

DIFFERENTIABLE SEMIGROUPS ARE LIE GROUPS

JOHN P. HOLMES

Department of Mathematics (FAT)
Auburn University
Alabama 36849-5310

MITCH ANDERSON

Department of Mathematics
University of Hawaii at Hilo
Hilo, HI 96720-4091

(Received May 18, 1992 and in revised form April 25, 1994)

ABSTRACT. We present here a modern, detailed proof to the following theorem which was introduced by Garrett Birkhoff [1] in 1938. *If S is a local semigroup with neighborhood of 1 homeomorphic to a Banach space and with multiplication strongly differentiable at 1, then S is a local Lie Group.* Although this theorem is more than 50 years old and remains the strongest result relating to Hilbert's fifth problem in the infinite dimensional setting, it is frequently overlooked in favor of weaker results. Therefore, it is the goal of the authors here to clarify its importance and to demonstrate a proof which is more accessible to contemporary readers than the one offered by Birkhoff.

KEY WORDS AND PHRASES: Lie groups, Lie algebras, strong differentiability, canonical parameters, Campbell-Baker-Hausdorff formula.

1980 AMS SUBJECT CLASSIFICATION CODE: 22A20

0. INTRODUCTION

In the year 1900, Hilbert presented a list of twenty-three problems at the Second International Congress of Mathematics in Paris. The central idea behind Hilbert's fifth problem is whether every C^0 group is in fact a Lie group. More explicitly, is continuity of the multiplication enough to ensure analyticity of the multiplication? In his comprehensive book on Lie groups, Price [2] gives credit to Schur [3], in 1889, for showing that if G is a C^k -Lie group ($k \geq 2$), then G is a Lie group. Price also points out that an outline of the methods used by Schur may be found in Montgomery and Zippin [4], while details for the case $k \geq 3$ may be found in Pontrjagin [5]. It is interesting to note the absence of Birkhoff's result in Price's book, a result which requires less differentiability (indeed, less than C^1), and which appeared one year prior to Pontrjagin's paper of 1939. In 1929, Von Neumann [6] showed that each topologically closed subgroup of the group of invertible matrices is a manifold and that the multiplication is analytic. In fact, he showed that the exponential series maps a neighborhood of 0 in the tangent space onto a neighborhood of I in the group. In 1952, Gleason [7] and Montgomery and Zippin [8] improved on this result by showing that every finite dimensional, locally connected, locally compact group is a Lie group. Finally, in 1957 Jacobi [9] showed that if G is a local group with Euclidean neighborhood at 1, then G is a local Lie group; while shortly thereafter, Mostert and Shields [10] showed that if S is a local semigroup with Euclidean neighborhood at 1, then S is a local group. Hence, combining these last two results we have if S is a local semigroup with Euclidean neighborhood at 1, then S is a local Lie group. This seems to close the book on Hilbert's fifth problem, at least in the finite dimensional case. However, this result is not true in the general setting of a Banach space. Consider, for example, infinite dimensional Hilbert space, which is homeomorphic to R^N , the countable product of the real numbers. R^N admits

a continuous multiplication, namely coordinate-wise multiplication. Therefore, Hilbert space admits a continuous semigroup multiplication with identity 1. However, there are points arbitrarily close to 1 which are non-invertible. Hence, this semigroup is not even a local group. This indicates that something stronger than continuity is required in the infinite dimensional case, particularly if G is required only to be a semigroup and not a group. It is at this point that the significance of Birkhauff's paper, which holds in the infinite dimensional case, becomes apparent.

DEFINITION. The statement that the function f , with domain the open set U in the Banach space X , and with codomain contained in the Banach space Y , is *strongly differentiable* at the point p in U means there is a continuous linear function $T : X \rightarrow Y$ such that if $\varepsilon > 0$ there is a $\delta > 0$ such that if each of x and y is in U and is within δ of p then $|f(x) - f(y) - T(x - y)| \leq \varepsilon|x - y|$.

The statement in Birkhoff's Theorem that S is a local Lie group means that under canonical parameters and in an appropriate neighborhood of 1, S is representable by a power series. More precisely, suppose S has a neighborhood of 1 homeomorphic to the Banach space X , and V is the corresponding multiplication, in X , defined on a neighborhood of 0. Then there are open sets C and D containing 0 and a transformation $T : D \xrightarrow{\text{onto}} C$ such that T is strongly differentiable at 0, $T'(0) = I$, and if $x \in D$ and each of s, t , and $s + t \in [-1, 1]$, then $V(T(sx), T(tx)) = T((s + t)x)$. This then implies the existence of a multiplication $W : D \times D \rightarrow X$ defined by $W(x, y) = T^{-1}(V(T(x), T(y)))$ which is also strongly differentiable at $(0, 0)$, is associative, and has the property that if $x \in D$ and each of s, t , and $s + t \in [-1, 1]$ then $W(sx, tx) = (s + t)x$. That is, the restriction of W to a straight ray through 0 in D is a local one parameter subgroup. It is the multiplication W that is representable by a power series.

The first appearance of the power series representation of the group multiplication, a particular series given in terms of the associated Lie algebra multiplication, is attributed to Campbell [11] in 1898, by Price, although the key question regarding convergence of the series was not settled there. This problem of convergence was later avoided independently by Baker [12], in 1905, and by Hausdorff [13], in 1906, by giving a purely algebraic version of the series which relies on the algebra of matrices. Consequently, this Lie group power series is often referred to as the Campbell-Baker-Hausdorff (CBH) formula. However, Birkhoff attributes the series originally to Schur and refers to it as the SCH series, inadvertently overlooking Baker altogether. In any event, the convergence of the CBH series is of primary importance to these authors and a detailed proof is given towards the end of this paper.

1. THE CANONICAL TRANSFORMATION

The primary goal of section 1 is to present a proof of the following theorem.

THEOREM 1.1. *If S is a local semigroup with neighborhood of 1 homeomorphic to the Banach space X and with multiplication strongly differentiable at 1, then S can be transformed into canonical parameters.*

Before proceeding directly to a proof of Theorem 1.1, we develop some necessary machinery. Notice that in the theorems which follow, strong differentiability is required at a single point only. The following versions of the chain rule, the Inverse Function Theorem, and the Implicit Function Theorem are included here for completeness, and may be found in numerous functional analysis texts.

THEOREM 1.2. *If each of $X, Y,$ and Z is a Banach space, $f : X \rightarrow Y, g : Y \rightarrow Z,$ and g is strongly differentiable at $f(x_0)$ and f is strongly differentiable at x_0 then $g \circ f$ is strongly differentiable at x_0 at x_0 and $(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0).$*

THEOREM 1.3. *Inverse Function Theorem. If $f : X \rightarrow Y$ is strongly differentiable at x_0 and $f'(x_0)$ is invertible, then there are neighborhoods U and V of x_0 and $f(x_0)$ respectively so that $f|U$ is a homeomorphism onto V and $(f|U)^{-1}$ is strongly differentiable at $f(x_0).$*

THEOREM 1.4. *Implicit Function Theorem. If $f : X \times Y \rightarrow Z$ is strongly differentiable at (x_0, y_0) and $f(x_0, \cdot)'(y_0)$ is invertible, then there is a neighborhood U of x_0 and a unique continuous function $u : U \rightarrow Y$ such that $u(x_0) = y_0$ and $f(x, u(x)) = f(x_0, y_0)$ for each x in $U.$ Moreover, u is strongly differentiable at x_0 and $u'(x_0) = [f(x_0, \cdot)'(y_0)]^{-1} \circ f(\cdot, y_0)'(x_0).$*

Next, let D be an open neighborhood of 0 in the Banach space, $X,$ and $V : D \times D \rightarrow X$ such that V is strongly differentiable at $(0,0)$ and $V(x, 0) = V(0, x) = x$ for each $x \in D.$

APPLICATIONS TO V. $V'(0,0)(x, y) = x + y$ implies that $V(0, \cdot)'(0) = I.$ Therefore, Theorem 1.4 implies the existence of u so that $V(x, u(x)) = 0$ for each x in some neighborhood D' of 0. Similarly, a v exists such that $V(v(x), x) = 0$ for each x in some neighborhood D'' of 0. By associativity of $V,$ $u(x) = v(x)$ for each x close enough to 0. Thus, without loss of generality we may assume that for each $x \in D$ there is a unique x^{-1} and moreover that $x \rightarrow x^{-1}$ is strongly differentiable at 0 and has derivative $-I$ there.

Furthermore, if $T : X \rightarrow X$ is strongly differentiable at 0 and $T(0) = 0$ and $T'(0) = I,$ then the function $W : X \times X \rightarrow X$ defined by $W(x, y) = T^{-1}(V(T(x), T(y)))$ is associative and strongly differentiable at $(0,0)$ by Theorems 1.2 and 1.3.

NOTATION: $\prod_{i=1}^n x_i \equiv V(x_n, V(x_{n-1}, V(\dots, V(x_2, x_1))\dots))$ if this product makes sense, and $x^n \equiv V(x, x^{n-1})$ where $x^2 = V(x, x)$ again whenever each of these makes sense.

Theorems 1.5 and 1.6, which follow, are extremely powerful and will be used throughout the remainder of the paper.

THEOREM 1.5. *If $\epsilon > 0$ there is a $\delta > 0$ so that if $\sum_{i=1}^n |x_i| < \delta$ then*

- i. $\prod_{i=1}^n x_i$ exists,
- ii. $\left| \prod_{i=1}^n x_i - \sum_{i=1}^n x_i \right| \leq \epsilon \sum_{i=1}^n |x_i|,$ and
- iii. $\left| \prod_{i=1}^n x_i \right| < 2 \cdot \sum_{i=1}^n |x_i|$

PROOF. Suppose $0 < \epsilon < 1.$ Choose $\delta > 0$ so that if $|x| + |y|, |a| + |b| < \delta$ then $|V(x, y) - V(a, b) - (x - a) - (y - b)| \leq \epsilon(|x - a| + |y - b|).$ If $|x_1| + |x_2| < \frac{\delta}{2}$ then $|V(x_1, x_2) - V(0, 0) - x_1 - x_2| \leq \epsilon(|x_1| + |x_2|),$ and $|V(x_1, x_2)| \leq (\epsilon + 1)(|x_1| + |x_2|) < \delta.$ In order to use induction on $n,$ suppose that if $\sum_{i=1}^n |x_i| < \frac{\delta}{2}$ then $\prod_{i=1}^n x_i$ exists, has norm $\leq \delta,$ and $\left| \prod_{i=1}^n x_i - \sum_{i=1}^n x_i \right| \leq \epsilon \sum_{i=1}^n |x_i|.$ Now suppose $\sum_{i=1}^{n+1} |x_i| < \frac{\delta}{2}.$ It then follows that

$$\left| \prod_{i=1}^{n+1} x_i - \sum_{i=1}^{n+1} x_i \right| \leq \left| V\left(x_{n+1}, \prod_{i=1}^n x_i\right) - V\left(0, \prod_{i=1}^n x_i\right) - x_{n+1} \right| + \left| \prod_{i=1}^n x_i - \sum_{i=1}^n x_i \right| < \epsilon \sum_{i=1}^{n+1} |x_i|$$

and $\left| \prod_{i=1}^{n+1} x_i \right| \leq (\epsilon + 1) \sum_{i=1}^{n+1} |x_i| < \delta$. Thus, by induction, a proof is complete.

THEOREM 1.6. *If $\epsilon > 0$ there is a $\delta > 0$ so that if $\sum_{i=1}^n |x_i| < \delta$ and $\sum_{i=1}^n |y_i| < \delta$ then each of $\prod_{i=1}^n x_i$ and $\prod_{i=1}^n y_i$ exists and*

$$\left| \prod_{i=1}^n x_i - \prod_{i=1}^n y_i - \sum_{i=1}^n (x_i - y_i) \right| \leq \epsilon \cdot \sum_{i=1}^n |x_i - y_i| .$$

PROOF. Note that $(x, y, z) \rightarrow V(x, V(y, z))$ is strongly differentiable at $(0,0,0)$ by Theorem 1.2, so we may choose $\delta_1 > 0$ so that if $|x| + |y| + |z| < \delta_1$ and $|a| + |b| + |c| < \delta$, then

$$|V(x, V(y, z)) - V(a, V(b, c)) - (x - a) - (y - b) - (z - c)| < \epsilon(|x - a| + |y - b| + |z - c|) .$$

By Theorem 1.5, choose $\delta_2 > 0$ so that if $\sum_{i=1}^n |x_i| < \delta_2$ then $\prod_{i=1}^n x_i$ exists, $\left| \prod_{i=1}^n x_i \right| < 2 \cdot \sum_{i=1}^n |x_i|$, and $\left| \prod_{i=1}^n x_i - \sum_{i=1}^n x_i \right| \leq \epsilon \cdot \sum_{i=1}^n |x_i|$. Let $\delta = \min\left\{\frac{\delta_1}{6}, \frac{\delta_2}{3}\right\}$, and suppose $\sum_{i=1}^n |x_i| < \delta$ and $\sum_{i=1}^n |y_i| < \delta$. Thus, the choice of δ and δ_2 implies that each of $\left| \prod_{i=j+1}^n x_i \right|$, $\left| \prod_{i=1}^j y_i \right|$, $|x_j|$, and $|y_j|$ has norm less than $\frac{\delta_1}{3}$, for each $j \in \{1, 2, \dots, n\}$. It then follows from the choice of δ_1 that

$$\begin{aligned} \left| \prod_{i=1}^n x_i - \prod_{i=1}^n y_i - \sum_{i=1}^n (x_i - y_i) \right| &\leq \epsilon \sum_{j=1}^n \left| V\left(\prod_{i=j+1}^n x_i, V\left(x_j, \prod_{i=1}^{j-1} y_i\right)\right) - V\left(\prod_{i=j+1}^n x_i, V\left(y_j, \prod_{i=1}^{j-1} y_i\right)\right) - (x_j - y_j) \right| \\ &\leq \epsilon \sum_{i=1}^n |x_i - y_i| . \end{aligned}$$

In the following discussion we will consider functions from $[a, b]$ into the Banach space X . If $D > 0$, the function f is said to be D -admissible provided that f is continuous, of bounded variation, $|f(x)| \leq D$ for each $x \in [a, b]$, and $\text{Var}_{[a,b]} f \leq D$. If $e = [\alpha, \beta]$ is a subinterval of $[a, b]$ then $f_e = f(\beta) - f(\alpha)$ and $f^e = f(\beta)f(\alpha)^{-1}$ if it exists. A subdivision of an interval I is a finite collection of nonoverlapping intervals whose union is I .

If P is a subdivision of $[a, b]$ and $t \in [a, b]$ then let P_t denote the subdivision of $[a, t]$ consisting of those members e of P which are contained in $[a, t]$ together with $[\alpha, t]$ if $t \in [\alpha, \beta] \in P$. Define $\prod_P f : [a, b] \rightarrow X$ by $(\prod_P f)(t) = f_e f_{e_{n-1}} \dots f_{e_1}$ if $P_t = \{e_1, \dots, e_n\}$ with $e_i < e_{i+1}$ for each $i = 1, \dots, n-1$, provided the product exists; and define $\sum_P f : [a, b] \rightarrow X$ by $(\sum_P f)(t) = \sum_{e \in P_t} f^e$ provided f^e exists for each $e \in P_t$.

THEOREM 1.7. *There is a $D > 0$ so that if f is D -admissible then each of $\prod_P f$ and $\sum_P f$ exist on $[a, b]$ for each subdivision P . Moreover, if $\epsilon > 0$ there is a $\delta > 0$ so that if each of P and Q is a subdivision of $[a, b]$ with mesh less than δ then*

- i. $|(\prod_P f)(b) - (\prod_Q f)(b)| \leq \epsilon \cdot \text{Var}_{[a,b]} f$, and
- ii. $|(\sum_P f)(b) - (\sum_Q f)(b)| \leq \epsilon \cdot \text{Var}_{[a,b]} f$.

PROOF. Using the strong differentiability at $(0,0)$ of $(x, y) \rightarrow xy^{-1}$ and Theorems 1.2 and 1.6, choose $D > 0$ so that

- i. if each of $|x|$ and $|y|$ is less than $2D$ then y^{-1} exists and $|xy^{-1} - (x - y)| \leq |x - y|$, and
- ii. if each of $\sum_{i=1}^n |x_i|$ and $\sum_{i=1}^n |y_i|$ is less than $6D$ then each of $\prod_{i=1}^n x_i$ and $\prod_{i=1}^n y_i$ exist and

$$\left| \prod_{i=1}^n x_i - \prod_{i=1}^n y_i - \sum_{i=1}^n (x_i - y_i) \right| \leq \sum_{i=1}^n |x_i - y_i| .$$

Next, suppose $0 < \varepsilon < 1$. Using Theorem 1.5, choose $\delta' > 0$ so that if $\sum_{i=1}^n |x_i| < 2\delta'$ then $\prod_{i=1}^n x_i$ exists and $\left| \prod_{i=1}^n x_i - \sum_{i=1}^n x_i \right| \leq \frac{\varepsilon}{4} \cdot \sum_{i=1}^n |x_i|$. Now, assuming f is D -admissible, choose $\delta > 0$ so that if e is a subinterval of $[a, b]$ with length less than δ then $Var_e f \leq \delta'$. Suppose each of P and Q is a subdivision of $[a, b]$ with mesh less than δ and Q refines P . For each $e \in P$ let $Q^e = \{d \in Q : d \subseteq e\}$. Note that if $e \in P$ and $e = [\alpha, \beta]$, then $\sum_{d \in Q^e} |f_d| \leq Var_e f \leq \delta'$ since the length of e is less than δ . Therefore, $\prod_{Q^e} f$ exists and $\left| \left(\prod_{Q^e} f \right) (\beta) - f_e \right| = \left| \left(\prod_{Q^e} f \right) (\beta) - \sum_{d \in Q^e} f_d \right| < \frac{\varepsilon}{4} \cdot \sum_{d \in Q^e} |f_d|$. Thus, $\left| \left(\prod_{Q^e} f \right) (\beta) \right| \leq \left(1 + \frac{\varepsilon}{4} \right) \cdot \sum_{d \in Q^e} |f_d|$ which implies $\sum_{e \in P} \left| \left(\prod_{Q^e} f \right) (\beta) \right| \leq 2 \cdot Var_{[a,b]} f \leq 2D$. Moreover, $\sum_{e \in P} |f_e| \leq D$ and $\sum_{e \in Q} |f_e| \leq D$; hence, each of $(\prod_P f)(b)$ and $(\prod_Q f)(b)$ exist. It then follows, since $(\prod_Q f)(b) = \prod_{e \in P} \left(\prod_{Q^e} f \right) (\beta)$, that

$$\begin{aligned} \left| (\prod_P f)(b) - (\prod_Q f)(b) \right| &\leq \left| (\prod_P f)(b) - \prod_{e \in P} \left(\prod_{Q^e} f \right) (\beta) - \sum_{e \in P} \left(f_e - \left(\prod_{Q^e} f \right) (\beta) \right) \right| + \sum_{e \in P} \left| f_e - \left(\prod_{Q^e} f \right) (\beta) \right| \\ &\leq 2 \cdot \sum_{e \in P} \left| f_e - \left(\prod_{Q^e} f \right) (\beta) \right| \\ &\leq 2 \cdot \frac{\varepsilon}{4} \sum_{e \in P} \sum_{d \in Q^e} |f_d| \\ &\leq \frac{\varepsilon}{2} \cdot Var_{[a,b]} f. \end{aligned}$$

Now, since $|f(t)| \leq 2D$ for each $t \in [a, b]$, which implies $|f(t)^{-1}| \leq 4D$, it follows that f^e exists for each subinterval e of $[a, b]$ and $|f^e| \leq 2|f_e|$. Therefore, $\sum_{e \in P} |f^e| \leq 2 \cdot Var_{[a,b]} f \leq 2D$ for each subdivision P of $[a, b]$. If each of P and Q is defined as above, it follows that $\sum_{d \in Q^e} |f^d| \leq 2 \cdot Var_e f \leq \delta'$ and thus, $\left| \sum_{d \in Q^e} f^d - \prod_{d \in Q^e} f^d \right| \leq \frac{\varepsilon}{4} \cdot \sum_{d \in Q^e} |f^d|$. However, since $\prod_{d \in Q^e} f^d = f^e$, we have

$$\left| \sum_{e \in P} \left(f^e - \sum_{d \in Q^e} f^d \right) \right| \leq \sum_{e \in P} \left| f^e - \sum_{d \in Q^e} f^d \right| \leq \frac{\varepsilon}{4} \cdot \sum_{e \in P} \sum_{d \in Q^e} |f^d| \leq \frac{\varepsilon}{2} \cdot Var_{[a,b]} f.$$

The result now follows easily. Moreover, Theorem 1.7 leads us to the following.

COROLLARY 1.8. *There is a $D > 0$ so that if f is D -admissible and $\varepsilon > 0$ then there is a $\delta > 0$ so that if $e = [\alpha, \beta]$ is a subinterval of $[a, b]$, then there are points $\prod_e f$ and $\sum_e f$ in X so that if P is a subdivision of $[\alpha, \beta]$ with mesh less than δ then*

- i. $|\prod_e f - (\prod_P f)(\beta)| < \varepsilon \cdot Var_{[a,b]} f$, and
- ii. $|\sum_e f - (\sum_P f)(\beta)| < \varepsilon \cdot Var_e f$.

Let $\prod f$ and $\sum f$ be defined for such f by $(\prod f)(t) = \prod_{[a,b]} f$ and $(\sum f)(t) = \prod_{[a,b]} f$. We now return to a proof of Theorem 1.1.

PROOF OF THEOREM 1.1. Choose $D > 0$ so that $6D$ satisfies Theorem 1.6 for $\varepsilon = 1$ and so that D satisfies Corollary 1.8. Let $B(0, D)$ denote the D -neighborhood of 0 in X . For each x in $B(0, D)$ define $f_x : [0, 1] \rightarrow X$ by $f_x(s) = sx$. Since f_x is D -admissible for each $x \in B(0, D)$ we can define $T : B(0, D) \rightarrow X$ by $T(x) = \prod f_x(1)$. Notice that $T(0) = 0$ and that if P is a partition of $[0, 1]$ then

$\sum_{e \in P} (f_x)_e = x$ for each $x \in B(0, D)$. It then follows from Theorem 1.6 that T is strongly differentiable at 0 and $T'(0) = I$. Therefore, applying Theorem 1.3 to T , let U and V be open neighborhoods of 0 in X so that $T|U$ is a homeomorphism. We can now define $W : U \times U \rightarrow V$ by $W(x, y) = T^{-1}(V(T(x), T(y)))$. In order to complete a proof of Theorem 1.1, it remains only to be shown that if $x \in U$ and $s, t, s + t \in [-1, 1]$, then $W(sx, tx) = (s + t)x$. In the case where $s, t \in (0, 1]$ this is clear by construction. Therefore, first consider the case where $t = -s$. Let x be nonzero in $U, s \in (0, 1]$, and $\epsilon > 0$. Choose a positive integer N so that if n is a positive integer, $n \geq N$, then

$$\left| V\left(\frac{sx}{n}, \frac{-sx}{n}\right) - \frac{sx}{n} + \frac{sx}{n} \right| < \frac{\epsilon \cdot \left|\frac{x}{n}\right|}{2 \cdot |sx|} = \frac{\epsilon}{2n}.$$

Let n be an integer, $n \geq N$. The choice of D guarantees that if j is a positive integer, $j \leq n$, then $\left|\left(\frac{sx}{n}\right)^j\right| < 2 \cdot |sx| < 2D$, $\left|\left(\frac{-sx}{n}\right)^j\right| < 2D$, and $\left|\frac{sx}{n} \cdot \frac{-sx}{n}\right| < 2D$. Therefore, since the sum of these three is less than $6D$, it follows that

$$\left| \left(\frac{sx}{n}\right)^{j-1} \left(\frac{sx}{n} \cdot \frac{-sx}{n}\right) \left(\frac{-sx}{n}\right)^{j-1} - \left(\frac{sx}{n}\right)^{j-1} \left(\frac{-sx}{n}\right)^{j-1} - \left(\frac{sx}{n} \cdot \frac{-sx}{n}\right) \right| < \left| \frac{sx}{n} \cdot \frac{-sx}{n} \right|,$$

which implies that

$$\left| \left(\frac{sx}{n}\right)^j \left(\frac{-sx}{n}\right)^j \right| \leq \left| \left(\frac{sx}{n}\right)^{j-1} \left(\frac{-sx}{n}\right)^{j-1} \right| + 2 \cdot \left| \frac{sx}{n} \cdot \frac{-sx}{n} \right|,$$

which, by induction on j is less than $2j \cdot \left|\frac{sx}{n} \cdot \frac{-sx}{n}\right| \leq \epsilon$. Therefore, since $T(sx) = \lim_{N \rightarrow \infty} \left(\frac{sx}{N}\right)^N$, it follows that $V(T(sx), T(-sx)) = 0$ which implies that $W(sx, -sx) = 0 - W(-sx, sx)$. The situation where $t \neq -s$ can be considered in various cases depending on which of s, t , and $s + t$ is positive. For example, in the case where $s > 0, t < 0$, and $s + t > 0$, we have, by construction, $T(sx) = V(T((s + t)x), T(-tx))$. However, since $T(-tx)$ is the inverse of $T(tx)$, it follows that $V(T(sx), T(tx)) = T((s + t)x)$ and hence, $W(sx, tx) = (s + t)x$. The other cases follow similarly, which completes a proof of Theorem 1.1.

Theorems 1.9-1.13 which follow, although not required for a proof of Theorem 1.1, will be used in the main construction of section 3.

THEOREM 1.9. *There is a $D > 0$ so that if $f : [a, b] \rightarrow X$ and f is D -admissible and P is a subdivision of $[a, b]$, then each of $\sum_p f, \sum f, \prod_p f$, and $\prod f$ is $6D$ -admissible.*

PROOF. Choose D' small enough to satisfy Theorem 1.6 for $\epsilon = 1$, Theorem 1.7 and Corollary 1.8, and so that if $|x|, |y|, |z|, |w| \leq D'$ then $|xy^{-1} - zw^{-1} - (x - z) + (y - w)| \leq |x - z| + |y - w|$ and $|xy - zw - (x - z) - (y - w)| \leq |x - z| + |y - w|$. Let $D = D'/3$ and suppose f is D -admissible. Choose $\delta > 0$ so that if $e = [\alpha, \beta]$ is a subinterval of $[a, b]$ and P is a subdivision of e with mesh less than δ then

- i. $|\prod_p f - (\prod_p f)(\beta)| < Var_e f$ and
- ii. $|\sum_e f - (\sum_p f)(\beta)| < Var_e f$.

Note that $\sum_{d \in P} |f_d| \leq Var_e f \leq D$. Therefore, by the choice of D we have $\left| (\prod_p f)(\beta) - \sum_{d \in P} f_d \right| \leq \sum_{d \in P} |f_d|$ which implies $|(\prod_p f)(\beta)| \leq 2 \cdot Var_e f$. Thus, $|\prod_p f| \leq |(\prod_p f)(\beta)| + Var_e f \leq 3 \cdot Var_e f$. Similarly, since $|f^d| \leq 2|f_d|$ for any subinterval d of $[a, b]$ we have $|\sum_e f| \leq \sum_{d \in P} |f^d| + Var_e f \leq 3 \cdot Var_e f$. Note that $|(\prod f)(\beta) - (\prod f)(\alpha)| = |\prod_p f \cdot (\prod f)(\alpha) - (\prod f)(\alpha)|$; but, we have just seen that $|(\prod f)(\alpha)| \leq 3D$

and $|\prod_p f| \leq 3D$. Therefore, we have that

$$|(\prod_p f)_e| \leq |\prod_p f \cdot (\prod_p f)(\alpha) - (\prod_p f)(\alpha) - \prod_p f| + |\prod_p f| \leq 2 \cdot |\prod_p f| \leq 6 \cdot \text{Var}_e f.$$

It follows that $\text{Var}_e \prod_p f \leq 6 \cdot \text{Var}_e f$ for any subinterval e of $[a, b]$ and, hence, that $\prod_p f$ is $6D$ -admissible. Next, $|(\sum_p f)_e| = |(\sum_p f)(\beta) - (\sum_p f)(\alpha)| = |\sum_p f|$, which as noted above, is less than or equal to $3 \cdot \text{Var}_e f$; so $\sum_p f$ is $3D$ -admissible. Now suppose P is an arbitrary subdivision of $[a, b]$. If $e = [\alpha, \beta]$ is in P and $d = [\gamma, \delta]$ is contained in e then

$$\left| \left(\sum_p f \right)_d \right| = |(\sum_p f)(\delta) - (\sum_p f)(\gamma)| = |f(\delta)f(\alpha)^{-1} - f(\gamma)f(\alpha)^{-1}|.$$

But, each of $|f(\delta)|$, $|f(\alpha)|$, and $|f(\gamma)| \leq 2D \leq D'$, which implies $\left| \left(\sum_p f \right)_d \right| \leq 2|f_d|$. It follows that $\text{Var}_d(\sum_p f) \leq 2 \cdot \text{Var}_d f$ for any subinterval d of $[a, b]$. Hence, $\sum_p f$ is $6D$ -admissible. Finally, note that

$$|(\prod_p f)(\delta) - (\prod_p f)(\gamma)| = |(f(\delta) - f(\alpha))(\prod_{[a, \alpha]} f)(\alpha) - (f(\gamma) - f(\alpha))(\prod_{[a, \alpha]} f)(\alpha)|.$$

Therefore, since $|(\prod_{[a, \alpha]} f)(\alpha)| \leq 3D$ and $|f(\delta) - f(\alpha)|, |f(\gamma) - f(\alpha)| \leq D$, we have

$$|(\prod_{[a, \alpha]} f)(\delta) - (\prod_p f)(\gamma)| \leq 2 \cdot |(f(\delta) - f(\alpha)) - (f(\gamma) - f(\alpha))| \leq 2 \cdot \text{Var}_d f.$$

It follows that $\text{Var}_d \prod_p f \leq 2 \cdot \text{Var}_d f$ for any interval d of $[a, b]$ and $\prod_p f$ is $6D$ -admissible.

THEOREM 1.10. *There is a $D > 0$ so that if $f, g : [a, b] \rightarrow X$ are D -admissible, then $|(\prod_p f)(\beta) - (\prod_p g)(\beta)| \leq 2 \cdot \text{Var}_e(f - g)$ for any subinterval e of $[a, b]$ and any subdivision P of e .*

PROOF. Choose $D > 0$ so that D satisfies Theorem 1.7 and Theorem 1.6 for $\varepsilon = 1$. It then follows that $\prod_p f$ and $\prod_p g$ exist for any subdivision P of e and that

$$|\prod_p f - \prod_p g| \leq \sum_{d \in P} |f_d - g_d| \leq 2 \cdot \text{Var}_e(f - g).$$

We now have the following.

COROLLARY 1.11. *If D is as in Theorem 1.10 and each of f and g is D -admissible, then $|\prod_p f - \prod_p g| \leq 2 \cdot \text{Var}_e(f - g)$ for each subinterval e of $[a, b]$.*

THEOREM 1.12. *There is a $D > 0$ so that if $f : [a, b] \rightarrow X$ is D -admissible and $\varepsilon > 0$ there is a $\delta > 0$ so that if P is a subdivision of $[a, b]$ with mesh less than δ then $\text{Var}_{[a, b]}(\sum f - \sum_p f) \leq \varepsilon \cdot \text{Var}_{[a, b]} f$.*

PROOF. Choose $D_1 > 0$ as in Corollary 1.8 and suppose $\varepsilon > 0$. Choose δ as in Corollary 1.8 so that if $e = [\alpha, \beta]$ is a subinterval of $[a, b]$ and P is a subdivision of e with mesh less than δ then $|\sum_p f - (\sum_p f)(\beta)| < \varepsilon \cdot \text{Var}_e f$. Suppose P is a subdivision of $[a, b]$ with mesh less than δ . If $e = [\alpha, \beta]$ is in P we have that $\{e\}$ is a subdivision of e with mesh less than δ . Therefore, since $\sum_p f = (\sum_p f)(\beta) - (\sum_p f)(\alpha)$ and $f^e = (\sum_p f)(\beta) - (\sum_p f)(\alpha) = (\sum_{[e]} f)(\beta)$, we have

$$\left| (\sum_p f - \sum_p f)_e \right| = \left| \sum_p f - (\sum_{[e]} f)(\beta) \right| \leq \varepsilon \cdot \text{Var}_e f.$$

THEOREM 1.13. *There is a $D > 0$ so that if f is D -admissible then each of $\sum f$ and $\prod(\sum f)$ exist and $(\prod(\sum f))(t) = f(t)f(a)^{-1}$.*

PROOF. Choose D_1 small enough to satisfy Theorems 1.9, 1.10, and 1.12 and let $D = D_1/6$. If f is D -admissible, then $\sum f$ and $\sum_p f$ are D_1 -admissible for any subdivision P of $[a, b]$. Also, if $\epsilon > 0$ then there is a $\delta > 0$ so that if P is a subdivision of $[a, b]$ with mesh less than δ then $|\left(\prod(\sum f)\right)(b) - \left(\prod_p(\sum f)\right)(b)| < \epsilon \cdot \text{Var}_{[a,b]} \sum f \leq 6\epsilon \cdot \text{Var}_{[a,b]} f$. Therefore, by Theorem 1.10 we have that $|\left(\prod_p(\sum f)\right)(b) - \left(\prod_p(\sum_p f)\right)(b)| < 2 \cdot \text{Var}_{[a,b]}(\sum f - \sum_p f)$, which, by Theorem 1.12 can be made smaller than $2\epsilon \cdot \text{Var}_{[a,b]} f$ by making the mesh of P small. Thus, there is a P so that

$$\begin{aligned} |\left(\prod(\sum f)\right)(b) - \left(\prod_p(\sum_p f)\right)(b)| &\leq |\left(\prod(\sum f)\right)(b) - \left(\sum_p(\sum f)\right)(b)| + |\left(\sum_p(\sum f)\right)(b) - \left(\sum_p f\right)(b)| \\ &\leq \epsilon \cdot \text{Var}_{[a,b]} f + 4\epsilon \cdot \text{Var}_{[a,b]} f. \end{aligned}$$

However, $\left(\prod_p(\sum_p f)\right)(b) = f(b)f(a)^{-1}$, since $\left(\sum_p f\right)_e = f^e$ for each $e \in P$. Thus, $\left(\prod(\sum f)\right)(b) = f(b)f(a)^{-1}$ and by extension $\left(\prod(\sum f)\right)(t) = f(t)f(a)^{-1}$ for each $t \in [a, b]$.

2. THE DEVELOPMENT OF THE LIE ALGEBRA MULTIPLICATION

In this section we show that if S is a local semigroup with neighborhood of 1 homeomorphic to the Banach space X , then there is a bilinear "bracket" function $[] : X \times X \rightarrow X$ which is a Lie-Algebra multiplication on X . We first define the commutator function $K : S \times S \rightarrow X$ on $S \times S$ by $K(x, y) = x^{-1}y^{-1}xy$ and then after a sequence of theorems show that bracket can be defined on all of $X \times X$ by $[x, y] = \lim_{a,b \rightarrow 0} \frac{1}{ab} K(ax, by)$ for real a and b .

THEOREM 2.1. *If $c > 0$ there is a $d > 0$ so that if $|x| + |y| + |u| < d$ then $|K(ux, y) - K(x, y) - K(u, y)| \leq c |K(u, y)|$ and $|K(x, uy) - K(x, y) - K(x, u)| \leq c |K(x, u)|$.*

PROOF. Suppose $0 < c \leq 1$. Let $d_1 > 0$ such that:

- i. If $|w| + |z| < d_1$ then $|wz - w - z| \leq \frac{c}{4} |w|$, using strong differentiability of the multiplication at $(0,0)$, and
- ii. if $|x| + |a| + |y| < d_1$ and $|x| + |b| + |y| < d_1$ then $|xay - xby - a - b| < \frac{c}{2} |a - b|$, by Theorem 1.2. Choose $d > 0$ so that if $|x| + |u| + |y| < d$ then $|x^{-1}u^{-1}y^{-1}uyx| + |x^{-1}y^{-1}xy| < \frac{d_1}{2}$ and $|x| + |x^{-1}| < \frac{d_1}{2}$, using continuity of the multiplication and inversion functions.

Next, let $|x| + |u| + |y| < d$. It then follows from the choice of d_1 that

$$\begin{aligned} |K(ux, y) - K(x, y) - K(u, y)| &< |K(ux, y) - K(x, y) - x^{-1}k(u, y)x| \\ &\quad + |x^{-1}K(u, y)x - K(u, y)| \\ &\leq |(x^{-1}u^{-1}y^{-1}uyx)(x^{-1}y^{-1}xy) - x^{-1}u - 1y - 1uyx - x^{-1}y - 1xy| \\ &\quad + |x^{-1}u - 1y^{-1}uyx - u^{-1}y^{-1}1uy| \\ &\leq \frac{c}{4} |x^{-1}u^{-1}y^{-1}uyx| + \frac{c}{2} |u - 1y^{-1}uy| \\ &\leq c |K(u, y)|. \end{aligned}$$

Similarly, $|K(x, uy) - K(x, y) - K(x, u)| \leq c |K(x, u)|$.

THEOREM 2.2. *There is a $d > 0$ such that if each of m and n is a positive integer and $m|x| + n|y| < d$ then $|K(x, y)| \leq \frac{2|K(x^m, y^n)|}{mn}$.*

PROOF. Let $d_1 > 0$ such that if $|x| + |u| + |y| < d_1$ then $|K(ux, y) - K(x, y)| \leq \frac{1}{8}|K(u, y)|$. Using Theorem 1.5 for $\varepsilon = 1$, let $d_2 > 0$ such that if $\sum_{i=1}^n |x_i| < d_2$ then $\prod_{i=1}^n x_i$ exists and $\left| \prod_{i=1}^n x_i - \sum_{i=1}^n x_i \right| \leq \sum_{i=1}^n |x_i|$, which of course implies $\left| \prod_{i=1}^n x_i \right| \leq 2 \sum_{i=1}^n |x_i|$. Now suppose m is a positive integer such that $m|x| + |y| < \frac{d_2}{6}$. Then it follows that $|x^t| + |y| + |x| < d_2$ for each $t = 1, 2, \dots, m - 1$. Therefore, the choice of d_1 yields $|K(x^m, y) - mK(x, y)| \leq \sum_{i=1}^m |K(x^i, y) - K(x^{i-1}, y) - K(x, y)| \leq \frac{m}{8}|K(x, y)|$. Similarly, there is a $0 < d_3 < \frac{d_2}{6}$ such that if n is a positive integer and $|x| + n|y| < d_3$ then $|K(x, y^n) - nK(x, y)| \leq \frac{n}{8}|K(x, y)|$. Now let $0 < d < d_3$ and suppose $m|x| + n|y| < d$. It then follows that

$$\begin{aligned} |K(x^m, y^n) - mnK(x, y)| &\leq |K(x^m, y^n) - nK(x^m, y)| + |nK(x^m, y) - K(x, y)| \\ &\leq \frac{n}{8}|K(x^m, y)| + \frac{mn}{8}|K(x, y)| \\ &\leq \frac{mn}{2}|K(x, y)|. \end{aligned}$$

The conclusion now follows from the triangle inequality.

THEOREM 2.3. *There are $d > 0$ and $M > 0$ such that if $|x| + |y| < d$ then $|K(x, y)| \leq M|x||y|$.*

PROOF. Let $d_1 > 0$ such that:

- i. if $m|x| + n|y| < d_1$ then $|K(x, y)| \leq \frac{2|K(x^m, y^n)|}{mn}$, using Theorem 2.2,
- ii. if $m|x| + n|y| < d_1$ then $|x^m| \leq 2m|x|$ and $|y^n| \leq 2n|y|$, using Theorem 1.5 for $\varepsilon = 1$, and
- iii. if $|x^m| + |y^n| < 2d_1$ then $|K(x^m, y^n)| < 1$, using the continuity of K at $(0, 0)$.

Let $0 < d < \frac{d_1}{4}$ and choose $M > \frac{32}{d^2}$. Next, suppose $|x| + |y| < d$. Choose m and n so that $\frac{d_1}{4} < |mx| < \frac{d_1}{2}$ and $\frac{d_1}{4} < |ny| < \frac{d_1}{2}$, which implies $|mx||ny| > \frac{d_1^2}{16}$. It then follows from the choice of d_1 that

$$|K(x, y)| \leq \frac{2|K(x^m, y^n)|}{mn} \leq \frac{2}{mn} \leq \frac{2|mx||ny|}{mn \frac{d_1^2}{16}} \leq M|x||y|.$$

THEOREM 2.4. *There are $d > 0$ and $M > 0$ such that if $|x| + |h| + |g| < d$ then $|g^{-1}xg - h^{-1}xh| \leq M|g - h||x|$.*

PROOF. Choose a positive number d_1 so that if $|a^{-1}| + |b| < d_1$ then $|a^{-1}b - a^{-1}a - (b - a)| \leq \frac{1}{2}|b - a|$. Using Theorem 1.2, choose a positive number d_2 so that if $|a| + |b| < d_2$ then $|a^{-1}ba - a^{-1}a - b| < |b|$. Choose positive numbers M and d so that if $|x| + |h| + |g| < d$ then:

- i. $|(g^{-1}xg)^{-1}| + |h^{-1}xh| < d_1$, which implies $|g^{-1}xg - h^{-1}xh| \leq 2|g^{-1}x^{-1}gh^{-1}xh|$,
- ii. $|h| + |K(gh^{-1}, x)| \leq d_2$, using the continuity of K at $(0, 0)$,
- iii. $|K(gh^{-1}, x)| \leq \frac{M}{4}|gh^{-1}||x|$, using Theorem 2.3 and continuity, and
- iv. $|gh^{-1} - hh^{-1} - (g - h)| \leq |g - h|$, which implies $|gh^{-1}| \leq 2|g - h|$.

Now suppose $|x| + |h| + |g| < d$. It then follows from the choice of d that

$$\begin{aligned} |g^{-1}xg - h^{-1}xh| &\leq 2|g^{-1}x^{-1}gh^{-1}xh| \\ &= 2|h^{-1}K(gh^{-1}, x)h| \\ &\leq 2|K(gh^{-1}, x)| \\ &\leq \frac{M}{2}|gh^{-1}||x| \\ &\leq M|g - h||x|. \end{aligned}$$

The following is a direct consequence of Theorem 2.4 and strong differentiability of the multiplication at $(0,0)$.

COROLLARY 2.5. *There are $d > 0$ and $M > 0$ such that if $|x| + |y| + |u| < d$ then $|K(x, y) - K(u, y)| \leq M|x - u||y|$.*

THEOREM 2.6. *If $c > 0$ there is a $d > 0$ such that if $|x| + |w| + |y| < d$ then $|K(x + w, y) - K(x, y) - K(w, y)| \leq c|w||y|$ and $|K(x, w + y) - K(x, w) - K(x, y)| \leq c|w||x|$.*

PROOF. Suppose $c > 0$. Let $d_1 > 0$ such that if $|x| + |w| < d_1$ then there is a u such that $x + w = ux$, using Theorem 1.4. Let $0 < d_2 < d_1$ such that if $|x| + |y| + |u| < d_2$ then $|K(ux, y) - K(x, y) - K(u, y)| \leq \frac{c}{4}|u||y|$, using Theorems 2.1 and 2.3. Let $0 < d_3 < d_2$ and $M > 0$ such that if $|u| + |y| + |w| < d_3$ then $|K(u, u) - K(w, y)| < M|u - w||y|$, by Corollary 2.5. Finally, let $0 < d < \frac{d_3}{3}$ such that if $|x| + |u| < d$ then $|ux - u - x| \leq \frac{1}{2}|u|$ and $|ux - u - x| \leq \frac{c}{4M}|u|$. Now suppose $|x| + |w| + |y| < d$. Let u be in X such that $x + w = ux$, by choice of d_1 . Notice, since $w = ux - x$, the choice of d yields $|u| \leq 2|w|$, $|u - x| \leq \frac{c}{4M}|u| \leq \frac{c}{2M}|w|$, $|x| + |y| + |u| \leq |x| + |y| + 2|w| \leq 2d < d_2$ and $|u| + |y| + |w| \leq |y| + 3|w| \leq 3d < d_3$. It then follows that

$$\begin{aligned} |K(x + w, y) - K(x, y) - K(w, y)| &\leq |K(ux, y) - K(x, y) - K(u, y)| + |K(u, y) - K(w, y)| \\ &\leq \frac{c}{4}|u||y| + M|u - w||y| \\ &\leq c|w||y|. \end{aligned}$$

Similarly, $|K(x, y + w) - K(x, y) - K(x, w)| \leq c|x||w|$.

Theorem 2.6 and the triangle inequality give the following.

COROLLARY 2.7. *If $c > 0$ there are $d > 0$ and $M > 0$ such that if $|x| + |y| + |u| + |v| < d$ then $|K(x + u, y + v) - K(x, y)| \leq M(|u| + |v|)$.*

THEOREM 2.8. *If $c > 0$ there is a $d > 0$ such that if each of m and n is a positive integer and $m|x| + n|y| < d$ then $|K(mx, v) - mnK(x, y)| \leq cmn|x||y|$.*

PROOF. Suppose $c > 0$. We first show the case where $n = 1$. Let $d > 0$ satisfying Theorem 2.6, and suppose $m|x| + |y| < d$. Then since $(m - j)|x| + |x| + |y| < d$ for each $j = 1, 2, \dots, m$ it follows that $|K(mx, y) - mK(x, y)| \leq \sum_{j=1}^m |K((m - j)x + x, y) - K((m - j)x, y) - K(x, y)| \leq cM|x||y|$. Similarly, there is a $d > 0$ such that if $|x| + n|y| < d$ then $|K(x, v) - nK(x, y)| \leq cn|x||y|$. Next, choose a $d > 0$ so that if $m|x| + |y| < d$ then $|K(mx, y) - mK(x, y)| \leq \frac{c}{2}m|x||y|$ and so that if $|x| + n|y| < d$ then $|K(x, v) - nK(x, y)| \leq \frac{c}{2}n|x||y|$. Suppose $m|x| + n|y| < d$. Since this implies

that $|x| + n|y| < d$ and $m|x| + |v| < d$, it follows that

$$\begin{aligned} |K(mx, v) - mnK(x, y)| &\leq |K(mx, v) - mK(x, v)| + m|K(x, v) - nK(x, y)| \\ &\leq cmn|x||y|. \end{aligned}$$

THEOREM 2.9. *If $c > 0$ there is a $d > 0$ such that if $|x| + |y| < d$ and each of h, l, m , and n is a positive integer such that $0 < \frac{h}{m}, \frac{l}{n} < 1$ then*

$$\left| K\left(\frac{h}{m}x, \frac{l}{n}y\right) - \frac{hl}{mn}K(x, y) \right| \leq c \frac{hl}{mn}|x||y|.$$

PROOF. Suppose $c > 0$. Using Theorem 2.8, choose $d > 0$ so that if $m|x| + n|y| < d$ then $|K(mx, v) - mnK(x, y)| \leq \frac{c}{2}mn|x||y|$. Suppose $|x| + |y| < d$ and $0 < \frac{h}{m}, \frac{l}{n} < 1$. Then, since $h\left|\frac{x}{m}\right| + l\left|\frac{y}{n}\right| < d$, it follows that

$$\begin{aligned} \left| K\left(\frac{h}{m}x, \frac{l}{n}y\right) - \frac{hl}{mn}K(x, y) \right| &\leq \left| K\left(h\left(\frac{x}{m}\right), l\left(\frac{y}{n}\right)\right) - hlK\left(\frac{x}{m}, \frac{y}{n}\right) \right| + \left| hlK\left(\frac{x}{m}, \frac{y}{n}\right) - \frac{hl}{mn}K(x, y) \right| \\ &\leq \frac{c}{2}hl \left| \frac{x}{m} \right| \left| \frac{y}{n} \right| + \frac{hl}{mn} \left| mnK\left(\frac{x}{m}, \frac{y}{n}\right) - K\left(m\left(\frac{x}{m}\right), n\left(\frac{y}{n}\right)\right) \right| \\ &\leq c \frac{hl}{mn}|x||y|. \end{aligned}$$

We now show Theorem 2.9 holds for real numbers s and t between 0 and 1.

THEOREM 2.10. *If $c > 0$ there is a $d > 0$ such that if $|x| + |y| < d$ and $0 < s, t < 1$ then*

$$\left| \frac{1}{st}K(sx, ty) - K(x, y) \right| \leq c|x||y|.$$

PROOF. Suppose $c > 0$. Let $d_1 > 0$ and $M_1 > 0$ such that if $|x| + |y| + |u| + |v| < d_1$ then $|K(x + u, y + v) - K(x, y)| \leq M_1(|u| + |v|)$, by Corollary 2.7. Using Theorem 2.9, choose $0 < d_2 < d_1$ so that if $|x| + |y| < d_2$ and $0 < \frac{h}{m}, \frac{l}{n} < 1$ then $\left| K\left(\frac{h}{m}x, \frac{l}{n}y\right) - \frac{hl}{mn}K(x, y) \right| \leq \frac{c}{6}hl|x||y|$. Now suppose $|x| + |y| < d$ and $0 < s, t < 1$. Choose rational $\frac{h}{m} < s$ and $\frac{l}{n} < t$ so that

- i. $\frac{M}{st} \left(\left| s - \frac{h}{m} \right| |x| + \left| t - \frac{l}{n} \right| |y| \right) \leq \frac{c}{3}|x||y|$,
- ii. $\frac{1}{st} \cdot \frac{hl}{mn} < 2$, and
- iii. $\left(st - \frac{hl}{mn} \right) \frac{M}{st} < \frac{c}{3}|x||y|$.

It now follows that

$$\begin{aligned} \left| \frac{1}{st}K(sx, ty) - K(x, y) \right| &\leq \frac{1}{st} \left| K\left(\frac{h}{m}x + \left(s - \frac{h}{m}\right)x, \frac{l}{n}y + \left(t - \frac{l}{n}\right)y\right) - K\left(\frac{h}{m}x, \frac{l}{n}y\right) \right| \\ &\quad + \frac{1}{st} \left| K\left(\frac{h}{m}x, \frac{l}{n}y\right) - \frac{hl}{mn}K(x, y) \right| + \frac{1}{st} \left| \frac{hl}{mn} - st \right| |K(x, y)| \\ &\leq c|x||y|. \end{aligned}$$

It is clear that we can now define the bracket function on all of $X \times X$ by $[x, y] = \lim_{s, t \downarrow 0} \frac{1}{st}K(sx, ty)$. Furthermore, since $K(-x + x, y) = K(0, y) = 0$, Theorem 2.6 can be used to show that if $c > 0$ there is a $d > 0$ such that if $|x| + |y| < d$ then $|K(-x, y) + K(x, y)| \leq c|x||y|$, which in turn can be used to show $[x, y] = \lim_{s, t \rightarrow 0} \frac{1}{st}K(sx, ty)$. The next theorem follows immediately from the definition of $[x, y]$ and Theorem 2.10.

THEOREM 2.11. *If $c > 0$ there is a $d > 0$ such that if $|x| + |y| < d$ then $|[x, y] - K(x, y)| \leq c|x||y|$.*

The following theorem shows that the bracket function is bilinear.

THEOREM 2.12. $[x + w, y] = [x, y] + [w, y]$ and $[x, y + z] = [x, y] + [x, z]$.

PROOF. Suppose $c > 0$ and each of x, w , and y is in X . Choose $0 < s, t < 1$ so that

- i. $\left|[x + w, y] - \frac{1}{st}K(s(x + w), ty)\right| < \frac{c}{4} \left|[x, y] - \frac{1}{st}K(sx, ty)\right| < \frac{c}{4} \left|[w, y] - \frac{1}{st}K(sw, ty)\right| < \frac{c}{4}$, using the definition of $[x, y]$, and
- ii. $|K(sx + sw, ty) - K(sx, ty) - K(sw, ty)| \leq \frac{c|sw||ty|}{4|w||y|}$, using Theorem 2.6.

The conclusion now follows easily from the triangle inequality.

We now combine some of the previous results into a single theorem. Notice that part ii, which follows from part i and Theorem 2.11, together with Theorem 2.12 implies that the bracket function is continuous.

THEOREM 2.13. *There are $d > 0$ and $M > 0$ such that if $|x| + |y| + |v| < d$ then*

- i. $|K(x, y)| \leq M|m||y|$,
- ii. $|[x, y]| \leq M|x||y|$, and
- iii. $|y^{-1}xy - v^{-1}xv| \leq M|x||y - v|$.

We now show that $[\]$ is a Lie-Algebra.

THEOREM 2.14. $[x, y] = -[y, x]$.

PROOF. Suppose $c > 0$ and each of x and y is in X . Using Theorem 2.13, choose $d > 0$ and $M > 0$ such that if $|a| + |b| < d$ then $|K(a, b)| \leq M|a||b|$. Choose $0 < s < 1$ so that

- i. $|sx| + |sy| < d$,
- ii. $|K(sx, sy) + K(sy, sx)| = |K(sx, sy) - K(sx, sy)K(sy, sx) + K(sy, sx)| \leq \frac{c}{3M}K(sx, sy)$, using strong differentiability of the multiplication at $(0,0)$, and
- iii. $\left|[x, y] - \frac{1}{s^2}K(sx, sy)\right| \leq \frac{c}{3}$ and $\left|[y, x] - \frac{1}{s^2}K(sy, sx)\right| \leq \frac{c}{3}$, by definition of $[x, y]$.

The conclusion now follows from the triangle inequality.

The next two theorems will be used in Theorem 2.17 to show that the bracket function satisfies Jacobi's identity.

THEOREM 2.15. *If $c > 0$ there is a $d > 0$ such that if $|x| + |y| + |z| < d$ then*

$$|[[x, y], z] - K(K(x, y), z)| \leq c|x||y||z|.$$

PROOF. Suppose $c > 0$ and each of x, y , and z , is in X . Using Theorem 2.13, choose $d_1 > 0$ and $M_1 > 1$ so that if $|x| + |y| < d_1$ then $|[x, y]| \leq M_1|x||y|$. Using Corollary 2.5, choose $0 < d_2 < d_1$ and $M > M_1$ so that if $|x| + |y| + |z| < d_2$ then $|K(x, z) - K(y, z)| \leq M|x - y||z|$. Using Theorem 2.11, choose $0 < d_3 < d_2$ so that if $|x| + |y| + |z| < d$ then $|K(x, y)| + |[x, y]| + |z| < d_3$. Now suppose $|x| + |y| + |z| < d$. It then follows that

$$\begin{aligned}
|K(K(x, y), z) - [[x, y], z]| &\leq |K(K(x, y), z) - K([x, y], z)| + |K([x, y], z) - [[x, y], z]| \\
&\leq M|K(x, y) - [x, y]| |z| + \frac{c}{2M}|[x, y]| |z| \\
&\leq \frac{Mc}{2M}|x| |y| |z| + \frac{c}{2M}M_1|x| |y| |z| \\
&\leq c|x| |y| |z|.
\end{aligned}$$

THEOREM 2.16. *If $c > 0$ there is a $d > 0$ such that if $|x| + |y| + |z| < d$ then*

$$|K(K(x, y), z) \cdot K(K(z, x), y) \cdot K(K(y, z), x)| \leq c|x| |y| |z|.$$

PROOF. First note that it is easy to verify * below by direct computation that

$$K(x, y)K(K(x, y), z)K(z, y)K(z, x)K(K(z, x), y)K(y, x)K(y, z)K(K(y, z), x)K(x, z) = 0.$$

Now suppose $c > 0$. Using strong differentiability of the multiplication at $(0, 0)$ and Theorem 2.13, choose $d_1 > 0$ and $M > 1$ so that if $|a| + |b| + |g| + |h| < d_1$ then

$$|gabh - gabh + K(b, a)| \leq |K(b, a)| \leq M|a| |b|,$$

which implies $|gabh - gabh| \leq 2M|a| |b|$, since $abK(b, a) = ba$. Choose $0 < d < d_1$ so that

- i. $d < \frac{c}{8M^4}$, and
- ii. if $|x| + |y| + |z| < d$ and $a = \prod_{i=1}^j a_i$, where $j \in \{1, 2, \dots, 9\}$ and a_i is one of the nine terms in (*) above, then $|a| < \frac{d_1}{4}$.

Now suppose $|x| + |y| + |z| < d$. Then, since

$$K(x, y)K(y, x) = K(z, y)K(y, z) = K(z, x)K(x, z) = 0,$$

the conclusion follows from the triangle inequality.

We now give Jacobi's identity.

THEOREM 2.17. $[[x, y], z] + [[z, x], y] + [[y, z], x] = 0$.

PROOF. Suppose each of x, y , and z is in X and $c > 0$. First choose $d_1 > 0$ and $M > 0$ so that if $|a| + |b| < d_1$ then $|K(a, b)| \leq M|a| |b|$, by Theorem 2.13. Next choose $0 < d < d_1$ so that if $|x_1| + |x_2| + |x_3| < d$ then

- i. $|x_1x_2x_3 - (x_1 + x_2 + x_3)| \leq \frac{c}{7M^2}(|x_1| + |x_2| + |x_3|)$, using Theorem 1.5,
- ii. $|K(K(x_1, x_2), x_3) \cdot K(K(x_3, x_1), x_2) \cdot K(K(x_2, x_3), x_1)| \leq \frac{c}{7}|x_1| |x_2| |x_3|$, using Theorem 2.16, and
- iii. $|[[x_1, x_2], x_3] + [[x_3, x_1], x_2] + [[x_2, x_3], x_1] - (K(K(x_1, x_2), x_3) + K(K(x_3, x_1), x_2) + K(K(x_2, x_3), x_1))| \leq \frac{3c}{7}(|x_1| |x_2| |x_3|)$, using Theorem 2.15 and the triangle inequality.

Now choose $0 < \lambda < 1$ so that $|\lambda x| + |\lambda y| + |\lambda z| < d$, and

$$|K(K(\lambda x, \lambda y), \lambda z)| + |K(K(\lambda z, \lambda x), \lambda y)| + |K(K(\lambda y, \lambda z), \lambda x)| < d.$$

It then follows that

$$\begin{aligned}
|[x, y], z] + [[z, x], y] + [[y, z], x] &= \frac{1}{\lambda^3}([[\lambda x, \lambda y], \lambda z] + [[\lambda z, \lambda x], \lambda y] + [[\lambda y, \lambda z], \lambda x]) \\
&\leq c|x| |y| |z|,
\end{aligned}$$

which completes a proof.

3. THE MAIN CONSTRUCTION

The purpose of this section is to prove that if an analytical group is in canonical coordinates then the multiplication function is determined by the bilinear function $[\ , \]$ in a canonical way.

THEOREM 3.1. *If (X, V) is an analytical group in canonical parameters then there is a $d > 0$ so that if $|x|, |y| < d$ then*

$$y^{-1}xy = x - [y, z] + (1/2!)[y, [y, x]] - \dots = \exp(ad(-y))(x).$$

Notice in canonical parameters this also implies that

$$xyx^{-1} = x + [y, x] + (1/2!)[y, [y, x]] + \dots = \exp(ad(y))(x).$$

PROOF. Suppose $\varepsilon > 0$. By Theorem 2.13 choose M and $d > 0$ so that if $|x|, |y|, |v| < d$ then

$$|K(x, y)| < M|x||y|,$$

$$|[x, y]| < M|x||y|,$$

and

$$|y^{-1}xy - v^{-1}xv| < M|x||y - v|.$$

Suppose $|y| < 1$ and choose n so that $|\frac{1}{n}y| < d$. Let T be defined by

$$Tx = \left(\frac{1}{n}y\right)^{-1}x\left(\frac{1}{n}y\right)$$

and S by

$$Sx = x + \left[x, \frac{1}{n}y\right].$$

Suppose $|x| < d/e^M$. We then have

$$|Tx - 0^{-1}x0| < M|x| \left|\frac{1}{n}y\right|$$

since $|x| < d$ and $|\frac{1}{n}y| < d$. Thus we have

$$|Tx| < \left(1 + \frac{1}{n}M|y|\right)|x|.$$

Assume that if $0 \leq l \leq k - 1 < n$ then $(T^l)x$ exists and

$$|(T^l)x| < \left(1 + \frac{1}{n}M|y|\right)^l |x|.$$

Since $\left(1 + \frac{1}{n}M|y|\right)^{k-1} < e^M$ we have $|T^{k-1}x| < d$ so $T(T^{k-1})x$ exists and

$$|T^k x| \leq \left(1 + \frac{1}{n}M|y|\right) |T^{k-1}x| \leq \left(1 + \frac{1}{n}M|y|\right)^k |x|.$$

A similar induction shows that

$$|S^k x| \leq \left(1 + \frac{1}{n}M|y|\right)^k |x|$$

for $0 \leq k \leq n$.

By Theorem 2.11 choose $0 < \delta < 1$ so that if $|a|, |b| < \delta$ then

$$|K(a, b) - [a, b]| < \varepsilon|a||b|$$

and so that

$$|ab - a - b| < \varepsilon|b| .$$

Suppose $|y| < 1$ and choose an integer n so that $\frac{1}{n}|y| < \delta$ and $M\frac{1}{n}|y| < \delta$. Suppose $|x| < (\min(\delta, d))/e^M$.

Then $|T^k x| < \min(\delta, d)$ and $|K(T^k x, \frac{1}{n}y)| < M|T^k x| \frac{1}{n}|y| < \delta$ for $k = 0, 1, \dots, n - 1$.

Suppose $0 \leq k \leq n$ and consider

$$\begin{aligned} |T(T^k x) - S(T^k x)| &= \left| T^k x K\left(T^k x, \frac{1}{n}y\right) - T^k x - \left[T^k x, \frac{1}{n}y\right] \right| \leq \\ &\left| T^k x K\left(T^k x, \frac{1}{n}y\right) - T^k x - K\left(T^k x, \frac{1}{n}y\right) \right| + \left| K\left(T^k x, \frac{1}{n}y\right) - \left[T^k x, \frac{1}{n}y\right] \right| \leq \\ &\varepsilon \left| K\left(T^k x, \frac{1}{n}y\right) \right| + \varepsilon |T^k x| \frac{1}{n}|y| \leq \varepsilon(M + 1) \left(1 + \frac{1}{n}M|y|\right)^k \frac{1}{n}|y| |x| . \end{aligned}$$

For the same x, y, n , and range of k we have

$$|S(T^k x) - S(S^k x)| \leq \left(1 + \frac{1}{n}M|y|\right) |T^k x - S^k x| .$$

It follows that

$$|T^k x - S^k x| \leq \varepsilon(M + 1) \left(1 + \frac{1}{n}M|y|\right)^k \frac{1}{n}|y| |x| + \left(1 + \frac{1}{n}M|y|\right) |T^{k-1} x - S^{k-1} x|$$

for $k = 1, \dots, n$. By induction then we have

$$|T^n x - S^n x| \leq n\varepsilon(M + 1) \left(1 + \frac{1}{n}M|y|\right)^n |x| \frac{1}{n}|y| \leq \varepsilon(M + 1)|x| |y| e^{M|y|} .$$

If we are in canonical parameters then $T^n x = y^{-1}xy$. Moreover, $S^n x$ approaches $\exp(ad(-y))(x)$ as n increases without bound. Thus we have shown that if $\varepsilon > 0$ then there is a $\delta > 0$ so that if $|x| < \delta$ and $y^{-1}xy$ exists then

$$|y^{-1}xy - \exp(ad(-y))(x)| < \varepsilon|x| |y| .$$

However, in canonical parameters, we have

$$n \left(y^{-1}\left(\frac{1}{n}x\right)y\right) = \left(y^{-1}\left(\frac{1}{n}x\right)y\right)^n = \left(y^{-1}\left(\frac{1}{n}x\right)^n y\right) = y^{-1}xy$$

so $\left(y^{-1}\left(\frac{1}{n}x\right)y\right) = \frac{1}{n}(y^{-1}xy)$. Thus, if x and y are such that $y^{-1}(tx)y$ exists for each $t \in [0, 1]$, we have that if n is sufficiently large then

$$\left|y^{-1}\left(\frac{1}{n}x\right)y - \exp(ad(-y))\left(\frac{1}{n}x\right)\right| < \varepsilon \frac{1}{n}|x| |y|$$

and the fact that we are in canonical parameters implies that we can cancel out $\frac{1}{n}$ in the inequality to obtain

$$|y^{-1}xy - \exp(ad(-y))(x)| < \varepsilon|x| |y| .$$

Hence we have

$$y^{-1}xy = \exp(ad(-y))(x)$$

and we are done with a proof of Theorem 3.1.

Recall the definition of D admissible following the proof of Theorem 1.6.

LEMMA 3.2. *There is a $\delta > 0$ so that if $\delta > D > 0$ and each of p and q is D admissible then the product pq is $4D$ admissible.*

PROOF. Choose $\delta > 0$ so that if each of x, y, a , and b has norm less than δ then

$$|xy - ab - (x - a) - (y - b)| < |x - a| + |y - b| .$$

Let $0 < D < \delta$ and suppose each of p and q is D admissible. Let $[a, b] = \text{dom}(p) = \text{dom}(q)$ and suppose $[c, d] \subset [a, b]$. Note that each of $p(c), p(d), q(c)$, and $q(d)$ has norm less than δ and hence

$$|p(d)q(d) - p(c)q(c) - (p(d) - p(c)) - (q(d) - q(c))| < |p(d) - p(c)| + |q(d) - q(c)| .$$

It follows that

$$|p(d)q(d) - p(c)q(c)| < 2(|p(d) - p(c)| + |q(d) - q(c)|)$$

and hence that pq has variation not exceeding $2(\text{Var}_{[a,b]}p + \text{Var}_{[a,b]}q) < 4D$.

We have from our choice of δ that

$$|p(d)q(d) - 00 - (p(d) + q(d))| < |p(d)| + |q(d)|$$

or that

$$|p(d)q(d)| < 2(|p(d)| + |q(d)|) < 4D .$$

Thus pq is $4D$ admissible.

The following lemma will be used to prove Theorem 3.4.

LEMMA 3.3. *There is $\delta > 0$ so that if $\delta > D > 0$ and p is D admissible then p^{-1} is $2D$ admissible.*

Here p^{-1} denotes the multiplicative inverse; that is, $p^{-1} : [a, b] \rightarrow X$ by $p^{-1}(x) = (p(x))^{-1}$.

PROOF. Since the derivative of $x \rightarrow x^{-1}$ at 0 is $-Id$, choose $\delta > 0$ so that if each of x and y has norm less than δ then

$$|x^{-1} - y^{-1} - (y - x)| < |x - y| .$$

If $0 < D < \delta$ and p is D admissible then the last inequality shows that for each x and y in $\text{dom}(p)$ we have

$$|p(x)^{-1} - p(y)^{-1}| < 2|p(x) - p(y)|$$

and

$$|p(x)^{-1} - (-p(x))| < |p(x)| .$$

It follows that p^{-1} is $2D$ admissible.

LEMMA 3.4. *There is a $D > 0$ such that if each of p and q is a D admissible function from $[0, \Lambda]$ into X and $\varepsilon > 0$ then there is a $\delta > 0$ such that if $[\alpha, \beta] \subset [0, \Lambda]$ has length less than δ then*

$$|p(\beta) - p(\alpha) - [u(\alpha)(q(\beta) - q(\alpha))u(\alpha)^{-1} + u(\beta) - u(\alpha)]| \leq \varepsilon(\text{Var}_{[\alpha,\beta]}q + \text{Var}_{[\alpha,\beta]}u) ,$$

where $u(\lambda) = (\prod p(\lambda))[\prod q(\lambda)]^{-1}$.

PROOF. Choose $D_1 > 0$ as in Lemmas 3.2 and 3.3 and suppose $0 < D_2 < D_1$ satisfies Corollary 1.8 and Theorems 1.9 and 1.13. Choose $0 < D < D_2/(48 * 6)$ and so that if $|x| + |a| + |b| < D$ then $|xax^{-1} - xbx^{-1}| < 2|a - b|$. Now suppose $\varepsilon > 0$ and each of p and q is D admissible. It then follows from the two previous theorems that each of u and $\sum u$ is D_2 admissible. Choose $\delta > 0$ to satisfy Corollary 1.8 and also small enough so that if $e = [\alpha, \beta] \subseteq [0, \Lambda]$ with length less than δ then

$$|\prod(\sum_e u)u(\alpha)\prod_e q[u(\alpha)]^{-1} - \prod(\sum_e u) - u(\alpha)\prod_e q[u(\alpha)]^{-1}| \leq \varepsilon|\prod(\sum_e u)| .$$

Here $\prod(\sum_e u)$ denotes $\prod(\sum u|_{[\alpha,\beta]})(\beta)$ and similarly $\prod_e q$ denotes $\prod(q|_{[\alpha,\beta]})(\beta)$. Now assume $e = [\alpha, \beta]$ has length less than δ and consider the following inequality.

$$\begin{aligned}
 & |p(\beta) - p(\alpha) - [u(\alpha)(q(\beta) - q(\alpha))u(\alpha)^{-1} + u(\beta) - u(\alpha)]| \\
 & \leq |p(\beta) - p(\alpha) - \prod_r p| + |\prod_r p - \prod(\sum_r u[u(\alpha)])\prod_r q[u(\alpha)]^{-1}| \\
 & \quad + |(\prod(\sum_r u)) (u(\alpha)\prod_r q[u(\alpha)]^{-1}) - (\prod(\sum_r u)) - (u(\alpha)\prod_r q[u(\alpha)]^{-1})| \\
 & \quad + |\prod(\sum_r u) - (u(\beta) - u(\alpha))| + |u(\alpha)\prod_r q[u(\alpha)]^{-1} - u(\alpha)(q(\beta) - q(\alpha))u(\alpha)^{-1}|.
 \end{aligned}$$

Each of the first and fourth summands is less than $\epsilon Var_{[\alpha,\beta]}p$ by choice of D_2 . The second summand is identically 0, again by choice of D_2 , since $\prod_r p = \prod p(\beta)[\prod p(\alpha)]^{-1}$, $\prod_r q = \prod q(\beta)[\prod q(\alpha)]^{-1}$, and $\prod(\sum_r u) = u(\beta)[u(\alpha)]^{-1}$. The third summand is less than $\epsilon |u(\beta)[u(\alpha)]^{-1}|$, by choice of δ , which is less than $2\epsilon Var_{[\alpha,\beta]}u$. Finally, the fifth summand is less than $2|\prod_r q - q(\beta) - q(\alpha)|$ by choice of D , which in turn is less than $2\epsilon Var_{[\alpha,\beta]}q$ by choice of D_2 . Therefore, if D is chosen small enough to also insure that $Var_{[\alpha,\beta]}p \leq 2Var_{[\alpha,\beta]}\prod p \leq 4(Var_{[\alpha,\beta]}u + Var_{[\alpha,\beta]}[\prod q]^{-1}) \leq 48(Var_{[\alpha,\beta]}q)$, it follows that the original quantities can be made to differ by less than $51\epsilon(Var_{[\alpha,\beta]}u + Var_{[\alpha,\beta]}q)$, which completes a proof of Theorem 3.4.

THEOREM 3.5. *There is a $D > 0$ such that if each of $p, q,$ and q' is a continuously differentiable, D admissible function from $[0, \Lambda]$ into X and $u = (\prod p)(\prod q)^{-1}$, then $p' = uq'u^{-1} + u'$.*

PROOF. Choose D to satisfy Theorem 3.1 and Lemma 3.4 and suppose each of $p, q,$ and q' is D admissible. Let $\alpha \in [0, \Lambda]$ and $\epsilon > 0$. Choose $\delta > 0$ satisfying Lemma 3.4 and the definition of differentiability at α for u and q . If $\beta \in [0, \Lambda]$ and $|\alpha - \beta| < \delta$, it follows that

$$\begin{aligned}
 & |p(\beta) - p(\alpha) - (uq'u^{-1} + u')(\alpha) \cdot (\beta - \alpha)| \\
 & \leq |p(\beta) - p(\alpha) - [u(\alpha)(q(\beta) - q(\alpha))u(\alpha)^{-1} + u(\beta) - u(\alpha)]| \\
 & \quad + |u(\alpha)(q(\beta) - q(\alpha))u(\alpha)^{-1} - u(\alpha)(q'(\alpha)(\beta - \alpha))u(\alpha)^{-1}| \\
 & \quad + |u(\alpha)(q'(\alpha)(\beta - \alpha))u(\alpha)^{-1} - u\alpha q'(\alpha)u(\alpha)^{-1} \cdot (\beta - \alpha)| + |u(\beta - u(\alpha) - u'(\alpha)(\beta - \alpha))|.
 \end{aligned}$$

The first summand is less than $\epsilon(Var_{[\alpha,\beta]}q + Var_{[\alpha,\beta]}u)$ by choice of D and δ , while the fourth summand is less than $\epsilon|\beta - \alpha|$ by choice of δ . If D and δ are chosen small enough, the second summand is less than $2|q(\beta) - q(\alpha) - q'(\alpha)(\beta - \alpha)|$, which in turn is less than $2\epsilon|\beta - \alpha|$, and the third summand is identically 0 by Theorem 3.1. This is enough to complete a proof of Theorem 3.5.

Suppose each of x and y is near 0 and consider the continuously differentiable function q defined on $[0,1]$ by

$$q(t) = tx + ty + \frac{t^2}{2!}[x, y] + \frac{t^3}{3!}[x, [x, y]] + \dots$$

The derivative of q is given by the series

$$q'(t) = x + y + \frac{t}{1!}[x, y] + \frac{t^2}{2!}[x, [x, y]] + \dots = x + \exp(tad(x))(y)$$

by Theorem 3.1.

Now consider the function r defined by $r(t) = (tx)(ty)$. If x and y are close to 0 then $\prod(\sum r)(1) = r(1)r(0)^{-1} = xy$. The next lemma shows that $\sum r = q$ so that we have $\prod q(1) = xy$.

LEMMA 3.6. *There is a $d > 0$ so that if each of x and y is within d of 0 and each of q and r is defined as above then $q = \sum r$.*

PROOF. By Lemma 3.2 we have that r is $4d$ admissible if $|x|, |y| < d$ if d is sufficiently small. By Theorem 3.1 we have the existence of a $d > 0$ so that if $|x|, |y| < d$ then $q'(t) = x + (tx)y(tx)^{-1}$ for $t \in [0,1]$. Suppose d is small enough to make both of these happen and so that Corollary 1.8 holds for $D = 4d$ and $|x|, |y| < d$.

Suppose each of t and $t + h$ is in $[0,1]$ and consider

$$\begin{aligned} |r(t+h)r(t)^{-1} - (q(t+h) - q(t))| &\leq |r(t+h)r(t)^{-1} - hq'(t)| + |hq'(t) - (q(t+h) - q(t))| \\ &\leq |r(t+h)r(t)^{-1} - hq'(t)| + Mh^2 \end{aligned}$$

for some $M > 0$.

Continuing the chain of inequalities we have the last sum equals

$$|(hx)(tx)(hy)(ty)(ty)^{-1}(tx)^{-1} - hx - (tx)(hy)(tx)^{-1}| + Mh^2.$$

We can force each of hx and $(tx)(hy)(tx)^{-1}$ as close as we like to 0 by choosing h small so if $\epsilon > 0$ we can find a $\delta > 0$ so that if $|h| < \delta$ then the last quantity is dominated by $\epsilon|h||x| + Mh^2$. On the other hand, if we apply Corollary 1.8 to r we have

$$|r(t+h)r(t)^{-1} - ((\sum r)(t+h) - (\sum r)(t))| < \epsilon \text{Var}_{[t, t+h]} r.$$

But, using strong differentiability of multiplication, we have the existence of a $B > 0$ so that $|r(b) - r(a)| < B|b - a|$ and hence that $\text{Var}_{[t, t+h]} r < B|h|$.

Combining all the inequalities we have if each of $|x| < d$ and $|y| < d$ and $|h| < \delta$ then

$$|(\sum r)(t+h) - (\sum r)(t) - (q(t+h) - q(t))| < \epsilon B|h| + \epsilon|x| + Mh^2$$

and hence that $(\sum r)$ is differentiable and that $(\sum r)' = q'$ on $[0,1]$. The lemma now follows since $(\sum r)(0) = q(0)$.

The strategy for building a series expression for xy is now apparent. We have the series q whose product integral is xy . We will modify q into a succession of p 's, each having product integral xy . Indeed, noting that each of the first two terms of q is already of the form tz for some $z \in X$, we will alter the third term $\frac{t^2}{2!}[x, y]$ into this form while at the same time introducing a new infinite series of terms. Each of these new terms, together with each of the terms of the original series, will be altered one at a time in succeeding steps, each step introducing a new series and in the limit approaching a P of the form $P(t) = tZ$, where here Z denotes the Campbell-Hausdorff series. Due to the complexity of the construction, rather than choosing a D at this time, we will simply assume that each of x and y is chosen small enough for each intermediate step and show in the end that it is indeed possible to choose such a D .

We first choose a norm on X so that if each of x and y is in X then $|[x, y]| \leq |x| |y|$. Next, let b be a denumeration of the set of all brackets involving x and y such that $b_1 = x, b_2 = y, b_3 = [x, y]$ and if n is a positive integer greater than 2 then $b_n = [b_i, b_j]$, where each of i and j is a distinct positive integer less than n . Furthermore, we define the length of each of b_1 and b_2 as 1 and the length of b_n inductively as $l_i + l_j$ where l_k denotes the length of b_k and we require the sequence l to be non-decreasing. In other words, b is a sequence of brackets in x and y of non-decreasing length. Next, define $p_1 : [0, 1] \rightarrow X$ by

$$p_1(t) = tx + ty + \frac{t^2}{2!}[x, y] + \frac{t^3}{3!}[x, [x, y]] + \dots = \sum_{i=1}^{\infty} p_{1,i}(t)b_i.$$

Notice that $p_{1,i} = 0$ for infinitely many i and that Lemma 3.6 gives $\prod p_i(1) = xy$. In order to define p_{n+1} inductively, assume $p_n : [0, 1] \rightarrow X$ is defined as $p_n(t) = \sum_{i=1}^{\infty} p_{n,i}(t)b_i$ and define $u_n : [0, 1] \rightarrow X$ by

$$u_n(t) = (tp_{n,n}(1) - p_{n,n}(t)) \cdot b_n \equiv \beta_n(t) \cdot b_n.$$

Notice that $u_1(t) = u_2(t) = 0$ for each t and are included only for completeness. Define $p_{n+1} : [0, 1] \rightarrow X$ as

$$p_{n+1} = \sum(u_n \prod p_n).$$

If D is chosen small enough so that Theorem 1.13 applies, it follows that

$$\prod p_{n+1}(t) = u_n(t) \prod p_n(t) \prod p_n(0)^{-1} u(0)^{-1} = u_n(t) \prod p_n(t),$$

since each of $\prod p_n(0)$ and $u(0)$ is 0. Therefore $\prod p_{n+1}(1) = u(1) = \prod p_n(1)$, which by induction equals xy . Moreover, since $u_n = \prod p_{n+1}(\prod p_n)^{-1}$, it follows from Theorems 3.5 and 3.1, that if D is chosen small enough,

$$\begin{aligned} p'_{n+1} &= u'_n u_n^{-1} + u'_n \\ &= (p_{n,n}(1) - p'_{n,n})b_n + p'_{n,n} + [u_n, p'_{n,n}] + \frac{1}{2!}[u_n, [u_n, p'_{n,n}]] + \dots \\ &= (p_{n,n}(1) - p'_{n,n})b_n + \sum_{i=1}^{\infty} p'_{n,i} b_i + \left[\beta_n b_n, \sum_{i=1}^{\infty} p'_{n,i} b_i \right] + \frac{1}{2!} \dots \\ &= (p_{n,n}(1) - p'_{n,n})b_n + \sum_{i=1}^{\infty} p'_{n,i} b_i + \sum_{k=1}^{\infty} \sum_{i=1}^{\infty} p'_{n,i} \frac{\beta_n^k}{k!} [b_n, b_{i,k}], \end{aligned}$$

where $b_{i,1} = b_i$ and $b_{i,k+1} = [b_n, b_{i,k}]$.

Next, notice that each bracket in $p_1 = \sum_{i=1}^{\infty} p_{1,i} b_i$, which has a non-zero polynomial coefficient $p_{1,i}$, is of the form $[x, b_j] = [b_1, b_j]$ and hence has a first term of length 1. Recall the length of b_k is denoted by l_k . Now suppose n is a positive integer and each bracket in p_n with non-zero coefficient is of the form $[b_i, b_j]$, where $l_i \leq l_n$ and if $l_i = l_n$ then $b_i \neq b_n$. Then, ignoring simplifications resulting from bracket algebra such as $[x, y] = -[y, x]$ and the Jacobian identity, it follows that if $p_{n,m}$ is non-zero and $b_m = [b_i, b_j]$, then each bracket of the form $[b_n, b_{m,k}]$, as defined above for p'_{n+1} , has a first term of value b_n and hence of length l_n and therefore corresponds to a polynomial $p_{n,i}$ which is constantly 0.

We now define $q_{n+1} : [0, 1] \rightarrow X$ by $q_{n+1}(t) = \sum_{j=1}^{\infty} q_{n+1,j}(t)b_j$, where for each j , $q_{n+1,j}$ is a real valued polynomial defined on $[0, 1]$ as

$$q_{n+1,j}(t) = \begin{cases} tp_{n,n}(1), & \text{if } j = n \\ p_{n,j}(t), & \text{if } p_{n,j} \text{ is non-zero, } j \neq n \\ \int_0^t \frac{(\beta_n(s))^k}{k!} p'_{n,j}(s) ds, & \text{if } b_j = [b_n, b_{i,k}] \text{ for some } i \text{ and } k \\ 0, & \text{otherwise} \end{cases}$$

It follows from induction on n and the argument above that q_{n+1} is well defined, at least coordinate-wise, and that $q_{n+1,j}(0) = 0$ for each j . It will be shown below that each of $\sum_{i=1}^{\infty} \sup_{[0,1]} |q_{n,i} b_i|$, $\sum_{i=1}^{\infty} \sup_{[0,1]} |q'_{n,i} b_i|$, and $\sum_{i=1}^{\infty} \sup_{[0,1]} |q''_{n,i} b_i|$ is bounded for each n and hence each of q_n and q'_n exist. Moreover, this also implies that $q'_n = \sum_{i=1}^{\infty} q'_{n,i} b_i$.

Furthermore, since $q'_{n+1}(t) = p'_{n+1}(t)$ for each $t \in [0, 1]$ and $q_{n+1}(0) = p_{n+1}(0) = 0$, it follows that $p_{n+1} = q_{n+1}$, and hence that $\prod q_{n+1}(1) = \prod p_{n+1}(1) = xy$. In other words, even though p_{n+1} was defined earlier in terms of p_n and u_n to obtain $\prod p_{n+1}(1) = xy$, it follows that p_{n+1} is also given by the infinite series q_{n+1} . Notice, by induction on n , that the first n terms of q_{n+1} , and hence of p_{n+1} , are each of the form tz for some z in X .

We now show that $\sum_{i=1}^{\infty} \sup_{[0,1]} |q_{n,i} b_i|$ is bounded by showing that $\sum_{i=1}^{\infty} \text{Var}_{[0,1]}(q_{n,i} b_i)$ is bounded and in fact can be kept as small as desired, provided D is chosen accordingly. First, we define $v_i = \text{Var}_{[0,1]}(q_{2,i} b_i)$ and notice that if D is chosen small enough we have $\sum_{i=1}^{\infty} v_i < 1/20$, since $q_2 = p_2 = p_1$. Also $b_1 = x$ and $b_2 = y$ gives $q_3 = q_2$. However, since $b_3 = [x, y]$ and hence the coefficient of b_3 is non-zero, it follows that

$$\text{Var}_{[0,1]} q_4 = \sum_{i=1}^{\infty} \text{Var}_{[0,1]}(q_{4,i} b_i) = \sum_{i=1}^{\infty} v_i + |q_{3,3}(1) b_3| + \sum_{i=1}^{\infty} w_i \leq \sum_{i=1}^{\infty} v_i + \sum_{i=1}^{\infty} w_i,$$

where

$$w_i = \begin{cases} \text{Var}_{[0,1]}(q_{4,i} b_i), & \text{if } b_i = [b_3, b_j] \text{ for some } j \\ 0, & \text{otherwise.} \end{cases}$$

Therefore, using the fact that

$$\begin{aligned} \text{Var}_{[0,1]} \left(\int_0^1 \frac{(\beta_3(s))^k}{k!} p'_{3,i}(s) ds \cdot [b_3, b_{i,k}] \right) &\leq \int_0^1 (|\beta_3(s)|^k |p'_{3,i}(s)| ds) \cdot |[b_3, b_{i,k}]| \\ &\leq (\text{Var}_{[0,1]} \beta_3)^k |b_3|^k \cdot \text{Var}_{[0,1]} p_{3,i} \cdot |b_i| = (\text{Var}_{[0,1]} \beta_3 b_3)^k \cdot \text{Var}_{[0,1]} p_{3,i} b_i, \end{aligned}$$

and the fact that $\text{Var}_{[0,1]} \beta_3 < 2 \text{Var}_{[0,1]} p_{3,3}$, we have that

$$\begin{aligned} \text{Var}_{[0,1]} q_4 &\leq \sum_{i=1}^{\infty} v_i + \sum_{i=1}^{\infty} \text{Var}_{[0,1]} p_{3,i} b_i \cdot \sum_{k=1}^{\infty} (\text{Var}_{[0,1]} \beta_3 b_3)^k \\ &\leq \sum_{i=1}^{\infty} v_i + \frac{1}{10} \cdot 2 \text{Var}_{[0,1]} p_{3,3} b_3 \\ &\leq \sum_{i=1}^{\infty} v_i + \frac{1}{10} \cdot \frac{2}{1-2(1/10)} \text{Var}_{[0,1]} p_{3,3} b_3 \\ &= \sum_{i=1}^{\infty} v_i + \frac{1}{4} v_3. \end{aligned}$$

The reader should note that if w_i is non-zero, then b_i is not of the form $[x, b_j]$ and hence the coefficient of b_i in p_i is zero. Thus, the variation of these previously zero coefficients form and infinite series, namely $\sum_{i=1}^{\infty} w_i$, which sums to less than $\frac{1}{4}$ of a single previous summand, v_3 . Clearly, $\text{Var}_{[0,1]} q_4 < \frac{1}{20} + \frac{1}{4} \cdot \frac{1}{20} < \frac{1}{10}$. Subsequent steps in the construction yield similar results, where v_3 is replaced by either v_i for some $i > 3$, or w_i for some positive integer i , or some other variation introduced at an earlier stage while altering a term such as w_i . Thus, if n is a positive integer and $n-1$ new series have been introduced, each with their corresponding series of variations, it follows that

$$\text{Var}_{[0,1]} q_n \leq \sum_{i=1}^{\infty} v_i + \frac{1}{4} \sum_{i=1}^{\infty} v_i + \frac{1}{4^2} \sum_{i=1}^{\infty} v_i + \dots < \frac{1}{20} \sum_{i=1}^{\infty} \frac{1}{4^i} = \frac{1}{20} \cdot \frac{4}{3} < \frac{1}{10}.$$

Thus, just as before, we have

$$\text{Var}_{[0,1]q_{n+1}} \leq \sum_{i=1}^{\infty} \text{Var}_{[0,1]q_{n,i}b_i} + \sum_{i=1}^{\infty} u_i,$$

where here

$$u_i = \begin{cases} \text{Var}_{[0,1]q_{n,i}b_i}, & \text{if } b_i = [b_n, b_j] \text{ for some } j \\ 0, & \text{otherwise.} \end{cases}$$

Therefore, using a similar argument as before, we have

$$\begin{aligned} \text{Var}_{[0,1]q_{n+1}} &\leq \sum_{i=1}^{\infty} q_{n,i}b_i + \sum_{i=1}^{\infty} \text{Var}_{[0,1]p_{n,i}b_i} \cdot \sum_{k=1}^{\infty} (\text{Var}_{[0,1]\beta_n b_n})^k \\ &\leq \sum_{i=1}^{\infty} q_{n,i} + \frac{1}{4} \text{Var}_{[0,1]p_{n,n}b_n} \\ &\leq \sum_{i=1}^{\infty} \text{Var}_{[0,1]q_{n,i}b_i} + \frac{1}{4} \text{Var}_{[0,1]q_{n,n}b_n}. \end{aligned}$$

Thus, the variation of q_n is less than $\frac{1}{10}$ for each n and, except for showing the convergence of the various series and their derivatives, as promised above, we are done. However, in order to drive home the idea that the tails of the series must approach 0, suppose $\epsilon > 0$. Choose a positive integer N_1 such that $\sum_{k=0}^{\infty} \frac{1}{4^k} \sum_{i=N_1}^{\infty} v_i < \frac{\epsilon}{2}$. This implies that the sum of the variations of the series introduced to

alter the $v_i, i > N_1$, plus the series introduced to alter these, etc. sum to less than $\frac{\epsilon}{2}$. However, we still need to take care of the fact that each of the at most $N_1 - 1$ variation series introduced to alter bracket terms with index less than N_1 , such as the W series used to alter b_3 , contain infinitely many terms corresponding to brackets with index greater than N_1 . For each of these series, say W_2, W_3, \dots, W_{N_1} choose a positive integer N_j satisfying $\sum_{k=0}^{\infty} \frac{1}{4^k} \sum_{i=N_j}^{\infty} W_{j,i} < \frac{\epsilon}{2}$. Thus, if N is chosen greater than $\max\{N_1, N_2, \dots, N_{N_1}\}$ it follows that the sum of the variations of all series introduced after the N^{th} stage is less than ϵ .

We now show that $\sum_{i=1}^{\infty} \sup_{[0,1]} |q'_{n+1,i} b_i|$ can be made arbitrarily small by an appropriate choice of D . But this argument is very similar to the one above.

For if

$$q_{n+1,i}(t) = \int_0^t \frac{(\beta_n(s))^k}{k!} p'_{n,i}(s) ds$$

as defined above, and hence $b_j = [b_n, b_{i,k}]$, then

$$\sup_{[0,1]} |q'_{n+1,i} b_i| \leq (\text{Var}_{[0,1]\beta_n b_n})^k \cdot \sup_{[0,1]} |q'_{n,i} b_i|.$$

Thus, for example, if we denote $\sup_{[0,1]} |q'_{3,i} b_i|$ as v'_i we have that

$$\begin{aligned} \sum_{i=1}^{\infty} \sup_{[0,1]} |q'_{4,i} b_i| &\leq \sum_{i=1}^{\infty} v'_i + |p_{3,3}(1)b_3| + \sum_{k=1}^{\infty} (\text{Var}_{[0,1]\beta_3 b_3})^k \cdot \sum_{i=1}^{\infty} |p'_{3,i} b_i| \\ &\leq \sum_{i=1}^{\infty} v'_i + v_3 + \frac{1}{10} \cdot \frac{2v_3}{1 - 2(1/10)} = \sum_{i=1}^{\infty} v'_i + \frac{5}{4} v_3. \end{aligned}$$

Therefore, if we again use the property that $\sum_{i=1}^{\infty} \sup_{|0,1|} |p'_{n,i} b_i| \leq \frac{1}{10}$ as our induction step we see that $\sum_{i=1}^{\infty} \sup_{|0,1|} |q'_{n,i} b_i| < \sum_{i=1}^{\infty} v'_i + \frac{\epsilon}{4} \sum_{i=1}^{\infty} \text{Var}_{[0,1]} q_{n,i} b_i$. But we've already shown this last sum can be made as small as desired, since $1/10$ may clearly be replaced by $(1/10)\epsilon$; hence we are done.

The property that $\sum_{i=1}^{\infty} \sup_{|0,1|} |q''_{n+1,i} b_i|$ is bounded for each n follows from the fact that

$$q''_{n+1,i} b_i = \frac{k(\beta_n)^{k-1}}{k!} \beta'_n p'_{n,i} [b_n, b_{i,k}] + \frac{(\beta_n)^k}{k!} p''_{n,i} [b_n, b_{i,k}].$$

Thus, since $\sup_{|0,1|} |\beta'_n| \leq |p_{n,n}(1)| + \sup_{|0,1|} |p'_{n,n}| = M_n$, we have

$$\begin{aligned} \sum_{i=1}^{\infty} \sup_{|0,1|} |q''_{n+1,i} b_i| &\leq \sum_{i=1}^{\infty} \sup_{|0,1|} |q''_{n,i} b_i| + M_n \cdot \sum_{k=1}^{\infty} \sum_{i=1}^{\infty} k (\text{Var}_{[0,1]} \beta_n b_n)^{k-1} \cdot \sup_{|0,1|} |p'_{n,i} b_i| \\ &\quad + (\text{Var}_{[0,1]} \beta_n b_n)^k \cdot \sup_{|0,1|} |p''_{n,i} b_i|. \end{aligned}$$

But the first summand of the double sum is bounded due to the previous arguments for $\sum_{i=1}^{\infty} \text{Var}_{[0,1]} q_{n,i} b_i$ and $\sum_{i=1}^{\infty} \sup_{|0,1|} |q'_{n,i} b_i|$ and the fact that $\sum_{k=1}^{\infty} k \alpha^{k-1} < \infty$ for $|\alpha| < 1$. Furthermore, the second summand is also bounded due to the previous arguments, and by induction as is the first sum. Thus, $\sum_{i=1}^{\infty} \sup_{|0,1|} |q''_{n,i} b_i|$ is bounded for each n .

Therefore, we may now conclude that it is possible to choose a $D > 0$ small enough so that Theorem 1.13 holds for u_n [Theorem 3.5 holds for p_{n+1} and p'_n , and Theorem 3.1 holds for u_n and p'_n , for each n .

REFERENCES

- [1] BIRKHOFF, G., Analytical groups, *Trans. Amer. Math. Soc.* **43** (1938), 61-101.
- [2] PRICE, J. F., Lie groups and compact groups, Cambridge University Press, Cambridge, England and London, 1977.
- [3] SCHUR, F., Neue begründung der theorie der endlichen transformations Gruppen, *Mathematische Annalen* **35** (1889), 161-197.
- [4] MONTGOMERY, D. and ZIPPIN, L., Topological transformation groups, Interscience, New York, London, 1955.
- [5] PONTRJAGIN, L. S., Topological groups, Princeton University Press, 1939.
- [6] VON NEUMANN, J., Über die analytischen Eigenschaften von Gruppen linearer transformationen und ihrer darsellungen, *Mathematische Zeitschrift* **30** (1929), 3-42.
- [7] GLEASON, A., Groups without small subgroups, *Ann. of Math.* **56** (1952), 193-212.
- [8] MONTGOMERY, D. and ZIPPIN, L., Small subgroups of finite-dimensional groups, *Ann. of Math.* **56** (1952), 213-41.
- [9] JACOBY, R., Some theorems on the structure of locally compact local groups, *Ann. of Math.* **66** (1957), 36-69.
- [10] MOSTERT, P. S. and SHIELDS, A. L., Semigroups with identity on a manifold, *Trans. Amer. Math. Soc.* **91** (1959), 380-389.
- [11] CAMPBELL, J. E., On a law of operators, *Proc. Lon. Math. Soc.* **28** (1897), 381-90.
- [12] BAKER, H. F., Alternants and continuous groups, *Proc. Lon. Math. Soc.* **3** (1905), 24-47.
- [13] HAUSDORFF, F., Die symbolische exponentialformel in der Gruppen-theorie, *Berichte der Schsischen Akad. Wiss.*, Leipzig **58** (1906), 19-48.