorem 3.** *Let $k \geq 1$. If $u \in G_k$ and $0 < i \leq k$, then ${}_u{}^k\phi_i{}^0 = 0$. Moreover, for each $s_0 \in G$ there is exactly one $u \in G_k$ with ${}_u{}^k\phi_0{}^0 = s_0$. For any $u \in K_k$, ${}_u{}^k\phi_0{}^0$ is the identity element $I \in G$.*

*Proof.* For any $u \in {}^kG$, ${}_u{}^k\phi_0{}^0 = {}_0\pi \circ \cdots \circ {}_{k-1}\pi(u)$. So the assertion for $u \in K_u$ is immediate from the definition of $K_k$, while the assertion about ${}_u{}^k\phi_0{}^0$ for $u \in G_k$ follows because (by Theorem 1) ${}_0\pi \circ \cdots \circ {}_{k-1}\pi$ carries $G_k$ isomorphically onto $G$. For $i > 0$ and $u = (s_0, \cdots, s_k)$ in $GL(B) \times gl(B)^k = {}^kGL(B)$, $u \in G_k$ implies $s_1 = \cdots = s_k = 0$. Thus ${}_u{}^k\phi_i{}^0 = 0$ follows from Theorem 2 and more precisely from the formula

$$ {}_u{}^k\phi(r) = (C(0,0)s_0(q_1), C(1,0)s_1(q_1) + C(1,1)s_0(q_2), \cdots, \Sigma_j C(k, j)s_{k-j}(q_{j+1})) \ . \tag{9} $$

Now let $\phi_k$ denote the realization of ${}^kG$ in $GL(B^{k+1})$, i.e., let $\phi_k(x) = [{}_x{}^k\phi_i{}^j]$, and let $\mathscr{L}(\phi_k)\colon \mathscr{L}({}^kG) \to gl(B^{k+1})$ denote the corresponding realization of the Lie algebra $\mathscr{L}({}^kG)$.

**Theorem 4.** *The image of $\mathscr{L}(\phi_k)$ consists of all matrices $[A_i{}^j]$ with each $A_i{}^j \in \mathscr{L}(G) \subseteq gl(B)$ satisfying $A_i{}^j = 0$ for $i < j$ and $A_i{}^j = C(i, j)A_{i-j}{}^0$ for $j \leq i$. Such a matrix corresponds to an element of $\mathscr{L}(K_k)$ iff $A_0{}^0 = 0$. Such a matrix corresponds to an element of $\mathscr{L}(G_k)$ iff $A_i{}^0 = 0$ for all $i > 0$.*

*Proof.* Let $U$ and $V$ be open neighborhoods of 0 and $I$, respectively, in $\mathscr{L}(G)$ and $G$ such that the exponential map $\exp\colon U \to V$ is a diffeomorphism. Then ${}^k\exp\colon {}^kU = U \times \mathscr{L}(G)^k \to {}^kV \subseteq {}^kG$ is also a diffeomorphism for all $k$. Viewing ${}^kG \subseteq {}^kGL(B) = GL(B) \times gl(B)^k$ one has $\phi_k(A_0, \cdots, A_k) = [A_i{}^j]$, where $A_i{}^0 = A_i$. Let $f\colon U \to GL(B)$ be given by $f(r_0) = \exp(r_0)$.

Simply because $f$ is a differentiable map from an open set in a Banach space to an open set in another, one readily checks that ${}^kf\colon {}^kU = U \times \mathscr{L}(G)^k \to {}^kGL(B) = GL(B) \times gl(B)^k$ is given by

$$ {}^kf(r_0, \cdots, r_k) = (f_0(r_0), f_1(r_0, r_1), \cdots, f_k(r_0, \cdots, r_k)) \ , \tag{10} $$

where $f_0 = f$ and $f_i(r_0, \cdots, r_i) = D(f_{i-1})(r_0, \cdots, r_{i-1})(r_1, \cdots, r_i)$ for all $i > 0$. Now the range of ${}^kf$ is actually ${}^kV$. The Lie algebra determinations are made by passing curves $\nu = (\nu_0, \cdots, \nu_k)$ through the origin in ${}^kU$ and differentiating $\phi_k \circ {}^kf \circ \nu$ at $O \in R$. Observe that, for all $i = O, \cdots, k$, one has

$$ D(f_i)(O, \cdots, O)(A_0, \cdots, A_i) = A_i \ . \tag{11} $$

(Actually, as one checks inductively $f_i(O, \cdots, O, A_j, O, \cdots, O) = O$ for $j < i$ while $f_i(O, \cdots, O, A_i) = A_i$. In particular, $D_{j+1}(f_i)(O, \cdots, O)A_j = 0$ for $j < i$

while $D_{i+1}(f_i)(\mathrm{O}, \cdots, \mathrm{O})A_i = A_i$. Thus $D(f_i)(\mathrm{O}, \cdots, \mathrm{O})(A_0, \cdots, A_i) = \Sigma_j D_{j+1}(f_i)(\mathrm{O}, \cdots, \mathrm{O})A_j = A_i$.) On the other hand, since $\phi_k$ is the restriction of a continuous linear map $gl(B)^{k+1} \to gl(B^{k+1})$, one always has

$$D(\phi_k)(X_0, \cdots, X_k) = \phi_k \ . \tag{12}$$

So, for any curve $\nu$ through the origin in ${}^kU$, the chain rule and (10), (11), and (12) yield

$$(\phi_k \circ {}^kf \circ \nu)'(\mathrm{O}) = \phi_k(\nu_0'(\mathrm{O}), \cdots, \nu_k'(\mathrm{O})) \ . \tag{13}$$

The matrix indicated in (13) is clearly of the required general type, and any such matrix $[A_i{}^j]$ can be obtained by letting $\nu = (\nu_0, \cdots, \nu_k)$ where the curves $\nu_i$ with values in $\mathscr{L}(G)$ are chosen such that $\nu_i'(\mathrm{O}) = A_i{}^0$ holds for each $i$. By Theorem 3, $\phi_k \circ {}^kf \circ \nu$ takes values strictly in $\phi_k({}^kV \cap G_k)$ provided $\nu_i = 0$ for all $i > 0$, while $\phi_k \circ {}^kf \circ \nu$ takes values strictly in $\phi_k({}^kV \cap K_k)$ provided $\nu_0 = 0$. This accounts for the splitting of the images of $\mathscr{L}(G_k)$ and $\mathscr{L}(K_k)$.

The task remains to fully describe the entries $x^k\phi_i{}^0$ $(i > 0)$ which can arise in $\phi_k(x)$ for $x \in K_k$. Letting $\exp_k : gl(B^{k+1}) \to GL(B^{k+1})$ denote the ordinary exponential map, one knows that its restriction to $\mathscr{L}(\phi_k)(\mathscr{L}(K_k))$ is just the exponential map over $\phi_k(K_k)$. We shall establish that this restriction is actually a diffeomorphism. This completes the task for, in view of Theorem 4, the restriction of $\exp_k$ to $\mathscr{L}(\phi_k)(\mathscr{L}(K_k))$ is quite easy to compute and yields a sharp description of the matrices ${}_x{}^k\phi_i{}^0$.

**Theorem 5.** *Treat* ${}^kgl(B) = gl(B)^{k+1}$ *as the tangent space at the identity in* ${}^kGL(B) = GL(B) \times gl(B)^k$ *via* $(A_0, \cdots, A_k) \to (I, \mathrm{O}, \cdots, ; A_0, \cdots, A_k)$. *Then the exponential map over* ${}^kGL(B)$ *is just* ${}^k\exp : {}^kgl(B) \to {}^kGL(B)$, *where* $\exp : gl(B) \to GL(B)$ *is the usual exponential map.*

*Proof.* The proof reduces inductively to the following result.

**Lemma.** *Let $V$ be any Banach space. Then* ${}^1\exp : {}^1gl(V) = gl(V)^2 \to {}^1GL(V) = GL(V) \times gl(V)$ *is the exponential map over* ${}^1GL(V)$. *Thus, for any Lie subgroup $H$ in* $GL(V)$, ${}^1(\exp|_{\mathscr{L}(H)}) : {}^1\mathscr{L}(H) \to {}^1H$ *is the exponential map over* ${}^1H$, *where* ${}^1\mathscr{L}(H)$ *is identified with its image in* ${}^1gl(V) = gl(V)^2$.

Indeed, the lemma handles the case $k = 1$ in the theorem. Moreover, assuming the conclusion of the theorem holds for $k$ arbitrary, the inductive step to $k + 1$ is accomplished by letting $H = {}^kGL(B) \subseteq gl(V)$ in the lemma with $V = B^{k+1}$.

*Proof of lemma.* Only the first assertion requires proof. For convenience, transfer ${}^1\exp : {}^1gl(V) = gl(V)^2 \to {}^1GL(V) = GL(V) \times gl(V)$ to the matrix map ${}^1\mathrm{Exp} : \mathscr{L}(\phi_1)(\mathscr{L}({}^1GL(V))) \to \phi_1({}^1GL(V))$ defined so as to satisfy $\phi_1 \circ {}^1\exp = {}^1\mathrm{Exp} \circ \mathscr{L}(\phi_1)$. Thus ${}^1\mathrm{Exp}$ is given by

$${}^1\mathrm{Exp}\left(\begin{bmatrix} X & 0 \\ Y & X \end{bmatrix}\right) = \begin{bmatrix} \exp(X) & 0 \\ D(\exp)(X)(Y) & \exp(X) \end{bmatrix} . \tag{14}$$

To see that ${}^1\mathrm{Exp}$ is the required exponential map, i.e., that the restriction of the standard exponential map $gl(V^2) \to GL(V^2)$, it suffices to check that each curve $\nu(t) = {}^1\mathrm{Exp}\left(\begin{bmatrix} tX & 0 \\ tY & tX \end{bmatrix}\right)$ satisfies the differential equation

$$\nu'(t) = \nu(t) \circ \nu'(\mathrm{O}) . \tag{15}$$

Now the curve $t \mapsto \exp(tX)$ satisfies the differential equation

$$(d/dt)(\exp(tX)) = \exp(tX) \circ (d/dt)|_{t=0}(\exp(tX)) = \exp(tX) \circ X . \tag{16}$$

Also by the chain rule one has

$$(d/dt)(D(\exp)(tX)(tY)) = D^2(\exp)(tX)(tY, X) + D(\exp)(tX)(Y) . \tag{17}$$

Thus

$$\nu'(t) = \begin{bmatrix} \exp(tX) \circ X & 0 \\ D^2(\exp)(tX)(tY, X) + D(\exp)(tX)(Y) & \exp(tX) \circ X \end{bmatrix} . \tag{18}$$

In particular, one has

$$\nu'(\mathrm{O}) = \begin{bmatrix} X & 0 \\ Y & X \end{bmatrix} \tag{19}$$

and therefore

$$\nu(t) \circ \nu'(\mathrm{O}) = \begin{bmatrix} \exp(tX) \circ X & 0 \\ D(\exp)(tX)(tY) \circ X + \exp(tX) \circ Y & \exp(tX) \circ X \end{bmatrix} . \tag{20}$$

Thus comparing (15), (18) and (20) we must show

$$\begin{aligned} &D^2(\exp)(tX)(tY, X) + D(\exp(tX)(Y) \\ &\qquad = D(\exp)(tX)(tY) \circ X + \exp(tX) \circ Y . \end{aligned} \tag{21}$$

To verify (21) we resort to the classical formula

$$D(\exp)(A)(B) = \exp(A) \circ \textstyle\sum_{j\geq 0}((-\mathrm{ad}(A))^j(B))/(j+1)! . \tag{22}$$

(Formula (22) is well known in the finite-dimensional case; see [12, p. 95] for instance. It is probably also standard in the general Banach setting; see [10, p. 89] for an indirect reference. At any rate, the formula can be checked directly.) By (22) one has

$$\begin{aligned} D(\exp)(tX)(Y) &= \exp(tX) \circ (Y + \textstyle\sum_{j\geq 1}((-\mathrm{ad}(tX))^j(Y))/(j+1)!)) \\ &= \exp(tX) \circ Y + \exp(tX) \circ (\textstyle\sum_{j\geq 1}((-\mathrm{ad}(tX))^j(Y))/(j+1)!) . \end{aligned} \tag{23}$$

Also by (22) one has

$$D(\exp)(tX)(tY) \circ X = (\exp(tX) \circ (\textstyle\sum_{j\geq 0} ((-\mathrm{ad}(tX))^j(tY))/(j+1)!)) \circ X . \tag{24}$$

Comparing (21), (23) and (24) we must show

$$D^2(\exp)(tX)(tY, X) = \exp(tX) \circ \sum_{j\geq 0} \left( \frac{(-\mathrm{ad}(tX))^j(tY)}{(j+1)!} \circ X - \sum_{j\geq 1} \frac{(-\mathrm{ad}(tX))^j(Y)}{(j+1)!} \right) . \tag{25}$$

Let $\rho = (\rho_1, \rho_2) : gl(V) \to \mathrm{End}_R(gl(V))^2$ be given by $\rho_1(A) = {}_{\exp(A)}\mu =$ left-multiplication by $\exp(A)$ and $\rho_2(A) = \sum_{j\geq 0}(-\mathrm{ad}(A))^j/(j+1)!$. Let $S: \mathrm{End}_R(gl(V))^2 \to \mathrm{End}_R(gl(V))$ be given by $S(L_1, L_2) = L_1 \circ L_2$. Then from (22) it follows that

$$D(\exp) = S \circ \rho . \tag{26}$$

Thus by the chain rule and bilinearity of $S$ one has

$$D^2(\exp)(A)(B, C) = (D(\rho_1)(A)(B))(\rho_2(A)(C)) + (\rho_2(A))((D(\rho_2)(A)(B))(C)) . \tag{27}$$

Now $D(\rho_1)(A)(B) = {}_{D(\exp)(A)(B)}\mu$, while term-by-term differentiation and the chain rule yield

$$D(\rho_2)(A)(B) = \sum_{j\geq 1} \frac{(-1)^j}{(j+1)!}(\mathrm{ad}(B) \circ (\mathrm{ad}(A))^{j-1} + \cdots + (\mathrm{ad}(A))^{j-1} \circ \mathrm{ad}(B)) ,$$

there being $j$ terms in each internal sum on the right corresponding to the various possible placements of $\mathrm{ad}(B)$. So (27) can be rewritten

$$\begin{aligned} D^2(\exp)(A)(B, C) &= D(\exp)(A)(B) \circ \sum_{j\geq 0} (-)^j \frac{(\mathrm{ad}(A))^j(C)}{(j+1)!} \\ &+ \exp(A) \circ \Big( \sum_{j\geq 1} \frac{(-1)^j}{(j+1)!} (\mathrm{ad}(B) \circ (\mathrm{ad}(A))^{j-1} + \cdots \\ &+ (\mathrm{ad}(A))^{j-1} \circ \mathrm{ad}(B))(C) \Big) . \end{aligned} \tag{28}$$

In particular, one has

$$
\begin{aligned}
D^2(\exp)(tX)(tY, X) &= D^2(\exp)(tX)(X, tY) \\
(29)\qquad &= \exp(tX) \circ \sum_{j\geq 0} \frac{(-\operatorname{ad}(tX))^j(tY)}{(j+1)!} \circ X \\
&\quad + \left[X, \sum_{j\geq 0} \frac{(-1)^j}{(j+1)!} (\operatorname{ad}(tX))^j(tY)\right] \\
&\quad + \sum_{j\geq 1} \frac{(-1)^j}{(j+1)!} (\operatorname{ad}(X) \circ (\operatorname{ad}(tX))^{j-1} + \cdots \\
&\qquad + (\operatorname{ad}(tX))^{j-1} \circ \operatorname{ad}(X))(tY)) ,
\end{aligned}
$$

where [ , ] denotes the usual Lie bracket. But (29) is just (25) as required since

$$
\begin{aligned}
&\left[X, \sum_{j\geq 0} \frac{(-1)^j}{(j+1)!} (\operatorname{ad}(tX))^j(tY)\right] = \left[X, \sum_{j\geq 1} \frac{(-1)^{j+1}}{j!} (\operatorname{ad}(tX))^{j-1}(tY)\right] \\
&= \operatorname{ad}(tX)\left(\textstyle\sum_{j\geq 1}((-1)^{j+1}j!)(\operatorname{ad}(tX))^{j-1}(Y)\right) \\
&= \sum_{j\geq 1} \frac{(-1)^{j+1}}{j!} (\operatorname{ad}(tX))^j(Y) = -\sum_{j\geq 1} \frac{(-1)^j}{j!} (\operatorname{ad}(tX))^j(Y)) \\
&= -\sum_{j\geq 1} \frac{(-1)^j}{(j+1)!} (\operatorname{ad}(tX))^j(Y) + \sum_{j\geq 1} \frac{(-1)^j}{(j+1)!} j (\operatorname{ad}(tX))^j(Y)) \\
&= -\left(\textstyle\sum_{j\geq 1}((-1)^j/(j+1)!)\right)(\operatorname{ad}(tX))^j(Y) \\
&\quad + \sum_{j\geq 1} \frac{(-1)^j}{(j+1)!} (\operatorname{ad}(X) \circ (\operatorname{ad}(tX))^{j-1} + \cdots \\
&\qquad + (\operatorname{ad}(tX))^{j-1} \circ \operatorname{ad}(X))(tY)) .
\end{aligned}
$$

So the proof of the lemma is complete.

**Theorem 6.** *Again let $G$ be an arbitrary Lie subgroup of $GL(B)$, let* exp *denote the restriction of the usual exponential map to $\mathscr{L}(G) \subseteq gl(B)$, and let* $\exp_k: gl(B^{k+1}) \to GL(B^{k+1})$ *denote the usual exponential map. Then the restriction of* $\exp_k$ *to $\mathscr{L}(\phi_k)(\mathscr{L}(K_k))$ is a diffeomorphism onto $\phi_k(K_k)$.*

*Proof.* Choose open sets $U$ and $V$ about the origin in $\mathscr{L}(G)$ and the identity $I$ in $G$, respectively, such that the restriction of exp to $U$ is a diffeomorphism onto $V$. Then ${}^k\exp$ is a diffeomorphism of ${}^kU = U \times \mathscr{L}(G)^k$ onto ${}^kV$. But $K_k \subseteq {}^kV$, whence $\mathscr{L}(K_k) \subseteq {}^kU$ because $\mathscr{L}(\phi_k)(\mathscr{L}(K_k)) \subseteq \mathscr{L}(\phi_k)({}^kU)$. Thus ${}^k\exp$ restricts to a diffeomorphism from $\mathscr{L}(K_k)$ onto $K_k$. Now by Theorem 5 and the naturality of exponential maps one knows $\phi_k \circ {}^k\exp = \exp_k \circ \mathscr{L}(\phi_k)$. Thus the proof is complete.

**Example 1.** Let $G$ be an open subgroup of $GL(B)$. Using the functions $f_0, \cdots, f_k$ introduced in the proof of Theorem 4, one can sharpen the results of Theorem 2 directly to conclude that $\phi_k({}^kG)$ consist of all matrices $[H_i{}^j]$ such that $H_0{}^0 \in G$, $H_i{}^0 \in \mathscr{L}(G)$ for $i > 0$, $H_i{}^j = 0$ for $i < j$, and $H_i{}^j = C(i, j)H_{i-j}{}^0$ for $j \leq i$. Still, $G$ serves to illustrate the rest of the machinery developed. Let

$k = 2$ for instance. By Theorem 3 one knows $\phi_2(G_2) = \left\{\begin{bmatrix} g & 0 & 0 \\ 0 & g & 0 \\ 0 & 0 & g \end{bmatrix} : g \in G\right\}$.

From Theorem 4 it follows $\mathscr{L}(\phi_2)(\mathscr{L}(K_2)) = \left\{\begin{bmatrix} 0 & 0 & 0 \\ E & 0 & 0 \\ F & 2E & 0 \end{bmatrix} : E, F \in \mathscr{L}(G)\right\}$. By Theorem 6, $\phi_2(K_2) = \exp_2 (\mathscr{L}(\phi_2)(\mathscr{L}(K_2)))$. Now

$$\exp_2 \begin{bmatrix} 0 & 0 & 0 \\ E & 0 & 0 \\ F & 2E & 0 \end{bmatrix} = \sum_{j \geq 0} \frac{1}{j!} \begin{bmatrix} 0 & 0 & 0 \\ E & 0 & 0 \\ F & 2E & 0 \end{bmatrix}^j = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ E & 0 & 0 \\ F & 2E & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ E^2 & 0 & 0 \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ E & I & 0 \\ F + E^2 & 2E & I \end{bmatrix},$$

since $j \geq 3$ implies $\begin{bmatrix} 0 & 0 & 0 \\ E & 0 & 0 \\ F & 2E & 0 \end{bmatrix}^j = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. Thus

$$\phi_2(K_2) = \left\{\begin{bmatrix} I & 0 & 0 \\ E & I & 0 \\ F + E^2 & 2E & I \end{bmatrix} : E, F \in \mathscr{L}(G)\right\}.$$

By Theorem 1,

$$\phi_2(^2G) = \left\{\begin{bmatrix} g & 0 & 0 \\ E \circ g & g & 0 \\ (F + E^2) \circ g & 2E \circ g & g \end{bmatrix} : E, F \in \mathscr{L}(G), g \in G\right\}.$$

**Example 2.** If $G$ is not an open subgroup of $GL(B)$, the subdiagonals may contain entries in neither $G$ nor $\mathscr{L}(G)$.

Let $G$ be the group of proper orthogonal transformations in the plane. ($G = 0^+(2) = SO(2) = T^1 = S^1 = \cdots$.) Graphically, one can interpret $G$ as the unit circle and $^1G = T(G)$ as the family of all tangent lines at points on the unit circle. Now $\mathscr{L}(G)$ consists of all skew-symmetric matrices $A = \begin{bmatrix} 0 & -a \\ a & 0 \end{bmatrix}$ where $a \in R$. Of course

$$\phi_1(G_1) = \left\{\begin{bmatrix} B & 0 \\ 0 & B \end{bmatrix} : B = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}, \theta \in [0, 2\pi)\right\}.$$

Since $\mathscr{L}(\phi_1)(\mathscr{L}(K_1)) = \left\{\begin{bmatrix} 0 & 0 \\ A & 0 \end{bmatrix} : A \in \mathscr{L}(G)\right\}$,

$$\phi_1(K_1) = \exp_1 (\mathscr{L}(\phi_1)(\mathscr{L}K_1))) = \left\{\begin{bmatrix} A & 0 \\ I & I \end{bmatrix} : A \in \mathscr{L}(G)\right\} .$$

Thus $\phi_1({}^1G) = \phi_1(K_1)\phi_1(G_1) = \left\{\begin{bmatrix} B & 0 \\ AB & B \end{bmatrix} : A \in \mathscr{L}(G), B \in G\right\}$. Treat the points in the plane (i.e., the complex numbers) as all $2 \times 2$ real matrices of the form $\begin{bmatrix} x & -y \\ y & x \end{bmatrix}$ (corresponding to the number $x + iy$). Then, for fixed $B = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \in G$, $\{AB : A \in \mathscr{L}(G)\}$ is just the line through the origin parallel to the tangent line at $B$, and hence the subdiagonal term can be any complex number.

## References


[1] R. Abraham & J. Robbins, *Transversal mappings and flows,* Benjamin, New York, 1967.

[2] R. H. Bowman, *Differentiable extension,* Tensor (N.S.) **21** (1970) 139–150.

[3] ——, *Differentiable extensions.* II, Tensor (N.S.) **21** (1970) 261–264.

[4] ——, *Second order connections,* I, II, J. Differential Geometry **7** (1972) 549–561, **8** (1973) 75–84.

[5] ——, *High order dissections,* J. Differential Geometry **5** (1971) 169–173.

[6] ——, *Almost submanifold structures,* J. Differential Geometry **9** (1974) 537–546.

[7] H. V. Craig, *Vector and tensor analysis,* McGraw-Hill, New York, 1943.

[8] J. Diudonné, *Foundations of modern analysis,* Academic Press, New York, 1969.

[9] B. Eckmann (ed.), *Seminar on triples and categorical homology theory,* Lecture Notes in Math. No. 80, Springer, Berlin, 1969.

[10] P. de La Harpe, *Classical Banach-Lie algebras and Banach-Lie groups of operators in Hilbert space,* Lecture Notes in Math. No. 285, Springer, Berlin, 1972.

[11] S. Lang, *Differentiable manifolds,* Addison-Wesley, Reading, 1972.

[12] S. Helgason, *Differential geometry and symmetric spaces,* Academic Press, New York, 1972.

[13] S. Sternberg, *Lectures on differential geometry,* Addison-Wesley, Reading, 1972.


ARKANSAS STATE UNIVERSITY
UNIVERSITY OF TEXAS, AUSTIN