Università degli studi di Padova

DIPARTIMENTO DI MATEMATICA "TULLIO LEVI-CIVITA"

Corso di Laurea Magistrale in Matematica

New perspectives in Baker-Campbell-Hausdorff problem and allied topics

Relatore:
Prof. Franco Cardin

Laureanda:
Laura Ballarini, 1159800

14 Dicembre 2018

Anno Accademico 2017-2018

Contents

Introduction

At the end of 19th century the need of studying the formal identity $e^X e^Y = e^Z$ in an autonomous way in relation to the theory of Lie groups became prominent. One of the first scientists who studied the problem was J. E. Campbell (cf. [12]); in 1897 he solved the question of existence of Z which satisfies the identity $e^X e^Y = e^Z$, but his paper, written in a very concise style, lacks clarity. Later the problem was studied by a lot of scientists, like J. H. Poincaré and E. Pascal. On the dawning days of the 20th century, the question was solved by H. F. Baker and F. Hausdorff, who independently published their articles [4, 19], in which there were not any references to theory of Lie groups: in fact they used analytic techniques. In this way it is proved the relation between the exponential of X, Y, non commutative indeterminates, described by

$$e^X e^Y = e^Z \tag{1}$$

which is usually called the Baker-Campbell-Hausdorff theorem, or BCH formula.

An important contribution was given later by E. B. Dynkin, for this reason the identity (1) was also called the Baker-Campbell-Hausdorff-Dynkin theorem, or BCHD formula. Dynkin gave the first estimate of the convergence domain and he studied how to write the representation of Z explicitly, using only the commutators of X and Y (cf. [14])

$$Z = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}([X, [X, Y]] - [Y, [X, Y]]) + \dots$$

The importance of this formula lies in the many applications of it, both in mathematics and physics; in fact it is used in various field, for example in the theory of Lie groups and Lie algebras, in linear partial differential equations, in quantum mechanics, in numerical analysis, and more.

The aim of this thesis is to study the using of the BCH formula in the dynamical systems; hence we want to understand how the formula behaves if we use the vector fields as non commutative indeterminates. The idea of facing this question derives from the fact that e^X represents the flow at time 1 of vector field X, so the relation, expressed by BCH formula, leads us to think that the product of the flows at time 1 of the vector fields X and Y is the flow at time 1 of a vector field Z. The topic is not so easy: it is necessary to deal with an infinite-dimensional Lie group, and this causes some problems because, as compared to finite-dimensional case, Lie groups of infinite dimension lose some important properties. The literature, in generally, debates the application of BCH formula to matrix finite-dimensional case, instead few information are given for infinite dimension Lie groups. Hence we would like to clarify the application of it on the particular example of vector fields and flows.

The main question, debated in this thesis, about BCH formula is linked to the more complex and deep problem of finding, given a diffeomorphism, the vector field whose flow is exactly the diffeomorphism. In fact, if it is always true that the flow of a vector field is a diffeomorphism, the inverse statement is certainly true only under specific hypothesis; for example we are able to construct this correspondence using perturbation theory, such as in [6, 21]. Hence we would like to construct this correspondence without using the perturbation theory.

The thesis is structured as following.

Chapter 1 contains a short review of background knowledge about smooth manifolds, Lie groups and Lie algebras.

In *Chapter 2* we work on the BCH formula. Starting from its proof, we focus our attention on the formal series

$$Zf(x) = \log(e^X e^Y)f(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}(e^X e^Y - \mathbb{I})^n f(x) \tag{2}$$

which formally allows us to determine the vector field Z whose flow at time 1 is the product of the flows at time 1 of the vector fields X, Y, in particular we have to find its domain of convergence. This is precisely what M. Postnikov does in [29, Lecture 4], but in his treatise he estimates the above series supposing the smallness of X and Y with the operator norm

$$\|XYf\| \leq |||XY|||\,\|f\| \tag{3}$$

for any f analytic in a suitable domain. But this is based on the wrong hypothesis that the vector fields are continuous operators. Our attempt is to find a new suitable norm such that it satisfies a relation like (3), but unfortunately the series that we obtain does not converge.

Another solution, which is given by S. Biagi and A. Bonfiglioli, is to reduce

the problem to finite dimension, so we take the Lie subalgebra of vector fields, and we can find a norm defined on it which satisfies also the sub-multiplicative property, like (3). With these hypotheses we can find the domain of convergence of (2). This reduction is de facto also a reduction to integrable system, in which the BCH formula has an easier formulation.
In *Chapter 3* we briefly face the link between diffeomorphisms and vector fields. We start presenting two results using the perturbation theory, then we try to construct an exactly correspondence between vector fields and diffeomorphisms. Working with some estimates linked to BCH formula, we are able to construct a series which describes a vector field X whose flow generates a diffeomorphism ψ. This series that we obtain is not convergent but the first terms decrease, so this is sufficient to say that a series is "convergent for astronomers". As Poincaré writes in [27, Chapter VIII], between geometers and astronomers there is a misunderstanding about the meaning of the word *convergence*; the first ones say that a series is convergent if its partial sums tend to a finite limit, on the other hand, the astronomers maintain that a series is convergent if the first, for example, twenty terms decrease very rapidly, even if the other terms increase indefinitely.

Chapter 1

Background knowledge

In this chapter we want to recall some basic concepts about differentiable manifolds and Lie groups. So we briefly summarize the main notions, recalling only some proofs. For more information the reader could see the following references: [2, 22] for manifolds and [31, 33] for Lie groups and Lie algebras. This part is inspired especially by the book of B. C. Hall [18, Appendix C].

1.1 Manifolds

1.1.1 Definitions

A *topological manifold* M of dimension n is a topological space that is locally homeomorphic to $\mathbb{R}^n$. This means that, for each point $m \in M$, there are a neighbourhood U of m and a continuous map $\phi\colon U \to \mathbb{R}^n$, where $\phi(U)$ is an open set of $\mathbb{R}^n$, such that the inverse map $\phi^{-1}\colon \phi(U) \to U$ is also continuous. From this map ϕ we can define local coordinate functions $x_1, \dots, x_n$, where each x_k is the continuous function from U to $\mathbb{R}$ given by $x_k(m) = \phi(m)_k$, which is the k-th component of $\phi(m)$. If ψ is another homeomorphism of another neighbourhood V of m and $y_k(m) = \psi(m)_k$ is the associated coordinate system, then both coordinate systems are defined in the neighbourhood $U \cap V$ of m. The link between these coordinate systems is given by the map $\psi \circ \phi^{-1}$; it maps the set $\phi(U \cap V)$ onto the set $\psi(U \cap V)$ and it is the change coordinates map

$$(y_1(m), \dots, y_n(m)) = (\psi \circ \phi^{-1})(x_1(m), \dots, x_n(m))$$

This change coordinates map is obviously continuous, since both ψ and ϕ^{-1} are continuous.

Definition 1.1. (Smooth manifold)
A *smooth manifold* of dimension n is a topological manifold M together with a distinguished family of local coordinate systems $(U_\alpha, \phi_\alpha)_{\alpha \in \mathcal{A}}$ with the following properties:

1. every point in M is contained in at least one of the U_α's;

2. for any two of these coordinate systems (U_α, ϕ_α) and (U_β, ϕ_β), the change coordinates map $\phi_\beta \circ \phi_\alpha^{-1}$ is a smooth map from the set $\phi_\alpha(U_\alpha \cap U_\beta) \subset \mathbb{R}^n$ onto the set $\phi_\beta(U_\alpha \cap U_\beta) \subset \mathbb{R}^n$.

Now we can introduce the functions and the maps on smooth manifolds.
A function $f\colon M \to \mathbb{R}$ is said smooth if in each local coordinate system (U, ϕ), $f \circ \phi^{-1}$ is a smooth function.
A map $f\colon M \to N$, where M, N are smooth manifolds of dimension m, n, respectively, is said smooth if it is smooth in local coordinates; so, if ϕ_α is a local coordinate system on M and ϕ_β is a local coordinate system on N, then $\phi_\beta \circ f \circ \phi_\alpha^{-1}$ is a smooth map from an open subset of $\mathbb{R}^m$ onto $\mathbb{R}^n$.

1.1.2 The tangent space

There are more than one construction, both intrinsic and extrinsic, for the tangent space, here we see one of the first type. A possible definition for tangent space of a general manifold, not necessarily embedded in $\mathbb{R}^n$, is given using the derivation

Definition 1.2. (Derivation)
Given an algebra A over a field $\mathbb{K}$, a $\mathbb{K}$-*derivation* is a $\mathbb{K}$-linear map $D\colon A \to \mathbb{K}$ satisfying the following properties:

- $D(k) = 0$, for all $k \in \mathbb{K}$;

- product rule, $D(ab) = aD(b) + bD(a)$, for all $a, b \in A$.

Now we can give the definition of tangent space

Definition 1.3. (Tangent space)
The *tangent space* at m to M, denoted by T_mM, is the set of all derivations $C^\infty \to \mathbb{R}$ in m.

An element $X \in T_mM$ is said *tangent vector* at m if $X\colon C^\infty(M) \to \mathbb{R}$ is a linear map with the properties of Definition 1.3. This means that if M is a manifold of dimension n, then, for each $m \in M$, T_mM is a real vector space of dimension n. Moreover, since $(\frac{\partial}{\partial x^1}|_m, \ldots, \frac{\partial}{\partial x^n}|_m)$ is a basis for T_mM, for each $X \in T_mM$ we have

$$X = \sum_{i=1}^{n} X(x^i) \frac{\partial}{\partial x^i}\bigg|_m$$

where $(x^1, \ldots, x^n)$ is a local coordinate system in a neighbourhood of m.

Let $F\colon M \to N$ be a map between smooth manifolds M, N, we could define

Definition 1.4. (Differential)
If F is a map between manifolds as above, given $m \in M$, the *differential* of F in m, $dF_m\colon T_mM \to T_{F(m)}N$, is a linear map defined as

$$dF_m(X)(g) = X(g \circ F) \qquad \forall\ X \in T_mM$$

for all $g \in C^\infty(F(m))$.

Definition 1.5. (Pull-back)
If F is a map between manifolds as above, given $m \in M$, we say that $F_m^*\colon C^\infty(F(m)) \to C^\infty(m)$ defined as

$$F_m^*(g) = g \circ F$$

is the *pull-back map* .

Remark 1.6. It is possible to use the definition of pull-back map to rewrite, in equivalent way, the Definition 1.4, so the differential is the linear map such that

$$dF_m(X) = X \circ F_m^*$$

1.1.3 Vector fields and flows

Given a smooth manifold M of dimension n, the disjoint union of all tangent spaces T_mM, $m \in M$, is a smooth manifold of dimension $2n$ and it is called *tangent bundle*

$$TM = \bigcup_{m \in M} T_mM$$

It is defined the natural projection

$$\pi\colon TM \to M$$

which is the map whose fibers are precisely the tangent spaces, $T_mM = \pi^{-1}(m)$.

Definition 1.7. (Vector field)
A smooth *vector field* on M is a smooth map $X\colon M \to TM$ such that

$$X(m) \in T_m M$$

for all $m \in M$. In other words, a vector field X on M is a smooth section of the tangent bundle, that is $\pi \circ X = \mathrm{id}$, where $\pi\colon TM \to M$.

The collection of all vector fields on M is a vector space of infinite dimension and it is denoted by $\mathfrak{X}(M)$.
Given any smooth map $F\colon M \to N$, its differential $dF_m\colon T_m M \to T_{F(m)}N$ defines a smooth map $TF\colon TM \to TN$ such that the following diagram is commutative

$$\begin{array}{ccc} TM & \xrightarrow{TF} & TN \\ \Big\downarrow{\scriptstyle\pi} & & \Big\downarrow{\scriptstyle\pi} \\ M & \xrightarrow{F} & N \end{array}$$

Given a local coordinate system, a vector field X could be written as

$$X(m) = \sum_{k=1}^{n} X^k(m) \frac{\partial}{\partial x^k}\bigg|_m$$

where X^k are real functions. A vector field is said smooth if the functions X^k are smooth in each coordinate system.

A different way to see the vector fields is as derivation of the algebra of smooth functions:

$$X\colon C^\infty(M) \to C^\infty(M)$$

is linear and satisfy the product rule

$$X(gf) = fX(g) + X(f)g$$

Between the two interpretation there exists an isomorphism: the vector space $\mathfrak{X}(M)$ of all vector fields on the manifold M is isomorphic to the space of derivations of $\mathbb{R}$-algebra C^∞ (for the proof see [2, Proposition 3.3.2]).
It is quite clear that there is a link between ordinary differential equations and vector fields; now we have to rewrite the existence and uniqueness theorem for ODE on the manifold, but before we have to introduce some objects.

Definition 1.8. (Integral curve)
Let $X \in \mathfrak{X}(M)$ be a vector field on the manifold M and $m \in M$. A curve

$\sigma\colon I \to M$, where $I \subseteq \mathbb{R}$ is an open interval containing 0, such that $\sigma(0) = m$ and

$$\sigma'(t) = X(\sigma(t))$$

for all $t \in I$, is said *integral curve* of X starting at m.

If we consider a local coordinate system in a point $m \in M$, a vector field $X \in \mathfrak{X}(M)$ could be written as $X = \sum_i X^i \frac{\partial}{\partial x^i}$. If $\sigma\colon (-\varepsilon, \varepsilon) \to M$ is a curve starting at m, with $\sigma(0) = m$, we can say that σ is integral curve of X if and only if it solves the Cauchy problem

$$\begin{cases} \frac{d}{dt}\sigma^i = X^i(\phi \circ \sigma(t)) & i = 1, \dots, n \\ \phi \circ \sigma(0) = 0 \end{cases}$$

where $\phi \circ \sigma = (\sigma^1, \dots, \sigma^n)$.
It is true the following result

Theorem 1.9. *Let $X \in \mathfrak{X}(M)$ be a vector field on manifold M. Then there exist an unique open neighbourhood U of $\{0\} \times M$ in $\mathbb{R} \times M$ and an unique smooth map $\Phi\colon U \to M$ satisfying the following properties:*

i) for all $m \in M$ the set $U^m = \{t \in \mathbb{R} | (t, m) \in U\}$ is an open interval containing 0;

ii) for all $m \in M$ the curve $\phi^m\colon U^m \to M$ defined by $\phi^m(t) = \Theta(t, m)$ is the unique maximal integral curve of X starting at m;

iii) for all $t \in \mathbb{R}$ the set $U_t = \{m \in M | (t, m) \in U\}$ is open in M;

iv) if $m \in U_t$, then $m \in U_{t+s}$ if and only if $\Phi(t, m) \in U_s$, moreover, in this case

$$\phi_s(\phi_t(m)) = \phi_{s+t}(m)$$

where $\phi_t\colon U_t \to M$ is defined by $\phi_t(m) = \Phi(t, m)$. In particular, $\phi_0 = id$ e $\phi_t\colon U_t \to U_{-t}$ is a diffeomorphism with inverse map ϕ_{-t};

v) for all $(t, m) \in U$ it is

$$d(\phi_t)_m(X) = X_{\phi_t(m)}$$

vi) for all $f \in C^\infty(M)$ and $m \in M$ it is

$$\frac{d}{dt}(f \circ \phi^m)\bigg|_{t=0} = (Xf)(m)$$

The proof of this result is omitted, but the reader could see it in [2, Theorem 3.3.5].

Definition 1.10. (Local flow)
The map $\phi_X^t \colon U \to M$ introduced in the above theorem is said *local flow* of the vector field $X \in \mathfrak{X}(M)$.
The vector field X is said *complete* if all the integral curves of X are defined for any time.

Let X be a vector field on a manifold M, we can associate a new operator

Definition 1.11. (Lie derivative)
It is defined the *Lie derivative* as the operator $\mathcal{L}_X \colon C^\infty(M) \to C^\infty(M)$ such that

$$\mathcal{L}_X f := \frac{d}{dt}(f \circ \phi_X^t)\bigg|_{t=0}$$

for any $f \in C^\infty(M)$.

Remark 1.12. An equivalent way to write the above definition is using the pull-back

$$\mathcal{L}_X f = \frac{d}{dt}(\phi_X^t)^* f\bigg|_{t=0}$$

It is true the equality

$$\mathcal{L}_X f = Xf$$

The Lie derivative $\mathcal{L}_X$ measures how the function f moves respect to the integral curve of X at $t = 0$, but it could be calculated at any t.
Moreover, it holds the following proposition

Proposition 1.13. *Let X be a vector field on a manifold M, it holds*

$$\frac{d}{dt}(\phi_X^t)^* f = (\phi_X^t)^* \mathcal{L}_X f$$

for any t.

Proof.

$$\begin{aligned}
\frac{d}{dt}(\phi_X^t)^* f\bigg|_{t=\bar{t}} &= \frac{d}{dt} f \circ \phi_X^t\bigg|_{t=\bar{t}} = \frac{d}{ds} f \circ \phi_X^{\bar{t}+s}\bigg|_{s=0} \\
&= \frac{d}{dt} f \circ \phi_X^s \circ \phi_X^{\bar{t}}\bigg|_{s=0} = \frac{d}{dt} f \circ \phi_X^s\bigg|_{s=0} \phi_X^{\bar{t}} \\
&= \mathcal{L}_X f \circ \phi_X^{\bar{t}} = (\phi_X^{\bar{t}})^* \mathcal{L}_X f
\end{aligned}$$

□

The interpretation of vector field as derivation of algebra of smooth function allows us to introduce another operation

Definition 1.14. (Lie bracket)
The *Lie bracket* between two vector fields X, Y is the vector field $[X, Y] = XY - YX$ defined by

$$[X, Y](f) = X(Yf) - Y(Xf)$$

for all $f \in C^\infty(M)$.
In particular, we say that two vector fields commute if $[X, Y] = 0$.

Remark 1.15. The commutator of vector fields, as defined in 1.14, is also a vector field because it is a derivation. The proof of this fact is the easy calculation of $[X, Y](fg)$.

The most important properties for Lie bracket are resumed in the following proposition:

Proposition 1.16. *If X, Y, Z are vector fields on the manifold M, $a, b \in \mathbb{R}$ and $f, g \in C^\infty(M)$, it is true*

1. $[X, Y] = -[Y, X]$;
2. $[aX + bY, Z] = a[X, Z] + b[Y, Z]$;
3. $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$;
4. $[fX, gY] = fg[X, Y] + f(Xg)Y - g(Yf)X$.

The proof of this proposition is the application of the Definition 1.14.

1.2 Lie groups and Lie algebras

1.2.1 Lie groups

A Lie group is a smooth manifold that is also a group. More precisely, the definition is the following

Definition 1.17. (Lie group)
A *Lie group* is a smooth manifold G together with a smooth map

$$G \times G \to G \qquad (a, b) \mapsto ab$$

that makes G a group and such that the inverse map

$$a \mapsto a^{-1}$$

is a smooth map of G to itself.

Example 1.18. The simplest example of Lie group is $G = \mathbb{R}^n$ with the product map given by $(x, y) \mapsto x + y$.
A more interesting examples are the matrix Lie groups, such as $GL(n), SO(n), \ldots$.

1.2.2 Lie algebras

If G is a Lie group and g is its element, we want to define the map

$$L_g \colon G \to G \quad \text{such that} \quad L_g(h) = gh \tag{1.19}$$

called *left multiplication by g*. Similarly it is possible to define the *right multiplication by g* as

$$R_g \colon G \to G \quad \text{such that} \quad R_g(h) = hg \tag{1.20}$$

Since the product defined on G is a smooth map, both maps L_g and R_g are smooth.

Definition 1.21. (Left-invariant vector field)
A vector field X on a Lie group G is said *left-invariant* if $d_e L_g(X) = X$ for all $g \in G$ or, equivalently, if

$$d_h(L_g)(X_h) = X_{gh}$$

for all $g, h \in G$.

Remark 1.22. In the same way, it is possible to define the *right-invariant* vector field X, such as

$$d_h(R_g)(X_h) = X_{hg} \text{ or } d_e R_g(X) = X$$

for all $g, h \in G$.
So, all the results, which we will prove for left-invariant vector fields, are true also for right-invariant vector fields.

Lemma 1.23. *Let G be a Lie group of identity element $e \in G$, then*

1. *the map $X \mapsto X(e)$ is an isomorphism between the subset of $\mathfrak{X}(G)$ of left-invariant vector field and the tangent space T_eG;*

2. *if $X, Y \in \mathfrak{X}(G)$ are left-invariant, then also $[X, Y]$ is left-invariant.*

Proof. 1. If $X \in \mathfrak{X}(G)$ is left-invariant, then for definition

$$d_e(L_g)(X_e) = X_{ge} = X_g$$

for all $g \in G$, so X is completely determined by its value in e.
On the other side, if we choose $v \in T_eG$ and we define $X \in \mathfrak{X}(G)$ as

$$X_g = d_e(L_g)(v) \in T_gG$$

for all $g \in G$, we have a left-invariant vector field because

$$d_h(L_g)(X_h) = d_h(L_g)d_e(L_h)(v) = d_e(L_{gh})(v) = X_{gh}$$

and, moreover, its value at the identity element is v.

2. If X, Y are two left-invariant vector fields then it is easy to demonstrate that[1]

$$dL_g[X,Y] = [dL_gX, dL_gY] = [X,Y]$$

for all $g \in G$, therefore also $[X,Y]$ is left-invariant.

□

Definition 1.24. (Lie algebra)
A vector space V together with the further operation $[\cdot,\cdot]\colon V \times V \to V$ satisfying the following properties

a) anti-symmetry, $[v,w] = -[w,v]$;

b) bilinearity, $[au + bv, w] = a[u,w] + b[v,w]$;

c) Jacobi identity, $[u,[v,w]] + [v,[w,u]] + [w,[u,v]] = 0$;

is said *Lie algebra*.

If V, W are Lie algebras, a *morphism of Lie algebras* is a linear map $L\colon V \to W$ such that $[L(v_1), L(v_2)] = L[v_1, v_2]$ for all $v_1, v_2 \in V$.

Example 1.25. Let A be a non-commutative algebra on the field $\mathbb{K}$. Then we can put on A a structure of Lie algebra with the commutator $[\cdot,\cdot]\colon A \times A \to A$ defined as

$$\forall\, X, Y \in A \qquad [X,Y] = XY - YX$$

It is easy to show that the Lie bracket satisfies the properties of Definition 1.24.

[1] To demonstrate it we use that if $F\colon M \to M$ is a diffeomorphism and $X, Y \in \mathfrak{X}(M)$ then it holds $[dF(X), dF(Y)] = dF[X,Y]$ (see [2, Lemma 3.4.9]).

The space of all vector fields $\mathfrak{X}(M)$ on a manifold M with the Lie bracket is a Lie algebra.

Definition 1.26. (Lie Algebra $\mathfrak{g}$)
Let G be a Lie group of identity element $e \in G$. For all $v \in T_eG$, we write $X^v \in \mathfrak{X}(G)$ as the unique left-invariant vector field such that $X^v(e) = v$. Then the tangent space at the identity element with the structure of vector space and the operation $[\cdot,\cdot]\colon T_eG \times T_eG \to T_eG$ defined by $[v,w] = [X^v, X^w](e)$, is said *Lie algebra* $\mathfrak{g}$ of Lie group G.

We see very briefly the relation between two Lie groups and their two Lie algebras.

Lemma 1.27. *Let G, H be Lie groups of Lie algebras $\mathfrak{g}, \mathfrak{h}$, respectively, and let $F\colon G \to H$ be an homomorphism of Lie groups. Then $d_eF\colon \mathfrak{g} \to \mathfrak{h}$ is a morphism of Lie algebras.*

1.2.3 Lie subgroups and exponential map

Now we study briefly the structure of a particular Lie subgroup:

Definition 1.28. (One-parameter subgroup)
Let G be a connected Lie group. An *one-parameter subgroup* of G is the smooth map $\beta\colon \mathbb{R} \to G$ which is also an homomorphism of groups. In other words, we say that $\beta(0) = e$ is the identity of G and $\beta(t+s) = \beta(t)\beta(s)$ for all $s, t \in \mathbb{R}$.

The integral curve of left-invariant vector fields are one-parameter subgroups.

Lemma 1.29. *Let G be a Lie group of Lie algebra $\mathfrak{g}$. Given $X \in \mathfrak{g}$, let $\widetilde{X} \in \mathfrak{X}(G)$ be the left-invariant vector field associated to X. Then:*

1. *the integral curve of $\widetilde{X}$ starting at e is an one-parameter subgroup of G;*

2. *on the other side, if $\beta\colon \mathbb{R} \to G$ is an one-parameter subgroup with $\beta'(0) = X$, then β is the integral curve of $\widetilde{X}$ starting at e.*

Proof. 1. Let $\sigma\colon (-\varepsilon, \varepsilon) \to G$ be the maximal integral curve of $\widetilde{X}$ starting at e. We want to show that for all $t_0 \in (-\varepsilon, \varepsilon)$, the curve $\gamma\colon (-\varepsilon,\varepsilon) \to G$ defined as $\gamma(t) := \sigma(t_0)\sigma(t)$ is an integral curve of $\widetilde{X}$ starting at $\sigma(t_0)$; in fact we have

$$\gamma'(t) = d_{\sigma(t)}(L_{\sigma(t_0)})(\sigma'(t)) = d_{\sigma(t)}(L_{\sigma(t_0)})(\widetilde{X}(\sigma(t)) = \widetilde{X}(\gamma(t))$$

For the uniqueness of integral curve we have $\gamma(t) = \sigma(t_0 + t)$, therefore

$$\sigma(t_0 + t) = \sigma(t_0)\sigma(t)$$

for all $t_0, t \in (-\varepsilon, \varepsilon)$. It means that ε must be infinite and σ must be one-parameter subgroup.

2. We suppose that β is an one-parameter subgroup with $\beta'(0) = X$. Then $\beta(t_0 + t) = L_{\beta(t_0)}\beta(t)$, so

$$\beta'(t_0) = \frac{d}{dt}(L_{\beta(t_0)}\beta(t))\Big|_{t=0} = d_e L_{\beta(t_0)}(\beta'(0)) = d_e L_{\beta(t_0)}(X) = \widetilde{X}(\beta(t_0))$$

therefore β is the integral curve of $\widetilde{X}$ starting at e.

□

In particular for all $X \in \mathfrak{g}$ there exists an unique one-parameter subgroup $\beta_X \colon \mathbb{R} \to G$ such that $\beta_X'(0) = X$, which is the integral curve of $\widetilde{X}$ starting at e.

Proposition 1.30. *If G is a Lie group, then every left-invariant vector field on G is complete.*

Proof. Let $\mathfrak{g}$ be the Lie algebra, given $X \in \mathfrak{g}$, there exists a maximal integral curve $\gamma\colon (a, b) \to G$ of X with $0 \in (a, b)$ and $\gamma(0) = e$; namely $\gamma'(t) = X(\gamma(t))$. Since

$$\frac{d}{dt}L_g(\gamma(t))\Big|_{t=t_0} = d_e L_g(X(\gamma(t_0))) = X(L_g(\gamma(t_0)))$$

we have that $L_g \circ \gamma$ is an integral curve of X starting at g. In particular, if $b < \infty$, by taking $g = \gamma(s)$ with s very close to b, this shows that γ can be extended beyond b, leading by contraction. Similarly, one sees that $a = -\infty$. Hence X is complete. □

This result justifies the following definition

Definition 1.31. (Exponential map)
Let G be a Lie group. Given $X \in \mathfrak{g}$, the integral curve $\beta_X \colon \mathbb{R} \to G$ of the left-invariant vector field $\widetilde{X}$ starting at e is said one-parameter subgroup of X. The *exponential map* of G is the map

$$\exp\colon \mathfrak{g} \to G \quad \text{such that} \quad \exp(X) = \beta_X(1)$$

Remark 1.32. If γ_X is the integral curve of any $X \in \mathfrak{g}$, we note that

$$\frac{d}{ds}\gamma_X(ts)\bigg|_{s=s_0} = t\gamma'_X(ts_0) = tX(\gamma_X(ts_0)$$

which implies $\gamma_X(ts) = \gamma_{tX}(s)$. Therefore

$$\begin{aligned}\gamma_X(t) &= \gamma_{tX}(e)\\ &= \exp(tX)\end{aligned}$$

We resume in the following proposition the main properties of exponential map

Proposition 1.33. *Let G be the Lie group of the Lie algebra $\mathfrak{g}$, then:*

1. *the exponential map* $\exp\colon \mathfrak{g} \to G$ *is smooth;*
2. *the differential* $d_O \exp\colon \mathfrak{g} \to \mathfrak{g}$ *is the identity (where we have identify the tangent space to* $\mathfrak{g}$ *in* e *with* $\mathfrak{g}$ *itself);*
3. $\exp$ *is a diffeomorphism between a neighbourhood of* $O \in \mathfrak{g}$ *and a neighbourhood of* $e \in G$*;*
4. *if* $F\colon G \to H$ *is an homomorphism of Lie groups, then* $\exp \circ d_e F = F \circ \exp$*, so the following diagram is commutative*

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{d_e F} & \mathfrak{h} \\ \Big\downarrow \exp & & \Big\downarrow \exp \\ G & \xrightarrow{F} & H \end{array}$$

   where $\mathfrak{h}$ *is the Lie algebra of* H*;*

Proof. 1. For all $X \in \mathfrak{g}$, we write Φ_X for the flow of the left-invariant field $\widetilde{X}$. For the Definition 1.31 and the Theorem 1.9 we have that $\exp(X) = \Phi_X(1, e)$ so we have only to demonstrate that the map $X \mapsto \Phi_X(1, e)$ is smooth.
We introduce the vector field $\boldsymbol{X}$ on the product manifold $G \times \mathfrak{g}$ as

$$\boldsymbol{X}_{g,X} = (X_g, O) \in T_g G \oplus T_X \mathfrak{g} \cong T_{(g,X)}(G \times \mathfrak{g})$$

whose flow is

$$\boldsymbol{\Phi}(t, (g, X)) = (\Phi_X(t, g), X)$$

In particular, $\exp(X) = \pi_1(\Phi(1, (e, X))$, where $\pi_1\colon G \times \mathfrak{g} \to G$ is the projection on the first coordinate; therefore exp depends smoothly on X, because the flow is smooth for the Theorem 1.9.

2. Fixed $X \in \mathfrak{g}$, let $\sigma \colon \mathbb{R} \to \mathfrak{g}$ be the curve $\sigma(t) = tX$. Then $\sigma'(0) = X$ and

$$d_O \exp(X) = (\exp \circ \sigma)'(0) = \frac{d}{dt} \exp(\sigma(t))\bigg|_{t=0} = \frac{d}{dt} \exp(tX)\bigg|_{t=0} = X$$

where the last equivalence comes from the Proposition 1.29 and the Remark 1.32.

3. It follows from the above result and the inverse function theorem.

4. It is sufficient to demonstrate that $\sigma(t) = F(\exp(tX))$ is the one-parameter subgroup in H coming from $d_eF(X)$ for $X \in \mathfrak{g}$. We see that σ is one-parameter subgroup:

$$\begin{aligned}\sigma(s)\sigma(t) &= F(\exp(sX))F(\exp(tX)) = F(\exp(sX)\exp(tX)) \\ &= F(\exp((s+t)X)) = \sigma(s+t)\end{aligned}$$

where we have used that F is a group homomorphism and $t \mapsto \exp(tX)$ is an one-parameter subgroup. Moreover, $\sigma'(0) = d_eF(X)$ is deduced from the fact that $\exp(tX)$ is one-parameter subgroup.

□

We have already defined the exponential map (see 1.31)

$$\exp \colon \mathfrak{g} \to G$$

Now we would study the logarithm map

$$\log \colon G \to \mathfrak{g}$$

and it is quite natural to consider it as the inverse of exponential map, so we have to prove that.
First of all, we have to remind the formal power series

$$\exp(X) = \sum_{n=0}^{\infty} \frac{X^{[n]}}{n!} \tag{1.34}$$

where $X \in \mathfrak{g}$ and we use $[\cdot]$ to indicate the power of vector fields: $X^{[n]} = X(X^{[n-1]})$, and

$$\log(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}(x-1)^n \tag{1.35}$$

where $x \in G$.
Before we demonstrate our aim, we have to recall some results

Lemma 1.36. *Let $X \in \mathfrak{g}$, then*

$$\underbrace{\exp(X)\dots\exp(X)}_{k-times} = \exp(kX)$$

where $k \in \mathbb{N}$.

Proof. We prove it for $k = 2$

$$\begin{aligned}
\exp(X)\exp(X) &= \sum_{p,q=0}^{\infty} \frac{X^{[p]}X^{[q]}}{p!q!} = \sum_{p,q=0}^{\infty} \frac{X^{[p+q]}}{p!q!} \\
&= \sum_{q=0}^{\infty}\sum_{n=q}^{\infty} \frac{X^{[n]}}{(n-q)!q!} = \sum_{n=0}^{\infty} \frac{X^{[n]}}{n!} \sum_{q=0}^{n} \frac{n!}{(n-q)!q!} \\
&= \sum_{n=0}^{\infty} \frac{2^n X^{[n]}}{n!} = \exp(2X)
\end{aligned}$$

where we have defined $n = p + q$, so $p = n - q$. □

We recall some well-known identities (they could be demonstrated by induction)

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0 \tag{1.37}$$

$$\sum_{n=k}^{m} \binom{n}{k} = \binom{n+1}{k+1} \tag{1.38}$$

and we will use also this identity

Lemma 1.39. *For $m, n \in \mathbb{Z}$ with $0 \leq m < n$,*

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} k^m = 0$$

Proof. Firstly, we notice that

$$x\frac{d}{dx}x^k = kx^k$$

and we remind the well-known Newton binomial theorem

$$(x-1)^n = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} x^k \tag{1.40}$$

If we apply the operator $x\frac{d}{dx}$ to both side of (1.40) m-times, with $0 \leq m < n$, we have

$$(x-1)^{n-m}p(x) = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} k^m x^k$$

where $p(x)$ is a polynomial in x of degree m. Putting $x = 1$ in the above equation, the terms on left vanishing, so we obtain

$$\sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} k^m = 0$$

□

With this result, we are able to demonstrate the following proposition

Proposition 1.41. *The exponential map and logarithm map are inverse to each other.*

Proof. Let $X \in \mathfrak{g}$, we want to prove

$$\log(\exp(X)) = X$$

Using the formal series (1.34) and (1.35), and the Lemma 1.36

$$\begin{aligned}
\log(\exp(X)) &= \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} (\exp(X) - 1)^n \\
&= \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} (\exp(X))^k \\
&= \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} (\exp(kX)) \\
&= \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} \sum_{m=0}^{\infty} \frac{k^m X^{[m]}}{m!} \\
&= -\sum_{m=0}^{\infty} \frac{X^{[m]}}{m!} \sum_{n=1}^{\infty} \frac{1}{n} \sum_{k=0}^{n} \binom{n}{k} (-1)^k k^m \\
&= -\sum_{m=0}^{\infty} \frac{X^{[m]}}{m!} a_m
\end{aligned}$$

where a_m is defined as

$$a_m = \sum_{n=1}^{\infty} \frac{1}{n} \sum_{k=0}^{n} \binom{n}{k} (-1)^k k^m$$

Now we have to study the term a_m and we split the sum over n in two parts: the first from 1 to m and the second from $m+1$ to ∞, so we have

$$a_m = \sum_{n=1}^{m} \frac{1}{n} \sum_{k=0}^{n} \binom{n}{k} (-1)^k k^m + \sum_{n=m+1}^{\infty} \frac{1}{n} \sum_{k=0}^{n} \binom{n}{k} (-1)^k k^m$$

In the second part, the summation over k is zero by Lemma 1.39 because $n > m$. So

$$\begin{aligned} a_m &= \sum_{n=1}^{m} \frac{1}{n} \sum_{k=0}^{n} \binom{n}{k} (-1)^k k^m = \sum_{n=1}^{m} \sum_{k=0}^{n} \frac{1}{n} \binom{n}{k} (-1)^k k^m \\ &= \sum_{k=1}^{m} (-1)^k k^m \sum_{n=k}^{m} \frac{1}{k} \binom{n-1}{k-1} = \sum_{k=1}^{m} (-1)^k k^{m-1} \sum_{n=k}^{m} \binom{n-1}{k-1} \\ &= \sum_{k=1}^{m} (-1)^k k^{m-1} \binom{m}{k} \end{aligned}$$

where we used the identities (1.37) and (1.38). Now we would like to use the (1.37) but it requires the $k = 0$ terms, so we replace it by including $k = 0$ in the sum and subtracting it outside the sum

$$a_m = \sum_{k=0}^{m} (-1)^k k^{m-1} \binom{m}{k} - 0^{m-1}$$

The first term is zero for the Lemma 1.39, using the convention that $0^0 = 1$ we can conclude that

$$a_m = -\delta_{m,1}$$

where δ is the Kronecker delta. So

$$\log(\exp(X)) = -\sum_{m=0}^{\infty} \frac{X^{[m]}}{m!} a_m = X$$

□

Remark 1.42. (Note of the proof)
It is fundamental to remind that all the evaluations about the exponential and the logarithm map must be done only under hypothesis of convergence. This question of convergence and the use of the norm, which will be central in the next chapter when we will speak about BCH formula, are not simple, as it seems. The main problem is the existence of a suitable norm on the Lie algebra $\mathfrak{g}$, for which it looks like it will be necessary the hypothesis of $\mathfrak{g}$ Banach Lie algebra.

1.2.4 The adjoint representation of a Lie group

Let G be a Lie group. For every $g \in G$, we consider the inner automorphism of G

$$\Psi_g \colon G \to G \qquad h \mapsto ghg^{-1}$$

This is a smooth map from G in itself that fixes e. Then its differential at the point e is

$$d_e\Psi_g = Ad_g \colon \mathfrak{g} \to \mathfrak{g}$$

where we have identify T_eG as $\mathfrak{g}$.
In this way, since Ad_g is a Lie algebra automorphism, we can give the following definition

Definition 1.43. (Adjoint representation of G)
The linear representation $\mathrm{Ad}\colon G \to \mathrm{Aut}(\mathfrak{g})$ defined by the differential of Ψ_g

$$g \mapsto \mathrm{Ad}_g$$

is called the *adjoint representation of G*.

From the definition above it follows this lemma

Lemma 1.44. *Using the above notation, it is true that*

$$Ad(g)X = \frac{d}{dt}(g\exp(tX)g^{-1})\Big|_{t=0}$$

for $X \in \mathfrak{g}$.

Remark 1.45. Applying the fourth statement of Proposition 1.33 to the automorphism Ψ_g, it follows that

$$\exp(t\mathrm{Ad}(g)X) = \Psi_g(\exp(tX)) = g\exp(tX)g^{-1}$$

and, in particular, for $t = 1$,

$$g\exp(X)g^{-1} = \exp(\mathrm{Ad}(g)X)$$

The differential of the adjoint representation Ad is denoted ad,

$$\mathrm{ad}\colon \mathfrak{g} \to \mathrm{End}(\mathfrak{g}) \qquad X \mapsto d_e\mathrm{Ad}X \tag{1.46}$$

$$\mathrm{ad}(X)Y = \frac{d}{dt}(\mathrm{Ad}(\exp(tX))Y)\Big|_{t=0}$$

Another way to define it is with the commutator, i.e.

$$\mathrm{ad}(X)Y = [X, Y]$$

Once more, it follows from the forth statement of Proposition 1.33 that

$$\mathrm{Ad}(\exp(tX)) = \exp(t\mathrm{ad}(X))$$

that is, the following diagram is commutative

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\mathrm{ad}"] \arrow[d, "\exp"'] & \mathrm{End}(\mathfrak{g}) \arrow[d, "\exp"] \\ G \arrow[r, "\mathrm{Ad}"'] & \mathrm{Aut}(\mathfrak{g}) \end{tikzcd}$$

In particular, for $t = 1$

$$\mathrm{Ad}(\exp(X)) = \exp(\mathrm{ad}(X)) \tag{1.47}$$

1.2.5 Infinite dimensional Lie groups

In the following chapter we like to concentrate on dynamical systems, so the Lie group that we use is the group of smooth diffeomorphisms on a manifold M, $\mathrm{Diff}(M)$. Therefore it is necessary to speak briefly about the infinite-dimensional Lie groups; we will highlight the differences from the finite-dimensional case. This topic is not debated in all books about Lie groups and Lie algebras generally, a specific reference could be the article of R. Schmid [30] or the slides of P. W. Michor [23].

Almost all the definitions given in finite dimension are still true in infinite dimension. The main differences between the two cases is that infinite-dimensional Lie groups are not locally compact. For this reason, some results are not still true in infinite dimension, for example if G is a finite-dimensional Lie group, we have yet seen that the exponential map $\exp\colon \mathfrak{g} \to G$ is a local diffeomorphism from a neighbourhood of zero in $\mathfrak{g}$ onto a neighbourhood of the identity in G. This is not true in infinite dimension: the exponential map $\exp\colon \mathfrak{X}(M) \to \mathrm{Diff}_C(M)$ satisfies $T_0 \exp = \mathrm{id}$ but it is not locally surjective near the identity, so, as we will see in the next chapter, this causes some problems.
Another false result in infinite dimension is the connection between Lie algebras and Lie groups; in fact, if $\mathfrak{g}$ is any finite-dimensional Lie algebra, then there exists a connected finite-dimensional Lie group G such that $\mathfrak{g} \cong T_eG$,

in general this is not true in infinite dimension. Moreover, as it is shown in [23], Lie subalgebras do not correspond to Lie subgroup.

Now, we want to study very briefly the example of diffeomorphisms group

$$\mathrm{Diff}(M) = \{f \colon M \to M | f \in C^{\infty}\}$$

on a compact manifold M. This is an infinite-dimensional Lie group and the Lie algebra is given by $\mathfrak{g} \cong \mathfrak{X}(M)$ (to see a complete treatise of this argument, the reader could read the book of A. Banyaga [5]).
We get the family of diffeomorphism ϕ_t as the trajectories of the differential equation

$$\begin{cases} \dot{\phi}_t(x) = X(\phi_t(x)) \\ \phi_0(x) = x \end{cases}$$

The diffeomorphism ϕ_1 is called the time one map of the flow and the correspondence $X \mapsto \phi_1$ is the well-known exponential map

$$\exp \colon \mathfrak{X}(M) \to \mathrm{Diff}_C(M)$$

which is the analogue of the exponential map of finite-dimensional Lie groups, with the big difference that it fails to be surjective near the identity.

Remark 1.48. We consider the set of all smooth diffeomorphisms on a manifold M, which is a smooth Fréchet-Lie group. The hypothesis of smooth diffeomorphisms is fundamental, because the group of $\mathbb{C}^n$-diffeomorphisms is a smooth Banach manifold and a topological group, but not a Lie group. The problem is that the group operations are continuous but not differentiable.

The diffeomorphism, that we get as time 1 maps of flows, are said embeddable into a flow and we will denote by $\exp(\mathfrak{X}(M))$ the set of all diffeomorphisms with compact supports which are embeddable in some flow. We have to note that $\exp(\mathfrak{X}(M))$ is not a subgroup of $\mathrm{Diff}_c(M)$. For example in [26] J. Palis has shown that $\exp(\mathfrak{X}(M))$ is only a small piece of $\mathrm{Diff}_c(M)$. This is a reason why it is so difficult to link a diffeomorphism with the flow of vector field, but we will speak about this more thoroughly in the third chapter.

Chapter 2

BCH formula

The BCH formula is an useful and powerful instrument to write the product of exponential of two non-commutative indeterminates X, Y; it holds

$$e^X e^Y = e^Z$$

where Z has a certain form depending only on X, Y. The formula is used in many different applications, and the reader could find a large amount of references, for example [9, 10, 11].
Our goal is to use this formula to solve a physical mathematics problem: given two vector fields $X, Y \colon \mathbb{R}^n \to \mathbb{R}^n$, exists there a vector field $Z \colon \mathbb{R}^n \to \mathbb{R}^n$ such that $\phi^1_X \circ \phi^1_Y = \phi^1_Z$? The answer to the question is not easy and it is related to another topic: given a diffeomorphism $\psi \colon \mathbb{R}^n \to \mathbb{R}^n$, exists there a vector field X such that its flow at time 1, ϕ^1_X, is exactly the diffeomorphism ψ?

2.1 Formulation of the BCH formula

There exists a large amount of different proofs of BCH formula, we present one of these (cf. the lecture notes [1]).
Let G be a Lie group of Lie algebra $\mathfrak{g}$. We consider the exponential map $\exp \colon \mathfrak{g} \to G$, there exist $U' \subset \mathfrak{g}$ and $U \subset G$ such that the exponential map is a diffeomorphism between U' and U. The inverse map of exponential is naturally the logarithm map, denoted by $\log \colon G \to \mathfrak{g}$, as we have shown in the Proposition 1.41.
Let ψ be the complex function

$$\psi(z) = \frac{z \log z}{z - 1} \tag{2.1}$$

which is analytic in the open disk $\{z | |z-1| < 1\}$ since $\log z$ is analytic there and has a zero at $z = 1$. The function ψ has a power series expansion in the disk looking like

$$\psi(z) = \frac{z \log z}{z-1} = z \sum_{n=0}^{\infty} \frac{(-1)^n (z-1)^n}{n+1}$$

In particular, if we substitute $z = 1 + u$ we have

$$\psi(1+u) = (1+u) \sum_{n=0}^{\infty} \frac{(-1)^n u^n}{n+1} = 1 + \frac{u}{2} - \frac{u^2}{6} + \dots$$

and this series is absolutely convergent for $|u| < 1$.
With these hypotheses we can enunciate the following theorem

Theorem 2.2. *(BCH formula)*
Let G be a Lie group of Lie algebra $\mathfrak{g}$ and let X, Y be elements of the Lie algebra $\mathfrak{g}$. Then

$$\log(\exp(X)\exp(Y)) = X + \int_0^1 \psi(\exp(ad_X)\exp(t ad_Y)) Y\, dt \tag{2.3}$$

whenever the $\log(\exp(X)\exp(Y))$ *is defined.*

Before the proof of the theorem, we have to give the following results about the derivative of the exponential map

Proposition 2.4. *Let G be a Lie group and let $X, Y \in \mathfrak{g}$ as above. Then*

$$d_X \exp Y = d_e L_{\exp(X)} \int_0^1 \exp(-s ad_X Y) ds = d_e L_{\exp(X)} \frac{1 - \exp(-ad_X)}{ad_X} Y \tag{2.5}$$

Proof. We study the following map

$$\exp(-sX)\exp(s(X+tY))$$

Deriving with respect to t and putting $t = 0$, and subsequently taking derivative with respect s, we obtain this expression

$$d_X \exp Y = \frac{\partial}{\partial t} \exp(X+tY)\bigg|_{t=0} = d_e L_{\exp(X)} \int_0^1 \frac{\partial}{\partial s}\frac{\partial}{\partial t} \exp(-tX)\exp(s(X+tY))\bigg|_{t=0} ds$$

The order of the two differential operations can be swapped, so we firstly derive with respect to s

$$\frac{\partial}{\partial s} \exp(-sX)\exp(s(X+tY)) = d_e L_{\exp(-sX)} d_e R_{\exp(s(X+tY))}(-X) + \\ d_e L_{\exp(-sX)} d_e R_{\exp(s(X+tY))}(X+tY) = d_e L_{\exp(-sX)} d_e R_{\exp(s(X+tY))} tY$$

Taking the derivative with respect to t and putting $t = 0$, we see that

$$\frac{\partial}{\partial t}\frac{\partial}{\partial s}\exp(-sX)\exp(s(X+tY))\Big|_{t=0} = d_e L_{\exp(-sX)} tY\, d_e R_{\exp(s(X+tY))} tY +$$
$$d_e L_{\exp(-sX)} d_e R_{\exp(s(X+tY))} Y\Big|_{t=0} = d_e L_{\exp(-sX)} d_e R_{\exp(sX))} Y =$$
$$\exp(-s\mathrm{ad}_X)Y$$

where we use that

$$(\exp(X))Y(\exp(X))^{-1} = \exp(\mathrm{ad}_X Y)$$

In this way we demonstrate the first equality.
To understand the second equality, it is sufficient to see that the function $\frac{1-e^{-z}}{z}$ has the following power series expansion

$$\frac{1-e^{-z}}{z} = \sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n!}z^{n-1} = \sum_{n=0}^{\infty}\frac{(-1)^n}{n!}\int_0^1 s^n z^n ds = \int_0^1 \sum_{n=1}^{\infty}\frac{(-1)^n}{n!} s^n z^n ds = \int_0^1 e^{sz}$$

converging for any z. So if we put ad_X instead of z, we can conclude. □

Corollary 2.6. *If $v(t)$ is a smooth function defined on an interval I taking values in $\mathfrak{g}$, then*

$$d_{v(t)}\exp v'(t) = d_e L_{\exp(v(t))}\frac{1-\exp(-ad_{v(t)})}{ad_{v(t)}}v'(t) \tag{2.7}$$

Now we are able to demonstrate the BCH formula

Proof. (Theorem 2.2)
Let $z(t)$ be the curve $z(t) := \log(\exp(X)\exp(tY))$, that is $\exp(z(t)) = \exp(X)\exp(tY)$, since the exponential map and the logarithm map are inverse to each other.
Computing the derivation of z we get on one hand

$$\begin{aligned} d_{z(t)}\exp(z'(t)) &= d_{\exp(tY)}L_{\exp(X)}d_t\exp(tY) = d_{\exp(tY)}L_{\exp(X)}d_e L_{\exp(tY)}Y \\ &= d_e L_{\exp(X)\exp(tY)}Y = d_e L_{\exp(z)}Y \end{aligned} \tag{2.8}$$

On the other hand, for the Corollary 2.6

$$d_{z(t)}\exp z'(t) = d_e L_{\exp(z(t))}\frac{1-\exp(-\mathrm{ad}_{z(t)})}{\mathrm{ad}_{z(t)}}z'(t) \tag{2.9}$$

Combining the equation (2.8) with the equation (2.9) and using the definition of function ψ, as defined in (2.1), we have

$$Y = \psi(\exp(\mathrm{ad}_{z(t)})^{-1} z'(t)$$

so

$$z'(t) = \psi(\exp(\mathrm{ad}_{z(t)}))Y$$

Using the fundamental theorem of calculus, we obtain

$$\log(\exp(X)\exp(Y)) = z(1) = X + \int_0^1 \psi(\exp(\mathrm{ad}_{z(t)}))Y\,dt$$

Now we have to study better the term $\exp(\mathrm{ad}_{z(t)})$: firstly we recall that adjoint map and the exponential map commute (see (1.47)), i.e.

$$\mathrm{Ad}_{\exp} = \exp \mathrm{ad}$$

using it, we see that

$$\begin{aligned}\exp(\mathrm{ad}(z(t))) &= \mathrm{Ad}_{\exp(z(t))} = \mathrm{Ad}_{\exp(X)\exp(tY)} = \mathrm{Ad}_{\exp(X)}\mathrm{Ad}_{\exp(tY)} \\ &= \exp(\mathrm{ad}_X)\exp(t\mathrm{ad}_Y)\end{aligned}$$

so we can conclude the proof. □

Remark 2.10. As we have seen in the Section 1.2.4, ad_x and ad_y are endomorphism of $\mathfrak{g}$, given by

$$\mathrm{ad}_X V = [X, V] \qquad \mathrm{ad}_Y U = [Y, U]$$

The expressions $\exp(\mathrm{ad}_X)$ e $\exp(t\mathrm{ad}_Y)$ are the usual exponentials of endomorphism, so they are also endomorphisms of $\mathfrak{g}$ and they could written as power series

$$\exp(\mathrm{ad}_X) = 1 + \mathrm{ad}_X + \frac{\mathrm{ad}_X^2}{2} + \frac{\mathrm{ad}_X^3}{6} + \dots$$

$$\exp(t\mathrm{ad}_Y) = 1 + t\mathrm{ad}_Y + t^2\frac{\mathrm{ad}_Y^2}{2} + t^3\frac{\mathrm{ad}_Y^3}{6} + \dots$$

So, if we calculate the first terms described in the formula (2.3), we have

$$U = \exp(\mathrm{ad}_X)\exp(t\mathrm{ad}_Y) - 1 = \mathrm{ad}_X + \frac{1}{2}\mathrm{ad}_X^2 + t\mathrm{ad}_Y + \frac{t^2}{2}\mathrm{ad}_Y^2 + \dots$$

$$U^2 = \mathrm{ad}_X^2 + t\mathrm{ad}_Y\mathrm{ad}_X + \dots$$

and substituting these terms in the integral

$$\int_0^1 \left(1+\frac{U}{2}-\frac{U^2}{6}+\dots\right)Y dt =$$
$$= \int_0^1 \left[Y+\frac{\mathrm{ad}_X Y}{2}+\frac{\mathrm{ad}_X^2 Y}{4}+\frac{t\mathrm{ad}_Y\mathrm{ad}_X Y}{2}-\frac{\mathrm{ad}_X^2 Y}{6}-\frac{\mathrm{ad}_Y\mathrm{ad}_X Y}{6}+\dots\right]dt$$
$$= \int_0^1 \left[Y+\frac{1}{2}\mathrm{ad}_X Y+\frac{1}{12}\mathrm{ad}_X^2 Y+\frac{t}{2}\mathrm{ad}_Y\mathrm{ad}_X Y-\frac{1}{6}\mathrm{ad}_Y\mathrm{ad}_X Y+\dots\right]dt$$
$$= Y+\frac{1}{2}\mathrm{ad}_X Y+\frac{1}{12}\mathrm{ad}_X^2 Y+\frac{1}{4}\mathrm{ad}_Y\mathrm{ad}_X Y-\frac{1}{6}\mathrm{ad}_Y\mathrm{ad}_X Y+\dots$$
$$= Y+\frac{1}{2}[X,Y]+\frac{1}{12}[X,[X,Y]]+\frac{1}{12}[Y,[Y,X]]+\dots$$

So we can conclude that, given $X, Y \in \mathfrak{g}$, it follows from the above observation that

$$e^X e^Y = e^{Y+\frac{1}{2}[X,Y]+\frac{1}{12}([X,[X,Y]]+[Y,[Y,X]])+\dots}$$

which is the usual form of the BCH formula.

2.2 The problem of convergence

We have seen in the previous section the formulation of BCH formula, now we have to study the validity of it. The main problem is that the formula is defined as a formal series so it is true only under condition of convergence: we have to control the domain of convergence of logarithm series. This problem has a solution in finite dimension, but the convergence of the BCH formula in infinite dimension is a very difficult topic.

2.2.1 Postnikov's work

Now we analyse the demonstration of convergence given by M. Postnikov in his book [29, Lecture 4]. This proof is valid only in finite-dimensional case, so we want to underline where are the difficulties to extend the result to infinite dimension and the inaccuracies of Postnikov's work.

Let G be an analytic Lie group and let $X \in \mathfrak{g}$ be an analytic vector field. Let f be an analytic function defined in a neighbourhood of e, moreover, it is necessary to introduce into the domain of f, $W(f)$, a topology, because we shall use a norm.

If we apply the exponential map to the function f, as defined in (1.34), we can write

$$e^X f = f + Xf + \frac{X^{[2]}f}{2} + \ldots + \frac{X^{[n]}f}{n} + \ldots = \sum_{n=0}^{\infty} \frac{X^{[n]}f}{n!}$$

where $X^{[n]}$ designates the n-th fold iteration of the operator X, $X^{[n]}f = X(X^{[n-1]}f)$. For the sake of brevity, we use the notation e^X to indicate the exponential map. This operator applies only to such functions for which series above has a non-empty domain of convergence, hence it is fundamental to determine it.
We will not think the operator e^X only as a differential operator, but also as the flow of vector field X, i.e. a diffeomorphism. This only requires that integral curves $t \mapsto \phi_a(t)$ of X should be defined for $|t| \leq 1$ and there should be points $a \in W(f)$ such that $\phi_a(t) \in W(f)$ for $|t| \leq 1$.
Supposing that the vector field and the function are analytic, it is true a very useful property of the exponential operator: the exchange theorem by Gröbner (see [17, 1.2]).

Proposition 2.11. *(Exchange theorem)*
If X is an analytic vector field and f is an analytic function, as defined above, it holds

$$(e^X f)(x) = f(\phi_X^1(x))$$

for $x \in G$.

Proof. To demonstrate the equality we follow a different way from Grobner. Starting from the Proposition 1.13

$$\frac{d}{dt}(\phi_X^t)^* f = (\phi_X^t)^* \mathcal{L}_X f$$

we see that

$$\left.\frac{d^2}{dt^2}(\phi_X^t)^* f\right|_{t=0} = \left.\frac{d}{dt}\left[\frac{d}{dt}(\phi_X^t)^* f\right]\right|_{t=0} = \left.\frac{d}{dt}[(\phi_X^t)^* \mathcal{L}_X f]\right|_{t=0}$$
$$= (\phi_X^t)^* \mathcal{L}_X(\mathcal{L}_X f) = (\phi_X^t)^* \mathcal{L}_X^2 f$$

and for any $n \in \mathbb{N}$

$$\left.\frac{d^n}{dt^n}(\phi_X^t)^* f\right|_{t=0} = (\phi_X^t)^* \mathcal{L}_X^n f \tag{2.12}$$

Since both X and f are assumed to be analytic, we can expand $(\phi_X^1)^* f$ in Taylor series at $t = 0$, so

$$(\phi_X^1)^* f = (\phi_X^0)^* f + \left.\frac{d}{dt}(\phi_X^t)^* f\right|_{t=0} \cdot 1 + \frac{1}{2}\left.\frac{d^2}{dt^2}(\phi_X^t)^* f\right|_{t=0} \cdot 1^2 + \ldots$$

Using the (2.12), since $\phi^0_X(x) = x$ we obtain

$$\frac{d}{dt}(\phi^t_X)^* f\Big|_{t=0} = \mathcal{L}_X f$$

$$\frac{d^2}{dt^2}(\phi^t_X)^* f\Big|_{t=0} = \mathcal{L}^2_X f$$

and, in general,

$$\frac{d^n}{dt^n}(\phi^t_X)^* f\Big|_{t=0} = \mathcal{L}^n_X f$$

Hence

$$f(\phi^1_X) = (\phi^1_X)^* f = \sum_{n=0}^{\infty} \frac{\mathcal{L}^n_X}{n!} f = \sum_{n=0}^{\infty} \frac{X^n}{n!} f = e^X f$$

□

We can give the following definition

Definition 2.13. (Normal neighbourhood)
A neighbourhood $\mathring{U}$ of the null vector in $\mathfrak{g}$ is said to be *normal* if:

1. it has the property of being starlike, i.e. along some vector X it contains all vectors of the form tX for $|t| \leq 1$;

2. the mapping exp transforms $\mathring{U}$ diffeomorphically onto some neighbourhood U of the identity of the group G.

The neighbourhood U will also be called a *normal neighbourhood.*

Since the mapping exp is a diffeomorphism at the identity, there are arbitrary small normal neighbourhoods.
It follows from statement (1) of Definition 2.13 that the condition on the domain of f is satisfied if this domain is a normal neighbourhood U of the point e. Since any point in U is of the form e^X, where $X \in \mathring{U}$, formula (2.11) defines the function f on the entire neighbourhood U.
Using (2.11), we want to extend it for more than one vector field: let X, Y be two vector fields in $\mathfrak{g}$ we want to calculate $f(e^X e^Y)$, assuming $e^X e^Y \in U$. Hence

$$f(e^X e^Y(x)) = e^X f(e^Y(x)) = e^X e^Y f(x)$$

where we use the (2.11) 2-times. In general, it holds

$$f(e^{X_1} \dots e^{X_n}) = e^{X_1} \dots e^{X_n} f \tag{2.14}$$

for any $X_1, \ldots, X_n$ in the corresponding normal neighbourhood of the zero of the Lie algebra $\mathfrak{g}$.
This hypothesis of exponential map as diffeomorphism around the identity is important because on it Postnikov constructs the proof of convergence and we now that it is true only in finite dimension. This is the first problem to extend the result to infinite dimension.

By definition of exponential map we have

$$e^X e^Y = \sum_{p=0}^{\infty} \frac{X^{[p]}}{p!} \left(\sum_{q=0}^{\infty} \frac{Y^{[q]}}{q!} \right) = \sum_{p,q=0}^{\infty} \frac{X^{[p]} Y^{[q]}}{p!q!}$$

and substituting it in the logarithm series we obtain

$$\begin{aligned} \log(e^X e^Y) &= \sum_{n=1}^{\infty} \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} \left(\sum_{p,q=0}^{\infty} \frac{X^{[p]} Y^{[q]}}{p!q!} - \mathbb{I} \right)^k \\ &= \sum_{n=1}^{\infty} \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} \left(\sum_{\substack{p,q=0 \\ p+q>0}}^{\infty} \frac{X^{[p]} Y^{[q]}}{p!q!} \right)^k \\ &= \sum_{n=1}^{\infty} \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} \sum \frac{X^{[p_1]} Y^{[q_1]} \ldots X^{[p_k]} Y^{[q_k]}}{p_1! q_1! \ldots p_k! q_k!} \end{aligned}$$

where in the internal sum the summation is taken over all possible collection $(p_1, \ldots, p_k, q_1, \ldots, q_k)$ of non-negative integer, subject to the conditions

$$p_1 + q_1 > 0, \ldots, p_k + q_k > 0$$

$$p_1 + \ldots + p_k + q_1 + \ldots + q_k = n$$

Now it arises the big problem of convergence of the logarithm series. Following Postnikov's work, we have to introduce a norm to estimate the series.
First of all we assume that there exists a norm $\|\cdot\|$ on $\mathfrak{g}$. We will show that there exists a number $\delta > 0$ such that for $|||X|||, |||Y||| < \delta$ the operator $\ln(e^X e^Y)$ is applicable to any smooth function f in G. Since every element of G in the domain of f has a coordinate neighbourhood U with a compact closure $\bar{U}$, which is contained in this domain, it suffices to prove that for any coordinate neighbourhood U with compact closure $\bar{U}$ the operator $\log(e^X e^Y)$ is defined on the space of all smooth functions on $\bar{U}$.
So, using the hypothesis of compactness of the set $\bar{U}$, Postnikov decided to introduce the following norm

$$\|f\| = \max\left(|f|, \left| \frac{\partial f}{\partial x^1} \right|, \ldots, \left| \frac{\partial f}{\partial x^n} \right| \right)$$

The choice of this norm looks not so clear, because it considers only the first derivative of the function f. There is not any reason to suppose that the second derivative remains bounded, although the function and its first derivative are bounded. For example, we could introduce the norm

$$\|f\|_m := \max_{x\in\mathbb{R}^n}\Big(\sum_{\alpha\leq m}|\partial^\alpha f(x)|^2\Big)^{\frac{1}{2}}$$

for $m \in \mathbb{N}$ and $f \in C_c^\infty(\mathbb{R}^n,\mathbb{R})$. So

$$\|Xf\|_1 = \max_{x\in\mathbb{R}^n}\Bigg(\Big(\sum_{i=1}^n X^i(x)\partial_i f(x)\Big)^2 + \sum_{k=1}^n\Big(\sum_{i=1}^n \partial_k X^i(x)\partial_i f(x) + X^i(x)\partial_k\partial_i f(x)\Big)^2\Bigg)^{\frac{1}{2}}$$

It is clear that, although $\|f\|_1$ is bounded, it can happen that the second derivative $\partial_k\partial_i f(x)$ is not bounded. Therefore it seems that a better choice is a norm which considers all the derivatives of the function, not only the first.

Now we arrive at the main problem of Postnikov's work. Since $Xf = X^i\frac{\partial f}{\partial x^i}$ and $|X^i| \leq \|X^i\|$ for all i, for every vector field $X \in \mathfrak{g}$ with $|||X||| < \delta$ it holds

$$\|Xf\| \leq \delta\|f\| \tag{2.15}$$

This inequality comes from the definition of operator norm

$$|||X||| := \sup_{\substack{f\in G\\ \|f\|=1}}\|Xf\| = \sup_{\substack{f\in G\\ \|f\|\neq 0}}\frac{\|Xf\|}{\|f\|}$$

therefore

$$\|Xf\| \leq |||X|||\,\|f\|$$

But the operator norm is defined if and only if the operator is continuous or, equivalently, bounded. This is a big problem in our case, because the differential operator is not continuous, so we cannot define the operator norm of a vector field and the above estimate is not true.
The estimate remains efficient when the operator norm is defined (for example in finite dimension). Otherwise it is necessary to introduce a further hypothesis, that is the Lie algebra $\mathfrak{g}$ is a Banach Lie algebra, therefore it holds the sub-multiplicative property

$$\|[X,Y]\| \leq M\|X\|\|Y\| \tag{2.16}$$

for $X, Y \in \mathfrak{g}$ and $M > 0$. This hypothesis is requested, for example, in the work of S. Blanes and F. Casas [9], in the book of N. Bourbaki [11] or in the article of S. Biagi and A. Bonfiglioli [8], but unfortunately they do not show any example of such Lie algebra and our case does not satisfy the hypothesis.

From the result (2.15), if we suppose $|||X|||\,, |||Y||| < \delta$, by induction it could be derived that

$$\begin{aligned}\|X^{[p_1]}Y^{[q_1]}1\dots X^{[p_k]}Y^{[q_k]}f\| &= \|X^{[p_1]}(Y^{[q_1]}1\dots X^{[p_k]}Y^{[q_k]}f)\| \\ &\le \underbrace{\left|\left|\left|X^{[p_1]}\right|\right|\right|}_{\le\delta^{p_1}} \|Y^{[q_1]}1\dots X^{[p_k]}Y^{[q_k]}f\| \\ &\le \delta^n\|f\|\end{aligned}$$

where $n = p_1 + \dots + p_k + q_1 + \dots + q_k$. Using this estimates it follows

$$\begin{aligned}\|\log(e^Xe^Y)f\| &= \left\|\sum_{n=1}^{\infty}\sum_{k=1}^{n}\frac{(-1)^{k-1}}{n}\sum\frac{X^{[p_1]}Y^{[q_1]}\dots X^{[p_k]}Y^{[q_k]}f}{p_1!q_1!\dots p_k!q_k!}\right\| \\ &\le \sum_{n=1}^{\infty}\sum_{k=1}^{n}\frac{|-1|^{k-1}}{k}\sum\frac{\|X^{[p_1]}Y^{[q_1]}\dots X^{[p_k]}Y^{[q_k]}f\|}{p_1!q_1!\dots p_k!q_k!} \\ &\le \sum_{n=1}^{\infty}\sum_{k=1}^{n}\frac{1}{k}\sum\frac{\left|\left|\left|X^{[p_1]}Y^{[q_1]}\dots X^{[p_k]}Y^{[q_k]}\right|\right|\right|\,\|f\|}{p_1!q_1!\dots p_k!q_k!} \\ &\le |\log(e^{2\delta})|\|f\| \\ &= \sum_{n=1}^{\infty}\frac{(e^{2\delta}-1)^n}{n}\|f\|\end{aligned}$$

The series converges for $|e^{2\delta}-1| < 1$, i.e. for $\delta < \frac{\log 2}{2}$. Hence for $\delta < \frac{\log 2}{2}$ the series will uniformly converge in $\bar{U}$ to some smooth function, i.e. the operator $\log(e^Xe^Y)$ will be applicable to the function f.

2.2.2 Introduction of a new norm

We have seen in the previous section the proof for the convergence of the series done by Postnikov and the reasons for which we cannot apply to vector fields. Now we would to show our attempt to adapt the previous proof to our case, unfortunately we do not achieve our goal.
We have seen that the main problem is to introduce a suitable norm, so initially our research has as objective to find this norm.

We arouse our interest in hamiltonian mechanics: in particular in a Giorgilli's work (see [16, Appendix A]) we can find a suitable norm. For a given generating function $\chi(p,q)$ the Lie series operator is defined as the exponential of Lie derivative $\mathcal{L}_\chi \cdot = \{\cdot, \chi\}$

$$\exp(\varepsilon\mathcal{L}_\chi) = \sum_{n\geq 0} \frac{\varepsilon^n}{n!}\mathcal{L}_\chi^n$$

for $\varepsilon > 0$. This is, as we have already known, the autonomous flow of the canonical vector field generated by $\chi(p,q)$. Moreover, if X is the vector field related to the generating function χ, it holds the following series of equality

$$\{f, \chi\} = \mathcal{L}\chi f = \mathcal{L}_X f = X(f) \tag{2.17}$$

for a smooth function f.
Let introduce the basic elements to discuss the converge of Lie series. We restrict the attention to the case of a phase space endowed with action-angle variables (p,q) and we use a family of complex domain

$$\mathcal{D}_{(1-d)(\rho,\sigma)} := \Delta_{(1-d)\rho} \times \mathbb{T}^n_{(1-d)\sigma} \tag{2.18}$$

for fixed $\rho, \sigma > 0$ and $0 \leq d < 1$, where

$$\Delta_\rho = \{p \in \mathbb{C}^n : |p| \leq \rho\} \quad \text{and} \quad \mathbb{T}^n_\sigma = \{q \in \mathbb{C}^n : |\Im(q)| \leq \sigma\}$$

Certainly the whole argument may be extended to the complex domain by making the union of all complex disks of radius ρ.
For an holomorphic function $f(p,q)$ in the domain $\mathcal{D}_{(\rho,\sigma)}$, we shall use the supremum norm

$$\|f\|_{(\rho,\sigma)} = \sup_{(p,q)\in\mathcal{D}_{(\rho,\sigma)}} |f(p,q)|$$

which is assumed finite.
So, assuming that we know the norm $\|\chi\|_{(\rho,\sigma)}$ and the norm $\|f\|_{(1-d')(\rho,\sigma)}$, then for $d' < d < 1$ it is true the estimates

$$\|\mathcal{L}_\chi f\|_{(1-d)(\rho,\sigma)} \leq \frac{C}{d(d-d')\rho\sigma}\|\chi\|_{(\rho,\sigma)}\|f\|_{(1-d')(\rho,\sigma)}$$

with some constant $C \geq 1$.
From this estimates in [15] C. Efthymiopoulos obtains a similar norm, more interesting for our application. In the same domain (2.18), let $\chi_1, \ldots, \chi_n$ and f be analytic functions with Fourier-weighted norms $\|\chi_i\|_{(\rho,\sigma)}$, $\|f\|_{(\rho,\sigma)}$. Then

$$\|(\mathcal{L}_{\chi_1} \circ \ldots \circ \mathcal{L}_{\chi_n})f\|_{(\rho,\sigma)} \leq \frac{n!}{e}\left(\frac{e}{d^2\rho\sigma}\right)^n \|\chi_n\|_{(\rho,\sigma)} \ldots \|\chi_1\|_{(\rho,\sigma)}\|f\|_{(\rho,\sigma)} \tag{2.19}$$

for every d with $0 < d < 1$ and for every positive integer n.
This estimate is precisely what we would like to use; in fact it is quite similar to the (2.16), except for the presence of some constant, but in this case we have to estimate the norm of generating functions, not directly the norm of vector fields.
Let H_1, H_2 be two Hamiltonians with vector fields X_1, X_2 respectively. Let suppose $\|H_1\|_{(\rho,\sigma)}, \|H_2\|_{(\rho,\sigma)} < \delta$, for an analytic function f, $\|f\|_{(\rho,\sigma)} < \infty$, we can rewrite the well-known series

$$\begin{aligned}\log(e^{X_1}e^{X_2})f &= \sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}(e^{X_1}e^{X_2}-\mathbb{I})^{[n]}f\\ &= \sum_{n=1}^{\infty}\sum_{k=1}^{n}\frac{(-1)^{k-1}}{k}\left(\sum\frac{X_1^{[p_1]}X_2^{[q_1]}\ldots X_1^{[p_k]}X_2^{[q_k]}}{p_1!q_1!\ldots p_k!q_k!}\right)f\end{aligned}$$

where the inner summation is taken over all the possible collections of positive integers $(p_1,\ldots,p_k,q_1,\ldots,q_k)\in\mathbb{N}^{2k}$ such that $p_1+q_1>0,\ldots,p_k+q_k>0$ and $p_1+\ldots+p_k+q_1+\ldots+q_k=n$. For the sake of brevity, we suppose that $\frac{e}{d^2\rho\sigma}\le 1$, so we can ignore the constants. Using the norm (2.19) we have

$$\begin{aligned}\|\log(e^{X_1}e^{X_2})f\|_{(1-d)(\rho,\sigma)} &\le \sum_{n=1}^{\infty}\sum_{k=1}^{n}\frac{1}{k}\left\|\sum\left(\frac{X_1^{[p_1]}X_2^{[q_1]}\ldots X_1^{[p_k]}X_2^{[q_k]}}{p_1!q_1!\ldots p_k!q_k!}\right)f\right\|_{(\rho,\sigma)}\\ &\le \sum_{n=1}^{\infty}\sum_{k=1}^{n}\frac{1}{k}\sum n!\frac{\|H_1\|_{(\rho,\sigma)}^{p_1}\ldots\|H_2\|_{(\rho,\sigma)}^{q_k}}{p_1!q_1!\ldots p_k!q_k!}\|f\|_{(\rho,\sigma)}\\ &\le \sum_{n=1}^{\infty}\sum_{k=1}^{n}\frac{1}{k}\delta^n\sum\frac{n!}{p_1!q_1!\ldots p_k!q_k!}\|f\|_{(\rho,\sigma)}\\ &= \sum_{n=1}^{\infty}\sum_{k=1}^{n}\frac{1}{k}\delta^n(2k)^n\|f\|_{(\rho,\sigma)}\\ &= \sum_{n=1}^{\infty}\sum_{k=1}^{n}\frac{1}{k}(2k\delta)^n\|f\|_{(\rho,\sigma)}\end{aligned}$$

where we have used the multinomial theorem

$$(a_1+\ldots+a_{2k})^n = \sum_{\substack{(p_1,\ldots,p_{2k})\in\mathbb{N}^{2k}\\ p_1+\ldots+p_{2k}=n}}\frac{n!}{p_1!\ldots p_{2k}!}a_1^{p_1}\ldots a_{2k}^{p_{2k}}$$

in particular for $a_1=\ldots=a_{2k}=1$ it results

$$(2k)^n = \sum_{\substack{(p_1,\ldots,p_{2k})\in\mathbb{N}^{2k}\\ p_1+\ldots+p_{2k}=n}}\frac{n!}{p_1!\ldots p_{2k}!}$$

If we look at the above estimates, it is clearly that it may not converge, neither the partial sum is convergent because the series increases very rapidly. Unfortunately our research does not give the hoped result, but this is in line with the present literature, in which the topic of application of the BCH formula with the vector fields is barely debated.

2.2.3 A solution for the problem

A possibly solution, which is presented in the unpublished article of A. Bonfiglioli and S. Biagi [7], is to reduce the problem to finite dimension. More clearly, the idea is to restrict the Lie algebra of all vector fields to a finite-dimensional Lie subalgebra $V \subset \mathfrak{X}(\Omega)$. In this way, we are sure that the exponential map is well defined; moreover, we can solve the problem of convergence because all the norms are equivalent in finite dimension and they satisfy the sub-multiplicative property. With these premises, it is true the following theorem (cf. [7, Theorem 13.9])

Theorem 2.20. *(BCH theorem for ODEs)*
Let V be a finite-dimensional Lie subalgebra of $\mathfrak{X}(\Omega)$, the smooth vector fields on the open set $\Omega \subset \mathbb{R}^n$, and let $\|\cdot\|$ be a fixed norm on V. Then there exists a positive number ε, depending on $\|\cdot\|$, such that the homogeneous BCH series

$$Z(X,Y) = \log(\exp(X)\exp(Y))$$

is convergent for every $X, Y \in V$ with $\|X\|, \|Y\| < \varepsilon$.

We want to study more carefully this reduction to finite dimension subalgebra.
Let V be a m-dimensional Lie subalgebra of $\mathfrak{X}(\Omega)$, the Lie algebra of smooth vector fields on the open set $\Omega \subset \mathbb{R}^n$. Since $\dim(V) = m < n$, we can fix a basis $\{X_1, \dots, X_m\}$.
We recall some notions about the distribution and Frobenius' theorem (for a detailed treatise of the topic, the reader could see [22, Chapter 11]).

Definition 2.21. (Distribution)
A *smooth rank m distribution* on an n-dimensional manifold M is a smooth rank m vector subbundle $E \to M$ of the tangent bundle.

It is not difficult to see that a smooth rank m distribution on a n-manifold M gives a m-dimensional subspace $E_p \subset T_pM$ for each $p \in M$ such that for each fixed $p \in M$ there is a family of smooth vector fields $X_1, \dots, X_m$ defined on a neighbourhood U of p and such that $X_{1|q}, \dots, X_{m|q}$ are linearly independent and span E_q for each $q \in U$. In other words, $X_1, \dots, X_m$ can be

viewed as a local frame field for the subbundle E.
An important property of the distribution is to be involutive,

Definition 2.22. (Involutive)
If for every pair of locally defined vector fields X, Y with common domain that lie in a distribution $E \to M$, the bracket $[X, Y]$ also lies in the distribution, then we say that the distribution is *involutive*.

Moreover, a vector field X lies in a distribution if and only if for any spanning local frame field $X_1, \dots, X_m$ we have $X = \sum_{i=1}^m f^i X_i$ for smooth functions f^i defined on the common domain of $X_1, \dots, X_m$.
From these premises, it is clear that the basis of V, $\Delta = \{X_1, \dots, X_m\}$, is a smooth rank m distribution which is also involutive. This condition is very important because it lets us to use the proof of Frobenius' Theorem.
Let $\Omega \subset \mathbb{R}^n$, as in the Theorem 2.20. Given $p \in \Omega$, let U be a neighbourhood of p and let ϕ be a continuous map which defines local coordinate functions, such that $\{X_1, \dots, X_m\}$ is a local frame for Δ. Up to a reorganisation, we may suppose that $\{X_{1|_p}, \dots, X_{m|_p}, \partial_{m+1|_p}, \dots, \partial_{n|_p}\}$ is a basis for $T_p\Omega$, where we use the notation $\partial_i := \frac{\partial}{\partial x^i}$. Since $\{\partial_1, \dots, \partial_n\}$ is a local frame for $T\Omega$, there exist functions $a_i^j \in C^\infty(U)$ such that

$$X_i = \sum_{i=1}^n a_i^j \partial_j \qquad i = 1, \dots, n$$

where we indicate $X_j = \partial_j$ also for $j = m+1, \dots, n$.
Since $X_{1|_p}, \dots, X_{n|_p}$ is a basis for $T_p\Omega$, the matrix $A(p) := (a_i^j(p))$ is non-degenerate, with $\det(A) \neq 0$ in each point of U. So for every point of U the matrix A is invertible and its inverse is the matrix $B = (b_j^i)$, hence for $j = 1, \dots, n$

$$\partial_j = \sum_{i=1}^n b_j^i X_i = \sum_{i=1}^m b_j^i X_i + \sum_{i=m+1}^n b_j^i X_i$$

We can define new vector fields as

$$Y_j = \sum_{i=1}^m b_j^i X_i \qquad j = 1, \dots, m$$

Now we want to demonstrate that these new vector fields are a local frame for Δ and they are commutating.
Let $F\colon U \to \mathbb{R}^m$ be the function defined as $F = \pi \circ \phi$, where π is the projection onto the first m-coordinates. For every $q \in U$ it is true that

$$\partial_{j|_{F(q)}} = dF_q(\partial_j) = dF_q(Y_j) + \sum_{i=m+1}^n b_j^i(q) \underbrace{dF_q(\partial_i)}_{=0} = dF_q(Y_j)$$

Since $dF_q(Y_j) = \partial_{j|_{F(q)}}$ for each $q \in U$ and $\partial_{1|_{F(q)}}, \dots, \partial_{m|_{F(q)}}$ are linearly independent, then also $Y_1, \dots, Y_m$ must be linearly independent for each $q \in U$, hence they are a local frame for Δ. Moreover, it results that $dF_{q|_{\Delta_q}}$ is injective for each $q \in U$. So it holds that

$$dF_q([Y_i, Y_j]) = [dF_q(Y_i), dF_q(Y_j)] = [\partial_i, \partial_j]_{|_{F(q)}} = 0$$

for every $q \in U$ and for all $i, j = 1, \dots, m$. Since Δ is involutive, i.e. $[Y_i, Y_j]_{|_q} \in \Delta_q$ and $dF_{q|_{\Delta_q}}$ is injective, we deduce that $[Y_i, Y_j]_{|_q} = 0$ for each $q \in U$.

Hence, from this construction, we can conclude that, starting from a basis $\{X_1, \dots, X_m\}$ for the Lie subalgebra V, we are able to construct another basis $\{Y_1, \dots, Y_m\}$ such that

$$Y_i = \sum_{i=1}^{m} b_i^j X_j \qquad \text{for } j = 1, \dots, m$$

where b_i^j are C^∞-functions defined as above. The particularity of this new basis is that the vector fields are commutating, so the BCH formula is clearly convergent for vector fields sufficiently close to the zero vector field and becomes

$$\exp(Y_1)\exp(Y_2) = \exp(Y_1 + Y_2)$$

Chapter 3

Connection between flows and diffeomorphisms

Now we want to return to the question: given a diffeomorphism ψ, exists there a vector field X such that the flow at time 1, ϕ_X^1, is precisely ψ? The answer is not so clear and it is negative often.
We want to investigate very briefly the reason why the answer is not always positive, and, after that, to examine the way to construct this connection.

3.1 *Vector fields generate few diffeomorphisms*

One of the first papers in which this problem was studied is Palis's work [26], whose title is exactly *Vector fields generate few diffeomorphisms*. In fact he demonstrates that few diffeomorphisms, in sense of Baire category, embed in flows or are generated by vector fields. Before we enunciate the main result, we introduce the Baire category

Definition 3.1. (First category)
A subset E of a topological space S is said to be of *first category* in S if E can be written as the countable union of subsets which are nowhere dense in S, i.e. if E is expressible as a union

$$E = \bigcup_{n\in\mathbb{N}} E_n$$

where each subset $E_n \subset S$ is nowhere dense in S.
Sets which are not of first category are of second category.

So we can cite the Palis's result

Theorem 3.2. *The subset of $Diff^1(M)$ of diffeomorphisms that embed in C^1 flows is of first category.*

Now it is clear that the sets of a first category are "small" subsets of the host space, and sets of first category are sometimes referred to as meager. Therefore, it appears clear why it is not so easy to connect diffeomorphisms with flows.
After Palis's work the topic has been debated several times and some results have been reached. Nowadays, for example, it is known that non-autonomous C^1-smooth symplectomorphisms form a C^1-open and dense set in the group of C^1 diffeomorphisms (see [3]), and a similar statement is true also for C^∞-smooth hamiltonian diffeomorphisms of certain symplectic manifolds. Another result is due to L. Polterovich and E. Shelukhin, they prove that for a closed symplectically aspherical manifold the set Ham $\setminus$ Aut contains a C^∞-dense subset which is open in the topology induced by Hofer's metric (see [28]).

3.2 Diffeomorphisms and autonomous Hamiltonian

As it is written in [21], we know that the 1-time flow of an hamiltonian vector field is a symplectic diffeomorphism homotopic to the identity. So we want to investigate the inverse statement: which exact symplectic diffeomorphisms homotopic to the identity can be the flow of an hamiltonian vector field? The answer to this very general question is quite unknown. But if we restrict ourselves to a perturbative situation, we are able to answer to the previous question and to construct the hamiltonian vector field, so we collect some results using this techniques.

3.2.1 Restriction to perturbative situation

In [6] G. Benettin and A. Giorgilli prove that for any mapping Ψ_ε, analytic and ε-close to the identity, there exists an analytic autonomous Hamiltonian, H_ε, such that its time-one mapping Φ_{H_ε} differs from Ψ_ε by a quantity exponential small in $\frac{1}{\varepsilon}$. The idea is firstly to introduce a map Ψ_ε smoothly depending on a small parameter ε, such that $\Psi_0(p,q) = (p,q)$. After that the aim is to construct an Hamiltonian $H_\varepsilon = \varepsilon h_1 + \varepsilon^2 h_2 + \ldots$, as a formal series, whose time-one flow agrees, order by order, with Ψ_ε. This idea comes from a Moser's work in which, even if there is not properly the presence

of a small parameter ε, the author infers the existence of a formal power series $H(p,q)$ such that the flow of its vector field is precisely the solution of hamiltonian system of equation (for the detailed treatise see [24]). Some quite similar techniques could be found in the works [13, 20, 32], where the canonical transformation are defined as power series depending on a small parameter ε.
On the other hand, a second reference to Benettin and Giorgilli's work is the article of A. Neishtdat [25]. He adopts an indirect method to obtain the Hamiltonian; the idea is to preliminary introduce a non-autonomous Hamiltonian $\widetilde{H}_\varepsilon(p,q,t)$ which interpolates the diffeomorphism Ψ_ε exactly and then, using a suitable canonical transformation, eliminate from $\widetilde{H}_\varepsilon$ the variable t, which it turns out to be a fast variable.
After all this foreword, we want to study the method used by Benettin and Giorgilli to find the proper hamiltonian function. Let $\Psi_\varepsilon\colon \mathbb{R}^n \to \mathbb{R}^n$ be a mapping near the identity which could be written as a power series in ε

$$\Psi_\varepsilon(x) = x + \varepsilon\psi_1(x) + \varepsilon^2\psi_2(x) + \dots \tag{3.3}$$

where each ψ_k, for $k \geq 1$, is assumed to be real analytic in a suitable complex neighbourhood $\mathcal{D}_\rho$ defined as

$$\mathcal{D}_\rho = \bigcup_{x\in\mathcal{D}} \Delta_\rho$$

$$\Delta_\rho = \{y \in \mathbb{C}^n | |y^i - x^i| \leq \rho^i, 1 \leq i \leq n\}$$

for $\rho = (\rho^1, \dots, \rho^n)$ and $\mathcal{D} \subset \mathbb{R}^n$. Now, for $\rho' \leq \rho$, we introduce the norm

$$\|f\|_{\rho'} = \sup_{x\in\mathcal{D}_{\rho'}} |f(x)|$$

for f real analytic function in $\mathcal{D}_{\rho'}$. Given a vector field $X = (X^1, \dots, X^n)$, the differential operator acting on scalar function f is denoted by $L_X f$.
With these notations, it holds the following proposition

Proposition 3.4. *Consider the mappping Ψ_ε as defined in* (3.3) *and assume the functions ψ_k are real analytic in $\mathcal{D}_{\rho'}$ and satisfy the estimate*

$$\|\psi_k\|_{\rho'} \leq \gamma^{k-1}\Gamma$$

for some positive constants γ, Γ. Then there exists a formal series of vector fields

$$F_\varepsilon^\infty = \varepsilon X_1 + \varepsilon^2 X_2 + \dots$$

analytic in $\mathcal{D}_\rho$, such that

1. *one has formally,*
$$\exp(L_{F_\varepsilon^\infty})\xi = \Psi_\varepsilon \tag{3.5}$$
where ξ is the identity function in $\mathbb{R}^n$.

2. *The vector field X_k, for $k \geq 1$, satisfy the estimates*
$$\|X_1\|_\rho \leq \Gamma$$
$$\|X_k\|_{\rho/2} < \frac{1}{2} k^{k-1} \beta^{k-1} \Gamma \quad \textit{for } k \geq 2$$
with $\beta = 4\max(\gamma, \Gamma)$. Moreover, as long as $1 \leq r \leq \frac{1}{2\beta\varepsilon}$, the finite sum $F_\varepsilon^r = \varepsilon X_1 + \varepsilon^2 X_2 + \ldots + \varepsilon^r X_r$ satisfies the estimates
$$\|F_\varepsilon^r\|_{\rho/2} \leq \frac{3}{2}\varepsilon\Gamma$$
$$\|\Phi_{F_\varepsilon^r} - \Psi_\varepsilon\|_{\rho/4} < 3\varepsilon\Gamma(2r\beta\varepsilon)^r$$

3. *If Ψ_ε is symplectic, then all vector fields $X_1, X_2, \ldots$ are locally hamiltonian.*

Proof. We restrict ourselves to show the proof for the first part of the proposition, the rest of the proof could be found in [6, 3.1].
For the sake of brevity, we will use a short notation, $L_k := L_{X_k}$. By formally expanding, the exponential of $L_{F_\varepsilon^\infty}$ is
$$\exp(L_{F_\varepsilon^\infty}) = \mathbb{I} + \sum_{n\geq 1} \frac{1}{n!}\left(\sum_{k\geq 1} \varepsilon^k L_k\right)^n = \mathbb{I} + \sum_{n\geq 1} \varepsilon^n \sum_{m=1}^{n} \frac{1}{n!} A_{m,n} \tag{3.6}$$
where
$$A_{m,n} = \sum_{\substack{k_1+\ldots+k_m \geq 1 \\ k_1+\ldots+k_m = n}} L_{k_m} \ldots L_{k_n}$$
From the definition one could get the recursive relations for $A_{m,n}$
$$A_{1,n} = L_n$$
$$A_{m,n} = \sum_{k=1}^{n-m+1} L_k A_{m-1,n-k}$$
By comparing the expression (3.6) with (3.3), it results
$$X_1 = \psi_1$$

$$X_n = \psi_n - \sum_{m=2}^{n} \frac{1}{m!} A_{m,n}\xi$$

In this way we have a recursive scheme to calculate the vector field X_k such that its flow $\exp(L_{F_\varepsilon^\infty})$ coincides with Ψ_ε at any order in ε. □

Another important work for the topic is the article of S. Kuksin and J. Pöschel [21]. The authors work on the perturbations of two kinds of integrable maps: we see to the first example. Let $D \subset \mathbb{R}^n$ be bounded and convex domain and let $D \times \mathbb{T}^n$ be the symplectic manifold endowed with the standard exact symplectic structure $\upsilon = d\alpha$, where $\alpha = \sum_j I_j d\phi_j$. If we suppose the map

$$F_0 \colon D \times \mathbb{T}^n \to D \times \mathbb{T}^n \qquad (I, \phi) \mapsto (I, \phi + \omega(I))$$

is real analytic and symplectic, then

$$0 = \sum_j d\omega_j(I) \wedge dI_j = d\Big(\sum_j \omega_j dI_j\Big)$$

so, by the convexity, there exists a real analytic function h on D such that

$$\omega(I) = \frac{\partial h}{\partial I}$$

The function h defines an integrable hamiltonian system on $D \times \mathbb{T}^n$ with equations of motion $\dot{I} = 0, \dot{\phi} = \omega(I)$, so its flow interpolates exactly F_0

$$X_h^1 = F_0 \tag{3.7}$$

on $D \times \mathbb{T}^n$.

Now we interpolate the function F_0 to study if there exists a relation like the (3.7) which will depend on ε. We consider a real analytic family of real analytic maps

$$F_\varepsilon \colon D \times \mathbb{T}^n \to \mathbb{R}^n \times \mathbb{T}^n \qquad -\varepsilon < \varepsilon_0 < \varepsilon$$

perturbing the integrable map $F_0 \colon (I, \phi) \mapsto (I, \phi + \omega(I))$, where $\omega(I) = \frac{\partial h}{\partial I}$ and h is a real analytic function on D. Moreover, the F_ε are assumed exact symplectic. Each F_ε is assumed to be real analytic on a fixed complex neighbourhood $V_r D \times V_r \mathbb{T}^n$ of $D \times \mathbb{T}^n$, where

$$V_r D = \bigcup_{I_0 \in D} \{I \in \mathbb{C}^n \,||I - I_0| < r\} \subset \mathbb{C}^n$$

The same domain of analyticity is assumed for h. Possibly decreasing r and ε_0 we also have a uniform bound for the sup-norm $|F_\varepsilon|_{V_rD\times V_r\mathbb{T}^n}$ for all ε. Moreover,

$$|F_\varepsilon - F_0|_{V_rD\times V_r\mathbb{T}^n} = O(\varepsilon)$$

by Cauchy's estimates.

Theorem 3.8. *Suppose F_ε satisfies the preceding assumptions. Then for all sufficiently small ε there exists a real analytic, 1-periodic time dependent Hamiltonian H_ε on $D\times\mathbb{T}^{n+1}$, such that*

$$X^t_{H_\varepsilon|_{t=1}} = F_\varepsilon$$

on $D\times\mathbb{T}^n$. Moreover, there exists a $\rho>0$ such that H_ε is real analytic in $V_\rho D\times V_\rho\mathbb{T}^{n+1}$ for all small ε and satisfies

$$|H_\varepsilon - h|_{V_\rho D\times V_\rho\mathbb{T}^{n+1}} = O(\varepsilon)$$

as $\varepsilon\to 0$.

3.2.2 Exact correspondence between diffeomorphisms and flows

If we leave the previous strategy and we try to search an exact correspondence between a diffeomorphism and the flow of a vector field, not necessarily hamiltonian, the problems grow. What we like to do is to reuse the thoughts about BCH formula for this allied question.
Let $X\colon\mathbb{R}^n\to\mathbb{R}^n$ be an analytic vector field, using the definitions of logarithm and exponential series, for every $f\colon\mathbb{R}^n\to\mathbb{R}$ analytic we have

$$Xf(x) = \log(e^X)f(x) = \sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}(e^X-\mathbb{I})^{[n]}f(x)$$

If we use the exchange theorem (2.11) to write the term $(e^X-\mathbb{I})^{[n]}f(x)$ of the above equation, we obtain

$$(e^X-\mathbb{I})^{[1]}f(x) = e^Xf(x) - f(x) = f(e^X)(x) - f(x)$$

$$\begin{aligned}(e^X-\mathbb{I})^{[2]}f(x) &= (e^X-\mathbb{I})(e^X-\mathbb{I})f(x) = (e^X-\mathbb{I})(e^Xf(x)-f(x))\\ &= (e^X-\mathbb{I})(f(e^X)(x)-f(x)) = e^Xf(e^X)(x) - f(e^X)(x) - e^Xf(x) + f(x)\\ &= f(e^{2X})(x) - 2f(e^X)(x) + f(x)\end{aligned}$$

where we use the (2.14). If we take the projection $f = \pi_j$, i.e. $f(x) = x_j$, for any n it is true

$$(e^X - \mathbb{I})^{[n]}(x) = \sum_{k=0}^{n} \binom{n}{k} e^{kX}(x)(-1)^{n-k}$$

Hence we have

$$X(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \sum_{k=0}^{n} \binom{n}{k} (\phi_X^1)^k(x)(-1)^{n-k} \tag{3.9}$$

where, since ϕ_X^1 is a diffeomorphism, the k-th power means $\phi_X^1 \circ \ldots \phi_X^1$ k-times.
Let consider an analytic diffeomorphism ψ, we want to investigate if there exists a vector field X such that its 1-time flow generates the diffeomorphism. The idea is to start from the above formula (3.9) substituting the flow ϕ_X^1 for the diffeomorphism ψ

$$\begin{aligned} X(x) &= \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \sum_{k=0}^{n} \binom{n}{k} \psi^k(x)(-1)^{n-k} \\ &= -\sum_{n=1}^{\infty} \frac{1}{n} \sum_{k=0}^{n} \binom{n}{k} (-1)^k \psi^k(x) \end{aligned} \tag{3.10}$$

The series above is not so easily worked, even if we suppose the best hypotheses. For example, let suppose that x is a point which belongs to the attractive set of a fixed point $\bar{x}$, hence the composition of ψ k-times moves away little from $\bar{x}$ and we can estimate it easily. Despite this hypothesis, the series is divergent because of the presence of binomial coefficient which grows very rapidly.

We have a better result if we take care only of the first terms, not of the convergence of the entire series. This view, as it is said by Poincaré in [27, Chapter VIII], is typical of the astronomers, who pay essentially attention to the way in which the first terms decrease or increase. We want to find a vector field X such that its flow at time 1, ϕ_X, generates the analytic diffeomorphism ψ, which is near the identity using the C^0 norm.
Let start from the well-known identity (3.10), where we have used the definitions of logarithm and exponential map and the exchange theorem (2.11). We study in a different way from the (3.10) the second summation; noticing the similarity with the binomial theorem, we define a new operation which

is quite similar to the power of real number

$$(\psi - \mathbb{I})^{\{n\}}(x) := \sum_{k=0}^{n} \binom{n}{k} \psi^k(x)(-1)^{n-k}$$

Using this new operation we can define the vector field X as

$$X(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} (\psi - \mathbb{I})^{\{n\}}(x) \tag{3.11}$$

Now we want to study the term $(\psi - \mathbb{I})^{\{n\}}$, we notice that

$$(\psi - \mathbb{I})^{\{1\}} = \psi - \mathbb{I}$$

$$(\psi - \mathbb{I})^{\{2\}} = \psi^2 - 2\psi + \mathbb{I} = (\psi - \mathbb{I}) \circ \psi - (\psi - \mathbb{I})$$
$$(\psi - \mathbb{I})^{\{3\}} = \psi^3 - 3\psi^2 + 3\psi - \mathbb{I} = (\psi - \mathbb{I})^{\{2\}} \circ \psi - (\psi - \mathbb{I})^{\{2\}}$$

so we can generalize

$$(\psi - \mathbb{I})^{\{n\}} = (\psi - \mathbb{I})^{\{n-1\}} \circ \psi - (\psi - \mathbb{I})^{\{n-1\}}$$

Let $\rho > 0$ be the radius of the open ball $B_\rho = \{x||x| \leq \rho\}$; it is defined the sup-norm $|\cdot|_\rho$, which we use to estimate $(\psi - \mathbb{I})^{\{n\}}$

$$|(\psi - \mathbb{I})^{\{n\}}|_\rho = |(\psi - \mathbb{I})^{\{n-1\}} \circ \psi - (\psi - \mathbb{I})^{\{n-1\}}|_\rho \leq |\nabla(\psi - \mathbb{I})^{\{n-1\}}|_\rho |(\psi - \mathbb{I})|_\rho$$

and, using the Cauchy estimate for $\delta > 0$, we have

$$|\nabla(\psi - \mathbb{I})^{\{n-1\}}|_{\rho-2\delta} \leq \frac{1}{\delta} |(\psi - \mathbb{I})^{\{n-1\}}|_\rho$$

Using the above estimates we can write

$$|(\psi - \mathbb{I})^{\{n\}}|_{\rho - \sum_{k=1}^{n} \frac{\delta}{k^2}} \leq |\nabla(\psi - \mathbb{I})^{\{n-1\}}|_{\rho - \sum_{k=1}^{n} \frac{\delta}{k^2}} |(\psi - \mathbb{I})|_{\rho - \sum_{k=1}^{n} \frac{\delta}{k^2}}$$

and

$$|\nabla(\psi - \mathbb{I})^{\{n-1\}}|_{\rho - \sum_{k=1}^{n} \frac{\delta}{k^2}} \leq 2\frac{1^2}{\delta} |(\psi - \mathbb{I})^{\{n-1\}}|_{\rho - \sum_{k=2}^{n} \frac{\delta}{k^2}}$$

Repeating the same reasoning, we can estimate $|(\psi - \mathbb{I})^{\{n-1\}}|_{\rho - \sum_{k=2}^{n} \frac{\delta}{k^2}}$

$$|(\psi - \mathbb{I})^{\{n-1\}}|_{\rho - \sum_{k=2}^{n} \frac{\delta}{k^2}} \leq |\nabla(\psi - \mathbb{I})^{\{n-2\}}|_{\rho - \sum_{k=2}^{n} \frac{\delta}{k^2}} |(\psi - \mathbb{I})|_{\rho - \sum_{k=2}^{n} \frac{\delta}{k^2}}$$

and

$$|\nabla(\psi - \mathbb{I})^{\{n-2\}}|_{\rho - \sum_{k=2}^{n} \frac{\delta}{k^2}} \leq 2\frac{2^2}{\delta} |(\psi - \mathbb{I})^{\{n-2\}}|_{\rho - \sum_{k=3}^{n} \frac{\delta}{k^2}}$$

If we join all the above estimates, we obtain

$$\begin{aligned}|(\psi-\mathbb{I})^{\{n\}}|_{\rho-\sum_{k=1}^{n}\frac{\delta}{k^2}} \leq & 2^2\frac{1^2\cdot 2^2}{\delta^2}|(\psi-\mathbb{I})^{\{n-2\}}|_{\rho-\sum_{k=3}^{n}\frac{\delta}{k^2}} \\ & |(\psi-\mathbb{I})|_{\rho-\sum_{k=2}^{n}\frac{\delta}{k^2}}|(\psi-\mathbb{I})|_{\rho-\sum_{k=1}^{n}\frac{\delta}{k^2}}\end{aligned}$$

Iterating the same estimates, we have

$$|(\psi-\mathbb{I})^{\{n\}}|_{\rho-\frac{\pi^2}{6}\delta} \leq (n!)^2\left(\frac{2}{\delta}|\psi-\mathbb{I}|_\rho\right)^n$$

If we define $\varepsilon^2 := \frac{2}{\delta}|\psi-\mathbb{I}|_\rho$, using the Stirling's approximation we obtain that

$$n!\varepsilon^n \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n\varepsilon^n = \sqrt{2\pi}(n\varepsilon)^{n+\frac{1}{2}}\frac{e^{-n}}{\sqrt{\varepsilon}}$$

hence, if we suppose $\varepsilon^2 < 1$, i.e. the diffeomorphism ψ is near the identity for the sup-norm in the ball B_ρ, then for $n^* := \left[\frac{1}{\varepsilon}\right]$ we have

$$|(\psi-\mathbb{I})^{\{n^*\}}|_{\rho-\frac{\pi^2}{6}\delta} \leq 2\pi\frac{e^{-\frac{2}{\varepsilon}}}{\varepsilon}$$

In conclusion, we can say that the series (3.11), which defines the vector field X, is certainly divergent, but the first n^* terms decrease, so the series is "convergent for astronomers".

Bibliography

[1] Notes 5: The Baker-Campbell-Hausdorff formula, Autumn 2012. https://www.uio.no/studier/emner/matnat/math/MAT4270/h12 (accessed October 3, 2018).

[2] M. Abate and F. Tovena. *Geometria differenziale.* Number 54 in UNITEXT-La Matematica per il 3+2. Springer, Milan, 2011.

[3] M.-C. Arnaud, C. Bonatti, and S. Crovisier. Dynamiques symplectiques génériques. *Ergodic Theory Dynam. Systems*, 25(5):1401–1436, 2005.

[4] H. F. Baker. Alternants and Continuous Groups. *Proc. London Math. Soc. (2)*, 3:24–47, 1905.

[5] A. Banyaga. *The structure of classical diffeomorphism groups*, volume 400 of *Mathematics and its applications.* Kluwer Academic Publisher Group, Dordrecht, 1997.

[6] G. Benettin and A. Giorgilli. On the Hamiltonian interpolation of near-to-the identity symplectic mappings with application to symplectic integration algorithms. *Journal of Statistical Physics*, 74(5-6):1117–1143, Mar. 1994.

[7] S. Biagi and A. Bonfiglioli. An introduction to the geometrical analysis of vector fields. with applications to maximum principles and lie groups. to appear (2019).

[8] S. Biagi and A. Bonfiglioli. On the convergence of the Campbell-Baker-Hausdorff-Dynkin series in infinite-dimensional Banach-Lie algebras. *Linear Multilinear Algebra*, 62(12):1591–1615, 2014.

[9] S. Blanes and F. Casas. On the convergence and optimization of the Baker-Campbell-Hausdorff formula. *Linear Algebra and its Applications*, 378:135–158, Feb. 2004.

[10] A. Bonfiglioli and R. Fulci. *Topics in noncommutative algebra. The theorem of Campbell, Baker, Hausdorff and Dynkin.* Number 2034 in Lecture notes in mathematics. Springer, Heidelberg, 2012.

[11] N. Bourbaki. *Lie groups and Lie algebras: chapters 1 - 3.* Elements of mathematics. Springer-Verlag, Berlin, Berlin, softcover ed. of the 2. printing 1989 edition, 1998. Translated from the French, Reprint of the 1989 English translation.

[12] J. E. Campbell. On a Law of Combination of Operators (second paper). *Proc. Lond. Math. Soc.*, 29:14–32, 1898.

[13] A. Deprit. Canonical transformations depending on a small parameter. *Celestial Mechanics*, 1:12–30, Mar. 1969/1970.

[14] E. B. Dynkin. On the representation by means of commutators of the series $\log(e^x e^y)$ for noncommutative x and y. *Mat. Sbornik N.S.*, 25(67):155–162, 1949.

[15] C. Efthymiopoulos. Computational perturbative methods for hamiltonian systems and their applications. October 15, 2018.

[16] A. Giorgilli. Perturbation methods in celestial mechanics. August 28-September 2, 2017. Lectures at "SDSM 2017 – Satellite Dynamics and Space Missions: Theory and Applications of Celestial Mechanics".

[17] W. Gröbner. *Serie di Lie e loro applicazioni.* Poliedro. Cremonese, 1975.

[18] B. C. Hall. *Lie groups, Lie algebras, and representations.* Number 222 in Graduate Texts in Mathematics. Springer-Verlag, New York, 2003.

[19] F. Hausdorff. Die symbolische Exponentialformel in der Gruppentheorie. *Leipz. Ber.*, 58:19–48, 1906.

[20] J. Henrard and J. Roels. Equivalence for Lie transforms. *Celestial Mech.*, 10:497–512, 1974.

[21] S. Kuksin and J. Pöschel. On the inclusion of analytic symplectic maps in analytic Hamiltonian flows and its applications. In *Seminar on Dynamical Systems (St. Petersburg, 1991)*, volume 12 of *Progr. Nonlinear Differential Equations Appl.*, pages 96–116. Birkhäuser, Basel, 1994.

[22] J. M. Lee. *Manifolds and differential geometry*, volume 107 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2009.

[23] P. Michor. Old and new on diffeomorphism groups too. 25 September - 1 October 2016. 50th Seminar ”Sophus Lie”.

[24] J. Moser. *Lectures on Hamiltonian systems*. Number 81 in Mem. Amer. Math. Soc. Amer. Math. Soc., Providence, 1968.

[25] A. I. Neĭshtadt. The separation of motions in systems with rapidly rotating phase. *Prikl. Mat. Mekh.*, 48(2):197–204, 1984.

[26] J. Palis. Vector fields generate few diffeomorphisms. *Bulletin of the American Mathematical Society*, 80(3):503–506, May 1974.

[27] H. Poincaré. *Les méthodes nouvelles de la mécanique céleste. Tome II. Méthodes de MM. Newcomb, Gyldén, Lindstedt et Bohlin.* Dover Publications, Inc., New York, N.Y., 1957.

[28] L. Polterovich and E. Shelukhin. Autonomous Hamiltonian flows, Hofer's geometry and persistence modules. *Selecta Math. (N.S.)*, 22(1):227–296, 2016.

[29] M. Postnikov. *Lie groups and Lie algebras*. "Mir", Moscow, 1986. Lectures in geometry. Semester V, Translated from the Russian by Vladimir Shokurov.

[30] R. Schmid. Infinite-Dimensional Lie Groups and Algebras in Mathematical Physics. *Advances in Mathematical Physics*, 2010:1–35, 2010.

[31] J.-P. Serre. *Lie algebras and Lie groups*. Number 1500 in Lecture notes in mathematics. Springer-Verlag, Berlin, 2nd. edition, 2006. 1964 Lectures given at Harvard University, Corrected fifth printing of the second (1992) edition.

[32] H. Shniad. The equivalence of Von Zeipel mappings and Lie transformations. *Celestial Mech.*, 2:114–120, 1970/1971.

[33] V. S. Varadarajan. *Lie groups, Lie algebras, and their representations*. Number 102 in Graduate Texts in mathematics. Springer-Verlag, New York, 1984.