Alma Mater Studiorum - Università di Bologna
in cotutela con Sorbonne Université

DOTTORATO DI RICERCA IN

MATEMATICA

Ciclo 35

Settore Concorsuale: 01/A3 - ANALISI MATEMATICA, PROBABILITÀ' E STATISTICA MATEMATICA

Settore Scientifico Disciplinare: MAT/05 - ANALISI MATEMATICA

NEUROGEOMETRY OF STEREO VISION

Presentata da: Maria Virginia Bolelli

Coordinatore Dottorato

Valeria Simoncini

Supervisore

Giovanna Citti

Supervisore

ALESSANDRO SARTI

Esame finale anno 2023

Abstract

This work aims to develop a neurogeometric model of stereo vision, based on cortical architectures involved in the problem of 3D perception and neural mechanisms generated by retinal disparities. We apply it to reproduce phenomenological experiments as well as to process 3D images, identifying three-dimensional visual percepts in space, solving the correspondence problem.

First, we provide a sub-Riemannian geometry for stereo vision. This proposal is inspired by the work on the stereo problem by Li and Zucker in [LZ06], and it uses sub-Riemannian tools introduced by Citti and Sarti in [CS06] for monocular vision. In particular, we present a mathematical interpretation of the neural mechanisms underlying the behavior of binocular cells, that integrate monocular inputs, by introducing a suitable cortical fiber bundle. The natural compatibility between stereo geometry (stereo triangulation [Fau93]) and neurophysiological models (binocular energy model [AOF99b]) shows that these binocular cells are sensitive to position and orientation. Therefore, we model their action in the space $\mathcal{M} = \mathbb{R}^3 \times \mathbb{S}^2_{\theta,\varphi}$ equipped with a sub-Riemannian metric.

Integral curves of the sub-Riemannian structure shed light on the computations underlying the correspondence problem. They encode not only the variables of the space, but also curvature and torsion, which are necessary for the 3D matching. Moreover, a fan of these curves can model the 3D analog of the psychophysical association fields of Field, Heyes and Hess [FHH93], for the 3D process of regular contour formation, studied by psychophysical experiments in [HF95, HHK97, KHK16]. This illustrates how good continuation in the world generalizes good continuation in the plane, as conjectured in [LZ06].

As a second step, we study the constitution of 3D perceptual units in the three-dimensional scene generated from the sub-Riemannian geometry. Following

[BCCS14], these perceptual units emerge as a consequence of the random cortico-cortical connection of binocular cells. We present the stochastic process at the basis of this phenomenon, by considering an opportune stochastic version of the integral curves. We generate a family of kernels through the forward Kolmogorov operator associated to the generated stochastic process, and approximate them using a Monte Carlo simulation-based method.

These kernels represent the probability of interaction between binocular cells and are implemented as facilitation patterns to define the evolution in time of neural population activity at a point $\xi \in \mathcal{M}$. This activity is usually modeled through a mean field equation.

The existence and uniqueness of a solution classically follow from the Cauchy problem in Banach spaces, while the stability analysis is performed using the Lyapunov method, leading to the consideration of the associated eigenvalue problem. We show that three-dimensional perceptual units naturally arise from the discrete version of the eigenvalue problem associated to the integro-differential equation of the population activity.

Résumé en français

Ce travail vise à développer un modèle neurogéométrique de la vision stéréo, basé sur les architectures corticales impliquées dans le problème de la perception 3D et les mécanismes neuronaux générés par les disparités rétiniennes. Nous l'appliquons pour reproduire des expériences phénoménologiques ainsi que pour traiter des images 3D, en identifiant des percepts visuels tridimensionnels dans l'espace, en résolvant le problème de correspondance (*the correspondence problem*).

Tout d'abord, nous proposons une géométrie Sous-Riemannienne pour la vision stéréo. Cette proposition s'inspire du travail sur le problème de la stéréo effectué par Li et Zucker dans [LZ06], et elle utilise les outils Sous-Riemanniens introduits par Citti et Sarti dans [CS06] pour la vision monoculaire. En particulier, nous présentons une interprétation mathématique des mécanismes neuronaux qui sous-tendent le comportement des cellules binoculaires, qui intègrent les entrées monoculaires, en introduisant un espace fibré corticales approprié. La compatibilité naturelle entre la géométrie stéréo (triangulation stéréo [Fau93]) et les modèles neurophysiologiques (modèle d'énergie binoculaire [AOF99b]) montre que ces cellules binoculaires sont sensibles soit à la position et soit à l'orientation. Nous modélisons donc leur action dans l'espace $\mathcal{M} = \mathbb{R}^3 \times \mathbb{S}^2_{\theta,\varphi}$ équipé d'une métrique Sous-Riemannienne.

Les courbes intégrales de la structure Sous-Riemannienne éclairent les calculs qui sous-tendent le problème de correspondance. Elles codent non seulement les variables de l'espace, mais aussi la courbure et la torsion, qui sont nécessaires pour résoudre la correspondance 3D. De plus, un éventail de ces courbes peut modéliser l'analogue 3D des champs d'association psychophysiques de Field, Heyes et Hess [FHH93] ; il s'agit d'un processus de formation de contours réguliers, étudié par des expériences psychophysiques dans [HF95, HHK97, KHK16]. Ceci illustre comment une bonne continuation dans le monde généralise une bonne continuation en 2D,

comme conjecturé dans [LZ06].

Ensuite, nous étudions la constitution d'unités perceptuelles 3D dans la scène tridimensionnelle générée à partir de la géométrie sub-riemannienne. Suivant [BCCS14], ces unités perceptuelles émergent comme une conséquence de la connexion cortico-corticale aléatoire des cellules binoculaires. Nous présentons le processus stochastique à la base de ce phénomène, en considérant une version stochastique opportune des courbes intégrales. Nous générons une famille de solutions fondamentales pour l'opérateur de Kolmogorov direct associé au processus stochastique généré, en les approximant à l'aide d'une méthode basée sur la simulation de Monte Carlo.

Cette famille représente la probabilité d'interaction entre les cellules binoculaires et est mise en œuvre comme un modèle de facilitation pour définir l'évolution dans le temps de l'activité de la population neuronale à $\xi \in \mathcal{M}$. Cette activité est généralement modélisée par une équation de champ moyen.

L'existence et l'unicité d'une solution découlent classiquement du problème de Cauchy dans les espaces de Banach. En revanche, l'analyse de stabilité est effectuée à l'aide de la méthode de Lyapunov, ce qui conduit à la prise en compte du problème des valeurs propres associé. Nous démontrons que les unités perceptuelles tridimensionnelles apparaissent de façon naturelle, à partir de la version discrète du problème des valeurs propres, associé à l'équation intégro-différentielle de l'activité de la population.

Contents

Chapter 1

Introduction

In this thesis, we present a mathematical model of the neural mechanisms underlying the perception of three-dimensional visual stimuli. Since differential geometric tools are used in this interpretation, we speak about *neurogeometry of stereo vision*, as the word *neurogeometry* was introduced by Jean Petitot in [PT99] to refer to geometrical models of the functional architecture of primary visual areas, which include mathematical models concerning the neural origin of the perceived space.

Precisely, stereo vision systems recover the 3D images from projections onto the two retinae, projections which are slightly different, due to the horizontal separation of the eyes. These differences are usually referred to as disparities and they are the main clues for 3D visual perception. Our main focus will be on *horizontal disparity* d, defined by introducing a set of coordinates:

$$x = \frac{x_L + x_R}{2}, \quad y = y_L, \quad d = \frac{x_L - x_R}{2} \tag{1.1}$$

with (x_L, y_L) left retinal coordinates, (x_R, y_R) right retinal coordinates, and noting that $y_L = y_R$ since we are projecting onto aligned parallel planes. This set of coordinates is usually referred to as the *cyclopean coordinate*, because we perceive the world like we have a single eye in the middle of the head, [Jul71].

The three-dimensional world must be reconstructed from two-dimensional retinal projections. The major difficulty is understanding how to pair points in the left retinal plane with corresponding points in the right retinal plane. This inverse optics problem is fundamentally underdetermined because a given pair of retinal images coincides with many distal stimuli [Piz01, Bru21], leading to the *stereo cor-*

respondence problem. The principal hints used (by the brain or computer vision algorithms) to reconstruct the 3D world and thus perceive the visual scene are precisely the disparities.

The brain processes the visual signals coming from the two eyes along the visual pathway. In particular, the integration of the (visual) binocular signals is generally thought to occur in the primary visual cortex. By studying cells stimulated by the depth feature in V1, Ohzawa et al. in '99 [AOF99b] found that these cells perform a non-linear integration of left and right monocular cells that are selective for orientation. A classic model of binocular cell behavior based on physiological experiments is the binocular energy model (BEM) first presented in [AOF99b], which encodes disparities through the receptive profiles of simple cells, leading to the notion of position and phase disparities [JR15]. However, Read and Cumming proposed in [RC07], building upon [AOF99a], that phase disparity neurons tend to be strongly activated by false correspondence pairs. Therefore, it is widely concluded, the most relevant disparity in the receptive fields is the position alone. This, however, neglects the orientation difference between the two eyes [NKB77], disregarding the orientation disparity. Although there are attempts to extend the energy model to incorporate binocular differences in receptive-field orientation [BCP01], they are limited. The geometrical model we will present incorporates orientation differences directly.

The goal of this work is to propose a neurogeometric model for the cortical-inspired geometry underlying the stereo vision, based on the encoding of position and orientation differences in the two eyes. The encoding of features in the early stages of the visual process is the focus of neuromathematical models, developed from the work of Hoffmann [Hof89] and Koenderink-van Doorn [KvD87]. Our neuro-geometrical model considers essential the geometric frameworks used for the descriptions of the functional architecture of V1 proposed by Petitot and Tondut [PT99], Bresslof and Cowan [BC03], Citti and Sarti [CS06], Zucker [Zuc06], and many others works [Pet08, DBRS13, SCS10a, SC15, SCP07].

The starting point is the study of the problem from both technical (stereo triangulation) and neural (binocular energy model) perspectives. If we assume that a three-dimensional stimulus is a curve $\gamma : \mathbb{R} \longrightarrow \mathbb{R}^3$, we can use the classical tools of stereo geometry to determine both the position and the direction of γ given the

corresponding retinal points. Similarly, binocular neurons encode the orientation of the three-dimensional scene as a first-order information of the stimulus: the binocular cells naturally recognize the unitary tangent direction $\dot{\gamma}$ to the 3D spatial stimulus γ, which can be described by the parameters $n = n(\theta, \varphi)$ of $\mathbb{S}^2$. This natural compatibility between neural elements and stereo geometry allows us to suitably define the space of stereo cells as the manifold of positions and orientations $\mathcal{M} = \mathbb{R}^3 \rtimes \mathbb{S}^2$.

The coupling of positions and orientations, as already noted by Duits and Franken in [DF11], allows to endow the manifold with a sub-Riemannian structure by using left-invariant vector fields of $SE(3)$ acting on $\mathcal{M}$, thus defining the horizontal connectivity for the space. This action is quite natural since $\mathcal{M}$ is a quotient of $SE(3)$ by the one-parameter subgroup $\{0\} \times SO(2)$. Under these conditions, the sub-Riemannian structure on $\mathcal{M}$ can be expressed locally by considering a local orthonormal frame $\{Y_3, Y_\theta, Y_\varphi\}$ for the distribution $\mathcal{D}$, where the vector field Y_3 encodes the tangent of the stimulus, while Y_θ and Y_φ involve orientation directions.

Integral curves with constant coefficients of the sub-Riemannian structure, defined by the differential equation

$$\mathrm{d}\,\Gamma(t) = Y_{\mathbb{R}^3,\Gamma(t)}\,\mathrm{d}\,t + c_1 Y_{\theta,\Gamma(t)}\,\mathrm{d}\,t + c_2 Y_{\varphi,\Gamma(t)}\,\mathrm{d}\,t, \quad c_1, c_2 \in \mathbb{R}, \tag{1.2}$$

shed light on the computations underlying the correspondence problem. They encode not only the variables of the space, but also curvature and torsion, which are necessary for the 3D matching, as favored by Zucker et al. in [AZ00, LZ03, LZ06]. Curvature and torsion of the restriction of the integral curves (1.2) in $\mathbb{R}^3$ are expressed in terms of the coefficients $c_1, c_2 \in \mathbb{R}$:

$$k = \sqrt{c_1^2 + c_2^2}, \qquad \tau = \frac{c_1^2 - c_2^2}{k^2} \operatorname{cotan} \varphi \tag{1.3}$$

Planar curves of this family reduce to circular arcs, and are compatible with the integral curves of the monocular model proposed by Citti-Sarti in [CS06]. Also helices in the direction of depth belong to the family (1.2), introducing a geometrical setting in which the curves conjectured by Alibhai-Zucker in [AZ00] are natural.

These curves generalize to the three-dimensional case the Gestalt law of *good continuation*, one of the laws underlying the formation of percepts in space and introduced in seminal psychological works such as [Kan79, Wer23, Kof35, Kof63,

Koh67. The law of good continuation states that elements that are aligned in space tend to be perceived as belonging to the same perceptual unit. How the visual system encodes this regular contour formation principle was theorized by Field, Hayes and Hess in FHH93 through psychophysical experiments for the bi-dimensional case. They introduced the concept of *association field*, which essentially describes the pattern that determines whether stimuli can be associated with the same perceptual unit in terms of position and orientation: a stronger correlation characterizes mainly aligned or co-circular elements.

The generalization of the 3D process of regular contour formation has been studied through psychophysical experiments in HF95, HHK97, KHK16, DW15, KGS05a, KGS$^+$05b. Kellman, Garrigan, Shipley et al. KGS05a, KGS$^+$05b generalize to the three-dimensional case the concept of alignment and co-circularity introduced by Field, Hayes and Hess FHH93 in the two-dimensional case. They introduce a theory called 3D relatability which basically imposes conditions on edges' orientations in space. Moreover, they propose that the strength of the relatable edges in co-planar planes with the initial edge must meet the relations of the bi-dimensional association fields. The properties of the curves that are suitable for connecting these relatable points have been investigated in HHK97, KHK16, DW14, DW15, and they are described as being smooth and monotonic. We model associations underlying the 3D perceptive organization with the family of integral curves (1.2). These curves satisfy the above properties and locally connect the fan of associations generated by the geometry of 3D relatability.

The extension of the law of good continuation and its formalization through the concept of association field makes it possible to identify the visual units in 3D space. The first step for understanding this identification is to consider the stochastic counterpart of the involved integral curves. These are assumed to describe the interactions between the (binocular) cells involved. The stochastic process underlying this phenomenon is described by a stochastic differential equation:

$$\mathrm{d}\,\Gamma(t) = Y_{\mathbb{R}^3,\Gamma(t)}\,\mathrm{d}\,t + \lambda(Y_{\theta,\Gamma(t)}, Y_{\varphi,\Gamma(t)})\,\mathrm{d}\,B(t), \qquad \lambda \in \mathbb{R}, \tag{1.4}$$

with $B(t)$ 2-dimensional Brownian motion. The probability of reaching a point ξ, starting from a point ξ' along these curves has a (time-independent) density:

$$\{J_\lambda(\xi, \xi')\}_{\lambda\in\mathbb{R}}, \tag{1.5}$$

which coincides with the (time-integrated) fundamental solution of the forward Kolmogorov equation associated to (1.4) with operator

$$\mathcal{L} = -Y_3 + \lambda(Y_\theta^2 + Y_\varphi^2) \tag{1.6}$$

written in terms of the chosen vector fields. These PDEs, in principle very difficult to deal with, have been solved exactly in [PD17, DBM19] using Fourier transforms and validated with numerical Fourier methods and Monte-Carlo simulations in [DBM19]. The latter one is the method we implement, in order to follow the approach presented in [BCCS14] based on the stochastic integral curves of the sub-Riemannian structure.

The evolution of a state of a population of cells has been first modeled through mean field equations by Wilson-Cowan [WC72, WC73] and Amari [Ama72], then extended by Ermentrout and Cowan in [EC79, EC80] and by Bresslof and Cowan in [BCG+02, BC03]. Other modified models that take into account the delay of the cortex are proposed by Faye and Faugeras in [FF10].

Following the work of Sarti and Citti [SC15], the result of the propagation along the connectivity of the visual impulse can be described, without considering the delays for the transmission of the signal, by the following integro-differential equation in the 3D perceptive space of positions and orientations:

$$\frac{da(\xi,t)}{dt} = -a(\xi,t) + \int \mu J_\lambda(\xi,\xi')a(\xi',t)d\xi' + h(\xi,t), \text{ in } \mathcal{M} \tag{1.7}$$

where $\xi \in \mathcal{M}$, $h(\xi,t)$ is the feedforward input, $\mu \in \mathbb{R}$. The existence and uniqueness of a solution classically follow using instruments of functional analysis, recasting equation (1.7) as a Cauchy Problem.

Stability analysis is performed through the Lyapunov method applying the results of Faye-Faugeras [FF10] in our setting. The existence of a stable steady state of (1.7) leads to an eigenvalue problem. In analogy with the work of Perona and Freeman in [PF98], we show that the eigenvectors associated with the largest eigenvalues encode the greater salience of objects in the analyzed visual scene, leading to the identification of the perceptual units in space. An opportune normalization has been introduced to improve the performance of the affinity matrix in [MS01], with the introduction of the transition matrix of a Markov chain. In this setting, stable solutions are computed as eigenvectors of this matrix, and they

represent perceptual units in the three-dimensional visual scene.

The thesis is organized as follows:

Chapter 2 starts by briefly reviewing the basis underlying visual sensory perception. We first describe the visual cortex, introducing the visual pathways and concentrating on the monocular and binocular organization of neurons in the primary visual cortex V1. Then, we present the perceptual phenomena that govern the formation of perceptual units in space, investigating the Gestalt laws, with particular attention to the law of good continuation and its neural formalization through the introduction of the notion of association field by Field, Hayes and Hess in [FHH93].

Chapter 3 provides a background for the stereo problem. First, we introduce classical stereo geometry. Then, we study neural models proposed for describing the behavior of binocular cells, introducing the classical binocular energy model. Finally, presenting the constraints that are usually imposed for solving the stereo correspondence problem, we conclude with the description of our main reference model for the solution of the stereo problem: the proposal of Zucker et al. in [AZ00, LZ03, LZ06].

Chapter 4 gives a presentation of the main mathematical tools, definitions, and notions of sub-Riemannian geometry used to develop geometrical models for the functional architecture of the primary visual cortex. We present the notions of fiber bundles, integral curves, Lie groups and algebras, sub-Riemannian manifolds. Finally, we describe its applications to vision, presenting the sub-Riemannian model for monocular cells selective for orientation introduced by Citti and Sarti in [CS06].

Chapter 5 contains our first original contribution. Starting from binocular receptive profiles, we introduce the neuro-mathematical model for binocular cells. First, we define the cortical fiber bundle of binocular neurons, followed by the differential interpretation of the binocular profiles in terms of the neurogeometry of the simple cells. Then, we give a mathematical definition of the manifold $\mathbb{R}^3 \rtimes \mathbb{S}^2$ with the sub-Riemannian structure we choose and we study its integral curves and the properties of these curves in relation to the notion of torsion and curvature. Finally, we proceed to the validation of our geometry with respect to biological and psychophysical phenomena present in literature. Integral curves model neural connectivity, and they can be related to the geometric relationships from psychophysical experiments on perceptual organization of oriented elements in $\mathbb{R}^3$, generalizing the concept of an association field in 3D.

Chapter 6 deals with the individuation of three-dimensional perceptual units in the visual scene. We start with the introduction of a stochastic model for cortical connectivity, considering the stochastic differential equation (1.4). We compute its probability density through the fundamental solution of the generator of the associated forward Kolmogorov operator. Numerical implementation of the kernel is performed via a Monte Carlo simulation. Then, we introduce the neural activity of the population via an integro-differential equation, studying the existence and uniqueness of the solution through a Cauchy problem in Banach spaces, and we perform stability analysis using a variation of the Lyapunov method introduced in [FF10]. Finally, we see how the discretization of the problem leads to spectral clustering. We perform some numerical experiments to validate our algorithm, and we conclude with a comparison with Gaussian type kernel to emphasize the role of our chosen metric.

The last Chapter is devoted to conclusions.

Chapter 2

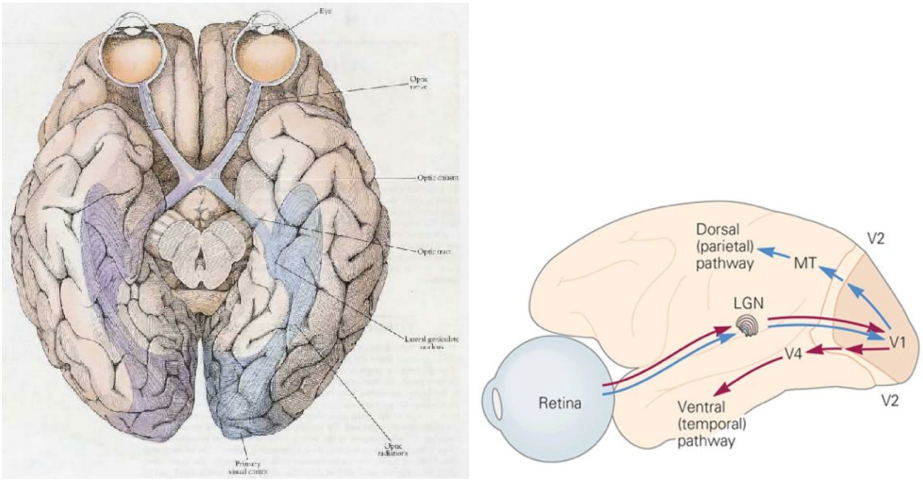
Neurophysiology and psychophysics of vision

In this chapter we briefly give an introduction to the basis of visual sensory perception. Let us first clarify that perception does not necessarily coincide with sensation: sensation *is the registration of the physical stimuli on the sensory receptors and* perception *is the process of creating a conscious perceptual experience from sensory input. We first summarize a background on the main structures of the cortical areas involved in the vision process, with particular attention devoted to the early visual pathways and the primary visual cortex. Then, we investigate how the stimuli in the outside world become a perceptual experience.*

Contents

2.1 The cerebral cortex and the (early) visual pathway

The visual system begins to process the real world by converting the light entering the eye into a neural signal. This process is called *transduction* and can be interpreted in terms of pixelation of the external stimulus. The generated action potentials are conveyed through the optical nerve in the lateral geniculate nucleus (LGN), and then they spread in the main visual sensory areas. In the following, we review this procedure.



(a) (b)

Figure 2.1.1: Visual pathway in the brain. (a) Left (violet) and right (blue) cerebral areas involved in the visual pathway, image taken from [Hub95]. (b) Cortical regions implicated in the process.

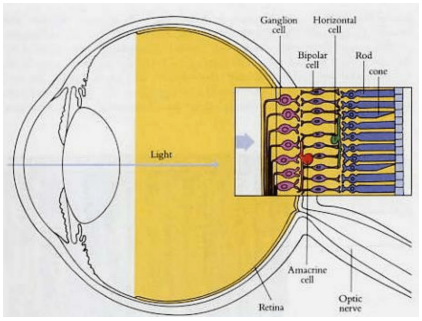
2.1.1 Retina and LGN

The neural mechanisms involved in the visual system originate in the retina, the first region where the light arrives from entering through the eye.

Anatomically, the retina has many layers of brain tissue, in particular, the vertebrate retina is inverted so light has to pass through layers of neurons (as for example ganglion and bipolar cells), before reaching the photosensitive sections of rods and cones (visual sensory photoreceptors), which are the closest to the optic nerve. The ganglion cells, whose axons form the optic nerve, are at the front of

the retina: the optic nerve must cross through the retina to reach the brain, see for example Figure 2.1.2.

Functionally, two main phenomena take place in the retina: the transduction and the transmission of the signal.



(a)

Figure 2.1.2: Anatomic structure of the retina. Image adapted from [Hub95].

The transduction process translates the arriving light into action potentials by visual sensory receptors, measuring the physical energy of light signals and pixelating them, converting the light into electrical impulses. Ganglion (or bipolar) cells receive input from a group of adjacent photoreceptors.

Then, the signals coming from ganglion (or bipolar) cells are fed to the brain, broadcasted into the optical nerve, which transmits the visual information to the lateral geniculate nucleus (LGN). The LGN is a bilateral[1] structure in the thalamus that relays information from the optic nerve to (visual) sensory processing cortical areas. The most important area is the primary visual cortex, where most of the retinal output first arrives.

2.1.2 The primary visual cortex V1

The primary visual cortex receives input from the LGN, it is responsible for the early visual processing, and it is the most widely studied visual area: this is why it is referred to by different names in different contexts. For instance, we recall here

- *primary visual cortex V1*: the position of the area along the flow of information in the visual system;

[1] It means that one is present in each hemisphere of the brain.

- *striate cortex*: the way the brain cells in this area look under certain conditions;

- *area 17 and BA 17* : area's position on Brodmann charts.

Anatomically, as shown in Figure 2.1.1, the primary visual cortex is a bilateral structure in the occipital lobe of the brain: there is one in the left hemisphere, which receives inputs from the right visual field, and there is one in the right hemisphere, which receives inputs from the left visual field. The two sides of V1 are connected via the corpus callosum. Moreover, V1 is divided into 6 horizontal layers with a characteristic distribution of inputs and outputs across layers [DM98].

The most important typical feature it is its high organization: the anatomical structure can be correlated directly with its functions, as we will see in the next paragraphs.

2.1.2.1 Retinotopic organization

The propagation of the visual signal from the retina to the visual cortices is performed through the optical nerve. The induced map from the retina to the cortex respect the topographic organization of the retina: adjacent locations on the retina project to adjacent neurons in V1 by a point-by-point relationship, through a topographic map, which is maintained in the LGN. This phenomenon goes under the name of *retinotopy*. This map is not isotropic: some regions of the retina get

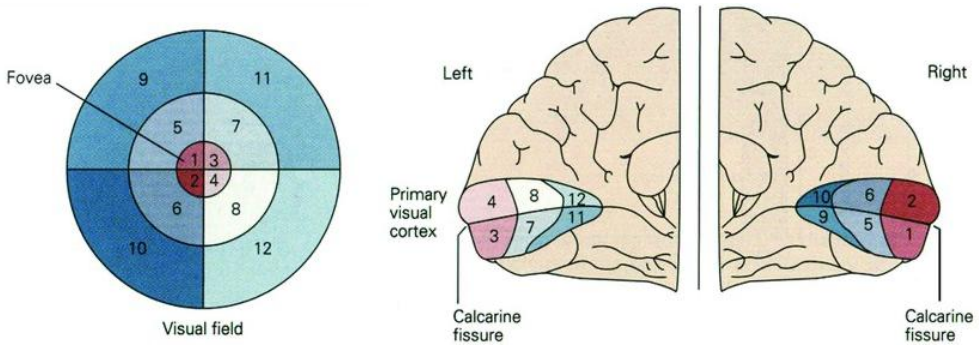


Figure 2.1.3: Retinotopic mapping of the visual field (left) in the primary visual cortices (right). Image adapted from [PWK18].

to take up a much greater proportion of V1 than others. This feature is called *cortical magnification*: there is more space in the cortex devoted to some sensory receptors than to others.

2.1.2.2 Different types of neurons

Neurons that populate V1 can be of different types. Hubel and Wiesel in [HW62] discovered particular families:

- *simple cells*: neurons characterized by a strong push-pull antagonism built in terms of elongated, adjacent ON and OFF subregions that have a mutually suppressive influence. They respond to stimuli with particular orientations within their receptive field.

- *complex cells*: neurons that also respond optimally to stimuli with particular orientations. But, unlike simple cells, they respond to a variety of stimuli across different locations, regardless of where it is within their receptive field.

- *hypercomplex cells*: neurons that display orientation selectivity but are confined to stimuli of a limited size. This property is referred to as *end-stopping*. If for example, the pattern continues beyond the receptive field, these cells do not respond as greatly.

- *binocular cells*: neurons that receive inputs from both the left and right eyes and integrate the signals to create a perception of depth. They respond excellently to binocular disparity and are thought to be the first stage for the processing of 3D vision.

2.1.2.2.1 Cell's receptive profile Each cell reacts in an excitatory or inhibitory way to stimuli present in a precise portion of the retina called receptive field (RF), which differs from a cell to the other. The response of the cell when a stimulus is presented on a point (x, y) of their RF, is quantified in terms of a function $\varphi(x, y)$, called receptive profile (RP)

$$\varphi : D \to \mathbb{R}$$

where D is the receptive field and (x, y) are retinal coordinates. This function describes the neural output of the cell in response to a punctual stimulus in the two-dimensional point (x, y). More precisely, the area is considered ON if the cell spikes in a positive way responding to a signal (excitatory response to a light stimulus) and OFF if the cell spikes negatively to a signal (inhibitory response to a light signal).

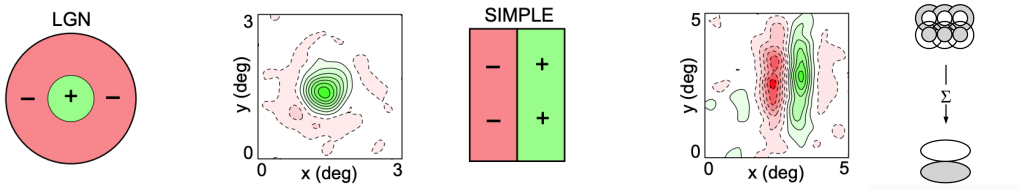
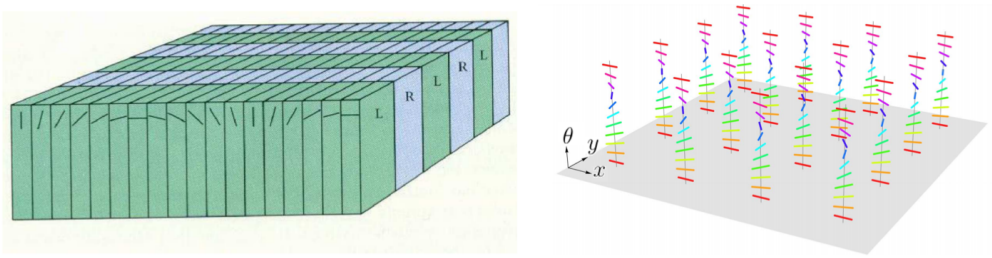


Figure 2.1.4: Receptive fields and profiles of LGN and simple cells. Simple cells arise from LGN cells. Images adapted from [DAOF95, Wan95, Pet08].

De Angelis et al. in [DAOF95] performed a mapping for the receptive fields and profiles of LGN and simple cells, and results are shown in Figure 2.1.4. For example, simple receptive fields arise from multiple isotropic LGN receptive fields by converging in a line [Hub95, Wan95].

2.1.2.3 Monocular organization

The primary visual area organizes its monocular neurons in an *hypercolumnar structure*, where columns correspond to parameters representing the cell's features such as orientation, ocular dominance, color, etc.



(a) (b)

Figure 2.1.5: (a) Hubel's "Ice Cube" model of the functional architecture of the visual cortex. Image adapted from [Hub95]. (b) Orientation columns.

This sensational discovery was made by Hubel and Wiesel, who showed that the preferred orientation of V1 neurons is roughly constant moving perpendicularly to the cortical surface [HW62]. These groups of neurons with similar orientation selectivity are called orientation columns. On the other hand, the preferred orientation varies gradually in the directions parallel to the surface, in a way that different columns are sensitive to different orientations, see Figure 2.1.5.

Moreover, V1 neurons alternate ocular dominance columns [LHW75], containing cells that are responsive only to input from the left or right eye at the center

of these bands (and binocular neurons at the boundaries), and there are also blobs [LH84], columns for a group of neurons sensitive to color.

All these columnar structures are organized in hypercolumns, a block of V1 containing both the ocular dominance and orientation columns for a particular region in the visual space, together with blob structure. This is the so-called classical "ice cube" model presented in [HW62], shown in Figure 2.1.5.

2.1.2.3.1 Orientation maps Orientation maps are the bi-dimensional implementation of the functional organization for orientation preference, and this is another example of the topographic organization besides retinotopy, introduced in Section 2.1.2.1. Since tangential penetration in the superficial layers of the cortex

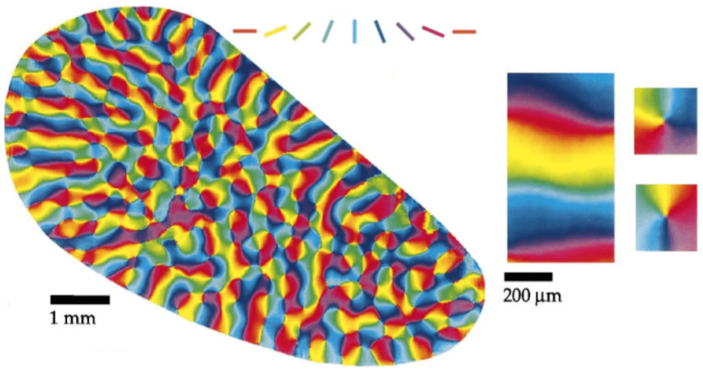
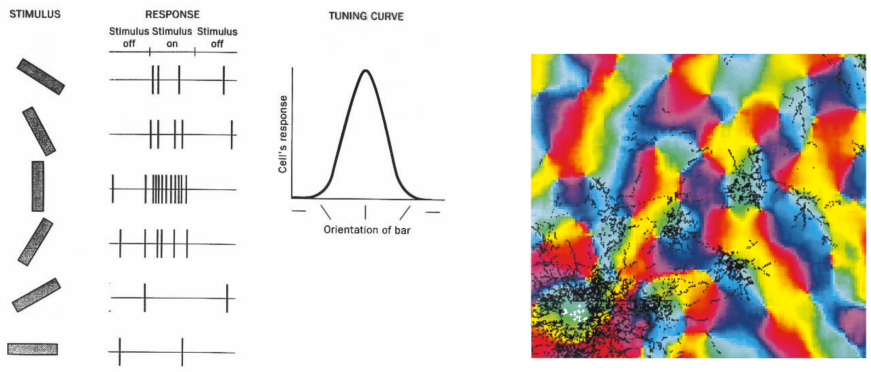


Figure 2.1.6: Different orientations are coded on a two-dimensional V1 map with different colors. On the right, there is an example of a smooth change of orientation and singular point, also called *pinwheel*. Image adapted from [BZSF97].

done by Hubel and Wiesel in [HW62, HW77] reveals that RPs of cells close to each other strongly overlap while the orientation preference varies smoothly, these maps present regions where the value changes smoothly. However, there are also points of singularity, with all orientations arranged around them, and for this reason, they are commonly called pinwheels, displayed in Figure 2.1.6. In particular, each pinwheel corresponds to an orientation column of Figure 2.1.5, and so orientation is engrafted onto the positional map [Hub95]. This two-dimensional reduction was presented by Bosking et al. (Fitzpatrick group) [BZSF97] using optical imaging techniques in which the cells' orientation preference is color-coded, see Figure 2.1.6.

2.1.2.3.2 Cortical connectivity To conclude our review of the functional architecture of monocular V1 cells, we discuss now the connectivity between neurons inside the hypercolumnar structure. Two types of communications play a central role: short and long-range connectivities.

The *short range connectivity* involves a local circuit operating at sub-millimeters dimensions consisting of a mixture of intracortical excitation and inhibition: this is why it is usually called local, short-range or intracortical connectivity. It has been suggested in [BC03] that such circuitry provides the response of cells to local features of visual stimuli. More precisely, the action of the intracortical circuitry is to select within the column the cell which gives the maximal response to a feature of the visual stimulus and suppresses all the others. The mechanism able to produce this selection is called *non-maximal suppression* or feature selection. These local interactions generate, for example, orientation and spatial frequency tuning curves, see image (a) of Figure 2.1.7.



(a) (b)

Figure 2.1.7: Examples of cortical connectivity. (a) Response of a cell to bars presented at various orientations. Orientation tuning curve comes out from the short-range connectivity. (b) Diffusion of a tracer of biocytin along horizontal connections. Image adapted from [BZSF97].

On the other hand, there exists another type of connectivity of the primary visual cortex, modulating interactions between different columns. Since this circuit operates over a range of several millimeters, it usually goes under the name of *long range* or *horizontal connectivity*. Neurophysiological experiments such as [BZSF97] and [GDI+96] reveal the existence of these connections along the cortical surface:

they correlate cells with the same orientation belonging to different hypercolumns, with non-overlapping receptive fields. This is shown in image (b) of Figure 2.1.7: the injection of a chemical tracer into a small area of the visual cortex propagates through lateral connections and this result is combined with the orientation map. It can be noticed that in the immediate neighborhood of each neuron, the connections are relatively isotropic, but over larger distances, they follow the orientation preferences, preferably by connecting neurons with co-oriented aligned receptive profiles.

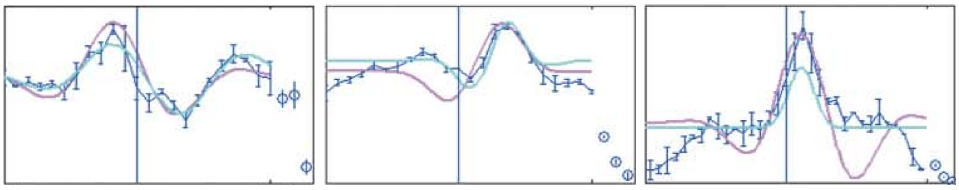
2.1.2.4 Binocular organization

A generally accepted theory is that binocularly driven neurons in the primary visual cortex form the first stage for the processing of binocular disparity ([PNB68, BBP67, PF77, SK17]). Indeed, cells at lower levels in the visual pathway, such as LGN cells or monocular simple cells, are not driven by stimulation through both eyes, as remarked by Hubel and Wiesel in [HW62]. They account for disparity-tuned neurons in a brief report in [HW70] and they observe single units driven from both eyes from which it is possible to determine separate left and right RFs.

The architecture and the neural connections of the visual cortex underlying binocular vision have recently been studied in [PSK16], while binocular interactions characterized by binocular and monocular synaptic inputs have been investigated in [STR+22]. In general, binocular domains form a separate set of connections in V1, suggesting (if there exists) a distinct columnar system beside the monocular one, [MAHG93]. Anyways, in V1 the binocularly-driven cells tend to be closer to the ocular dominance boundaries, while the monocular cells are toward the centers [LHW75].

Moreover, as already said, binocular cells in V1 encode horizontal positional disparity, usually referred to as disparity, and orientation features, directly linked with the monocular orientation preferences. We will investigate these properties in the next paragraphs.

2.1.2.4.1 Disparity tuning cells The positional disparity is encoded by binocular cells through monocular receptive fields, in two different ways: considering the difference between receptive profile center locations (position-shift model) of the two eyes, or by differences in receptive profile organization (phase-shift model). Binocular cells in V1 respond to their preferred disparity value, generating disparity tuning curves, as shown in Figure 2.1.8.



(a) Near (b) Far (c) Zero

Figure 2.1.8: Disparity tuning curves for different types of cells. The horizontal axis represents the disparity while the vertical axis represents the response of the cell to the value of disparity. Image (a) involves a cell responding to negative values of disparity, recognizing objects located in front of the plane of fixation. Image (b) represents cells responding to positive values of disparity, identifying objects located behind the plane of fixation. Image (c) displays cells responding to zero values of disparity, describing objects on the plane of fixation. Images adapted from [TCL03].

.

The tuning curves presented in Figure 2.1.8 involve different disparity-tuned cells. Poggio and colleagues in [PF77, PGK88, Pog95] have classified them into several categories. We recall here three main groups:

- *tuned excitatory neurons* displayed binocular facilitation to stimuli over a narrow range of small disparities, including zero disparity. Since zero disparity is a characterization of the fixation point, they respond best to objects that are located within the plane of fixation, namely when the image has zero disparity.

- *far neurons* are inhibited by stimuli in front of the fixation point and activated by stimuli behind the fixation point, so they respond optimally to objects behind the plane of fixation.

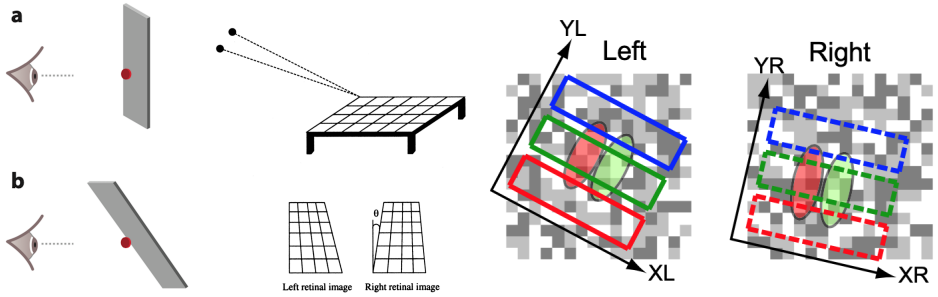
- *near neurons* show a reciprocal disparity sensitivity with respect to far neurons. Indeed, they respond best to objects located in front of the plane of fixation.

2.1.2.4.2 Interocular orientation differences from monocular outputs

The standard neurophysological technique applied for the measurement of binocular profiles assumes the mono dimensionality of the monocular RPs, ignoring information about the orientation selectivity of monocular simple cells. This assumption is made because the response of the cell should change quickly when

disparities are measured orthogonally to the RP's orientation, as the structure of the monocular RPs changes most rapidly in that direction [CD01]. Differences between left and right structures which only involve the (absolute) locations of the RPs are usually called absolute (horizontal) positional disparity, or *zero order disparity*.

However, tilt and slant of objects in space are coded in the concept of interocular orientation differences, namely differences between left and right monocular preferred orientations. This difference in orientation generates when stimuli are not located on fronto-parallel planes, introducing the concept of *first order disparity*, involving the difference in orientation of the two left and right retinal projections.



(a) (b) (c)

Figure 2.1.9: (a) Top image illustrates zero-order stimulus involving the frontoparallel plane. The bottom image shows a first-order stimulus on a surface lying in depth. Image adapted from [Orb11]. (b) Orientations differences of two projections. Image adapted from [BCP01]. (c) Differences in orientations receptive fields. Image adapted from [STO10].

The existence of different preferred monocular orientations as inputs of a binocular cell, confirmed by neural works such as [STO10, BCP01, BFM72, BC01, STR+22], has led to the question of whether neurons characterized by this interocular difference are selective for orientation disparities, and hence forming another neural mechanism for depth perception, [BFM72], or not. Until now, no one has found evidence for this mechanism, see for example [BC01], attributing the role of physiological variables to the couple of monocular orientations.

2.1.2.4.3 Long range connectivity for binocular cells Equivalently to the monocular case, also for binocular cells it is possible to speak of long-range horizontal connections. The existence of lateral connections which spread horizontally across the cortex was detected in [MAHG93] using biocytin injection targeted at

a binocular site, see Figure 2.1.10. They observed that the interactions only propagate between binocular regions, and the distribution avoids the centers of ocular dominance columns.

Synaptic interactions between binocular neurons have recently been studied also in [STR+22], discovering that binocular synapses varied in degree of the response correlation between left and right visual stimulation. Binocular congruent input consists of interocular orientation preference difference up to ≈ 19 degrees, and they are highly orientation selective, maintaining the aligned orientation preference, similarly to the monocular case.

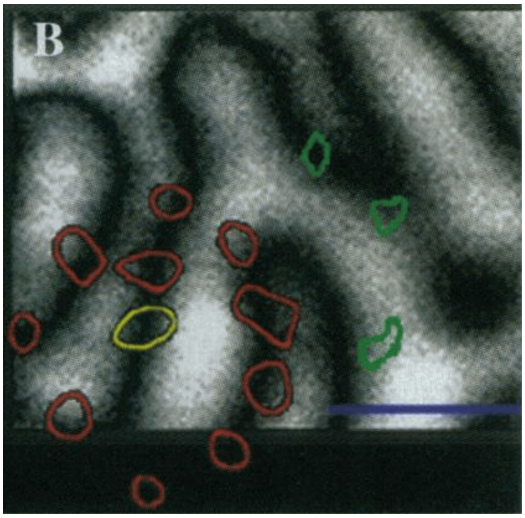


Figure 2.1.10: A biocytin injection superimposed on a map of ocular dominance columns, image result from the work in [MAHG93]. Binocular zones are coded black, while monocular zones are coded white. The injection site (yellow circle) was centered on a binocular zone. The patches' propagation (red corresponds to dense while green to sparsely labeled) tended to avoid highly monocular sites and were located in binocular zones.

2.1.3 Beyond V1

Physiological and psychophysical neural evidence [SK17, Hub95] states that V1 is just the beginning of the visual processing. Although it is not really clear how the visual pathway continues, the second major area involved in processing the visual signal is V2, the secondary visual cortex (also known as Brodmann area 18, or prestriate cortex).

V2 receives strong feedforward input from and gives feedback to V1, with

which it shares many properties: cells are tuned to simple characteristics such as orientation, spatial frequency, and color but also cells' responses are modulated by more complex properties such as orientation of illusory contours [VdHPB84, APVE07], and of course, binocular disparity [VDHZF00].

Moreover, V2 sends strong connections to the upper layers of the visual cortex. The signal travels in multiple areas creating functional pathways that feed visual information from V1 into associational cortical areas, where the processing of visual information continues. These areas are situated in the parietal and temporal lobes, and following the "two-streams hypothesis", they give rise to two different pathways, [SK17]:

- *dorsal (lateral parietal) pathway*: it is also known as the "where pathway" and it is involved with processing the object's spatial location relative to the viewer through the environment.

- *ventral (inferior temporal) pathway*: it is also known as the "what pathway" and it is responsible for a high-resolution form of vision, meaning that it is involved with objects and visual identification and recognition, and color processing.

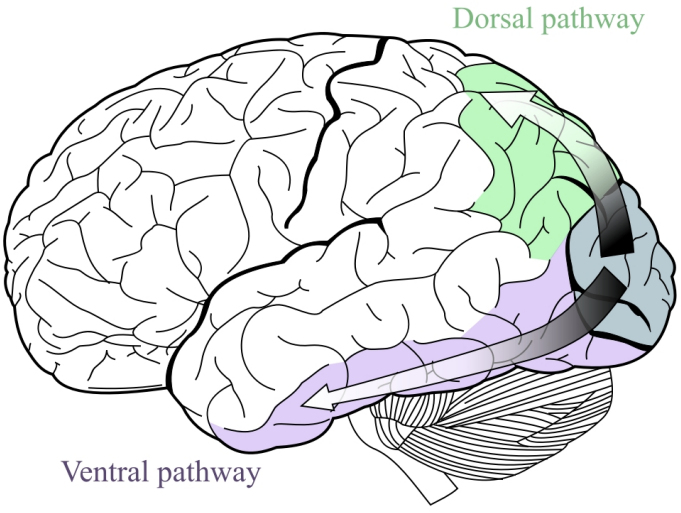


Figure 2.1.11: Two streams hypothesis. Green: dorsal pathway. Violet: ventral pathway. Image adapted from Wikipedia.

These two streams are supposed to communicate with each other, to ensure that vision comes together. In fact, the only sure thing is that V1 and V2 are the

last areas to have all the visual information. And so, conscious vision happens when simultaneous or approximately simultaneous activations across all visual areas happen, [SK17].

What appears to be really important is the existence of feedback loops that stand in the brain. At every level, there is a return of input from higher levels of the visual pathways to earlier ones. Forward and backward connections are thought to play a role in synchronizing the responses in the brain so that our perceptual experiences are whole objects and not just fragmented parts. Thus, our experience of vision is distributed across the cortex and constant feedback loops from higher to lower levels integrate our perception.

2.2 Perceptual phenomena

Visual perception is the result of a series of complex mechanisms carried out by the visual system, which mediate between the acquisition of the real stimulus and its interpretation, to obtain a coherent description of the visual scene. This process takes place through a perceptual organization, whose principles were first expressed by Gestalt psychologists on a phenomenological level, using only subjective experience without any neurophysiological evidence.

Many psychological experiments suggest that the principles underlying Gestalt laws are implemented within the mammalian primary visual cortex [BZSF97].

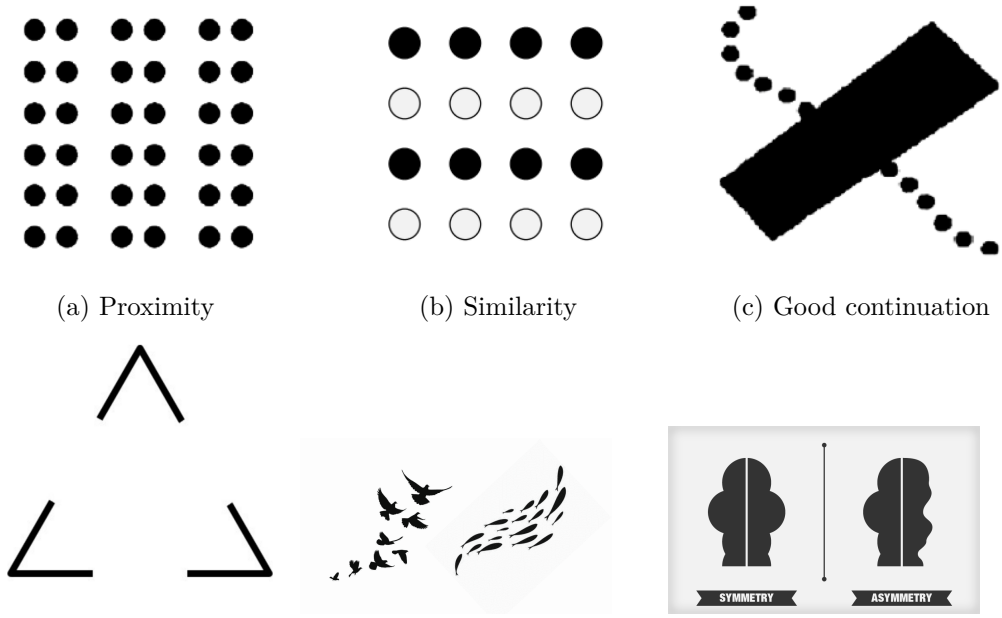
2.2.1 Gestalt theory

The first studies concerning the perceptual organization of the visual scene were conducted in the 1920s by German psychologists Max Wertheimer, Wolfgang Köhler, and Kurt Koffka, who founded the school of thought known as Gestalt psychology. The classical references for these topics are [Kof35, Koh29, Wer23] and it is also possible to refer to [WEK+12] for a more recent review. The word *Gestalt* in German means "organized whole" and refers to the general concept that the perceived stimulus is something more than the simple sum of all its parts.

The purpose of Gestalt theory is to understand the laws behind the human mind's ability to construct meaningful perceptions from disorganized real stimuli. A pattern for self-organization is sought that can combine smaller elements to form larger objects.

The Gestalt approach is based on the development of a set of principles known

as the *laws of perceptual organization*, which can be summarized by the following points, visually represented in Figure 2.2.1:



(a) Proximity (b) Similarity (c) Good continuation

(d) Closure (e) Common fate (f) Simmetry

Figure 2.2.1: Visual representation of the laws of perceptual organization. Images adapted from [Kan79].

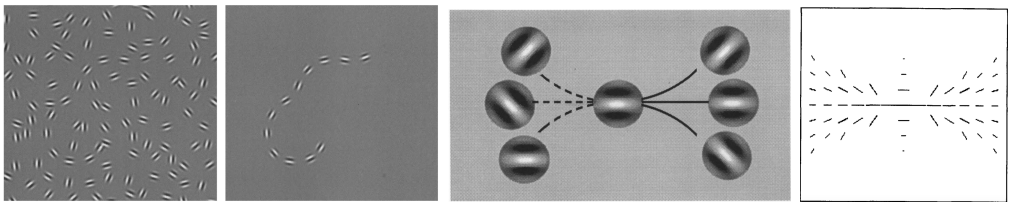
- *Law of proximity:* elements that are close to each other and apart enough from the rest of the elements form a group or a cluster.
- *Law of similarity:* elements similar in color, texture, shape, or orientation are grouped together.
- *Law of good continuation:* elements aligned (or with comparable alignment) and which tend to form continuous curves are grouped together.
- *Law of closure:* the perception of forms and figures in their complete appearance happens despite the absence of some of their parts.
- *Law of common fate:* pieces which are subjected to the same spatial or temporal change are grouped together.
- *Law of symmetry:* spatially symmetric pieces are more likely to be perceived together rather than non-symmetric ones.

2.2.2 Association field and good continuation

One of the most studied laws of perceptual organization, which plays a fundamental role within the problem of perceptual grouping, is the law of good continuation, quantitatively studied in many works [FHH93, GM91, GM85, PZ89, KS91, SK94].

In [FHH93] Field et al. investigated through psychophysical experiments how the visual system encodes the good continuation of the contour. In doing so, they introduced the notion of associative fields, thus formalizing the Gestalt principle in terms of the position and mutual orientation of elements or patches in two-dimensional space.

During the experiment, an observer was shown an image composed of aligned Gabor patches (oriented elements in space), which form a path, against the background of a large number of randomly distributed patches, image (a) of Figure 2.2.2. The purpose of the experiment was to test the subject's ability to detect the perceptual units present in the visual stimulus.



(a) (b) (c)

Figure 2.2.2: Pshychophysical introduction of the association field. (a) Field-Hayes and Hess experiment. Image adapted from [FHH93](b) Pattern describing the association field. Image adapted from [FHH93](c) Ladder effect. Image adapted from [YF98].

The results of this experiment can be summarized in the concept of an association field. This is the description of a pattern that allows determining the stimuli that can be associated with the same perceptual unit, representing the elements of the stimulus that are associated with a central point. In particular, correlations between elements are stronger between segments that are collinear or co-circular. In fact, the solid lines indicate the elements that can be connected with the central component, while the dotted lines indicate unrelated elements, image (b) of Figure 2.2.2.

An interesting result that comes out of these psychophysical experiments concerns the correlation between the association fields and the horizontal connections of the monocular cells of V1. In this sense, lateral connectivity is seen as a potential

implementation of the perceptual phenomenon of the law of good continuation.

Moreover, while the classical concept of association fields involves only coaxial-oriented elements, a similar correlation is also observed between parallel segments, or more precisely orthogonal with respect to the central-oriented element. This effect is usually called *ladder effect* [MC82, FHH93, YF98], see image (c) of Figure 2.2.2.

Historically, good continuation in depth is much less well developed than good continuation in the plane. Despite that, the origins of the problem are quite ancient: in [Kof63], Koffka noticed that the perception of a perspective drawing is strengthened when viewed with stereoscopic vision, compared to the monocular case, observing the introduction of mutual reinforcement rather than a conflict of forces. One of the objectives of this thesis is to realize a mutual reinforcement model, developing good continuation in depth analogously to the models of contour organization in two dimensions. Psychophysical investigations such as [DW14, DW15, KHK16, Utt83] suggest that this should be feasible. Indeed, Hess and colleagues in [HF95, HHK97] extend the good regularity in 3D, a property that is in accordance with the more recent works of Deas-Wilcox and Khuu-Honson-Kim in [DW15, KHK16]. We will use a more mathematical approach, in agreement with the study of 3D-circularity applied to crossing-preserving enhancement DW-MRI introduced in [MOW+17, PFS+15, RDRV10], or for mapping connectivity in diffusion MRI [MS09, MS12]. We will study this in more detail in Chapter 5 and Chapter 6.

Chapter 3

The stereo problem

In this chapter, we introduce a brief review of classical geometry and reference models for the stereo problem. First, we recall the geometrical properties of perspective projections which are the basis of stereo triangulation. Then, we concentrate on models proposed for describing the stereo vision. At the beginning we examine the neural model for the integration of monocular outputs performed by binocular cells, then we study mathematical models for the stereo correspondence problem, illustrating typical constraints and finally presenting the differential model of reference.

Contents

3.1 Classical stereo geometry

It is generally accepted that the first scientist to give mathematical attention to 3D vision is the greek mathematician Euclid around 300 BC, in his work Optics, which deals almost entirely with the geometry of vision. These preliminary notions laid the foundation for the development of the configuration of the stereo vision problem over the centuries, and a complete historical background can be found in [How12, HR95]. Nowadays, extensive textbooks on the subject have been published, as for example [Fau93, FL01, HZ03, SR11]. In this subsection we briefly recall this geometrical configuration, to define the variables that we will use, mainly referring to [Fau93, Ch. 6].

3.1.1 Stereo variables

Consider the global reference system (O, i, j, k) in $\mathbb{R}^3$, with $O = (0, 0, 0)$ the origin and with coordinates (r_1, r_2, r_3). We introduce the optical centers $C_L = (-c, 0, 0)$ and $C_R = (c, 0, 0)$, with c real positive element, and we define two reference systems: (C_L, i_L, j_L), (C_R, i_R, j_R), the reference systems of the retinal planes $\mathcal{R}_L$ and $\mathcal{R}_R$ with coordinates respectively (x_L, y), (x_R, y). In the global system we suppose the retinal planes to be parallel and to have equation $r_3 = f$, with f denoting the focal length. This geometrical set-up is shown in Figure 3.1.1.

Remark 3.1.1. *A point $Q = (r_1, r_2, r_3)$ in $\mathbb{R}^3$ projects in the two retinal planes via perspective projection through optical centers C_L and C_R. This computation defines two projective maps π_L and π_R respectively for the left and right retinal planes:*

$$\begin{array}{ccccc} \pi_L: & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 & \\ & \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} & \mapsto & \begin{pmatrix} \frac{f(r_1+c)}{r_3} \\ \frac{fr_2}{r_3} \end{pmatrix}, & \end{array} \qquad \begin{array}{ccccc} \pi_R: & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 & \\ & \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} & \mapsto & \begin{pmatrix} \frac{f(r_1-c)}{r_3} \\ \frac{fr_2}{r_3} \end{pmatrix}. & \end{array} \tag{3.1}$$

We call the pair of points Q_L and Q_R projections of Q as **corresponding points**, since this definition is usually referred to points on the retinae of the two eyes which when simultaneously stimulated normally produce a single visual impression.

Proof. A point on the left retinal plane of local coordinates (x_L, y) has global

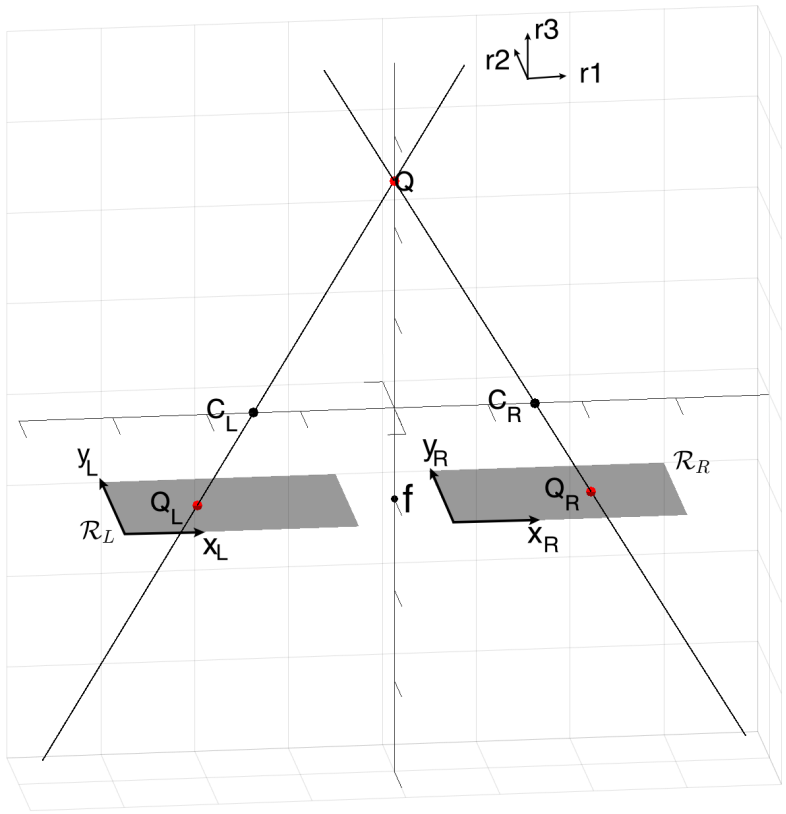


Figure 3.1.1: Stereo triangulation. A three-dimensional point Q projects into the left and right parallel and aligned retinal planes $\mathcal{R}_L$, $\mathcal{R}_R$ through optical centers C_L and C_R. The bidimensional projections on the retinal planes characterized by equation $r_3 = f$ are called Q_L and Q_R.

coordinates $Q_L = (-c + x_L, y, f)$, and it corresponds to a point $Q = (r_1, r_2, r_3)$ in the Euclidean $\mathbb{R}^3$ such that C_L, Q_L and Q are aligned. This means that the vectors $Q_L - C_L = (x_L, y, f)$ and $Q - C_L = (r_1 + c, r_2, r_3)$ are parallel, obtaining the following relationships:

$$x_L = f\frac{r_1 + c}{r_3}, \quad y = f\frac{r_2}{r_3}. \tag{3.2}$$

Analogously, considering Q_R and C_R, we get:

$$x_R = f\frac{r_1 - c}{r_3}, \quad y = f\frac{r_2}{r_3}. \tag{3.3}$$

□

In a standard way, the **horizontal disparity** is defined as the differences between left and right retinal abscissa coordinates

$$d := \frac{x_L - x_R}{2}, \tag{3.4}$$

up to a scalar factor. Moreover, it is also possible to define the coordinate x as the average of the two retinal coordinates $x := \frac{x_L + x_R}{2}$, leading to the following change of variables:

$$\begin{cases} x = \frac{fr_1}{r_3} \\ y = \frac{fr_2}{r_3} \\ d = \frac{fc}{r_3} \end{cases} \longleftrightarrow \begin{cases} r_1 = \frac{xc}{d} \\ r_2 = \frac{yc}{d} \\ r_3 = \frac{fc}{d} \end{cases}, \tag{3.5}$$

where the set of coordinates (x, y, d) is known as **cyclopean coordinates**, because, in a sense, we perceive the world with a single eye in the middle of the head. For further details see [Jul71].

3.1.2 Tangent estimation

Corresponding points in the retinal planes allow to project back in $\mathbb{R}^3$. An analogous reasoning can be done for the tangent structure: starting from tangent vectors of corresponding curves in the retinal planes, it is possible to project them back and recover the 3D tangent vector. We recall here this result, while for a detailed explanation the reader can consult [Fau93].

Remark 3.1.2. *Consider γ_L and γ_R corresponding left and right retinal curves, perspective projections of a curve $\gamma \in \mathbb{R}^3$. Knowing the left and right retinal tangent structures, then it is possible to recover the direction of the tangent vector $\dot{\gamma}$.*

Proof. Starting from a curve $\gamma \in \mathbb{R}^3$, we can project it in the two retinal planes obtaining $\gamma_L = \pi_L(\gamma)$ and $\gamma_R = \pi_R(\gamma)$ from equations (3.1). The retinal tangent vectors are obtained through the differential $\dot{\gamma}_{L,R}(t) = D\pi_{\gamma(t)}\dot{\gamma}(t)$:

$$\dot{\gamma}_R(t) = \begin{pmatrix} \frac{f(\gamma_3\dot{\gamma}_1 + (c-\gamma_1)\dot{\gamma}_3)}{\gamma_3(t)^2} \\ \frac{f(\gamma_3\dot{\gamma}_2 - \gamma_2\dot{\gamma}_3)}{\gamma_3^2} \end{pmatrix}, \dot{\gamma}_L(t) = \begin{pmatrix} \frac{f(\gamma_3\dot{\gamma}_1 - (c+\gamma_1)\dot{\gamma}_3)}{\gamma_3(t)^2} \\ \frac{f(\gamma_3\dot{\gamma}_2 - \gamma_2\dot{\gamma}_3)}{\gamma_3^2} \end{pmatrix}. \tag{3.6}$$

Extending the tangent vectors and the points into $\mathbb{R}^3$, we get $\tilde{t}_L = (\dot{\gamma}_{L1}, \dot{\gamma}_{L2}, 0)^T$, and $\tilde{m}_L = (\gamma_{L1}, \gamma_{L2}, f)^T$, and $U_{t_L} = (P_L)^{-1}\tilde{m}_L \times (P_L^{-1})\tilde{t}_L$, with $P_L = \begin{pmatrix} 1 & 0 & -c/f \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ projection matrix. The same reasoning holds for the right structure, with projection matrix $P_R = \begin{pmatrix} 1 & 0 & c/f \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Then $U_{t_R} \times U_{t_L}$ is a vector parallel to the tangent vector $\dot{\gamma}$:

$$\begin{aligned} U_{t_R} \times U_{t_L} &= \left(\underbrace{\frac{f^4 2c(\dot{\gamma}_2\gamma_3 - \dot{\gamma}_3\gamma_2)}{\gamma_3^4}}_{\lambda(t)} \dot{\gamma}_1, \frac{f^4 2c(\dot{\gamma}_2\gamma_3 - \dot{\gamma}_3\gamma_2)}{\gamma_3^4}\dot{\gamma}_2, \frac{f^4 2c(\dot{\gamma}_2\gamma_3 - \dot{\gamma}_3\gamma_2)}{\gamma_3^4}\dot{\gamma}_3 \right) \\ &= \lambda(t)\left(\dot{\gamma}_1(t), \dot{\gamma}_2(t), \dot{\gamma}_3(t)\right) \\ &= \lambda(t)\dot{\gamma}(t). \end{aligned} \tag{3.7}$$

□

3.2 Neural model for stereo vision

In this section, we recall the classical model used to explain the behavior of binocular neurons in V1. We will focus in particular on binocular simple cells.

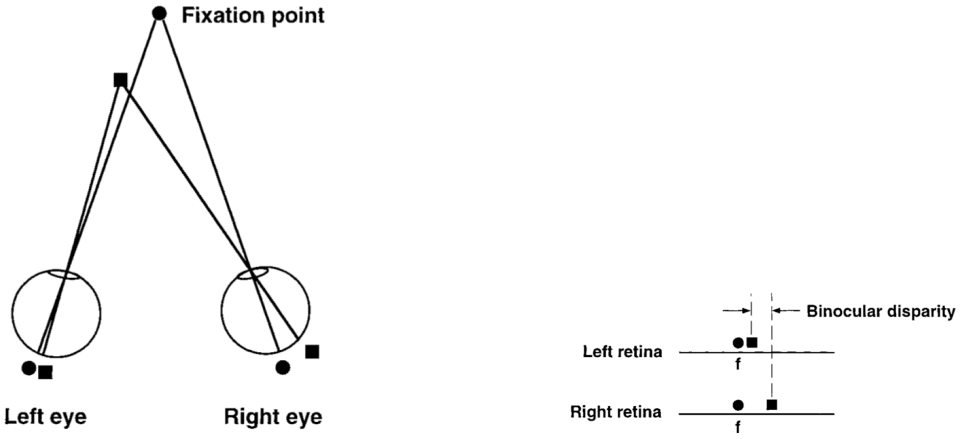
3.2.1 Binocular energy model (BEM)

The typical model, based on physiological experiments, which describes the output of binocular cells is the *binocular energy model (BEM)*, first introduced by Anzai, Ohzawa and Freeman in [AOF99b]. This model encodes disparities through the receptive profiles of simple cells, leading to the definition of position and phase disparities [JR15].

3.2.1.1 Position and phase disparities

The characterization of *position disparity* is based on the idea that the binocular visual system infers different positions in the retinae. The encoding mechanism of this phenomenon is developed through position differences between the centers of right and left eye receptive profiles, of which a graphic representation can be seen

in Figure 3.2.1. Formally, this definition coincides with the definition of horizontal disparity introduced in (3.4).



(a) (b)

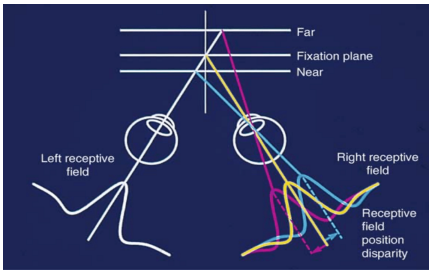
Figure 3.2.1: Positional disparity. (a) Geometry of binocular projection. The fixation point projects to the two corresponding foveas and has zero disparity by definition. Other points of the visual scene project in different positions. (b) Position disparity definition on the retinae. Circle and square points represent the centers of the RPs. Images adapted from [Qia97].

The notion of *phase disparity* involves different shapes of left and right receptive profiles while maintaining constant receptive profile positions. Denoting with ϕ the phase of the RP, fitted as usual with a Gabor function (we will see this with more details in Chapter 4, subsection 4.4.2), it is possible to define the phase disparity as:

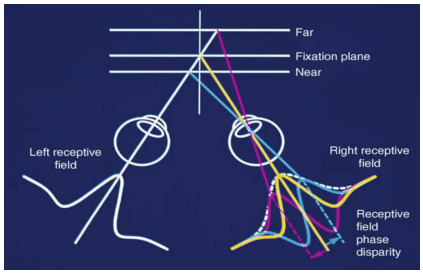
$$d_P = \phi_R - \phi_L \text{ (deg PA)} \qquad d_P = -(\frac{\phi_R}{360 f_R} - \frac{\phi_L}{360 f_L}) \text{ (deg VA)} \tag{3.8}$$

with PA=phase angle in the retina, and VA= visual angle in the 3D world, f_R, and f_L frequency of the right and left RPs.

Read and Cumming in [RC07], building upon [AOF99a], proposed that phase disparity neurons tend to be strongly activated by *false correspondence pairs*, namely pairs of points in projection on the retinae that do not correspond to a three-dimensional percept. Therefore the most relevant disparity in the receptive fields is the position alone.



(a)



(b)

Figure 3.2.2: Behavior of RPs in the definition of position and phase disparity. (a) Position disparity. (b) Phase disparity. Images adapted from [JR15]

3.2.1.2 Description of the model

The BEM model, introduced at the beginning of the 90s in [AOF99b], involves the behavior of binocular simple cells. The basic idea behind it, it is a binocular neuron with a receptive field in each eye. The output of each eye depends on the inner product of each eye's image with the corresponding receptive field: for example for the left eye the output is

$$O_L = \int_{\mathcal{R}} \varphi_L(x,y) I_L(x,y) dxdy, \tag{3.9}$$

where the function I_L represents the (left) retinal image and φ_L represents the (left) classical receptive profile. More details concerning the monocular case are given in Chapter 4, subsection 4.4.2.

More precisely, the energy model assumes that monocular left and right outputs are summed linearly; if the result is negative, the neuron is inhibited and does not fire, while if positive, the neuron fires at a rate proportional to the square of the inputs:

$$Pos(O_L + O_R)^2, \tag{3.10}$$

where $Pos(x) := max\{x, 0\}$.

A relationship exists between the binocular interaction RP and monocular RPs, and it can be explained observing that, if $O_L + O_R > 0$ then the output of the binocular simple cell is the following:

$$O_L^2 + O_R^2 + 2O_L O_R.$$

The first two terms represent responses due to monocular stimulation while the third term $O_B := O_L O_R$ is the one that depends on binocular disparity, so it is called *binocular interaction* term. In this sense, although the model is mainly linear, the binocular interaction introduces a non-linear factor (multiplicative model).

The binocular receptive profile's computation of the interaction part is performed in [AOF99b] and it is in accordance with the prediction of the BEM, as displayed in Figure 3.2.3.

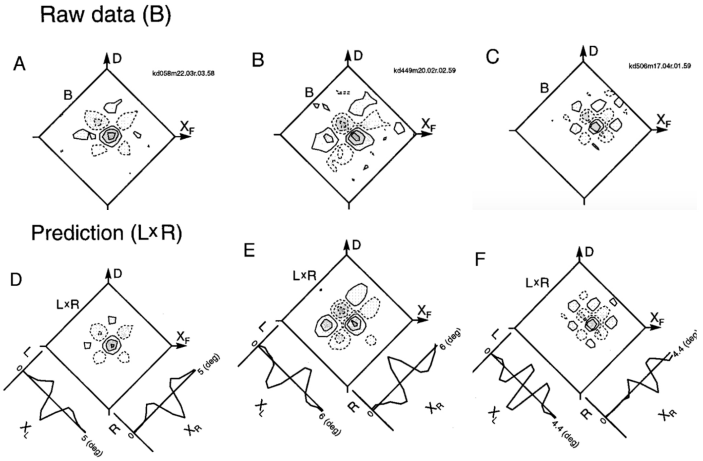


Figure 3.2.3: Comparisons between binocular interaction RPs and the product of left and right eye RPs. First row: binocular interaction RPs of 3 cells (A-C) are shown on top (Raw data). Second raw: contour plots for the product of left and right eye RPs (LR) are shown for each cell (D-F) on the right (Prediction) along with 1-dimensional (1D) profiles of the left (L) and right (R) eye RPs. Contour plots for the prediction are scaled so that each has the same peak as that of the corresponding plot for the raw data. Predictions are qualitatively very similar to the raw data. Image adapted from [AOF99b].

3.3 Mathematical model for stereo correspondence

Three-dimensional vision involves the use of two eyes (or at least two cameras) to encode visual scene information. This information is projected into the aligned retinal planes starting from parallel optical centers. Going from the 3D world to the retinae is a simple process, characterized by central projections; on the other hand, the 3D reconstruction problem requires to solve the stereo matching problem.

The stereo matching problem, also known as **stereo correspondence problem**, refers to the problem of ascertaining which parts of one image correspond to which parts of the other image. Indeed, the horizontal separation of the eyes (different points of view) causes the two retinal images to be slightly different and so the reconstruction forces to find the set of points in one image which can be identified as the same points in the other image. This is usually done by considering points or features in one image and matching them with the points or features in the other image, establishing corresponding points or corresponding features.

Many works in literature have addressed the stereo problem using different viewpoints. Here we recall the work of Vaillant and Faugeras [VF92] in relation to extremal boundaries for 3D curved objects, and, very important from our perspective, the works of Zucker et al. [AZ00, LZ06, LZ03] which deal with curves stereo geometry using tools of differential geometry.

In the next subsection, we first recall the classical constraints imposed to facilitate the coupling problem, and then we introduce the model proposed by Zucker et al. [AZ00, LZ06, LZ03] to solve the correspondence.

3.3.1 Constraints on the coupling problem

Modeling stereo vision has made significant progress in the field of computer vision in the last 20 years and surveys for this literature review can be found for example in [KK16, HI16].

The central point is the introduction of algorithms able to find a disparity map that solves the matching. However, none of the proposed algorithms has optimal performance on all occasions and this happens because they often conflict with the presence of constraints, which can be inserted or naturally emerge from the geometry of stereo images.

The classical constraints that often appear are:

- *Epipolar constraint*: the imposition of the ordinate coordinate of the two images to be the same. It is always possible to return in this case by applying the rectification process. For more details on this topic see [Fau93].

- *Uniqueness*: the requirement that a feature in one stereo half-image has to be matched to, at most, one similar feature in the other half-image.

- *Continuity (or smoothness)*: the disparity of the matches should vary smoothly almost everywhere over the image.

- *Ordering*: the ordering of features is preserved across images.

- *The disparity gradient*: the disparity gradient is defined as the derivative of the disparity with respect to the Cyclopean coordinate x, namely $\partial d/\partial x$. A disparity gradient limit requires that the absolute value of the disparity gradient never exceeds a given value.

- *Geometric constraints*: punctual constraints, arising from the geometric relationships underlying the problem embedded in the Euclidean three-dimensional space. The constraints imposed by the projection are described in Section 3.1.

In general, several models have been proposed for vision problems which incorporate different types of constraints: for instance, we cite the model introduced by Marr and Poggio in [MP79] to describe human stereo vision, integrating the constraints of uniqueness and continuity. They build it considering orientational differences in the two eyes to compute disparity. Or, the work of Faugeras and Robert [FR96], which relates information concerning two images with a third one, based on constraints emerging from epipolar geometry.

At the beginning of the $21st$ century, Zucker et al. [AZ00, LZ03, Zuc14] employed position, orientations, and curvature in 2D retinal planes to relate the 2D Frenet differential structure in the left and right retinal planes with the Frenet geometry of the 3D space, integrating geometric constraints. We investigate better this model in the following subsection.

3.3.2 The Zucker model

The model introduced by Zucker, Alibhai and Li for stereo vision in [AZ00, LZ03, LZ06], aims at generalizing the fundamental concept of co-circularity in three dimensional space. In particular, co-circularity in $\mathbb{R}^2$ has been firstly developed by Parent and Zucker in [PZ89] observing that a bi-dimensional curve γ can be locally approximated at 0 via the osculating circle and so its generalization involves the Frenet differential geometry of a three-dimensional curve.

While in the two-dimensional case the approximation of the curve using the Frenet 2D basis causes the curvature to appear in the coefficient of the Taylor series development ($1st$ order), in the three-dimensional case the coefficients involve both the curvature and torsion. So, in [AZ00] the authors propose heuristically to generalize the osculating circle for space curves with an osculating helix, with a

preference for r_3-helices (helices whose spiral develops along the r_3-axis) to provide stability in terms of camera calibration. In this way the orientation disparity is encoded in the behavior of the helix in the 3D space: there is no difference in orientation in the retinal planes if the helix is confined to be in the fronto-parallel plane (the helix in this plane becomes a circle and the torsion of the curve is equal to zero), otherwise as the magnitude of the torsion increases, moving along the 3D curves, the retinal projections have different orientations.

In [LZ03, LZ06] they observe that with the introduction of the curvature variable as a feature in the two monocular structures, assuming having corresponding points, it is possible to reconstruct the 3D Frenet geometry of the curve, starting from the two-dimensional Frenet geometry, with the exception of the torsion parameter.

The success of this model is due to the simultaneous use of the notions of disparity of position and orientation, which are naturally tied together by differential geometry, as we will see in the next paragraphs.

3.3.2.1 Geometry of the model

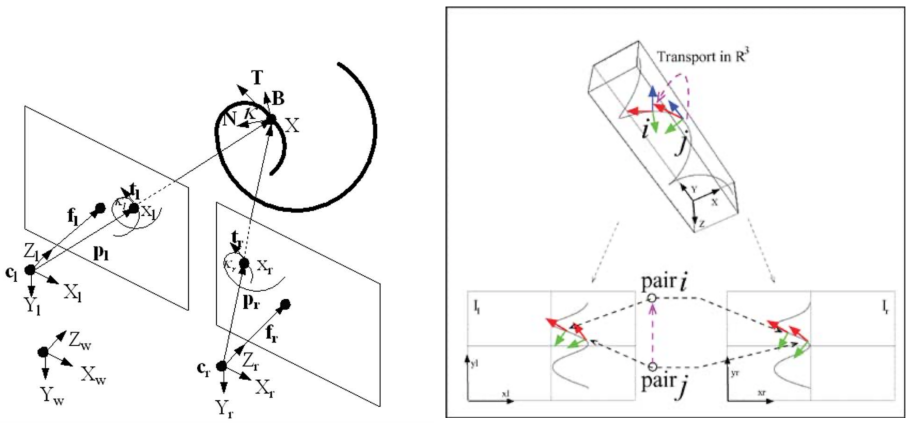
In [LZ03, LZ06] the authors choose in the external and in the retinal space coordinate systems with the same orientation up to a translation. They suppose to know for every (left and right) retinal image its position, tangent, and curvature (x, y, θ, k), and they prove the following Proposition, of which a graphic representation can be found in Figure 3.3.1.

Proposition 3.3.1. *Given two perspective views of a 3D space curve with full calibration, the normal N and curvature k at a curve space point are uniquely determined from the positions, tangents, and curvatures of its projections in two images. Thus the Frenet frame $\{T, N, B\}$ and curvature k at the space point can be uniquely determined.*

Hence, using the knowledge of the Frenet basis together with the fundamental addition of the curvature variable, they used the concept of *transport* to build a stereo correspondence algorithm.

3.3.2.2 Transport in $\mathbb{R}^3$

The proposed idea is to move the 3D Frenet frame consistently with the 2D Frenet structure of the left and right retinal planes, to match the stereo correspondence



(a) (b)

Figure 3.3.1: (a) Geometrical setup of Proposition 3.3.1. Given points $(x_L, y_L, \theta_L, k_L)$ and $(x_R, y_R, \theta_R, k_R)$ it is possible to determine the position, the Frenet frame $\{T, N, B\}$ and the curvature k but not the torsion τ. (b) Transport of the Frenet frames in $\mathbb{R}^3$ helps find consistent left-right image tangent pairs. Images adapted from [LZ06].

between pairs of (left and right) pairs of tangents. More specifically, they compare the measurements obtained from transporting the differential structure with the information on the measured point. This is shown in Figure 3.3.1, image (a). Since the osculating plane is at disposition, thanks to Proposition 3.3.1, it is possible to define a distance on the osculating plane and, for everything to remain compatible, a minimum torsion constraint is (heuristically) imposed.

Thus, the notion of transport translates into the definition of a distance, or rather a probability measure, which tells how likely it is that a pair of points is made up of corresponding elements and it is not a false match. This goes under the name of *compatibility*.

More precisely, the compatibility from a point i to a point j is defined starting from the reconstructed tangents T_i and T_j, considering the projection of T_i^* of T_i in the osculating plane at point j. They call s the transport distance from T_i^* to T_j on the osculating plane , c the projection distance from T_i to T_i^* and $\Delta\phi$ the difference of rotation between T_i^* and T_j in the osculating plane, and define the compatibility matrix:

$$R = [R_{ij}] = \mathcal{G}_{\sigma_t}(s)\mathcal{G}_{\sigma_c}(c)\cos\Delta\phi, \tag{3.11}$$

with $\mathcal{G}_\sigma$ Gaussian kernel of parameter σ.

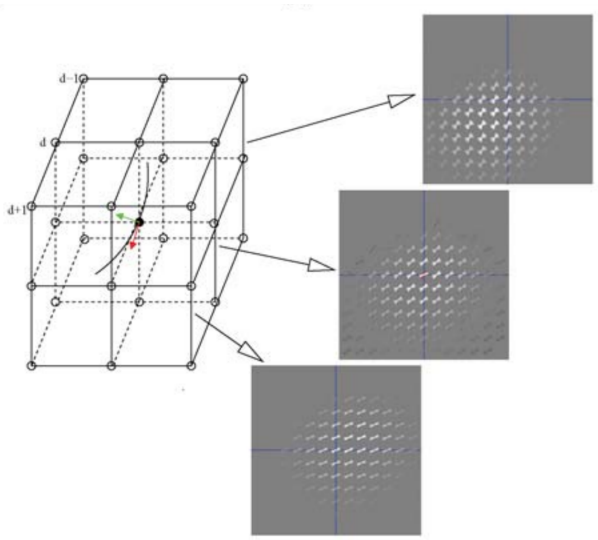


Figure 3.3.2: Sample compatibility fields around j. Brightness encodes the positive support. Source [LZ06].

An intuitive demonstration of compatibility field is shown in Figure 3.3.2. The image shows slices of compatibility R_{ij} around the point j by displaying the degree of support a neighboring node i receives from j. From these few slices, it is possible to see that both the position and the orientation disparities affect the compatibility and the emergence of the bending and twisting of the structure.

Chapter 4

Elements of sub-Riemannian geometry and application to vision

This chapter includes a brief introduction to the mathematical tools that are used to describe the neural mechanisms underlying the visual system. We will introduce the notions of smooth manifold, tangent bundle, Lie group and sub-Riemannian manifold. Finally, we present an application of this mathematical theory to the field of perceptual vision, describing the sub-Riemannian model proposed by Citti-Sarti in [CS06] for neural mechanisms underlying monocular vision.

Contents

4.1 Basic differential geometry

In this section, we review basic notions of differential geometry, mainly following [Tu11].

Definition 4.1.1. *A topological space M is a **manifold** of dimension n if every point on M has a neighborhood U homeomorphic to an open subset of $\mathbb{R}^n$. The homeomorphism $\phi : U \to \mathbb{R}^n$ is called coordinate system or chart. If $p \in U$ and $\phi(p) = 0$, then we say that the coordinate system is centered at p.*

Definition 4.1.2. *Let M be a topological manifold of dimension n. Then, a **smooth structure** $\mathcal{F}$ on M consists of a maximal atlas on M, namely $\mathcal{F} = \{\phi_\alpha; \alpha \in A\}$ is a collection of coordinate systems $\phi_\alpha : U_\alpha \to \mathbb{R}^m$ such that:*

- $\cup_{\alpha \in A} U_\alpha = M$;
- *the transition map $\phi_\alpha \circ \phi_\beta^{-1} : \phi_\beta(U_\alpha \cap U_\beta) \to \phi_\alpha(U_\alpha \cap U_\beta)$ between local coordinate systems is C^∞ for every $\alpha, \beta \in A$;*
- *if a coordinate system ϕ satisfies that $\phi \circ \phi_\alpha$ and $\phi_\alpha \circ \phi^{-1}$ are C^∞ for every $\alpha \in A$, then $\phi \in \mathcal{F}$.*

*A manifold M together with a smooth structure $\mathcal{F}$ on it is called **smooth manifold** (or differentiable manifold).*

Definition 4.1.3. *Let M and N be smooth manifolds. A map $F : M \longrightarrow N$ is C^∞ at a point $p \in M$ if there exist charts (U, ϕ), (V, ψ), around $p \in M$ and $F(p) \in N$ respectively, such that the composition $\psi \circ F \circ \phi^{-1}$ is C^∞ at p. We say that F is smooth if it is C^∞ at every point of M, and write $F \in C^\infty(M, N)$. If $N = \mathbb{R}$, we simply write $F \in C^\infty(M)$.*

For smooth maps between manifolds, it is possible to give the notion of differential, which will be described as a linear map acting on germs of smooth functions.

Definition 4.1.4. *Let $U, V \subset M$ be neighborhoods of $p \in M$ and let $f \in C^\infty(U)$ and $g \in C^\infty(V)$. We say that f and g are **equivalent** if there is a neighborhood $W \subset U \cap V$ of p such that f agrees with g on W. The equivalence class of $f : U \to \mathbb{R}$ is called the **germ** of f at p.*

The set of germs of smooth functions at p in M is denoted by $C^\infty_p(M)$ and it is a unitary commutative ring.

Definition 4.1.5. *Let M be a smooth manifold and let $p \in M$. A **derivation** at p is a linear map $D : C^\infty_p(M) \to \mathbb{R}$ such that, for any $f, g \in C^\infty_p(M)$,*

$$D_p(fg) = (D_p f)g(p) + f(p)(D_p g) \qquad \textit{(Leibniz rule).} \tag{4.1}$$

*A derivation at p is also called a **tangent vector** at p. We denote by T_pM the **tangent space** of M at p, i.e. the set of all tangent vectors at p.*

Definition 4.1.6. *Let $F \in C^\infty(M, N)$ and let $p \in M$. The **differential** of F at p is the linear map $D_pF : T_pM \to T_{F(p)}N$ defined as:*

$$D_pF(X_p)(f) = X_p(f \circ F), \qquad \forall f \in C^\infty_{F(p)}(N),\ X_p \in T_pM. \tag{4.2}$$

When the target space is $N = \mathbb{R}$, we also use the notation d_pF for the differential of F at p.

Note that the above definition is well-posed since $f \circ F \in C^\infty_p(M)$ for any $f \in C^\infty_{F(p)}(N)$. Using the notion of differential, one can prove that T_pM is a vector space of dimension $n = \dim M$.

Definition 4.1.7. *Let M and N be smooth manifolds of dimension n. We say that $F \in C^\infty(M, N)$ is a **diffeomorphism** if F is a homeomorphism with smooth inverse. We say that F is a **local diffeomorphism** at p if there exist neighborhoods $U \subset M$, $V \subset N$ of p and $F(p)$ respectively such that $F_{|U} \colon U \to V$ is a diffeomorphism.*

4.1.1 Vector fields and covariant tensors

Definition 4.1.8. *Let E, M, N be smooth manifolds and let $\pi \colon E \to M$ be a smooth surjection. We say that the structure (E, M, π, N) is a **fiber bundle** if, for every $x \in E$, there exists an open neighborhood $U \subset M$ of $\pi(x)$ and a diffeomorphism $\phi \colon \pi^{-1}(U) \to U \times N$, such that the following diagram commutes:*

$$\begin{array}{ccc} \pi^{-1}(U) & \xrightarrow{\phi} & U \times N \\ & {\scriptstyle \pi} \searrow & \downarrow {\scriptstyle \mathrm{proj}_1} \\ & & U \end{array} \tag{4.3}$$

where $\mathrm{proj}_1 : U \times N \longrightarrow U$ is the projection onto the first factor. The manifold M is called the base space, E the total space, and N the fiber. In particular, for any

$p \in M$, the preimage $\pi^{-1}(p)$ is diffeomorphic to the fiber over p, namely $\{p\} \times N$.

Definition 4.1.9. *Let (E, M, π, N) be a fiber bundle. We say that $e : M \longrightarrow E$ is a **smooth section** of the fiber bundle if $e \in C^\infty(M, E)$ and $\pi \circ e = id_M$.*

Note that, every fiber bundle $\pi : E \longrightarrow M$ is an open map, since projections of products are open maps. A particular case of fiber bundle is given by *vector bundles*, that is when the fiber $N = \mathbb{R}^k$ for some $k \in \mathbb{N}$. In this case, we say that the triple (E, M, π) is a vector bundle of rank k. A notable example of vector bundle is the tangent bundle of a manifold.

Example 4.1.1. *Let M be a smooth manifold of dimension n. Denote by TM the disjoint union of all tangent spaces of M, namely*

$$TM = \bigsqcup_{p \in M} T_pM. \tag{4.4}$$

*TM has the structure of a differentiable manifold. Moreover, denote by $\pi : TM \to M$ the projection such that, for every $p \in M$, $\pi^{-1}(p) = T_pM$. The triple (TM, M, π) is a vector bundle of rank n, called **tangent bundle**. A **vector field** on manifold M is a smooth section of the tangent bundle. The set of vector fields on a manifold M is denoted by $\mathrm{Vec}(M)$.*

*In an analogous way, one can define the **cotangent bundle** $T * M$, which is a vector bundle of rank n whose fibers are given by $T_p^*M := (T_pM)^*$, for every $p \in M$. Sections of the cotangent bundle are called **differential 1-forms**.*

We remark that, by the very definition of tangent vector, a vector field can be seen as a derivation on $C^\infty(M)$, namely a linear map on $C^\infty(M)$ satisfying the Leibniz rule. Moreover, $\mathrm{Vec}(M)$ is a module over the ring of smooth functions.

We introduce the notion of differential form on a manifold, which is a particular case of a covariant tensor. We refrain to give the general definition of tensor bundles and (k, r)-tensors on a manifold since we will not need it.

Definition 4.1.10. *Let M be a smooth manifold. A **covariant tensor** T of order $k \in \mathbb{N}$ is a $C^\infty(M)$-multilinear mapping of the form:*

$$T\colon \underbrace{\mathrm{Vec}(M) \times \ldots \times \mathrm{Vec}(M)}_{k \text{ times}} \longrightarrow C^\infty(M). \tag{4.5}$$

Namely, this means that T is a homomorphism of $C^\infty(M)$-modules in each argument.

Remark 4.1.1. *A covariant tensor T of order k on a smooth manifold is a* pointwise *object, in the following sense: let $p \in M$ and fix a set of coordinates $(x_1, \dots, x_n)$, on $U \subset M$, where $p \in U$. These coordinates induce a basis of the tangent space T_qM, for any $q \in U$, denoted by $\{\partial_{x_1}, \dots, \partial_{x_n}\}$. Then, letting $X_1, \dots, X_k \in \mathrm{Vec}(M)$, we may write*

$$X_{i|q} = \sum_{j=1}^{n} a_{ij}(q)\partial_{x_j|q}, \qquad \forall\, q \in U,\ i = 1, \dots, k. \tag{4.6}$$

Therefore, by linearity we have

$$T(X_1, \dots, X_k)(q) = \sum_{j_1,\dots,j_k=1}^{n} a_{ij_1}(q) \dots a_{ij_k}(q) \underbrace{T(\partial_{x_{j_1}}, \dots, \partial_{x_{j_k}})}_{T_{j_1,\dots,j_k}(q)}. \tag{4.7}$$

The functions $T_{j_1,\dots,j_k}$ on U are called components of T. The expression (4.7) implies that the value of $T(X_1, \dots, X_k)$ at p depends only *on the values of $X_1, \dots, X_k$ at p and of the components of T at p. Moreover, we may also consider T_p as a multilinear map on $(T_pM)^k$.*

Definition 4.1.11. *Let M be a smooth manifold. A* ***Riemannian metric*** *g on M is a* symmetric *covariant tensor of order* 2*. In particular, this means that, for any $p \in M$,*

$$g_p \colon T_pM \times T_pM \longrightarrow \mathbb{R} \tag{4.8}$$

is a bilinear, symmetric map and the dependence $p \mapsto g_p$ is smooth. We say that the pair (M, g) is a ***Riemannian manifold****.*

One can define notions of length and distance on M, using the Riemannian metric. We will do this in the more general framework of sub-Riemannian manifolds in Section 4.3.

Remark 4.1.2. *Let (M, g) be a Riemannian manifold. Then, we can define a* canonical isomorphism *between the tangent space at a point and its dual:*

$$\alpha \colon T_pM \to T_p^*M; \qquad \langle \alpha(v), w \rangle = g_p(v, w), \quad \forall\, w \in T_pM. \tag{4.9}$$

4.1.2 Differential forms and Hodge star operator

Definition 4.1.12. *Let M be a smooth manifold. We say that ω is a* ***differential form*** *of order k, and write $\omega \in \Omega^k(M)$, if it is an* alternating *covariant tensor of*

order k, such that for any permutation $\sigma \in \mathfrak{S}_k$,

$$\omega\Big(X_{\sigma(1)}, \dots, X_{\sigma(k)}\Big) = \operatorname{sgn}(\sigma)\omega(X_1, \dots, X_k), \qquad \forall\, X_1, \dots, X_k \in \operatorname{Vec}(M). \tag{4.10}$$

By construction, $\Omega^k(M)$ is a real vector space of finite dimension.

Note that the notion of differential forms of order 1 coincides with the one given in Definition 4.1.1, as covariant tensors are pointwise objects.

Definition 4.1.13. *Let M be a smooth manifold and let $\omega \in \Omega^k(M)$, $\eta \in \Omega^l(M)$. Then, the **exterior product** between ω and η is the differential form $\omega \wedge \eta \in \Omega^{k+l}(M)$, defined by: for any $X_1, \dots, X_{k+l} \in \operatorname{Vec}(M)$,*

$$(\omega \wedge \eta)(X_1, \dots, X_{k+l}) = \sum_{\sigma \in \mathfrak{S}(k,l)} \operatorname{sgn}(\sigma)\omega\big(X_{\sigma(1)}, \dots, X_{\sigma(k)}\big)\eta\big(X_{\sigma(k+1)}, \dots, X_{\sigma(k+l)}\big), \tag{4.11}$$

where $\mathfrak{S}(k,l) = \{\sigma \in \mathfrak{S}_{k+l} \mid \sigma(1) < \dots < \sigma(k),\ \sigma(k+1) < \dots < \sigma(k+l)\}$.

Remark 4.1.3. *Let us consider a smooth manifold M of dimension n. Then, according to Definition 4.1.6, for any smooth function $f \in C^\infty(M)$, the differential df defines a 1-form. Indeed,*

$$df(X)(p) \coloneqq d_p f(X_p), \qquad \forall\, p \in M,\ X \in \operatorname{Vec}(M). \tag{4.12}$$

Moreover, since covariant tensors are pointwise objects, we may define local differential forms. In coordinates $(x_1, \dots, x_n)$ on $U \subset M$, (4.12) allows to make sense of the (local) 1-forms $dx_1, \dots, dx_n$. By construction, we also have

$$dx_i(\partial_{x_j}) = \delta_{ij}, \qquad \forall\, i, j = 1, \dots, n. \tag{4.13}$$

Reasoning as in Remark 4.1.1, one can prove that any 1-form is a linear combination of the set $\{dx_1, \dots, dx_n\}$, thus it is a basis of $\Omega^1(U)$. In an analogous way, one can prove that $\dim(\Omega^k(M)) = \binom{n}{k}$ with basis given by

$$dx_{i_1} \wedge \dots \wedge dx_{i_k}, \qquad \forall\, 1 \le i_1 < \dots < i_k \le n. \tag{4.14}$$

Definition 4.1.14. *Let (M, g) be a Riemannian manifold. Then, the **Riemannian volume form** vol_g is the n-form defined, in local coordinates $(x_1, \dots, x_n)$*

on $U \subset M$, as

$$\mathrm{vol}_{g|q} = \sqrt{|\det g_q|}dx_1 \wedge \ldots \wedge dx_n, \qquad \forall\, q \in U. \tag{4.15}$$

The definition of vol_g at a point is independent of the choice of local coordinates. Moreover, if M is orientable than $\mathrm{vol}_g \in \Omega^n(M)$.

Remark 4.1.4. *Let (M, g) be a Riemannian manifold and let $(x_1, \ldots, x_n)$ be local coordinates on $U \subset M$. The canonical ismoprphism (4.9) and the description through (4.14) allow to extend the Riemannian metric to k-forms, namely for 1-forms:*

$$g_q^1(dx_i, dx_j) \coloneqq g_q(\alpha(\partial_{x_i}), \alpha(\partial_{x_j})), \qquad \forall\, q \in U,\ i, j = 1, \ldots, n. \tag{4.16}$$

Then, for k-forms, for any $\omega_1, \ldots, \omega_k, \eta_1, \ldots, \eta_k \in \Omega^1(M)$, we have

$$g_q^k(\omega_1 \wedge \ldots \wedge \omega_k, \eta_1 \wedge \ldots \wedge \eta_k) \coloneqq \det\left(g_q^1(\omega_i, \eta_j)\right)_{i,j=1}^{k}, \qquad \forall\, q \in U. \tag{4.17}$$

Definition 4.1.15. *Let (M, g) be an orientable Riemannian manifold of dimension n. Then, the **Hodge star operator** $\star\colon \Omega^k(M) \to \Omega^{n-k}(M)$ is the linear operator defined by*

$$\omega \wedge \star\eta = g^k(\omega, \eta)\mathrm{vol}_g, \qquad \forall\, \omega, \eta \in \Omega^k(M). \tag{4.18}$$

A common example of the Hodge star operator is the case n = 3 for the Euclidean space $M = \mathbb{R}^3$.

Example 4.1.2. *Let $M = \mathbb{R}^3$ with the basis $\{dx, dy, dz\}$ of one-forms. One finds that:*

$$\begin{aligned} \star\, dx &= dy \wedge dz \\ \star\, dy &= dz \wedge dx \\ \star\, dz &= dx \wedge dy. \end{aligned} \tag{4.19}$$

Moreover, the Hodge star relates the exterior and cross-product in three dimensions:

$$\star(\mathbf{u} \wedge \mathbf{v}) = \mathbf{u} \times \mathbf{v} \qquad \star(\mathbf{u} \times \mathbf{v}) = \mathbf{u} \wedge \mathbf{v}. \tag{4.20}$$

where u and v are 1-forms, or vectors, according to the opportune interpretation.

4.1.3 Integral curves on a manifold

Definition 4.1.16. *Let M be a smooth manifold. A **smooth curve** is defined by a map $\gamma : I \to M$, where $I \subset \mathbb{R}$ is an open interval. The tangent vector $\gamma'(t)$ to the curve γ is defined as its differential at t, namely*

$$\gamma'(t) = D_t\gamma\,(\partial_t) \in T_{\gamma(t)}M, \qquad t \in I, \tag{4.21}$$

where ∂_t denotes the tangent vector to the interval I at t.

Definition 4.1.17. *Let M be a smooth manifold and $X \in \mathrm{Vec}(M)$. We say that $\gamma\colon I \to M$ is an **integral curve** of X if*

$$\gamma'(t) = X_{\gamma(t)}, \qquad \forall t \in I. \tag{4.22}$$

*The **flow** $\phi\colon I \times M \to M$ of the vector field X is defined in such a way $\phi(\cdot, p)$ is the unique integral curve of X starting at p, i.e. such that $\phi(0, p) = p$.*

In local coordinates $(x_1, \ldots, x_n)$, we may write a smooth curve as $\gamma(t) = (\gamma_1(t), \ldots, \gamma_n(t))$. In the basis of the tangent space $\{\partial_{x_1}, \ldots, \partial_{x_n}\}$, we have

$$\gamma'(t) = (\gamma_1'(t), \ldots, \gamma_n'(t)) = \sum \gamma_i'(t)\partial_{x_i}. \tag{4.23}$$

Analogously, we can write the vector field $X \in \mathrm{Vec}(M)$ in coordinates:

$$X_p = \sum a_i(x_1, \ldots, x_n)\partial_{x_i}, \tag{4.24}$$

where a_i's are smooth and $p \in M$ has coordinates $(x_1, \ldots, x_n)$. Therefore, we can rewrite equation (4.22) as

$$\sum_i \gamma_i'(t)\frac{\partial}{\partial_{x_i}} = \sum a_i(\gamma_1, \ldots, \gamma_m)\partial_{x_i}, \tag{4.25}$$

Thus, γ is integral curve of X if and only if, in coordinates, is a solution of the previous system of autonomous first order ODEs. In particular, this ensures that, locally, integral curves of a vector field always exist.

4.2 Lie algebras and Lie groups

Definition 4.2.1. *Let M be a smooth manifold and let $X, Y \in \mathrm{Vec}(M)$. Their* ***Lie bracket****, or* commutator*, is defined by their action on smooth functions $f \in C^\infty(M)$ as:*

$$[X, Y](f) = X(Y(f)) - Y(X(f)). \tag{4.26}$$

This quantity is a measure of non-commutativity of the vector fields. Indeed, if $[X, Y] = 0$, the flows of X and Y commute, in the sense that $\phi_X(t, \phi_Y(p)) = \phi_Y(t, \phi_X(p))$, for any $p \in M$. In this case, it is said that X and Y *commute* with each other.

Definition 4.2.2. *A* ***Lie algebra*** $\mathfrak{g}$ *on $\mathbb{R}$ is a real vector space together with a bilinear operation $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, called* Lie bracket*, which satisfy:*

- anti-commutativity property*:* $[X, Y] = -[Y, X]$
- Jacobi identity*:* $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0$ *for all* $X, Y, Z \in \mathfrak{g}$.

Note that the bracket operation defined in Definition 4.2.1 is actually a bilinear form on $\mathrm{Vec}(M)$ satisfying the requirements of the previous definition. Thus, $(\mathrm{Vec}(M), [\cdot, \cdot])$ is a Lie algebra of infinite dimension.

Definition 4.2.3. *Let M be a smooth manifold and let $\mathcal{F} = \{X_1, \ldots, X_k\} \subset \mathrm{Vec}(M)$ be a family of vector fields. The* ***Lie algebra generated by*** $\mathcal{F}$ *is the smallest sub-algebra of $\mathrm{Vec}(M)$ containing the family itself, i.e.*

$$\mathrm{Lie}(\mathcal{F}) = \mathrm{span}\{[X_{j_1}, [\ldots [X_{j_{l-1}}, X_{j_l}]]] \mid j_i \in \{1, \ldots, k\},\ l \in \mathbb{N}\}. \tag{4.27}$$

Definition 4.2.4. *Let $(G, *)$ be a be a group. We say that G is a* ***Lie Group*** *if it carries the structure of a differentiable manifold in such a way that the group operation*

$$* : G \times G \to G; \qquad (g, h) \mapsto g * h, \quad \forall\, g, h, \in G \tag{4.28}$$

and the inversion

$$i : G \to G; \qquad i(g) = g^{-1}, \quad \forall\, g \in G \tag{4.29}$$

are smooth maps.

Definition 4.2.5. *Let $(G, *)$ be a Lie group and let $g \in G$. The* ***left multiplica-***

tion, *or* left translation, $L_g : G \to G$ *is the map :*

$$L_g(h) = g * h, \quad \forall\, h \in G. \tag{4.30}$$

A vector field X *on* G *is called* ***left-invariant*** *if*

$$X(f \circ L_g) = (Xf) \circ L_g, \qquad \forall\, f \in C^\infty(G),\ g \in G. \tag{4.31}$$

Definition 4.2.6. *Let* $(G, *)$ *be a Lie group. The* ***Lie algebra of a Lie group*** *is the vector space of all left-invariant vector fields on* G*:*

$$Lie(G) = \{X \in \mathrm{Vec}(G) \mid X \text{ left-invariant } \}. \tag{4.32}$$

The previous definition makes sense, since $\mathrm{Vec}(G)$, with the bracket, is a Lie algebra. The Lie algebra of a Lie group can be identified as the tangent space at the identity e of the group G :

$$Lie(G) \cong T_eG. \tag{4.33}$$

4.3 Sub-Riemannian geometry

We recall some basic facts about sub-Riemannian geometry, following [ABB19].

Definition 4.3.1. *A sub-Riemannian manifold is defined as a triple* $(M, \mathcal{D}, g)$*, where* M *is a smooth, connected finite-dimensional manifold,* $\mathcal{D}$ *is a sub-bundle of rank* $k \in \mathbb{N}$ *of the tangent bundle, called* ***distribution****, and* g *is a metric on* $\mathcal{D}$*. This implies that for any* $p \in M$*, there exist a neighborhood* $U \subset M$ *of* p *and* k *orthonormal vector fields* $\{X_1, \dots, X_k\}$ *defined on* U*, such that*

$$\mathcal{D}_p = \mathrm{span}\{X_{1|_p}, \dots, X_{k|_p}\} \subseteq T_pM, \qquad \forall\, p \in U. \tag{4.34}$$

We always assume that the distribution is ***bracket-generating****, namely the Lie algebra generated by* $\mathcal{D}$ *at the point* p *coincides with* T_pM*, for every* $p \in M$*, i.e.*

$$\mathrm{Lie}_p(\mathcal{D}) = T_pM, \qquad \forall\, p \in M. \tag{4.35}$$

Recall that, according to Definition 4.2.3, the Lie algebra generated by $\mathcal{D}$ *at* p *can*

be expressed by means of (4.34), namely

$$\mathrm{Lie}_p(\mathcal{D}) = \mathrm{Lie}_p(\{X_1, \dots, X_k\}). \tag{4.36}$$

We sometimes refer to (4.35) as ***Hörmander condition****. With a slight abuse of notation, we say that M is a sub-Riemannian manifold.*

Note that the previous definition includes Riemannian manifolds, indeed if $\mathcal{D} = TM$, then g must be a Riemannian metric on M. Thus, sub-Riemannian geometry is a generalization of Riemannian geometry.

Definition 4.3.2. *Let M be a sub-Riemannian manifold. We say that $\gamma : [0, T] \to M$ is a* ***horizontal curve****, if it is absolutely continuous in coordinates and*

$$\dot{\gamma}(t) \in \mathcal{D}_{\gamma(t)}, \qquad \text{for a.e. } t \in [0, T]. \tag{4.37}$$

This implies there exists a function $u : [0, T] \to \mathbb{R}^k$, such that

$$\dot{\gamma}(t) = \sum_{i=1}^{k} u_i(t) X_{i|_{\gamma(t)}}, \qquad \text{for a.e. } t \in [0, T]. \tag{4.38}$$

where $\{X_1, \dots, X_k\}$ is a local orthonormal frame for $\mathcal{D}$ as in (4.34). We require that $u \in L^2([0, T], \mathbb{R}^k)$.

By construction, $\|\dot{\gamma}(t)\|_{|\gamma(t)} = |u(t)|$ for a.e. $t \in [0, T]$, where $|\cdot|$ is the Euclidean norm in $\mathbb{R}^k$. Therefore the map $t \mapsto \|\dot{\gamma}(t)\|_{|\gamma(t)}$ is square integrable on $[0, T]$ and we can evaluate the length of horizontal curves. This allows us to define the sub-Riemannian distance on M.

Definition 4.3.3. *Let M be a sub-Riemannian manifold, and let $\gamma \colon [0, T] \to M$ be a horizontal curve. Its* ***length*** *is defined as follows:*

$$\ell(\gamma) = \int_0^T \|\dot{\gamma}(t)\|_{|\gamma(t)} dt. \tag{4.39}$$

Then, the ***Carnot–Carathéodory distance****, or sub-Riemannian distance, is defined by length-minimization procedure: for any $p, q \in M$, we set*

$$d_{\mathrm{SR}}(p, q) = \inf\{\ell(\gamma) \mid \gamma \colon [0, T] \to M, \ \text{horizontal curve joining } p \text{ and } q\}. \tag{4.40}$$

According to Definition 4.3.2, not every curve is admissible and has finite

length, thus it is not guaranteed that any couple of points is at finite distance. The next theorem ensures that, under the Hörmander condition for the distribution, the sub-Riemannian distance is indeed well-defined.

Theorem 4.3.1 (Chow-Rashevsky)**.** *Let M be a sub-Riemannian manifold. Then, the distance $d_{\mathrm{SR}} \colon M \times M \to \mathbb{R}$ is finite, continuous and it induces the same topology as that of the manifold. Furthermore, the resulting metric space (M, d_{SR}) is locally compact.*

We report here a seminal result by Hörmander [Hör67] ensuring the hypoellipticity of operators obtained as sum of squares of vector fields plus possibly a drift.

Theorem 4.3.2. *Let M be a smooth manifold and let $\{X_1, \dots, X_k\}$ be a set of bracket-generating vector fields on M. Then the operator*

$$L = X_1 + \sum_{i=2}^{k} X_i^2 \tag{4.41}$$

is hypoelliptic, which means that if φ is a distribution defined on an open set $\Omega \subset M$, such that $L\varphi \in C^\infty(\Omega)$, then $\varphi \in C^\infty(\Omega)$.

In a sub-Riemannian manifold M, with generating frame (4.34) satisfying the Hörmander condition, we can immediately apply the above theorem to obtain regularity for the solution to $Lu = 0$.

4.4 Sub-Riemannian model of monocular vision

Sub-Riemannian models of visual perception have been introduced in the work of Citti and Sarti in [CS06], enriching the field of neurogeometry of vision. In this section, we briefly review the history underlying the development of this discipline, and then we concentrate on the study of the monocular Citti-Sarti model, one of the starting points for our work.

4.4.1 Neurogeometry of vision

Families of cells involved in the visual sensory perceptual process are sensitive to different properties of the image: position, orientation, contrast, curvature, scale, phase, motion and stereo. In general, most of these families have a hypercolumnar structure, as shown in Section 2.1.2.3: over every point (x, y) of the retinal plane

there is an entire set of cells, each one sensible to a particular feature. Neuromathematical models of V1 often adopt a differential geometry approach to modeling the functional architecture of this area. We summarize the main steps characterizing the evolution of neurogeometry, principally following [CS14, Chapter 1]:

- **Introduction of differential geometry to vision**
  The pioneering idea underlying the application of differential geometry to vision is introduced in the seminal works of perceptual psychology specialists Koenderink and van Doorn in [KvD87] and Hoffmann in [Hof89] in the mid-80s. In particular, Hoffman had the breakthrough idea of representing the hypercolumnar organization of V1 as a fiber bundle with the space of retinal locations as a basis, introducing Lie group theory to model the functional architecture of the primary visual cortex. Almost simultaneously Mumford in [Mum94], announced a variational approach to describe smooth edges, in terms of the elastica functional, developing a stochastical explanation.

- **The concept of neurogeometry**
  Results dealing with differential geometry and vision were unified under the name of neurogeometry by Petitot and Tondut [PT99], who related the association fields of Field Hayes and Hess [FHH93] with the contact geometry introduced by Hoffmann [Hof89] and the elastica of Mumford [Mum94]. By describing the set of simple cells as a fiber bundle, they identified the structure of this layer of cortical cells with the Heisenberg group, performing contour completion in this structure by minimizing a suitable functional.

- **Developments after 2000**
  The proposal of Petitot in [PT99], with the introduction of the term neurogeometry, is at the basis of a large class of models, which uses instruments of differential geometry or group theory to explain the behavior of the visual cortex starting from its functional architecture.
  In this framework, Citti and Sarti observed that stratified Lie groups structures more appropriately describe the functional architecture of the visual cortex, proposing a description for the set of simple cells requiring invariance under roto-translations, namely introducing the Lie group $SE(2)$. The important idea was to endow it with a sub-Riemannian metric, which well describes the behavior of the neural and psychophysical mechanisms at the

basis of visual perception. We will study this model in more detail in Section 4.4.2.

This mathematical description of V1 has been used in many other works to define a symplectic structure [SCP07], to study cortical orientation maps [BCSS12, BCS14], to study different features selectivity such as scale [BCS18] and spatio-temporal visual stimuli [BCCS14], to group perceptual stimuli [BCCS14, ASFC+17], to study visual illusions [FMCS19, BCF+20]. A semi-discrete variant of [CS06] was proposed in [BCGR14], and a relation between association field curves and sub-Riemannian geodesics in $SE(2)$ has been carried out in [DBRS14]. Moreover, a spherical extension of the contour perception model has been proposed in [MD16].

4.4.2 The Citti-Sarti model

Formally, it is possible to model the primary visual cortex as $\mathbb{R}^2 \times \mathbb{S}^1$ (position-orientation space), and thus naturally encoding the hypercolumnar structure discovered by Hubel and Wiesel in [HW62]. An example of this structure is displayed in image (a) of Figure 4.4.3, where the model proposed by Ben-Shahar and Zucker in [BSZ04] is shown.

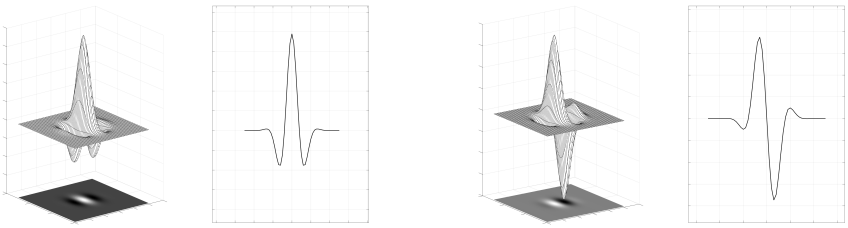
We have seen that the receptive field (RF) of a cortical neuron is the portion of the retina to which the neuron reacts, and the receptive profile (RP) φ is the function that models the activation of a cortical neuron. In particular, the receptive profiles of simple cells in the primary visual cortex V1, sensitive to position (x, y) and orientation θ, has classically been modeled through a bank of Gabor filters $\varphi_{\{x,y,\theta\}}$, which act on a visual stimulus ([BCS14, Dau85, JP87]).

Following the neuro-mathematical model of Citti and Sarti [CS06], the set of simple cells RPs can be obtained via translations of vector (x, y) and rotation of angle θ from a unique mother profile $\varphi_0(\xi, \eta)$

$$\varphi_0(\xi, \eta) = \exp(2\pi i \xi) \exp\left(-\frac{\xi^2 + \eta^2}{2\sigma^2}\right), \tag{4.42}$$

a Gabor function with real (even) and imaginary (odd) parts, phase[1] ϕ equal to 0 and frequency f equal to 1, displayed in Figure 4.4.1. Translations and rotations

[1]The general expression for the Gabor function with phase ϕ and frequency f is $\varphi_0(\xi, \eta) = \exp(i(2\pi f \xi + \phi)) \exp\left(-\frac{\xi^2+\eta^2}{2\sigma^2}\right)$. We assume $f = 1$ and $\phi = 0$. For the reader interested in a model of phase variation, we recommend [BCS18].



(a) (b)

Figure 4.4.1: Even (a) and odd (b) part of Gabor function: the surface of the two-dimensional filters, their common bi-dimensional representation and a mono-dimensional section.

can be expressed as:

$$T_{(x,y,\theta)}(\xi,\eta) = \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix}, \tag{4.43}$$

and $T_{(x,y,\theta)}$ denotes the action of the group of rotation and translation $SE(2)$ on $\mathbb{R}^2$, which associate to every vector (ξ,η) a new vector $(\tilde{x},\tilde{y})$, according to the law $(\tilde{x},\tilde{y}) = T_{(x,y,\theta)}(\xi,\eta)$. Hence a general RP can be expressed as

$$\varphi_{(x,y,\theta)}(\xi,\eta) = \varphi_0(T^{-1}_{(x,y,\theta)}(\xi,\eta)), \tag{4.44}$$

and this represents the action of the group $SE(2)$ on the set of receptive profiles.

The retinal plane $\mathcal{R}$ is identified with the $\mathbb{R}^2$ plane, whose coordinates are denoted by (x,y). When a visual stimulus $I : \mathcal{R} \longrightarrow \mathbb{R}^+$ activates the retinal layer, the neurons centered at every point (x,y) produce an output $O(x,y,\theta)$, which can be modeled as the integral of the signal I with the set of Gabor filters:

$$O(x,y,\theta) = \int_{\mathcal{R}} \varphi_{\{x,y,\theta\}}(\xi,\eta) I(\xi,\eta) d\xi d\eta, \tag{4.45}$$

where the function I represents the retinal image.

For (x,y) fixed, we will denote $\bar{\theta}(x,y)$ the angle of maximal response:

$$\max_{\theta} \|O(x,y,\theta)\| = \|O(x,y,\bar{\theta}(x,y))\|; \tag{4.46}$$

we will then say that the point (x,y) is lifted to the point $(x,y,\bar{\theta}(x,y))$. If all the points of the image are lifted in the same way, the level lines of the $2D$ image I

are lifted to new curves in the $3D$ cortical space (x, y, θ).

In the right-hand side of the equation (4.45) the integral of the signal with the real and imaginary part of the Gabor filter is expressed. The two families of cells have different shapes, hence they detect different features. In particular odd cells are responsible for boundary detection, and we concentrate on these ones. The output of simple cells (4.45) can then be locally approximated as $O(x, y, \theta) = -X_{3,\theta}(I_\sigma)(x, y)$ at every point (x, y, θ) where I_σ is a smoothed version of I, obtained by convolving it with a Gaussian kernel, and

$$X_{3,\theta} = -\sin\theta\partial_x + \cos\theta\partial_y. \tag{4.47}$$

Switching to dual, the action of simple cells induces a choice of contact form separately on each cell:

$$\omega_\theta = -\sin\theta\,\mathrm{d}\,x + \cos\theta\,\mathrm{d}\,y. \tag{4.48}$$

The vector fields lying on its kernel, namely

$$X_1 = \cos\theta\partial_x + \sin\theta\partial_y, \quad X_2 = \partial_\theta \tag{4.49}$$

define a bi-dimensional sub-bundle of the tangent bundle to $\mathbb{R}^2 \times \mathbb{S}^1$, referred to as the *horizontal* tangent bundle. One can define a scalar product on this sub-bundle by imposing the orthonormality of X_1 and X_2: this determines a sub-Riemannian structure on $\mathbb{R}^2 \times \mathbb{S}^1$.

Example 4.4.1. (Imposing orthonormality of vector fields) *The enforcement of the orthonormality of the vector fields X_1 and X_2 results in the imposition of a degenerate metric for the space $\mathbb{R}^2 \times \mathbb{S}^1$ with coordinates (x, y, θ). The metric matrix in these coordinates can be obtained in two steps: first by considering the metric matrix in the basis determined by the involved vector fields, and then by applying a classic change of basis.*

The tangent space basis determined by the involved vector fields, more precisely X_1, X_2 defined in (4.49) and X_3 introduced in (4.47), is:

$$\mathcal{B} = \{X_1, X_3, X_2\}. \tag{4.50}$$

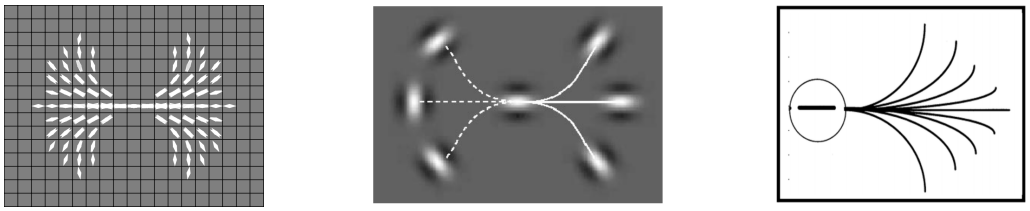
Using this basis, the $g_{\mathcal{B}}$ matrix of the metric takes the form

$$g_{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{4.51}$$

since we are imposing just the orthonormality of X_1 and X_2. Let us now consider the canonical basis of the space tangent to $\mathbb{R}^2 \times \mathbb{S}^1$ with respect to the coordinates (x, y, θ) i.e. $\mathcal{C} = \{\partial_x, \partial_y, \partial_\theta\}$. The metric matrix in this basis is given by

$$g_{\mathcal{C}} = \begin{pmatrix} \cos^2\theta & \cos\theta\sin\theta & 0 \\ \cos\theta\sin\theta & \sin^2\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4.52}$$

obtained via $g_{\mathcal{C}} = M^T_{\mathcal{C}\to\mathcal{B}}(g_{\mathcal{B}})M_{\mathcal{C}\to\mathcal{B}}$, where $M_{\mathcal{C}\to\mathcal{B}}$ is the basic change matrix from $\mathcal{C}$ to $\mathcal{B}$.

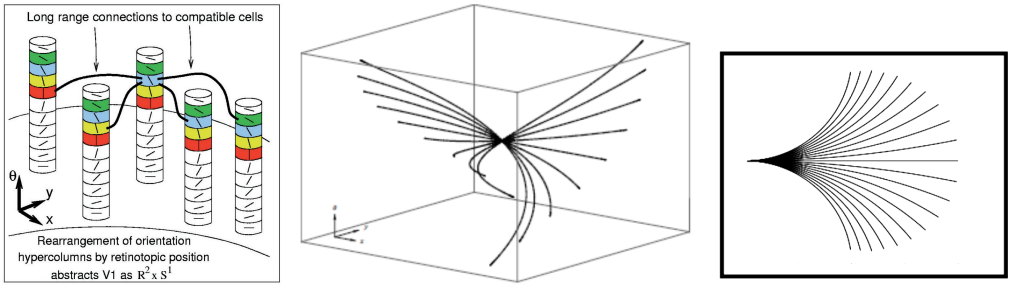


(a) (b) (c)

Figure 4.4.2: (a) Examples of the compatibilities around the central point of the image, derived from planar co-circularity. Brightness encodes compatibility values. Figure adapted from [BSZ04]. (b) Starting from the central initial oriented point, the solid line indicates a configuration between the patches where the association exists while the dashed line indicates a configuration where it does not. Figure adapted from [FHH93]. (c) Association field of Field, Hayes and Hess. Figure adapted from [FHH93].

The visual signal propagates, in an anisotropic way, along cortical connectivity and connects more strongly cells with comparable orientations. As we have seen in Section 2.2.2 the relevance of alignment in bi-dimensional contour integration has been first stated in the Gestalt law of good continuation [Kof63, Koh67]. The co-circularity in the bi-dimensional planes has been studied by Zucker et al in [PZ89] to describe the consistency and the compatibility of neighboring oriented points, in accordance with specific values of curvature. An example of this compatibilities

can be found in Figure 4.4.2, image (a). Psychophysical experiments were further conducted, see for example [IBR89, SV87, Utt83]. In particular, Field and al. in [FHH93] describe the association rules for bi-dimensional contour integration, introducing the concept of association fields. A representation of these connections can be found in Figure 4.4.2, images (b) and (c). Neurophysiological studies [Bla92, BZSF97, HMD14, MAHG93, SGLS97] show that the cortical correlate of the psychophysically defined association field is the cortical connectivity among cells of similar orientation preference, the so-called long-range connectivity.



(a) (b) (c)

Figure 4.4.3: (a) Orientation columns of cells in (x, y, θ) coordinates and long range horizontal connections. Figure adapted from [BSZ04]. (b) Horizontal integral curves in $\mathbb{R}^2 \times \mathbb{S}^1$ generated by the sub-Riemannian model geometry proposed by Citti and Sarti in [CS06]. (c) Projection of the fan of the integral curves in the (x, y) plane. Figure adapted from [CS06].

Based on these findings, Citti and Sarti in [CS06] modeled the propagation in the cortex as a propagation along integral curves of the vector fields X_1 and X_2, namely

$$\dot{\gamma}(t) = X_{1,\gamma(t)} + kX_{2,\gamma(t)}, \tag{4.53}$$

obtained by varying the parameter k in $\mathbb{R}$, see Figure 4.4.3. Their 2D projection is shown to produce a close approximation of the association fields, as shown in [CS06] displayed in Figure 4.4.3, image (b).

Other models have been proposed, as for example the work of Boscain, Duits, Rossi and Sachkov [BDRS14] or the one of Duits, Boscain, Rossi and Sachkov [DBRS13]. They consider the SE(2) sub-Riemannian point of view presented in [CS06], and then, they study the geodesics of the sub-Riemannian structure, to take into account all appropriate end-conditions of association fields. More precisely, the set of possible orientations in a point (x, y) that can be reached and connected with an initial segment in $(0, 0)$ oriented along the horizontal axis is usually described

by an angle θ, which satisfies the following inequality: $\tan^{-1}\left(\frac{y}{x}\right) \leq \theta \leq \frac{\pi}{2}$. This condition is sometimes referred to as *2D relatability*, see for example [KGS05a]. The authors of [BDRS14, DBRS13] observe that in 2D the set of end-points can be obtained by minimizing geodesics whose spatial projections do not reveal cusps, proposing then a model for 2D relatability and association fields.

It is worth making a final consideration regarding the fact that, in the model proposed by Citti-Sarti, the retina is treated locally as if it were a plane, using retinotopic coordinates (x, y). We will adopt this approach throughout the thesis, as it describes well the local behavior, even if globally the retina can be viewed as a sphere. A model that has adopted this point of view has been introduced in [MD16] in which spherical association fields arise from the $SO(3)$ sub-Riemannian geometry acting on the spherical frame bundle. However, the comparison between sub-Riemannian geodesics and wavefronts in $SE(2)$ and $SO(3)$ studied by the authors of [MDS+17], shows that locally the two geometries are very similar.

Chapter 5

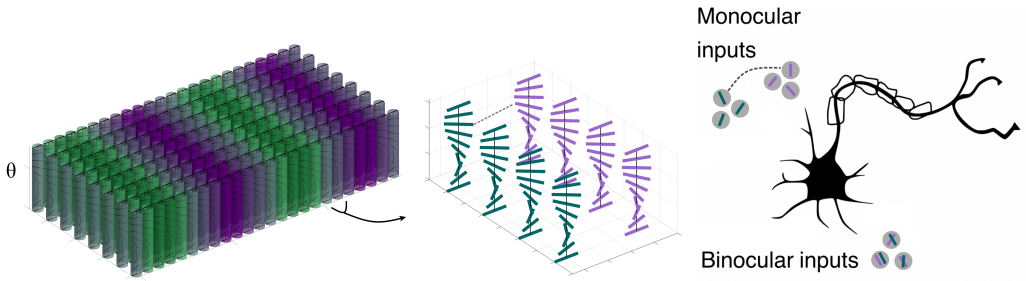
The sub-Riemannian model for stereo vision

In this chapter we present the neurogeometric model for stereo vision proposed in [BCSZ23], in which the parametrization involves both spatial and orientation disparities. This model is based on a sub-Riemannian structure on the three-dimensional space of position-orientation and provides insight into neurobiology, suggesting an implicit organization for neural interactions and a well-defined 3D association field. Integral curves of the sub-Riemannian structure shed light on the computations underlying the correspondence problem, and the model illustrates how good continuation in the world generalizes good continuation in the plane.

Contents

5.1 Motivation for the mathematical approach

Results presented in previous chapters show that, for the cortical-inspired geometry underlying stereo vision, it makes sense to encode positional disparities and orientation differences in the information coming from the two eyes.



(a) (b) (c)

Figure 5.1.1: Cartoon of visual cortex, V1, superficial layers. (a) Macroscopic organization: a number of (abstracted) orientation hypercolumns, colored by left-eye (green)/right-eye (purple) dominant inputs. The color grading emphasizes that at the center of the ocular dominance bands the cells are strongly monocular, while at the boundaries they become binocularly-driven. (b) A zoom in to a few orientation columns showing left and right monocular cells at the border of ocular dominance bands. Cells in these nearby columns will provide the anatomical substrate for our model. (c) More recent work shows that both monocular and binocular inputs matter to these cells (redrawn from [STR+22], using data from ferret). This more advanced wiring suggests the connection structures in our model.

Monocular information is segregated into ocular dominance bands [LHW75] in layer 4, as cartooned in Figure 5.1.1. Image (a) shows an array of orientation hypercolumns arranged over retinotopic positions, image (b) accentuates the orientation distribution along a few of the columns near each position: horizontal connections (not shown) affect the interactions between these units (subsection 2.1.2.3). Images (a) and (b) in Figure 5.1.1 are colored by dominant eye inputs: while the monocular cells are toward the centers of the bands, the binocularly-driven cells

tend to be closer to the ocular dominance boundaries. This configuration leads to wonder about the nature of the interaction among groups of cells representing different orientations at nearby positions and innervated by inputs from the left and right eyes. The physiology suggests (Figure 5.1.1 image (c)) the answer lies in the interactions among both monocular and binocular cells.

In this chapter we specify these interactions through the proposition of a neurogeometric model for binocular cells, building on neuromathematical models, with particular emphasis on the neurogeometry of monocular simple cells in V1 ([CS06, Pet08, PT99, SCS10a, SC15, SCP07]).

5.2 Differential interpretation of binocular neural mechanisms

The first step to tie biological results, pre-existing neuromathematical models and classical stereo geometry is to give a differential interpretation of binocular neural mechanisms underlying the stereo problem. We start with binocular receptive profiles and we realize a natural fiber bundle for modeling stereo, then, we identify the natural direction emerging from these binocular RPs, which will be the direction of the tangent to a potential 3D curvilinear object in the scene.

5.2.1 Binocular profiles

Binocular neurons receive inputs from both the left and right eyes, and to facilitate calculations, we assume these inputs are first combined in simple cells in the primary visual cortex. This is a widely studied approach, see for example [AOF99b, CD01, KBSO16, MF04]. In particular, it provides a first approximation in which binocular RPs are described as the product of monocular RPs, see Figure 3.2.3.

This binocular model allows us to define disparity and frontoparallel coordinates as

$$\begin{cases} d = \frac{x_L - x_R}{2} \\ x = \frac{x_R + x_L}{2}, \end{cases} \tag{5.1}$$

perfectly in accordance with the introduction of cyclopean coordinates in (3.4). In this way (x, y, d) correspond to the neural correlate of (r_1, r_2, r_3), via the change of variables (3.5).

5.2.2 The cortical fiber bundle of binocular cells

The hypercolumnar structure of monocular simple cells (orientation selective) has been described as a jet fiber bundle in the works of Petitot and Tondut [PT99], among many others. We concentrate on the fiber bundle $\mathbb{R}^2 \times \mathbb{S}^1$, with fiber $\mathbb{S}^1$, see e.g. [BSZ04].

In our setting, the binocular structure is based on monocular ones. In particular, for each cell on the left eye there is an entire fiber of cells on the right, and vice versa, for each cell on the right there is an entire fiber of cells on the left. This implies that the binocular space is equipped with a symmetry that involves the left and right structures, allowing us to use the cyclopean coordinates (x, y, d) defined in (5.1).

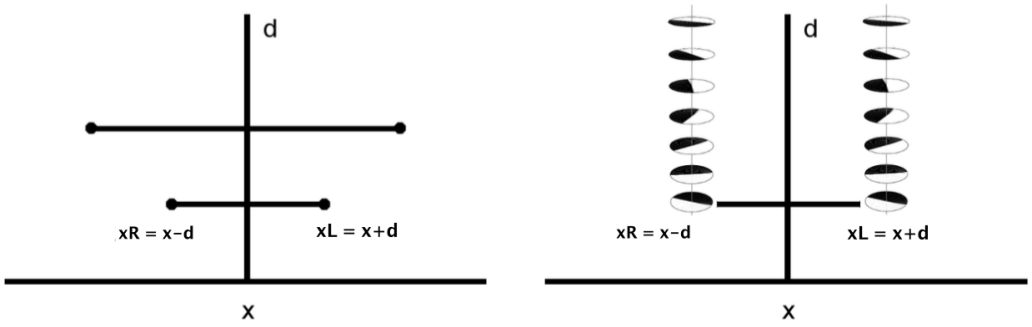


Figure 5.2.1: Left: schematic representation of the fiber bundle in two dimension, with relationships between left and right retinal coordinates. Right: representation of the selection of a whole fiber of left and right simple cells, for every x and for every d.

Hence, we define the cyclopean retina $\mathcal{R}$, identified with $\mathbb{R}^2$, endowed with coordinates (x, y). The structure of the fiber is $\mathcal{F} = \mathbb{R} \times \mathbb{S}^1 \times \mathbb{S}^1$, with coordinates $(d, \theta_L, \theta_R) \in \mathcal{F}$. The total space is defined in a trivial way, $\mathcal{E} = \mathcal{R} \times \mathcal{F} = \mathbb{R}^2 \times \mathbb{R} \times \mathbb{S}^1 \times \mathbb{S}^1$, and the projection $\pi : \mathcal{E} \longrightarrow \mathcal{R}$ is the trivial projection $\pi(x, y, d, \theta_L, \theta_R) = (x, y)$. The preimage of the projection $\mathcal{E}_{(x,y)} := \pi^{-1}(\{(x, y)\})$, for every $(x, y) \in \mathcal{R}$, is isomorphic to the fiber $\mathcal{F}$, and the local trivialization property is naturally satisfied.

A schematic representation can be found in Figure 5.2.1. The base has been depicted as 1-dimensional, considering the restriction $\mathcal{R}_{|x}$ of the cyclopean retina $\mathcal{R}$ on the coordinate x. The left image displays only the disparity component of the

fiber $\mathcal{F}$, encoding the relationships between left and right retinal coordinates. The right image shows the presence of the left and right monodimensional orientational fibers.

5.2.3 Beyond the binocular energy model

We have recalled in subsection 3.2.1 that the classical model for expressing the left/right-eye receptive field combination is the binocular energy model (BEM) introduced in [AOF99b] for binocular RPs. However, this proposal typically ignores the orientational feature of monocular simple cells, even if we have seen that this information is needed to encode tilt and slant of objects in the space: orientation disparity happens in the projections in the left and right eye. This orientational inequality has mathematical but also neurophysiological evidence, as shown for example in [BCP01], when the authors extended the BEM to incorporate these differences, or in the many works presented in paragraph 2.1.2.4.2.

So, even though the binocular energy model is a type of minimal model, it serves as a starting point, allowing the combination of monocular inputs. However, it is not sufficient to solve the stereo-matching problem. It is argued in [PSK16, SPTL13] that, in addition to the neural mechanisms that couple characteristics (such as signals, stimuli, or particular features) relating the left and right monocular structures, there must be a system of connections between binocular cells, which characterizes the processing mechanism of stereo vision.

In the following, binocular RPs and differential forms will allow us to identify a natural direction along which we can establish an opportune set of connections.

5.2.4 Differential forms and binocular RPs

The BEM model, according to equation (3.10), generates the interaction term $O_L O_R$, which can be expressed using left and right monocular RPs:

$$\begin{aligned} O_R O_L &= \int \varphi_{\theta_R,x_R,y}(\tilde{x}_R,\tilde{y}_R) I_R(\tilde{x}_R,\tilde{y}_R) d\tilde{x}_R d\tilde{y}_R \int \varphi_{\theta_L,x_L,y}(\tilde{x}_L,\tilde{y}_L) I_L(\tilde{x}_L,\tilde{y}_L) d\tilde{x}_L d\tilde{y}_L \\ &= \int\int \varphi_{\theta_R,x_R,y}(\tilde{x}_R,\tilde{y}_R)\varphi_{\theta_L,x_L,y}(\tilde{x}_L,\tilde{y}_L) I_R(\tilde{x}_R,\tilde{y}_R) I_L(\tilde{x}_L,\tilde{y}_L) d\tilde{x}_R d\tilde{y}_R d\tilde{x}_L d\tilde{y}_L. \end{aligned} \tag{5.2}$$

If we fix $\tilde{y}_i$ for $i = L, R$, we derive the expression of the binocular profiles $\varphi_{R,L} = \varphi_{\theta_R,x_R,y}\varphi_{\theta_L,x_L,y}$ as product of monocular left and right ones, so that $O_R O_L = \int\int \varphi_{R,L} I_L I_R$. This is in accordance with the measured profiles of Figure 3.2.3.

Since we know from the monocular case ([CS06]) that there exists a relationship between RPs and 1-forms, we can see that there is an analogous correspondence in the binocular case with a 2-form of the space expressed in cyclopean coordinates.

Proposition 5.2.1. *The binocular interaction term can be recast as wedge product of the two monocular 1-forms ω_{θ_L} and ω_{θ_R} defined in (4.48):*

$$O_R O_L = \omega_{\theta_R} \wedge \omega_{\theta_L}. \tag{5.3}$$

Proof. It follows from [CS06] that the output of simple cells (4.45) in $SE(2)$ can be locally approximated as $O(x, y, \theta) = -X_{3,\theta}(I_\sigma)(x, y)$. 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. Accordingly, it is possible to re-write the binocular interaction term as:

$$O_R O_L = X_{3,\theta_R}(I_{\sigma_R})(x_R, y) X_{3,\theta_L}(I_{\sigma_L})(x_L, y). \tag{5.4}$$

We define $v_R := (D_{(x_R,y)} I_{\sigma_R} \vec{X}_{3,\theta_R}) X_{3,\theta_R}$ with $D_{(x_R,y)} I_{\sigma_R}$ differential (Jacobian) of the smoothed version of the image I_{σ_R} at the point (x_R, y), in such a way that we have

$$O_R = X_{3,\theta_R}(I_{\sigma_R})(x_R, y) = (D_{(x_R,y)} I_{\sigma_R} \vec{X}_{3,\theta_R}) = \omega_{\theta_R}(v_R) \tag{5.5}$$

since $\omega_{\theta_R}(X_{3,\theta_R}) = 1$ and $D_{(x_R,y)} I_{\sigma_R} \vec{X}_{3,\theta_R} \in \mathbb{R}$; the same reasoning holds for the left structure. It is then possible to recast (5.4) in the retinal coordinates as:

$$\begin{aligned} O_R O_L =& \omega_{\theta_R}(v_R) \omega_{\theta_L}(v_L) \\ =& \omega_{\theta_R} \wedge \omega_{\theta_L}(v_R, v_L) + \underbrace{\omega_{\theta_L}(v_R) \omega_{\theta_R}(v_L)}_{=0}, \\ =& \omega_{\theta_R} \wedge \omega_{\theta_L}(v_R, v_L), \end{aligned} \tag{5.6}$$

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

Expressing the retinal coordinates in terms of cyclopean ones (3.4) we have

$$x_R = x - d \text{ and } x_L = x + d, \tag{5.7}$$

and the extended left and right 1-form can be written as:

$$\begin{aligned}\omega_{\theta_R} &= -\sin\theta_R\,\mathrm{d}\,x + \cos\theta_R\,\mathrm{d}\,y + \sin\theta_R\,\mathrm{d}\,d\\ \omega_{\theta_L} &= -\sin\theta_L\,\mathrm{d}\,x + \cos\theta_L\,\mathrm{d}\,y - \sin\theta_L\,\mathrm{d}\,d.\end{aligned} \tag{5.8}$$

We define $\omega_{bin} := \omega_{\theta_R} \wedge \omega_{\theta_L}$ as the natural 2-form characterizing the binocular structure. By an abuse of notation, identifying 1-forms with vectors, it is possible to use the Hodge duality in three dimension to relate the wedge product with the cross product $\omega_{bin} = \star(\omega_{\theta_R} \wedge \omega_{\theta_L}) = \omega_{\theta_R} \times \omega_{\theta_L}$. In this way we get:

$$\omega_{bin} = \begin{pmatrix}\sin(\theta_R + \theta_L)\\ 2\sin\theta_R \sin\theta_L\\ \sin(\theta_L - \theta_R)\end{pmatrix}. \tag{5.9}$$

Remark 5.2.1. *The vector ω_{bin} of equation (5.9) can be interpreted as the intersection of the kernels generated from $\omega_{\theta_R}, \omega_{\theta_L}$ expressed in in cyclopean coordinates (x, y, d):*

$$\begin{aligned}&\ker\omega_{\theta_L} = \operatorname{span}\left\{\begin{pmatrix}\cos\theta_L\\ \sin\theta_L\\ 0\end{pmatrix}, \begin{pmatrix}1\\ 0\\ 1\end{pmatrix}\right\}, \qquad \ker\omega_{\theta_R} = \operatorname{span}\left\{\begin{pmatrix}\cos\theta_R\\ \sin\theta_R\\ 0\end{pmatrix}, \begin{pmatrix}-1\\ 0\\ 1\end{pmatrix}\right\}\\ &\Rightarrow \omega_{bin} = \ker\omega_{\theta_R} \cap \ker\omega_{\theta_L}.\end{aligned} \tag{5.10}$$

The result of the intersection of these structures identifies a direction, as shown in Figure 5.2.2.

In other words, we earlier recalled (subsection 4.4.2) that the result of the action of a monocular odd simple cell is to select directions for the propagation of information. We now have combined these, for the two eyes, to show that in the three-dimensional case the binocular neural mechanisms also lead to a direction. We will see in the next sections that this direction is the direction of the tangent vector to the 3D stimulus, provided points are corresponding.

5.2.5 Compatibility with stereo geometry

We consider the 2-form characterizing the binocular structure ω_{bin} defined in (5.9) and we show that it can be associated with the tangent vector of the 3D stimulus, mathematically modeled as a curve $\gamma : \mathbb{R} \longrightarrow \mathbb{R}^3$. The idea is that this tangent

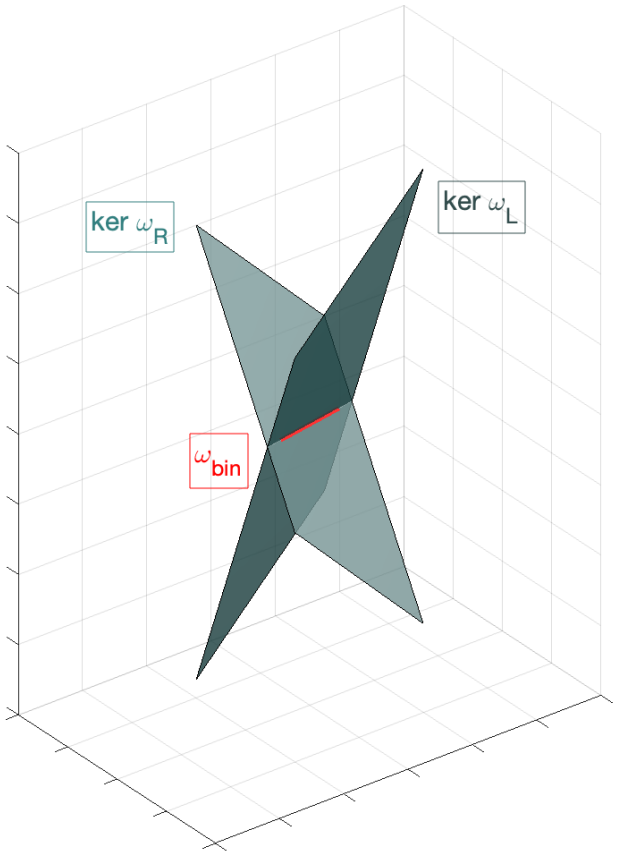


Figure 5.2.2: Direction detected by ω_{bin} through the intersection of left and right planes generated by $\ker \omega_{\theta_R}$ and $\ker \omega_{\theta_L}$. Red vector corresponds to the associated 2-form ω_{bin}.

vector is orthogonal both to ω_{θ_R} and to ω_{θ_L} , and therefore it has the direction of the vector product $\omega_{\theta_R} \wedge \omega_{\theta_L}$.

Precisely, we consider the normalized tangent vector t_L and t_R on retinal planes

$$t_R = (\cos\theta_R, \sin\theta_R) \quad t_L = (\cos\theta_L, \sin\theta_L), \tag{5.11}$$

applied to the points (x_R, y) and (x_L, y), retinal projections of $\gamma \in \mathbb{R}^3$, according to equations (3.1). Taking into account that f is the focal coordinate of the retinal planes in $\mathbb{R}^3$, we associate to these planar points the corresponding points $\tilde{m}_L = (x_L - c, y, f)$, $\tilde{m}_R = (x_R + c, y, f)$ in $\mathbb{R}^3$, with $c \in \mathbb{R}$ coordinate of the optical centers as defined in section 3.1. Applying equation (3.7) to this setting,

it is possible to derive the tangent vector of the three-dimensional contour:

$$\begin{aligned} U_{t_L} = P_L^{-1}\tilde{m}_L \times P_L^{-1}\tilde{t}_L &= \begin{pmatrix} x_L \\ y_L \\ f \end{pmatrix} \times \begin{pmatrix} \cos\theta_L \\ \sin\theta_L \\ 0 \end{pmatrix} = \begin{pmatrix} -f\sin\theta_L \\ f\cos\theta_L \\ x_L\sin\theta_L - y_L\cos\theta_L \end{pmatrix}, \\ U_{t_R} = P_R^{-1}\tilde{m}_R \times P_R^{-1}\tilde{t}_R &= \begin{pmatrix} x_R \\ y_R \\ f \end{pmatrix} \times \begin{pmatrix} \cos\theta_R \\ \sin\theta_R \\ 0 \end{pmatrix} = \begin{pmatrix} -f\sin\theta_R \\ f\cos\theta_R \\ x_R\sin\theta_R - y_R\cos\theta_R \end{pmatrix}, \end{aligned} \tag{5.12}$$

and recovering the tangential direction via the cross product

$$U_{t_R} \times U_{t_L} = f \begin{pmatrix} \cos\theta_R(x_L\sin\theta_L - y_L\cos\theta_L) - \cos\theta_L(x_R\sin\theta_R - y_R\cos\theta_R) \\ \sin\theta_R(x_L\sin\theta_L - y_L\cos\theta_L) - \sin\theta_L(x_R\sin\theta_R - y_R\cos\theta_R) \\ f(\sin\theta_L\cos\theta_R - \sin\theta_R\cos\theta_L) \end{pmatrix}. \tag{5.13}$$

Proposition 5.2.2. *If we define*

$$\tilde{\omega}_{\theta_L} := U_{t_L}, \quad \tilde{\omega}_{\theta_R} := U_{t_R} \tag{5.14}$$

and the corresponding 2-form $\omega_{\mathbb{R}^3} := \tilde{\omega}_{\theta_R} \wedge \tilde{\omega}_{\theta_L}$*, then, up to a multiplicative factor, the following equalities hold:*

$$\tilde{\omega}_{\theta_L} = \omega_{\theta_L}, \quad \tilde{\omega}_{\theta_R} = \omega_{\theta_R}, \quad \omega_{\mathbb{R}^3} = \omega_{bin}. \tag{5.15}$$

Proof. Let us recover the expression of the 1-forms $\tilde{\omega}_{\theta_L} := U_{t_L}$ and $\tilde{\omega}_{\theta_R} := U_{t_R}$. Recall here the change of variable (3.5):

$$\begin{cases} r_1 = \frac{xc}{d} \\ r_2 = \frac{yc}{d} \\ r_3 = \frac{fc}{d} \end{cases}, \tag{5.16}$$

and its differential:

$$\begin{cases} \mathrm{d}\, r_1 = \frac{c}{d}\,\mathrm{d}\, x - \frac{cx}{d^2}\,\mathrm{d}\, d \\ \mathrm{d}\, r_2 = \frac{c}{d}\,\mathrm{d}\, y - \frac{cy}{d^2}\,\mathrm{d}\, d \\ \mathrm{d}\, r_3 = -\frac{fc}{d^2}\,\mathrm{d}\, d \end{cases}. \tag{5.17}$$

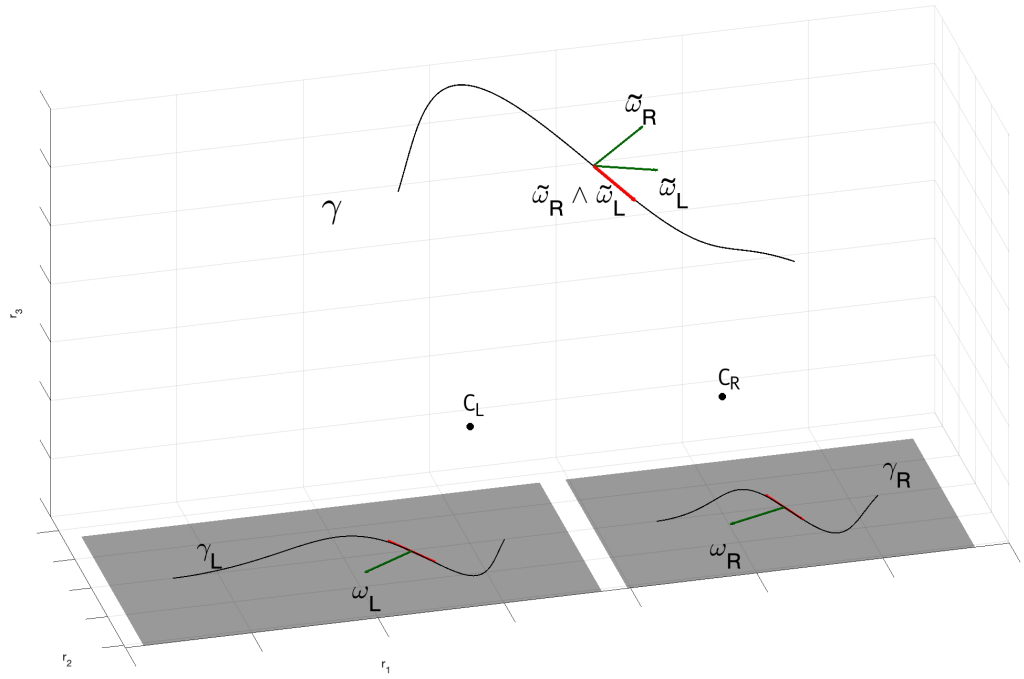


Figure 5.2.3: Three-dimensional reconstruction of the space from retinal planes. The one-dimensional form ω_{θ_L} and ω_{θ_R} are identified with the normal to the curves γ_L and γ_R. Their three dimensional counterpart $\tilde{\omega}_{\theta_L}$ and $\tilde{\omega}_{\theta_R}$ identify the tangent vector to the curve $\gamma : \mathbb{R} \to \mathbb{R}^3$ by the wedge product $\tilde{\omega}_{\theta_L} \wedge \tilde{\omega}_{\theta_R}$.

Writing the quantity U_{t_L} of equation (5.12), in term of a 1-form in the variables (r_1, r_2, r_3), we have:

$$\tilde{\omega}_{\theta_L} = -f \sin\theta_L \, \mathrm{d}\, r_1 + f \cos\theta_L \, \mathrm{d}\, r_2 + (x_L \sin\theta_L - y \cos\theta_L) \, \mathrm{d}\, r_3. \tag{5.18}$$

Changing coordinates:

$$\begin{aligned} \tilde{\omega}_{\theta_L} =& -f \sin\theta_L \left(\frac{c}{d} \, \mathrm{d}\, x - \frac{cx}{d^2} \, \mathrm{d}\, d \right) + f \cos\theta_L \left(\frac{c}{d} \, \mathrm{d}\, y - \frac{cy}{d^2} \, \mathrm{d}\, d \right) \\ & + (x_L \sin\theta_L - y \cos\theta_L) \left(-\frac{fc}{d^2} \, \mathrm{d}\, d \right) \\ =& \frac{fc}{d} \left(-\sin\theta_L \, \mathrm{d}\, x + \cos\theta_L \, \mathrm{d}\, y - \sin\theta_L \, \mathrm{d}\, d \right) \\ =& \frac{fc}{d} \omega_{\theta_L}. \end{aligned} \tag{5.19}$$

So, up to the multiplicative factor $\frac{fc}{d}$, we have that $\tilde{\omega}_{\theta_L} = \omega_{\theta_L}$ in the variables

(x, y, d). The same reasoning holds for the right structure. □

In this way, the disparity binocular cells couple in a natural way positions, identified with points in $\mathbb{R}^3$, and orientations in $\mathbb{S}^2$, identified with three-dimensional unitary tangent vectors. As already observed in subsection 5.2.3, the geometry of the stereo vision is not solved only with these punctual and directional arguments, but there is the need to take into account suitable types of connections. In [AZ00, LZ03, LZ06], Zucker et al. proposed a model that considered the curvature of monocular structures as an additional variable. Instead, we propose to consider simple monocular cells selective for orientation and to insert the notion of curvature directly into the definition of connection. It is therefore natural to introduce the perceptual space via the manifold $\mathbb{R}^3 \rtimes \mathbb{S}^2$, and look for appropriate curves.

5.3 A perceptual model in the space of 3D position-orientation $\mathbb{R}^3 \rtimes \mathbb{S}^2$

We have just clarified that binocular cells are parametrized by points in $\mathbb{R}^3$, and orientations in $\mathbb{S}^2$. An element ξ of the space $\mathbb{R}^3 \rtimes \mathbb{S}^2$ it is defined by a point $p = (p_1, p_2, p_3)$ in $\mathbb{R}^3$ and an unitary vector $n \in \mathbb{S}^2$. It is usual to define the 2 dimensional sphere as the set of points in $\mathbb{R}^3$ having (Euclidean) distance from the origin equal to 1, formally $\mathbb{S}^2 = \{n \in \mathbb{R}^3; \|n\| = 1\}$. Since the topological dimension of this geometric object is 2, we introduce the classical spherical coordinates (θ, φ) such that $n = (n_1, n_2, n_3) \in \mathbb{S}^2$ can be parametrized as:

$$\begin{aligned} n_1 &= \cos\theta \sin\varphi \\ n_2 &= \sin\theta \sin\varphi \\ n_3 &= \cos\varphi \end{aligned} \tag{5.20}$$

with $\theta \in [0, 2\pi]$ and $\varphi \in (0, \pi)$. The ambiguity that arises using a local coordinate chart is overcome by the introduction of a second chart (for example, see Remark A.1.1, in Appendix A), covering the singular points.

Since $\mathbb{S}^2$ is not a Lie group (Hairy Ball Theorem), we cannot set a product on $\mathbb{R}^3 \rtimes \mathbb{S}^2$. However, we can use the Euclidean group of rigid body motions in dimension three, $SE(3)$, and consider a group action. An element $g = (p, R)$ in $SE(3)$ is defined by a point $p \in \mathbb{R}^3$ and a rotation $R \in SO(3)$. The group law expresses the fact that the composition of two body motions is again a body

motion, and so for two elements $g = (p, R)$ and $g' = (p', R')$ in $SE(3)$ their product is defined as

$$gg' = (p + Rp', RR'). \tag{5.21}$$

The left group action L_g of $g \in SE(3)$ onto $(q, n) \in \mathbb{R}^3 \rtimes \mathbb{S}^2$ is:

$$L_g(q, n) = (Rq + p, Rn). \tag{5.22}$$

It is then clear that the group product (5.21) influences the group action (5.22) in $\mathbb{R}^3 \rtimes \mathbb{S}^2$; for this reason then we use the notation $\rtimes$ even though this is usually reserved for the semi-direct product of groups. For further information regarding the relationships between $SE(3)$ and $\mathbb{R}^3 \rtimes \mathbb{S}^2$ from a differential point of view, see Duits and Franken in [DF11] and also Duits, Creusen, Ghosh and Dela Haije in [DCGDH11], with application to HARDI/DTI images. In the following, we adapt these mathematical results to apply them to the problem of stereo perception. More details about these mathematical spaces can be found in Appendix A.

5.3.1 Stereo sub-Riemannian geometry

The emergence of a privileged direction in $\mathbb{R}^3$, associated with the tangent vector to the three-dimensional curvilinear stimulus, is the reason why we endow $\mathbb{R}^3 \rtimes \mathbb{S}^2$ with a sub-Riemannian structure that favors the direction in 3D identified by the 2-form ω_{bin}.

To do so, we consider the following basis for the tangent space

$$\partial_1, \ \partial_2, \ \partial_3, \ \partial_\theta, \ \partial_\varphi. \tag{5.23}$$

and we move them with the push-forward of the group action (5.22): $Y_i|_g \phi = (L_g)_* \partial_i \phi = \partial_i(\phi \circ L_g)$, with ϕ smooth function. Noticing that $R \in SO(3)$ can be decomposed as a product of rotations around the Cartesian axes (Euler coordinates in (A.1)), and choosing $g \in SE(3)$ as the element defined by a rotation $R \in SO(3)$ with the first rotation around the third axis equal to zero, the push forwarded

vectors can be expressed in the coordinate chart as:

$$\begin{aligned} Y_1 &= \cos\varphi\cos\theta\partial_1 + \cos\varphi\sin\theta\partial_2 - \sin\varphi\partial_3 \\ Y_2 &= -\sin\theta\partial_1 + \cos\theta\partial_2 \\ Y_3 &= (\sin\varphi\cos\theta)\partial_1 + (\sin\varphi\sin\theta)\partial_2 + \cos\varphi\partial_3 \\ Y_\theta &= -\frac{1}{\sin\varphi}\partial_\theta \\ Y_\varphi &= \partial_\varphi \end{aligned} \tag{5.24}$$

for $\varphi \neq 0, \varphi \neq \pi$.

We define the distribution $\mathcal{D}$ on $\mathbb{R}^3 \rtimes \mathbb{S}^2$ in such a way that it encodes the coupling between positions and orientations. As proved by Duits and Franken in [DF11], $\mathcal{D}$ is well-defined and a local frame in the coordinate chart is given by:

$$\mathcal{D} = \text{span}\{Y_3, Y_\theta, Y_\varphi\}. \tag{5.25}$$

In particular, the vector field Y_3 identifies the privileged direction in $\mathbb{R}^3$, while Y_θ and Y_φ involve just orientation variables of $\mathbb{S}^2$. The commutation rules of the vector fields $\{Y_3, Y_\theta, Y_\varphi\}$ make them verify the *Hörmander condition*, defining a sub-Riemannian structure if we also choose a metric on $\mathcal{D}$. The metric is not unique and the simplest choice is the metric which makes the vector fields $\{Y_3, Y_\theta, Y_\varphi\}$ orthonormal. More precisely, the metric that we consider restricted to $\mathbb{S}^2$ is the classical Riemannian metric of the sphere, induced by the Euclidean metric of $\mathbb{R}^3$. The reader interested in more details can refer to Appendix A.

5.3.2 Change of variables

We have set the sub-Riemannian model in $\mathbb{R}^3 \rtimes \mathbb{S}^2$ starting from a cortical parametrization involving position and orientation disparity. We have already expressed the change of variable in the variables (x, y, d) to (r_1, r_2, r_3) in equations (3.5). However, the cortical coordinates also contain the angular variables θ_R and θ_L related in the perceptual space to the spherical angles θ, φ.

To identify a change of variable among these variables, we first introduce the function $(r_1, r_2, r_3, \theta, \varphi) \xrightarrow{F} (x, y, d, \theta_L, \theta_R)$:

$$
\begin{array}{rccc}
F: & \mathbb{R}^3 \rtimes \mathbb{S}^2 & \longrightarrow & \mathbb{R}^3 \times \mathbb{S}^1 \times \mathbb{S}^1 \\
& \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ \theta \\ \varphi \end{pmatrix} & \mapsto & \begin{pmatrix} \frac{fr_1}{r_3} \\ \frac{fr_2}{r_3} \\ \frac{cf}{r_3} \\ \tan^{-1}\left(\frac{r_3 \sin\theta \sin\varphi - r_2 \cos\varphi}{r_3 \cos\theta \sin\varphi - (c+r_1)\cos\varphi}\right) \\ \tan^{-1}\left(\frac{r_3 \sin\theta \sin\varphi - r_2 \cos\varphi}{r_3 \cos\theta \sin\varphi - (c-r_1)\cos\varphi}\right) \end{pmatrix},
\end{array}
\tag{5.26}
$$

expressed locally in chart $\theta \in [0, 2\pi], \varphi \in (0, \pi)$ and where the retinal right and left angles $\theta_R = \tan^{-1}\left(\frac{r_3 \sin\theta \sin\varphi - r_2 \cos\varphi}{r_3 \cos\theta \sin\varphi - (c+r_1)\cos\varphi}\right)$ and $\theta_L = \tan^{-1}\left(\frac{r_3 \sin\theta \sin\varphi - r_2 \cos\varphi}{r_3 \cos\theta \sin\varphi - (c-r_1)\cos\varphi}\right)$ are obtained considering equation (3.6).

Remark 5.3.1. *Three-dimensional stimuli lying in fronto parallel planes project in retinal planes with the same orientation. The angle $\varphi = k\frac{\pi}{2}$ with $k \in \mathbb{N}$ identifies fronto-parallel planes, and substituting this value in (5.26), it follows that $\theta_L = \theta_R = \theta$.*

Analogously, it is possible to define the change of variable $(x, y, d, \theta_L, \theta_R) \xrightarrow{G} (r_1, r_2, r_3, \theta, \varphi)$:

$$
\begin{array}{rccc}
G: & \mathbb{R}^3 \times \mathbb{S}^1 \times \mathbb{S}^1 & \longrightarrow & \mathbb{R}^3 \rtimes \mathbb{S}^2 \\
& \begin{pmatrix} x \\ y \\ d \\ \theta_R \\ \theta_L \end{pmatrix} & \mapsto & \begin{pmatrix} \frac{cx}{d} \\ \frac{cy}{d} \\ \frac{cf}{d} \\ \tan^{-1}\left(\frac{2\sin\theta_R \sin\theta_L}{\sin(\theta_R+\theta_L)}\right) \\ \tan^{-1}\left(\frac{\sqrt{\sin^2(\theta_R+\theta_L) + 4\sin^2\theta_R \sin^2\theta_L}}{\sin(\theta_R-\theta_L)}\right) \end{pmatrix}.
\end{array}
\tag{5.27}
$$

The angles $\theta = \tan^{-1}\left(\frac{2\sin\theta_R \sin\theta_L}{\sin(\theta_R+\theta_L)}\right)$ and $\varphi = \tan^{-1}\left(\frac{\sqrt{\sin^2(\theta_R+\theta_L) + 4\sin^2\theta_R \sin^2\theta_L}}{\sin(\theta_R-\theta_L)}\right)$ are obtained considering that $\tan\theta = \frac{(\dot{\gamma})_2}{(\dot{\gamma})_1}$ and $\tan\varphi = \frac{\sqrt{(\dot{\gamma})_1^2 + (\dot{\gamma})_2^2}}{(\dot{\gamma})_3}$, applying these relations to the spatial components of vector field Y_3.

Remark 5.3.2. *In accordance with Remark 5.3.1, projections on retinal planes in which the tangents are characterized by angles $\theta_L = \theta_R$ correspond to three dimensional stimuli lying on fronto-parallel planes, identified with an integer multiple of $\varphi = \frac{\pi}{2}$.*

5.3.3 Integral curves

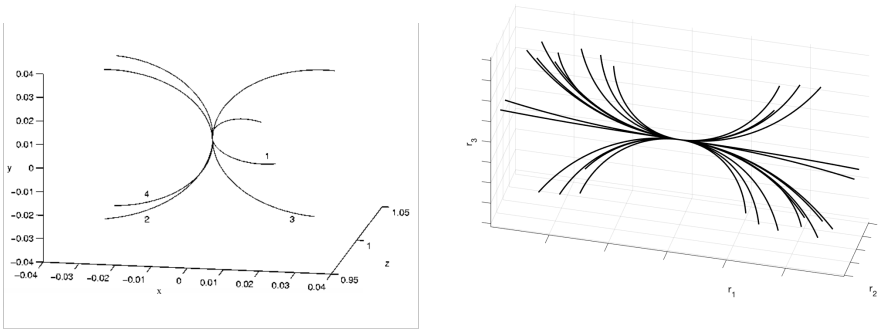
We will assume that the connectivity of the space is described by horizontal curves of the distribution $\mathcal{D}$, and thanks to the Hörmander condition it defines a metric. According to definition (4.3.2), a curve $\Gamma : [0, T] \longrightarrow \mathbb{R}^3 \rtimes \mathbb{S}^2$ is said to be *admissible* (or horizontal) if :

$$\dot{\Gamma}(t) \in \mathcal{D}_{\Gamma(t)}, \leftrightarrow \dot{\Gamma}(t) = a(t)Y_{3,\Gamma(t)} + b(t)Y_{\theta,\Gamma(t)} + c(t)Y_{\varphi,\Gamma(t)}, \tag{5.28}$$

where a, b, c are sufficiently smooth functions on $t \in [0, T]$. We consider the particular case of constant coefficient admissible curves, with $a(t) = 1$, since the vector field Y_3 represents the tangent direction of the stimulus and so it never vanishes:

$$\dot{\Gamma}(t) = Y_{3,\Gamma(t)} + c_1 Y_{\theta,\Gamma(t)} + c_2 Y_{\varphi,\Gamma(t)}, \tag{5.29}$$

with c_1 and c_2 varying in $\mathbb{R}$. An example of the fan of integral curves with constant coefficients satisfying (5.29) is shown in Figure 5.3.1, where a strong resemblance with the fan of helices proposed in [AZ00] can be observed.



(a) (b)

Figure 5.3.1: Fans of integral curves proposed for the geometry of the stereo vision. (a) Fan of helices from the work of Zucker et al. [AZ00]. (b) Fan of sub-Riemannian integral curves defined by equation (5.29) with varying c_1 and c_2 in $\mathbb{R}$.

These curves can be thought of in terms of trajectories in $\mathbb{R}^3$ describing a movement in the Y_3 direction, which can eventually change according to Y_θ and Y_φ. It is worth noting that in the case described by coefficients c_1 and c_2 equal to zero, the 3D trajectories would be straight lines in $\mathbb{R}^3$; by varying the coefficients c_1 and c_2 in $\mathbb{R}$, we allow the integral curves to follow curved trajectories, twisting

and bending in all space directions.

Formally, the amount of "twisting and bending" in space is measured by introducing the notions of curvature and torsion. We then investigate how these measurements are encoded in the parameters of the family of integral curves, and what constraints have to be imposed to obtain different typologies of curves.

Remark 5.3.3. *The 3D projection of the integral curves (5.29) will be denoted γ and satisfy $\dot{\gamma}(t) = (\cos\theta(t)\sin\varphi(t), \sin\theta(t)\sin\varphi(t), \cos\varphi(t))^T$. Classical instruments of differential geometry let us compute the curvature and the torsion of the curve $\gamma(t)$:*

$$\begin{aligned} k &= \sqrt{(\dot{\varphi})^2 + \sin^2\theta(\dot{\theta})^2}, \\ \tau &= \frac{1}{k^2}(-\cos\varphi\sin^2\varphi(\dot{\theta})^3 - \sin\varphi\dot{\varphi}\ddot{\theta} + \dot{\theta}(-2\cos\varphi(\dot{\varphi})^2 + \sin\varphi\ddot{\varphi})). \end{aligned} \tag{5.30}$$

Using the explicit expression of the vector fields Y_θ and Y_φ in equation (5.29), we get

$$\dot{\theta} = -\frac{c_1}{\sin\varphi}, \qquad \dot{\varphi} = c_2, \tag{5.31}$$

from which it follows that:

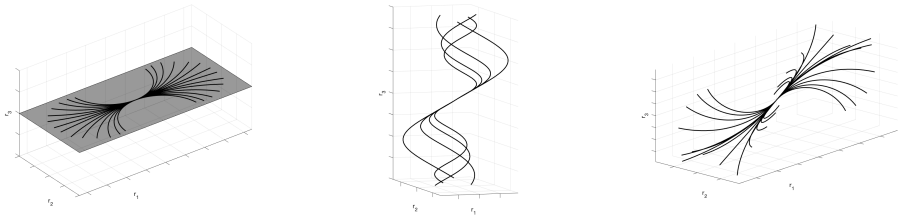
$$\begin{aligned} k &= \sqrt{c_1^2 + c_2^2} \\ \tau &= \frac{c_1^2 - c_2^2}{k^2} c_1 \operatorname{cotan}\varphi. \end{aligned} \tag{5.32}$$

Proposition 5.3.1. *By varying the parameters c_1 and c_2, or equivalently imposing conditions on the angles θ and φ of the integral curves (5.29), we have:*

1. *If $\varphi = \frac{\pi}{2}$ then $k = \sqrt{c_1^2}, \tau = 0$, and so the family of curves (5.29) are circles of radius $1/c_1^2$ on the fronto-parallel plane $r_3 = cost$.*
2. *If $\varphi = \varphi_0$, with $\varphi_0 \neq \pi/2$, then $k = \sqrt{c_1^2}$ and $\tau = c_1 \operatorname{cotan}\varphi_0$, and so the family of curves (5.29) are r_3-helices.*
3. *If $\theta = \theta_0$ then $k = \sqrt{c_2^2}$, $\tau = 0$, and so the family of curves (5.29) are circles of radius $1/c_2^2$ in the osculating planes.*
4. *If $c_1 = \pm c_2$ then $\tau = 0$, and so the family of curves (5.29) are circles of radius $1/c_2^2$ in the osculating planes.*

Proof. The computation follows immediately from the values of curvature and torsion (5.32) and classical results of differential geometry. □

This result allows us to naturally go back to the work of the Zucker group on the problem of stereovision. Indeed, the second point of Proposition 5.3.1 upholds the conjecture formulated by Alibhai and Zucker in [AZ00], proposing helices with spiral along the depth direction to generalize the concept of co-circularity in dimension 3. Moreover, the following remark show compatibility with the results presented in [LZ03].



(a) (b) (c)

Figure 5.3.2: Different types of curves generated by (5.29). (a) Arc of circles for $\varphi = \pi/2$. (b) r_3-helices for $\varphi = \pi/3$. (c) Family of curves with constant curvature k and varying torsion parameter.

Remark 5.3.4. *If the value of the curvature k is fixed, and one parameter, suppose c_2, is free in the definition of the integral curves (5.29), then we are in the setting of Proposition 3.3.1. In fact, the coefficient c_1 is obtained by imposing $c_1 = \pm\sqrt{k^2 - c_2^2}$, and in particular the component that remains to be determined is the torsion.*

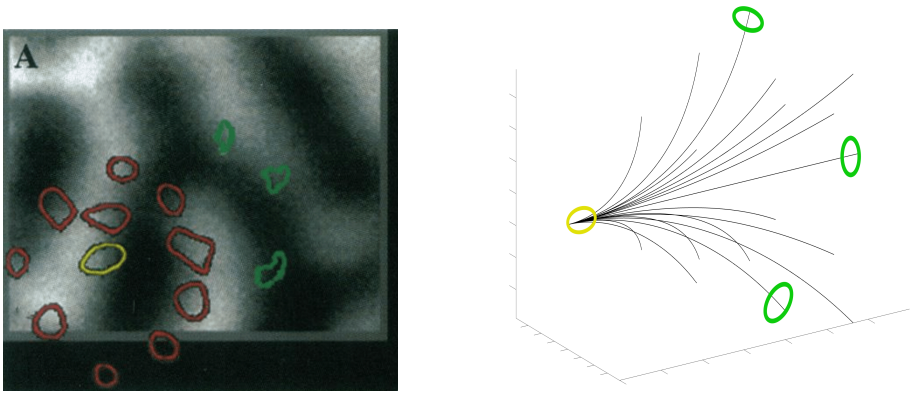
Examples of particular cases of the integral curves (5.29) according to Proposition 5.3.1 and Remark 5.3.4 are visualized in Figure 5.3.2.

5.4 Comparison with experimental data

In this section, we present results of compatibility between the proposed sub-Riemannian model and biological and psychophysical phenomena present in literature. The integral curve viewpoint suggests underlying biological circuits by providing the direction of information flow. Moreover, the integral curves formalize association field models: their parameters describe the spray of curves that is well in accordance with 3D curves as studied in psychophysical experiments in [HF95, HHK97, KHK16].

5.4.1 Biological connections

The foundation for building our sub-Riemannian model of stereo was a sub- Riemannian model of curve continuation. This was motivated by the orientation column at each position, and the connections between cells in nearby columns. These connections were, in turn, a direct model of the long-range horizontal connections in visual cortex, for which there is biological data (e.g. [BZSF97]). Aspects of the cortical architecture that support binocular processing are illustrated in 2.1.2.4. Although the inputs from each eye are organized into ocular dominance bands, there is no direct evidence for "stereo columns" analogous to the monocular orientation columns. But, there is evidence of long-range connections between binocular cells, and our model informs, concretely, what information should be carried by these long range connections. Thus, an organization for stereo is suggested, but it is implicit in the architecture. Nevertheless, there is evidence in support of it.



(a) (b)

Figure 5.4.1: (a) A biocytin injection superimposed on a map of ocular dominance columns, image result from the work in [MAHG93]. Binocular zones are in the middle of monocular zones (coded in black and white). Starting from the injection site (yellow circle in the center of a binocular zone) the patches' propagation (red corresponds to dense while green to sparsely labeled) tends to avoid highly monocular sites, bypassing the centers of ocular dominance columns, and are located in binocular zones. (b) $3D$ interpretation of the physiological image (a).

Just as information propagates to enforce monocular curve continuation, the binocular signal propagates to form a coherent binocular representation. The Grinwald group established this for stereo [MAHG93] (see also Figure 5.4.1 image (a)), using biocytin injections, that propagate directly along neuronal processes and are deposited at excitatory synapses. Thus, this technique demonstrates the

presence of long-range connections between binocular cells. These results were refined, more recently, by the Fitzpatrick group [STR+22], using in vivo calcium imaging. As shown in Fig. 5.1.1 image (c) the authors demonstrated both the monocular and the binocular inputs for stereo, and the dependence on orientations.

More precisely, [MAHG93] showed selective anisotropic connectivity among binocular regions: the biocytin tracer does not spread uniformly, but rather is highly directional with distance from the injection point.[1] Putting this together with [STR+22], we interpret the anisotropy as being related to (binocular) orientation ([STR+22]), which is exactly the behavior of the integral curves of our vector fields. Our 3D association fields are strongly directional, and information propagates preferentially in the direction of (the starting point of) the curve. An example can be seen in Figure 5.4.1, image (b), where the fan of integral curves (5.29) is represented, superimposed with colored patches, following the experiment proposed in [MAHG93].

5.4.2 Psychophysics and association fields

The connections described by the integral curves in our model can be related to the geometric relationships from psychophysical experiments on perceptual organization of oriented elements in $\mathbb{R}^3$. The goal here is to establish that our connections serve as a generalization of the concept of an association field in 3D.

5.4.2.1 Towards a notion of *association field* for 3D contours

The perception of continuity between two elements of position-orientation in $\mathbb{R}^3$ has been studied experimentally: in the works of Kellman, Garrigan, Shipley et al. [KGS05a, KGS+05b], by introducing a theory called *3D relatability*; in the works of Hess and colleagues in [HF95, HHK97] by extending the good regularity/continuity in depth; and the same topic has more recently been addressed by Deas-Wilcox and Khuu-Honson-Kim in [DW15, KHK16].

5.4.2.1.1 3D relatability Kellman, Garrigan, Shipley [KGS05a, KGS+05b] introduce a theory called *3D relatability* extending to 3D the fundamental work of Field, Heyes and Hess [FHH93], and determining the geometrical affinities between orientations under which a pair of position-orientation elements are perceived as connected in a 3D scene.

[1] This was the case with monocular biocytin injections as well.

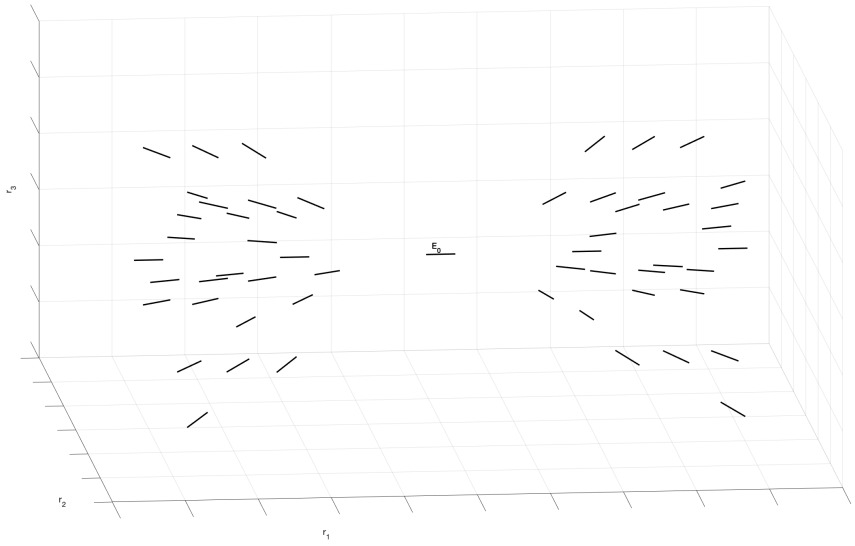


Figure 5.4.2: Example of the fan of the 3D relatable edges with initial point E_0.

Particularly, in a system of 3D Cartesian coordinates, it is possible to introduce an oriented edge E, defining the application point (r_1, r_2, r_3) and its orientation identified with the angles θ and φ. This orientation can be read through the direction expressed by the vector $(\cos\theta \sin\varphi, \sin\theta \sin\varphi, \cos\varphi)^T$. Consider an initial edge E_0, with application point on the origin of the coordinate system $(0, 0, 0)$ and orientation lying on the r_1-axis, described by $\theta = 0, \varphi = \pi/2$. The range of possible orientations (θ, φ) [2] for 3D-relatable edges with E_0 is given by:

$$\tan^{-1}\left(\frac{r_2}{r_1}\right) \leq \theta \leq \frac{\pi}{2} \quad \text{and} \quad \frac{\pi}{2} \leq \frac{3\pi}{2} - \varphi \leq \tan^{-1}\left(\frac{r_3}{r_1}\right). \tag{5.33}$$

The bound of these equations identified with the quantity $\frac{\pi}{2}$ incorporate the 90 degrees constraint in three dimensions, while the bounds defined by the inverse of the tangent express the absolute orientation difference between the reference edge E_0 and an edge positioned at the arbitrary oriented point $E_{(r_1,r_2,r_3)}$ so that its linear extension intersects E_0, see [KGS05a, KGS+05b] for further details.

In Figure 5.4.2 we visually represent an example of 3D positions and orientations that meet the 3D relatable criteria. Starting from initial edge E_0 applied

[2] The angle φ here has been modified to be compatible with our set of coordinates. The relationship between the angle $\tilde{\varphi}$ in works [KGS05a, KGS+05b] can be expressed as : $\tilde{\varphi} = \text{acos}(\sin\varphi) + \pi$.

in (p_{01}, p_{02}, p_{03}) and orientation on the e_1- axis, we display for an arbitrary point (p_1, p_2, p_3) the limit of the relatable orientation (θ, φ).

5.4.2.1.2 Good continuation properties in 3D Psychophysical studies, see [DW15, HF95, HHK97], have investigated the properties of the curves that are suitable for connecting these relatable points. These curves are well described by being smooth and monotonic. In particular, using non-oriented contour elements for contours, Hess et al. in [HHK97] indicate that contour elements can be effectively grouped based primarily on the good continuation of contour elements in depth. This statement is confirmed by the more recent work of Deas and Wilcox [DW15], who in addition observe that detection of contours defined by regular depth continuity is faster than detection of discontinuous contours. All these results support the existence of depth grouping operations, arguing for the extension of Gestalt principles of continuity and smoothness in three-dimensional space.

Finally, on the relationship of the three-dimensional curves to 2-dimensional association fields, see [KGS+05b, KHK16]. These authors have assumed that the strength of the relatable edges in the co-planar planes of E_0 must meet the relations of the bi-dimensional association fields of [FHH93].

5.4.2.2 Compatibility with the sub-Riemannian model

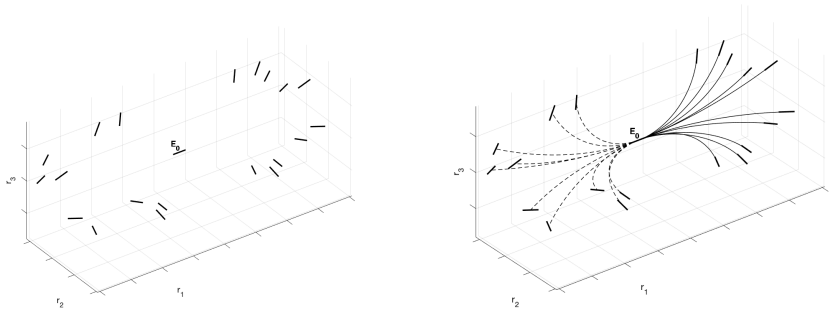
To model associations underlying the 3D perceptive organization previously introduced, we consider a family of integral curves (with constant coefficients) generated from the local orthonormal frame (5.25) of the distribution $\mathcal{D}$:

$$\dot{\Gamma}(t) = Y_{3,\Gamma(t)} + c_1 Y_{\theta,\Gamma(t)} + c_2 Y_{\varphi,\Gamma(t)}. \tag{5.34}$$

These curves locally connect the association fan generated by the geometry of 3D relatability. In particular, image (b) of Figure 5.4.3 shows the family of the horizontal curves connecting the initial point E_0 with 3D relatable edges.

In analogy with the experiment of Field, Hayes and Hess in [FHH93], we choose to represent non-relatable edges to the left of the starting point E_0, while on the right 3D relatable edges. So, filled lines of the integral curves indicate the correlation between the central horizontal element E_0 and the ones on its right, while dotted lines connect the starting point E_0 with elements not correlated with it, as represented on the left part of the image.

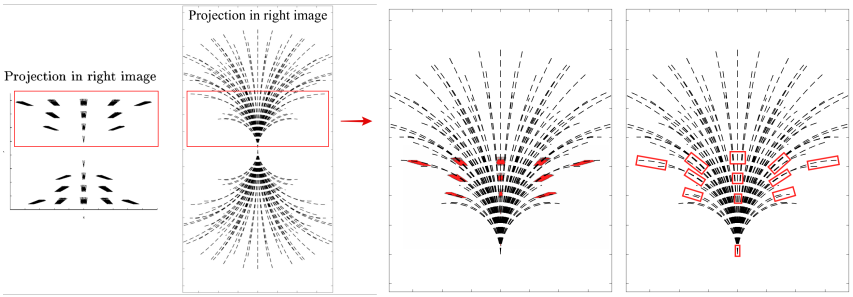
Points along the integral curves in $\mathbb{R}^3 \rtimes \mathbb{S}^2$ which satisfy the 3D relatability



(a) (b)

Figure 5.4.3: 3D relatable points and integral curves. (a) 3D relatable edges are displayed on the right of the initial edge E_0. Unrelatable 3D edges are displayed on the left. (b) Horizontal integral curves with filled lines connect 3D relatable edges with initial point E_0. Horizontal integral curves with dotted lines do not connect 3D unrelatable edges.

condition are then projected into the retinal planes. We obtain a picture that is in agreement with the projections of the stereo compatibility fields in retinal images, introduced by Alibhai and Zucker in [AZ00], as shown in Figure 5.4.4. In particular, image (a) displays the two right retinal projections (we omit the very similar results for the left structures) of the position-orientation elements. The superimposition of the upper part of the compatibility field with the upper



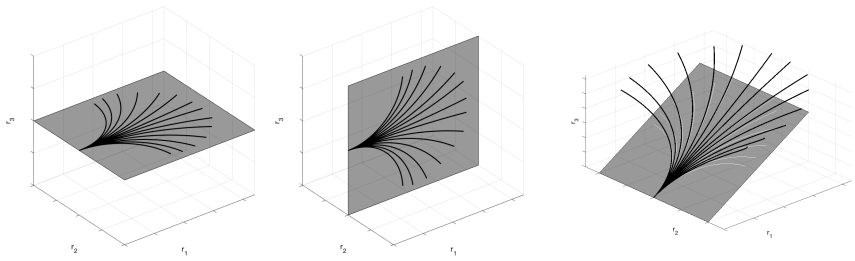
(a) (b) (c)

Figure 5.4.4: (a) Bidimensional projections of the compatibility fields (left, image adapted from [AZ00]) and of the 3D relatable points (right). (b)Superimposition of the compatibility field on the 3D relatable points projection. (c) Emphasizing the involved region on the 3D relatable points projection.

part of the 3D relatable points (Figure 5.4.4 image (b)) allows the selection of suitable points (image (c) of Figure 5.4.4) to analyze the existence of common

patterns. In particular, by remaining vertically aligned with the starting point (traveling along the y axis), the number of admissible orientations increases as the ordinate coordinate grows. Moreover, in the examined regions the directions that are taken into consideration are comparable, and in general, it is observed that these investigated orientations lead to the global idea of co-circularity.

Restricting the integral curves on the neighborhood of co-planar planes with an arbitrary edge E, we capture behaviors compatible with the 2D case. More precisely, on the r_1-r_2 plane (fronto-parallel) and the r_1-r_3 plane we have arc of circles, in accordance with Proposition 5.3.1. Furthermore, for an arbitrary plane in $\mathbb{R}^3$ containing an edge E, we observe that the curves generated with fixed angle φ are helices, locally satisfying the bi-dimensional constraint. Examples can be found in Figure 5.4.5. In particular, the curves displayed in images (a) and (b) of Figure 5.4.3 are well in accordance with the curves of the Citti-Sarti model, depicted in Figure 4.4.2.



(a) (b) (c)

Figure 5.4.5: Compatibility with bidimensional constraint. (a) Restriction of the fan of the integral curves on the e_1-e_2 plane. (b) Restriction of the fan of the integral curves on the e_1-e_3 plane. (c) Restriction of the fan at $\varphi = \varphi_0$. These curves (black lines) are not planar curves but helices. However, their projection (white lines) on the coplanar plane with the initial edge satisfies the bidimensional constraints.

Finally, observing the behavior of the fan (5.34) on a three-dimensional curvilinear stimulus, we can see in Figure 5.4.6 how the proposed 3D association field envelopes a space curve, in the same way that a 2D association field envelopes a planar curve. The integral curve viewpoint does provide a direction of information flow: curvature, implicitly described through coefficients c_1 and c_2, supplies a kind of glue that enable transitions from points to nearby points.

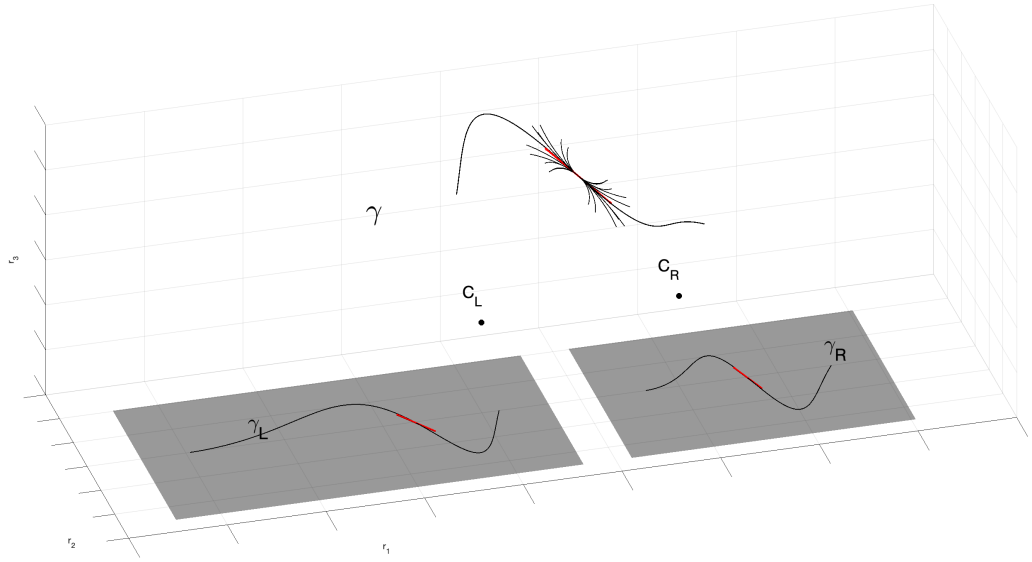


Figure 5.4.6: The three-dimensional space curve γ is enveloped by the $3D$ the association field centered at a point. Formally, this association field is a fan of integral curves in the sub-Riemmanian geometry described by equation (5.34) with varying c_1 and c_2 in $\mathbb{R}$.

As in all models, it is possible to admit alternative viewpoints. One of these, not covered in this thesis, involves the geodesics of the sub-Riemannian structure. These curves have already been proposed in the two-dimensional case, by the authors of [BDRS14, DBRS13], as a model for the geometry of 2D relatability (and the two-dimensional association fields). Geodesics of the sub-Riemannian structure of position-orientation in 3D have been the object of interesting papers, such as [DGDHM16, DMMP18]. In these works, the authors not only characterize geodesics with spatial cuspless projections but also their admissible boundary conditions, which, depending on the point, could be used to involve all orientational end-conditions according to the theory of 3D relatability.

5.4.3 Integration of contours and the stereo correspondence problem

The notion of 3D association field helps to understand how the brain integrates contours to perceive a three-dimensional stimulus from images captured by the eyes. We can see how the geometry we propose is a good starting point to understand how to match left and right points and features.

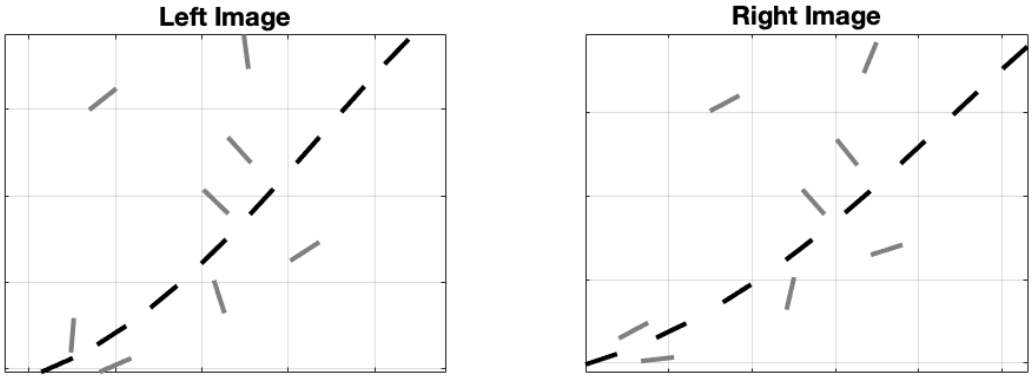


Figure 5.4.7: Left and right retinal images of the set Ω. Black points are the projection of the point of the curve γ, while gray points are background random noise.

Inspired by the experiment of Hess and Field in [HF95], we consider a path stimulus γ interpreted as a contour, embedded in a background of randomly oriented elements: left and right retinal visual stimuli are depicted in Figure 5.4.7. We perform a first simplified lift of the retinal images to a set Ω subset of $\mathbb{R}^3 \rtimes \mathbb{S}^2$. This set contains all the possible corresponding points, obtained by coupling left and right points which share the same y retinal coordinate, see Figure 5.4.8. The set Ω contains false matches, namely points that do not belong to the original stimulus. It is the task of correspondence to eliminate these false matches.

We compute for every lifted point the binocular output O_B of equation (3.10). This output can be seen as a probability measure that gives information on the correspondence of the couple of left and right points. We can simply evaluate which are the points with the highest probability of being in correspondence, applying a process of suppression of the non-maximal pairs over the fiber of disparity. In this way, noise points are removed (Figure 5.4.9, image (a)).

We now directly exploit good continuation in depth. The remaining noise elements are orthogonal to the directions of the elements of the curve that we would like to reconstruct. Calculating numerically the coefficients c_1 and c_2 of integral curves (5.29) that connect all the remaining pairs of points, we can obtain for every pair the value of curvature and torsion using (5.32).

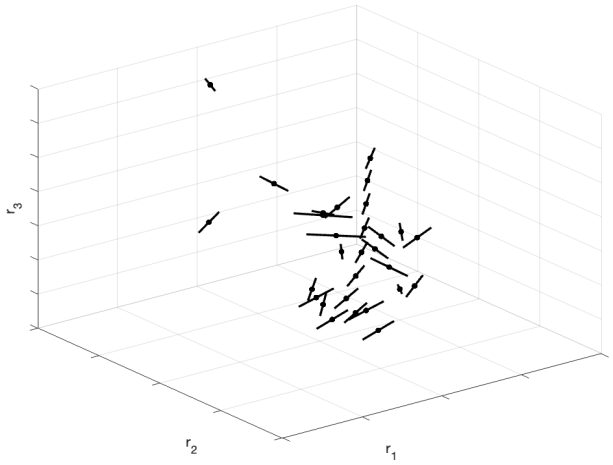
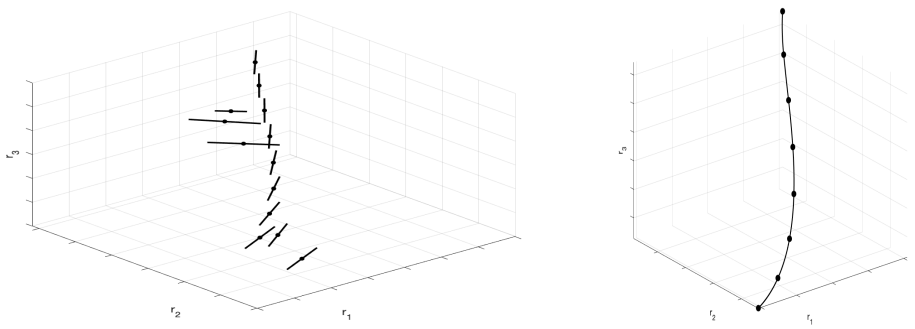


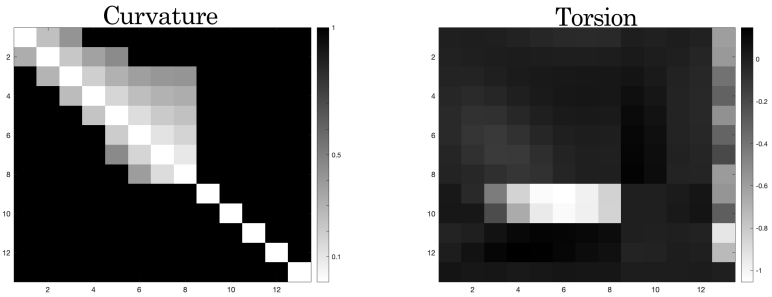
Figure 5.4.8: Lifting of the two left and right retinal images of Figure 5.4.7 in the space of position and orientation $\mathbb{R}^3 \times \mathbb{S}^2$.



(a) (b)

Figure 5.4.9: (a) Selection of lifted points according to the binocular output. (b) Points of the stimulus γ connected by integral curves (5.29).

Figure 5.4.10 read it in terms of matrices M representing the values of curvature or torsion for every couple of points ξ_i, ξ_j in the element M_{ij}. In particular, we observe that random points are characterized by a very high curvature and in general also the torsion deviates from minimum magnitudes. So, by discarding these high values, we select only the three-dimensional points of the curve γ, which are well connected by the integral curves, as shown in image (b) of Figure 5.4.9. This is in accordance with the idea developed in [AZ00, LZ03, LZ06], where curvature and torsion provide constraints for reconstruction in 3D.



(a) (b)

Figure 5.4.10: Matrices M which element M_{ij} represents the value of curvature or torsion for every couple of points ξ_i, ξ_j. The first eight points correspond to the stimulus curve γ while the others are random noise. (a) Curvature matrix. (b) Torsion matrix.

The proposed neuro-geometry of stereo vision allows us to be precise about the type of geometry that is relevant for understanding stereo abstractly, and concretely it is informative toward the physiology: even though a "stereo columnar architecture" is not obvious from the anatomy, it is well-formed computationally, reasonable matching (at a first stage) left and right points and features.

Chapter 6

Individuation of 3D percepts

In this chapter we study the emergence of perceptual units in the three-dimensional scene, starting from the stochastic counterpart of the 3D association fields. These curves model the interactions among binocular cells involved, and we use them to define a connectivity kernel inserted as a facilitator pattern in the activity equation of the population. Then, we see how the stability analysis of this equation leads to the identification of the three-dimensional perceptual units of the scene, through the proposition of a grouping algorithm. Finally, we examine the numerical results.

Contents

6.1 Stochastic model of 3D association fields

The probabilistic version of the association fields is used in the two-dimensional case as a model of connectivity for cortical cells, for example in the works [PT99, SC15, ASFC+17, DF10, SCS10b, BCCS14], based on the pioneering work of Mumford [Mum94]. We apply this approach to the three-dimensional case, probabilistically interpreting the 3D extension of the law of good continuation: entities described by similar local orientations are more likely to belong to the same perceptual unit. We will consider an equation predominated by transport in the Y_3 direction and diffusion in the orientation variables, as done for example in [MS09, MS12]. In general, these equations describe the motion of a particle moving with constant speed in a direction randomly changing accordingly with the stochastic process, and the effect is that particles tend to travel in straight lines, but over time, they drift in space.

6.1.1 Stochastic differential Langevin equation

We consider the trajectory described by a deterministic law given by the vector field Y_3, altered by a certain random noise in the directions Y_θ and Y_φ. The result is a Itô's stochastic process $\Gamma(t)$ satisfying the stochastic differential Langevin equation:

$$\mathrm{d}\,\Gamma(t) = Y_{3,\Gamma(t)}\,\mathrm{d}\,t + \lambda \underbrace{(Y_{\theta,\Gamma(t)}, Y_{\varphi,\Gamma(t)})}_{\sigma(\Gamma(t))}\,\mathrm{d}\,B(t), \tag{6.1}$$

with $B(t) = (B_1(t), B_2(t))^T$ two dimensional Brownian motion, and $\lambda \in \mathbb{R}$ scalar parameter. By abuse of notation, Y_3 is interpreted as a deterministic coefficient and it is called the *drift* coefficient, while σ is the *diffusion* coefficient of the stochastic differential equation (SDE) (6.1). In our case, locally in the chart of subsection

5.3.1, the drift term is given by

$$Y_{3,\Gamma} = (\cos\theta \sin\varphi, \sin\theta \sin\varphi, \cos\varphi, 0, 0)^T , \tag{6.2}$$

and the diffusion coefficient $\sigma(\Gamma) = (Y_{\theta,\Gamma}, Y_{\varphi,\Gamma})$ is

$$Y_{\theta,\Gamma} = (0, 0, 0, -1/\sin\varphi, 0)^T \text{ and } Y_{\varphi,\Gamma} = (0, 0, 0, 0, 1)^T. \tag{6.3}$$

The coefficients written in coordinates are smooth functions with locally bounded derivatives. Therefore, classical results on uniqueness and existence of a strong solution to the SDE (6.1) hold thanks to the fact that the terms of the drift and diffusion coefficients are Lipschitz continuous.

6.1.2 Forward Kolmogorov operator

Itô's formula allows us to formally associate to equation (6.1) its *characteristic operator*, a second order differential operator $\mathcal{L}$ with variable coefficients. Locally in the chart of subsection 5.3.1, $\mathcal{L}$ can be written as (see for example [Øks03]):

$$\mathcal{L} = \sum_{i=1}^{3} (Y_{3,\Gamma})_i \partial_i + \frac{\lambda}{2} \sum_{i,j \in \{\theta,\varphi\}} (\sigma(\Gamma)\sigma(\Gamma)^T)_{i,j} \partial_i \partial_j \tag{6.4}$$

Let's consider a starting point $\Gamma(t_0) = \xi_0$ at a time t_0 and a test function ϕ. After having applied the expected value to Itô's formula, we have

$$\int\int (\partial_t + \mathcal{L})\phi(t,\xi)\rho_\lambda(\xi, t; \xi_0, t_0) d\xi dt = 0, \tag{6.5}$$

where ρ_λ is the density law of the stochastic process Γ, and it is a distributional solution to the forward Kolmogorov (or Fokker-Planck) equation:

$$(\partial_t + \mathcal{L}^*)\rho_\lambda(\xi, t; \xi_0, t_0) = \delta_{\xi-\xi_0}\delta_{t-t_0} \tag{6.6}$$

with δ the delta distribution and $\mathcal{L}^*$ the formal adjoint of $\mathcal{L}$. Since the drift term and the diffusion coefficient are defined as (6.2) and (6.3), the operator $\mathcal{L}^*$ satisfies the following expression in coordinates:

$$\mathcal{L}^* f(t,\xi) = -\sum_{i=1}^{3} (Y_{3,\xi})_i \partial_i f(t,\xi) + \frac{\lambda}{2} \sum_{i,j \in \{\theta,\varphi\}} (\sigma(\xi)\sigma(\xi)^T)_{i,j} \partial_i \partial_j f(t,\xi). \tag{6.7}$$

Remark 6.1.1. *The forward Kolmogorov operator $\mathcal{L}^*$ can be written in terms of vector fields of the distribution $\mathcal{D}$ presented in (5.25):*

$$\mathcal{L}^* = -Y_3 + \lambda(Y_\theta^2 + Y_\varphi^2). \tag{6.8}$$

Since the vector fields satisfy the Hörmander's rank condition, the operator $\mathcal{L}^*$ is hypoelliptic, and so there exists a fundamental solution associated with the forward Kolmogorov operator, with the property of being smooth out of the pole (see Hörmander theorem [Hör67], recall here in Theorem 4.3.2). We will recall in the next subsection that this fundamental solution is none other that the probability density ρ_λ of the forward Kolmogorov equation (6.5), integrated over time.

6.1.3 Time independent kernel

The fundamental solution for the operator $\partial_t + \mathcal{L}^*$, with $\mathcal{L}^*$ defined in (6.8), corresponds to the probability density function $\rho_\lambda(\xi, t; \xi_0, t_0)$ associated with the stochastic process satisfying equation (6.1). More precisely, it expresses the probability of having reached the point ξ after having evolved the equation (6.1) up to time t, starting from ξ_0 at time t_0.

The main goal of this section is to identify a time-independent probability, to characterize each point of the space in terms of the paths (6.1) that reach it independently of the value of the evolution temporal parameter. It is therefore worth noting that these density functions ρ_λ locally admit exponential-type estimates, which decay to zero for very "evolutional" times.

Remark 6.1.2. (Gaussian estimates) *The operator $\mathcal{L}^*$ of (6.6) belongs to a class of hypoelliptic operators that locally admit estimates for their fundamental solution. Rotschild and Stein in [RS76] obtained optimal local estimates of the kernels, while Gaussian upper and lower bounds have been proven in [BLU02, BP07a, CP08].*

In literature, it is common to integrate over time to obtain such a time-independent fundamental solution as shown in [BLU02] or in [Coc14, FCS17]. We recall here this result, which we will adopt throughout the thesis.

Proposition 6.1.1. *Let ξ be a point in $\mathbb{R}^3 \rtimes \mathbb{S}^2$ and ρ_λ the fundamental solution of the operator $\partial_t + \mathcal{L}^*$ with $\mathcal{L}^*$ defined in (6.8), and pole in $(\xi_0, 0)$, with $t_0 = 0$*

without loss of generality. Then

$$J_\lambda(\xi, \xi_0) = \int_{\mathbb{R}_+} \rho_\lambda(\xi, t; \xi_0, 0) dt \tag{6.9}$$

is fundamental solution for the operator $\mathcal{L}^*$*:*

$$\mathcal{L}^* J_\lambda(\xi, \xi_0) = \delta_{\xi - \xi_0}. \tag{6.10}$$

Proof. Let u be a smooth function on $\mathbb{R}^3 \rtimes \mathbb{S}^2$ with compact support, independent from the time variable t. Then:

$$\begin{aligned} -u(\xi_0) &= \int_{\mathbb{R}_+ \times \mathbb{R}^3 \rtimes \mathbb{S}^2} \rho_\lambda(\xi, t; \xi_0, 0)(\partial_t + \mathcal{L}^*) u(\xi) d\xi dt \\ &= \int_{\mathbb{R}_+ \times \mathbb{R}^3 \rtimes \mathbb{S}^2} \rho_\lambda(\xi, t; \xi_0, 0) \mathcal{L}^* u(\xi) d\xi dt \\ &= \int_{\mathbb{R}^3 \rtimes \mathbb{S}^2} \left(\int_{\mathbb{R}_+} \rho_\lambda(\xi, t; \xi_0, 0) dt \right) \mathcal{L}^* u(\xi) d\xi \\ &= \int_{\mathbb{R}^3 \rtimes \mathbb{S}^2} J_\lambda(\xi, \xi_0) \mathcal{L}^* u(\xi) d\xi \end{aligned} \tag{6.11}$$

with $J_\lambda(\xi, \xi_0) = \int_{\mathbb{R}_+} \rho_\lambda(\xi, t; \xi_0, 0) dt$. □

Alternatively, another way to deal with the time parameter is to make a further assumption on the random walk, supposing having exponentially distributed traveling time. Therefore, the time-independent fundamental solution is obtained by taking the Laplace transform of the probability density, as done for example by the authors of [DF11, PD17].

6.2 Monte Carlo simulation of the kernel

Approximation techniques based on Monte Carlo simulation are commonly developed to numerically compute the probability density related to these diffusion processes but also to validate exact solutions obtained analytically, for example in terms of Fourier transform as done in the works [PD17, DBM19] for (convection-) diffusion Kolmogorov equations in the 3D space of position-orientations. In these papers, the authors compare the numerical Fourier methods with the Monte Carlo simulations. Although the approach is the same, we use this simulation with the purpose of locally traveling along the integral curves, assuming the possibility of drifting in space in accordance with some stochastic noise. In this way,

we approximate the fundamental solution (6.9) using a procedure developed in [SC15, BCCS14] adapted to our case. First, a simulation of the stochastic paths is carried out by applying the Euler-Maruyama method, then an appropriate average is made, using the Strong Law of Large Numbers.

6.2.1 Euler-Maruyama's type scheme

In order to generate the discrete version of the stochastic process $\Gamma_t \in \mathbb{R}^3 \rtimes \mathbb{S}^2$, in coordinates $\Gamma_t = (r_1(t), r_2(t), r_3(t), \varphi(t), \theta(t))$, satisfying the Langevin equation (6.1), we choose the Euler-Maruyama scheme approximation: this very simple method allows a convergence rate result very satisfying if $M \to \infty$. For the reader interested in this topic, further details can be found in [GT13].

We fix $T > 0$ and $\mathbb{N} \ni M \geq 1$ number of steps of the path. Then, we construct iteratively $\Gamma^M_{kT/M}$ with $0 \leq k \leq M$, by setting

$$\begin{cases} \Gamma^M_0 & = \Gamma_0, \\ \Gamma^M_{(k+1)T/M} & = \Gamma^M_{kT/M} + Y_3(\Gamma^M_{kT/M})\frac{T}{M} + \lambda\sigma(\Gamma^M_{kT/M})\sqrt{\frac{T}{M}}\delta_k \end{cases} \tag{6.12}$$

where $\sqrt{\frac{T}{M}}\delta_k$ comes from the fact that the Brownian increments $dW_i(t)$ follow the Gaussian law $N(0, \frac{T}{M})$, with zero mean and variance $\frac{T}{M}$, and so $\delta_k = (\delta^k_1, \delta^k_2)$ is defined from the fact that δ^k_i has normal distribution $N(0, 1)$. In coordinates we have the following system:

$$\begin{cases} r_1(k+1) = r_1(k) + \frac{T}{M}\cos\theta(k)\sin\varphi(k) \\ r_2(k+1) = r_2(k) + \frac{T}{M}\sin\theta(k)\sin\varphi(k) \\ r_3(k+1) = r_3(k) + \frac{T}{M}\cos\varphi(k) \\ \theta(k+1) = \theta(k) - \sqrt{\frac{T}{M}}\frac{\lambda\delta^k_1}{\sin\varphi(k)} \\ \varphi(k+1) = \varphi(k) + \sqrt{\frac{T}{M}}\lambda\delta^k_2 \end{cases} \qquad \text{for } k = 0, 1, \ldots, M-1, \tag{6.13}$$

with $\Gamma_0 = (r_1(0), r_2(0), r_3(0), \varphi(0), \theta(0))$ initial point.

Remark 6.2.1. *A similar formulation has been introduced in [DBM19] to compute numerically the fundamental solution associated with the Kolmogorov equation of a diffusion process (the operator is written as sum of squares of vector fields generating the distribution) on the position-orientation space. The proposed discretization [DBM19, eq. (72)] is based on the fact that a point in $\mathbb{S}^2$ can be thought of in*

terms of three-dimensional rotations starting from an initial axis (see Appendix A.2, equation (A.18)). For example, $\mathbb{S}^2 \ni n(\theta,\varphi) = R_3^\theta R_2^\varphi e_3$*, where* e_3 *is the third Cartesian reference axis,* R_3^θ *is the rotation around the third axis of angle* θ *and* R_2^φ *is the rotation around the second axis of angle* φ*. If we recast (6.13) using rotations, we have:*

$$\begin{cases} r(M) = r(0) + \sum_{k=0}^{M-1} \frac{T}{M} n(k+1) \\ n(k+1) = \left(R_3^{\theta_{(k+1)}} R_2^{\varphi_{(k+1)}} \right) e_3 \end{cases} \tag{6.14}$$

where $n(k+1) = (\sin\theta_{(k+1)} \cos\varphi_{(k+1)}, \cos\theta_{(k+1)} \sin\varphi_{(k+1)}, \cos\varphi_{(k+1)}) \in \mathbb{S}^2$*, the point* $r(k) = (r_1(k), r_2(k), r_3(k)) \in \mathbb{R}^3, k = 0, M$*, and*

$$\begin{cases} \theta(k+1) = \theta(k) - \sqrt{\frac{T}{M}} \frac{\lambda \delta_1^k}{\sin\varphi(k)} \\ \varphi(k+1) = \varphi(k) + \sqrt{\frac{T}{M}} \lambda \delta_2^k \end{cases} \tag{6.15}$$

for $k = 0 \dots M-1$*. This differs from the discretization proposed in [DBM19, eq. (72)], where the spherical term* $n(k+1)$ *is computed as a consecutive rotations' product on the index* k*.*

6.2.2 Strong Law of Large Numbers

Let us denote with $\Gamma_{(t,\xi_0,t_0)}$ the stochastic process associated to (6.1) starting from the point ξ_0 at time t_0. Its density law ρ_λ is defined through a probability measure P as

$$\rho_\lambda(\mathcal{S}, t; \xi_0, t_0) := P[\Gamma_{(t,\xi_0,t_0)} \in \mathcal{S}] = \mathbb{E}[\mathbb{1}_{\mathcal{S}}(\Gamma_{(t,\xi_0,t_0)})], \tag{6.16}$$

with $\mathbb{E}$ the average in probability and $\mathcal{S}$ a subset of $\mathbb{R}^3 \rtimes \mathbb{S}^2$. To recover the density from the approximated stochastic process (6.13), it is possible to use the Strong Law of Large Numbers.

Theorem 6.2.1. *(Strong Law of Large Numbers) Let* $(\Gamma_t^{(i)}), i \geq 0$ *be a sequence of independent and identically distributed random variables. Assume that* $\mathbb{E}[\Gamma_t^{(i)}] < \infty$*. For* $N \geq 1$*, denote the empirical mean of* $(\Gamma_t^{(1)}, \dots, \Gamma_t^{(N)})$ *by* $\tilde{S}_N = \frac{1}{N} \sum_{i=1}^N \Gamma_t^{(i)}$*. Then, the Strong Law of Large Numbers holds true:*

$$\lim_{N\to\infty} \tilde{S}_N = \mathbb{E}[\Gamma_t^{(1)}], \qquad P - a.s. \tag{6.17}$$

To apply the Theorem to our case, we first concentrate on the set $\mathcal{S}$ of interest.

We fix a discrete covering grid $\{\xi_j\}_{j\in\mathbb{N},j\leq J}$ of a suitable subset V of $\mathbb{R}^3 \rtimes \mathbb{S}^2$: this is a collection of subsets $\{\Omega_i\}$ satisfying $\Omega_i \cap \Omega_j = \emptyset$ if $i \neq j$ and $\cup_{j=1}^{J} \Omega_j = V$. If we perform a discretization on $\mathbb{R}^3$ with step size $\Delta_1, \Delta_2, \Delta_3$ and on $\mathbb{S}^2$ with step size $\Delta_\varphi, \Delta_\theta$, we assign the element ξ_j to be a representative of the j^{th} box. In this sense we define $\mathcal{S} := \{\xi_j\}_{j\in\mathbb{N},j\leq J}$.

Then, applying Theorem 6.17 to the stochastic process $\mathbb{1}_{\xi_j}(\Gamma^{(i)}_{(t,x_0,t_0)})$, we get:

$$\rho_\lambda(\xi_j, t; \xi_0, t_0) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{\xi_j}(\Gamma^{(i)}_{(t,x_0,t_0)}). \tag{6.18}$$

In other words, for a given $\xi_0 \in \mathbb{R}^3 \rtimes \mathbb{S}^2$ we simulate N several discrete-time random paths and assign to each region ξ_j a value between 0 and 1 corresponding to the number of paths that passed through it at the final time t divided by N. This provides a distribution over the cells ξ_j that, up to a multiplicative constant, for large values of N gives a discrete approximation of the kernel.

What we have just obtained is a probability density relative to the Kolmogorov equation (6.6). We have seen in subsection 6.1.3 that a time integration is sufficient to obtain the time-independent fundamental solution. In the discrete case, this corresponds to a sum over t:

$$J^T_\lambda(\xi_j, \xi_0) = \sum_{t=t_0}^{T} \rho_\lambda(\xi_j, t; \xi_0, t_0). \tag{6.19}$$

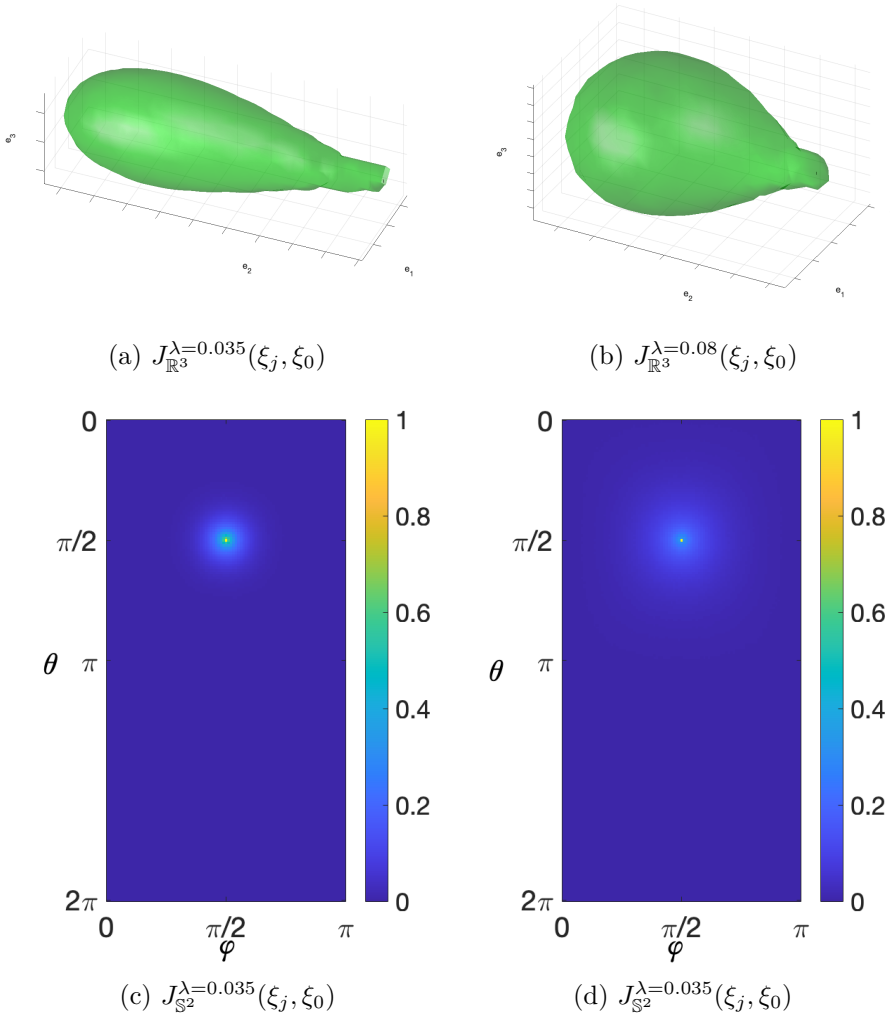
In other words, by applying the numerical method just described not only on the final points of the random walk (6.13) but on all the points of the path (relative to different evolutions times $t < T$), we obtain a numerical approximation of J^T_λ. With this notation we explicitly underline the dependence of the kernel on the parameters of diffusion λ and time T. We will see in subsection 6.2.4 how these parameters affect the shape of the kernel.

Examples of the numerical kernels are shown in Figure 6.2.1. The iso-surfaces in $\mathbb{R}^3$, images (a) and (b), are obtained considering the marginal distribution

$$J_{\mathbb{R}^3}(\xi_j, \xi_0) = \int_{\mathbb{S}^2} J^T_\lambda(\xi_j, \xi_0) d\sigma, \tag{6.20}$$

where $d\sigma = \sin\varphi d\theta d\varphi$ is the spherical measure on $\mathbb{S}^2$. On the other hand, the marginal distributions on the sphere $J^\lambda_{\mathbb{S}^2}(\xi_j, \xi_0)$ are obtained by integrating $J_\lambda(\xi_j, \xi_0)$ over $\mathbb{R}^3$. They are shown using a chart of $\mathbb{S}^2$ in images (c) and (d): the

greater probability is that of remaining in a neighborhood of the initial point of angles (θ_0, φ_0).



(a) $J_{\mathbb{R}^3}^{\lambda=0.035}(\xi_j, \xi_0)$ (b) $J_{\mathbb{R}^3}^{\lambda=0.08}(\xi_j, \xi_0)$

(c) $J_{\mathbb{S}^2}^{\lambda=0.035}(\xi_j, \xi_0)$ (d) $J_{\mathbb{S}^2}^{\lambda=0.035}(\xi_j, \xi_0)$

Figure 6.2.1: Top: Display of $J_{\mathbb{R}^3}(\xi_j, \xi_0)$ for two different diffusion coefficients λ (left $\lambda = 0.035$, right $\lambda = 0.08$), at isovalue 0.1. Euler-Maruyama scheme parameters are $M = 400$, $T = 100$, $N = 10^6$. Bottom: Display of the marginal distribution $J_{\mathbb{S}^2}$ obtained after an integration of J_λ^T over $\mathbb{R}^3$, for the same parameters as above. The starting point ξ_0 has orientation defined by $(\theta_0, \varphi_0) = (\pi/2, \pi/2)$.

This simulated kernel approximates well the fundamental solution thanks to the Strong Law of Large Numbers, recalled here in Theorem 6.17. And, even though this Theorem can be applied to strengthen numerical convergence results in this setting, as, for instance, those proposed by Duits, Bekkers and Mashtakov

in [DBM19, eq. (76)], a necessary condition is that N is large enough. To better understand how precise the numerical estimation is, it is useful to study the error uncertainty regarding the reported measurement (6.18), as we will see in the next section.

6.2.3 Error estimate

Monte Carlo method is based on the Strong Law of Large numbers and clearly yields approximate results whose accuracy depends on the number of values N used in equation (6.17). Indeed, the numerical mean $J_\lambda^T(\xi, \xi_0)$ computed via (6.18) and (6.19) estimates the real mean $\mathbb{E}[\mathbb{1}_{\xi_j}(\Gamma^{(1)})]$. The difference between these two quantities is usually called **statistical error**. Since it is not numerically possible to know the true value of $\mathbb{E}[\mathbb{1}_{\xi_j}(\Gamma^{(1)})]$, a possible measurement of the error corresponds to studying the probability with which the true mean $\mathbb{E}[\mathbb{1}_{\xi_j}(\Gamma^{(1)})]$ lies in the interval $I = [J_\lambda^T - r, J_\lambda^T + r]$, where r is a suitable confidence radius around the estimated value J_λ^T. In probabilistic terms, this is equivalent to calculate

$$P(|J_\lambda^T - \mathbb{E}[\mathbb{1}_{\xi_j}(\Gamma^{(1)})]| \leq r). \tag{6.21}$$

The interval I is usually called **confidence interval**, because there is a probability equal to (6.21) to know if the real mean lies within the interval itself.

Taking advantage, first, of the fact that we want this convergence radius to depend on the parameter N (more specifically we require that this radius decreases as N increases), and the fact that we can use the central limit theorem to compute (6.21), it makes sense to consider $r = \frac{\sigma}{\sqrt{N}}$, with $\sigma = \sqrt{Var((\mathbb{1}_\xi(\Gamma^{(1)}))}$ standard deviation of the random variable. Since the actual variance of the process cannot be computed, it is standard to consider its approximation:

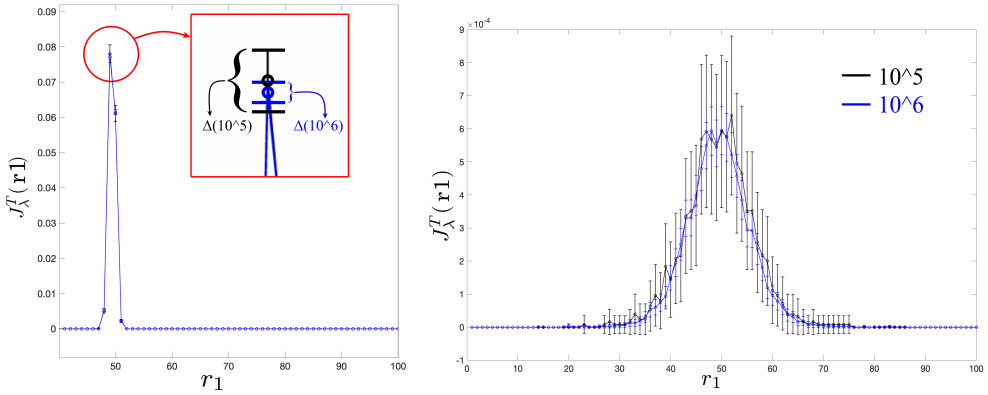
$$\bar{\sigma}^2 = \frac{1}{N} \sum_{i=1}^{N} \left(\mathbb{1}_\xi \left(\Gamma^i \right) - J_\lambda^T \right)^2, \tag{6.22}$$

also known as the variance of Monte Carlo's estimates.

So, with a probability determined by (6.21), we can say that the true mean is in a neighborhood of the approximated mean. Monte Carlo practitioners usually use one standard deviation to determine the confidence radius, and thus the range considered is $[J_\lambda^T - \frac{\bar{\sigma}}{\sqrt{N}}, J_\lambda^T + \frac{\bar{\sigma}}{\sqrt{N}}]$: this corresponds to know with a probability of 33 percent that the estimated mean lies within the interval. Another widely

used value (the one used in our simulation) is to consider 2.57 standard deviations, having $r = \frac{2.57\bar{\sigma}}{\sqrt{N}}$, which turns into a probability of 95 percent. Statisticians usually display the confidence interval using an **error bar** of symmetric size r around the empirical mean. For the reader interested in more information, we refer to [Nic14].

We display error bars for some sections of our kernel, and the results are shown in Figure 6.2.2 and Figure 6.2.3. More precisely, keeping fixed the initial point $\xi_0 \in \mathbb{R}^3 \rtimes \mathbb{S}^2$, and integrating over $\mathbb{S}^2$ as in (6.20), the kernel can be viewed as a function of the variable $(r_1, r_2, r_3) \in \mathbb{R}^3$, namely $J^T_\lambda = J^T_\lambda(r_1, r_2, r_3)$. For simplicity, we focus on a section of the kernel: for example, fixing $(\bar{r}_2, \bar{r}_3)$ we consider $J^T_\lambda(r_1) = J^T_\lambda(r_1, \bar{r}_2, \bar{r}_3)$. We estimate the variance $\bar{\sigma}$ through the equation (6.22), and an error bar corresponding to the confidence interval determined by $r = \frac{2.57\bar{\sigma}}{\sqrt{N}}$ at the point of coordinates $(r_1, J^T_\lambda(r_1))$ is then placed along the vertical axis. The same procedure can be applied to the others coordinates.

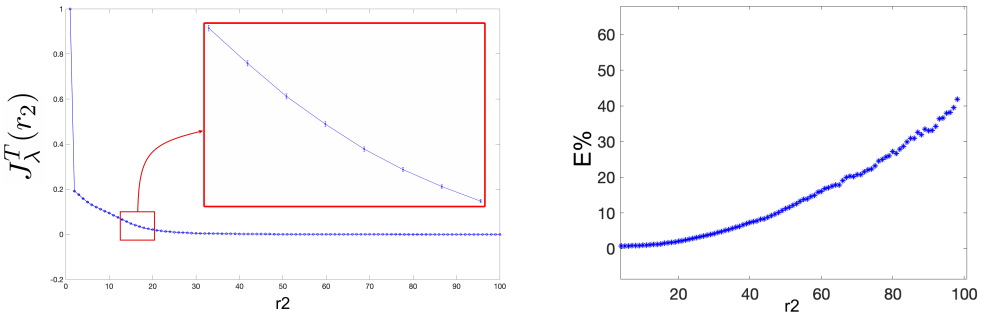


(a) (b)

Figure 6.2.2: Display of error bars on $J^T_\lambda(r_1)$ sections with $\lambda = 0.035$ and $T = 100$. The initial point ξ_0 is described by spatial indices $(r_{1_0}, r_{2_0}, r_{3_0}) = (50, 1, 50)$, while $(\theta_0, \varphi_0) = (\pi/2, \pi/2)$. (a) Display of $J^T_\lambda(r_1)$ identified by $(\bar{r}_2, \bar{r}_3) = (10, 50)$ for two different numbers of paths N: blue color corresponds to $N = 10^6$, the black one to $N = 10^5$. Emphasized in the red square: the amplitude $(\Delta(N))$ of the two confidence intervals. The width of the blue interval is about half of the black one. (b) $J^T_\lambda(r_1)$ identified by $(\bar{r}_2, \bar{r}_3) = (50, 50)$ for two different number of paths: blue correspond to $N = 10^6$, black to $N = 10^5$. Moving away from the pole, both the kernel value and the confidence interval decrease.

Figure 6.2.2 shows error bars for $J^T_\lambda(r_1)$, namely a section perpendicular to the r_1-axis, related to kernels simulated with a different number of paths. Image (a) shows the comparison between the values of a kernel generated by a number

of paths $N = 10^5$ with one generated considering $N = 10^6$, in a neighborhood of the pole ξ_0: the magnitude of the confidence interval, indicated in the picture as $\Delta(N)$, decreases as N increases. Image (b) shows that, as we get further and further away from ξ_0 (both along the r_1-axis and on the r_2-axis), the kernel value decays to zero, as does the standard deviation. This is well in agreement with Remark 6.1.2: as time becomes larger, the kernel value tends to zero.



(a) (b)

Figure 6.2.3: Display of error bars and percentage error for $J_\lambda^T(r_2)$, with kernel parameters $\lambda = 0.035$ and $T = 100$, $N = 10^6$. The initial point ξ_0 is described by $(r_{1_0}, r_{2_0}, r_{3_0}) = (50, 1, 50)$ and $(\theta_0, \varphi_0) = (\pi/2, \pi/2)$. The section has indices $(r_{1_0}, r_{3_0}) = (50, 50)$. (a) Display of error bars on $J_\lambda^T(r_2)$. (b) Display of the percentage error $E_\% = 100\frac{r}{J_\lambda^T(r_2)}$, with $r = \frac{2.57\bar{\sigma}}{\sqrt{N}}$.

Figure 6.2.3 shows the error along the section perpendicular to the r_2-axis and passing through ξ_0. This section is orthogonal to the ones studied in Figure 6.2.2. Image (a) displays the trend of the error bars on the values $J_\lambda^T(r_2)$: moving along the r_2-axis the function decreases to zero, as does the error. This is well in accordance with results of Figure 6.2.2 and Remark 6.1.2. On the other hand, image (b) of Figure 6.2.3 shows the behavior of the percent error, defined here as $E_\% := 100\frac{r}{J_\lambda^T(r_2)}$, with $r = \frac{2.57\bar{\sigma}}{\sqrt{N}}$. In this case, the tendency is the opposite compared to image (a): we have a good percentage estimate near ξ_0, while this goodness decreases as the distance from the ξ_0 increases. This happens because near the pole the points are reached by a high number of stochastic paths, and this number decreases as we move away from ξ_0.

6.2.4 Dependence on parameters

The kernel defined by numerical approximation mainly depends on 4 parameters: M the number of steps of the simulated path, N the number of paths considered, λ the diffusion coefficient, and T the final evolution time of the generated path. We will always consider the first two parameters quite large to ensure the convergence of the method. The last two, on the other hand, affect the behavior of the kernel, see Figure 6.2.4.

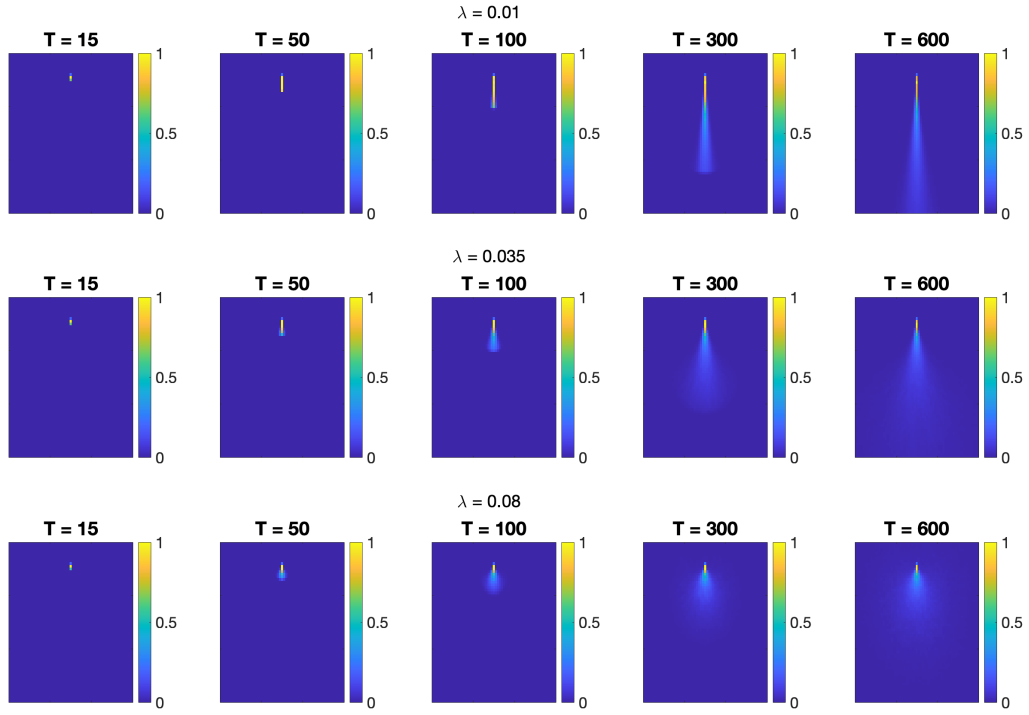


Figure 6.2.4: Marginal projections on the plane $r_1 - r_2$ to display dependence on the parameters λ and T. The horizontal axis of the pictures corresponds to r_1-axis, while the vertical one corresponds to r_2-axis. Columns describe the scale parameter $T = 15, 50, 100, 300, 600$, while rows correspond to different values of diffusion $\lambda = 0.01, 0.035, 0.08$.

This diffusion parameter operates a modification on the thickness of the kernel: the bigger λ is, the thicker the kernel, making the diffusion terms prevail. On the other hand, the smaller λ is, the thinner the kernel, mainly concentrated on the initial direction $n(\varphi_0, \theta_0)$, making the transport term the leading term characterizing the equation.

On the other hand, the temporal parameter T can be seen as a scale parameter. The effect of the variation of the parameter T on the shape of the kernel is shown in Figure 6.2.4: the columns show a proportional relationship between the increase in the final time of the stochastic path, and the amplitude of the kernel. During numerical experiments for visual grouping, this parameter will be taken in accordance with the image dimension.

6.2.5 Comparison with stereo compatibility fields

In [LZ06] the authors transport the differential structure in a way that nearby points are comparable with the measured information, and in particular, the set of all compatible points is called the compatibility field around the starting point j. The correlation between a neighboring point i and the starting point j shows the degree of support in term of positions and orientations, and moreover, the different slices of Figure 3.3.2 bring out the bending and twisting in space encoded by the differential structure.

The kernels presented in this section, defined through the transition probabilities of equation (6.19), are the counterpart of the compatibility fields in our work. Indeed, these kernels express the correlation rate between neighbors of a starting point p, which in this case we can think of as a position-orientation element whose orientation is determined by the direction $(0, 1, 0)$ in the 3D Cartesian space. Positions and orientations in 3D space affect the compatibilities, as shown in Figure 6.2.5. More precisely, the image shows a collection of position-orientation elements selected by the kernel: we represent for points in $\mathbb{R}^3$ the preferred orientation picked out from the kernel, together with its intensity in $\mathbb{R}^3$, computed via equation (6.19). The intensity is characterized both by color and length of the three-dimensional segment: the darker and longer the segment is, the more it correlates with the starting point p. The bending and twisting information are coded in the measurements since the different position-orientation elements are not constrained to be in fronto-parallel planes. The dependence on position is well displayed: as we move away from the initial position, the points in space lose intensity. The same thing happens for orientations compatible with the initial direction.

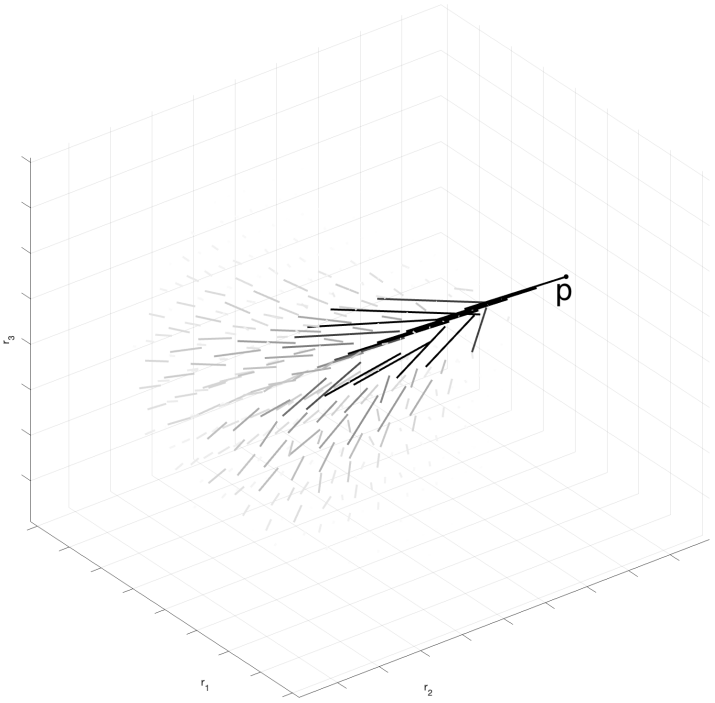


Figure 6.2.5: Position-orientation elements correlated with initial point $p \in \mathbb{R}^3$ and initial direction $n(\theta, \varphi) = n(\pi/2, \pi/2)$ in the first coordinate chart. Intensity decay depends on position and orientation: the darker and longer the segment is, the more it correlates with the starting point.

6.3 Neural activity via mean field equation

The evolution of a state of a cells' population has been modeled through a mean field equation in many works, for example see [EC80, BC03, FF10, SC15], firstly proposed in the works of Amari [Ama72] and Wilson and Cowan [WC72]. According to these works, the result of the propagation along the connectivity generated by the visual scene can be described, without considering the delays for the transmission of the signal, by the following integer differential equation in the 3D perceptive space of positions and orientations:

$$\frac{da(\xi,t)}{dt} = -\alpha a(\xi,t) + \varrho\left(\int_\Omega \mu J_\lambda(\xi,\xi')a(\xi',t)d\xi' + h(\xi,t)\right), \text{ in } \mathbb{R}^3 \rtimes \mathbb{S}^2 \quad (6.23)$$

where ξ is a point of $\mathbb{R}^3 \rtimes \mathbb{S}^2$, $t > 0$, $d\xi = dxd\sigma(n) = dx\sin\varphi d\theta d\varphi$, with dx Lebesgue measure in $\mathbb{R}^3$ and $d\sigma$ spherical measure on $\mathbb{S}^2$, the coefficient α repre-

sents the decay of activity, and $h(\xi, t)$ is the feedforward input.

The function ϱ is the transfer function of the population and it has a piecewise linear behavior, as proposed in [SC15]:

$$\varrho(s) = \begin{cases} 0, & s \in]-\infty, c - \frac{1}{2\gamma}[\\ \gamma(s-c) + \frac{1}{2} & s \in [c - \frac{1}{2\gamma}, c + \frac{1}{2\gamma}] \\ 1, & s \in]c - \frac{1}{2\gamma}, +\infty[\end{cases} \tag{6.24}$$

where γ is a real number that represents the slope of the linear regime and c is the half-height threshold.

The kernel $J_\lambda(\xi, \xi')$ is the contribution of the cortico-cortical connectivity introduced in the previous sections through equation (6.9), while the parameter μ is a coefficient of short-term synaptic facilitation.

Remark 6.3.1. *In the bi-dimensional case, it is customary when dealing with kernels with the purpose of describing pure orientation, to consider reciprocal connections [SCS10b, BCCS14]. This assumption is justified by neural studies, see for example [KE92], where it is reasonable to consider two cells to be symmetrically connected. We then take as kernels a symmetrization of the corresponding connectivity kernels. The chosen symmetrization consists of taking*

$$J_S(\xi_i, \xi_j) = \frac{J_\lambda(\xi_i, \xi_j) + J_\lambda(\xi_j, \xi_i)}{2}. \tag{6.25}$$

This symmetrization is equivalent to sum the fundamental solution associated to (6.1) with the fundamental solution of the same operator under an angular shift in the opposite three-dimensional selected direction. Such rotation turns the drift term Y_3 into $-Y_3$, hence transforming the forward Kolmogorov equation into the corresponding backward equation, so the sum of the two solutions is symmetric.

Remark 6.3.2. *As firstly remarked in [SC15], the set Ω is the restriction on a domain on the space of points activated by the visual stimulus. In particular, by assuming that the function of the feedforward input h can attain only two values, namely 0 and a constant value c, we define the set of points activated by the presence of the input:*

$$\Omega = \{\xi \in \mathbb{R}^3 \rtimes \mathbb{S}^2; h(\xi) = c\}. \tag{6.26}$$

We suppose that the set Ω has a finite measure with respect to the introduced

measure $d\xi$. *Hence the mean field activity equation reduces to:*

$$\frac{da(\xi,t)}{dt} = -\alpha a(\xi,t) + \gamma\left(\int_\Omega \mu J_S(\xi,\xi')a(\xi',t)d\xi' + c\right), \ \textit{in } \Omega. \tag{6.27}$$

6.3.1 Existence and uniqueness of a solution

In this subsection, we deal with the property of existence and uniqueness of a solution to equation (6.23). We use classical instruments of functional analysis, assuming to work in $L^2(\Omega,\mathbb{R})$, the space of the square-integrable functions from Ω to $\mathbb{R}$. This is a Hilbert space for the usual inner product

$$< a_1, a_2 >= \int_\Omega a_1(\xi)a_2(\xi)d\xi, \tag{6.28}$$

inducing the norm $\|a\|_{L^2(\Omega,\mathbb{R})} := \left(\int_\Omega a^2(\xi)d\xi\right)^{1/2}$. We define $f(a) := -\alpha a + A(a) + c$, with

$$A(a) = \gamma\int_\Omega \mu J_S(\xi,\xi')a(\xi',t)d\xi', \tag{6.29}$$

and we denote with I a closed interval on the real line containing 0. We consider a mapping $a : I \longrightarrow L^2(\Omega,\mathbb{R})$, and equation (6.23) is recast as a Cauchy Problem:

$$\begin{cases} a'(t) = f(a) \\ a(t_0) = a_0 \end{cases} \tag{6.30}$$

with initial value a_0.

To apply the results that guarantee the existence and uniqueness of the solution to the Cauchy Problem (6.30) in Banach spaces, we need to show that the function f is well defined on $L^2(\Omega,\mathbb{R})$ and that f is Lipschitz, in accordance with the results presented in [HL13].

Proposition 6.3.1. *If* $J_S(\xi,\xi') \in L^2(\Omega\times\Omega,\mathbb{R})$, *the function* f *is well defined and* $f(a) \in L^2(\Omega,\mathbb{R})$ *for all* $a \in L^2(\Omega,\mathbb{R})$.

Proof. Performing a direct computation:

$$\begin{aligned} &\|f(a)\|^2_{L^2(\Omega,\mathbb{R})} \leq \\ &\|a\|^2 + \|A(a)\|^2 + \|c\|^2 + 2\|a\|^2\|A(a)\|^2 + 2\|c\|^2\|A(a)\|^2 + 2\|c\|^2\|a\|^2, \end{aligned} \tag{6.31}$$

exploiting the Cauchy-Schwarz inequality. Since a and $c \in L^2(\Omega,\mathbb{R})$, we only need

to prove that $\|A(a)\|_{L^2(\Omega,\mathbb{R})} < \infty$:

$$\begin{aligned}\|A(a)\|^2_{L^2(\Omega,\mathbb{R})} &= \int_\Omega \left(\int_\Omega J_S(\xi,\xi')a(\xi',t)d\xi' \right)^2 d\xi \\ &\leq \int_\Omega \left(\int_\Omega J_S^2(\xi,\xi')d\xi' \int_\Omega a^2(\xi',t)d\xi' \right) d\xi \\ &\leq \|J_S\|^2_{L^2(\Omega\times\Omega,\mathbb{R})}\|a\|^2_{L^2(\Omega,\mathbb{R})}\end{aligned} \tag{6.32}$$

using again the Cauchy-Schwarz inequality. □

Remark 6.3.3. *The kernel $J_S(\xi,\xi')$ is obtained as the fundamental solution of the hypoelliptic operator (6.8). Local estimates in terms of Gaussian upper and lower bounds are provided for example in [BP07b], recalled here in Remark 6.1.2. It follows that this kernel is locally integrable, in particular locally square integrable. Furthermore, the operator (6.29) is linear, bounded, and compact on the measure space of definition, see [Con19].*

Proposition 6.3.2. *Let $J_S \in L^2(\Omega\times\Omega,\mathbb{R})$, then we have that the function f is Lipschitz in the variable a.*

Proof. We have

$$\begin{aligned}\|f(a_1) - f(a_2)\|_{L^2(\Omega,\mathbb{R})} =& \|-a_1 + a_2 + A(a_1) - A(a_2)\|_{L^2(\Omega,\mathbb{R})} \\ \leq& \|a_1 - a_2\|_{L^2(\Omega,\mathbb{R})} + \|A(a_1) - A(a_2)\|_{L^2(\Omega,\mathbb{R})} \\ \leq& (1+M)\|a_1 - a_2\|_{L^2(\Omega,\mathbb{R})}\end{aligned} \tag{6.33}$$

where we first use the triangular inequality, and then the properties of Remark 6.3.3. In particular, M is the Lipschitz constant for the operator A. □

6.3.2 Stability analysis

We perform a stability analysis for the equation (6.23), exploiting the Lyapunov method in Banach spaces, by applying the results of Faye and Faugeras in [FF10] to our case of study.

6.3.2.1 General results

Definition 6.3.1. *Let $\phi \in C^1([0,+\infty), L^2(\Omega,\mathbb{R}))$ be a solution to the Cauchy problem (6.30). The function ϕ is said to be* **Lyapunov stable** *if for all $\varepsilon > 0$ there exists $\delta > 0$ such that for every u_0 s.t. $\|a_0 - u_0\| < \delta$, then for the solution*

ψ of the Cauchy problem with initial datum u_0, it holds $\|\psi(t) - \phi(t)\| < \varepsilon$, for all $t \in [0, +\infty)$. If then $\lim_{t\to+\infty}\|\psi(t) - \phi(t)\| = 0$ then ϕ is said **asymptotically stable**.

Definition 6.3.2. *Let $\mathcal{F} \subseteq L^2(\Omega, \mathbb{R})$ an open subset such that $0 \in \mathcal{F}$, and let $V : \mathcal{F} \longrightarrow \mathbb{R}$ be a continuous functional. Then V is a* **Lyapunov functional** *for the problem (6.30) if the following conditions hold:*

- *$V(0) = 0$ and $V(a) > 0$ for all $a \in \mathcal{F}, a \neq 0$,*
- *$dV(a)(f(a)) \leq 0$ for all $a \in \mathcal{F}$,*

where the derivative of the functional is defined as:

$$dV(a)(f(a)) = \lim_{\varepsilon\to 0} \frac{V(a + \varepsilon f(a)) - V(a)}{\varepsilon}. \tag{6.34}$$

Corollary 6.3.1. *Let us assume without loss of generality that the initial datum of the problem (6.30) is $a_0 = 0$. Let V be a Lyapunov functional defined on $L^2(\Omega, \mathbb{R})$ such that the following conditions hold:*

- *$u_1(\|a\|) \leq V(a) \leq u_2(\|a\|)$, where u_1 and u_2 are real, continuous, non decreasing functions such that $u_1(s), u_2(s)$ positive for $s > 0$ and $u_1(0) = u_2(0) = 0$.*
- *$u_1(t) \to +\infty, t \to +\infty$*
- *$dV(a)(f(a)) < -u_3(\|a\|)$ where u_3 is real continuous non decreasing positive for $s > 0$ and $u_3(0) = 0$.*

Then, the point 0 is asymptotically stable for (6.30).

In the following, we are interested in the study of stationary solutions, namely time-independent solutions.

Definition 6.3.3. *An element $a_0 \in L^2(\Omega, \mathbb{R})$ is said to be a* **stationary (or steady) state** *for (6.30) if $f(a_0) = 0$.*

In the case of a steady state a_0, the solution $\phi(t) = a_0$ is a constant solution to the Cauchy problem (6.30). It is possible to write the Taylor expansion of the function f in a neighborhood of the steady state a_0, by defining $\tilde{A} := df(a_0)$, in the following way

$$f(a) = \tilde{A}(a - a_0) + o(\|a - a_0\|), \text{ for } a \to a_0. \tag{6.35}$$

Definition 6.3.4. *We call* **linearized system** *associated to* (6.30) *in* a_0 *the following*

$$a' = \tilde{A}a. \tag{6.36}$$

6.3.2.2 Lyapunov method in the space of position and orientation

To study the stability of the stationary solution a_0, we consider a small perturbation around the stationary state $u = a - a_0$, obtaining the linearized system

$$\tilde{A}(u) = -\alpha u + \gamma\mu \int_\Omega J_S(\xi, \xi')u(\xi')d\xi'. \tag{6.37}$$

In particular, the linearized equation around a_0, associated to (6.23) has the following expression:

$$\partial_t u = -\alpha u + \gamma\mu \int_\Omega J_S(\xi, \xi')u(\xi')d\xi'. \tag{6.38}$$

If 0 is asymptotically stable for for (6.38), then a_0 is asymptotically stable for (6.23).

The Lyapunov functional that we consider is defined as the quadratic form:

$$V(u) = \frac{1}{2}\int_\Omega u^2(\xi)d\xi. \tag{6.39}$$

For (6.39) to be a Lyapunov functional, we need to check the properties of Definition 6.3.2. The first property is trivially satisfied. Regarding the derivative, we have:

$$\begin{aligned} dV(u)(\tilde{A}(u)) &= -\int_\Omega \alpha u(\xi)u(\xi)d\xi + \int_\Omega u(\xi)\mu\gamma \int_\Omega J_S(\xi, \xi')u(\xi')d\xi' d\xi \\ &\leq -\alpha\|u\|^2_{L^2(\Omega,\mathbb{R})} + \mu\gamma\|J_S\|_{L^2(\Omega^2,\mathbb{R})}\|u\|^2_{L^2(\Omega,\mathbb{R})} \\ &= (-\alpha + \mu\gamma\|J_S\|_{L^2(\Omega^2,\mathbb{R})})\|u\|^2_{L^2(\Omega,\mathbb{R})}. \end{aligned} \tag{6.40}$$

In particular, it follows that if $\|J_S\|_{L^2(\Omega^2,\mathbb{R})} < \frac{\alpha}{\mu\gamma}$ then a_0 it is stable.

6.3.2.3 Stability via Eigenvalues problem

The equation satisfied by small perturbation around the stationary state $u = a - a_0$, is given by:

$$\frac{du(\xi, t)}{dt} = -\alpha u(\xi, t) + \gamma \int_\Omega \mu J_S(\xi, \xi')u(\xi', t)d\xi', \text{ in } \Omega. \tag{6.41}$$

In particular, it is worth noting that in this framework, u is a solution of the homogeneous equation associated to equation (6.23). Then, the stability of the solution of this linear equation can be studied by mean of the eigenvalue problem associated to the linear operator

$$Lu = -\alpha u + \gamma\mu \int_\Omega J_S(\xi, \xi')u(\xi')d\xi' = \lambda u, \tag{6.42}$$

which is equivalent to the following expression

$$\int J_S(\xi, \xi')u(\xi')d\xi' = \tilde{\lambda}u, \tag{6.43}$$

with $\tilde{\lambda} = \frac{\lambda+\alpha}{\gamma\mu}$. In general, the system is stable if $\lambda < 0$ and so imposing it to be negative we get $\tilde{\lambda} < \frac{\alpha}{\gamma\mu}$. Moreover, if we consider all parameters equal to 1, we obtain that the system is stable if the eigenvalue $\tilde{\lambda} < 1$.

It follows from the Spectral Theorem on Hilbert spaces that a solution ϕ can be written in terms of eigenvectors u_i and eigenvalues λ_i of the operator (6.29) in the following way:

$$\phi(t) = \sum c_i u_i e^{\lambda_i t} \tag{6.44}$$

and in particular, if $t = 0$ we have $a_0 = \sum c_i u_i$.

6.4 Discrete mean field equation and spectral clustering

The spectral analysis implemented by the neural population translates into the identification of perceptual units, leading to a grouping performance strictly linked with arguments of spectral clustering and dimensionality reduction, where the salient objects in the scene correspond to the eigenvectors with the largest eigenvalues.

More precisely, in application it is common to work with a finite number of points, hence having a discrete structure of the input. The configurations are then constituted by a finite number N of elements in the lifted space with coordinates $\xi_i = (p_i, n(\theta_i, \varphi_i)), p_i \in \mathbb{R}^3, n(\theta_i, \varphi_i) \in \mathbb{S}^2$, which are in relationship with the binocular cells activated by the stimulus. We call the activated set Ω_d, and so the

mean field equation (6.41) becomes

$$\frac{da(\xi_i,t)}{dt} = -\alpha a(\xi_i,t) + \gamma\mu \sum_{j=1}^{N} J_S(\xi_i,\xi_j)a(\xi_j,t), \text{ in } \Omega_d. \tag{6.45}$$

In particular, the kernel J_S is reduced to a symmetric matrix $\mathbf{J}$ of dimension $N \times N$, whose entries (i,j) are

$$\mathbf{J}_{ij} = \gamma\mu J_S(\xi_i,\xi_j),$$

and the eigenvalue problem (6.43) becomes

$$\mathbf{J}a = \tilde{\lambda}a.$$

This matrix can be considered as the equivalent of the *affinity matrix* introduced in many works for spectral clustering and dimensionality reduction problems, see for example [RS00, BN03, CL06, PF98, MS01, SM00].

6.4.1 Spectral Clustering and dimensionality reduction

In this paragraph, we explain how (discrete) spectral analysis is connected with the problem of perceptual grouping, linked with results of dimensionality reduction.

Let us suppose that the visual setting is composed of N discrete elements $\xi_i, i = 1,\ldots,N$ that are present in the visual scenery. It is possible to describe the visual scene in terms of an affinity matrix $\mathbf{J}$, whose elements $\mathbf{J}_{i,j}$ describe the affinities between elements ξ_i and ξ_j. It is then clear that the complexity of the matrix $\mathbf{J}$ is of order $O(N^2)$.

The idea of Perona and Freeman in [PF98] is to characterize the scene by approximating the matrix $\mathbf{J}$ by matrices of rank 1 and complexity N, each of which will identify a perceptual unit in the scene. In particular, a rank 1 matrix is represented as the external product of a vector p with itself: the element of place i,j of the matrix is $p_i p_j$, for every $i,j = 1,\ldots,N$. So, the first approximating matrix is computed as the best approximation of $\mathbf{J}$ minimizing the Frobenius norm as follows:

$$p_1 = \text{argmin}_p \sum_{i,j=1}^{N} (\mathbf{J}_{ij} - p_i p_j)^2. \tag{6.46}$$

It has been proved in [PF98] that the minimizer p_1 is the first eigenvector v_1 of the matrix $\mathbf{J}x$ with largest eigenvalue λ_1: $p_1 = \lambda_1^{1/2} v_1$.

This procedure is applied iteratively until all significant eigenvectors are deter-

mined. The salience of an eigenvector is determined starting from the magnitude of the corresponding eigenvalue: the process ends when this is sufficiently small. In general, this method selects only n eigenvectors, with $n < N$, reducing the dimensionality of the problem.

6.4.2 Normalization with the transition matrix of a Markov chain

Spectral analysis techniques, proposed in [MS01] and used for example in [BCCS14], have shown that there exists a normalization that turns the real symmetric affinity matrix $\mathbf{J}$ into the transition matrix $\mathbf{P}$ of a Markov process, via row-wise normalization. More precisely, if $\mathbf{D}$ is the diagonal degree matrix, having elements

$$\mathbf{D}_{i,i} = \sum_{j=1}^{N} \mathbf{J}_{ij}, \text{ for } i = 1 \dots N$$

the normalized affinity matrix $\mathbf{P}$ is given by

$$\mathbf{P} = \mathbf{D}^{-1}\mathbf{J}.$$

Remark 6.4.1. *This matrix in general is not symmetric. However, it can be shown (see for example [MS01]) that its eigenvalues λ_i are real and satisfy $0 \leq \lambda_i \leq 1$, and its eigenvectors u_i accordingly can be chosen with real components.*

Remark 6.4.2. *This grouping mechanism arises from a spectral clustering of the associated graph Laplacian, see for example [SM00, SB98, CG97, DlPHHVDW08]. Then it is possible to interpret the stimulus as a weighted graph.*

Remark 6.4.3. *We recall here that the normalization introduced in this paragraph is supposed to occur at neural level throughout the neural visual system ([CH12]), and in particular, in the primary visual cortex, see for example [Hee92]. In general, a normalization divides the response of a neuron and the sum of the activity of a pool of neurons. This is in agreement with the so-called Tononi normalization, introduced in [TSE94], where the sum of the weights of the connections is required to be equal to a constant value. This is consistent with the behavior of biological neural systems in which the weight of connections is adjusted in such a way that a neuron with few connections will weigh these inputs more heavily than a neuron with many connections. Further discussion regarding this type of normalization can be found in [BBB09].*

6.5 Grouping results

In this section, we develop the ideas illustrated so far by numerical examples: first, we describe the stimulus used, and then we proceed with the spectral analysis, adapting the spectral clustering algorithm introduced in [BCCS14, SC15] to our case of study. In particular, eigenmodes are computed as eigenvectors of the $\mathbf{P}$ matrix, and we show how they represent perceptual units in the 3D space. The main steps of the algorithm are summarized in Remark 6.5.1.

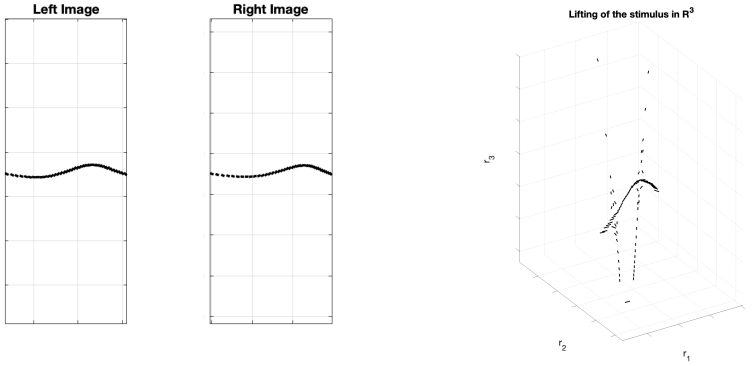
Remark 6.5.1. (Main steps of the algorithm)

- *Recover the domain* $\Omega_d \subset \mathbb{R}^3 \rtimes \mathbb{S}^2$, $\xi_k \in \Omega_d, k = 1, \dots n$, *from the coupling of retinal images: the set of input elements appearing in the visual scene, plus some (random) noise.*
- *Build the affinity matrix* $\boldsymbol{J}_{ij} = J(\xi_i, \xi_j)$.
- *Compute the normalization* $\boldsymbol{P} = \boldsymbol{D}^{-1}\boldsymbol{J}$ *and solve the eigenvalue problem* $\boldsymbol{P}u = \lambda u$.
- *Find the* q *eigenvalues* $\{\lambda_i\}_{i=1}^q > 1 - \varepsilon$ *and take the corresponding eigenvectors* $\{u_i\}_{i=1}^q$.
- *For* $k = 1, \dots, n$ *assign the point* ξ_k *to the clustered labeled by* $\max_i\{u_i(k)\}_{i=1}^q$.
- *Join together the clusters with less than* Q *elements.*

6.5.1 Parameterized curve

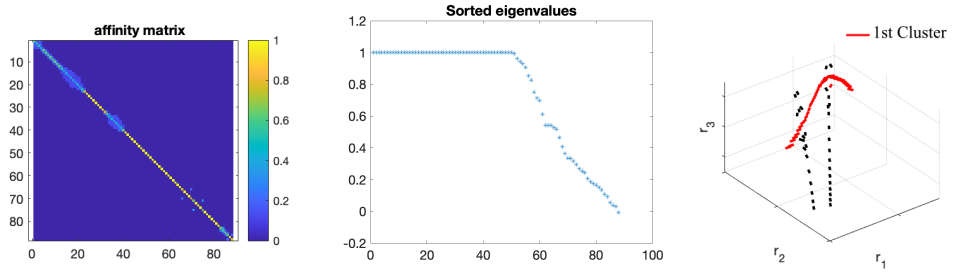
Starting from two retinal images which are the projection of a curve $\gamma : [0, T] \to \mathbb{R}^3$ of $n_{\text{curve}} = 40$ points, we recover the lifting of the stimulus in the space $\mathbb{R}^3 \rtimes \mathbb{S}^2$, as shown in Figure 6.5.1.

Then, we apply the grouping algorithm to resolve both the corresponding problem and the segmentation of the tridimensional scene. The results derived from the spectral analysis, in this case, can be seen in Figure 6.5.2. We used the time interval of evolution $[0, 300]$ for generating random paths for the computation of the kernel. We compute the affinity matrix $\mathbf{P}$ and, after having thresholded and selected the eigenvalues, we notice that there is only one main cluster that group together more than $Q = 25$ points, described fully by the first eigenvector. The first eigenvector in this case recovers satisfyingly the real curve, except for a ladder point.



(a) (b)

Figure 6.5.1: (a) Left and right retinal images. (b) Lifting of the stimulus in $\mathbb{R}^3 \rtimes \mathbb{S}^2$: the coupling generates $n = 88$ corresponding points.



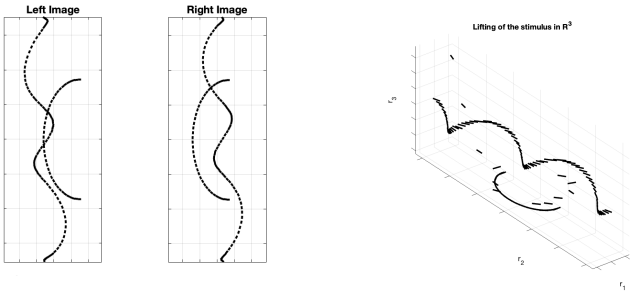
(a) (b) (c)

Figure 6.5.2: (a) Display of the affinity matrix $\mathbf{P}$ obtained using the kernel with diffusion parameter $\lambda = 0.02$, $M = 300$, $N = 10^6$. (b) Plot of the sorted eigenvalues of the affinity matrix. (c) Colored in red are the points recovered from the first eigenvector, associated with the only main cluster.

6.5.2 Helix and arc

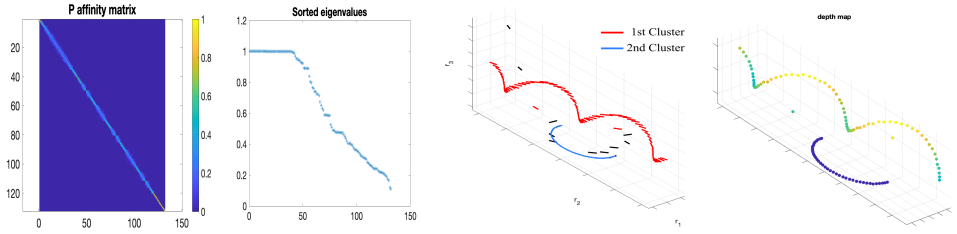
Then, we tested our algorithm on a synthetic stimulus similar to the one presented in the work of Alibhai and Zucker [AZ00]. The stimulus is a couple of synthetic images, composed of an arc and an y-helix. In our example the arc has $n_{\text{arc}} = 40$ points, and the helix is composed by $n_{\text{helix}} = 80$ points, as shown in Figure 6.5.3.

We then apply our grouping algorithm in the interval of time $[0, 57]$ and the spectral analysis is displayed in Figure 6.5.4. The algorithm segments well in two parts the stimulus, recovering the two main clusters of the set: the first cluster corresponds to the helix, red points in image (c) of Figure 6.5.4. The second cluster recovers the point characterizing the arc, they are the blue ones in image



(a) (b)

Figure 6.5.3: (a) Left and right retinal images. (b) Lifting of the stimulus in $\mathbb{R}^3 \rtimes \mathbb{S}^2$: the coupling generates $n = 132$ corresponding points.



(a) (b) (c) (d)

Figure 6.5.4: (a) Display of the affinity matrix $\mathbf{P}$ obtained using the kernel with diffusion parameter $\lambda = 0.0275$, $M = 300$, $N = 10^6$. (b) Plot of the sorted eigenvalues of the affinity matrix. (c) Points clustered together are marked by the same color. The number of minimum elements to form a cluster is $Q = 35$. (d) Depth map.

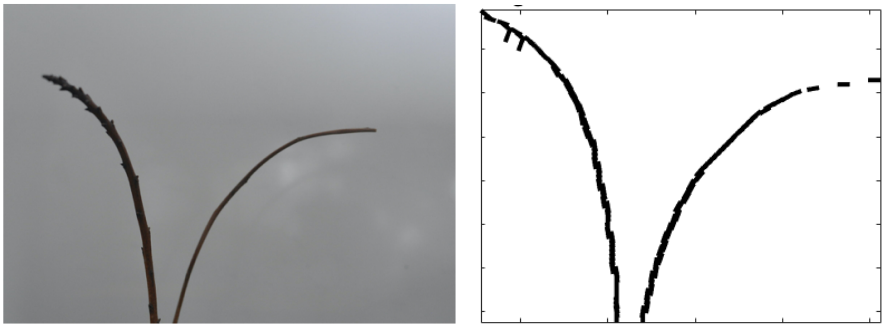
(c) of Figure 6.5.4. So, these eigenvectors are well suited to describe the two main perceptual units in the visual scene, even if two ladder points are recovered from the helix.

We also generate a depth map, which allows us to display the points recovered in $\mathbb{R}^3$ with different colors following depth (blue is closest to the observer, and yellow is the farthest away).

The stereo reconstruction algorithm of [AZ00] recovers the helix and the arc, which result separate in depth. Even our algorithm allows well this kind of reconstruction, with the additional segmentation of the tridimensional scene in two perceptual objects.

6.5.3 Natural images

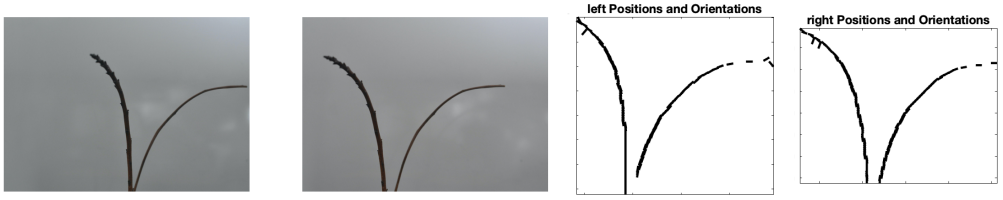
A black and white image I is typically described with a function $I : \mathbb{R}^2 \to [0, 1]$, where the value 0 corresponds to the color white, while the value 1 corresponds to the black. In particular, since $I(x, y)$ describes the intensity of the retinal image on the retinal point (x, y), we assume that points of maximum intensity are the points characterized by the presence of the stimulus.



(a) (b)

Figure 6.5.5: (a) Natural image imported in Matlab. (b) Segmentation of the image through the process of maximal selectivity for orientation.

Applying the maximal selectivity process which characterized the action of simple cells, as described in [CS06], we can extract the cartoon image, formally expressed as a piecewise constant function, which takes only the values 0 or 1, and the corresponding direction θ. Results of this process are shown in Figure 6.5.5.

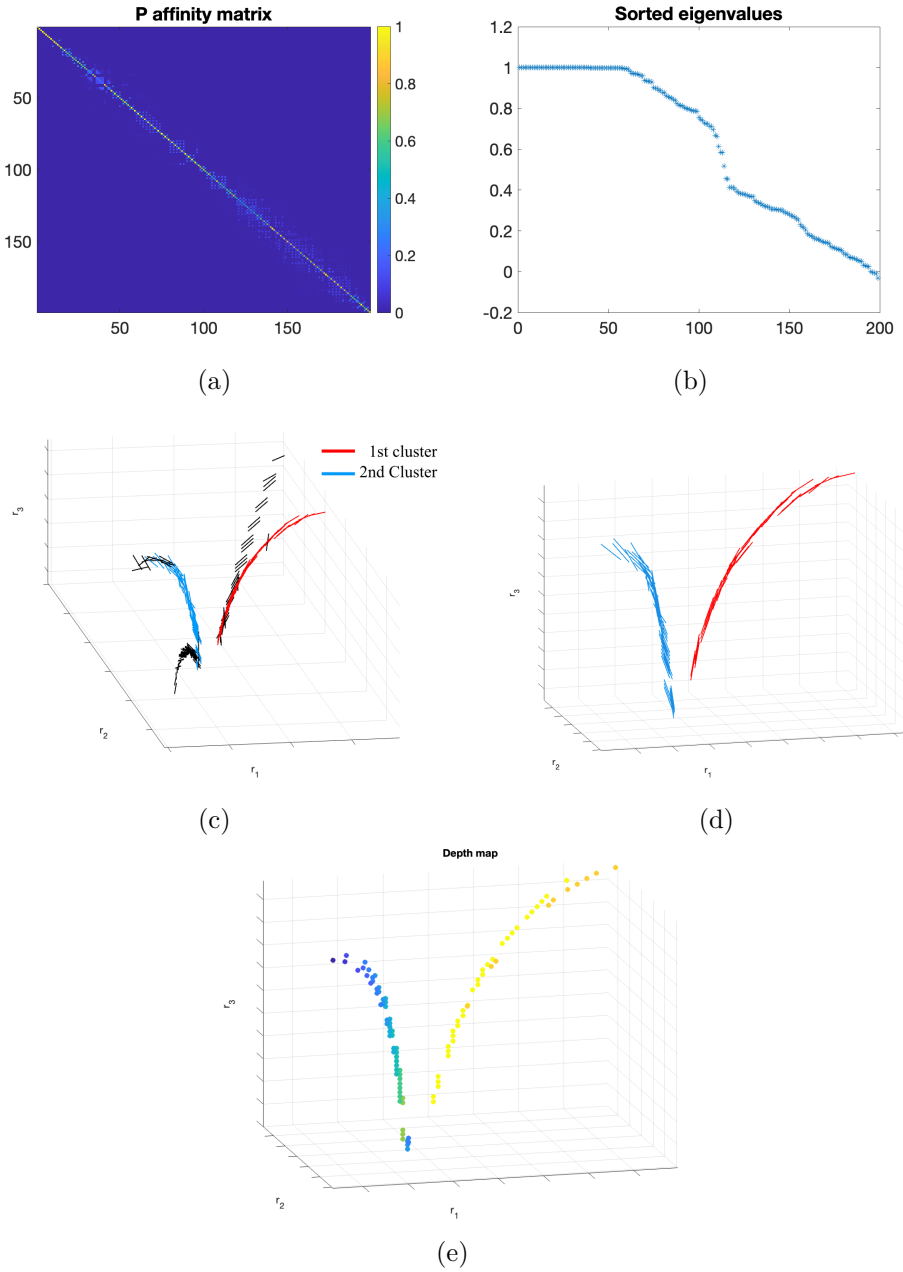


(a) (b)

Figure 6.5.6: (a) Left and right natural images. (b) Lifting in the space $\mathbb{R}^2 \times \mathbb{S}^1$.

We work with a couple of twigs images, generated from cameras with a baseline of 6 cm, with the aim to well simulate the distance between the left and right eyes. The couple of images imported in Matlab can be seen on the left of Figure 6.5.6. The rescaled images have dimensions 89×161. After having applied the maximal selectivity orientation process, we obtain 121 points for the left image, and 127 for

the right one, both displayed in the middle of Figure 6.5.6. These points are lifted in the space $\mathbb{R}^3 \rtimes \mathbb{S}^2$, generating 200 corresponding points.



(a) (b) (c) (d) (e)

Figure 6.5.7: (a) Display of the affinity matrix $\mathbf{P}$ obtain using the kernel with diffusion parameter $\lambda = 0.035$, $M = 300$, $N = 10^6$. (b) Plot of the sorted eigenvalues of the affinity matrix. (c-d) Points clustered together. The number of minimum elements to form a cluster is $Q = 40$. (e) Depth map.

The spectral analysis of the stimulus is shown in Figure 6.5.7. The kernel for computing the affinity matrix has been generated using an evolution interval of time $[0, 100]$. There are two main clusters of points that correspond to the two twigs represented in the figures: the first cluster (displayed in red) is composed by 46 points, while the second main cluster (blu) is formed by 64. And so 110 points are recovered from the tridimensional visual scene. We consider also a depth map, that allows us to better understand the configuration of the stimulus in the space, image (e) of Figure 6.5.7. The points near the observer are marked with dark blue, while the points farther away are characterized by yellow.

6.6 Comparison with Riemannian distance

In this section, we focus on the comparison between the sub-Riemannian distance-based kernel we have introduced, and the kernel obtained by the classical exponential function whose argument depends on the Riemannian distance of the space. The results presented here show the benefits of grouping using sub-Riemannian distances rather than using isotropic distances.

6.6.1 Gaussian kernel

It is typical in graph-based clustering [SM00, CG97, CL06] to use a Gaussian kernel as a similarity measure to group clouds of points. In particular, if we consider a manifold M, this kernel is usually a non-linear function of a Euclidean-type distance:

$$k_{\xi_0,E}(\xi) = \frac{1}{4\pi\sigma} \exp\left(-\frac{d_E(\xi - \xi_0)^2}{4\sigma}\right), \tag{6.47}$$

with d_E Euclidean-type distance on M, $\xi_0, \xi \in M$, with ξ_0 fixed, and $\sigma \in \mathbb{R}_+$.

When we talk about Euclidean-type distance, we are talking about Riemannian distance on the manifold we are considering, in our case, $M = \mathbb{R}^3 \rtimes \mathbb{S}^2$. The distance is defined as the classic Euclidean distance $d_{\mathbb{R}^3}$ on $\mathbb{R}^3$, plus the classical Riemannian distance on the sphere $d_{\mathbb{S}^2}$:

$$d_E(\xi, \xi_0) = d_{\mathbb{R}^3}(p, p_0) + d_{\mathbb{S}^2}(n, n_0), \text{ with } \xi = (p, n) \in \mathbb{R}^3 \rtimes \mathbb{S}^2. \tag{6.48}$$

This connectivity kernel can be seen as a fundamental solution of the classical heat operator in M. The parameters involved here are only position $\xi \in M$ and the variance $\sqrt{\sigma} \in \mathbb{R}$, so this kernel associates points that are close in the

typical euclidean sense, and decorrelates points that are far. The parameter σ has the function of a scale parameter, and so its function is to enlarge or shrink the correlation area: if σ is small, the correlation is higher between points near ξ_0, while if σ is large the correlation involves also points far away. This is basically due to the behavior of the Gaussian function.

It is then clear that the Riemannian distance of the space, core of the Gaussian similarity kernel of equation (6.47), is not associated with a process of alignment towards a principal direction, but it is at the basis of correlation between points characterized by closeness. We can therefore think that this kernel encodes the Gestalt principle of proximity, rather than the one of good continuation. In fact, the euclidean kernel is usually used for grouping clouds of points instead of elements characterized by specific features.

6.6.2 Numerical simulation

First of all, to visually understand the differences between the introduced sub-Riemannian kernel (6.18) and the Riemannian-based Gaussian kernel, we picture its iso-surfaces by displaying the intensity of (6.47) in $\mathbb{R}^3$, integrating on the sphere in a similar way to what we have done in (6.20). The three-dimensional Gaussian kernel surface is shown in Figure 6.6.1, image (a): they are Euclidean balls, characterized by different radii as the parameter $\sigma \in \mathbb{R}_+$ varies.

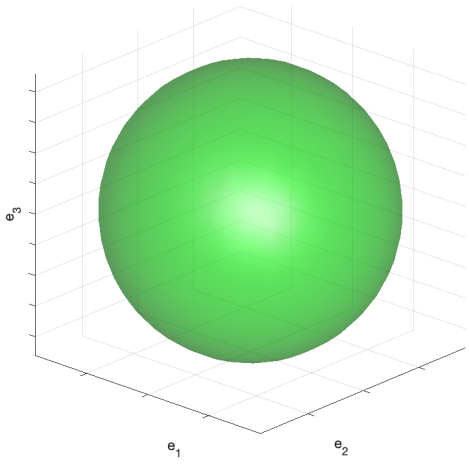
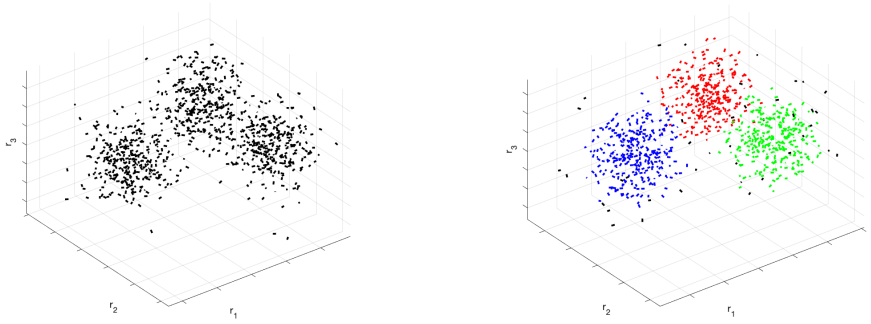


Figure 6.6.1: Iso-surface of Gaussian kernel (6.47).

This kernel is used to segment the cloud points of image (a) of Figure 6.6.2, by applying the same algorithm proposed in Remark 6.5.1: we generate the affinity matrix with the similarity measure induced by (6.47), and then we diagonalize it, recovering the eigenvectors. We find that the three clouds are represented by the three first eigenvectors of the affinity matrix, as displayed in image (a) of Figure 6.6.2. In particular, each color corresponds to elements belonging to the same eigenvector.

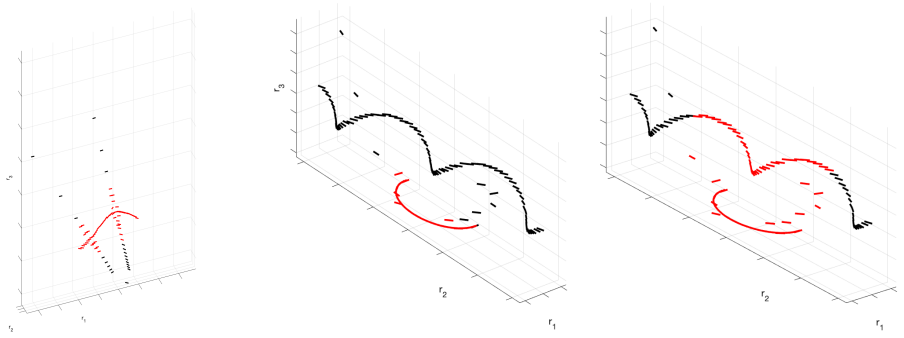


(a) (b)

Figure 6.6.2: (a) Cloud of points in $\mathbb{R}^3 \rtimes \mathbb{S}^2$. (b) Grouping of the three clouds with the first three eigenvectors, using the algorithm proposed in Table 6.5.1 with the Gaussian kernel (6.47).

On the other hand, if we try to use this kernel on a three-dimensional scene characterized by perceptual units in $\mathbb{R}^3 \rtimes \mathbb{S}^2$ surrounded by noise elements, we observe that the eigenvectors associated with the Riemannian affinity matrix do not represent the contours of the stimulus, but the areas with the highest density of elements. We try this on the introduced stimuli of Section 6.5.

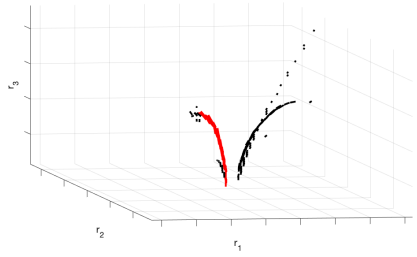
First, we concentrate on the three-dimensional synthetic images. Performing the grouping algorithm on the 3D image representing the parametrize curve, we observe that the first eigenvector (red points) does encode the area with the higher number of elements, but it does not perform 3D reconstruction, result is shown in image (a) of Figure 6.6.3. In this case, the only parameter we can modify to get a better performance is the σ scale parameter, but it is not enough to retrieve the 3D object. This observation is even clearer in the case of the three-dimensional image representing the arc of circle and y-helix. If we group with a suitable value of sigma, for example equal to 3, we notice that the first eigenvector is able to select the points that form the arc of circle, image (b) of Figure 6.6.3. However,



(a) (b) (c)

Figure 6.6.3: Grouping with Gaussian kernel (6.47) on different stimuli. The red points correspond to points selected by the first eigenvector. (a) Grouping algorithm performed on the $3D$ curve of Figure 6.5.1. (b) Grouping algorithm performed on the $3D$ images generated from Figure 6.5.3, with $\sigma = 3$. (c) The same algorithm performed with $\sigma = 9$.

the second eigenvector is not able to reconstruct the three-dimensional helix. We can modify the σ value to enlarge the area of influence, and results are displayed in image (c) of Figure 6.6.3,. In this case even the first eigenvector is not able to reconstruct the 3D arc but it groups only the elements forming the area with the higher number of points.



(a)

Figure 6.6.4: Grouping with Gaussian kernel (6.47) on twig stimuli. The red points correspond to points selected by the first eigenvector.

Then, we perform the same algorithm on the twig images, which an example can be found in Figure 6.6.4. Even in this case, we find results similar to the previous case, observing that the perceptual law underlying the grouping is based on proximity, and so the closeness of points, without properly segmenting and reconstructing the visual scene.

We have showed that, sub-Riemannian geometry, in comparison with Riemannian distances, achieves top performance in grouping the (three-dimensional) visual scenes to solve the stereo-matching problem. Similar results were also found by Bekkers, Chen and Portegies in [BCP18], for perceptual grouping of blood vessels in 2D and 3D, comparing sub-Riemannian and Riemannian distances in $SE(2)$ and $SE(3)$. So, for problems characterized by a strong (three-dimensional) directionality, at the basis of the law of good continuation, sub-Riemannian distances achieve an excellent performance compared to distances of the Euclidean-type.

Chapter 7

Conclusion

In this thesis, a neuro-geometrical model for stereo vision is presented. Stereo geometry has a mathematical structure that is a direct extension of plane curve geometry: instead of (orientation, position) neural columns, these become abstract fibers. The collection of fibers across position (and disparity) is a fiber bundle, and elements of the fiber can be viewed as neurons. The natural coordinates are position, positional disparity and orientations from the left and right eyes, respectively.

The behavior of binocular cells is usually described by the binocular RPs, the product of the left and right monocular RPs. We note that starting from binocular receptive profiles, it is possible to reconstruct the three-dimensional space using only the position and orientation of the visual stimulus in the retinal planes (assuming having corresponding points). This differs from previous models [AZ00, LZ03, LZ06] which assume the existence of monocular cells with curvature-sensitive receptive profiles.

We propose a sub-Riemannian model on the space of position and orientation $\mathbb{R}^3 \rtimes \mathbb{S}^2$ for describing the perceptual space of the involved neurons. This geometrical structure favors the tangent direction of a 3D curve stimulus as a natural consequence of neurophysiological models and classical stereo geometry. The integral curves of the sub-Riemannian structure encode the concepts of curvature and torsion in their coefficients and are introduced to describe the connections between elements. This model can be considered as an extension of the two-dimensional model proposed by Citti and Sarti in [CS06] in the three-dimensional scene.

Integral curves model neural connectivity and they formalize association field

models. They are well in accordance with the formation of regular 3D curves as studied by psychophysical experiments in [HF95, HHK97, KHK16, DW15], with the generalization of co-circularity in 3D [AZ00, LZ03, LZ06], and with mathematical conjectures [KGS05a, KGS+05b] describing the ability to perceive two contour elements in the three-dimensional space as belonging to the same perceptual unit. This provides a new framework for specifying the correspondence problem, by illustrating how good continuation in the 3-D world generalizes good continuation in the 2-D plane.

The stochastic counterpart of the 3D association fields provides a kind of glue that allows transitions from points of the space to nearby points, and is thought of as the underlying circuits of binocular neurons. This leads to the introduction of connectivity kernels that encode the 3D geometric position-orientation information and describe the probability of co-occurrence of elements of the space. These probability densities are inserted as facilitation inducers in a neural population activity model, whose stability analysis leads to the emergence of three-dimensional perceptual units.

We implement a clustering algorithm to test the visual grouping properties of the (discrete) connectivity kernels. We find that our algorithm is able to cluster elements belonging to a single object in the 3D space. At the same time, the algorithm is able to solve the stereo correspondence problem by determining the appropriate match between left and right pixels, while discarding false matches. A comparison with the same algorithm with the classical Gaussian kernel equipped with a Euclidean-type distance explains and justifies the introduction of the proposed sub-Riemannian metric.

Appendices

Appendix A

Position-orientation manifold $\mathbb{R}^3 \rtimes \mathbb{S}^2$ and Lie group of rigid body motions $SE(3)$

In this appendix, we recall the detailed definition of the position-orientation space that we used throughout the thesis to describe the behavior of binocular cells. We mainly refer to [DF11, DCGDH11, DGDHM16] to remind how $\mathbb{R}^3 \rtimes \mathbb{S}^2$ can be thought of homogeneous space for the Lie group of rigid body motions $SE(3)$, by taking the quotient of $SO(3)$, the three-dimensional rotation group, on a subgroup identified with $SO(2)$, the bi-dimensional rotation group.

A.1 3D rotation group $SO(3)$

The 3D rotation group $SO(3)$, also known as the special orthogonal group, is the group of all rotations about the origin of three-dimensional Euclidean space $\mathbb{R}^3$ under the operation of composition.

In particular, if R_i^Ψ denotes a counterclockwise rotation around the i-th axis by an angle Ψ, then every rotation $R \in SO(3)$ can be written uniquely in Euler angle form:

$$R = R_3^\theta R_2^\varphi R_3^\alpha, \tag{A.1}$$

where $\theta, \alpha \in [0, 2\pi]$, while $\varphi \in (0, \pi)$. Thus, the angles θ, φ, α can be taken as coordinates in $SO(3)$.

Remark A.1.1. *Since $SO(3)$ is diffeomorphic to real projective space $P3(\mathbb{R})$, which is a quotient of $\mathbb{S}^3$ by identifying antipodal points, inevitable problems always exist in local charts. In (A.1) the ambiguity arises at the intersection of the equator with the r_3-axis. So, it is possible to introduce a second chart, where every counterclockwise rotation is written as $R = R_1^{\theta} R_2^{\varphi} R_3^{\alpha}$ with $\varphi \in (-\pi/2, \pi/2)$, $\theta, \alpha \in [-\pi, \pi]$. In this second chart, the ambiguity arises at the intersection between the equator and the r_1-axis.*

The group of 3D rotation is a Lie group, and associated with every Lie group is its Lie algebra. The Lie algebra of $SO(3)$ is denoted as $so(3)$ and it is the set spanned by the skew symmetric matrices:

$$A_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{A.2}$$

each of these represents respectively the vector fields $\partial_\alpha, \partial_\varphi, \partial_\theta$ in the neutral element[1] $e \in SO(3)$, and so $so(3) = \text{span}\{\partial_\alpha, \partial_\varphi, \partial_\theta\}_e$.

We have recalled in Chapter 4, section 4.2, that it is possible to identify the Lie algebra of a Lie group with the set of left-invariant vector fields. The left-invariant vector fields can be obtained from the basis $\{\partial_i\}_i$ of the tangent space at the neutral element through the pushforward of the left multiplication $L_P Q := PQ, P$ and $Q \in SO(3)$ by:

$$(L_P)_* \partial_i \phi = \partial_i (\phi \circ L_P)$$

for all smooth $\phi : U_P \to \mathbb{R}$, which are locally defined on some neighborhood $U_P \subseteq SO(3)$ of P.

Example A.1.1. *We recall here the computation of left-invariant vector fields, considering the Lie algebra basis of $so(3)$ in the neutral element $e \in SO(3)$, starting from $\partial_{\varphi,e}$:*

$$(L_R)_* \partial_{\varphi,e} = (DL_R)_e \partial_{\varphi,e} = \frac{d}{dt}_{|t=0} L_R(\gamma(t)) \tag{A.3}$$

with $\gamma : \mathbb{R} \longrightarrow SO(3)$ such that $\gamma(0) = e$ and $\dot{\gamma}(0) = \partial_{\varphi,e}$. We explicitly remind

[1] The neutral element e in $SO(3)$ corresponds to the identity matrix I_{d} of dimension 3.

that $\partial_{\varphi,e} = A_2$ *and so it is natural to choose the curve* γ *as :*

$$\gamma(t) = R_2^t = \begin{pmatrix} \cos t & 0 & \sin t \\ 0 & 1 & 0 \\ -\sin t & 0 & \cos t \end{pmatrix}. \tag{A.4}$$

This leads to:

$$\frac{d}{dt}_{|t=0} L_R(\gamma(t)) = R\frac{d}{dt}_{|t=0} \begin{pmatrix} \cos t & 0 & \sin t \\ 0 & 1 & 0 \\ -\sin t & 0 & \cos t \end{pmatrix} = RA_2, \tag{A.5}$$

the left-invariant vector fields at point $R \in SO(3)$*, which we can call* $\tilde{Y}_{\varphi,R}$*.*

Then, we need to express this quantity in a basis for the tangent space of $SO(3)$ *in a generic point* R*. The matrices that represents the vector fields* $\{\partial_\alpha, \partial_\varphi, \partial_\theta\}$ *in a generic point* R *of coordinates* $(\alpha, \varphi, \theta)$ *in the second chart, are:*

$$\begin{aligned} \partial_{\alpha,R} &= R_1^\theta R_2^\varphi R_3^\alpha A_3 \\ \partial_{\varphi,R} &= R_1^\theta R_2^\varphi A_2 R_3^\alpha \\ \partial_{\theta,R} &= R_1^\theta A_1 R_2^\varphi R_3^\alpha \end{aligned} \tag{A.6}$$

computed using a curve η *such that* $\eta(0) = R$ *and* $\eta'(0)$ *equal to the chosen vector field. Expliciting the left invariant vector field* $\tilde{Y}_{\varphi,R}$ *in the basis* $\{\partial_\alpha, \partial_\varphi, \partial_\theta\}_{|R}$ *we have:*

$$\tilde{Y}_{\varphi,R} = \sin\alpha \tan\varphi \partial_{\alpha,R} + \cos\alpha \partial_{\varphi,R} - \frac{\sin\alpha}{\cos\varphi}\partial_{\theta,R}. \tag{A.7}$$

According to [DF11], left-invariant vector fields express in the second chart, forgetting the dependence on the point $R \in SO(3)$, are:

$$\begin{aligned} \tilde{Y}_\theta &= -\cos\alpha \tan\varphi \partial_\alpha + \sin\alpha \partial_\varphi + \frac{\cos\alpha}{\cos\varphi}\partial_\theta \\ \tilde{Y}_\varphi &= \sin\alpha \tan\varphi \partial_\alpha + \cos\alpha \partial_\varphi - \frac{\sin\alpha}{\cos\varphi}\partial_\theta \\ \tilde{Y}_\alpha &= \partial_\alpha; \end{aligned} \tag{A.8}$$

while in the first chart (A.1) are:

$$\begin{aligned}\tilde{Y}_\theta &= (\cos\alpha\cot\varphi)\partial_\alpha + (\sin\alpha)\partial_\varphi - \frac{\cos\alpha}{\sin\varphi}\partial_\theta \\ \tilde{Y}_\varphi &= -(\sin\alpha\cot\varphi)\partial_\alpha + (\cos\alpha)\partial_\varphi + \frac{\sin\alpha}{\sin\varphi}\partial_\theta \\ \tilde{Y}_\alpha &= \partial_\alpha.\end{aligned} \tag{A.9}$$

Since we are interested in working with the two-dimensional sphere, it is possible to define a sub-Riemannian metric $g_{SO(3)}$ with respect to vector fields $\tilde{Y}_\theta$ and $\tilde{Y}_\varphi$ (belonging to the tangent space of the sphere) by imposing their orthonormality. In this setting, we say that a curve[2] $\gamma : I \longrightarrow SO(3)$ is horizontal if $\dot{\gamma}(s) = b(s)\tilde{Y}_\theta + c(s)\tilde{Y}_\varphi \in \text{span}\{\tilde{Y}_\theta, \tilde{Y}_\varphi\}$ for all $s \in I$, and its length is defined as:

$$\ell_{SO(3)}(\gamma) = \int_I \sqrt{g_{SO(3)}(\dot{\gamma}(s), \dot{\gamma}(s))}ds = \int_I \sqrt{b^2(s) + c^2(s)}ds. \tag{A.10}$$

The Carnot-Charatheodory distance between two points P and $Q \in SO(3)$, is defined as the infimum of the $SO(3)$-lengths of the horizontal curves joining P and Q:

$$d_{SO(3)}(P, Q) = \inf_\gamma\{\ell_{SO(3)}(\gamma); \gamma(0) = P, \gamma(1) = Q\}, \tag{A.11}$$

and such a curve exists because of the Hörmander condition satisfied by the generating vector fields $\tilde{Y}_\theta$ and $\tilde{Y}_\varphi$. In particular, for the reader interested in the curves (geodesics) that solve(A.11), we refer to [MDS+17] .

A.2 Two dimensional manifold $\mathbb{S}^2$

It is usual to define the two dimensional sphere as the set of points in $\mathbb{R}^3$ having (Euclidean) distance from the origin equal to 1, namely $\mathbb{S}^2 = \{x \in \mathbb{R}^3; \|x\| = 1\}$. However, since the topological dimension of this geometric object is 2, we introduce the classical spherical coordinates (θ, φ) such that $n = (n_1, n_2, n_3) \in \mathbb{R}^3$ can be parametrized as:

$$\begin{aligned}n_1 &= \cos\theta\sin\varphi \\ n_2 &= \sin\theta\sin\varphi \\ n_3 &= \cos\varphi\end{aligned} \tag{A.12}$$

with $\theta \in [0, 2\pi]$ and $\varphi \in (0, \pi)$.

[2] In the following, unless stated otherwise, I is identified with the interval $[0, 1]$.

The ambiguity that arises using a local coordinate chart is overcome by the introduction of a second chart, covering the singular points similarly as already explained in Remark A.1.1.

The Euclidean metric induces on the sphere the classical intrinsic Riemannian metric:

$$g_{\mathbb{S}^2} := d\varphi^2 + \sin^2\varphi d\theta^2, \tag{A.13}$$

and a set of local vector fields that are orthonormal with respect to this metric are

$$\begin{aligned} Y_\theta &= -\frac{1}{\sin\varphi}\partial_\theta, \\ Y_\varphi &= \partial_\varphi, \end{aligned} \tag{A.14}$$

well defined for $\theta \in [0, 2\pi], \varphi \in (0, \pi)$.

If we take a curve $\gamma : I \longrightarrow \mathbb{S}^2$, such that $\dot\gamma(s) = (\dot\theta(s), \dot\varphi(s))$, it is possible to define its length with respect to the Riemannian metric introduced in (A.13):

$$\ell_{\mathbb{S}^2}(\gamma) = \int_I \sqrt{\dot\varphi^2(s) + \dot\theta(s)^2 \sin\varphi^2(s)}ds \tag{A.15}$$

The Carnot-Charatheodory distance associated with this metric is defined by taking the infimum of the $\mathbb{S}^2$-length of a curve joining the two points:

$$d_{\mathbb{S}^2}(n_0, n_1) = \inf_\gamma\{\ell_{\mathbb{S}^2}(\gamma); \gamma(0) = n_0, \gamma(1) = n_1\}, \tag{A.16}$$

for every couple of points $n_i = (\theta_i, \varphi_i) \in \mathbb{S}^2, i = 0, 1$.

Remark A.2.1. *The two-dimensional sphere $\mathbb{S}^2$ can be interpreted in terms of the rotation group $SO(3)$ if we choose to apply a rotation R to the r_3-axis. Formally, this corresponds to associate to $\mathbb{S}^2$ the Lie group quotient of $SO(3)$ over H, where H is a one-parameter subgroup of $SO(3)$. The subgroup H is defined as:*

$$H = (0, 0, \alpha) = \{R \in SO(3); R = R_3^\alpha, \forall\alpha \in [0, 2\pi]\}, \tag{A.17}$$

the set of rotations around the third axis, usually identified with $SO(2)$. The quotient space is then defined as $SO(3)/H = \{RH, R \in SO(3)\}$, namely the class of the left cosets of H. The equivalence relation reads as follow : $R1, R2 \in SO(3)$, $R1 \sim R2 \iff (R1)^{-1}R2 \in SO(2)$. In particular, the correspondence between

$SO(3)/SO(2)$ and $\mathbb{S}^2$ is given by the following map:

$$\begin{aligned} SO(3)/SO(2) \ni [R_3^\theta R_2^\varphi] = \{R_3^\theta R_2^\varphi R_3^\alpha; \alpha \in [0,\pi]\} \iff \\ n(\theta,\varphi) := (\cos\theta\sin\varphi, \sin\theta\sin\varphi, \cos\varphi)^T = R_3^\theta R_2^\varphi R_3^\alpha e_3 \in \mathbb{S}^2, \end{aligned} \tag{A.18}$$

using the first chart, and with $e_3 = (0,0,1) \in \mathbb{R}^3$. Analogous reasoning can be done considering chart (A.1.1).

Basically, it is possible to identify the element $n \in \mathbb{S}^2$ with the group element $R_n \in SO(3)$, where R_n is any rotation matrix such that $R_n e_3 = n$.

A.3 3D Euclidean motion group $SE(3)$

An element $g = (p, R)$ in $SE(3)$ is defined by a point $p \in \mathbb{R}^3$ and a rotation $R \in SO(3)$. The law group expresses the fact that the composition of two body motions is again a body motion, and so for two elements $g = (p, R)$ and $g' = (p', R')$ in $SE(3)$ their product is defined as

$$gg' = (p + Rp', RR'). \tag{A.19}$$

It is possible to notice that the group $SE(3)$ is a semi-direct product of the translation group $\mathbb{R}^3$ and the rotation group $SO(3)$, since it uses an isomorphism $R \mapsto (p \mapsto Rp)$ from the rotation group onto the automorphism on $\mathbb{R}^3$. That is why it is more appropriate to write $SE(3) = \mathbb{R}^3 \rtimes SO(3)$.

We consider the following basis for the tangent space $T_e(SE(3))$ at the unity element $e = (0_{\mathbb{R}^3}, I_{\mathrm{d}}) \in SE(3)$

$$\mathcal{B}_{T_e(SE(3))} = \{\partial_1, \partial_2, \partial_3, \partial_\theta, \partial_\varphi, \partial_\alpha\} \tag{A.20}$$

with table of Lie brackets (see [DF11]):

$$([\partial_i, \partial_j])_{i,j=1\ldots3,\theta,\varphi,\alpha} = \begin{pmatrix} 0 & 0 & 0 & 0 & \partial_3 & -\partial_2 \\ 0 & 0 & 0 & -\partial_3 & 0 & \partial_1 \\ 0 & 0 & 0 & \partial_2 & -\partial_1 & 0 \\ 0 & \partial_3 & -\partial_2 & 0 & \partial_\alpha & -\partial_\varphi \\ -\partial_3 & 0 & \partial_1 & -\partial_\alpha & 0 & \partial_\theta \\ \partial_2 & -\partial_1 & 0 & \partial_\varphi & -\partial_\theta & 0 \end{pmatrix}. \tag{A.21}$$

The corresponding left-invariant vector field $\{\tilde{Y}_i\}_i$ are obtained by the push-forward of the left multiplication, and they can be expressed in the coordinate chart (A.1) for $\varphi \neq 0, \varphi \neq \pi$:

$$\begin{aligned}
\tilde{Y}_1 =&(\cos\alpha\cos\varphi\cos\theta - \sin\alpha\sin\theta)\partial_1 + (\sin\alpha\cos\theta + \cos\alpha\cos\varphi\sin\theta)\partial_2 + \\
&- \cos\alpha\sin\varphi\partial_3 \\
\tilde{Y}_2 =&(-\sin\alpha\cos\varphi\cos\theta - \cos\alpha\sin\theta)\partial_1 + (\cos\alpha\cos\theta + \sin\alpha\cos\varphi\sin\theta)\partial_2 + \\
&\sin\alpha\sin\varphi\partial_3 \\
\tilde{Y}_3 =&(\sin\varphi\cos\theta)\partial_1 + (\sin\varphi\sin\theta)\partial_2 + \cos\varphi\partial_3 \\
\tilde{Y}_\theta =&(\cos\alpha\cot\varphi)\partial_\alpha + (\sin\alpha)\partial_\varphi - \frac{\cos\alpha}{\sin\varphi}\partial_\theta \\
\tilde{Y}_\varphi =&-(\sin\alpha\cot\varphi)\partial_\alpha + (\cos\alpha)\partial_\varphi + \frac{\sin\alpha}{\sin\varphi}\partial_\theta \\
\tilde{Y}_\alpha =&\partial_\alpha
\end{aligned} \tag{A.22}$$

or using the second chart (A.1.1), for $\varphi \neq -\pi/2, \varphi \neq \pi/2$:

$$\begin{aligned}
\tilde{Y}_1 &= (\cos\alpha\cos\varphi)\partial_1 + (\cos\theta\sin\alpha + \cos\alpha\sin\varphi\sin\theta)\partial_2 + (\sin\alpha\sin\varphi - \cos\alpha\cos\theta\sin\varphi)\partial_3 \\
\tilde{Y}_2 &= (-\sin\alpha\cos\varphi)\partial_1 + (\cos\alpha\cos\theta - \sin\alpha\cos\varphi\sin\theta)\partial_2 + (\sin\alpha\sin\varphi\cos\theta + \cos\alpha\sin\theta)\partial_3, \\
\tilde{Y}_3 &= (\sin\varphi)\partial_1 - (\cos\varphi\sin\theta)\partial_2 + (\cos\varphi\cos\theta)\partial_3
\end{aligned} \tag{A.23}$$

with $\tilde{Y}_\theta, \tilde{Y}_\varphi$ and $\tilde{Y}_\alpha$ are the same of (A.8), as computed in [DF11].

As already noticed by Duits and Franken in [DF11], interpreting a point on the sphere (A.12) as a direction, then the vector field $\tilde{Y}_3$ expresses the coupling position-orientation. So, by noticing that $\tilde{Y}_\theta$ and $\tilde{Y}_\varphi$ belong to the tangent space of the sphere, it is possible to consider the distribution that naturally couples positions and orientations generated by these vector fields, namely

$$\mathcal{D} = \text{span}\{\tilde{Y}_3, \tilde{Y}_\theta, \tilde{Y}_\varphi\}, \tag{A.24}$$

and define a sub-Riemannian metric tensor $g_{SE(3)}$ with respect to these vector fields by imposing their orthonormality. In this setting, a curve $\gamma : I \longrightarrow SO(3)$ is horizontal if $\dot{\gamma}(s) = a(s)\tilde{Y}_3 + b(s)\tilde{Y}_\theta + c(s)\tilde{Y}_\varphi$ for all $s \in I$, and its length is defined as:

$$\ell_{SE(3)}(\gamma) = \int_I \sqrt{g_{SE(3)}(\dot{\gamma}(s), \dot{\gamma}(s))}ds = \int_I \sqrt{a^2(s) + b^2(s) + c^2(s)}ds. \tag{A.25}$$

Moreover, if we consider g_1 and $g_2 \in SE(3)$, then the Carnot-Caratheodory distance between g_1 and g_2 is the infimum of the $SE(3)$-lengths of the horizontal curves joining g_1 and g_2:

$$d_{SE(3)}(g_1, g_2) = \inf_{\gamma}\{\ell_{SE(3)}(\gamma); \gamma(0) = g_1, \gamma(1) = g_2\}. \tag{A.26}$$

Such a curve exists because of the Hörmander condition satisfied by vector fields $\tilde{Y}_3, \tilde{Y}_\theta$ and $\tilde{Y}_\varphi$, since their commutation rules are the same as the ones shown in Table (A.21). For the reader interested in the curves (geodesics) that solve(A.11), we refer to the works [DGDHM16, MP20] .

A.4 3D space of positions and orientations $\mathbb{R}^3 \rtimes \mathbb{S}^2$

The 3D position-orientation space $\mathbb{R}^3 \rtimes \mathbb{S}^2$ and the associated local orthonormal frame used to describe the behavior of binocular cells have been introduced in section 5.3, exploiting the action of the group of $SE(3)$ on $\mathbb{R}^3 \rtimes \mathbb{S}^2$. However, it is commonly known that a deeper relationship exists between these spaces: $\mathbb{R}^3 \rtimes \mathbb{S}^2$ can be outlined in terms of a quotient of $SE(3)$, basically taking advantage of Remark A.2.1, namely

$$\mathbb{R}^3 \rtimes \mathbb{S}^2 := SE(3)/(\{0\} \times SO(2)). \tag{A.27}$$

In other words, $\mathbb{R}^3$ remains unchanged, while the quotient identifies $\mathbb{S}^2$ elements with cosets of $SO(3)/SO(2)$.

Remark A.4.1. *Within the quotient structure $SE(3)/(\{0\}\times SO(2))$ two elements $\xi = (p, [R])$ and $\xi' = (p', [R]')$ are equivalent if:*

$$\begin{aligned} \xi' \sim \xi &\iff (\xi')^{-1}\xi \in \{0\} \times SO(2) \\ &\iff p = p' \text{ and } \exists\ \alpha : (R')^{-1}R = R_3^\alpha \in SO(2), \end{aligned} \tag{A.28}$$

where it holds the identification of Remark A.2.1, and the relationship between the elements of $\mathbb{S}^2$ and $SO(3)/SO(2)$ is defined according to (A.18).

The group action $\cdot$ of $g = (p, R) \in SE(3)$ onto $(q, n) \in \mathbb{R}^3 \rtimes \mathbb{S}^2$ is defined by

$$g \cdot (q, n) = (Rq + p, Rn); \tag{A.29}$$

and so, because this group product influences the group action in $\mathbb{R}^3 \rtimes \mathbb{S}^2$, we use

the semi-product notation $\rtimes$ even though this is usually reserved for the semi-direct product of groups.

Remark A.4.2. *Via the group action, the equivalence relation* (A.28) *amounts to:*

$$g' \sim g \iff g' \cdot (0, e_3) = g \cdot (0, e_3), \tag{A.30}$$

where $e_3 = (0, 0, 1)$ *is the third reference axis in* $\mathbb{R}^3$.

Thereby, an arbitrary element ξ in $\mathbb{R}^3 \rtimes \mathbb{S}^2$ can be considered as the equivalence class of all rigid body motions that map reference position and orientation $(0, e_3)$ onto (p, n) with arbitrary $p \in \mathbb{R}^3$ and $n \in \mathbb{S}^2$ defined considering the spherical coordinates introduced in equation (A.12).

It is quite natural now to consider the vector field $\tilde{Y}_3$ of (5.24) which couples positions and orientations and vector field on the sphere Y_θ and Y_φ defined in (A.14). (When we work in the position-orientation space, the vector field $\tilde{Y}_3$ will be denoted with Y_3 for clarity.) The local frame $\{Y_3, Y_\theta, Y_\varphi\}$ satisfies the Hörmander's condition in the chart of definition, and in this setting we extend the classical Riemannian metric on $\mathbb{S}^2$ to a sub-Riemannian metric (A.13) on $\mathbb{R}^3 \rtimes \mathbb{S}^2$, by imposing the orthonormality of these vector fields.

Naturally, the length of a horizontal curve $\gamma : I \longrightarrow \mathbb{R}^3 \rtimes \mathbb{S}^2$, is defined as

$$\ell_{\mathbb{R}^3 \rtimes \mathbb{S}^2}(\gamma) = \int_I \sqrt{a(s)^2 + \dot{\varphi}^2(s) + \dot{\theta}(s)^2 \sin \varphi^2(s)} ds, \tag{A.31}$$

where $a(s)$ is the coefficient associated to vector field Y_3.

The Carnot-Charatheodory distance between two points $\xi_1 = (p_1, n_1)$ and $\xi_2 = (p_2, n_2)$ is:

$$d_{\mathbb{R}^3 \rtimes \mathbb{S}^2}(\xi_1, \xi_2) = \inf_\gamma \{\ell_{\mathbb{R}^3 \rtimes \mathbb{S}^2}(\gamma); \gamma(0) = \xi_1, \gamma(1) = \xi_2\}. \tag{A.32}$$

Appendix B

Fundamental solutions in $SE(3)$ and $\mathbb{R}^3 \rtimes \mathbb{S}^2$

In this appendix, we briefly introduce a first investigation on the parametric relationship between the fundamental solution of the forward Kolmogorov operator (6.8), (computed numerically in 6.2), and the fundamental solution of the corresponding hypo-elliptic convection-diffusion operator in $SE(3)$, already solved exactly and with numerical, analytical and stochastical approximation in [DBM19, PD17, PSMD15].

First, we find a (local) isometry between $\mathbb{R}^3 \rtimes \mathbb{S}^2$ and its $SE(3)$ counterparts, and then, we relate the behavior of fundamental solutions for operators defined in these spaces, using the notion of lifting of operators, embedding in the work of Bonfiglioli-Biagi [BB17].

B.1 Metric property between $\mathbb{R}^3 \rtimes \mathbb{S}^2$ and $SE(3)$

The manifold $\mathbb{R}^3 \rtimes \mathbb{S}^2$ is a left quotient of the Lie group of rigid body motion $SE(3)$, as recalled in subsection A.4. In this section we show that we can pass the metric defined in subsection A.3 on $SE(3)$ to the quotient on $\mathbb{R}^3 \rtimes \mathbb{S}^2$, investigating the existence of an isometry between these two spaces.

B.1.1 Local isometry from $SE(3)$ to $\mathbb{R}^3 \rtimes \mathbb{S}^2$

The existence of a local isometry, between $SE(3)$ and $\mathbb{R}^3 \rtimes \mathbb{S}^2$ together with the distances defined in the previous sections, is proved by adapting a technique pro-

posed by Arcozzi-Baldi in [AB08] to our setting. In particular, this technique is developed locally, providing the appropriate lifting relating $\mathbb{R}^3 \rtimes \mathbb{S}^2$ to $SE(3)$.

Theorem B.1.1. *Let $\tilde{H} = \{0\} \times H$ be a one parameter subgroup of $SE(3)$, H defined in (A.17). Then $X = SE(3)/\tilde{H}$ is locally (in charts) isometric to $\mathbb{R}^3 \rtimes \mathbb{S}^2$.*

Proof. The subgroup $\tilde{H}$ has a unique intersection with the plane $\{\alpha = 0\}$ and so we can identify X with $\mathbb{R}^3 \rtimes \mathbb{S}^2$ via the following map:

$$\phi : (p_{\mathbb{R}^3}, \theta, \varphi, 0) \cdot \tilde{H} \mapsto (p_{\mathbb{R}^3}, \theta, \varphi). \tag{B.1}$$

We will see that ϕ is a surjective isometry: the surjective property follows immediately from the transitive action of $SE(3)$ on $\mathbb{R}^3 \rtimes \mathbb{S}^2$.

On the other hand, to show that ϕ is an isometry, we need to prove the equality of the Carnot-Caratheodory distances:

$$d_X(g_1\tilde{H}, g_2\tilde{H}) = d_{\mathbb{R}^3 \rtimes \mathbb{S}^2}(\phi(g_1\tilde{H}), \phi(g_2\tilde{H})). \tag{B.2}$$

It is sufficient to show that every horizontal curve γ in $\mathbb{R}^3 \rtimes \mathbb{S}^2$ has a horizontal lifting $\tilde{\gamma}$ in $SE(3)$. In particular, we say that $\tilde{\gamma}$ is an horizontal lift if $\tilde{\gamma}$ is horizontal (in $SE(3)$) and we have $\phi \circ \pi(\tilde{\gamma}) = \gamma$, where π is the usual projection. So, we will have $\ell_{SE(3)}(\tilde{\gamma}) = \ell_{\mathbb{R}^3 \rtimes \mathbb{S}^2}(\gamma)$ and we can conclude.

First, we need to determine the existence of the opportune lifting: if we consider a curve $\gamma = (p, (\theta, \varphi))$ in $\mathbb{R}^3 \rtimes \mathbb{S}^2$ in a way that its length satisfy (A.31), we can define α such that:

$$\alpha(s) = \alpha_0 - \int_0^s \dot{\theta} \cos\varphi ds. \tag{B.3}$$

This integral is well defined, and so taking the horizontal lifting $\tilde{\gamma} = (p, \theta, \varphi, \alpha)$, we get that $\tilde{\gamma}$ is an horizontal curve for $SE(3)$:

$$\dot{\tilde{\gamma}} = a(s)\tilde{Y}_3 + b(s)\tilde{Y}_\theta + c(s)\tilde{Y}_\varphi \tag{B.4}$$

with $b(s) = \dot{\varphi}\sin\alpha - \dot{\theta}\sin\varphi\cos\alpha$ and $c(s) = \dot{\varphi}\cos\alpha + \dot{\theta}\sin\varphi\sin\alpha$. With these

coefficients it is possible to compute the length of the curve in $SE(3)$:

$$\begin{aligned}\ell_{SE(3)}(\tilde{\gamma}) &= \int \sqrt{a^2(s) + b^2(s) + c^2(s)}ds \\ &= \int \sqrt{a^2(s) + (\dot{\varphi}\sin\alpha - \dot{\theta}\sin\varphi\cos\alpha)^2(s) + (\dot{\varphi}\cos\alpha + \dot{\theta}\sin\varphi\sin\alpha)^2(s)}ds \\ &= \int \sqrt{a^2(s) + \dot{\varphi}^2(s) + \dot{\theta}(s)^2\sin\varphi^2(s)}ds \\ &= \ell_{\mathbb{R}^3\rtimes\mathbb{S}^2}(\gamma).\end{aligned} \tag{B.5}$$

So, the equality of the Carnot-Caratheodory distances (B.2) follows immediately. □

Corollary B.1.1. *Let H be a one-parameter subgroup of $SO(3)$, as defined in (A.17). Then $X = SO(3)/H$ is isometric to $\mathbb{S}^2$.*

Proof. The proof follows directly from the proof of Theorem B.1.1, by restriction to the spaces $SO(3)$ and $\mathbb{S}^2$. □

In particular, there exists a relationship between the left-invariant vector fields defined in (A.22) and the spherical vector fields (A.14).

Remark B.1.1. *The spherical vector fields Y_i presented in (A.14) can be obtained from left invariant vector fields $\tilde{Y}_i$ defined in (A.22) through a rotation described by the matrix R_3^α and quoting with respect to the α variable:*

$$\begin{pmatrix} Y_\theta \\ Y_\varphi \\ 0 \end{pmatrix} = R_3^\alpha \begin{pmatrix} \tilde{Y}_\theta \\ \tilde{Y}_\varphi \\ \tilde{Y}_\alpha \end{pmatrix}_{|\mathbb{R}^3\rtimes\mathbb{S}^2} = \begin{pmatrix} \cos\alpha\tilde{Y}_\theta - \sin\alpha\tilde{Y}_\varphi \\ \sin\alpha\tilde{Y}_\theta + \cos\alpha\tilde{Y}_\varphi \\ \tilde{Y}_\alpha \end{pmatrix}_{|\mathbb{R}^3\rtimes\mathbb{S}^2}. \tag{B.6}$$

Actually, there is also another way to get the same set of vector fields in $\mathbb{R}^3 \rtimes \mathbb{S}^2$, and this is done by choosing the representative of α in the quotient space as $\alpha = 0$. This can be thought of as the implicit function theorem proved in homogeneous Lie groups by Franchi, Serapioni and Serra Cassano in [FSSC01] and then in Carnot-Carathéodory spaces by Citti and Manfredini in [CM06]: if a function f is of class C^1 with respect to the general vector fields, then $f_{|\alpha=0}$ is of class C^1 with respect to vector fields restricted to $\alpha = 0$.

This existing relationship that links the vector fields of $SE(3)$ with the vector fields defined in $\mathbb{R}^3 \rtimes \mathbb{S}^2$, thanks to the formula (B.1), allows us to understand the definition of the lifting proposed in Theorem B.1.1.

Remark B.1.2. *The lifting of* (B.3) *has been determined expressing the length of a curve* $\gamma : I \longrightarrow \mathbb{S}^2$ *defined in* (A.15) *in terms of the local orthonormal vector fields* (A.14)*:*

$$\dot{\gamma} = -\dot{\theta}\sin\varphi Y_\theta + \dot{\varphi} Y_\varphi. \tag{B.7}$$

Using result (B.6)*, thus substituting the expression of the vector fields* Y_θ, Y_φ *in terms of the vector fields defined in* $SO(3)$*, we get:*

$$\begin{aligned}\dot{\gamma} &= -\dot{\theta}\sin\varphi(\cos\alpha\tilde{Y}_\theta - \sin\alpha\tilde{Y}_\varphi) + \dot{\theta}(\sin\alpha\tilde{Y}_\theta + \cos\alpha\tilde{Y}_\varphi)\\ &= (\dot{\varphi}\sin\alpha - \dot{\theta}\sin\varphi\cos\alpha)\tilde{Y}_\theta + (\dot{\theta}\sin\varphi\sin\alpha + \dot{\varphi}\cos\alpha)\tilde{Y}_\varphi\\ &= -\dot{\theta}\cos\varphi\partial_\alpha + \dot{\varphi}\partial_\varphi + \dot{\theta}\partial_\theta,\end{aligned} \tag{B.8}$$

from which we choose to define $\dot{\alpha}$ *as the coefficient of* ∂_α*.*

B.2 Fundamental solutions for lifted operators

Since it is possible to switch from a set of vector fields in $\mathbb{R}^3 \rtimes \mathbb{S}^2$ to a set in $SE(3)$, through a change of variable and a passage to the quotient, in this section we study the existence of a relationship between the fundamental solutions of operators in $SE(3)$ and in $\mathbb{R}^3 \rtimes \mathbb{S}^2$, operators written in terms of vector fields.

B.2.1 Involved operators

Thanks to Remark B.1.1, it is possible to consider a change of variable $F : SE(3) \longrightarrow SE(3)$ such that its differential acts on the tangent space via (B.6), introducing the following set of vector fields:

$$\begin{aligned}Z_3 =& DF(\tilde{Y}_3) = \tilde{Y}_3\\ Z_\theta =& DF(\tilde{Y}_\theta) = \cot\varphi\partial_\alpha - \frac{1}{\sin\varphi}\partial_\theta\\ Z_\varphi =& DF(\tilde{Y}_\varphi) = \partial_\varphi.\end{aligned} \tag{B.9}$$

Two hypoelliptic operators, one on $\mathbb{R}^3 \rtimes \mathbb{S}^2$ and the other one on $SE(3)$, can be defined in terms of the vector fields considered:

$$L_{SE(3)} := Z_3 + Z_\theta^2 + Z_\varphi^2 \tag{B.10}$$

and

$$L_{\mathbb{R}^3 \rtimes \mathbb{S}^2} := Y_3 + Y_\theta^2 + Y_\varphi^2, \tag{B.11}$$

and we observe that one can be written in terms of the other, differing by a remainder.

Remark B.2.1. *The operator $L_{SE(3)}$ can be expressed in terms of the operator $L_{\mathbb{R}^3 \rtimes \mathbb{S}^2}$:*

$$L_{SE(3)} = L_{\mathbb{R}^3 \rtimes \mathbb{S}^2} + R, \tag{B.12}$$

with R remainder that involves the α directions, namely in coordinates

$$R = \cot\varphi \partial_\alpha^2 - 2\frac{\cot\varphi}{\sin\varphi}\partial_\theta\partial_\alpha. \tag{B.13}$$

B.2.2 Lifting of operators

In the following we introduce the definition of the lifting of an operator and the results it implies, applying to this setting the introductive results of the work presented by Biagi-Bonfiglioli in [BB17].

Definition B.2.1. *Let L be a smooth linear partial differential operator on a manifold M and $\tilde{L} = L + R$ another second order differential operator defined in an higher manifold $M \times N$. We say that $\tilde{L}$ is a **lifting** of L if the followings conditions hold:*

i) $\tilde{L}$ has a smooth coefficients, possibly depending on $\xi \in M$ and $\alpha \in N$;

ii) for every fixed $f \in C^\infty(M)$, one has

$$\tilde{L}(f \circ \pi)(\xi, \alpha) = (Lf)(\xi), \text{ for every } (\xi, \alpha) \in M \times N, \tag{B.14}$$

where $\pi(\xi, \alpha) = \xi$ is the canonical projection of $M \times N$ onto M.

Remark B.2.2. *The partial differential operator $L_{SE(3)}$ introduce in (B.10) is a lifting of $L_{\mathbb{R}^3 \rtimes \mathbb{S}^2}$ introduced in equation (B.11) by considering $M = \mathbb{R}^3 \rtimes \mathbb{S}^2$ and $N = \mathbb{S}^1$.*

This definition allows us to recover the relationship between the fundamental solutions, obtaining the solution defined on the quotient space simply by integrating the fundamental solution associated to the operator on $SE(3)$, with respect to the additional variable.

Theorem B.2.1. *Let $L_{\mathbb{R}^3 \rtimes \mathbb{S}^2}$ be the smooth operator defined in (B.11) on $\mathbb{R}^3 \rtimes \mathbb{S}^2$ and let $L_{SE(3)}$ be the lifting on $SE(3)$ introduced in (B.10), according to Definition (B.2.1). We suppose that there exists a fundamental solution $\tilde{\Gamma}$ for $L_{SE(3)}$ satisfying the following properties:*

i) for every fixed $\xi, \xi' \in \mathbb{R}^3 \rtimes \mathbb{S}^2$ with $\xi \neq \xi'$, one has

$$\alpha \mapsto \tilde{\Gamma}(\xi, 0; \xi', \alpha) \text{ belongs to } L^1(S^1); \tag{B.15}$$

ii) for every fixed $\xi \in \mathbb{R}^3 \rtimes \mathbb{S}^2$ and every compact set $K \subseteq \mathbb{R}^3 \rtimes \mathbb{S}^2$, one has

$$(\xi', \alpha) \mapsto \tilde{\Gamma}(\xi, 0; \xi', \alpha) \text{ belongs to } L^1(K \times S^1). \tag{B.16}$$

Then the function $\Gamma : \{(\xi, \xi') \in \mathbb{R}^3 \rtimes \mathbb{S}^2 \times \mathbb{R}^3 \rtimes \mathbb{S}^2 : \xi \neq \xi'\} \longrightarrow \mathbb{R}$ defined by

$$\Gamma(\xi, \xi') = \int_{S^1} \tilde{\Gamma}(\xi, 0; \xi', \alpha) d\alpha, \tag{B.17}$$

is a fundamental solution for $L_{\mathbb{R}^3 \rtimes \mathbb{S}^2}$ on $\mathbb{R}^3 \rtimes \mathbb{S}^2$.

Proof. It follows from i) that Γ is well-posed. Then, to have Γ fundamental solution, we need

- $\Gamma(\xi, \cdot) \in L^1_{loc}(\mathbb{R}^3 \rtimes \mathbb{S}^2)$. This follows from ii);

- $L_{\mathbb{R}^3 \rtimes \mathbb{S}^2} \Gamma(\xi, \cdot) = -\mathrm{Dir}_\xi \in \mathcal{D}'(\mathbb{R}^3 \rtimes \mathbb{S}^2)$.

In order to prove the second property, we consider $f \in C_0^\infty(\mathbb{R}^3 \rtimes \mathbb{S}^2)$. This function can be consider as a function of $C_0^\infty(SE(3))$ such that $\partial_\alpha^k f = 0$ for all $k \in \mathbb{N}$. Then, we have:

$$\begin{aligned} -f(\xi) = -f \underbrace{(\xi, 0)}_{g \in SE(3)} &= \int_{SE(3)} \tilde{\Gamma}(g, g') L^*_{SE(3)} f(g') dg' = \\ &= \int_{SE(3)} \tilde{\Gamma}(\xi, 0; \xi', \alpha) L^*_{SE(3)} f(\xi') d\xi' d\alpha' \end{aligned}, \tag{B.18}$$

with $L^*_{SE(3)}$ formal adjoint of $L_{SE(3)}$. Thanks to (B.12), we know that $L^*_{SE(3)} =$

$L^*_{\mathbb{R}^3\rtimes\mathbb{S}^2} + R^*$, and so

$$\begin{aligned}
&= \int_{SE(3)} \tilde{\Gamma}(\xi, 0; \xi, \alpha')(L^*_{\mathbb{R}^3\rtimes\mathbb{S}^2} + R^*) f(\xi') d\xi' d\alpha' \\
&= \int_{\mathbb{R}^3\rtimes\mathbb{S}^2\times\mathbb{S}^1} \tilde{\Gamma}(\xi, 0; \xi', \alpha')(L^*_{\mathbb{R}^3\rtimes\mathbb{S}^2}) f(\xi') d\xi' d\alpha' \\
&= \int_{\mathbb{R}^3\rtimes\mathbb{S}^2} \Gamma(\xi, \xi')(L^*_{\mathbb{R}^3\rtimes\mathbb{S}^2}) f(\xi') d\xi'.
\end{aligned} \tag{B.19}$$

The second equality follows from the fact that the formal adjoint R^* of the operator R has the form:

$$\begin{aligned}
R^*\Psi =& (-1)^2 \cot\varphi \partial^2_\alpha \Psi(\xi, \alpha) + (-1)^2 \partial_\theta \left(-2 \frac{\cot\varphi}{\sin\varphi} \partial_\alpha \Psi(\xi, \alpha) \right) \\
=& \cot\varphi \partial^2_\alpha \Psi(\xi, \alpha) - 2 \frac{\cot\varphi}{\sin\varphi} \partial_\theta \partial_\alpha \Psi(\xi, \alpha) \\
=& R\Psi,
\end{aligned} \tag{B.20}$$

and so we can conclude that

$$\int_{\mathbb{R}^3\rtimes\mathbb{S}^2\times\mathbb{S}^1} \tilde{\Gamma}(\xi, 0; \xi', \alpha')(R^*) f(\xi') d\xi' d\alpha' = 0. \tag{B.21}$$

□

Remark B.2.3. *It would be interesting to consider the same operator of (B.10) in $SE(3)$ written in terms of left-invariant vector fields $\tilde{Y}_{i,SE(3)}$ and searching for a relationship between its fundamental solution and the one defined on $\mathbb{R}^3 \rtimes \mathbb{S}^2$ for the operator defined in (B.11). In particular, the left-invariant operator in $SE(3)$ has an extra first-order term not involving the direction identified by α, so there is the need to introduce a different strategy. We may use a parametrix method, with the idea of choosing the fundamental solution in $\mathbb{R}^3 \rtimes \mathbb{S}^2$ as parametrix, to obtain an approximation via iteration for the left-invariant operator. This will be the object of future research.*

List of Figures

Bibliography

[AB08] N. Arcozzi and A. Baldi. From Grushin to Heisenberg via an isoperimetric problem. *Journal of mathematical analysis and applications*, 340(1):165–174, 2008. doi:10.1016/j.jmaa.2007.07.053.

[ABB19] A. Agrachev, D. Barilari, and U. Boscain. *A comprehensive introduction to sub-Riemannian geometry*, volume 181. Cambridge University Press, 2019.

[Ama72] S. I. Amari. Characteristics of random nets of analog neuron-like elements. *IEEE Transactions on Systems, Man, and Cybernetics*, 2(5):643–657, 1972. doi:10.1109/TSMC.1972.4309193.

[AOF99a] A. Anzai, I. Ohzawa, and R.D. Freeman. Neural mechanisms for encoding binocular disparity: Receptive field position versus phase. *Journal of Neurophysiology*, 82(2):874–890, Aug 1999. doi:10.1152/jn.1999.82.2.874.

[AOF99b] A. Anzai, I. Ohzawa, and R.D. Freeman. Neural mechanisms for processing binocular information i. simple cells. *Journal of Neurophysiology*, 82(2):891–908, Aug 1999. doi:10.1152/jn.1999.82.2.891.

[APVE07] A. Anzai, X. Peng, and D. C. Van Essen. Neurons in monkey visual area V2 encode combinations of orientations. *Nature Neuroscience*, 10(10):1313–1321, 2007. doi:10.1038/nn1975.

[ASFC+17] 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. doi:10.1109/TIP.2017.2761543.

[AZ00] S. Alibhai and S. W. Zucker. Contour-based correspondence for stereo. In *Computer Vision - ECCV 2000*, pages 314–330. Springer Berlin Heidelberg, 2000. doi:10.1007/3-540-45054-8_21.

[BB17] S. Biagi and A. Bonfiglioli. The existence of a global fundamental solution for homogeneous Hörmander operators via a global lifting method. *Proceedings of the London Mathematical Society*, 114(5):855–889, 2017. doi:10.1112/plms.12024.

[BBB09] L. Barnett, C.L. Buckley, and S. Bullock. Neural complexity and structural connectivity. *Physical Review E*, 79(5), 2009. doi:10.1103/PhysRevE.79.051914.

[BBP67] H. B. Barlow, C. Blakemore, and J. D. Pettigrew. The neural mechanism of binocular depth discrimination. *The Journal of physiology*, 193(2):327, 1967. doi:10.1113/jphysiol.1967.sp008360.

[BC01] H. Bridge and B.G. Cumming. Responses of macaque V1 neurons to binocular orientation differences. *Journal of Neuroscience*, 21(18):7293–7302, 2001. doi:10.1523/JNEUROSCI.21-18-07293.2001.

[BC03] P. C. Bressloff and J. D. Cowan. The functional geometry of local and horizontal connections in a model of V1. *Journal of Physiology-Paris*, 97(2):221–236, 2003. Neurogeometry and visual perception. doi:10.1016/j.jphysparis.2003.09.017.

[BCCS14] 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. doi:10.1007/s10851-013-0482-z.

[BCF+20] M. Bertalmio, L. Calatroni, V. Franceschi, B. Franceschiello, and D. Prandi. Cortical-inspired Wilson-Cowan-type equations for orientation-dependent contrast perception modelling. *Journal of Mathematical Imaging and Vision*, Jun 2020. URL: https://hal.science/hal-02316989, doi:10.1007/s10851-020-00960-x.

[BCG+02] P. C. Bressloff, J. D. Cowan, M. Golubitsky, P. J. Thomas, and M. C. Wiener. What geometric visual hallucinations tell us about the visual cortex. *Neural Computation*, 14(3):473–491, 2002. doi:10.1162/089976602317250861.

[BCGR14] U. Boscain, R. A. Chertovskih, J. P. Gauthier, and A.O. Remizov. Hypoelliptic diffusion and human vision: a semidiscrete new twist. *SIAM Journal on Imaging Sciences*, 7(2):669–695, 2014. doi:10.1137/130924731.

[BCP01] H. Bridge, B. G. Cumming, and A. J. Parker. Modeling V1 neuronal responses to orientation disparity. *Visual Neuroscience, Cambridge University Press*, 18:879–891, 2001. doi:10.1017/S0952523801186049.

[BCP18] E. J. Bekkers, D. Chen, and J. M . Portegies. Nilpotent approximations of sub-riemannian distances for fast perceptual grouping of blood vessels in 2D and 3D. *Journal of Mathematical Imaging and Vision*, 60:882–899, 2018. doi:10.1007/s10851-018-0787-z.

[BCS14] D. Barbieri, G. Citti, and A. Sarti. How uncertainty bounds the shape index of simple cells. *The Journal of Mathematical Neuroscience*, 4(1):5, 2014.

[BCS18] E. Baspinar, G. Citti, and A.. Sarti. A geometric model of multi-scale orientation preference maps via gabor functions. *Journal of Mathematical Imaging and Vision*, 60(6):900–912, 2018. doi:10.1007/s10851-018-0803-3.

[BCSS12] D. Barbieri, G. Citti, G. Sanguinetti, and A. Sarti. An uncertainty principle underlying the functional architecture of V1. *Journal of Physiology-Paris*, 106(5-6):183–193, 2012. doi:10.1016/j.jphysparis.2012.03.001.

[BCSZ23] M. V. Bolelli, G. Citti, A. Sarti, and S. W. Zucker. Good continuation in 3D: the neurogeometry of stereo vision. *arXiv preprint arXiv:2301.04542*, 2023.

[BDRS14] U. Boscain, R. Duits, F. Rossi, and Y. Sachkov. Curve cuspless reconstruction via sub-riemannian geometry. *ESAIM: Control, Optimisation and Calculus of Variations*, 20(3):748–770, 2014. doi:10.1051/cocv/2013082.

[BFM72] C. Blakemore, A. Fiorentini, and L. Maffei. A second neural mechanism of binocular depth discrimination. *The Journal of Physiology*, 226(3):725–749, 1972. doi:10.1113/jphysiol.1972.sp010006.

[Bla92] G.G. Blasdel. Orientation selectivity, preference, and continuity in monkey striate cortex. *Journal of Neuroscience*, 12(8):3139–3161, 1992. doi:10.1523/JNEUROSCI.12-08-03139.1992.

[BLU02] A. Bonfiglioli, E. Lanconelli, and F. Uguzzoni. Uniform Gaussian estimates for the fundamental solutions for heat operators on Carnot groups. *Advances in Differential Equations*, 7(10):1153 – 1192, 2002. doi:10.57262/ade/1356651633.

[BN03] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6):1373–1396, 2003. doi:10.1162/089976603321780317.

[BP07a] U. Boscain and S. Polidoro. Gaussian estimates for hypoelliptic operators via optimal control. *Rendiconti Lincei-Matematica e Applicazioni*, 18(4):333–342, 2007. doi:10.4171/RLM/499.

[BP07b] U. Boscain and S. Polidoro. Gaussian estimates for hypoelliptic operators via optimal control. *Rendiconti Lincei - Matematica e Applicazioni*, pages 333–342, 2007. doi:10.4171/rlm/499.

[Bru21] N. Bruno. *Introduzione alla percezione visiva: come facciamo a vedere.* Bologna: Il Mulino., 2021.

[BSZ04] O. Ben-Shahar and S. Zucker. Geometrical computations explain projection patterns of long-range horizontal connections in visual cortex. *Neural computation*, 16(3):445–476, 2004. doi:10.1162/089976604772744866.

[BZSF97] W. H. Bosking, Y. Zhang, B. Schofield, and D. Fitzpatrick. Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *Journal of Neuroscience*, 17(6):2112–2127, 1997. doi:10.1523/JNEUROSCI.17-06-02112.1997.

[CD01] B. G. Cumming and G. C. DeAngelis. The physiology of stereopsis. *Annual Review of Neuroscience*, 24(1):203–238, Mar 2001. doi:10.1146/annurev.neuro.24.1.203.

[CG97] F. R. K. Chung and F. C. Graham. *Spectral graph theory.* American Mathematical Soc., 1997.

[CH12] M. Carandini and D. J. Heeger. Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, 13(1):51–62, 2012. doi:10.1038/nrn3136.

[CL06] R. R. Coifman and S. Lafon. Diffusion maps. *Applied and computational harmonic analysis*, 21(1):5–30, 2006. doi:10.1016/j.acha.2006.04.006.

[CM06] G. Citti and M. Manfredini. Implicit function theorem in Carnot–Carathéodory spaces. *Communications in Contemporary Mathematics*, 8(05):657–680, 2006. doi:10.1142/S0219199706002234.

[Coc14] G. Cocci. *Spatio-temporal models of the functional architecture of the visual cortex*. Phd thesis, University of Bologna, 2014. doi:10.6092/unibo/amsdottorato/6546.

[Con19] J. B. Conway. *A course in functional analysis*, volume 96. Springer, 2019.

[CP08] C. Cinti and S. Polidoro. Pointwise local estimates and gaussian upper bounds for a class of uniformly subelliptic ultraparabolic operators. *Journal of Mathematical Analysis and Applications*, 338(2):946–969, 2008. doi:10.1016/j.jmaa.2007.05.059.

[CS06] 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. doi:10.1007/s10851-005-3630-2.

[CS14] G. Citti and A. Sarti. *Neuromathematics of vision*, volume 32. Springer, 2014.

[DAOF95] G. C. De Angelis, I. Ohzawa, and R. D. Freeman. Receptive-field dynamics in the central visual pathways. *Trends in neurosciences*, 18(10):451–458, 1995. doi:10.1016/0166-2236(95)94496-r.

[Dau85] J. G. Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America A*, 2(7):1160, Jul 1985. doi:10.1364/josaa.2.001160.

[DBM19] R. Duits, E. J. Bekkers, and A. Mashtakov. Fourier transform on the homogeneous space of 3D positions and orientations for exact solutions to linear pdes. *Entropy*, 21(1):38, 2019. doi:10.3390/e21010038.

[DBRS13] R. Duits, U. Boscain, F. Rossi, and Y. Sachkov. Association fields via cuspless sub-Riemannian geodesics in SE(2). *Journal of Mathematical Imaging and Vision*, 49(2):384–417, Dec 2013. doi:10.1007/s10851-013-0475-y.

[DBRS14] R. Duits, U. Boscain, F. Rossi, and Y. Sachkov. Association fields via cuspless sub-riemannian geodesics in se (2). *Journal of Mathematical Imaging and Vision*, 49(2):384–417, 2014. doi:10.1007/s10851-013-0475-y.

[DCGDH11] R. Duits, E. Creusen, A. Ghosh, and T. Dela Haije. Diffusion, convection and erosion on R3xS2 and their application to the

enhancement of crossing fibers. *arXiv preprint arXiv:1103.0656*, 2011.

[DF10] R. Duits and E. Franken. Left-invariant parabolic evolutions on SE(2) and contour enhancement via invertible orientation scores part I: Linear left-invariant diffusion equations on SE(2). *Quarterly of Applied Mathematics*, 68(2):255–292, 2010. doi:10.1090/S0033-569X-10-01172-0.

[DF11] R. Duits and E. Franken. Left-invariant diffusions on the space of positions and orientations and their application to crossing-preserving smoothing of HARDI images. *International Journal of Computer Vision*, 2011. doi:10.1007/s11263-010-0332-z.

[DGDHM16] R. Duits, A. Ghosh, T. C. J. Dela Haije, and A. Mashtakov. On sub-riemannian geodesics in SE(3) whose spatial projections do not have cusps. *Journal of Dynamical and Control Systems*, 22(4):771–805, 2016. doi:10.1007/s10883-016-9329-4.

[DlPHHVDW08] J. De la Porte, B. M. Herbst, W. Hereman, and S.J. Van Der Walt. An introduction to diffusion maps. In *Proceedings of the 19th symposium of the pattern recognition association of South Africa (PRASA 2008), Cape Town, South Africa*, pages 15–25, 2008.

[DM98] R. Douglas and K. Martin. Neocortex. In G. M. Shepherd, editor, *The synaptic organization of the brain*, pages 459–509. Oxford University Press, 1998.

[DMMP18] R. Duits, S. P. L. Meesters, J. M. Mirebeau, and J. M . Portegies. Optimal paths for variants of the 2D and 3D Reeds–Shepp car with applications in image analysis. *Journal of Mathematical Imaging and Vision*, 60:816–848, 2018. doi:10.1007/s10851-018-0795-z.

[DW14] L. M. Deas and L. M. Wilcox. Gestalt grouping via closure degrades suprathreshold depth percepts. *Journal of Vision*, 14(9), 2014. doi:10.1167/14.9.14.

[DW15] L. M. Deas and L. M. Wilcox. Perceptual grouping via binocular disparity: The impact of stereoscopic good continuation. *Journal of Vision*, 15(11):11, Aug 2015. doi:10.1167/15.11.11.

[EC79] G. B. Ermentrout and J. D. Cowan. Temporal oscillations in neuronal nets. *Journal of Mathematical Biology*, 7(3):265–280, 1979. doi:10.1007/BF00275728.

[EC80] G. B. Ermentrout and J. D. Cowan. Large scale spatially organized activity in neural nets. *SIAM Journal on Applied Mathematics*, 38(1):1–21, 1980. doi:10.1137/0138001.

[Fau93] O. Faugeras. *Three-dimensional computer vision: a geometric viewpoint*. MIT press, 1993.

[FCS17] M. Favali, G. Citti, and A. Sarti. Local and global gestalt laws: A neurally based spectral approach. *Neural Computation*, 29(2):394–422, 2017. doi:10.1162/NECO_a_00921.

[FF10] G. Faye and O. Faugeras. Some theoretical and numerical results for delayed neural field equations. *Physica D: Nonlinear Phenomena*, 239(9):561–578, 2010. doi:10.1016/j.physd.2010.01.010.

[FHH93] D. J. Field, A. Hayes, and R. F. Hess. Contour integration by the human visual system: Evidence for a local "association field". *Vision Research*, 33(2):173–193, 1993. doi:10.1016/0042-6989(93)90156-Q.

[FL01] O. Faugeras and Q. T. Luong. *The geometry of multiple images: the laws that govern the formation of multiple images of a scene and some of their applications*. MIT press, 2001. doi:10.7551/mitpress/3259.001.0001.

[FMCS19] B. Franceschiello, A. Mashtakov, G. Citti, and A. Sarti. Geometrical optical illusion via sub-riemannian geodesics in the roto-translation group. *Differential Geometry and its Applications*, 65:55–77, 2019. doi:10.1016/j.difgeo.2019.03.007.

[FR96] O. Faugeras and L. Robert. What can two images tell us about a third one? *International Journal of Computer Vision*, 18(1):5–19, 1996. doi:10.1007/BF00126137.

[FSSC01] B. Franchi, R. Serapioni, and F. Serra Cassano. Rectifiability and perimeter in the heisenberg group. *Mathematische Annalen*, 321(3):479–531, 2001. doi:10.1007/s002080100228.

[GDI+96] C. D. Gilbert, A. Das, M. Ito, M. Kapadia, and G. Westheimer. Spatial integration and cortical dynamics. *Proceedings of the National Academy of Sciences*, 93(2):615–622, 1996. doi:10.1073/pnas.93.2.615.

[GM85] S. Grossberg and E. Mingolla. Neural dynamics of form perception: boundary completion, illusory figures, and neon color spreading. *Psychological review*, 92(2):173, 1985. doi:10.1016/S0166-4115(08)61758-6.

[GM91] Z. Gigus and J. Malik. *Detecting curvilinear structure in images.* University of California, Berkeley, Computer Science Division, 1991.

[GT13] C. Graham and D. Talay. *Stochastic simulation and Monte Carlo methods.* Stochastic Modelling and Applied Probability. Springer, Heidelberg, 2013.

[Hee92] D. J. Heeger. Normalization of cell responses in cat striate cortex. *Visual neuroscience*, 9(2):181–197, 1992. doi:10.1017/s0952523800009640.

[HF95] R. F. Hess and D. J. Field. Contour integration across depth. *Vision Research*, 35(12):1699–1711, 1995. doi:10.1016/0042-6989(94)00261-J.

[HHK97] R. F. Hess, A. Hayes, and F. A. A. Kingdom. Integrating contours within and through depth. *Vision Research*, 37(6):691–696, 1997. doi:10.1016/S0042-6989(96)00215-5.

[HI16] R. A. Hamzah and H. Ibrahim. Literature survey on stereo vision disparity map algorithms. *Journal of Sensors*, 2016, 2016. doi:10.1155/2016/8742920.

[HL13] J. K. Hale and S. M. V. Lunel. *Introduction to functional differential equations*, volume 99. Springer Science & Business Media, 2013.

[HMD14] R. F. Hess, K. A. May, and S. O. Dumoulin. Contour integration: Psychophysical, neurophysiological, and computational perspectives. *The Oxford Handbook of Perceptual Organization.*, 2014.

[Hof89] W. C. Hoffman. The visual cortex is a contact bundle. *Applied Mathematics and Computation*, 32(2):137–167, 1989. doi:10.1016/0096-3003(89)90091-X.

[Hör67] Lars Hörmander. Hypoelliptic second order differential equations. *Acta Mathematica*, 119:147–171, 1967.

[How12] I. P. Howard. *Perceiving in depth, volume 1: basic mechanisms.* Oxford University Press, 2012.

[HR95] I. P. Howard and B. J. Rogers. *Binocular vision and stereopsis.* Oxford University Press, USA, 1995.

[Hub95] D. H. Hubel. *Eye, brain, and vision.* Scientific American Library/Scientific American Books, 1995.

[HW62] D. H. Hubel and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1):106–154, Jan 1962. doi:10.1113/jphysiol.1962.sp006837.

[HW70] D. H. Hubel and T. N. Wiesel. Stereoscopic vision in macaque monkey: cells sensitive to binocular depth in area 18 of the macaque monkey cortex. *Nature*, 225(5227):41–42, 1970. doi:10.1038/225041a0.

[HW77] D. H. Hubel and T. N. Wiesel. Ferrier lecture-functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 198(1130):1–59, 1977. doi:10.1098/rspb.1977.0085.

[HZ03] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, USA, 2 edition, 2003.

[IBR89] R. Ivry, J. Beck, and A. Rosenfeld. Line segregation. *Spatial Vision*, 4(2-3):75 – 101, 1989. doi:10.1163/156856889X00068.

[JP87] J. P. Jones and L. A. Palmer. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, 58(6):1233–1258, Dec 1987. doi:10.1152/jn.1987.58.6.1233.

[JR15] D Jaeger and J. Ranu. *Encyclopedia of Computational Neuroscience*. Springer New York, 2015.

[Jul71] B. Julesz. *Foundations of cyclopean perception*. Chicago: The University of Chicago Press, 1971.

[Kan79] G. Kanizsa. *Organization in vision: Essays on Gestalt perception*. Praeger Publishers, 1979.

[KBSO16] D. Kato, M. Baba, K. S. Sasaki, and I. Ohzawa. Effects of generalized pooling on binocular disparity selectivity of neurons in the early visual cortex. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1697):20150266, Jun 2016. doi:10.1098/rstb.2015.0266.

[KE92] Z.F. Kisvarday and U. T. Eysel. Cellular organization of reciprocal patchy networks in layer iii of cat visual cortex (area 17). *Neuroscience*, 46(2):275–286, 1992. doi:10.1016/0306-4522(92)90050-c.

[KGS05a] P. J. Kellman, P. Garrigan, and T. F. Shipley. Object interpolation in three dimensions. *Psychological Review*, 112(3):586–609, 2005. doi:10.1037/0033-295x.112.3.586.

[KGS+05b] P. J. Kellman, P. Garrigan, T. F. Shipley, C. Yin, and L. Machado. 3-d interpolation in object perception: Evidence from an objective performance paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, 31(3):558–583, 2005. doi:10.1037/0096-1523.31.3.558.

[KHK16] S. K. Khuu, V. Honson, and J. Kim. The perception of three-dimensional contours and the effect of luminance polarity and color change on their detection. *Journal of Vision*, 16(3):31, Feb 2016. doi:10.1167/16.3.31.

[KK16] D. Kumari and K. Kaur. A survey on stereo matching techniques for 3D vision in image processing. *International Journal of Engineering and Manufacturing*, 4:40–49, 2016. doi:10.5815/ijem.2016.04.05.

[Kof35] K. Koffka. Principles of gestalt psychology (nueva york, harcourt, brace &co). 1935.

[Kof63] K. Koffka. Principles of Gestalt psychology. New York, NY, USA: A Harbinger Book, 1963.

[Koh29] W. Kohler. Gestalt psychology. new york: H. liveright, 1929.

[Koh67] W. Kohler. Gestalt psychology. *Psychologische Forschung*, 31(1):XVIII–XXX, 1967. doi:10.1007/bf00422382.

[KS91] P. J. J Kellman and T. F. Shipley. A theory of visual interpolation in object perception. *Cognitive Psychology*, 23(2):141–221, 1991. doi:10.1016/0010-0285(91)90009-d.

[KvD87] J. J. Koenderink and A. J. van Doorn. Representation of local geometry in the visual system. *Biological Cybernetics*, 55(6):367–375, Mar 1987. doi:10.1007/bf00318371.

[LH84] Margaret S Livingstone and David H Hubel. Anatomy and physiology of a color system in the primate visual cortex. *Journal of Neuroscience*, 4(1):309–356, 1984. doi:10.1523/JNEUROSCI.04-01-00309.1984.

[LHW75] S. LeVay, D. H. Hubel, and T. N. Wiesel. The pattern of ocular dominance columns in macaque visual cortex revealed by a reduced silver stain. *Journal of Comparative Neurology*, 159(4):559–575, 1975. doi:10.1002/cne.901590408.

[LZ03] G. Li and S. W. Zucker. A differential geometrical model for contour-based stereo correspondence. In *Proc. of IEEE Workshop on Variational, Geometric and Level set Methods in Computer Vision, Nice, France*, 2003.

[LZ06] G. Li and S. W. Zucker. Contextual inference in contour-based stereo correspondence. *International Journal of Computer Vision*, 69(1):59–75, 2006. doi:10.1007/s11263-006-6853-9.

[MAHG93] R. Malach, Y. Amir, M. Harel, and A. Grinvald. Relationship between intrinsic connections and functional architecture revealed by optical imaging and in vivo targeted biocytin injections in primate striate cortex. *Proceedings of the National Academy of Sciences*, 90(22):10469–10473, 1993. doi:10.1073/pnas.90.22.10469.

[MC82] G. Mitchison and F. Crick. Long axons within the striate cortex: their distribution, orientation, and patterns of connection. *Proceedings of the National Academy of Sciences*, 79(11):3661–3665, 1982. doi:10.1073/pnas.79.11.3661.

[MD16] A. P. Mashtakov and R. Duits. A cortical based model for contour completion on the retinal sphere. *Program Systems: Theory and Applications*, 7(4):231–247, 2016.

[MDS+17] A. Mashtakov, R. Duits, Yu. Sachkov, E. J. Bekkers, and I. Beschastnyi. Tracking of lines in spherical images via sub-Riemannian geodesics in SO(3). *Journal of Mathematical Imaging and Vision*, 58(2):239–264, 2017. doi:10.1007/s10851-017-0705-9.

[MF04] M. D. Menz and R. D. Freeman. Functional connectivity of disparity-tuned neurons in the visual cortex. *Journal of Neurophysiology*, 91(4):1794–1807, Apr 2004. doi:10.1152/jn.00574.2003.

[MOW+17] S. Meesters, P. Ossenblok, L. Wagner, O. Schijns, P. Boon, L. Florack, and R. Vilanova, A.and Duits. Stability metrics for optic radiation tractography: Towards damage prediction after resective surgery. *Journal of neuroscience methods*, 288:34–44, 2017. doi:10.1016/j.jneumeth.2017.05.029.

[MP79] D. Marr and T. Poggio. A computational theory of human stereo vision. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 204(1156):301–328, May 1979. doi:10.1098/rspb.1979.0029.

[MP20] A. P. Mashtakov and A. Yu. Popov. Asymptotics of extremal controls in the sub-riemannian problem on the group of motions of Euclidean space. *Russian Journal of Nonlinear Dynamics*, 16(1):195–208, 2020. doi:10.20537/nd200115.

[MS01] M. Meilă and J. Shi. A random walks view of spectral segmentation. In *International Workshop on Artificial Intelligence and Statistics*, pages 203–208. PMLR, 2001.

[MS09] P. Momayyez and K. Siddiqi. 3D stochastic completion fields for fiber tractography. In *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 178–185. IEEE, 2009. doi:10.1109/CVPRW.2009.5204044.

[MS12] P. MomayyezSiahkal and K. Siddiqi. 3D stochastic completion fields for mapping connectivity in diffusion MRI. *IEEE transactions on pattern analysis and machine intelligence*, 35(4):983–995, 2012. doi:10.1109/TPAMI.2012.184.

[Mum94] D. Mumford. Elastica and computer vision. In *Algebraic geometry and its applications*, pages 491–506. Springer, 1994. doi:10.1007/978-1-4612-2628-4_31.

[Nic14] A. Nicholls. Confidence limits, error bars and method comparison in molecular modeling. Part 1: the calculation of confidence intervals. *Journal of Computer-aided Molecular Design*, 28(9):887–918, 2014. doi:10.1007/s10822-016-9904-5.

[NKB77] J. I. Nelson, H. Kato, and P. O. Bishop. Discrimination of orientation and position disparities by binocularly activated neurons in cat straite cortex. *Journal of Neurophysiology*, 40(2):260–283, Mar 1977. doi:10.1152/jn.1977.40.2.260.

[Øks03] B. Øksendal. *Stochastic differential equations*. Springer, 2003.

[Orb11] G. A. Orban. The extraction of 3D shape in the visual system of human and nonhuman primates. *Annual Review of Neuroscience*, 34(1):361–388, Jul 2011. doi:10.1146/annurev-neuro-061010-113819.

[PD17] J. M. Portegies and R. Duits. New exact and numerical solutions of the (convection-)diffusion kernels on SE(3). *Differential Geom. Appl.*, 53:182–219, 2017. doi:10.1016/j.difgeo.2017.06.004.

[Pet08] J. Petitot. *Neurogéométrie de la vision: modeles mathematiques et physiques des architectures fonctionnelles*. Editions Ecole Polytechnique, 2008.

[PF77] GF Poggio and B Fischer. Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *Journal of Neurophysiology*, 40(6):1392–1405, 1977. doi:10.1152/jn.1977.40.6.1392.

[PF98] P. Perona and W. Freeman. A factorization approach to grouping. In *European Conference on Computer Vision*, pages 655–670. Springer, 1998. doi:10.1007/BFb0055696.

[PFS+15] J. M. Portegies, R. H. J. Fick, G. R. Sanguinetti, S. P. L. Meesters, G. Girard, and R. Duits. Improving fiber alignment in HARDI by combining contextual pde flow with constrained spherical deconvolution. *PloS one*, 10(10), 2015. doi:10.1371/journal.pone.0138122.

[PGK88] G. F. Poggio, F. Gonzalez, and F. Krause. Stereoscopic mechanisms in monkey visual cortex: binocular correlation and disparity selectivity. *Journal of Neuroscience*, 8(12):4531–4550, 1988. doi:10.1523/JNEUROSCI.08-12-04531.1988.

[Piz01] Z. Pizlo. Perception viewed as an inverse problem. *Vision Research*, 41(24):3145–3161, 2001. doi:10.1016/s0042-6989(01)00173-0.

[PNB68] J.D. Pettigrew, T. Nikara, and P.O. Bishop. Binocular interaction on single units in cat striate cortex: simultaneous stimulation by single moving slit with receptive fields in correspondence. *Experimental Brain Research*, 6(4):391–410, 1968. doi:10.1007/BF00233186.

[Pog95] G. F. Poggio. Mechanisms of stereopsis in monkey visual cortex. *Cerebral Cortex*, 5(3):193–204, 1995. doi:0.1093/cercor/5.3.193.

[PSK16] A. J. Parker, J. E. T. Smith, and K. Krug. Neural architectures for stereo vision. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1697):20150261, Jun 2016. doi:10.1098/rstb.2015.0261.

[PSMD15] J. Portegies, G. Sanguinetti, S. Meesters, and R. Duits. New approximation of a scale space kernel on SE(3) and applications in neuroimaging. In *Scale Space and Variational Methods in Computer Vision: 5th International Conference, SSVM 2015, Lège-Cap Ferret, France, May 31-June 4, 2015, Proceedings 5*, pages 40–52. Springer, 2015. doi:10.1007/978-3-319-18461-6_4.

[PT99] J. Petitot and Y. Tondut. Vers une neurogéométrie. Fibrations corticales, structures de contact et contours subjectifs modaux. *Mathématiques et Sciences humaines*, 145:5–101, 1999. URL: http://www.numdam.org/item/MSH_1999__145__5_0/.

[PWK18] L. Paulun, A. Wendt, and N. Kasabov. A retinotopic spiking neural network system for accurate recognition of moving objects using neucube and dynamic vision sensors. *Frontiers in Computational Neuroscience*, 12:42, 2018. doi:10.3389/fncom.2018.00042.

[PZ89] P. Parent and S. W. Zucker. Trace inference, curvature consistency, and curve detection. *IEEE Transactions on pattern analysis and machine intelligence*, 11(8):823–839, 1989. doi:10.1109/34.31445.

[Qia97] N. Qian. Binocular disparity and the perception of depth. *Neuron*, 18(3):359–368, 1997. doi:10.1016/s0896-6273(00)81238-6.

[RC07] J. C. Read and B. G. Cumming. Sensors for impossible stimuli may solve the stereo correspondence problem. *Nature Neuroscience*, Oct 2007. doi:10.1038/nn1951.

[RDRV10] P. Rodrigues, R. Duits, B. M. ter Haar Romeny, and A. Vilanova. Accelerated Diffusion Operators for Enhancing DW-MRI. In Dirk Bartz, Charl Botha, Joachim Hornegger, Raghu Machiraju, Alexander Wiebel, and Bernhard Preim, editors, *Eurographics Workshop on Visual Computing for Biology and Medicine*. The Eurographics Association, 2010. doi:10.2312/VCBM/VCBM10/049-056.

[RS76] Linda Preiss Rothschild and Elias M Stein. Hypoelliptic differential operators and nilpotent groups. *Acta Mathematica*, 137:247–320, 1976.

[RS00] S. T. Roweis and L. K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *science*, 290(5500):2323–2326, 2000. doi:10.1126/science.290.5500.2323.

[SB98] S. Sarkar and K. L. Boyer. Quantitative measures of change based on feature organization: Eigenvalues and eigenvectors. *Computer vision and image understanding*, 71(1):110–136, 1998.

[SC15] A. Sarti and G. Citti. The constitution of visual perceptual units in the functional architecture of V1. *Journal of computational neuroscience*, 38(2):285–300, 2015. doi:10.1007/s10827-014-0540-6.

[SCP07] A. Sarti, G. Citti, and J. Petitot. The symplectic structure of the primary visual cortex. *Biological Cybernetics*, 98(1):33–48, Nov 2007. doi:10.1007/s00422-007-0194-9.

[SCS10a] G. Sanguinetti, G. Citti, and A. Sarti. A model of natural image edge co-occurrence in the rototranslation group. *Journal of vision*, 2010. doi:10.1167/10.14.37.

[SCS10b] G. Sanguinetti, G. Citti, and A. Sarti. A model of natural image edge co-occurrence in the rototranslation group. *Journal of Vision*, 10(14), 2010. doi:doi.org/10.1167/10.14.37.

[SGLS97] K. E. Schmidt, R. Goebel, S. Löwel, and W. Singer. The perceptual grouping criterion of colinearity is reflected by anisotropies of connections in the primary visual cortex. *European Journal of Neuroscience*, 9(5):1083–1089, 1997. doi:10.1111/j.1460-9568.1997.tb01459.x.

[SK94] T. F. Shipley and P. J. Kellman. Spatiotemporal boundary formation: Boundary, form, and motion perception from transformations of surface elements. *Journal of Experimental Psychology: General*, 123(1):3, 1994. doi:10.1037/0096-3445.123.1.3.

[SK17] B. L. Schwartz and J. H. Krantz. *Sensation and perception*. Sage Publications, 2017.

[SM00] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22(8):888–905, 2000. doi:10.1109/34.868688.

[SPTL13] J. M. Samonds, B. R. Potetz, C. W. Tyler, and T. S. Lee. Recurrent connectivity can account for the dynamics of disparity processing in V1. *Journal of Neuroscience*, 33(7):2934–2946, 2013. doi:10.1523/JNEUROSCI.2952-12.2013.

[SR11] P. Sturm and S. Ramalingam. *Camera models and fundamental concepts used in geometric computer vision*. Now Publishers Inc, 2011.

[STO10] K. S. Sasaki, Y. Tabuchi, and I. Ohzawa. Complex cells in the cat striate cortex have multiple disparity detectors in the three-dimensional binocular receptive fields. *Journal of Neuroscience*, 30(41):13826–13837, 2010. doi:10.1523/JNEUROSCI.1135-10.2010.

[STR+22] B. Scholl, C. Tepohl, M. A. Ryan, C. I. Thomas, N. Kamasawa, and D. Fitzpatrick. A binocular synaptic network supports interocular response alignment in visual cortical neurons. *Neuron*, 110(9):1573–1584, 2022. doi:10.1113/jphysiol.1972.sp010006.

[SV87] J. T. S. Smits and P. G. Vos. The perception of continuous curves in dot stimuli. *Perception*, 16(1):121–131, 1987. doi:10.1068/p160121.

[TCL03] D. Y. Tsao, B. R. Conway, and M. S. Livingstone. Receptive fields of disparity-tuned simple cells in macaque V1. *Neuron*, 38(1):103–114, 2003. doi:10.1016/s0896-6273(03)00150-8.

[TSE94] G. Tononi, O. Sporns, and G.M. Edelman. A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Sciences*, 91(11):5033–5037, 1994. doi:10.1073/pnas.91.11.503.

[Tu11] L. W. Tu. Manifolds. In *An Introduction to Manifolds*, pages 47–83. Springer, 2011.

[Utt83] W. R. Uttal. *Visual form detection in 3-dimensional space*. Hillsdale, N.J : L. Erlbaum Associates, 1983. doi:10.4324/9780203781166.

[VdHPB84] R. Von der Heydt, E. Peterhans, and G. Baumgartner. Illusory contours and cortical neuron responses. *Science*, 224(4654):1260–1262, 1984. doi:10.1126/science.6539501.

[VDHZF00] R. Von Der Heydt, H. Zhou, and H. S. Friedman. Representation of stereoscopic edges in monkey visual cortex. *Vision Research*, 40(15):1955–1967, 2000. doi:10.1016/s0042-6989(00)00044-4.

[VF92] R. Vaillant and O. Faugeras. Using extremal boundaries for 3-d object modeling. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 14(02):157–173, 1992. doi:10.1109/34.121787.

[Wan95] B. A. Wandell. *Foundations of vision*. Sinauer Associates, 1995.

[WC72] H. R. Wilson and J. D. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, 12(1):1–24, 1972. doi:10.1016/S0006-3495(72)86068-5.

[WC73] H. R. Wilson and J. D. Cowan. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, 13(2):55–80, 1973. doi:10.1007/BF00288786.

[WEK+12] J. Wagemans, J. H. Elder, M. Kubovy, S. E. Palmer, M. A. Peterson, M. Singh, and R. von der Heydt. A century of gestalt psychology in visual perception: I. perceptual grouping and figure–ground organization. *Psychological Bulletin*, 138(6):1172, 2012. doi:10.1037/a0029333.

[Wer23] M. Wertheimer. Laws of organization in perceptual forms. *Psycologische Forschung*, 4, 1923. doi:10.1037/11496-005.

[YF98] S. C. Yen and L. H. Finkel. Extraction of perceptually salient contours by striate cortical networks. *Vision Research*, 38(5):719–741, 1998. doi:0.1016/s0042-6989(97)00197-1.

[Zuc06] S. Zucker. Differential geometry from the Frenet point of view: boundary detection, stereo, texture and color. In *Handbook of mathematical models in computer vision*, pages 357–373. Springer, 2006.

[Zuc14] S. W. Zucker. Stereo, shading, and surfaces: Curvature constraints couple neural computations. *Proceedings of the IEEE*, 102(5):812–829, 2014. doi:10.1109/JPROC.2014.2314723.