Chapter 4
Functional Architecture I: The Pinwheels of $V1$

We have just seen in the previous chapter how the mechanisms used by visual neurons to filter the optical signal can implement localized differential data. But this is still insufficient; the visual cortex can enable the perception of forms only because it can pass from local to global and *integrate* local differential data into global geometric forms. This extraordinary performance has intrigued all vision specialists, even well before the famous psychophysical experiments of the Gestalt theory at the beginning of the last century.

In this chapter and the next, we will explore, among other things, the neural infrastructures that underlie the simplest of the Gestalt principles, that of so-called good continuation, which is associated with the ideal geometric notion of a line. To do so, we must introduce the *functional architecture* of the visual areas. It is indeed the very particularity and global coherence of this architecture that create the geometry we want to model. We should be clear that the remarkable geometric properties of perception result from very specific properties of the generating neural dynamics and that this specificity requires equally specific functional architectures. Architectures with connections that are either too local or global, either too isotropic or random, *could not* create such geometric structures.

As we have said, we will give particular importance to $V1$, the first of the primary visual areas. Such a restriction might seem too drastic, insofar as the later areas like $V2$ or $V4$ have many 'top-down' feedback connections to $V1$. But $V1$ is already very important, more important than one might think. Let us reiterate that we are adopting Mumford and Lee's high-resolution buffer hypothesis (Lee et al. [1]), according to which $V1$ is not a simple 'bottom-up early module', but participates in all visual processes requiring fine resolution, whence its functional architecture is essential for the totality of the visual system. As William Beaudot and Kathy Mullen argue in [2, p. 688]:

> All higher aspects of form perception rely on this early orientation-selective processing stage.

J. Petitot, *Elements of Neurogeometry*, Lecture Notes in Morphogenesis,
DOI 10.1007/978-3-319-65591-8_4

The literature in this area is immense. We therefore run into an obvious problem of presentation. We have tried to resolve this problem by presenting, although certainly in too brief a manner, enough data to give the reader some idea of the variety, the richness, and the difficulty of the questions. This research requires innumerable experimental feats, so we must be aware of the fact that, despite the mass of remarkable experimental results already available, all of the data presented here is 'work in progress', subject to debate.

The functional architecture includes two main components, that is two main classes of connections. The first is the class of retino-geniculo-cortical 'vertical' connections, which will be studied in this chapter. The second is the class of cortico-cortical 'horizontal' connections, which will be discussed in the next chapter.

4.1 The Areas of the Visual Cortex

In Sect. 3.1 of Chap. 3, we very briefly mentioned the visual pathways and areas. Here, we give a few additional clarifications. In humans, there are about fifty such areas in the cortex, totalling an area of about 2 500 cm^2. The thickness of the cortex varies between 2 and 3 mm, which makes a volume of about 625 cm^3. Estimating the average number of neurons per mm^3 to be from 20 000–50 000 yields on the order of 10 to 30 billion neurons and 60 to 240 $\times$ 10^{12} synapses, although these estimates vary slightly among specialists.

We have seen in Sect. 3.1 of Chap. 3 (see Fig. 3.1) the general structure of the retino-geniculo-cortical pathways. Area $V1$ corresponds to area 17 in the classification of Korbinian Brodmann, represented in Fig. 4.1. A large part is located in the calcarine fissure, with the retina's 'fovea $\rightarrow$ periphery' gradient corresponding to the 'occipital pole $\rightarrow$ anterior fissure' gradient with a magnification of the fovea (see Fig. 4.12 below). It includes about 100 million neurons (recall that the optical nerve contains about a million fibres). Figures 4.2 and 4.3 show the localization of the visual areas around $V1$. Figure 4.4 shows a planar flattening of the structure of the visual areas.

Area $V2$ consists of a ventral part and a dorsal part in each hemisphere; as we shall see in Sect. 5.8 of Chap. 5, it plays an important role in the detection of illusory contours and the determination of an edge of a figure in front of a background. It is the same for $V3$. $V2$ projects onto the dorsal $V3$, which in turn projects onto the parietal cortex. The ventral $V3$ projects onto the inferotemporal (IT) cortex. Area $V4$, the third ventral area after $V2$ and $V3$, receives direct projections from $V1$ and $V2$. As we shall see, it is important for colour processing and recognition of simple forms. Area $V5$, also called MT for 'medio-temporal', is sensitive not only to orientations but also to directions, and it is essential for detecting local movement (e.g. the motion of points) as well as for controlling eye movements. When $V5$ is damaged, vision is altered by giving the subject the impression of perceiving static images. It is equally sensitive to binocular disparity and therefore plays a part in stereopsis. It receives

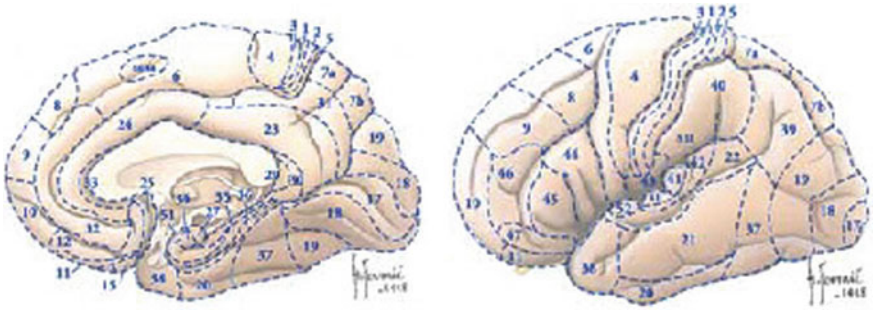


Fig. 4.1 The Brodmann areas (1909). •*1, 2, 3* Primary sensory cortex. Postcentral gyrus. *Sensitivity*. •*4* Primary motor cortex. Precentral gyrus. *Motricity*. •*5* Posterior parietal association area. Superior parietal gyrus. *Stereognosis*. •*6* Premotor cortex and supplementary motor area. Precentral gyrus and adjacent rostral cortex. *Programming of movements*. •*7* Posterior parietal association area. Superior parietal gyrus. *Visuomotor coordination, perception*. •*8* Frontal oculomotor field. Superior and middle frontal gyrus, internal face. *Saccades*. •*9–12* Prefrontal association cortex. Superior and middle frontal gyrus, internal face. *Cognitive areas, programming of movements*. •*13–16* Vegetative areas. Insular cortex. •*17* Primary visual area. Calcarine fissure. *Vision*. •*18* Secondary visual area. Around area *17*. •*19* Tertiary visual area. Around area *18*. •*20* Inferotemporal visual area. Inferior temporal gyrus. *Recognition of forms*. •*21* Inferotemporal visual area. Middle temporal gyrus. *Recognition of forms*. •*22* Association auditory area. Superior temporal gyrus. *Hearing*. •*23–27* Limbic association cortex. Subcallosal, cingulate, retrosplenial, parahippocampal cortex. *Emotion, memory*. •*28* Olfactory cortex, limbic association cortex. Parahippocampal gyrus. *Smell, emotions*. •*29–33* Limbic association cortex. Cingulate and retrosplenial gyrus. *Emotions*. •*34–36* Olfactory cortex, limbic association cortex. Parahippocampal gyrus. *Scents, emotions*. •*37* Parietal-temporal-occipital association cortex, middle temporal visual area. Middle and inferior temporal gyrus (T-O junction). *Perception, vision, reading, language*. •*38* Olfactory cortex, limbic association cortex. Temporal pole. *Scents, emotions*. •*39* T-P-O association cortex. Temporo-parieto-occipital junction (angular gyrus). *Perception, vision, reading, language*. •*40* T-P-O association cortex. Temporo-parieto-occipital junction (supramarginal gyrus). *Perception, vision, reading, language*. •*41–42* Primary auditory cortex. Heschl's gyrus and superior temporal gyrus. *Hearing*. •*43* Olfactory cortex. Insular cortex, frontal parietal operculum. •*44* Broca's area, lateral premotor cortex. Inferior frontal gyrus (frontal operculum). *Language, planning of movement*. •*45* Prefrontal association cortex. Inferior frontal gyrus. •*46* Prefrontal association cortex. Middle frontal gyrus. •*47* Inferior frontal gyrus. Taken from Hasboun [3]

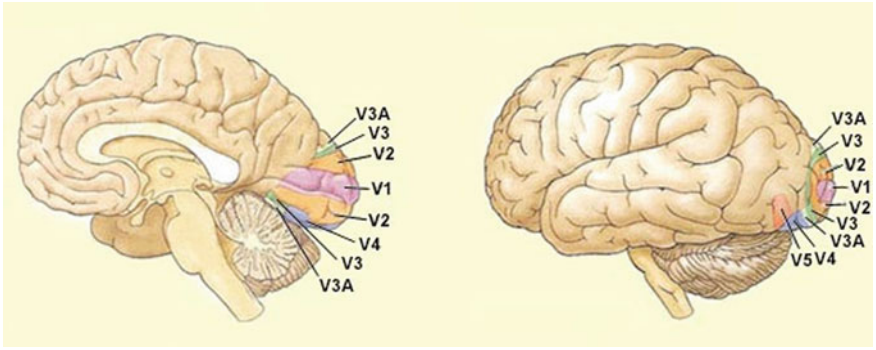


Fig. 4.2 Localization of visual areas around *V* 1. Taken from [4]

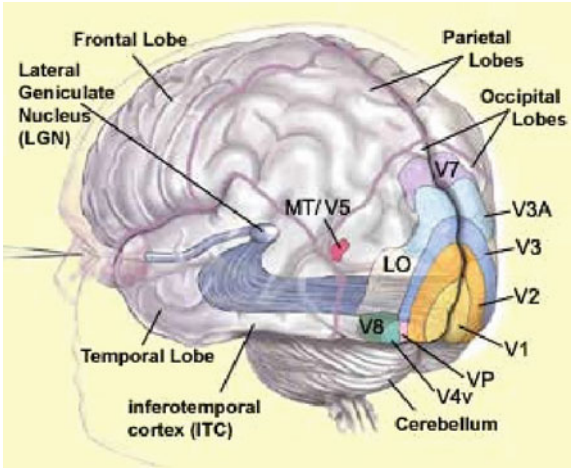


Fig. 4.3 Three-dimensional view of the localization of visual areas. Taken from [5]

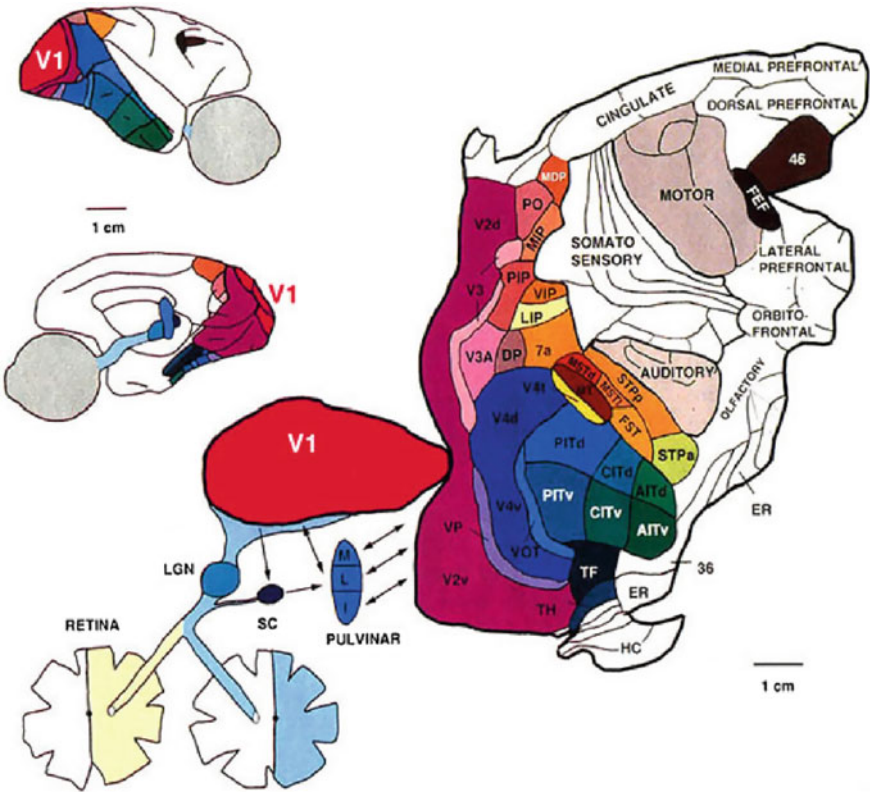


Fig. 4.4 Planar flattening of the structure of the visual areas. Taken from [6]

many connections from $V1$, but also from $V2$ and the dorsal $V3$. All of these areas possess numerous feedback connections coming back down towards the LGN.

The dorsal and ventral pathways that we have encountered are, respectively:

1. The magnocellular pathway dealing with spatial localization and movement (the 'Where' pathway of Ungerleider and Mishkin [7]), viz.

$$LGN\,(\text{magno}) \rightarrow V1\ (4C\alpha) \rightarrow V1\ (4B) \rightarrow V2\,(\text{thick stripes}) \rightarrow MT\ (V5)$$

   which leads, after the parietal cortex, all the way to the frontal cortex, where the frontal oculomotor field is located (among others). The latter directs the gaze.
2. The parvocellular pathway dealing with forms and colour (the 'What' pathway of Ungerleider and Mishkin [7]), viz.

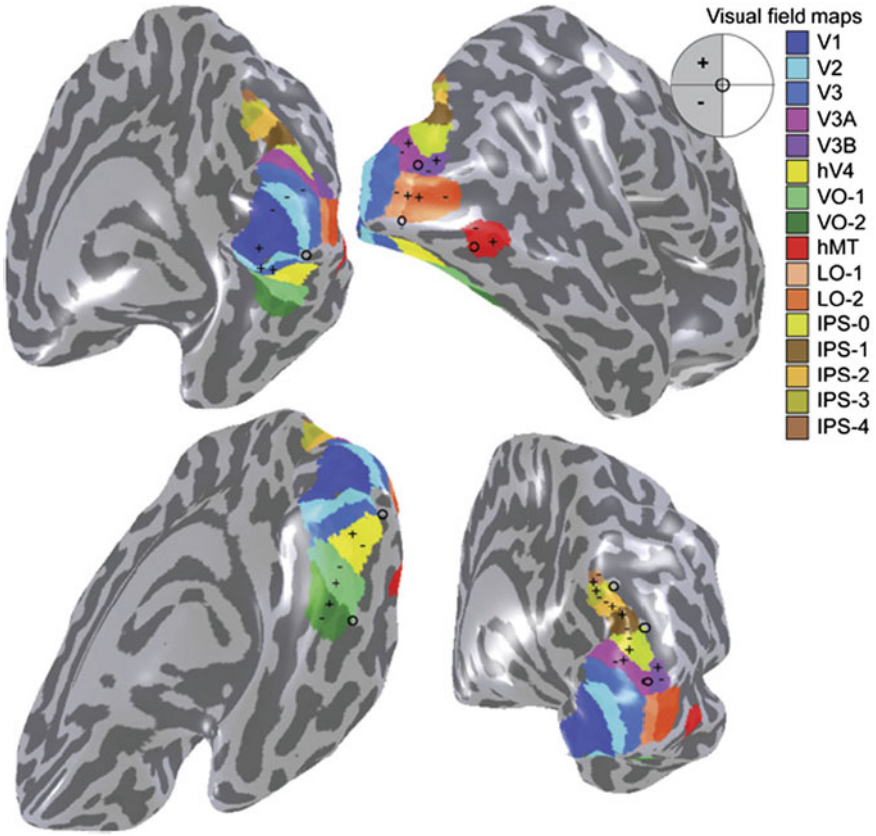


Fig. 4.5 Visual areas in humans (see text for acronyms). The prefix h stands for 'human', while ∘, +, − denote the fovea and the superior and inferior visual fields, respectively. Taken from [8]

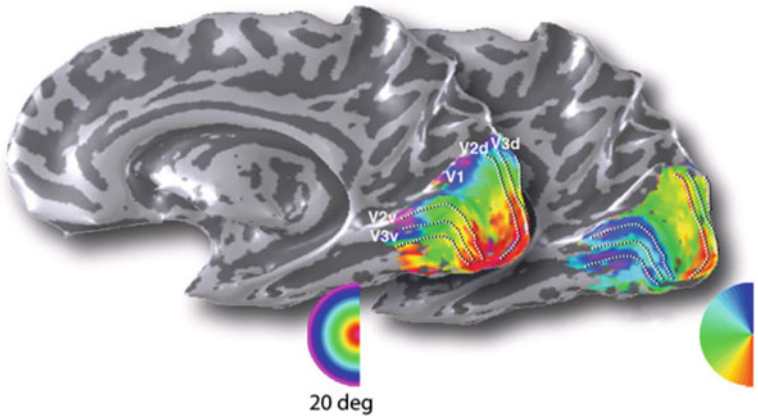


Fig. 4.6 fMRI of the retinotopic projection of a visual hemifield onto the corresponding visual hemisphere. The concentric circles of varying eccentricity and the rays of varying orientation on the half-disc target are colour-coded. Taken from [8]

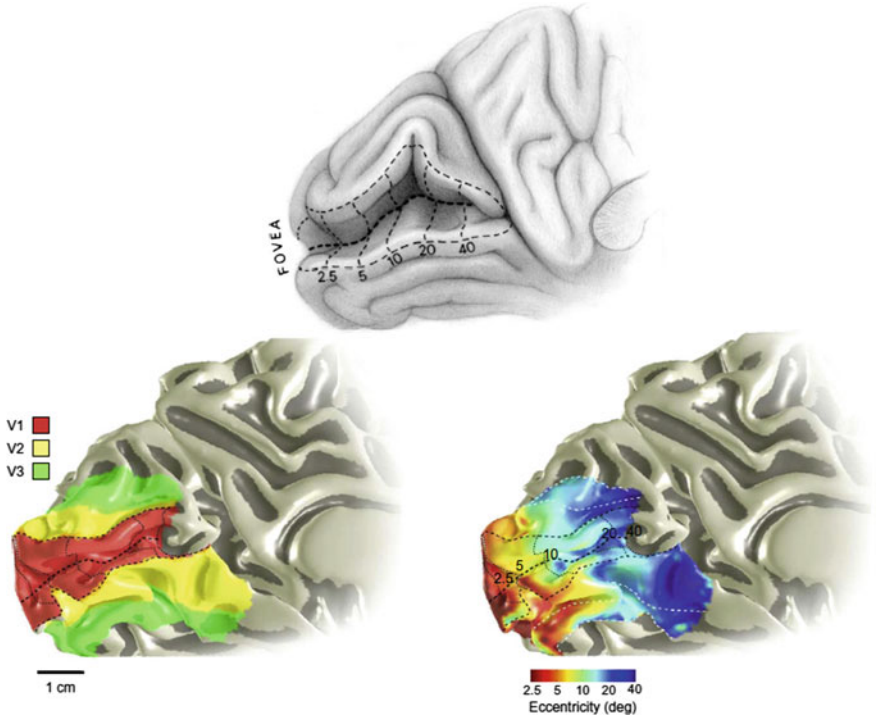


Fig. 4.7 Position of $V1$ (and $V2$, $V3$) in the cortex. The *upper edge* corresponds to the inferior vertical semi-meridian, and the *lower edge* to the superior vertical semi-meridian. The horizontal semi-meridian corresponding to the represented hemisphere is situated along the median fissure. Note that the scale of the eccentricities (radii of the circles of the retinal half-target) is logarithmic. Taken from [11]

$$LGN \text{ (parvo)} \rightarrow V1 \ (4C\beta) \rightarrow V1 \ (2/3 \text{ blobs})$$
$$\rightarrow V2 \text{ (thin stripes)} \rightarrow V3 \rightarrow V4 \text{ and } IT$$

With fMRI methods, already mentioned in Sect. 3.2.6.4 of Chap. 3 (phase-encoded retinotopy on travelling wave) which use checkered dynamic stimuli (expanding rings and rotating wedges) and record the induced cortical activity waves, we can get a better idea of what these areas do in a human brain. In Fig. 4.5, taken from the article [8] by Brian Wandell and Serge Dumoulin, we see the medial occipital areas ($V1$, $V2$, $V3$), lateral occipital areas (LO- 1 and LO-2, see [9], $hMT+$, corresponding in humans to the macaque's area $MT = V5$), the ventral occipital areas ($hV4$, VO-1, VO-2), the dorsal occipital areas ($V3A$, $V3B$), and the posterior occipital areas (from IPS-0 $= V7$ to IPS-4, $IPS =$ intraparietal sulcus).

Figure 4.6 is an fMRI version of the classic Fig. 4.12 (see below) first obtained by Tootell et al. [10] in the macaque using lethal methods. More detail is shown in Fig. 4.7, which is a more accurate representation of the position of $V1$ and its boundaries, showing how the two hemispherical $V1$ areas are glued along the vertical meridian by the callosal connections.

The definition of the visual areas is often a delicate matter. To give just one example, it remains an open debate as to whether, in humans, an area $V8$ should be

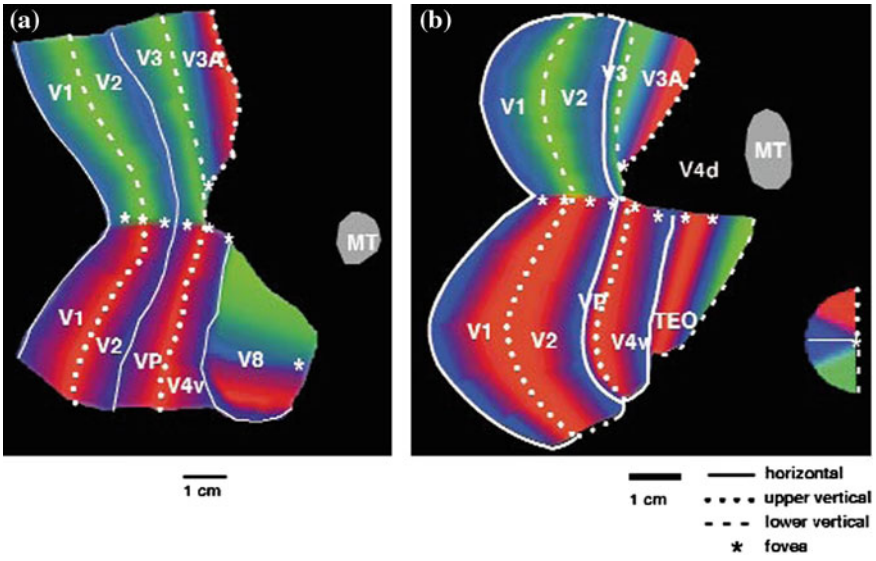


Fig. 4.8 Comparison of the polar angle retinotopy in humans (**a**) and in macaques (**b**). The stimuli are radial gratings rotating at low spatial frequency (the colour code of the angles is given *bottom right*). The boundaries of visual areas $V1$ to $V4$ can thus be established, and the mechanism for representing angles. The orthogonality of the angular distributions between area $V8$ in humans and area TEO in macaques is clearly visible. $V4v$ is the ventral $V4$ and $V4d$ is the dorsal $V4$. Taken from [13, Fig. 1]

introduced—possessing its own retinotopic representation—that would correspond to the area *TEO* (posterior inferotemporal cortex) in macaques (see [12]), or whether it would be more properly considered as a sub-area of $V4$, the problem being that, with respect to the retinotopy of the already well-identified $V4$, the retinotopy of this hypothetical area would lie at an angle of 90° (see Fig. 4.8 of [13]).

All these areas develop and stabilize with great plasticity through (i) critical periods, (ii) exposure to an enormous amount of stimuli, and (iii) the spontaneous cortical activity that induces neuronal waves (see, e.g., Hooks and Chen [14]).

4.2 Hypercolumnar Structure of the $V1$ Area

In $V1$, the density of neurons is on the order of $2 \times 10^5/\text{mm}^2$, while the receptive fields are on the order of a few degrees. Neurophysiological studies have made it possible to distinguish three types of structures in $V1$. These are layered, retinotopic, and (hyper)columnar, respectively.

Layered Structure

This is about 1.8 mm thick and composed of 6 'horizontal' layers, i.e., parallel to the surface of the cortex. The most important for us is layer 4 (traditionally called 'granular') and more specifically the sub-layer $4C$. Most of the fibres coming from the lateral geniculate nucleus project onto the latter, with sub-layer $4C\alpha$ receiving in particular the magnocellular projections and sub-layer $4C\beta$ the parvocellular projections (see Figs. 4.9 and 4.10). Layer 4 relays these inputs towards the 'supragranular' layers 2 and 3 and also projects in the other direction, onto the 'infragranular' layers, connecting them to the thalamus (layer 6) and the spinal cord (layer 5). Figure 4.11 shows for the macaque a small part of this complex network of projections containing several loops (see also Sect. 5.10 of Chap. 5). This structure does depend on the species, but we shall not go into the details here, apart from a brief note in Sect. 4.9.4, despite the importance of this observation.

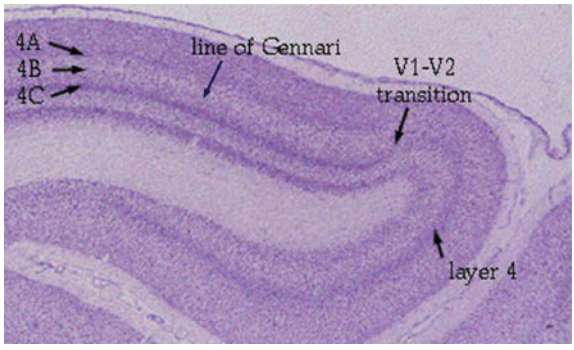


Fig. 4.9 Layered structure of $V1$ and layer 4. Taken from [6]

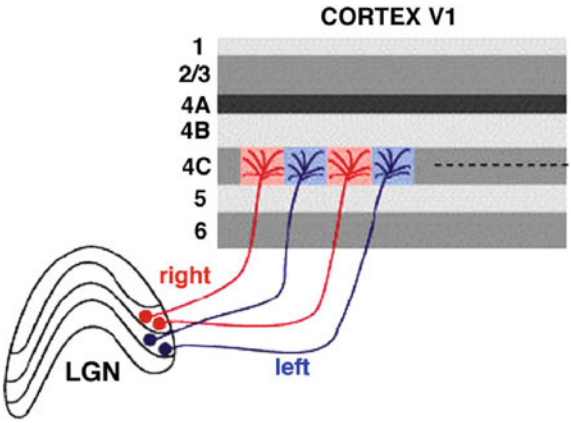


Fig. 4.10 Projection of the lateral geniculate nucleus onto layer $4C$ of $V1$. Taken from [6]

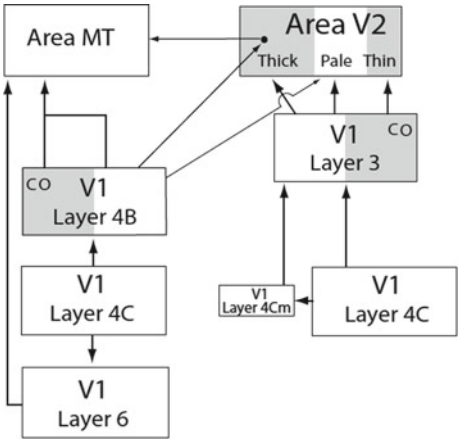


Fig. 4.11 Part of the network of projections between the layers of $V1$ in the macaque. Taken from [15]

Retinotopic Structure

Retinotopy refers to the fact that the projections (in the neurophysiological sense) of the retina onto the cortical layers are in fact mappings (in the mathematical sense) preserving the retinal topography. A typical example is the logarithmic conformal mapping between the retina and the $4C$ sub-layer of layer 4 where, as we have just seen, it is mainly fibres from the lateral geniculate nucleus that project. If R is the plane of the retina and M the cortical layer, the retinotopy is described by a mapping $\chi : R \to M$ which is a non-isometric isomorphism for a certain level of geometric structure that is not as rigid as the metric level (see Fig. 4.12).[1]

There are several models for this retinotopic mapping. The first was a monopole model $\mathrm{Log}(z + a)$, where z is a complex variable varying in R, but better is a dipole model like

[1] When we don't need to distinguish between R and M, we shall set $R = M$ and $\chi = Id$.

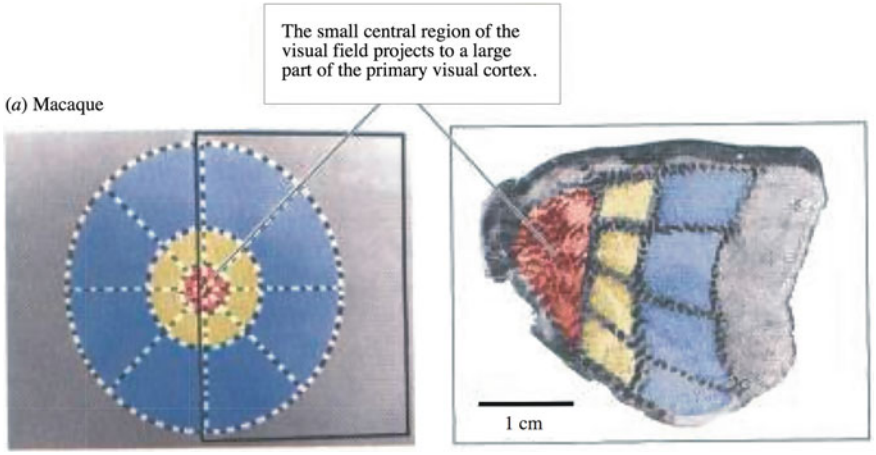
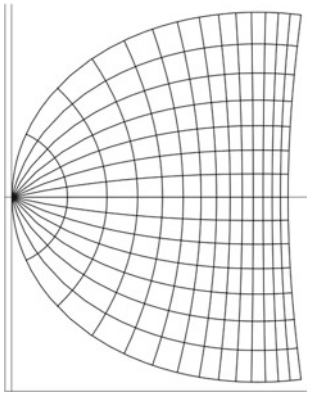


Fig. 4.12 Retinotopic projection of the retina onto layer $4C$ of area $V1$. The retinal hemitarget is transformed by a diffeomorphism which is not an isometry but a conformal map. From [10]. See also [16]

Fig. 4.13 Mathematical model of the retinotopic projection of the retina onto $V1$ using a conformal map that is a logarithm of a homography



$$\mathrm{Log}\frac{z+0.333}{z+6.66}\ .$$

This agrees quite well with empirical data [17], as shown in Fig. 4.13. A still better model is a wedge-dipole model common to the three areas $V1$, $V2$, and $V3$, such as

$$\mathrm{Log}\frac{w(z)+a}{w(z)+b}\ ,$$

illustrated in Fig. 4.14.

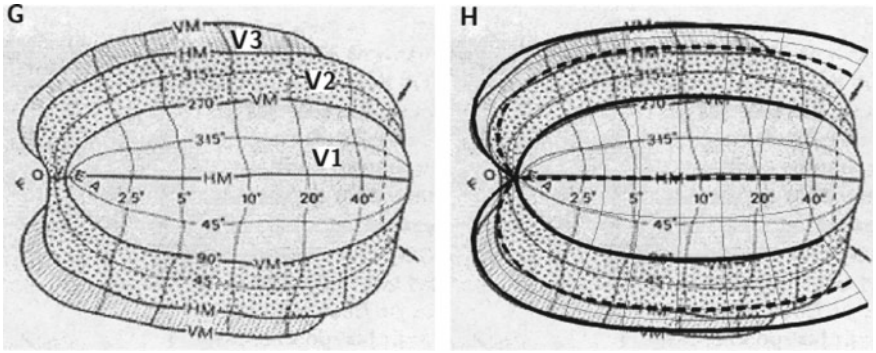


Fig. 4.14 Model of the form $\mathrm{Log}\big[(w(z)+a)/(w(z)+b)\big]$ for $V1$, $V2$, and $V3$. *Left* Areas $V1$, $V2$, and $V3$. *Right* Fit with the model. From Balasubramanian et al. [18]

Column and Hypercolumn Structure

This was the great discovery of Nobel prizewinners David Hubel and Torsten Wiesel in the early 1960s, following up work by their PhD supervisor Stephen Kuffler on retinal GCs (see Sect. 3.2.6.1). This was preceded in the late 1950s by the work of Vernon Mountcastle on the somatosensory cortex of the cat, and after Hubel and Wiesel, it was also found in the motor cortex and the auditory cortex. As we saw in Sect. 3.2.6 of Chap. 3, there are 'simple neurons' in the $V1$ area (as opposed to the 'complex' and 'hypercomplex neurons' discussed in Sect. 3.2.6.3) which are sensitive to orientation, ocular dominance, and colour. These are the ones of interest to us here.

We have already considered the structure of their receptive fields and receptive profiles. We shall focus particularly on those having the form of a second derivative of a Gaussian. If to begin with we simplify the situation as far as possible by not taking into account either the scale (the resolution and the spatial frequency) or the phase, we may say that these neurons detect pairs (a, p) of retinal positions a and orientations p at a. In Sect. 2.7 of the *Introduction*, these were referred to as mesoscopic *contact elements* of the visual plane. Indeed, as we said in Sect. 3.3.4 of Chap. 3, if we consider an edge crossing the receptive field, the response of the neuron will be maximal when the edge is aligned with the preferred orientation.

Using the methods to record responses to appropriate stimuli discussed in the last chapter (oriented bars crossing the RF of the neurons, etc.), it can be shown that, perpendicular to the surface of the cortex, the retinal position a and the preferred orientation p remain roughly constant. This 'vertical' redundancy—which specifies a population coding of the position—defines *orientation columns* of about 20 μm. As shown by DeAngelis [19], it is the *phase variation* that dominates in the columns:

> Spatial phase is the single parameter that accounts for most of the difference between receptive fields of nearby neurons.

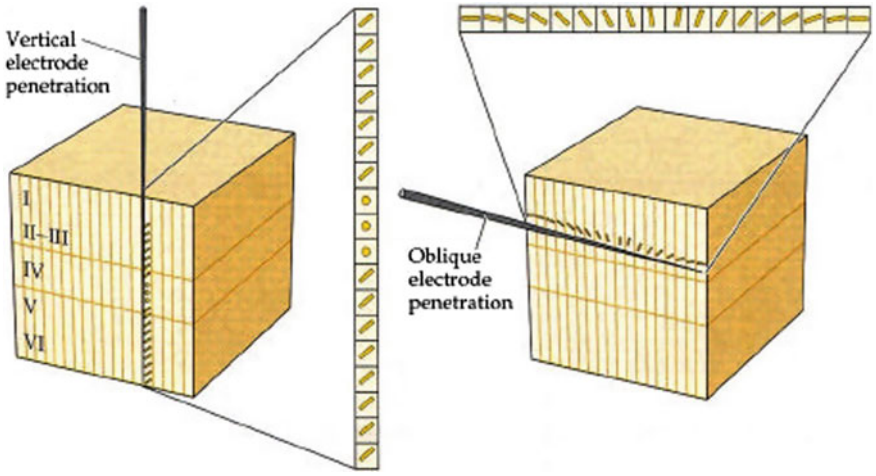


Fig. 4.15 Classic experiments by Hubel and Wiesel which led to the discovery of the orientation hypercolumns in $V1$. Taken from [24]

Moreover (see Sect. 3.3.5 of Chap. 3), the population coding gives the system a higher resolution than single neurons. The reader is referred to Beaudot and Mullen [2, 20], who suggest a mechanism to explain this, and also Snippe and Koenderink [21], and Ringach [22]. The latter studies the dependence of the resolution on the Fourier spectrum of the tuning curves of the neurons.

On the other hand, along lines parallel to the surface of the cortex, the preferred orientation p varies in steps of about 10°. To a first approximation, it can be considered to vary monotonically along line segments. In fact, this is not really the case, and for subtle topological reasons which we shall explain in Sect. 4.4.3 when we comment on Braitenberg's paper [23], the variation of p is not necessarily monotonic. However, we shall suppose here that it is. A 'horizontal' grouping of columns whose orientations vary over a range of π defines an *orientation hypercolumn* which is a broad neural micromodule measuring between 200 μm and 1 mm (see Figs. 4.15, 4.16 and 4.17).

The idea of a column just discussed clearly needs to be fleshed out, and a great deal of work has been devoted to this, dealing with problems such as establishing the link between anatomical and functional definitions, investigating the genuinely columnar nature of different ways of processing the stimuli (i.e. asking to what extent they cross the layers of $V1$), analyzing the strengthening of the edges of the columns by lateral inhibition. Among the various reviews available on these matters, the reader is referred to the one by Lund et al. [25].

We shall discuss different species of mammals, such as the tree shrew (tupaia), cats, primates, and humans, where the concept of orientation column is of key importance. This is quite restrictive because, as we saw in Sect. 3.8 of Chap. 3, in other species of mammals, and in particular rodents like the rat, this idea is less relevant.

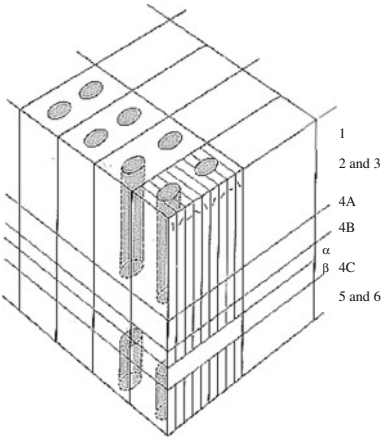
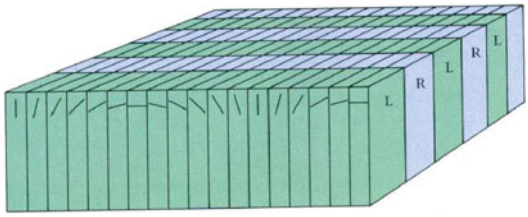


Fig. 4.16 Hypercolumn of $V1$ from Hubel and Wiesel

Fig. 4.17 Classic diagram of hypercolumns. From [6]



There we encounter other types of spatial processing, as carried out, for example, by the grid and place cells (see Sect. 3.8 of Chap. 3). There is orientation selectivity, but it is *dispersed* throughout $V1$; i.e., it is not brought together anatomically in columns and hypercolumns by a functional architecture.

However, even restricting to species for which this idea is relevant, there is nevertheless a certain diversity in the structure of $V1$. The general structure, concerning retinotopy, orientation selectivity, spatial frequency tuning, etc., remains the same, but the fine structure varies. As noted by Stephen Van Hooser:

> There is considerable diversity in the abundance of different cell classes, laminar organization, functional architecture, and functional connectivity. [26]

Figure 4.18 compares several different species: the macaque (primate, diurnal, frontally placed eyes), the cat (carnivorous, crepuscular, frontally placed eyes), the tree shrew or tupaia (Scandentia, diurnal, laterally placed eyes), and the gray squirrel (rodent, diurnal, laterally placed eyes).[2] We find that there is indeed orientation selectivity everywhere, but that it is not always present in all layers, as in the cat or the

[2] Van Hooser's paper also discusses the rat (rodent, nocturnal, laterally placed eyes), the night monkey, also known as the owl monkey or douroucoulis (New World primate, nocturnal, frontally placed eyes), the ferret, etc.

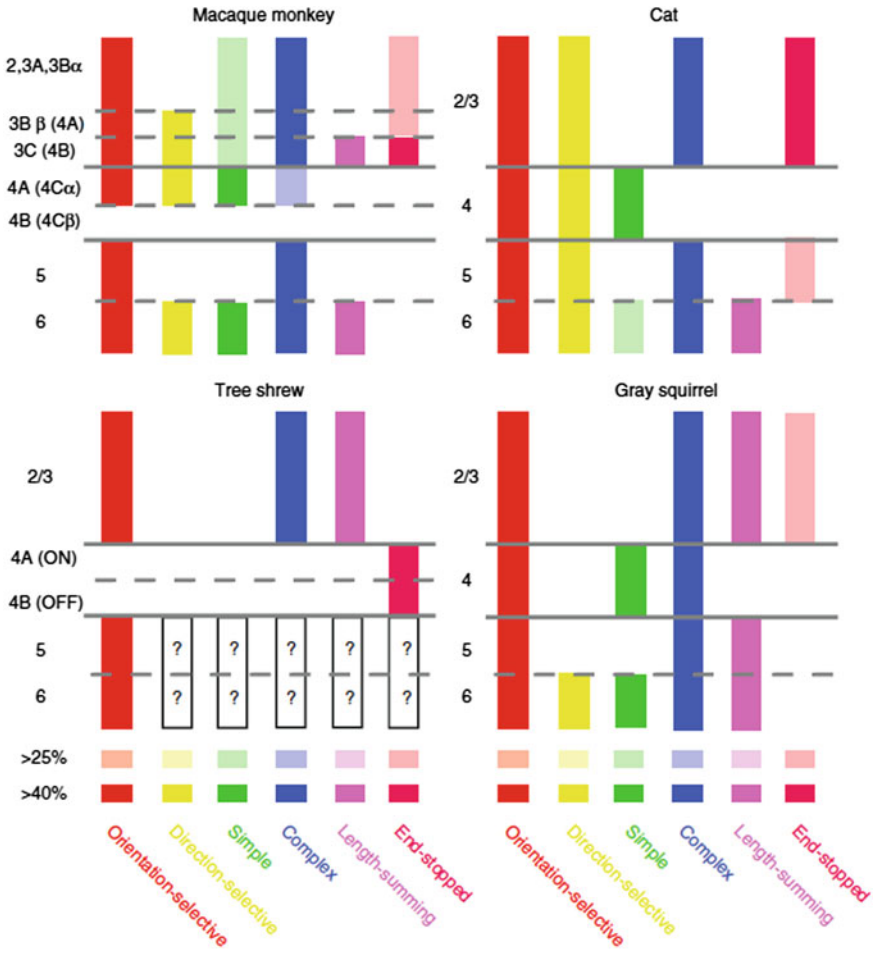


Fig. 4.18 Similarities and differences in a certain number of geometric features for different species of mammals. Taken from Van Hooser [26]

squirrel. Direction selectivity occurs throughout all layers in the cat (and also the ferret), but less in primates and scarcely at all in the tree shrew. The hypercomplex (end-stopped) cells occur in layers 2 and 3, but not in layer 4, except in the tree shrew, where they are in fact present only in layer 4. The cells responding most strongly when the bars on the gratings are longer (length-summing cells) are rare in carnivores, etc.

We thus see just how important the hypercolumn structure of $V1$ is and how crucial it is to find a suitable mathematical structure to model it. We shall show that, on the *mesoscopic* scale that is our own, the right concept, which is altogether natural, is the geometric concept of a *fibration*, so well known to mathematicians.

4.3 $V1$ as a Mesoscopic Fibration

4.3.1 'Bridging Scales': The Mesoscopic Level

We should stress, as we have already done in Sect. 2.7 of the *Introduction*, that the geometric structures we have introduced refer to the *mesoscopic* scale. Indeed, realistic 'micro' simulations of an orientation column are already unbelievably complex and have featured among some of the biggest computational projects in the world.

Launched in 2005 by Henry Markram at the École Polytechnique Fédérale de Lausanne (EPFL), the *Blue Brain Project* (BBP) aimed to simulate a rat cortical column containing some 10,000 neurons and 30 million synapses. It had access to a computational power (supplied by IBM) of 20 teraflops (one teraflop is 10^{12} operations per second). Since 2013, the *Human Brain Project* (rival of the American project *Brain*) has picked up where Blue Brain left off, using supercomputers of up to 10^6 teraflops. The HBP aims to simulate complete brain areas, and in particular, the visual cortex. Bringing together many universities and research institutes, it has been chosen as one of the two Future and Emerging Technologies (FET) Flagships of the European Union and will receive a total of a billion euros over 10 years. It explicitly exploits the possibilities of the information and communication technologies (ICT) and involves six platforms: neuroinformatics, brain simulation, high-performance computing, medical informatics, neuromorphic computing, and neurorobotics. Mathematical models do not play a major role because, as noted by those who designed the project, they are just 'toy models', drastically simplifying the biophysical data to make them amenable to mathematical analysis.

However, certain modelling problems have been identified, in particular regarding the four themes:

- bridging scales,
- synaptic plasticity, learning, and memory,
- large-scale models,
- principles of brain computation.

An internal debate has sprung up between massively computational approaches and more model-oriented and structural approaches. Our own model is an element in this debate.

4.3.2 Fibrations and Engrafted Variables

On the mesoscopic level, through the hypercolumn functional architecture of $V1$, a (discretized) copy of the space P of directions p in the plane is associated (in a retinotopic and anatomically observable way) with each retinal position $a \in R$. As a consequence, there is a *neural implementation* of the projection $\pi : R \times P \to R$

of the Cartesian product $R \times P$ onto its first factor R, a projection which is in fact a (trivial) fibration with base the retinal space R and fibre the manifold P.

In these mesoscopic models, a column is thus viewed as a simple contact element (a, p). This may look like an unacceptable reduction with regard to the computational projects just mentioned. However, we shall set out to show here that, despite the enormous reduction in complexity, the functional organization of these elements leads to a significant geometric complexity. This means that the complexity of a realistic 'micro' model of a functional architecture only becomes manageable by using highly elaborate multiscale bridges.

According to the materialist principle that 'a structure only exists if it is implemented', we must inquire into the neural materiality of the mesoscopic projection $\pi : R \times P \to R$ over and above its formal ideality. Now, if the projection π is mathematically (ideally) trivial, it is not at all so from the neurophysiological (material) point of view:

(i) The receptive fields of the ganglion cells and cortical neurons, or rather the cortical columns, are very small *local charts* which overlap and can be glued together on the overlap.
(ii) We shall see in Sect. 4.9.1 that it is no easy matter to test the direct product structure experimentally, i.e. the independence of the position and orientation variables. There, too, there are significant differences between species, as we saw above: for some species, it is a single layer of $V1$ which implements the product $R \times P$, while for others, several layers are involved.
(iii) The projection π (in the geometrical sense) is implemented by the fine circuitry of the retino-geniculo-cortical pathways, which project the retina onto $V1$ (in the neurophysiological sense).[3]

The geometric structure of the product of R as base space with a space of 'secondary' variables such as orientation, ocular dominance, direction of motion was well expressed by David Hubel when he spoke of 'engrafted variables':

> What the cortex does is map not just two but many variables on its two-dimensional surface. It does so by selecting as the basic parameters the two variables that specify the visual field coordinates (distance out and up or down from the fovea), and on this map it engrafts other variables, such as orientation and eye preference, by finer subdivisions. [27, p. 131]

A hypercolumn can thus be modelled as the Cartesian product of the RFs with the space of secondary variables which are 'grafted on', so to speak. The overlaps of the RFs are then interpreted geometrically as a *gluing together* of these local models, and the projection π is for its part implemented in the vertical connections:

- from the retina to the hypercolumns,
- within the hypercolumns themselves.

[3] In the geometric models of neural functional architectures, there are many problems of terminology. Lexical items such as 'fibre', 'projection', 'connection' are used in different ways by mathematicians and neurophysiologists. In general, the meaning should be clear from the context.

This is the idealized model of $V1$ proposed by Hubel, leading to the fundamental geometrical concept of a *fibration* or *fibre bundle*.

4.3.3 Fibre Bundles

The key idea of a fibration or fibre bundle was developed by mathematicians, and more recently by theoretical physicists, for deep reasons. The problem was to associate with each point in some base space M an entity of a certain type F, such as a scalar, a vector, a covector, a tensor, an exterior form, a direction, a phase, or a quantum number, which depends smoothly on this point. An obvious solution for modelling such a *field* on M would be to use maps $\varphi : M \to F$. But in many cases, we must take into account the fact that the *whole set* F of possible values of such maps φ is associated with each point of M. To take a concrete technical example, in each pixel of a computer screen, all grey levels (1 byte) or colours RGB (3 bytes) are represented.

Intuitively, a fibre bundle comprises a base space M (a differentiable manifold) and copies of a manifold F called the fibre 'above' each point of M. Globally, the space E of the fibre bundle, with the fibres glued together, is not necessarily a Cartesian product $M \times F$. It results from gluing together several Cartesian products $U_i \times F$ defined on local open domains U_i of M. Up to now, this *local triviality* has only been of interest to geometers in cases where the base space M is not a globally trivial space like $\mathbb{R}^n$ but rather a manifold that may not be simply connected and thus may have a non-trivial homology. In our case, the fibre bundles are globally trivial, but their local structure is imposed by neurophysiology (the receptive fields).

By definition, a fibre bundle is a 4-uple (E, M, F, π) such that:

1. E, M, and F are differentiable manifolds, called the total space, the base space, and the fibre of the fibre bundle, respectively.
2. $\pi : E \to M$ is a surjective differentiable map called the projection of the fibre bundle.
3. The inverse images $E_x = \pi^{-1}(x)$ $(x \in M)$ are isomorphic to F and $E_x \cong F$ is called the fibre at the point x (see Fig. 4.19).
4. For any point $x \in M$, there is a neighbourhood U of x such that $\pi^{-1}(U)$ is diffeomorphic to the Cartesian product $U \times F$ equipped with its canonical projection $\pi : U \times F \to U$, $(x, q) \longmapsto x$ (local triviality, see Fig. 4.20).

A *section* of a fibre bundle is a differentiable map that lifts the projection π, associating an element of the fibre E_x with each point x of the base space M. If $s : M \to E$ is a section, we thus have $\pi \circ s = Id_M$ (see Fig. 4.21). Sections can be defined simply locally on open sets $U \subset M$. In the case of a globally trivial fibre bundle $\pi : E = M \times F \to M$, a section above U is nothing other than a map $s : U \to F$.

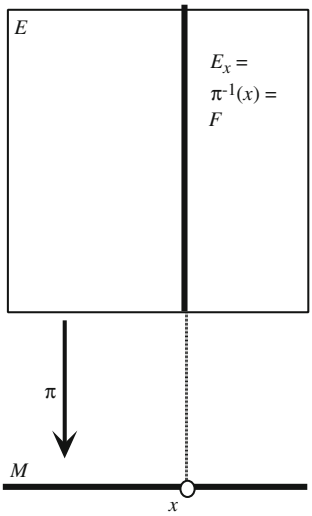


Fig. 4.19 Fibre bundle with base space M, fibre F, and total space E. Above each point x of M, the fibre $\pi^{-1}(x) = E_x$ is isomorphic to F

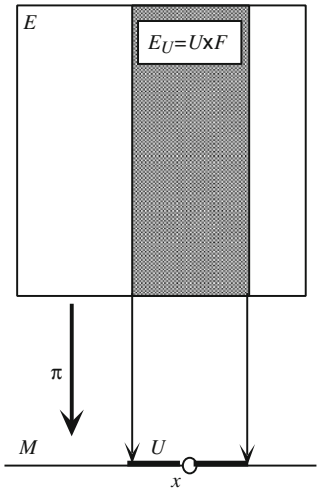


Fig. 4.20 Local triviality of a fibre bundle. For any point x in M, there is a neighbourhood U of x whose inverse image $\pi^{-1}(U) = E_U$ is the direct product $U \times F$, where π is the projection onto the first factor

4.3.4 V1 as a Geometric Fibre Bundle

If we idealize the functional architecture of the retino-geniculo-cortical pathway mathematically on the mesoscopic scale, the retinotopic and hypercolumn structures of $V1$ can be naturally modelled by the bundle $\pi : V \to R$ associating a copy P_a of the space P of directions in the plane with each point a of the retina R. The total space V of these copies P_a of P, glued together by local coordinate changes in the base space R, is a fibre bundle. We shall see in the next section that this is in fact

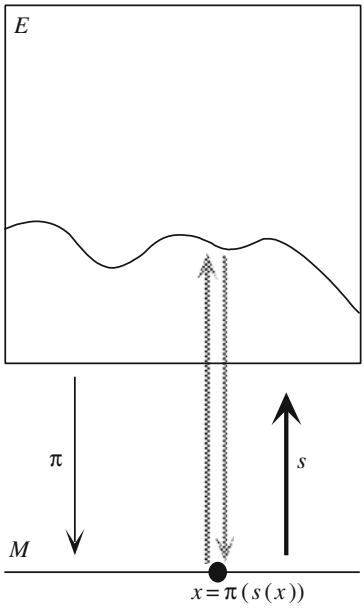


Fig. 4.21 A section of a fibre bundle defined on an open set U of M associates a value $s(x)$ in the fibre E_x above x with each point x of U

the *contact fibre bundle* CR of R, in other words, the projectivization of the tangent bundle TR of R. The points of V, that is, the pairs (a, p) comprising a point a of R and an orientation p at a, are the contact elements of R, already mentioned several times (see Fig. 4.22).

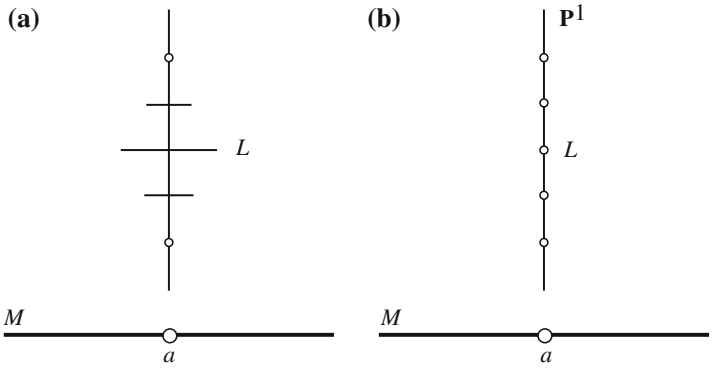


Fig. 4.22 The fibre bundle $E = V$ with base space the retinal plane $M = R$ (represented by a line to simplify) and fibre the projective line $\mathbb{P}^1$ of directions in M. **a** Elements of the fibre above a are represented by rotating horizontal line segments viewed in perspective. **b** Elements of the fibre above a are represented as points (the coordinate in the fibre encodes the angle θ of the direction p)

Through this functional architecture, a (discretized) copy of the orientation space P is associated retinotopically with each retinal position a. There thus exists a neural implementation of the structure $\pi : R \times P \to R$, the set of feed-forward projections (in the neurophysiological sense) of the retino-geniculo-cortical pathways implementing the projection π (in the geometric sense).

4.3.5 $V1$ as a 1-jet Fibre Bundle

Few specialists grasp the fundamental importance of an abstract geometrical model with dimension equal to the number of degrees of freedom of the given empirical structure, in this case 3. This theoretical requirement leads us to introduce the fibre bundle $\pi : V = R \times P \to R$, and this can be interpreted in different ways:

(i) As $\pi : R \times \mathbb{P}^1 \to R$ if the fibre P is identified with the projective line of orientations in the plane specified modulo π (P is topologically a circle).
(ii) As $\pi : R \times \mathbb{S}^1 \to R$, where $\mathbb{S}^1$ is the unit circle, if the fibre P is described by an angular coordinate θ modulo 2π, taking into account sense as well as orientation.
(iii) As $\pi : R \times \mathbb{R} \to R$ if the fibre P is described by the tangent $p = \tan\theta$.

An excellent example of interpretation (i) is provided by the work of Steve Zucker and Ohad Ben-Shahar. In their paper [28], we find Fig. 4.23.

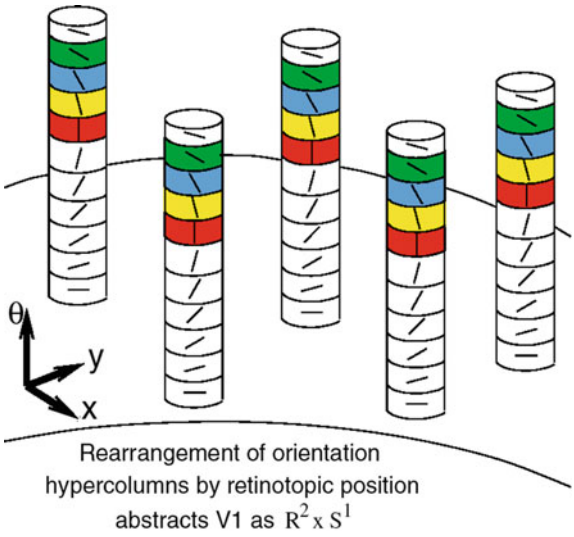


Fig. 4.23 Fibre bundle $\pi : \mathbb{R}^2 \times \mathbb{S}^1 \to \mathbb{R}^2$ introduced by Ben-Shahar and Zucker [28] to model the abstract structure of the orientation hypercolumns of $V1$

However, in interpretation (iii), $V1$ is identified with what is known as the 1-*jet bundle* of curves in R. The idea of jets generalizes the classical notion of Taylor[4] expansion, thereby giving it an intrinsic geometric meaning, i.e., independent of coordinates. Suppose that, in a certain coordinate system (x, y) of R, a smooth curve γ is the graph $\{x, f(x)\}$ of a real-valued function f on $\mathbb{R}$. The first-order jet of f at x, denoted by $j^1 f(x)$, is characterized by 3 arguments: the coordinate x, the value $y = f(x)$ of f at x, and the value $p = f'(x)$ of the derivative of f at x, i.e., the slope of the tangent to the graph of f at the point $a = (x, f(x))$ of R. So if we identify R with a domain of $\mathbb{R}^2$, a 1-jet is just a pair $c = (a, p)$, i.e. a contact element. Conversely, with each contact element $c = (a, p)$, we may associate the set of regular functions f whose graph is tangent to c at a. Then, $J^1 R$ will denote the fibre bundle with base space R, usually denoted $J^1(\mathbb{R}, \mathbb{R})$, of 1-jets of curves in R.

These 1-jets are feature detectors specialized in the detection of tangents. The fact that the $V1$ area can be ideally identified with $J^1 R$ in the case of 'simple neurons' explains why it is functionally essential for *contour integrations*. In the 2D manifold R, the determination of the direction p tangent to a contour γ at a point a requires one to compare the values of γ in the neighbourhood of this point. However, the neural system can access this *local* geometric information directly in the *pointlike format* of a simple numerical value, provided that it calculates in the jet space $V1$, a three-dimensional space. This saves having to carry out a local computation that would be costly in terms of wiring.

Jan Koenderink in [29] stressed the importance of the jet concept for theories of vision. Without jets, it is hard to understand how the visual system could extract geometrical features like the tangent or the curvature of a curve at a point:

> Geometrical features become multilocal objects, i.e., in order to compute [boundary or curvature] the processor would have to look at different positions simultaneously, whereas in the case of jets it could establish a format that provides the information by addressing a single location. Routines accessing a single location may aptly be called *point processors*, those accessing multiple locations *array processors*. The difference is crucial in the sense that point processors need no geometrical expertise at all, whereas array processors do (e.g. they have to know the environment or neighbours of a given location). Koenderink [29, p. 374]

This is indeed the key point: $V1$ must carry out geometrical tasks, but without being availed of any 'geometrical know-how'. However, this is possible *only if the geometry is neurally hard-wired* and, as we shall see, this is precisely the role of a functional architecture.

We can already say to some extent why functional architectures are so important. In the brain, there are what are essentially *temporal correlations*. So how can temporal coherence transform itself into spatial morphologies, Gestalten, or patterns? As we shall see in detail in the second volume, the key phenomenon is *synchronization*. So how can synchronizations be transformed into patterns? To do this, we require highly constrained connections and hence a functional architecture.

[4]Named after Brook Taylor.

The deepest point, and without doubt the most difficult to understand for a non-mathematician, is that the introduction of a further independent variable p provides a way to replace 'geometrical know-how' by a functional architecture. Let us say once again that neurons are point processors (on the scale specified by the size of their RFs) and as a consequence can only measure a quantity at a point. But in order to do differential geometry with point processors rather than local processors, the only way is to add supplementary variables evaluating partial derivatives of appropriate degree. This is why Jan Koenderink stresses the fact that the (hyper)columns implement jet spaces neurobiologically:

> The modules (like 'cortical columns' in the physiological domain) of the sensorium are local approximations (*Nth order jets*) of the retinal illuminance that can be addressed as a *single datum* by the point processors. [29, p. 374]

Let us specify how $V1$ can be interpreted as a neural realization of the space of 1-jets of curves in R. At each point a in M, consider not the tangent space T_aR, but the set of its hyperplanes (its vector subspaces of codimension 1 and hence its straight lines), denoted by C_aR. This is isomorphic to the projective space $\mathbb{P}^1$. The total space in which these fibres are glued together is called the *contact fibre bundle* of R, denoted by CR.[5]

CR is almost the space of 1-jets $J^1(\mathbb{R},\mathbb{R}) = J^1R$ associated with the choice of coordinates (x, y). To see this, we only have to interpret the coordinate on its fibres C_aR in terms of T_aR. Equipping R with local coordinates (x, y) at a and the tangent plane T_aR with the natural coordinates (ξ, η) associated with it in the basis $(\partial/\partial x, \partial/\partial y)$, then on an open set not containing the 'vertical' straight line $\xi = 0$, a local coordinate on C_aR is $p = \eta/\xi$, and in the neighbourhood of $\xi = 0$, we may take the coordinate $p = \xi/\eta$. An element c of CR is thus attributed the coordinates $(x, y, p) = (a, p)$. It is straightforward to check that changes of chart associated with these natural coordinates on CR are diffeomorphisms. Hence, CR is a 3D differentiable manifold isomorphic to $V = R \times \mathbb{P}^1$.

The difference between CR and J^1R is that the fibre of J^1R is not the whole of $\mathbb{P}^1$ but the $\mathbb{R}$ given by the values of $\tan\theta$, the angle θ being specified modulo π, measured given a choice of x-axis and never taken equal to $\pi/2$. To obtain the fibre $\mathbb{P}^1$ of CR, we must compactify $\mathbb{R}$ by adding a point at infinity. CR is the compactification at infinity of J^1R, and its fibre corresponds to that of J^1R via the stereographic projection $\mathbb{P}^1 \to \mathbb{R}, \theta \to \tan\theta$. In the language of algebraic geometry, J^1R is the open affine subset of CR complementary to the section at infinity, the choice of this section corresponding to the choice of an x-axis in the plane R.

Specifying a section s of the fibre bundle $\pi : V = CR \to R$ above a subset U of R means associating an element $s(a)$ of the fibre C_aR above a, i.e., an orientation, with each point a of U. The sections are thus fields of pairs $(a, p) =$ (position, orientation). A fundamental special case is the sections restricted to differentiable

[5] We could distinguish between the retinal plane R and the cortical layer M (the base space of V) to which it projects. However, to simplify, we shall not do so, considering the retinotopic map $\chi : R \to M$ as the identity.

curves γ in R. They are obtained by choosing, above each point a of γ, the orientation of the *tangent* to γ at a.

4.3.6 Legendrian Lifts

By the implicit function theorem, any differentiable curve γ is locally the graph of a function $f(x)$, except at points with a vertical tangent. Let $j^1\gamma(a)$ be its 1-jet. It can be identified with the 1-jet $j^1 f(x)$ comprising the abscissa x of a, the value $y = f(x)$ of f at x, and the value $p = f'(x)$ of its derivative there. Rather than considering the *plane* equipped with coordinates (x, y) and calculating $y' = \mathrm{d}y/\mathrm{d}x$—as we have seen, this requires knowing not only the value $y = f(x)$ of f at x, but also the values of f in a neighbourhood of x—we work in the space with three dimensions spanned by coordinates (x, y, p), imposing the constraint $y' = p$. This very profound idea goes back to William Hamilton who, by introducing the conjugate momenta p_i of the position variables q_i of a mechanical system as independent variables, replaced the Lagrangian formulation of mechanics by what is now known as the Hamiltonian formulation.

Quite generally, if γ is a differentiable curve in R, parametrized by the equations $x(s)$ and $y(s)$, the 1-jet $j^1\gamma(a(s))$ of γ at $a(s) = (x(s), y(s))$ is the contact element $(a(s), p(s))$, where $p(s) = y'(s)/x'(s)$ is the slope of the tangent to γ at $a(s)$.

The image of $j^1\gamma$ is called the *Legendrian lift* of γ. The Legendrian lifts of curves γ in R no longer represent these curves as sets of points in R, but in a dual sense, i.e., what is called *projective duality*, as envelopes of their tangents. It is remarkable that biological evolution should have created two neurophysiological structures, the retina and the $V1$ area, in order to implement projective duality for contours.

4.3.7 Integrability Condition

We can thus associate a Legendrian lift $\Gamma = j^1\gamma$ with any smooth curve γ in R. However, these lifts Γ are rather specific so we need to characterize them carefully. Indeed, let $\Gamma = v(s) = (a(s), p(s)) = (x(s), y(s), p(s))$ be any skew curve in V. Its projection $a(s) = (x(s), y(s))$ is indeed a curve γ in R. However, there is no reason why Γ should be the Legendrian lift $j^1\gamma$ of its projection γ. This is only the case if $p(s) = p_{a(s)}(s)$. Put another way, a curve Γ defined locally by equations $y = f(x)$, $p = g(x)$ is the lift of a curve γ in V if and only if $g(x) = f'(x)$, i.e. $p = y'$. This condition is called an *integrability condition*. It is crucial, and we shall return to this at length.

4.3.8 $SE(2)$ *Invariance of* 1*-jets*

To simplify, set $R = \mathbb{R}^2$. The structure of $J^1\mathbb{R}^2$ is invariant under the action of the Euclidean group $SE(2) = \mathbb{R}^2 \rtimes SO(2)$ of isometries of the plane which is the semi-direct product $\rtimes$ of the group of translations $\mathbb{R}^2$ and the group of rotations $SO(2)$. Generally speaking, if G is a group and if H is a subgroup of G which operates on another, normal subgroup N of G, then G is the semi-direct product $N \rtimes H$ if its product law $\circ$ is $\left(n', h'\right) \circ (n, h) = \left(n'h'\,(n)\,, h'h\right)$.

Let (q, r_θ) be an element of $SE(2)$, where q is a point in $\mathbb{R}^2$ and r_θ the rotation through angle θ. Then, (q, r_θ) acts on the points a of $\mathbb{R}^2$ according to

$$(q, r_\theta)(a) = q + r_\theta(a)\ .$$

If (q, r_θ) and (s, r_φ) are 2 elements of $SE(2)$, their (non-commutative) product is given by

$$(s, r_\varphi) \circ (q, r_\theta) = \left(s + r_\varphi(q), r_{\varphi+\theta}\right)\ .$$

The product is non-commutative because $(q, r_\theta) \circ (s, r_\varphi) = \left(q + r_\theta(s), r_{\theta+\varphi}\right)$. Naturally, $r_{\varphi+\theta} = r_{\theta+\varphi}$, but $s + r_\varphi(q) \neq q + r_\theta(s)$ (see Fig. 4.24).

The rotation r_θ acts on the fibre bundle $J^1\mathbb{R}^2 \to \mathbb{R}^2$ by

$$r_\theta(a, \psi) = (r_\theta(a), \psi + \theta)\ ,$$

where ψ is the angular coordinate in the fibre. This further action on the fibre ensures that the alignment of the preferred orientations is itself also $SE(2)$-invariant (see

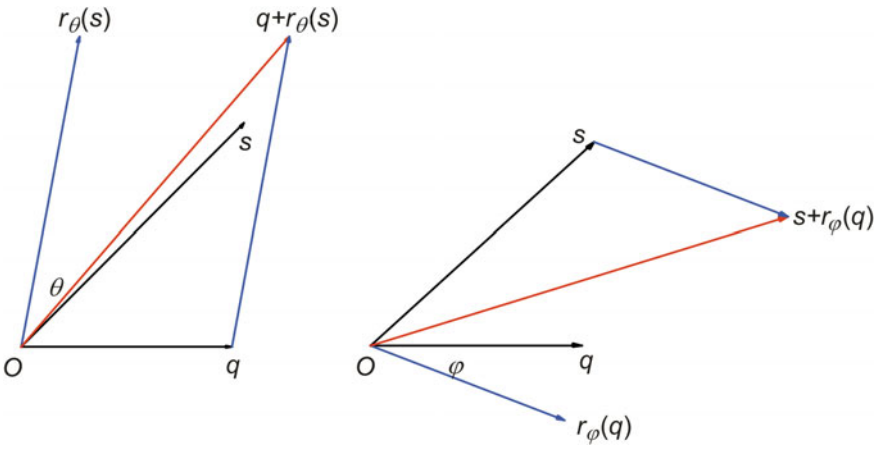


Fig. 4.24 Non-commutativity of the Euclidean group $SE(2)$. The vectors $s + r_\varphi(q)$ and $q + r_\theta(s)$ are completely different

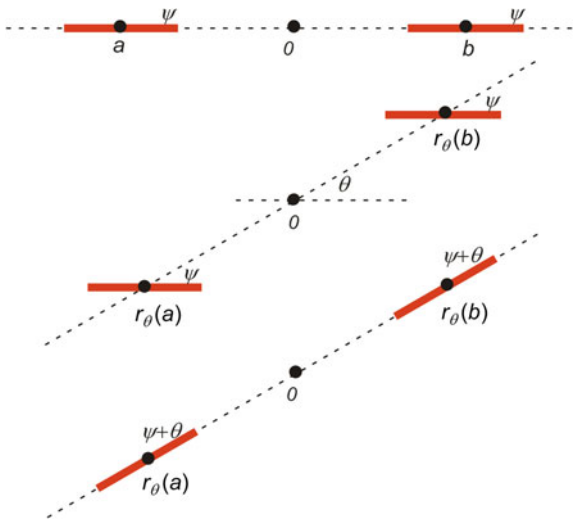


Fig. 4.25 $SE(2)$ invariance of the jet space

Fig. 4.25). This makes the action interesting from the point of view of group representation theory.

Regarding the group $G = SE(2)$, it is interesting to note a general phenomenon that is quite fundamental here. By defining $SE(2)$ as the group of isometries in the plane $\mathbb{R}^2$, we assume that $\mathbb{R}^2$ is given and we consider G only as a consequence. But this approach can be reversed. Indeed, $\mathbb{R}^2$ and the action of G can be reconstructed from the group structure of G. For if $H = SO(2) \simeq \mathbb{S}^1$ is the compact commutative subgroup of rotations around 0, its conjugates gHg^{-1} for $g = (q, r_\theta)$ in G give the rotation subgroups about different points q. We may then identify $\mathbb{R}^2$ with the quotient G/H and the action of G on $\mathbb{R}^2$ with the quotient of the action of G on itself.

4.3.9 Generalizing the Model

In accordance with Hubel's idea of engrafted variables, this model can be extended to other characteristic variables of the visual signal which are represented in the hypercolumns. To do this, we consider spaces in which the new variables vary, viz. the interval [0, 1] for the level of ocular dominance, the projective plane $\mathbb{P}^2$ for the colour, the circle $\mathbb{S}^1$ for directions of motion (the reader is referred to, e.g. Zhang and Wu [30] or Weliky et al. [31]). Then, $V1$ will be modelled by a fibre bundle with base space R and fibre the Cartesian product of the spaces of secondary variables.

Such generalizations might also lead us to consider k-jets of order $k > 1$, as we shall see in Sect. 5.6 of Chap. 5 when we discuss *curvature detectors*.

4.3.10 Neurophysiology and Its Geometrical Idealization

There are significant differences between the neurophysiological data and the geometrical idealization of the fibre bundle. Here, we mention three of these:

(i) To begin with, the RFs introduce a resolution scale so we require a *multiscale* theory of fibre bundles. Moreover, the RFs are adaptive and modulated by stimuli.
(ii) Then, there is a significant *redundancy* in the columns. Indeed, a 'point' (a, p) of the fibre bundle actually corresponds to a whole column. As we have seen, this so-called population coding is essential for adaptive capacity and refining the resolution. Among other things, it allows *oscillatory responses* (the columns can become oscillators through Hopf bifurcation) and hence synchronizing effects through phase-locking.
(iii) Finally, there is a fundamental *dimensional* constraint. From an abstract point of view, the fibre bundle $\pi : R \times P \to R$ has dimension 3, i.e. 2 degrees of freedom for the retinal position $a = (x, y)$ and 1 degree of freedom for the orientation p, whereas the cortical layers are essentially 2D. There is therefore a problem of 'dimensional collapse'. The visual systems produced by evolution that interest us here have solved this problem through the fascinating structure of 'pinwheels' in the $V1$ area, and these have received much attention since the pioneering work of Tobias Bonhöffer, Gary Blasdel, and Amiram Grinvald.

4.4 The Pinwheel Structure of $V1$

The model of $V1$ as a 1-jet space is a *continuous* model which is in fact the limit of a *discrete* model defined on a lattice in R. We shall return at length to the idea of continuous models in the second volume. But in the present chapter, we shall begin by gathering together some experimental data regarding discrete models.

4.4.1 Observation of Pinwheels

4.4.1.1 Functional Orientation Maps

Fundamental experiments made possible by recent progress in brain imaging have shown that the hypercolumns are arranged geometrically in little wheels called *pinwheels*. The observed cortical layer is covered with a lattice of singular points (about 1 200 μm apart in the cat and about 600 μm apart in primates): the centres of local pinwheels which join up to form a global structure. The imaging method used here was developed at the beginning of the 1990s, by Bonhöffer and Grinvald [32] among others, and it is referred to as in vivo optical imaging based on activity-dependent

intrinsic signals. It exploits the fact that the metabolic activity of the nerve tissue changes its optical properties, whence it can acquire images of the activity of the surface cortical layers. More precisely, it exploits the differential absorption of oxyhemoglobin or deoxyhemoglobin, or of dyes whose fluorescence indicates local depolarization of neurons.

The method has profoundly transformed the observation of neural activity. Previously, there were only multielectrode methods for recording the activity of a few individual neurons,[6] or post-mortem visualization of cortical activity using 2-deoxyglucose maps. Both methods were drastically inadequate. In vivo optical imaging made it possible to visualize functional organization; in other words, it made the cerebral black box 'transparent'. As pointed out by Ohki and Reid [33]:

> Optical imaging revolutionized the study of functional architecture by showing the overall geometry of functional maps.

However, the experimental challenge is enormous. First, the signal-to-noise ratio is very small, viz. $\sim 10^{-3}$, because the background noise is huge. The intrinsic signal comes from the hemodynamic properties of the cortical tissue and thus from an area of vascular metabolism greater than the area of the activated neurons. Then, millions of neurons are connected together, each with hundreds or even thousands of synapses, and this imaging operates on a *mesoscale* defined by averaging (in Sect. 4.7.3, we shall discuss the methods of two-photon confocal microscopy on the microlevel). One 'neuron' is in fact a cortical position at which a bunch of neurons is located. Moreover, the in vivo optical imaging methods with their good mesoscopic spatial resolution (50 μm) actually have a rather poor temporal resolution, because the intrinsic signals are slow, so they can only analyze slow intrinsic changes in the optical properties of the cortical layer. Other methods are required to visualize the cortical dynamics, such as voltage-sensitive dyes, which colour the active cells. The dye molecules bind to the neural membranes and act:

> [...] as molecular transducers that transform changes in membrane potential into optical signals. [34]

As the changes in the membrane potential are correlated with millisecond changes in the absorbed or emitted fluorescence, we may thus obtain temporal resolutions of millisecond order.

For a discussion of these new techniques for investigating mammalian brains, the reader is referred to the review [34], presented by their inventor Amiram Grinvald. Figure 4.26 shows the accuracy that can be reached in the topography of the $V1$ area of the owl monkey (douroucoulis). Such maps became available from the end of the 1980s. Figure 4.27 produced by Blasdel and Salama in 1986 already shows pinwheels in the $V1$ area of the macaque.

---

[6] The transition from recordings of a few isolated neurons to a visualization of the overall activity of a piece of brain area is analogous to the leap forward in meteorology when recordings made by weather balloons were replaced by satellite imaging. No need for further comment.

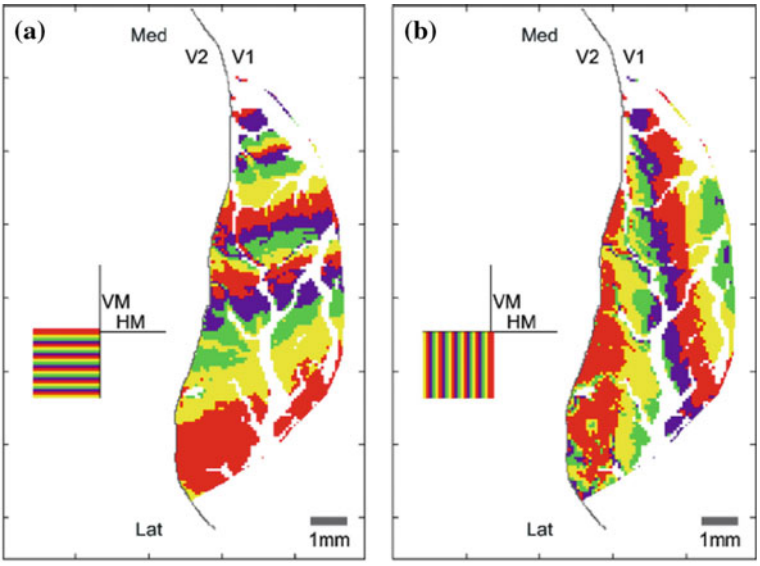


Fig. 4.26 Topography of the $V1$ area of the New World owl monkey as measured by in vivo optical imaging. A target made of two gratings, one horizontal and the other vertical, is projected onto $V1$. VM vertical meridian, HM horizontal meridian. From Grinvald [34]

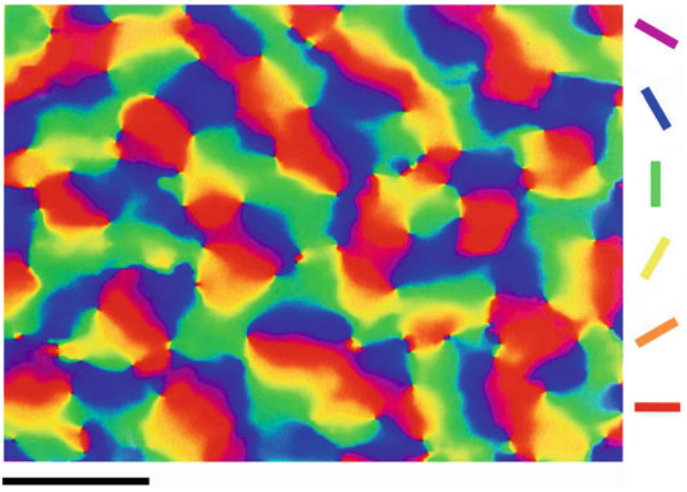


Fig. 4.27 Pinwheels in the $V1$ area of the macaque. From Blasdel and Salama [35]

According to a standard experimental protocol, the animal is shown high-contrast gratings between 20 and 80 times. These are made of black bands (e.g. 6.25°) alternating with white bands (e.g. 1.25°), with several different orientations (e.g. 8) and an angular speed of say 22.5°/s. A cranial window is opened above $V1$, and the cortex is illuminated with orange light (605 nm). The orientation maps have very low amplitude relative to the light intensity of the recorded cortical images. We thus subtract the average intensity of the responses for all orientations, which is known as the cocktail blank. Then, depending on the orientation of the gratings, differential absorption patterns are observed, due to local spatial non-uniformities in the ratio of deoxyhemoglobin to oxyhemoglobin. We subsequently sum the images of $V1$ activity obtained for different gratings and construct differential maps. These are normalized by dividing the relative deviations from the average of each pixel by the global average deviation, and low-frequency noise is also eliminated.

We thus obtain *functional maps* like those in Figs. 4.28 and 4.29 obtained by Crair et al. [36] and Bosking et al. [37], respectively, which concern the 2 and 3 layers of a tree shrew (tupaia): LGN $\rightarrow$ layer 4 $\rightarrow$ (strictly feed-forward) $\rightarrow$ layers 2 and 3. (The tree shrew looks like a primate but with the difference that there is

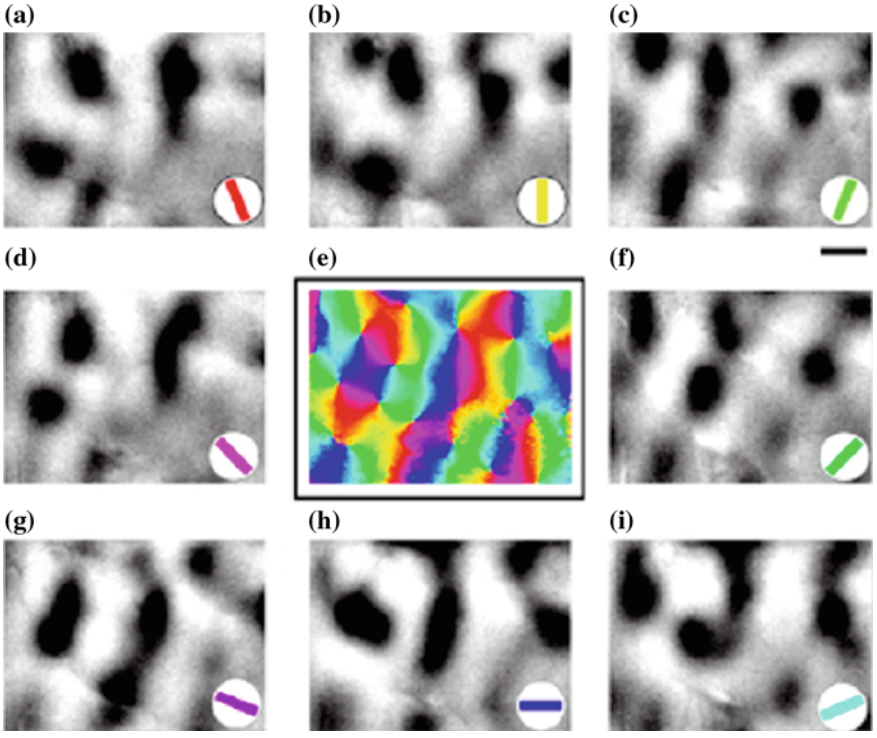


Fig. 4.28 Method for producing an orientation map of $V1$. Preferred orientations are conventionally colour-coded. Taken from Crair et al. [36]

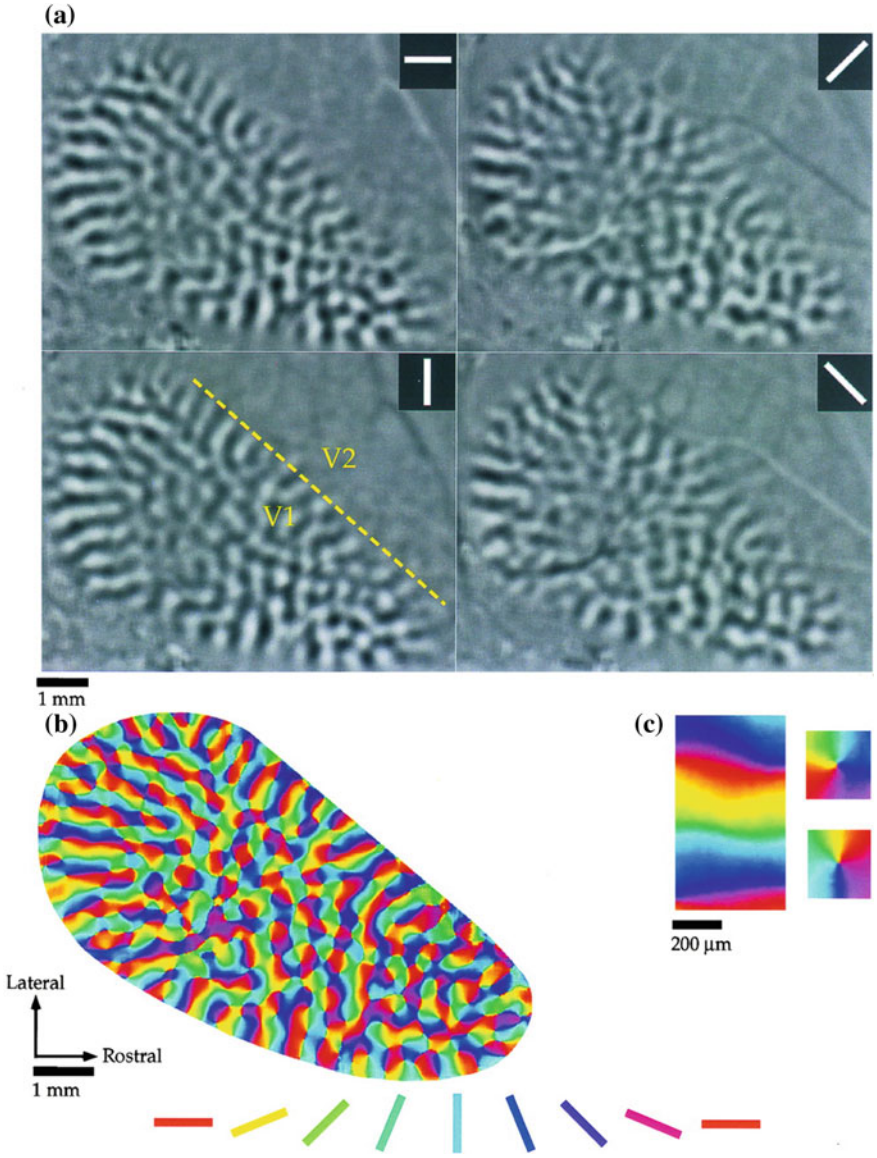


Fig. 4.29 *V*1 of a tree shrew (tupaia). The different orientations are colour-coded. *Right* Zoom on examples of regular points and singular points of opposite chirality, as explained in Sect. 4.4.1.2. From Bosking et al. [37]

no fovea/periphery distinction, and this simplifies observation. For a survey of the functional organization of his visual cortex, see Fitzpatrick [38] and Lund et al.

[39].) The orientations are colour-coded, and the iso-orientation lines are thus the monochromatic lines.

As emphasized by Yves Frégnac et al. [40], these mesoscale optical functional imaging methods can be supplemented by microscale recordings of individual neurons, either extracellular and simultaneous recordings of several cells using multielectrodes able to measure correlations in the activities of the various neurons involved, or intracellular recordings that can measure both spiking activity and sub-threshold activity, and by inverse analysis can be used to reconstruct the afferent network of the given neuron (functional synaptic imaging). Figure 4.30 from [40] shows the relationship between these three methods schematically.

Optical imaging data helps us to understand what is meant by the population coding of a stimulus. Figure 4.31, which we shall consider again in Sect. 4.9.1, shows the neurons in $V1$ which are activated by a long vertical bar. We see that it is a rather thick band, also very patchy, and nothing like a line. As pointed out by Ulf Eysel [41, p. 641]:

> A continuous line across the whole visual field would be cortically depicted in a patchy discontinuous fashion.

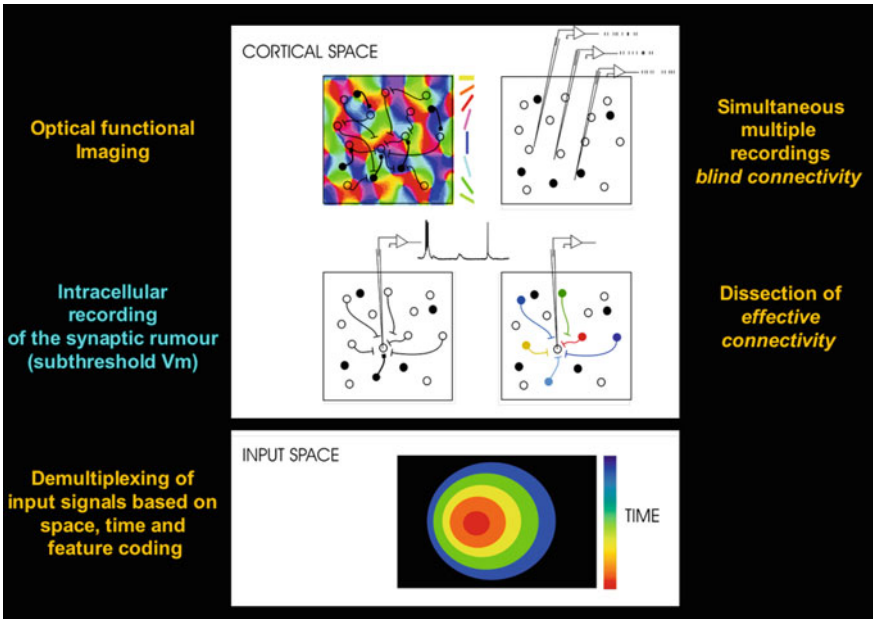


Fig. 4.30 Three methods for microlevel recording of the activity of individual neurons. Simultaneous extracellular recording of several cells, intracellular recording, and reconstruction of the afferent network of a neuron by functional synaptic imaging. Taken from Frégnac et al. [40]

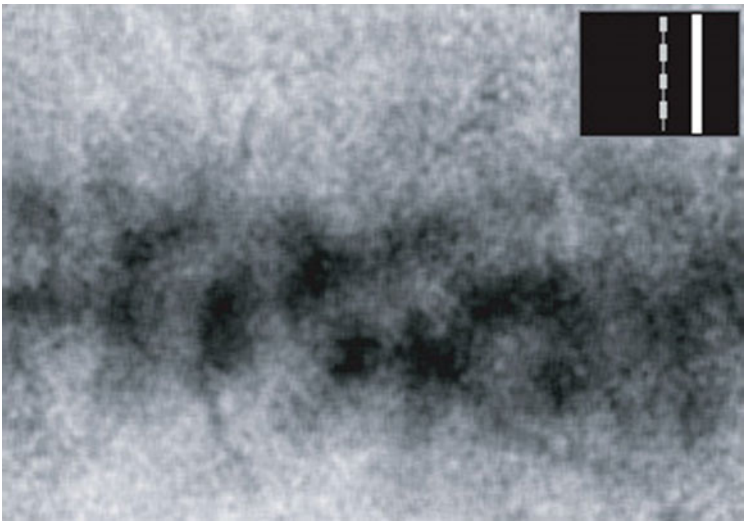


Fig. 4.31 Neurons of $V1$ activated by a long vertical bar. Note that we obtain a thick and rather patchy band. From [42]

4.4.1.2 Periodicity, Retinotopy, and Singularities

We return here to Fig. 4.29. Note to begin with that the lattice has a kind of *characteristic length*—a mesh or *wave number*. This periodicity can be measured accurately by taking the map of the pinwheels, translating it by $t = (u, v)$, and calculating the correlation between the two maps. The autocorrelation obviously has a primary peak at $t = (0, 0)$ (the map correlates perfectly with itself) and secondary peaks giving the periodicity. Figure 4.32 is taken from the paper [43] by McLoughlin and Schiessl and gives the example of the marmoset monkey.

We shall see in Sect. 4.6.6 that, under a Fourier transform, an orientation map can be interpreted as a superposition of plane wave solutions of the Helmholtz equation and that the mesh of the pinwheel lattice comes from the fact that these plane waves have almost the same wave number. More precisely, let $\Phi(z) = e^{i\varphi(z)}$ be the orientation map [or $\Phi(z) = e^{i2\varphi(z)}$, if we prefer to work modulo 2π and not take into account the chirality of the pinwheels]. The autocorrelation function of $\Phi(z)$ is given by $C(u) = \int_{\mathbb{C}} \Phi^*(z)\,\Phi(z-u)\,dz$, where $\Phi^*(z)$ is the complex conjugate of $\Phi(z)$. Isotropy implies that $C(u) = C(r)$ depends only on the modulus $r = |u|$ of u. So let us consider the Fourier transform $P(k)$ of $C(r)$ (known as the power spectrum), viz.

$$P(k) = \int_{\mathbb{R}^+} e^{-ikr} C(r)\,dr\ .$$

$P(k)$ is maximal on a ring of radius $k_0 = 2\pi/\Lambda_0$, and Λ_0 specifies the mesh of the pinwheel lattice. This structure in $P(k)$ arises in the following way:

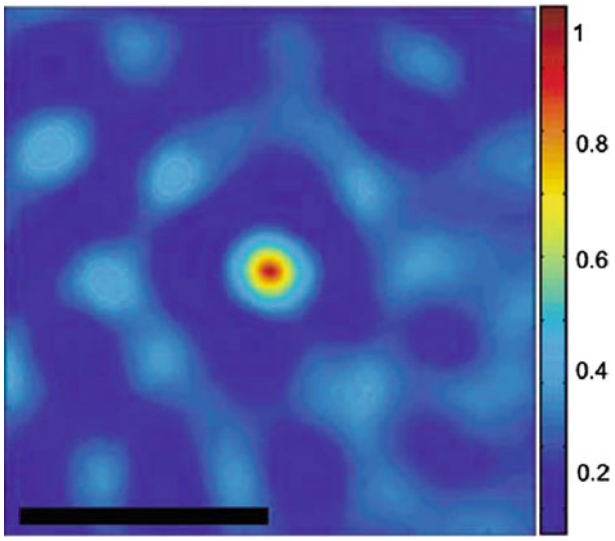


Fig. 4.32 Autocorrelation with its translates of the pinwheel map of the marmoset monkey $V1$ area. It is colour-coded from 0 to 1. At the *centre*, the *red* point means that the autocorrelation is 1 for $t = (0, 0)$. The first *dark blue* ring around the centre corresponds to the distance at which the pinwheels of opposite chirality are superposed. Then, the first *light blue ring* corresponds to the basic period. The average distance between pinwheels is 575 μm in this example. Taken from [43]

(i) Local triviality outside singularities indicates a small local variation in orientations, and this low local correlation implies that there are no Fourier components with high spatial frequencies $k \gg k_0$, because these would induce large local variations in the orientations everywhere.
(ii) Uniformity, i.e. the fact that all orientations are represented in equal amounts, implies that there are no Fourier components with low spatial frequencies $k \ll k_0$, because these would induce long-range non-uniformities.

Figure 4.33, produced by Niebur and Wörgötter [44], shows an orientation map for a macaque (area 18) and the power spectrum concentrated on a ring of average radius $k_0 = 2\pi/\Lambda_0$. As already mentioned, we shall return in Sect. 4.6.6 to the limiting case in which the power spectrum is totally concentrated on a *circle* of radius k_0.

Note that in Fig. 4.29, the orientations (represented by colours) are distributed in a globally homogeneous way. If the sectors of different colours seem to be distributed inhomogeneously in the pinwheels, this is because of the perceptual structure of the continuous wheel of colours; in contrast with what happens for discretized wheels, the names of the colours correspond to sectors of a different angular width (see Fig. 4.34).

Note also that there are three classes of points (they are represented in the idealized diagram of Fig. 4.35):

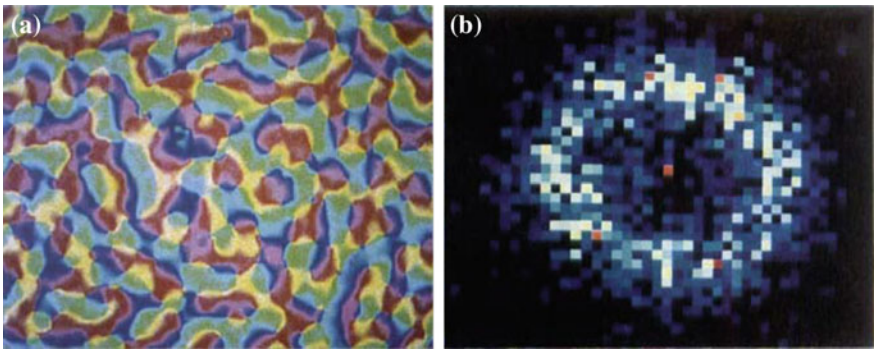


Fig. 4.33 Periodicity of an orientation map (*right*). The power spectrum (Fourier transform of the autocorrelation function of the map) is concentrated on a ring of average radius $k_0 = 2\pi/\Lambda_0$, and Λ_0 specifies the periodicity of the map. From Niebur and Wörgötter [44]

Fig. 4.34 In the standard continuous colour wheel, the names of the colours correspond in fact to angular sectors of different angles. The primary colours in additive colour mixing, viz. *blue*, *green*, and *red*, occupy large sectors, whereas their complementary colours, which are the primary colours in subtractive colour mixing, viz. *yellow*, *magenta*, *cyan*, each occupy narrow sectors

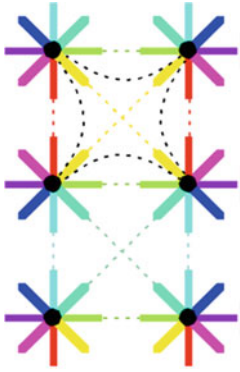


Fig. 4.35 Idealized 'crystalline' pinwheel structure on a regular square lattice of singularities. There are three classes of points: regular, singular (pinwheel), and saddle points

(i) Regular points where the orientation field is locally trivial in the sense that the iso-orientation lines are approximately parallel.
(ii) Singular points at the centres of the pinwheels where all the orientations converge. They have a positive or negative chirality depending on whether, when we move around the centre, the orientations rotate in the same sense or not. They have opposite chiralities when adjacent.
(iii) Saddle points at the centres of the lattice cells, points where the iso-orientation lines bifurcate: two neighbouring iso-orientation lines leave the same singular point but end up at two opposite singular points.

Xu et al. [45] studied analogous structures for the prosimian primate[7] known as the bush baby (galago), which is a monkey ancestor with an analogous $V1$ to the monkey primates, but very different $V2$. Figure 4.36, taken from [5], shows the isochromatic lines more clearly, these being the iso-orientation field lines of $V1$. Figure 4.37, produced by Hongbo Yu et al. [46], shows the receptive fields corresponding to positions along a line segment in $V1$. We see the superposition of receptive fields for neighbouring positions in a pinwheel and their slow displacement when we move from one pinwheel to another. This explains the pioneering experiments by Hubel and Wiesel.

4.4.2 Limitations of This Analysis

As noted by Amit Basole et al. [47], optical imaging methods are very powerful, but the result must be interpreted cautiously because they depend among other things on the type of inputs used. For example, they often use gratings which cross the receptive fields (RFs), so we can only test the motion selectivity of the neurons in the orthogonal direction (the aperture problem). If the input patterns are changed by using shorter segments than the RFs in order to be able to test other directions of motion, we obtain the same pinwheel structure for the orthogonal motions, but a *distortion* of this structure for oblique motions.

4.4.3 Functional Maps as Fields

It is interesting to note that the pinwheel lattices with their iso-orientation lines look like 'field models'. The singularities c_i with their chiralities are analogous to positive and negative charges producing field lines in $\mathbb{R}^2$. This kind of field structure was introduced early on by Valentino Braitenberg, in fact in 1979, in a paper entitled *Geometry of orientation columns in the visual cortex* [23], well before the development of in vivo optical imaging methods. Starting with Hubel and Wiesel's results

[7] There are two classes of primates: on the one hand, monkeys and humans, and on the other, the prosimians.

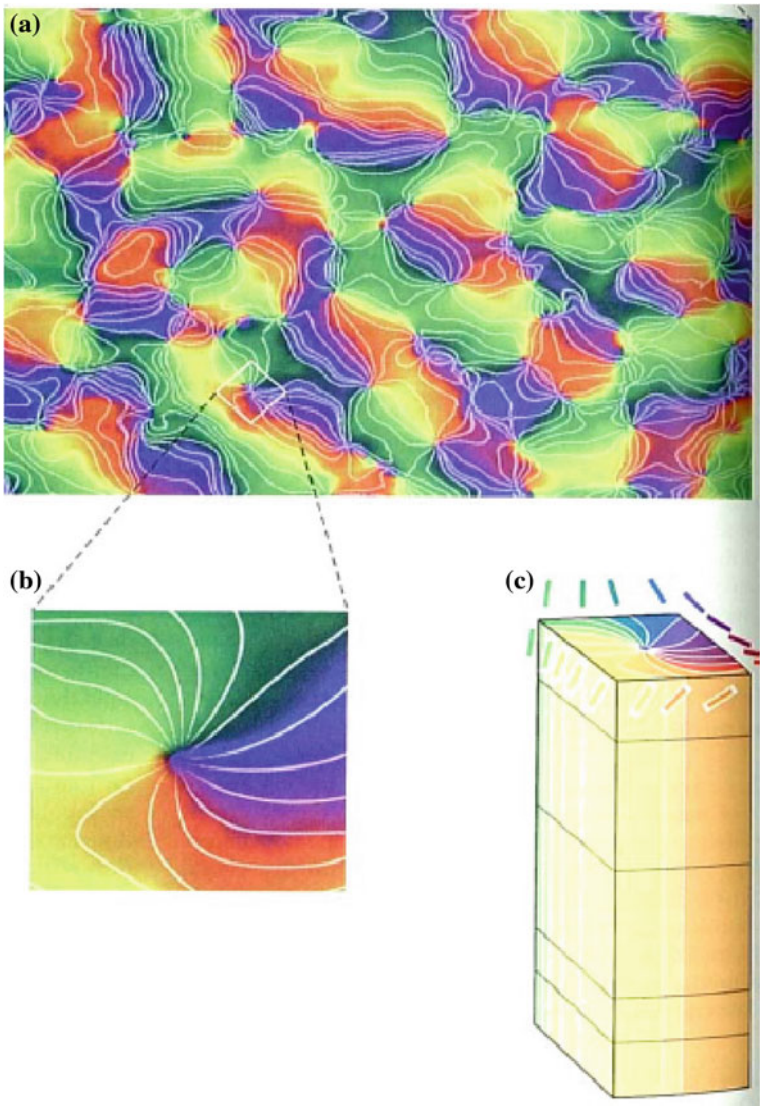


Fig. 4.36 The isochromatic *lines* are the iso-orientation lines of *V*1. From [5]

on monkeys and cats in 1962, and their discoveries (i) that the preferred orientation of neurons depends smoothly on the tangential penetrations of the electrodes, (ii) that there is chirality, and (iii) that the latter can reverse along a penetration, the Braitenbergs tried to build orientation fields. Through quite remarkable abduction, they came to the following conclusion:

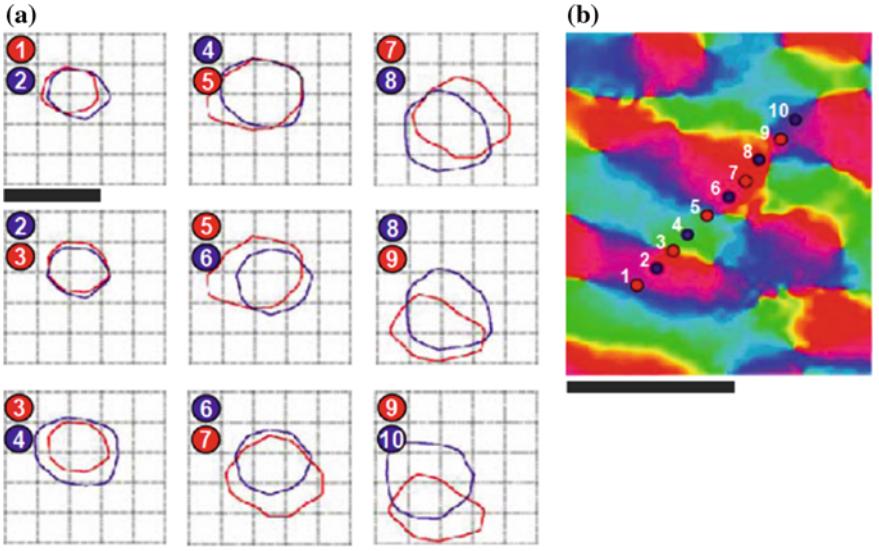


Fig. 4.37 Receptive fields of simple $V1$ neurons at different positions along a line segment. We see the superposition of the receptive fields for neighbouring positions in a pinwheel and their slow displacement when we move from one pinwheel to another. From Yu et al. [46]

> We believe that the most natural explanation of the facts observed would be in terms of orientations arranged with circular symmetry around centres, either radially or along concentric circles.

They clearly anticipated by clever abduction the pinwheel structure discovered later experimentally. Figure 4.38 of [23] shows how a straight line penetration L encounters the pinwheels and can reverse the chirality when the singularities lie on either side of L.

Following the work by the Braitenbergs, other specialists like William Baxter and Bruce Dow carried out further investigations of these inferences. In their paper [48] on pinwheels in the macaque, they explicitly constructed the flow of the orientation field and introduced singularities of a different chirality and different topological index, i.e., singularities of index ± 1 giving sources and sinks and those of index $\pm 1/2$ giving end points and triple points, which we shall discuss shortly. Figure 4.39 gives an example.

It would be interesting to reconstruct more or less explicit models of orientation fields that can be observed empirically in the pinwheels of different species. To do this, we could use the relationship between the fields of straight lines provided by the orientation fields and vector fields carried by these straight lines. Away from singularities where the vectors vanish or diverge, a vector field automatically induces an orientation field, simply by taking the orientations of the vectors. Supposing that it is meaningful to consider the singularities with their chirality as positive and

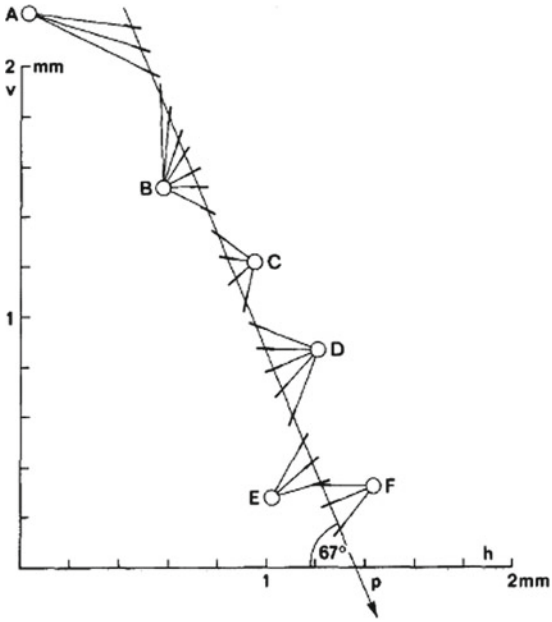


Fig. 4.38 The Braitenbergs' theoretical reconstruction by abduction of the pinwheel structure of $V1$. We observe the orientations along the straight line tangential penetration L of an electrode. The pinwheels are reconstructed to explain the data, and in particular the reversal of chirality when the singularities lie on either side of L. From Braitenberg [23]

negative charges, we could then consider the field lines of an associated 'physical' field deriving from a potential.

Consider, for example, the field of Fig. 4.40 produced by Geoffrey Goodhill in [49]. An arrangement of charges ± 1 is constructed by applying the rule that two neighbouring singularities always have opposite charges, together with several supplementary terms to take into account edge effects. By inspection, we find that the 'physical' analogy is roughly correct. In Fig. 4.41, the field lines are the orange lines from the sources $(+)$ towards the sinks $(-)$, while the lines orthogonal to the field line are equipotentials, i.e. lines of equal potential, for the potential whose gradient gives the field.

To give another example, consider the fragment shown in Fig. 4.42 of the image in Fig. 4.97 that we will comment on in Sect. 4.9.3. Figure 4.43 shows the physical analogy with the field induced by the $+1$ and -1 charges located at singular points and with a certain number of equipotential curves which are by definition orthogonal to the field.

The equipotential line of mean level is a separatrix between the influence zones of charges $+1$ and -1 (see Fig. 4.44).

For a regular 'crystalline' pinwheel lattice with square unit cell like the one in Fig. 4.35 of Sect. 4.4, we obtain Figs. 4.45 and 4.46, whose idealized geometry is shown in Fig. 4.47.

In Figs. 4.48 and 4.49, we show how the ideal crystal model can be distorted.

With the notion of phase field, we shall discuss a deeper way to understand pinwheels as 'physical fields' in Sect. 4.6.

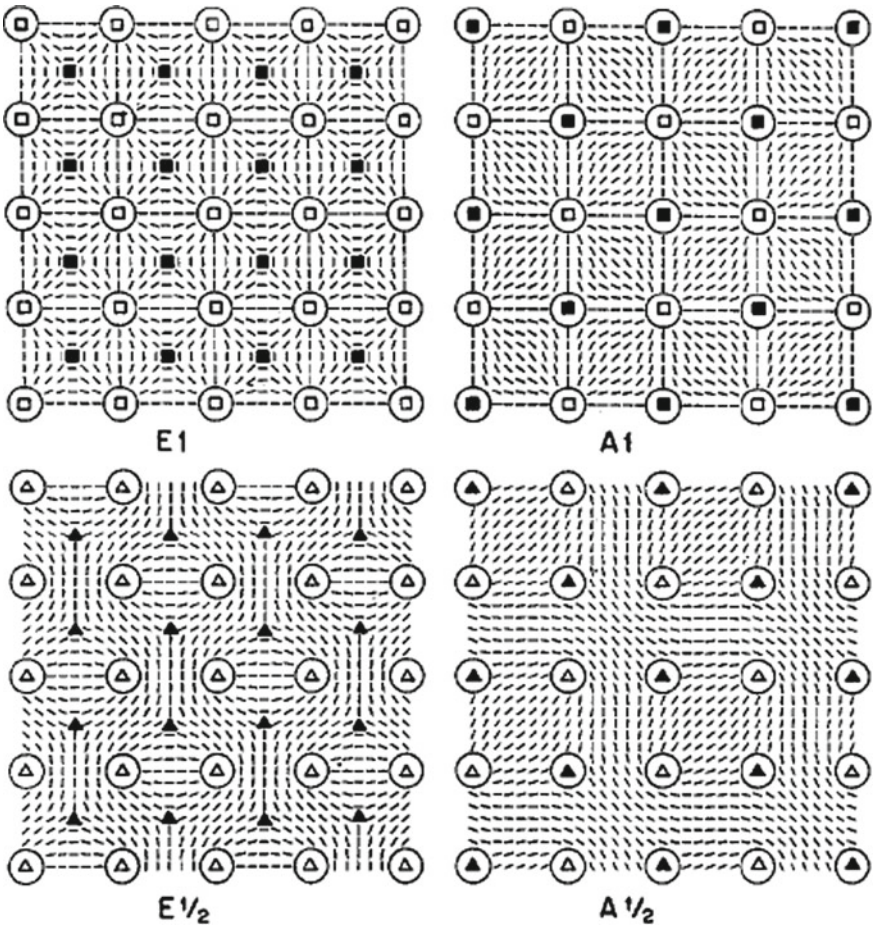


Fig. 4.39 Four examples of abstract 'crystal' models of pinwheels put forward by William Baxter and Bruce Dow. *White* and *black squares* represent singularities with topological indices $+1$ and -1, respectively, while *white* and *black triangles* represent singularities with topological indices $+1/2$ and $-1/2$, respectively. Taken from [48]

4.4.4 Development of Pinwheels

There is some experimental data concerning the evolution of orientation maps and their pinwheels while the $V1$ area is developing. For example, in the mid-1990s, Barbara Chapman, Michael Stryker, and Tobias Bonhöffer [51] adapted the optical imaging method to carry out what they called *chronic* optical imaging of intrinsic signals. The aim was to study the emergence and development of the pinwheel structure. (They were working on the ferret, whose visual system is similar to the cat's.) They showed that, starting from a barely structured initial orientation map,

Fig. 4.40 Example of an orientation field reconstructed from neural network models, with its pinwheels and field lines. The *black lines* are the boundaries of ocular dominance domains, discussed in Sect. 4.10.1. From Goodhill [49]

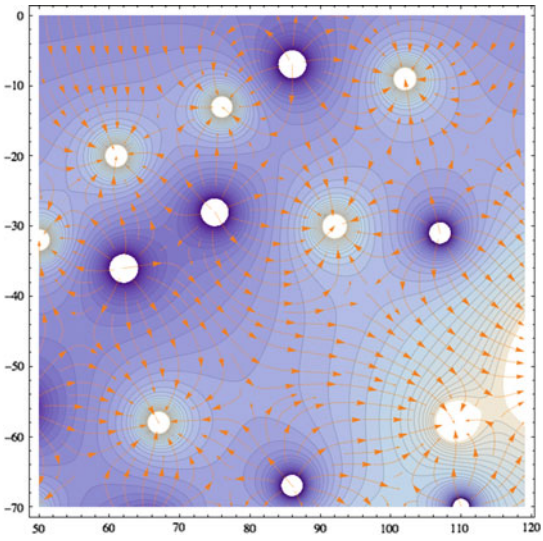


Fig. 4.41 'Physical' field with charges $+1$ and -1 distributed like the pinwheels in Fig. 4.40 (ordinate $-y$). Field lines are orange lines going from the sources $(+)$ to the sinks $(-)$. Lines orthogonal to the field lines are equipotentials, i.e. lines of equal potential, for the potential whose gradient gives the field. Plot drawn with *Mathematica* [50]

Fig. 4.42 Fragment of the image in Fig. 4.97

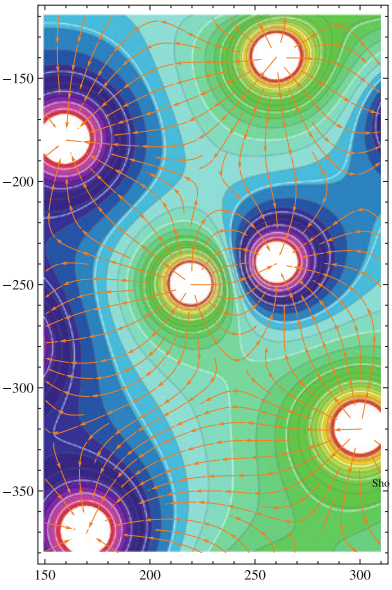


Fig. 4.43 Physical analogue of the field in Fig. 4.42 induced by $+1$ and -1 charges located at singular points. A certain number of equipotential curves orthogonal to the field are also shown

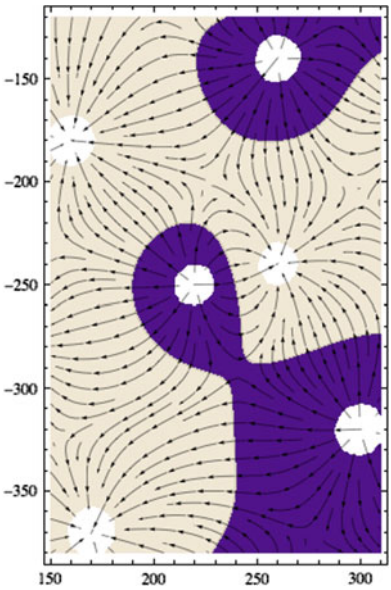


Fig. 4.44 If we take the mean level equipotential of the field in Fig. 4.43, we get a separatrix between the influence zones of charges $+1$ and -1. The small white discs around the singular points are due to the fact that the divergences of the potential at these points are clipped

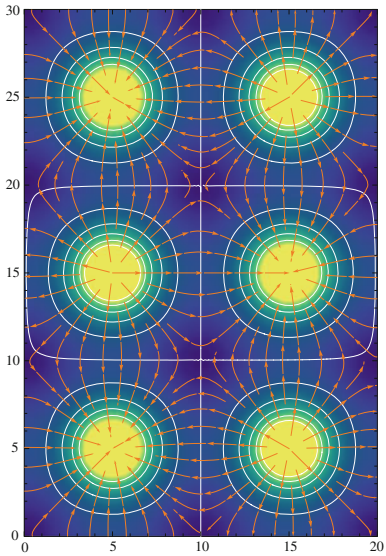


Fig. 4.45 Field of the regular 'crystalline' pinwheel lattice with square unit cell in Fig. 4.35 of Sect. 4.4

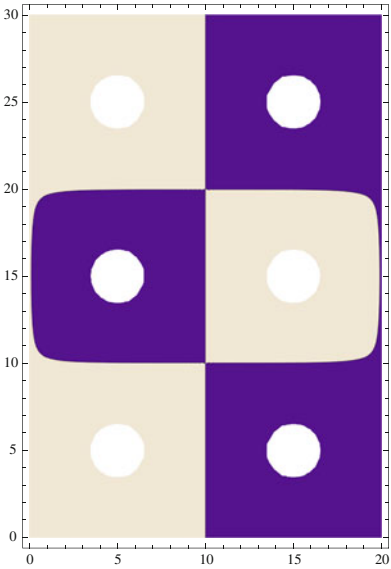


Fig. 4.46 Equipotential separating the influence zones of $+1$ and -1 charges of the field in Fig. 4.45

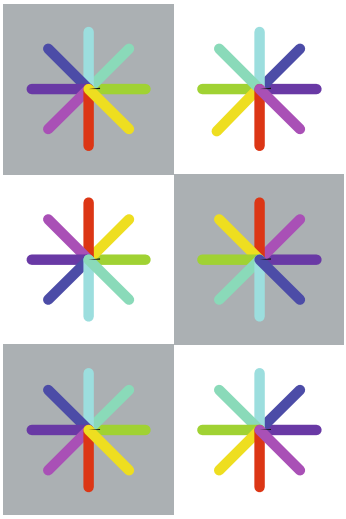


Fig. 4.47 $+/-$ dominance domains in the 'crystal' model idealized in Fig. 4.35

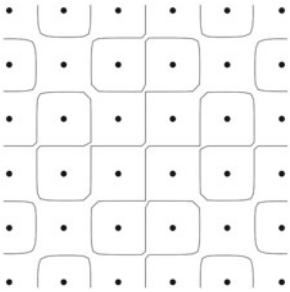


Fig. 4.48 Idealized crystalline geometric model of Fig. 4.47, interpreted using a physical analogy in which the pinwheel singularities generate a field. Gradient lines (corresponding to iso-orientation lines) and equipotentials are shown

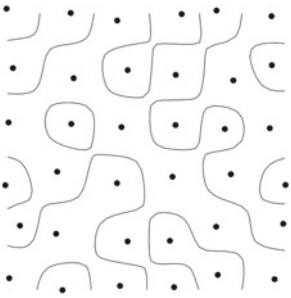


Fig. 4.49 Distortion of the physical analogy in Fig. 4.48

an organized map gradually comes into being and eventually becomes remarkably stable.

We shall return to pinwheel morphogenesis and the associated learning processes in Sect. 4.7.1.2, but this time in terms of models.

4.4.5 Pinwheels and Evolution

Regarding the pinwheel structure of $V1$, as for any other biological structure, we may wonder how it evolved, and what exactly was being optimized as a result of some evolutionary pressure. We already discussed this question in Sect. 3.6 of Chap. 3, in relation to the shape of the receptive profiles of ganglion cells and LGN and $V1$ neurons.

4.4.5.1 Minimizing the Wiring

Alexei Koulakov and Dmitri Chklovskii [52] investigated the hypothesis that the pinwheel structure *minimizes the total length L* of intracortical connections, while maintaining a good representation of the properties of the stimuli. To do this, they con-

sidered connection functions $c(\theta)$ giving the number of connections between a neuron $\mathscr{A}$ with preferred orientation 0 and neurons with orientations $\theta \in [-\pi/2, \pi/2]$. The functions $c(\theta)$ considered have the form $c + G(\theta)$, where c is a constant favouring a number of connections independent of θ, and $G(\theta)$ is a Gaussian with mean zero favouring connections with neurons having the same orientation as $\mathscr{A}$. In the limit, $G(\theta)$ gives the Dirac distribution at 0. These two models c and $G(\theta)$ compete to minimize L, and when the Gaussian G is narrow enough, pinwheels appear.

Seen from this point of view, the advantage of the pinwheels is that a neuron does not have to go further than the nearest pinwheel in order to connect itself to a neuron of any given orientation. Figure 4.50a of [52] shows an example function $c(\theta)$ and Fig. 4.50b the pinwheel map obtained by minimizing L. The lattice contains $N = 50 \times 50 = 2500$ neurons. We consider the set $\mathfrak{M}$ of 2500×2500 matrices $M = (M_{ij})_{i,j=1,\ldots,N}$ of connections between neurons $i = (x_i, y_i)$ and $j = (x_j, y_j)$ which satisfy the constraint $c(\theta)$ with $M_{ij} = 1$ if the neurons i and j are connected and 0 otherwise. We have $L = \sum_{i,j=1}^{i,j=N} d_{ij} M_{ij}$, where d_{ij} is the distance between i and j, and we minimize L in $\mathfrak{M}$, which is difficult, using simulated annealing algorithms.[8]

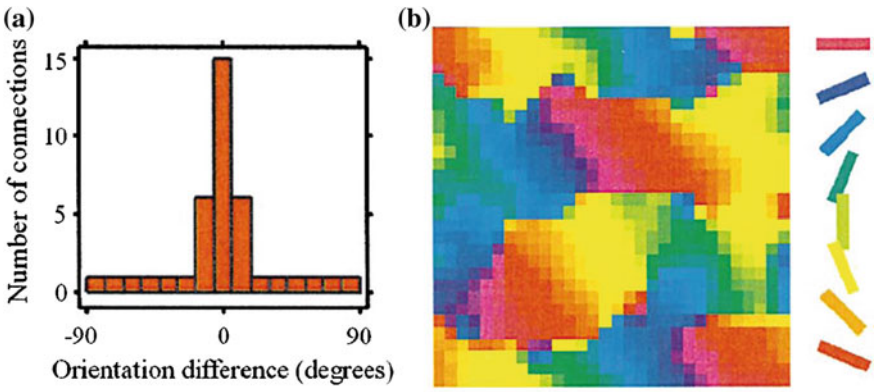


Fig. 4.50 Minimizing the total length L of intracortical connections for the connection function $c(\theta)$ shown in (**a**). **b** Pinwheel map for a square lattice of $N = 50 \times 50 = 2500$ neurons. We consider 2500×2500 matrices of connections $M = (M_{ij})_{i,j=1,\ldots,N}$ between neurons $i = (x_i, y_i)$ and $j = (x_j, y_j)$ which satisfy the constraint $c(\theta)$ with $M_{ij} = 1$ if neurons i and j are connected and 0 otherwise. We have $L = \sum_{i,j=1}^{i,j=N} d_{ij} M_{ij}$, where d_{ij} is the distance between i and j. We minimize L using simulated annealing algorithms. From Koulakov and Chklovskii [52]

[8] For an introduction to simulated annealing, see, for example, the Bourbaki lecture by Robert Azencott [53].

4.4.5.2 Column Structure and Sparse Representations

For their part, Karol Gregor, Arthur Szlam, and Yann LeCun [54] related the column and hypercolumn structure of $V1$ to the sparse code representation of statistical regularities in natural images. They used the model we discussed in Sect. 3.6.2 of Chap. 3. Natural images I with $P = p^2$ pixels form a subset $\Sigma \subset \mathbb{R}^P$ and are thus described by vectors $I = (I_r)_{r=1,\dots,P}$ in the canonical basis of $\mathbb{R}^P$. We seek representations $I = \sum_{i=1}^{i=N} s_i \varphi_i$ using filters (RPs of neurons) $\varphi_i \in \mathbb{R}^P, i = 1, \dots, N$. The latter constitute a 'dictionary' Φ which is a $P \times N$ matrix and we wish these representations to be sparse; that is, for each image I, only a limited number of filters respond, most of the s_i being zero. As we have seen, the basic idea is to minimize an energy of the form

$$E(I, s) = \frac{1}{2} \|I - \Phi s\|^2 + \lambda \sum_{i=1}^{i=N} |s_i| \ .$$

But we can also try to *structure* the lattice of filters by introducing, rather like Hyvärinen and Hoyer [55], a set of lateral neural connections between the filters and by penalizing the simultaneous activity of two neurons by means of *inhibitory connections*. If $U = \{(i_1, j_1), \dots, (i_k, j_k)\}$ is the set of these connections, we cannot have both $s_i \neq 0$ and $s_j \neq 0$ in the coding of an image I when $(i, j) \in U$.

In order to impose this constraint, we introduce a quadratic term $s^{\mathrm{T}} W s$ into the energy E, where W is a matrix of weights for the inhibitory connections. If U is given, learning will reinforce W. However, if U has to be learnt by exposure to natural images, then learning will weaken the weights of the connections (i, j) between filters that are often activated together. The authors of [54] show that the lattice of filters gets structured into columns and hypercolumns. We shall return to this point in the second volume.

4.4.6 End Points and Triple Points

All the pinwheels observed experimentally have topological index ± 1. However, singularities with topological index $\pm 1/2$ enter the scene when we look at the way the orientation itself varies. Figure 4.51 was produced by Shmuel. We have included the orientation field lines in the neighbourhood of two singularities of opposite chirality.

We see that dextrorotatory and levorotatory pinwheels are associated with the two types of generic singularity in the orientation fields in the plane. This is due to the fact that, when the ray rotates through an angle θ about the centre of the pinwheel, the associated orientation rotates through $\theta/2$. Hence, two diametrically opposite rays correspond to orthogonal orientations.

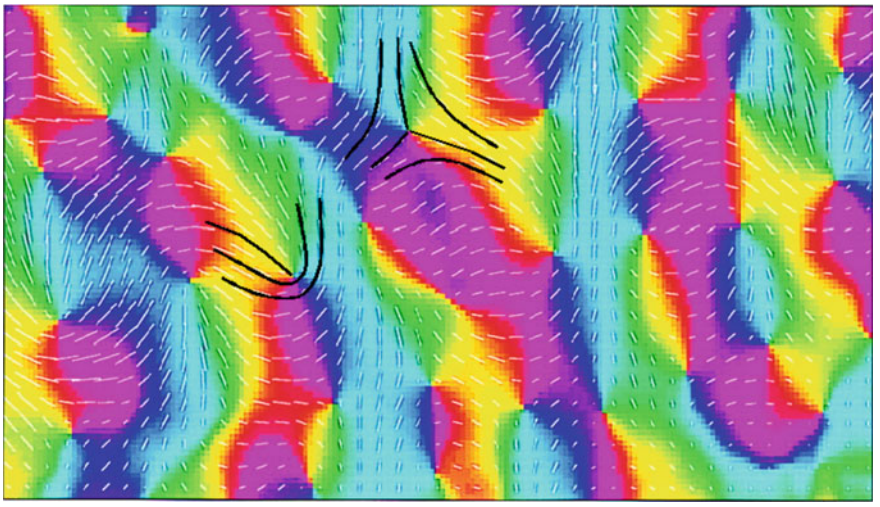


Fig. 4.51 Orientation maps and pinwheels in the $V1$ area of a tree shrew. Here, we see the relation between pinwheels (*colours*) and the preferred orientations. We have represented the orientation field lines in the neighbourhood of two singularities with opposite chirality. Adapted from Shmuel [56]. Copyright (2000) National Academy of Sciences, USA

If the orientation ψ_θ associated with the ray at angle θ is

$$\psi_\theta^+ = \alpha + \frac{\theta}{2} = \frac{\varphi_\theta^+}{2} ,$$

the two directions will be the same for $\psi_{\theta_0}^+ = \alpha + \theta_0/2 = \theta_0$, i.e. for $\theta_0 = 2\alpha$, and then $\varphi_\theta^+ = 4\alpha$. As α is specified modulo π, there is only one solution and we obtain the local model in Fig. 4.52 (an end point). If on the other hand the orientation ψ_θ associated with the ray at angle θ is

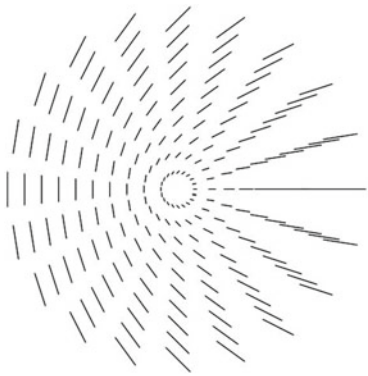


Fig. 4.52 End point singularity

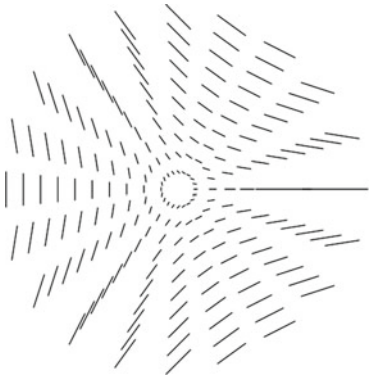


Fig. 4.53 Triple point singularity

$$\psi_\theta^- = \alpha - \frac{\theta}{2} = \frac{\varphi_\theta^-}{2} ,$$

the two directions will be the same for $\psi_{\theta_0}^- = \alpha - \theta_0/2 = \theta_0$, i.e., for $\theta_0 = 2\alpha/3$, and then $\varphi_\theta^+ = 4\alpha/3$. There are three solutions, and we obtain the local model shown in Fig. 4.53 (triple point).

It is very easy to calculate the integral curves of the orientation field in the neighbourhood of these singularities, for example for $\alpha = 0$. Adopting polar coordinates (ρ, θ), we have $x = \rho \cos(\theta)$ and $y = \rho \sin(\theta)$, whence

$$\begin{cases} \mathrm{d}x = \cos(\theta)\,\mathrm{d}\rho - \rho \sin(\theta)\,\mathrm{d}\theta , \\ \mathrm{d}y = \sin(\theta)\,\mathrm{d}\rho + \rho \cos(\theta)\,\mathrm{d}\theta . \end{cases}$$

For end points, the constraint is $\mathrm{d}y/\mathrm{d}x = \tan(\theta/2)$, i.e.

$$\sin\left(\frac{\theta}{2}\right)\mathrm{d}x - \cos\left(\frac{\theta}{2}\right)\mathrm{d}y = 0 ,$$

or

$$\begin{aligned} \sin\left(\frac{\theta}{2}\right)\left[\cos(\theta)\,\mathrm{d}\rho - \rho \sin(\theta)\,\mathrm{d}\theta\right] - \cos\left(\frac{\theta}{2}\right)\left[\sin(\theta)\,\mathrm{d}\rho + \rho \cos(\theta)\,\mathrm{d}\theta\right] \\ = \left[\sin\left(\frac{\theta}{2}\right)\cos(\theta) - \cos\left(\frac{\theta}{2}\right)\sin(\theta)\right]\mathrm{d}\rho \\ -r\left[\sin\left(\frac{\theta}{2}\right)\sin(\theta) + \cos\left(\frac{\theta}{2}\right)\cos(\theta)\right]\mathrm{d}\theta = 0 . \end{aligned}$$

But

$$\cos(\theta) = \cos^2\left(\frac{\theta}{2}\right) - \sin^2\left(\frac{\theta}{2}\right) , \quad \sin(\theta) = 2\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) ,$$

and

$$\cos^2\left(\frac{\theta}{2}\right) + \sin^2\left(\frac{\theta}{2}\right) = 1 ,$$

so we obtain the differential equation

$$\sin\left(\frac{\theta}{2}\right) d\rho + \rho\cos\left(\frac{\theta}{2}\right) d\theta = 0 ,$$

and hence,

$$\sin\left(\frac{\theta}{2}\right)\frac{d\rho}{d\theta} + \rho\cos\left(\frac{\theta}{2}\right) = 0 .$$

The solutions are

$$\rho = \frac{\rho_\pi}{\sin^2(\theta/2)} = 2\frac{\rho_\pi}{1-\cos(\theta)} ,$$

where the constant of integration ρ_π is the value $\rho(\pi)$. This is a parabola. Its axis of symmetry is the x-axis, it has vertical tangent at $(\rho = \rho_\pi, \theta = \pi)$, and its branches tend to infinity for $\theta \to 0$. Figure 4.54 shows the case $\rho_\pi = 1$.

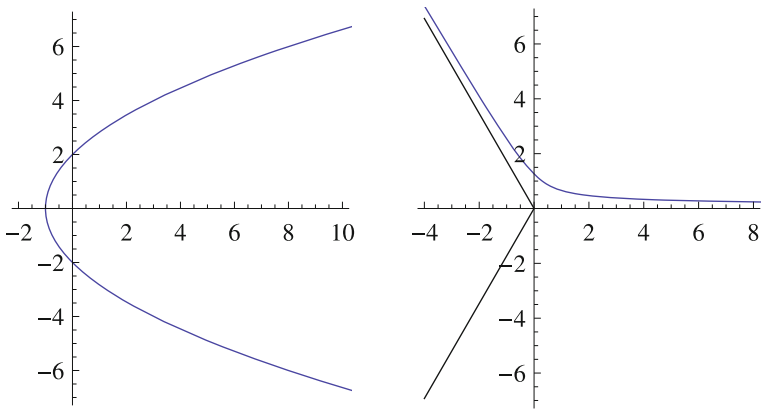


Fig. 4.54 Integral curves of the orientation field in the neighbourhood of the pinwheels. *Left* End point singularity. *Right* Triple point singularity

For the triple points, the constraint is $\mathrm{d}y/\mathrm{d}x = -\tan(\theta/2)$ and an analogous calculation leads to

$$\sin\left(\frac{\theta}{2}\right)\left[3\cos^2\left(\frac{\theta}{2}\right) - \sin^2\left(\frac{\theta}{2}\right)\right]\mathrm{d}\rho + \rho\cos\left(\frac{\theta}{2}\right)\left[3\sin^2\left(\frac{\theta}{2}\right) - \cos^2\left(\frac{\theta}{2}\right)\right]\mathrm{d}\theta = 0\,,$$

which has solutions

$$\rho = \frac{\rho_\pi}{\sin^{2/3}(3\theta/2)}\,.$$

The symmetry $\theta \to \theta + 2\pi/3$ is obvious, since it changes $3\theta/2$ to $3\theta/2 + \pi$ and thus leaves $\sin^{2/3}(3\theta/2)$ invariant. Figure 4.54 shows one of the trajectories in the case $\rho_\pi = 1$.

When $\alpha \neq 0$, the solutions become

$$\rho = \frac{\rho_{\pi+2\alpha}}{\sin^2\left(\frac{\theta}{2} - \alpha\right)}\,, \qquad \rho = \frac{\rho_{(\pi+2\alpha)/3}}{\sin^{2/3}\left(\frac{3\theta}{2} - \alpha\right)}\,,$$

respectively, and the orientation fields rotate through α. However, it should be noted that the fields of the $\varphi = 2\psi$ depend in a more subtle way on α. The singularity of the field φ^+ varies from a node with central symmetry ($\alpha = 0$), first to stable foci, then to a centre ($\alpha = \pi/4$), then to unstable foci, whereas the singularity of the field φ^- gives different kinds of saddle point. Figure 4.55 (obtained using Mathematica's StreamPlot) shows the field lines of φ^+, φ^-, ψ^+, and ψ^- for four values of the parameter α: 0, $\pi/8$, $\pi/4$, and $3\pi/8$. These different prototype models of field lines all correspond to a prototype pinwheel like the one in Fig. 4.56.

4.4.7 *Distortions and Defects in the Neighbourhood of the V1/V2 Boundary*

Given the chirality of the pinwheels, the most standard $V1$ structure is a distortion of a square lattice (the 'crystal' structure in Fig. 4.35). However, like any distortion of a regular lattice, it may contain defects inducing, for example triangular or pentagonal domains (see Fig. 4.57).

The distortion becomes significant at the boundary between $V1$ and $V2$ because, as has been shown by Ohki et al. [57] among others, pinwheels of the same chirality are aligned along this boundary (see Fig. 4.58). Their field lines and level sets can be studied and simulated in detail, as shown in Fig. 4.59. Distortion is manifested by the fact that the $V1/V2$ boundary resembles a 'cliff', as can be seen from the tightly bunched parallel level sets.

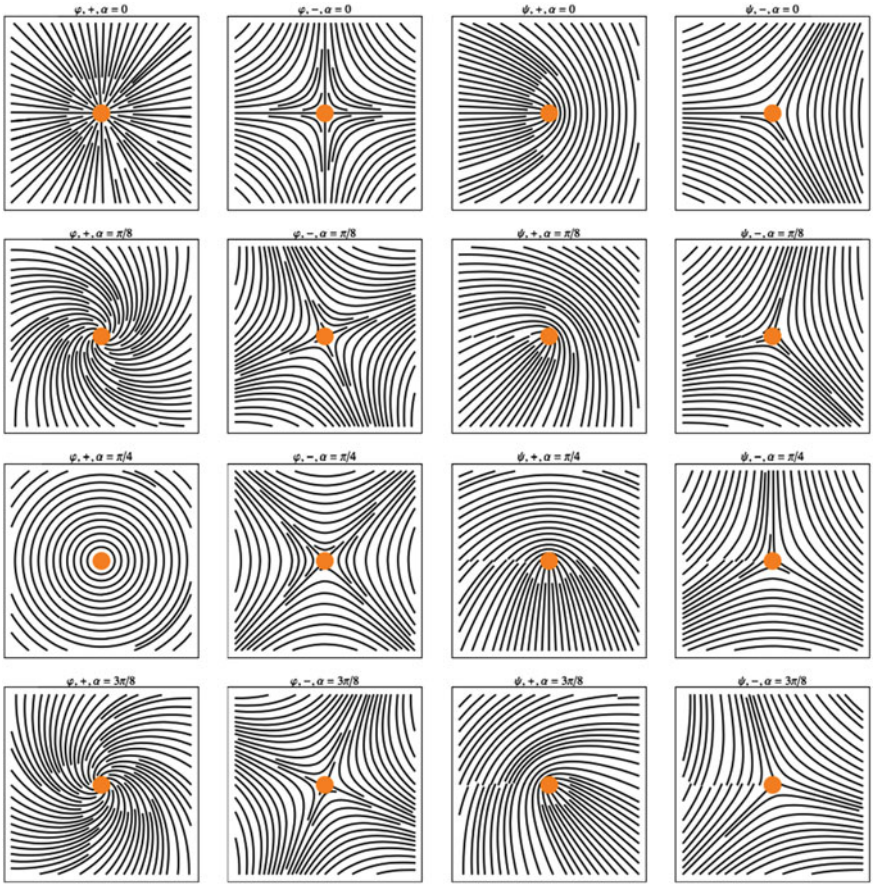


Fig. 4.55 *Left to right columns* Field lines of the fields φ^+, φ^-, ψ^+, ψ^- (see text) for the 4 values (*top to bottom*) 0, $\pi/8$, $\pi/4$, and $3\pi/8$ of the parameter α

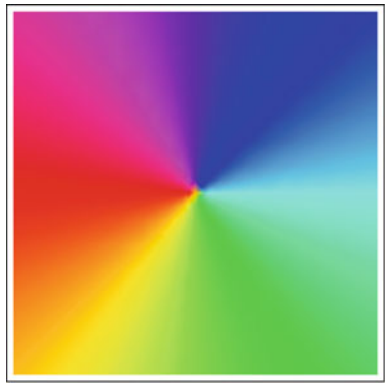


Fig. 4.56 Typical pinwheel with $\psi_\theta = \theta/2$

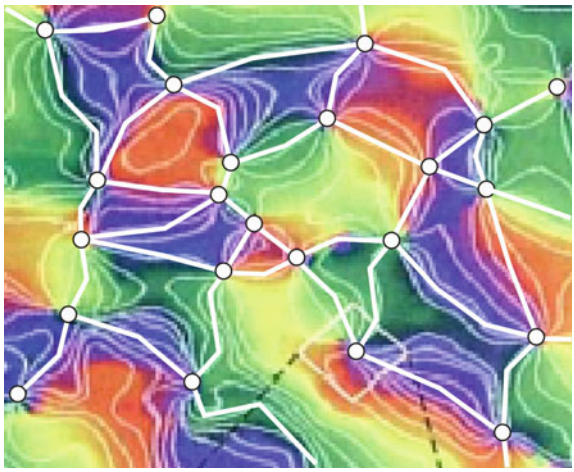


Fig. 4.57 Part of Fig. 4.36 showing the distortion relative to the 'crystal' lattice in Fig. 4.35 and the appearance of defects in the ordered structure. We have not indicated the connections to pinwheels outside the frame

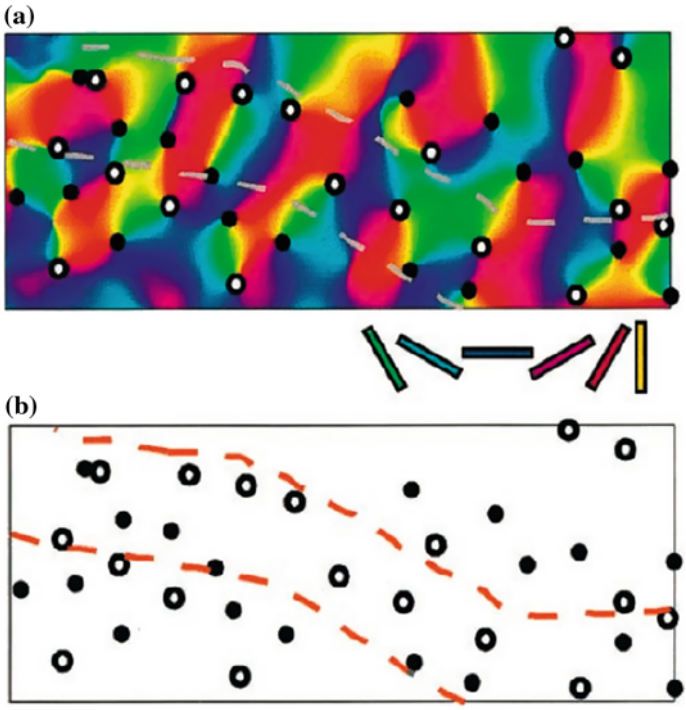


Fig. 4.58 Arrangement of pinwheels near the boundary between $V1$ and $V2$. It is highly distorted because it includes alignments of several pinwheels with the same chirality. *Small white and black discs* represent singularities with levorotatory and dextrorotatory chiralities, respectively. From Ohki et al. [57]

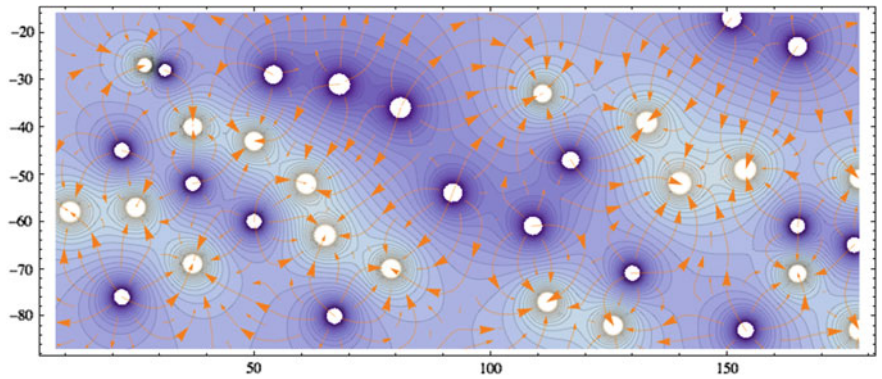


Fig. 4.59 Physical field model of the distortion of the pinwheel lattice at the $V1/V2$ boundary. This boundary resembles a 'cliff', as can be seen from the tightly spaced bundles of parallel level sets

4.5 Topological Universality of Pinwheels

We shall see in Sect. 4.9.4 that the pinwheel structure is widespread in visual systems. Note, however, that this concrete empirical omnipresence is matched by an abstract *topological universality* which helps to explain it. In a very elegant paper [58], Daniel Bennequin and coworkers Alberto Romagnoni, Jérôme Ribot, and Jonathan Touboul provided a topological demonstration. The point of this theorem is to show just how far apparently very general and barely restrictive conditions can effectively amount to drastic constraints. It is one of the great advantages of mathematics to be able to identify this kind of 'Platonic' constraint.

The idea is to consider the way cells parametrized by $z = (x, y)$ in a domain $\Omega \subset \mathbb{C}$ of the visual plane identified with the plane $\mathbb{C} = \mathbb{R}^2$ can encode orientations (modulo π) $\psi \in \mathbb{S}^1_\pi$ (where $\mathbb{S}^1_\pi$ is $[0, \pi]$, with π identified with 0), i.e., a *periodic quantity*. Such a coding can be described by a mapping $f : \Omega \to \mathbb{S}^1_\pi$ which may have singularities, i.e. points where it is not defined because multivalued there.

Since the idea is to model hypercolumn functional architectures, we assume that f is *local*, i.e. defined apart from a few possible singular points on a domain Ω which is a small disc centred on 0 and having a circle as boundary $\Gamma = \partial\Omega$. But at the same time, since the notion of size is relative to a *scale*, we seek local models which would ideally be *scale invariant*. This means that these are models centred on a singularity and with a qualitative structure that does not vary when we zoom in or out.

The authors begin by defining a rather general class $\mathscr{F}$ of sufficiently smooth maps f satisfying natural conditions. They then assume that f is exhaustive, in the sense that it codes all orientations, i.e. it is *surjective*. Clearly, the standard model of the pinwheel, viz.

$$f_0 : \Delta \to \mathbb{S}^1_\pi \ , \qquad z = \rho e^{i\theta} \longmapsto \psi = \frac{\theta}{2} \ ,$$

defined over the unit disc Δ, must belong to $\mathscr{F}$. Now, f_0 has several properties:

(i) f_0 is defined away from 0, with 0 being a singularity at which it is multivalued (the 'image' of 0 is the whole of $\mathbb{S}^1_\pi$).
(ii) Away from 0, f_0 is regular (smooth, infinitely differentiable).
(iii) The level sets $C_\psi = f_0^{-1}(\psi)$ on which f_0 is constant are rays and hence smooth curves joining the boundary Γ to the singular point 0.

The fact that the level sets C_ψ are curves is to be expected. Indeed, suppose that $f : \Omega \to \mathbb{S}^1_\pi$ is smooth (apart from singularities) and *generic*, i.e. with a topology that remains qualitatively the same when f is slightly deformed. Then, the level sets C_ψ must be sub-manifolds of Ω with codimension $c = \dim\left(\mathbb{S}^1_\pi\right) = 1$ so that there exists in Ω a degree of freedom transverse to the C_ψ, corresponding to changes in ψ. Hence, since Ω has dimension $n = 2$, the level sets C_ψ must be sub-manifolds of dimension $n - c = 2 - 1 = 1$.[9] As curves, the C_ψ can have singularities (which are not necessarily singularities of f, where f is not defined). Generically, these are end points or triple points.

The authors then introduce a *minimal complexity* constraint known as 'parsimony', which stipulates that the topological redundancy, that is the number of connected components of the level sets C_ψ, should be *minimal*, i.e., equal to unity (as for the rays of the standard pinwheel). The class $\mathscr{F}$ defined in this way is thus the class of maps $f : \Omega \to \mathbb{S}^1_\pi$ which are smooth away from singularities in the interior of Ω, surjective (exhaustivity constraint), and whose level sets C_ψ are smooth connected curves (parsimony constraint). The authors then prove the following theorem:

Theorem *Topological universality of pinwheels. The elements $f : \Omega \to \mathbb{S}^1_\pi$ of $\mathscr{F}$ have the pinwheel topology, i.e., an isolated singularity in the interior of Ω with the C_ψ joining the different points of the boundary $\Gamma = \partial\Omega$ to this singularity.*

The proof is based on purely topological arguments. Here, we give a heuristic account. We use a fundamental topological property of any circle, e.g. the standard circle $\mathbb{S}^1$: if we remove any point or interval, it will not be disconnected. So if we remove an arc $J = [\theta_1, \theta_2]$ containing $(\theta_1 + \theta_2)/2$, we can still move from θ_1 to θ_2 by following the other arc containing $\pi + (\theta_1 + \theta_2)/2$ [the point opposite to $(\theta_1 + \theta_2)/2$ on $\mathbb{S}^1$]. This is not the case with $\mathbb{R} = (-\infty, +\infty)$: if we remove an interval $I = [x_1, x_2]$, we disconnect $\mathbb{R}$ into two disjoint intervals $(-\infty, x_1)$ and $(x_2, +\infty)$ (see Fig. 4.60).

Let $\psi_0 \in f(\Gamma)$ be an orientation coded by a point on the boundary Γ. By hypothesis, $C_{\psi_0} = C$ is connected and intersects Γ. Now, consider the complement $\Omega - C$ of C in Ω and let c be the number of its connected components (connected by arcs[10]).

[9] For more on such questions of dimensions, the reader could consult my 1982 review [59] and references therein.

[10] Given the assumptions of continuity and smoothness, we need consider only connectedness by arcs.

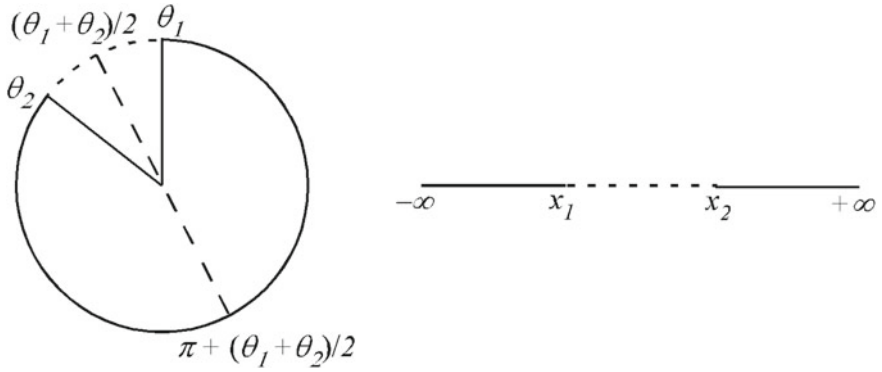


Fig. 4.60 A circle remains connected if we remove an arc $[\theta_1, \theta_2]$, whereas a straight line becomes disconnected if we remove an interval $[x_1, x_2]$

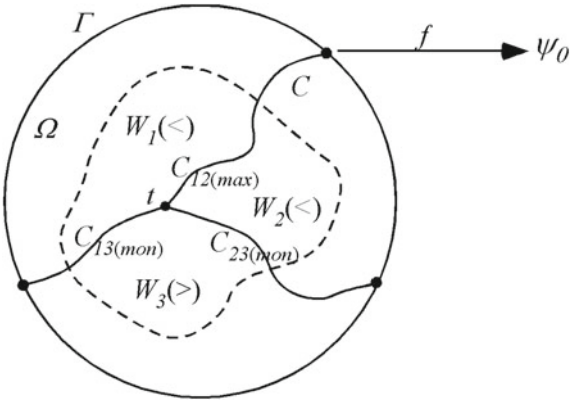


Fig. 4.61 Level curves of f cannot have triple points decomposing Ω into three connected components. Symbols $<$ and $>$ indicate $f < \psi_0$ or $f > \psi_0$, respectively

It can be shown that we cannot have $c \geq 3$ (see Fig. 4.61). Indeed, suppose that C has a triple point t and consider a small enough open neighbourhood T of t and a small enough open neighbourhood V of ψ_0. Let $W = f^{-1}(V) \cap T$. Then, W is made up of points close to t, where f takes values close to ψ_0. Let W_i, $i = 1, 2, 3$, be the three connected components of $W - C$. We can then arrange for f to be $> \psi_0$ or $< \psi_0$ in each W_i. As this makes two possibilities for three components, there must therefore be two components with the same property. Suppose, for example, that $f < \psi_0$ on W_1 and W_2 and $f > \psi_0$ on W_3. This means that (i) the part C_{12} of C which is the shared boundary of W_1 and W_2 is a curve along which f is transversally maximal, and (ii) the parts C_{23} and C_{31} of C which are shared boundaries of W_2 and W_3 and of W_3 and W_1 are curves along which f is transversally monotonic. Figure 4.62 shows that, in this case, the level lines C_ψ for $\psi \neq \psi_0$, $\psi \in V$ are no longer connected, which is forbidden by parsimony.

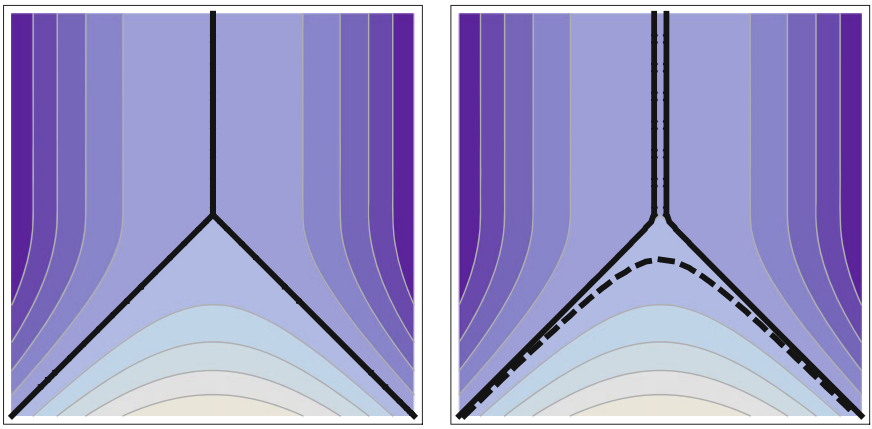


Fig. 4.62 *Left* Structure of level lines in the neighbourhood of a triple point of C. *Right* A level line close to C (*continuous line*) which is disconnected and another level line close to C (*dashed line*) which is connected

But neither is it possible to have $c = 2$, $\Omega - C$ being decomposed into two disjoint connected components U_1 and U_2 on which we would have $f > \psi_0$ and $f < \psi_0$, respectively (see Fig. 4.63). Let $I_1 \subset \mathbb{S}^1_\pi$ be the values taken by f on $U_1 \cup C$. Since $U_1 \cup C$ is compact and connected and f is continuous, the image I_1 is compact and connected[11] and is thus an interval $I_1 = [\psi_0, \psi_1]$ (by continuity, we can have $\psi_1 > \pi$). Likewise, the image of $U_2 \cup C$ is an interval $I_2 = [\psi_2, \psi_0]$ (by continuity, we can have $\psi_2 < 0$). If $\psi_1 - \psi_2 < \pi$, then part of $\mathbb{S}^1_\pi$ is not covered by f, so f is not surjective, a situation excluded by hypothesis. If, on the other hand, $\psi_1 - \psi_2 \geq \pi$, then there is at least one value $\psi \neq \psi_0$ for which the level curve C_ψ intersects both U_1 and U_2, which are disjoint (see Fig. 4.64). These two pieces of C_ψ cannot be connected by C_ψ because, for this to happen, C_ψ would have to cross the boundaries of U_1 and U_2, and hence C_{ψ_0}. But this is impossible because $\psi \neq \psi_0$, whence C_ψ and C_{ψ_0} are disjoint. Therefore, C_ψ is not connected and the topological redundancy is ≥ 2, which is forbidden by hypothesis.

So the connected level line $C = C_{\psi_0}$ does not disconnect Ω. It can be shown that it cannot be entirely contained within the boundary Γ and that it enters into Ω and ends at a singular point s. We then choose two points z_1 and z_2 in $\Omega - C$ and connect then by two paths: one of them $\gamma : z_1 \to z_2$ crossing C and the other $\gamma' : z_1 \to z_2$ going round C, in such a way as to give the situation in Fig. 4.65.

Let $\psi_1 = f(z_1)$ and $\psi_2 = f(z_2)$. Since $z_1, z_2 \notin C$, we have $\psi_1, \psi_2 \neq \psi_0$. Since γ and γ' are compact and connected, their images are intervals $[\psi_1, \psi_2]$ of $\mathbb{S}^1_\pi$. Since γ crosses C, its image is the interval $[\psi_1, \psi_2]$ containing ψ_0, and since γ' goes around C, its image is the interval $[\psi_1, \psi_2]$ not containing ψ_0. Hence, the

[11] Here, the authors use basic theorems of general topology going back to Bolzano, Weierstrass, Heine, Borel, and Lebesgue. For separated topological spaces, the continuous image of a compact (connected) set is compact (connected).

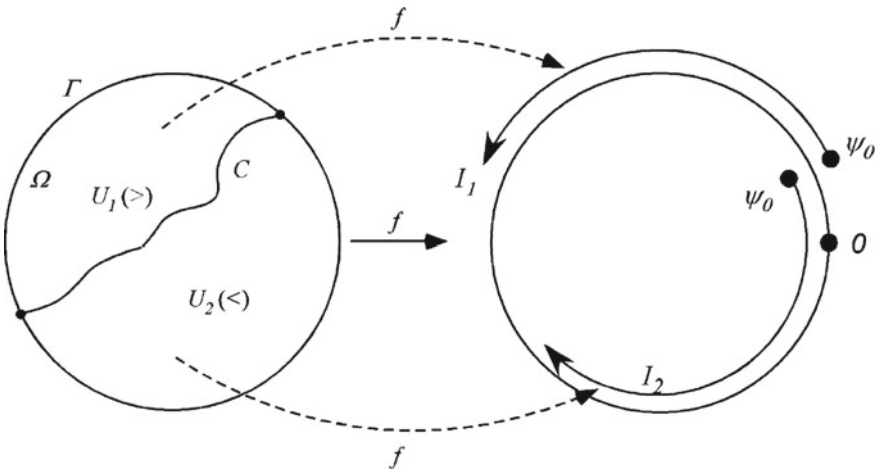


Fig. 4.63 Level curves of f cannot decompose Ω into two connected components. Symbols $<$ and $>$ indicate $f < \psi_0$ or $f > \psi_0$, respectively

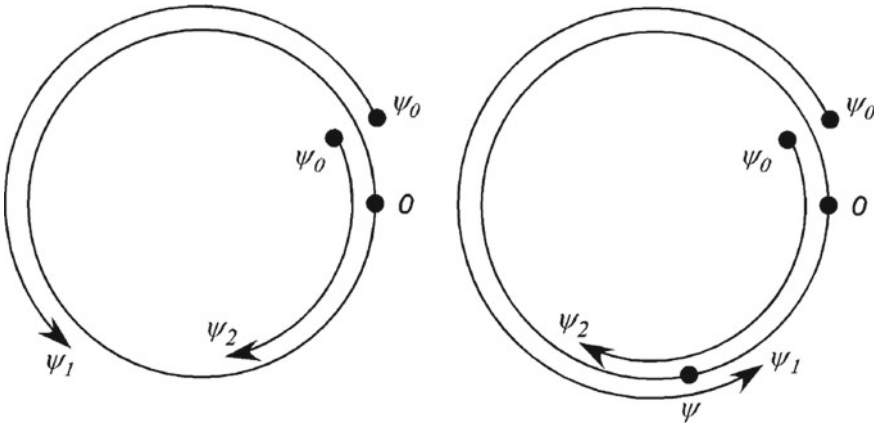


Fig. 4.64 *Left* If $(\psi_1 - \psi_0) - (\psi_2 - \psi_0) = \psi_1 - \psi_2 < \pi$, the interval (ψ_1, ψ_2) (*bottom*) is not covered by f. *Right* If, on the other hand, $(\psi_1 - \psi_0) - (\psi_2 - \psi_0) = \psi_1 - \psi_2 \geq \pi$, the interval (ψ_2, ψ_1) (*bottom*) is covered twice (example given by angle ψ)

image of $\gamma \cup \gamma'$ under f is the whole circle $\mathbb{S}^1_\pi$. But the configuration in Fig. 4.65 can be shrunk towards s. Thus, all the level lines C_ψ must converge towards the point s.

Finally, it can be shown that f is already surjective on the boundary Γ. Indeed, $f\big|_\Gamma$ is a continuous map of a circle onto a circle. If it is not surjective, it then takes each of its values generically at least twice (counted with multiplicity). This means that there are at least two level curves C_ψ for each value ψ in the range, which contradicts the assumption of minimal topological redundancy (equal to unity). Therefore, f is

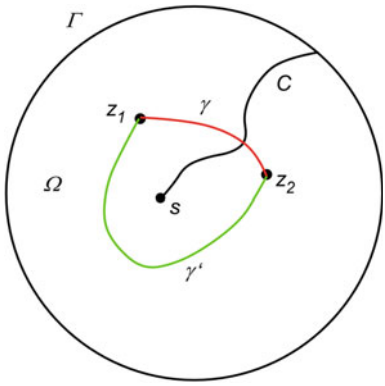


Fig. 4.65 The level line C does not disconnect Ω, enters Ω, and ends at a singular point s. z_1 and z_2 are two points in $\Omega - C$ joined by two paths, one of them $\gamma : z_1 \to z_2$ (*red*) crossing C and the other $\gamma' : z_1 \to z_2$ (*green*) going around C

surjective. In fact, by the same argument of minimal topological redundancy, it is also injective.

4.6 Pinwheels as Phase Fields

Having discussed the topological universality of pinwheels as local models of hypercolumns, we shall now turn to their global lattice structure. On a certain scale, functional orientation maps assign an orientation $\psi\,(a) \bmod \pi$ to each point a of the cortical surface of $V1$. For simplicity, we shall treat this surface as a plane $\mathbb{R}^2$ with coordinates (x, y), identified with a complex plane $\mathbb{C}$ with coordinate z. By treating $V1$ in Sect. 4.3.5 as an implementation of the space of 1-jets of curves in $\mathbb{R}^2$, we represented ψ by its tangent $\tan(\psi)$. It was associated with a function $e^{i\varphi}$, where φ is the *phase* $\varphi = 2\psi$, defined $\bmod 2\pi$.

These maps are thus *phase fields* in which the pinwheels are singularities. This kind of field turns up in many types of physical phenomenon, and in particular in *optical* and *liquid crystal* structures. There is a vast literature about this, and especially about the singularities. In this section, we shall import a certain number of the associated formalisms and models to neurogeometry.

We do this with a certain amount of mathematical detail because we shall not return to this model in the second volume (where we shall in fact develop an alternative model in considerable mathematical detail).

4.6.1 Fields and Coordinates

Since we shall be considering several spaces and several fields, we begin by defining the notation. The cortical layer is modelled by $\mathbb{R}^2$. If $a = (x, y)$ is a point of $\mathbb{R}^2$,

it will often be useful to consider it as a complex number $z = x + \mathrm{i}y$ using the standard $\mathbb{R}$-linear isomorphism between $\mathbb{C}$ and $\mathbb{R}^2$. To study the structure of the field in the neighbourhood of a singular point a_0, it will often be useful to transport the coordinate system on $\mathbb{R}^2$ to a_0 by a translation carrying 0 to a_0 and an appropriate rotation of the axes, and then to take polar coordinates (ρ, θ) around $a_0 = 0$, i.e. $z = \rho \mathrm{e}^{\mathrm{i}\theta}$, $x = \rho \cos(\theta)$, $y = \rho \sin(\theta)$. A phase field assigning the phase $\varphi(a)$ to each point $a \in \mathbb{R}^2$ is thus a map $\Phi : \mathbb{R}^2 \to \mathbb{S}^1$, $\Phi(a) = \mathrm{e}^{\mathrm{i}\varphi(a)}$, in other words a section of the fibre bundle $\pi : \mathbb{R}^2 \times \mathbb{S}^1 \to \mathbb{R}^2$, with singularities of a particular type in places where the phase $\varphi(a)$ *is not defined*. Clearly, as $\varphi(a)$ is a function defined on $\mathbb{R}^2$, if it is differentiable, it can have singularities in the classical sense, i.e. *critical points* where the gradient $\nabla\varphi = 0$. Generically, there are three types: extrema, i.e. maxima or minima, and saddle points. Experimentally, there do not seem to be any extrema (where the isochromatic curves would be concentric circles locally), whereas there are many saddle points occupying the 'centres' of the cells defined by the pinwheels.

The experimental data show that the orientation maps $\psi(a) = \varphi(a)/2$ can also have lines of discontinuity called fractures, across which the orientation jumps from one value to another. If there are no fractures, $\psi(a)$ is smooth away from singular points, where it is not specified. In the local pinwheel models involving end point and triple point singularities, discussed in Sect. 4.4.6, we took $\psi = \alpha \pm \theta/2$.

It is often possible and natural to associate an *amplitude*, that is a modulus $r(a)$, with the phase of a phase field. For functional orientation maps, Fred Wolf and Theo Geisel suggested introducing the strength of the orientation selectivity, i.e. the width of the tuning curve. We shall return to this key idea in Sect. 4.7.1. Under such an assumption, the phase field Φ becomes the 'phase part' of a *complex scalar field* $Z : \mathbb{C} \to \mathbb{C}$, $z = \rho \mathrm{e}^{\mathrm{i}\theta} \mapsto r(z)\, \mathrm{e}^{\mathrm{i}\varphi(z)}$, also denoted $r(a)\, \mathrm{e}^{\mathrm{i}\varphi(a)}$. In Cartesian coordinates, $Z(a)$ will be written $Z(a) = X(a) + \mathrm{i}Y(a)$, where X and Y are two real functions of the variables (x, y).

The map Z can be viewed as a section of the bundle $\pi_1 : \mathbb{R}^2 \times \mathbb{C} \to \mathbb{R}^2$ whose fibre is no longer $\mathbb{S}^1$ but $\mathbb{C}$. The group $E(2)$ acts on π_1 as it acts on π: in the standard way in the base space $\mathbb{R}^2$ and by rotation of the phase in the fibre.

4.6.2 Singularities of a Phase Field

Given a phase field $Z(a)$—which may be denoted $Z(z)$, $X(a) + \mathrm{i}Y(a)$, $X(z) + \mathrm{i}Y(z)$, $r(a)\, \mathrm{e}^{\mathrm{i}\varphi(a)}$, or $r(z)\, \mathrm{e}^{\mathrm{i}\varphi(z)}$, depending on the context—the geometry can be analyzed using the standard tools of vector analysis, viz. gradients ∇, divergences, curls, Laplacians Δ, and so on.

The singularities of Z are points where the phase φ is not defined, but where Z is for its part *well defined*, although with a particular value, 0 or ∞. They correspond to *zeros and poles* of Z. If $|Z|$ is bounded, then they can only be zeros. We go from zeros to poles by allowing $|Z|$ to diverge and taking for the space of values of Z, not only the complex plane $\mathbb{C}$, but its completion $\widehat{\mathbb{C}}$ by a point at infinity (called the

Riemann sphere or the complex projective line). The reciprocal

$$\frac{1}{Z} = \frac{1}{r\,(a)} \mathrm{e}^{-\mathrm{i}\varphi(a)}$$

of Z then has a pole at a if Z has a zero at a. For a zero, the simplest local model is $\pm z$ if Z is a *holomorphic* function, i.e., a smooth function of the *complex* variable z, hence *differentiable with respect to* z (and not only with respect to x and y separately, see below), and for a pole, it is $\pm 1/z$ if Z is a meromorphic function (see below). However, there is no reason a priori why Z should be holomorphic. In the neighbourhood of a point a_0 of the base space $\mathbb{R}^2$ taken as the origin 0, we then have to first order

$$Z\,(x,\,y) \approx X\,(0) + x\frac{\partial X}{\partial x}\,(0) + y\frac{\partial X}{\partial y}\,(0) + \mathrm{i}\left[Y\,(0) + x\frac{\partial Y}{\partial x}\,(0) + y\frac{\partial Y}{\partial y}\,(0)\right] ,$$

whence

$$Z\,(a) \approx Z\,(0) + a \cdot \nabla_0 X + \mathrm{i}a \cdot \nabla_0 Y \ ,$$

where $\nabla_0 X$ is the value of the gradient of X at 0, $\nabla X = (\partial X/\partial x,\, \partial X/\partial y)$, and the same for Y. Hence,

$$\left|Z\,(a) - Z\,(0)\right|^2 \approx R^2 = (a \cdot \nabla_0 X)^2 + (a \cdot \nabla_0 Y)^2 \ ,$$

and the level lines $R = \text{const.}$ are *ellipses* to this order of approximation. They are only *circles* if $Z\,(a)$ can be written as a function $Z\,(z)$ of z, in other words if on the one hand $x\partial X/\partial x + \mathrm{i}y\partial Y/\partial y$ is proportional to z, which requires $\partial X/\partial x = \partial Y/\partial y$, and if on the other hand $y\partial X/\partial y + \mathrm{i}x\partial Y/\partial x$ is proportional to $\mathrm{i}z$, which requires $\partial X/\partial y = -\partial Y/\partial x$. These fundamental conditions, known as the *Cauchy–Riemann equations*, express the fact that the gradients ∇X and ∇Y of Z are *orthogonal*. They characterize holomorphic functions.

Now in all the situations where we analyze fields, the *singularities* play a determining structural role and contain most of the essential morphological information. As noted by Berry [60], the eminent specialist in optics (whether it be geometrical, wave, or quantum optics):

> Each singularity is a window to a deeper theory.

Regarding lines of phase singularities for 3D waves,[12] Berry also stresses that [61, p. 724]:

> Wave vortex lines can be regarded as a skeleton, characterizing and supporting the full structure of the wave.

[12] Phase singularities are generically points in 2D and lines in 3D because they are specified by two conditions and so have codimension 2 (see below).

This is a general 'philosophy' which geometers have clearly established since René Thom (see Sect. 1.3 of the Preface). We shall take this as a starting point, and all the more so in that, for a long time now, we have been following Michael Berry's work in optics very closely, and in particular his work on *caustics*, because these represent one of the main applications of the theory of singularities in fundamental physics. More recent applications of these studies in phase fields are now also used in the theory of pinwheels by Fred Wolf and Theo Geisel (we shall return to this), Bennequin [62], Afgoustidis [63, 64], and Giovanna Citti and Sarti [65].

Many of the concepts used in the theory of singularities, such as genericity, codimension, bifurcations, unfolding, or normal forms, have thus turned out to be highly relevant in neurogeometry. The reader will find much more about this in our surveys [59] and [66] on the theory of singularities and critical phenomena, and in particular, further reading (but see also our *Landmarks* [67]).

Let us assume that the field Z is smooth outside singular points, where the phase φ is indeterminate and where Z vanishes. Since $Z = X + \mathrm{i}Y$, these points are intersections between curves with equations $X = 0$ and $Y = 0$. The condition $X = 0$ corresponds to $r\cos(\varphi) = 0$, i.e. $\varphi = \pi/2 \bmod \pi$ if $r \neq 0$, and $Y = 0$ corresponds to $r\sin(\varphi) = 0$, i.e. $\varphi = 0 \bmod \pi$ if $r \neq 0$. If $X = Y = 0$, we necessarily have $r = 0$ because the two conditions on φ are incompatible. Generically, the curves $X = 0$ and $Y = 0$ cross transversally at isolated points. This means that the points which satisfy both conditions have codimension 2 and, as the surrounding space $\mathbb{R}^2$ is 2D, are isolated points (whereas in a 3D surrounding space, they would be lines). In the theory of mesophases (liquid crystals), these phase singularities are called *dislocations*. The same term could be used here to say that the pinwheels are dislocations in the orientation field implemented in $V1$.

Note that these dislocations $Z = 0$ are *invariant* under gauge transformation $Z \to \mathrm{e}^{\mathrm{i}\sigma} Z$ in the target space and under change of differentiable coordinates $(x, y) \to (\xi, \eta)$ in the source space.

4.6.3 Orientation and Iso-orientation Fields

There are several different fields here. The field Z is the phase field $\varphi(a) = 2\psi(a)$. It has *isophase lines*, called wave fronts by analogy with optics. As isophase means iso-orientation, they are represented by isochromatic lines in the pinwheel maps.

In addition, there is the orientation field $\psi(a) = \varphi(a)/2$. As such, it defines a foliation of the plane $\mathbb{R}^2$ by its integral curves. With the end point and triple point models in Sect. 4.4.6, we identified the local geometry of these foliations at the singular points. Let us denote this field by $W(a) = s(a)\,\mathrm{e}^{\mathrm{i}\psi(a)}$, assuming that meaning can be attributed to the amplitude $s(a)$. W also has field lines and isophase lines. However, in contrast to the phase φ, which is defined modulo 2π, the angle ψ is only defined modulo π. For pinwheels without distortion, φ can in fact be identified with $2\alpha \pm \theta$ and ψ with $\alpha \pm \theta/2$. In this case, we have $\mathrm{e}^{\mathrm{i}\varphi} = \left(\mathrm{e}^{\mathrm{i}\psi}\right)^2$. The simplest

way to satisfy this condition is to take $Z = W^2$. Since $W = \sqrt{Z}$, there is a Riemann cut, because when φ changes by 2π, which leaves Z invariant, ψ changes by π and W changes sign, the square root function changing determination.

4.6.4 Topological Charge and Index

Let a_0 be a singularity of Z taken as origin. The *topological charge* of this singularity is defined to be $q = \oint_\gamma \mathrm{d}\varphi/2\pi$, where $\oint_\gamma \mathrm{d}\varphi$ denotes the integral of the differential 1-form $\mathrm{d}\varphi$, i.e. the change in φ around a small closed path γ going once around a_0 in the right-handed sense. Since a_0 is an isolated singularity, there are such paths γ enclosing only this singularity and it can be shown that the integral is independent of the choice of γ. From the Euclidean structure of $\mathbb{R}^2$, the differential 1-form

$$\mathrm{d}\varphi = \frac{\partial \varphi}{\partial x}\mathrm{d}x + \frac{\partial \varphi}{\partial y}\mathrm{d}y$$

corresponds to the vector gradient $\nabla\varphi$ with components $\nabla\varphi = (\partial\varphi/\partial x, \partial\varphi/\partial y)$, and if the differentials are interpreted as in the past as infinitesimal variations, we have $\mathrm{d}\varphi = \nabla\varphi \cdot \mathrm{d}a$ (scalar product) and $\oint_\gamma \mathrm{d}\varphi = \oint_\gamma \nabla\varphi \cdot \mathrm{d}a$ becomes what is known as the circulation of the gradient field $\nabla\varphi$ around the path γ. The topological charge $q = \oint_\gamma \nabla\varphi \cdot \mathrm{d}a/2\pi$ can then be interpreted as the *topological index* of the field $\nabla\varphi$.

For the field Z, the phase φ varies as $\pm\theta$ and the index is ± 1. However, for the field W, the orientation ψ varies as $\pm\theta/2$ and the index is $\pm 1/2$. We already noted this for the pinwheels in Sects. 4.4.3 and 4.4.6.

In the vicinity of a pinwheel, the isophase lines $\varphi = \text{const.}$, i.e. the wave fronts, are rays of the pinwheel. Along these wave fronts, $\mathrm{d}\varphi = 0$ and hence $\nabla\varphi \cdot \mathrm{d}a = 0$, which means that the field $\nabla\varphi$ is orthogonal to the rays and that its trajectories are therefore locally qualitatively like concentric circles centred on the singular point. Quite generally, the trajectories of $\nabla\varphi$ are orthogonal to the wave fronts.

4.6.5 Current, Vorticity, and Divergence

To understand the behaviour of $\nabla\varphi$ in the neighbourhood of a singularity, the best thing is to use polar coordinates. Since $x = \rho\cos(\theta)$ and $y = \rho\sin(\theta)$, differentiating we obtain

$$\begin{cases} \mathrm{d}x &= \cos(\theta)\,\mathrm{d}\rho - \rho\sin(\theta)\,\mathrm{d}\theta\ , \\ \mathrm{d}y &= \sin(\theta)\,\mathrm{d}\rho + \rho\cos(\theta)\,\mathrm{d}\theta\ , \\ \mathrm{d}\rho &= \cos(\theta)\,\mathrm{d}x + \sin(\theta)\,\mathrm{d}y\ , \\ \rho\mathrm{d}\theta &= -\sin(\theta)\,\mathrm{d}x + \cos(\theta)\,\mathrm{d}y\ . \end{cases}$$

Then, since

$$\mathrm{d}\varphi = \frac{\partial \varphi}{\partial x}\mathrm{d}x + \frac{\partial \varphi}{\partial y}\mathrm{d}y \ ,$$

we have

$$\begin{cases} \dfrac{\partial \varphi}{\partial x} = \cos(\theta)\,\dfrac{\partial \varphi}{\partial \rho} - \dfrac{\sin(\theta)}{\rho}\dfrac{\partial \varphi}{\partial \theta} \ , \\ \dfrac{\partial \varphi}{\partial y} = \sin(\theta)\,\dfrac{\partial \varphi}{\partial \rho} + \dfrac{\cos(\theta)}{\rho}\dfrac{\partial \varphi}{\partial \theta} \ . \end{cases}$$

At the singular point $\rho = 0$, the gradient $\nabla\varphi$ is not defined and diverges.

To regularize this situation, physicists usually consider the *current* $\mathscr{J}$ of the field, which is the vector pointing in the direction of the gradient $\nabla\varphi$ when it does not vanish, defined by

$$\mathscr{J} = r^2 \nabla\varphi \ .$$

Note that if $Z = X + \mathrm{i}Y$, then

$$\mathscr{J} = X\nabla Y - Y\nabla X \ ,$$

and hence that $\mathscr{J}$ is well defined, even at singular points of the phase φ of Z. Indeed, since $X = r\cos(\varphi)$ and $Y = r\sin(\varphi)$,

$$\begin{cases} \nabla X = \left[\cos(\varphi)\,\dfrac{\partial r}{\partial x} - r\sin(\varphi)\,\dfrac{\partial \varphi}{\partial x}, \cos(\varphi)\,\dfrac{\partial r}{\partial y} - r\sin(\varphi)\,\dfrac{\partial \varphi}{\partial y}\right] , \\ \nabla Y = \left[\sin(\varphi)\,\dfrac{\partial r}{\partial x} + r\cos(\varphi)\,\dfrac{\partial \varphi}{\partial x}, \sin(\varphi)\,\dfrac{\partial r}{\partial y} + r\cos(\varphi)\,\dfrac{\partial \varphi}{\partial y}\right] , \\ X\nabla Y = \left[r\cos(\varphi)\sin(\varphi)\,\dfrac{\partial r}{\partial x} + r^2\cos^2(\varphi)\,\dfrac{\partial \varphi}{\partial x}, \right. \\ \qquad\qquad \left. r\cos(\varphi)\sin(\varphi)\,\dfrac{\partial r}{\partial y} + r^2\cos^2(\varphi)\,\dfrac{\partial \varphi}{\partial y}\right] , \\ Y\nabla X = \left[r\sin(\varphi)\cos(\varphi)\,\dfrac{\partial r}{\partial x} - r^2\sin^2(\varphi)\,\dfrac{\partial \varphi}{\partial x}, \right. \\ \qquad\qquad \left. r\sin(\varphi)\cos(\varphi)\,\dfrac{\partial r}{\partial y} - r^2\sin^2(\varphi)\,\dfrac{\partial \varphi}{\partial y}\right] , \\ X\nabla Y - Y\nabla X = \left[r^2\dfrac{\partial \varphi}{\partial x}, r^2\dfrac{\partial \varphi}{\partial y}\right] = r^2\nabla\varphi \ . \end{cases}$$

Note also that, in terms of the complex conjugate values Z and $\overline{Z}$, the current $\mathscr{J}$ can be written

$$\mathscr{J} = \mathrm{Im}\left(\overline{Z}\nabla Z\right) \ .$$

Indeed,

$$\overline{Z}\nabla Z = (X - \mathrm{i}Y)\,(\nabla X + \mathrm{i}\nabla Y) = X\nabla X + Y\nabla Y + \mathrm{i}\,(X\nabla Y - Y\nabla X) \ .$$

We shall see an example of a current in Sect. 4.6.7.

Another vector, in fact a pseudovector, used by physicists is the *vorticity* Ω of the current $\mathscr{J}$, i.e. its *curl*, up to a factor. By definition,

$$\Omega = \frac{1}{2}\nabla \times \mathscr{J} = \nabla X \times \nabla Y \ ,$$

where the symbol $\times$ stands for the *exterior product* of two vectors in $\mathbb{R}^2$. If

$$u = \left(u_x, u_y\right) = u_x e_x + u_y e_y \ , \quad v = \left(v_x, v_y\right) = v_x e_x + v_y e_y \ ,$$

are two vectors in $\mathbb{R}^2$, where e_x and e_y are unit vectors associated with the x- and y-axis, the exterior product $u \times v$ is a vector of magnitude

$$\det \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = u_x v_y - u_y v_x = \omega$$

along an axis orthogonal to $\mathbb{R}^2$, with unit vector e_3, such that the frame $\left\{e_x, e_y, e_3\right\}$ is right-handed. It is the area of the parallelogram constructed from u and v and oriented normal to the plane they lie in. For $\nabla X \times \nabla Y$,

$$\omega = \det \begin{pmatrix} \partial X/\partial x & \partial X/\partial y \\ \partial Y/\partial x & \partial Y/\partial y \end{pmatrix}$$

is the determinant of the Jacobian[13] of Z considered as a map from $\mathbb{R}^2$ into $\mathbb{R}^2$. In fact, the right interpretation of $\Omega = \omega e_3$ is as a differential 2-form. It can be checked that $\Omega = \nabla \times \mathscr{J}/2$, the components $\mathscr{J}_x$ and $\mathscr{J}_y$ of $\mathscr{J}$ being

$$\mathscr{J}_x = X\frac{\partial Y}{\partial x} - Y\frac{\partial X}{\partial x} \ , \qquad \mathscr{J}_y = X\frac{\partial Y}{\partial y} - Y\frac{\partial X}{\partial y} \ .$$

[13] Named after Carl Gustav Jacob Jacobi.

Indeed,

$$\begin{aligned}\Omega &= \frac{1}{2}\nabla\times\mathscr{J} = \frac{1}{2}\left(\frac{\partial\mathscr{J}_y}{\partial x}-\frac{\partial\mathscr{J}_x}{\partial y}\right)e_3\\ &= \frac{1}{2}\left(\frac{\partial X}{\partial x}\frac{\partial Y}{\partial y}+X\frac{\partial^2 Y}{\partial x\partial y}-\frac{\partial Y}{\partial x}\frac{\partial X}{\partial y}-Y\frac{\partial^2 X}{\partial y\partial x}-\frac{\partial X}{\partial y}\frac{\partial Y}{\partial x}-X\frac{\partial^2 Y}{\partial x\partial y}\right.\\ &\qquad\left.+\frac{\partial Y}{\partial y}\frac{\partial X}{\partial x}+Y\frac{\partial^2 X}{\partial y\partial x}\right)e_3\\ &= \det\begin{pmatrix}\partial X/\partial x & \partial X/\partial y\\ \partial Y/\partial x & \partial Y/\partial y\end{pmatrix}e_3=\omega e_3\ .\end{aligned}$$

Note that when $\Omega = 0$, either $\nabla X = 0$ or $\nabla Y = 0$ (isolated points with codimension 2), or the real gradients ∇X and ∇Y have the same orientation (lines of codimension 1), whence $\nabla Y = \alpha\nabla X$ for $\alpha\in\mathbb{R}$. The complex gradient $\nabla Z = \nabla X + \mathrm{i}\nabla Y \in \mathbb{R}^2\oplus\mathrm{i}\mathbb{R}^2$ is thus the real vector ∇X multiplied by a factor $1+\mathrm{i}\alpha$. The condition $\Omega = 0$ saying that ∇X and ∇Y are parallel is the opposite of the Cauchy–Riemann equations saying that ∇X and ∇Y are orthogonal. We shall also see an example of vorticity in Sect. 4.6.7.

Note also that, in terms of the values of the complex conjugates Z and $\overline{Z}$, the vorticity Ω can be written

$$\Omega = \frac{1}{2}\mathrm{Im}\left(\nabla\overline{Z}\times\nabla Z\right)\ .$$

Indeed,

$$\begin{aligned}\nabla\overline{Z}\times\nabla Z &= \frac{\partial\overline{Z}}{\partial x}\frac{\partial Z}{\partial y}-\frac{\partial\overline{Z}}{\partial y}\frac{\partial Z}{\partial x}\\ &= \left(\frac{\partial X}{\partial x}-\mathrm{i}\frac{\partial Y}{\partial x}\right)\left(\frac{\partial X}{\partial y}+\mathrm{i}\frac{\partial Y}{\partial y}\right)-\left(\frac{\partial X}{\partial y}-\mathrm{i}\frac{\partial Y}{\partial y}\right)\left(\frac{\partial X}{\partial x}+\mathrm{i}\frac{\partial Y}{\partial x}\right)\\ &= 2\mathrm{i}\left(\frac{\partial X}{\partial x}\frac{\partial Y}{\partial y}-\frac{\partial X}{\partial y}\frac{\partial Y}{\partial x}\right)\ .\end{aligned}$$

Points where $\Omega = 0$ are invariant under the gauge transformation $Z\to \mathrm{e}^{\mathrm{i}\sigma}Z$ (multiplication by $\mathrm{e}^{\mathrm{i}\sigma}$ is effectively a rotation in the Z plane so parallel vectors are mapped to parallel vectors) and coordinate changes $(x, y)\to(\xi,\eta)$ [because ω becomes $\omega' = \omega\det(\mathrm{Jac})$, where Jac is the Jacobian of the change of variables and, since by definition $\det(\mathrm{Jac})\neq 0$, $\omega = 0 \Longleftrightarrow \omega' = 0$].

The vorticity of $\mathscr{J}$ is not generally trivial. This is not the case for $\nabla\varphi$ because the curl of a gradient $\nabla\times\nabla f$ is always zero due to the fact that, for any vector u, we have $u\times u = 0$. Away from pinwheels where it is not defined, the field $\nabla\varphi$ is thus curl-free.

The idea of vorticity can be used to resolve a slight difficulty regarding the difference between the topological charge and index of a singularity. For the charge, the right-handed orientation of a small closed path γ about a singular point is the right-handed orientation of $\mathbb{R}^2$. The charges of the field Z (whose trajectories are isophase lines, i.e. wave fronts) are then ± 1. For the index, we often orient the γ with respect to Ω and then the index is always $+1$.

We saw above that, to first order, in the neighbourhood of a point a_0 taken as origin 0, the modulus of Z is given by

$$|Z(a) - Z(0)|^2 = R^2 = (a \cdot \nabla_0 X)^2 + (a \cdot \nabla_0 Y)^2 \ .$$

The current $\mathscr{J}$ is given to first order by the 2-column vector

$$\begin{aligned}
\mathscr{J}(a) = \mathscr{J}(x, y) &\approx \begin{pmatrix}
\left[X(0) + x\frac{\partial X}{\partial x}(0) + y\frac{\partial X}{\partial y}(0)\right]\left[\frac{\partial Y}{\partial x}(0) + x\frac{\partial^2 Y}{\partial x^2}(0) + y\frac{\partial^2 Y}{\partial x \partial y}(0)\right] \\
\quad - \left[Y(0) + x\frac{\partial Y}{\partial x}(0) + y\frac{\partial Y}{\partial y}(0)\right]\left[\frac{\partial X}{\partial x}(0) + x\frac{\partial^2 X}{\partial x^2}(0) + y\frac{\partial^2 X}{\partial x \partial y}(0)\right] \\
\left[X(0) + x\frac{\partial X}{\partial x}(0) + y\frac{\partial X}{\partial y}(0)\right]\left[\frac{\partial Y}{\partial y}(0) + x\frac{\partial^2 Y}{\partial x \partial y}(0) + y\frac{\partial^2 Y}{\partial y^2}(0)\right] \\
\quad - \left[Y(0) + x\frac{\partial Y}{\partial x}(0) + y\frac{\partial Y}{\partial y}(0)\right]\left[\frac{\partial X}{\partial y}(0) + x\frac{\partial^2 X}{\partial x \partial y}(0) + y\frac{\partial^2 X}{\partial y^2}(0)\right]
\end{pmatrix} \\
&= \begin{pmatrix}
X(0)\frac{\partial Y}{\partial x}(0) - Y(0)\frac{\partial X}{\partial x}(0) + x\left[X(0)\frac{\partial^2 Y}{\partial x^2}(0) - Y(0)\frac{\partial^2 X}{\partial x^2}(0)\right] \\
\quad + y\left[X(0)\frac{\partial^2 Y}{\partial x \partial y}(0) - Y(0)\frac{\partial^2 X}{\partial x \partial y}(0) + \frac{\partial X}{\partial y}(0)\frac{\partial Y}{\partial x}(0) - \frac{\partial X}{\partial x}(0)\frac{\partial Y}{\partial y}(0)\right] \\
X(0)\frac{\partial Y}{\partial y}(0) - Y(0)\frac{\partial X}{\partial y}(0) + y\left[X(0)\frac{\partial^2 Y}{\partial y^2}(0) - Y(0)\frac{\partial^2 X}{\partial y^2}(0)\right] \\
\quad + x\left[X(0)\frac{\partial^2 Y}{\partial x \partial y}(0) - Y(0)\frac{\partial^2 X}{\partial x \partial y}(0) - \frac{\partial X}{\partial y}(0)\frac{\partial Y}{\partial x}(0) + \frac{\partial X}{\partial x}(0)\frac{\partial Y}{\partial y}(0)\right]
\end{pmatrix}.
\end{aligned}$$

If $X(0) = 0$ and $Y(0) = 0$ (pinwheel, i.e., dislocation), then these formulas simplify enormously to give

$$\mathscr{J}(x,y) \approx \begin{pmatrix} y\left[\frac{\partial X}{\partial y}(0)\frac{\partial Y}{\partial x}(0) - \frac{\partial X}{\partial x}(0)\frac{\partial Y}{\partial y}(0)\right] \\ x\left[-\frac{\partial X}{\partial y}(0)\frac{\partial Y}{\partial x}(0) + \frac{\partial X}{\partial x}(0)\frac{\partial Y}{\partial y}(0)\right] \end{pmatrix}$$
$$= (a\cdot\nabla_0 X)\nabla_0 Y - (a\cdot\nabla_0 Y)\nabla_0 X = \Omega_0 \times a = \omega_0 \begin{pmatrix} -y \\ x \end{pmatrix} .$$

We can thus evaluate $\left|\mathscr{J}\right| = r^2|\nabla\varphi| \approx |\omega|\rho$ in the vicinity of dislocations where $\omega \neq 0$. However, locally, φ is constant on rays from such a singular point and $\nabla\varphi$ is orthogonal to the rays, and in polar coordinates

$$\nabla\varphi = \frac{\partial\varphi}{\partial\rho}e_\rho + \frac{1}{\rho}\frac{\partial\varphi}{\partial\theta}e_\theta ,$$

where e_ρ is the unit vector along the radius at a and e_θ is the unit vector orthogonal to e_ρ, i.e. e_ρ rotated through $+\pi/2$, so we have

$$\nabla\varphi \approx \frac{1}{\rho}\frac{\partial\varphi}{\partial\theta}e_\theta ,$$

and hence,

$$r^2\left|\frac{\partial\varphi}{\partial\theta}\right| \approx \rho^2|\omega| .$$

This tells us that, whereas r is locally constant on the *ellipses*

$$r^2 = (a\cdot\nabla X)^2 + (a\cdot\nabla Y)^2 ,$$

the quantity $r^2|\partial\varphi/\partial\theta|$ is constant on the *circles* $\rho = \text{const}$. As noted by Mark Richard Dennis, this is a kind of Kepler's law for $r^2|\partial\varphi/\partial\theta|$, which is analogous to the angular momentum [68, p. 41]:

> Equal area vectors of the core anisotropy ellipse [$r^2 = \text{const.}$] are swept out in equal intervals of phase.

Note that the eccentricity of the ellipses measures the anisotropy of the vorticity. As we have seen, there is only isotropy (the ellipses are only circles) if the Cauchy–Riemann equations are satisfied.

4.6.6 Helmholtz Equation

When we discussed the experimental results for pinwheels, we saw that these appear when we superpose maps showing the response to different orientations. We also

saw that the pinwheel lattice has a *characteristic length* (see Sect. 4.4.1). These two empirical facts suggest considering the field Z as a superposition of simpler fields with a characteristic mesh. Moreover, on the mathematical level, any field can be considered through its Fourier transform as a superposition of *plane waves*. The latter are the simplest fields with a characteristic length. They have the form $Ae^{i\kappa\cdot a}$, where A is a complex amplitude $Ee^{i\phi}$ and $\kappa = (\kappa_x, \kappa_y)$ is a vector called the *wave vector*, whose magnitude $k = |\kappa|$, called the *wave number*, is analogous to a momentum and associated with a wavelength $\Lambda = 2\pi/k$ (the smaller the wavelength, the greater the wave number). When they evolve in time, their prototype is $Ae^{i(\kappa\cdot a-\omega t)}$, where ω is an angular frequency associated with a frequency $\nu = \omega/2\pi$ and a period $T = 1/\nu = 2\pi/\omega$.

It is straightforward to check that the plane waves $U = Ae^{i\kappa\cdot a}$ satisfy a fundamental equation called the *Helmholtz equation*, viz. $\Delta U + k^2 U = 0$. Indeed, $\kappa \cdot a = x\kappa_x + y\kappa_y$, whence

$$\Delta U = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} = -A\kappa_x^2 e^{i\kappa\cdot a} - A\kappa_y^2 e^{i\kappa\cdot a} = -k^2 U \ .$$

As the Helmholtz equation is linear, any linear superposition of solutions with different values of κ but the *same* magnitude k is also a solution. It is thus natural to assume that the field Z satisfies the Helmholtz equation for a certain wave number k:

$$\Delta Z + k^2 Z = 0 \ .$$

Let $\mathscr{H}_k$ be the space of solutions C^∞ of the Helmholtz equation with wave number k. It can be shown that $\mathscr{H}_k$ is $SE(2)$-invariant; that is, if $Z(a) \in \mathscr{H}_k$ and if $g \in G = SE(2)$, then $gZ(a) = Z\left(g^{-1}(a)\right)$ is also an element of $\mathscr{H}_k$. The action of G on $\mathscr{H}_k$ therefore defines a representation of G, and this has the property of *irreducibility* in the usual sense that there is no *closed* subspace which is G-invariant.

Figure 4.66, recomputed from the data of Michael Berry's work [69] on optical currents, shows a superposition of 10 plane waves with the same value of k. We see how closely this phase field resembles our pinwheels, with its isophase lines, i.e. iso-orientation lines, its orthogonal gradient lines, and its saddle points.

4.6.7 Illustration

It is interesting to consider Michael Berry's example, which is a superposition of 10 plane waves, viz.

$$Z = \sum_{j=1}^{j=10} E_j \exp\Big[i\big[\phi_j + 2\pi x\cos(\alpha_j) + 2\pi y\sin(\alpha_j)\big]\Big] = re^{i\varphi} \ ,$$

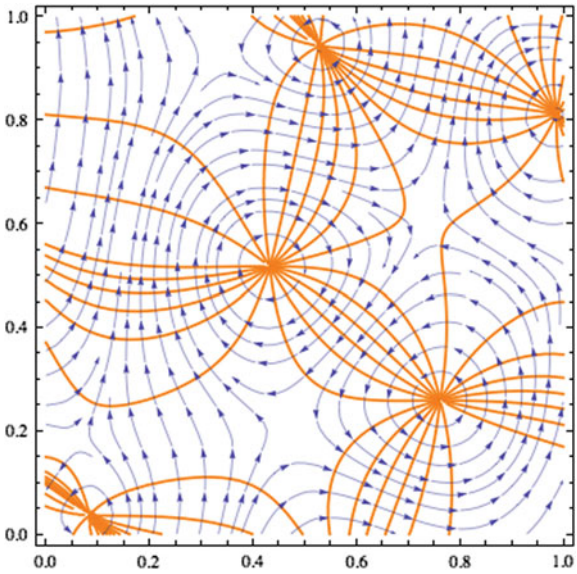


Fig. 4.66 Superposition in the *rectangle* $[0, 1] \times [0, 1]$ of 10 plane waves with the same wave number $k = 2\pi$. The wave vectors $\kappa_j = \left(2\pi \cos\left(\alpha_j\right), 2\pi \sin\left(\alpha_j\right)\right)$ are given below in Table 4.1. Recomputed from the data of Berry [69]

Table 4.1 Values used in Berry's example [69]

	1	2	3	4	5	6	7	8	9	10
α_j	5.971	2.666	0.939	4.629	1.023	1.537	2.710	3.273	4.356	5.032
ϕ_j	3.846	0.777	5.008	2.916	6.274	4.344	2.411	5.688	1.734	0.214
E_j	0.337	0.015	0.762	0.785	0.625	0.442	0.688	0.065	0.064	0.035

with the same wave number $k = 2\pi$ (and hence wavelength $\Lambda = 2\pi/k = 1$) and wave vectors $\kappa_j = \left(2\pi \cos\left(\alpha_j\right), 2\pi \sin\left(\alpha_j\right)\right)$. The angles α_j are chosen randomly in $[0, 2\pi]$, the phase shift ϕ_j randomly in $[0, 2\pi]$, and the amplitudes E_j randomly in $[0, 1]$. The values used by Berry are shown in Table 4.1. If Φ_j are the phases $\phi_j + 2\pi x \cos\left(\alpha_j\right) + 2\pi y \sin\left(\alpha_j\right)$, we get an expression $Z = \sum_{j=1}^{j=10} E_j \exp\left(\mathrm{i}\Phi_j\right)$ that is easy to calculate with. Figure 4.67 shows the pinwheels of Z. The white lines are cuts where φ jumps by 2π, due to the fact that φ takes values in $\mathbb{S}^1$ but is represented as having values in $\mathbb{R}$. The coordinates of the 5 pinwheels are

$$\{0.528545, 0.942654\}\ ,\quad \{0.988124, 0.811337\}\ ,\quad \{0.433271, 0.516137\}\ ,$$

$$\{0.761954, 0.258734\}\ ,\quad \{0.0838329, 0.0359263\}\ .$$

They are given by the intersections of the level lines $X = 0$, $Y = 0$. Figure 4.68 represents the lines $X = 0$ in red and the lines $Y = 0$ in blue.

Figure 4.69 shows the structure of the phase field Z on the square $x, y \in [0, 1]$. The first line shows the modulus r of Z (i) with its level lines and (ii) as a function

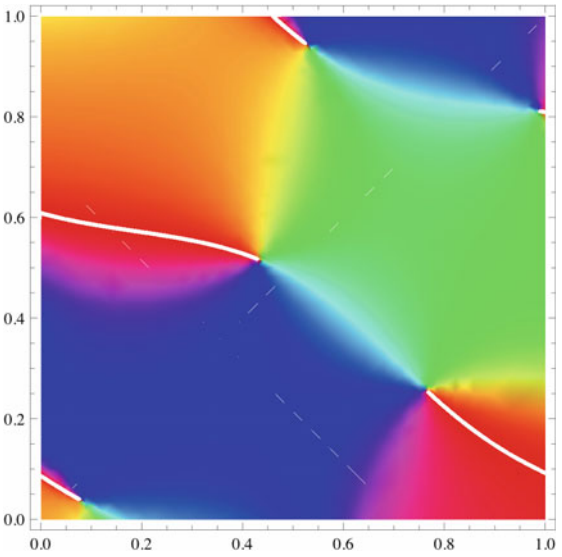


Fig. 4.67 Pinwheels of the phase field Z for $x, y \in [0, 1]$. *White lines* are cuts where φ jumps by 2π

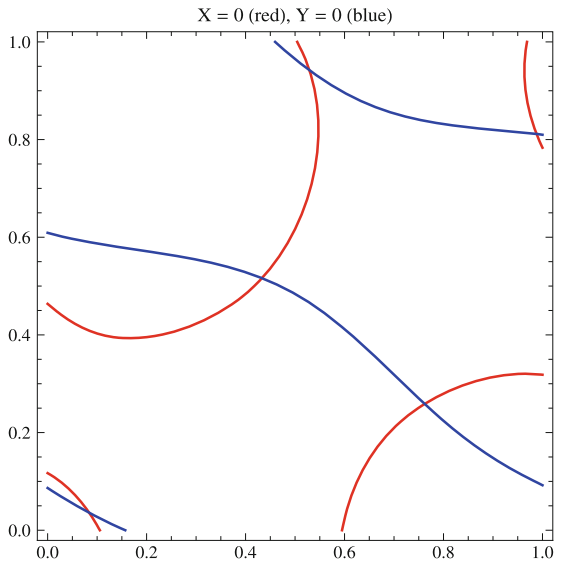


Fig. 4.68 Lines $X = 0$ (*red*) and $Y = 0$ (*blue*) of the phase field Z

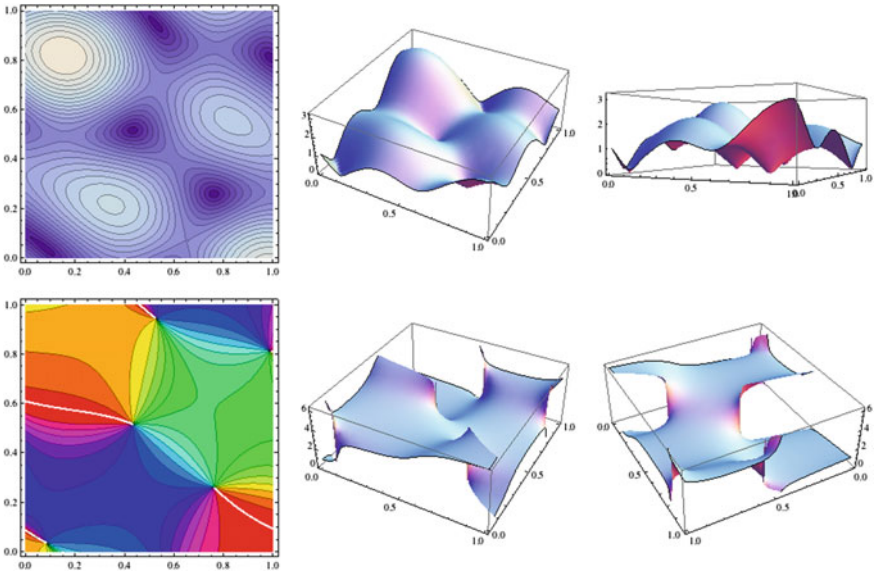


Fig. 4.69 Structure of the phase field Z for $x, y \in [0, 1]$. *First line* modulus r of Z. (i) Level lines. (ii) The function $r\,(x, y)$ from two perspectives, the second giving a good view of the singular points (dislocations) where $r = 0$ and also the maxima of r. *Second line* argument φ of Z. (i) Level lines. (ii) Function $\varphi\,(x, y)$ from two perspectives, the second giving a good view of the cuts where φ jumps by 2π

$r\,(a) = r\,(x, y)$ from two perspectives, the second giving a good view of the singular points (dislocations), where $r = 0$, and the maxima of r. The second line shows the argument φ of Z (i) with its level lines (we recover those in Fig. 4.66) and (ii) as a function $\varphi\,(a) = \varphi\,(x, y)$ from two perspectives, the second giving a good view of the cuts, where φ jumps by 2π.

Figure 4.70 shows more pinwheels of Z ($x, y \in [0, 3]$). The white cuts represent $\varphi = 0 = 2\pi$. Note that there are 29 pinwheels in an area of $3^2 = 9$, giving a *density* d equal to $29/9 \sim 3.2$. We shall explain in Sect. 4.6.11 a formula giving $d = \pi/\Lambda^2$, which means $d = \pi$ in our case, since $\Lambda = 1$. We see that the approximation is excellent. Concerning Fig. 4.71, it represents the phase φ for partial sums of Z.

In Fig. 4.72, we also show the orientation lines, that is the field lines of W, and the field lines of Z (not to be confused with the isophase lines). We do indeed see examples of the typical singularities illustrated in Fig. 4.55 of Sect. 4.4.6.

It is straightforward to calculate the current

$$\mathscr{J} = r^2 \nabla\varphi = X\nabla Y - Y\nabla X = \mathrm{Im}\left(\overline{Z}\nabla Z\right) \ .$$

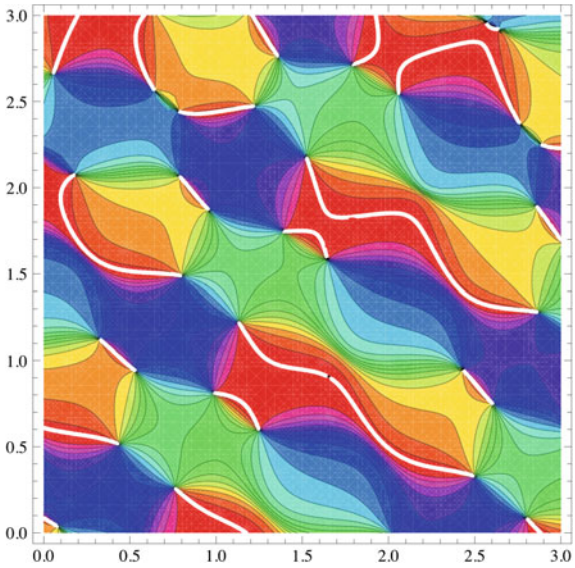


Fig. 4.70 Pinwheels of Z for $x, y \in [0, 3]$. *White cuts* represent $\varphi = 0 = 2\pi$

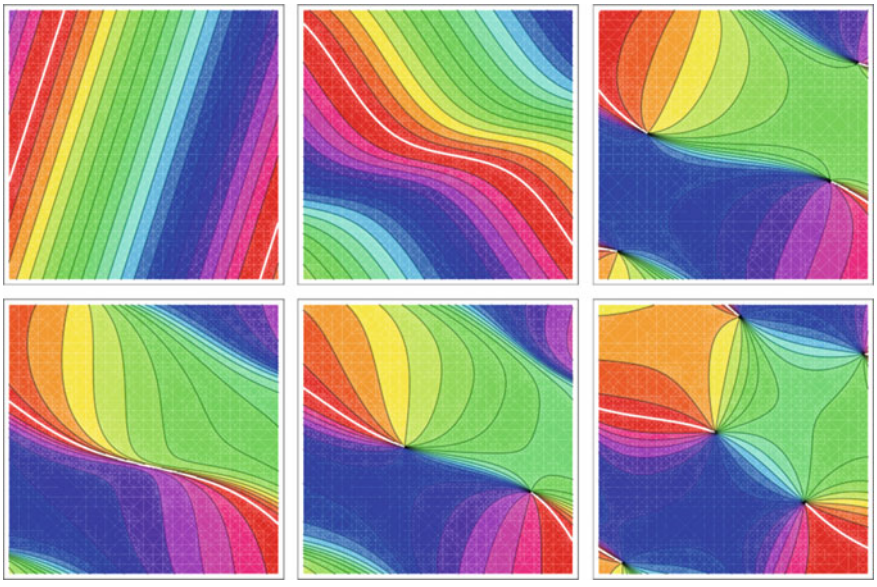


Fig. 4.71 Phase φ for partial sums of Z

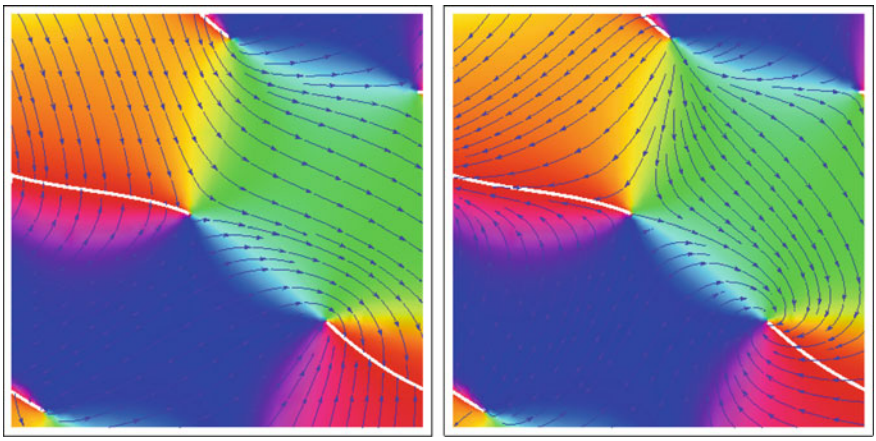


Fig. 4.72 Field lines of W and Z for the example of Fig. 4.69

Indeed,

$$\begin{cases} X = \sum_{k=1}^{k=10} E_k \cos(\Phi_k) \ , \\ Y = \sum_{k=1}^{k=10} E_k \sin(\Phi_k) \ , \\ \dfrac{\partial X}{\partial x} = -\sum_{j=1}^{j=10} 2\pi \cos\left(\alpha_j\right) E_j \sin(\Phi_j) \ , \\ \dfrac{\partial X}{\partial y} = -\sum_{j=1}^{j=10} 2\pi \sin\left(\alpha_j\right) E_j \sin(\Phi_j) \ , \\ \dfrac{\partial Y}{\partial x} = \sum_{j=1}^{j=10} 2\pi \cos\left(\alpha_j\right) E_j \cos(\Phi_j) \ , \\ \dfrac{\partial Y}{\partial y} = \sum_{j=1}^{j=10} 2\pi \sin\left(\alpha_j\right) E_j \cos(\Phi_j) \ , \end{cases}$$

and hence the components $\mathscr{J}_x$ and $\mathscr{J}_y$ of $\mathscr{J}$ are

$$\mathscr{J}_x = \sum_{j,k=1}^{j,k=10} \left(\mathscr{J}_x\right)_{j;k} \ , \qquad \mathscr{J}_y = \sum_{j,k=1}^{j,k=10} \left(\mathscr{J}_y\right)_{j;k} \ ,$$

with

$$\begin{cases} \left(\mathscr{J}_x\right)_{j;k} = 2\pi E_j E_k \cos(\alpha_j) \cos\left[\phi_j + 2\pi\left[x\cos(\alpha_j) + y\sin(\alpha_j)\right]\right] \\ \qquad\qquad \times \cos\left[\phi_k + 2\pi\left[x\cos(\alpha_k) + y\sin(\alpha_k)\right]\right] \\ \qquad + 2\pi E_j E_k \cos(\alpha_j) \sin\left[\phi_j + 2\pi\left(x\cos(\alpha_j) + y\sin(\alpha_j)\right)\right] \\ \qquad\qquad \times \sin\left[\phi_k + 2\pi\left[x\cos(\alpha_k) + y\sin(\alpha_k)\right]\right], \\ \left(\mathscr{J}_y\right)_{j;k} = 2\pi E_j E_k \sin(\alpha_j) \cos\left[\phi_j + 2\pi\left[x\cos(\alpha_j) + y\sin(\alpha_j)\right]\right] \\ \qquad\qquad \times \cos\left[\phi_k + 2\pi\left[x\cos(\alpha_k) + y\sin(\alpha_k)\right]\right] \\ \qquad + 2\pi E_j E_k \sin(\alpha_j) \sin\left[\phi_j + 2\pi\left(x\cos(\alpha_j) + y\sin(\alpha_j)\right)\right] \\ \qquad\qquad \times \sin\left[\phi_k + 2\pi\left[x\cos(\alpha_k) + y\sin(\alpha_k)\right]\right]. \end{cases}$$

Figure 4.66 already displayed above recovers the structure of the phase field Z with its isophase lines in orange and the current lines orthogonal to them, since the current $\mathscr{J}$ is parallel to the gradient $\nabla\varphi$ of φ and hence orthogonal to the lines $\varphi = \text{const}$.

The vorticity $\Omega = \nabla \times \mathscr{J}/2 = \nabla X \times \nabla Y$ is shown in Fig. 4.73 by the graph of the function

$$\omega(x, y) = \frac{\partial X}{\partial x}\frac{\partial Y}{\partial y} - \frac{\partial X}{\partial y}\frac{\partial Y}{\partial x},$$

for $x, y \in [0, 1]$. Figure 4.74 shows the position of the lines $\Omega = 0$ for $x, y \in [0, 3]$. We see that adjacent pinwheels do indeed have opposite chirality, since they belong to regions where Ω has opposite sign.

We know that, when $\Omega = 0$, either $\nabla X = 0$ or $\nabla Y = 0$ (isolated points of codimension 2), or the real gradients ∇X and ∇Y have the same orientation, i.e. $\nabla Y = \alpha \nabla X$ (see Sect. 4.6.5). Figure 4.75 shows the curves $\partial X/\partial x = \partial_x X = 0$ and $\partial_y X = 0$ in red and $\partial_x Y = 0$ and $\partial_y Y = 0$ in blue [the $\partial_x X$, etc., are denoted by $X(x)$, etc., in the label]. Note that points where $\nabla X = 0$ (intersection of red lines at top left) and $\nabla Y = 0$ (intersection of blue lines at top right) do indeed lie on the lines $\Omega = 0$. Likewise, the points where ∇X and ∇Y are both vertical ($\partial_x X = 0$ and $\partial_x Y = 0$) or both horizontal ($\partial_y X = 0$ and $\partial_y Y = 0$) also lie on the lines $\Omega = 0$ (intersections of red and blue lines at bottom left and bottom right).

Figure 4.76 shows the curves $\partial_x X = \alpha \partial_x Y$ and $\partial_y X = \alpha \partial_y Y$ (and their intersections, where $\nabla X = \alpha \nabla Y$) for different values of α. We see that these points do indeed lie on the lines $\Omega = 0$. Figure 4.77 shows the lines $\Omega = 0$ for $x, y \in [0, 1]$. It is interesting to look at the details. The typical situation is when two adjacent pinwheels are 'close', a line $\Omega = 0$ intersecting the isophase line that connects them

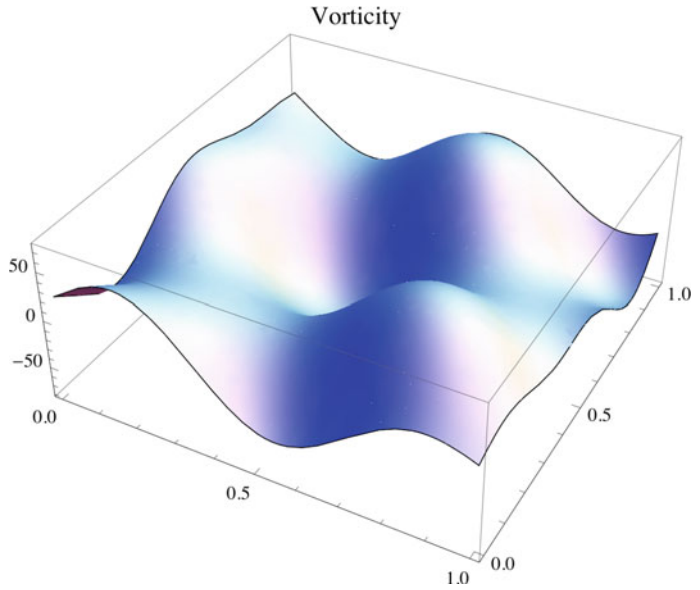


Fig. 4.73 Vorticity Ω of the phase field Z for $x, y \in [0, 1]$, i.e. the graph of the function $\omega(x, y) = \frac{\partial X}{\partial x}\frac{\partial Y}{\partial y} - \frac{\partial X}{\partial y}\frac{\partial Y}{\partial x}$

the most directly in a strongly transverse manner somewhere near the middle. This typical situation can be highly distorted for 'distant' adjacent pinwheels.

4.6.8 Current Conservation

Assuming that Z solves the Helmholtz equation, we consider the divergence of the current $\mathscr{J}$ given by

$$\begin{aligned}\operatorname{div}\left(\mathscr{J}\right) &= \frac{\partial \mathscr{J}_x}{\partial x} + \frac{\partial \mathscr{J}_y}{\partial y} = \frac{\partial}{\partial x}\left(r^2\frac{\partial\varphi}{\partial x}\right) + \frac{\partial}{\partial y}\left(r^2\frac{\partial\varphi}{\partial y}\right)\\ &= r^2\frac{\partial^2\varphi}{\partial x^2} + 2r\frac{\partial r}{\partial x}\frac{\partial\varphi}{\partial x} + r^2\frac{\partial^2\varphi}{\partial y^2} + 2r\frac{\partial r}{\partial y}\frac{\partial\varphi}{\partial y}\\ &= r^2\Delta\varphi + 2r\nabla r\cdot\nabla\varphi\ ,\end{aligned}$$

where $\nabla r \cdot \nabla\varphi$ is the scalar product and the Laplacian operator $\Delta\varphi$ appears because it is by definition the divergence of the gradient. The Laplacian ΔZ of $Z = re^{i\varphi}$ is given by

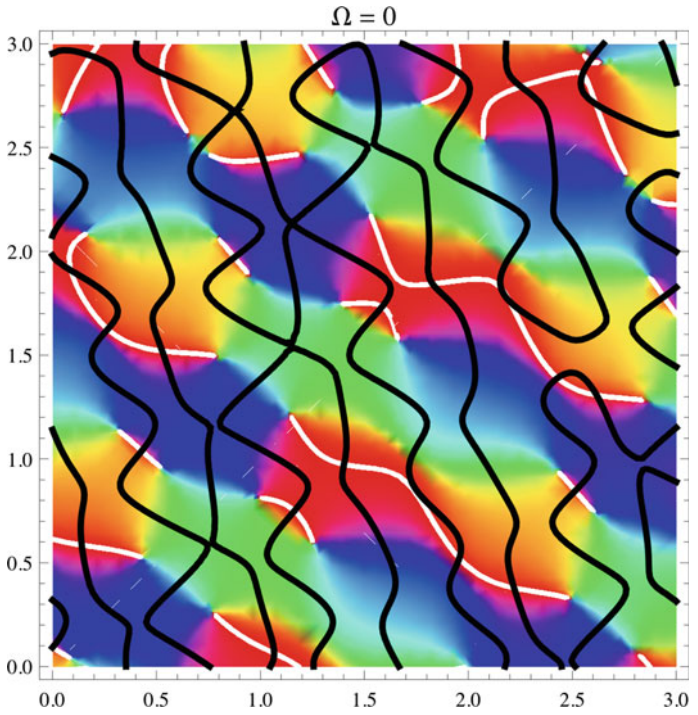


Fig. 4.74 Position of the vorticity lines $\Omega = 0$ for the phase field Z when $x, y \in [0, 3]$

$$\begin{aligned}\Delta Z &= \frac{\partial^2 Z}{\partial x^2} + \frac{\partial^2 Z}{\partial y^2} = \frac{\partial^2 Z}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial Z}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2 Z}{\partial \theta^2} \\ &= \mathrm{e}^{\mathrm{i}\varphi}\Big[\Delta r - r\,|\nabla\varphi|^2 + \mathrm{i}\,(r\Delta\varphi + 2\nabla r\cdot\nabla\varphi)\Big]\,,\end{aligned}$$

so if $\Delta Z + k^2 Z = 0$, we must have

$$\begin{cases}\Delta r + r\big(k^2 - |\nabla\varphi|^2\big) = 0\,, \\ r\Delta\varphi + 2\nabla r\cdot\nabla\varphi = 0\,.\end{cases}$$

The second equation expresses the fact that the divergence of the current is zero, that is $\mathrm{div}\,(\mathscr{J}) = 0$. This is a *conservation law*. It implies that $\mathscr{J}$ can be written in the form $e_3 \times \nabla S = (-\partial S/\partial y, \partial S/\partial x)$. Indeed, since $\mathbb{R}^2$ is simply connected, a necessary and sufficient condition for this is $\partial^2 S/\partial x \partial y = \partial^2 S/\partial y \partial x$, i.e., $\partial \mathscr{J}/\partial x = -\partial \mathscr{J}/\partial y$, which says precisely $\mathrm{div}\,(\mathscr{J}) = 0$. We then have

$$\omega = \frac{1}{2}\left(\frac{\partial \mathscr{J}_y}{\partial x} - \frac{\partial \mathscr{J}_x}{\partial y}\right) = \frac{1}{2}\left[\frac{\partial}{\partial x}\left(\frac{\partial S}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial S}{\partial y}\right)\right] = \frac{1}{2}\Delta S\,.$$

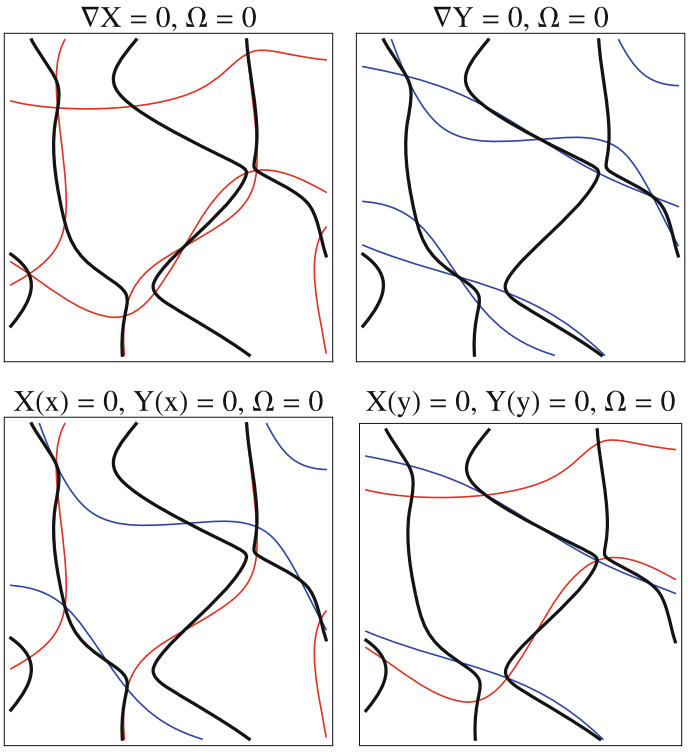


Fig. 4.75 Curves $\partial_x X = 0$ and $\partial_y X = 0$ (*red*) and $\partial_x Y = 0$ and $\partial_y Y = 0$ (*blue*). *Top left* points where $\nabla X = 0$ (*intersection of red lines*). *Top right* points where $\nabla Y = 0$ (*intersection of blue lines*). *Bottom left* points where ∇X and ∇Y are both vertical [*intersections of red and blue lines* $\partial_x X = X(x) = 0$ and $\partial_x Y = Y(x) = 0$]. *Bottom right* points where ∇X and ∇Y are both horizontal [*intersections of red and blue lines* $\partial_y X = X(y) = 0$ and $\partial_y Y = Y(y) = 0$]. All these points lie on the lines $\Omega = 0$

Regarding the first equation, it turns up everywhere in optics in a form where k is part of the phase, i.e. where $Z = re^{ik\varphi}$. It then becomes $\Delta r = k^2 r\left(|\nabla\varphi|^2 - 1\right)$. When $k \to \infty$, which corresponds to the geometrical optics approximation, it becomes the well-known *eikonal equation* $|\nabla\varphi|^2 = 1$, which expresses the fact that 'light rays', i.e. the trajectories of the field gradient $\nabla\varphi$, go around the singularity at a constant rate, while the wave fronts $\varphi = \text{const.}$ are the rays coming from the singularity (leading to some confusion between the two meanings of the word 'ray'). To see this, we expand r asymptotically in powers of k, viz.

$$r(a) \sim k^{\mu} \sum_{n=0}^{n=\infty} k^{-n} r_n(a) \ ,$$

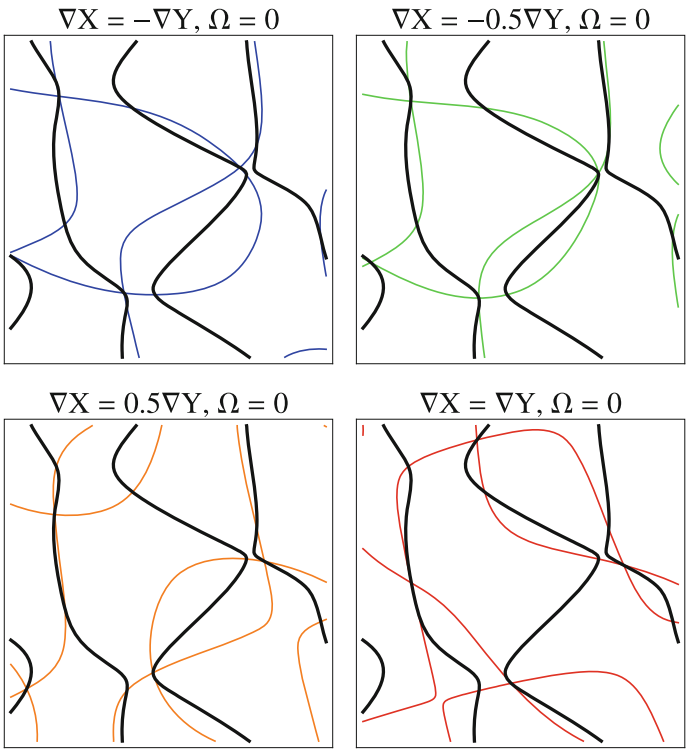


Fig. 4.76 Four examples of curves $\partial_x X = \alpha \partial_x Y$ and $\partial_y X = \alpha \partial_y Y$ (and their intersections where $\nabla X = \alpha \nabla Y$) for different values of α. *Top left* $\alpha = -1$ (*blue*). *Top right* $\alpha = -1/2$ (*green*). *Bottom left* $\alpha = 1/2$ (*orange*). *Bottom right* $\alpha = 1$ (*red*). We see that all points where $\nabla X = \alpha \nabla Y$ do indeed lie on the lines $\Omega = 0$

with $r_0(a) \neq 0$. When $k \to \infty$, we have $r(a) \to k^{\mu} r_0(a)$ and we deduce that $r_0\left(1 - |\nabla\varphi|^2\right) = 0$, and hence $|\nabla\varphi|^2 = 1$.

The quantity S can be calculated explicitly for superpositions of plane waves. Indeed, S is a sum of integrals like

$$\int 2\pi E_j E_k \cos(\alpha_j) \cos\Big[\phi_j + 2\pi\big[x\cos(\alpha_j) + y'\sin(\alpha_j)\big]\Big] \times \cos\Big[\phi_k + 2\pi\big[x\cos(\alpha_k) + y'\sin(\alpha_k)\big]\Big] \mathrm{d}y' \,,$$

which gives

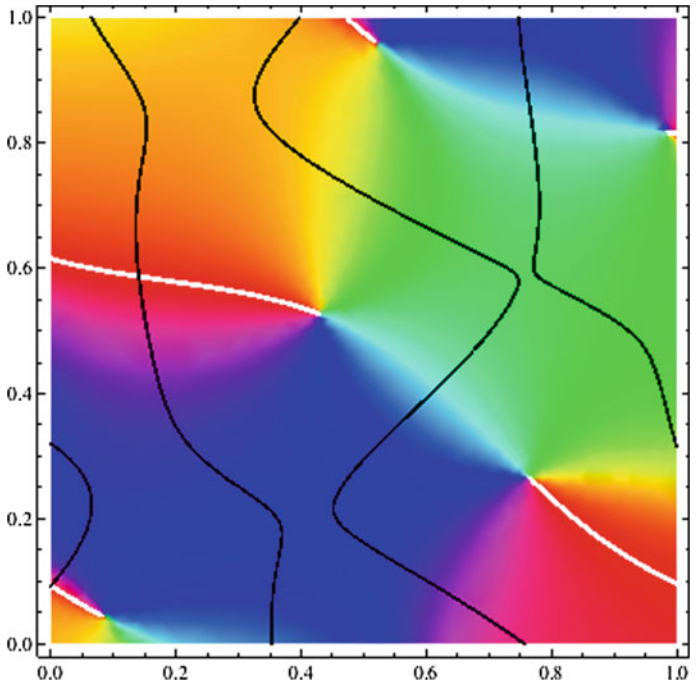


Fig. 4.77 Lines of vorticity $\Omega = 0$ for the phase field Z when $x, y \in [0, 1]$

$$\frac{1}{2}E_j E_k \cos(\alpha_j)\left\{\frac{\sin\left[2\pi\left[x\cos(\alpha_j)+y\sin(\alpha_j)-x\cos(\alpha_k)-y\sin(\alpha_k)\right]+\phi_j-\phi_k\right]}{\sin(\alpha_j)-\sin(\alpha_k)}\right.$$
$$+\frac{\sin\left[2\pi\left[y\sin(\alpha_j)+y\sin(\alpha_k)+x\cos(\alpha_j)+x\cos(\alpha_k)\right]+\phi_j+\phi_k\right]}{\sin(\alpha_j)+\sin(\alpha_k)}$$
$$-\frac{\sin\left[2\pi x\left[\cos(\alpha_j)+\cos(\alpha_k)\right]+\phi_j+\phi_k\right]}{\sin(\alpha_j)+\sin(\alpha_k)}$$
$$\left.-\frac{\sin\left[2\pi x\left[\cos(\alpha_j)-\cos(\alpha_k)\right]+\phi_j-\phi_k\right]}{\sin(\alpha_j)-\sin(\alpha_k)}\right\},$$

for $k \neq j$ and

$$\tfrac{1}{4}E_j^2\cot(\alpha_j)\Big[\sin\left[2\phi_j+4\pi x\cos(\alpha_j)+4\pi y\sin(\alpha_j)\right]$$
$$-\sin\left[2\phi_j+4\pi x\cos(\alpha_j)+4\pi y\sin(\alpha_j)\right]\Big],$$

for $k = j$.

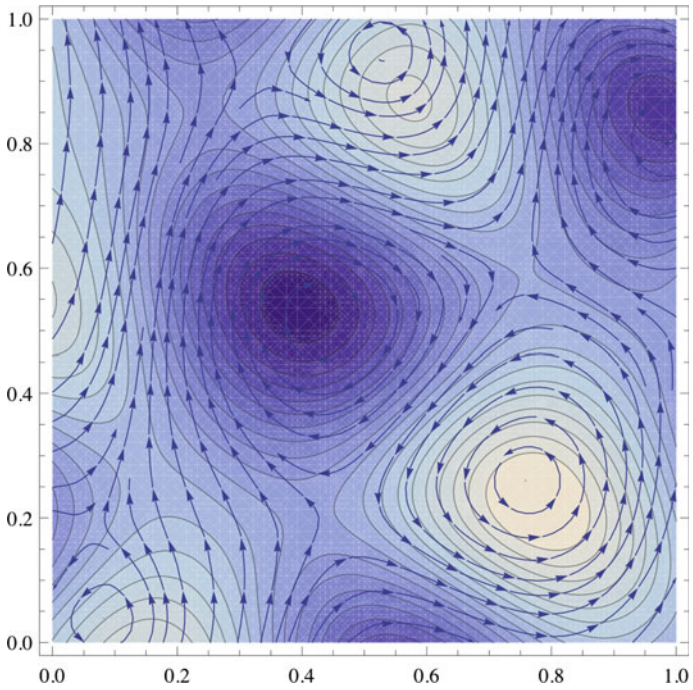


Fig. 4.78 Flow of the current $\mathscr{J}$ shown with the level lines of S

Figure 4.78 shows the flow of the current $\mathscr{J}$ with the level lines of S. We note that the dislocations do indeed occur in regions of highest vorticity (> 0 or < 0), but this does not mean that they coincide with the extrema of S. The definition of the latter involves second derivatives of the field Z:

$$\begin{cases} \dfrac{\partial \omega}{\partial x} = \dfrac{\partial Y}{\partial y}\dfrac{\partial^2 X}{\partial x^2} + \dfrac{\partial X}{\partial x}\dfrac{\partial^2 Y}{\partial x \partial y} - \dfrac{\partial Y}{\partial x}\dfrac{\partial^2 X}{\partial x \partial y} - \dfrac{\partial X}{\partial y}\dfrac{\partial^2 Y}{\partial x^2}\,, \\ \dfrac{\partial \omega}{\partial y} = \dfrac{\partial Y}{\partial y}\dfrac{\partial^2 X}{\partial x \partial y} + \dfrac{\partial X}{\partial x}\dfrac{\partial^2 Y}{\partial y^2} - \dfrac{\partial Y}{\partial x}\dfrac{\partial^2 X}{\partial y^2} - \dfrac{\partial X}{\partial y}\dfrac{\partial^2 Y}{\partial x \partial y}\,. \end{cases}$$

4.6.9 Critical Points

Apart from dislocation singularities where the phase φ is not defined, there are other geometrically interesting points giving structure to the field Z. As already noted, these are the *critical points* of φ where the gradient $\nabla\varphi$ vanishes, and hence also the current $\mathscr{J} = r^2\nabla\varphi$. At such a point a_0, the first term in the Taylor expansion of φ

is second order. If δa is an increment in a, we have to second order $\varphi(a_0 + \delta a) \approx \varphi(a_0) + \delta a^{\mathrm{T}} H_\varphi \delta a$, where δa is treated as a column vector, δa^{T} is its transpose, and H_φ is the symmetric 2×2 matrix of second partial derivatives of φ at a_0, a matrix known as the *Hessian* of φ:

$$H_\varphi = \begin{pmatrix} \partial^2\varphi/\partial x^2 & \partial^2\varphi/\partial x \partial y \\ \partial^2\varphi/\partial x \partial y & \partial^2\varphi/\partial y^2 \end{pmatrix}.$$

The eigenvalues of the Hessian H_φ determine the type of critical point. They are real because of the symmetry of H_φ, generically distinct, and nonzero.[14] If they are both strictly positive, then $\delta a^{\mathrm{T}} H_\varphi \delta a > 0$ and a_0 is a minimum of φ. If they are both strictly negative, a_0 is a maximum. If they have opposite signs, then a_0 is a saddle point of φ. Pinwheel maps would not appear to have orientation maxima or minima. However, we have seen that there are many saddle points, mainly located at the centre of the cells of the pinwheel lattice.

There can also be critical points in the *amplitude* r^2, where $\nabla r^2 = 0$. Since $r^2 = X^2 + Y^2$, we have $\nabla r^2 = 2(X\nabla X + Y\nabla Y)$. At a singularity that is a zero of Z, $X = Y = 0$ and the point is critical. However, there are other critical points where $X, Y \neq 0$. At these points, ∇X and ∇Y are necessarily parallel, so the vorticity $\Omega = \nabla \times \mathscr{J}/2 = \nabla X \times \nabla Y$ is necessarily zero. We thus see that the critical points in the intensity include dislocations and the points where $\Omega = 0$. The critical points in the amplitude are invariant under gauge transformations and coordinate transformations.

4.6.10 Mesogeometry and Microphysics

The optical analogy which suggests treating orientation maps as phase fields and pinwheels as dislocation singularities in such fields is also useful for understanding the relations between different levels. In optics, there are three levels: geometric, wave, and quantum. In our analogy, the geometric level corresponds to the mesogeometric level involving the contact structures, symplectic structures, and sub-Riemannian structures that we shall discuss here in detail.

The wave level corresponds to what we have just been doing, identifying pinwheels with singularities in the phase fields. However, as noted by Berry [69], wave optics is an average over microphysical interactions described by quantum optics. In particular, the optical current is an energy flow whose trajectories are the level lines of S, in a certain sense a momentum density, giving the classical force on a small particle placed at a:

---

[14] Let $H_\varphi = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. The eigenvalues are solutions of the quadratic equation $\mathrm{Det}\left(H_\varphi - \lambda I\right) = 0$, which can be written $\lambda^2 - \lambda(a + c) + \left(ac - b^2\right) = 0$. The discriminant $(a + c)^2 - 4\left(ac - b^2\right) = (a - c)^2 + 4b^2$ is always non-negative, and the roots are therefore real.

The current gives the time-averaged force on small particles.

For its part, the phase gradient $\nabla\varphi$ gives the momentum imparted to the particle by impacts from individual photons. And as the probability of these impacts is r^2, the average momentum is indeed $\mathscr{J} = r^2\nabla\varphi$.

We thus recover the assumption (see Sect. 4.3.2 of Chap. 4) that there is a microphysics of elementary events for which the mesogeometry of the orientation maps is a kind of morphological skeleton. It is presumably the action potentials that play the role of the tiny particles.

4.6.11 Statistics of Pinwheels as Phase Singularities

The pinwheel maps like phase fields can come in many forms. It is thus interesting to carry out a *statistical investigation* on the basis of certain simplifying assumptions. Such studies have already been carried out in optics, in particular by Michael Berry and Mark Richard Dennis (see, e.g. [70, 71]). This is a topical subject bringing together work by Wolf and Geisel [72, 73], studies by Daniel Bennequin and coworkers, and also recent work by Citti, Sarti, and one of their doctoral students, Davide Barbieri [74].

In his thesis [68], Dennis gives precise results for superpositions of plane waves

$$Z = \sum_{\kappa} A_\kappa \mathrm{e}^{\mathrm{i}\kappa\cdot a}$$

with complex amplitudes $A_\kappa = E_\kappa \mathrm{e}^{\mathrm{i}\phi_\kappa}$, in particular in the *isotropic case*, i.e. where there is rotation invariance, so that the E_κ have a distribution depending only on the magnitude $k = |\kappa|$ of the wave vectors (the wave number) and where the spatial phases ϕ_κ are random variables uniformly distributed on $[0, 2\pi]$. If the sampling of the κ in the given sums Z is fine enough, we can consider that the statistics of the components X and Y of Z and their partial derivatives are circular Gaussian distributions, which makes the calculations more accessible. In particular, we define the energy spectrum by

$$\frac{1}{2}\sum_{\kappa} E_\kappa^2 = \int \mathscr{E}\,(\kappa)^2\,\mathrm{d}\kappa\ ,$$

and the radial energy spectrum by

$$\frac{1}{2}\sum_{\kappa} E_\kappa^2 = \int \frac{\mathscr{R}\,(k)^2}{2\pi k}\mathrm{d}k\ .$$

A further simplification comes by considering 'monochromatic' waves, with the same wave number k, where the wave vector κ thus varies over a circle of radius k.

In this case, $\mathscr{R}(u)$ becomes the Dirac delta $\delta(u-k)$. This hypothesis corresponds to the fact that Z is a solution of the Helmholtz equation.

We can then calculate the average density d of phase dislocations. As these are defined by the conditions $X=0, Y=0$, it will be given by the average of $\delta(X)\,\delta(Y)$ with respect to the measure $\mathrm{d}X\mathrm{d}Y$. Relative to the measure $\mathrm{d}x\mathrm{d}y$, we need to insert the Jacobian of $Z(x,y)=X(x,y)+\mathrm{i}Y(x,y)$, i.e.

$$|\omega| = |\nabla X \wedge \nabla X| = \left| \frac{\partial X}{\partial x}\frac{\partial Y}{\partial y} - \frac{\partial X}{\partial y}\frac{\partial Y}{\partial x} \right| .$$

We must therefore calculate the average

$$\left\langle \delta(X)\,\delta(Y) \left| \frac{\partial X}{\partial x}\frac{\partial Y}{\partial y} - \frac{\partial X}{\partial y}\frac{\partial Y}{\partial x} \right| \right\rangle .$$

Assuming that X, Y, and their partial derivatives are independent Gaussian random variables, we can do this using the integrals

$$\delta(u) = \frac{1}{2\pi}\int \mathrm{e}^{\mathrm{i}tu}\mathrm{d}t\ , \qquad |u| = -\frac{1}{\pi}\int \frac{\partial}{\partial s}\left(\mathrm{e}^{\mathrm{i}su}\right)\frac{\mathrm{d}s}{s}\ ,$$

taking the latter in the sense of the Cauchy principal value. We thus obtain

$$d = \frac{K}{4\pi}\ , \quad K = \int_0^\infty k^2 \mathscr{R}(k)\,\mathrm{d}k = \left\langle k^2 \right\rangle_{\mathscr{R}} \text{ for the measure } \mathscr{R}(k)\,\mathrm{d}k\ .$$

As noted by Michael Berry, the fact that Z is a superposition of waves barely comes into the calculation, and [70, p. 2076]:

> The results apply to any complex scalar random function. […] The geometry thus revealed is extraordinarily complicated and occasionally counterintuitive.

The wave number k is proportional to the reciprocal of a wavelength $\Lambda = 2\pi/k$, whence $\Lambda^2 = 4\pi^2/k^2$ and $k^2/4\pi = \pi/\Lambda^2$. Consequently, the density of singularities d is the average $\left\langle \pi/\Lambda^2 \right\rangle_{\mathscr{R}}$. In Sect. 4.7.1.2, we shall once again come across this term π/Λ^2, also found by Fred Wolf and Theo Geisel.

4.6.12 Pinwheels and Gaussian Fields

Quite generally, the orientation maps can be treated as *random sections* of the fibre bundle $\mathbb{R}^2 \times P \to \mathbb{R}^2$ satisfying a set of constraints explaining their pinwheel geometrical structure. The problem then is to calculate their dislocation distribution in statistical terms. At each point a of the base space $\mathbb{R}^2$, we thus consider a random variable $\mathscr{Z}_a$, which defines a random field $\mathscr{Z}$ whose orientation maps $Z(a)$ are

samples. To simplify, we generally assume that the field $\mathscr{Z}$ is *Gaussian* (RGF), i.e. that the $\mathscr{Z}_a$ are Gaussians with mean $m_a = \mathbb{E}\{\mathscr{Z}_a\}$, where $\mathbb{E}$ is the expectation, and variance $\sigma_a^2 = \mathbb{E}\left\{(\mathscr{Z}_a - m_a)^2\right\}$, and that all the joint distributions $\sum_i \alpha_i \mathscr{Z}_{a_i}$ for a finite number of points a_i are also Gaussian. Moreover, it is natural to assume that the distribution of the $\mathscr{Z}_a$ is $SE(2)$-invariant [the distribution of the $\mathscr{Z}_a$ and obviously not the samples $Z(a)$]. Translation invariance is known as *stationarity* and rotation invariance as *isotropy*. For an introduction to random Gaussian fields, the reader could consult, for example, Petter Abrahamsen's review [75].

In the paper [63] on Gaussian models of pinwheels, Alexandre Afgoustidis carried out a numerical calculation of the *variance* of the number $N(J, \theta)$ of neurons with preferred orientation θ along a line segment J in the plane $V1$. His calculation used a formula due to Cramer and Leadbetter which contains oscillating integrals that are rather difficult to calculate. The main result is that the variance is minimal when the power spectrum is concentrated on a circle.

While preparing the paper, the author had an interesting discussion with one of the referees (I was aware of this, being a referee myself). For the initial orientation field, the assumption of an RGF model is plausible. However, we may wonder whether it remains so for the stable orientation fields obtained by learning. Indeed, the role of long-range horizontal connections becomes crucial and changes the statistical properties of the fields. In his thesis [64], Afgoustidis discusses this point with Fred Wolf.

The random variables $\mathscr{Z}_a$ cannot be decorrelated (i.e. independent, since for Gaussian random variables, independence and decorrelation are equivalent), because otherwise there would only be Gaussian noise and no geometrical structure. Furthermore, the very definition of a *continuum* of independent Gaussian random variables raises some tricky questions. What characterizes the field $\mathscr{Z}$ is the *correlation function* $C(a, b) = \mathbb{E}\{(\mathscr{Z}_a - m_a)(\mathscr{Z}_b - m_b)\}$. Dividing by the variances, we obtain the normalized correlation function $\Gamma(a, b) = C(a, b)/\sigma_a\sigma_b$. Stationarity implies that $C(a, b) = C(a - b)$, and isotropy implies in addition that $C(a, b) = C(a - b) = C(\|a - b\|) = C(r)$. The means are all equal, i.e. $m_a = m$, and so are the variances, i.e., $\sigma_a^2 = \sigma^2 = C(0)$, and $\Gamma(r) = C(r)/C(0)$.

The correlation functions are rather special, being symmetric and non-negative definite:

(i) In the case of stationarity, a theorem due to Salomon Bochner tells us that they admit a spectral representation which is a generalized Fourier transform. This implies that $\Gamma(a) = \int_{\mathbb{R}^2} \mathrm{e}^{\mathrm{i}\langle a,\kappa\rangle} \mathrm{d}F(\kappa)$[15], where F is a bounded non-negative measure on the space of wave vectors κ dual to the positions a. We have $F\left(\mathbb{R}^2\right) = \Gamma(0) = 1$. In polar coordinates (k, α) for κ, we have

$$\Gamma(a) = \int_{\mathbb{R}^+ \times \mathbb{S}^1} \mathrm{e}^{\mathrm{i}k\left[x\cos(\alpha) + y\sin(\alpha)\right]} \mathrm{d}\alpha \, \mathrm{d}P(k) \ ,$$

for a measure $P(k)$ on $\mathbb{R}^+$.

[15] $\langle\ ,\ \rangle$ is the natural pairing between vectors and dual covectors. It can also be expressed by a dot.

(ii) If the measure F is smooth enough relative to the Lebesgue measure $\mathrm{d}\kappa$, then it has a spectral density $f(\kappa)$ and the generalized FT in Bochner's theorem reduces to a Fourier transform

$$\Gamma(a) = \int_{\mathbb{R}^2} \mathrm{e}^{\mathrm{i}\langle a,\kappa\rangle} f(\kappa)\,\mathrm{d}\kappa \ ,$$

with inverse transform

$$f(\kappa) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \mathrm{e}^{-\mathrm{i}\langle a,\kappa\rangle} \Gamma(a)\,\mathrm{d}a \ .$$

(iii) If there is also isotropy, then $\Gamma(r) = \int_0^\infty J_0(kr)\,kf(k)\,\mathrm{d}k$ with $k = \|\kappa\|$, where J_0 is the Bessel function.

(iv) If in addition, we consider the solutions of the Helmholtz equation with wave number k_0, then $f(k)$ is proportional to $\delta(k - k_0)$ and $\Gamma(r)$ is proportional to $J_0(kr)\,k_0$.

In this context, the formula for the statistics of the dislocations in the fields $Z(a)$ sampling the random Gaussian field $\mathscr{Z}$ is a special case of a fundamental formula called the *Kac–Rice formula* (from Kac and Rice) [76]. We would like to calculate the average $d = \mathbb{E}\{\#\{a \in T : Z(a) = 0\}\}$ (recall that # indicates the cardinality of a set) of the number of zeros of Z in a unit square T. Let N_T be this number. We thus have $d = \mathbb{E}\{N_T\}$. The Rice formula tells us that

$$d = \int_T \mathbb{E}\Big\{ |\det(\mathrm{Jac}(\mathscr{Z}_a))| : \mathscr{Z}_a = 0 \Big\} p_{\mathscr{Z}_a}(0)\,\mathrm{d}a \ ,$$

where Jac is the Jacobian and $p_{\mathscr{Z}_a}$ the density of $\mathscr{Z}_a$. The calculation done by Dennis [68] involved working out Gaussian integrals. Let $\mathscr{X}_a$ and $\mathscr{Y}_a$ be the components of $\mathscr{Z}_a$ and $\mathscr{J}_a = \big|\det(\mathrm{Jac}(\mathscr{Z}_a))\big|$. It can be shown that the 6 random variables

$$\mathscr{X}_a \ , \quad \mathscr{Y}_a \ , \quad \left(\frac{\partial \mathscr{X}}{\partial x}\right)_a \ , \quad \left(\frac{\partial \mathscr{X}}{\partial y}\right)_a \ , \quad \left(\frac{\partial \mathscr{Y}}{\partial x}\right)_a \ , \quad \left(\frac{\partial \mathscr{Y}}{\partial y}\right)_a \ ,$$

are independent Gaussian variables, each with a distribution of the form

$$\frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\xi^2/2\sigma^2} \ .$$

As we have seen, we then have to evaluate an integral of the form

$$\int \delta(X)\,\delta(Y)\,Jp\,\mathrm{d}(X)\,\mathrm{d}(Y)\,\mathrm{d}\left(\frac{\partial X}{\partial x}\right)\mathrm{d}\left(\frac{\partial X}{\partial y}\right)\mathrm{d}\left(\frac{\partial Y}{\partial x}\right)\mathrm{d}\left(\frac{\partial Y}{\partial y}\right) \ ,$$

where p is the product of the distributions. The first two variables $\mathscr{X}_a$ and $\mathscr{Y}_a$ have variance 1, which introduces a factor $\left(1/\sqrt{2\pi}\right)^2 = 1/2\pi$ in the integral, and the other four have variance $\sigma^2 = K/2$, which introduces a factor $\left(1/\sqrt{2\pi}\sigma\right)^4 = \left(1/2\pi\sigma^2\right)^2 = 1/(\pi K)^2$. The condition $\mathscr{Z}_a = 0$ amounts to putting $\delta(X)\,\delta(Y)$ in the integral, but $\int \delta(X)\,\mathrm{e}^{-X^2/2} = 1$ and the same for Y. It thus remains to evaluate

$$\frac{1}{2\pi}\frac{1}{(\pi K)^2}\int J\mathrm{e}^{-\left(\|\nabla X\|^2+\|\nabla Y\|^2\right)/K}\,\mathrm{d}\left(\frac{\partial X}{\partial x}\right)\mathrm{d}\left(\frac{\partial X}{\partial y}\right)\mathrm{d}\left(\frac{\partial Y}{\partial x}\right)\mathrm{d}\left(\frac{\partial Y}{\partial y}\right) .$$

If we change to polar coordinates by writing $\nabla X = R_X\mathrm{e}^{\mathrm{i}\psi_X}$ and $\nabla Y = R_Y\mathrm{e}^{\mathrm{i}\psi_Y}$, the integral becomes

$$\frac{1}{2\pi}\frac{1}{(\pi K)^2}\int_{R_X=0}^{R_X=\infty}\int_{R_Y=0}^{R_Y=\infty}\int_{\psi_X=0}^{\psi_X=2\pi}\int_{\psi_Y=0}^{\psi_Y=2\pi} (R_X)^2(R_Y)^2\left|\sin(\psi_Y-\psi_X)\right|\mathrm{e}^{-\left(R_X^2+R_Y^2\right)/K}\mathrm{d}R_X\,\mathrm{d}R_Y\,\mathrm{d}\psi_X\,\mathrm{d}\psi_Y .$$

The integral of the sine gives 8π and the integrals of R_X and R_Y each give $K^{3/2}\sqrt{\pi}/4$. So finally,

$$d = \frac{1}{2\pi}\frac{1}{(\pi K)^2}8\pi\frac{1}{16}K^3\pi = \frac{K}{4\pi} .$$

When the wave vectors κ are concentrated on a circle of radius $k_0 = 2\pi/\Lambda$ in the Fourier space, the Gaussian random variables

$$\left(\frac{\partial\mathscr{X}}{\partial x}\right)_a , \quad \left(\frac{\partial\mathscr{X}}{\partial y}\right)_a , \quad \left(\frac{\partial\mathscr{Y}}{\partial x}\right)_a , \quad \left(\frac{\partial\mathscr{Y}}{\partial y}\right)_a ,$$

with their distribution

$$\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\xi^2/2\sigma^2} ,$$

which has variance $\sigma^2 = K/2$ satisfy $K = k_0^2$, and we thus obtain, as stated above,

$$d = \frac{K}{4\pi} = \frac{k_0^2}{4\pi} = \left(\frac{2\pi}{\Lambda}\right)^2\frac{1}{4\pi} = \frac{\pi}{\Lambda^2} .$$

These statistical calculations, which only give one particularly simple example of the connection between statistics and geometry, are especially interesting from a theoretical point of view (and not only numerically) for the following reason. In their 2005 reference book *Random Fields and Geometry* [77], Robert Adler and Jonathan

Taylor studied in great depth the generalizations of the Kac–Rice formula for random fields $\mathscr{F}_a$ defined on a base space M and with values in $\mathbb{R}^k$.

Let us take, for example, $k = 1$. One of the main problems, and an extremely difficult one, is to calculate $\mathbb{P}\left\{\sup_{a\in M} \mathscr{F}_a \geq u\right\}$ for large u. This so-called probability of excursion in the interval $[u, \infty)$ is well approximated by $\mathbb{E}\left\{\chi\left(A_{[u,\infty)}\right)\right\}$, where quite generally, if D is a domain of $\mathbb{R}^k$, $A_D := \{a \in M : \mathscr{F}_a \in D\}$, and where χ is the *Euler–Poincaré characteristic*. Assuming Gaussianity, stationarity, isotropy, and smoothness of the correlation functions C, we can obtain explicit but complicated formulas for the $\mathbb{E}\{\chi(A_D)\}$. What is interesting is that the proof of these formulas requires all the fundamental tools of the Morse–Whitney–Thom 'philosophy' discussed in Sect. 1.3 of the *Preface*.

To begin with, in order to be able to handle enough cases, we assume that M is a manifold with boundary of dimension N, equipped with a 'good' stratification $M = \cup_{k=0}^{k=N} \partial_k M$, called a Whitney stratification, satisfying Whitney's properties A and B. Here, k is the dimension of the strata making up $\partial_k M$, with $\partial_N M = \mathring{M}$, where $\mathring{M}$ is the interior of M, and $\partial_0 M = \{\text{vertices of } M\}$. We assume that the smoothness properties of C imply that the samples F of the field $\mathscr{F}$ are *Morse functions* on M, using the generalization of Morse theory to stratified manifolds due to Robert MacPherson. In addition, the field $\mathscr{F}$ defines a *natural metric* $d_{\mathscr{F}}(a, b)$ on the base space M through the formula

$$d^2_{\mathscr{F}}(a, b) = \mathbb{E}\left\{\|\mathscr{F}_a - \mathscr{F}_b\|^2\right\} .$$

We can thus also make use of the resources of Riemannian geometry, such as curvature tensor, Levi-Civita connection, covariant derivative, Lipschitz–Killing curvatures, and so on.

For $k = 1$, we thus apply to $A_{[u,\infty)}$ the formulas relating Morse theory to the Euler–Poincaré characteristic. If the sample F of $\mathscr{F}$ is Morse and if u is a regular, i.e. non-critical, value of F, then $A_{[u,\infty)}$ is a sub-manifold with boundary of M that is 'well stratified' by the intersection strata $\mathring{A}_{[u,\infty)} \cap \partial_k M$ and $\partial A_{[u,\infty)} \cap \partial_k M$. F is not necessarily Morse on $A_{[u,\infty)}$, but it can be approximated by a Morse function $\widetilde{F}$ whose critical points correspond to the critical points of F situated above u.

The Rice–Kac formula corresponds to a rectangle $M = T$ of $\mathbb{R}^N$ and to $\mathscr{F}$ with values in $\mathbb{R}^N$. Let J be the Jacobian of a sample F (an $N \times N$ matrix). Let N_u be the number of points of T for which $F(a) = u \in \mathbb{R}^N$. Then, the formula tells us that

$$\mathbb{E}\{N_u\} = \int_T \mathbb{E}\Big\{|\det(J)| : F(a) = u\Big\} p_a(u)\, \mathrm{d}t .$$

In our case $N = 2$, $\mathscr{F} = \mathscr{Z}$, $u = 0$, and T is a unit square.

4.6.13 Evolution of Pinwheels as Phase Singularities

We can also study the temporal evolution of the pinwheels, that is, the singularities in the field Z, e.g. by assuming that Z is a superposition of plane waves $Z = \sum_\kappa A_\kappa e^{i(\kappa\cdot a-\omega t)}$, where, as we saw in Sect. 4.6.6, ω is an angular frequency associated with the frequency $\nu = \omega/2\pi$ and the period $T = 1/\nu = 2\pi/\omega$, the corresponding wavelength being $\lambda = cT = c/\nu = 2\pi c/\omega$, where c is the speed of propagation.

Topological accidents can occur during this evolution, in particular when a singularity $Z = 0$ also satisfies $\Omega = \omega e_3 = 0$. We saw in Sect. 4.6.5 that, when $\Omega = 0$, the real ∇X et ∇Y have the same orientation, $\nabla Y = \alpha \nabla X$, $\alpha \in \mathbb{R}$, and hence that the complex gradient ∇Z is the real vector ∇X multiplied by a factor $1 + i\alpha$. As explained by Michael Berry and Mark Richard Dennis in [71], by a gauge transformation and a coordinate change, we can reduce to the case in which the field $Z(a)$ has the form

$$Z(a) = i\beta y + \frac{1}{2} a^T H a \ , \qquad \beta \in \mathbb{R} \ ,$$

where the complex Hessian

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12} & H_{22} \end{pmatrix}$$

is a symmetric 2×2 complex matrix. We do indeed have $Z(0) = 0$, and since $\nabla X(0) = 0$, $\Omega(0) = 0$.

This unstable normal form can be unfolded using a time variable t, giving the dynamical model

$$\begin{aligned} Z(a,t) &= t + i\beta y + \frac{1}{2} a^T H a \ , \\ X(a,t) &= t + \frac{1}{2}\left[x^2 \mathrm{Re}(H_{11}) + 2xy\mathrm{Re}(H_{12}) + y^2 \mathrm{Re}(H_{22})\right] , \\ Y(a,t) &= \beta y + \frac{1}{2}\left[x^2 \mathrm{Im}(H_{11}) + 2xy\mathrm{Im}(H_{12}) + y^2 \mathrm{Im}(H_{22})\right] . \end{aligned}$$

For the singularities $Z = 0$, the equation $Y = 0$ gives, for the lowest order terms, $y \sim -x^2 \mathrm{Im}(H_{11})/2\beta$ and the equation $X = 0$ gives $x^2 \mathrm{Re}(H_{11}) \sim -2t$. If $\mathrm{Re}(H_{11}) > 0$, then for $t < 0$, there are two solutions with abscissa values $x = \pm\left[-2t/\mathrm{Re}(H_{11})\right]^{1/2}$ on the parabola $y = -x^2 \mathrm{Im}(H_{11})/2\beta$ which coalesce and disappear for $t \geq 0$: two singularities with opposite chiralities $x\beta \mathrm{Re}(H_{11})$ annihilate in what is known as a fold catastrophe. If on the other hand $\mathrm{Re}(H_{11}) < 0$, then when t becomes positive, two singularities of opposite chirality appear.

We shall return to these pinwheel bifurcations in more detail in Sect. 4.7.1.2.

4.7 Pinwheel Singularities

We have seen that pinwheels are centred on singularities. In this section, we shall discuss some of the experimental data which supports the hypothesis that the functional orientation maps can be treated as phase fields.

4.7.1 Structure in the Vicinity of Singularities

4.7.1.1 Pinwheels as Zeros of the Orientation Field

If we enter the cortex 'vertically' at a regular point, we come across the columns of Hubel and Wiesel with their redundancy and their population coding. However, when we enter at a singular point, we come across neurons of every orientation. It is tricky but crucial to understand this phenomenon correctly. Orientation selectivity is highly variable in $V1$, and clearly, on the level of the mesoscopic resolution of the images presented so far, it is necessarily rather low or even zero on average at singular points, since all orientations are present in a roughly uniform manner. In fact, we may assume that the meso- or 'coarse-grained' level is one where the preferred orientation corresponds to what is known as an *order parameter* in statistical physics, and that the models of this level are analogous to the so-called *mean field* models.

As an example of *mesoscopic models*, let us return to those proposed by Wolf and Geisel [72] to understand the *learning process* behind orientational selectivity (see Sect. 4.6.1). They are phase field models in the sense of Sect. 4.6. The field of preferred orientations $\Phi(a) = \mathrm{e}^{\mathrm{i}\varphi(a)}$ of simple $V1$ neurons can be treated as a section of the fibre bundle $\pi : R \times P \to R$ defined outside the lattice L of the pinwheel centres of R. Identifying R with $\mathbb{R}^2$ and $\mathbb{R}^2$ with $\mathbb{C}$, Wolf and Geisel suggested modelling the columns by a continuous *complex field* $Z(a) = r(a)\mathrm{e}^{\mathrm{i}\varphi(a)} = r(a)\Phi(a)$ on the complex variable $a = \rho\mathrm{e}^{\mathrm{i}\theta}$, where the spatial phase $\varphi(a)$ thus encodes the orientation preference $\psi = \varphi/2$ and where the modulus $r(a) = |Z(a)|$ encodes the strength of selectivity, i.e. the width of the response curve, which is Gaussian. The complex value Z is thus a section of the fibre bundle $\pi : \mathbb{C} \times \mathbb{D} \to \mathbb{C}$, which has base space $\mathbb{C}$, and if the maximal selectivity is normalized to unity, its fibre is the unit disc $\mathbb{D}$. As we have seen, the singular points (pinwheel centres) then correspond to the *zeros* of Z. To obtain a naive and simple illustration of the behaviour of the field Z in the neighbourhood of a zero, we take $\varphi = \theta$ and $r = \rho/2$. Above the points a of a small circle C_ρ with centre 0 and radius ρ fixed in the base space $\mathbb{C}$, we have in the fibre the circles $\Sigma_{\rho/2}$ of fibres $\mathbb{D}_a$ of radius $\rho/2$, and the inverse image of C_ρ under π is thus the torus $\pi : C_\rho \times \Sigma_{\rho/2} \to C_\rho$. The lift of C_ρ on this torus is the curve Γ_ρ (see Fig. 4.79) given parametrically by

$$\Gamma_\rho = \left(\frac{\rho}{2} \sin(\theta)\,,\, \rho\left[1 - \frac{1}{2}\cos(\theta)\right]\cos(\theta)\,,\, \rho\left[1 - \frac{1}{2}\cos(\theta)\right]\sin(\theta) \right) .$$

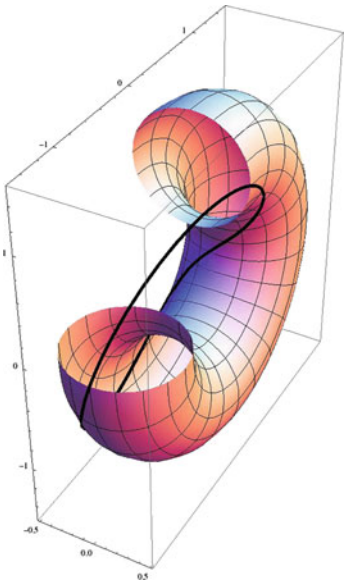


Fig. 4.79 The torus $\pi : C_\rho \times \Sigma_{\rho/2} \to C_\rho$ and the lift Γ_ρ of C_ρ (see text)

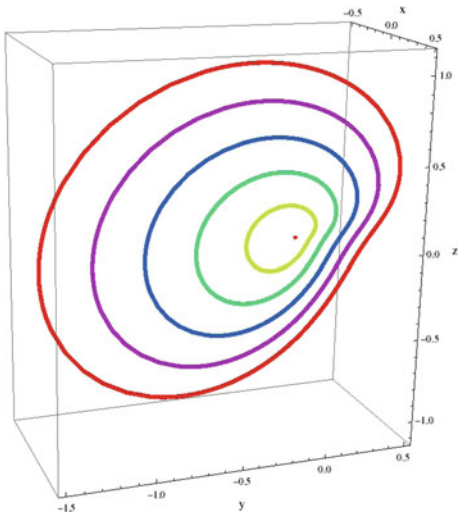


Fig. 4.80 When $\rho \to 0$, the section $\Gamma_\rho \to 0$

As the orientation selectivity vanishes at 0, when $\rho \to 0$, the section Γ_ρ also tends to 0 and the projection π is locally a diffeomorphism (see Fig. 4.80).

4.7.1.2 Pinwheel Morphogenesis and Learning

One of the advantages of this approach is that it provides a good theory of the learning process. The paper [78] by Nicholas Swindale contains a summary of these development models. As the author puts it:

> The repetitive stochastic patterns of eye dominance and orientation preference [are of an] intriguing nature. Many aspects of their development seem likely to be dependent upon both spontaneous and visually driven patterns of neural activity. […] Remarkably simple models, based on Hebbian synaptic plasticity, intracortical interactions and competitive interactions between cells and growing axons, have been able to explain much of the phenomenology.

A good example of such models is the one developed by James Bednar and Risto Miikkulainen in [79, 80] using self-organized map (SOM) models (see Sect. 4.10.1.4), which are[16]:

> […] networks of simple artificial neurons with initially unspecific connections that are modified by Hebbian learning and homeostatic plasticity.

They show quite clearly how the double constraints imposed on the one hand by the geometry and statistics of *external stimuli* and, on the other, by the geometry and statistics of the *internal structure* of the cortical areas lead in a self-organized way to pinwheel maps that depend only weakly on the initial synaptic weights of the neurons, but heavily on the way the stimuli arrive. Figure 4.81 gives an example.

The dynamical self-organization models proposed by Fred Wolf, Theo Geisel, Matthias Kashube, and Michael Schnabel in [72, 81, 82] are models of pattern formation, starting from an unstructured initial state, through *Turing instabilities*. They

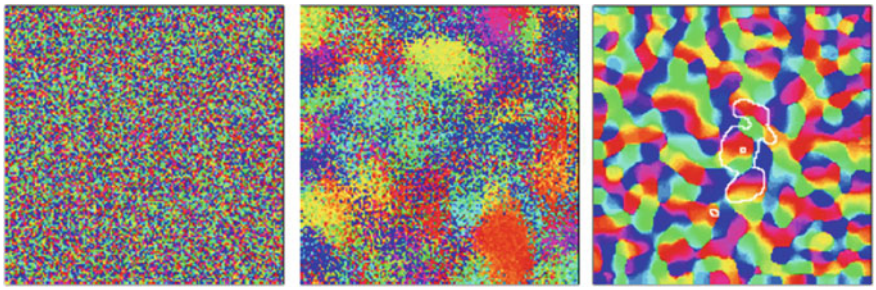


Fig. 4.81 Pinwheel map built from GCAL self-organized maps (gain control, adaptive, laterally connected SOMs). Starting from random synaptic weights, the neural network is subjected to a flow of stimuli of a certain type. The weights are modified by Hebbian learning and homeostatic plasticity, gradually producing a pinwheel structure which stabilizes. It can be shown that such maps depend little on the initial synaptic weights of the neurons, but a great deal on the flow of stimuli. In the last figure, *white lines* indicate the inhibitory lateral connections of a neuron. Figure adapted from Bednar [79]

[16] For Hebb's law, see Sect. 3.6.2.

adopt the development model first introduced by Nicholas Swindale and describe the morphogenesis of the field by a partial differential equation (PDE) of the form

$$\frac{\partial Z(a,t)}{\partial t} = F\big(Z(a,t)\big) + \eta\,(a,t)\ ,$$

where F is a linear operator and η a stochastic term representing the intrinsic fluctuations associated with this activity. This kind of dynamics can be induced by a Hebbian learning process, with $F\,(Z(a))$ the average of $Z(a)$ for rapidly changing stimuli A with a certain probability distribution.

The authors also take into account ocular dominance and distortion of retinotopy (see Figs. 4.12 and 4.13 in Sect. 4.2). They show that, starting from an initial state in which the field $Z(a,0)$ is weak (with little orientational selectivity), there is first a period of proliferation of pinwheels, with a growth in selectivity and the appearance of a characteristic wavelength Λ, the pinwheel density[17] becoming greater than π/Λ^2. Then, the nonlinearities and long-range lateral interactions stabilize the process and the number of pinwheels through their displacements, collisions, and annihilations of pairs of pinwheels of opposite chirality (see Fig. 4.82). The reader could consult Ha Youn Lee et al. [83].

More precisely, the standard technique for analyzing such a PDE is to expand $F\,(Z)$ in a series

$$F\,(Z) = L\,(Z) + F_2\left(Z, \overline{Z}\right) + F_3\left(Z, Z, \overline{Z}\right) + \cdots\ ,$$

where $L\,(Z)$ is linear, $F_2\left(Z, \overline{Z}\right)$ is bilinear, $F_3\left(Z, Z, \overline{Z}\right)$ is trilinear, and so on.[18] In the given model, it can be shown that the linearized $L\big(Z(a,t)\big)$ whose spectral analysis is to be carried out is of fourth order and can be written $L = \mu - \left(k_c^2 + \Delta\right)^2$, where μ is a bifurcation parameter such that, for $\mu < 0$, the homogeneous base state $Z \equiv 0$ is stable, while for $\mu > 0$, it becomes unstable and bifurcates towards

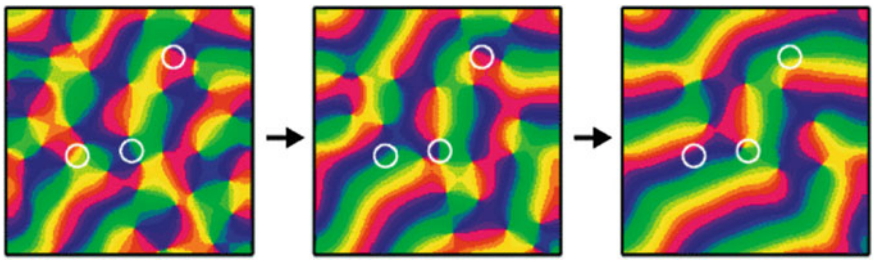


Fig. 4.82 Annihilation of pinwheels of opposite chirality simplifies and stabilizes the pinwheel geometry. From Wolf and Geisel [72]

[17] We have already encountered this density π/Λ^2 in Sect. 4.6.11.

[18] We have to use the two variables Z and $\overline{Z}$ because the functions F_j are not necessarily analytic functions depending only on Z.

patterns with characteristic wavelength $\Lambda = 2\pi/k_c$ for critical eigenvalues $\lambda(k_c)$ ($k_c \neq 0$) of the Fourier representation of L given by $\mu = \lambda(k_c)$.[19] To third order, the dynamical PDE becomes analogous to the equation that Jack Swift and Pierre Hohenberg introduced into the study of convection in hydrodynamics (see [84]).

It is important to note that the PDE must be *equivariant* under the action of the group of symmetries of the plane $SE(2)$, and even $E(2)$. We shall return at length in the second volume on the structure of the group $SE(2)$ and the importance of symmetries. Let us just say for the moment that they impose strong constraints on the form of the functional F. Indeed, F must be equivariant under the action of $SE(2)$, acting by translations $a \to a + b$ and rotations $a \to R_\theta(a)$ on the base plane $R = \mathbb{C}$ of the a, and also by rotations $Z \to e^{i\psi} Z$ in the fibres. Now recall that, if $Z(a)$ is a function of a variable $a \in R$ and if T is an element of a group G of transformations of R, T acts on $Z(a)$ by $\widetilde{T}(Z(a)) = Z\left(T^{-1}(a)\right)$. In addition, we say that a function or functional $F(Z)$ on which G operates is G-equivariant if $T\big(F(Z)\big) = F\big(\widetilde{T}(Z)\big)$, i.e. if F commutes with the actions of G on Z and on the space of values of F. Consequently, we must have

$$\widetilde{b}\big(F(Z(a))\big) = F\big(\widetilde{b}(Z(a))\big) = F(Z(a-b)) \ ,$$

for any translation $\widetilde{b}$ by a vector b in the plane R,

$$\widetilde{R_\theta}\big(F(Z(a))\big) = F\big(\widetilde{R_\theta}(Z(a))\big) = F\big(Z(R_{-\theta}(a))\big) \ ,$$

for any rotation $\widetilde{R_\theta}$ through angle θ of the plane R, and $F\left(e^{i\psi} Z\right) = e^{i\psi} F(Z)$ for any rotation of Z. Note that the last symmetry implies that $F(0) = 0$, since if $Z = 0$, then $e^{i\psi} Z = 0$ for any ψ, whence $e^{i\psi} F(0) = 0$ for all ψ.

When F is expanded in the neighbourhood of the equilibrium state $Z \equiv 0$, its linear part L commutes with the symmetries, and the Fourier representation is therefore diagonal with eigenvalues $\lambda(k_c)$ which depend only on the magnitude k_c. The authors show that the second-order term F_2 can be neglected and that we may choose a third-order term F_3 depending on $|Z|^2 Z$ and long-range lateral connections. We shall see in the second volume that this cubic normal form $|Z|^2 Z$ often arises in problems with $E(2)$ symmetry.

The analysis of patterns $Z(a)$, i.e. pinwheel maps, that can emerge by bifurcation thus reduces to a classic problem that has been widely studied in physics, the problem of *symmetry breaking*. If a modulus (wave number) k is selected by a bifurcation at μ, we may consider superpositions of eigenstates (Fourier modes) with eigenvectors (wave vectors) κ_j with magnitude $|\kappa_j| = k$. We thus obtain *planforms*:

[19] In the expansion of L, the powers of partial derivatives like $(\partial/\partial x)^n$ mean repeated differentiation $\partial^n/\partial x^n$. The linear PDE $\partial Z/\partial t = L(Z)$ for $L = \mu$ describes exponential damping towards 0 for $\mu < 0$ (stability) and exponential growth for $\mu > 0$ (instability). The PDE $\partial Z/\partial t = L(Z)$ for $L = -\Delta$ is a diffusion equation.

$$Z(a) = \sum_{j=0}^{j=n-1} A_j e^{i\kappa_j \cdot a} , \quad |\kappa_j| = k , \quad \kappa_j = k\left(\cos\left(\frac{2\pi j}{n}\right), \sin\left(\frac{2\pi j}{n}\right)\right) .$$

The symmetries impose a precise structure on the amplitudes A_j. By expressing the fact that they are solutions of the PDE satisfied by Z, we show that they must satisfy differential equations of the form

$$\frac{dA_i}{dt} = A_i - \sum_{j=1}^{j=n} g_{ij} |A_j|^2 A_i - \sum_{j=1}^{j=n} f_{ij} A_i A_{j^*} \overline{A_{i^*}} ,$$

where j^*, the mode antiparallel to the mode j, is defined by $\kappa_{j^*} = -\kappa_j$, and where the coefficients g_{ij} and f_{ij} satisfy the properties

$$\begin{cases} g_{ij} = \left(1 - \frac{1}{2}\delta_{ij}\right) g\left(|\alpha_i - \alpha_j|\right) , \\ f_{ij} = \left(1 - \delta_{ij} - \delta_{i^* j}\right) f\left(|\alpha_i - \alpha_j|\right) . \end{cases}$$

Here, δ_{ij} is the Kronecker symbol ($\delta_{ii} = 1$ and $\delta_{ij} = 0$ if $i \neq j$), α_i is the angle of mode i, and $g(\alpha)$ and $f(\alpha)$ are π-periodic angle functions calculated from F. A refined version of this model can be found in the paper Schnabel et al. [85].

Figure 4.83, taken from [81, 82], shows the planforms for $n = 1, 2, 3, 5, 15$, with the positions of the wave vectors κ_j and the density of the pinwheels as a function of n. An example of a pinwheel map constructed as a linear combination of planforms is also shown.

4.7.1.3 Pinwheels as Genuine Singularities

These models based on the parallel between pinwheels and dislocations are quite remarkable. However, two points should be emphasized:

1. The model assumes that the orientation selectivity vanishes on the dislocations located at the centres of the pinwheels.
2. No meaning can be attached to the limit as the mesh of the pinwheel lattice tends to 0, since it then gives a field Z that is identically zero.

It thus requires closer examination because it may be that the selectivity remains good at singular points on the *micro* level of individual neurons and only vanishes at the *meso* level due to averaging. As noted by Jonathan Polimeni et al. when commenting on a paper by Pedro Maldonado et al. [86] which we shall discuss in a moment [87, p. 4158]:

> The optical recording signal is a population average, and so individual neurons in the pinwheel centres—even if they were strongly tuned to orientation—would average to a weak population response at the pinwheel centres imaged through optical recording.

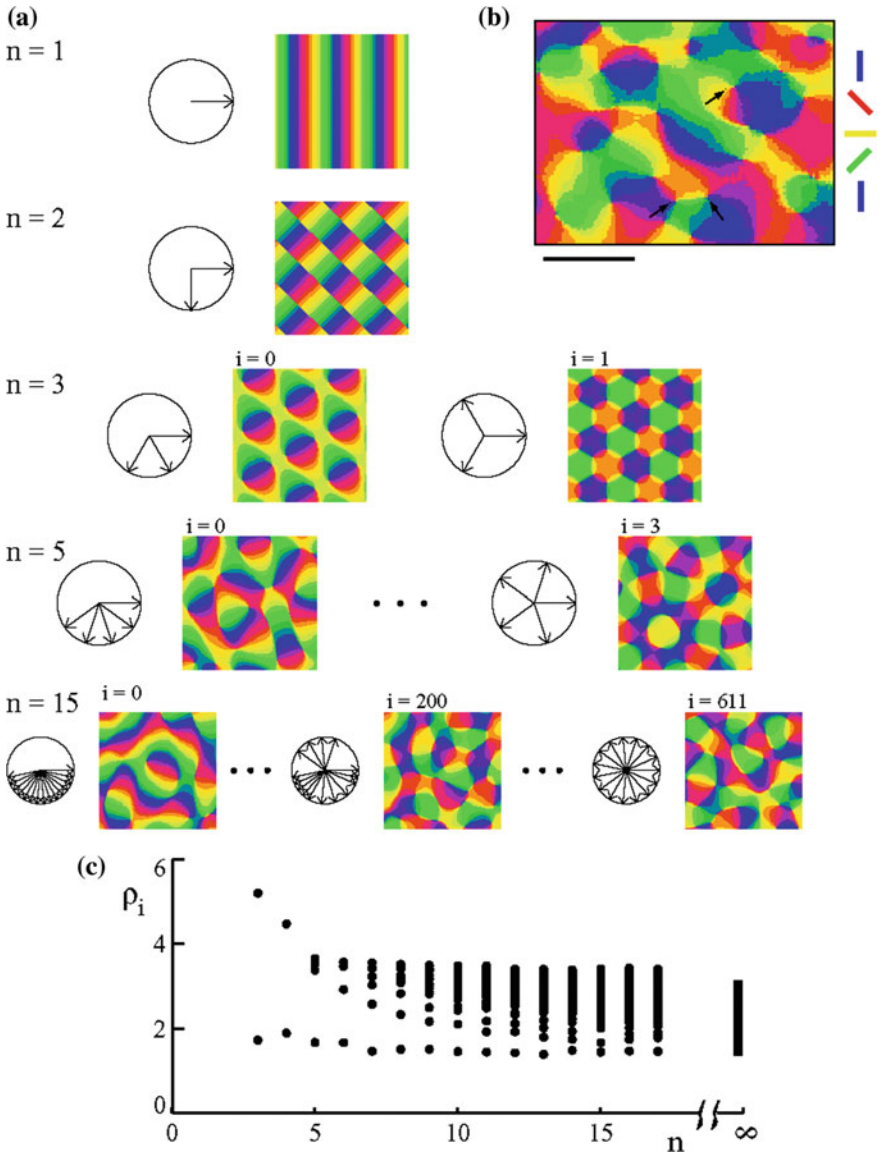


Fig. 4.83 Model by Wolf and Geisel. **a** Planforms for $n = 1, 2, 3, 5, 15$, with the position of the wave vectors κ_j. **b** Example of a pinwheel map built as a linear combination of planforms. **c** Diagram giving the density ρ of the pinwheels as a function of n : for a given n, each point is the density of one of the patterns corresponding to n. From [82]

The difference is clear on the level of the geometry of the models. Let us come back to Fig. 4.79 and suppose, for example, that the selectivity r is everywhere maximal, viz. $r = 1$, even at the origin. Then, the lift of C_ρ is the curve Γ_ρ given by

$$\Gamma_\rho = \left(\frac{\rho}{2}\sin(\theta), \left(1 - \frac{1}{2}\rho\cos(\theta)\right)\cos(\theta), \left(1 - \frac{1}{2}\rho\cos(\theta)\right)\sin(\theta)\right),$$

and when $\rho \to 0$, the section Γ_ρ no longer tends to 0 but to the 'vertical' unit circle $\Gamma_0 = (0, \cos(\theta), \sin(\theta))$, and the projection π is not at all a local diffeomorphism at 0. It is such away from 0, even if it turns out to be a highly twisted diffeomorphism, but at 0 itself, it has an 'exceptional fibre' Γ_0 of dimension 1 (see Fig. 4.84).

It is thus well worth carrying out precise experiments on the structure of the orientation field in the neighbourhood of singularities. They are difficult to do but rich in results. By combining imaging methods with methods for intracellular recording of spikes triggered by the synaptic inputs, but also recording of membrane potentials of single neurons, Pedro Maldonado, Imke Gödecke, Charles Gray, and Tobias Bonhöffer were able to analyze the fine structure of the orientation maps at singularities, observing that [86, p. 969]:

> Orientation columns contain sharply tuned neurons of different orientation preference lying in close proximity.

Put another way, column redundancy seems to disappear at singular points.

For further discussion of these results, the reader is referred to David McLaughlin et al. [88] and Michael Shelley et al. [89], who model the cortical processes generating orientation selectivity in the $4C\alpha$ layer of the macaque $V1$ area. They start from the microlevel with a network of 16 000 integrate-and-fire neurons and reduce it at the meso level to what they refer to as 'a spatially coarse-grained system for firing rates of

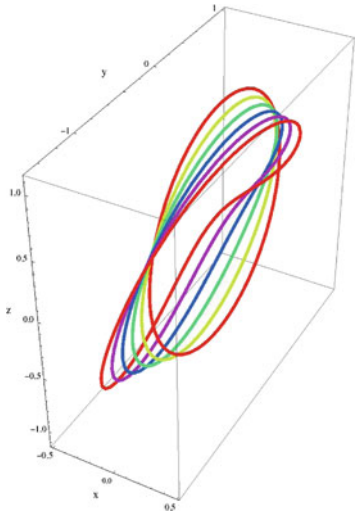


Fig. 4.84 When $\rho \to 0$, the section Γ_ρ no longer tends to 0 but to the 'vertical' unit circle Γ_0 (see text)

neuronal subpopulations' [89, p. 97]. They then introduce a plausible connectivity for the 'vertical' connections coming from the LGN and for the 'horizontal' intracortical inhibitory and excitatory connections (which we shall return to at length in Sect. 5.1 of Chap. 5). In the second volume, we shall discuss cortical models comprising a huge number of differential equations describing the membrane potentials of each neuron in the network, along with coarse-grained mean field methods which allow us to study them on the meso level. What is important here is the result from these authors, formulated here for an orientation hypercolumn, i.e. a pinwheel, including isotropic internal inhibitory and excitatory connections, but no anisotropic long-range lateral connections with other hypercolumns:

> This analysis showed it is an interaction between the pinwheel structure of the preferred orientation mapping and the isotropic architecture that produces greater orientation selectivity near pinwheel centres [89, p. 121].

In other words, these models show that the pinwheels are 'true' singularities where all the orientations are present, *each with strong selectivity*.

Many experimental results confirm this idea. For example, using moving gratings as stimuli, James Schummers et al. have shown that:

> Neurons near pinwheel centres have sub-threshold responses to all stimulus orientations but spike responses to only a narrow range of orientations. [90, p. 969]

They assess the selectivity of spike responses and sub-threshold responses as a function of position relative to the pinwheels. Far from the pinwheels:

> [Cells] show a strong membrane depolarization response only for a limited range of stimulus orientation, and this selectivity is reflected in their spike responses. [90, p. 970]

At the centre of a pinwheel, on the other hand, only the spike response is selective and the membrane is highly depolarized for all orientations (see Fig. 4.85).

These results show that, at singular points, all orientations are indeed present, but using a novel solution with regard to the connectivity of the relevant neural microcircuits:

> These examples indicate that both simple and complex cells located near pinwheel centres receive synaptic inputs over a broad range of stimulus orientations, although not all of these inputs are represented in the spike outputs. [90, p. 971]

Schummers [91] obtained further results with Jorge Mariño (on cats). The local neural circuits are very different at regular points and at singular points. If we look at the total conductance g, we observe that it is strongly peaked at a regular point and flatter at a singular point. It is the different ways in which it is distributed between inhibitory and excitatory conductances g_i/g_e and the interaction between the excitatory and inhibitory connections which show us how, 'despite the diversity of local environments', there can nevertheless be 'a sharp orientation tuning at all locations in the orientation map' [91, p. 195]. With highly localized injections of retrograde tracer (where 'localized' refers to a region <100 μm), the difference between excitatory and inhibitory can be clearly distinguished. Figure 4.86 shows

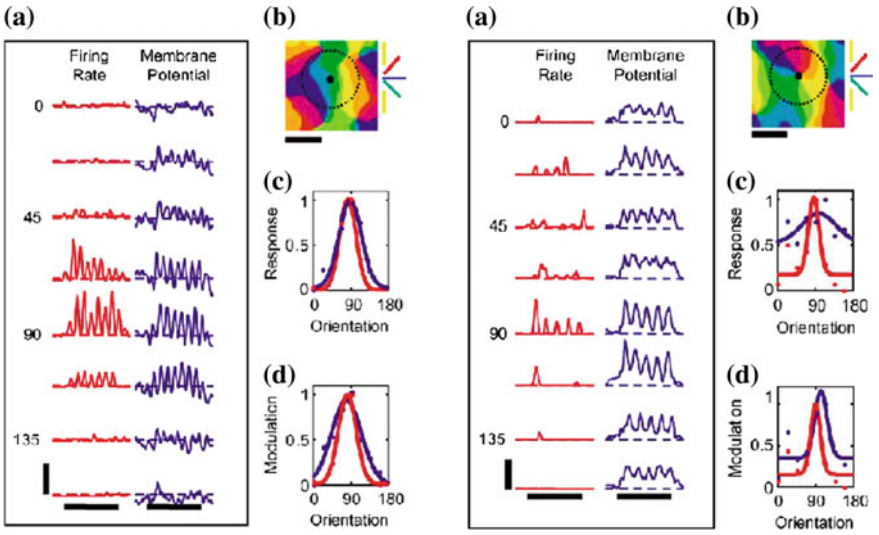


Fig. 4.85 Responses to stimuli in the form of moving periodic gratings with different orientations. *Left* Far from singularities. Scales 8 spikes/s, 10 mV, 2 s. *Right* At the centre of a pinwheel. Scales 3 spikes/s, 8 mV, 2 s. **a** Comparison between spike emissions (*red*) and the membrane potential (*blue*). **c** The two curves representing the average amplitude of the responses for different orientations. **d** Modulation of responses. From Schummers [90]

an injection site with the inhibitory and excitatory connections. We observe that inhibition is more localized and more isotropic, whereas excitation propagates over greater distances, but in a highly *anisotropic* way, clustering in regions of the same colour and hence of the *same orientation* as the injection site. We shall return to this crucial point, the key to the functional architecture of $V1$, in Sect. 5.1 of Chap. 5.

These measurements confirm that the difference between regular and singular points is not reflected in the level of the spikes (there is good orientation selectivity in both cases), but at the level of the sub-threshold signals. This is because:

> […] as a result of the appropriate inhibitory balance at orthogonal orientations, which keeps the membrane potential below threshold, it is not reflected in the spike responses.

4.7.2 The Problem of Resolution

As already mentioned, it is important to note that the question of *resolution* plays an important role in the definition of the pinwheels. In an optical device, there is always a resolution transforming points (Dirac measures) into Gaussians of a certain width σ. In the optical imaging of intrinsic signals in vivo, there are systematic statistical errors. As noted by Jonathan Polimeni et al., these are due to:

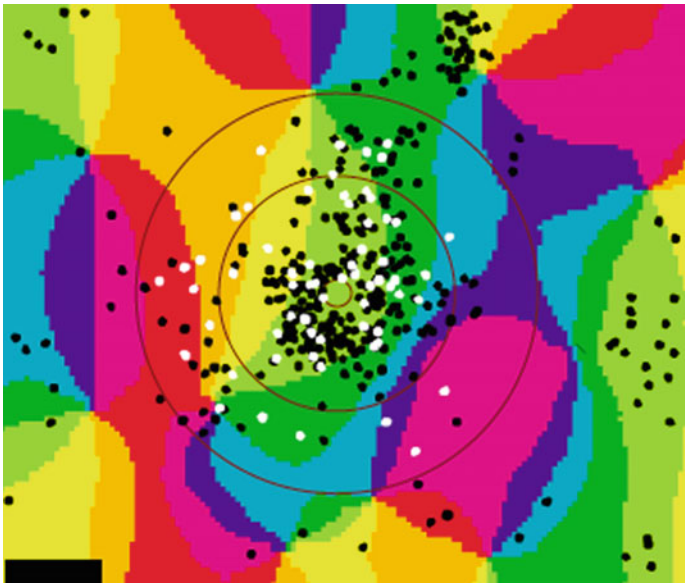


Fig. 4.86 The injection site is the *small central circle*. The *outer circle C* is at 250 μm from the *intermediate circle* (see the 250 μm *scale bar* in the *bottom left corner*). *White dots* represent inhibitory connections and *black dots* excitatory connections. We observe that inhibition is more localized, bounded by the circle *C*, and more isotropic, whereas excitation propagates well outside the circle *C*, but in a highly anisotropic way, clustering in regions of the same colour and hence of the same orientation as the injection site. From Mariño and Schummers [91]

> […] photon scatter and absorption in brain tissue combined with the blurring introduced by the optics of the imaging system. [87, p. 4158]

In this sense:

> Optical recording, as it has been used to date, has insufficient spatial resolution to accurately locate pinwheel centers. [87, p. 4158]

Fortunately, as we shall soon see, there are now much higher resolution methods which provide a way around these difficulties.

However, it remains important to take into account the fact that the functional orientation maps are obtained by superposing single-orientation maps and are thus vector-valued images. They have a resolution corresponding to a width σ of about 250 μm. Figure 4.87, taken from [87], shows what happens to a pinwheel map when it is convolved with a Gaussian kernel of width 250 μm. Obviously, we observe that the map changes significantly. In other words, good *mesoscopic* models must always be *multiscale* models.

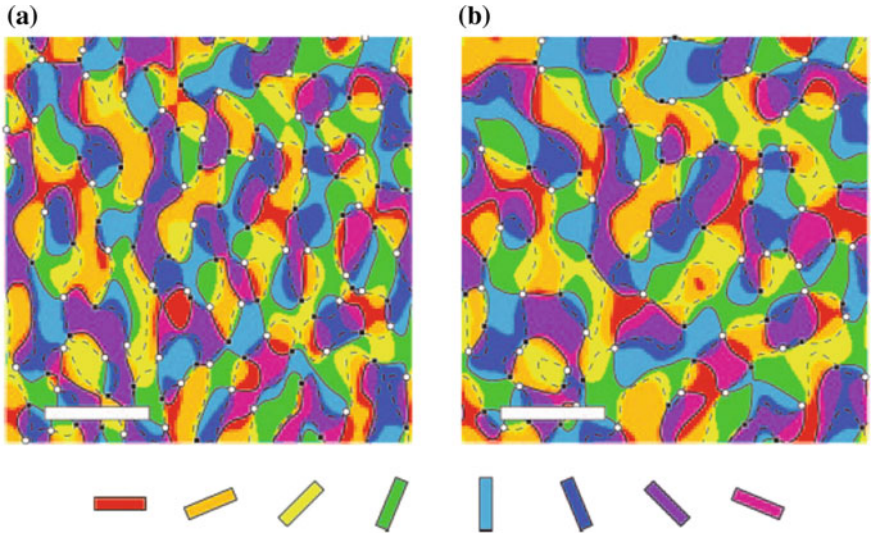


Fig. 4.87 Transformation of a pinwheel map by convolution with a Gaussian kernel of width 250 μm. *White pinwheels* are (+), *black pinwheels* are (−), and *continuous* and *dashed lines* are the zero-crossings of the map. The *scale bar* represents 1 mm. We note how much the map changes. **a** Initial map with 124 pinwheels. **b** Transformed map with only 93 pinwheels. From Polimeni et al. [87]. Copyright (2005) National Academy of Sciences, USA

4.7.3 Two-Photon Confocal Microscopy

The imaging techniques mentioned up to now are not accurate enough. However, recently, new techniques like *in vivo* imaging by *two-photon confocal microscopy* have given us functional maps with a resolution down to the level of single neurons. Kenichi Ohki and coworkers [92] have shown that, in cats, pinwheels defined on the mesolevel remain highly ordered on the microlevel. Consequently:

> Pinwheel centers truly represent singularities in the cortical map.

The idea of the method is to inject calcium-sensitive indicators (Oregon Green BAPTA-1 acetoxylmethyl ester) which label a few thousand neurons in regions of 300–600 μm. We simultaneously measure calcium signals triggered by visual stimuli in hundreds of neurons at different depths (from 130 to 290 μm, in steps of 20 μm), and we find pinwheels with the same orientation wheel (see Fig. 4.88):

> This demonstrates the columnar structure of the orientation map at a very fine spatial scale.

Whence the problem of connectivity implementing the fine selection of orientations near singularities; several hypotheses have been put forward by the authors about the dendritic tree near the centre C (a few tens of μm) in an iso-orientation domain D (see Fig. 4.89):

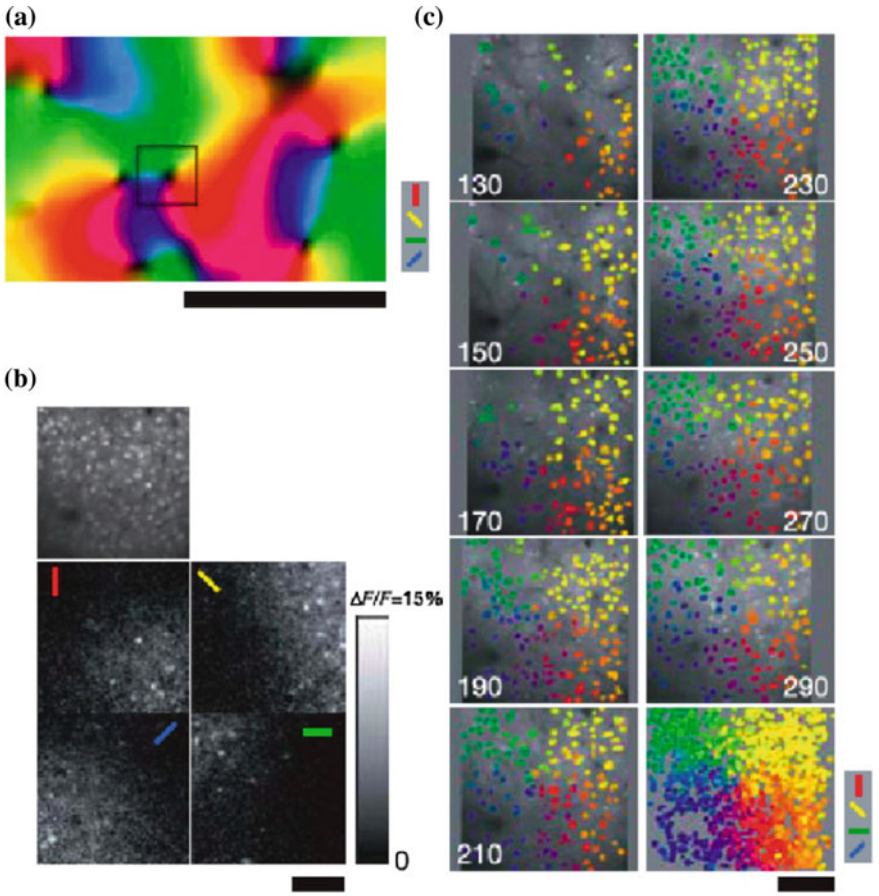


Fig. 4.88 Two-photon optical imaging can simultaneously measure calcium signals triggered by visual stimuli in hundreds of neurons at different depths (from 130 to 290 μm in steps of 20 μm). We find analogous pinwheels with the same orientation wheel at these different depths. From Ohki et al. [92]

(a) Unbalanced dendritic tree towards D.
(b) Symmetric dendritic tree, but excitatory inputs unbalanced towards D.
(c) Symmetric dendritic tree, symmetric excitatory inputs, but local and within D (good segregation in the vicinity of C).
(d) Symmetric dendritic tree, symmetric excitatory inputs, integrated over a wide dendritic area.

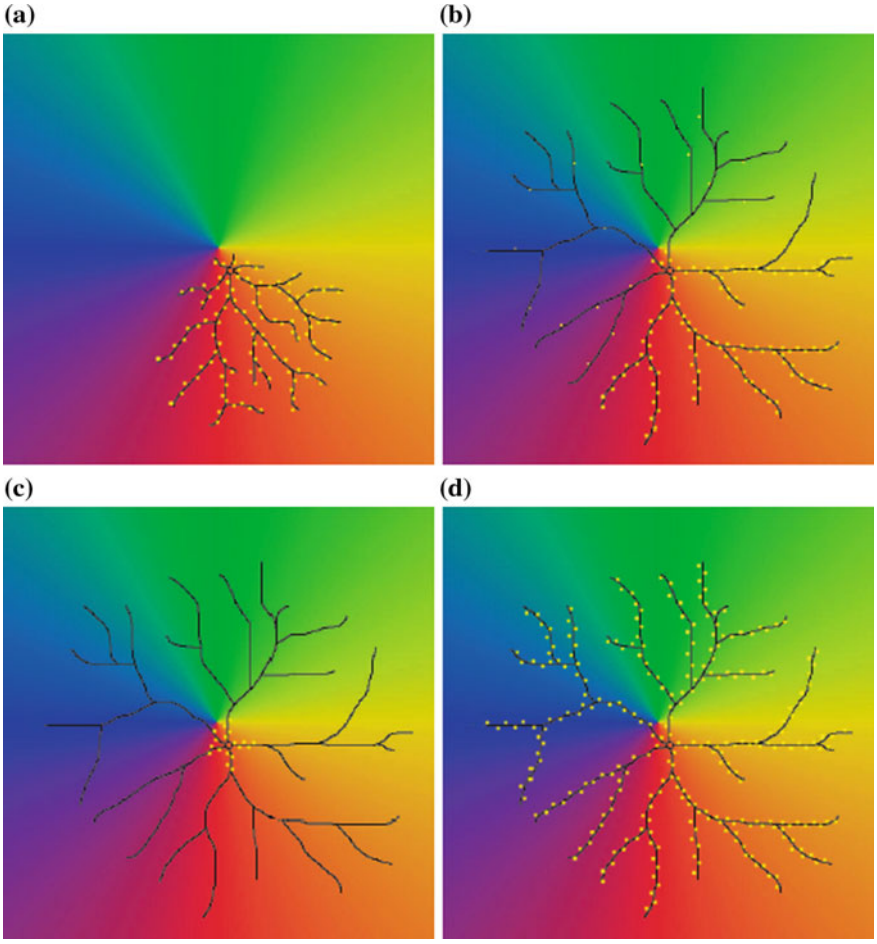


Fig. 4.89 Four dendritic tree structures in the vicinity of a singularity, as proposed by Kenichi Ohki et al. [92] (see text)

4.8 Pinwheels and Blow-ups

The results of Schummers and Maldonado lead to two competing idealized models. In both cases, all the orientations are indeed present with good selectivity in the neighbourhood of singularities, but at the singularity itself (insofar as this has a precise meaning on the meso level), either the orientation selectivity is sub-threshold (Schummers), or there is a loss of column redundancy (Maldonado). Further experiments will be needed to elucidate this, but we may nevertheless make a suggestion for Maldonado's results. The theoretical problem we have to solve concerns the dimensional collapse of an ideally 3D structure onto 2D neural layers.

4.8.1 The Geometric Concept of Blow-up

So how can the concrete 2D structure of pinwheels fit with the abstract 3D structure of the fibre bundle? To understand this, we shall use the geometric operation known as 'blowing up' of singularities. The intuitive idea, which needs to be carefully tested, is that the pinwheels could be like local blow-ups of singular points, and the pinwheel lattice like a gluing together of these local blow-ups to give a discrete approximation of the projection $\pi : R \times \mathbb{P}^1 \to R$. In this picture, π would correspond to a limiting situation where in some sense all the points of R would blow up in parallel.

In algebraic geometry, the blow-up of a variety such as the plane $M = \mathbb{R}^2$, at a point, e.g. the origin $O = (0, 0)$, is defined in the following way. Let $a = (x, y) \neq (0, 0)$ be a point of $\mathbb{R}^2$. We associate the direction p of Oa with this point and thus specify a map δ :

$$\begin{aligned} \delta : \mathbb{R}^2 - \{O\} &\longrightarrow \mathbb{P}^1 \\ a = (x, y) &\longmapsto \delta(a) = p = y/x \end{aligned}$$

The graph of δ is a helicoidal surface H in the 3D fibre bundle $V = \mathbb{R}^2 \times \mathbb{P}^1$, and the topological closure of H in V is a helicoid $\overline{H}$ with $\pi^{-1}(O) = \Delta \cong \mathbb{P}^1$. This copy of $\mathbb{P}^1$ above O is an exceptional fibre which, for historical and technical reasons, is called the *exceptional divisor* of the blow-up. The restriction to $\overline{H}$ of the projection $\pi : \mathbb{R}^2 \times \mathbb{P}^1 \to \mathbb{R}^2$ is an isomorphism of H on $\mathbb{R}^2 - \{O\}$. If d is the straight line generated by Oa in $\mathbb{R}^2$, the closure of the inverse image $\pi^{-1}(d - \{O\})$ is made up of points $(\lambda a,\ \delta(a) = p)$ of $V = \mathbb{R}^2 \times \mathbb{P}^1$, i.e., by the straight line d' at height $\delta(a) = p = y/x$. When the straight line d rotates in the plane $\mathbb{R}^2$, d' also rotates, but while translating through Δ, whence the helicoidal motion.

Since the inverse image of the point O under π is the projective line $\Delta = \mathbb{P}^1$, π is not at all an isomorphism at O, but a projection collapsing a 1D fibre to a zero-dimensional point. In this sense, the blow-up of the plane at a point generates a geometric structure which is somehow 'intermediate' between the 2D plane and the 3D fibre bundle. It is the fibre bundle V above O and the (twisted) plane $\mathbb{R}^2$ away from O. One might say that the blow-up $\pi : \overline{H} \to \mathbb{R}^2$ unfolds the orientation wheel centred on $O = (0, 0)$ in a third dimension (see Figs. 4.90 and 4.91).

This construction can be understood as an interpretation of *polar coordinates* in terms of a *field of directions*. Indeed, consider the fibre bundle $\pi_1 : \mathbb{R}^2 \times \mathbb{S}^1 \to \mathbb{R}^2$. On the plane with one point removed $\mathbb{R}^2 - \{O\}$, the argument $\theta(a) \in [0, 2\pi]$ of a point a is well defined and we may thus consider the section Θ_1 of π_1 defined by $\Theta_1 : a \to \Theta_1(a) = \left(a, \mathrm{e}^{\mathrm{i}\theta(a)}\right)$. The fibre bundle $\pi : \mathbb{R}^2 \times \mathbb{P}^1 \to \mathbb{R}^2$ is the quotient of the fibre bundle $\pi_1 : \mathbb{R}^2 \times \mathbb{S}^1 \to \mathbb{R}^2$ obtained by identifying θ with $\theta + \pi$ (i.e., by identifying $\mathrm{e}^{\mathrm{i}\theta}$ with $-\mathrm{e}^{\mathrm{i}\theta}$) and Θ_1 lifts to π_1 the section of π defined by $\Theta : a \to \Theta(a) = \left(a, \mathrm{e}^{\mathrm{i}\theta(a)}\right)$, where $\theta(a) \in [0, \pi]$ is now considered modulo π. $\Theta_1(a)$ is constant on the rays $\theta = \text{const.}$, and when it is lifted from $\mathbb{R}^2 \times \mathbb{P}^1$ to $\mathbb{R}^2 \times \mathbb{S}^1$, the surface H of $\mathbb{R}^2 \times \mathbb{P}^1$ becomes the image of Θ_1.

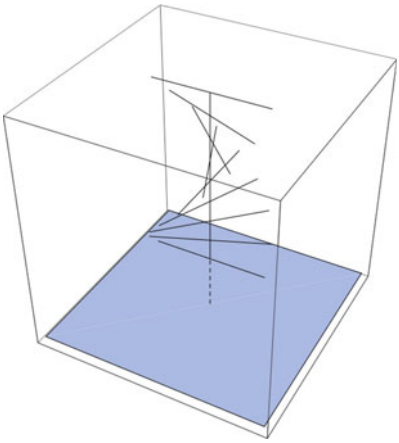


Fig. 4.90 Blow-up of the plane at a point a. The directions at a are unfolded in a third dimension

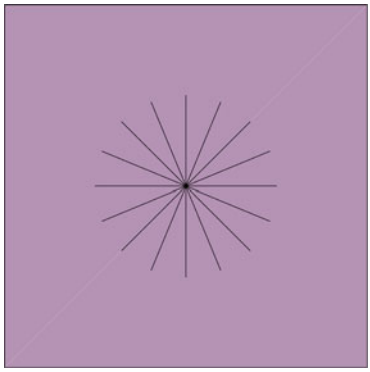


Fig. 4.91 When the third dimension collapses, the blow-up becomes a pinwheel

The concept of blow-up was introduced under the name of 'quadratic transformation' at the end of the nineteenth century by specialists of projective algebraic geometry. It represents the simplest case of what are known as birational transformations, and it is fundamental for desingularizing singular curves. If a curve γ in $\mathbb{R}^2$ has a singular point at O where several branches intersect with different tangents, then by lifting γ to $\overline{H}$, we obtain a curve $\Gamma = \pi^{-1}(\gamma)$ with various branches at different heights, thereby eliminating the intersections.

We can now *localize* this algebraic model and even consider an *infinitesimal version*, where we restrict to points $a = (\mathrm{d}x, \mathrm{d}y)$ infinitely close to the blow-up point $O = (0, 0)$. To do this, we take what is known as the 'germ' of the structure in the neighbourhood of (Δ, O). In the local model, which is no longer algebraic but differential, we have $p = \mathrm{d}y/\mathrm{d}x$ and the surface H is thus included in the kernel of the differential form $\omega = \mathrm{d}y - p\mathrm{d}x$ defined on $V = \mathbb{R}^2 \times \mathbb{P}^1$ (we shall return to this key point in Sect. 5.4.1 of Chap. 5). Conversely, the algebraic model can be considered as the 'tangent' structure to the local model, where infinitesimal segments are replaced by tangent vectors.

The blow-up model can be used to understand orientation singularities. We can neglect the redundancy of the cortical columns outside the singularities insofar as, at regular points, simple $V1$ neurons mainly detect the same pair (a, p). However, we have seen that, according to Maldonado et al. [86], this is not the case at singular points, where all orientations are in fact present in the column. The fact that the orientation associated with a ray of a pinwheel selects this orientation at the centre of the pinwheel can then be expressed by saying that the operation yielding the topological closure $\overline{H}$ of H is implemented neurally and that the singular point is thereby blown up throughout the thickness of the cortical layer.

When the pinwheel is modelled in this way, it corresponds to the section Ψ of the projection $\pi : \left(\mathbb{R}^2 - \{O\}\right) \times \mathbb{P}^1 \to \left(\mathbb{R}^2 - \{O\}\right)$ given by $\Psi(a) = \left(a, \mathrm{e}^{\mathrm{i}[\alpha+\theta(a)/2]}\right)$, where α is the orientation encoded by the ray $\theta(a) = 0$. We thus require a double rotation of θ through angle π (hence two round trips of $\mathbb{P}^1$) to get back to the same ray, and this is why, as explained above, two diametrically opposite rays at angles θ and $\theta + \pi$ correspond to orthogonal orientations $\alpha + \theta/2$ and $\alpha + \theta/2 + \pi/2$. This means that the pinwheels implement what is known in geometry as a *spin structure*. The section Ψ is not the section $\Theta : a \to \Theta(a) = \left(a, \mathrm{e}^{\mathrm{i}\theta(a)}\right)$, $\theta(a) \in [0, \pi]$, of the bundle $\pi : \left(\mathbb{R}^2 - \{O\}\right) \times \mathbb{P}^1 \to \mathbb{R}^2 - \{O\}$, but rather a field of orientations of the same order, and this justifies modelling it by a blow-up.

4.8.2 Blow-ups and Lines of Dislocations

A link can be made between these blow-ups and the singularities of the phase fields discussed in Sect. 4.6. On the one hand, if we rotate the x-axis by θ_0 without changing the origin in $\mathbb{P}^1$, the helicoid H is translated by θ_0 to the helicoid H_{θ_0} and the blow-up can be considered as the set of all the H_{θ_0}. On the other hand, we can consider the field of phases $Z = a\mathrm{e}^{-\mathrm{i}p} = \rho\mathrm{e}^{\mathrm{i}(\theta-p)}$ in $\mathbb{R}^2 \times \mathbb{P}^1$, where p varies in the fibre $\mathbb{P}^1$. Surfaces of constant phase are then the helicoids H_{θ_0} and the fibre $\mathbb{P}^1$ above 0 is a line of dislocations.

4.8.3 From Blow-up to Fibre Bundle

To construct a fibre bundle model for a global pinwheel structure from the local blow-up model, one must blow up all the points of an (irregular) lattice L of the plane, the lattice of singular points, *in parallel*. This is not possible in the framework of algebraic geometry, because the local structures determine the global structures, and to iterate the algebraic model for the blow-up of a point, one would have to immerse the successive blow-ups in spaces of ever higher dimension. However, it can be done in the context of differential geometry by gluing together the local models of different points in the lattice. We thus obtain a model for the pinwheel structure (see Figs. 4.92 and 4.93).

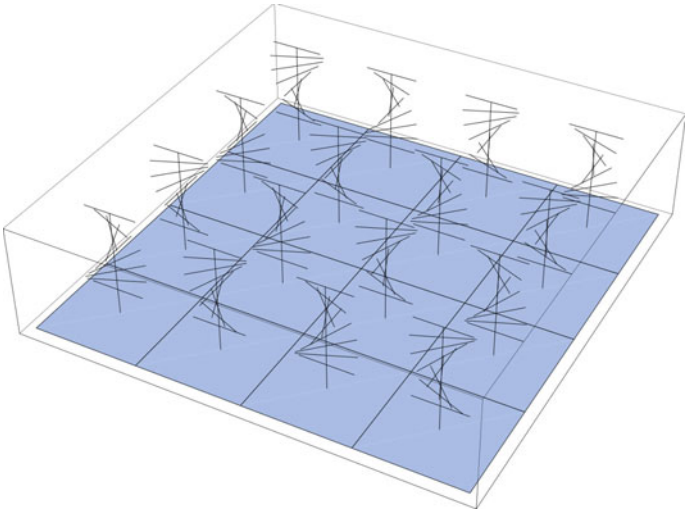


Fig. 4.92 Simultaneous blow-up of a lattice of points

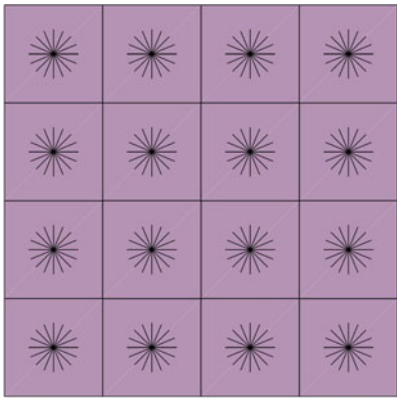


Fig. 4.93 When the third dimension collapses, a simultaneous blow-up of a lattice of points yields a lattice of pinwheels

In a field model like the one in Sect. 4.4.3, the field can be lifted from $\mathbb{R}^2$ to $V = \mathbb{R}^2 \times \mathbb{P}^1$ by blowing up the singularities c_i and lifting the field lines like curves: if a field line γ goes from a source c_1 with angle θ_1 to reach a sink c_2 with angle θ_2, γ lifts to a curve Γ in V starting at height θ_1 in the fibre above c_1 and arriving at height θ_2 in the fibre above c_2. Only the fibres above c_i are involved, so everything happens as though the c_i had been blown up in parallel. The structures tangent to this multiple blow-up in the neighbourhoods of the fibres $\Delta_i \to c_i$ are all isomorphic to the local algebraic model.

We could therefore consider that, when the mesh of the lattice L tends to 0, the limit of this multiple blow-up gives back the bundle $\pi : \mathbb{R}^2 \times \mathbb{P}^1 \to \mathbb{R}^2$, gluing back the infinitesimal models at all the points of $\mathbb{R}^2$. In this sense, the pinwheel structure

could effectively be interpreted as a discrete approximation of the bundle π, and conversely, π could be considered as the simultaneous blow-up of all the points in the plane. A good way could be to use a non-standard model[20] $(\mathbb{R}^*)^2$ of $\mathbb{R}^2$ where, around each standard point $a = (x, y)$, there would be a 'halo' or a 'monad' of infinitesimals $\mu(a) = \{(x + \mathrm{d}x, y + \mathrm{d}y)\}$. In the blow-up, the exceptional fibre Δ^* would then be $\left(\mathbb{P}^1\right)^*$ and $\mathrm{d}y/\mathrm{d}x$ a non-standard real number $p^* = p + \mathrm{d}p$, equivalent to $p \in \mathbb{P}^1$. The field lines would all lie within the monads, and so would be non-standard, and to first order, there would only remain the segments $a \to (a + \mathrm{d}a)$. The standard part of the structure at $\Delta^* \to a$ would thus be the tangent structure. In this way, we would obtain not only the bundle $\pi : \mathbb{R}^2 \times \mathbb{P}^1 \to \mathbb{R}^2$, but also the infinitesimal structure defined on it by the differential 1-form $\omega = \mathrm{d}y - p\mathrm{d}x$, i.e. the *contact structure*, to which we shall return.

The advantage of a non-standard model is that it provides an intuitive picture of the characteristic dimensional reduction of the pinwheel structure. We take the bundle $\pi : \mathbb{R}^2 \times \mathbb{P}^1 \to \mathbb{R}^2$, compactify the fibres[21] to make them infinitesimal, then project them into the monads $\mu(a)$. I used to think that the idea of blow-up with the exceptional fibre made infinitesimal and projected back onto the base plane was original. However, reading the correspondence between Pierre Deligne, Bernard Malgrange, and Jean-Pierre Ramis on 'irregular singularities', I found a letter dated 7 January 1986 on the singularities of analytic functions and Gevrey classes, where Deligne [94] introduced the concept of 'thick points'. The idea is to replace a point $a \in \mathbb{C}$, say $a = 0$, by a small disc D with boundary Δ, consider the space $\widetilde{\mathbb{C}} = \mathbb{C}^* \cup D$, union of $\mathbb{C}^* = \mathbb{C} - \{0\}$ and D, and equip it with the topology of the blow-up of 0 in $\mathbb{C}$ along $\mathbb{C}^* \cup \Delta$. Moreover, in his last paper on Gevrey classes (edited by Jean-Pierre Ramis [95]), Martinet [96] used this construction with the discs D that are infinitesimal in the sense of non-standard analysis.[22]

We could thus say that, in the continuous limit, a pinwheel lattice model amounts to treating the points of the plane as infinitesimal 'thick points' in the sense of Deligne and Martinet, with the standard part of such a structure giving back the bundle $\pi : V = \mathbb{R}^2 \times \mathbb{P}^1 \to \mathbb{R}^2$.

4.8.4 Discrete Versus Continuous Models

We thus see that there are two complementary ways to model pinwheels and one which combines both:

1. We can go to the continuous limit and work in $\pi : V = \mathbb{R}^2 \times \mathbb{P}^1 \to \mathbb{R}^2$. This is what we shall do at great length in the second volume.

[20] For a didactic introduction to non-standard analysis, see, for example, Petitot [93] and the references therein.

[21] Rather as in the Kaluza–Klein field theories of physics.

[22] I thank Guy Wallet and Michel Berthier for this reference.

2. Alternatively, we do not go to the limit, but keep the finite mesh size of the lattice L and consider the orientation field $\Psi(a)$ of simple $V1$ neurons as a section of the bundle π defined on the open set $\mathbb{R}^2 - L$. This is the approach we shall pursue in the following sections of this chapter.
3. However, it is also possible to adopt a mixed approach and introduce continuous models that get discretized by breaking their symmetry groups. We shall also pursue this idea in the second volume.

4.9 Different Aspects of Pinwheels

To illustrate our models with experimental data, we shall discuss here some other aspects of the pinwheel maps.

4.9.1 Position–Orientation Independence and Local Triviality

The pinwheel structure is a good example of the way neurophysiology requires us to rethink the most basic geometric structures used to model it. In the above, we have assumed that the direct product structure $U \times \mathbb{P}^1$ ($U \subset R$) in the fibre bundle $\pi : U \times \mathbb{P}^1 \to U$ (local triviality) raises no particular problems. However, it assumes an independence between the position and orientation variables which is not at all obvious from the neurophysiological point of view (see, e.g. Das, Gilbert [97]) and must be carefully checked. This is what has been done by Bosking et al. [42] by analyzing the pattern of neural activity elicited by a long line crossing the visual field.

Figure 4.94, part of which appears as Fig. 4.31 in Sect. 4.4.1, shows the following:

(a) The band of neurons in $V1$ which are activated by a long line located at a precise (vertical) position x (scale 1 mm).
(b) The way this band is situated within the population of $V1$ neurons responding to the same vertical orientation but at different positions.

By carefully analyzing the way these bands and the response peaks change as the stimulus moves (10 positions x at intervals of 1°),[23] the authors have shown that the maps of the positions of the stimulus and the orientations (pinwheels) are essentially independent. Figure 4.95 shows the relationship between a $1° \times 1°$ lattice of vertical and horizontal positions of the bar and the pinwheel structure. From a scale of $4° \times 4°$ the coverage becomes uniform.

In short, Bosking shows that:

> The map of visual space in $V1$ is orderly at a fine scale and has uniform coverage of position and orientation without local relationships in the mapping of these features.

[23] Lengths are measured in degrees of the visual field.

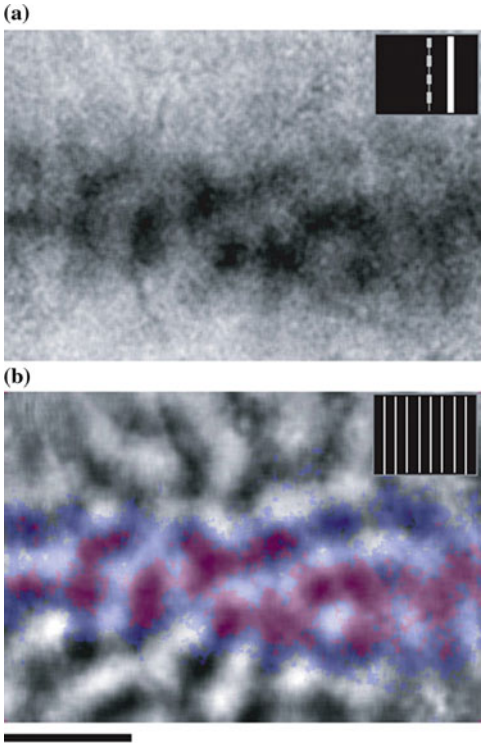


Fig. 4.94 Neurons in $V1$ activated by a long line located in a precise vertical position. From Bosking et al. [42]

In geometrical terms, this means that the Cartesian product structure (the triviality) of the bundle $\pi : R \times \mathbb{P}^1 \to R$ is indeed neurally implemented.

4.9.2 Other Engrafted Variables

Other variables are engrafted in Hubel's sense in the pinwheel structure, for example the direction of motion, phase, spatial frequency, and ocular dominance.

4.9.2.1 Direction of Motion

Simple $V1$ neurons are orientation selective and detect edges. However, they are often also sensitive to the direction of motion of the edges, and in general, this is optimal in the orthogonal direction. This is the case, for example, in the cat, as shown in Fig. 4.18 taken from Van Hooser [26]. In [98], Ohki et al. confirmed this on the level of individual cells in area 18 of the cat (see Fig. 4.96).

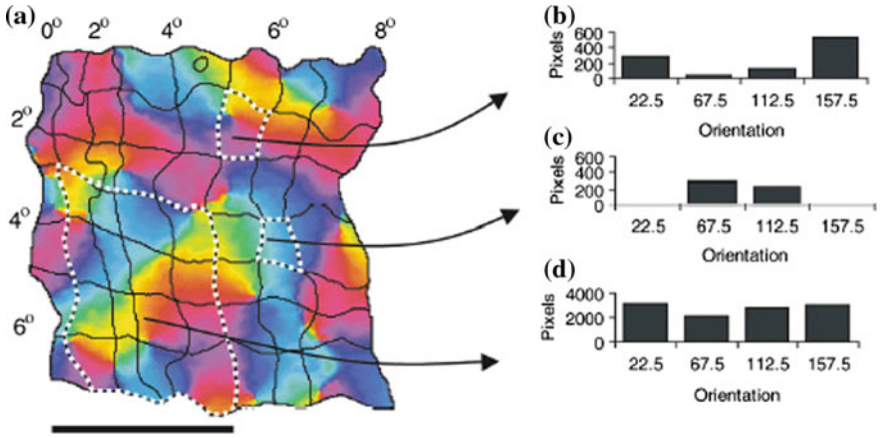


Fig. 4.95 Maps of the positions of the stimulus and the pinwheels. The coordinate system (x, y) reconstructed from the iso-azimuth contours of a $1° \times 1°$ lattice of vertical and horizontal positions of the bar is independent of the pinwheel structure. The tiny $1° \times 1°$ domains clearly have dominant orientations [histograms (**b**) and (**c**)]. But from a scale of $4° \times 4°$, the coverage is uniform [histogram (**d**)]. From Bosking et al. [42]

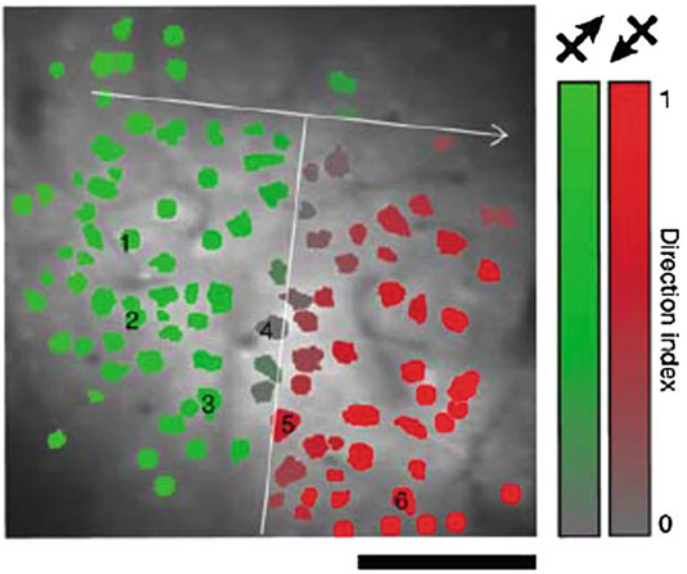


Fig. 4.96 Direction selectivity in cats. The colours *red* and *green* encode the two directions orthogonal to a given orientation. From Ohki et al. [98]

In Sect. 5.11 of Chap. 5, we shall return to the problem of direction in the context of other visual areas. We shall also discuss this a little more in Sect. 4.10.1.4.

4.9.2.2 Phase

Concerning the variation of the *phase* in a given column, DeAngelis et al. [19] compared the *spatiotemporal* RF (X, Y = space and T = time = correlation delay) of two neighbouring cells in the same column. The stimuli were small randomly flashed bars (length 1.5°, 40 ms flashes) with the preferred orientations of the two cells, these being recorded simultaneously. The authors measured the cross-correlation between the sequence of stimuli and the response (spike trains) with different correlation delays. They observed that the visuotopy, orientations, and spatial frequencies are the same, but not the phases.

4.9.3 Spatial Frequency

Here, we attribute a little more importance to another engrafted variable, viz. the spatial frequency, where interesting new models have recently been developed.

4.9.3.1 Some Preparatory Notes

Note that pinwheel rays leave room for a further parameter. Recent work has conjectured that this could be the *spatial frequency* (SF, see, e.g., DeAngelis et al. [19] and Bressloff-Cowan [99]). The experiments are not easy to carry out. Care must be taken to use stimuli that only select a single SF, and the results are still controversial.

Figure 4.97 due to Hübener et al. [100] shows the edges of domains of low spatial frequency. Statistically, the pinwheels tend to be located towards the centre of the frequency domains and the iso-orientation lines are rather strongly *transverse*, sometimes almost orthogonal, to the edges. This is a strong transversality condition.

In their work [101], Naoum Issa, Christopher Trepel, and Michael Stryker made a detailed analysis of this spatial frequency (SF) distribution in cats (see Fig. 4.98). Their results are slightly different, which shows that this is only the beginning of such investigations. They have suggested that, for this parameter, there may also be a column structure comprising columns of about 0.7 mm with ends close to the pinwheels. As they note:

> The organization of cortical maps permits nearly all combinations of orientation and SF preference to be represented in $V1$. [101, p. 8504]

Put another way, it would seem that there is a neural implementation of the bundle $\pi : \mathbb{R}^2 \times P \times F \to \mathbb{R}^2$, where F is the interval of the SFs observed (roughly from

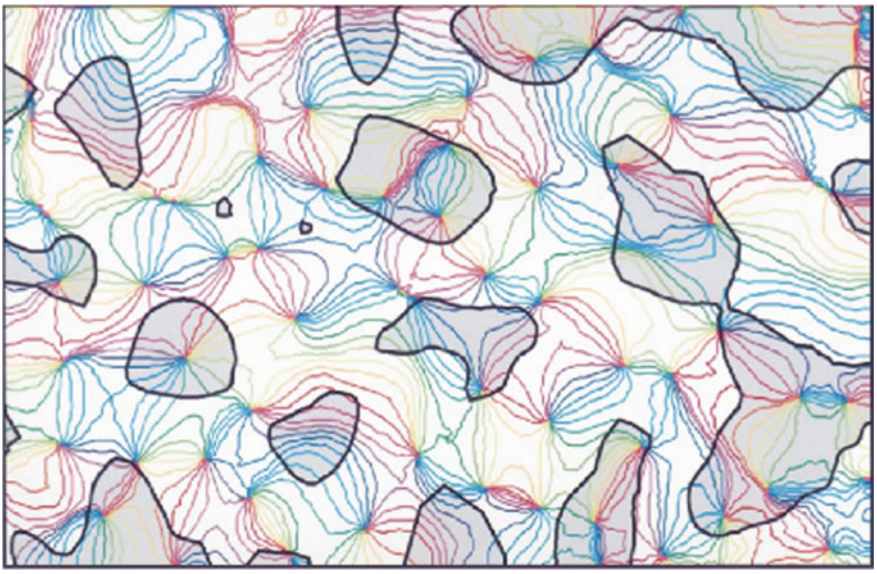


Fig. 4.97 Edges of domains of low spatial frequency (*gray*). From Hübener et al. [100]

0.2 to 1.8 cycle/degree). In fact, the binarization of the SFs into high and low frequencies may be correlated on the one hand with the segregation of the parvocellular (X) and magnocellular (Y) pathways and on the other with the chirality of the pinwheels. It often happens that two adjacent pinwheels (hence of opposite chirality) are in frequency columns of opposite colours (hence high and low frequencies).

We see immediately the kind of problems that such results raise for modelling. We may for example pair together the orientation θ with the SF f in a fibre space H and thus consider a field with values in H defined on the cortical surface, i.e. a section of the bundle $\mathbb{R}^2 \times H$. But how should H be defined? If θ and f are treated as independent variables, then $H = P \times F$. But we may also pair θ and f. This is what is done by Paul Bressloff and Jack Cowan in [99] by introducing a local 'spherical' model in which they glue together a disc (pinwheel) of maximal f with a disc (pinwheel) of minimal f along their boundaries. The common boundary is the equator, $f_{\min}$ and $f_{\max}$ are the poles, and a point of the sphere is thus a pair (θ, f) located at a distance from the singular point (the centre of the sphere) defined by f.

However, these results from Issa et al. [101] were taken up by other specialists who came to different conclusions. For example, Sirovich and Uglesich [102] detected a bias due to vascular artefacts in the hemodynamic measurements and, when they eliminated it, they concluded that there was no column organization in the SFs; for them, the SF was more like a parameter *distorting* the orientation map. Figure 4.99 shows, for two pinwheels, the difference between the maps obtained with high SFs (left) and low SFs (right). See also Born and Tootell [103].

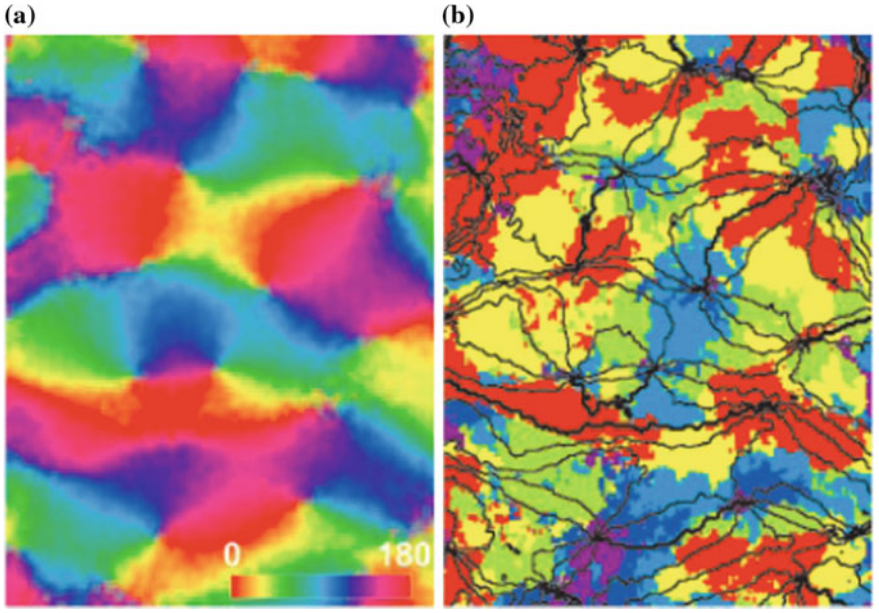


Fig. 4.98 **a** Orientation columns. **b** Spatial frequency columns. High frequencies are coded in *blue* and low frequencies in *red*. Note that they correspond well to the two chiralities of the pinwheels, i.e. two adjacent pinwheels (hence of opposite chirality) are in frequency columns of opposite colours. The *scale bar* is 1 mm. From Issa et al. [101]

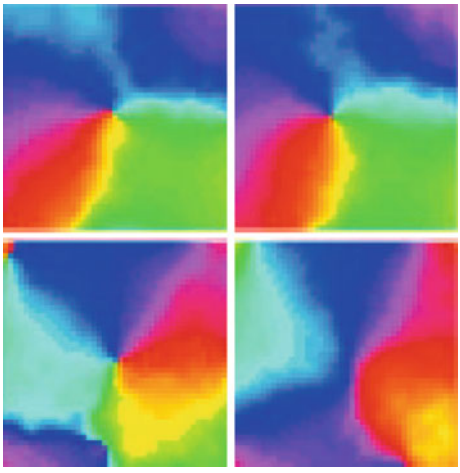


Fig. 4.99 Spatial frequency (SF) dependence of orientation maps. For two pinwheels (*top and bottom*), we observe the difference between the map obtained with high SFs (*left*) and low SFs (*right*). From Sirovich and Uglesich [102]. Copyright (2004) National Academy of Sciences, USA

These critical results were themselves repeated with autofluorescence imaging techniques using oxidation and reduction of flavoproteins in the metabolism of mitochondria. These techniques can correct for the vascular artefact pointed out by Sirovich, and it would seem that there is in fact no column organization (see Issa et al. [104]).

There are many other studies of this issue. We may cite for example the results of Zhu et al. [105] on the correlated variation of the OR and the SF.

4.9.3.2 First Dipole Model

If we accept this idea, the basic module is no longer the isolated pinwheel as a hypercolumn, but a *pair* of adjacent pinwheels with opposite chiralities. Near the singularities, i.e. the pinwheel centres (PCs), only a small part of the SFs (low or high) would be represented. A natural model would then be a *dipole–dipinwheel model*.

In Sect. 4.4.3, we discussed the analogy between the functional pinwheel maps and the topological charge fields encountered in physics. A pinwheel is locally like the field produced by a $+$ or $-$ charge. A pair of pinwheels with opposite chiralities thus corresponds to the field produced by a pair of opposite charges $(+, -)$. Such a field source is known in physics as a dipole. It will be useful to dwell on this for a moment.

In electrostatics, a dipole is made by placing opposite charges $-q$ and $+q$, e.g. -1 and $+1$, at two points A and B. Let d be the distance between A and B and place A at $(-d/2, 0)$ and B at $(d/2, 0)$, as in Fig. 4.100. Let (r, θ) be the polar coordinates of M. Up to a multiplicative constant which we ignore, we have the potential $V = 1/BM - 1/AM$. Since

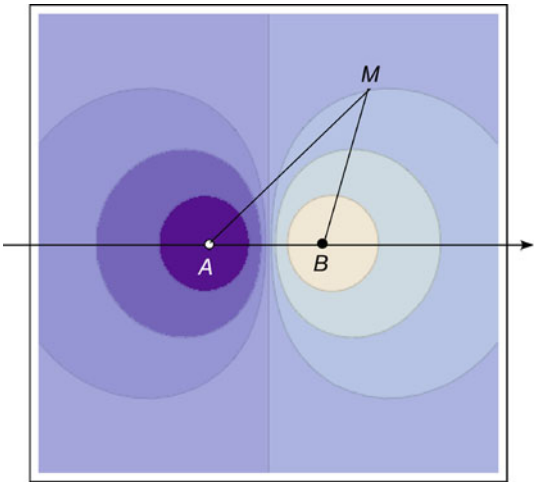


Fig. 4.100 Dipole structure. Opposite charges $-q$ and $+q$ are placed at two points A and B and we calculate the potential V at different points M. The figure shows several equipotentials. The values of V range from $-\infty$ (*black*) at A to $+\infty$ (*white*) at B. The field is $-\text{grad}\,(V)$, and its field lines are orthogonal to the equipotentials

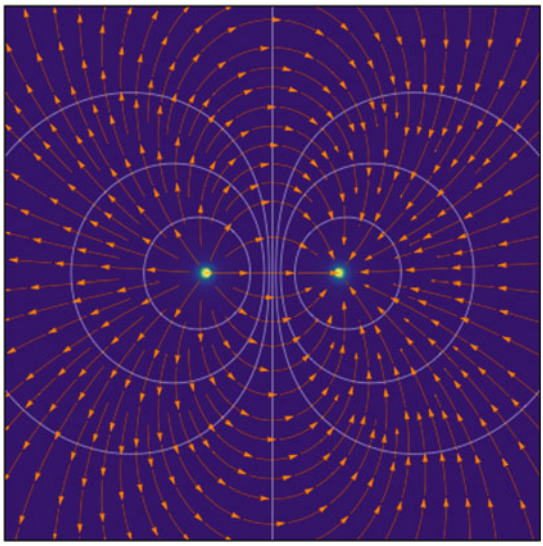


Fig. 4.101 Dipole field with equipotentials in *light blue* and field lines in *orange*

$$\begin{cases} BM^2 = r^2 + \dfrac{d^2}{4} - dr\cos(\theta) \ , \\ AM^2 = r^2 + \dfrac{d^2}{4} + dr\cos(\theta) \ , \end{cases}$$

we obtain the explicit formula

$$V(M) = \frac{1}{r}\left[\frac{1}{\sqrt{1 + \dfrac{d^2}{4r^2} - \dfrac{d}{r}\cos(\theta)}} - \frac{1}{\sqrt{1 + \dfrac{d^2}{4r^2} + \dfrac{d}{r}\cos(\theta)}}\right] .$$

In polar coordinates, the field $E = -\text{grad}(V)$ has components

$$E_r = -\frac{\partial V}{\partial r} \ , \quad E_\theta = -\frac{1}{r}\frac{\partial V}{\partial \theta} \ .$$

The $+$ charge is attractive and the $-$ charge repulsive. This is illustrated in Fig. 4.101, which is a special case of the field configurations of pinwheel lattices considered in Sect. 4.4.3.

In this dipole–dipinwheel model of an orientation–spatial frequency module comprising two adjacent pinwheels of opposite chirality, the field lines are iso-orientation curves going from low SFs centred on one of the pinwheels to high SFs centred on the other pinwheel, while the equipotentials are level lines of the SF. These two families of curves are orthogonal, i.e. maximally transverse, thus satisfying the Hübener hypothesis. Locally, in the neighbourhood of each of the PCs, we have a classical

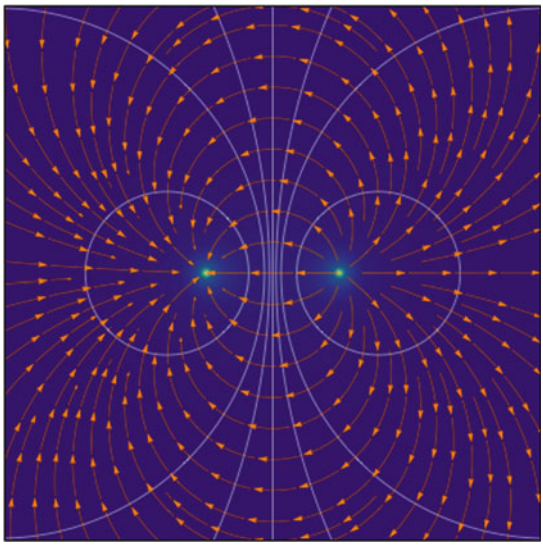


Fig. 4.102 Equipotentials and field lines of an intrinsically 2D dipole. If we compare with the field of Fig. 4.101 with potential going as $1/r$, we see that the geometries are qualitatively similar, but that the directions of the field lines are reversed, so $+$ charges are repulsive and $-$ charges attractive

orthogonal model made up of rays for the orientations and nested circles for the SFs, the latter varying smoothly near the PC.

Even though what interests us in these physical analogies is the geometry of the fields, not the physics they actually represent, we should nevertheless note that potentials of the form $1/r$ are related to forces of the form $1/r^2$ which depend intrinsically on the 3D nature of the space. Since we are working on phenomena that are intrinsically 2D, we should use 2D physical analogies where the forces go as $1/r$ and the potentials as $\mathrm{Log}\,(r)$. Such situations have been investigated, but they are rather strange, because positive charges become repulsive: $1/r$ decreases when r increases, whereas $\mathrm{Log}\,(r)$ increases. However, this does not change much in our case since we consider pairs of opposite charges, and as shown in Fig. 4.102, the geometry of the field barely changes qualitatively.

In fact, we may even consider forces of the form $1/r^{\sigma}$, where the exponent σ is a parameter to be determined experimentally.

4.9.3.3 Second Dipole Model

A certain number of more refined experiments using functional imaging techniques with better resolution have shown that the standard dipole–dipinwheel orthogonal model is not corroborated experimentally, in ways that go beyond the criticisms made by Issa [101]:

1. There are *both maxima and minima* of the SF in the vicinity of many PCs.
2. The strong transversality condition of Hübener is not satisfied at all.

In a paper in 2016, Ribot et al. [106] (see also [107]) analyzed in great detail the fine structure of the functional maps of the SFs in area 17 of the cat. They used sophisticated high-resolution imaging methods, essential from the methodological point of view if we are to test the possibility of large variations in the SF in the vicinity of a PC. They discovered that the SF generally has *two extrema*, a minimum and a maximum, in the vicinity of a given PC. This is a functional organization because it means that all the ORs and all the FSs can be represented at the same time within *the same* hypercolumn. However, it must then have a completely different topology to that of the standard orthogonal circular model.

Figure 4.103 shows the results of the data analysis carried out by Jérôme Ribot. The orders of magnitude are: (i) 1 mm for the size of a hypercolumn, (ii) 300 μm for the extent of the neighbourhood of the PC analyzed (so this remains well within the hypercolumn), (iii) 150 μm for the extent of the little neighbourhood of the PC where measurement errors remain too great. Three examples are shown. The two extrema are perfectly clear.

Hence the idea of keeping the pinwheel–dipole model, but *dissociating* the size d of the dipole from the distance between two adjacent pinwheels. We thus assume that the dipole size d is *small* compared with the size of the pinwheel as hypercolumn module. In the limit, we would then have an *infinitesimal* dipole, a situation that is easy to model.

Indeed, if the dimension d of the dipole is small enough relative to the distances r where the field is measured, then by expanding up to a few terms in d/r, we may use the approximations

$$\frac{1}{BM} = \frac{1}{r}\left[1 + \frac{d}{2r}\cos(\theta)\right], \quad \frac{1}{AM} = \frac{1}{r}\left[1 - \frac{d}{2r}\cos(\theta)\right], \quad V(M) = \frac{d}{r^2}\cos(\theta).$$

The equipotentials are then given by the equation $r^2 = k\cos(\theta)$. The components of the field in polar coordinates are

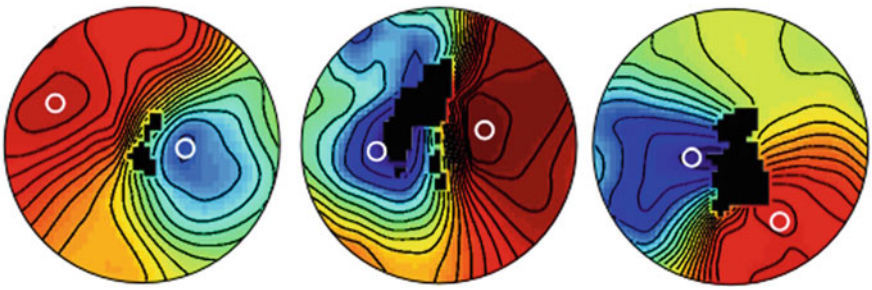


Fig. 4.103 Level lines of the SFs and their extrema in the neighbourhood of three PCs. They lie within a single hypercolumn. *Black pixels* are those where measurements are not accurate enough. From Ribot et al. [106, Fig. 5]

$$E_r = -\frac{\partial V}{\partial r} = \frac{2d}{r^3}\cos(\theta) \ , \quad E_\theta = -\frac{1}{r}\frac{\partial V}{\partial \theta} = \frac{d}{r^3}\sin(\theta) \ .$$

The equations of the field lines are therefore

$$\frac{\mathrm{d}r}{r} = \frac{E_r}{E_\theta}\mathrm{d}\theta = 2\frac{\cos(\theta)\,\mathrm{d}\theta}{\sin(\theta)} \ ,$$

that is $\mathrm{Log}(r) = 2\mathrm{Log}(\sin(\theta)) + c$, whence $r = k\sin^2(\theta)$.

For the intrinsically 2D dipole, the formulas change. In the limit, we obtain an approximation of the potential $V = -(d/r)\cos(\theta)$, equipotentials with equation $r = 2R\cos(\theta)$, and field lines with equation $r = 2R\sin(\theta)$. The first constitute a family of circles of radius R centred on the x-axis and with tangent the y-axis at 0, since

$$x = r\cos(\theta) \ , \quad y = r\sin(\theta) \ , \quad r^2 = x^2 + y^2 \ , \quad (x-R)^2 + y^2 = R^2 \ ,$$

whence $r^2 = 2Rx = 2Rr\cos(\theta)$, i.e. $r = 2R\cos(\theta)$. The second constitute the orthogonal family of circles centred on the y-axis and with tangent the x-axis at 0 (see Fig. 4.104). Figure 4.105 shows the potential $V(M)$ in the neighbourhood of 0.

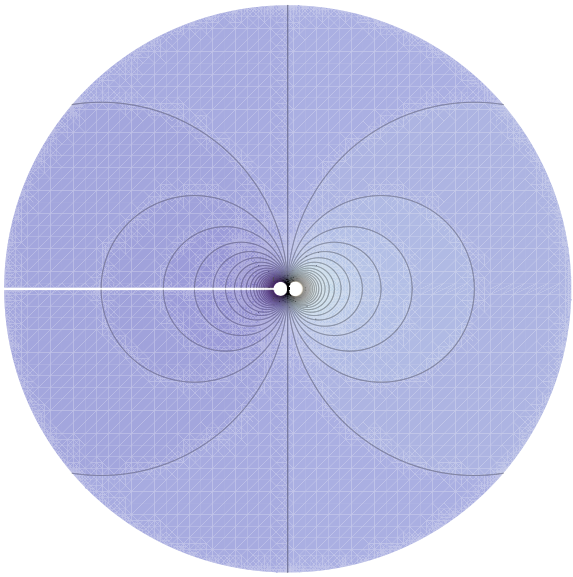


Fig. 4.104 Equipotentials of an intrinsically 2D infinitesimal dipole. This is a family of circles centred on the x-axis and with tangent the y-axis at 0

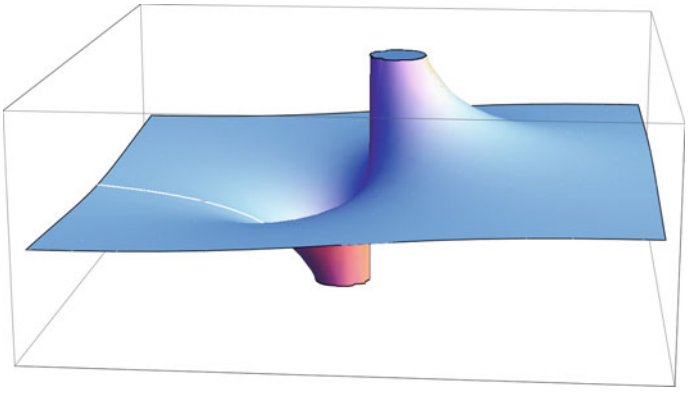


Fig. 4.105 Potential $V(M)$ of the infinitesimal dipole in Fig. 4.104. We see how V diverges to $+\infty$ at 0^+ and $-\infty$ at 0^-

4.9.3.4 Topological Universality of Dipoles

In their elegant paper [58], already cited in Sect. 4.5, Alberto Romagnoni, Jérôme Ribot, Daniel Bennequin, and Jonathan Touboul showed that just as the topological structure of the pinwheel is universal for the 'exhaustive coding' and the 'parsimonious coding' of periodic quantities, so the topological structure of the dipole is universal for the 'exhaustive coding' and the 'parsimonious coding' of non-periodic quantities.

As for the pinwheels, the idea is to consider the way the cells parametrized by $z = (x, y)$ in a small disc $\Omega \subset \mathbb{C}$ (with boundary $\Gamma = \partial\Omega$) of the visual plane (identified with the plane $\mathbb{C} = \mathbb{R}^2$) can encode spatial frequencies $\nu \in \mathbb{F}$ in a *local* and, ideally, *scale invariant* way, where $\mathbb{F}$ is the allowed spatial frequency domain. As we are only interested in topological structure, we can take $\mathbb{F} = \mathbb{R}$, introducing saturation if necessary. Such a coding can thus be described by a map $g : \Omega \to \mathbb{R}$ which may have singularities, i.e. points where it is not defined because it is multivalued or divergent.

We use the standard topological properties of $\mathbb{R}$:

- $\mathbb{R}$ is *not compact*, i.e. its points escape to infinity.
- In contrast to $\mathbb{S}^1$, if we remove a point or a finite interval from $\mathbb{R}$, it disconnects into two connected components.
- *Intermediate value theorem* (which fails disastrously in $\mathbb{S}^1$): if $h : I \to [h_{\min}, h_{\max}]$ is continuous for an interval I and if $\nu \in]h_{\min}, h_{\max}[$ is an intermediate value, then there exists $x \in I$ with image $h(x) = \nu$.

As for the pinwheels, the authors define a rather general class $\mathcal{G}$ of sufficiently smooth maps g satisfying natural constraints. They assume to begin with that g is continuous away from its singularities and 'exhaustive', i.e. it codes all the SFs. In other words, it is *surjective*. This implies that g must have at least one singularity. Indeed, if g

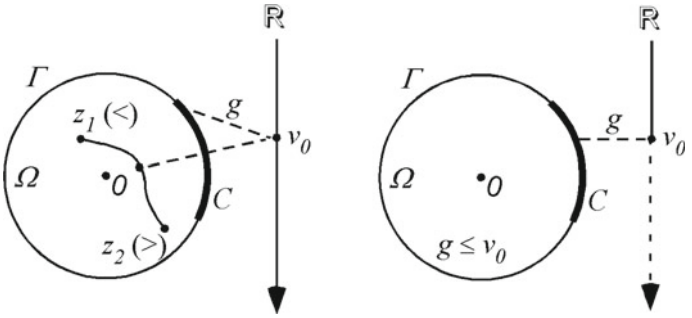


Fig. 4.106 A connected level line $C = C_{v_0} = g^{-1}(v_0)$ of a surjective g cannot be entirely contained within the boundary Γ of Ω unless $g \geq v_0$ or $g \leq v_0$ everywhere in Ω. But then g is not surjective. So C must enter Ω. As usual, $(\gtrless)$ means $z \gtrless v_0$

had no singularities, it would be a continuous map on a compact set Ω and its image $g(\Omega)$ would therefore be compact, hence contained in a closed *bounded* interval which could not then be equal to $\mathbb{R}$. Since we work locally and topologically, we shall assume that g has just one singularity and that it is at 0. We write $\Omega^* = \Omega - \{0\}$.

The authors then introduce three assumptions to define $\mathcal{G}$:

H1 Smoothness condition: the map $g : \Omega^* \to \mathbb{R}$ is smooth (infinitely differentiable) and its level lines $C_v = g^{-1}(v)$ are smooth curves (as we saw in Sect. 4.5, this is a generic property [59]).

H2 Exhaustivity condition: since the model must have scale-invariant properties, g must be exhaustive (surjective) in *any* neighbourhood of 0.

H3 Minimal complexity or parsimony condition: topological redundancy must be minimal on *all* scales. There are arbitrarily small discs Ω' centred on 0 where the topological redundancy is minimal.

Concerning assumption H3, the authors use a lemma according to which *there is no* continuous surjective $g : \Omega^* \to \mathbb{R}$ with topological redundancy 1. Indeed, let $v_0 \in g(\Gamma)$ be a value of g taken on the boundary Γ of Ω and consider the level line $C = C_{v_0} = g^{-1}(v_0)$. To begin with, C cannot be completely included in the boundary Γ (see Fig. 4.106). Indeed, let $z_1 \neq z_2 \in \mathring{\Omega}^*$ be two points in the interior $\mathring{\Omega}^*$ of Ω^* and assume that $g(z_1) < v_0$ and $g(z_2) > v_0$. Let $\gamma : z_1 \to z_2$ be a continuous path joining z_1 and z_2 in $\mathring{\Omega}^*$. Then, γ is an embedding of an interval I in Ω^* and $h = g \circ \gamma : I \to \mathbb{R}$ is a continuous map, so according to the intermediate value theorem, there is a point z of γ where $g(z) = v_0$ and hence $z \in C \cap \mathring{\Omega}^*$. But this is impossible since this intersection is empty by the assumption about C. Hence, on $\mathring{\Omega}^*$, either $g \leq v_0$ everywhere or $g \geq v_0$ everywhere and g is not surjective, contradicting H1.

Thus, C enters the interior $\mathring{\Omega}^*$ of Ω^*. If we assume that the topological redundancy is 1, then C is *connected*. If C stops in $\mathring{\Omega}^*$, then we can find a topological disc $\Omega' \subset \Omega$ around 0 (possibly highly twisted) which is decomposed into two parts by C, i.e.,

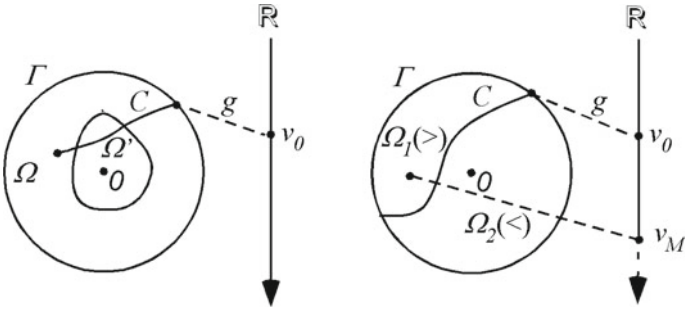


Fig. 4.107 If the connected level line $C = C_{v_0} = g^{-1}(v_0)$ of g starts from the boundary Γ of Ω and enters Ω without reaching 0, then C decomposes a neighbourhood Ω' of 0 into two subdomains and g is not surjective on the new $\Omega = \Omega'$, because $g \leq v_M$ everywhere on Ω^*. As usual, $(\gtrless)$ means $g \gtrless v_0$

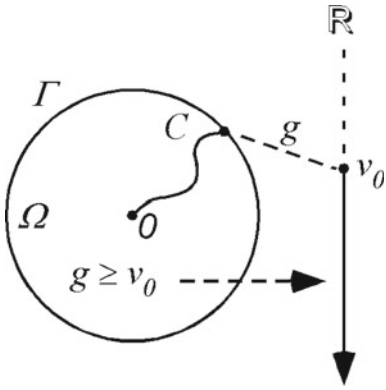


Fig. 4.108 If the connected level line $C = C_{v_0} = g^{-1}(v_0)$ starting from the boundary Γ of Ω reaches the singular point 0, then either $g \geq v_0$ or $g \leq v_0$ everywhere on Ω^* and g is not surjective

C joins two points on the boundary Γ'. Replacing Ω by Ω', we may thus assume that C decomposes Ω into two closed domains Ω_1 and Ω_2 with a common boundary C (see Fig. 4.107). Assume that $0 \notin \Omega_1$, whence $0 \in \Omega_2 - C$. By the intermediate value theorem once again, we have $g > v_0$ everywhere on $\Omega_1 - C$ and $g < v_0$ everywhere on $\Omega_2^* - C$, or the opposite. Assume the first case. Since Ω_1 is compact and g is defined on the whole of Ω_1, g reaches a maximum $v_M \geq v_0$ on Ω_1 and hence $g \leq v_M$ everywhere on Ω^*. Consequently, g cannot be surjective.

So if C is connected, C must reach the singular point 0. But then $\Omega^* - C$ is *connected*, and g is continuous on $\Omega^* - C$ and does not take the value v_0. As its image $g(\Omega^* - C)$ is connected, being the continuous image of a connected set, then from the topological properties of $\mathbb{R}$, either $g \geq v_0$, or $g \leq v_0$ everywhere on Ω^*, so g cannot be surjective (see Fig. 4.108).

In short, level curves C_v with points on the boundary Γ necessarily have at least *two* connected components and the topological redundancy of g is ≥ 2. However, $C_v \subset \mathring{\Omega}^*$ lying completely in the interior of Ω^* can be connected.

Clearly, the standard model of the 2D infinitesimal dipole $g_0 : \Delta^* \to \mathbb{R}$ defined on the unit disc $\Delta^* = \Delta - \{0\}$ with the origin removed,

$$z = \rho e^{i\theta} \longmapsto \nu = \frac{\cos(\theta)}{r},$$

belongs to $\mathscr{G}$:

(i) g_0 is well defined away from 0, 0 being a singularity where it is completely undefined, since θ is not defined and $1/r$ diverges.
(ii) Away from 0, g_0 is smooth (infinitely differentiable).
(iii) The level sets $C_\nu = g_0^{-1}(\nu)$, where g_0 is constant, are smooth curves.
(iv) g_0 is surjective on any neighbourhood of 0 (and even on any small circle surrounding 0).
(v) The topological redundancy of g_0 is 2, and it is therefore minimal.

Note that (see Fig. 4.104), insofar as the level lines constitute a bundle of circles (with the origin 0 removed), small enough circles of radius $\leq 1/2$ lie within Δ and are connected level lines, whereas circles of radius $> 1/2$ stop at Γ and are level lines with two connected components. The authors refer to the configuration of internal circles as a 'bouquet' of circles. They prove the following theorem:

Theorem *Topological universality of dipoles. The elements $g : \Omega^* \to \mathbb{R}$ of $\mathscr{G}$, i.e., satisfying hypotheses H1, H2, and H3 on all scales, have the dipole topology.*

As for the theorem on the topological universality of pinwheels, the ideas used in the proof are purely topological. We can give a heuristic insight (incomplete and not rigorous). Topologically, the connected components C of the level lines C_ν of $g \in \mathscr{G}$ can be of 4 types: (i′) a closed loop not going through 0 and not going around zero, (i″) a closed loop not going through 0 but going around 0, (ii) a path from Γ to Γ, (iii) a path connecting Γ to 0, and (iv) a loop going through 0 (see Fig. 4.109).

As g is continuous on the boundary Γ and the boundary is compact, $g(\Gamma)$ is a compact subset K of $\mathbb{R}$ and hence bounded. As g has its values in K for paths of

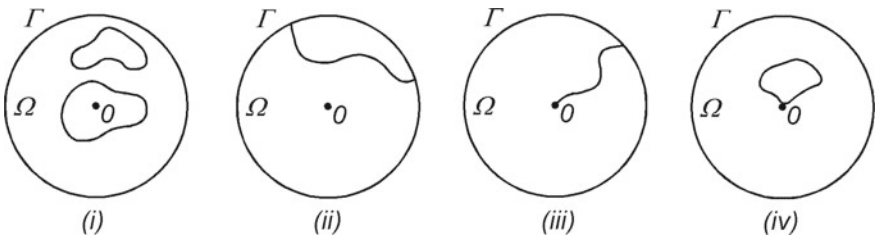


Fig. 4.109 The four types of connected components of level lines for a $g \in \mathscr{G}$. *i* Loops in the interior of Ω and not going through 0 but either encircling it or not. *ii* Paths connecting two points on the boundary Γ. *iii* Paths connecting a point on the boundary Γ with the singularity 0. *iv* Loops in the interior of Ω and passing through 0

types (ii) and (iii) g, there must necessarily be level lines of types (i) or (iv). Let C be of type (i′), i.e. $C \subset \mathring{\Omega}^*$, with 0 outside C. Let Λ be the closed topological disc with boundary C and $\Omega^{*\prime}$ a small disc centred on 0 and contained in $\Omega^* - \Lambda$. From the scale invariance hypothesis, we can localize in $\Omega^{*\prime}$ and Λ only adds redundancy. We can thus eliminate the case (i′), and also case (ii) for the same reason (we can remove the connected component of $\Omega - C$ which does not contain 0).

There remain the cases (i″), (iii), and (iv). Let C be of type (i″) and $\Lambda \ni 0$ the topological disc with C as boundary; since C is a level curve, according to the smoothness assumptions, no level line in the interior of Λ^* can meet C. There can thus only be C' of types (i″) and (iv) in Λ^*. However, there cannot only be C' of type (i″) because they would be 'concentric', i.e. nested, and localizing even more closely around 0, g would no longer be surjective, thus violating H2. So there is at least one C' of type (iv) in the interior of Λ^*.

Let C be a level curve C_ν of type (iv) on which $g = \nu$. Then, C is the boundary of a disc Λ and $0 \in C = \partial\Lambda$. Since 0 is a singularity of g, the latter can diverge as we approach 0 in Λ^*, but it can diverge only to either $+\infty$ or $-\infty$ and not both. There will thus be curves C^+ and C^- bounding regions Λ^+ and Λ^-. Consider the case where g diverges to $+\infty$ and let $\mathring{\Lambda}^+_{\mathrm{m}}$ be the interior of the Λ^+ of the largest component of type (iv) diverging to $+\infty$. We obtain a lobe of concentric regions Λ^+ on which g tends to ν when the Λ^+ expands to Λ^+_{m} and tends to $+\infty$ when Λ^+ contracts to 0. If ν^+_{m} is the value of g on the boundary C^+_{m} of Λ^+_{m}, g takes the values $\left[\nu^+_{\mathrm{m}}, +\infty\right]$ on Λ^+_{m}.

The complement $\Omega - \mathring{\Lambda}^+_{\mathrm{m}}$ is topologically a closed ring on which g must be able to diverge to $-\infty$ (because this divergence is not realized in $\mathring{\Lambda}^+_{\mathrm{m}}$ by construction). This requires curves C of type (iv) in the interior of $\Omega - \mathring{\Lambda}^+_{\mathrm{m}}$ and bounding regions Λ^-. We thus have, as for the standard dipole (see Fig. 4.104), a lobe of concentric regions Λ^+ up to Λ^+_{m} and a lobe of concentric regions Λ^- expanding out to Λ^-_{m}. If ν^-_{m} is the value of g on the boundary C^-_{m} of Λ^-_{m}, then g takes the values $\left[-\infty, \nu^-_{\mathrm{m}}\right]$ on Λ^-_{m}.

From the minimal complexity hypothesis, i.e. that the topological redundancy is equal to 2, there cannot be more than two lobes of this kind. Indeed, consider a level line of type (iv). If we have restricted it to a small enough disc Ω' around 0, we obtain two connected components (see Fig. 4.110i). The values of g on two lobes cannot be the same because, for small enough Ω', that would give four connected components, thereby violating hypothesis H3 (see Fig. 4.110ii). Suppose then that there are three lobes. Then, g will take different values on each of them. Suppose that $\nu_1 < \nu_3 < \nu_2$ with $g(a) = \nu_1$ for a point a of the lobe L_1, $g(b) = \nu_2$ for a point b of the lobe L_2, and $g(c) = \nu_3$ for a point c of the lobe L_3. Let $\gamma : a \to b$ be a path passing outside the lobe L_3 (see Fig. 4.110iii). According to the intermediate value theorem, there is a point d on γ, where g takes the value ν_3. But then there is a connected component of C_{ν_3} outside L_3 and reaching 0. Since the part of C_{ν_3} inside L_3 itself has two connected components if Ω is small enough, C_{ν_3} has a topological redundancy ≥ 3, thereby violating the hypothesis H3.

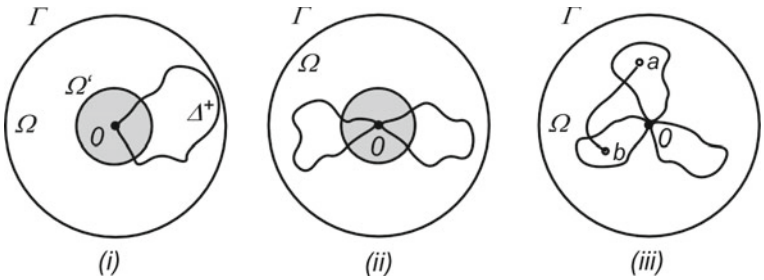


Fig. 4.110 Diagrams illustrating the proof of the topological universality of the dipole

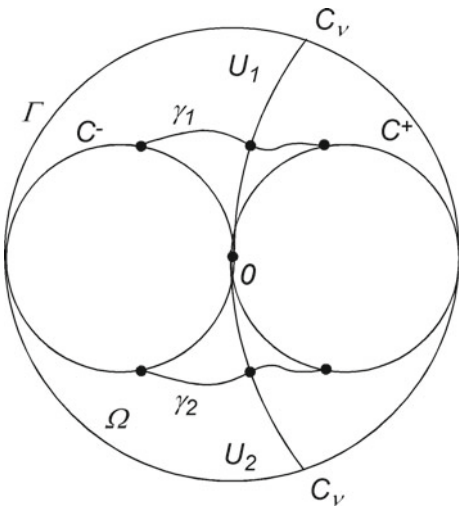


Fig. 4.111 Illustration of the last part of the proof of the topological universality of the dipole

We thus end up with a 'bouquet' comprising a lobe Λ_m^+ of concentric regions Λ^+ on which g takes the values $\left[\nu_m^+, +\infty\right]$ and a lobe Λ_m^- of concentric regions Λ^- on which g takes the values $\left[-\infty, \nu_m^-\right]$. On the respective boundaries C^+ and C^- of Λ_m^+ and Λ_m^-, g takes the values ν_m^+ and ν_m^-, respectively. It remains to consider the values in the interval $\left(\nu_m^-, \nu_m^+\right)$. The complement of the union $\Lambda_m^+ \cup \Lambda_m^-$ of the two lobes in Ω comprises two disjoint open sets U_1 and U_2. Let γ_1 be a path from C^+ to C^- in U_1 and γ_2 a path from C^+ to C^- in U_2. Using the intermediate value theorem once again, if $\nu \in \left(\nu_m^-, \nu_m^+\right)$, g takes the value ν on both γ_1 and γ_2. The level line C_ν then comprises two connected components of type (iii) joining Γ to 0 (see Fig. 4.111).

4.9.3.5 Pinwheel–Dipole Model

Daniel Bennequin and coworkers thus arrive at a model of the orientation–spatial frequency module of the form (f, g) with $f(z) = \theta/2$ and $g(z) = \cos(\theta)/r^\alpha$, with

$z = re^{i\theta}$ in the unit disc with the centre removed Δ^* and α an exponent used to fit the model to the data. Here, 0 is a singularity of both f and g:

> [This model] is the unique scale-invariant topology achieving minimal representation of a pair (θ, ν) for the circular variable θ and the variable ν taking values in an open interval. [58, p. 9]

In this model, the orthogonality of the level lines of f and g is no longer satisfied at all. Indeed, let $z_0 = r_0 e^{i\theta_0}$ be a point of Δ^*. The level line F_0 of f corresponding to $f(z_0) = \theta_0/2 = \text{const.}$ is the ray through z_0. The level line G_0 of g corresponding to $g(z_0) = \cos(\theta_0)/r_0^\alpha = \text{const.}$ has equation $\cos(\theta)/r^\alpha = \cos(\theta_0)/r_0^\alpha = C$, i.e. $\cos(\theta) - Cr^\alpha = 0$. Considering the symmetries of the configuration, we can restrict to $\theta \in [0, \pi/2]$.

Suppose to begin with that $\alpha = 1$. Then, G_0 has equation $r = 2R\cos(\theta)$, with $2R = r_0/\cos(\theta_0)$ so the tangent at z satisfies $\mathrm{d}r + 2R\sin(\theta)\,\mathrm{d}\theta = 0$. Using

$$x = r\cos(\theta) = 2R\cos^2(\theta) \ , \quad y = r\sin(\theta) = 2R\cos(\theta)\sin(\theta) \ ,$$

we obtain

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}\theta} = -4R\cos(\theta)\sin(\theta) = -2R\sin(2\theta) \ , \\ \dfrac{\mathrm{d}y}{\mathrm{d}\theta} = 2R\big[\cos^2(\theta) - \sin^2(\theta)\big] = 2R\cos(2\theta) \ , \\ \dfrac{\mathrm{d}y}{\mathrm{d}x} = -\dfrac{1}{\tan(2\theta)} = \tan\left(\dfrac{\pi}{2} + 2\theta\right) \ . \end{cases}$$

Hence, the tangents to F_0 and G_0 at z_0 make angles θ_0 and $\varphi_0 = 2\theta_0 + \pi/2$, so the angle between them is $\psi_0 = \varphi_0 - \theta_0 = \theta_0 + \pi/2$. In particular, they are orthogonal for $\theta_0 = 0$, i.e. at the point $(2R, 0)$, and become tangents when θ_0 tends to $\pi/2$. The normalized distribution $P(\psi)$ of the angles ψ is thus uniform and equal to $P(\psi) = 2/\pi$, so that $\int_{\pi/2}^{\pi} P(\psi)\,\mathrm{d}\psi = 1$.

For $\alpha \neq 1$, an exact calculation is still possible, but more complicated. We write G_0 in the form $\cos(\theta) = Cr^\alpha$, with $C = \cos(\theta_0)/r_0^\alpha$. So along C, we have $\sin(\theta)\,\mathrm{d}\theta + C\alpha r^{\alpha-1}\mathrm{d}r = 0$, i.e.,

$$r\frac{\mathrm{d}\theta}{\mathrm{d}r} = -C\frac{\alpha r^\alpha}{\sin(\theta)} \ .$$

The tangent is thus

$$\begin{aligned} \frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{\sin(\theta)\,\mathrm{d}r + r\cos(\theta)\,\mathrm{d}\theta}{\cos(\theta)\,\mathrm{d}r - r\sin(\theta)\,\mathrm{d}\theta} = \frac{\sin(\theta) - C\dfrac{\alpha r^\alpha\cos(\theta)}{\sin(\theta)}}{\cos(\theta) + C\alpha r^\alpha} \\ &= \frac{\sin^2(\theta) - C\alpha r^\alpha\cos(\theta)}{\sin(\theta)\cos(\theta) + C\alpha r^\alpha\sin(\theta)} \ . \end{aligned}$$

With $\theta = \theta_0$ and $C = \cos(\theta_0)/r_0^\alpha$, we thus obtain

$$\tan(\varphi_0) = \tan(\psi_0 + \theta_0) = \frac{\tan(\psi_0) + \tan(\theta_0)}{1 - \tan(\psi_0)\tan(\theta_0)} = \frac{\mathrm{d}y}{\mathrm{d}x}$$
$$= \frac{\sin^2(\theta_0) - \dfrac{\cos(\theta_0)}{r_0^\alpha}\alpha r_0^\alpha \cos(\theta_0)}{\sin(\theta_0)\cos(\theta_0) + \dfrac{\cos(\theta_0)}{r_0^\alpha}\alpha r_0^\alpha \sin(\theta_0)} = \frac{\tan^2(\theta_0) - \alpha}{(1+\alpha)\tan(\theta_0)} .$$

We conclude that $\tan(\psi_0) = -\alpha/\tan(\theta_0)$, so the angular distribution is

$$\mathrm{d}\theta = \frac{\alpha}{1 + \left(\alpha^2 - 1\right)\cos^2(\psi)}\mathrm{d}\psi = P_1(\psi)\,\mathrm{d}\psi .$$

Normalizing $P_1(\psi)$ to $P(\psi)$ so that $\int_{\pi/2}^{\pi} P(\psi)\,\mathrm{d}\psi = 1$, and since ψ varies from $\pi/2$ to π when θ varies from 0 to $\pi/2$ and $\int_0^{\pi/2} \mathrm{d}\theta = \pi/2$, we obtain the distribution

$$P(\psi) = \frac{2\alpha}{\pi\left[1 + \left(\alpha^2 - 1\right)\cos^2(\psi)\right]} .$$

Figure 4.112 shows several distributions $P(\psi)$ for different values of the exponent α.

The high-resolution experimental data for the cat reported in [58] and [106] show that there is a certain advantage for parallelism over orthogonality due to saturation blocking the dipole divergence at the PC. This can be modelled by taking an exponent $\alpha < 1$. The data in Fig. 4.113 from [58] fit particularly well with the value $\alpha = 0.73$.

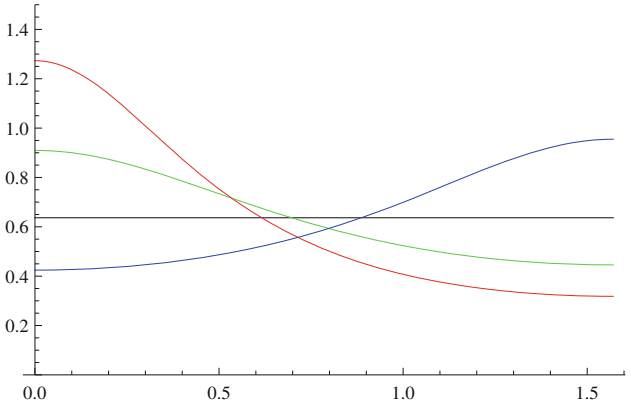


Fig. 4.112 Examples of the angular distribution $P(\psi)$ for different values of α. The angle ψ varies from 0 to $\pi/2 \sim 1.57$. *Black* $\alpha = 1$ and $P(\psi) = 2/\pi \sim 0.64$. *Green* $\alpha = 0.7$. *Red* $\alpha = 0.5$. *Blue* $\alpha = 1.5$

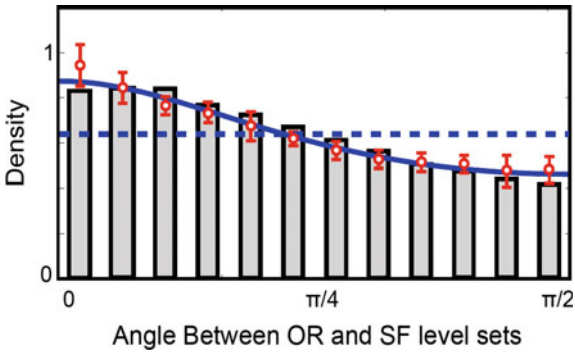


Fig. 4.113 Observed distribution of the angles between the level lines of the ORs and the SFs. From [58]

4.9.3.6 Orientation–Spatial Frequency Dependence

In the classic orthogonal circular model centred on a PC where the SF diverges, for example, to $-\infty$ and grows to, let's say, ν_0, there is for each value of $\nu \in (-\infty, \nu_0)$ a full circle of orientations $\omega = \theta/2$. Thus, Ω^* has a *direct product* structure $\mathbb{S}^1_\pi \times (-\infty, \nu_0]$ which is a subset of the direct product $\mathbb{S}^1_\pi \times \mathbb{R}$. Combined with orthogonality, i.e. maximal transversality, this shows that, in this coding of the (ω, ν), the variables ω and ν are *independent*. We shall return to the connection between transversality and independence in Sect. 4.10.2. However, note that, in the pinwheel–dipole model, ω and ν are *not* coded independently.

Indeed, in Ω^*, r varies from 0 to 1 and hence, for θ fixed, $\nu = \cos(\theta)/r$ varies between $\pm\infty$ and $\cos(\theta)$, $\pm$ being the sign of $\cos(\theta)$. Figure 4.114 shows the set of pairs (ω, ν) represented. They are distributed throughout the full direct product $\mathbb{S}^1_\pi \times \mathbb{R}$, but without covering it and without having a direct product structure. This agrees with results like those of Tani et al. [108], showing that the ORs and SFs develop *interdependently*, at least in cats.

We thus see that the two coding strategies of the dipole–dipinwheel and pinwheel–dipole models are very different.

4.9.4 Generality of Pinwheels

Since it is so functional, the pinwheel structure is extremely general. It is not only found in $V1$, but also in the other areas of the primary visual cortex. As shown, e.g., by Xiangmin Xu et al. [109] for the New World owl monkey (douroucoulis), as we go from $V1$ to $V2$ and then $V3$, the pinwheel structure remains, but the mesh of the singularity lattice increases with the size of the receptive fields.

Furthermore, the structure of $V1$ depends on the species. However, the existence of pinwheels is a robust interspecies phenomenon. Here, we give just one example in Fig. 4.115, produced by Liu and Pettigrew [110], which compares orientation maps

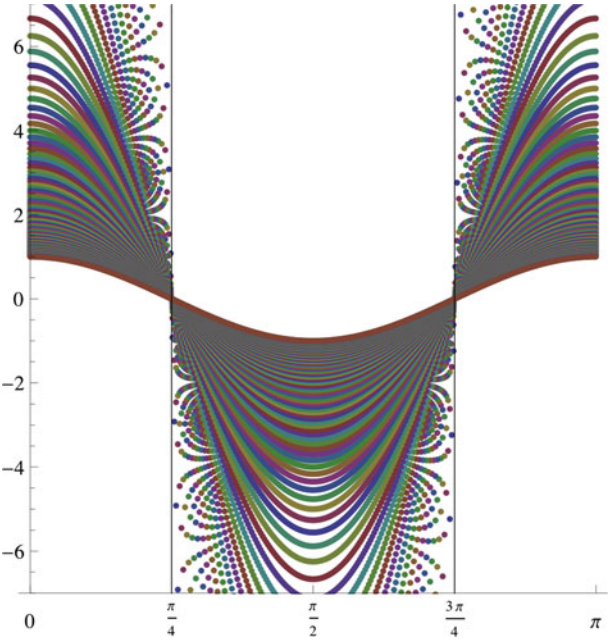


Fig. 4.114 Image of a regular lattice of $(\theta, r) \in [0, 2\pi] \times [0, 1]$ in the cylinder $\mathbb{S}^1_\pi \times \mathbb{R}$ of the $(\omega = \theta/2, \nu = \cos(\theta)/r)$ (represented by the band $[0, \pi] \times \mathbb{R}$). The divergence of ν is bounded by ± 7. The image of $[0, 2\pi] \times [0, 1]$ is clearly visible in $\mathbb{S}^1_\pi \times \mathbb{R}$

of the $V1$ and $V2$ areas in the cat and the marmoset monkey with the equivalent in the tawny owl. We also observe a pinwheel structure in the tawny owl with an interpatch distance of about 0.9 mm.

4.10 Retinotopic Maps and Their Transversality

With ever more sophisticated experimental techniques, specialists have been able to build up several retinotopic maps corresponding to different geometric features and in several different species. We have already discussed orientation, direction, phase, and spatial frequency. Now, we shall go on to consider ocular dominance, colour, temporal frequency, and motion. With so many different issues, we may ask the following questions:

1. Are the geometric features studied with the kind of stimuli generally used, viz. gratings, etc., sufficient to describe the cortical activity induced by *complex and natural* visual stimuli?
2. What are the *relations* between the maps? Are the features independent of one another, or not?

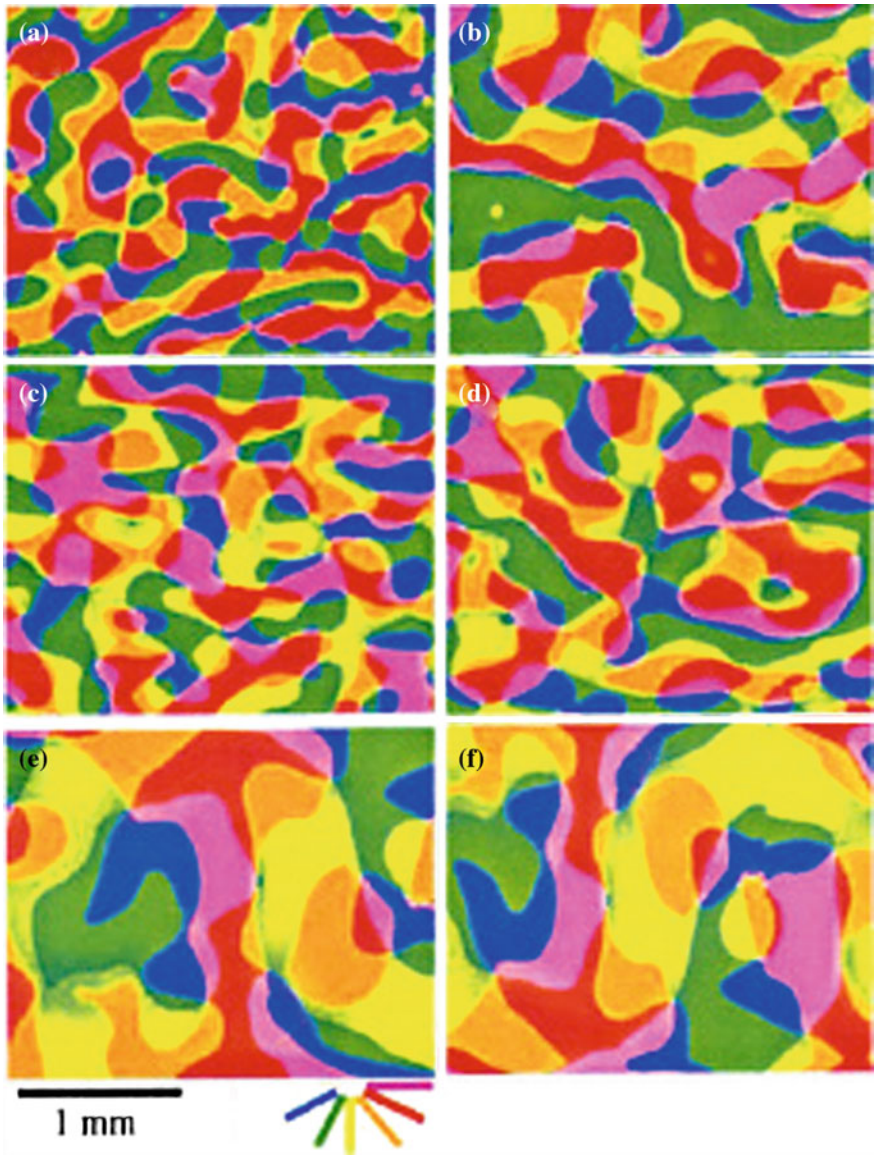


Fig. 4.115 Pinwheels in the $V1$ and $V2$ areas of the cat, (**a**) and (**b**), respectively, and the marmoset (**c**) and (**d**), and equivalent areas in the visual cortex of the tawny owl (**e**) and (**f**). From Liu and Pettigrew [110]

3. What is the mechanism of the *dimensional reduction* which implements abstract $(2 + N)$-dimensional fibre bundles in the 2D cortical layers (where N is the number of features)?

Regarding the first question, it seems that the feature detections are adequate if we take into account, for each feature, the preferred values, intensities, and variances, i.e. the complete tuning curves, and any possible dependences between the features (see, e.g. Issa et al. [104]).

Regarding the second question, we have seen several examples, and we shall discuss several more.

For the third question, it should be emphasized that the maps and their relations are the solutions to a problem with two opposing constraints. One concerns the uniformity of coverage of the features in the 2D cortical layer. For example, in the standard circular model of the SFs, when a pinwheel lies at the centre of a domain of high spatial frequency, all orientations are automatically combined locally with this SF (see Sect. 4.9.3.6). The other is a continuity constraint: the features are continuous functions of the retinotopic position, although with the possibility of singularities.

4.10.1 Pinwheels and Ocular Dominance

4.10.1.1 Some Experimental Data

We have seen the relations between the pinwheels and the spatial frequencies. There are also quite remarkable relations between the pinwheel structure of $V1$ and its organization into ocular dominance domains (ODDs), i.e. dominance of the left or right eye, or more precisely, ipsilateral or controlateral dominance relative to the given hemisphere. The ocular dominance bands measure about 1 mm in the monkey and the cat and about 2 mm in humans. As shown by Hübener et al. [100], the iso-orientation lines are essentially *transverse*, and even almost orthogonal, to the boundaries of the ODDs (see Figs. 4.116 and 4.117):

> Many iso-orientation lines cross the borders between ocular dominance domains close to right angles, and the pinwheel centers are preferentially located in the middle of these ocular dominance domains. [100]

Xiangmin Xu et al. [45] found the same relationship in the bush baby or galago, a small nocturnal African primate, already encountered in Sect. 4.4.1. The histogram of the angles of intersection between the iso-orientation lines and the boundaries of the ODDs shows that transversality is statistically well represented: the angles between $3\pi/8$ and $\pi/2$ represent more than 50%. Moreover, as clearly shown by Crair et al. [36], ocular dominance peaks are situated very close to pinwheel centres in the middle of the ODDs (see Fig. 4.118).

One can also compare the ocular dominance domains with the spatial frequency domains (see, e.g., Hübener et al. [100]).

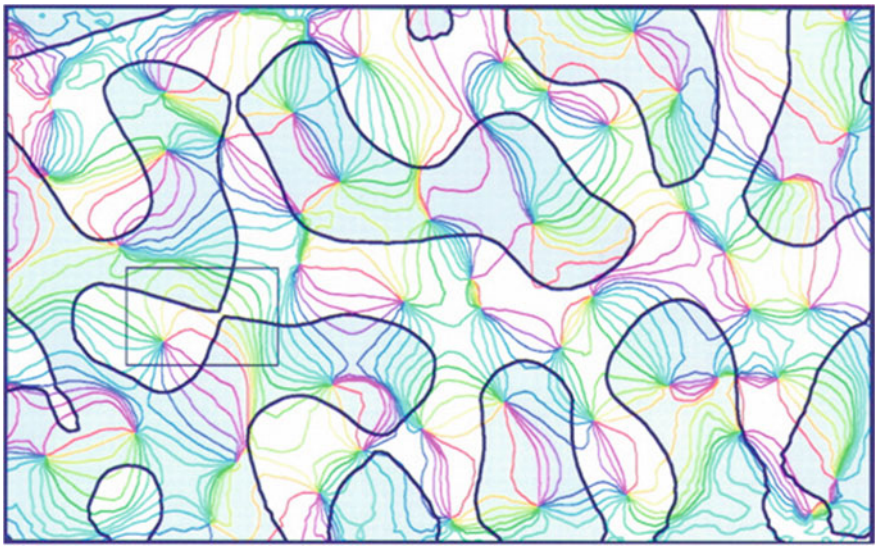


Fig. 4.116 Relations between pinwheels and ocular dominance domains. Several iso-orientation lines cross the boundaries of the ODDs almost at *right angles*. From Hübener et al. [100]

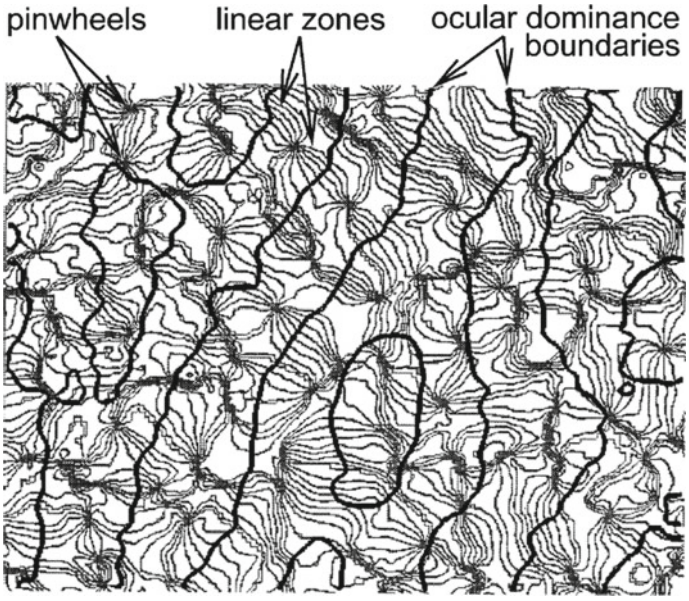


Fig. 4.117 Relations between pinwheels and ocular dominance domains in the macaque. From Obermayer, Blasdel [111]

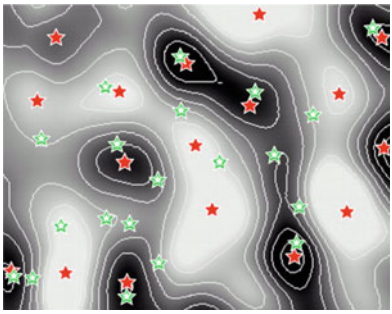


Fig. 4.118 Pinwheel centres (*green stars*) and ocular dominance peaks (*red stars*) are neighbours and located on ODD ridges. From Crair et al. [36]

4.10.1.2 Fields and Equipotentials

As for the pinwheels, the geometrical configuration of the OD peaks and the ODDs can be made intuitive using a physical field deriving from a potential V (see Sect. 4.4.3), the peaks corresponding to the extrema of V. Then, the ODD boundaries can be identified with the level lines of the potential segregating the influence zones of the maxima and the minima. If the peaks of OD would be close to the pinwheels (but it is only partially the case), then the two OR and OD fields would be close and, as in a field equipotentials are orthogonal to field lines, this could explain the strong transversality between the iso-orientation lines of the OR field and the boundaries of the ODDs.

4.10.1.3 An Elastic Net Model

It is easy enough to simulate these field structures with orientation columns and ODDs using computational models. For example, Carreiro-Perpiñán and Goodhill [112] used elastic net models which minimize the total length of wiring while satisfying a compromise between the uniformity U and the continuity C of the cortical representation of the various features of the stimuli.[24] To do this, they minimize an energy $E = U + \beta C/2$. We work in the space of features

$$\{\text{position } a = (x, y), \text{ ocular dominance, orientation (angle and selectivity) in polar coordinates}\},$$

corresponding to a lattice of $N = N_x \times N_y \times N_{\text{OD}} \times N_{\text{OR}} \times 1$ stimuli in the space

$$T = [0, 1] \times [0, 1] \times [-\ell, \ell] \times [-\pi/2, \pi/2] \times [0, \rho] \ .$$

Let y_m be the centre of the receptive field of neuron m, $m = 1, \ldots, M$, in the space T and x_n the vector of stimulus n, $n = 1, \ldots, N$. The uniformity U is defined by

[24] We have already encountered this problem of minimizing the wiring in Sect. 4.4.5.1.

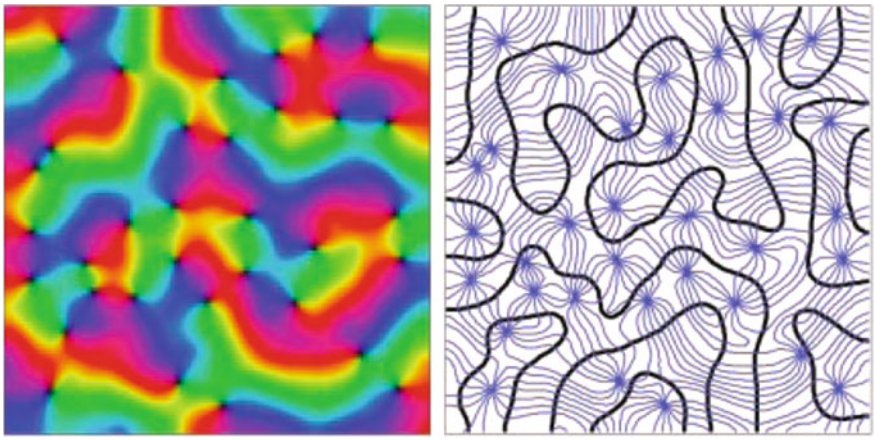


Fig. 4.119 Field obtained by minimizing the energy E (see text) for $p = 1$ and $\beta = 10$. From [112]

$$U(y_1, \ldots y_M, K) = -K \sum_{n=1}^{n=N} \log \left[\sum_{m=1}^{m=M} \exp\left(-\frac{1}{2} \left\| \frac{x_n - y_m}{K} \right\|^2 \right) \right] ,$$

where K fixes the size of the receptive fields. The continuity C is defined by

$$C(y_1, \ldots y_M) = \sum_{m=1}^{m=M} \| f(y_m) \|^2 ,$$

where f is a linear combination of the neurons neighbouring neuron m which approximates an order p differential operator. Figure 4.119, already shown in Sect. 4.4.3 (see Fig. 4.40), illustrates the result of minimizing for $p = 1$ and $\beta = 10$. Figure 4.120 gives the result for $p = 3$. We see that $p = 1$ looks like the empirical maps.

4.10.1.4 A LISSOM Model

Other models of joint self-organization of maps of orientation, direction of motion, and ocular dominance can be found in the work of Bednar and Miikkulainen [113, 114]. They are based on the laterally interconnected synergetically self-organizing map (LISSOM) model. Their 2005 book entitled *Computational Maps in the Visual Cortex* [80], written with Yoonsuck Choe and Joseph Sirosh is particularly interesting:

> [It] presents a unified computational approach to understanding the structure, development and function of the visual cortex.

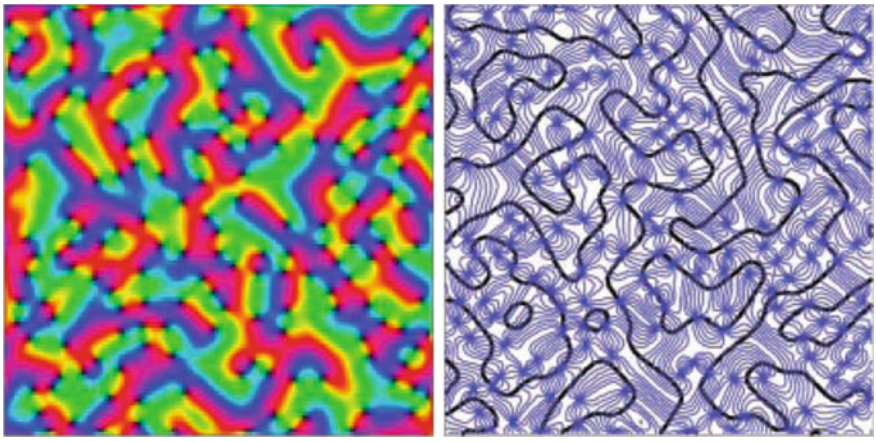


Fig. 4.120 Field obtained by minimizing the energy E (see text) for $p = 3$ and $\beta = 10$. From [112]

For example, Fig. 4.121 shows the result of joint learning of orientation (OR, colours), ocular dominance (OD, boundaries), and direction of motion (DR, arrows) in a LISSOM model on the basis of a flux of simple stimuli. Note that the DRs are reasonably orthogonal to the ORs and that in each orientation patch there are two sub-patches with opposite directions (orthogonal to the preferred orientation of the patch). This synthesis provides a good model for the cat. We shall return to direction maps in Sect. 5.11 of Chap. 5.

4.10.2 Independent Maps and Transversality Principle

We see that, after the spatial frequency, the ocular dominance is another parameter that is implemented in the 2D neural layers. With the orientation, an abstract 3D structure collapsed to two dimensions. Now, we have an abstract 5D structure. Such a drastic reduction in dimension obviously raises questions about how the *independence* of the parameters can be represented in two dimensions. It seems that the solution discovered by evolution was to maximize a *transversality condition*: the boundaries of the frequency domains and those of the ODDs are strongly transverse to the lines of the iso-orientation field. Clearly, if there is more than one extra parameter, the transversality cannot be strong everywhere, but there can nevertheless be optimization of two opposing constraints on transversality.

Nicholas Swindale explicitly addressed the question of how to understand the optimization of the interactions between different maps in his paper [115] entitled *How many maps are there in visual cortex?* (see also [116]). To simplify, Swindale started with N binary variables (e.g. binarizing continuous variables like the orien-

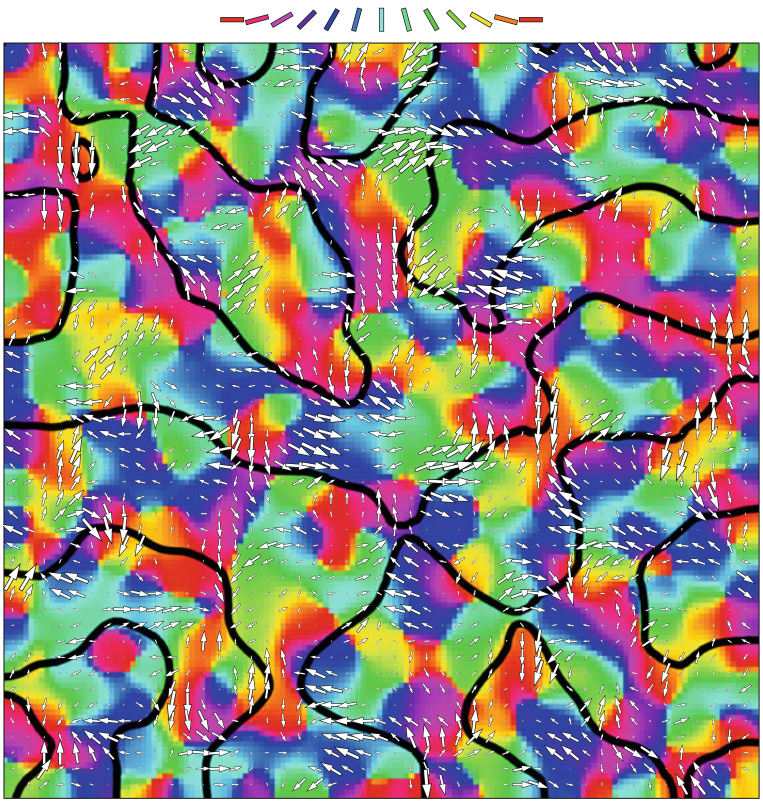


Fig. 4.121 Orientation maps (*colours*), ocular dominance (*boundaries*), and direction of motion (*arrows*) in a LISSOM model. The DRs are strongly transverse to the ORs. In each orientation patch there are two sub-patches with opposite directions (orthogonal to the preferred orientation of the patch: for example, *vertical arrows* in the *red regions* correspond to the horizontal orientation, etc.). From [80, Fig. 5.29]

tation, ocular dominance, or spatial frequency), regions where a variable is constant corresponding to the different values of the features of the given stimuli. He then asked about the maximal N ensuring good coding efficiency, the effect of introducing a new map on the other maps, and the information provided by a map about the presence of other maps. To answer these questions, he used the well-known algorithm called self-organizing feature maps due to Kohonen, which is analogous to the algorithms used in the previous section.

For this, we take a retinal grid $(i, j), i, j = 0, 1, \ldots, M$ $(M = 150)$ and associate with the neuron (i, j) the vector $w_{ij} = (x, y, t_1, \ldots, t_N)$ defined as the set of central values of its receptive field and the features t_k it detects. This is therefore a discrete and generalized version of what we introduced at the outset when we said that a simple $V1$ neuron codes a contact element $(a = (x, y)\ , p)$. In other words, we work in the fibre bundle $\pi_N : V = R \times T \to R$, where T is a fibre of dimension N. For

Fig. 4.122 Swindale's model for 6 maps. From Swindale [115]

their part, the stimuli are also coded by vectors v of V, $v_s = (x_s, y_s, u_1, \ldots, u_N)$ and we take enough of them (2.5×10^6) to ensure that the learning is significant. We select random sequences of them and, at each step v, the w_{ij} are incremented by Δw_{ij} applying the following standard rule: if (i_0, j_0) is the neuron in the network which is closest to the v, then $\Delta w_{ij} = \varepsilon h(r)\left(v - w_{ij}\right)$, where ε is the learning rate ($\varepsilon = 0.01$), r is the distance between (i_0, j_0) and (i, j), and $h(r) = \exp\left(-r^2/2\sigma_c\right)$ is the learning Gaussian ($\sigma_c = 2.5$). To minimize, we then carry out simulated annealing, the algorithm already discussed in Sect. 4.4.5.1 when minimizing the total length of wiring in $V1$. For example, Fig. 4.122 shows a piece for $N = 6$. There are $2^N = 64$ feature values, each coded by a colour.

As the variables are binary, each map comprises blobs and stripes whose intersections encode the structural relations between the maps. The morphology of the individual maps does not change much qualitatively when other maps are added, but their structural relations change much more. Figure 4.123 shows the boundaries for $N = 6$. We observe strong transversality relations.

Hongbo Yu et al. [46] also examined the relations between the three maps, viz. orientation, ocular dominance, and spatial frequency, but using continuous variables and focusing on the way strong transversality codes the independence of the associated variables. They considered the *gradients* of the variables and showed first that the gradients are maximal in disjoint regions and then that the transversality is *maximal* when the two gradients are jointly high enough:

> Two features are mapped orthogonally in their high-gradient overlap regions. [46, p. 277]

Figure 4.124 shows the orientation field lines (the level lines of the 'orientation' variable) and the OD level lines. The pinwheel centres and the ODD boundaries

Fig. 4.123 Boundaries of the domains of the $N = 6$ binary variables in Fig. 4.122. From Swindale [115]

avoid one another since the pinwheel centres are located on the ridges of the ODDs. The grey zones are those where the two gradients are jointly highest. We observe that transversality is indeed very strong; in fact, we almost have orthogonality.

One interesting result obtained by Swindale in [115] concerns the retinotopic map. Figure 4.125 shows the projection of the cortical network on the retinal space for the case $N = 7$. Note that there are many folds and pleats. Two remarks are in order:

1. Firstly, there should be no surprise in finding folds and pleats. Indeed, according to the well-known theorem due to Hassler Whitney and generalized by René Thom, the only singularities possible in a generic differentiable map between 2D manifolds are fold lines and isolated pleat points.
2. Secondly, these singularities are not those of the conformal retinotopic map (a complex logarithm, see Sect. 4.2) between the retinal positions and the cortical positions. They are induced by what happens in the fibres of the bundle $\pi_N : V = R \times T \to R$ and show that the values of the features $w_{ij} = (x, y, t_1, \ldots, t_N)$ do not generate a single-valued section of π_N, but rather a multivalued section whose image is a surface in V that is not everywhere transverse to the fibres.[25]

[25] For an adequate treatment of this point, one must introduce the rather technical geometric notion of a Lagrangian sub-manifold. We shall say a little more about this in the second volume. Here we only make elementary remarks about the geometry.

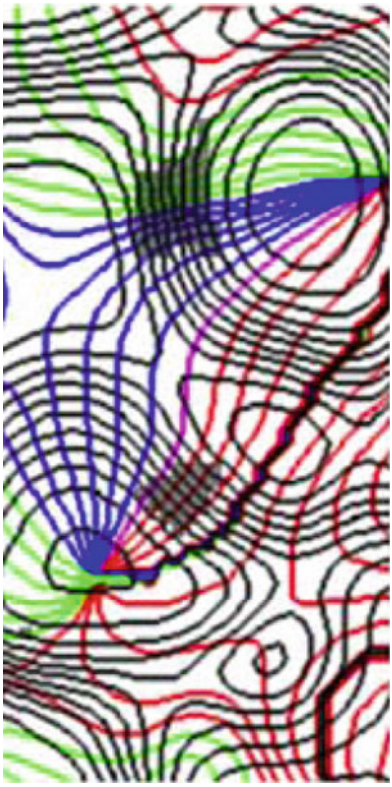


Fig. 4.124 Iso-orientation lines and ODD level lines. *Grey regions* are those where the two gradients are at their highest. We observe that transversality is very strong there. From Yu et al. [46]

4.10.3 Binocularity

4.10.3.1 Ocular Dominance and Binocular Disparity

The ocular dominance maps and ODDs of the controlateral and ipsilateral monocular cells must be studied in relation to *binocular disparity* (BD) of the binocular cells. Among others, Prakash Kara and Jamie Boyd have studied the functional architecture of the BD and its relationship with OD in a piece of the $V2$ area (18) of the cat [117]. Using two-photon confocal microscopy, method already discussed in Sect. 4.7.3 in the context of Ohki's work on pinwheels and which can be used to measure the activity of several hundred individual neurons in layer 2/3 (a region of about 300 μm), Kara and Boyd were able to show that there is a BD selectivity map and that OD and BD are *mutually independent* at the level of individual neurons. Once again, this independence is manifested through strong transversality properties: in regions where the gradients of the two variables are both high, they are orthogonal.

The stimuli used are gratings with the preferred direction of motion and orientation of the cell being measured. Either one of the eyes is stimulated by such a grating

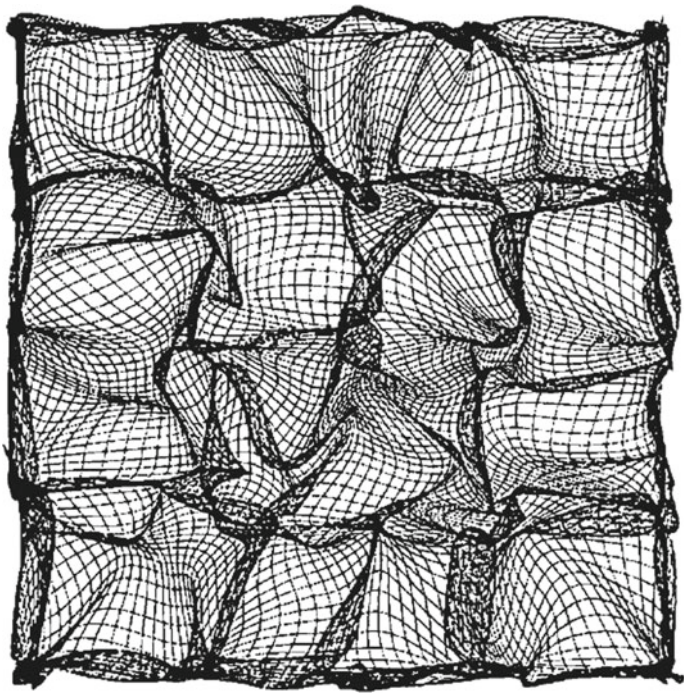


Fig. 4.125 Projection of the cortical network on the retinal space in the Swindale $N = 7$ model. From Swindale [115]

and the other by a uniformly grey stimulus, or both eyes are stimulated by such gratings, with the same orientation and direction, but out of phase (8 values of the phase difference are used). Figure 4.126 shows the responses of several cells to the right eye (R) and the left eye (L) and to the phase of the BD. Figure 4.127 shows maps of the BD and the OD, confirming that, in regions where both gradients are high, the level lines are close to orthogonal.

4.10.3.2 Bistability and Binocular Rivalry

Since we have been discussing binocularity, let us say a word about the phenomena of *binocular rivalry* and *bistable images*, important subjects when studying the relationship between high and low levels of visual cognition.

Ambiguous bistable images like the Necker cube, which can be perceived in two different ways, are well known. When we stare at the image for long enough, we observe regular switches between the conflicting interpretations. It is thus natural to assume that there are two competing populations of neurons implementing the two interpretations, the dominant population selecting the perceived interpretation at a given time and inhibiting the other. When the inhibition falls sufficiently or the cues

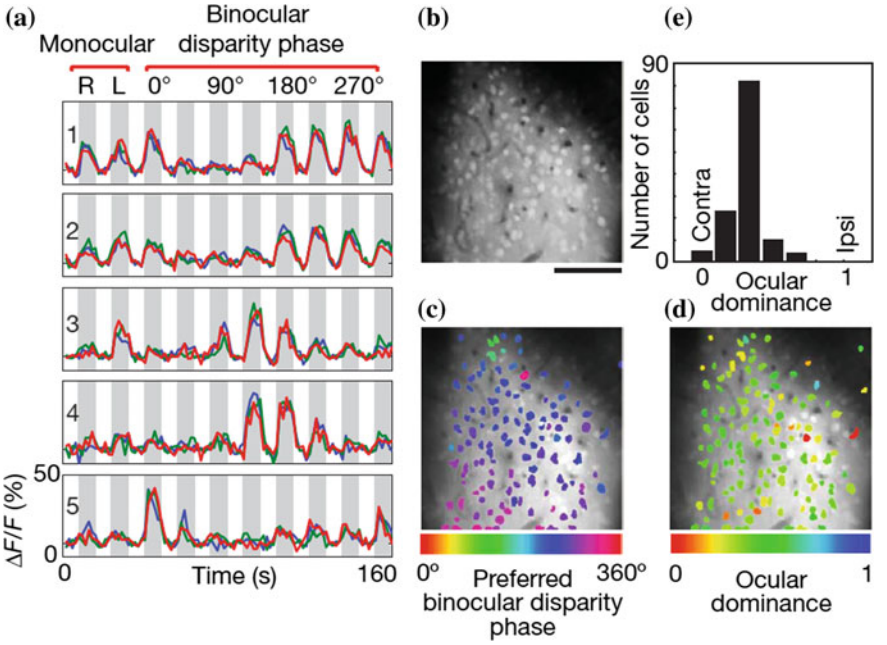


Fig. 4.126 Responses of 5 cells to monocular stimuli and the phase of the binocular disparity. For example, cell 1 responds to L and R and to phase differences of 0, π, $5\pi/4$, $3\pi/2$, and $7\pi/4$, but not to phase differences of $\pi/4$, $\pi/2$, and $3\pi/4$. **c** gives the BD, **d** the OD, and **e** the histogram of the OD. From Kara and Boyd [117]

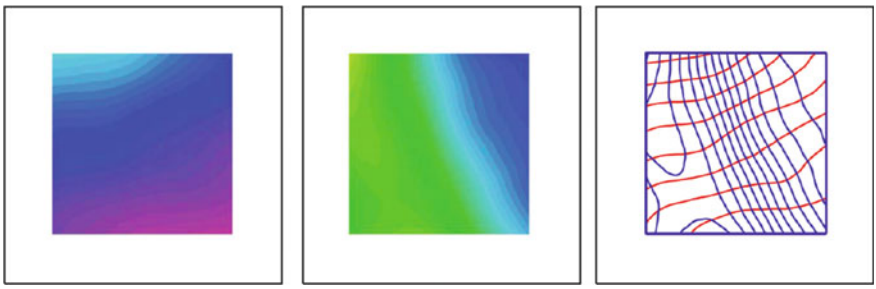


Fig. 4.127 Maps of the BD and the OD. In regions where both gradients are high, level lines are close to orthogonal. From Kara and Boyd [117]

favouring the other interpretation evolve sufficiently, a switch can occur. Here, we should ask at what level in the elaboration of percepts an interpretation is selected. Is it a basically low-level mechanism or rather some high-level cognitive mechanism involving knowledge? In actual fact, it almost certainly involves both because any interpretation involves *inference*.

Spontaneous bifurcation phenomena are well known in every field. We shall meet some fine examples in the second volume. They can be modelled using the following guiding idea introduced by René Thom in [118] and [119], an idea already discussed in Sect. 1.3 of the *Preface*. Let S be a system satisfying the following hypotheses:

1. Inside S, there is an internal process X which specifies the internal states that S can occupy.
2. The internal process X specifies all the internal states of S.
3. There is a selection criterion I which selects the current state among the various possible internal states on the basis of certain criteria specific to the system, and which may vary significantly.
4. Finally, the system S is controlled, with good regularity properties, by a certain number of control parameters. These vary in a space W called the *external space* (or control space, or substrate space) of S, to distinguish it from the internal space of S. The internal process X is thus a process X_w depending on w.

If $\mathfrak{X}$ is the space of possible internal processes, the system S will then be described by the field $\sigma : W \to \mathfrak{X}$ associating the corresponding internal process X_w with $w \in W$, and also by the selection criterion I. As w varies, the actual state will vary and, for certain critical values of w, it may bifurcate towards another internal state.

The best known model of this kind is known as the cusp model. The states of the system are represented by the minima of a potential function $f_w : \mathbb{R} \to \mathbb{R}$ parametrized by a parameter w varying in a 2D external space W, and the internal dynamics X_w are the gradient dynamics of the f_w. Depending on the value of w, f_w has either one minimum or two minima A and B separated by a threshold, and when there are two minima, these are in competition. Hence, there are two kinds of possible bifurcation corresponding to the exceptional values of w varying along three lines of W: along two of these lines, A or B disappear, and along the third, A and B compete and balance each other (see Fig. 4.128).

We shall return to this model in Sects. 5.9.2 and 5.11.3 of Chap. 5. It was used to model bistable images in the 1970s by Christopher Zeeman, the minima of f_w corresponding to attractors of the relevant neural dynamics.

Binocular rivalry does not concern the conflict between two possible interpretations of a given stimulus, but something deeper and more enigmatic, namely the conflict between two different stimuli. We consider two stimuli A and B, e.g. a chessboard and a face, and we present one to each eye using appropriate apparatus. The two stimuli are thus both present in $V1$. However, the subject does not perceive a superposition $A + B$, but in each case only one image, the two images alternating regularly. In other words, there are successive intervals of time in which the subject perceives only A or B (exclusivity and uniqueness), and between these intervals, short transition periods during which the stimulus that is actually present, say A, is destabilized and bifurcates towards the other stimulus B. Temporal series of such spontaneous shifts have been studied in detail.

There are roughly speaking two classes of theory to explain the phenomena of binocular rivalry: on the one hand, low-level theories which say that the suppression

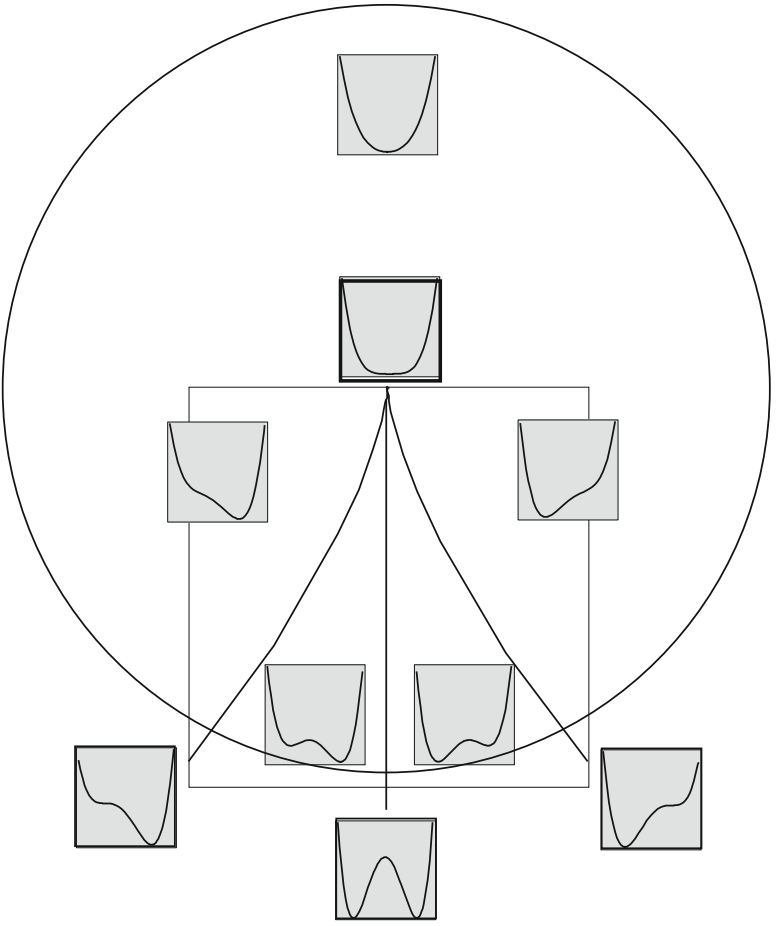


Fig. 4.128 Universal unfolding f_w of the cusp singularity. The parameter w varies in a 2D external space W and the internal dynamics is the gradient descent $-\text{grad}\,(f_w)$. Depending on the value of w, f_w has either one minimum or two competing minima separated by a threshold. There are two kinds of catastrophe along three strata in W : either one of the minima bifurcates by collision with the maximum, or the two minima compete and balance each other. This figure will be taken up again in Fig. 5.39 and explained further in Sect. 5.11.3 of Chap. 5

of one of the competing stimuli already begins in the $V1$ area, or even before $V1$, and on the other hand, high-level theories which say that the suppression is made through cognitive inferences. According to the low-level hypothesis known as interocular competition, rivalry arises because the two images cannot be merged by binocular stereopsis and this mismatch blocks one or other of the monocular pathways. But the two kinds of theory are probably both partly correct. Indeed, fMRI studies show that high-level non-visual areas (frontal and parietal) are active during bifurcations. This

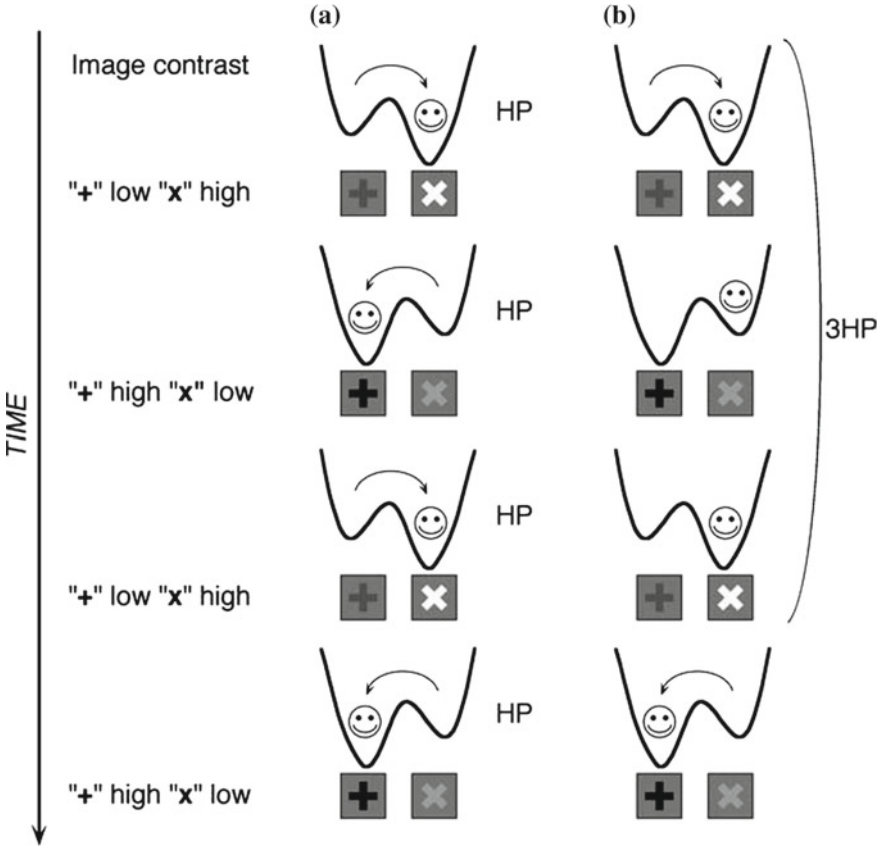


Fig. 4.129 Cusp model for binocular rivalry. The two percepts are symbolized by + and × and the depths of their respective potential wells are alternately modulated by the image contrast. Noise provides the means to overcome the threshold (the maximum). Stochastic resonance is manifested by the fact that the distribution of the dominance periods has resonance peaks for the values $(2k + 1)\,HP$, where HP is the half period of contrast modulation. **a** Peak for HP. **b** Peak for $3HP$. From Kim et al. [121]

is therefore a complex phenomenon. For a good summary, the reader is referred to the paper [120] by Randolph Blake and Hugh Wilson.

The cusp dynamical model turns up in certain studies of binocular rivalry. For example, Yee-Joon Kim et al. [121] use it to apply the methods of stochastic resonance. In Fig. 4.129, the two percepts are symbolized by + and × and the depths of their respective potential wells are alternately modulated by the *image contrast*. Noise is introduced to overcome the threshold represented by the maximum separating the two minima. Stochastic resonance is manifested by the fact that the distribution of the dominance periods, i.e. the time series of the shifts, has resonance peaks for the values $(2k + 1)\,HP$, where HP is the half period of contrast modulation.

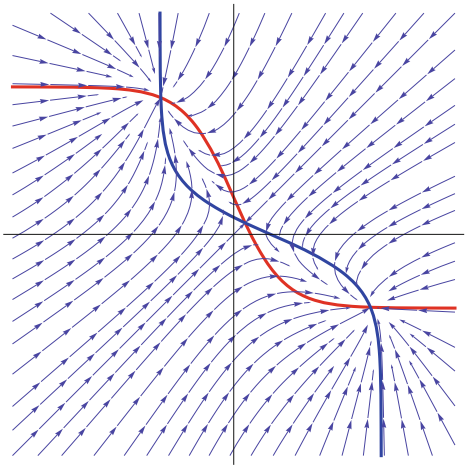


Fig. 4.130 Initial fast field (h_1, h_2) with two nullclines $\partial h_1/\partial t = 0$ (*red*) and $\partial h_2/\partial t = 0$ (*blue*) intersecting at three fixed points (two attractive nodes and one partially repulsive saddle point). Recomputed from de Jong [122]

In his 2008 master's thesis entitled *The Dynamics of Visual Rivalry* [122], Ties Marijn de Jong investigated a simple but explicit model of bifurcation with slow/fast dynamics. The competing stimuli were orthogonal gratings. By averaging the dynamics of single neurons, we obtain, for the two populations $i = 1, 2$:

- The local activity fields h_{i} of the percepts (average membrane potentials), these being the fast variables.
- Slow adaptation variables a_{i} expressing the fact that the activity of the dominant percept gradually weakens.

Introducing a scale change parameter $\tau \gg 1$ between slow and fast times and averaging the standard neural equations, the author produced the following model:

$$\begin{cases} \tau \dfrac{\partial h_i}{\partial t} = X_i - (1 + a_i)\, h_i - \gamma\sigma\left(h_j\right) \ , \\ \dfrac{\partial a_i}{\partial t} = -a_i + \alpha\sigma\left(h_i\right) \ , \end{cases}$$

where α and γ are constants arising from the underlying neural interactions, the X_i are the initial forces of the two stimuli as inputs, and where $\sigma\,(h)$ is a sigmoid curve, i.e. a C^∞ approximation to the Heaviside step function, equal to zero for $h < 0$ and unity for $h \geq 0$. As the mean firing rate is $\sigma\,(h)$, $h_i < 0$ implies that the stimulus i is suppressed by the perceptual system.

If we begin with a symmetric rivalry, that is two stimuli of the same strength and equal adaptation coefficients, and examine the field (h_1, h_2), a classic bifurcation scenario comes to light (see Fig. 4.130). The slow manifolds are the nullclines $\partial h_i/\partial t = 0$ with a fixed. At the beginning, they intersect at three points which are equilibrium points since the field vanishes there. Two of these points are attractive (nodes), on either side of a third which is partially repulsive (a saddle point). For each

attractor, one of the h_i is strictly positive while the other (h_j) is strictly negative, implying that only the stimulus i is perceived. If we begin in a state with no activity, close to 0, and if we activate the system with the stimuli, the fast dynamics projects the system onto one of the attractors. From this point, the slow dynamics modifies the field and adaptation breaks the initial symmetry $a_1 = a_2$. It displaces the nullclines, thereby leading to a bifurcation by the merging of the initial attractor and the saddle point. The system then jumps into the other attractor. We shall return to this kind of fast/slow dynamics and bifurcations in the second volume.

The most profound causes of perceptual bistabilities are still debated. For example, in [123], David Leopold and Nikos Logothetis criticize the favoured hypothesis according to which the bifurcations (spontaneous reversals) result from antagonistic connections. In their view:

> [Alternations] reflect responses to active, programmed events initiated by brain areas that integrate sensory and non-sensory information to coordinate a diversity of behaviours. [123, p. 254]

4.10.4 Blobs and Colour

We should also say a word about 'blobs' of cytochrome oxidase (CO) in $V1$ and $V2$ which are sensitive to *colour* and already process it in $V1$. By imaging, we obtain colour-selective response maps in those regions which, in $V1$, are centred on the ODDs, but which are in fact barely selective to orientation and which do not overlap with the central regions of the pinwheels. In $V2$, the colour regions coincide with the thin stripes and the orientation regions with the pale and thick stripes. There is therefore a kind of functional segregation of colour and orientation corresponding in part (this point is debated) to the distinction between blobs and interblob regions.

The structure of the blob map has been studied by many specialists, such as Haidong Lu and Anna Roe [124], using optical imaging methods on the macaque. In primates, the blobs in $V1$ are found mainly in layers 2 and 3, but also in layers 1, $4B$, 5, and 6. They measure in the range 150–250 μm, and there are about 5/mm^2. They tend to be centred on the ODD axes with an average period of about 350 μm (see Polimeni et al. [87]). Figure 4.131 shows their distribution. It is valid for a monkey like the macaque, but in a prosimian primate like the bush baby, the constraints on the positions of the pinwheels and the blobs within the ODDs are less stringent (see Xiangmin Xu et al. [45]). Finally, Fig. 4.132 shows a diagram of $V1$ with pinwheels, ODDs, and blobs.

We shall return to the band structure of $V2$ in Sect. 5.8 of Chap. 5.

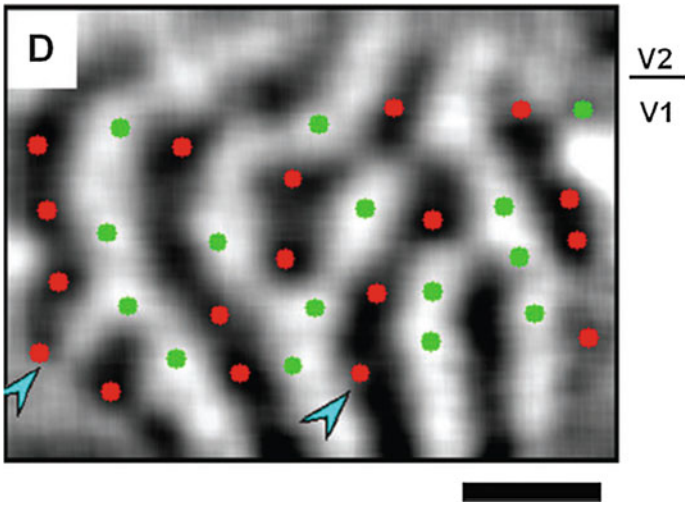


Fig. 4.131 Distribution of blobs of cytochrome oxidase in the ODDs of a macaque. *Red*: left eye. *Green*: right eye. They are centred preferentially on the ODD axes. *Arrows* indicate blobs furthest from the centres. *Scale bar*: 1 mm. From Lu and Roe [124]

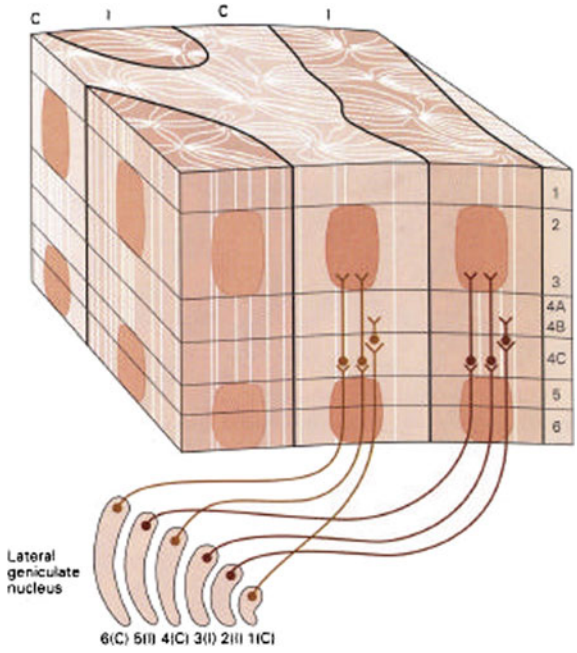


Fig. 4.132 Diagram of *V*1 with pinwheels, ODDs, and blobs. From [24]

4.10.5 *Functionality of Maps*

Here, we see just how rich the structure of $V1$ is. All the various 'secondary' variables, engrafted in the sense of Hubel onto the primary spatial variables, are processed by micromodules measuring $400\,\mu m \times 800\,\mu m$ and containing some 60 000 neurons. With their functional architectures, they constitute very precise maps that are just as precisely correlated with one another. This is how, according to Yu et al. [46], dimensional reduction can:

> […] smoothly map several (more than two) response properties onto a two-dimensional cortical surface.

It is also how it can solve the fundamental functional problem of ensuring the uniformity and continuity of features with respect to their positions by the design of the neural hardware. To minimize the complexity of this structure and make a 'simplex' structure in the sense of Alain Berthoz (who generalized what we shall say about 1-jets in Sect. 5.4.6 of Chap. 5), evolution grouped together functionally related neurons locally through the spatial extent of the cortical layers.

4.11 Hemispheres and Callosal Connections

Here, we shall say something about the way the two halves of $V1$ in the two hemispheres are connected up. This provides a remarkable example of a geometrical *gluing* process. It is the *callosal connections* of the corpus callosum that do this. The corpus callosum is the biggest bundle of nerve fibres in mammals. In humans, it comprises some 200 million axons.[26] The region of the visual field located close to the vertical meridian, called the transition zone (TZ) or visual midline (VM), is projected onto the two parts of $V1$ in the vicinity of the $V1$–$V2$ boundary.

The structure of the gluing map is fascinating. Figure 4.133 shows this map for the cat (areas 17/18). The sub-zones A, B, C, D, E of a hemisphere are connected to the zones with the same label in the other hemisphere: zones within the TZ are connected to zones outside it and conversely.

If we cut through the fibres in the optical chiasm coming from the nasal hemiretina, while keeping those from the temporal hemiretina (split-chiasm preparation), the right visual hemifield projects onto the left $V1$ area via the left eye and the activity of the right $V1$ area becomes entirely due to the callosal connections (see Fig. 4.134). Depending on which eye is stimulated, we can thus activate either the geniculo-cortical pathway or the transcallosal pathway.

The problem is to understand the distribution of the callosal connections in relation to the pinwheel structure in $V1$. This is a very delicate matter. An experiment by William Bosking (Bosking et al. [126]) shows (see Fig. 4.135) that, for the tree shrew

[26] Recall that the optic nerve contains about 1.5 million axons, so less than a hundredth of the number.

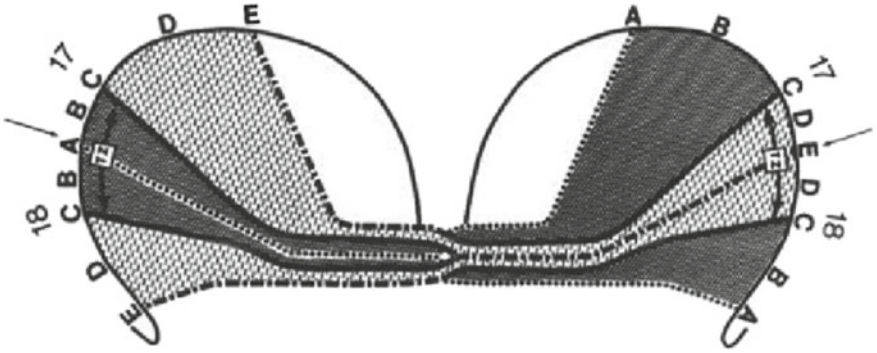


Fig. 4.133 Structure of the callosal gluing map between the two hemispherical parts of $V1$ and $V2$ in the cat. Zones A, B, C, D, E of one hemisphere are connected to zones with the same label in the other hemisphere. From Rochefort [125]

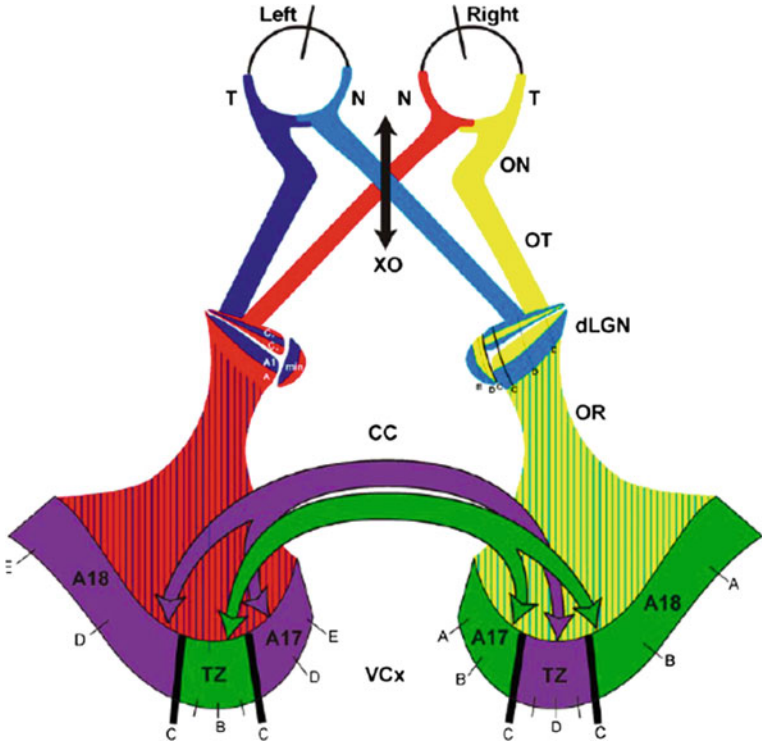


Fig. 4.134 If we cut through the optical chiasm, the right visual hemifield projects onto the left $V1$ area via the left eye and the activity of the right $V1$ area is entirely due to the callosal connections. From Rochefort [125]

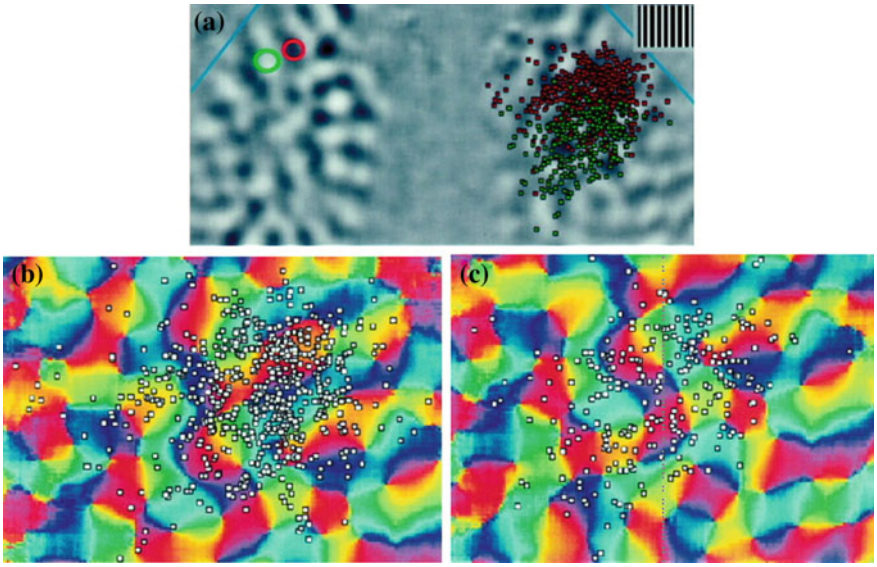


Fig. 4.135 Callosal connections in the tupaia (tree shrew). The two *red*/*green* sites of $V1_L$ in **a** project onto $V1_R$ regardless of orientations as shown in **b** and **c**. In $V1_R$, we find a rather uniform distribution of activated neurons: they are distributed across several regions of different preferred orientation (*colour*). From Bosking et al. [126]

(tupaia), if we inject rhodamine into a small region of $V1_L$ with vertical preferred orientation (red circle in a black region) and fluorescein into another small region with horizontal preferred orientation (green circle in a neighbouring white region), the callosal projections onto $V1_R$ do not exhibit any orientation specificity:

> Callosal connections appear to terminate without regard for the map of orientation preference, showing little sign of the orientation-specific modular and axial specificity that is characteristic of long-range horizontal connections. (Bosking et al. [126])

For the cat, the situation seems to be different. The classic work by Olavarria [127] showed that the distribution of retrogradely labelled callosal cells is polarized by the ocular dominance domains:

> Callosal cells correlate preferentially with contralateral ODCs (ocular dominance columns) within the 17/18 transition zone (TZ), and with ipsilateral ODCs in regions of areas 17 and 18 located outside the TZ.

Milleret and Rochefort [128, 129] made further investigations using in vivo optical imaging of intrinsic signals coupled with a method for reconstructing and labelling the callosal axons. They showed that the callosal connections *do conserve* orientation selectivity. More precisely, using the split-chiasm preparation with an injection site in the left hemisphere, they were able to reconstruct the distribution, in layers II, III, and IV (upper part) of the transition zone (ZT) of the right hemisphere, of the synaptic buttons of several labelled callosal axons. They thus observed that these axons

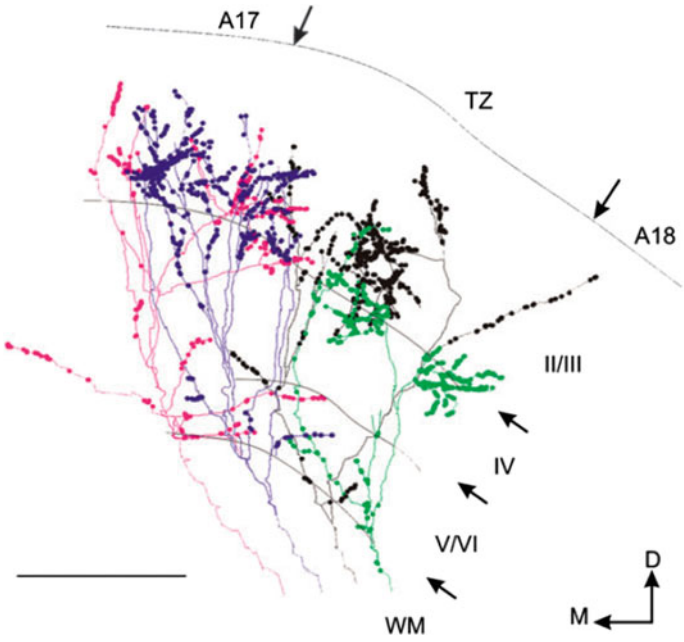


Fig. 4.136 Front view of the reconstruction of four callosal axons in cats (two neighbouring injection sites and two axons per site) and their clusters of synaptic buttons in layers II, III, and IV (*upper part*) of the transition zone *TZ* between area 17 and area 18. *WM* white matter, *D* dorsal, M medial. *Scale bar* 500 μm. From Rochefort et al. [129]

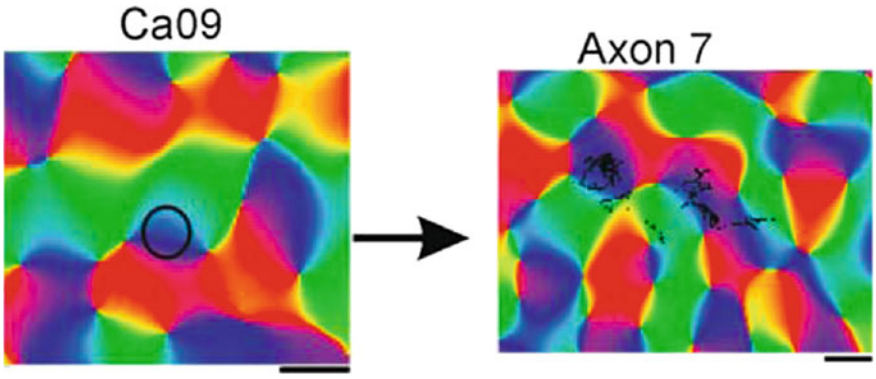


Fig. 4.137 Conservation of orientation by transcallosal connections in the cat. *Left* Injection site in the left hemisphere. *Right* Reconstruction of the distribution of the synaptic buttons of a labelled callosal axon in the right hemisphere. The axon projects into isochromatic zones (*blue*) of the same colour as the injection site. From Rochefort [125]

project into isochromatic zones relative to the colour of the injection site, thereby showing that the orientation is conserved. Figure 4.136 shows four reconstructed axons. Figure 4.137 shows the conservation of orientation by the transcallosal connections.

4.12 Homogeneous and Inhomogeneous Qualities

4.12.1 Responses to Homogeneous Surfaces

Up to now the basic geometric feature we have studied is the orientation of contour elements. However, if we consider in the most qualitative phenomenological way a natural image made up of identifiable objects, we see immediately that there are two fundamentally different kinds of local structure:

1. Locally *homogeneous* points near which features vary only slightly and in a continuous way.
2. Locally *heterogeneous* points near which certain features reveal discontinuities.

This dichotomy between homogeneous and heterogeneous is absolutely fundamental and very general, going back at least to Aristotle, who, himself referring to Anaxagoras, made a distinction in biology between homoeomeric parts ('made up' of similar [omios] parts [meros]) and anhomoeomeric parts.[27] As we shall see in the second volume, this distinction is of key importance in Husserl's phenomenology of perception. It also underpins René Thom's morphological models. Thom referred to the locally homogeneous points of the substrate of a form as its *regular points* and the locally inhomogeneous points as its *singular points*. The latter are the phenomenological manifestation of symmetry breaking in the underlying processes, and this is why the theory of singularities, the theory of bifurcations, and the theory of symmetry breaking formed the backbone of all his models of natural morphology, not just in physics, chemistry, and biology, but also in psychology and the social sciences.

These models are based on Thom's guiding idea as discussed in Sect. 4.10.3.2. Here, we apply to the case where the external space W is that of the visual field and where the attractors of the internal dynamics X_w implement sensory qualities. In this case, when the internal dynamics X_w undergoes a bifurcation, the point w is singular.[28]

The dichotomy between regular and singular points was also important in the visual theories of Grossberg [131], which assume that there are two fundamental systems of visual perception:

---

[27] *On generation and corruption*, I, 5, 312b.

[28] All the details can be found in René Thom's two books [118] and [119]. For a didactic introduction, see [130].

1. The *boundary contour system* (BCS), which detects, enhances, and completes edges using a 'spatially long-range cooperative process'.
2. The *featural contour system* (FCS), which fills in the regions bounded by the BCS with qualities ('featural filling-in'): 'These filling-in processes lead to visible percepts of colour-and-form-in-depth at the final stage of the FCS'.

According to Grossberg [131, p. 35]:

> Boundary contours activate a boundary completion process that synthesizes the boundaries that define perceptual domains. Feature contours activate a diffusion filling-in process that spreads featural qualities, such as brightness or color, across these perceptual domains.

It is thus interesting to know whether the dichotomy between homogeneous-regular and heterogeneous-singular is relevant in the coding of features in $V1$ or $V2$.[29] Toshiki Tani et al. [132] reached this conclusion by exhibiting in the 17/18 areas of the cat a map of neurons that detect the *interior* of homogeneous domains and respond to spots which completely cover their receptive fields. These neurons are located in the $V2$ area near the boundary with the $V1$ area. The authors conclude that, in the primary visual cortex, there is not only a representation of contours but also a representation of *surfaces*. They stress that these regions of $V2$ are small and centred on the pinwheels where all the orientations are present.

As these zones for local detection of homogeneous surfaces are located close to the boundary between $V2$ and $V1$, which is the transition zone (TZ) of the vertical meridian where the corpus callosum comes into play, as we have seen, the authors suggest that these neurons may transmit information about the interior of homogeneous zones from one hemifield and one hemisphere to the other:

> One possible function of the surface-responsive regions may be to link visual information about large surfaces extending across both the right and left hemifields. [132, p. 1123]

We shall now specify the problem of regular and singular points by returning to the processing of *colour*.

4.12.2 Colour Processing

4.12.2.1 First Steps: Goethe, Helmholtz, Hering

We discussed colour processing in the retina in Sect. 3.2.5 of Chap. 3, and we have just been talking about blobs in $V1$ and $V2$. Let us now say a little more about this.

The trichromatic theory developed by Thomas Young and Hermann von Helmholtz in the first half of the nineteenth century involved a remarkable anticipation of the

[29] The pinwheel structure is based on the dichotomy between regular and singular points, but this concerns field lines of the orientation field. Here we are talking about something quite different.

way the retina actually processes colour. However, it was clearly inadequate. Goethe had already recorded decades of phenomenological observation of colour perception in his *Farbenlehre* of 1810. He was interested in the phenomenon of colour as an experienced quality and sought to describe this experience in as precise a way as possible, since accurate description was for him essential. From the details of these observations (which are experimental in the scientific sense and also in the subjective sense of experience), he hoped to infer more explicative underlying mechanisms. In particular, he noted that the persistent image of a coloured figure has another colour: pieces of white paper on a yellow wall become tinted with purple, and when we remove an orange sheet of paper from a white wall, this induces a blue image and an orange background. As he notes in § 60:

> These phenomena are of the greatest importance, because they point to laws of vision.

Indeed, they point to the antagonisms between complementary colours, namely yellow/purple (Y/R+B), blue/orange (B/Y+R), and red/green (R/B+Y), where R = red, G = green, B = blue, and Y = yellow.

Later, in 1892, the great Austrian physiologist Ewald Hering (1834–1918) developed an antagonistic theory, known as the opponent process theory, in a simplified tetrachromatic form, assuming that there are in fact four fundamental colours RGBY, organized in two pairs of complementary colours R/G and B/Y. With this theory, he was able to solve the problem of yellow, which seemed to be an elementary colour rather than a mixture. Hence, the idea that there were indeed three RGB detectors, but four simple primary physiological colours.

Hering's theory of pairs of complementary colours anticipated in a quite remarkable way what we know today about colour processing in the cortex, between the LGN and *V* 1.

4.12.2.2 Chromatic Opponency: Single Versus Double

In the LGN, colour-tuned cells are essentially parvocellular 'single-opponent' cells detecting R/G contrasts, i.e. L/M.[30] Magnocellular cells are insensitive to colour. However, there are also cells in the koniocellular pathway, discovered in 1994, which detect B/Y contrasts, i.e. S/(L+M). In *V* 1, the colour cells located in blobs are of two kinds. The first calculate *chromatic contrasts* R/G, i.e., L/M, and B/Y, i.e. S/(L+M). The second calculate *spatial contrasts* between the complementary colours R/G and B/Y. These are therefore doubly antagonistic (double-opponent cells) which calculate opponency in both the external space and the chromatic space. They can be denoted $\pm$L/ $\mp$ M. In other words, these are orientation cells that also detect chromatic contrasts. They are essential because, as emphasized by Shapley and Hawken [133, p. 701]:

[30] Recall (see Sect. 3.2.5 of Chap. 3) that the three kinds of cones in humans are L/M/S, *L* red, *M* green, *S* blue.

Color, form and motion are inextricably linked as properties of objects in visual perception and in the visual cortex.

So from the lowest levels of the $V1$ and the $V2$ areas, there is a functional entanglement between the spatiality of perceived scenes and the colours of objects making them up. The brain must reconstruct by inference from the colour the objective reflectance of the surfaces perceived and this independently of a host of extremely variable factors such as the illumination of sources, indirect irradiation, angles of incidence, and reflection. This *inverse problem* is exceedingly hard to solve. The *direct* problem is this: for each position $a = (x, y)$ of the visual field, given the reflectances $\rho(a, \lambda)$ and an illumination spectrum $\sigma(a, \lambda)$, to calculate the spectrum $\gamma(a, \lambda) = \sigma(a, \lambda)\rho(a, \lambda)$, where $\gamma(a, \lambda)$ is itself encoded locally by the excitation of the L/M/S cones at a. The inverse problem is not well posed, like most inverse problems, and can only be solved if the system has priors at its disposal, i.e. priors in the Bayesian sense, regarding the $\sigma(a, \lambda)$. An examination of these difficulties can be found in the paper by Foster [134].

In [133], Robert Shapley and Michael Hawken study the receptive profiles $\varphi(a, \lambda)$ of colour-sensitive $V1$ neurons which, in the linear case, act by convolution on the signal $I(a, \lambda)$. For a trichromatic species like the macaque or humans, they have the form

$$\varphi(a, \lambda) = \alpha_{\mathrm{L}} L(\lambda) r_{\mathrm{L}}(a) + \alpha_{\mathrm{M}} M(\lambda) r_{\mathrm{M}}(a) + \alpha_{\mathrm{S}} S(\lambda) r_{\mathrm{S}}(a) \ ,$$

where the α are coefficients, the $r(a)$ are spatial receptive profiles, and $L(\lambda)$, $M(\lambda)$, and $S(\lambda)$ are the spectral responses of the different kinds of cone. For example, the opponent cells R/G, i.e. L/M, correspond to $\alpha_{\mathrm{S}} = 0$ and $\alpha_{\mathrm{L}} = -\alpha_{\mathrm{M}}$, while the B/Y, i.e. S/(L+M) correspond to $\alpha_{\mathrm{S}} = -(\alpha_{\mathrm{L}} + \alpha_{\mathrm{M}})$. Figure 4.138, produced by Elizabeth Johnson in [135] and reproduced in [133], represents the receptive profile of a double-opponent cell R/G, where the ON and OFF regions have positive and negative heights, respectively. We see the level lines corresponding to L-cones and M-cones. Figure 4.139 shows schematically single- and double-opponent cell models with oriented or circular spatial profile. Note in particular the orientation cells $\mathrm{L} \pm /\mathrm{M} \mp$.

The difference between single- and double-opponent cells is crucial from the theoretical point of view, but also for the purposes of modelling. Indeed, single-opponent cells respond to *interior* regions of continuously varying colour domains (with no detectable spatial contrast), whereas double-opponent cells respond to the edges of such regions where the colour undergoes a *qualitative discontinuity*. This confirms neurophysiologically the phenomenological models suggested by René Thom at the end of the 1960s [118, 119], models which we have already related to Stephen Grossberg's work in [131].

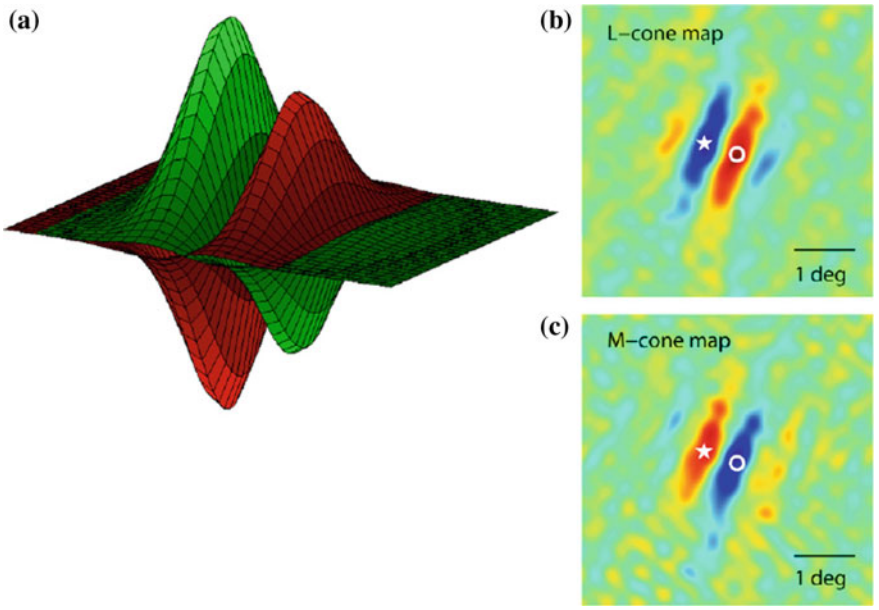


Fig. 4.138 **a** Complete receptive profile of a double-opponent cell R/G of *V* 1. This is the superposition of an L profile and an M profile. The ON and OFF regions (where an increase or a decrease in luminosity bring about a greater response) have positive and negative heights, respectively. **b** and **c** Level lines of the L-cones and the M-cones obtained by the technique of inverse correlation. The scale of growth goes from *blue* (low) to *red* (high). At the position marked with a *star*, the L map is OFF (*blue*) and the M map is ON (*red*), and at the position marked with a *circle*, the opposite. From Johnson [135] and Shapley and Hawken [133]

4.12.2.3 Double Opponency and Natural Images

As pointed out in Sect. 3.6.2 of Chap. 3, the receptive profiles (RPs) can be deduced from the statistical properties of natural images using the techniques of independent component analysis (ICA), insofar as they minimize redundancy in the neural representations of sense data. Likewise for the RPs of colour processing neurons. For example, Dharmesh Tailor, Leif Finkel, and Gershon Buchsbaum showed in [136] that independent spatiochromatic filters extracted from databases of natural colour images look much like those of double-opponent colour cells observed, for example, in the macaque *V* 1 area. For statistical correlations between colour and orientation in natural images, see also [137].

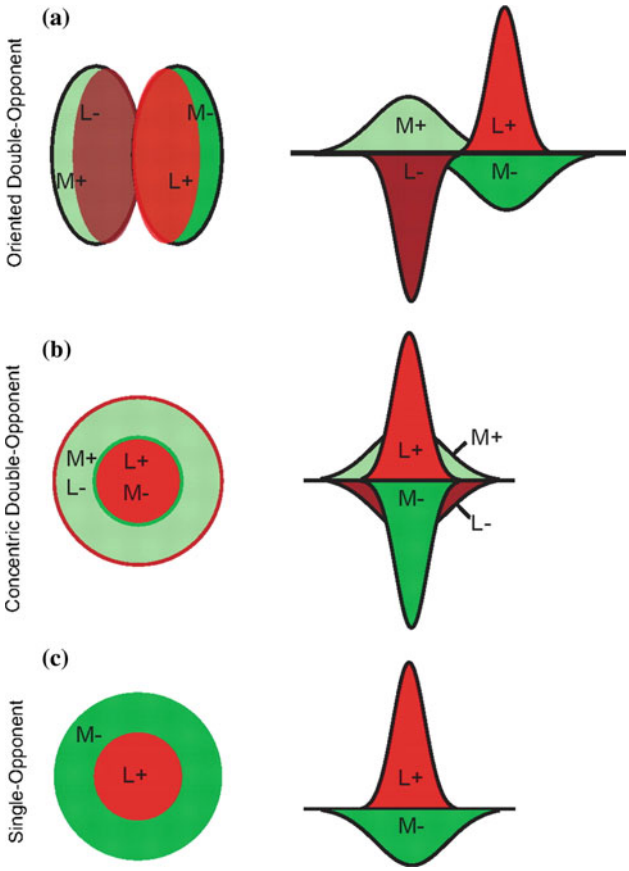


Fig. 4.139 Diagrams of single- and double-opponent L/M cells with oriented or circular spatial profile. From Johnson [135]

References

1. Lee, T.S., Mumford, D., Romero, R., Lamme, V.A.F.: The role of primary visual cortex in higher level vision. Vision Res. **38**, 2429–2454 (1998)
2. Beaudot, W.H.A., Mullen, K.T.: Orientation selectivity in luminance and color vision assessed using 2D band-pass filtered spatial noise. Vision Res. **45**, 687–695 (2005)
3. Hasboun, D.: neur@nat. http://www.chups.jussieu.fr/ext/neuranat/
4. Brain. http://thebrain.mcgill.ca
5. Visual Cortex. http://www.vision.ee.ethz.ch/en/
6. Vision Web. http://webvision.med.utah.edu
7. Ungerleider, L.G., Mishkin, M.: Two cortical visual systems. In: Ingle, D.J., Goodale, M.A., Mansfield, R.J.W. (eds). Analysis of Visual Behavior, MIT Press, pp. 549–586. Cambridge, MA (1982)
8. Wandell, B.A., Dumoulin, S.O., Brewer, A.A.: Visual field maps in human cortex. Neuron **56**, 366–383 (2007)

9. Larsson, J., Heeger, D.J.: Two retinotopic visual areas in human lateral occipital cortex. J. Neurosci. **26**(51), 13128–13142 (2006)
10. Tootell, R.B., Switkes, E., Silverman, M.S., Hamilton, S.L.: Functional anatomy of macaque striate cortex. II retinotopic organization. J. Neurosci. **8**(5), 1531–1568 (1988)
11. Dumoulin, S.O., Wandell, B.A.: Population receptive field estimates in human visual cortex. NeuroImage **39**, 647–660 (2008)
12. Brewer, A.A., Press, W.A., Logothetis, N.K., Wandell, B.A.: Visual areas in macaque cortex measured using functional magnetic resonance imaging. J. Neurosci. **22**(23), 10416–10426 (2002)
13. Hadjikhani, N., Liu, A.K., Dale, A.D., Cavanagh, P., Tootell, R.B.H.: Retinotopy and color sensitivity in human visual cortical area *V*8. Nat. Neurosci. **1**(3), 235–241 (1998)
14. Hooks, B.M., Chen, C.: Critical periods in the visual system: changing views for a model of experience-dependence plasticity. Neuron **56**, 312–326 (2007)
15. Gur, M., Snodderly, D.M.: Direction selectivity in *V*1 of alert monkeys: evidence for parallel pathways for motion processing. J. Physiol. **585**(2), 383–400 (2007)
16. Vision. Accessible at http://jeanpetitot.com/Woodruff-Pak_vision.ppt
17. Alexander, D.M., Sheridan, P., Bourke, P.D., Konstandatos, O., Wright, J.J.: Global and local symmetry of the primary visual cortex: derivation of orientation preference. http://paulbourke.net/papers/visualneuro/
18. Balasubramanian, M., Polimeni, J., Schwartz, E.L.: The *V*1-*V*2-*V*3 complex: quasi conformal dipole maps in primate striate and extra-striate cortex. Neural Networks **15**(10), 1157–1163 (2002)
19. DeAngelis, G.C., Ghose, G.M., Ohzawa, I., Freeman, R.D.: Functional micro-organization of primary visual cortex: receptive field analysis of nearby neurons. J. Neurosci. **19**(9), 4046–4064 (1999)
20. Beaudot, W.H.A., Mullen, K.T.: Orientation discrimination in human vision: psychophysics and modeling. Vision Res. **46**, 26–46 (2006)
21. Snippe, H.P., Koenderink, J.J.: Discrimination thresholds for channel-coded systems. Biol. Cybern. **66**, 543–551 (1992)
22. Ringach, D.L.: Population coding under normalization. Vision Res. **50**, 2223–2232 (2010)
23. Braitenberg, V., Braitenberg, C.: Geometry of orientation columns in the visual cortex. Biol. Cybern. **33**, 179–186 (1979)
24. Visual system. http://brain.phgy.queensu.ca/pare/assets/Higher%20Processing%20handout.pdf
25. Lund, J.S., Angelucci, A., Bressloff, P.C.: Anatomical substrates for functional columns in macaque monkey primary visual cortex. Cereb. Cortex **13**(1), 15–24 (2003)
26. Van Hooser, S.D.: Similarity and diversity in visual cortex: is there a unifying theory of cortical computation? Neuroscientist **13**(6), 639–656 (2007)
27. Hubel, D.H.: Eye, Brain and Vision. Scientific American Library, W.H. Freeman & Co, New York (1988)
28. Ben-Shahar, O., Zucker, S.: Geometrical computations explain projection patterns of long-range horizontal connections in visual cortex. Neural Comput. **16**(3), 445–476 (2004)
29. Koenderink, J.J., Van Doorn, A.J.: Representation of local geometry in the visual system. Biol. Cybern. **55**, 367–375 (1987)
30. Zhang, J., Wu, S.: Structure of visual perception. Proc. Natl. Acad. Sci. USA **87**, 7819–7823 (1990)
31. Weliky, M., Bosking, W., Fitzpatrick, D.: A systematic map of direction preference in primary visual cortex. Nature **379**, 725–728 (1996)
32. Bonhöffer, T., Grinvald, A.: Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns. Nature **353**, 429–431 (1991)
33. Ohki, K., Reid, R.C.: Specificity and randomness in the visual cortex. Curr. Opin. Neurobiol. **17**(4), 401–407 (2007)

34. Grinvald, A., Shoham, D., Shmuel, A., Glaser, D., Vanzetta, I., Shtoyerman, E., Slovin, H., Sterkin, A., Wijnbergen, C., Hildesheim, R., Arieli, A.: In-vivo optical imaging of cortical architecture and dynamics, Technical Report GC-AG/99-6, The Weizmann Institute of Science (2001)
35. Blasdel, G.G., Salama, G.: Voltage sensitive dyes reveal a modular organization in monkey striate cortex. Nature **321**, 579–585 (1986)
36. Crair, M.C., Ruthazer, E.S., Gillespie, D.C., Stryker, M.P.: Ocular dominance peaks at pinwheel center singularities of the orientation map in cat visual cortex. J. Neurophysiol. **77**, 3381–3385 (1997)
37. Bosking, W.H., Zhang, Y., Schofield, B., Fitzpatrick, D.: Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. J. Neurosci. **17**(6), 2112–2127 (1997)
38. Fitzpatrick, D.: The functional organization of local circuits in visual cortex: insights from the study of tree shrew striate cortex. Cereb. Cortex **6**, 329–341 (1996)
39. Lund, J.S., Fitzpatrick, D., Humphrey, A.L.: The striate visual cortex of the tree shrew. In: Jones, E.G., Peters, A. (eds). Cerebral Cortex, pp. 157–205. Plenum, New York (1985)
40. Frégnac, Y., Baudot, P., Chavane, F., Lorenceau, J., Marre, O., Monier, C., Pananceau, M., Carelli, P., Sadoc, G.: Multiscale functional imaging in $V1$ and cortical correlates of apparent motion. In: Masson, G.S., Ilg, U.J. (eds). Dynamics of Visual Motion Processing. Neuronal, Behavioral, and Computational Approaches, pp. 73–94. Springer, Berlin, New York (2010)
41. Eysel, U.: Turning a corner in vision research. Nature **399**, 641–644 (1999)
42. Bosking, W.H., Crowley, J.C., Fitzpatrick, D.: Spatial coding of position and orientation in primary visual cortex. Nat. Neurosci. **5**(9), 874–882 (2002)
43. McLoughlin, N., Schiessl, I.: Orientation selectivity in the common marmoset (Callithrix jacchus): The periodicity of orientation columns in $V1$ and $V2$. NeuroImage **31**, 76–85 (2006)
44. Niebur, E., Wörgötter, F.: Design principles of columnar organization in visual cortex. Neural Comput. **6**, 602–614 (1994)
45. Xu, X., Bosking, W.H., White, L.E., Fitzpatrick, D., Casagrande, V.A.: Functional organization of visual cortex in the prosimian bush baby revealed by optical imaging of intrinsic signals. J. Neurophysiol. **94**, 2748–2762 (2005)
46. Yu, H., Farley, B.J., Jin, D.Z., Sur, M.: The coordinated mapping of visual space and response features in visual cortex. Neuron **47**, 267–280 (2005)
47. Basole, A., White, L.E., Fitzpatrick, D.: Mapping multiple features in the population response of visual cortex. Lett. Nat. **423**(26), 986–990 (2003)
48. Baxter, W.T., Dow, B.M.: Horizontal organization of orientation-sensitive cells in primate visual cortex. Biol. Cybern. **61**, 171–182 (1989)
49. Goodhill, G.J.: Contributions of theoretical modeling to the understanding of neural map development. Neuron **56**, 301–311 (2007)
50. Nöckel, J.: Gradient field plots in Mathematica. http://pages.uoregon.edu/noeckel/computernotes/Mathematica/fieldPlots.html
51. Chapman, B., Stryker, M.P., Bonhöffer, T.: Development of orientation preference maps in ferret primary visual cortex. J. Neurosci. **16**(20), 6443–6453 (1996)
52. Koulakov, A.A., Chklovskii, D.B.: Orientation preference patterns in mammalian visual cortex: a wire length minimization approach. Neuron **29**, 519–527 (2001)
53. Azencott, R.: Simulated Annealing, Séminaire Bourbaki 697, Astérisque 161–162, Paris (1988)
54. Gregor, K., Szlam, A., LeCun, Y.: Structured sparse coding via lateral inhibition. http://books.nips.cc/nips24.html
55. Hyvärinen, A., Hoyer, P.O.: A two-layer sparse coding model learns simple and complex cell receptive fields and topography from natural images. Vision Res. **41**(18), 2413–2423 (2001)
56. Shmuel, A., Grinvald, A.: Coexistence of linear zones and pinwheels within orientation maps in cat visual cortex. Proc. Natl. Acad. Sci. **97**(10), 5568–5573 (2000)
57. Ohki, K., Matsuda, Y., Ajima, A., Kim, D.-S., Tanaka, S.: Arrangement of orientation pinwheel centers around area 17/18 transition zone in cat visual cortex. Cereb. Cortex **10**, 593–601 (2000)

58. Romagnoni, A., Ribot, J., Bennequin, D., Touboul, J.: Parsimony, exhaustivity and balanced detection in neocortex. PLoS Comput. Biol. 11(11), e1004623
59. Petitot, J.: Eléments de théorie des singularités. http://jean.petitot.pagesperso-orange.fr/ArticlesPDF_new/Petitot_Sing.pdf (1982)
60. Berry, M.V.: Much ado about nothing: optical dislocation lines (phase singularities, zeros, vortices ...). In: Soskin, M.S. (ed.) Proceedings of International Conference on Singular Optics, vol. 3487, pp. 1–5. SPIE (Society of Photographic Instrumentation Engineers) (1998)
61. Berry, M.V., Dennis, M.R.: Reconnections of wave vortex lines. Eur. J. Physics **33**, 723–731 (2012)
62. Bennequin, D.: Remarks on invariance in the primary visual systems of mammals. In: Citti, G., Sarti, A. (eds). Neuromathematics of Vision, pp. 243–333. Springer, Berlin (2014)
63. Afgoustidis, A.: Monochromaticity of orientation maps in $V1$ implies minimum variance for hypercolumn size. J. Math. Neurosci. **5**, 10 (2015)
64. Afgoustidis, A.: Représentations de groupes de Lie et fonctionnement géométrique du cerveau, *Thèse* (D. Bennequin dir.), Université Paris VII (2016)
65. Citti, G., Sarti, A.: From functional architectures to percepts: a neuromathematical approach. In: Citti, G., Sarti, A. (eds.) Neuromathematics of Vision, pp. 131–171. Springer (2014)
66. Petitot, J.: Introduction aux phénomènes critiques. http://jean.petitot.pagesperso-orange.fr/ArticlesPDF_new/Petitot_CritPh.pdf (1982)
67. Petitot, J.: Landmarks for neurogeometry. In: Citti, G., Sarti, A. (eds). Neuromathematics of Vision, pp. 1–85. Springer, Berlin (2014)
68. Dennis, M.R.: *Topological Singularities in Wave Fields* , PhD Thesis, H.H. Wills Laboratory, University of Bristol (2001)
69. Berry, M.V.: Optical currents. J. Opt. A: Pure Appl. Opt. **11**(9), 094001 (2009)
70. Berry, M.V., Dennis, M.R.: Phase singularities in isotropic random waves. Proc. Roy. Soc. London A **456**, 2059–2079 (2000)
71. Berry, M.V., Dennis, M.R.: Topological events on wave dislocation lines: birth and death of loops, and reconnection. J. Physics A: Math. Theor. **40**, 65–74 (2007)
72. Wolf, F., Geisel, T.: Spontaneous pinwheel annihilation during visual development. Nature **395**, 73–78 (1998)
73. Wolf, F., Geisel, T.: Universality in visual cortical pattern formation. In: Petitot, J., Lorenceau, J. (eds). Neurogeometry and Visual Perception, pp. 253–264. J. Physiol. Paris **97**, 2–3 (2003)
74. Barbieri, D., Citti, G., Sanguinetti, G., Sarti, A.: An uncertainty principle underlying the functional architecture of V1. J. Physiol. Paris **106**(5–6), 183–193 (2012)
75. Abrahamsen, P.: A Review of Gaussian Random Fields and Correlation Functions. http://publications.nr.no/917_Rapport.pdf (1997)
76. Azaïs, J.-M., León, J.R., Wschebor, M.: Rice formulae and Gaussian waves. Bernoulli **17**(1), 170–193 (2011)
77. Adler, R.J., Taylor, J.E.: Random Fields and Geometry. Springer, Berlin (2007)
78. Swindale, N.V.: The development of topography in the visual cortex: a review of models. Netw. Comput. Neural Syst. **7**(2), 161–247 (1996)
79. Bednar, J.A.: Hebbian learning of the statistical and geometrical structure of visual inputs. In: Citti, G., Sarti, A. (eds). Neuromathematics of Vision, pp. 335–366. Springer, Berlin (2014)
80. Miikkulainen, R., Bednar, J.A., Choe, Y., Sirosh, J.: Computational Maps in the Visual Cortex. Springer, New York (2005)
81. Wolf, F.: Symmetry, multistability, and long-range interaction in brain development. Phys. Rev. Lett. **95**, 208701, 1–4 (2005). Erratum: 103, 20 (2009)
82. Kaschube, M., Schnabel, M., Wolf, F.: Self-organization and the selection of pinwheel density in visual cortical development. arXiv:0801.3651v1 (2008)
83. Lee, H.Y., Yahyanejad, M., Kardar, M.: Symmetry considerations and development of pinwheels in visual maps, arXiv:cond-mat/0312539v1
84. Swift, J.B., Hohenberg, P.C.: Hydrodynamic fluctuations at the convective instability. Phys. Rev. A **15**, 319–328 (1977)

85. Schnabel, M., Kaschube, M., Wolf, F.: Pinwheel stability, pattern selection and the geometry of visual space. arXiv:0801.3832v2 (2008)
86. Maldonado, P.E., Gödecke, I., Gray, C.M., Bonhöffer, T.: Orientation selectivity in pinwheel centers in cat striate cortex. Science **276**, 1551–1555 (1997)
87. Polimeni, J.R., Granquist-Fraser, D., Wood, R.J., Schwartz, E.L.: Physical limits to spatial resolution of optical recording: clarifying the spatial structure of cortical hypercolumns. Proc. Natl. Acad. Sci. **102**(11), 4158–4163 (2005)
88. McLaughlin, D., Shapley, R., Shelley, M., Wielaard, D.J.: A neuronal network model of macaque primary visual cortex ($V1$): orientation selectivity and dynamics in the input layer 4Cα. Proc. Natl. Acad. Sci. **97**(14), 8087–8092 (2000)
89. Shelley, M., McLaughlin, D.: Coarse-grained reduction and analysis of a network model of cortical response: I. drifting grating stimuli. J. Comput. Neurosci. **12**, 97–122 (2002)
90. Schummers, J., Mariño, J., Sur, M.: Synaptic integration by $V1$ neurons depends on location within the orientation map. Neuron **36**, 969–978 (2002)
91. Mariño, J., Schummers, J., Lyon, D.C., Schwabe, L., Beck, O., Wiesing, P., Obermayer, K., Sur, M.: Invariant computations in local cortical networks with balanced excitation and inhibition. Nat. Neurosci. **8**(2), 194–201 (2005)
92. Ohki, K., Chung, S., Kara, P., Hübener, M., Bonhöffer, T., Reid, R.C.: Highly ordered arrangement of single neurons in orientation pinwheels. Nature **442**, 925–928 (2006)
93. Petitot, J.: Rappels sur l'Analyse non standard. In: Salanskis, J.-M., Barreau, H. (eds). La Mathématique non standard, pp. 187–209. Éditions du CNRS, Paris (1989)
94. Deligne, P., Malgrange, B., Ramis, J.-P.: Singularités irrégulières. Correspondance et documents, Société Mathématique de France (2007)
95. Ramis, J.-P.: Les derniers travaux de Jean Martinet. Annales de l'Institut Fourier **42**(1–2), 15–47 (1992)
96. Martinet, J.: Introduction à la théorie de Cauchy sauvage (unpublished). See Ramis above
97. Das, A., Gilbert, C.D.: Distortions of visuotopic map match orientation singularities in primary visual cortex. Nature **387**, 594–598 (1997)
98. Ohki, K., Chung, S., Ch'ng, Y.H., Kara, P., Reid, R.C.: Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. Nature **433**, 597–603 (2005)
99. Bressloff, P., Cowan, J.: The functional geometry of local and horizontal connections in a model of $V1$. In: Petitot, J., Lorenceau, J. (eds). Neurogeometry and Visual Perception, pp. 221–236. J. Physiol. Paris **97**, 2–3 (2003)
100. Hübener, M., Shoham, D., Grinvald, A., Bonhöffer, T.: Spatial relationships among three columnar systems in cat area 17. J. Neurosci. **17**, 9270–9284 (1997)
101. Issa, N.P., Trepel, C., Stryker, M.P.: Spatial frequency maps in cat visual cortex. J. Neurosci. **22**, 8504–8514 (2000)
102. Sirovich, L., Uglesich, R.: The organization of orientation and spatial frequency in primary visual cortex. Proc. Natl. Acad. Sci. **101**(48), 16941–16946 (2004)
103. Born, R.T., Tootell, R.B.: Spatial frequency tuning of single units in macaque supragranular striate cortex. Proc. Natl. Acad. Sci. **88**, 7066–7070 (1990)
104. Issa, N.P., Rosenberg, A., Husson, T.R.: Models and measurements of functional maps in $V1$. J. Neurophysiol. **99**, 2745–2754 (2008)
105. Zhu, W., Xing, D., Shelley, S., Shapley, R.: Correlation between spatial frequency and orientation selectivity in $V1$ cortex: implications of a network model. Vision Res. **50**, 2261–2273 (2010)
106. Ribot, J., Romagnoni, A., Milleret, C., Bennequin, D., Touboul, J.: Pinwheel-dipole configuration in cat early visual cortex. NeuroImage **128**, 63–73 (2016)
107. Ribot, J., Aushana, Y., Bui-Quoc, E., Milleret, C.: Organization and origin of spatial frequency maps in cat visual cortex. J. Neurosci. **33**, 13326–13343 (2013)
108. Tani, T., Ribot, J., O'Hashi, K., Tanaka, S.: Parallel development of orientation maps and spatial frequency selectivity in cat visual cortex. Eur. J. Neurosci. **35**, 44–55 (2012)
109. Xu, X., Bosking, W.H., Sáry, G., Stefansic, J., Shima, D., Casagrande, V.: Functional organization of visual cortex in the owl monkey. J. Neurosci. **24**(28), 6237–6247 (2004)

110. Liu, G.B., Pettigrew, J.D.: Orientation mosaic in barn owl's visual Wulst revealed by optical imaging: comparison with cat and monkey striate and extra-striate areas. Brain Res. **961**, 153–158 (2003)
111. Obermayer, K., Blasdel, G.G.: Geometry of orientation and ocular dominance columns in monkey striate cortex. J. Neurosci. **13**, 4114–4129 (1993)
112. Carreira-Perpiñán, M.A., Goodhill, G.J.: Influence of lateral connections on the structure of cortical maps. J. Neurophysiol. **92**, 2947–2959 (2004)
113. Bednar, J.A., Miikkulainen, R.: Joint maps for orientation, eye, and direction preference in a self-organizing model of $V1$. Neurocomputing **69**, 1272–1276 (2006)
114. Miikkulainen, R., Bednar, J.A., Choe, Y., Sirosh, J.: A Self-organizing neural network model of the primary visual cortex. In: Proceedings of the Fifth International Conference on Neural Information Processing (ICONIP'98) (1998)
115. Swindale, N.V.: How many maps are there in visual cortex? Cereb. Cortex **7**, 633–643 (2000)
116. Swindale, N.V.: How different feature spaces may be represented in cortical maps? Network **15**, 217–242 (2004)
117. Kara, P., Boyd, J.D.: A micro-architecture for binocular disparity and ocular dominance in visual cortex. Nat. Lett. **458**, 627–631 (2009)
118. Thom, R.: Stabilité structurelle et Morphogenèse. Benjamin, New York, Ediscience, Paris (1972)
119. Thom, R.: Modèles mathématiques de la morphogenèse. Bourgois, Paris (1980)
120. Blake, R., Wilson, H.: Binocular vision. Vision Res. **51**, 754–770 (2011)
121. Kim, Y.-J., Grabowecky, M., Suzuki, S.: Stochastic resonance in binocular rivalry. Vision Res. **46**, 392–406 (2006)
122. De Jong, T.M.: The Dynamics of Visual Rivalry. University of Utrecht, Thesis (2008)
123. Leopold, D.A., Logothetis, N.K.: Multistable phenomena: changing views in perception. Trends Cogn. Sci. **3**(7), 254–264 (1999)
124. Lu, H.D., Roe, A.W.: Functional organization of color domains in $V1$ and $V2$ of macaque monkey revealed by optical imaging. Cereb. Cortex **18**(3), 516–533 (2008)
125. Rochefort, N.L.: Functional Specificity of Callosal Connections in the Cat Visual Cortex. PhD Thesis, Ruhr-Universität, Bochum and Université de Paris VI, Paris (2007)
126. Bosking, W.H., Kretz, R., Pucak, M.L., Fitzpatrick, D.: Functional specificity of callosal connections in tree shrew striate cortex. J. Neurosci. **20**(6), 2346–2359 (2000)
127. Olavarria, J.F.: Callosal connections correlate preferentially with ipsilateral cortical domains in cat areas 17 and 18, and with contralateral domains in the 17/18 transition zone. J. Comp. Neurol. **433**(4), 441–457 (2001)
128. Rochefort, N.L., Buzás, P., Kisvárday, Z.F., Eysel, U.T., Milleret, C.: Layout of transcallosal activity in cat visual cortex revealed by optical imaging. NeuroImage **36**(3), 1804–1821 (2007)
129. Rochefort, N.L., Buzás, P., Quenech'du, N., Koza, A., Eysel, U.T., Milleret, C., Kisvárday, Z.F.: Functional selectivity of interhemispheric connections in cat visual cortex. Cereb. Cortex **19**, 2451–2465 (2009)
130. Petitot, J.: Physique du Sens. Éditions du CNRS, Paris (1992)
131. Grossberg, S.: Neural Networks and Natural Intelligence. MIT Press, Cambridge, MA (1988)
132. Tani, T., Yokoi, I., Ito, M., Tanaka, S., Komatsu, H.: Functional organization of the cat visual cortex in relation to the representation of a uniform surface. J. Neurophysiol. **89**, 1112–1125 (2003)
133. Shapley, R., Hawken, M.J.: Color in the cortex: single- and double-opponents cells. Vision Res. **51**, 701–717 (2011)
134. Foster, D.H.: Color constancy. Vision Res. **51**, 674–700 (2011)
135. Johnson, E.N., Hawken, M.J., Shapley, R.: The orientation selectivity of color-responsive neurons in macaque $V1$. J. Neurosci. **28**, 8096–8106 (2008)
136. Tailor, D.R., Finkel, L.H., Buchsbaum, G.: Color-opponent receptive fields derived from independent component analysis of natural images. Vision Res. **40**, 2671–2676 (2000)
137. Cecchi, G.A., Rao, A.R., Xiao, Y., Kaplan, E.: Statistics of natural scenes and cortical color processing. J. Vision **10**(11), 1–13 (2010)