sciendo
Control and Cybernetics
vol. **53** (2024) No. 1
pages: 43-75
DOI: 10.2478/candc-2024-004

Shape and topological derivatives as Hadamard semidifferentials[1]

by

Michel C. Delfour

Centre de recherches mathématiques and
Département de mathématiques et de statistique,
Université de Montréal, C.P. 6128, succ. Centre-ville, Montreal,
Canada H3C 3J7,
michel.delfour@umontreal.ca

Abstract: The object of this paper is to further investigate the notion of *shape* and *topological derivatives* in the light of the general notion of *Hadamard semidifferential* for a function defined on a subset of a topological vector space.

The use of semitrajectories and the characterization of the *adjacent tangent cone* provide simple tools for defining Hadamard semi-differentials and differentials without a priori introduction of geometric structures such as, for instance, a differential manifold. Such a simple notion retains all the operations of the classical differential calculus, including the chain rule, for a large class of non-differentiable functions, in particular, the norms and the convex functions. It also provides a direct access to functions defined on a lousy set or a manifold with boundary.

This direct approach is first illustrated in the context of the classical *matrix subgroups* of the general linear group $\mathrm{GL}(n)$ of invertible $n \times n$ matrices, which are the prototypes of *Lie groups*. For the shape derivative we have groups of diffeomorphisms of the Euclidean space $\mathbb{R}^n$ with the *composition* operation, and the adjacent tangent cone is a linear space; for the topological derivative we have the group of *characteristic functions* with the *symmetric difference* operation and the adjacent tangent cone is only a cone at some points.

Keywords: Hadamard semidifferential, adjacent tangent cone, shape derivative, topological derivative, matrix group, diffeomorphism group, characteristic function group, automatic differentiation, deep learning

[1]Submitted: March 2024; Accepted: July 2024.

1. Introduction

The object of this paper is to further investigate the notion of *shape* and *topological derivatives* in the light of the general notion of *Hadamard semidifferential* for a function defined on a subset of a topological vector space (Delfour, 2020a,b)).

In Differential Geometry, defining the differential of a function, which is defined on a subset of a topological vector space, requires the *a priori* specification of some differential structure with a linear tangent space at each point of the subset.

The use of semitrajectories in a subset of a topological vector space and the characterization of the *adjacent tangent cone* provide simple tools to define Hadamard semidifferentials and differentials without *a priori* introducing a differential manifold structure and without the requirment that the semitrajectories be smooth or even continuous. This can be seen as a relaxation of the notion of *Lie derivative* (Marsden and Ratiu, 1994, Sec. 4.3, pp. 120–121). The notion of Hadamard semidifferential retains all the operations of the classical differential calculus, including the chain rule for a large class of non-differentiable functions, such as the norms and convex functions (Delfour, 2023b). It also provides a direct access to functions, defined on a lousy set or a manifold with boundary such as a finite line in dimension two, where the adjacent tangent cone is a half line at both ends.

The approach is first illustrated in the context of *matrix subgroups* of the general linear group $\mathrm{GL}(n)$ of invertible $n \times n$ matrices, which are the prototypes of *Lie groups*. This simple, quick, and direct approach was advocated in Lange (2024).

For shape derivatives, the subset is a group of diffeomorphisms of mappings from the Euclidean space $\mathbb{R}^n$ into itself with the *composition* operation, and the adjacent tangent cone is a linear space. For the topological derivative, the subset is the group of *characteristic functions* $\mathrm{X}(\mathbb{R}^n)$ with the *symmetric difference* operation, but the adjacent tangent cone is only a cone at some points (Delfour, 2018a).

The emerging point of view is to consider the elements of the group $\mathrm{X}(\mathbb{R}^n)$ of characteristic functions χ_Ω of Lebesgue measurable subsets $\Omega \subset \mathbb{R}^n$ as a subset of the space of distributions $\mathcal{D}(\mathbb{R}^n)'$. It is conjectured that the whole adjacent tangent cone $T_{\chi_\Omega}(\mathrm{X}(\mathbb{R}^n))$ is contained in the ambient space of distributions $\mathcal{D}(\mathbb{R}^n)'$. In Section 6.2 the compact d-rectifiable subsets E generate *semitangents* that are Radon measures; in Section 6.3 the *Velocity Method* generates *tangents* that are distributions in $H^1_0(\mathbb{R}^n)'$ (Delfour, 2016, sec 3.1. pp. 234–235, Delfour 2018a, sec. 4.1, pp. 967–968). As a result, the adjacent tangent cone $T_{\chi_\Omega}(\mathrm{X}(\mathbb{R}^n))$ is not a linear space and it does not only contain measures, but we do not know how large it is.

Section 2 of the present paper recalls the main definitions and theorems from Delfour (2020b, sec. 4) for the Hadamard semidifferential of a function defined on a subset of a topological vector space (TVS). It also briefly discusses the relation between the Hadamard semidifferential and numerical aspects, such as, for instance, *automatic differentiation*. Section 3 considers linear mappings from $\mathbb{R}^n$ to $\mathbb{R}^n$, that is, $n \times n$ matrices, as a preliminary to nonlinear mappings. Section 4 considers the *Stiefel manifold* of rectangular $n \times d$ matrices. Section 5 considers *diffeomorphism groups* of nonlinear mappings and revisits the notion of shape derivative in Section 5.2. Section 6 is devoted to the group of characteristic functions and its application to both topological and shape derivatives.

2. Hadamard semidifferential on subsets of a TVS

2.1. Definitions and notation

The use of semitrajectories in a subset of a topological vector space and the characterization of the adjacent tangent cone provide a simple tool to define Hadamard semidifferentials and differentials without a priori introducing the geometric structures such as, for instance, a differential manifold. Such notions retain all the operations of the classical differential calculus, including the chain rule for a large class of non-differentiable functions, such as the norms and the convex functions. It is remarkable that a few relatively simple notions can embrace a whole range of problems arising in the applications of Mathematics and Statistics to Science and Engineering. We recall definitions from Delfour (2020b, sec. 4.)

Definition 1 *Let $A \neq \varnothing$ be a subset of a topological vector space (TVS) X. An* admissible semitrajectory[1] *at $x \in A$ in A is a function $h : [0, \tau) \to A$, $\tau > 0$, such that*

$$h(0) = x \quad \textit{and} \quad h'(0^+) \overset{\text{def}}{=} \lim_{t \searrow 0} \frac{h(t) - h(0)}{t} \textit{ exists in } X.$$

$h'(0^+)$ is the semitangent *to the trajectory h in A at $h(0) = x$.*

Using semitrajectories with a semitangent instead of trajectories with a tangent allows for the sets A, which are not smooth or not connected or manifolds with boundaries.

[1] The term *semitrajectory* in $[0, \tau]$ is preferred over *path*, which is often assumed to be continuous or differentiable in an open interval $(-\tau, \tau)$ around 0. Such paths would not be appropriate to obtain the semidifferential of the absolute value or the norm at the origin.

Definition 2 *Let $A \neq \varnothing$ be a subset of a topological vector space X. The* adjacent (or intermediary) tangent cone[2] *to A at $x \in A$ is defined as*

$$T_x^\flat(A) \overset{\text{def}}{=} \left\{ v \in X : \forall \{t_n \searrow 0\},\, \exists \{x_n\} \subset A \text{ such that } \lim_{n\to\infty} \frac{x_n - x}{t_n} = v \right\}.$$

We use the notation $\mathbb{R}_+ \overset{\text{def}}{=} \{\lambda \in \mathbb{R} : \lambda \geq 0\}$, and $T_x^\flat(A) = \overline{\mathbb{R}_+(A-x)}$ for a convex subset A of a *Fréchet space* (that is, a topological vector space, which is complete, metrizable, and locally convex).

$T_x^\flat(A)$ is related to the notion of admissible semitrajectories at x in A and this provides an equivalent alternative way to characterize $T_x^\flat(A)$.

Theorem 1 *Let A be a subset of a topological vector space X. For $x \in A$,*

$$T_x^\flat(A) = \left\{ h'(0^+) : h \text{ an admissible semitrajectory in } A \text{ at } x \right\}.$$

We now have all the elements, which are needed to extend the definition of the Hadamard semidifferential to a subset of a TVS.

Definition 3 *Let X and Y be TVS, A, $\varnothing \neq A \subset X$, and $f : A \to Y$.*

(i) *The function f is* Hadamard semidifferentiable at $x \in A$ in the direction $v \in T_x^\flat(A)$ *if there exists $g(x,v) \in Y$ such that, for all admissible semitrajectories h in A at x such that $h'(0^+) = v$,*

$$(f \circ h)'(0^+) \overset{\text{def}}{=} \lim_{t \searrow 0} \frac{f(h(t)) - f(h(0))}{t} = g(x,v) \text{ in } Y. \tag{1}$$

The element $g(x,v)$ will be denoted $d_H f(x;v)$.

(ii) *f is* Hadamard semidifferentiable at $x \in A$ *if f is Hadamard semidifferentiable at x in all directions $v \in T_x^\flat(A)$.*

(iii) *f is* Hadamard differentiable[3] at $x \in A$ *if $T_x^\flat(A)$ is a linear subspace, f is Hadamard semidifferentiable at $x \in A$, and the function $v \mapsto d_H f(x;v) : T_x^\flat(A) \to Y$ is linear, in which case it will be denoted $Df(x)$.*

In finite dimensions Definition 3 applies to sets A, which are not differentiable manifolds, and to non-differentiable functions, as this is seen from the next example.

[2]The natural tangent cone, associated with a Hadamard semidifferentiable function, defined on a subset of a topological vector space (see Delfour, 2020b) is the *adjacent tangent cone* defined in the book of Aubin and Frankowska (1990, p. 128) in terms of sequences rather than its equivalent definition (Aubin and Frankowska, 1990, Dfn. 4.1.5, p. 127) in a normed vector space via the distance function.

[3]In finite dimension, Hadamard diferentiability is equivalent to Fréchet differentiability, but in infinite dimensions it is weaker, as recognized by Fréchet (1937).

EXAMPLE 1 *Consider the boundary* ∂S *of the unit square* $S = \{(x_1, x_2) : 0 \leq x_i \leq 1\}$. *It is not a differentiable manifold at the four corners. At* $(0,0)$ *one cannot find a trajectory* $h : (-\tau, \tau) \to \partial S$ *through* $h(0) = (0,0)$ *such that* $h'(0)$ *exists, but you can find semitrajectories* $h : [0, \tau) \to \partial S$ *through* $h(0) = (0,0)$ *such that* $h'(0^+)$ *exists. This one-sided derivative* $h'(0^+)$ *defines a semitangent at* $h(0) = x$. *The set of all such semitangents is the non-convex adjacent tangent cone* $T^\flat_{(0,0)}(\partial S)$. *If the calculation is repeated at a point* x, *which is not a corner, the computation will reveal that* $T^\flat_x(\partial S)$ *is the linear subspace* $\mathbb{R}$.

As an example, the non-differentiable function $f(x_1, x_2) = |x_1| + |x_2|$ *on* ∂S *is Hadamard semidiferentiable at* $(0,0)$ *for directions in the adjacent tangent cone* $T^\flat_{(0,0)}(\partial S) = \{(0, v_2) : v_2 \geq 0\} \cup \{(v_1.0) : v_1 \geq 0\}$ *at* $(0,0)$.

Definition 3 also applies to embedded smooth submanifolds, groups and subgroups of matrices, and groups of diffeomorphisms and characteristic functions encountered in shape and topological optimization. It can be seen as a relaxation of the *Lie derivative* (Marsden and Ratiu, 1994, Sec. 4.3, pp. 120–121) which *a priori* requires some form of differential manifold structure.

REMARK 1 *For a* C^k, $k \geq 1$, *submanifold* S *of* $\mathbb{R}^n$ *of dimension* $d \leq n$, *the tangent space is usually defined from trajectories* $h : (-\tau, \tau) \to S$ *through the point* $h(0) = x \in S$ *such that the derivative* $h'(0)$ *exists. This derivative (when it exists) defines a tangent line through* $h(0) = x$. *But, in general, this is not sufficient to characterize the whole tangent space and to show that it is linear. The information that* S *is a submanifold of* $\mathbb{R}^n$ *is necessary to conclude that* $T^\flat_x(S)$ *is a linear subspace of* $\mathbb{R}^n$. *The issue is the same for a differentiable manifold, but its definition and the definition of the tangent space are more complicated.*

The Hadamard semidifferentiability enjoys all the nice properties of the classical finite dimensional differential calculus, including the chain rule.

THEOREM 2 *Let* X *and* Y *be topological vector spaces and* A, $\varnothing \neq A \subset X$.

(i) *If* $f : A \to Y$ *is Hadamard semidifferentiable at* $x \in A$ *in the direction* $v \in T^\flat_A(x)$, *then for all admissible semitrajectory* h *in* A, *such that* $h'(0^+) = v$, $f \circ h$ *is an admissible trajectory in* $f(A)$ *such that* $(f \circ h)'(0^+) = d_H f(x; v) \in T^\flat_{f(A)}(f(x))$. *The mapping*

$$v \mapsto d_H f(x; v) : T^\flat_A(x) \to T^\flat_{f(A)}(f(x)) \subset Y \tag{2}$$

is sequentially continuous for the induced topologies.

(ii) *If* $f_1 : A \to Y$ *and* $f_2 : A \to Y$ *are Hadamard semidifferentiable at* $x \in A$ *in the direction* $v \in T^\flat_A(x)$, *then for all* α *and* β *in* $\mathbb{R}$,

$$d_H(\alpha f_1 + \beta f_2)(x; v) = \alpha\, d_H f_1(x; v) + (1 - \alpha)\, d_H f_2(x; v), \tag{3}$$

and $\alpha f_1 + \beta f_2$ is Hadamard semidifferentiable at x in the direction v.

(iii) *(Chain rule) Let X, Y, Z be topological vector spaces, $A \subset X$, $g : A \to Y$, and $f : g(A) \to Z$ be functions such that g is Hadamard semidifferentiable at x in the direction $v \in T_A^\flat(x)$ and f is Hadamard semidifferentiable at $g(x)$ in $g(A)$ in the direction $d_H g(x;v)$. Then, $d_H g(x;v) \in T_{g(A)}^\flat(x)$, $f \circ g$ is Hadamard semidifferentiable at x in the direction $v \in T_A^\flat(x)$, and*

$$d_H(f \circ g)(x;v) = d_H f(g(x); d_H g(x;v)). \tag{4}$$

The next question is to address the continuity of a Hadamard semidifferentiable function.

Theorem 3 *Let X and Y be topological vector spaces, A a non-empty subset of X, and $f : A \to Y$ a function. Assume that f is Hadamard semidifferentiable at $x \in A$.*

(i) *If there exists a bounded*[4] *neighborhood $U(0) \in \mathscr{R}$ in X, then f is sequentially continuous*[5] *at x in A for the induced topology on A.*

(ii) *If X is a Fréchet space, then $v \mapsto d_H f(x;v) : T_A^\flat(x) \to T_{f(A)}^\flat(f(x))$ is positively homogeneous and continuous for the induced topologies. If X and Y are Fréchet spaces, then f is continuous at x.*

2.2. Some comments and perspectives

2.2.1. Numerical global optimization

In the context of semidifferentials and tangent cones, numerical methods have to be adapted, but we are not starting from scratch. Working with a cone means that some directions will be "taboo" as in global optimization (see, for instance, Ji and Klinowski, 2006). Descent methods for Lipschitz continuous functions, which are only Hadamard semidifferentiable at the minimum, work well, since they are differentiable almost everywhere and the probability that the computer

[4]Recall that in a *topological vector space* there is a *fundamental system* $\mathscr{R}$ of neighborhoods of the origin, for which (i) every V in $\mathscr{R}$ is absorbing and balanced, and (ii) for every $V \in \mathscr{R}$, there exists $U \in \mathscr{R}$ such that $U + U \subset V$. In this paper, we always assume that the *neighborhoods* are elements of $\mathscr{R}$. A set A is bounded if, for all $V(0) \in \mathscr{R}$, there exists $\alpha > 0$ such that $A \subset \lambda V(0)$ for all $\lambda \geq \alpha$ (Horváth, 1966, Dfn. 1, p. 108).

[5]Note the following natural equivalence for the semicontinuity in terms of semitrajectories. Let X and Y be topological spaces and A a subset of X. A function $f : A \to Y$ is *sequentially continuous* at $a \in A$ if and only if for all *semitrajectories* $h : [0, \tau) \to A$

$$\lim_{t \searrow 0} h(t) = a \quad \Rightarrow \quad \lim_{t \searrow 0} f(h(t)) = f(a), \tag{5}$$

where A is endowed with the topology induced by X.

hits a point of non-differentiability is very very low (see, for instance, Delfour and Huot-Chantal, 2019, and Huot-Chantal, 2018, "On the figure of columns of Lagrange").

2.2.2. Automatic differentiation

By its very definition, the Hadamard semidifferentiabilty is closely related to differentiation along trajectories as in *automatic differentiation*. Quoted from Lange (2024, p. 58):

> Numerical computation of semidifferentials can be accomplished via automatic differentiation, which implements the rules of the differential calculus, including the chain rule, at specific points in parameter space (see Baydin et al., 2018, and Neidinger, 2010). Automatic differentiation is computationally fast and relieves humans of the tedium of manual differentiation. It is hardly surprising that the machine learning community has embraced automatic differentiation.
>
> The forward version of automatic differentiation is better adapted to semidifferentials than the backward version. The backward version is geared toward the computation of gradients, but these no longer exist once linearity is abandoned. Computation of forward directional derivatives is straightforward because it dispenses with gradients. Modern computer languages such as Julia simplify the coding of forward mode automatic semidifferentiation.

2.2.3. Stochastic gradient methods and deep learning

Some authors have introduced generalized derivatives to provide nonsmooth approaches with a flexible calculus in the context of stochastic gradient methods and deep learning. For instance, Bolte and Pauwels (2021) introduced generalized derivatives called *conservative fields*, for which they develop a calculus and provide representation formulas. Functions having a conservative field are called *path differentiable*: convex, concave, Clarke regular and any semialgebraic Lipschitz continuous functions are path differentiable. But such functions are all Hadamard semidifferentiable and the chain rule of Theorem 2 is built in the Definition 3 of a Hadamard semidifferential.

3. Square matrices and their groups and subgroups

This section considers linear mappings from $\mathbb{R}^n$ to $\mathbb{R}^n$ as a preliminary to nonlinear mappings $F : \mathbb{R}^n \to \mathbb{R}^n$, which belong to some topological vector space

Θ, such as the spaces of m-times differentiable mappings with compact support $C_c^m(\mathbb{R}^n, \mathbb{R}^n)$, m-times continuously differentiable mappings $C^m(\mathbb{R}^n, \mathbb{R}^n)$, $0 \le m \le \infty$, (m, ℓ)-times Hölderian mappings $C^{m,\ell}(\mathbb{R}^n, \mathbb{R}^n)$, $m \ge 0$, $0 \le \ell \le 1$ (see Delfour and Zolésio, 2011, Chpt. 3).

3.1. The $n \times n$ matrices and their determinant

The vector space of $n \times n$ matrices with the matrix addition and multiplication by a scalar is denoted $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$. Endowed with the *Frobenius inner product* and *Frobenius norm*

$$\forall A, B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n), \quad A \cdot\cdot B \stackrel{\text{def}}{=} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, b_{ij}, \quad \|A\|_2 = \sqrt{A \cdot\cdot A}, \tag{6}$$

tt is a Hilbert space identified with the n^2-dimensional Euclidean space $\mathbb{R}^{n^2}$.

The *matrix multiplication* is a well-defined operation in that space. It corresponds to the composition $(A \circ B)(x) = A(B(x))$ of the linear mappings from $\mathbb{R}^n$ into $\mathbb{R}^n$, generated by $n \times n$ matrices. It contains the identity matrix I_n as the neutral element.

All convergences in the construction of the adjacent tangent cone take place in $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ and not in some metric topology that can be defined on a group or subgroup of matrices to make it complete.

3.1.1. Determinant and matrix of cofactors

Let A be an $n \times n$ matrix whose entries are denoted $\{a_{i,j}\}$. In the sequel it will be useful to write the matrix A in terms of its column vectors a_j, $(a_j)_i = a_{i,j}$, $1 \le j \le n$: $A = [a_1, \ldots, a_n]$. Recall the definition of the $n \times n$ *cofactor matrix* $\operatorname{Cof} A$

$$(\operatorname{Cof} A)_{ij} = (-1)^{i+j} \det A_{i,j}, , \quad 1 \le i \le n,\, 1 \le j \le n, \tag{7}$$

where $A_{i,j}$ is the matrix obtained by removing its ith line and its jth column:

$$A_{i,j} \stackrel{\text{def}}{=} \begin{bmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{bmatrix}.$$

With the above notation

$$\det A \overset{\text{def}}{=} \sum_{i=1}^{n} (a_j)_i \, (\mathrm{Cof} A)_{ij} = \sum_{i=1}^{n} a_{i,j} (\mathrm{Cof} A)_{ij} = A \cdot\cdot \, \mathrm{Cof} A = \mathrm{Cof} A \cdot\cdot \, A \quad (8)$$

by using the *Frobenius scalar product*. Finally, the *trace* of A is defined as

$$\mathrm{tr}(A) \overset{\text{def}}{=} \sum_{i=1}^{n} a_{i,i} = A \cdot\cdot \, I_n. \quad (9)$$

With this notation

$$\det A = \mathrm{tr}(A^\top \, \mathrm{Cof} A). \quad (10)$$

3.1.2. The differential of the determinant and of its absolute value

As the determinant $\det(A)$ is a polynomial function of the entries, it is (Hadamard) differentiable, that is, $d_H \det(A; B)$ exists for any $n \times n$ matrix B and the function $B \mapsto d_H \det(A; B) : \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \to \mathbb{R}$ is linear. As a result, it is sufficient to find its semidifferential at A in the direction B

$$\forall B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n), \quad d \det(A; B) \overset{\text{def}}{=} \lim_{t \searrow 0} \frac{\det(A + tB) - \det A}{t}. \quad (11)$$

THEOREM 4 (DELFOUR, 2020A, EXAMPLE 3.10, CHAP. 3, P. 108–109)

(i) *The function $A \mapsto \det A : \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \to \mathbb{R}$ is (Hadamard) differentiable and*

$$\forall B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n), \quad d_H \det(A; B) = (\mathrm{Cof} A) \cdot\cdot \, B = \mathrm{tr}((\mathrm{Cof} A)^\top B). \quad (12)$$

(ii) *The function $A \mapsto |\det A| : \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \to \mathbb{R}$ is Hadamard semidifferentiable at A and, for all $B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$,*

$$d_H |\det|(A; B) = \begin{cases} \dfrac{\det(A)}{|\det(A)|} \, (\mathrm{Cof} A) \cdot\cdot \, B, & \text{if } \det(A) \neq 0, \\ |(\mathrm{Cof} A) \cdot\cdot \, B|, & \text{if } \det(A) = 0. \end{cases} \quad (13)$$

REMARK 2 *In the change of variable formula $x \mapsto Ax$ for an integral over $\mathbb{R}^n$, it is not $\det(A)$ but $|\det(A)|$ that appears in the formula.*

REMARK 3 *If the matrix A is invertible, then*

$$A^{-1} = (\det A)^{-1} (\mathrm{Cof} A)^\top$$

$$d_H \det(A; B) = \det(A) \, [A^{-1}]^\top \cdot\cdot \, B = \det(A) \, I_n \cdot\cdot \, A^{-1} B = \det(A) \, \mathrm{tr}(A^{-1} B).$$

Therefore, if A is invertible, the absolute value of the determinant at A is Hadamard differentiable.

The proof of Theorem 4 requires the following useful lemma.

LEMMA 1 *Given the matrices* $A = \{a_{ij}\}$ *and* $B = \{b_{ij}\}$ *in* $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ *and the column vectors* a_j, $(a_j)_i = a_{ij}$, *and* b_j, $(b_j)_i = b_{ij}$, $1 \le j \le n$,

$$\det B - \det A = \sum_{j=1}^{n} \det[a_1, \ldots, a_{j-1}, b_j - a_j, b_{j+1}, \ldots, b_n]. \tag{14}$$

PROOF OF LEMMA 1 The proof rests on the fact that the function

$$a_j \mapsto \det[a_1, \ldots a_{j-1}, a_j, a_{j+1}, \ldots, a_n]$$

is linear. The formula is true for $n = 2$:

$$\begin{aligned} \det B - \det A &= \det[b_1, b_2] - \det[a_1, a_2] \\ &= \det[b_1 - a_1, b_2] + \det[a_1, b_2] - \det[a_1, a_2] \\ &= \det[b_1 - a_1, b_2] + \det[a_1, b_2 - a_2]. \end{aligned}$$

Assume that it is true for $n - 1$ and compute the difference

$$\begin{aligned} &\det B - \det A \\ &= \sum_{i=1}^{n} (-1)^{n+i} b_{in} \det[b_1, \ldots, b_{n-1}] - (-1)^{n+i} a_{in} \det[a_1, \ldots, a_{n-1}] \\ &= \sum_{i=1}^{n} (-1)^{n+i} [b_{in} - a_{in}] \det[a_1, \ldots, a_{n-1}] \\ &\quad + \sum_{i=1}^{n} (-1)^{n+i} b_{in} [\underbrace{\det[b_1, \ldots, b_{n-1}] - \det[a_1, \ldots, a_{n-1}]}_{=\sum_{j=1}^{n-1} \det[a_1, \ldots, a_{j-1}, b_j - a_j, b_{j+1}, \ldots, b_{n-1}]}] \end{aligned}$$

since the property is valid for $n - 1$. Now, rearranging the two pieces

$$\begin{aligned} &\det B - \det A \\ &= \det[b_1, \ldots, b_{n-1}, b_n - a_n] \\ &+ \sum_{j=1}^{n-1} \sum_{i=1}^{n} (-1)^{n+i} b_{in} \det[a_1, \ldots, a_{j-1}, b_j - a_j, b_{j+1}, \ldots, b_{n-1}] \\ &= \det[b_1, \ldots, b_{n-1}, b_n - a_n] + \sum_{j=1}^{n-1} \det[a_1, \ldots, a_{j-1}, b_j - a_j, b_{j+1}, \ldots, b_n] \\ &= \sum_{j=1}^{n} \det[a_1, \ldots, a_{j-1}, b_j - a_j, b_{j+1}, \ldots, b_n]. \end{aligned}$$

■

Proof of Theorem 4 (i) Since the function $A \mapsto \det(A)$ is polynomial, it is (Hadamard) differentiable and it is sufficient to find the one-sided directional derivative $d\det(A;B)$ at A in the direction B in (11). From Lemma 1

$$\det(A+tB) - \det A = \sum_{j=1}^{n} \det[a_1, \dots, a_{j-1}, t\, b_j, (a+tb)_{j+1}, \dots, (a+tb)_n].$$

Upon dividing by $t > 0$ and letting $t \to 0$

$$\begin{aligned}
&\frac{\det(A+tB) - \det A}{t} \\
&= \frac{1}{t} \sum_{j=1}^{n} \sum_{i=1}^{n} (-1)^{i+j}\, t\, b_{ij}\, \det[a_1, \dots, a_{j-1}, (a+tb)_{j+1}, \dots, (a+tb)_n] \\
&= \sum_{j=1}^{n} \sum_{i=1}^{n} b_{ij}\, (-1)^{i+j} \det[a_1, \dots, a_{j-1}, (a+tb)_{j+1}, \dots, (a+tb)_n] \\
&\to \sum_{j=1}^{n} \sum_{i=1}^{n} b_{ij}\, (-1)^{i+j} \det[a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_n] \text{ as } t \searrow 0
\end{aligned}$$

and we obtain the formula

$$d\det(A;B) = \sum_{j=1}^{n} \sum_{i=1}^{n} b_{ij}\, (\mathrm{Cof} A)_{ij} = B \cdot\cdot\, \mathrm{Cof} A = \mathrm{tr}((\mathrm{Cof} A)^\top B).$$

(ii) By using the chain rule for the composition, we get

$$d_H |\det|(A;B) = \begin{cases} \dfrac{\det(A)}{|\det(A)|} \mathrm{Cof} A \cdot\cdot\, B, & \text{if } \det(A) \neq 0, \\ |\mathrm{Cof} A \cdot\cdot\, B|, & \text{if } \det(A) = 0. \end{cases}$$

■

3.2. The *General Linear Group* GL(n) and the matrix groups

Definition 4 *The* General Linear Group

$$\mathrm{GL}(n) \overset{\text{def}}{=} \{X \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) : X \text{ is invertible}\} \tag{15}$$

is the set of all $n \times n$ *invertible matrices in* $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$.

By Theorem 4, for the determinant at $X \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$

$$\forall Y \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n), \quad d_H \det(X;Y) = (\mathrm{Cof} X) \cdot\cdot\, Y = \mathrm{tr}((\mathrm{Cof} X)^\top Y)$$

and for the determinant at $X \in \mathrm{GL}(n)$ in the direction $Y \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$

$$d_H \det(X;Y) = \det(X)\, [X^{-1}]^\top \cdot\cdot\, Y = \det(X)\, \mathrm{tr}(X^{-1} Y).$$

Theorem 5 (i) $\mathrm{GL}(n)$ *is an open subset of* $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ *in the Frobenius matrix norm, the mapping* $X \mapsto X^{-1} : \mathrm{GL}(n) \to \mathrm{GL}(n)$ *is continuous, and* $\mathrm{GL}(n)$ *is a group under matrix multiplication.*

(ii) *For all* $F \in \mathrm{GL}(n)$,

$$T_F^\flat(\mathrm{GL}(n)) = \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \quad \Rightarrow T_{I_n}^\flat(\mathrm{GL}(n)) = \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n). \tag{16}$$

Proof (i) From Rudin (1976, Thm. 9.8(b), Chpt. 9, p. 209).

(ii) Since $\mathrm{GL}(n)$ is an open subset of $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, $T_F^\flat(\mathrm{GL}(n)) = \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ for all $F \in \mathrm{GL}(n)$. In particular, $T_{I_n}^\flat(\mathrm{GL}(n)) = \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) = T_{I_n}^\flat(\mathrm{GL}(n))$. ■

Definition 5 *A* Matrix Group G *is an (algebraically) closed subgroup of* $\mathrm{GL}(n)$, *that is, for all* $F_1, F_2 \in G$, $F_1F_2 \in G$, $I_n \in G$ *and for each* $F \in G$, $F^{-1} \in G$.

By definition, $\mathrm{GL}(n)$ is a matrix group, but $\mathrm{GL}(n)$ is not topologically closed in $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$.

Theorem 6 *Let* G *be a matrix group.*

(i) $T_{I_n}^\flat(G)$ *is a closed linear subspace of* $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ *and*

$$\forall X \in G, \quad \begin{aligned} &T_X^\flat(G) = XT_{I_n}^\flat(G) = T_{I_n}^\flat(G)X \\ &XT_{I_n}^\flat(G)X^{-1} = T_{I_n}^\flat(G). \end{aligned} \tag{17}$$

(ii) *The* Lie bracket

$$(A, B) \mapsto [A, B] : T_{I_n}^\flat(G) \times T_{I_n}^\flat(G) \to T_{I_n}^\flat(G)$$

is well-defined and continuous.

(iii) *For each* $X \in \mathrm{GL}(n)$, *the mappings*

$$Y \mapsto YX^{-1} : T_X^\flat(G) \to T_{I_n}^\flat(G) \text{ and } Y \mapsto X^{-1}Y : T_X^\flat(G) \to T_{I_n}^\flat(G) \tag{18}$$

are bijective and bi-continuous.

Corollary 1 *As a closed subgroup of itself, the conclusions of the theorem are true for* $\mathrm{GL}(n)$, *where, for all* $X \in \mathrm{GL}(n)$, $T_X^\flat(\mathrm{GL}(n)) = T_{I_n}^\flat(\mathrm{GL}(n)) = \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$.

Remark 4 *For a matrix group* G, $T_{I_n}^\flat(G)$ *with the Lie bracket operation*

$$(Z_1, Z_2) \mapsto [Z_1, Z_2]; T_{I_n}^\flat(G) \times T_{I_n}^\flat(G) \to T_{I_n}^\flat(G) \tag{19}$$

is the Lie algebra $\mathfrak{g}$ *of* G *and* G *is a* Lie group. *So, we have obtained those properties without a priori introducing the notion of* differentiable manifold, *which serves to guarantee that the tangent space is a linear subspace and legitimates the use of full (smooth) trajectories or paths over the less demanding semitrajectories.*

PROOF (i) Let $Z_1, Z_2 \in T^\flat_{I_n}(G)$ with admissible semitrajectories h_1 and h_2. Since G is a closed subgroup of $\mathrm{GL}(n)$, the semitrajectory $h(t) = h_1(t)h_2(t)$, $h(0) = h_1(0)h_2(0) = I_n$, belongs to G and

$$h'(0^+) = h_1'(0^+)h_2(0) + h_1(0)h_2(0^+) = Z_1 I_n + I_n Z_2 = Z_1 + Z_2 \in T^\flat_{I_n}(G).$$

For $\alpha \geq 0$ and $Z \in T^\flat_{I_n}(G)$ with admissible semitrajectory h, the semitrajectory $h_\alpha(t) = h(\alpha t)$ belongs to G, and

$$h_\alpha(0) = I_n \text{ and } h_\alpha'(0^+) = \alpha h'(0^+) = \alpha Z \in T^\flat_{I_n}(G).$$

Since G is a closed subgroup of $\mathrm{GL}(n)$, for any admissible semitrajectory $h : [0,\tau[\to G$, the semitrajectory $t \mapsto h^{-1}(t) = h(t)^{-1}$ in $[0,\tau]$ belongs to G,

$$h^{-1}(t) \to h^{-1}(0) = I_n^{-1} = I_n \text{ and } (h^{-1})'(0^+) = -h'(0^+) = -Z \in T^\flat_{I_n}(G).$$

As a result, for $Z_1, Z_2 \in T^\flat_{I_n}(G)$ and $a, b \in \mathbb{R}$, $aZ_1 + bZ_2 \in T^\flat_{I_n}(G)$. So $T^\flat_{I_n}(G)$ is closed as a linear subspace of the finite dimensional space $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$.

For $X \in G$, $Y \in T^\flat_X(G)$, let $h : [0,\tau] \to G$ be an admissible trajectory such that $h(0) = X$ and $h'(0^+) = Y$. Since X is invertible and $(h(t) - X)/t \to Y$,

$$\frac{X^{-1}h(t) - I_n}{t} \to X^{-1}Y \in T^\flat_{I_n}(G) \text{ and } X^{-1}T^\flat_X(G) \subset T^\flat_{I_n}(G)$$
$$\frac{h(t)X^{-1} - I_n}{t} \to YX^{-1} \in T^\flat_{I_n}(G) \text{ and } T^\flat_X(G)X^{-1} \subset T^\flat_{I_n}(G).$$

Conversely, for $X \in G$, $Z \in T^\flat_{I_n}(G)$, and and all admissible semitrajectories $k; [0,\tau] \to G$ such that $k(0) = I_n$ and $k'(0^+) = Z$,

$$\frac{k(t) - I_n}{t} \to Z, \quad \frac{Xk(t) - X}{t} \to XZ \in T^\flat_X(G), \text{ and } XT^\flat_{I_n}(G) \subset T^\flat_X(G)$$
$$\frac{k(t)X - X}{t} \to ZX \in T^\flat_X(G) \text{ and } T^\flat_{I_n}(G)X \subset T^\flat_X(G),$$

which implies that

$$XT^\flat_{I_n}(G) = T^\flat_X(G) \text{ and } T^\flat_{I_n}(G)X = T^\flat_X(G).$$

Therefore, for all $X \in G$, $XT^\flat_{I_n}(G)X^{-1} = T^\flat_{I_n}(G)$ and

$$\forall Z \in T^\flat_{I_n}(G) \quad \Rightarrow \; XZX^{-1} \in T^\flat_{I_n}(G).$$

(ii) Take A and B in $T^\flat_{I_n}(G)$ and an admissible semitrajectory $\alpha : [0,\tau] \to G$ such that $\alpha(0) = I_n$ and $\alpha'(0^+) = A$. Consider the new admissible semitrajectory $h(t) = \alpha(t)B\alpha(t)^{-1}$, $h(0) = B$. From the last identity (17) in part (i),

$h(t) \in T^\flat_{I_n}(G)$ and since $T^\flat_{I_n}(G)$ is a closed linear subspace

$$\frac{\alpha(t)B\alpha(t)^{-1} - B}{t} \in T^\flat_{I_n}(G) \text{ and } h'(0^+) = \lim_{t\searrow 0} \frac{\alpha(t)B\alpha(t)^{-1} - B}{t} \in T^\flat_{I_n}(G)$$
$$\Rightarrow h'(0^+) = \alpha'(0^+)B\alpha(0)^{-1} + \alpha(0)B(\alpha^{-1})'(0^+) = AB - BA \in T^\flat_{I_n}(G),$$

since $(\alpha^{-1})'(0^+) = -\alpha'(0^+) = -A$. Therefore, $[A, B] \in T^\flat_{I_n}(G)$ and the mapping

$$(A, B) \mapsto [A, B] : T^\flat_{I_n}(G) \times T^\flat_{I_n}(G) \to T^\flat_{I_n}(G)$$

is well-defined and continuous.

(iii) Direct consequence from (17) in part (i). ∎

3.3. Examples of matrix groups

We consider three classical examples.

3.3.1. The *Special Linear Group* $SL(n)$

The *special linear group*

$$\mathrm{SL}(n) \overset{\text{def}}{=} \{X \in \mathrm{GL}(n) : \det X = 1\} \tag{20}$$

is a closed subset of $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, which is not bounded for $n > 1$.[6]

Theorem 7 (i) $\mathrm{SL}(n)$ *is a closed subgroup of* $\mathrm{GL}(n)$.

(ii) *The adjacent tangent cone to* $\mathrm{SL}(n)$ *at* $X \in \mathrm{SL}(n)$ *is the closed linear subspace*

$$T^\flat_X(\mathrm{SL}(n)) = \{Y \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) : \mathrm{tr}\,(X^{-1}Y) = 0\} \tag{21}$$

of $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ *and*

$$T^\flat_{I_n}(\mathrm{SL}(n)) = \{Y \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) : \mathrm{tr}\, Y = 0\}. \tag{22}$$

[6]Given a Cauchy sequence $\{X_m\}$ in $\mathrm{SL}(n)$ converging to X in $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, $1 = \det(X_m) \to \det(X)$ by continuity of $A \mapsto \det(A)$. So, $\mathrm{SL}(n)$ is closed. It is bounded in the Frobenius norm for $n = 1$, but not for $n \geq 2$. For instance, for $n = 2$

$$\forall k \in \mathbb{R}, \quad X_k \overset{\text{def}}{=} \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix} \in \mathrm{SL}(2), \quad X_k \cdot\cdot X_k = 1 + k^2 \to \infty \text{ as } k \to \infty$$

and $\mathrm{SL}(n)$ is not bounded and not compact for $n \geq 2$.

(iii) *For each* $X \in \mathrm{SL}(n)$,

$$T_X^\flat(\mathrm{SL}(n)) = X T_{I_n}^\flat(\mathrm{SL}(n)) = T_{I_n}^\flat(\mathrm{SL}(n)) X \tag{23}$$

$$\text{and} \quad X T_{I_n}^\flat(\mathrm{SL}(n)) X^{-1} = T_{I_n}^\flat(\mathrm{SL}(n)). \tag{24}$$

(iv) *The* Lie bracket *mapping*

$$(A, B) \mapsto [A, B] : T_{I_n}^\flat(\mathrm{SL}(n)) \times T_{I_n}^\flat(\mathrm{SL}(n)) \to T_{I_n}^\flat(\mathrm{SL}(n))$$

is well-defined and continuous.

(v) *For each* $X \in \mathrm{SL}(n)$, *the mappings*

$$Y \mapsto Y X^{-1} : T_{\mathrm{SL}(n)}^\flat(X) \to T_{I_n}^\flat(\mathrm{SL}(n))$$
$$\text{and } Y \mapsto X^{-1} Y : T_X^\flat(\mathrm{SL}(n)) \to T_{I_n}^\flat(\mathrm{SL}(n))$$

are bijective and bi-continuous.

Proof (i) $\mathrm{SL}(n)$ is a closed subgroup: $\det I_n = 1$; for each $X \in \mathrm{SL}(n)$ $\det X = 1$ implies $\det X^{-1} = 1$; for $X_1, X_2 \in \mathrm{SL}(n)$, $\det(X_1 X_2) = \det X_1 \, \det X_2 = 1$.

(ii) If $Y \in T_X^\flat(\mathrm{SL}(n))$, then there exists an admissible semitrajectory $X(t) \in \mathrm{SL}(n)$ such that

$$\frac{X(t) - X}{t} \to Y, \quad \text{where } \det(X(t) = \det X = 1.$$

For the column vectors of $X_j(t)$, $[X_j(t)]_i = X(t)_{ij}$, and of X, X_j, $[X_j]_i = X_{ij}$,

$$\det X(t) = \det[X(t)_1, \dots X(t)_{j-1}, X(t)_j, X(t)_{j+1}, \dots, X(t)_n]$$
$$\det X = \det[X_1, \dots X_{j-1}, X_j, X_{j+1}, \dots, X_n].$$

From Lemma 1 we have the formula

$$\det X(t) - \det X = \sum_{j=1}^{n} [\det[X_1, \dots X_{j-i}; X(t)_j - X_j, X(t)_{j+1}, \dots, X(t)_n].$$

In particular

$$\begin{aligned} &\det X(t) - \det X \\ &= \sum_{j=1}^{n} \sum_{i=1}^{n} (-1)^{i+j} (X(t)_{ij} - X_{ij}) \, \det[X_1, \dots X_{j-1}, X(t)_{j+1}, \dots, X(t)_n] \\ &= \sum_{j=1}^{n} \sum_{i=1}^{n} (X(t)_{ij} - X_{ij}) \, (-1)^{i+j} \det[X_1, \dots X_{j-1}, X(t)_{j+1}, \dots, X(t)_n]. \end{aligned}$$

Dividing by $t>0$ and letting n go to infinity

$$\begin{aligned}
&\frac{\det X(t)-\det X}{t}\\
&=\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{X(t)_{ij}-X_{ij}}{t}\,(-1)^{i+j}\det[X_1,\dots X_{j-1},X(t)_{j+1},\dots,X(t)_n]\\
&\quad\to\sum_{j=1}^{n}\sum_{i=1}^{n}Y_{ij}\,(-1)^{i+j}\det[X_1,\dots X_{j-1},X_{j+1},\dots,X_n] \text{ as } n\to\infty\\
&\quad=\sum_{j=1}^{n}\sum_{i=1}^{n}Y_{ij}\,(\mathrm{Cof}X)_{ij}=Y\cdot\cdot\,\mathrm{Cof}X=\mathrm{tr}(Y^\top\mathrm{Cof}X)
\end{aligned}$$

and we obtain the general formula

$$\lim_{n\to\infty}\frac{\det X(t)-\det X}{t}=(\mathrm{Cof}X)\cdot\cdot\,Y=\mathrm{tr}((\mathrm{Cof}X)^\top Y).$$

In our case X and $X(t)$ are invertible, $\det X(t)=\det X=1$, and

$$\begin{aligned}
&X^{-1}=(\mathrm{Cof}X)^\top\det X=(\mathrm{Cof}X)^\top\\
&\Rightarrow\ 0=\mathrm{tr}(Y^\top\mathrm{Cof}X)=\mathrm{tr}(Y^\top X^{-\top})=\mathrm{tr}(X^{-1}Y)
\end{aligned}$$

and for $X\in\mathrm{SL}(n)$

$$T_X^\flat(\mathrm{SL}(n))\subset\{Y\in\mathcal{L}(\mathbb{R}^n,\mathbb{R}^n):\mathrm{tr}(X^{-1}Y)=0\}.$$

Conversely, take $X\in\mathrm{SL}(n)=\{X\in\mathrm{GL}(n):\det X=1\}$ and $Y\in\mathcal{L}(\mathbb{R}^n,\mathbb{R}^n)$ such that $\mathrm{tr}(X^{-1}Y)=0$. For $t\geq 0$, consider the invertible matrix and its determinant

$$\begin{aligned}
&X(t)\stackrel{\text{def}}{=}Xe^{X^{-1}Yt},\quad \det(Xe^{X^{-1}Yt})=\det(X)\det(e^{X^{-1}Yt})\\
&\frac{d}{dt}Xe^{X^{-1}Yt}=XX^{-1}Ye^{X^{-1}Yt}\to Y \text{ as } t\searrow 0\\
&\frac{d}{dt}\det(Xe^{X^{-1}Yt})=\det(X)\mathrm{tr}(X^{-1}Y)\det(e^{X^{-1}Yt})\\
&\qquad\qquad=\mathrm{tr}(X^{-1}Y)\det(Xe^{X^{-1}Yt})=0\\
&\Rightarrow\ \forall t\geq 0,\quad \det(Xe^{X^{-1}Yt})=\det(X)=1
\end{aligned}$$

and $X(t)\in\mathrm{SL}(n)$, Therefore, $Y\in T_X^\flat(\mathrm{SL}(n))$.

(iii), (iv), and (v). From Theorem 6. ∎

3.3.2. (General) Orthogonal Group $\mathrm{O}(n)$

The *(General) Orthogonal Group*

$$\mathrm{O}(n) \stackrel{\text{def}}{=} \left\{X \in \mathrm{GL}_n(\mathbb{R}) : X^\top X = XX^\top = I\right\} \tag{25}$$

is a compact subset of $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$.[7]

Since $1 = \det(XX^\top) = \det(X)\det(X^\top) = \det(X)^2$, then $|\det(X)| = 1$ and $\mathrm{O}(n)$ has two closed connected components ($\det X = 1$ and $\det X = -1$). The one that contains the identity matrix I_n is a normal subgroup called the *Special Orthogonal Group*:

$$\mathrm{SO}(n) \stackrel{\text{def}}{=} \{X \in \mathrm{O}(n) : \det X = 1\}. \tag{26}$$

Theorem 8 (i) *$\mathrm{O}(n)$ is a closed subgroup of $\mathrm{GL}(n)$.*

(ii) *The adjacent tangent cone to $\mathrm{O}(n)$ at $X \in \mathrm{O}(n)$ is the closed linear subspace*

$$T^\flat_X(\mathrm{O}(n))) = \left\{Y \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) : Y^\top X + X^\top Y = 0\right\} \tag{27}$$

of $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ and

$$T^\flat_{I_n}(\mathrm{O}(n))) = \left\{Y \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) : Y^\top + Y = 0\right\}. \tag{28}$$

(iii) *For each $X \in \mathrm{O}(n)$,*

$$\begin{aligned} &T^\flat_X(\mathrm{O}(n)) = X T^\flat_{I_n}(\mathrm{O}(n)) = T^\flat_{I_n}(\mathrm{O}(n)) X \\ &\text{and } X T^\flat_{I_n}(\mathrm{O}(n)) X^{-1} = T^\flat_{I_n}(\mathrm{O}(n)). \end{aligned} \tag{29}$$

(iv) *The* Lie bracket *mapping*

$$(A, B) \mapsto [A, B] : T^\flat_{I_n}(\mathrm{O}(n)) \times T^\flat_{I_n}(\mathrm{O}(n)) \to T^\flat_{I_n}(\mathrm{O}(n))$$

is well-defined and continuous.

(v) *For each $X \in \mathrm{O}(n)$, the mappings*

$$Y \mapsto YX^{-1} : T^\flat_X(\mathrm{O}(n)) \to T^\flat_{I_n}(\mathrm{O}(n)) \tag{30}$$

$$\text{and } Y \mapsto X^{-1}Y : T^\flat_X(\mathrm{O}(n)) \to T^\flat_{I_n}(\mathrm{O}(n)) \tag{31}$$

are bijective and bi-continuous.

[7] Given a Cauchy sequence $\{X_m\}$ converging to X in $\mathrm{O}(n)$, $X_m^\top \to X^\top$ and,

$$I_n = X_m X_m^\top = (X_m - X)X_m^\top + XX_m^\top \to XX^\top$$
$$I_n = X_m^\top X_m = (X_m - X)^\top X_m + X^\top X_m \to X^\top X,$$

and $\mathrm{O}(n)$ is closed. It is bounded in the Frobenius norm, since $\|X\|^2 = X \cdot\cdot X = I_n \cdot\cdot X^\top X = I_n \cdot\cdot I_n = n$, and hence $\mathrm{O}(n)$ is compact.

Proof (i) Clearly, $I_n \in \mathrm{O}(n)$. For $X_1, X_2 \in \mathrm{O}(n)$

$$
\begin{aligned}
(X_1X_2)^\top X_1X_2 &= X_2^\top X_1^\top X_1X_2 = X_2^\top X_2 = 1\\
X_1X_2(X_1X_2)^\top &= X_1X_2X_2^\top X_1^\top = X_1^\top X_1 = 1
\end{aligned}
$$

and $X_1X_2 \in \mathrm{O}(n)(I_n)$. For $X \in \mathrm{O}(n)$, $X^{-1} = X^\top$ and $X^\top \in \mathrm{O}(n)$.

(ii) By definition, for $Y \in T_X^\flat(\mathrm{O}(n)))$, there exists an admissible semitrajectory $X(t) \in \mathrm{O}(n)$ such that

$$
\begin{aligned}
&\frac{X(t)-X}{t} \to Y, \quad \text{where } X(t)X(t)^\top = X(t)^\top X(t) = I_n = XX^\top = X^\top X\\
&\Rightarrow X(t)^\top X(t)) - X^\top X = 0\\
&\Rightarrow 0 = [X(t)-X]^\top X + X(t)^\top [X(t)-X] \quad \Rightarrow Y^\top X + X^\top Y = 0.
\end{aligned}
$$

Conversely, assume $X, Y \in \mathrm{O}(n)$ such that $Y^\top X + X^\top Y = 0$. Then

$$
(X+tY)^\top (X+tY) = I_n + t(Y^\top X + X^\top Y) + t^2 Y^\top Y = I_n + t^2 Y^\top Y.
$$

Since $I_n + t^2Y^\top Y$ is symmetric and positive definite for $t \geq 0$ small, its square root $[I_n + t^2Y^\top Y]^{1/2}$ has the same property. Choose the semi-trajectory

$$
X(t) \stackrel{\text{def}}{=} (X+tY)[I_n + t^2Y^\top Y]^{-1/2} \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)
$$

and show that $X(t)^\top X(t) = X(t)X(t)^\top = I_n$:

$$
\begin{aligned}
X(t)^\top X(t) &= [I_n + t^2Y^\top Y]^{-1/2}(X+tY)^\top (X+tY)[I_n + t^2Y^\top Y]^{-1/2} = I_n\\
X(t)X(t)^\top &= (X+tY)[I_n + t^2Y^\top Y]^{-1/2}[I_n + t^2Y^\top Y]^{-1/2}(X+tY)^\top\\
&= (X+tY)[I_n + t^2Y^\top Y]^{-1}(X+tY)^\top\\
&= (X+tY)\left[(X+tY)^\top (X+tY)\right]^{-1}(X+tY)^\top\\
&= (X+tY)(X+tY)^{-1}(X+tY)^{-\top}(X+tY)^\top = I_n.
\end{aligned}
$$

Hence, $X(t) \in \mathrm{O}(n)$. Consider the differential quotient

$$
\begin{aligned}
&\frac{X(t)-X}{t}\\
&= \frac{(X+tY)[I_n + t^2Y^\top Y]^{-1/2} - X}{t}\\
&= Y[I_n + t^2Y^\top Y]^{-1/2} + X\frac{[I_n + t^2Y^\top Y]^{-1/2} - I_n}{t}\\
&= Y[I_n + t^2Y^\top Y]^{-1/2} + X[I_n + t^2Y^\top Y]^{-1/2}\frac{I_n - [I_n + t^2Y^\top Y]^{1/2}}{t}.
\end{aligned}
$$

But

$$\left(I_n - [I_n + t^2 Y^\top Y]^{1/2}\right)\left(I_n + [I_n - t^2 Y^\top Y]^{1/2}\right)$$
$$= I_n - [I_n - t^2 Y^\top Y] = t^2 Y^\top Y$$
$$I_n - [I_n + t^2 Y^\top Y]^{1/2} = t^2 Y^\top Y \left(I_n + [I_n - t^2 Y^\top Y]^{1/2}\right)^{-1}.$$

Finally, as $t \searrow 0$, the differential quotient

$$\frac{X(t) - X}{t}$$
$$= Y[I_n + t^2 Y^\top Y]^{-1/2} + tX[I_n + t^2 Y^\top Y]^{-1/2} Y^\top Y \left(I_n + [I_n - t^2 Y^\top Y]^{1/2}\right)^{-1}$$

goes to Y and $\left\{Y \in \mathrm{O}(n) : X^\top Y + Y^\top X = 0\right\} \subset T^\flat_X(\mathrm{O}(n))$.

(iii), (iv), and (v). From Theorem 6. ∎

3.3.3. Special Orthogonal Group

The *Special Orthogonal Group* is the intersection $\mathrm{SO}(n) = \mathrm{O}(n) \cap \mathrm{SL}(n)$ of the compact group $\mathrm{O}(n)$ and the matrix group $\mathrm{SL}(n)$:

$$\mathrm{SO}(n) \overset{\text{def}}{=} \{X \in \mathrm{O}(n) : \det X = 1\} = \left\{X \in \mathrm{GL}(n) : \begin{array}{c} XX^\top = X^\top X = I_n \\ \det X = 1 \end{array}\right\}. \tag{32}$$

It is a compact matrix group. The adjacent tangent cone at $X \in \mathrm{SO}(n)$ is

$$\begin{aligned} T^\flat_X(\mathrm{SO}(n)) &= \left\{Y \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) : Y^\top X + X^\top Y = 0 \text{ and } \mathrm{tr}(X^\top Y) = 0\right\} \\ &= \left\{Y \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) : Y^\top X + X^\top Y = 0\right\} = T^\flat_X(\mathrm{O}(n)), \end{aligned} \tag{33}$$

since $\mathrm{SO}(n)$ is one of the two compact connected components of $\mathrm{O}(n)$.

4. Stiefel manifold

For integers $0 < d \le n$, let $\mathcal{L}(\mathbb{R}^d, \mathbb{R}^n)$ be endowed with the Frobenius inner product Frobenius and norm

$$\forall A, B \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^n), \quad A\cdot\cdot B \overset{\text{def}}{=} \sum_{i=1}^{d}\sum_{j=1}^{n} a_{ij}\, b_{ij}, \quad \|A\|_2 = \sqrt{A\cdot\cdot A}, \tag{34}$$

. The *Stiefel Manifold* is the subset

$$\mathrm{V}_d(\mathbb{R}^n) \overset{\text{def}}{=} \left\{\mathcal{L}(\mathbb{R}^d, \mathbb{R}^n) : X^\top X = I_d\right\}. \tag{35}$$

THEOREM 9 *Let $0 < d \le n$ be integers.* $\mathrm{V}_d(\mathbb{R}^n)$ *is compact and the adjacent tangent cone at* $X \in \mathrm{V}_d(\mathbb{R}^n)$*,*

$$T^\flat_X(\mathrm{V}_d(\mathbb{R}^n)) = \left\{V \in \mathrm{V}_d(\mathbb{R}^n) : X^\top V + V^\top X = 0\right\}, \tag{36}$$

is a closed linear subspace of $\mathcal{L}(\mathbb{R}^d, \mathbb{R}^n)$ *of dimension* $nd - d(d+1)/2$*.*

PROOF (i) For a Cauchy sequence $\{X_n\}$ in $T^\flat_X(\mathrm{V}_d(\mathbb{R}^n))$, converging to X in $\mathcal{L}(\mathbb{R}^d, \mathbb{R}^n)$, $X_n^\top \to X^\top$,

$$I_d = X_n^\top X_n = (X_n - X)^\top X_n + X^\top X_n \to X^\top X,$$

and $\mathrm{V}_d(\mathbb{R}^n)$ is closed. It is bounded in the Frobenius norm, since

$$\|X\|^2 = X \cdot\cdot\, X = I_d \cdot\cdot\, X^\top X = I_d \cdot\cdot\, I_d = d$$

and hence $\mathrm{V}_d(\mathbb{R}^n)$ is compact.

(ii) By definition of an element V of the adjacent tangent cone at X, $T^\flat_X(\mathrm{V}_d(\mathbb{R}^n))$: for all $t_n \searrow 0$, there exists $\{X_n\}$ in $\mathrm{V}_d(\mathbb{R}^n)$ such that

$$\lim_{n\to\infty} \frac{X_n - X}{t_n} = V.$$

Since $X_n^\top X_n = I_d = X^\top X$, $0 = X_n^\top X_n - X^\top X = (X_n - X)^\top X_n - X^\top (X - X_n)$ and

$$0 = \frac{(X_n - X)^\top}{t_n} X_n - X^\top \frac{X - X_n}{t_n} \to V^\top X + X^\top V.$$

So, $T^\flat_X(\mathrm{V}_d(\mathbb{R}^n)) \subset \left\{V \in \mathrm{V}_d(\mathbb{R}^n) : X^\top V + V^\top X = 0\right\}$.

Conversely, take $X, V \in \mathrm{V}_d(\mathbb{R}^n)$ such that $V^\top X + X^\top V = 0$. Then

$$(X + tV)^\top (X + tV) = I_d + t(V^\top X + X^\top V) + t^2 V^\top V = I_d + t^2 V^\top V.$$

Since $I_d + t^2 V^\top V$ is symmetric and positive definite for $t \ge 0$, its square root $[I_d + t^2 V^\top V]^{1/2}$ has the same property. Choose the semi-trajectory

$$[0, \tau] \mapsto X(t) \stackrel{\text{def}}{=} (X + tV)[I_d + t^2 V^\top V]^{-1/2} \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^n)$$

and show that $X(t)^\top X(t) = I_d$:

$$\begin{aligned}
&X(t)^\top X(t) \\
&= [I_d + t^2 V^\top V]^{-1/2} (X + tV)^\top (X + tV) [I_d + t^2 V^\top V]^{-1/2} \\
&= [I_d + t^2 V^\top V]^{-1/2} \left[X^\top X + t(X^\top V + V^\top X) + t^2 V^\top V\right] [I_d + t^2 V^\top V]^{-1/2} \\
&= [I_d + t^2 V^\top V]^{-1/2} \left[I_d + t^2 V^\top V\right] [I_d + t^2 V^\top V]^{-1/2} = I_d.
\end{aligned}$$

Hence, $X(t) \in \mathrm{V}_d(\mathbb{R}^n)$. Now, consider the differential quotient

$$\begin{aligned}
&\frac{X(t)-X}{t}\\
&= \frac{(X+tV)[I_d+t^2V^\top V]^{-1/2}-X}{t}\\
&= V[I_d+t^2V^\top V]^{-1/2} + X\frac{[I_d+t^2V^\top V]^{-1/2}-I_d}{t}\\
&= V[I_d+t^2V^\top V]^{-1/2} + X[I_d+t^2V^\top V]^{-1/2}\frac{I_d-[I_d+t^2V^\top V]^{1/2}}{t}.
\end{aligned}$$

But

$$\begin{aligned}
&\left(I_d-[I_d+t^2V^\top V]^{1/2}\right)\left(I_d+[I_d-t^2V^\top V]^{1/2}\right)\\
&= I_d-[I_d-t^2V^\top V] = t^2V^\top V\\
&I_d-[I_d+t^2V^\top V]^{1/2} = t^2V^\top V\left(I_d+[I_d-t^2V^\top V]^{1/2}\right)^{-1}.
\end{aligned}$$

Finally, as $t \searrow 0$, the differential quotient

$$\begin{aligned}
&\frac{X(t)-X}{t}\\
&= V[I_d+t^2V^\top V]^{-1/2} + tX[I_d+t^2V^\top V]^{-1/2}V^\top V\left(I_d+[I_d-t^2V^\top V]^{1/2}\right)^{-1}
\end{aligned}$$

goes to V and $\left\{V \in \mathrm{V}_d(\mathbb{R}^n) : X^\top V + V^\top X = 0\right\} \subset T_X^\flat(\mathrm{V}_d(\mathbb{R}^n))$. ■

5. Diffeomorphism groups of nonlinear mappings

5.1. The ambient space θ and the group of diffeomorphisms $\mathcal{F}(\theta)$

For the matrix groups in the previous section, the common ambient space was $L(R^n, R^n)$; for nonlinear mappings from R^n to R^n there is a zoo,[8] that is, a broad choice, of ambient spaces, which is application dependent, and for each choice there are subgroups similar to the subgroups $\mathrm{SL}(n)$, $\mathrm{O}(n)$, or $\mathrm{SO}(n)$ of $L(R^n, R^n)$.

To construct diffeomorphism groups of nonlinear mappings we start with an ambient space θ of mappings $F : R^n \mapsto R^n$, which contains the identity ϵ, that is, for all $x \in Rn, e(x) = x$, closed under addition, scalar multiplication, and composition, and endowed with a Fréchet space structure (that is, a topological vector space, which is complete, metrizable, and locally convex). Denote by *De*

[8]See, for instance, Michor and Mumford, (2013) and Poinsot (2017).

the Jacobian matrix, which can be identified with the $n \times n$ identity matrix $I \in R^n$ previously denoted I_n. For instance, consider the following choices of θ:[9]

(i) the vector space $C^0(\mathbb{R}^n, \mathbb{R}^n)$ of continuous mappings from $\mathbb{R}^n$ to $\mathbb{R}^n$ endowed with the family of seminorms indexed by the compact subsets K of $\mathbb{R}^n$ (see Horvath, 1966, pp. 89, 110, 136)

$$q_K(f) \stackrel{\text{def}}{=} \sup_{x \in K} |f(x)|; \tag{37}$$

(ii) the m-times continuously differentiable mappings $C^m(\mathbb{R}^n, \mathbb{R}^n)$, $m \geq 1$,

$$C^m(\mathbb{R}^n, \mathbb{R}^n) \stackrel{\text{def}}{=} \left\{ f \in C^{m-1}(\mathbb{R}^n, \mathbb{R}^n) : \begin{array}{c} \forall \alpha,\ |\alpha| = m \\ \partial^\alpha f \in C^0(\mathbb{R}^n, \mathbb{R}^n) \end{array} \right\} \tag{38}$$

with the family of seminorms indexed by the compact subsets K of $\mathbb{R}^n$ and the multi-index $\alpha \in \mathbb{N}^n$, $|\alpha| \leq m$ (see Horvath, 1966, pp. 89, 110, 136)

$$q_{K,\alpha}(f) \stackrel{\text{def}}{=} \sup_{x \in K} |\partial^\alpha f(x)|; \tag{39}$$

(iii) the Lipschitzian mappings

$$C^{0,1}(\mathbb{R}^n, \mathbb{R}^n) \stackrel{\text{def}}{=} \left\{ f \in C^0(\mathbb{R}^n, \mathbb{R}^n) : \begin{array}{c} \forall K \text{ compact } \exists c_K > 0, \\ \displaystyle\sup_{\substack{x,y \in K \\ x \neq y}} \frac{|f(y) - f(x)|}{|x - y|} \leq c_K \end{array} \right\} \tag{40}$$

with the seminorms indexed by the compact subsets K

$$q_K(f) \stackrel{\text{def}}{=} \sup_{x \in K} |f(x)| + \sup_{\substack{y,z \in K \\ y \neq z}} \frac{|f(y) - f(z)|}{|y - z|}; \tag{41}$$

(iv) the m-Lipschitzian mappings, $m \geq 1$,

$$C^{m,1}(\mathbb{R}^n, \mathbb{R}^n) \stackrel{\text{def}}{=} \left\{ f \in C^{m-1}(\mathbb{R}^n, \mathbb{R}^n) : \begin{array}{c} \forall \alpha,\ |\alpha| = m \\ \partial^\alpha f \in C^{0,1}(\mathbb{R}^n, \mathbb{R}^n) \end{array} \right\} \tag{42}$$

with the family of seminorms indexed by the compact subsets K of $\mathbb{R}^n$ and the multi-index $\alpha \in \mathbb{N}^n$, $|\alpha| \leq m$,

$$q_{K,\alpha}(f) \stackrel{\text{def}}{=} \left\{ \sup_{x \in K} |\partial^\alpha f(x)| + \sup_{\substack{y,z \in K \\ y \neq z}} \frac{|\partial^\alpha f(y) - \partial^\alpha f(z)|}{|y - z|} \right\}. \tag{43}$$

[9] See Horvath (1966), Delfour and Zolésio (2011, Chpt. 3), for other examples.

Note that they all contain the identity mapping e.

Given a Fréchet space Θ of mappings from $\mathbb{R}^n$ to $\mathbb{R}^n$ that satisfies the previous hypotheses, we are interested in the group

$$\boxed{\mathcal{F}(\Theta) \overset{\text{def}}{=} \left\{F \in \Theta : F^{-1} \text{ exists and } F^{-1} \in \Theta\right\}.} \tag{44}$$

Theorem 10 *Let $\Theta = C^m(\mathbb{R}^n, \mathbb{R}^n)$, $m \geq 1$, be the Fréchet space of m-times continuously differentiable mappings endowed with the Fréchet topology of uniform convergence on compact subsets of $\mathbb{R}^n$.*

(i) *The set $\left\{F \in \Theta : F^{-1} \text{ exists}\right\}$ is a group and*

$$\left\{F \in \Theta : F^{-1} \text{ exists}\right\} = \mathcal{F}(\Theta). \tag{45}$$

(ii) *The adjacent tangent cone $T^\flat_F(\mathcal{F}(\Theta))$ at $F \in \mathcal{F}(\Theta)$ is the linear space Θ.*

(iii) *For each $F \in \mathcal{F}(\Theta)$*

$$T^\flat_F(\mathcal{F}(\Theta)) = F \circ T^\flat_e(\mathcal{F}(\Theta)) = T^\flat_e(\mathcal{F}(\Theta)) \circ F, \tag{46}$$

$$T^\flat_e(\mathcal{F}(\Theta)) = F \circ T^\flat_e(\mathcal{F}(\Theta)) \circ F^{-1}. \tag{47}$$

(iv) *$T^\flat_e(\mathcal{F}(\Theta))$ is the Lie algebra of the group $\mathcal{F}(\Theta)$, that is,*

$$\forall V, W \in T^\flat_e(\mathcal{F}(\Theta)), \quad [V, W] \overset{\text{def}}{=} V \circ W - W \circ V \in T^\flat_e(\mathcal{F}(\Theta)). \tag{48}$$

Proof (i) To show that $\left\{F \in \Theta : F^{-1} \text{ exists}\right\} \subset \Theta$ is a group it is sufficient to show that $F^{-1} \in \Theta$. At each point x, $DF(x)$ is invertible and continuous in a neighborhood of x. By the inverse function theorem, F is invertible and F^{-1} is C^1 in a neighborhood of x, and

$$DF^{-1} = (DF \circ F^{-1})^{-1} = (DF)^{-1} \circ F^{-1}.$$

Moreover, if $F \in C^m(\mathbb{R}^n, \mathbb{R}^n)$ for $m \geq 1$, then $F^{-1} \in C^m(\mathbb{R}^n, \mathbb{R}^n)$ (cf. Lang, 1969, Chpt. VI, Sec. 1, p. 122). From this local property, $F^{-1} \in \Theta$ and $\left\{F \in \Theta : F^{-1} \text{ exists}\right\} = \mathcal{F}(\Theta)$.

(ii) For $F \in \mathcal{F}(\Theta)$, $G \in \Theta$, and $t \geq 0$, $F + tG \in \Theta$ and

$$F + tG = [e + tG \circ F^{-1}] \circ F$$
$$\Rightarrow \det D(F + tG) = \det[I + tDG \circ F^{-1}] \det DF^{-1} \to \det DF^{-1} \neq 0.$$

Since $\det DF^{-1} \neq 0$, $(F + tG)^{-1}$ exists for $t > 0$ small, $F + tG \in \mathcal{F}(\Theta)$, and

$$\lim_{t \searrow 0} \frac{F + tG \ - F}{t} = \lim_{t \searrow 0} G = G \in T^\flat_F(\mathcal{F}(\Theta)) \quad \Rightarrow \Theta \subset T^\flat_F(\mathcal{F}(\Theta)),$$

and, since $\mathcal{F}(\Theta) \subset \Theta$, $T_F^\flat(\mathcal{F}(\Theta)) \subset \Theta$ and $T_F^\flat(\mathcal{F}(\Theta)) = \Theta$.

(iii) Given $F \in \mathcal{F}(\Theta)$ and $V \in T_e^\flat(\mathcal{F}(\Theta))$, for all sequences $t_n \searrow 0$, there exists $\{G_n\} \subset \mathcal{F}(\Theta)$ such that

$$\frac{G_n - e}{t_n} \to V \text{ in } \Theta \quad \Rightarrow \begin{cases} \dfrac{G_n \circ F - F}{t_n} \to V \circ F \text{ in } \Theta \\ \dfrac{F \circ G_n - F}{t_n} \to F \circ V \text{ in } \Theta \end{cases}$$

and $T_e^\flat(\mathcal{F}(\Theta)) \circ F \subset T_F^\flat(\mathcal{F}(\Theta))$ and $F \circ T_e^\flat(\mathcal{F}(\Theta)) \subset T_F^\flat(\mathcal{F}(\Theta))$. Conversely, given $F \in \mathcal{F}(\Theta)$ and $W \in T_F^\flat(\mathcal{F}(\Theta))$, for all sequences $t_n \searrow 0$, there exists $\{F_n\} \subset \mathcal{F}(\Theta)$ such that

$$\frac{F_n - F}{t_n} \to W \text{ in } \Theta \quad \Rightarrow \begin{cases} \dfrac{F_n \circ F^{-1} - e}{t_n} \to W \circ F^{-1} \text{ in } \Theta \\ \dfrac{F^{-1} \circ F_n - e}{t_n} \to F^{-1} \circ W \text{ in } \Theta \end{cases}$$

and $T_F^\flat(\mathcal{F}(\Theta)) \subset T_e^\flat(\mathcal{F}(\Theta)) \circ F^{-1}$ and $F^{-1} \circ T_F^\flat(\mathcal{F}(\Theta)) \subset T_e^\flat(\mathcal{F}(\Theta))$. Combining the last two relations with the previous ones we get (46).

(iv) Let $V, W \in T_e(\mathcal{F}(\Theta))$. For all sequences $t_n \searrow 0$, there exists $\{G_n\} \subset \mathcal{F}(\Theta)$ such that $(G_n - e)/t_n \to W$ in Θ. Since $G_n^{-1} \to e$ in Θ,

$$G_n \circ G_n^{-1} = e \quad \Rightarrow \frac{G_n^{-1} - e}{t_n} = -\frac{G_n - e}{t_n} \circ G_n^{-1} \to -W \text{ in } \Theta.$$

From (47) in part (iii), the mapping $G_n \circ V \circ G_n^{-1} \in T_e^\flat(\mathcal{F}(\Theta)) = \Theta$ and $G_n \circ V \circ G_n^{-1} \to V \in \Theta$ since Θ is closed. Similarly, the quotient

$$\frac{G_n \circ V \circ G_n^{-1} - V}{t_n}$$

belongs to the closed linear space $T_e^\flat(\mathcal{F}(\Theta)) = \Theta$ and the following limit necessarily belongs to Θ

$$\begin{aligned} \frac{G_n \circ V \circ G_n^{-1} - V}{t_n} &= \frac{G_n - e}{t_n} \circ V \circ G_n^{-1} + V \circ \frac{G_n^{-1} - e}{t_n} \\ &= \frac{G_n - e}{t_n} \circ V \circ G_n^{-1} + V \circ \frac{G_n^{-1} - e}{t_n} \\ &\to W \circ V - V \circ W \in \Theta. \end{aligned}$$

We conclude that $T_e^\flat(\mathcal{F}(\Theta)) = \Theta$ is the Lie algebra of the Lie group $\mathcal{F}(\Theta)$. ∎

The proof and the conclusions of Theorem 10 for $\Theta = C^m(\mathbb{R}^n,\mathbb{R}^n)$, $m \geq 1$, will be the same for Θ equal to $C^{m,1}(\mathbb{R}^n,\mathbb{R}^n)$ (see, for instance, Poinsot, 2017, for Lipschitz groups[10] and his references to Assouad, 1983; Gromov, 1991; Pansu, 1989). Diffeomorphism subgroups analogous to SL(n), O(n), or SO(n) in $\mathcal{L}(\mathbb{R}^n,\mathbb{R}^n)$ can be introduced:

$$\left\{F \in \Theta : F^{-1} \in \Theta \text{ and } \det(DF) = 1\right\} \text{ (incompressibility)}$$

$$\left\{F \in \Theta : F^{-1} \in \Theta \text{ and } \begin{array}{c} DF(DF)^\top = (DF)^\top DF = I \\ \text{and } \det(DF) = 1 \end{array}\right\} \begin{array}{l}\text{(rigid} \\ \text{displacements} \\ \text{and small} \\ \text{deformations)}\end{array}$$

for Θ equal to $C^m(\mathbb{R}^n,\mathbb{R}^n)$, $m \geq 0$, or $C^{m,1}(\mathbb{R}^n,\mathbb{R}^n)$, $m \geq 1$. See, for instance, Michor and Mumford (2018) for a tour in a zoo of diffeomorphism groups on $\mathbb{R}^n$.

5.2. Formula for the shape derivative of the integral of a function

Let Ω be a bounded open subset of $\mathbb{R}^n$ with Lipschitz boundary, $f \in W^{1,1}(\mathbb{R}^n)$ and

$$J(\Omega) \stackrel{\text{def}}{=} \int_\Omega f\, dx. \tag{49}$$

Let $\Theta = C^1(\mathbb{R}^n,\mathbb{R}^n)$ and the family $\{\Omega_F = F(\Omega) : F \in \mathcal{F}(\Theta)\}$. Since $C^1(\mathbb{R}^n,\mathbb{R}^n) \subset C^{0,1}(\mathbb{R}^n,\mathbb{R}^n)$, F is locallly bi-Lipschitzian. A bi-Lipschitzian bijection transports open Lipschitzian domains onto open Lipschitzian domains and Sobolev spaces onto Sobolev spaces

$$H^1(F(\Omega)) = \{f \circ F : f \in H^1(\Omega)\}. \tag{50}$$

Associate with our family of Lipschitzian open domains the volume integral

$$J(\Omega_F) \stackrel{\text{def}}{=} \int_{\Omega_F} f\, dx = \int_\Omega f \circ F\, |\det DF|\, dx, \quad j(F) \stackrel{\text{def}}{=} \int_\Omega f \circ F\, |\det DF|\, dx. \tag{51}$$

Recall that for $V \in T_e^\flat(\mathcal{F}(\Theta))$, for all $t_n \searrow 0$, there exists $\{F_n\} \subset \mathcal{F}(\Theta)$ such that $(F_n - e)/t_n \to V$. We want to compute for $V \in T_e^\flat(\mathcal{F}(\Theta))$

$$dj(F;V) \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{j(F_n) - j(e)}{t_n} \tag{52}$$

[10] See Evans and Gariepy (1992, Crl. 1 (ii), p. 84). Let $f, g : \mathbb{R}^n \to \mathbb{R}^n$ be locally Lipschitz and $Y = \{x \in \mathbb{R}^n : g(f(x)) = x\}$. Then $Dg(f(x))Df(x) = I$ for m_n a.e. $x \in Y$.

and show that the limit is independent of the choice of $\{F_n\}$. Since $\det I = 1$, we can drop the absolute value of $\det DF_n$ as n goes to infinity.

The functions involved are Hadamard differentiable

$$\begin{aligned}
&F \mapsto g(F) \stackrel{\text{def}}{=} f \circ G : \Theta \to L^1(\mathbb{R}^n), \quad d_H g(e;V) = \nabla f \cdot V \\
&F \mapsto DF : \Theta \to C^0(\mathbb{R}^n;\mathbb{R}^n), \quad d_H D(e;V) = DV \\
&A \mapsto \det A : C^0(\mathbb{R}^n;\mathbb{R}^n) \to C^0(\mathbb{R}^n;\mathbb{R}^n), \quad d_H \det(A;B) = (\operatorname{Cof} A) \cdot\cdot\, B \\
&F \mapsto h(F) = \det(DF), \quad d_H h(e;V) = \operatorname{div}(d_H D(e;V)) = \operatorname{div} V.
\end{aligned}$$

Then, using the chain rule for $V \in T_e^\flat(\mathcal{F}(\Theta)) = C^1(\mathbb{R}^n,\mathbb{R}^n)$

$$j(F) = \int_\Omega g(F)\, h(F)\, dx \quad \Rightarrow dj(e;V) = \int_\Omega \nabla f \cdot V + f \operatorname{div} V\, dx. \tag{53}$$

So, for all $V \in T_e^\flat(\mathcal{F}(\Theta))$ the domain integral is Hadamard differentiable

$$dJ(\Omega, V) \stackrel{\text{def}}{=} dj(e;V) = \int_\Omega \nabla f \cdot V + f \operatorname{div} V\, dx \tag{54}$$

and

$$V \mapsto dJ(\Omega, V) : C^1(\mathbb{R}^n,\mathbb{R}^n) \to \mathbb{R} \tag{55}$$

is linear and continuous.

If instead of perturbing Ω with $V \in T_e^\flat(\mathcal{F}(\Theta))$, we perturb Ω_F by some $W \in T_F^\flat(\mathcal{F}(\Theta))$: for all $t_n \searrow 0$, there exists $\{F_n\} \subset \mathcal{F}(\Theta)$ such that

$$\frac{F_n - F}{t_n} \to W \text{ in } \Theta \quad \Rightarrow \frac{F_n \circ F^{-1} - e}{t_n} \to W \circ F^{-1} \text{ in } \Theta.$$

Using the change of variable $F_n \circ F^{-1}$, for which $\det D(F_n \circ F^{-1})$ goes to $+1$ as n goes to infinity, we can drop the absolute value of the determinant and

$$\begin{aligned}
&\int_{\Omega_{F_n}} f\, dx - \int_{\Omega_F} f\, dx = \int_{\Omega_F} f \circ (F_n \circ F^{-1}) \det D(F_n \circ F^{-1})\, dx - \int_{\Omega_F} f\, dx \\
&\frac{\int_{\Omega_{F_n}} f\, dx - \int_{\Omega_F} f\, dx}{t_n} \to \int_{\Omega_F} \nabla f \cdot (W \circ F^{-1}) + f \operatorname{div}(W \circ F^{-1})\, dx.
\end{aligned}$$

For all $F \in \mathcal{F}(\Theta))$ and $W \in T_F^\flat(\mathcal{F}(\Theta))$

$$dJ(\Omega_F; W) \stackrel{\text{def}}{=} \int_{\Omega_F} \nabla f \cdot (W \circ F^{-1}) + f \operatorname{div}(W \circ F^{-1})\, dx. \tag{56}$$

6. Group of characteristic functions

6.1. From shape to topological changes

Since diffeomorphism groups used to define the shape derivative cannot induce topological changes, we have to enlarge the group of subsets of $\mathbb{R}^n$ and the ambient space, where the adjacent tangent cone is defined. There are many avenues to explore, but the *group of characteristic functions*

$$\mathrm{X}(\mathbb{R}^n) \overset{\text{def}}{=} \left\{\chi \in L^1(\mathbb{R}^n) : \chi(1-\chi) = 0 \quad \mathrm{m}_n \text{ a.e.}\right\} \tag{57}$$

with respect to the n-dimensional *Lebesgue measure* m_n has been successful for the topological derivative (Delfour, 2016, 2018a, 2023a). It is one of several spaces of set-parametrized functions that can be endowed with a metric space structure.

For $\mathrm{X}(\mathbb{R}^n)$ the group operation

$$\chi_A \mathbin{\triangle} \chi_B \overset{\text{def}}{=} |\chi_A - \chi_B| \tag{58}$$

corresponds to the *symmetric difference* of two sets

$$A \mathbin{\triangle} B \overset{\text{def}}{=} (A\backslash B) \cup (B\backslash A) \quad \Rightarrow \chi_{A\triangle B} = \chi_A \mathbin{\triangle} \chi_B = |\chi_A - \chi_B|. \tag{59}$$

This operation induces an *Abelian group* structure on $\mathrm{X}(\mathbb{R}^n)$:

$$\begin{aligned}&A \mathbin{\triangle} B = B \mathbin{\triangle} A, \quad (A \mathbin{\triangle} B) \mathbin{\triangle} C = A \mathbin{\triangle} (B \mathbin{\triangle} C)\\ &\Rightarrow \chi_A \mathbin{\triangle} \chi_B = \chi_B \mathbin{\triangle} \chi_A \quad \text{and} \quad (\chi_A \mathbin{\triangle} \chi_B) \mathbin{\triangle} \chi_C = \chi_A \mathbin{\triangle} (\chi_B \mathbin{\triangle} \chi_C),\end{aligned}$$

$\varnothing$ is the *neutral element*, $A \mathbin{\triangle} \varnothing = A$ and $\chi_A \mathbin{\triangle} \chi_\varnothing = \chi_A$, and every element χ_A is its own inverse, $A \mathbin{\triangle} A = \varnothing$ and $\chi_A \mathbin{\triangle} \chi_A = \chi_{A\triangle A} = \chi_\varnothing$. Given the n-dimensional Lebesgue measure m_n, the metric

$$\rho(\chi_A, \chi_B) \overset{\text{def}}{=} \int_{\mathbb{R}^n} |\chi_A(x) - \chi_B(x)| \, d\mathrm{m}_n$$

makes the Abelian group $\mathrm{X}(\mathbb{R}^n)$ a complete metric space. As in the case of images of a fixed set for the shape derivative, we have a group $\mathrm{X}(\mathbb{R}^n)$ in the space $L^1(\mathbb{R}^n)$.

But $L^1(\mathbb{R}^n)$ is too small to accomodate semitangents to $\mathrm{X}(\mathbb{R}^n)$ at a characteristic function χ_Ω of a bounded open subset of $\mathbb{R}^n$, since, in general, the quotient $(\chi_{\Omega_n} - \chi_\Omega)/t_n$ does not converge in $L^1(\mathbb{R}^n)$. But it can converge in the larger space of distributions $\mathcal{D}(\mathbb{R}^n)'$: there exists $T \in \mathcal{D}(\mathbb{R}^n)'$ such that for all $\varphi \in \mathcal{D}(\mathbb{R}^n)$

$$\int_{\mathbb{R}^n} \frac{\chi_{\Omega_n} - \chi_\Omega}{t_n} \varphi \, d\mathrm{m}_n = \frac{\int_{\Omega_n} \varphi \, d\mathrm{m}_n - \int_\Omega \varphi \, d\mathrm{m}_n}{t_n} = \frac{\int_{\Omega_n \backslash \Omega} \varphi \, d\mathrm{m}_n}{t_n} \to < T, \varphi > .$$

In general, the convergence occurs in one of the smaller intermediate spaces $\mathcal{D}^m(\mathbb{R}^n)'$, $m \geq 1$, down to the space of bounded measures

$$\mathfrak{M}(\mathbb{R}^n) = \mathcal{D}^0(\mathbb{R}^n)', \tag{60}$$

where $\mathcal{D}^0(\mathbb{R}^n)$, also denoted $C_c(\mathbb{R}^n)$, is the space of continuous functions on $\mathbb{R}^n$ with compact support in $\mathbb{R}^n$. Note that we also have

$$\mathfrak{M}(\mathbb{R}^n) \stackrel{\text{def}}{=} C_0(\mathbb{R}^n)', \tag{61}$$

where $C_0(\mathbb{R}^n)$ is the space of continuous functions on $\mathbb{R}^n$, which vanish at infinity.

In order to use $\mathcal{D}^m(\mathbb{R}^n)'$, $0 \leq m \leq \infty$, introduce the continuous injection

$$f \mapsto i(f) : L^1(\mathbb{R}^n) \to \mathcal{D}^m(\mathbb{R}^n)', \quad i(f)(\varphi) \stackrel{\text{def}}{=} \int_{\mathbb{R}^n} f\, \varphi\, dm_n, \tag{62}$$

and consider the group $\mathrm{X}(\mathbb{R}^n)$ as a subset of the ambient space $\mathcal{D}^m(\mathbb{R}^n)'$. Now, for a bounded open subset of $\mathbb{R}^n$ and by the definition of a semitangent $\nu \in T^\flat_{\chi_\Omega}(X(\mathbb{R}^n))$: for each $t_n \searrow 0$, there exists a sequence $\{\chi_n\} \subset \mathrm{X}(\mathbb{R}^n)$ such that

$$\forall \varphi \in \mathcal{D}^m(\mathbb{R}^n), \quad \int_{\mathbb{R}^n} \frac{\chi_n - \chi_\Omega}{t_n}\, \varphi\, dm_n \to < \nu, \varphi > . \tag{63}$$

We know how to construct a sequence of measurable subsets $\Omega_n = \{x \in \mathbb{R}^n : \chi_n(x) = 1\}$ for several examples including the d-dimensional topological derivative.

6.2. Formula for the topological derivative of the integral of a function

For the topological derivative a bounded open subset Ω of $\mathbb{R}^n$ is perturbed by removing an r-dilation of a d-dimensional closed subset E of Ω

$$E \subset \Omega, \quad \Omega_r = \Omega \backslash E_r, \quad E_r \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n : d_E(x) \leq r\}. \tag{64}$$

Given a continuous function $f : \mathbb{R}^n \to \mathbb{R}$, $0 \leq d < n$, and $0 \leq r < R$ such that $E_R \subset \Omega$, the *d-dimensional topological derivative* of the volume integral of a function f is defined (Delfour, 2016, 2018a, 2023a) as the limit as $r \searrow 0$ of the differential quotient

$$\frac{\int_{\mathbb{R}^n} (\chi_{\Omega_r} - \chi_\Omega)\, f\, \mathrm{dm}_n}{\alpha(n-d) r^{n-d}} = -\frac{1}{\alpha(n-d) r^{n-d}} \int_{E_r} f\, \mathrm{dm}_n, \tag{65}$$

where $\alpha(m)$ is the volume of the unit ball in dimension m, $0 \leq m \leq n$ ($\alpha(0) = 1$ in dimension $m = 0$).

For $0 \le d < n$ and a *d-rectifiable set* E with *positive reach* and finite d-dimensional Hausdorff measure H^d, $H^d(E) < \infty$, the restriction $H^d \llcorner E$ of H^d to E is a Radon measure and the limit exists and it is equal to

$$\lim_{r \searrow 0} -\frac{1}{\alpha(n-d)r^{n-d}} \int_{E_r} f\, dx = -\int_E f\, dH^d = -\int_{\mathbb{R}^n} f\, d(H^d \llcorner E), \tag{66}$$

where the convergence takes place in the ambient space of bounded measures

$$\mathfrak{M}(\mathbb{R}^n) = \mathcal{D}^0(\mathbb{R}^n)' = C_c(\mathbb{R}^n)'. \tag{67}$$

Remark 5 *For $d = n$ and E being a compact subset of $\mathbb{R}^n$, the differential quotient in formula* (65) *reduces to a difference and*

$$\forall f \in \mathcal{D}^0(\mathbb{R}^n), \quad \int_{\Omega_r} f\, \mathrm{dm}_n - \int_{\Omega} f\, \mathrm{dm}_n = -\int_{E_r} f\, \mathrm{dm}_n \to -\int_E f\, \mathrm{dm}_n,$$

which is consistent with the case of $0 \le d < n$, since $H^n = \mathrm{m}_n$ in $\mathbb{R}^n$.

If $\mathfrak{M}(\mathbb{R}^n)$ is chosen as the ambient space, then, by definition of $\nu \in T^\flat_\chi(\mathrm{X}(\mathbb{R}^n))$: for each $t_n \searrow 0$, there exists a sequence $\{\chi_n\} \subset \mathrm{X}(\mathbb{R}^n)$ such that

$$\forall \varphi \in \mathcal{D}^0(\mathbb{R}^n), \quad \int_{\mathbb{R}^n} \frac{\chi_n - \chi}{t_n} \varphi\, dm_n \to \int_{\mathbb{R}^n} \varphi d\nu. \tag{68}$$

In this case, we know the form of the limit by Riesz Representation Theorem.

Theorem 11 (Evans and Gariepy, 1992, Thm. 1, sec. 1.8, p. 49) *Let $L : C_c(\mathbb{R}^n, \mathbb{R}^m) \to \mathbb{R}$ be a linear functional, satisfying*

$$\sup\{L(f) : f \in C_c(\mathbb{R}^n, \mathbb{R}^m),\ \|f\|_{C^0} \le 1,\ \mathrm{supp}\,(f) \subset K\} < \infty \tag{69}$$

for each compact set $K \subset \mathbb{R}^n$. Then, there exist a Radon measure μ on $\mathbb{R}^n$ and a μ-measurable function $\sigma : \mathbb{R}^n \to \mathbb{R}^m$ such that

(i) $|\sigma(x)| = 1$ *for μ a.e. x, and*
(ii) $L(f) = \int_{\mathbb{R}^n} f \cdot \sigma\, d\mu$ *for all $f \in C_c(\mathbb{R}^n, \mathbb{R}^m)$.*

Therefore, for all $\varphi \in C_c(\mathbb{R}^n)$

$$\begin{aligned} &\int_{\mathbb{R}^n} \frac{\chi_n - \chi}{t_n} \varphi\, dm_n \to \int_{\mathbb{R}^n} \varphi d\nu = \int_{\mathbb{R}^n} \varphi\, \sigma\, d\mu \\ &\Rightarrow \int_{\mathbb{R}^n} \frac{|\chi_n - \chi|}{t_n} dm_n = \frac{\mathrm{m}_n(A_n \triangle A)}{t_n} \to \int_{\mathbb{R}^n} d|\nu| = \int_{\mathbb{R}^n} |\sigma|\, d\mu = \int_{\mathbb{R}^n} d\mu, \end{aligned}$$

where $A_n = \{x \in \mathbb{R}^n : \chi_n(x) = 1\}$ and $A = \{x \in \mathbb{R}^n : \chi(x) = 1\}$.

Going back to the topological derivative introduced in (66)

$$\lim_{r\searrow 0} \frac{\int_{\Omega_r} f\,dx - \int_\Omega f\,dx}{\alpha(n-d)r^{n-d}} = -\int_E f\,dH^d = -\int_{\mathbb{R}^n} f\,d(H^d \llcorner E),$$

$\mu = H^d \llcorner E)$ and $\sigma = -1$, by constructing the r-dilatations E_r of E.

Can other constructions of the sequence of sets Ω_n be envisioned? Since $\mu = H^d \llcorner E)$ only generates a half tangent, the adjacent tangent cone $T^\flat_{\chi_\Omega}(X(\mathbb{R}^n))$ is possibly convex, but definitely not equal to the whole space $\mathfrak{M}(\mathbb{R}^n)$ or a linear subspace.

6.3. Velocity method

Another point raised in Delfour (2016, sec 3.1. pp. 234–235), Delfour (2018a, sec. 4.1, pp. 967–968) is that there are semitangents in the larger space $(\mathcal{D}^1(\mathbb{R}^n))'$[11], which correspond to shape derivatives via the *Velocity Method*.

For the Velocity Method, consider the following *continuous trajectory* in $X(\mathbb{R}^n)$

$$t \mapsto \chi_{T_t(V)(\Omega)} : [0,1] \to X(\mathbb{R}^n), \quad \frac{dT_t(V)}{dt} = V(t) \circ T_t(V),\ T_0(V) = I.$$

The *semitangent* at χ_Ω is obtained by considering the limit of the differential quotient $\left(\chi_{T_t(V)(\Omega)} - \chi_\Omega\right)/t \in L^1(\mathbb{R}^n)$, which does not exist in $L^1(\mathbb{R}^n)$. Associate with $\chi_{T_t(V)(\Omega)}$ the *distribution*

$$\phi \mapsto \int_{\mathbb{R}^n} \chi_{T_t(V)(\Omega)}\,\phi\,dx = \int_{T_t(V)(\Omega)} \phi\,dx = \int_\Omega \phi \circ T_t\ \det DT_t\,dx : \mathcal{D}(\mathbb{R}^n) \to \mathbb{R}\,.$$

If $V \in C^{0,1}(\overline{\mathbb{R}^n}, \mathbb{R}^n)$, then

$$\begin{aligned} \left.\frac{d}{dt}\right|_{t=0^+} \int_\Omega \phi \circ T_t\ \det DT_t\,dx &= \int_\Omega [\operatorname{div} V(0)\phi + V(0)\cdot\nabla\phi]\ dx \\ &= \int_{\mathbb{R}^n} \chi_\Omega\,[\operatorname{div} V(0)\phi + V(0)\cdot\nabla\phi]\ dx \\ &= \int_{\mathbb{R}^n} \chi_\Omega \operatorname{div}\ (V(0)\,\phi)\ dx \\ &= -< \nabla\chi_\Omega, V(0)\,\phi >, \end{aligned}$$

[11] $\mathcal{D}^m(\mathbb{R}^n)$ is the space of m times continuously differentiable functions with compact support.

where $\nabla\chi_\Omega$ is the distributional gradient of χ_Ω (see, for instance, Delfour and Zolésio, 2011, Thm. 4.1, Chapter 9, p. 483). The bilinear function

$$(\phi, V) \mapsto \int_{\mathbb{R}^n} \chi_\Omega \,\mathrm{div}\, (V(0)\,\phi)\; dx : H_0^1(\mathbb{R}^n) \times C^{0,1}(\overline{\mathbb{R}^n}, \mathbb{R}^n) \to \mathbb{R}$$

is continuous. This generates the continuous linear mapping $V \mapsto \nabla\chi_\Omega \cdot V : C^{0,1}(\overline{\mathbb{R}^n}, \mathbb{R}^n) \to H^{-1}(\mathbb{R}^n)$

$$\forall \phi \in H_0^1(\mathbb{R}^n), \quad < (\nabla\chi_\Omega \cdot V), \phi > \stackrel{\text{def}}{=} \int_{\mathbb{R}^n} \chi_\Omega \,\mathrm{div}\, (V(0)\,\phi)\; dx.$$

The support of $\nabla\chi_\Omega \cdot V$ is in Γ, the boundary of Ω.

So, the adjacent tangent cone $T^\flat_{\chi_\Omega}(\mathrm{X}(\mathbb{R}^n))$ to $X(\mathbb{R}^n)$ at χ_Ω (considered as a subset of the space of distributions $\mathcal{D}(\mathbb{R}^n)'$) contains the linear subspace of functions in $H^{-1}(\mathbb{R}^n)$ of the form

$$\left\{ \nabla\chi_\Omega \cdot V : V \in C^{0,1}(\overline{\mathbb{R}^n}, \mathbb{R}^n) \right\} \subset H^{-1}(\mathbb{R}^n) \subset \mathcal{D}(\mathbb{R}^n)'.$$

Therefore, $T^\flat_{\chi_\Omega}(\mathrm{X}(\mathbb{R}^n))$ is not completely contained in the space of bounded measures $\mathfrak{M}(\mathbb{R}^n)$. The full characterization of all the elements of $T^\flat_{\chi_\Omega}(\mathrm{X}(\mathbb{R}^n))$ remains an open problem which might require new ways to perturb of a domain Ω.

References

Absil, P.-A., Mahony, R., and Sepulchre, R. (2008) *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, Princeton, N.J., Woodstock.

Assouad, P. (1983) Plongements Lipschitziens dans R^n. *Bull. Soc. Math. France* **111**, 429–448.

Aubin, J.P. and Frankowska, H. (1990) *Set-Valued Analysis*. Birkhäuser, Boston.

Baydin, A.G., Pearlmutter, B.A., Radul, A.A., and Siskind, J.M. (2018) Automatic differentiation in machine learning: A survey. *Journal of Machine Learning Research*, **18**, 1–43.

Bolte, J. and Pauwels, E. (2021) Conservative set valued fields, automatic differentiation, stochastic gradient methods and deep learning. *Mathematical Programming: Series A and B.* **188** (1), 19–51.

Delfour, M.C. (2016) Differentials and semidifferentials for metric spaces of shapes and geometries. In: *System Modeling and Optimization*, L. Bociu, J.-A. Désidéri and A. Habbal, eds., 230–239, Springer International Publishing AG, Switzerland.

Delfour, M.C. (2018a) Topological derivative: a semidifferential via the Minkowski content. *Journal of Convex Analysis* **25** (3), 957–982.

Delfour, M.C. (2018b) Control, shape, and topological derivatives via minimax differentiability of Lagrangians. In: *Numerical Methods for Optimal Control Problems*. M. Falcone, R. Ferretti, L. Grüne, W. McEneaney, eds., 137–164, Springer INdAM Series **29**, Springer, Cham, Switzerland.

Delfour, M.C. (2020a) *Introduction to Optimization and Hadamard Semidifferential Calculus*, 2nd ed. SIAM, Philadelphia, PA.

Delfour, M.C. (2020b) Hadamard semidifferential of functions on an unstructured subset of a TVS. *J. Pure Appl. Funct. Anal.* **5** (5), 1039–1072.

Delfour, M.C. (2023a) One-sided Derivative of Parametrized Minima for Shape and Topological Derivatives. *SIAM J. Control. Optim.* **61** (3), 1322–1349.

Delfour, M.C. (2023b) Hadamard semidifferential of continuous convex functions. *J. Pure and Applied Functional Analysis* **8** (5), 1341–1356.

Delfour, M.C. and Huot-Chantal, F. (2019) On the figure of columns of Lagrange revisited. *J. Convex Anal.* (3)**26**, 855–876.

Delfour, M.C. and Zolésio, J.P. (2011) *Shapes and Geometries: Metrics, Analysis, Differential Calculus and Optimization*, 2nd ed. SIAM, Philadelphia, PA.

Edelman, A., Arias, T.A. and Smith, S.T. (1998) The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.* **20** (2), 303–353.

Evans, L.C. and Gariepy, R.F. (1992) *Measure Theory and the Properties of Functions*. CRC Press, Boca Raton, FL.

Fréchet, M. (1937) Sur la notion de différentielle. *Journal de Mathématiques Pures et Appliquées* **16**, 233–250.

Gromov, M. (1991) *Geometric Group Theory*, volume 2: *Asymptotic Invariants of Infinite Groups*. *London Mathematical Society Lecture Note Series*, **182**, Cambridge University Press.

Horváth, J. (1966) *Topological Vector Spaces and Distributions*, Vol. I. Addison-Wesley, Reading, MA.

Huot-Chantal, F. (2018) *Sur la figure des colonnes de Lagrange revisité*. Mémoire, Dép. de Mathématiques et de Statistique, Université de Montréal, Canada.

Ji, M. and Klinowski, J. (2006) Taboo evolutionary programming: a new method of global optimization. *Proc. R. Soc. A* **462**, 3613–3627.

Lang, S. (1969) *Analysis II*. Addison–Wesley Publishing Company, Reading, Mass.

Lange, K. (2024) A tutorial on Hadamard semidifferentials. *Foundations and Trends in Optimization* **6** (1), 1–62.

Marsden, J.E. and Ratiu, T.S. (1994) *Introduction to Mechanics and Symmetry.* Springer-Verlag, New York, Berlin.

Michor, P.W. and Mumford, D. (2013) A zoo of diffeomorphism groups on $\mathbb{R}^n$. *Ann. Glob. Anal. Geom.* **44** (4), 529–540.

Neidinger, R. D. (2010) Introduction to automatic differentiation and matlab object-oriented programming. *SIAM Review*, **52**, 545–563.

Pansu, P. (1989) Métriques de Carnot-Carathéodory et quasiisométries des espaces symétriques de rang un. *Ann. of Math.* **129** (2), 1–60.

Poinsot, L. (2017) Lipschitz groups and Lipschitz maps. *International Journal of Group Theory* **6** (1), 9–16.

Rudin, W. (1976) *Principles of Mathematical Analysis.* McGraw–Hill, New York.