Xavier Pennec

Univ. Côte d'Azur and Inria, France

UNIVERSITÉ CÔTE D'AZUR

3iA Côte d'Azur
Interdisciplinary Institute for Artificial Intelligence

Geometric Statistics

Mathematical foundations and applications in computational anatomy

Freely adapted from "Women teaching geometry", in Adelard of Bath translation of Euclid's elements, 1310.

1/ Intrinsic Statistics on Riemannian Manifolds

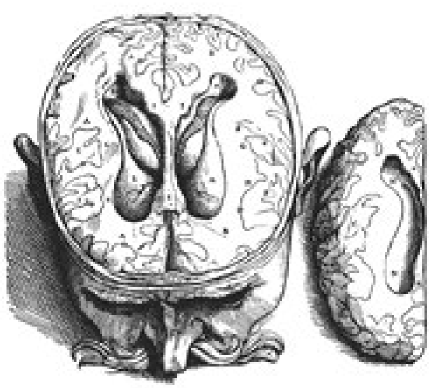
ESI semester Infinite-dimensional Geometry: Theory and Applications Week 5, 02/2025

Epione e-patient / e-medicine

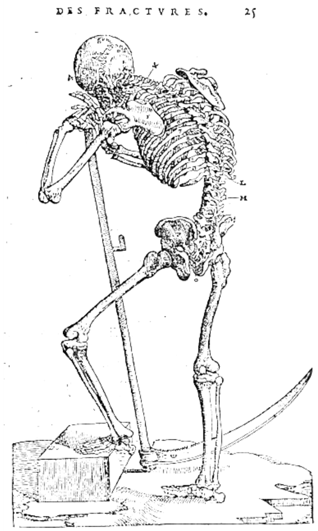
Inria

Anatomy

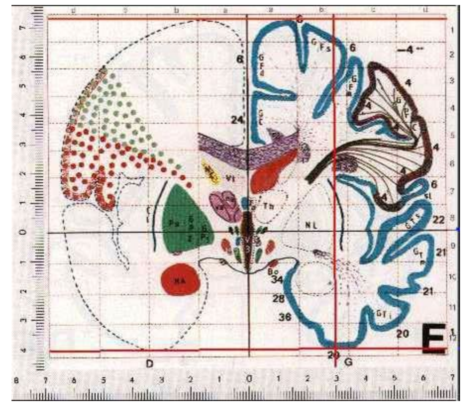
Science that studies the structure and the relationship in space of different organs and tissues in living systems [Hachette Dictionary]



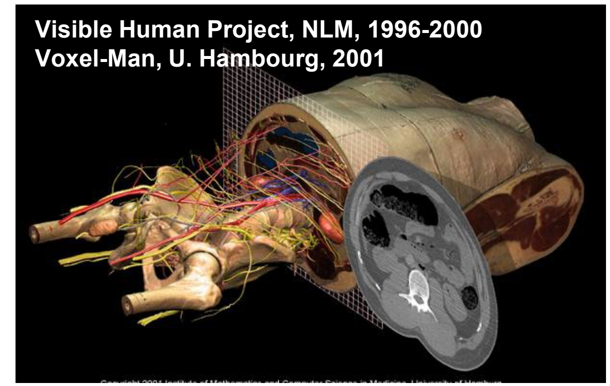
1er cerebral atlas, Vesale, 1543



Paré, 1585



Talairach & Tournoux, 1988



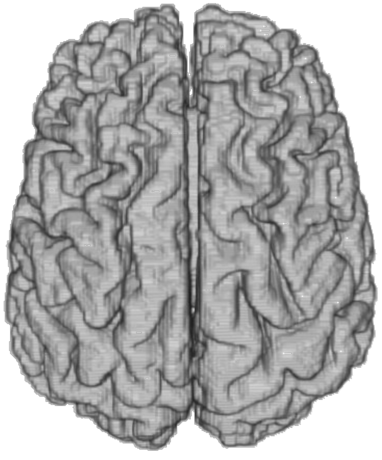
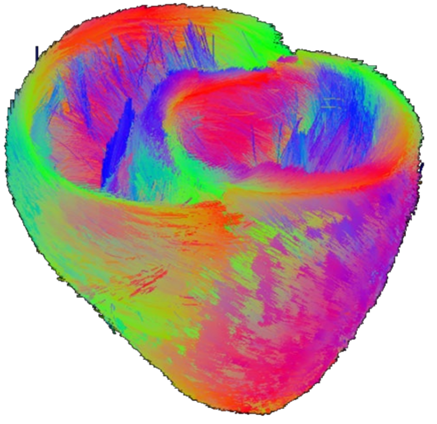
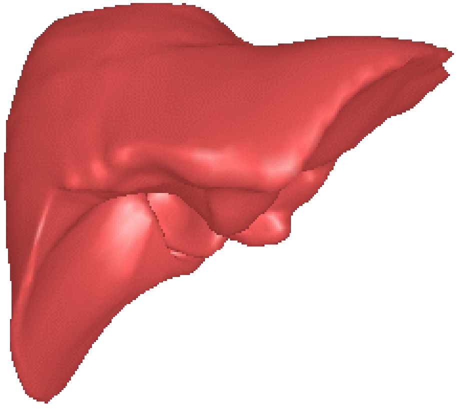
Visible Human Project, NLM, 1996-2000
Voxel-Man, U. Hambourg, 2001

Galien (131-201)
Vésale (1514-1564)
Paré (1509-1590)
Sylvius (1614-1672)
Willis (1621-1675)
Gall (1758-1828) : *Phrenology*
Talairach (1911-2007)

Revolution of observation means (~1990):

- From dissection to in-vivo in-situ imaging
- From the description of one representative individual to generative statistical models of the population

Computational Anatomy

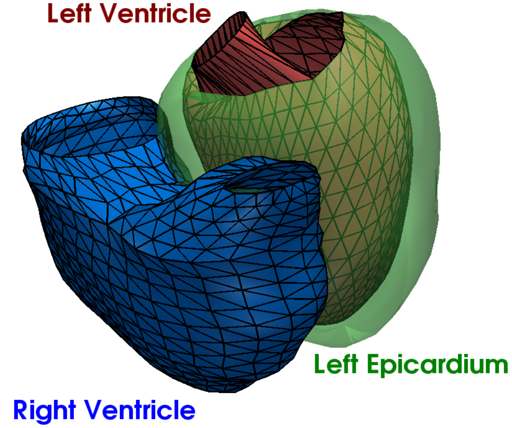
   

Statistics of organ shapes across subjects in species, populations, diseases…

- Mean shape (atlas), subspace of normal vs pathologic shapes
- Shape variability (Covariance)
- Model development across time (growth, ageing, ages…)

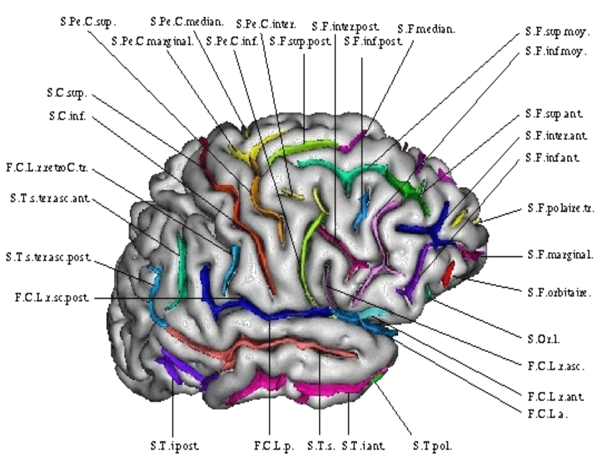
Use for personalized medicine (diagnostic, follow-up, etc)

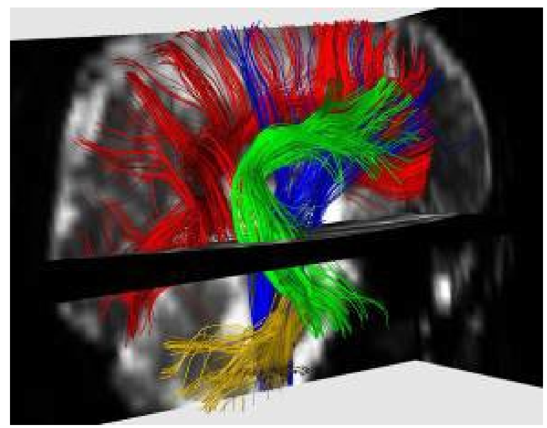
Geometric features in Computational Anatomy

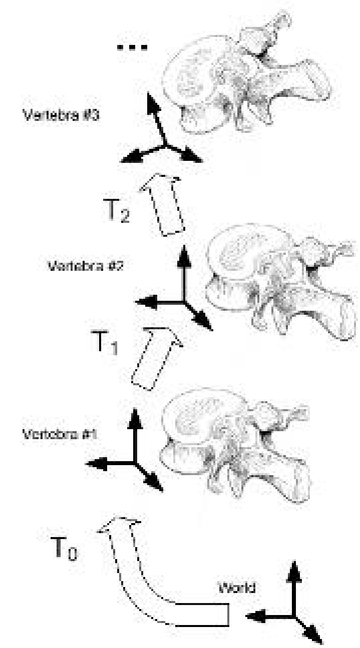


Noisy geometric features

- Curves, sets of curves (fiber tracts)
- Surfaces, SPD matrices
- Transformations







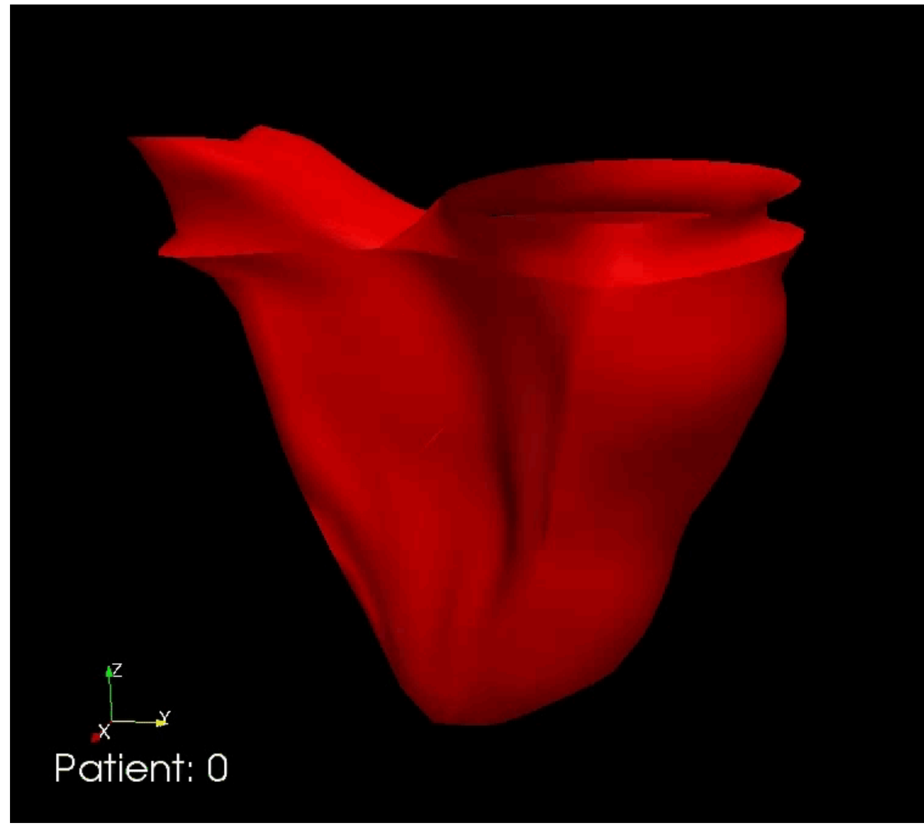
Statistical modeling at the population level

- **Simple Statistics on non-linear manifolds?**
  - Mean, covariance of its estimation, PCA, PLS, ICA
- **GS**: Statistics on manifolds vs **IG**: manifolds of statistical models

Methods of computational anatomy

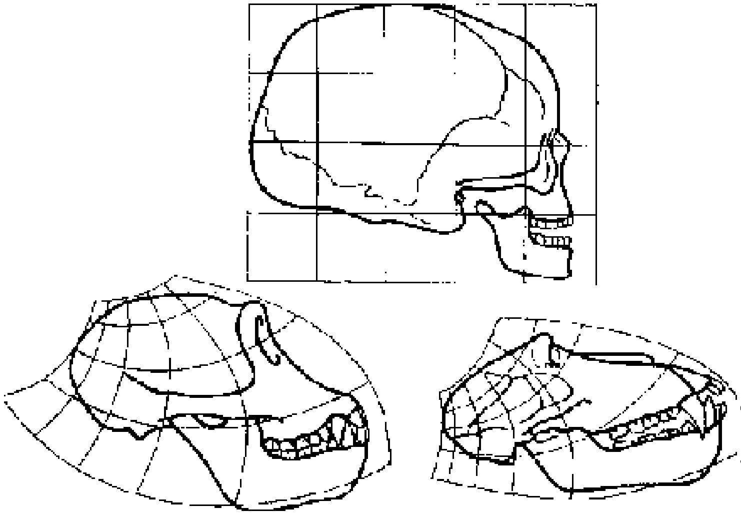
Remodeling of the right ventricle of the heart in tetralogy of Fallot

- Mean shape
- Shape variability
- Correlation with clinical variables
- Predicting remodeling effect

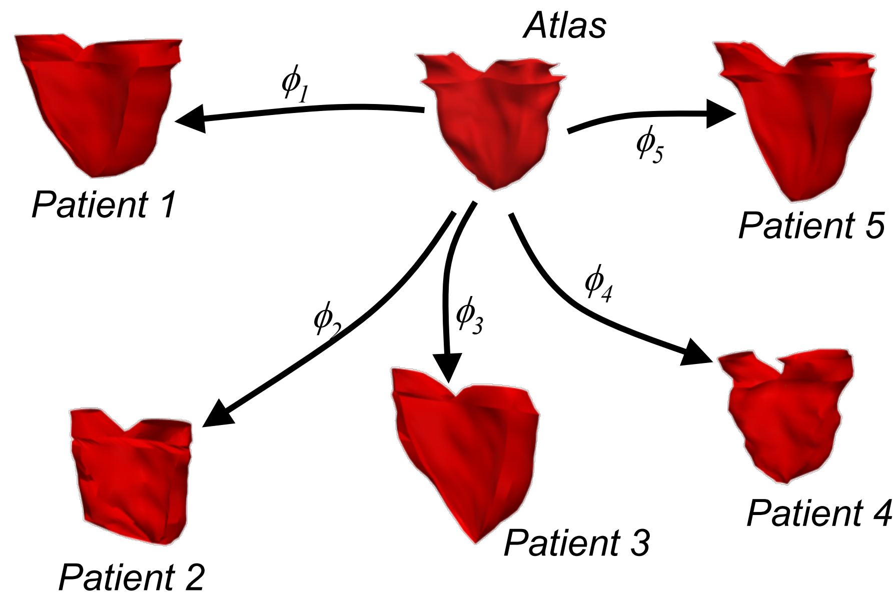


Shape of RV in 18 patients

Morphometry through Deformations



Skulls of a human, a chimpanzee and a baboon
and transformations between them



Measure of deformation [D'Arcy Thompson 1917, Grenander & Miller]

- Observation = "random" deformation of a reference template
- Reference template = Mean (atlas)
- Shape variability encoded by the deformations

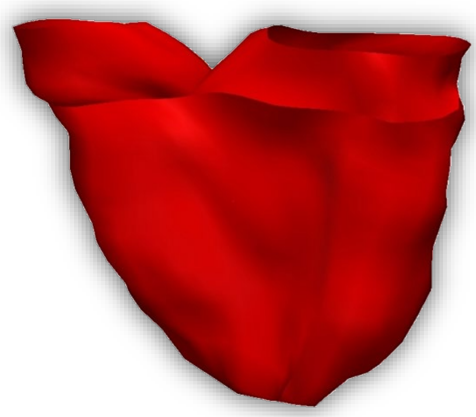
Statistics on groups of transformations (Lie groups, diffeomorphism)?

Consistency with group operations (non commutative)?

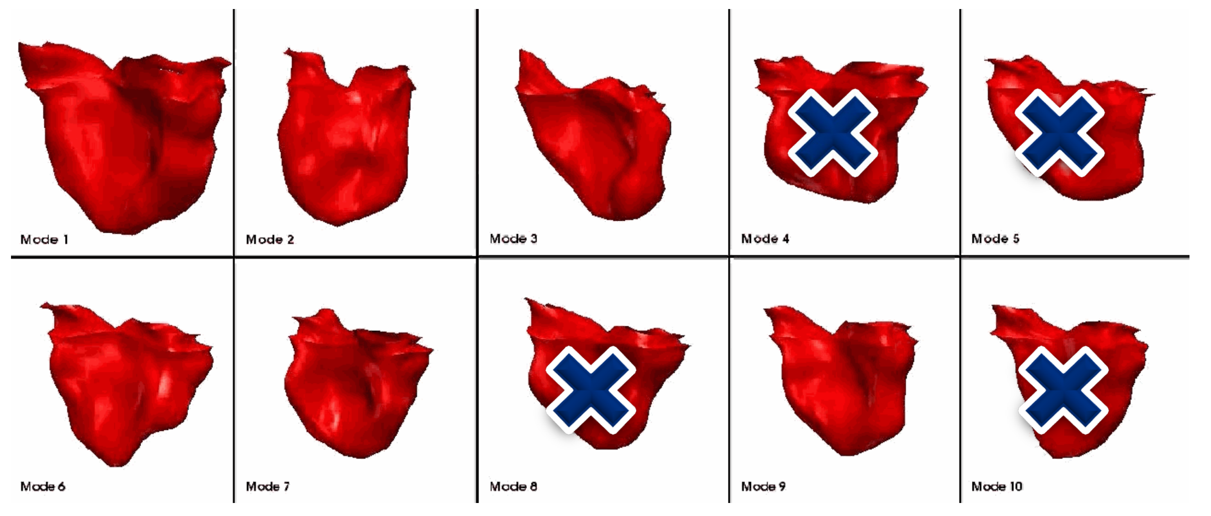
Atlas and Deformations Joint Estimation

Method: LDDMM to compute atlas + PLS on momentum maps

- Find modes that are significantly correlated to clinical variables (body surface area, tricuspid and pulmonary valve regurgitations).
- Create a generative model by regressing shape vs age (BSA)



Average RV anatomy of 18 ToF patients



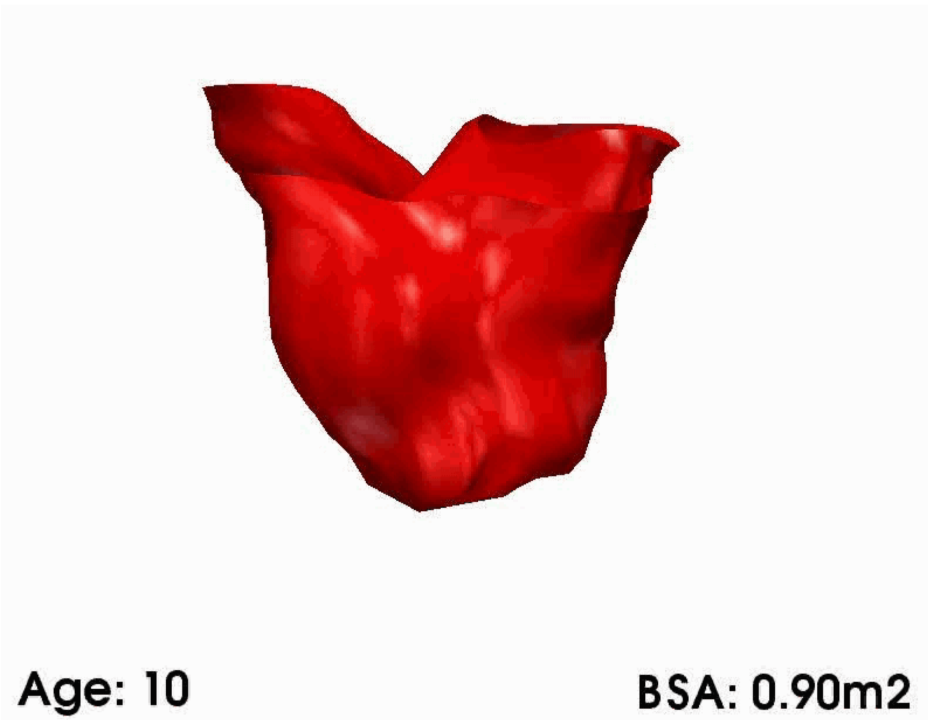
10 Deformation Modes = 90% of spectral energy

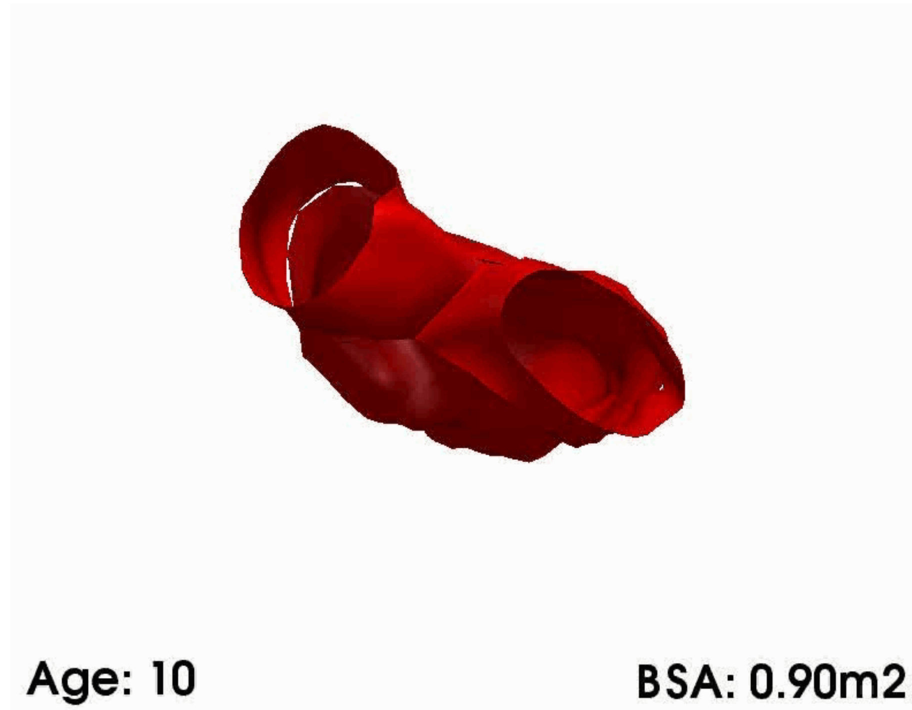
6 modes significantly correlated to BSA

[Mansi et al, MICCAI 2009, TMI 2011]

Statistical Remodeling of RV in Tetralogy of Fallot

[Mansi et al, MICCAI 2009, TMI 2011]





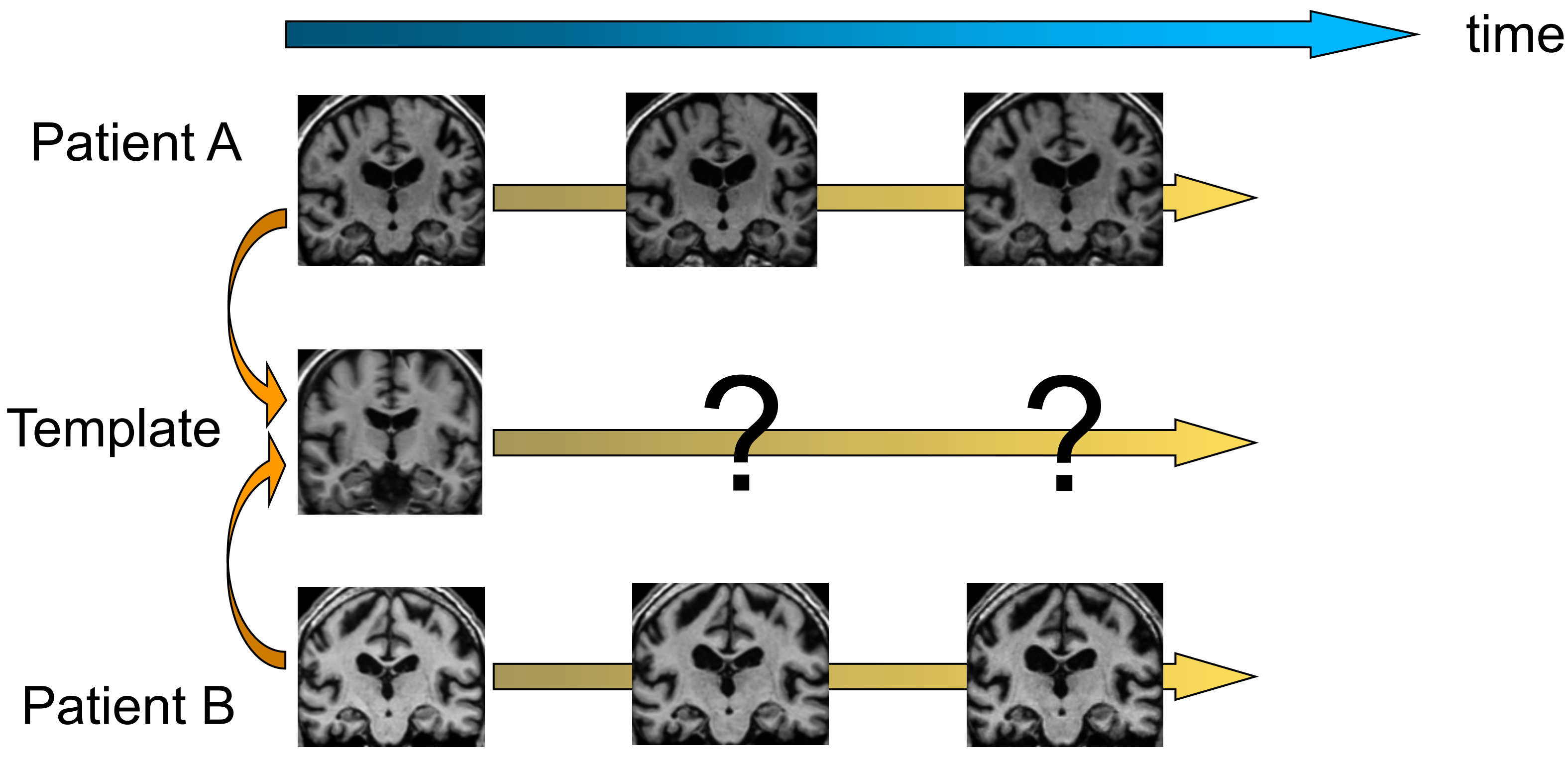
Predicted remodeling effect … **has a clinical interpretation**

Volume increases → Valve annuli deform → Pulmonary stenosis reduces → RV pressure decreases → Septum pushed inwards → RV free-wall outwards

[Mansi et al, MICCAI 2009, TMI 2011]

Longitudinal deformation analysis

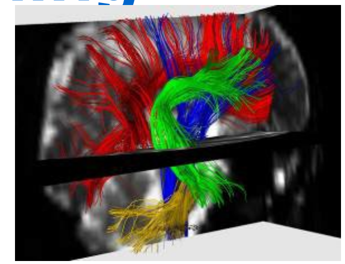
Dynamic obervations



How to transport longitudinal deformation across subjects?

What are the convenient mathematical settings?

Impact of geometry on statistical learning

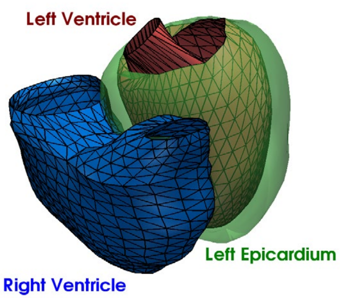


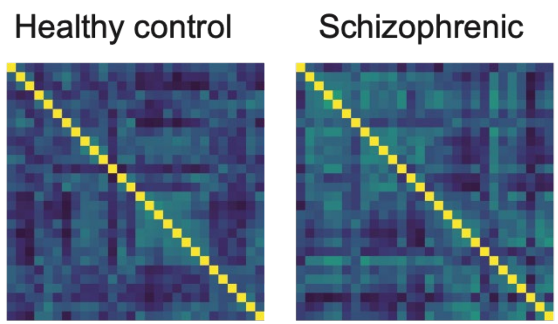
Data most often belong to non-linearity spaces

- Images, shapes, diffeomorphisms, texture, segmentations…
- Computational anatomy : Brain, heart, liver,
- Other applications: shape of molecules, OMICS correlation matrices…

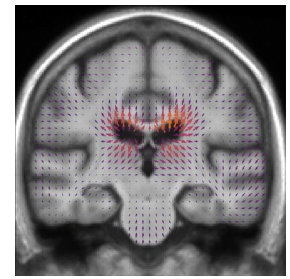
Non-linear structures: invariance → geometry

- Big data: locally flat (Euclidean)
- **Small data: geometry is the key to interpolate**

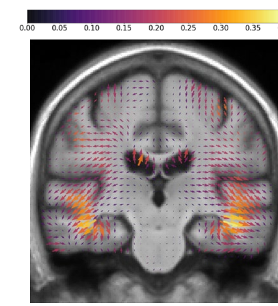




Connectomics



Normal aging



Additional AD specific

Bases of statistics in non-linear spaces

- **Simple Statistics on non-linear manifolds?**
- Mean, confidence region, PCA, PLS, ICA, transfer learning

Geometric Statistics: Mathematical foundations and applications in computational anatomy

Intrinsic Statistics on Riemannian Manifolds

- Introduction to computational anatomy
- The Riemannian manifold computational structure
- Simple statistics on Riemannian manifolds
- Applications to the spine shape and diffusion tensor imaging
- Conclusion

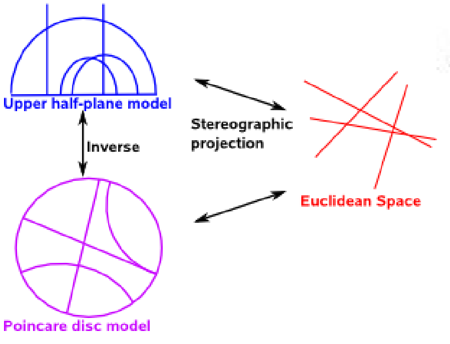
Metric and Affine Geometric Settings for Lie Groups

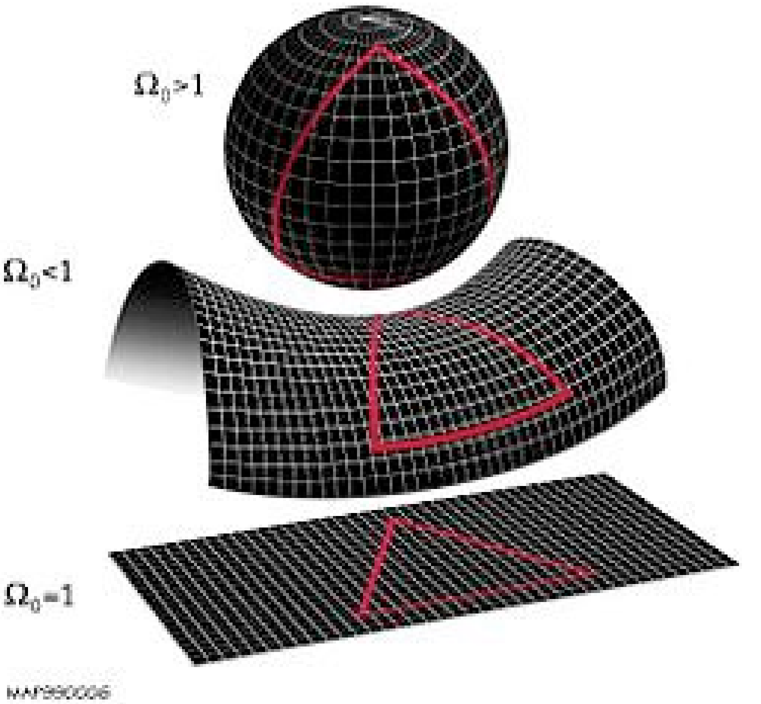
Advances Statistics: CLT & PCA

Which non-linear space?

Constant curvatures spaces

- Sphere,
- Euclidean,
- Hyperbolic

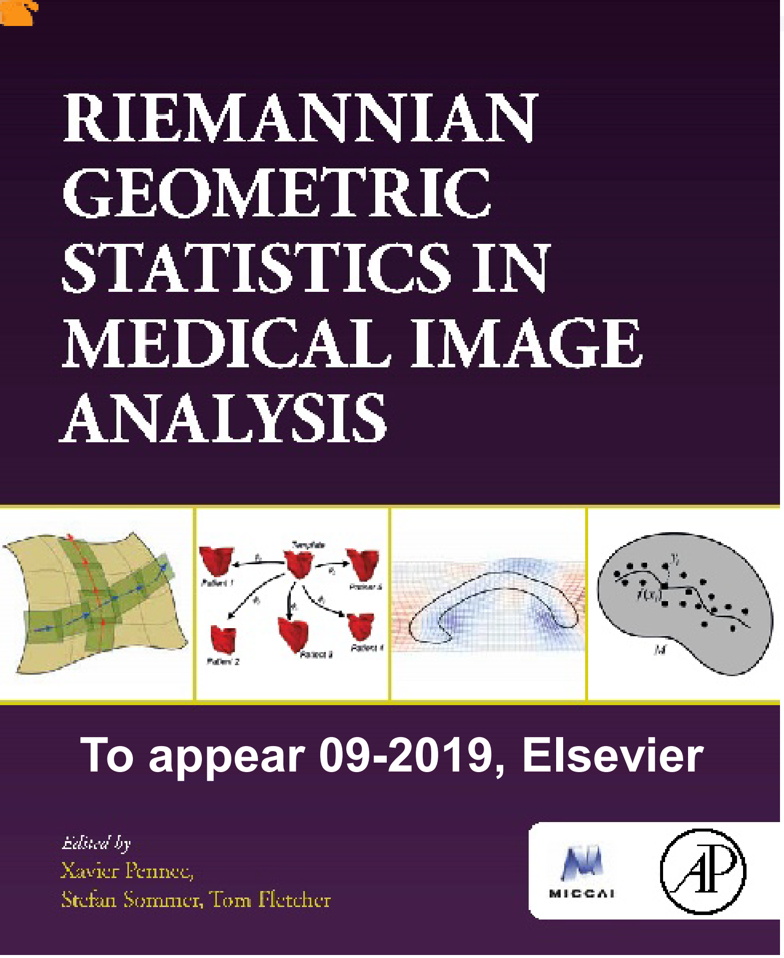




Homogeneous spaces, Lie groups and symmetric spaces

Riemannian or affine connection spaces

Towards non-smooth quotient and stratified spaces



Part 1: Foundations

- 1: Riemannian geometry [Sommer, Fetcher, Pennec]
- 2: Statistics on manifolds [Fletcher]
- 3: Manifold-valued image processing with SPD matrices [Pennec]
- 4: Riemannian Geometry on Shapes and Diffeomorphisms [Marsland, Sommer]
- 5: Beyond Riemannian: the affine connection setting for transformation groups [Pennec, Lorenzi]

Part 2: Statistics on Manifolds and Shape Spaces

- 6: Object Shape Representation via Skeletal Models (s-reps) and Statistical Analysis [Pizer, Maron]
- 7: Inductive Fréchet Mean Computation on S(n) and SO(n) with Applications [Chakraborty, Vemuri]
- 8: Statistics in stratified spaces [Ferage, Nye]
- 9: Bias in quotient space and its correction [Miolane, Devilier,Pennec]
- 10: Probabilistic Approaches to Statistics on Manifolds: Stochastic Processes, Transition Distributions, and Fiber Bundle Geometry [Sommer]
- 11: Elastic Shape Analysis, Square-Root Representations and Their Inverses [Zhang, Klassen, Srivastava]

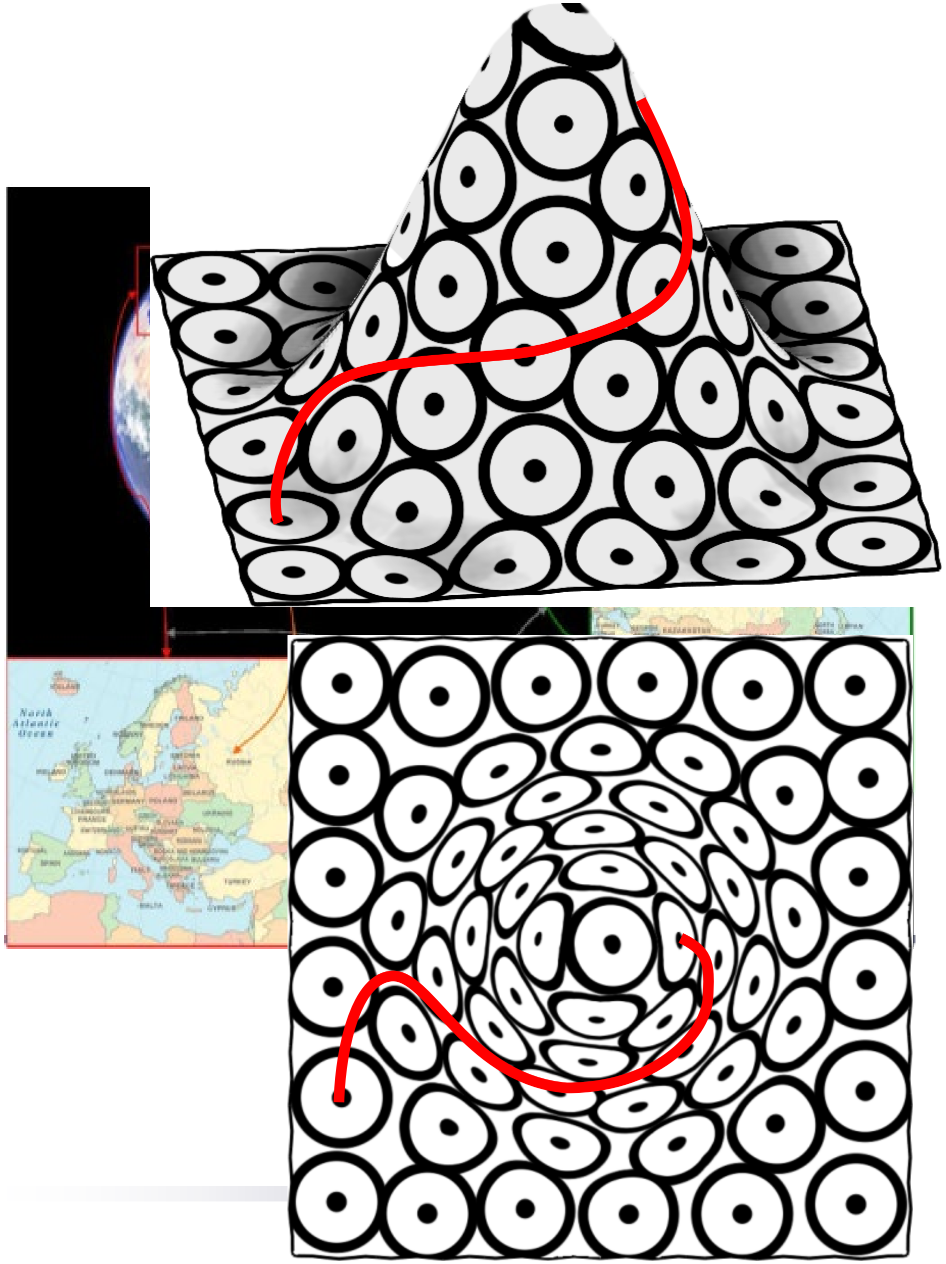
Part 3: Deformations, Diffeomorphisms and their Applications

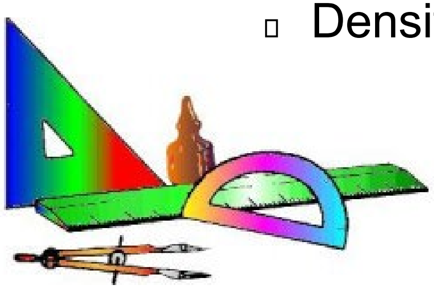
- 13: Geometric RKHS models for handling curves and surfaces in Computational Anatomy : currents, varifolds, f-shapes, normal cycles [Charlie, Charon, Glaunes, Gori, Roussillon]
- 14: A Discretize-Optimize Approach for LDDMM Registration [Polzin, Niethammer, Vialad, Modezitski]
- 15: Spatially varying metrics in the LDDMM framework [Vialard, Risser]
- 16: Low-dimensional Shape Analysis In the Space of Diffeomorphisms [Zhang, Fleche, Wells, Golland]
- 17: Diffeomorphic density matching, Bauer, Modin, Joshi]

Differentiable manifolds

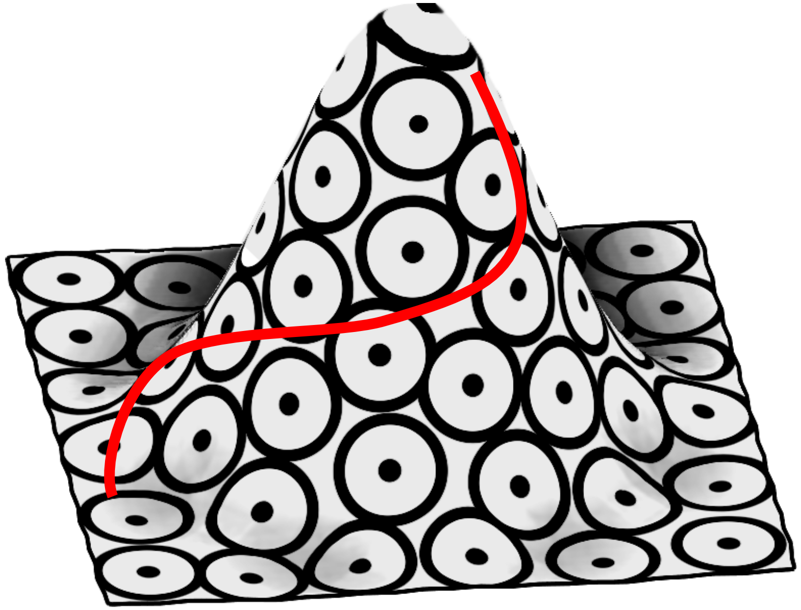
Computing on a manifold

- Extrinsic
  - Embedding in $\mathbb{R}^n$
- Intrinsic
  - Coordinates : charts
- Measuring?
  - Lengths
  - Straight lines
  - Density, volumes





Measuring extrinsic distances

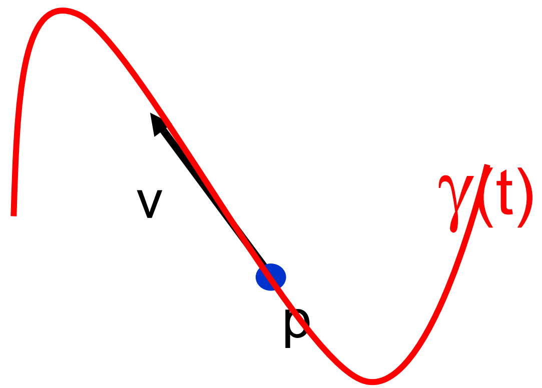


Basic tool: the scalar product

$$< v, w > = v^t w$$

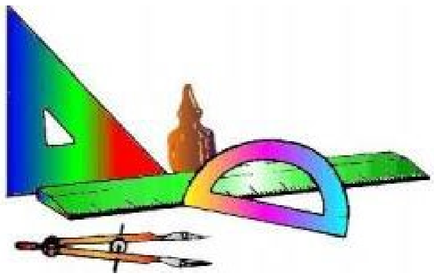
- Norm of a vector

$$\|v\| = \sqrt{< v, v >}$$



- Length of a curve

$$L(\gamma) = \int \|\dot{\gamma}(t)\| \, dt$$



Measuring extrinsic distances

Basic tool: the scalar product

Bernhard Riemann
1826-1866

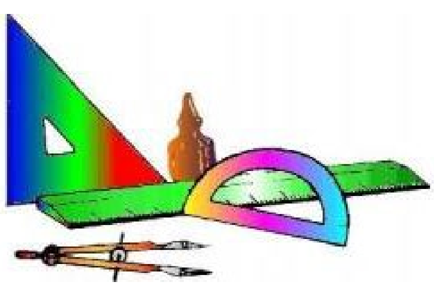
$$< v, w >_p = v^t G(p) w$$

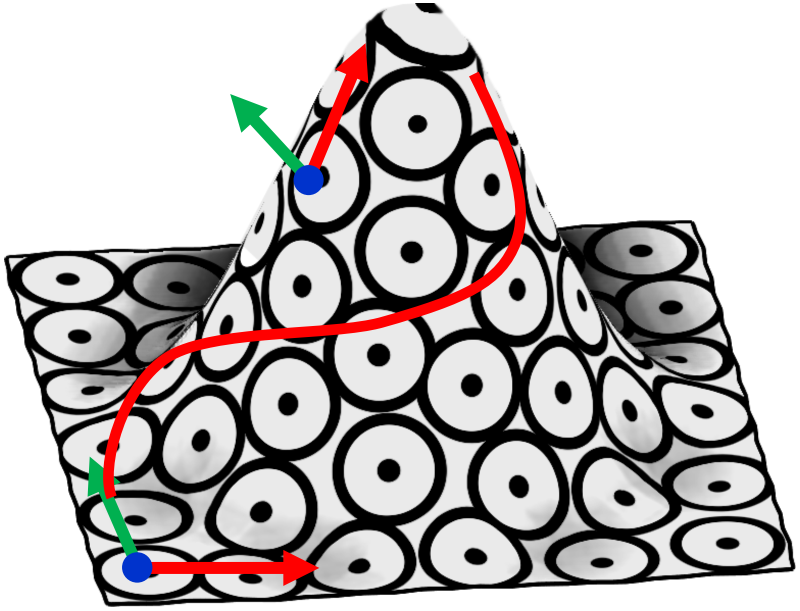
- Norm of a vector

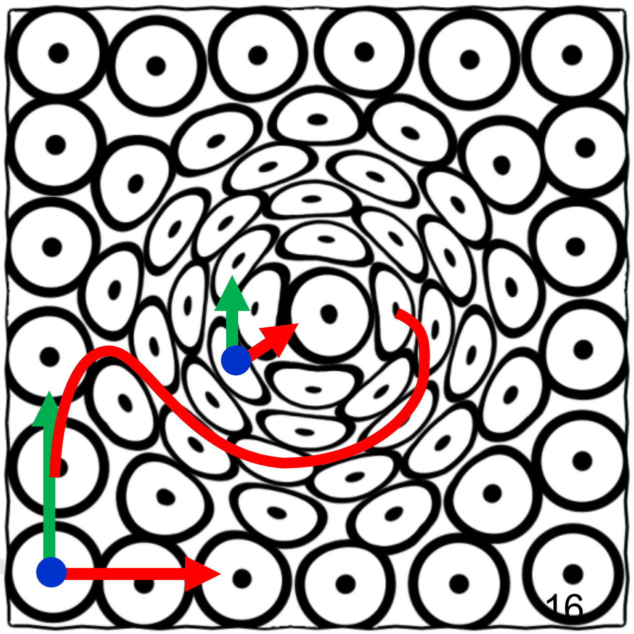
$$\|v\|_p = \sqrt{< v, v >_p}$$

- Length of a curve

$$L(\gamma) = \int \|\dot{\gamma}(t)\|_p \, dt$$





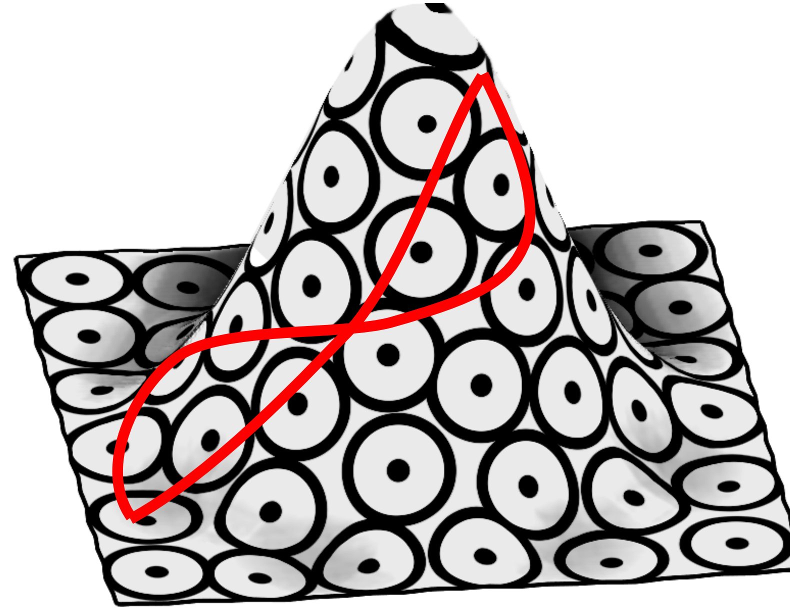


Riemannian manifolds

Basic tool: the scalar product

Bernhard Riemann
1826-1866

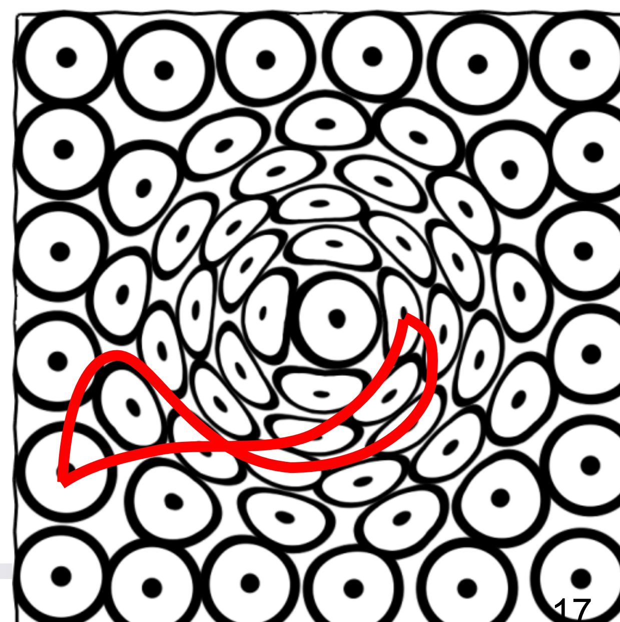
$$< v, w >_p = v^t G(p) w$$

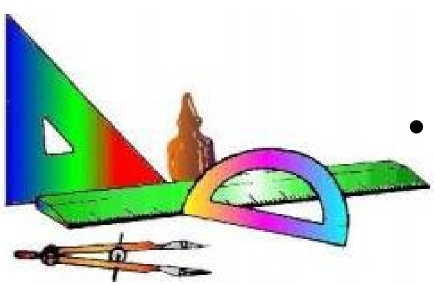


- Length of a curve

$$L(\gamma) = \int \| \dot{\gamma}(t) \|_p \, dt$$

- Geodesics
  - Shortest path between 2 points
  - Calculus of variations (E.L.) : 2nd order differential equation (specifies acceleration)
  - Free parameters: initial speed and starting point





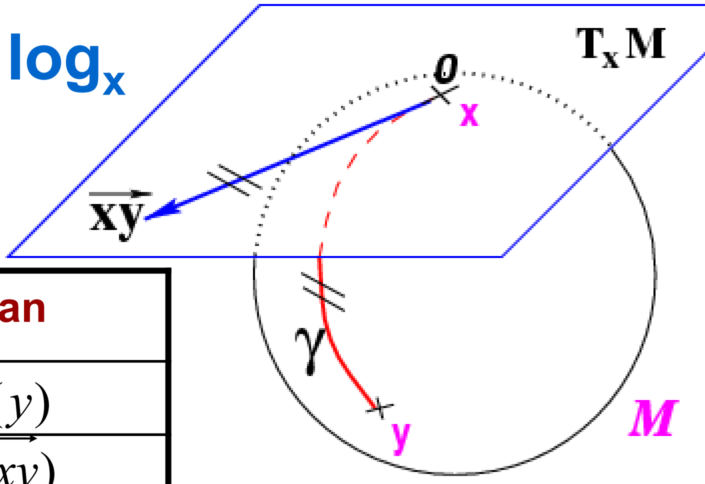
Bases of Algorithms in Riemannian Manifolds

Exponential map (Normal coordinate system):

- Exp_x = geodesic shooting parameterized by the initial tangent
- Log_x = unfolding the manifold in the tangent space along geodesics
  - Geodesics = straight lines with Euclidean distance
  - Geodesic completeness: covers M \ Cut(x)

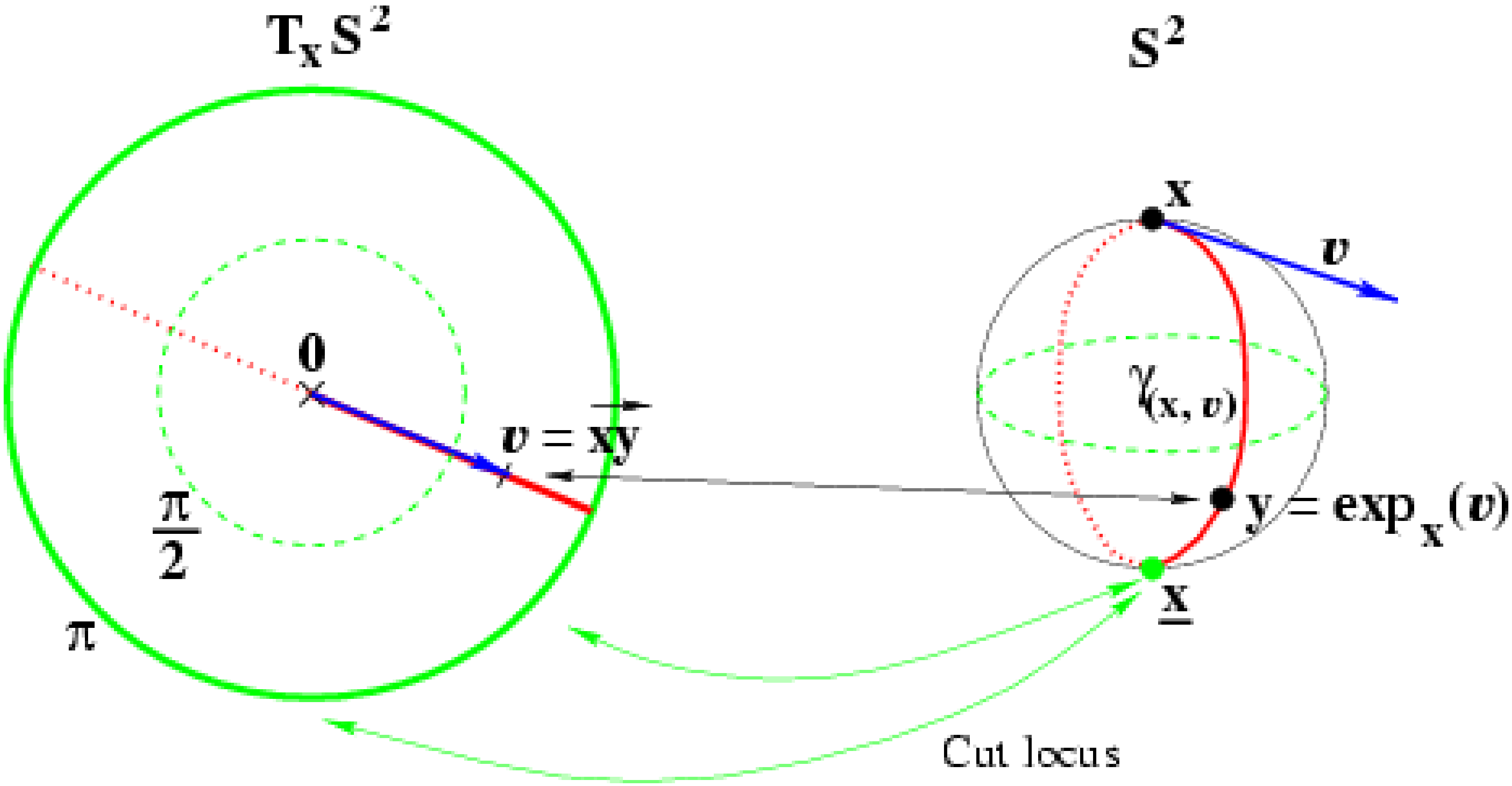
Reformulate algorithms with $\exp_x$ and $\log_x$

Vector -> Bi-point (no more equivalence classes)



Operation	Euclidean space	Riemannian
Subtraction	$\overrightarrow{xy} = y - x$	$\overrightarrow{xy} = Log_x(y)$
Addition	$y = x + \overrightarrow{xy}$	$y = Exp_x(\overrightarrow{xy})$
Distance	$\text{dist}(x, y) = \lVert y - x \rVert$	$\text{dist}(x, y) = \lVert \overrightarrow{xy} \rVert_x$
Gradient descent	$x_{t+\varepsilon} = x_t - \varepsilon \nabla C(x_t)$	$x_{t+\varepsilon} = Exp_{x_t}(-\varepsilon \nabla C(x_t))$

Cut locus



$T_x S^2$
S^2
0
$v = \overrightarrow{xy}$
$\frac{\pi}{2}$
π
x
v
$\gamma_{(x,v)}$
$y = \exp_x(v)$
$\underline{x}$
Cut locus

Geometric Statistics: Mathematical foundations and applications in computational anatomy

Intrinsic Statistics on Riemannian Manifolds

- Introduction to computational anatomy
- The Riemannian manifold computational structure
- Simple statistics on Riemannian manifolds
- Applications to the spine shape and diffusion tensor imaging
- Conclusion

Metric and Affine Geometric Settings for Lie Groups

Advances Statistics: CLT & PCA

Basic probabilities and statistics

Measure: random vector x of pdf $p_{\mathbf{x}}(z)$

Approximation: $\mathbf{x} \sim (\overline{\mathrm{x}}, \mathbf{\Sigma}_{\mathbf{xx}})$

- Mean: $\overline{\mathrm{x}} = \mathrm{E}(\mathbf{x}) = \int z . p_{\mathbf{x}}(z) . dz$
- Covariance: $\Sigma_{\mathbf{xx}} = \mathrm{E}\left[(\mathbf{x} - \overline{\mathrm{x}}).(\mathbf{x} - \overline{\mathrm{x}})^T\right]$

Propagation:

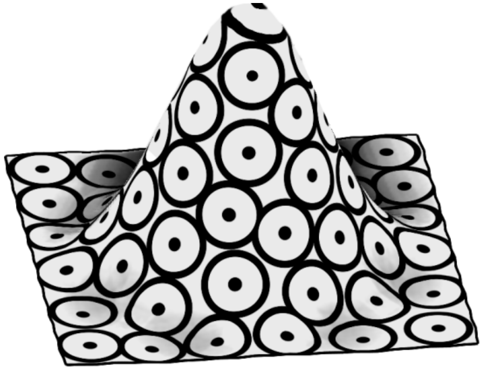
$$\mathbf{y} = h(\mathbf{x}) \sim \left(h(\overline{\mathrm{x}}),\ \frac{\partial h}{\partial \mathrm{x}} . \mathbf{\Sigma}_{\mathbf{xx}} . \frac{\partial h}{\partial \mathrm{x}}^{\mathrm{T}} \right)$$

Noise model: additive, Gaussian...

Principal component analysis

Statistical distance: Mahalanobis and χ^2

Random variable in a Riemannian Manifold



Intrinsic pdf of x

- For every set H

$$P(\mathbf{x} \in H) = \int_H p(y) dM(y)$$

- ~~Lebesgue's measure~~

→ Uniform Riemannian Mesure $dM(y) = \sqrt{\det(G(y))}\, dy$

Expectation of an observable in M

- $\boldsymbol{E}_{\mathbf{x}}[\phi] = \int_M \phi(y) p(y) dM(y)$
- $\phi = dist^2$ (variance) : $\boldsymbol{E}_{\mathbf{x}}[dist(.,y)^2] = \int_M dist(y,z)^2 p(z) dM(z)$
- $\phi = \log(p)$ (information) : $\boldsymbol{E}_{\mathbf{x}}[\log(p)] = \int_M p(y) \log(p(y)) dM(y)$
- ~~$\phi = x$ (mean) : $\boldsymbol{E}_{\mathbf{x}}[\mathbf{x}] = \int_M y\, p(y) dM(y)$~~

First statistical tools


Maurice Fréchet
(1878-1973)

From the mean to the Fréchet mean set

- Integral only valid in Hilbert/Wiener spaces [Fréchet 44]
- $\sigma^2(x) = Tr_g\big(\mathfrak{M}_2(x)\big) = \int_M dist^2(x,z)\, P(dz)$
- **Fréchet mean** [1948] = global minima of Mean Sq. Dev.
- **Exponential barycenters** [Emery & Mokobodzki 1991]

  $\mathfrak{M}_1(\bar{x}) = \int_M Log_{\bar{x}}(z)\, P(dz) = 0$ [critical points if P(C) =0]

Moments of a random variable: tensor fields

- $\mathfrak{M}_1(x) = \int_M Log_x(z)\, P(dz)$ Tangent mean: (0,1) tensor field
- $\mathfrak{M}_2(x) = \int_M Log_x(z) \otimes Log_x(z)\, P(dz)$ Second moment: (0,2) tensor field
  - Tangent covariance field: $Cov(x) = \mathfrak{M}_2(x) - \mathfrak{M}_1(x) \otimes \mathfrak{M}_1(x)$
- $\mathfrak{M}_k(x) = \int_M Log_x(z) \otimes Log_x(z) \ldots \otimes Log_x(z)\, P(dz)$ k-contravariant tensor field

Fréchet expectation (1944)

Minimizing the variance

$$E[\mathbf{x}] = \underset{y \in M}{\operatorname{argmin}} \left(E\left[\operatorname{dist}(y, \mathbf{x})^2\right]\right)$$

Existence

- Finite variance at one point

Characterization as an exponential barycenter (P(C)=0)

$$\operatorname{grad}\left(\sigma_{\mathbf{x}}^2(y)\right) = 0 \quad \Rightarrow \quad E\left[\overrightarrow{\bar{x}\mathbf{x}}\right] = \int_M \overrightarrow{\bar{x}\mathbf{x}}.p_{\mathbf{x}}(z).dM(z) = 0$$

Uniqueness **Karcher 77 / Kendall 90 / Afsari 10 / Le 10**

- Unique Karcher mean (thus Fréchet) if distribution has support in a regular geodesic ball with radius $r < r^* = \frac{1}{2}\min(inj(M), \pi/\sqrt{\kappa})$ (k upper bound on sectional curvatures on M)
- Empirical mean: a.s. uniqueness [Arnaudon & Miclo 2013]

Other central primitives

$$E^{\alpha}[\mathbf{x}] = \underset{y \in M}{\operatorname{argmin}} \left(E\left[\operatorname{dist}(y, \mathbf{x})^{\alpha}\right]\right)^{1/\alpha}$$

Algorithms to compute the mean

Karcher flow (gradient descent)

$$\bar{x}_{t+1} = \exp_{\bar{x}_t}(\epsilon_t \, v_t) \;\; with \;\; v_t = \mathrm{E}(\overrightarrow{\mathbf{yx}}) = \frac{1}{n} \sum_i \log_{\bar{x}_t}(x_i)$$

- Usual algorithm with $\epsilon_t = 1$ can diverge on SPD matrices
  [Bini & Iannazzo, Linear Algebra Appl., 438:4, 2013]
- Convergence for non-negative curvature (p-means)
  [Afsari, Tron and Vidal, SICON 2013]

Inductive / incremental weighted means

- $\bar{x}_{k+1} = \exp_{\bar{x}_k}\left(\frac{1}{k} \, v_k\right) \;\; with \;\; v_k = \log_{\bar{x}_k}(x_{k+1})$
- On negatively curved spaces [Sturm 2003],
  BHV centroid [Billera, Holmes, Vogtmann, 2001]
- On non-positive spaces [G. Cheng, J. Ho, H. Salehian, B. C. Vemuri 2016]

Stochastic algorithm

- [Arnaudon & Miclo, Stoch. Processes and App. 124, 2014]

A gradient descent (Gauss-Newton) algorithm

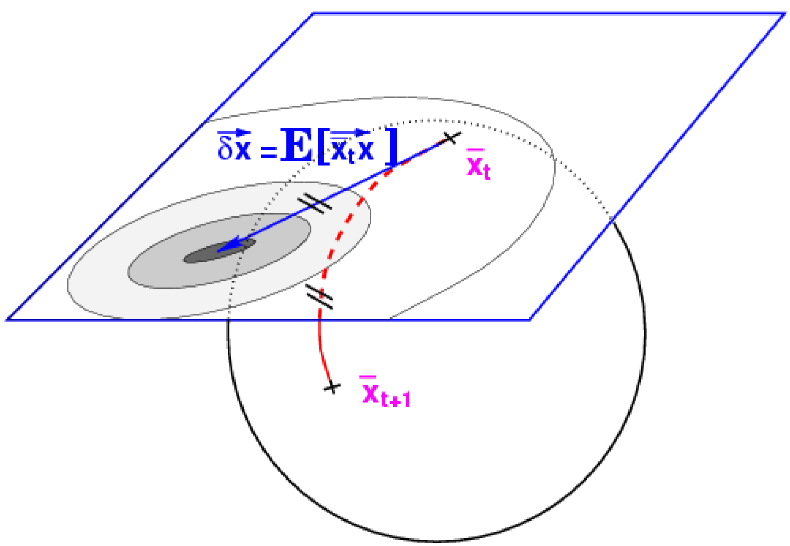
Vector space

$$f(x+v) = f(x) + \nabla f^T . v + \frac{1}{2} v^T . H_f . v$$

$$x_{t+1} = x_t + v \quad \text{with} \quad v = -H_f^{(-1)} . \nabla f$$

Manifold

$$f(\exp_x(v)) = f(x) + \nabla f(v) + \frac{1}{2} H_f(v,v)$$



$$\nabla(\sigma_{\mathbf{x}}^2(\mathrm{y})) = -2\, \mathrm{E}[\overrightarrow{\mathrm{y}\mathbf{x}}] = \frac{-2}{n} \sum_i \overrightarrow{\mathrm{y}\mathrm{x}_i}$$

$\boldsymbol{H_{\sigma_x^2}} \approx \boldsymbol{2\, I_d}$ **(for Euclidean spaces…)**

Geodesic marching

$$\overline{\mathrm{x}}_{t+1} = \exp_{\overline{\mathrm{x}}_t}(v) \quad \text{with} \quad v = \mathrm{E}\left[\overrightarrow{\mathrm{y}\mathbf{x}}\right]$$

Distributions for parametric tests

Uniform density:

- maximal entropy knowing X $\quad p_{\mathbf{x}}(z) = \mathrm{Ind}_X(z) / \mathrm{Vol}(X)$

Generalization of the Gaussian density:

- Stochastic heat kernel p(x,y,t) [complex time dependency]
- Wrapped Gaussian [Infinite series difficult to compute]
- **Maximal entropy knowing the mean and the covariance**

$$N(y) = k.\exp\left(\left(\overrightarrow{\bar{\mathrm{x}}\mathbf{x}}\right)^{\mathrm{T}} .\mathbf{\Gamma}. \left(\overrightarrow{\bar{\mathrm{x}}\mathbf{x}}\right) / 2 \right)$$

$$\mathbf{\Gamma} = \mathbf{\Sigma}^{(-1)} - \tfrac{1}{3}\mathrm{Ric} + O(\sigma) + \varepsilon(\sigma / r)$$

$$k = (2\pi)^{-n/2}.\det(\mathbf{\Sigma})^{-1/2}.\left(1 + O(\sigma^3) + \varepsilon(\sigma / r)\right)$$

Mahalanobis D2 distance / test: $\quad \mu_{\mathbf{x}}^2(\mathrm{y}) = \overrightarrow{\bar{\mathrm{x}}\mathrm{y}}^{\,t} .\Sigma_{\mathbf{xx}}^{(-1)}. \overrightarrow{\bar{\mathrm{x}}\mathrm{y}}$

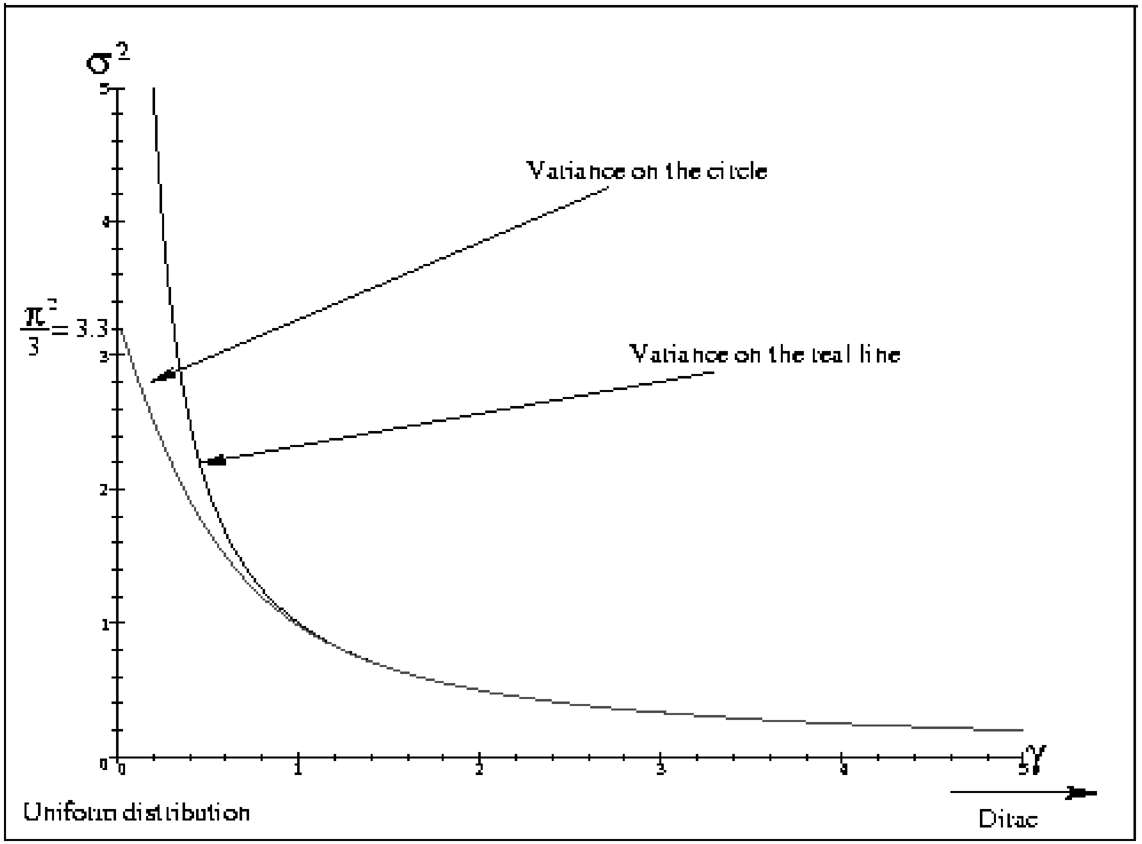
- Any distribution: $\quad \mathrm{E}\left[\mu_{\mathbf{x}}^2(\mathbf{x})\right] = n$
- Gaussian: $\quad \mu_{\mathbf{x}}^2(\mathbf{x}) \propto \chi_n^2 + O(\sigma^3) + \varepsilon(\sigma / r)$

[Pennec, JMIV06, NSIP'99]

Gaussian on the circle

Exponential chart: $x = r\theta \in \left]-\pi.r\,;\pi.r\right[$

Gaussian: truncated standard Gaussian



$r \to \infty$: standard Gaussian (Ricci curvature → 0)

$\gamma \to 0$: uniform pdf with

$$\sigma^2 = (\pi.r)^2 / 3$$

(compact manifolds)

$\gamma \to \infty$: Dirac

Extending PCA: tangent PCA vs PGA

Tangent PCA

- Generative model: Gaussian
- Find the subspace that best explains the variance
  → Maximize the squared distance to the mean

PGA (Fletcher 2004, Sommer 2014)

- Generative model:
  - Implicit uniform distribution within the subspace
  - Gaussian distribution in the vertical space
- Find a low dimensional subspace (geodesic subspaces?) that minimizes the error
  → Minimize the squared Riemannian distance from the measurements to that sub-manifold (no closed form)

Different models in curved spaces (no Pythagore thm)

Extension to BSA in course 3

Geometric Statistics: Mathematical foundations and applications in computational anatomy

Intrinsic Statistics on Riemannian Manifolds

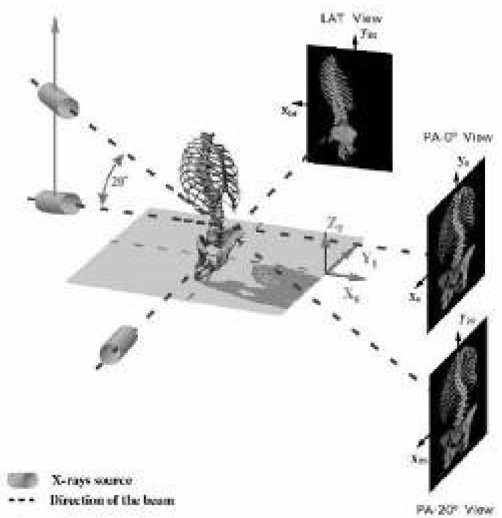
- Introduction to computational anatomy
- The Riemannian manifold computational structure
- Simple statistics on Riemannian manifolds
- Applications to the spine shape and diffusion tensor imaging
- Conclusion

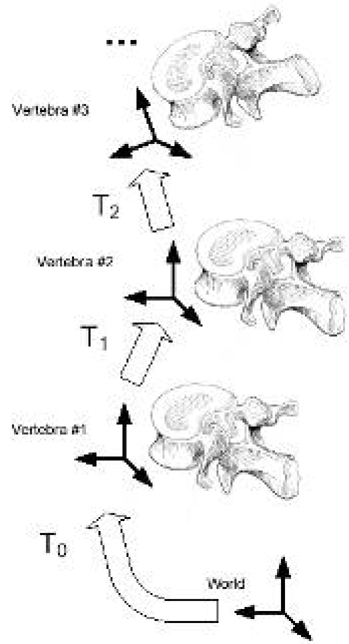
Metric and Affine Geometric Settings for Lie Groups

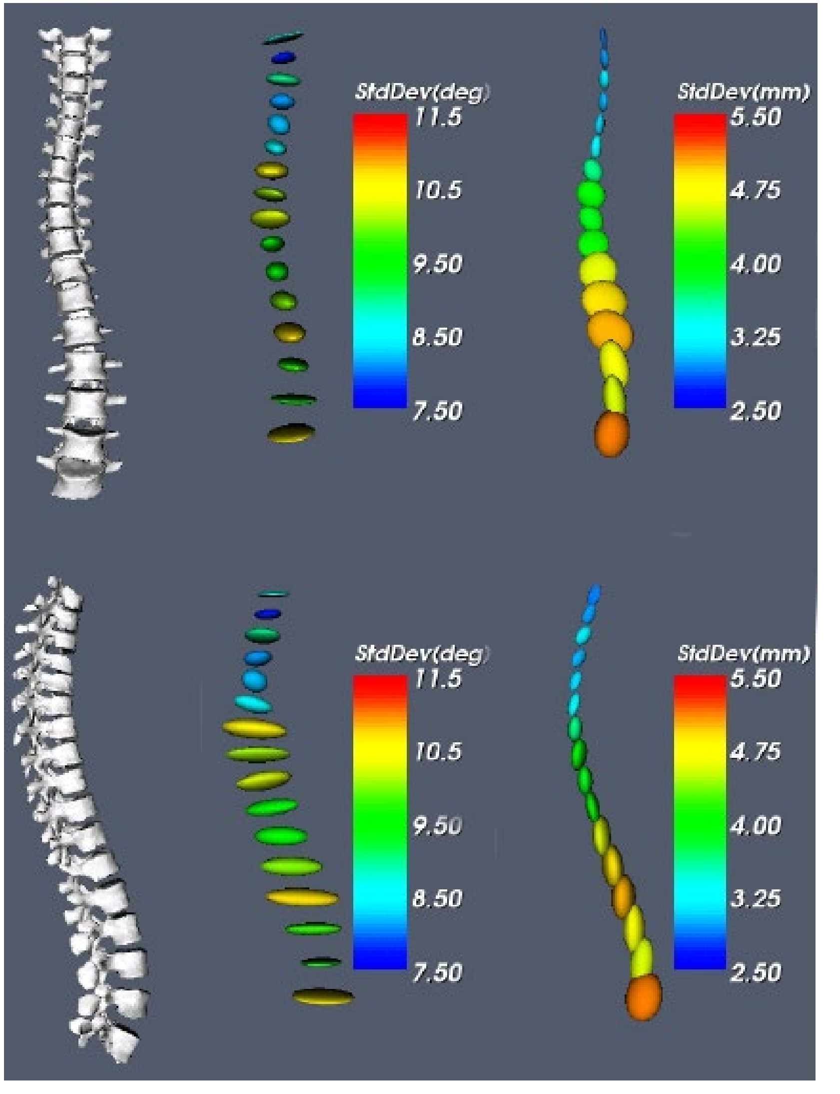
Advances Statistics: CLT & PCA

Statistical Analysis of the Scoliotic Spine

[J. Boisvert et al. ISBI'06, AMDO'06 and IEEE TMI 27(4), 2008]







Database

- 307 Scoliotic patients from the Montreal's Sainte-Justine Hospital.
- 3D Geometry from multi-planar X-rays

Mean

- Main translation variability is axial (growth?)
- Main rot. var. around anterior-posterior axis

Statistical Analysis of the Scoliotic Spine

[J. Boisvert et al. ISBI'06, AMDO'06 and IEEE TMI 27(4), 2008]
AMDO'06 best paper award, Best French-Quebec joint PhD 2009

PCA of the Covariance:
4 first variation modes have clinical meaning

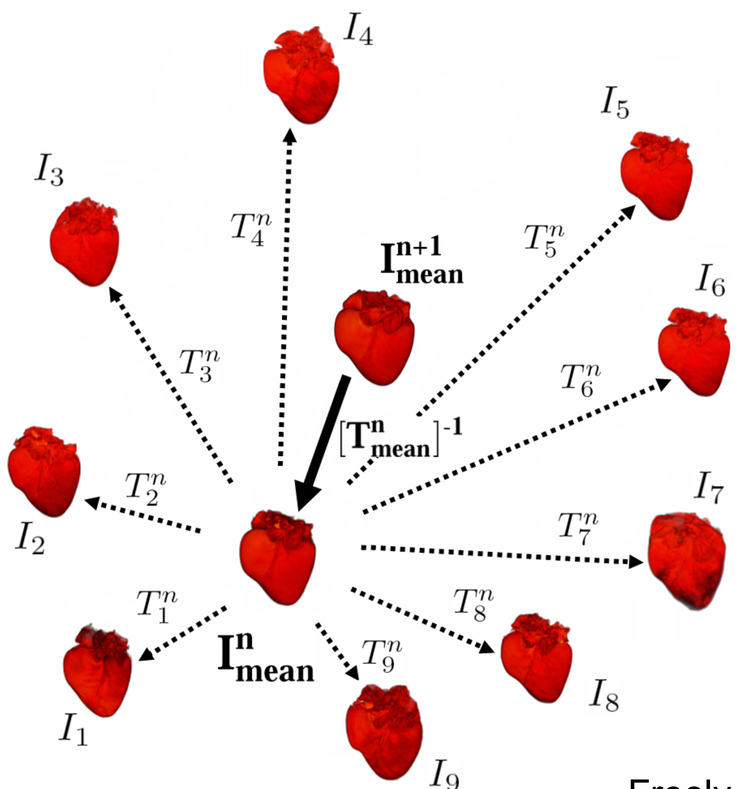
- Mode 1: King's class I or III
- Mode 2: King's class I, II, III
- Mode 3: King's class IV + V
- Mode 4: King's class V (+II)

A Statistical Atlas of the Cardiac Fiber Structure

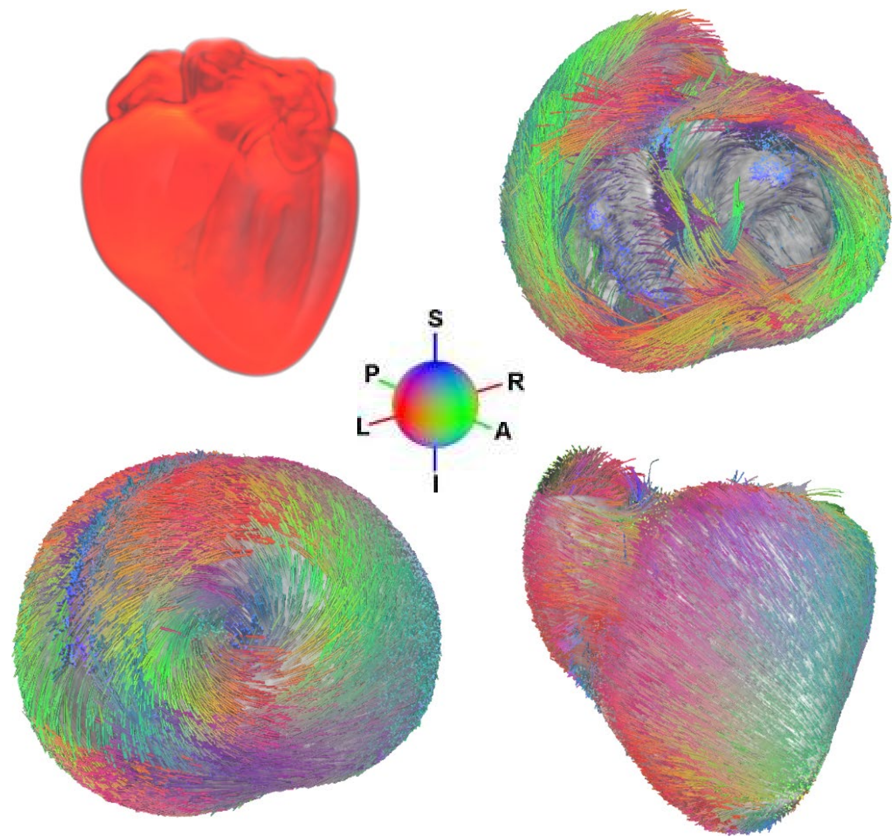
[J.M. Peyrat, et al., MICCAI'06, TMI 26(11), 2007]

Manifold data on a manifold

- Anatomical MRI and DTI
- Diffusion tensor on a 3D shape

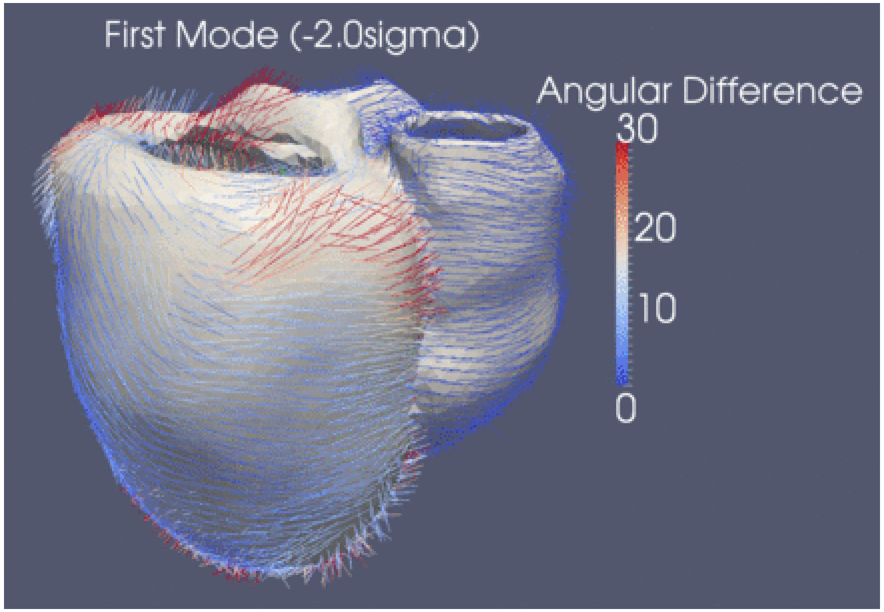


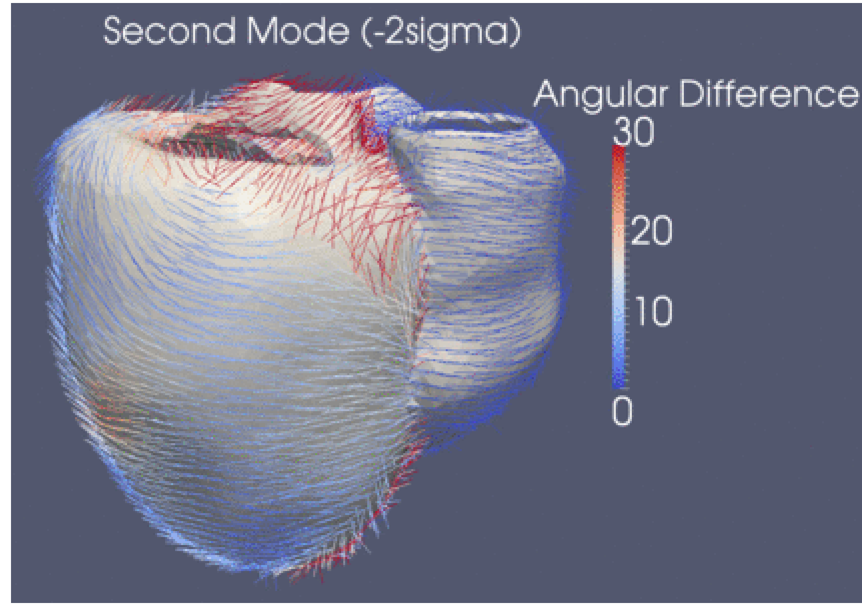
- *Average cardiac structure*
- Variability of fibers & collagen sheets



Freely available at **http://www-sop.inria.fr/asclepios/data/heart**

A Statistical Atlas of the Cardiac Fiber Structure

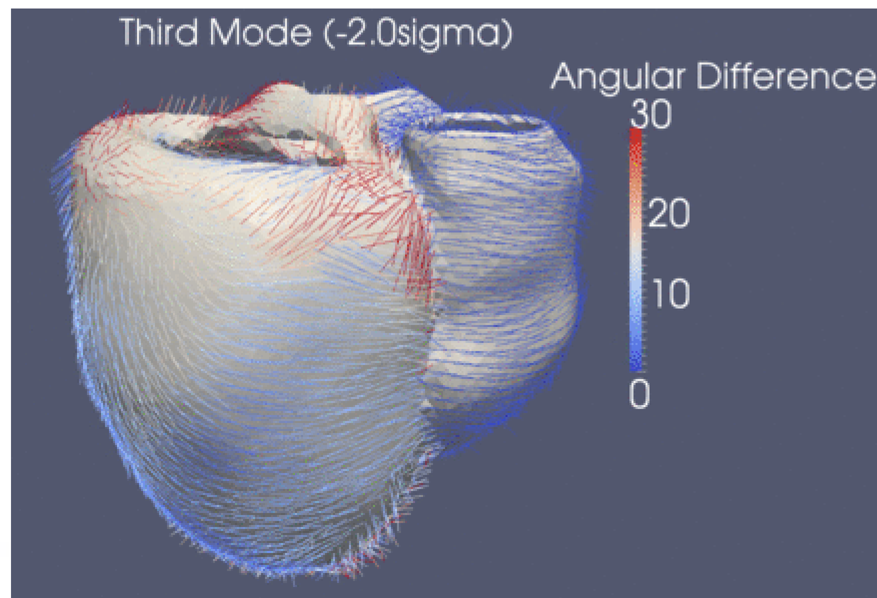




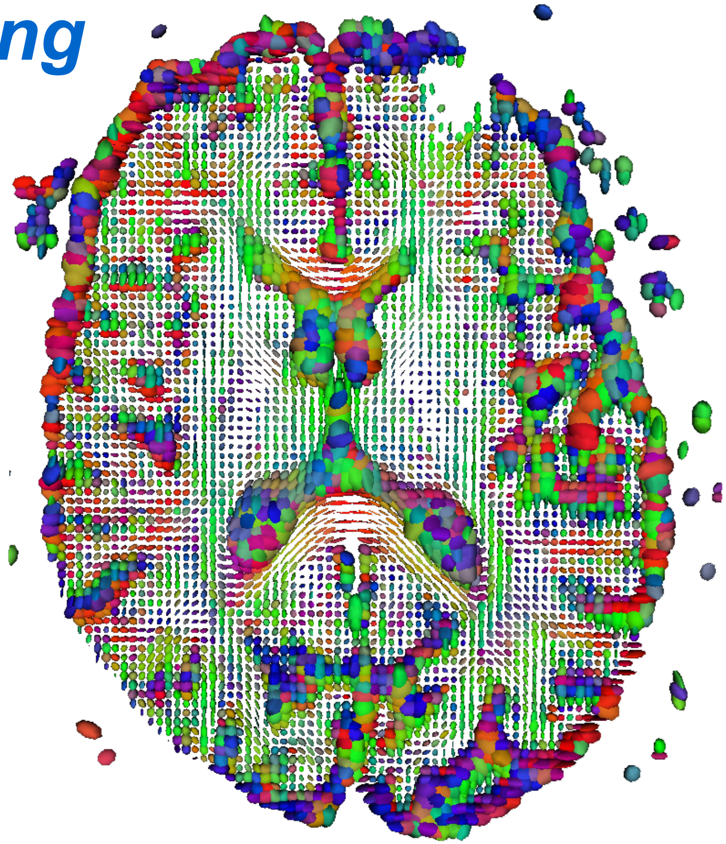
10 human ex vivo hearts (CREATIS-LRMN, France)

- Classified as healthy (controlling weight, septal thickness, pathology examination)
- Acquired on 1.5T MR Avento Siemens
  - bipolar echo planar imaging, 4 repetitions, 12 gradients
- Volume size: 128×128×52, 2 mm resolution

[R. Mollero, M.M Rohé, et al, FIMH 2015]



Manifold-valued image processing: Diffusion Tensor Imaging



Covariance of the Brownian motion of water

- Architecture of axonal fibers
- Filtering, regularization to remove noise
- Interpolation / extrapolation

Symmetric positive definite matrices

- Cone in Euclidean space (not complete)
- Convex operations are stable
  - mean, interpolation
- More complex operations are not
  - PDEs, gradient descent…

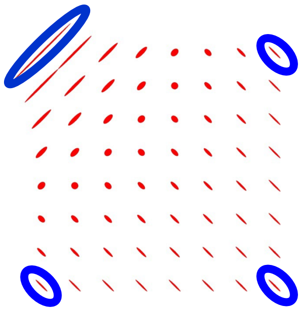
All invariant metrics under GL(n)

$$\langle W_1 \mid W_2 \rangle_{Id} = \mathrm{Tr}\left(W_1^T W_2\right) + \beta\, \mathrm{Tr}(W_1).\mathrm{Tr}(W_2) \quad (\beta > \text{-}1/\mathrm{n})$$

- Exponential map $Exp_\Sigma(\overrightarrow{\Sigma\Psi}) = \Sigma^{1/2} \exp(\Sigma^{-1/2}.\overrightarrow{\Sigma\Psi}.\Sigma^{-1/2})\Sigma^{1/2}$
- Log map $\overrightarrow{\Sigma\Psi} = Log_\Sigma(\Psi) = \Sigma^{1/2} \log(\Sigma^{-1/2}.\Psi.\Sigma^{-1/2})\Sigma^{1/2}$
- Distance $dist(\Sigma,\Psi)^2 = \left\langle \overrightarrow{\Sigma\Psi} \mid \overrightarrow{\Sigma\Psi} \right\rangle_\Sigma = \left\| \log(\Sigma^{-1/2}.\Psi.\Sigma^{-1/2}) \right\|_{Id}^2$

Manifold-valued image processing

Integral or sum in M: weighted Fréchet mean



- Interpolation
  - Linear between 2 elements: interpolation geodesic
  - Bi- or tri-linear or spline in images: weighted means
- Gaussian filtering: **convolution = weighted Fréchet mean**

  [Pennec, Fillard, Arsigny, IJCV 66(1), 2006]

$$\Sigma(x) = \min \sum_i G_\sigma(x - x_i)\ \operatorname{dist}^2(\Sigma, \Sigma_i)$$

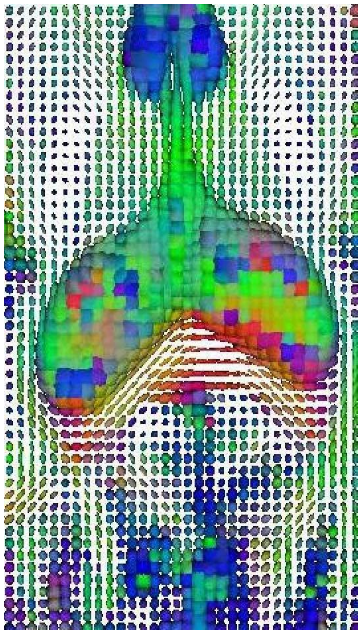
PDEs for regularization and extrapolation: the exponential map (partially) accounts for curvature

- Gradient of Harmonic energy = Laplace-Beltrami $\Delta\Sigma(x) \cong \frac{1}{\varepsilon}\sum_{u\in S}\overrightarrow{\Sigma(x)\Sigma(x+\varepsilon u)}$
- Anisotropic regularization using robust functions $\operatorname{Reg}(\Sigma) = \int \Phi\left(\|\nabla\Sigma(x)\|^2_{\Sigma(x)}\right)dx$
- Simple intrinsic numerical schemes thanks the exponential maps!

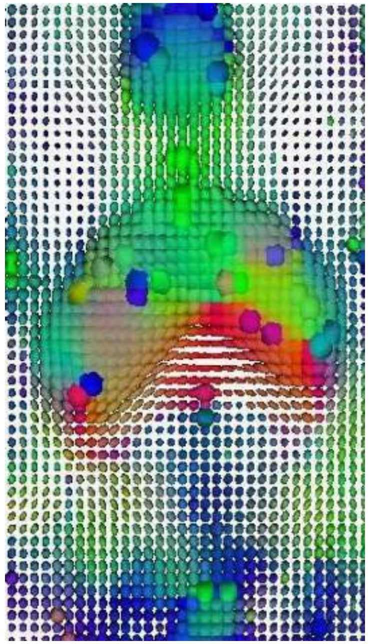
[Pennec, Fillard, Arsigny, IJCV 66(1), 2005, ISBI 2006]

Riemannian algorithms on SPD matrices

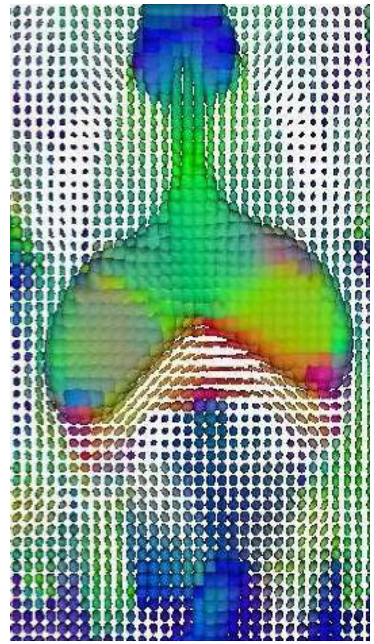
Impact of geometry on data analysis



Raw estimation



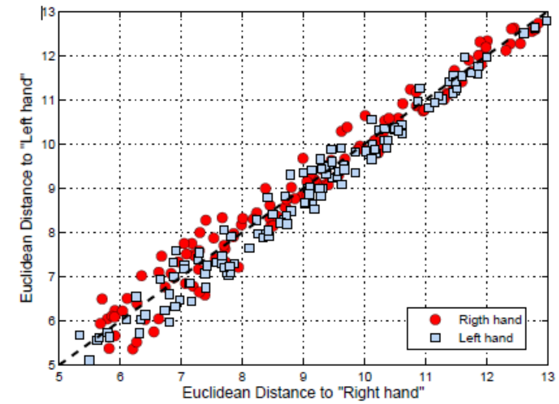
Euclidean regularization

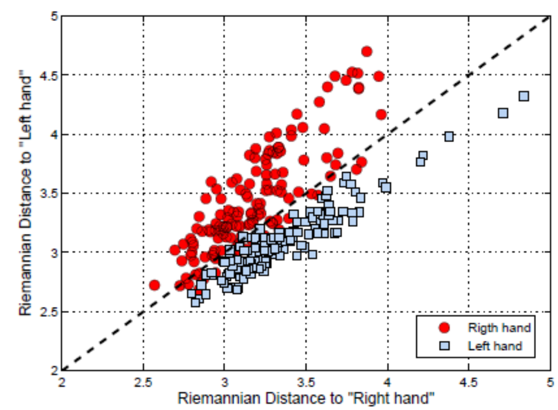


Affine-invariant anisotropic Riemannian regularization

Regularization of a DTI image

[Pennec, Fillard, Arsigny, IJCV 66(1), 2005, ISBI 2006]





Classification in BCI

[Barachant et al. 2012]

Geometric Statistics: Mathematical foundations and applications in computational anatomy

Intrinsic Statistics on Riemannian Manifolds

- Introduction to computational anatomy
- The Riemannian manifold computational structure
- Simple statistics on Riemannian manifolds
- Applications to the spine shape and diffusion tensor imaging
- Conclusion

Metric and Affine Geometric Settings for Lie Groups

Advances Statistics: CLT & PCA

Exp_x / Log_x and Fréchet mean are the basis of algorithms to compute on Riemannian/affine manifolds

Instead of a minimal # of non-linear charts, use a chart per point!

- Normal coordinate system = most linear chart at each point

Simple statistics

- Mean through an exponential barycenter iteration
- Covariance matrices and higher order moments
- Tangent PCA or more complex PGA / BSA

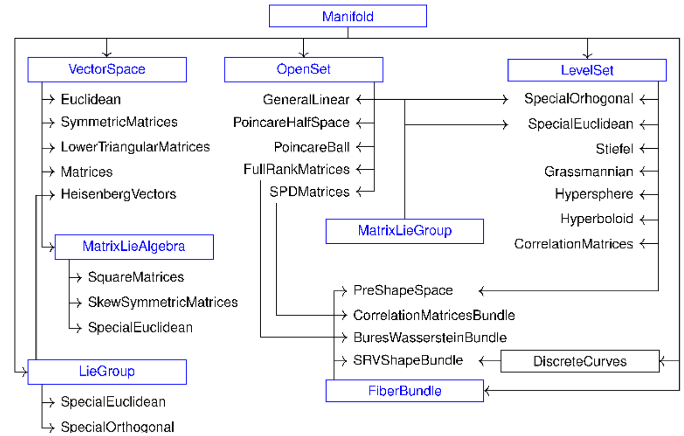
Manifold-valued image processing [XP, IJCV 2006]

- Interpolation / filtering / convolution: weighted means
- Diffusion, extrapolation:
  Discrete Laplacian in tangent space = Laplace-Beltrami

http://geomstats.ai : A Python Package for Geometry in Statistics and Machine Learning

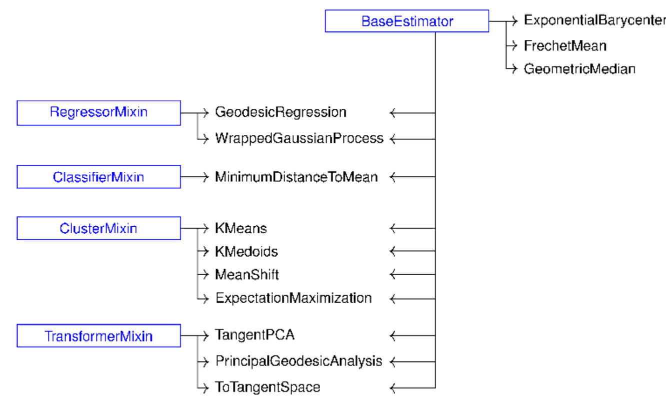
Specific & generic manifolds

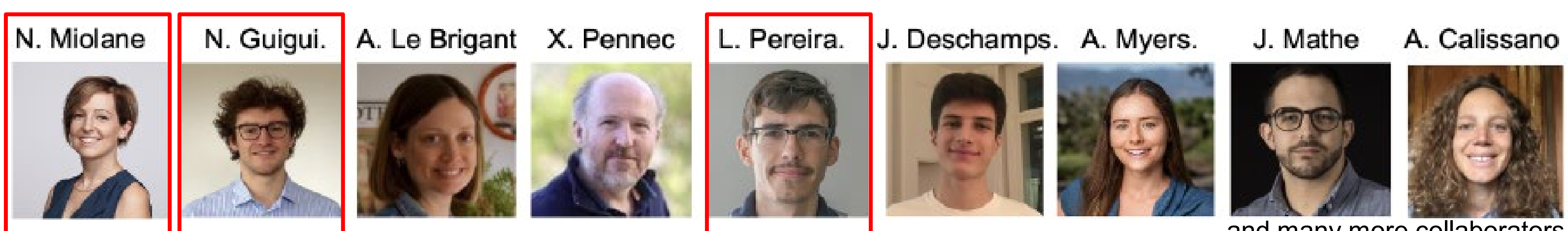
- Exp/Log map to generalize Euclidean tools
- **20+ specific manifolds / Lie groups** with closed-forms (SPD, H(n), SE(n), etc)
- Generic manifolds with **geodesics by integration / optimization**



Algorithms

- **Fréchet mean, geodesic regression, tangent / geodesic PCA, Riemannian k-means, mean-shift, parallel transport**
- scikit-learn API (GPU & learning tools)
- Collaboration with pyriemann for BCI





and many more collaborators

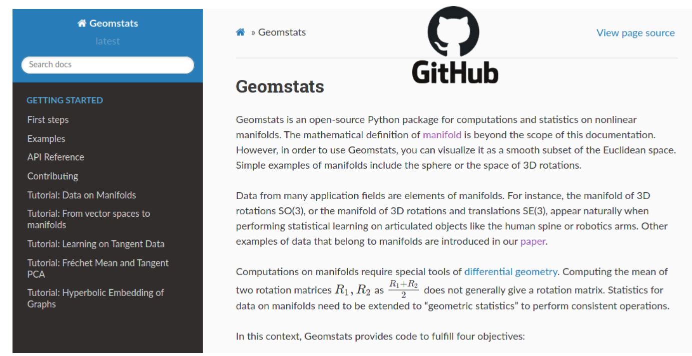
http://geomstats.ai : A Python Package for Geometry in Statistics and Machine Learning

pypi package 2.5.0 | Downloads 93k | DOI 10.5281/zenodo.6478729

Collaborative development

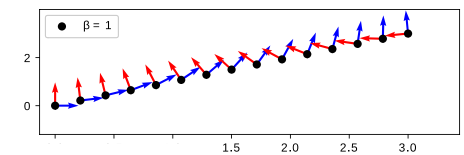
- 10 introductory tutorials
- ~ 35000 lines of code
- ~20 academic developers
- 10 hackathons in 2020-2024

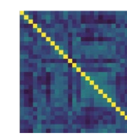
  Semestre thématique IHP Geometry and Statistics in Data Science
  Hackathon IHP Oct 17-21+ Journée Math & entreprises Nov 08, 2022

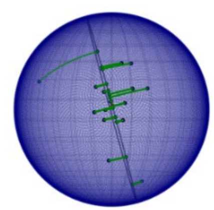


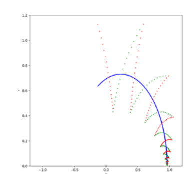
Pushing geometry in Machine Learning

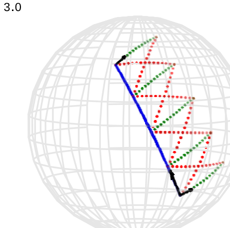
- Miolane, Guigui, et al. SciPy Int. Conf. (2020).
- Miolane et al. **Journal of Machine Learning Research (2020)**
- Guigui, Miolane, Pennec. Intro. to Riem. Geom. and Geom. Stats: from basic theory to implementation with Geomstats. Monography of 164 p. **Foundations and Trends in Machine Learning (2023, 16 (3):329-493).**











N. Miolane | N. Guigui. | A. Le Brigant | X. Pennec | L. Pereira. | J. Deschamps. | A. Myers. | J. Mathe | A. Calissano

and many more collaborators

The G-Statistics group

Yann Thanwerdas

erc

Geometric Statistics

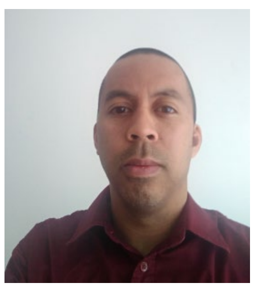
UNIVERSITÉ CÔTE D'AZUR

3iA Côte d'Azur
Interdisciplinary Institute for Artificial Intelligence

Nicolas Guigui

Morten Pedersen

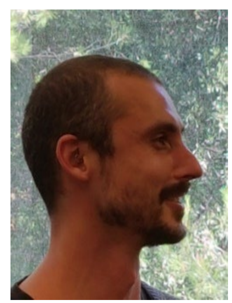
Elodie Maignant



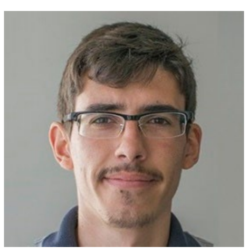
Dimbihery Rabenoro

Olivier Bisson

Anna Calissano



James Benn



Luis G. Pereira

Tom Szwagier

Collaborators

Researchers from Epione / Asclepios / Epidaure team

- Maxime Sermesant
- Nicholas Ayache

Former PhD students

- Jonathan Boisvert
- Pierre Fillard
- Vincent Arsigny
- Kristin McLeod
- Nina Miolane
- Loic Devillier
- Marc-Michel Rohé
- Tom Vercauteren
- Stanley Durrleman
- Marco Lorenzi
- Christof Seiler
- Yann Thanwerdas
- Nicolas Guigui
- Shuma Jia
- Elodie Maignant
- Morten Akhoj Pedersen
- …..

Current PhD students

- Tom Szwagier
- Olivier Bisson

References

1/ Intrinsic Statistics on Riemannian Manifolds

- XP. **Intrinsic Statistics on Riemannian Manifolds: Basic Tools for Geometric Measurements**. *J. Math. Imaging & Vision* 25 (1), 2006, pp.127-154. [DOI: 10.1007/s10851-006-6228-4][Preprint]
- XP, P. Fillard, N. Ayache. **A Riemannian Framework for Tensor Computing**. *Int. J. of Computer Vision*, 66(1):41-66, Jan. 2006. [DOI: 10.1007/s11263-005-3222-z]. [Preprint]
- XP. **Manifold-valued image processing with SPD matrices**. *In Riemannian Geometric Statistics in Medical Image Analysis*, Chap. 3, pp.75-134, Elsevier, 2020. [DOI: 10.1016/B978-0-12-814725-2.00010-8] [Preprint]

2/ Metric and affine geometric settings for Lie groups

- XP, M. Lorenzi. **Beyond Riemannian Geometry - The affine connection setting for transformation groups**. *In Riemannian Geometric Statistics in Medical Image Analysis*, Chap. 5, pp.169-229 RGSMIA. Elsevier, 2020. [DOI: 10.1016/B978-0-12-814725-2.00012-1] [Preprint] .
- XP, M. Lorenzi. **Geodesics, Parallel Transport and One-parameter Subgroups for Diffeomorphic Image Registration**. *Int. J. of Computer Vision* 105(2):111-127. 2013. [DOI: 10.1007/s11263-012-0598-4] [Preprint]
- XP. **Parallel Transport with Pole Ladder: a Third Order Scheme in Affine Connection Spaces which is Exact in Affine Symmetric Spaces**. 2018. [arXiv:1805.11436]
- N. Guigui, XP. **Numerical Accuracy of Ladder Schemes for Parallel Transport on Manifolds**. *Foundations of Computational Mathematics*, 22:757-790, 2022. [DOI: 10.1007/s10208-021-09515-x]

3/ Advanced statistics: central limit theorem and extension of PCA

- XP. **Curvature effects on the empirical mean in Riemannian and affine Manifolds**. 2019. [arXiv:1906.07418]
- XP. **Barycentric Subspace Analysis on Manifolds**. *Annals of Statistics.*46(6A):2711-2746, 2018. [DOI: 10.1214/17-AOS1636]