Séminaire Cerveau et Cognition
3 décembre 2009

Neurogéometrie de la vision

Jean Petitot
CREA
petitot@poly.polytechnique.fr

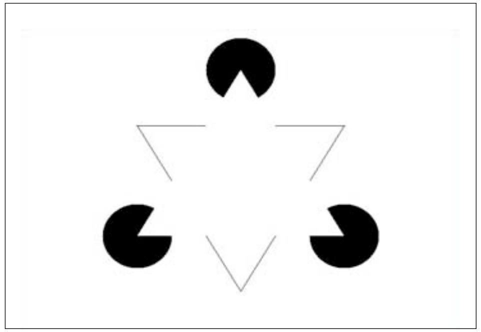
Introduction

- Neurogeometry concerns the neural implementation of the geometric structures of visual perception.
- They are very different from the Euclidean 3D structure of the objective external space which is the ouput of very sophisticated cognitive constructions.

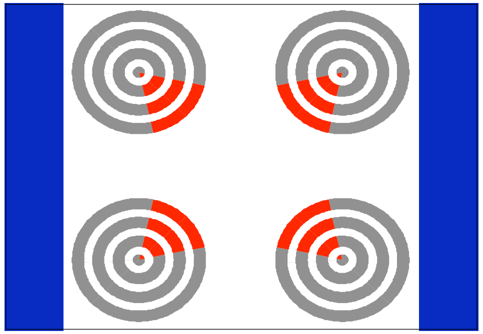
- Many non trivial mathematical structures have been introduced recently to explain this neural implementation of natural low level vision.
- I will focus on two of them:
  - Receptive fields of neural cells and wavelet analysis.
  - Differential (contact, symplectic, and sub-Riemannian) geometry and the functional architecture of area V1.

An example : Kanizsa illusory contours

- A typical example of the problems of neurogeometry is given by well known Gestalt phenomena such as Kanizsa illusory contours.
- The visual system (V1 with some feedback from V2) constructs very long range and sharp virtual contours.

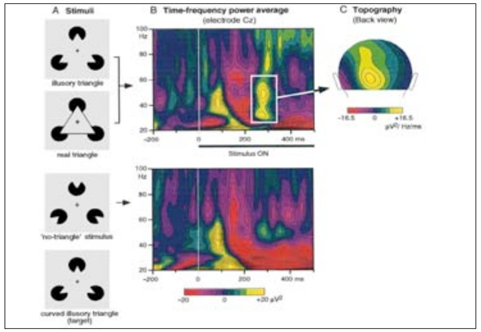


- They can even be curved.
- With the neon effect, virtual contours are boundaries for the diffusion of color inside them.

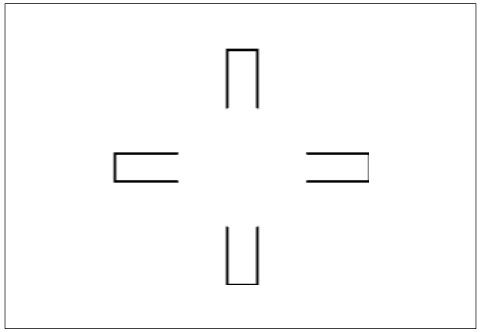


- Kanizsa subjective contours manifest a deep neurophysiological phenomenon.
- Here is a result of Catherine Tallon-Baudry in « Oscillatory gamma activity in humans and its role in object representation » (*Trends in Cognitive Science*, 3, 4, 1999).
- Subjects are presented with coherent stimuli (illusory and real triangles) « leading to a coherent percept through a bottom-up feature binding process ».

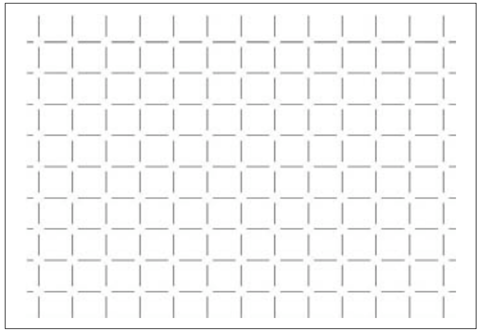
- « Time–frequency power of the EEG at electrode Cz (overall average of 8 subjects), in response to the illusory triangle (top) and to the no-triangle stimulus (bottom ».
- « Two successive bursts of oscillatory activities were observed.
  - A first burst at about 100 ms and 40 Hz. It showed no difference between stimulus types.
  - A second burst around 280 ms and 30-60 Hz. It is most prominent in response to coherent stimuli. »



- Many phenomena are striking. E.g. the change of strategy between a “diffusion of curvature” strategy and a “piecewise linear” strategy where the whole curvature is concentrated in a singular point.
- Bistability: the illusory contour is either a circle or a square.



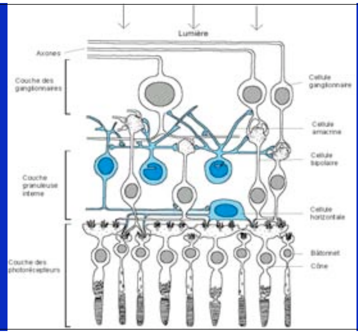
- The example of Ehrenstein illusion:

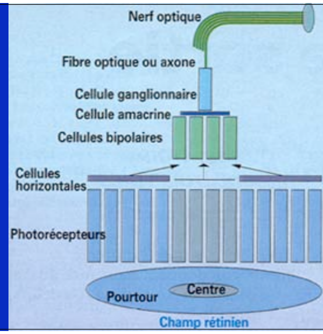


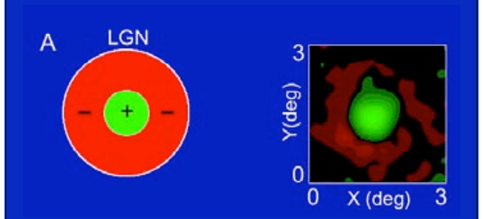
- The explanation of such phenomena is difficult because they are long range w.r.t. the size of individual neurons.
- They result from a <u>local to global integration</u> processing.
- We have therefore to understand
  - 1. the local detection of local features,
  - 2. Their integration into global morphologies.

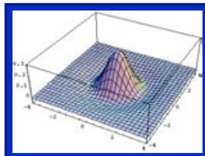
Retina and wavelets

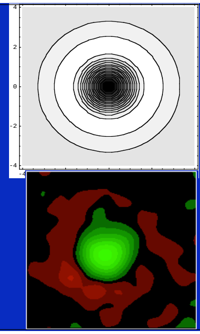
- Receptive fields (in the narrow sense of « minimal discharge field », see Y. Frégnac).
- Receptive profiles (linear approximation).











Good approximation by a Laplacian of Gaussian ΔG

- There is a lot of technical discussions concerning the exact form of RP.
- Richard Young. « The Gaussian Derivative model for spatio-temporal vision », *Spatial Vision*, 14, 3-4, 2001, 261-319.
  - « The initial stage of processing of receptive fields in the visual cortex approximates a 'derivative analyzer' that is capable of estimating the local spatial and temporal directional derivatives of the intensity profile in the visual environment. »
- How ?

- How do the RPs operate on the visual signal (linear approximation)?
- Let $I(x,y)$ be the visual signal (x,y are visual coordinates on the retina).
- Let $\varphi(x-x_0,y-y_0)$ be the RP of a neuron N whose receptive field is defined on a domain D of the retina centered on (x_0,y_0).

- N acts on the signal I as a filter :

$$I_\varphi(x_0,y_0)=\int_D I(x',y')\varphi(x'-x_0,y'-y_0)dx'dy'$$

- A field of such neurons act therefore by convolution on the signal

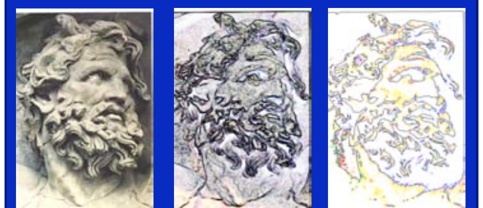
$$I_\varphi(x,y)=\int_D I(x',y')\varphi(x'-x,y'-y)dx'dy'=(I*\varphi)(x,y)$$

- But from the classical formula

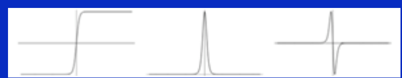
$$I*DG = D(I*G),$$

for G a Gaussian and D a differential operator, the convolution of the signal I with a DG-shaped RF amounts to apply D to the smoothing $I*G$ of the signal I at the scale defined by G.

- Hence a multiscale differential geometry which is a wavelet analysis.



- Zero-crossing (D. Marr).
- f : discontinuité
- f' : delta δ
- f'' : δ'



Wavelet analysis

- Signals $f\in L^2(\mathbb{R})$
- Fourier transform (analysis).

$$\widehat{f}(\omega)=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(x)e^{-i\omega x}dx=\langle f(x)|\,e^{i\omega x}\rangle$$

- Inverse transform (synthesis).

$$f(x)=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \widehat{f}(\omega)e^{i\omega x}d\omega=\left\langle \widehat{f}(\omega)\right|e^{-i\omega x}\rangle$$

- Isometry.
- Geometrical information is delocalized.

- Gabor transform (analysis).

$$Gf(\omega,u) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) g(x-u) e^{-i\omega x} dx = \langle f(x) \,|g_{\omega,u}(x)\rangle$$

- Inverse transform (synthesis).

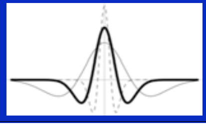
$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}^2} Gf(\omega,u) g(x-u) e^{i\omega x} d\omega du$$

- Isometry.
- Geometrical information is localized, but only at one scale.

- Multiscale wavelet transform (analysis).
- Mother wavelet and scaling :

$$\psi_s(x) = \frac{1}{\sqrt{s}} \psi\left(\frac{x}{s}\right)$$

- Typical example : ΔG



- Direct wavelet transform :

$$Wf(s,u) = \int_{\mathbb{R}} f(x) \psi_s(x-u) dx = \langle f(x) \,|\psi_{s,u}(x)\rangle$$

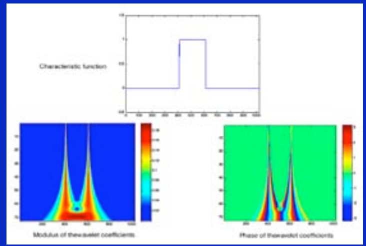
- $$(C) : \widehat{\psi}(0) = 0 \quad \text{et} \quad C_\psi = \int_{\mathbb{R}+} \frac{\left|\widehat{\psi}(\omega)\right|^2}{\omega} d\omega < \infty$$

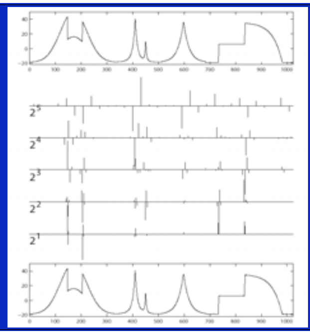
$$\psi(x) = (1-x^2)\, e^{-\frac{x^2}{2}} \qquad \widehat{\psi} = \omega^2 e^{-\frac{\omega^2}{2}}$$

$$C_\psi = \int_{\mathbb{R}+} \frac{|\widehat{\psi}(\omega)|^2}{\omega} d\omega \qquad C_\psi = \int_{\mathbb{R}+} \omega^3 e^{-\omega^2} d\omega = \frac{1}{2}$$

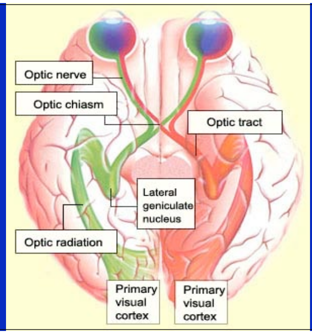
- $$f(x) = \frac{1}{C_\psi} \int_{\mathbb{R}+} \int_{\mathbb{R}} Wf(s,u) \psi_s(x-u) ds du$$

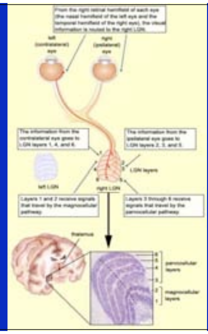
- Extraction of singularities :

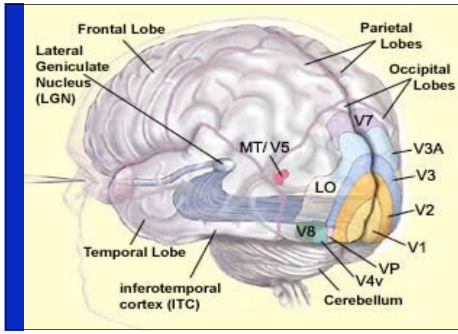


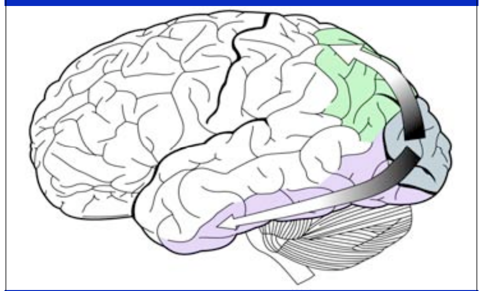


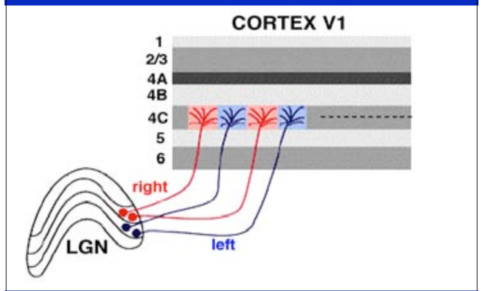
The primary visual cortex: area V1







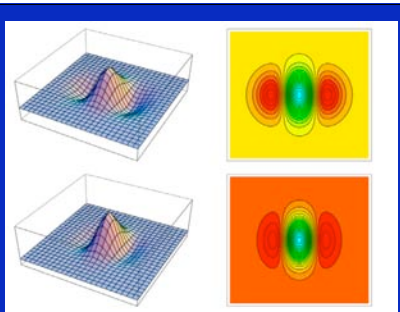




- Level curves of the receptive profiles of some simple cells of V1 can be modeled
  - by second order derivatives of Gaussians,
  - by Gabor wavelets

$$\exp(i2x)\exp\left(-\left(x^2+y^2\right)\right)$$

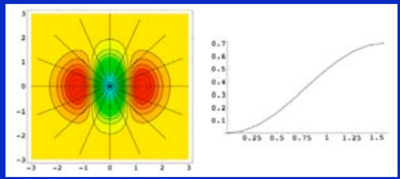
(real part).

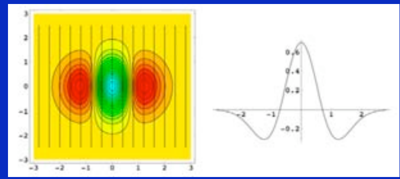


- The interest of Gabor wavelets is that they minimize uncertainty relations and are well adapted to harmonic analysis.
- The interest of Gaussian derivatives is that they explain how the brain can do differential geometry in a scale-space.

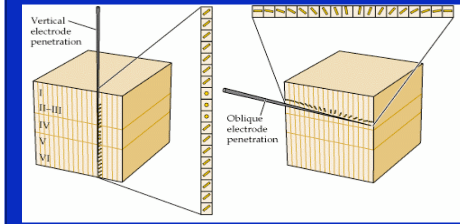
- We have seen how the RPs act upon the transduced optical signal $I(x,y)$.

$$I_\varphi(x,y)=\int_D I(x',y')\varphi(x'-x,y'-y)dx'dy'=(I*\varphi)(x,y)$$

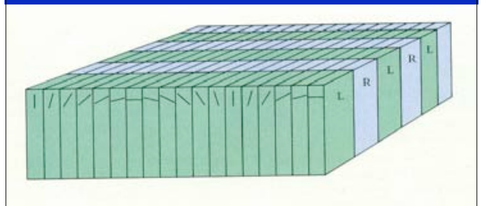




The functional architecture of area V1



- Hypercolumns (Hubel and Wiesel).



Engrafting variables : the fibration model

- The simple cells of V1 detect a preferential orientation (static or dynamic : moving gratings).
- They measure, at a certain scale, pairs (*a*, *p*) of a spatial (retinal) position *a* and of a local orientation *p* at *a*.
- Pairs (*a*, *p*) are contact elements.

- The hypercolumns associate retinotopically to each position *a* of the retina *R* a full exemplar P_a of the space *P* of orientations *p* at *a*.

- This functional architecture implements what is called in differential geometry the fibration $\pi : R \times P \rightarrow R$ with base *R*, fiber *P*, and total space $V = R \times P$.

- Fibration formalizes Hubel 's concept of "engrafting" "secundary" variables (orientation, ocular dominance, color, direction of movement, etc.) on the basic retinal variables (*x*,*y*) :
  - « 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 (…), and on this map it engrafts other variables, such as orientation and eye preference, by finer subdivisions. » (Hubel 1988, p. 131)

- How such cells with a prefered orientation can perform global tasks such as *contour integration in V1* ?

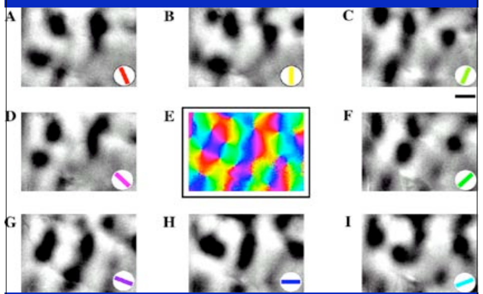
Pinwheels

- The fibration $\pi : R \times P \rightarrow R$ is of dimension 3 but is implemented in neural layers W of dimension 2.

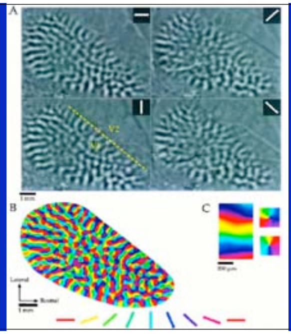
- Recent experiments have shown that the hypercolumns are geometrically organized in *pinwheels*.
- The cortical layer is reticulated by a network of singular points which are the centers of the pinwheels.
- Locally, around these singular points all the orientations are represented by the rays of a "wheel" and the local wheels are glued together in a global structure.

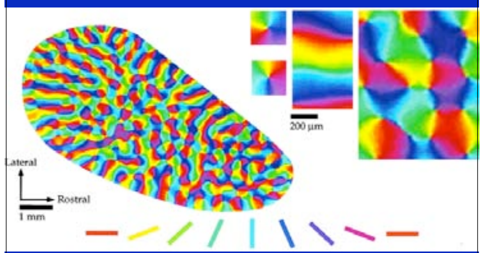
- The method (Bonhöffer & Grinvald, ~ 1990) of *in vivo optical imaging* based on activity-dependent intrinsic signals allows to acquire images of the activity of the superficial cortical layers.
- Gratings with high contrast are presented many times (20-80) with e.g. a width of 6.25° for the dark strips and of 1.25° for the light ones, a velocity of 22.5°/s, different (8) orientations.
- A window is opened above V1 and the cortex is illuminated with orange light.

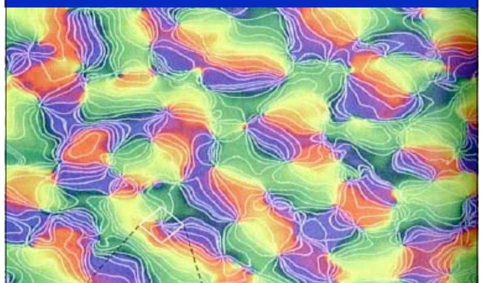
- One does the summation of the images of V1 's activity for the different gratings and constructs differential maps (differences between orthogonal gratings).
- The low frequency noise is eliminated.
- The maps are normalized (by dividing the deviation relative to the mean value at each pixel by the global mean deviation).



- In the following picture the orientations are coded by colors and iso-orientation lines are therefore coded by monocolor lines.
- William Bosking, Ying Zhang, Brett Schofield, David Fitzpatrick (Dpt of Neurobiology, Duke) 1997, « Orientation Selectivity and the Arrangement of Horizontal Connections in Tree Shrew Striate Cortex », *J. of Neuroscience*, 17, 6, 2112-2127.

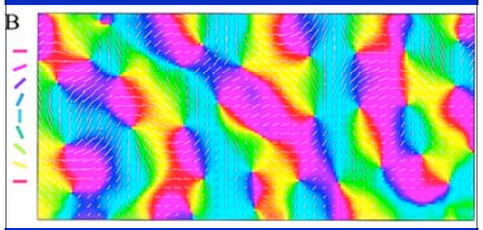


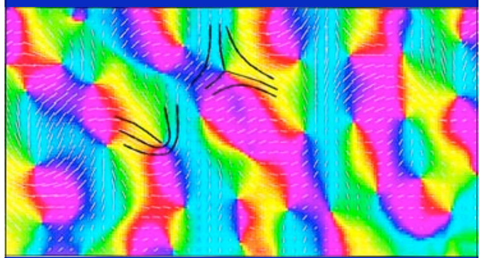




- There are 3 classes of points :
  - regular points where the orientation field is locally trivial;
  - singular points at the center of the pinwheels;
  - saddle-points localized near the centers of the cells of the network.
- Two adjacent singular points are of opposed chirality (CW and CCW).
- It is like a field in *W* generated by topological charges with « field lines » connecting charges of opposite sign.

- In the following picture due to Shmuel (cat's area 17), the orientations are coded by colors but are also represented by white segments.
- We observe very well the two types of generic singularities of 1D foliations in the plane.



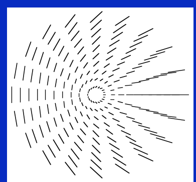


- They arise from the fact that, in general, the direction θ in V1 of a ray of a pinwheel is not the orientation p_θ associated to it in the visual field.
- When the ray spins around the singular point with an angle φ, the associated orientation rotates with an angle $\varphi/2$. Two diametrally opposed rays correspond to orthogonal orientations.
- There are two cases.

- If the orientation p_θ associated with the ray of angle θ is $p_\theta = \alpha + \theta/2$ (with $p_0 = \alpha$), the two orientations will be the same for

  $$p_\theta = \alpha + \theta/2 = \theta$$

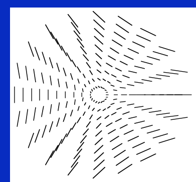
  that is for $\theta = 2\alpha$.
- As α is defined modulo π, there is only one solution : end point.



- If the orientation p_θ associated with the ray of angle θ is $p_\theta = \alpha - \theta/2$, the two orientations will be the same for

  $$p_\theta = \alpha - \theta/2 = \theta$$

  that is for $\theta = 2\alpha/3$.
- As α is defined modulo π, there are three solutions : triple point.

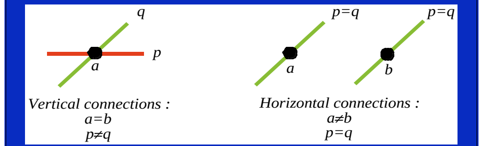


The horizontal structure

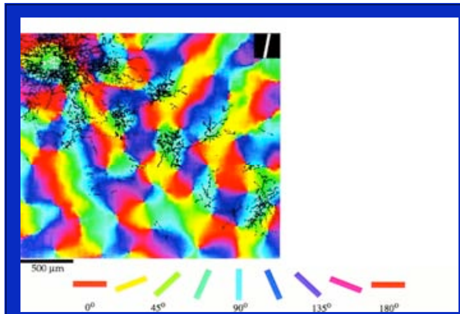
- Even if it is quite rich, such a "vertical" retinotopic structure is not sufficient.
- To implement a global coherence, the visual system must be able to *compare* two retinotopically neighboring fibers P_a et P_b over two neighboring points *a* and *b*.
- This is a problem of parallel transport. It has been solved at the empirical level by the discovery of "horizontal" cortico-cortical connections.

- Cortico-cortical connections are slow (≈ 0.2m/s) and weak.
- They connect neurons of approximately the *same* orientation in neighboring hypercolumns.
- This means that the system is able to know, for *b* near *a*, if the orientation *q* at *b* is the same as the orientation *p* at *a*.

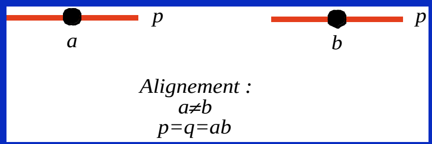
- The retino-geniculo-cortical "vertical" connections give an *internal* meaning for the relations between (*a*,*p*) and (*a*,*q*) (*different* orientations *p* and *q* at the *same* point *a*) .
- The "horizontal" cortico-cortical connec-tions give an *internal* meaning for the relations between (*a*,*p*) and (*b*,*p*) (*same* orientation *p* at *different* points *a* and *b*).



- The next slide shows how biocytin injected locally in a zone of specific orientation (green-blue) diffuses via horizontal cortico-cortical connections. The key fact is the following :
  - the *short range* diffusion is isotropic, but
  - the *long range* diffusion is on the contrary highly *anisotropic* and restricted to zones of the *same* orientation (the same color) as the initial one.



- Moreover cortico-cortical connections connect neurons coding pairs (*a*,*p*) and (*b*,*p*) *such that p is the orientation of the axis ab* (William Bosking).
  - « The system of long-range horizontal connections can be summarized as preferentially linking neurons with co-oriented, co-axially aligned receptive fields ».



- These results mean essentially that what geometers call the *contact structure* of the fibration

  $$\pi : R \times P \to R$$

  is *neurally implemented*.

The contact structure of V1

- We work in the fibration $\pi : V = R \times P \to R$ with base space R and fiber P = set of orientations p.
- Over every point $a = (x, y)$ of R, the fiber is the set $P_a = P$ of the orientations p at a.
- A local coordinate system for V is therefore given by triplets (x, y, p).

- The fibration π is an idealized model of the functional architecture of V1.
- Mathematically, it can be interpreted as the fibration $R \times \mathbf{P}^1$ ($\mathbf{P}^1$ = projective line), or as the fibration $R \times \mathbf{S}^1$ ($\mathbf{S}^1$ = unit circle), or as the space of *1-jets* of curves C in R.

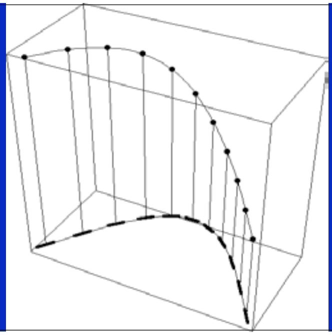
- If C is curve in R (a contour), it can be *lifted* to V. The lifting Γ is the map

$$j : C \to V = R \times P$$

wich associates to every point a of C the pair

(a, p_a) where p_a is the *tangent* of C at a.

- Γ represents C as the *enveloppe* of its tangents.



- If $a(s) = (x(s), y(s))$ is a parametrization of C, we have

$$p_a = y'(s) \,/\, x'(s) = dy \,/\, dx$$

and therefore

$$\Gamma = (a(s), p(s))$$

$$= (x(s), y(s), y'(s) \,/\, x'(s)) .$$

- If we can choose $s = x$, in terms of visual coordinates x and y, the equation of Γ writes

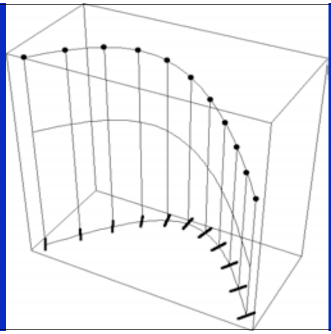
$$(x, y, p) = (x, y, y')。$$

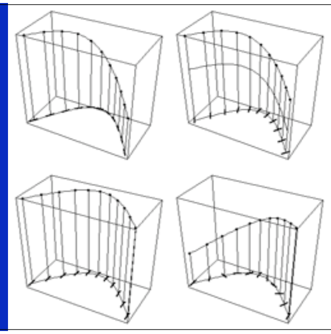
- Jan Koenderink (1987) strongly emphasized the importance of the concept of jet.
- Without jets, it is impossible to understand how the visual system could extract geometric features such as the tangent or the curvature of a curve.

– « geometrical features become *multilocal* objects, i.e. in order to compute boundary 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 *points 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). »

- To every curve C in R is associated a curve Γ in V. But the converse is false.
- Let $\Gamma = (a(s), p(s))$ be a (parametrized) curve in V. The projection $a(s)$ of Γ is a curve C in R. But Γ is the lifting of C *iff* $p(s) = y'(s) \,/\, x'(s)$.

- In differential geometry, this condition is called a Frobenius integrability condition. It says that to be a *coherent* curve in V, Γ must be an integral curve of the contact structure of the fibration π.





- Geometrically, the integrability condition means the following. Let (we suppose x is the basic variable)

$$t = (x, y, p; 1, y', p')$$

be a *tangent vector* to V at the point

$$(a, p) = (x, y, p).$$

If $y' = p$ we have

$$t = (x, y, p; 1, p, p').$$

- It is easy to show that this is equivalent to the fact that t is in the kernel of the 1-form

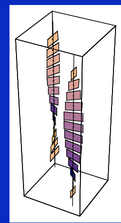
$$\omega = dy - p dx$$

$\omega = 0$ means simply $p = dy / dx$.

- But this kernel is in fact a plane called the contact plane of V at (a, p).

- The integrable curves are everywhere tangent to the field of contact planes.
- The vertical component p' of the tangent vector is then the *curvature* :

$$p = y' \Rightarrow p' = y''$$

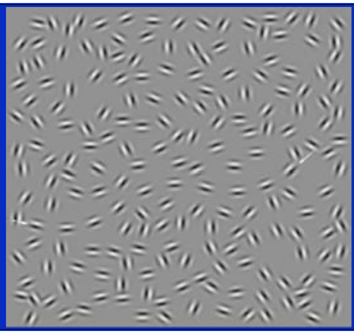


- The integrability condition for a curve Γ in V says that Γ is tangent at every of its point (a, p) to the contact plane at that point. It is in this sense that Γ is an *integral curve* of the contact structure of V.

Application to the association field

- The Frobenius integrability condition is a geometrical formulation of the Gestalt law of "good continuation" (J-M. Morel, Y. Frégnac, S. Mallat) .
- Its empirical counterpart has been studied psychophysically by David Field, Anthony Hayes and Robert Hess and explained via the concept of association field.

- Let (a_i, p_i) be a set of segments embedded in a background of distractors. The segments generate a perceptively salient curve (pop-out) iff the p_i are tangent to the curve C interpolating between the a_i.



- This is due to the fact that the activation of a simple cell detecting a pair (*a*, *p*) preactivates, via the horizontal cortico-cortical connections, cells (*b*, *q*) with *b* roughly aligned with *a* in the direction *p* and *q* close to *p*.

  – « Elements are associated according to joint constraints of position and orientation. »

  – « The orientation of the elements is locked to the orientation of the path; a smooth curve passing through the long axis can be drawn between any two successive elements. »

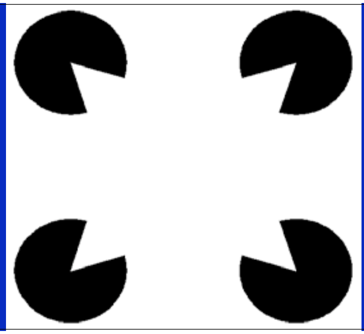
- This is a psychophysical formulation of the integrability condition.

- The pop-out of the global curve generated by the (a_i, p_i) is a typical translocal phenomenon resulting from a binding induced by the co-activation.
- Binding is a wave of activation along horizontal connections which synchronizes the cells (Singer, Gray, König).

Sub-Riemannian geometry and Kanizsa contours

- The contact structure $\mathcal{K}$ defines sub-Riemannian metrics on *V*.
- One considers metrics $g_{\mathcal{K}}$ defined only on the planes of $\mathcal{K}$ and only curves Γ in *V* which are integral curves of $\mathcal{K}$.
- We apply sub-Riemannian geometry to the analysis of Kanizsa illusory contours.

- We use curved Kanizsa contours where the sides of the internal angles of the pacmen are not aligned.



- Shimon Ullman (1976) introduced the key idea of *variational models*.

  « A network with the local property of trying to keep the contours " as straight as possible " can produce curves possessing the global property of minimizing total curvature. »

- Horn (1983) introduced the curves of least energy.
- David Mumford (1992, for *amodal* contours) used elastica: « Elastica and Computer Vision », *Algebraic Geometry and Applications*, Springer.

  Elastica are curves minimizing the integral of the square of the curvature κ, i.e. the energy

  $$E = \int (\alpha\kappa+\beta)^2 ds$$

- For natural vision, we have developped a slightly different variational model using the sub-Riemannian geometry associated to the contact structure.

- Two pacmen of respective centers *a* and *b* with a specific aperture angle define two elements (*a*, *p*) and (*b*, *q*) of *V*.
- A *K*-contour interpolating between (*a*, *p*) and (*b*, *q*) is
  - 1. a curve *C* from *a* to *b* in *R* with tangent *p* at *a* and tangent *q* at *b*;
  - 2. a curve minimizing an "energy" (variational problem).

- We lift the problem in *V*. We must find in *V* a curve Γ interpolating between (*a*, *p*) and (*b*, *q*) in *V*, wich is at the same time:
  - 1. "as straight as possible", that is "geodesic" ;
  - 2. an integral curve of the contact structure.
- In general Γ will not be a straight line because it will have to satisfy the Frobenius integrability condition.
- It is "geodesic" only in the class of integral curves of the contact structure.

- We have to solve constrained Euler-Lagrange equations for satisfying the condition of minimal length.
- It is a typical problem of sub-Riemannian geometry.
- Many very recent works on this problem.
- The natural framework is that of sub-Riemannian geometry on Lie groups.

Contact structure and Heisenberg group

- The contact structure on *V* is left-invariant for a group structure which is isomorphic to the Heisenberg group :

$$(x, y, p).(x', y', p') = (x + x', y + y' + px', p + p')$$

- If $t = (\xi, \eta, \pi)$ are the tangent vectors of T_0V, the Lie algebra of *V* has the Lie bracket

$$[t, t'] = [(\xi, \eta, \pi), (\xi', \eta', \pi')] = (0, \xi'\pi - \xi\pi', 0)$$

- It is generated by

$$\left\{t_1 = \frac{\partial}{\partial x} + p\frac{\partial}{\partial y} = (1, p, 0), t_2 = \frac{\partial}{\partial p} = (0, 0, 1)\right\}$$

(spanning the contact plane)

- We have

$$[t_1, t_2] = t_3 = -\frac{\partial}{\partial y} = (0, -1, 0)$$

(the other brackets = 0).

The Euclidean group

- But it is more natural to work with angles in the fibration $\pi : V = R \times P \rightarrow R$ with $P = \mathbb{S}^1$ and with the contact form

$$\omega = -\sin(\theta)dx + \cos(\theta)dy$$

- The contact planes are spanned by

$$X_1 = \cos(\theta)\,\partial_x + \sin(\theta)\,\partial_y$$
$$X_2 = \partial_\theta$$

with Lie bracket

$$[X_1, X_2] = \sin(\theta)\,\partial_x - \cos(\theta)\,\partial_y = -X_3$$

- *V* becomes a Lie group isomorphic to the Euclidean group (semi-direct product)

$$E(2) = SO(2) \ltimes \mathbb{R}^2$$

$$\begin{pmatrix} x_1 \\ y_1 \\ \theta_1 \end{pmatrix} \cdot \begin{pmatrix} x_2 \\ y_2 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2\cos(\theta_1) - y_2\sin(\theta_1) \\ y_1 + x_2\sin(\theta_1) + y_2\cos(\theta_1) \\ \theta_1 + \theta_2 \end{pmatrix}$$

- Inverse :

  $$\left(-x\cos(\theta) - y\sin(\theta)\,,\, x\sin(\theta) - y\cos(\theta)\,,\, -\theta\right)$$

- Left invariance

  $$\{\partial_x, \partial_y, \partial_\theta\}_0$$

  translates into ÷

$$\{\cos(\theta)\,\partial_x + \sin(\theta)\,\partial_y = X_1, -\sin(\theta)\,\partial_x + \cos(\theta)\,\partial_y = X_3, \partial_\theta = X_2\}_q$$

- And

  $$\omega_0 = dy$$

  translates into the contact form ω .

Sub-Riemannian geometry of the Euclidean group E(2)

- For the Heisenberg group there are explicit formulas for geodesics due to R. Beals, B. Gaveau, P. Greiner, A.M. Vershik, V.Y. Gershkovich.
- For the Euclidean group, after our work with Giovanna Citti and Alessandro Sarti, Andrei Agrachev and his group at the SISSA (Yuri Sachkov, Ugo Boscain, Igor Moiseev) solved the problem.

- The contact 1-form is

  $$\omega = -\sin(\theta)dx + \cos(\theta)dy$$

  and the metric makes $\{X_1, X_2, X_3\}$

  $$X_1 = \cos(\theta)\,\partial_x + \sin(\theta)\,\partial_y$$

  $$X_2 = \partial_\theta$$

  $$[X_1, X_2] = \sin(\theta)\,\partial_x - \cos(\theta)\,\partial_y = -X_3$$

  an orthonormal basis.

- Agrachev, Sachkov and Moiseev work in the fibration $V = R \times \mathbb{S}^1$ where the Legendrian lifts are solutions of the control system :

  $$\begin{cases} \dot{x} = u_1\cos(\theta) \\ \dot{y} = u_1\sin(\theta) \\ \dot{\theta} = u_2 \end{cases}$$

- They start with the kinetic energy defined on the tangent bundle TV

$$\dot{q}^2 = (u_1\cos(\theta))^2 + (u_1\sin(\theta))^2 + u_2^2 = u_1^2 + u_2^2$$

- They take the Legendre transform defined on the cotangent bundle T^*V

  $$h(p,q) = \langle p, \dot{q}\rangle - \frac{1}{2}\dot{q}^2$$

  $$p = (p_x, p_y, p_\theta) \in T_q^*V$$

- Then

$$h(p,q) = \langle p, u_1X_1(q) + u_2X_2(q)\rangle - \frac{1}{2}\left(u_1^2 + u_2^2\right)$$

- To get the Hamiltonian for geodesics they maximize $h(p, q)$ relatively to the controls u_1 and u_2. This yields

$$\begin{cases} u_1(p,q) = \langle p, X_1(q)\rangle = p_x\cos(\theta) + p_y\sin(\theta) \\ u_2(p,q) = \langle p, X_2(q)\rangle = p_\theta \end{cases}$$

- Hence the Hamiltonian on T^*V

$$H(p,q) = \frac{1}{2}\left(u_1^2 + u_2^2\right) = \frac{1}{2}\left((p_x\cos(\theta) + p_y\sin(\theta))^2 + p_\theta^2\right)$$

- Hamilton equations in the cotangent bundle T^*V are therefore :

  $$\begin{cases} \dot{x} = \frac{\partial H}{\partial p_x} = p_x\cos^2(\theta) + p_y\cos(\theta)\sin(\theta) \\ \dot{y} = \frac{\partial H}{\partial p_y} = p_y\sin^2(\theta) + p_x\cos(\theta)\sin(\theta) \\ \dot{\theta} = \frac{\partial H}{\partial p_\theta} = p_\theta \end{cases}$$

  $$\begin{cases} \dot{p}_x = -\frac{\partial H}{\partial x} = 0 \\ \dot{p}_y = -\frac{\partial H}{\partial y} = 0 \\ \dot{p}_\theta = -\frac{\partial H}{\partial \theta} = (p_x\cos(\theta) + p_y\sin(\theta))(-p_x\sin(\theta) + p_y\cos(\theta)) \end{cases}$$

  and the sub-Riemannian geodesics are the projections of the integral curves on V.

- p_x and p_y are constant. Write

  $(p_x, p_y) = \rho\exp(i\beta)$. Then

  $$\dot{p}_\theta = \frac{1}{2}\rho^2\sin(2(\theta - \beta))$$

  and H yields the first integral :

  $$\rho^2\cos^2(\theta - \beta) + p_\theta^2 = c$$

  and the ODE for θ (c , ρ and β are cst.) :

  $$\dot{\theta}^2 = p_\theta^2 = c - \rho^2\cos^2(\theta - \beta)$$

- For $\beta = 0$ (rotation invariance), the equations become :

$$\begin{cases} \dot{x} = \rho\cos^2(\theta) \\ \dot{y} = \rho\cos(\theta)\sin(\theta) = \frac{1}{2}\rho\sin(2\theta) \\ \dot{\theta} = p_\theta \\ \dot{p}_\theta = \frac{1}{2}\rho^2\sin(2\theta) \end{cases}$$

- For $\rho = 1$, $\varphi = \pi/2 - \theta$, and $\mu = 2\varphi = \pi - 2\theta$, we get a pendulum equation

$$\ddot{\mu} = -\sin(\mu)$$

with first integral

$$\dot{\varphi}^2 + \sin^2(\varphi) = c$$

- We show the trajectories in the $(\varphi, \dot{\varphi})$ plane :

- As

$$dt = \pm\frac{1}{\sqrt{c}}\frac{d\varphi}{\sqrt{1 - \frac{1}{c}\sin^2(\varphi)}}$$

the system can be integrated via elliptic functions.

- F elliptic integral of the first kind of module k

$$F(\psi, k) = \int_0^\psi \frac{1}{\sqrt{1 - k\sin^2(\theta)}}d\theta$$

- E elliptic integral of the second kind

$$E(\psi, k) = \int_0^\psi \sqrt{1 - k\sin^2(\theta)}d\theta$$

- am Jacobi amplitude, inverse of F : $\psi = am(u, k)$ iff $u = F(\psi, k)$,
- Jacobi functions $\mathrm{sn}(u) = \sin(\psi)$, $\mathrm{cn}(u) = \cos(\psi)$, $\mathrm{dn}(u) = (1 - k\sin^2(\psi))^{1/2}$.

- We get for t

$$t = \int_0^t dt = \frac{1}{\sqrt{c}}\int_{\varphi(0)}^{\varphi(t)} \frac{d\varphi}{\sqrt{1 - \frac{1}{c}\sin^2(\varphi)}}$$

$$= \frac{1}{\sqrt{c}}\left(F\left(\varphi(t), \frac{1}{c}\right) - F\left(\varphi(0), \frac{1}{c}\right)\right)$$

- For $\varphi(0) = 0$ ($\theta(0) = \pi/2$), and $c > 1$ (modulus $1/c < 1$), the pendulum makes complete turns.

$$\varphi(t) = \mathrm{am}\left(t\sqrt{c}, \frac{1}{c}\right) + k\pi$$

$$x(t) = ct - \sqrt{c}E\left(\varphi(t), \frac{1}{c}\right)$$

$$y(t) = \sqrt{c}\left(\mathrm{dn}\left(t\sqrt{c}, \frac{1}{c}\right) - 1\right)$$

-

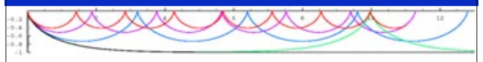


FIGURE 30. Différentes géodésiques sous-riemanniennes pour c décroissant de 2 à 1 : $c = 2$ (rouge), $c = 1.5$ (violet), $c = 1.1$ (bleu), $c = 1.001$ (vert), $c = 1$ (noir). (D'après les calculs d'A. Agrachev).

- For $c < 1$ (modulus $1/c > 1$), the pendulum oscillates between two extremal values $-\varphi_{ex}$ and $+\varphi_{ex}$ where with $\varphi_{ex} = \mathrm{Arcsin}(\sqrt{c})$

$$\theta(t) = \mathrm{Arccos}(\sqrt{c}\,\mathrm{am}(t, c))$$

$$x(t) = t - E(\mathrm{am}(t, c), c)$$

$$y(t) = \sqrt{c}(\mathrm{cn}(t, c) - 1)$$

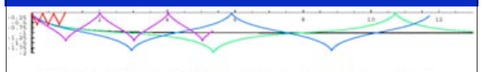


FIGURE 31. Différentes géodésiques sous-riemanniennes pour c croissant de 0.1 à 1 : $c = 0.1$ (rouge), $c = 0.5$ (violet), $c = 0.9$ (bleu), $c = 0.99$ (vert), $c = 1$ (noir). (D'après les calculs d'A. Agrachev).