GOOD CONTINUATION IN 3D: THE NEUROGEOMETRY OF STEREO VISION

M. V. BOLELLI[1,2,*], G. CITTI[1], A. SARTI[2], AND S. W. ZUCKER [3]

Appendices

A. A gentle introduction to sub-Riemannian geometry

In this paper we exploit techniques from differential geometry, and in particular sub - Riemannian geometry. In this appendix we provide an invitation to these ideas with a rather informal discussion. For the reader interested in a formal introduction on basic instruments of differential geometry (arguments of sections A.1 and A.2) please refer to [10]. For a complete and formal mathematical (comprehensive) introduction to sub-Riemannian geometry we refer to [2], while for a more informational point of view please consult [9, Ch. 4.2] and [5].

A.1. **Tangent bundle.** To start, imagine that you are standing at a point on a smooth surface in the world, far from any boundaries. Now, you can "walk away" from this point in any ($2D$) compass direction; for example, you could walk north or south or any direction in-between. If your steps were very very short, then the (flat) compass actually characterizes the $2D$ space of possible steps you might take. These same ideas are expressed more formally in differential geometry, as follows. One can attach to every point p of a differentiable manifold M (a generalized surface) a tangent space T_pM (the compass plus some algebra describing vector operations). That is, the tangent space is a real vector space that contains the possible *directions* in which one can tangentially pass through $p \in M$. If the manifold is connected, then the tangent space has, at every point, the same dimension as the manifold. So, if the manifold is a $2D$ surface, the tangent space at a point is a plane. In general, this tangent plane "approximates" the surface only locally.

The elements $\vec{X}_p$ of the tangent space T_pM at p are called *tangent vectors* at p. Attached to a point on the surface, as above, these tangent vectors define the directions in which one could walk away from the point. But modern differential geometry provides another interpretation: it is possible to think of the elements of the tangent space in terms of directional derivatives. Technically, for every smooth function f, $Xf(p) = \vec{X}_p \cdot \nabla f(p)$ will denote the directional derivative of f in the direction of the vector $\vec{X}_p$, with ∇ denoting the gradient vector (expressed in an appropriate coordinate system) and $\cdot$ scalar product between these vectors. We will also denote $X_p = \vec{X}_p \cdot \nabla_p$, omitting the function f.

[1]Department of Mathematics, University of Bologna, Italy.
[2]CAMS, CNRS - EHESS, Paris, France.
[3]Departments of Computer Science and Biomedical Engineering, Yale University, New Haven, CT, United States.
[*]Correspondence: M. V. Bolelli, maria.bolelli2@unibo.it.

We now consider pairs of directional derivatives X and Y. If X and Y are partial derivatives, for every regular function f one has $XYf = YXf$. If X and Y are directional derivatives, in general $XYf \neq YXf$. Explicit computation tell us that at every point p

$$[X,Y]f(p) = (XY - YX)f(p) = (J_{\vec{Y}_p}\vec{X}_p - J_{\vec{X}_p}\vec{Y}_p) \cdot \nabla f(p),$$

with $J_{\vec{X}_p}$ and $J_{\vec{Y}_p}$ Jacobian matrices of $\vec{X}_p$ and $\vec{Y}_p$. The quantity $[X,Y]f$ is called *commutator* since it expresses the fact that the two derivatives do not commute. The same notion can be expressed in terms of increments: one might visualize an increment from a point p as the head of a vector $\vec{X}_p$ applied at the point p. Then the expression $XY - YX$ will be geometrically obtained as follows: place X down at a point, then the other Y at its head, then the first one backward finally the second one backward. The issue is whether the quadrilateral is closed. Formally this is captured by the *commutator* of two elements X and Y at the point p.

In order to compute the second derivative XYf, we need to know Yf at every point near p. This lead to the more general notion of *vector fields*, which are abstractions of the velocity field of points moving in the manifold. A vector field X attaches to every point p of the manifold M a vector $\vec{X}_p$ from the tangent space at that point, in a smooth manner. There are no abrupt jumps between points.

Since we related each tangent vector with a derivation above, we can now go further; see Fig. A.1, images (A-C). Each vector field can be associated with an ordinary differential equation, whose solutions are called *integral curves* of the vector field: they are parametric curves that represent specific solutions to the ordinary differential equation depicted by the vector field. Think of it as follows: imagine you are starting at a point, and take an infinitesmimal step in the direction of a tangent vector at that point; you will now be at a neighboring point. So, again, you can take a step from this neighboring point in (possibly) another tangent direction. Continuing this process for a while, you geometrically trace out integral curves $\gamma : [t_1, t_2] \subseteq \mathbb{R} \longrightarrow M$. Importantly, the given vector field X at the point $\gamma(t)$ is the tangent vector to the curve at that point. Importantly, this holds true everywhere along the curve, so that the integral curve satisfies a differential equation:

$$\dot{\gamma}(t) = \vec{X}_{\gamma(t)}.$$

All the tangent spaces of a manifold may be "glued together" to form a new differentiable manifold with twice the dimension of the original manifold, called the *tangent bundle* of the manifold. As a set, it is given by the disjoint union of the tangent spaces of M, that is:

$$TM = \bigsqcup_{p \in M} T_pM = \{(p, X_p) \mid p \in M,\ X_p \in T_pM\}.$$

In particular, an element of TM can be thought of as a pair (p, X_p), where p is a point in M and X_p is a tangent vector to M at p. There exists a natural projection $\pi : TM \to M$ defined by $\pi(p, X_p) = p$. which maps each element of the tangent space T_pM to the single point p.

A.2. **Group action on a manifold.** The operation of adding (real) numbers has an important algebraic structure, called a group. It requires, for example, that the sum of any two numbers is

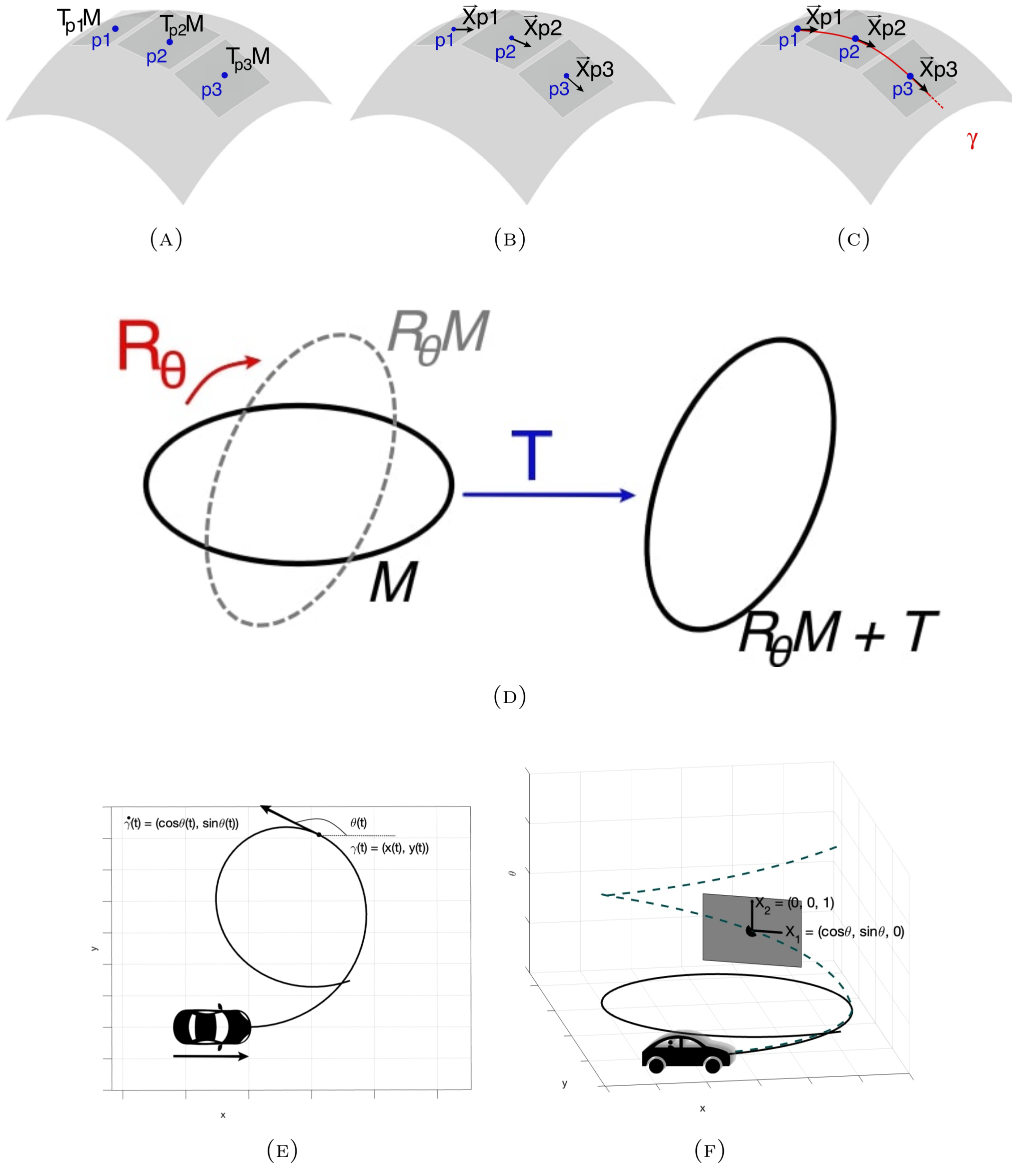

Figure A.1. (A) Tangent planes $T_{p_i}M$ (darker planes) at points $p_i, i = 1, 2, 3$ in the manifold M. (B) Vector field X defined on M: to every point $p_i, i = 1, 2, 3$ of the manifold M we have a vector $\vec{X}_{p_i}$ of the tangent space at that point. (C) Integral curve γ associated with the vector field X starting from $p_1 \in M$. (D) Group action of the roto-translation group $SE(2)$ on the manifold M (black ellipse): first, the manifold is rotated through a rotation of angle θ obtaining $R_\theta M$, and then a translation is applied, moving the rotated manifold in space realizing $R_\theta M + T$. (E) Geometric set-up of the motion of a car moving on a plane. (F)Sub-Riemannian formalization in $SE(2)$. Tangent vector of the path is constrained to be in the gray plane, span of $\vec{X}_{1,p}$ and $\vec{X}_{2,p}$, admissible directions of movement.

again a number; that there is an inverse operation "-"; and that there is an identity operation "0" that is, adding to any number yields the same number.

When a group G acts on a manifold (e.g. the real numbers, above), it means that each of its elements performs a certain operation on all the elements of the manifold in a way that is compatible with the manifold itself. More precisely, this action is described by a map $\sigma : G \times M \to M, (g, x) \mapsto g \cdot x$ which is the (left) group action of a group G on a smooth manifold M, if the map σ is differentiable.

For example, we can take the bidimensional roto-translation group $SE(2) = \mathbb{R}^2 \times \mathbb{S}^1$ and define its action on a smooth manifold $M \subseteq \mathbb{R}^2$ following the group law: first we apply a rotation and then a translation of the manifold itself. This is formalized through the map $\sigma : SE(2) \times M \longrightarrow M$, $\sigma(g, p) = (Rp + q)$, with $g = (q, R) \in SE(2)$, namely a point $q \in \mathbb{R}^2$ and R bidimensional rotation of angle $\theta \in \mathbb{S}^1$. A graphical example is shown in Figure A.1, image (D).

We are now ready to generalize these familiar ideas to cortical space, with its special position × orientation structure, or to stereo space.

A.3. **Sub-Riemannian geometry.** A point constraint to move on a manifold, illustrated above, dictates that one can move only along directions tangent to the manifold, since moving in the normal direction would leave the manifold. This means that, for every point p, the set of admissible directions of displacement coincides with the tangent plane T_pM. In the presence of further constraints, some tangent directions could be forbidden. This leads to introducing, at every point p, the admissible tangent space $\mathcal{A}_p$, which is the subspace of T_pM of admissible directions of movement. If the tangent space T_pM has dimension n, the admissible tangent space $\mathcal{A}_p$ will have dimension $m \leq n$. Repeating the same construction for every point of the manifold, we call the *admissible tangent bundle* the union of admissible tangent spaces at every point: $\mathcal{A} = \bigsqcup_{p \in M} \mathcal{A}_p$.

If we introduce a scalar product on $\mathcal{A}_p$, then we are able to define a norm on vectors with the aim to measure the length of such vectors and the distance between points. The manifold with these properties is usually called sub-Riemannian manifold, while manifolds where movements are allowed in any direction are called Riemannian manifolds.

Let us explicitly note that while Riemannian geometry arises in presence of a physical constraints, sub-Riemannian geometry arises in presence of differential constraints, as for example in the description of the motion of vehicles. A car moves on a bidimensional plane, but it can only move in its current direction or it can change its current orientation by rotating the steering wheel. These are the admissible directions. Moreover, the car cannot move "sideways" (forbidden direction): this prevents one from directly reaching any other direction while remaining in the initial position, restricting the allowable motions to a simultaneous combination of the two admissible movements. The trajectory described by the vehicle will therefore be a curve, whose tangent is constrained to follow the two admissible directions. The formalization of this sub-Riemannian problem takes place in $SE(2)$, considering for every $p \in SE(2)$ as admissible tangent space $A_pSE(2)$ the subspace generated by the current direction $\vec{X}_{1,p} = (\cos\theta, \sin\theta, 0)^T$ and the direction of rotation $\vec{X}_{2,p} = (0, 0, 1)^T$. See Figure A.1, images (E-F).

Similarly, we can move from a retinotopic (x, y) position to another retinotopic position, (x', y'), moving "up" or "down" through orientation columns from θ to θ', but we cannot reach θ' from θ maintaining the same initial position (running through the same orientation column): in order to reach the "forbidden direction" we have to walk simultaneously through positions and orientations.

This restriction of movement is what distinguishes a Euclidean (or Riemannian) geometry from a sub-Riemannian geometry.

B. Proof of Proposition 3.1

In this appendix, we show ho to prove Proposition 3.1 using tools of differential geometry, and in particular the concept of differential $k-$form.

B.1. **Differential forms.** A differential k-form on an n-dimensional smooth manifold M is any multilinear function $\omega : TM^k \longrightarrow \mathbb{R}$ which takes as input k smooth vector fields and outputs a scalar element, satisfying the antisymmetry property:

$$\omega(X_1, \ldots, X_i, \ldots, X_j, \ldots, X_k) = -\omega(X_1, \ldots, X_j, \ldots, X_i, \ldots, X_k),$$

with $k \leq n$ and $k, n \in \mathbb{N}$.

In the special case where ω is a 1-form, it is worth noting that this is an element of the dual space to TM (*cotangent space*): $\omega \in TM^* \iff \omega : TM \longrightarrow \mathbb{R}$. If we have coordinates $(x_1, \ldots, x_n)$ on M, we can express the 1-forms using the dual basis $\{\mathrm{d}\, x_1, \ldots, \mathrm{d}\, x_n\}$ of TM^*:

$$\omega_p = f_1(\bar{x}_1, \ldots, \bar{x}_n)\, \mathrm{d}\, x_1 + \ldots + f_n(\bar{x}_1, \ldots, \bar{x}_n)\, \mathrm{d}\, x_n, \text{ with } p = (\bar{x}_1, \ldots, \bar{x}_n),$$

with f_i scalar smooth functions.

Furthermore, it is possible to multiply via the wedge product $\wedge$ a differential k-form, ω, with a differential l- form, η, obtaining a differential $k + l$-form $\omega \wedge \eta$. More precisely, we are interested in the wedge product of 1-forms ω and η, where the wedge product can be computed as: $\omega \wedge \eta(X, Y) = \omega(X)\eta(Y) - \omega(Y)\eta(X)$, with X and Y vector fields on M.

B.2. **Development of the proof.**

Proposition B.1. *The binocular interaction term $O_L O_R$ can be associated with the cross product of the left and right directions defined through (13), namely $\omega^\star_{p_L}$ and $\omega^\star_{p_R}$ of monocular simple cells:*

$$O_L O_R = \omega^\star_{p_L} \times \omega^\star_{p_R}.$$

Proof. As noted in subsubsection 2.2.1, the output of simple cells (11) in $SE(2)$ can then be locally approximated as $O(x, y, \theta) = -X_{3,p}(I_\sigma)(x, y)$ where I_σ is a smoothed version of I, obtained by convolving it with a Gaussian kernel, the vector field

$$X_{3,p} = -\sin\theta \partial_x + \cos\theta \partial_y,$$

with $p = (x, y, \theta) \in SE(2)$. Switching to the dual space, the action of simple cells induces a choice of a 1-form separately on each cell:

$$\omega_p = -\sin\theta\, \mathrm{d}\, x + \cos\theta\, \mathrm{d}\, y.$$

Accordingly, it is possible to re-write the binocular interaction term as:

$$O_L O_R = \; X_{3,p_R}(I_{\sigma R})(x_R, y) X_{3,p_L}(I_{\sigma L})(x_L, y).$$

In the following, we will see that this binocular action can be described by a 2-form defined in terms of the two 1-forms of monocular simple cells.

We will denote with the subscript R the quantities corresponding to the right monocular structure, and we will use the subscript L for the left one. So, we define $v_R := (J_{I_{\sigma_R}}\vec{X}_{3,p_R})X_{3,p_R}$ using the Jacobian (differential) of the smoothed version of the image I, in such a way that we have $\omega_{p_R}(v_R) = X_{3,p_R}(I_{\sigma_R}) = (J_{I_{\sigma_R}}\vec{X}_{3,p_R})$ since $\omega_{p_R}(X_{3,p_R}) = 1$ and $J_{I_{\sigma_R}}\vec{X}_{3,p_R} \in \mathbb{R}$; the same reasoning holds for the left structure. It is then possible to recast (B.2) in the retinal coordinates as:

$$\begin{aligned} O_L O_R =& \omega_{p_L}(v_L)\omega_{p_R}(v_R) \\ =& \omega_{p_L} \wedge \omega_{p_R}(v_L, v_R) + \underbrace{\omega_{p_R}(v_L)\omega_{p_L}(v_R)}_{=0}, \\ =& \omega_{p_L} \wedge \omega_{p_R}(v_L, v_R), \end{aligned}$$

exploiting the properties of the wedge product and the left and right retinal coordinates.

The retinal coordinates can be expressed in terms of cyclopean coordinates (4) as $x_R = x - d$ and $x_L = x + d$; then, the extended left and right 1-form can be written as:

$$\begin{aligned} \omega_{p_R} =& -\sin\theta_R \,\mathrm{d}\,x + \cos\theta_R \,\mathrm{d}\,y + \sin\theta_R \,\mathrm{d}\,d \\ \omega_{p_L} =& -\sin\theta_L \,\mathrm{d}\,x + \cos\theta_L \,\mathrm{d}\,y - \sin\theta_L \,\mathrm{d}\,d. \end{aligned}$$

Taking advantage of the isomorphism provided by the Hodge star between vectors and 2-forms in $\mathbb{R}^3$, we relate the exterior and the cross product, using above notations [1], in the following way:

$$\star(\omega_{p_L} \wedge \omega_{p_R}) = \omega_{p_L}^{\star} \times \omega_{p_R}^{\star},$$

from which it follows the thesis. □

Throughout the paper, to lighten the notation, we will call $\omega_L = \omega_{p_L}$ and $\omega_R = \omega_{p_R}$.

B.2.1. *Meaning of the mathematical objects.* We conclude this section with a consideration on the mathematical tools introduced and used in this setting, to understand how the mathematical models proposed by Citti and Sarti, starting with [6], assign these different mathematical objects to the physical cell, to its action , and to the result of its action.

Remark B.1. *It is well known that an odd simple cell (selective for orientation) is activated as a result of the presence of a stimulus to select its direction (tangent vector to the perceptual curve). In this setting, the mathematical intuition behind the model proposed in* [6] *is to identify each cell with a* 1*-differential form, which is an element of the cotangent space. Roughly speaking, this differential form is able to grasp a vector that corresponds to the direction of the stimulus: this is the result of the action of the cell. Formally, this vector will be an element of the tangent space, and more precisely it will lay in the kernel of the* 1*-form. This vector space is then associated with the action of the cell.*

The same reasoning is applied to different families of cells in a series of papers ([8, 3, 1, 4]) even if these are characterized by distinct sub-Riemannian structures in various manifolds. The

[1] Using the notation $\omega^{\star}$ we identify the vector whose components are the coefficients of the 1-form ω with respect to the dual basis

interested reader could refer to [7] for a review. Similarly, we have found the same geometrical organization in the family of binocular cells.

Remark B.2. *In this paper, we have dealt with binocular cells which are a combination of monocular simple cells. To these coupled simple cells (one for the left and one for the right eye) we formally associate a 2-differential form, the wedge product of the two monocular left and right 1-forms. This 2-form can grasp again a vector, lying in the kernel of this mathematical object, identifying the three-dimensional stimulus direction. Thus, the same reasoning of Remark B.1 also applies here to the binocular family of cells.*

Translating the results of Remark B.1 into different spaces, with different dimensions, it is then possible to use the same mathematical objects to explain the behavior of families of different cells, identifying geometrically the mathematical objects at the basis of the functionality of the family of studied cells.

C. Change of variables

Let us recover the expression of the 1-forms $\tilde{\omega}_L := U_{t_L}$ and $\tilde{\omega}_R := U_{t_R}$. Recall here the change of variable (5):

$$\begin{cases} r_1 = \frac{xc}{d} \\ r_2 = \frac{yc}{d} \\ r_3 = \frac{fc}{d} \end{cases},$$

and its differential:

$$\begin{cases} \mathrm{d}\, r_1 = \frac{c}{d}\,\mathrm{d}\, x - \frac{cx}{d^2}\,\mathrm{d}\, d \\ \mathrm{d}\, r_2 = \frac{c}{d}\,\mathrm{d}\, y - \frac{cy}{d^2}\,\mathrm{d}\, d \\ \mathrm{d}\, r_3 = -\frac{fc}{d^2}\,\mathrm{d}\, d \end{cases}.$$

Writing the quantity U_{t_L}, defined in (25), in term of a 1-form in the variables (r_1, r_2, r_3) we have:

$$\tilde{\omega}_L = -f\sin\theta_L\,\mathrm{d}\, r_1 + f\cos\theta_L\,\mathrm{d}\, r_2 + (x_L\sin\theta_L - y\cos\theta_L)\,\mathrm{d}\, r_3.$$

Changing coordinates:

$$\begin{aligned} \tilde{\omega}_L = & -f\sin\theta_L\left(\frac{c}{d}\,\mathrm{d}\, x - \frac{cx}{d^2}\,\mathrm{d}\, d\right) + f\cos\theta_L\left(\frac{c}{d}\,\mathrm{d}\, y - \frac{cy}{d^2}\,\mathrm{d}\, d\right) \\ & + (x_L\sin\theta_L - y\cos\theta_L)\left(-\frac{fc}{d^2}\,\mathrm{d}\, d\right) \\ = & \frac{fc}{d}\left(-\sin\theta_L\,\mathrm{d}\, x + \cos\theta_L\,\mathrm{d}\, y - \sin\theta_L\,\mathrm{d}\, d\right) \\ = & \frac{fc}{d}\omega_L. \end{aligned}$$

So, up to a scalar factor, we have that $\tilde{\omega}_L = \omega_L$ in the variables (x, y, d). The same reasoning holds for the right structure.

References

[1] S. Abbasi-Sureshjani, M. Favali, G. Citti, A. Sarti, and B. M. ter Haar Romeny. Curvature integration in a 5d kernel for extracting vessel connections in retinal images. *IEEE Transactions on Image Processing*, 27(2):606–621, 2017.

[2] A. Agrachev, D. Barilari, and U. Boscain. *A comprehensive introduction to sub-Riemannian geometry*, volume 181. Cambridge University Press, 2019.

[3] D. Barbieri, G. Citti, G. Cocci, and A. Sarti. A cortical-inspired geometry for contour perception and motion integration. *Journal of mathematical imaging and vision*, 49(3):511–529, 2014.

[4] E. Baspinar, A. Sarti, and G. Citti. A sub-riemannian model of the visual cortex with frequency and phase. *The Journal of Mathematical Neuroscience*, 10(1):1–31, 2020.

[5] G. Citti, L. Grafakos, C. Pérez, A. Sarti, and X. Zhong. *Harmonic and geometric analysis*. Springer, 2015.

[6] G. Citti and A. Sarti. A cortical based model of perceptual completion in the roto-translation space. *Journal of Mathematical Imaging and Vision*, 24(3):307–326, feb 2006.

[7] G. Citti and A. Sarti. *Neuromathematics of vision*, volume 32. Springer, 2014.

[8] A. Sarti, G. Citti, and J. Petitot. The symplectic structure of the primary visual cortex. *Biological Cybernetics*, 98(1):33–48, nov 2007.

[9] A. Sarti, G. Citti, and D. Piotrowski. Differential heterogenesis and the emergence of semiotic function. *Semiotica*, 2019(230):1–34, 2019.

[10] L. W. Tu. *An Introduction to Manifolds*. Springer, 2011.