Alma Mater Studiorum · Università di Bologna

SCUOLA DI SCIENZE

Corso di Laurea Magistrale in Matematica

SUB-RIEMANNIAN MODELS OF DYNAMICAL OPTICAL ILLUSIONS

Tesi di Laurea in Biomedical Mathematical Models

Relatore:
Chiar.ma Prof.
GIOVANNA CITTI

Correlatore:
Chiar.mo Prof.
ALESSANDRO SARTI

Presentata da:
MARIA CLARA GIACCAGLIA

Sessione IV
Anno Accademico 2024-2025

Introduction

The objective of this thesis is to develop a mathematical model of optical perceptual illusions that incorporates time-dependent dynamics. Extending previous models of static geometric illusion developed by Franceschiello, Sarti, and Citti in [5, 6], we start from sub-Riemannian models of the brain, describe the action of the input as a deformation of the cortical space, and apply the model to time-dependent Ebbinghaus illusions. In this thesis, we propose a brain-inspired mathematical model of time-dependent optical illusions. Differential models of the visual cortex have been proposed by Petitot and Tondut [17], and by Citti and Sarti [20]. Each family of cells is typically modeled as a Lie group endowed with a sub-Riemannian metric. The visual input is lifted by the receptive profiles in V1, processed by simple cells, and sent to higher cortical areas for further processing. The result of this process is the perceived image, which, in general, does not coincide with the real one.

Geometric optical illusions are special images, in which the difference between perceived and real images is particularly clear, and are typically used to understand mechanisms of visual perception. The Ebbinghaus and Delboeuf effects (Figure 1) are examples of the so-called scale illusions, since the phenomenon is processed by cells responsible for scale perception.
Here we extend a model proposed in [6]. The visual input is lifted into a Lie group structure whose base space is $\mathbb{R}^2 \times S^1 \times \mathbb{R}^+$, where $\mathbb{R}^2$ represents spatial position, S^1 accounts for orientation, and $\mathbb{R}^+$ for scale. The metric of this space is sub-Riemannian: it models the cortical connectivity in V1 and modulates perception by introducing local distortions in feature space.

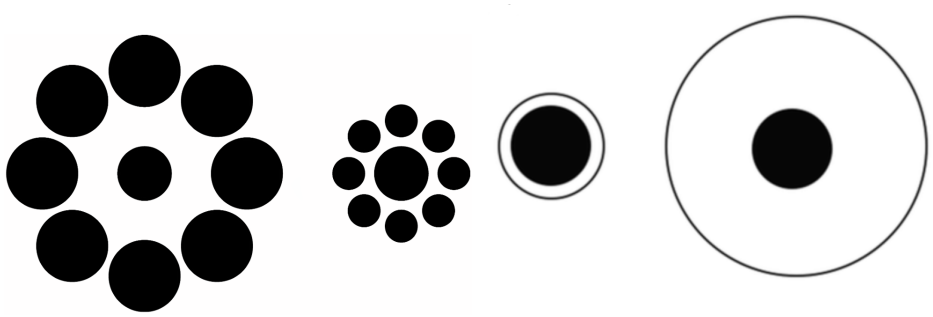

Figure 1: The Ebbinghaus illusion on the left, the Delbeuf illusion on the right.

Consequently, it induces a deformation of the input image, providing the perceived one.
In this work, we extend this research to geometrical optical illusions in motion, with the aim of showing that the illusory effect becomes significantly stronger when temporal dynamics are introduced. Specifically, we propose a model for the Dynamic Ebbinghaus illusion, studied by Mruczek, Blair, Strother and Caplovitz in [14], where the size and eccentricity of the inducers, surrounding a fixed central target, vary over time. Under such conditions, the size of the central circle is perceived to change in opposition to the size of the inducers, and the illusory effect is greatly enhanced compared to the static case studied in [5, 6]. For this reason, we aim to develop a time-dependent variant of the geometric model.

The thesis is organized as follows:

- **Chapter 1** briefly recalls the perceptual phenomena and the structure of the primary visual cortex (V1), the cortical area responsible for early-stage visual processing. We describe its functional organization in terms of hypercolumns and receptive fields, distinguishing between simple and complex cells, and explain how orientation, position, and scale selectivity can be modeled geometrically. This provides the biological foundation for representing cortical connectivity as a Lie group structure.

- **Chapter 2** recalls the necessary mathematical background. We introduce notions from differential and Riemannian geometry that lead to the definition of sub-Riemannian structures, which serve as the framework for modeling cortical geometry. We also review Hilbert spaces, frame theory, and wavelet analysis, preparing the ground for the introduction of Gabor filters as both mathematical and physiological models of receptive profiles.

- **Chapter 3** presents the model for static geometrical optical illusions. Receptive profiles of simple cells are modeled as two-dimensional Gabor filters, forming a frame in $L^2(\mathbb{R}^2)$. The cortical connectivity is described in the Lie group $\mathbb{R}^2 \times S^1 \times \mathbb{R}^+$, equipped with a sub-Riemannian metric that accounts for translation, rotation, and dilation invariance. From this setting, we develop a model for scale type illusions, specifically the Ebbinghaus effect, interpreting the illusion as a geometric deformation of the perceptual space. The displacement field responsible for the deformation is recovered by solving a system of elliptic PDEs derived from the cortical metric.

- **Chapter 4** extends the static framework to include the temporal dimension. We introduce a new sub-Riemannian geometry designed to model motion-sensitive cells, adding time and velocity variables to the feature space. This leads to a five-dimensional manifold $\mathcal{M} = \mathbb{R}^2 \times \mathbb{R}_t \times S^1 \times \mathbb{R}^+_\sigma \times \mathbb{R}_v$, equipped with a distribution generated by vector fields that satisfy the Hörmander condition. Receptive profiles are described by three-dimensional spatiotemporal Gabor functions, and the induced cortical metric is analyzed. We then study the existence of flat coordinates and show how non-flat cases can be approximated by solving a PDE system for the displacement vector field.

- **Chapter 5** presents the numerical simulations and results obtained through MATLAB. We first validate the static model by reproducing the Ebbinghaus illusion through multi-scale filtering and extraction of

the scale map, followed by the computation of the displacement vector field describing the perceptual deformation. We then extend the simulations to the dynamic case, constructing a stimulus that reproduces the stationary Dynamic Ebbinghaus illusion. The spatiotemporal Gabor filter bank captures the evolution of the perceived deformation.

Chapters 1 to 3 provide the theoretical and mathematical background necessary to understand the proposed model, while Chapters 4 and 5 constitute the original contribution of this work. In particular, they present the extension of the static geometric framework to the spatio-temporal domain and the corresponding numerical implementation and simulations.
In conclusion, this work contributes to the mathematical understanding of dynamic visual perception by formalizing the interaction between cortical connectivity, spatial structure, and temporal dynamics.

Contents

List of Figures

Chapter 1

Geometric optical illusions and their neural basis

1.1 Introduction

In order to develop a mathematical description of the visual phenomena that will be addressed in this work, it is first necessary to clarify the role of visual illusions to understand visual perception principles, and to examine the functional architecture of the cortex where this process of vision takes place: the primary visual cortex. We will concentrate only on the aspects that play a central role in the models introduced in the following chapters.

1.2 Visual perception and visual optical illusions

Visual perception is the result of the processing of the visual stimulus, which takes place in the visual cortex. Starting from the visual stimulus, it gives rise to a perceived image, which, in general, does not coincide with the real one. Visual perception has been studied by the Gestalt movement in psychology, and many concepts were introduced by its exponents. Gestalt psychology established basic "grouping laws", which are crucial in construct-

ing a phenomenological representation of the physical world: points having one or several characteristics in common, are grouped together to form a new and larger visual object. This approach provided computational models of Gestalt (for example [22]). One of the most studied perceptual phenomena is boundary completion, extensively analyzed by Kanizsa [23].
Geometric optical illusions are special images, in which the difference between perceived and real images is particularly clear, and are typically used to understand mechanisms of visual perception. They (GOIs) are a class of phenomena first discovered by German physicists and physiologists in the late XIX century ([15],[8]), and can be defined as situations where there is a perceptual mismatch between the visual stimulus and its geometrical properties. Those illusions are typically analyzed according to the main geometrical features of the stimulus, whether it is contours, orientation, contrast, context influence, size, or a combination of the above mentioned ones.

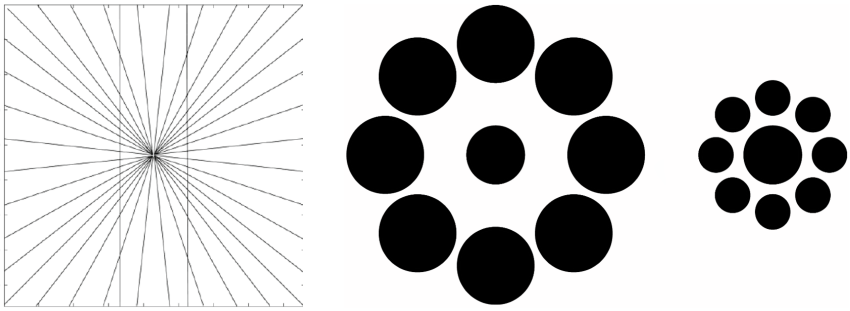

Figure 1.1: The Haring and the Ebbinghaus illusions

Notable examples include:

- the Hering illusion (Figure 3.1, left), which is due to a misperception of orientation. It was discovered by the German physiologist Ewald Hering in 1861. When two straight and parallel lines are presented in front of a radial background (like the spokes of a bicycle), the lines appear as if they were bowed outwards. The illusion was ascribed by Hering to an overestimation of the angle made at the points of intersection: this

angle overestimation results from process and estimate of angles which take in visual cortex.

- the Ebbinghaus illusion (Figure 3.1, right), which is due to a misperception of the scale of inducers. It was discovered by the German psychologist Hermann Ebbinghaus (1850–1909). Two circles of identical size are placed near to each other, and one is surrounded by large circles while the other is surrounded by small circles.

Several models have been proposed to explain geometric optical illusions. Hoffman [24] and Smith [25] described angle illusions through differential and Lie group formulations capturing perceptual invariants. Ehm and Wackermann [26] introduced a variational model minimizing curvature and orthogonality deviations, while Bayesian approaches, rooted in Helmholtz's theory of unconscious inference [27], interpret perception as a probabilistic estimation process [28, 29]. Other models, such as those by Fermüller and Malm [30] and Walker [31], link geometric distortions to visual noise or receptive field activity.

In this thesis, we focus on the time-dependent Ebbinghaus illusion presented in the physiological study by Mruczek, Blair, Strother, and Caplovitz [14], which shows that motion in the inducers of the classical illusion greatly amplifies the perceived size difference between identical central disks, revealing a strong interaction between spatial and temporal processing. To model this phenomenon, we extend the static geometric framework ([5, 6]) to the spatio-temporal domain by introducing motion-sensitive receptive profiles and a sub-Riemannian structure defined on spatial and temporal variables, following the approach of [2]. Neuronal connectivity is described by vector fields satisfying the Hörmander condition. We then simulate the dynamic Ebbignhaus illusion numerically through a MATLAB generated video and process it through spatio-temporal Gabor filters. The perceived deformation is recovered from the displacement vector field induced by the cortical metric. The results are compared with the static case, showing that motion enhances the illusory effect.

1.3 The visual pathway

The cerebral cortex is the outermost layer of neural tissue in the two cerebral hemispheres. It plays a central role in sensory and cognitive processing since most of the neurons responsible for these processes are located here. It is commonly divided in three parts: sensory, motor, and association. We are interested in the first of these, which is the part of the cortex that receives sensory inputs. In particular, the visual process is the result of several retinal and cortical mechanisms acting on the visual signal. The visual cortex is the area that serves the sense of vision and receives the optical information from the visual path.

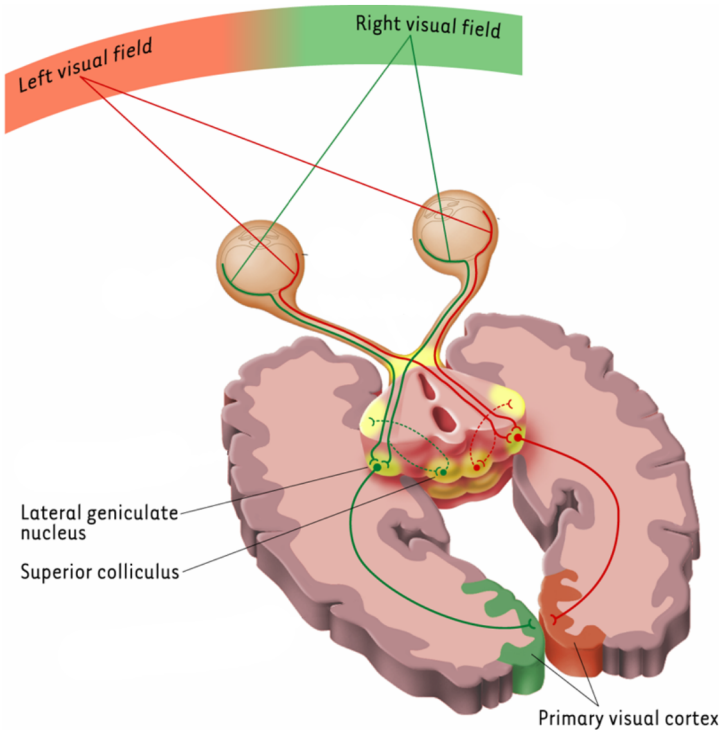

Figure 1.2: The visual path of the brain

First, light enters the eye and arrives to the retina, which is composed of ten thin layers of brain tissue where the neural processing of visual stimuli begins. Then the optic nerve brings the signal from the retina to the Lateral

Geniculate Nucleus (LGN) of the thalamus, a structure in the middle of the brain, which connects the sensory organs to their main sensory processing cortical areas. From the LGN the signal is sent to the visual cortex, situated in the back of the head, where the larger part of the visual processing is performed. The primary visual cortex (V1) is the area to which most of the retinal output first arrives and is the most widely studied visual area.

1.4 Receptive fields and receptive profiles

Focusing on how neurons react to the visual stimulus, we define Receptive Fields and Receptive Profiles. The Receptive Field of a visual neuron is the area of visual field in which visual stimulation influences neural responses. Every cell is linked to its Receptive Field, which can be seen as a specific domain D of the retina. Some of these cells are composed of ON areas, which react to the visual stimulus in a excitatory way, and OFF areas, which behave in the opposite way. Then we can define the Receptive Profile (RP) as a function $\psi : D \to \mathbb{R}$ which measures the reaction $\psi(x_1, x_2)$ of the neuron to the stimulus, where (x_1, x_2) are retinal coordinates.

1.5 The functional architecture of V1

We refer to the functional architecture as the spatial organization and the connectivity between neurons inside a cortical area. In V1 we can identify three structures:

1. the layered structure, which indicates that the cortex is formed of 6 horizontal layers.

2. the retinotopic structure has a particular kind of topographic organization implying that there exists a topology preserving mapping between the retina and the cortex. For this reason, cells in each structure can be seen as forming a map of the visual field: simple cell receptive fields form a mosaic that covers the retina.

3. the hypercolumnar structure organizes the cortical cells in columns corresponding to parameters such as orientation, ocular dominance, color, etc. For the simple cells, sensitive to orientation, columnar structure means that to every retinal position is associated a set of cells (hypercolumn) sensitive to all the possible orientations.

1.5.1 The layered structure

The layered structure indicates that the cortex is formed by 6 horizontal layers and a number of sublayers. The sublayer 4C is where most of the axons from the LGN arrive and where the concentration of oriented cells is higher.

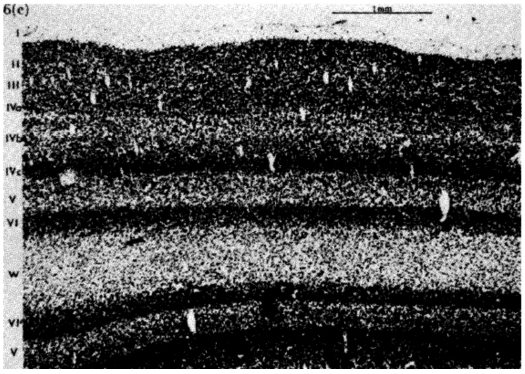

Figure 1.3: The layered structure of the primary visual cortex as photographed by Nissl (image taken from [9]). The picture reveals quite clearly the different layers I to VI.

1.5.2 The hypercolumnar structure

The hypercolumnar structure was discovered by the neurophysiologists Hubel and Wiesel in the 1960s. At a certain spatial scale and resolution, for each point of the retina (x_1, x_2) there exists a whole set of neurons in V1 that respond maximally to every possible local orientation θ. Since the position on the retina takes values in the plane $\mathbb{R}^2$ and the orientation preference lies

on the circle S^1, the visual cortex can be locally modeled as the product space $\mathbb{R}^2 \times S^1$. Each point (x_1, x_2, θ) of this three-dimensional space represents a column of cells in the cortex associated with a retinal position (x_1, x_2), all tuned to the orientation given by the angle θ.
This organization extends beyond orientation to other visual features such as scale, direction of motion, and color. Let the set of features be denoted by F. Over each retinal point we consider a whole hypercolumn of cells, each one sensitive to a specific instance of the considered feature $f \in F$. We can thus identify cortical cells by the triplet (x_1, x_2, f), where (x_1, x_2) represents the retinal position and f is the vector of extracted features. Consequently, the cortical space is modeled as $\mathbb{R}^2 \times F$.
When a visual stimulus $I : M \subset \mathbb{R}^2 \to \mathbb{R}^+$ of intensity $I(x_1, x_2)$ activates the retinal layer of photoreceptors, the neurons whose receptive fields intersect M spike, and their spike frequencies $O(x_1, x_2, f)$ can be modeled as a convolution between the stimulus and the corresponding receptive profiles:

$$O(x_1, x_2, f) = (I * \psi_f)(x_1, x_2).$$

1.5.3 The retinotopic mapping

The retinotopy is a particular kind of topographic organization implying that there exists a topology preserving mapping from the retina to the cortex. The position of the center of the simple cell receptive fields form an ordered mosaic that covers a portion of the visual field. Because of this arrangement, which emerges from the spatial specificity of connections between neurons in the retina and in the cortex, cells in each structure can be seen as forming a map of the visual field (also called a retinotopic map, or a visuotopic map). In other words, what is near in the retina is near in the cortex.

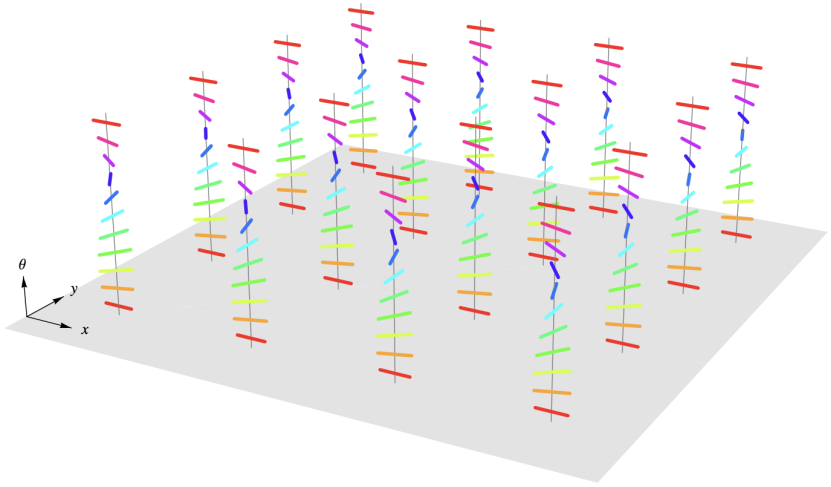

Figure 1.4: The visual cortex modeled as a set of hypercolumns. Over each retinal point there exists a whole set of neurons in V1 responding maximally to every possible local orientation θ. Each coloured bar represents a possible orientation $\theta \in S^1$

.

1.6 Cortical connectivity

We need to discuss the connectivity between neurons inside the structure we have seen. In the hypercolumnar structure we can identify two types of communication between neurons which play a central role in the model we want to present:

- The intracortical circuitry is able to select the hypercolumns orientation of maximum output in response to a visual stimulus and to suppress all the others. This mechanism is called non-maxima suppression. The intracortical neural circuitry acts within a single hypercolumn. In presence of a visual stimulus, at a point $x = (x_1, x_2)$, the whole hypercolumn over that point fires, but mechanism of non-maximal suppression acts, suppressing the output of cells that, within the same hypercolumn, are not maximal. This mechanism is also called orientation selection.

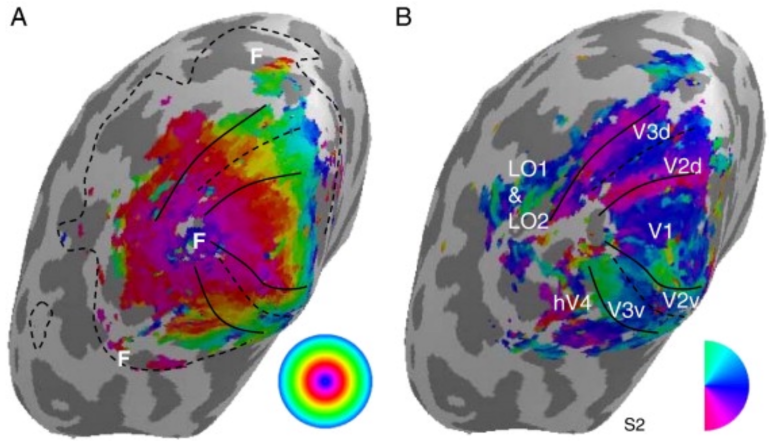

Figure 1.5: The retinotopic structure of the primary visual cortex recorded by FMRI (image taken from [18]). The mapping is visualized on a reconstructed anatomical volume. On the left the radial mapping is shown while on the right the polar angle is visualized.

- The horizontal or cortical connectivity takes place in the connectivity structure, the part of the visual cortex which ensures connectivity between hypercolumns. The horizontal connections connect cells with the same orientation belonging to different hypercolumns.

Chapter 2

Preliminaries

This section introduces the foundational concepts needed to understand the mathematical developments that follow. It is organized as follows: first we introduce Hilbert spaces and the properties that concern the convergence of the series. Then we recall some results of Riemannian and differential geometry needed for chapter 3.

2.1 Differential geometry

In this section, we first recall some fundamental notions from differential and Riemannian geometry, which are necessary to define the sub-Riemannian structure that will later be used to model perception.
We begin by introducing the concept of differential manifold.

Definition 2.1.1. *A topological space $\mathcal{M}$ is said to be locally euclidean of dimension n if $\forall$ p there exists U neighborhood of p that is homeomorphic to $\tilde{U}$ open set of $\mathbb{R}^n$ ($\phi : U \to \tilde{U}$ is a homeomorphism).*
$\phi(p) = \phi(x_1(p), ..., x_n(p))$ are local coordinates.

To be able to do calculus on $\mathcal{M}$, we need more structure than just a topological space.

Definition 2.1.2. *A topological space $\mathcal{M}$ is a topological manifold if $\mathcal{M}$ is second countable, Hausdorff and locally euclidean. Two charts (U, ϕ_u), (V, ϕ_v)*

are smoothly compatible if: $U \cap V = \emptyset$ *or* $\phi_v \circ \phi_u^{-1}$ *is a diffeomorphism.*
A smooth atlas for a topological manifold $\mathcal{M}$ *is a collection* $\{(U_i, \phi_i)\}_i$ *of smoothly compatible charts covering* $\mathcal{M}$*. The maximal atlas defines a smooth structure.*

Definition 2.1.3. *A differential manifold is a Hausdorff and second countable topological space* $\mathcal{M}$ *together with a smooth structure (maximal altlas) on* $\mathcal{M}$*.*

2.1.1 Riemannian geometry

Once we have a differential structure, we can introduce geometric concepts such as the distance by defining a Riemannian metric.

Definition 2.1.4. *Let* $\mathcal{M}$ *be a differential manifold. A Riemannian metric on* $\mathcal{M}$ *is defined by a scalar product on* $T_x\mathcal{M}$, $\forall x \in \mathcal{M}$ *which depends smoothly on* x*.*

In local coordinates, a scalar product on $T_x\mathcal{M}$ is expressed by a matrix $(g_{ij}(x))$ defined on $T_x\mathcal{M}$. For any $v, w \in T_x\mathcal{M}$ we have

$$v = (v_1, ..., v_n),\ w = (w_1, ..., w_n), \quad \langle v, w \rangle_{g(x)} = \sum_{i,j}^{n} g_{ij} v_i w_j(x)$$

The Riemannian metric allows us to define the length of curves.

Definition 2.1.5. *If* $\gamma : [0,1] \to \mathcal{M}$ *is a* C^1 *curve, if* M *is a Riemannian manifold,* $\forall t \in [0,1]$, $\gamma(t) \in M$ *and* $\gamma'(t) \in T_{\gamma(t)}M$*, then*

$$\exists\ ||\gamma'(t)||_{g(\gamma(t))} = \sqrt{\langle \dot{\gamma(t)}, \gamma'(t) \rangle}_{g(\gamma(t))}.$$

We then define the length of γ *in the metric* g *as*

$$l(\gamma) = \int_0^1 ||\gamma'(t)||_{g(\gamma(t))} dt.$$

We note that the integral exists since γ' and g are continuous, which implies that $||\gamma'(t)||_{g(\gamma(t))}$ is continuous.
Many paths are C^1 curves, but we can also face piecewise paths: definition 2.1.5 extends naturally to piecewise smooth curves.

Definition 2.1.6. *$\gamma : [0,1] \to \mathcal{M}$ is piecewise C^1 if $\exists\ t_0, ..., t_n \in [0,1]$ such that $t_0 = 0 < t_1 < ... < t_n = 1$ and $\gamma\Big|_{[t_{i-1},t_i]}$ is of class C^1.*

Definition 2.1.7. *If $\gamma : [0,1] \to \mathcal{M}$ is a piecewise C^1 curve, then we call*

$$l(\gamma) = \sum_{i=1}^{n} l\big(\gamma\big|_{[t_{i-1},t_i]}\big)$$

We can finally define the Riemannian distance between two points.

Definition 2.1.8. *Let $\mathcal{M}$ be a path-connected Riemannian manifold. For any $p, q \in \mathcal{M}$ we define*

$$d(p,q) = \inf\{l(\gamma),\ \gamma \text{ is piecewise } C^1 \text{ and } \gamma(0) = p, \gamma(1) = q\} \tag{2.1.1}$$

The fact that $\mathcal{M}$ is path-connected guarantees that $\forall\ (p,q)\ \exists\ \gamma$ with extreme p and q, thus the above set is non-empty and $l(\gamma) \geq 0$, so 2.1.1 is real.

2.1.2 Sub-Riemannian geometry

We now generalize the Riemannian framework by introducing constraints on the allowed directions of motion. This leads us to the sub-Riemannian setting, where only certain directions, defined by a distribution, are admissible.

Definition 2.1.9. *Let $\mathcal{M}$ be a differential manifold. We call distribution Δ a sub-bundle of the tangent bundle, i.e. $\forall\ p \in \mathcal{M}$, Δ_p is a subspace of $T_p\mathcal{M}$.*

Remark 2.1.10. *A choice of m vector fields $X_1, ..., X_m$ defines a distribution $\Delta_p = span(X_1\big|_p, ..., X_n\big|_p)$, $\forall\ p \in \mathcal{M}$.*

Definition 2.1.11. *We say that a sub-Riemannian manifold is a triple* $(\mathcal{M}, \Delta, g)$, *where* $\mathcal{M}$ *is a differential manifold,* Δ *is a distribution, and* g *is a metric on* Δ.

Definition 2.1.12. $\forall\, p \in \mathcal{M}$, Δ_p *is called horizontal or admissible distribution.*

Definition 2.1.13. *Let* $X_1, ..., X_m$ *be* C^1 *operators such that* $X_1 I, ..., X_m I$ *are the associated vector fields. We call integral curve of the vector fields* $X_1 I, ..., X_m I$ *a curve* $\gamma : [0,1] \to \mathcal{M}$ *such that* $\exists\ a_1, ..., a_m$ *continuous and* $\gamma' = a_1 X_1 I + ... + a_m X_m I$.

Proceeding analogously as before, γ is a piecewise C^1 integral curve of $X_1 I, ..., X_m I$ if γ is continuous and $t_0 = 0 < t_1 < ... < t_n = 1$ and $\gamma\big|_{[t_{i-1}, t_i]}$ is an integral curve of $X_1 I, ..., X_m I$.

Definition 2.1.14. *Let* $(\mathcal{M}, \Delta, g)$ *be a sub-Riemannian manifold. A curve* $\gamma : [0,1] \to \mathcal{M}$ *is horizontal (or admissible) if it is a piecewise integral curve of the generators of the distribution.*

Definition 2.1.15. *If* $\gamma : [0,1] \to \mathcal{M}$ *is horizontal, then we can define the length of a curve* $l(\gamma) = \int_0^1 ||\gamma'(t)||_{g(t)} dt$

Definition 2.1.16. *Let* $(\mathcal{M}, \Delta, g)$ *be a sub-Riemannian manifold path-connected, then, if* $p, q \in \mathcal{M}$, *we define the distance*

$$d(p,q) = \inf\{l(\gamma) : \gamma \text{ admissible},\ \gamma(0) = p,\ \gamma(1) = q\}. \tag{2.1.2}$$

Unlike the Riemannian setting, the set defined above can be empty (γ not admissible). In this case we define the distance 2.1.2 as $d(p,q) = +\infty$. Indeed, in a general sub-Riemannian manifold path-connected, there exist p, q which cannot be connected by an admissible curve. To address this problem, we express a condition, called the Hörmander condition, which ensures that any couple of points can be connected by a horizontal curve. To this aim we recall the following results.

Definition 2.1.17. *Let $X, Y \in C^\infty$ be vector fields. Then we call commutator (or Lie braket) of X, Y*

$$[X, Y] = XY - YX.$$

Definition 2.1.18. *Let $X_1, ..., X_n$ be vector fields with C^∞ coefficients. We call Lie algebra generated by $X_1, ..., X_n$ the closure of $X_1, ..., X_n$ with respect to linear combination and braket. We denote it as $Lie(X_1, ..., X_n)$.*

We note that, by definition, since the Lie algebra contains all linear combinations and all commutators, $span(X_1, ..., X_n) \subseteq Lie(X_1, ..., X_n)$.

Definition 2.1.19 (Hörmander condition)**.** *We say that a family of smooth vector fields $X_1, \cdots X_m$ defined on a manifold $\mathcal{M}$ satisfies the Hörmander condition if the generated Lie algebra has maximum rank at every point, equivalently if $Lie(X_1, ..., X_m)\big|_p = T_p\mathcal{M}\ \forall\ p$.*

This condition ensures controllability of the system: every point can be reached from every other via an admissible curve. We can state the following theorem.

Theorem 2.1.20 (Chow)**.** *If $(\mathcal{M}, \Delta, g)$ is a sub-Riemannian manifold connected by arch, and if the generators of Δ satisfy the Hörmander condition 2.1.19, then $\forall\ p, q \in \mathcal{M}$, (p, q) can be connected by an admissible curve.*

We now introduce some results about connections and curvature that will be used in Chapter 4

Definition 2.1.21. *An affine connection on a differentiable manifold $\mathcal{M}$ is an operator $\nabla : \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M}), \qquad (X, Y) \mapsto \nabla_X Y$, where $\mathfrak{X}(\mathcal{M})$ denotes the space of smooth vector fields on $\mathcal{M}$, such that for all $X, Y, Z \in \mathfrak{X}(\mathcal{M})$ and $f \in C^\infty(\mathcal{M})$:*

1. $\nabla_X(Y + Z) = \nabla_X Y + \nabla_X Z$,
2. $\nabla_X(fY) = X(f)Y + f\nabla_X Y$,

3. $\nabla_{fX}Y = f\nabla_X Y$.

Definition 2.1.22. *A connection ∇ on a Riemannian manifold $(\mathcal{M}, g)$ is said to be*

1. *symmetric (or torsion-free) if $\nabla_X Y - \nabla_Y X = [X, Y]$ for all X, Y.*

2. *metric-compatible if $X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle$ for all X, Y, Z.*

Theorem 2.1.23 (Levi Civita)**.** *Given a Riemannian manifold, there exists a unique affine connection that is symmetric and metric-compatible. This connection is called the Levi-Civita connection.*

Definition 2.1.24. *The Riemann curvature tensor associated with a connection ∇ is the $(1,3)$ tensor defined by $R(X,Y)Z = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z$.*

In local coordinates, the components of the curvature tensor are $R^i{}_{jk\ell} = \partial_k\Gamma^i_{j\ell} - \partial_\ell\Gamma^i_{jk} + \Gamma^i_{rk}\Gamma^r_{j\ell} - \Gamma^i_{r\ell}\Gamma^r_{jk}$, where Γ^i_{jk} are the Christoffel symbols of ∇.

Remark 2.1.25. *Since R is a tensor, its vanishing in one coordinate system implies its vanishing in any other.*

Definition 2.1.26. *Let X be a smooth vector field on a manifold M and $p \in M$. A maximal integral curve of X starting at p is a curve $\gamma_p : I_p \to M$ satisfying $\dot\gamma_p(t) = X(\gamma_p(t))$ for all $t \in I_p$, and such that γ_p is not contained in any larger integral curve of X. The interval I_p is called maximal interval.*

Definition 2.1.27. *Let X be a smooth vector field on a manifold M. If $\gamma_p : I_p \to M$ denotes the maximal integral curve of X starting at p, the flow of X is the map*

$$\phi : D \subset \mathbb{R} \times M \to M, \qquad \phi(t, p) = \gamma_p(t),$$

where $D = \{(t, p) \mid t \in I_p\}$ is an open subset of $\mathbb{R} \times M$.

Definition 2.1.28 (Lie derivative)**.** *Given a vector field X and a covariant tensor g, the Lie derivative is defined as*

$$\mathcal{L}_X g = \frac{d}{dt}\bigg|_{t=0} \phi_t^* g$$

.

Definition 2.1.29. *A vector field X on a Riemannian manifold $(\mathcal{M}, g)$ is called a killing vector field if $\mathcal{L}_X g = 0$, i.e. if the flow of X preserves the metric g.*

Remark 2.1.30. *If $\mathcal{L}_X g = 0$, then the flow ϕ_t of X preserves the metric, i.e. $\phi_t^* g = g$ for all t.*

Definition 2.1.31. *Let $(\mathcal{M}, g)$ be a Riemannian manifold with Levi–Civita connection ∇, and let $\gamma : [0,1] \to \mathcal{M}$ be a smooth curve. A vector field $V(t)$ along γ is said to be* parallel *if $\nabla_{\dot\gamma(t)} V(t) = 0$. The operation that assigns $V(1)$ to $V(0)$ is called* parallel transport *along γ.*

2.2 Hilbert spaces and frames

2.2.1 Hilbert spaces

We recall some general definitions and properties concerning Hilbert spaces.

Definition 2.2.1. *A Hilbert space is a vector space H endowed with an inner product*

$$H \times H \to \mathbb{C}$$

$$(x, y) \mapsto \langle x, y \rangle$$

satisfying the properties:

- $\langle x, y \rangle = \langle y, x \rangle$ *for any* $(x, y) \in H$*;*
- $\langle x, x \rangle \geq 0$ *for any* $x \in H$ *and if* $\langle x, x \rangle = 0$ *then* $x = 0$*;*

- *for any* $a, b \in \mathbb{C}$ *and for any* $x, y, z \in H$, $\langle ax + by, z\rangle = a\langle x, z\rangle + b\langle y, z\rangle$.

Definition 2.2.2. *Let* H *be a Hilbert space. A sequence* $\{x_n\}_n \subseteq H$ *is a basis for* H *if every element* $x \in H$ *can be written as*

$$x = \sum_{n=1}^{\infty} c_n(x) x_n$$

where $c_n(x)$ *is a unique choice of scalars.*

Definition 2.2.3. *A sequence* $\{x_n\}_n \subseteq H$ *is complete if* $\overline{Span\{x_n\}} = H$

Remark 2.2.4. *Note that every basis is a complete sequence, as for every* x *one can find a sequence* (c_n) *of scalars such that* $S_N = \sum_{n=1}^{N} c_n x_n \to x$ *for* $N \to \infty$. *Yet, not every complete sequence is a basis: to check it, it is sufficient to take a basis* x_n *and consider the sequence* $0 \cup x_n$.

Definition 2.2.5. *Consider a sequence* $\{x_n\} \subseteq H$, *then:*

- $\{x_n\}$ *is said to be orthogonal sequence if* $\langle x_m, x_n\rangle = 0$, *whenever* $m \neq n$,
- $\{x_n\}$ *is said to be orthonormal if* $\langle x_m, x_n\rangle = \delta_{m,n}$ *for every* m, n,
- $\{x_n\}$ *is an orthonormal basis if it is both orthonormal and a basis.*

Theorem 2.2.6. *Let* $\{x_n\}$ *be an orthonormal sequence in* H. *Then the following are equivalent:*

1. $\{x_n\}$ *is complete in* H,
2. $\{x_n\}$ *is a basis for* H,
3. $x = \sum_n \langle x, x_n\rangle x_n \quad \forall x \in H$,
4. $||x||^2 = \sum_n |\langle x, x_n\rangle|^2$ *(Plancherel's Equality),*
5. $\langle x, y\rangle = \sum_n \langle x, x_n\rangle \langle x_n, y\rangle$ *(Parseval's Equality).*

Let us now review a few notions about series convergence; we will exhibit some particular cases in which convergence comes with further properties and summarize the implications between them.

Definition 2.2.7. *Consider a sequence $\{x_n\}$ in H and an element $x \in H$.*

1. *The series $\sum_{n=1}^{\infty}\{x_n\}$ is convergent to x if the sequence $\{S_N\}_N = \{\sum_{n=1}^{N} x_n\}^N$ of the partial sums converges to x in the norm of H.*

2. *The series $\sum_{n=1}^{\infty}\{x_n\}$ is unconditionally convergent if it is convergent for every permutation σ of N.*

3. *The series $\sum_{n=1}^{\infty}\{x_n\}$ is absolutely convergent if $\sum_{n=1}^{\infty} ||x_n|| < +\infty$.*

4. *If the series $\sum_{n=1}^{\infty}\{x_n\}$ converges but not unconditionally, then it is said to be conditionally convergent.*

Proposition 2.2.8. *Let $\{x_n\}$ be a sequence in H. If $\sum_n x_n$ converges absolutely, then it converges unconditionally.*

Remark 2.2.9. *The converse to the above statement is not true in general. We have an equivalence of these two kinds of convergence only in special cases such as the one of sequences of scalars. Furthermore, we note that when defining unconditional convergence, we do not require that each permutation has the same limit x. Indeed, one shows that this property actually follows from the given definition.*

2.2.2 Frames in Hilbert spaces

Definition 2.2.10. *Let H be a (separable) Hilbert space and $\{x_n\}$ be a sequence in H. Then $\{x_n\}$ is a frame for H if there exist two constants $0 < A < B < \infty$ such that*

$$A||x||^2 \leq \sum_{n=1}^{\infty} |\langle x, x_n\rangle| \leq B||x||^2 \quad \text{for any } x \in H.$$

A and B are called lower frame bound and upper frame bound respectively. The largest possible lower frame bound and the smallest possible upper frame bound are said to be optimal. Frames generalize orthonormal bases.

Remark 2.2.11. *Note that the expression*

$$\sum_{n=1}^{\infty} |\langle x, x_n\rangle|^2 = ||(\langle x, x_n\rangle)_n||_{\ell^2}$$

defines an equivalent norm for H. Moreover, if $A = B = 1$, we have

$$||(\langle x, x_n\rangle)_n||_{\ell^2} = ||x||^2.$$

This shows a first similarity between frames and bases, as the latter equation represents a property of any orthonormal basis of a Hilbert space, namely Plancherel's Equality shown in 2.2.6.

We now exhibit some particular cases.

Definition 2.2.12. *Let $\{x_n\}$ be a frame for H.*

(i) If the optimal frame bounds are equal ($A = B$), then $\{x_n\}$ is said to be a tight, or A-tight frame.

(ii) If the optimal frame bounds are both equal to 1 ($A = B = 1$), then $\{x_n\}$ is called a Parseval frame.

(iii) If $\{x_n\}$ ceases to be a frame whenever any element is removed from the sequence, $\{x_n\}$ is called an exact frame.

It is useful at this point to highlight some remarks on frames.

1. In a frame repetition is allowed.
2. Taken a frame $\{x_n\}$, the sequences $\{0, x_1, x_2, ...\}$ and $\{x_1, x_1, x_2, ...\}$ are still frames.
3. An orthonormal basis $\{e_n\}$ is an exact Parseval frame.
4. Every frame is a complete sequence, since assuming $\langle x, xn\rangle$ yields, by 2.2.6
$$A||x||^2 \leq \sum_{n=1}^{\infty} |\langle x, x_n\rangle|^2 = 0 \Rightarrow x = 0.$$

5. If $\{x_n\}$ is a frame, then the series $\sum_n |\langle x, x_n\rangle|^2$ is absolutely convergent, therefore by 2.2.8 it also converges unconditionally. Thus, taken any permutation σ of N, we have that $\{x_\sigma(n)\}$ is still a frame. This implies that it does not matter which countable set we use to index a frame.

Frames generalize orthonormal basis. However, the following examples show that there exist frames which are not bases and bases that are not frames.

Example 2.2.13. *Let $\{e_n\}$ be an orthonormal basis for H. $\{e_1, e_1, e_2, e_2, ..., e_n, e_n, ...\}$ is a tight inexact frame (with bounds $A = B = 2$), but it is neither an orthogonal sequence nor a basis. Similarly, any sequence obtained as the union of two orthonormal bases is a frame.*

Example 2.2.14. *Let $\{e_n\}$ be an orthonormal basis for H. $\{e_1, e_2/2, e_3/3, ...\}$ is an orthogonal basis for H. However, it has no lower bound, hence it is not a frame.*

Remark 2.2.15. *One of the main advantages of frames over orthonormal bases is their redundancy. While redundancy implies that the coefficients of a frame expansion are not unique, it also ensures robustness: small perturbations of the coefficients do not compromise reconstruction. This property is particularly appealing from the modeling point of view, since biological systems such as the visual cortex are highly redundant.*

Moreover, every frame $\{x_n\}$ gives rise to a bounded, invertible operator $S : H \to H$ defined by

$$Sf = \sum_{n=1}^{\infty} \langle f, x_n\rangle x_n,$$

called the frame operator*. Its invertibility guarantees that any $f \in H$ can be reconstructed stably from its frame coefficients, thus generalizing the reconstruction formula valid for orthonormal bases.*

In the following section, we shall see how families of Gabor filters, obtained through translations, dilations, and rotations of a mother wavelet, provide concrete examples of frames in $L^2(\mathbb{R}^2)$. This will justify their use as a mathematical model for receptive profiles in the primary visual cortex.

2.3 Wavelets and linear filters

The Fourier transform provides a powerful tool for studying the frequency content of signals. In this chapter we review the main concepts and terminologies regarding the wavelet theory, limited to the cases that we are interested in treating later. Let us first specify the convention for the Fourier Transform which will be used throughout the chapter.

Definition 2.3.1. *For any $f \in L^2(\mathbb{R}^n)$, the Fourier Transform is defined as*

$$\hat{f}(\omega) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot\omega} f(x)dx, \quad \omega \in \mathbb{R}^n$$

where $x \cdot \omega$ is the scalar product in $\mathbb{R}^n$.

We aim to define the wavelet transform. We start by choosing a function ϕ, called mother wavelet, which in our case is a function in L^2. A set of functions is then obtained through transformations, specifically by dilations, rotations and translations, of the mother wavelet.
Let $\phi \in L^2(\mathbb{R}^2)$ be the 2D mother wavelet such that:

$$\phi_{\sigma,\theta,x_0,y_0}(x,y) := \frac{1}{|\sigma|}\phi^\theta\left(\frac{x-x_0}{\sigma}, \frac{y-y_0}{\sigma}\right)$$

where $\phi^\theta(x,y) = \phi(x\cos\theta + y\sin\theta, -x\sin\theta + y\cos\theta)$, σ is the dilation parameter, θ is the rotation parameter and (x_0, y_0) are the coordinates of the translation. We can now define the Continuous wavelet transform:

Definition 2.3.2. *Given the family $\{\phi_{\sigma,\theta,x_0,y_0}\}_{\sigma,\theta,x_0,y_0}$ obtained as before, given a signal $f \in L^2(\mathbb{R}^2)$, the Continuous Wavelet Transform is defined as*

$$WT(f)_{\sigma,\theta,x_0,y_0} := \langle f, \phi_{\sigma,\theta,x_0,y_0}\rangle_{L^2} = \int\int f(x,y)\overline{\phi_{\sigma,\theta,x_0,y_0}}(x,y)dxdy.$$

Given a suitable choice of the function ϕ we can recover the signal f from the coefficients of $WT(f)_{\sigma,\theta,x_0,y_0}$, through the inversion formula, which is

$$f(x,y) = \frac{1}{C_\phi}\int_0^{+\infty}\int_0^{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{\sigma^3}WT(f)_{\sigma,\theta,x_0,y_0}\phi_{\sigma,\theta,x_0,y_0}d\sigma d\theta dx_0 dy_0$$

with

$$C_\phi = 4\pi^2 \int_0^{+\infty} \int_0^{2\pi} \frac{|\hat{\phi}(\omega\cos\theta, \omega\sin\theta)|^2}{|\omega|} d\omega d\theta, \tag{2.3.1}$$

where $\hat{\phi}(u, v)$ is the Fourier transform of ϕ and $\omega = \sqrt{u^2 + v^2}$. Note that C_ϕ is not necessarily finite for any $\phi \in L^2(\mathbb{R}^2)$, that's why we can state the following admissibility condition, which expresses the property of a certain mother wavelet to produce a transform which is invertible:

Definition 2.3.3. *Given a mother wavelet $\phi \in L^2(\mathbb{R}^2)$, the constant C_ϕ defined above is called the admissibility constant of ϕ. The function ϕ is said to be admissible if $C_\phi < +\infty$.*

From a computational point of view, the continuous transform defined above is not adequate. In order to obtain a numerically efficient synthesis of a signal, a discretization of the set of parameters is required. One of the possible ways is the following.

$$\phi_{m,n,k,l} := a^{-m}\phi^l(a^{-m}x - nb, a^{-m}y - kb),$$

where

- ϕ is the mother wavelet;
- a, b and θ are fixed;
- $\phi^l(x, y) = \phi(x\cos(l\theta) + y\sin(l\theta), -xsin(l\theta) + y\cos(l\theta))$;
- θ denotes the stepsize of each angular rotation ($\theta = \frac{2\pi}{K}$ for some integer K);
- $a > 1, b > 0$;
- $(na^m b, ka^m b)$ represents the spatial translation and a^m represents the dilation:

$$(a^{-m}x - nb, a^{-m}y - kb) = \left(\frac{x - na^m b}{a^m}, \frac{y - ka^m b}{a^m}\right).$$

We note that parameters m, n, k have to vary in all $\mathbb{Z}$ in order to cover the whole space, while for the rotation parameter l we have $l = 0, ..., K-1$. If we denote $Q := \{0, ..., K-1\}$, then the family we are considering is

$$\{\phi_{m,n,k,l}\}_{m,n,k\in\mathbb{Z},l\in Q}.$$

An analysis performed through Continuous Wavelet Transform, but with a discrete wavelet family, produces the so-called Discrete Wavelet Transform. Specifically,

$$WT(f)_{m,n,k,l} := \langle f, \phi_{m,n,k,l}\rangle_{L^2(\mathbb{R}^2)}.$$

This transform provides a discrete set of coefficients instead of a continuous one. In this case, given an admissible mother wavelet ϕ, the synthesis of a signal f assumes the form of a discrete sum.
Wavelets provide an example of linear filters, i.e. linear operators (defined on certain spaces of functions) whose role is that of extracting specific pieces of information from the signals they are applied to. An example of linear filter can be obtained as follows: given $\psi \in L^2$, define

$$A : L^2 \to \mathbb{C}, \quad Af := \langle f, \psi\rangle_{L^2}$$

Note that a family of wavelets $\{\psi_N\}_N$ and their wavelet transforms determine a set of filters (usually referred to as a filter bank) of this kind

$$WT_N : L^2 \to \mathbb{C} \quad WT(f)_N := \langle f, \psi_N\rangle_{L^2}$$

Chapter 3

A model for static geometric optical illusions

3.1 Introduction

The goal of this chapter is to present a model for static geometric illusions. After briefly recalling a few static illusions, and in particular, the Ebbinghaus one, we focus on the processing performed by the visual cortex. The most common filters that model the receptive profiles (RPs) of simple cells in the primary visual cortex are Gabor filters. From both the mathematical and the neurophysiological perspectives, Gabor filters provide a natural representation of visual information. On the one hand, they belong to the family of wavelets, which are well-suited for multiscale signal analysis and can be organized into frames, allowing any signal to be decomposed and reconstructed without loss of information. On the other hand, experimental evidence shows that the structure of simple cell receptive fields is well captured by Gabor functions, which combine a Gaussian envelope with a sinusoidal carrier, reproducing the alternating excitatory and inhibitory subregions observed in physiology. In the following sections, we first recall some essential aspects of wavelet theory and its connection with linear filters. We then review how Gabor functions arise as a biologically meaningful subclass of wavelets, and

we illustrate why they provide an effective and widely accepted model for the set of receptive profiles in V1.

3.2 Ebbinghaus optical illusions

Here we focus in particular in the static optical illusion and more precisely on Ebbinghaus optical illusion. This is due to a misperception of the scale of inducers. In the best-known version of the illusion, two circles of identical size are placed near each other, and one is surrounded by large circles while the other is surrounded by small circles. As a result of the juxtaposition of circles, the central circle surrounded by large circles appears smaller than the central circle surrounded by small circles.

Recent work suggests that two other critical factors involved in the perception of the Ebbinghaus illusion are the distance of the surrounding circles from the central circle and the completeness of the annulus. Regardless of relative size, if the surrounding circles are closer to the central circle, the central circle appears larger and if the surrounding circles are far away, the central circle appears smaller. For this reason, the Ebbinghaus illusion is considered a scale illusion, as it affects the perceived size of the central target.

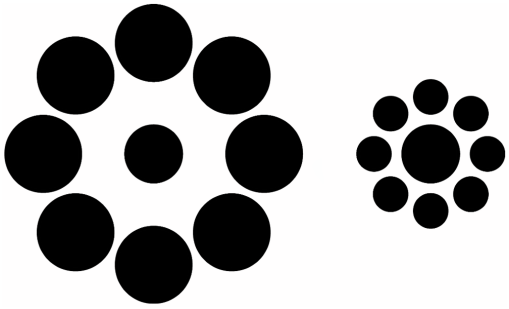

Figure 3.1: The Ebbinghaus illusions

Optical illusions can be studied starting from the functionality of the visual cortex.

In fact, differential models of the visual cortex have been developed by Petitot-Tondut [17], Citti-Sarti [20]: different areas of the visual cortex are

modeled as Lie groups, with a sub-Riemannian metric (see [2] for families of cells sensible to orientation and time). In this geometric setting, models of static geometric illusions were proposed by [5] and [6] in different Lie groups accounting for orientation or scale. In these models, the input visual stimulus induces a deformation in the geometry of the retina, which is at the basis of the illusion: as a result of this deformation, lines appear curved, and circles appear smaller or larger than they actually are.

3.3 Gabor wavelets as a model of receptive profiles

Since the process of vision is perform in the visual cortex, by the action of receptive profiles (RPs), our interest lies in wavelets those that provide a good model for them. Neurophysiological evidence shows that the receptive fields of visual neurons have a specific structure that is well described by Gabor filters. This section recalls the biological motivation for this choice and shows how Gabor wavelets naturally arise as models of simple cell RPs.

3.3.1 Receptive profiles in V1

It is a classical result in neurophysiology that ganglion cells in the retina, as well as LGN cells, have receptive profiles that are best modeled as Laplacians of Gaussians, as shown in Figure 3.2. Once the visual signal reaches the primary visual cortex (V1), it undergoes a more sophisticated processing.

V1 neurons can be divided into two main classes, as first discovered by Hubel and Wiesel in 1962: simple cells and complex cells. Simple cells are selective for the orientation of contours and behave as linear filters, while complex cells process higher-order features of the signal.

The basic structure of a simple cell's receptive field consists of alternating excitatory and inhibitory subregions, typically organized in a periodic fashion. This structure is well captured by Gabor functions, which are sinusoidal

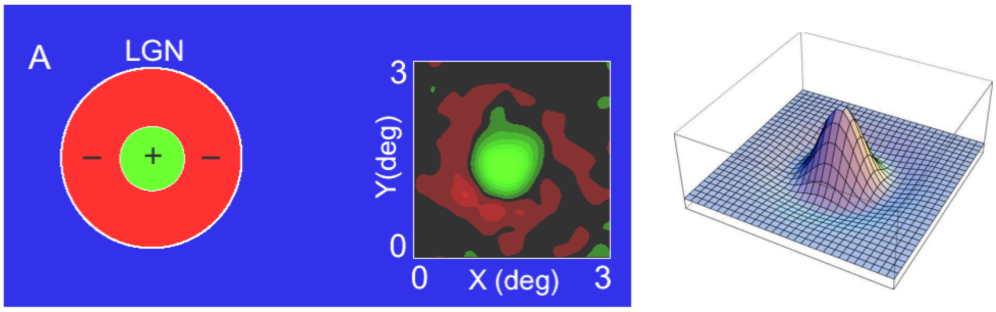

Figure 3.2: The receptive profile of an ON-center LGN cell. On the left, a schematic representation of the ON and OFF domains; in the middle, a visualization of the level sets of the RP. On the right, the Laplacian of Gaussian model for LGN RPs. Image taken from [16].

carriers modulated by Gaussian envelopes. In other words, Gabors combine the localization properties of Gaussians with the frequency selectivity of sinusoids, making them an excellent model of the RPs observed experimentally (Figure 3.3).

3.3.2 Gabor filters as receptive profiles

The most general family of 2D Gabor filters is

$$\psi_{x_0,y_0,u_0,v_0,\sigma,\beta}(x,y) = \frac{1}{2\pi\sigma\beta} e^{i(u_0(x-x_0)+v_0(y-y_0))} \cdot e^{-\left(\frac{(x-x_0)^2}{2\sigma^2}+\frac{(y-y_0)^2}{2\beta^2}\right)} \tag{3.3.1}$$

More precisely, in order to obtain real-valued filters, we consider the real and imaginary parts of the complex Gabor function, which correspond respectively to the even-symmetric and odd-symmetric receptive profiles observed in simple cells. Explicitly,

$$\psi^{\text{even}}_{x_0,y_0,u_0,v_0,\sigma,\beta}(x,y) = \frac{1}{2\pi\sigma\beta} e^{-\left(\frac{(x-x_0)^2}{2\sigma^2}+\frac{(y-y_0)^2}{2\beta^2}\right)} \cos(u_0(x-x_0)+v_0(y-y_0)),$$

$$\psi^{\text{odd}}_{x_0,y_0,u_0,v_0,\sigma,\beta}(x,y) = \frac{1}{2\pi\sigma\beta} e^{-\left(\frac{(x-x_0)^2}{2\sigma^2}+\frac{(y-y_0)^2}{2\beta^2}\right)} \sin(u_0(x-x_0)+v_0(y-y_0)).$$

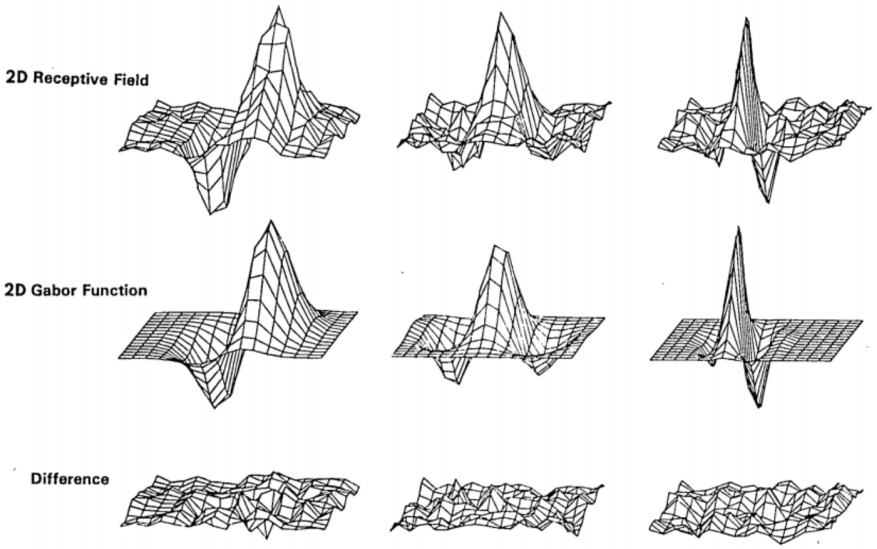

Figure 3.3: Illustration of experimentally measured 2D receptive profiles of three simple cells in cat striate cortex (first row). The plots displayed in the second row represent the best-fitting 2D Gabor function for them, with parameters fitted by least squares; the residual error, shown in the third row, was indistinguishable from random error for 33 of the 36 simple cells used in the experiments performed in the laboratory of L.A. Palmer and J.P. Jones (University of Pennsylvania Medical School). Image taken from [4]

An excellent description of the receptive profiles of simple cells in the visual cortex is provided by members of this family of optimal 2D filters, as one can see, for instance, in figure 3.3, which shows the comparison between the estimated receptive profiles of three simple cells in cat cortex and the best-fitting 2D Gabor functions for them.

By introducing a rotation parameter θ into the Gaussian envelope, one obtains oriented 2D Gabor filters that can capture directional selectivity. Furthermore, by including temporal and velocity variables, it is possible to construct 3D Gabor filters modeling time-dependent receptive profiles:

$$\psi_{q,v,\theta}(x,y,t) = e^{-2\pi i \frac{k}{\sigma}((x-q_x)\cos\theta+(y-q_y)\sin\theta+v(t-q_t))} e^{-\frac{(x-q_x)^2}{2\sigma^2}-\frac{(y-q_y)^2}{2\sigma^2}-\frac{(t-q_t)^2}{2\sigma^2}}.$$

In Chapter 4, we will make use of such extensions in order to describe spatiotemporal receptive profiles and to study time-dependent visual phenomena.

3.3.3 Gabor filters as a frame

The link between the mathematical and biological viewpoints is provided by frame theory. Families of Gabor filters generated by translations, dilations, and rotations can form frames in $L^2(\mathbb{R}^2)$, thus guaranteeing stable signal decomposition and reconstruction. This property makes them not only a biologically validated choice, but also a mathematically robust one.

Proposition 3.3.1. *Let us consider the family* $\{\phi_{m,n,k,l}\}_{m,n,k,l}$ *defined as* $\phi_{m,n,k,l} := a^{-m}\phi^l(a^{-m}x - nb, a^{-m}y - kb)$ *where* ϕ *is a Gabor mother wavelet defined as* $\phi^l(x,y) = \phi(x\cos(l\theta) + y\sin(l\theta), -x\sin(l\theta) + y\cos(l\theta))$ *and* $m, n, k \in \mathbb{Z}$, $l \in Q$, *with* $Q := 0, ..., K-1$.
The family $\{\phi_{m,n,k,l}\}_{m,n,k,l}$ *forms a frame for* $L^2(\mathbb{R}^2)$, *i.e. there exist constants* $A > 0$ *and* $B < +\infty$ *such that*

$$A\|f\|^2 \leq \sum_{m,n,k,l} |\langle f, \phi_{m,n,k,l}\rangle|^2 \leq B\|f\|^2, \quad \forall f \in L^2(\mathbb{R}^2).$$

The constants A and B, known as the frame bounds, can be computed explicitly under suitable assumptions on the parameters a, b, θ (see [7]). For our purposes, it's sufficient to recall that such bounds exist, ensuring that the family is suitable as a mathematical model for receptive profiles.
To conclude, Gabor functions emerge as a privileged choice for modeling receptive profiles of V1 simple cells. They combine the localization and multiscale properties of wavelets with the orientation and frequency selectivity required by neurophysiology, while at the same time forming frames in $L^2(\mathbb{R}^2)$. For these reasons, Gabor filters will serve as the foundation for the models developed in the following chapters, where the focus will shift from filter design to the underlying geometric structures of visual perception.

3.4 Neurogeometry of V1 and GOIs

In this section, we focus on the Ebbinghaus illusion (see Figure 3.1, right), and we present a mathematical model of the cortical processing underlying the perceptual deformation, following the approach of [6].

3.4.1 Output of receptive profiles

Let $x = (x_1, x_2)$ denote local coordinates on the retina $\mathbb{R}^2$, and let $M \subset \mathbb{R}^2$ be the visual field. When a visual stimulus I of intensity $I(x_1, x_2) : M \subset \mathbb{R}^2 \to \mathbb{R}^+$ activates the retinal layer of photoreceptors, the neurons whose RFs intersect M spike. As we saw previously, their spike frequencies $O(x_1, x_2, f)$ can be modeled (taking into account just linear contributions) as the integral of the signal $I(x_1, x_2)$ with the set ψ of Gabor filters. The expression for this output is

$$O(x_1, x_2, f) = \int_{\mathbb{R}^2} I(\xi_1, \xi_2)\, \psi_{x_1, x_2, f}(\xi_1, \xi_2)\, d\xi_1 d\xi_2 \tag{3.4.1}$$

3.4.2 Cortical connectivity

Since the output O is a higher dimensional function defined over the cortical space, the lateral connectivity propagates this output in the cortical space $\mathbb{R}^2 \times F$ giving rise to the cortical activity. A good model for the cortical connectivity can be obtained describing $\mathbb{R}^2 \times F$ as a Lie group, endowed with a sub-Riemannian metric. We are interested in scale models, we recall here the model of Sarti, Citti and Petitot in [21] and [17], who proposed a model of scale and orientation selectivity in $\mathbb{R}^2 \times S1 \times \mathbb{R}^+$, with the group laws of translation, rotation and dilation. Let $\sigma \in \mathbb{R}^+$ be the variable that accounts for size and $\theta \in S^1$ be the one accounting for orientation, then a basis of left invariant vector fields is

$$\begin{aligned} &X_1 = \sigma(\cos\theta \partial_{x_1} + \sin\theta \partial_{x_2}); \quad X_2 = \sigma \partial_\theta; \\ &X_3 = \sigma^2(-\sin\theta \partial_{x_1} + \cos\theta \partial_{x_2}); \quad X_4 = -\sigma^2 \partial_\sigma; \end{aligned} \tag{3.4.2}$$

In the present setting, as $\sigma \to 0$, the vector fields X_3 and X_4 vanish, while X_1 and X_2 have a non-zero limit. Moreover X_3 belongs to the Lie algebra generated by X_1 and X_2, in fact let $X_1 = \sigma(\cos\theta\,\partial_{x_1} + \sin\theta\,\partial_{x_2})$ and $X_2 = \sigma\,\partial_\theta$. For any smooth $\varphi(x_1, x_2, \theta, \sigma)$,

$$[X_1, X_2]\varphi \;=\; X_1(X_2\varphi) - X_2(X_1\varphi).$$

Since σ is independent of (x_1, x_2) and θ is independent of (x_1, x_2),

$$X_1(X_2\varphi) = \sigma(\cos\theta\,\partial_{x_1} + \sin\theta\,\partial_{x_2})(\sigma\,\partial_\theta\varphi) = \sigma^2(\cos\theta\,\partial_{x_1} + \sin\theta\,\partial_{x_2})(\partial_\theta\varphi).$$

Next,

$$\begin{aligned} X_2(X_1\varphi) &= \sigma\,\partial_\theta\Big(\sigma(\cos\theta\,\partial_{x_1} + \sin\theta\,\partial_{x_2})\varphi\Big) \\ &= \sigma^2\Big(-\sin\theta\,\partial_{x_1} + \cos\theta\,\partial_{x_2}\Big)\varphi \;+\; \sigma^2(\cos\theta\,\partial_{x_1} + \sin\theta\,\partial_{x_2})(\partial_\theta\varphi). \end{aligned}$$

Subtracting,

$$[X_1, X_2]\varphi = -\,\sigma^2\Big(-\sin\theta\,\partial_{x_1} + \cos\theta\,\partial_{x_2}\Big)\varphi = -\,X_3\,\varphi,$$

i.e.

$$[X_1, X_2] \;=\; -\,X_3, \quad X_3 = \sigma^2(-\sin\theta\,\partial_{x_1} + \cos\theta\,\partial_{x_2}).$$

The fourth vector field is $X_4 = -\sigma^2\,\partial_\sigma$. Therefore, the Lie algebra generated by $\{X_1, X_2\}$ already yields X_3, and together with X_4 it spans the full tangent space. Hence the Hörmander condition (2.1.19) is satisfied.

We now introduce the sub-Riemannian metric with which Citti and Sarti in [20] proposed to endow the $\mathbb{R}^2 \times S^1$ group to model the connectivity of the primary visual cortex V1. Starting from the vector fields X_1, X_2, X_3, X_4

we define the metric tensor g^{ij} as the product:

$$g^{ij}(x,y,\theta,\sigma) = \begin{pmatrix} \sigma\cos\theta & -\sigma^2\sin\theta & 0 & 0 \\ \sigma\sin\theta & \sigma^2\cos\theta & 0 & 0 \\ 0 & 0 & \sigma & 0 \\ 0 & 0 & 0 & -\sigma^2 \end{pmatrix} \begin{pmatrix} \sigma\cos\theta & \sigma\sin\theta & 0 & 0 \\ -\sigma^2\sin\theta & \sigma^2\cos\theta & 0 & 0 \\ 0 & 0 & \sigma & 0 \\ 0 & 0 & 0 & -\sigma^2 \end{pmatrix}$$

$$= \begin{pmatrix} \sigma^2\cos^2\theta + \sigma^4\sin^2\theta & \sigma^2\sin\theta\cos\theta - \sigma^4\sin\theta\cos\theta & 0 & 0 \\ \sigma^2\sin\theta\cos\theta - \sigma^4\sin\theta\cos\theta & \sigma^4\cos^2\theta + \sigma^2\sin^2\theta & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^4 \end{pmatrix}.$$

Therefore, the Lie algebra generated by $\{X_1, X_2\}$ spans the entire tangent space, and the Hörmander condition (2.1.19) is satisfied.
It is possible to recognize that g has a block structure,

$$g \;=\; g_1 \oplus g_2 \oplus g_3,$$

where

- g_1 is the 2×2 block depending on (x, y, θ, σ),
- $g_2 = (\sigma^2)$, corresponding to the orientation variable θ,
- $g_3 = (\sigma^4)$, corresponding to the scale variable σ.

For each point $(x, y, \theta, \sigma) \in \mathbb{R}^2 \times S^1 \times \mathbb{R}$, we restrict the connectivity tensor $g^{ij}(x, y, \theta, \sigma)$ with $i, j = 1, 2, 3, 4$ to the $\mathbb{R}^2$ plane generated by $\{\partial_x, \partial_y\}$ and obtain g_1.

3.5 The model for scale/size GOIs: static case

3.5.1 Scale selection in V1

We focus here on the detection of scale. This is an isotropic feature, and can be selected by isotropic cells. A good model for their receptive profiles

are Gabor filters.

Let (x, y) be the variables identifying position on the retina, (q_x, q_y) those identifying position on the cortex. Let k be a fixed parameter accounting for spatial frequency. Denoting T_{q_x,q_y} as a the translation by a vector (q_x, q_y) and D_σ as a dilation of amplitude σ, a good expression for the bank of filters is:

$$\psi_{q_x,q_y,\sigma,\theta}(x,y) = e^{-2\pi i \frac{k}{\sigma}((x-q_x)\cos\theta+(y-q_y)\sin\theta)} e^{-\frac{(x-q_x)^2}{2\sigma^2}-\frac{(y-q_y)^2}{2\sigma^2}}$$

or equivalently,

$$\psi_{q_x,q_y,\sigma,\theta} = D_\sigma T_{q_x,q_y}(\psi_0) \quad \text{where } \psi_0(\eta,\tau) = e^{2\pi i k(\eta+\tau)} e^{-\eta^2-\tau^2}.$$

This bank of filters acts on the initial stimulus $I : D \to \mathbb{R}$ and the hypercolumns response of simple cells in each hypercolumn provides an output. When considering only the scale feature, i.e. with $f = \sigma$, the output reduces to:

$$O(q_x, q_y, \sigma, \theta) = \int I(x,y)\psi_{(q_x,q_y,\sigma,\theta)}(x,y)dxdy$$

The non-maximal suppression mechanism selects the maxima over the orientation and scale hypercolumns, providing the selection of two maximal outputs for both features. In this setting we are interested in the scale parameter only, so we extract

$$O(q_x, q_y, \bar{\sigma}, \theta) = \max_\sigma O(q_x, q_y, \sigma, \theta)$$

The selected scale $\bar{\sigma}(q_x, q_y)$ corresponds to the value of σ that maximizes $O(q_x, q_y, \sigma, \theta)$ at each point:

$$\bar{\sigma}(q_x, q_y) = \arg\max_\sigma O(q_x, q_y, \sigma, \theta).$$

$\bar{\sigma}$ represents the distance from the nearest boundary, selected by the hypercolumns containing filters at all possible scales. In figure 3.5 we see the hypercolumnar structure for the scale model: a bank of filter with different scales is visualized over a gray circle which represents the visual stimulus. In Figure 3.4 we show the initial visual stimulus and the corresponding scale

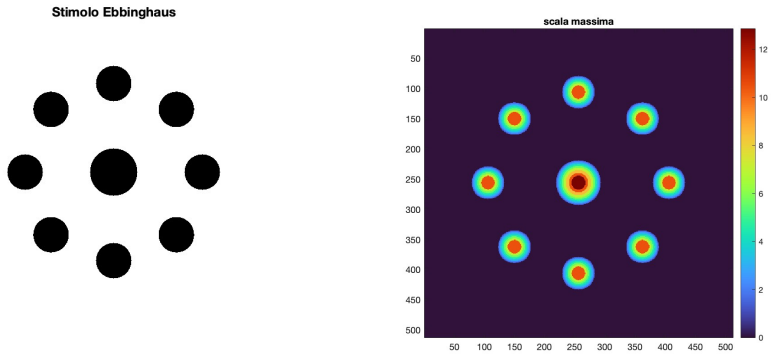

Figure 3.4: The visual stimulus and the maximum scale map

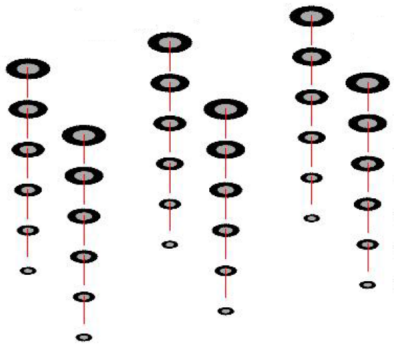

Figure 3.5: The hypercolumnar structure for the scale model.

map. The level lines of the function $\bar{\sigma}$ appear as circles, indicating regions of equal distance from the closest edge. This mechanism enables the visual system to extract spatial scale at each location, forming a cortical representation of size.

3.5.2 Size selection in V1

Simple cells in V1 are able to select the scale of an object, which corresponds approximately to the distance from its boundary. The size of the object, representing the spatial extent of the observed element, is strictly related to this scale.

Once the distance function from the boundary $\bar{\sigma}(q_x, q_y)$ has been defined,

we assume that cortical connectivity propagates the output within each perceptual unit. Since the perceived size $\rho(q_x, q_y)$ of an object can be identified with the maximum distance from the boundary within that unit, we define $\rho(q_x, q_y)$ by applying a non-maximal suppression procedure across each perceptual unit.

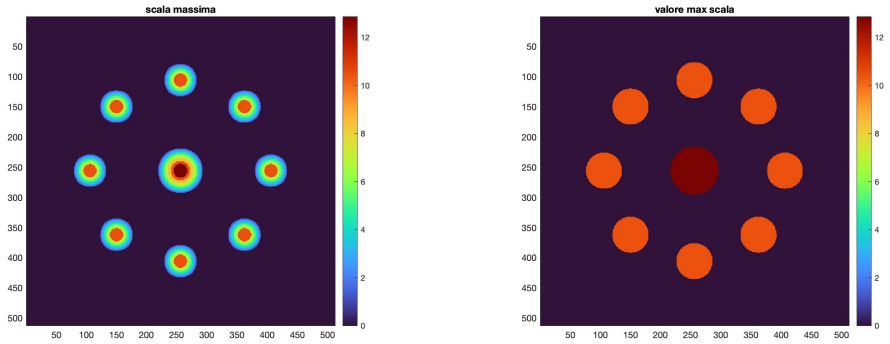

Figure 3.6: The uniform scale map

We directly assign to each point in the perceptual unit the maximum value of $\bar{\sigma}$ attained within that unit:

$$\rho(q_x, q_y) = \max_{(p_x,p_y)\in\mathcal{U}(q_x,q_y)} \bar{\sigma}(p_x, p_y),$$

where $\mathcal{U}(q_x, q_y)$ denotes the perceptual unit containing the point (q_x, q_y).

This construction yields a piecewise constant size map $\rho(q_x, q_y)$, where each unit is characterized by a single representative scale. In Figure 3.6, we visualize this process: starting from the left map showing $\bar{\sigma}(q_x, q_y)$, we extract a constant size value in each circle representing a perceptual unit, forming the size map on the right.

3.5.3 Displacement vector field

The perceived deformation induced by the illusion can be represented as a displacement vector field $u = (u_1, u_2)$, defined on the retinal domain

$M \subset \mathbb{R}^2$. It was proved by [6] that u is a solution of the following system:

$$\begin{cases} \Delta u_1 = \partial_1 p_{11} + 2\,\partial_2 p_{12} - \partial_1 p_{22} & \text{in } M, \\ \Delta u_2 = \partial_2 p_{22} + 2\,\partial_1 p_{12} - \partial_2 p_{11} & \text{in } M, \\ \partial_{\vec{n}} u_1 = \partial_{\vec{n}} u_2 = 0 & \text{on } \partial M, \end{cases} \tag{3.5.1}$$

∂M is Lipschitz continuous, with normal defined almost everywhere. Solutions for equation 3.5.1 are well defined up to an additive constant, which is recovered imposing $u(0,0) = 0$ for symmetry reasons, where $(0,0)$ is the center of the initial domain M. Let us explicitly note that p is obtained after convolution of Gabor filters, so that it is differentiable, allowing to write the system. Hence, we solve 3.5.1, recovering the displacement field.

Chapter 4

Introducing time in the model

4.1 From the geometry of the visual cortex to visual illusions

4.1.1 A sub-Riemannian geometry for movement

In this section, we look for a suitable sub-Riemannian structure able to model the behavior of cells sensitive to movement.

The retina receives the external visual stimulus. It is in general approximated as a 2D space. Here we will denote with (x, y) its variables and we will denote with t the time. The stimulus is then transmitted to the visual cortex: we are interested in the area of the cortex that processes movement. We will denote with (q_x, q_y) the cortical variables that identify position, with q_t the one identifying time, with θ the direction of the level lines of the stimulus image, with σ the perceived size and with v the perceived speed of movement. The whole set of variables defines a manifold

$$\mathcal{M} = \mathbb{R}^2 \times \mathbb{R}_t \times S^1 \times \mathbb{R}^+_\sigma \times \mathbb{R}_v. \tag{4.1.1}$$

Since we perceive movement only in the direction orthogonal to the level lines, the main differential relation between parameters is $\omega = e^{-\sigma}(\cos\theta dx + \sin\theta dy - vdt) = 0$. We will then consider a sub-Riemannian manifold, iden-

tified by $\mathcal{M}$ and $\ker(\omega)$ as a horizontal distribution. A basis of the kernel is

$$X_1 = -e^{\sigma}\sin\theta\partial_x + e^{\sigma}\cos\theta\partial_y \quad X_2 = \partial_\theta,$$
$$X_3 = \partial_\sigma \quad X_4 = \partial_v, \quad X_5 = e^{\sigma}(v\cos\theta\partial_x + v\sin\theta\partial_y + \partial_t).$$

From now on we will use the same symbol for a vector field and its associated first-order operator.

Remark 4.1.1. *Let us note that the Lie algebra generated by the vector fields X_1, X_2, X_3, X_4, X_5 satisfies the Hörmander condition (see definition 2.1.19 above).*

Definition 4.1.2. *We will consider in the sequel the manifold $(\mathcal{M}, \Delta, g)$, where $\mathcal{M}$ is defined in (4.1.1), $\Delta = span(X_1, X_2, X_3, X_4, X_5)$, and g is the metric which makes X_1, X_2, X_3, X_4, X_5 orthonormal.*

Let us explicitly note that, in order to describe dynamic illusion dependent on the scale instead of the orientation, we would need to change the definition of ω, and proceed in an analogous way for the definition of the space, and the metric g.

4.1.2 Cortical activity in presence of a stimulus

The intensity of the response is captured by the receptive profile (RPs). In this area, receptive profiles are modeled by 3-dimensional Gabor functions of the form

$$\psi_{q,v,\theta}(x,y,t) = e^{-2\pi i\frac{k}{\sigma}((x-q_x)\cos\theta+(y-q_y)\sin\theta+v(t-q_t))}e^{-\frac{(x-q_x)^2}{2\sigma^2}-\frac{(y-q_y)^2}{2\sigma^2}-\frac{(t-q_t)^2}{2\sigma^2}} \tag{4.1.2}$$

The action of each cell on the input stimulus can be described as a lifting from the space of the visual stimulus $L^2(\mathbb{R}^2\times\mathbb{R}^+)$ to the 6 dimensional feature space $\mathcal{M}$. The activity in V1 (primary visual cortex) is modeled by a map

$$F_{q,\theta,\sigma,v}(f) = \langle\psi_{q,\theta,\sigma,v,}, f\rangle_{L^2(\mathbb{R}^2\times R^+)} \tag{4.1.3}$$

where f is the visual stimulus. This inner product measures the similarity between the Gabor function and the stimulus.

4.1.3 The geometry induced by the stimulus and optical illusions

It is possible to recognize that the metric g is a block matrix, i.e. $g = g_1 \oplus g_2 \oplus g_3 \oplus g_4$, where g_1 is a 3×3 matrix depending on (x, y, t), g_2, g_3 and g_4 are 1×1 matrices depending, respectively, on θ, σ and v. Starting from the vector fields $\vec{X}_1, \vec{X}_2, \vec{X}_3, \vec{X}_4, \vec{X}_5$ we define a metric g^{ij}:

$$g^{i,j}(x,y,t,\theta,\sigma,v) = \begin{pmatrix} -e^{\sigma}\sin\theta & e^{\sigma}v\cos\theta & 0 & 0 & 0 \\ e^{\sigma}\cos\theta & e^{\sigma}v\sin\theta & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -e^{\sigma}\sin\theta & e^{\sigma}\cos\theta & 0 & 0 & 0 & 0 \\ e^{\sigma}v\cos\theta & e^{\sigma}v\sin\theta & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} =$$

$$\begin{pmatrix} e^{2\sigma}(\sin^2\theta + v^2\cos^2\theta) & -e^{2\sigma}(\sin\theta\cos\theta + v^2\cos\theta\sin\theta) & e^{\sigma}v\cos\theta & 0 & 0 & 0 \\ -e^{2\sigma}(\sin\theta\cos\theta + v^2\cos\theta\sin\theta) & e^{2\sigma}(\cos^2\theta + v^2\sin^2\theta) & e^{\sigma}v\sin\theta & 0 & 0 & 0 \\ e^{\sigma}v\cos\theta & e^{\sigma}v\sin\theta & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

with $i, j = 1, 2, 3, 4, 5$. It is possible to recognize that the metric g is a block matrix, i.e. $g = g_1 \oplus g_2 \oplus g_3 \oplus g_4$, where g_1 is a 3×3 matrix depending on (x, y, t), g_2, g_3 and g_4 are 1×1 matrices depending, respectively, on θ, σ and v. We restrict the tensor $g^{i,j}$ to $\mathbb{R}^2 \times \mathbb{R}^+$, obtaining the tensor

$$g_1 = \begin{pmatrix} e^{2\sigma}(\sin^2\theta + v^2\cos^2\theta) & -e^{2\sigma}(\sin\theta\cos\theta + v^2\cos\theta\sin\theta) & e^{\sigma}v\cos\theta \\ -e^{2\sigma}(\sin\theta\cos\theta + v^2\cos\theta\sin\theta) & e^{2\sigma}(\cos^2\theta + v^2\sin^2\theta) & e^{\sigma}v\sin\theta \\ e^{\sigma}v\cos\theta & e^{\sigma}v\sin\theta & 1 \end{pmatrix}. \tag{4.1.4}$$

Hence we will define the following metric on the space (x, y, t), induced by the integration of the variables θ, σ, v. Hence we will define the following metric

on the space (x, y, t), induced by the integration of the variables θ, σ, v

$$p(q) = \frac{\int |F_{q,\theta,\sigma,v}(f)| g_1 d\theta d\sigma dv}{\int |F_{q,\theta,\sigma,v}(f)| d\theta d\sigma dv}. \tag{4.1.5}$$

Let us explicitly note that, in general, since it is obtained by integration, $p(q)$ is Riemannian, and it is a metric induced on $\mathbb{R}^2 \times \mathbb{R}^+$ by the stimulus coded in the output $|F_{q,\theta,\sigma,v}(f)|$. We will interpret this change of metric in $\mathbb{R}^3$ as a deformation induced by the input images.

The next and final step of the algorithm will be to obtain a change of variable naturally associated to the problem

$$\Phi : (\mathbb{R}^2 \times \mathbb{R},\ \mathbf{p}) \longrightarrow (\mathbb{R}^2 \times \mathbb{R},\ \mathbf{Id}) \tag{4.1.6}$$

In this case, calling u the displacement, the perceived image will be computed as

$$f_p = f \circ \Phi = f \circ (u + Id) \tag{4.1.7}$$

4.2 Existence of flat coordinates and deformation of the space

4.2.1 Flat coordinates and deformations

We will now look for an isometry ϕ between the $\mathbb{R}^2$ plane equipped with the metric $\mathbf{p}$ and the $\mathbb{R}^2$ plane equipped with the metric $\mathbf{Id}$. The problem of finding the change of variables ϕ as in (4.1.6) has been solved by Riemann in the case of $\mathbf{p}$ being a Riemannian metric; in [1] the same problem has been solved for a large class of degenerate metrics. Let us first note that

Remark 4.2.1. *Consider a metric $p = p_{ij}$ on a domain in R^n with coordinates $x_1, ..., x_n$, consider a change of variables Φ and call y the new coordinates. Under this change of coordinates, the metric p takes the form*

$$\delta_{ij}(y) = \sum_{r,s} p_{rs}(x) \frac{\partial(\Phi^{-1})^r}{\partial y^i} \frac{\partial(\Phi^{-1})^s}{\partial y^j}. \tag{4.2.1}$$

Definition 4.2.2. *A coordinate system is called flat if, in this coordinate system, the metric p is given by a constant matrix.*

Definition 4.2.3. *We call Christoffel symbols of the first kind*

$$\Gamma_{ij,s} := \tfrac{1}{2}\left(\tfrac{\partial p_{js}}{\partial x^i} + \tfrac{\partial p_{is}}{\partial x^j} - \tfrac{\partial p_{ij}}{\partial x^s}\right). \tag{4.2.2}$$

We call Christoffel symbols of the second kind the numbers Γ^i_{jk}, with $\Gamma^i_{jk} = \Gamma^i_{kj}$, satisfying

$$\sum_s \left(\Gamma^s_{jk} p_{is} + \Gamma^s_{ik} p_{js}\right) = \frac{\partial p_{ij}}{\partial x^k}. \tag{4.2.3}$$

Theorem 4.2.4. *Christoffel symbols of the second kind exist if and only if the following condition holds:*

$$\sum_s \Gamma_{ij,s} v^s = 0 \quad \textit{for every } v^s \in \mathcal{R}, \tag{4.2.4}$$

where

$$\mathcal{R} := \mathcal{R}_p(x) := \mathrm{Kernel}(p) := \{v \in T_x M \mid p(v,\ \cdot) = 0\}. \tag{4.2.5}$$

If such numbers ijk exist, the "freedom" in choosing them is the addition of possibly several terms of the form

$$v^i T_{jk} \textit{ with } v \in \mathcal{R} \textit{ and } T_{jk} = T_{kj}. \tag{4.2.6}$$

Proof. Let's fix a point x and view 4.2.3 as a system of linear equations where the unknowns are Γ^i_{jk}. The coefficients of this system come from p and the derivatives of p.

The relation 4.2.3 is linear in the Christoffel symbols. To make this explicit, let us consider the collection of unknowns

$$\Gamma^s_{jk}, \qquad 1 \le s \le n,\ 1 \le j \le k \le n.$$

Since $\Gamma^s_{jk} = \Gamma^s_{kj}$, for each fixed s, there are $\frac{n(n+1)}{2}$ independent coefficients. Hence the total number of unknowns is

$$N \ = \ n \cdot \frac{n(n+1)}{2} \ = \ \frac{n^2(n+1)}{2}.$$

We therefore arrange all Christoffel symbols into a single vector

$$y \;=\; (y_1, \ldots, y_N) \in \mathbb{R}^N,$$

where each entry corresponds to one of the Γ^s_{jk}.

Equation (4.2.3), for fixed indices (i, j, k), then provides a linear equation in the unknowns y. The coefficients of this equation are given by the components of the metric p_{ij}, while the right-hand side is the known quantity $\partial_k p_{ij}$. Collecting all such relations for all triples (i, j, k) we get a linear system of the form

$$Ay = b,$$

where A is a matrix of size $N \times N$ with entries determined by the metric coefficients, y is the vector of Christoffel symbols, and b is the vector containing the partial derivatives of the metric coefficients.

Now we aim to reduce the equation 4.2.3 by algebraic manipulations. By lowering the free upper index [1] , we define the following quantities (employing Christoffel symbols of the first kind):

$$\Gamma_{jk,i} := \sum_{s=1}^{n} p_{is}\, \Gamma^s_{jk}, \qquad \Gamma_{ik,j} := \sum_{s=1}^{n} p_{js}\, \Gamma^s_{ik}.$$

We get

$$\Gamma_{jk,i} + \Gamma_{ik,j} = \partial_k p_{ij}, \tag{4.2.7}$$

we can solve for $\Gamma_{jk,i}$:

$$\Gamma_{jk,i} = \partial_k p_{ij} - \Gamma_{ik,j}.$$

We recall the explicit formula

$$\Gamma_{jk,i} = \frac{1}{2}\big(\partial_j p_{ik} + \partial_k p_{ij} - \partial_i p_{jk}\big),$$

[1] The operation of raising and lowering indices is a standard construction in Riemannian geometry. At each point of the manifold, the metric p defines a bilinear form $p : T_pM \times T_pM \to \mathbb{R}$, which induces a canonical isomorphism $\flat : T_pM \to T_p^*M$ between the tangent and cotangent spaces, given by $v \mapsto (w \mapsto g(v, w))$. In coordinates, this correspondence is expressed by $v_i = g_{ij} v^j$ and $v^i = g^{ij} v_j$, so any free index may be moved up or down.

expressing the first-kind symbols in terms of derivatives of the metric p.

By definition of the first-kind symbols,

$$\sum_s p_{is}\,\Gamma^s_{jk} = \Gamma_{jk,i} = \frac{1}{2}(\partial_j p_{ik} + \partial_k p_{ij} - \partial_i p_{jk}).$$

Define $\Gamma_{ij,k} := \Gamma_{jk,i}$ and swap indices i and k:

$$\sum_s p_{sk}\,\Gamma^s_{ij} = \Gamma_{ij,k} = \frac{1}{2}(\partial_i p_{jk} + \partial_j p_{ik} - \partial_k p_{ij}). \tag{4.2.8}$$

It is now a linear system for the second-kind Christoffel symbols Γ^s_{ij}. In this way all derivative terms are encoded in $\Gamma_{ij,k}$.

This shows that any solution Γ^s_{jk} of 4.2.3 corresponds to a solution $\Gamma_{jk,i}$ of 4.2.8.

Hence the two formulations are equivalent:

$$\sum_s \left(\Gamma^s_{jk}\,p_{is} + \Gamma^s_{ik}\,p_{js}\right) = \frac{\partial p_{ij}}{\partial x^k} \quad \Longleftrightarrow \quad \sum_s p_{sk}\,\Gamma^s_{ij} = \Gamma_{ij,k}.$$

It remains to observe that the condition $a^t b = 0$ applied to 4.2.8 is just the condition 4.2.4.

Concerning the "freedom", for a linear system of equations

$$Ay = b$$

such that the coefficient matrix A and the free terms b depend on x, one can find a smooth solution provided that a solution exists at every point and that the rank of A is constant. □

Even if the Christoffel symbols of the second kind are not unique, they allow to determine the Riemann tensor in a unique way.

Theorem 4.2.5. *Suppose that the rank of p is constant and assume that* (4.2.4) *is fulfilled at any point. Then, for any smooth functions Γ^i_{jk} with $\Gamma^i_{jk} = \Gamma^i_{kj}$ satisfying* (4.2.3)*, the functions*

$$R_{ijk\ell} := \sum_s p_{is}\left(\tfrac{\partial}{\partial x^k}\Gamma^s_{j\ell} - \tfrac{\partial}{\partial x^\ell}\Gamma^s_{jk} + \sum_a \left(\Gamma^s_{ka}\Gamma^a_{\ell j} - \Gamma^s_{\ell a}\Gamma^a_{jk}\right)\right) \tag{4.2.9}$$

do not depend on the "freedom" 4.2.6, i.e. they do not depend on the particular choice of Christoffel symbols of the second kind.

Proof. Since we want to show that the functions R_{ijkl} do not depend on the "freedom" 4.2.6, we plug in the formula 4.2.9 the term $\tilde{\Gamma}^i_{jk} = \Gamma^i_{jk} + v^i T_{jk}$ with $v \in \mathcal{R}$ instead of Γ. The terms of the form $v^s \frac{\partial}{\partial m} T_{jk}, v^s \frac{\partial}{\partial k} T_{jm}, v^s T_{ka} \tilde{\Gamma}^a_{lj}, v^s T_{la} \tilde{\Gamma}^a_{kj}$ vanish after contracting with p_{is} because of the definition 4.2.5 of $\mathcal{R}$. The result of the new equation differs from the original equation for R_{ijkl} by

$$\sum_s p_{is} T_{jl} \frac{\partial v^s}{\partial x^k} + \sum_a v^a \Gamma_{ka,i} T_{jl} - \sum_s p_{is} T_{kl} \frac{\partial v^s}{\partial x^j} - \sum_a v^a \Gamma_{la,i} T_{jk} \qquad (4.2.10)$$

Next, since by hypothesis we have that $v \in \mathcal{R}$ and, by definition, $\mathcal{R} := \{v \in T_x M | p(v, \cdot) = 0\}$, we have that

$$p(v, \cdot) = 0 \iff \sum_s p_{is} v^s = 0 \quad \forall i$$

Hence, we get the condition $\sum_s p_{is} v^s = 0$. This implies that $\sum_s p_{is} T_{jl} \frac{\partial v^s}{\partial x^k} = -\sum_s T_{jl} v^s \frac{\partial p_{is}}{\partial x^k}$. This, together with 4.2.2, implies that the sum of the first two terms of 4.2.10 is equal to $\sum_s T_{jl} \Gamma_{ki,s} v^s$, which is equal to zero by 4.2.4. Similarly, the sum of the last two terms is zero. Since we proved that 4.2.10 is equal to zero, we can conclude that the "freedom" in choosing Γ does not affect $R_{ijk\ell}$. □

Theorem 4.2.6. *Under the assumptions of the previous theorem, there exist flat coordinates for p if and only if the Christoffel symbols of the second kind exist and*

$$R_{ijk\ell} = 0 \qquad (4.2.11)$$

for every i, j, k, ℓ.

Proof. We are going to prove that the conditions stated in the theorem are necessary and sufficient. First of all, we know that $R_{ijk\ell}$ is a tensor, so if flat coordinates exist, then $R_{ijk\ell} = 0$, which is condition 4.2.11. This means that the conditions listed in the theorem are necessary. Now we aim to prove that they are sufficient. We first observe that for any $v \in \mathcal{R}$, the Lie derivative of

the metric p is given by

$$\begin{aligned}(\mathcal{L}_v p)_{ij} &= \sum_s \left(v^s \frac{\partial p_{ij}}{\partial x^s} + p_{is} \frac{\partial v^s}{\partial x^j} + p_{js} \frac{\partial v^s}{\partial x^i} \right) \\ &= \sum_s \left(v^s \frac{\partial p_{ij}}{\partial x^s} - v^s \frac{\partial p_{is}}{\partial x^j} - \frac{\partial p_{is}}{\partial x^i} \right) \\ &= -2 \sum_s v^s \Gamma_{ij,s} = 0\end{aligned} \tag{4.2.12}$$

where the last equality holds since, by hypothesis, the Christoffel symbols of the second kind exist. The metric p is then preserved by its flow.

Next we show that $\mathcal{R}$ is integrable, that is, for any two $u, v \in \mathcal{R}$ the commutator $[u, v]$ still lies in $\mathcal{R}$, i.e. $\mathcal{R}$ is closed under Lie brakets. We obtain this result by direct computation:

$$\begin{aligned}\sum_i p_{ij}[u,v]^i &= \sum_{s,i} \left(p_{ij} u^s \frac{\partial v^i}{\partial x^s} - p_{ij} v^s \frac{\partial u^i}{\partial x^s} \right) = -\sum_{s,i} (v^i u^s - u^i v^s) \frac{\partial p_{ij}}{\partial x^s} \\ &= \sum_{s,i} (v^i u^s - u^i v^s)(\Gamma_{is,j} + \Gamma_{js,i}) = 0\end{aligned}$$

We observe that, by the rank-nullity theorem [10], $\dim(\ker(p)) = \dim T_x M - rank(p)$. Since p is a degenerate metric, $rank(p)$ is not maximal. Let $k < n$ be the rank of p. Then $\dim(\mathcal{R}) = n-k$. By the implicit function theorem [11], there exist coordinates $(x^1, \cdots, x^k, y^1, \cdots y^{n-k})$ such that the distribution is spanned by $\frac{\partial}{\partial y^1}, \cdots, \frac{\partial}{\partial y^{n-k}}$. In these coordinates, the metric has the form

$$p = \sum_{ij}^{k} p_{ij} dx^i dx^j \tag{4.2.13}$$

By 4.2.12, we have that the components p_{ij} are independent of y-coordinates, indeed the vector fields $\frac{\partial}{\partial y^i} \in \mathcal{R}$ and therefore their flows preserve g. We can then view, by 4.2.13, g as a metric on a k-dimensional manifold with local coordinates system $x^1, \cdots, x^k$. Equation 4.2.3 implies that $\left(\Gamma^i_{jm}\right)_{i,j,m=1,\ldots,k}$ are coefficients of the Levi Civita connection of this metric (of dimension k). Without loss of generality, because of the freedom 4.2.6, we may assume that all Γ^i_{jm} with $i > j$ are equal to zero. Then, the formula for the components

$\mathcal{R}_{ij\ell m}$ of the curvature tensor (with lower indexes) of this k-dimensional metric coincides with 4.2.9 for $i, j, \ell, m < k$. Then, the problem is reduced to the case when p is nondegenerate, which was already solved by Riemann (see e.g. [19]). □

As the following example shows, the condition 4.2.11 almost everywhere does not imply that the rank of p is constant.

Example 4.2.7. *We consider the function*

$$\phi : \mathbb{R}^2 \to \mathbb{R}, \quad \phi(x, y) = x^2 + y^2$$

and as p we take $(d\phi)^2$ Computing $d\phi$ we get that

$$d\phi = 2xdx + 2ydy,$$

then

$$p = (d\phi) \otimes (d\phi) = (2xdx + 2ydy) \otimes (2xdx + 2ydy)$$

so in coordinates:

$$p_{11} = 4x^2 \quad p_{22} = 4y^2 \quad p_{12} = p_{21} = 4xy$$

that is

$$p(x, y) = \begin{pmatrix} 4x^2 & 4xy \\ 4xy & 4y^2 \end{pmatrix}.$$

This is a rank 1 matrix for $(x, y) \neq (0, 0)$. At $(x, y) = (0, 0)$ the rank is 0, meaning that the rank is not constant.

The example can easily be generalized for any dimension and any rank. On the other hand, the existence of continuous functions Γ^i_{jk} satisfying 4.2.3 implies that the rank of p is constant.

Corollary 4.2.8. *Assume p admits flat coordinates. Consider the following system of PDEs:*

$$0 = \nabla_j u_i := \frac{\partial u_i}{\partial x^j} - \sum_s \Gamma^s_{ij} u_s \tag{4.2.14}$$

on the unknown functions $u_1(x), ..., u_n(x)$, where Γ^s_{ij} is a (smooth) solution of 4.2.3. Then, for every point $\hat{x}$ and for any initial data $(\hat{u}_1, \cdots, \hat{u}_n) \in \mathbb{R}^n$ such that for every $v \in \mathcal{R}(\hat{x})$ we have $\sum_s v^s \hat{u}_s = 0$, there exists a unique solution $u_1, \cdots, u_n$ of 4.2.14 with the initial conditions $u_i(\hat{x}) = \hat{u}_i$. This solution has the property $\sum_s v^s \hat{u}_s = 0$ at every x and for every $v \in \mathcal{R}(x)$. Furthermore, for any such a solution $u_1, \cdots, u_n$ the 1-form $\omega = u_1 dx^1 + \cdots + u_n dx^n$ is closed so there exists locally a function f such that $\frac{\partial f}{\partial x^i} = u_i$. Moreover, if a solution vanishes at one point, it vanishes at every point.

Proof. The equation 4.2.14 means that the 1-form $\omega = u_1 dx^1 + \cdots + u_n dx^n$ is parallel with respect to the connection $\nabla = \Gamma^i_{jk}$. In particular, the equation is invariant with respect to the coordinate changes. $\mathcal{R}$ is invariant under parallel transport because the metric-compatibility $\nabla p = 0$ implies that parallel transport preserves inner products. Since both the 1-form ω and the subspace $\mathcal{R}$ are invariant under parallel transport, if $\mathcal{R} \subseteq \ker(u_1 dx^1 + \cdots + u_n dx^n)$ at the point $\hat{x}$, then $\mathcal{R} \subseteq \ker(u_1 dx^1 + \cdots + u_n dx^n)$ at every point.
In the flat coordinates $x_1, \cdots, x_n$ such that $p = \sum_{s=1}^r \epsilon_i (dx^i)^2$ (with $\epsilon_i \in \{-1, 1\}$) the equation 4.2.14 reads $\frac{du_i}{dx^j} = 0$. Then, if the initial data satisfy $\sum_s v^s \hat{u}_s = 0$ then for any solution we have $u_{r+1} = \ = u_n = 0$ and first r functions $u_1, ..., u_r$ satisfy $\frac{\partial u_i}{\partial x^j} = 0$ which implies that they are arbitrary constants. □

4.2.2 Approximated solution for non flat coordinates

In this section, we address the case in which the metric p defined in (4.1.5) does not satisfy the flatness condition of Theorem 4.2.6. In such cases, the intrinsic coordinates cannot be made Euclidean, and we approximate the deformation by introducing a displacement field u. We call displacement a map $u = \Phi(x_1, x_2) - (x_1, x_2)$, where $(x_1, x_2) \in \mathbb{R}^2$. Then :

$$p = (\nabla\Phi)^t \nabla\Phi = (\nabla u + Id)^t (\nabla u + Id) \approx \nabla u^t + \nabla u, \tag{4.2.15}$$

if $|\nabla u|$ is sufficiently small and ∇u^t is the transpose.
Let $M \subset \mathbb{R}^3$ denote the spatio–temporal domain on which p is defined. We have the following

Proposition 4.2.9. *Assume that M is an open subset of class C^1. Then a solution u of equation* (4.2.15) *with Neumann boundary satisfies the PDE system:*

$$\begin{cases} \Delta u_1 = \partial_1 p_{11} + 2\partial_2 p_{12} - \partial_1 p_{22} + 2\partial_3 p_{13} - \partial_1 p_{33} & \text{in} \quad M \\ \Delta u_2 = \partial_2 p_{22} + 2\partial_1 p_{21} - \partial_2 p_{11} + 2\partial_3 p_{23} - \partial_2 p_{33} & \\ \Delta u_3 = \partial_3 p_{33} + 2\partial_2 p_{32} - \partial_3 p_{22} + 2\partial_1 p_{31} - \partial_3 p_{11} & \\ \frac{\partial}{\partial \vec{n}} u_1 = \frac{\partial}{\partial \vec{n}} u_2 = \frac{\partial}{\partial \vec{n}} u_3 = 0 & \text{in } \partial M \end{cases}$$

Proof.

$$p \approx \nabla u^t + \nabla u = \begin{pmatrix} \partial_1 u_1 & \frac{1}{2}\Big(\partial_1 u_2 + \partial_2 u_1\Big) & \frac{1}{2}\Big(\partial_1 u_3 + \partial_3 u_1\Big) \\ \frac{1}{2}\Big(\partial_1 u_2 + \partial_2 u_1\Big) & \partial_2 u_2 & \frac{1}{2}\Big(\partial_2 u_3 + \partial_3 u_2\Big) \\ \frac{1}{2}\Big(\partial_1 u_2 + \partial_2 u_1\Big) & \frac{1}{2}\Big(\partial_2 u_3 + \partial_3 u_2\Big) & \partial_3 u_3 \end{pmatrix}$$

Let's compute the laplacian of u_1:

$$\partial_1 u_1 = p_{11} \Rightarrow \partial_1^2 u_1 = \partial_1 p_{11}$$

$$\frac{1}{2}\Big(\partial_1 u_2 + \partial_2 u_1\Big) = p_{12} \Rightarrow \partial_2 u_1 = 2p_{12} - \partial_1 u_2 \Rightarrow \partial_2^2 u_1 = 2\partial_2 p_{12} - \partial_{12}^2 u_2$$

So, by the fact that $\partial_2 u_2 = p_{22}$, which implies $\partial^2_{12} u_2 = \partial_1 p_{22}$, we get

$$\partial^2_2 u_1 = 2\partial_2 p_{12} - \partial_1 p_{22}$$

Analogously

$$\partial^2_3 u_1 = 2\partial_3 p_{13} - \partial_1 p_{33}$$

Then

$$\Delta u_1 = \partial_1 p_{11} + 2\partial_2 p_{12} - \partial_1 p_{22} + 2\partial_3 p_{13} - \partial_1 p_{33}$$

$$\Delta u_2 = \partial_2 p_{22} + 2\partial_1 p_{21} - \partial_2 p_{11} + 2\partial_3 p_{23} - \partial_2 p_{33}$$

$$\Delta u_3 = \partial_3 p_{33} + 2\partial_2 p_{32} - \partial_3 p_{22} + 2\partial_1 p_{31} - \partial_3 p_{11}$$

□

This proposition shows that, in the case of non-flat coordinates, the displacement formulation provides a practical way to linearize the metric and obtain an explicit PDE system for the components of u. The system, together with Neumann boundary conditions, can then be used to compute approximate solutions that describe small deformations, allowing us to analyze the behavior of the model even when the assumptions of Theorem 4.2.6 are not satisfied.

Chapter 5

Numerical simulations and results

5.1 Static Ebbinghaus illusion

The first step was to implement and visualize 2D even Gabor filters, defined as the product of a cosine carrier and a Gaussian envelope. The expression used was

$$\psi(x, y) = \cos\left(\frac{2\pi k}{\sigma}(x\cos\theta + y\sin\theta)\right) \cdot \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$$

We visualized the resulting filters both as 3D surface plots and 2D contour plots.

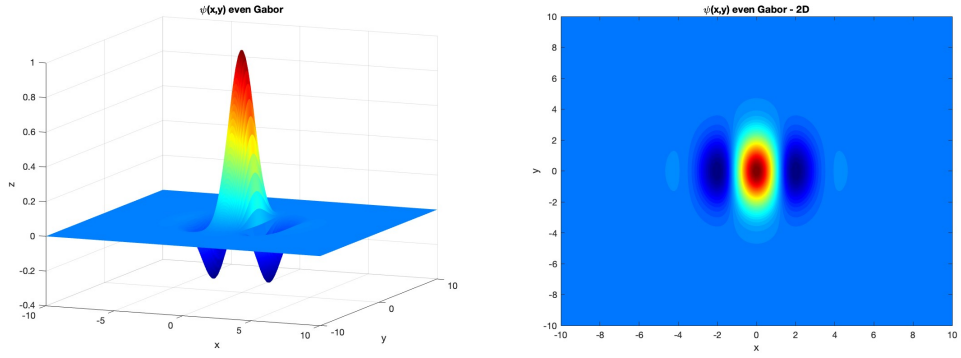

Figure 5.1: Even 2D Gabor filters

To analyze their behavior, the orientation θ and the scale σ were varied. This allowed to observe how the shape and spatial localization of the filter respond to changes in these parameters. In the following images, we see investigate how Gabor filters change as scale and orientation parameters vary.

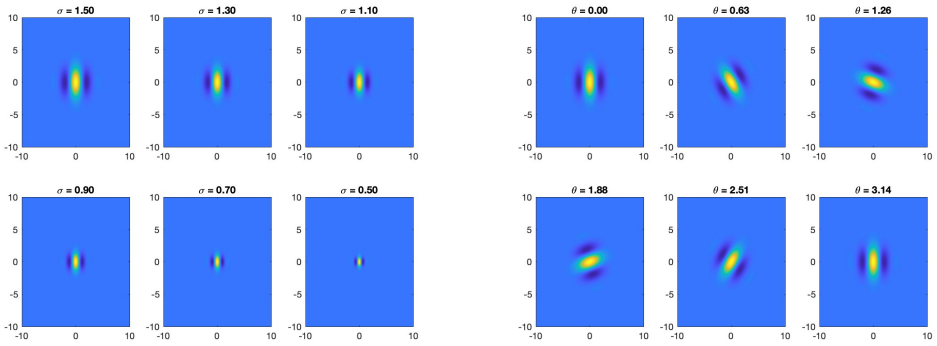

Figure 5.2: On the left: Gabor filters at different scales for a fixed orientation. On the right: different orientations for a fixed scale.

Then, a 3D scale-space representation was constructed using 2D even Gabor filters with a spatial scale σ, which varies continuously along the z axis, representing the filter's response across space. An isosurface at an arbitrary level is extracted and visualized, revealing how the Gabor filter changes as scale decreases.

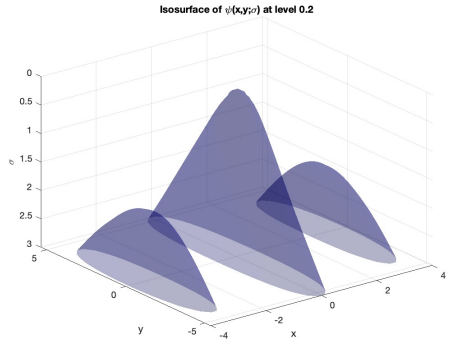

Figure 5.3: Isosurface of the Gabor filter at level $\psi = 0.2$

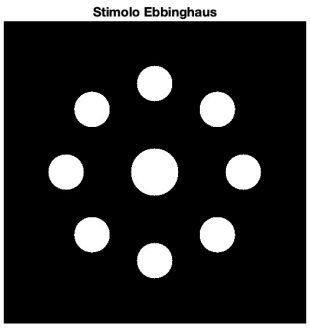

Figure 5.4: Visual stimulus: static Ebbinghaus illusion

Consequently, we created the image representing the classic Ebbinghaus illusion, with a central target circle surrounded by multiple smaller inducer circles arranged uniformly in a ring (Figure 5.4).

This image was processed by convolving it with a bank of 2D even Gabor filters, which vary across a range of spatial scales σ and orientations θ. The results of this operation were saved in a 4D array (position, scale, orientation) which contains the response of a Gabor filter at a specific scale and orientation.
We aimed to obtain a scale map by associating to every point of the image a scale value σ which maximizes the filter response. In order to do so, we selected a range of scale values adopting a logarithmic sampling of the form σ-values$= 0.7 * 2^{(0:0.3:4.3)}$, following the idea of a Gabor pyramid. The scale approximately doubles every two steps. With linearly spaced filters, the parameter of scale grew too slowly, causing the filters to be concentrated at small scales. Adding to this, we also normalized the Gabor filters to have a L1-norm equal to 1 and zero mean. Having L1-norm equal to 1 ensures consistent multi-scale comparisons and setting the mean of the filter to zero makes the filter insensitive to uniform regions in the image: the filter then responds only to local variations, such as edges.

We expected the scale map to be proportional to the distance from the boundary of the image stimulus. Reaching the correct tuning of the param-

eters we got the following result: the left panel shows the final scale map, while the right panel presents the same map with the contour of the original visual stimulus superimposed in white for reference.

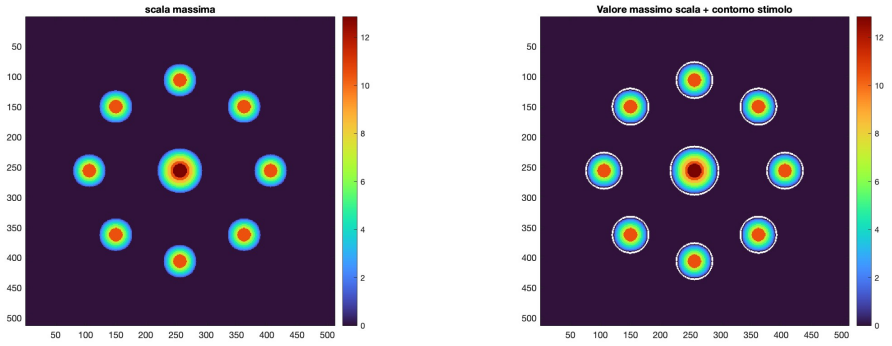

Figure 5.5: The maximum scale map.

We then inspected some of the Gabor filters to be sure that they were properly defined: in total there are 120 filters (number of sigma values times number of theta values), in Figure 5.6 we see 6 of them.

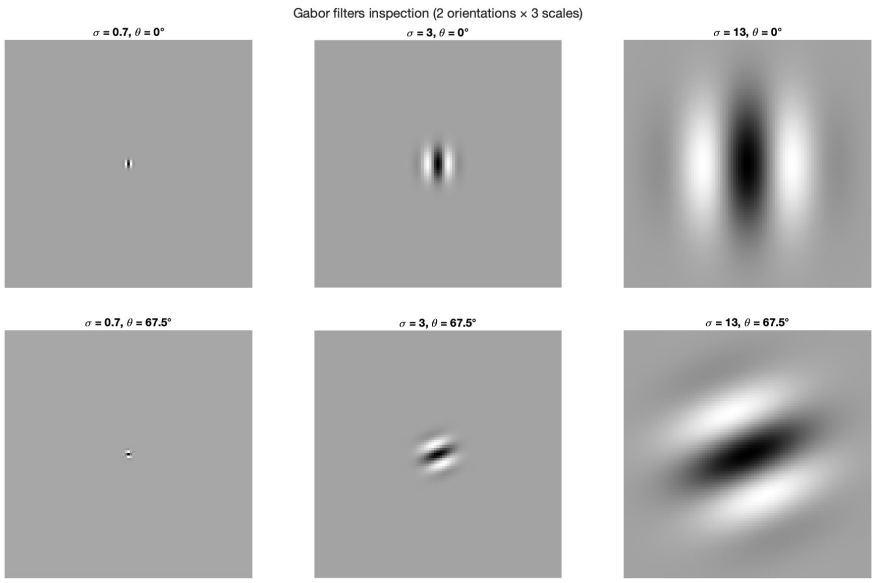

Figure 5.6: Gabor filter inspection for 2 chosen orientations and 3 chosen scales.

Figure 5.6 shows that the filters rotate correctly when θ varies and enlarge

when σ grows.

To check further, in Figure 5.7 we created a figure with three grayscale images displayed side by side, each showing the output of a Gabor filter response at different spatial scales and fixed orientation applied to the same frame of the stimulus. It served as a validation step to confirm that the filter bank behaves as expected across scales. The variation in responses demonstrates that the filters capture the visual stimulus at different resolutions.

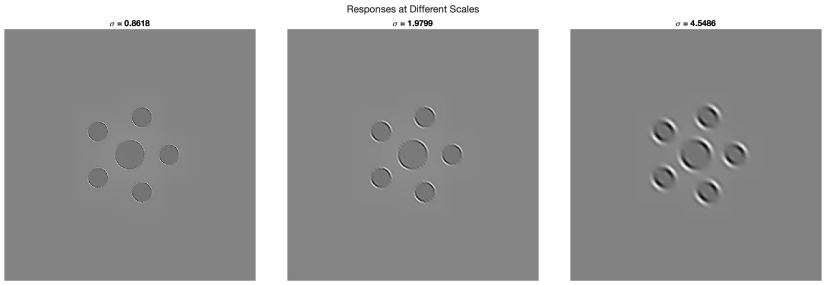

Figure 5.7: The output of a Gabor filter for different spatial scales

After obtaining a good map for the scale, we built a uniform scale map where each perceptual unit (i.e. the central target and the surrounding inducers) is assigned a constant scale value, which is the maximum scale value in each region. In Figure 5.8 we can see the specific case of the static Ebbinghaus illusion with 8 inducers being a bit smaller that the central target. Indeed we see that the colors in the uniform scale map are not distant in the colorbar. As an example, in Figure 5.9 we can see a simulation if the inducers were to be smaller: we see that the difference in size between the target and the inducers is bigger, which is coherent with the result obtained on the left panel of the figure.

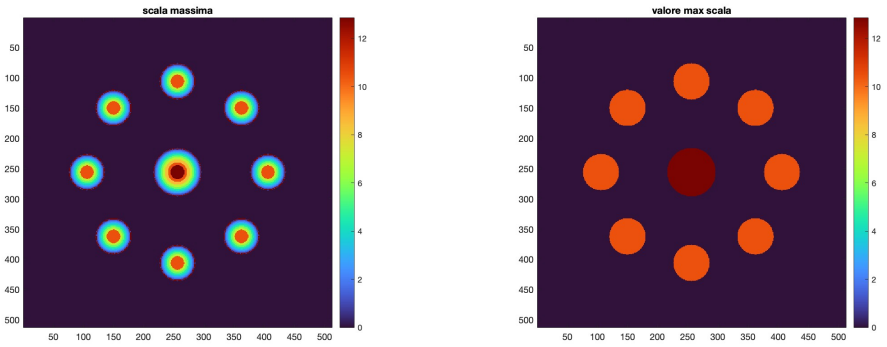

Figure 5.8: On the left, the scale map for 8 inducers, on the right the relative uniform map.

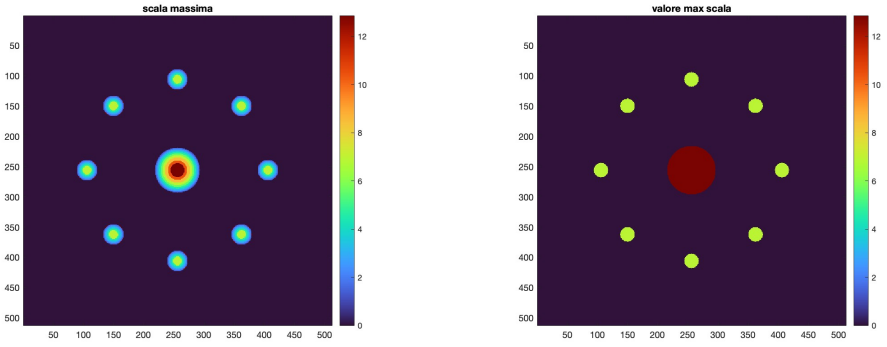

Figure 5.9: On the left, the scale map for 8 inducers smaller than before, on the right the relative uniform map.

From the uniform scale map, we extracted the displacement vector field that describes the local deformation induced by the illusion. In order to do this, we solved the PDE system 3.5.1 with Neumann boundary conditions. We computed the gradient of the uniform scale map using standard finite difference schemes: spatial derivatives were approximated via centered finite differences and then we employed the Jacobi iterative method, a classical technique for solving linear systems resulting from finite difference discretizations. The Jacobi method updates the solution at each grid point based on the previous iteration values of neighboring points.
We obtained a 2D vector field, which was then visualized using a quiver plot, restricted to the region corresponding to the central target circle. This allowed us to examine how the target appears to shrink or expand depending on the relative size of the surrounding inducers.
As we can see in Figure 5.10, the displacement vector field for inducers smaller than the central target is pointing outwards, meaning that the distortion consists in the target being perceived bigger than it actually is.

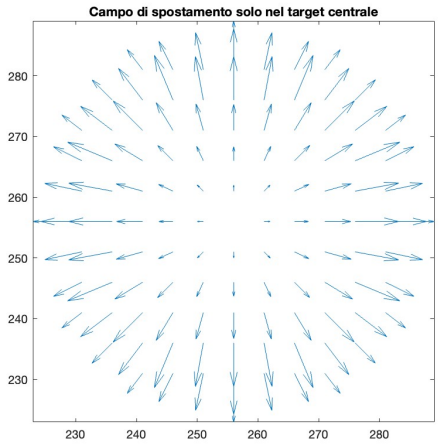

Figure 5.10: Outward pointing displacement vector field

The direction of the vectors reverses consistently with changes in inducer size, as one can see in Figure 5.11, which is the displacement field obtained in the case of inducers being bigger than the target.

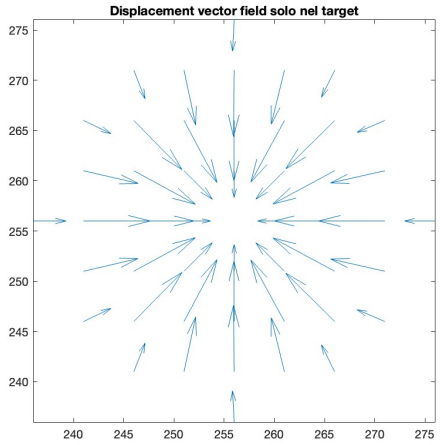

Figure 5.11: Inward pointing displacement vector field

5.2 Stationary dynamic Ebbinghaus illusion

Building on the previous results, we introduced time in the model. First we visualized a Gabor filter with the expression 4.1.2

$$\psi_{q,v,\theta}(x,y,t) = e^{-2\pi i \frac{k}{\sigma}((x-q_x)\cos\theta+(y-q_y)\sin\theta+v(t-q_t))} e^{-\frac{(x-q_x)^2}{2\sigma^2}-\frac{(y-q_y)^2}{2\sigma^2}-\frac{(t-q_t)^2}{2\sigma^2}}$$

In figures 5.12, 5.13, 5.14 we can see the evolution of the Gabor filters at time $t = 0, 1, 2$ with fixed velocity $v = 1$.

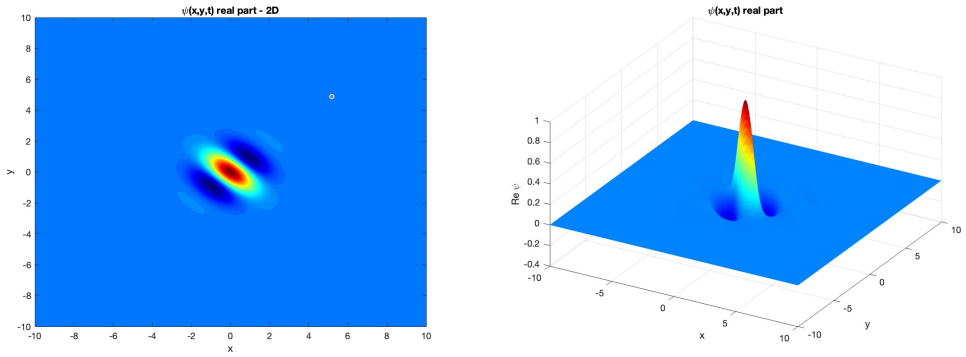

Figure 5.12: Gabor filter at time t=0.

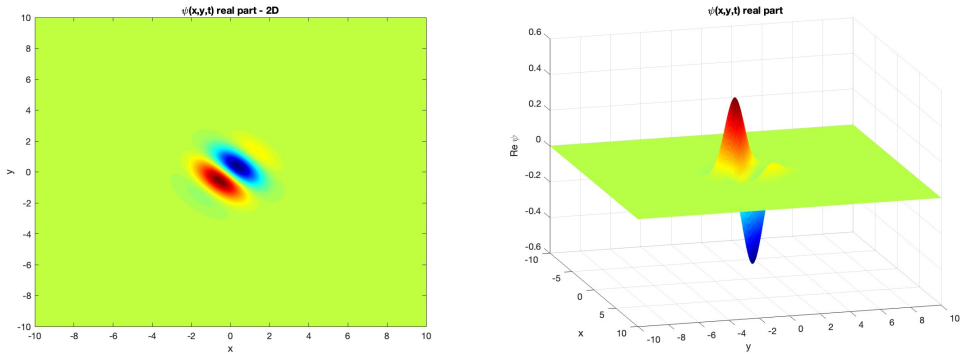

Figure 5.13: Gabor filter at time t=1.

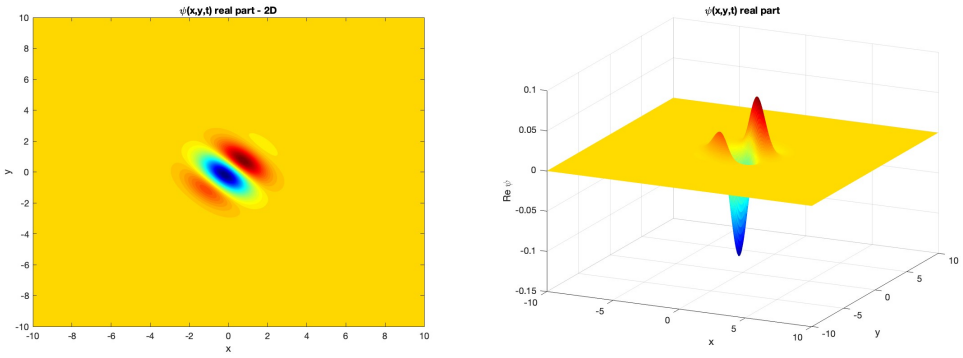

Figure 5.14: Gabor filter at time t=2.

As before, we started by building the image stimulus as a MATLAB generated video, following the approach of [14]. In figure 5.15 there are some of the frames of the dynamical stationary Ebbinghaus illusion.

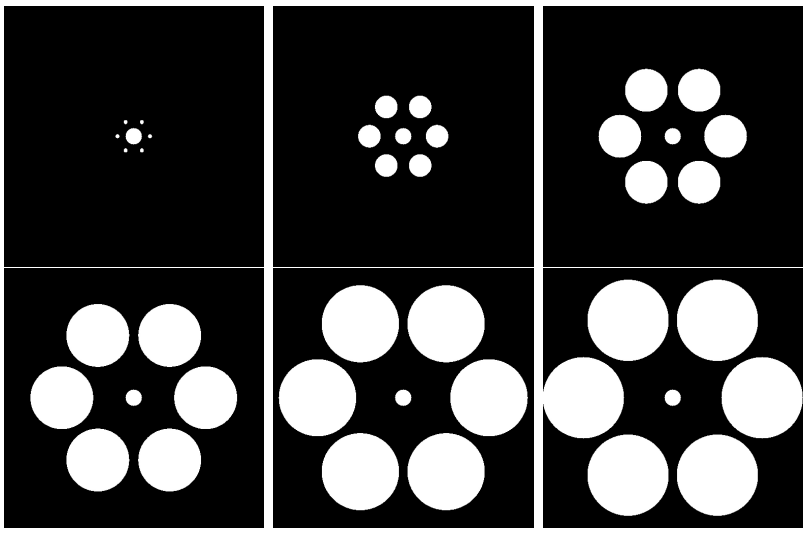

Figure 5.15: 6 frames of the generated video for the dynamical illusion.

We then created a bank of Gabor filters in order to process the video. In this case we employed a 3D bank of even Gabor filters that vary over spatial scale σ, orientation θ, and velocity v (i.e., temporal frequency given spatial frequency and speed). We then compute the responses through two different convolutions: first we convolve in space (2D) and then this result is convolved in time (1D).
Following the steps of the static case, we extracted the scale map through a soft-argmax on the result of the filtering. We inspected the scale map at 6 different times. In figure 5.16 we can see the results of the case of inducers growing in size.

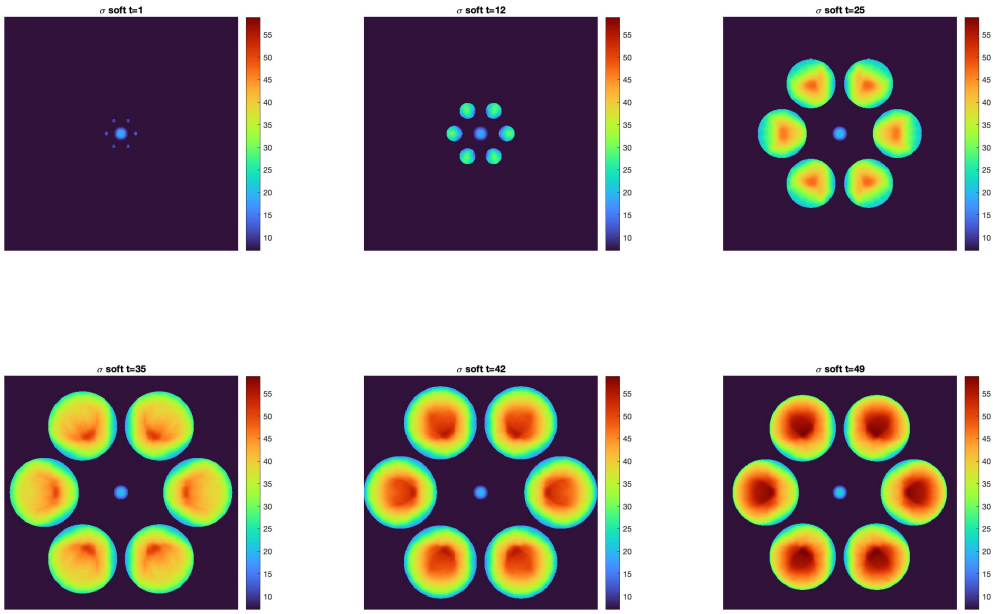

Figure 5.16: The sigma map at 6 different times in the case of growing inducers

We also explored the opposite case: we reversed the illusion in time, starting with big inducers which get smaller over time (Figure 5.17).

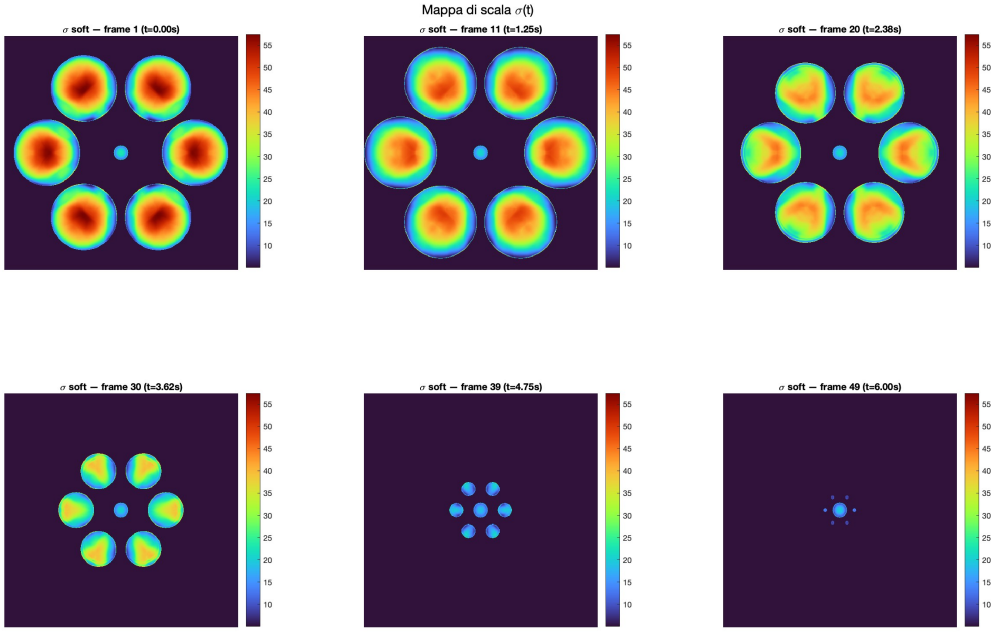

Figure 5.17: The sigma map at 6 different times in the case of shrinking inducers

We then proceeded to assign to each perceptual unit the maximum scale value (figure 5.18), in order to get the uniform scale map, which provides a compact representation of the perceived size at each time frame.

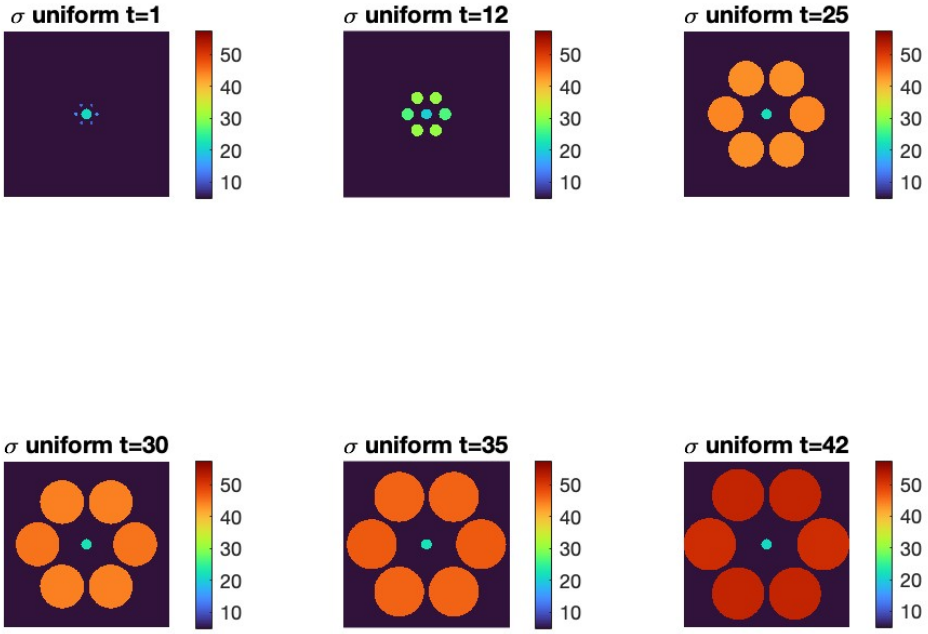

Figure 5.18: The uniform scale map at 6 different times, in the case of growing inducers.

From this map, we can extract a time-dependent displacement vector field $u(x, y, t)$ following the same procedure as in the static case, by solving the system in (3.5.1) for each frame. In Figure 5.19, two examples of the resulting displacement vector fields are shown.

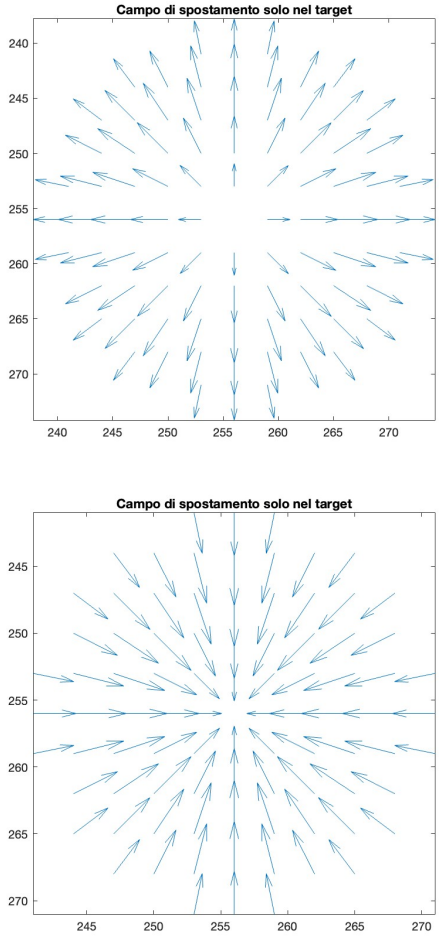

Figure 5.19: Above: normalized displacement vector field for an early frame of the stimulus, when the inducers are still smaller than the central target, producing an outward deformation. Below: vector field corresponding to a later frame, where the inducers are larger than the central target, resulting in an inward deformation.

In order to quantitatively compare the distortions represented by the displacement fields in the static and dynamic conditions, we adopted a common linear scaling of the vector magnitudes and explicitly set Autoscale to 'off' when plotting the vector fields with the "quiver" function. This ensures that the relative amplitudes of the vectors are preserved and directly comparable between the two configurations. We reproduced the same spatial configuration of the selected dynamic frame in the static condition and computed the corresponding static vector field, with the same procedure described in the previous section, using the same parameters. The comparison, reported in Figure 5.20, allows us to observe both the directional consistency of the deformation (outward or inward) and the relative intensity of the distortion, which increases when the inducers become larger than the target.

This result is consistent with the perceptual findings reported in [14], where the dynamic modulation of the inducers was shown to strongly amplify the perceived size distortion. In our model, this perceptual enhancement is reflected in the larger amplitudes of the displacement vectors, indicating a greater geometric deformation in the cortical representation as the inducers expand.

Temporal modulation plays a key role in amplifying contextual interactions within the visual system. Overall, the consistency between the model predictions and physiological evidence supports the hypothesis that perceptual distortions in dynamic visual illusions can be interpreted as geometric deformations emerging from a time-dependent cortical representation.

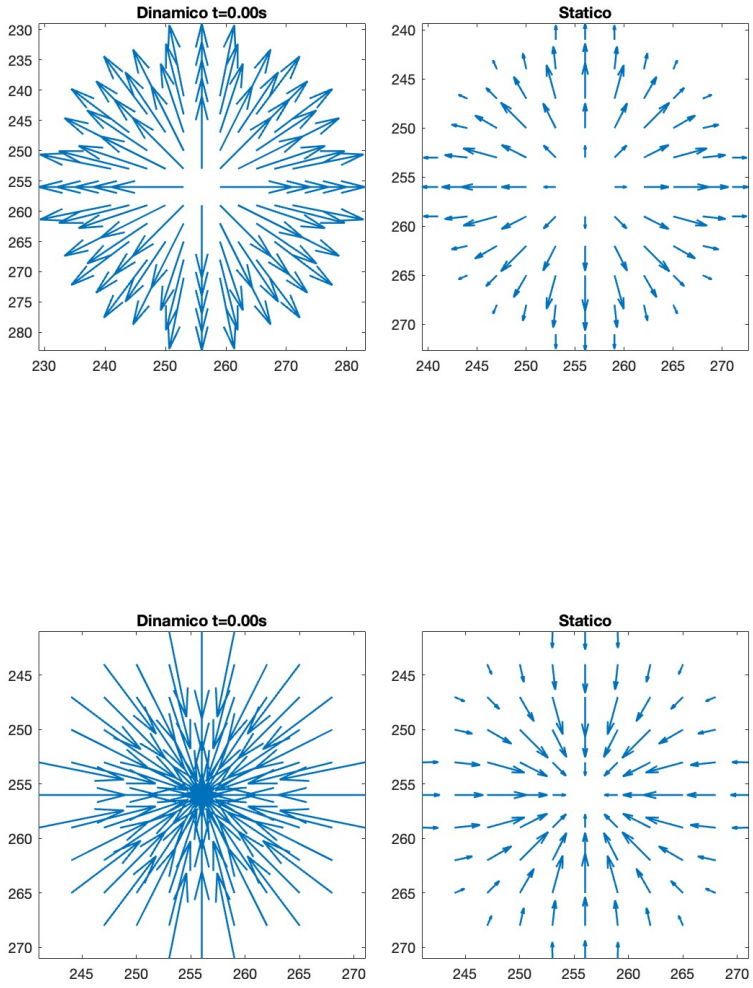

Figure 5.20: Displacement vector fields for the static (right) and dynamic (left) configurations of the Ebbinghaus illusion. Fields are computed using the same parameters and represented with a common linear scale (Autoscale off). The direction of the vectors indicates the local deformation of the perceived geometry, while their relative lengths reflect the magnitude of the distortion, which is notably stronger in the dynamic case.

Conclusions

In this thesis, we have developed a mathematical model of visual perceptual illusions that integrates spatial and temporal dynamics within a unified neurogeometric framework. Starting from sub-Riemannian models of the primary visual cortex (V1), we modeled the cortical processing of visual stimuli as a deformation of the underlying perceptual space induced by the activity of simple cells. Receptive profiles were described by Gabor functions, whose responses define, through the cortical metric, the geometry of perception.
In the static setting, we reproduced scale-type geometrical optical illusions such as the Ebbinghaus effect. The visual stimulus was convolved with a bank of 2D Gabor filters to obtain a scale map, from which we extracted the cortical deformation field by solving an elliptic PDE system. This allowed us to interpret the illusion as a metric deformation of the perceptual space, linking differential geometry and perception.
We then extended the model to include time-dependent dynamics. A new sub-Riemannian structure was introduced to describe the geometry of movement, adding temporal and velocity variables to the cortical space. Receptive profiles were modeled as 3D spatiotemporal Gabor functions, and the induced cortical metric was analyzed to describe dynamic distortions in perception. The resulting framework was applied to the Dynamic Ebbinghaus illusion, where the perceived size of a target changes over time according to the movement of surrounding inducers. Numerical simulations confirmed that the proposed model captures the qualitative behavior of the illusion. In particular the deformation in the time dependent model is stronger than in

the static one. The work presented here opens several directions for further research:

- The present model focuses on families of cells sensitive to a single feature (orientation, scale, or motion) at a time. A natural extension is to consider the interaction of multiple families of cells simultaneously. This would lead to a heterogeneous neurogeometric model, where coupled sub-Riemannian structures describe the joint activity of different cortical areas. The resulting system would couple the activity on distinct manifolds, leading to a richer description of perceptual integration.

- In the dynamic Ebbinghaus illusion, the growing or shrinking inducers also evoke depth perception, suggesting the involvement of binocular disparity processing. Extending the current model to a stereo-vision framework would introduce new variables encoding depth and disparity, expanding the feature space to a higher-dimensional manifold $\mathcal{M}_s$ that incorporates spatial, temporal, and binocular cues.

- From a computational standpoint, the structure of the model is inherently parallel, as each receptive profile acts independently on the input stimulus. This makes the model well suited for GPU-based implementations. Developing parallel architectures would enable real-time simulations and the study of large-scale cortical dynamics.

- Beyond its theoretical implications, understanding how the brain processes illusory stimuli can have direct applications in computational vision, robotics, and autonomous systems. Many current algorithms fail in the presence of illusion-like patterns or shadows. Incorporating neurogeometric mechanisms into artificial vision systems could improve their robustness to perceptual ambiguities.

In conclusion, this thesis provides a mathematical description of how time-dependent and spatial features combine to shape visual perception. The

framework developed here establishes a foundation for future models of dynamic vision, where motion, scale, and depth interact within a unified geometric representation of cortical activity.

Bibliography

[1] Bandyopadhyay, S., Dacorogna, B., Matveev, V. S., Troyanov M. (2023), Bernhard Riemann 1861 revisited: existence of flat coordinates for an arbitrary bilinear form, Mathematische Zeitschrift 305, pp. 1-25.

[2] Cocci G., Barbieri D., Citti G., Sarti A. (2015), Cortical spatiotemporal dimensionality reduction for visual grouping, Neural Computation 27 (6), pp. 1252-1293.

[3] do Carmo M. P. (1992), Riemannian Geometry, Birkhäuser, Boston.

[4] Daugman J. G. (1985), Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. J. Opt. Soc. Am. A 2, 1160–1169.

[5] Franceschiello B., Sarti A., Citti G. (2016), A Neuromathematical Model for Geometrical Optical Illusions. Journal of Mathematical Imaging and Vision 60 (1), 94-108.

[6] Franceschiello, B., Sarti, A., Citti, G. (2021), A Neuro-Mathematical Model for Size and Context Related Illusions, in Flash, T., Berthoz, A. (eds.) Lecture Notes in Morphogenesis. Springer, pp. 91-113.

[7] Gröchenig K. (2001), Foundations of Time-Frequency Analysis, Birkhäuser, Boston.

[8] Hering, H.E. (1861), Beitrage zur physiologie. 1–5. Leipzig, W. Engelmann.

[9] Hubel D. (1995), Eye, Brain, and Vision, Scientific American Library, No. 22. W. H. Freeman, 2nd edition.

[10] Lang, S. (1987), Linear Algebra. 3rd ed. New York, Springer.

[11] Lee, J.M. (2003), Introduction to Smooth Manifolds. New York, Springer.

[12] Lee T.S. (1996), Image Representation Using 2D Gabor Wavelets. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, No. 10.

[13] Meyer Y. (1992), Wavelet and Operators. New York: Cambridge Univ. Press.

[14] Mruczek R., Blair C., Strother L., Caplovitz G. (2015), The Dynamic Ebbinghaus: motion dynamics greatly enhance the classic contextual size illusion, Frontiers in Human Neuroscience 9, pp. 1-12.

[15] Oppel, J.J. (1855), Über geometrisch-optische Täuschungen, Jahresbericht des physikalischen Vereins zu Frankfurt am Main.

[16] Petitot, J. (2008), Neurogéométrie de la vision: modèles mathématiques et physiques des architectures fonctionnelles. Éditions École Polytechnique.

[17] Petitot J., Tondut Y. (1999), Vers une neurogéométrie. Fibrations corticales, structures de contact et contours subjectifs modaux, Mathématiques informatique et sciences humaines 145, pp. 5-101.

[18] Olman C., Van de Moortele P.-F., Schumacher J., Guy J., Ugurbil K., Yacoub E. (2010), Retinotopic mapping with spin echo BOLD at 7T. Magnetic Resonance Imaging, 28(9):1258–1269.

[19] Riemann, B. (2016), On the Hypotheses Which Lie at the Bases of Geometry. Edited and with Commentary by Jürgen Jost. Expanded En-

glish Translation of the German Original. Classic Texts in the Sciences. Birkhäuser–Springer.

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

[21] Sarti, A., Citti, G., Petitot, J. (2008), The symplectic structure of the primary visual cortex, Biological Cybernetics 98, pp. 33–48.

[22] Desolneux, A., Moisan, L., Morel, J.M., (2003), Computational gestalts and perception thresholds. Journal of Physiology-Paris, 97(2):311–324.

[23] Kanizsa, G., (1980) Grammatica del vedere: saggi su percezione e gestalt. Il mulino.

[24] Hoffman, W.C., (1971), Visual illusions of angle as an application of lie transformation groups. Siam Review, 13(2):169–184.

[25] Smith, D.A., (1978), A descriptive model for perception of optical illusions. Journal of Mathe- matical Psychology, 17(1):64–85.

[26] Ehm, W., Wackermann, J., (2012), Modeling geometric–optical illusions: A variational approach. Journal of Mathematical Psychology, 56(6):404–416.

[27] Helmholtz, H., Southall, J.P., (2005), Treatise on physiological optics, volume 3. Courier Corporation.

[28] Knill, C.D., Richards, W., (1996), Perception as Bayesian inference. Cambridge University Press.

[29] Weiss, Y, Simoncelli, E.P., Adelson, E.H, (2002), Motion illusions as optimal percepts. Nature neuroscience, 5(6):598–604.

[30] Fermüller, C., Malm, H., (2004), Uncertainty in visual processes predicts geometrical optical illusions. Vision research, 44(7):727–749.

[31] Walker, E.H, (1973), A mathematical theory of optical illusions and figural aftereffects. Perception & Psychophysics, 13(3):467–486.