

THALES



**Géométrie, médiane et diffusion de spectres de fréquences acoustiques :
Ou Comment "médianiser" ou
"diffuser" des morceaux de musiques
*F. Barbaresco***



■ Frédéric BARBARESCO (frederic.barbaresco@thalesgroup.com)

- Expert au département « **Strategy Technology & Innovation** » de la BL « Surface Radar » de la division « Air Systems » du groupe THALES (Domaine d'expertise : les **senseurs radar cognitifs** et le traitement avancé du signal radar, en particulier la refondation du traitement du signal radar sur la géométrie de l'information et la géométrie des matrices de covariance)
 - En charge de la coordination des thèses au sein de la BL
 - Membre du THALES KTD PCC (**Key Technology Domain, Processing Control & Cognition**) présidé par J.F. Marcotorchino (Directeur scientifique de THALES Land & Joint, enseignant à Paris 6)
 - Animateur du groupe **MATHERON** des experts THALES en traitement d'image
 - Membre Senior SEE, médaille Ampère 2007 et président du club SI²D (Signal, Image, Information & Décision) : www.see.asso.fr
 - Président du comité d'organisation de la conférence SEE Internationale COGIS (**COGNitive systems with Interactive Sensors**) avec Daniel Krob (X : Chaire Thales « Systèmes Complexes ») : www.cogis2009.org
 - Organisation de la journée 2009 SEE/Académie des Sciences «**Traitement de l'information : avancées et défis actuels** » (pour les 125 ans de la SEE) : http://www.academie-sciences.fr/conferences/colloques/colloque_html/colloque_27_01_09.htm
 - Membre du comité d'organisation de la conférence internationale MIA (**Mathematic & Image Analysis**) sponsorisée par THALES et présidé par L. Cohen & G. Peyré (CEREMADE, **GDR MSPC**), qui aura lieu à l'Institut Henri Poincaré en Décembre 09 : <http://www.ceremade.dauphine.fr/~peyre/mia09/>
 - Membre du club des clubs des partenaires industriels des GDRs :
 - **GDR ISIS** (Image Signal InformationS) : <http://gdr-isis.org/>
 - **GDR ONDES** : <http://gdr-ondes.lss.supelec.fr/>
 - Projet de REI DGA/MRIS « **Géométrie de l'information pour le traitement des signaux matriciels** » (Prof. Nielsen à l'X, Prof. Rouchon & Angulo à l'Ecole des Mines ParisTech, Prof. Marc Arnaudon à Université de Poitiers, Prof. Snoussi à l'UT Troyes, F. Barbaresco de Thales)
- ### ■ **Activité musicale :**
- **Choriste dans formation amateur de 80 chanteurs (La Brenadienne)**
 - **prochain concert Parisien « Benjamin Britten » (Rejoice in the Lamb, Jubilate Deo, Festival Te Deum op.32, Hymn to Saint Peter)**

THALES



**Géométrie, médiane et diffusion de spectres de fréquences acoustiques :
Ou Comment "médianiser" ou
"diffuser" des morceaux de musiques
*F. Barbaresco***

- La théorie de la Géométrie de l'Information introduite de façon parallèle par Rao et Chentsov, et la Géométrie Symplectique telle qu'introduite par Carl Ludwig Siegel permet de définir une métrique entre matrices de covariance d'une série temporelle (la matrice de covariance définissant le spectre de fréquences acoustiques présentes dans le signal).
- La géométrie naturelle de ces matrices de covariance symétriques (ou hermitiennes) définies positives est une géométrie Riemannienne symétrique. Il s'agit d'un espace métrique à courbure négative. Il est alors possible de calculer de façon explicite la géodésique (plus court chemin sur la variété) entre matrices de covariance.
- Sur la base de travaux du géomètre allemand Herman Karcher, on définit un flot de gradient qui converge vers la médiane de N matrices de covariances (appelée point de Fermat-Weber en Physique) en minimisant le critère égale à la somme des distances géodésiques à l'ensemble des N matrices (ceci étant l'approche classique du barycentre de Fréchet qui converge vers la moyenne lorsque qu'on prend les distances aux carrés).
- Sur la base de ce flot et par analogie avec l'équation de diffusion de Fourier, on définit une équation de diffusion sur un graphe de matrices de covariances et cherchons via l'équation de Campbell-Hausdorff à remonter à l'équation de diffusion à temps continu.
- Le problème de la "médiane dans un espace métrique" est un problème plus vaste qui recouvre des liens avec les sciences sociales par exemple (redécouverte des travaux de Condorcet sur la notion de "vote médian" par J.F. Marcotorchino et P. Michaux).

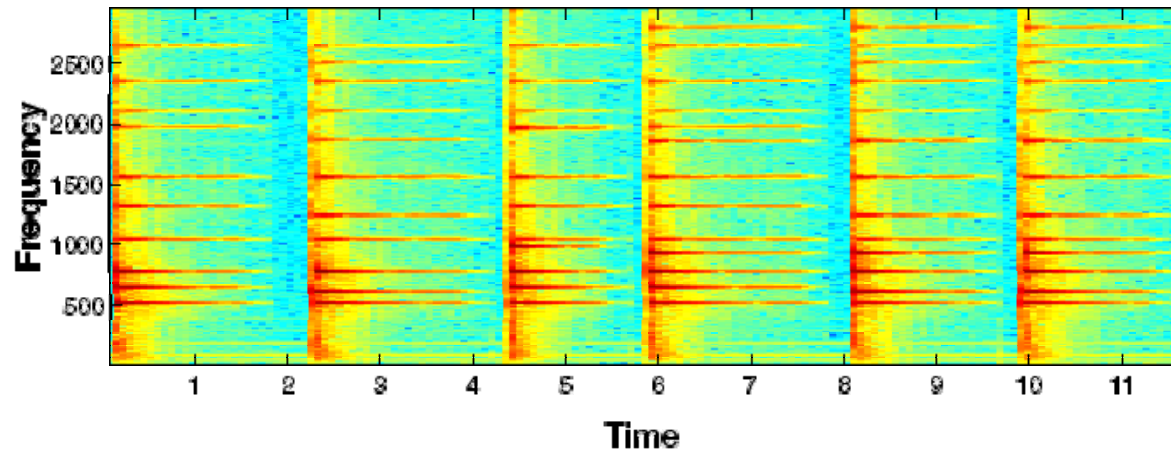
THALES



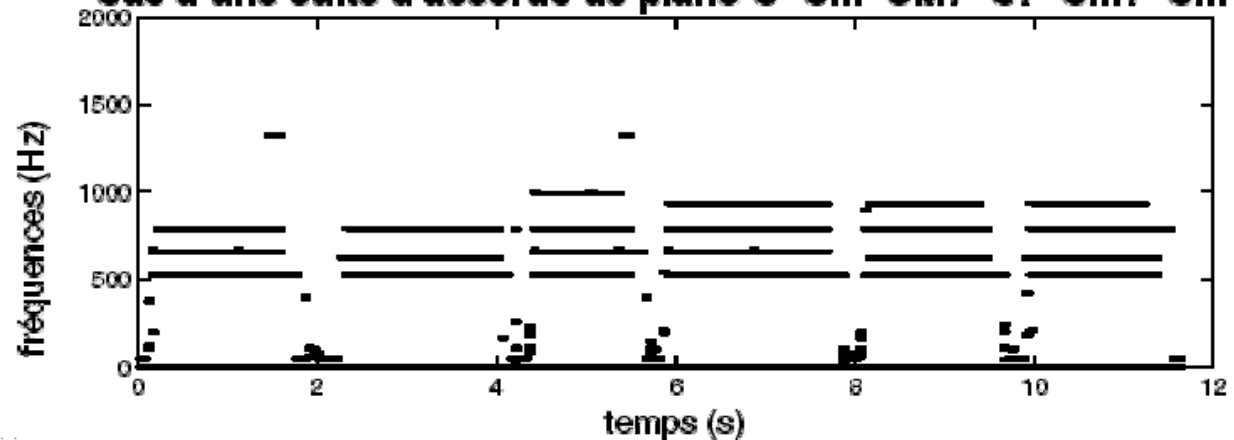
Prise de mesure de l'onde acoustique

■ Spectre d'un signal Acoustique

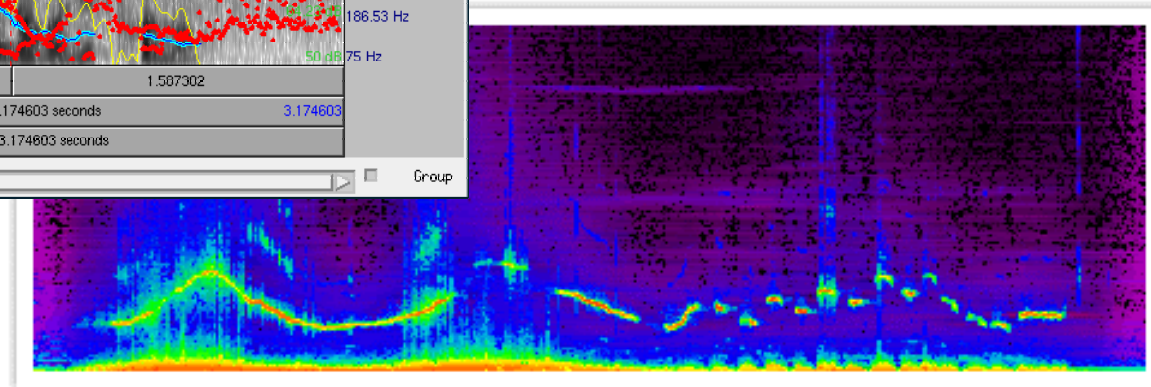
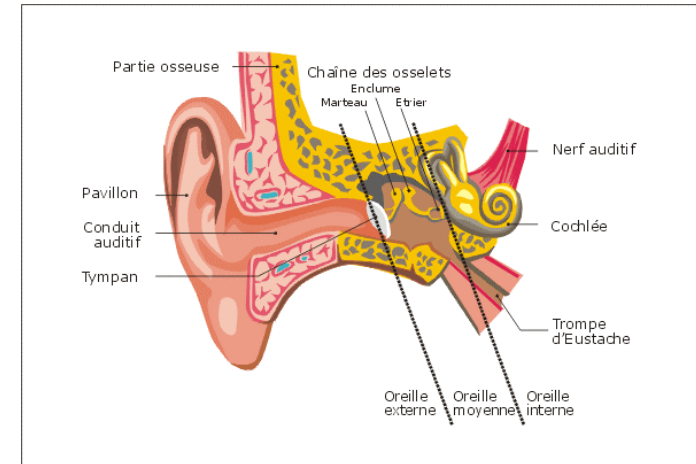
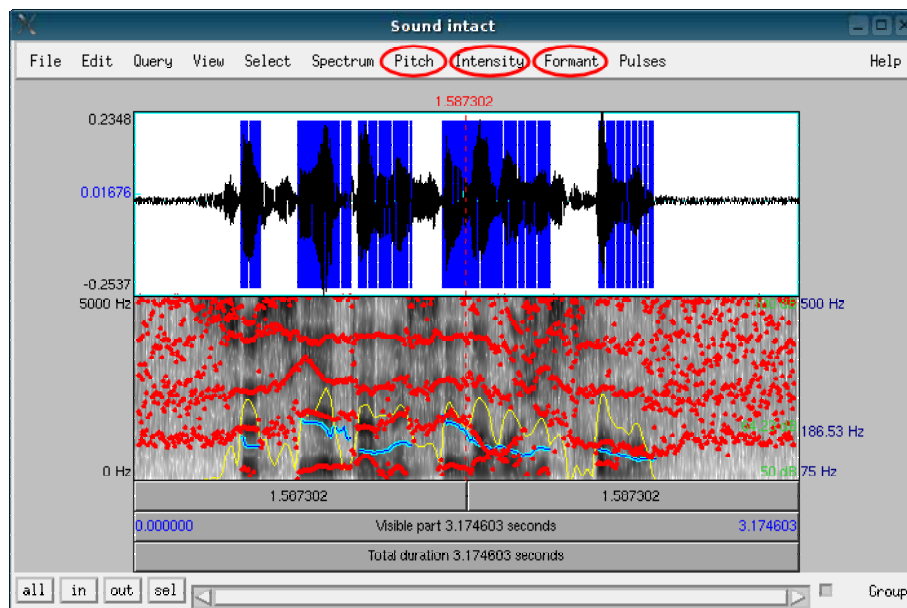
Spectrogramme d'une suite d'accords de piano C-Cm-CM7-C7-Cm7-Cm7



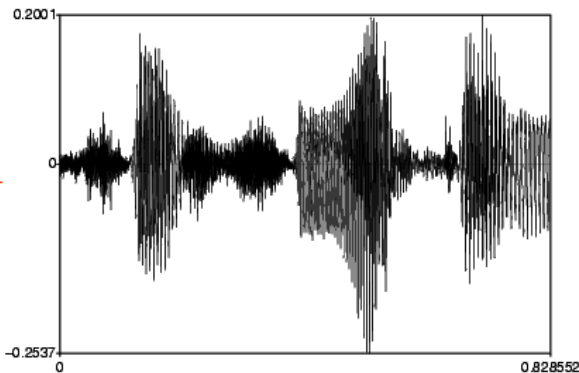
**fondamentales estimées par classification sur le modulo.
Cas d'une suite d'accords de piano C-Cm-CM7-C7-Cm7-Cm7**



■ Spectre d'un signal Acoustique



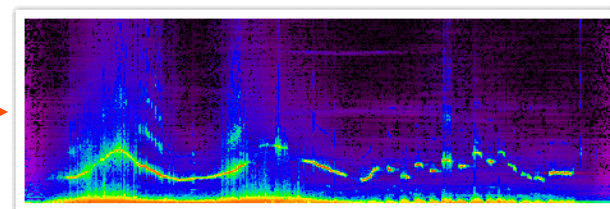
■ Spectre d'un signal Acoustique



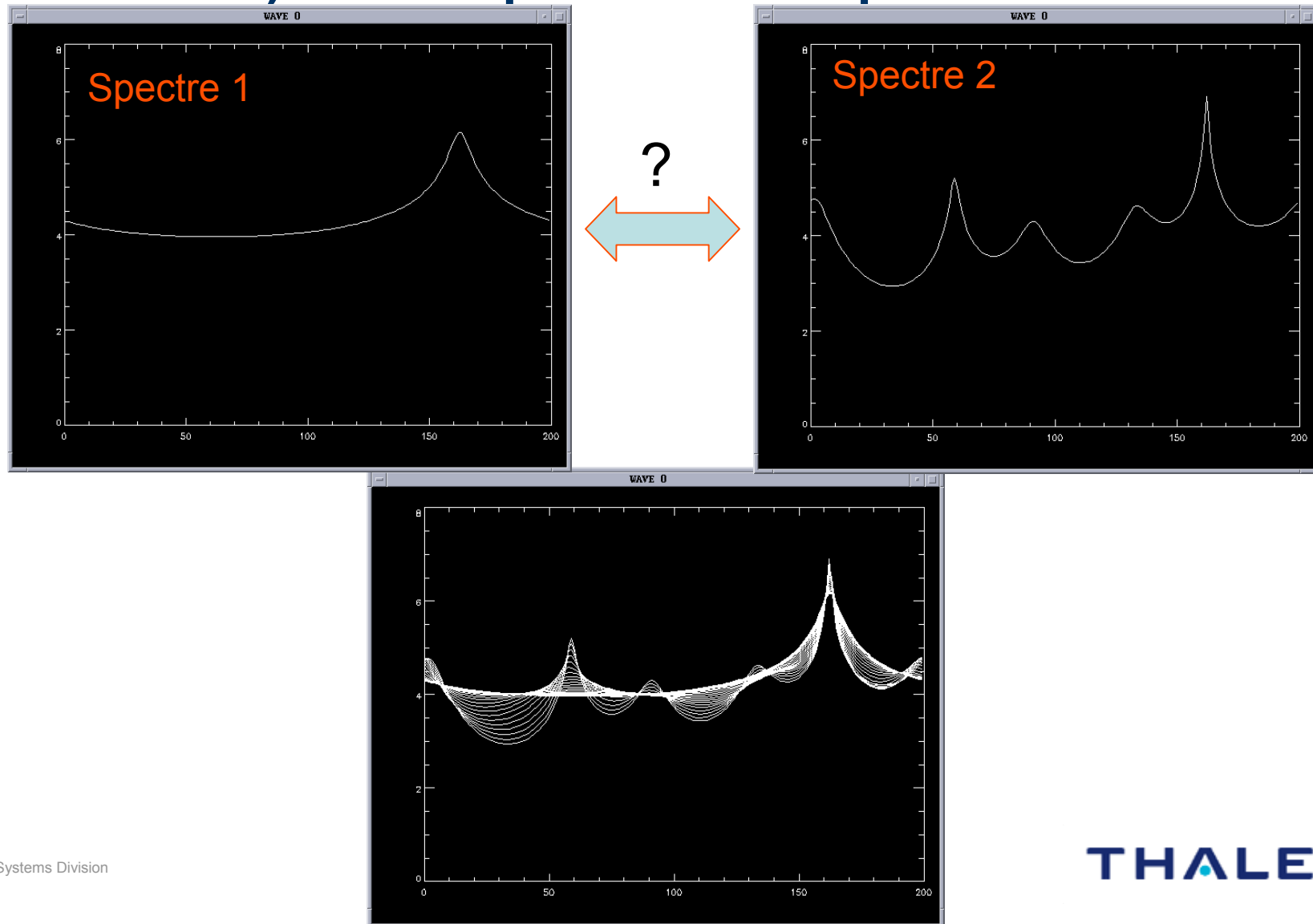
$$Z = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$$

$$R_n = E[ZZ^+] = \begin{bmatrix} c_0 & c_1^* & \cdots & c_{n-1}^* \\ c_1 & c_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_1^* \\ c_{n-1} & \cdots & c_1 & c_0 \end{bmatrix} \quad \text{avec } c_k = E[z_m z_{m-k}^*]$$

$$|S_Z(f)|^2 = \left| \sum_{k=-(n-1)}^{n-1} c_k e^{-j2\pi f k} \right|^2$$



- Comment définir la distance (et donc la géométrie associée) à deux spectres de fréquences ?





« Les animaux se divisent en : a) appartenant à l'Empereur, b) embaumés, c) apprivoisés, d) cochon de lait, e) sirènes, f) fabuleux, g) chiens en liberté, h) inclus dans la présente classification, i) qui s'agitent comme des fous, j) innombrables, k) dessinés avec un pinceau très fin en poil de chameau, l) et cetera, m) qui viennent de casser la cruche, n) qui de loin semblent des mouches »

Encyclopédie chinoise tiré d'un texte de Borges

(En préface de l'ouvrage « Les mots et les choses » de Michel Foucault)

« La monstruosité que Borges fait circuler dans son énumération consiste en ceci que l'espace commun des rencontres s'y trouve ruiné. **Ce qui est impossible, ce n'est pas le voisinage des choses, c'est le site lui-même où elles pourraient voisiner.** ... Les choses y sont 'couchées', 'posées', 'disposées' dans des sites à ce point différents **qu'il est impossible de trouver pour eux un espace d'accueil, de définir au dessous des uns et des autres un lieu commun.** ... Sur quelle 'table', selon quel espace d'identités, de similitudes, d'analogie, avons nous pris l'habitude de distribuer tant de choses différentes et pareilles ? »

Michel Foucault « Les mots et les choses »

« La trame sémantique de la ressemblance au XVIème siècle, est fort riche : Amicitia, Aequalitas (contractus, consensus, matrimonium, societas, pax et similia), Consonancia, Concertus, Continuum, Paritas, Proportio, Similitudo, Conjunctio, Copula .

Mais il y en a quatre qui sont, à coup sûr essentielles :

- La **CONVENIENTIA** : est une ressemblance liée à l'espace dans la forme du 'proche en proche'. Elle est de l'ordre de la conjonction et de l'ajustement. Elle appartient moins aux choses qu'au monde dans lequel elles se trouvent.
- L'**AEMULATIO** : sorte de convenance affranchie de la loi du lieu, qui jouerait immobile dans la distance. Les anneaux ne forment pas une chaîne comme les éléments de la convenance, mais plutôt des cercles concentriques, réfléchis et rivaux.
- L'**ANALOGIE** : affrontement des ressemblances à travers l'espace.
- le **JEUX DES SYMPATHIES** : nul chemin n'est déterminé à l'avance, nulle distance n'est supposée, nul enchaînement prescrit. Elle parcourt en un instant les espaces les plus vastes. Sa figure jumelle, l'antipathie maintient les choses en leur isolement et empêche les assimilations.

Enfin, Il n'y a pas de ressemblance sans signature. Le savoir des similitudes se fonde sur le relevé des signatures et sur leur déchiffrement. »

« Les mots et les choses » de Michel Foucault



■ Problème

- Le spectre des fréquences d'un signal acoustique mesuré via une série temporelle d'échantillons est décrit via la matrice de covariance de ce vecteur de mesure.
- Cette matrice de covariance est structurée et possède les propriétés de symétrie, définie positivité et le caractère Toeplitz
- Si l'on cherche la « structure géométrique » inhérente à ces matrices de covariances, les sources sont doubles

■ Les doubles sources de la géométrie des matrices de covariance

- **Source de la Géométrie de l'Information (Rao/Chentsov) :** Si on modélise le signal par une loi multivariée gaussienne de moyenne nulle, la métrique riemannienne est donnée par la matrice d'information de Fisher. Elle possède une propriété d'invariance par changement non-singulier de paramétrisation. Elle permet d'interpréter la métrique comme une prise en compte de la variance de la matrice de covariance qui déforme l'espace.
- **Source de la Géométrie Symplectique (C. L. Siegel) :** Si on cherche la géométrie intrinsèque à l'espace des matrices symétriques (hermitiennes) définies positives, il peut être introduit naturellement via l'espace de Siegel (généralisation au cas matriciel de l'espace de Poincaré). Dans cette approche, on cherche les automorphismes de cet espace et la métrique qui est invariante par ces automorphismes. Nous verrons que la métrique de la géométrie de l'information est un cas particulier.



- **Objectifs** : Ayant définie la métrique associée à ces matrices de covariances via les 2 sources de la géométrie de l'information et la géométrie intrinsèque, on s'intéresse aux problèmes suivants :
 - Définition de la distance entre 2 matrices de covariances
 - Définition de la géodésique entre 2 matrices de covariances
 - Caractérisation du fait qu'il s'agit d'un espace symétrique à courbure négative
 - Calcul de la moyenne/du barycentre de N matrices de covariance
 - Calcul de la médiane (Point de Fermat-Weber) de N matrices de covariance
 - Équation de diffusion de Fourier d'un graphe de Matrices
 - Équation de diffusion anisotrope d'un graphe de Matrices
 - Introduction d'une paramétrisation autorégressive pour conserver le caractère Toeplitz de la matrice de covariance
 - Modélisation variationnelle de la modélisation autorégressive
 - Calcul de moyenne et médiane de N coefficients de réflexion dans le disque de Poincaré

Métrieque et géométrie pour les matrices de covariance

- La géométrie de l'information pour les lois multivariées gaussiennes de moyenne nulle et la géométrie différentielle intrinsèque des matrices hermitiennes définies positives (cas particulier de la géométrie des espaces de Siegel) conduit à la même métrique :

- Géométrie de l'information :

$$p(Z_n / R_n) = (\pi)^{-n} \cdot |R_n|^{-1} \cdot e^{-Tr[\hat{R}_n \cdot R_n^{-1}]}$$

avec $\hat{R}_n = (Z_n - m_n) \cdot (Z_n - m_n)^+$

et $E[\hat{R}_n] = R_n$

$$\text{avec } g_{ij}(\theta) = -E \left[\frac{\partial^2 \ln p(Z_n / \theta_n)}{\partial \theta_i \cdot \partial \theta_j^*} \right]$$

$$m_n = 0 \implies ds^2 = Tr \left((R_n^{-1} (dR_n))^2 \right)$$

- Géométrie du demi-plan supérieur de C.L. Siegel :

$$SH_n = \{ Z = X + iY \in Sym(n, C) / \text{Im}(Z) = Y > 0 \}$$

$$ds_{Siegel}^2 = Tr \left(Y^{-1} (dZ) Y^{-1} (d\bar{Z}) \right) \text{ avec } Z = X + iY$$

$$\begin{cases} X = 0 \\ Y = i \cdot R_n \end{cases} \implies ds^2 = Tr \left((R_n^{-1} (dR_n))^2 \right)$$

Bonne métrique au sens de la géométrie de l'information



- La métrique est définie : $ds^2 = Tr\left(\left(R_n^{-1}(dR_n)\right)^2\right)$
 - Avantage vu de la Géométrie de l'information :
 - métrique invariante par changement non singulier de paramétrage

$$w = W(\theta) \Rightarrow ds^2(w) = ds^2(\theta)$$

- Métrique tenant compte de la statistique des paramètres

$$R_\theta = E\left[\left(\theta - \hat{\theta}\right)\left(\theta - \hat{\theta}\right)^+\right] \geq I[\theta]^{-1} : \text{borne de Cramer - Rao}$$

$$\text{et } ds^2 = d\theta^+ . I(\theta) . d\theta = d\theta^+ . R_\theta^{-1} . d\theta$$

Bonne métrique au sens de la géométrie des matrices

- La métrique est définie : $ds^2 = Tr\left(\left(R_n^{-1}(dR_n)\right)^2\right)$
- Avantage vu de la Géométrie de Siegel :
 - Les isométries de l'espace de Siegel SH_n sont données par le groupe quotient $PSp(n, R) \equiv Sp(n, R) / \{\pm I_{2n}\}$ avec $Sp(n, F)$ le groupe symplectique :

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \Rightarrow M(Z) = (AZ + B)(CZ + D)^{-1}$$

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in Sp(n, F) \Leftrightarrow \begin{cases} A^T C \text{ et } B^T D \text{ symétrique} \\ A^T D - C^T B = I_n \end{cases}$$

$$Sp(n, F) \equiv \left\{ M \in GL(2n, F) / M^T J M = J \right\}, J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \in SL(2n, R)$$

- La seule métrique invariante par $M(Z)$ est la métrique de Siegel :

$$ds_{Siegel}^2 = Tr\left(Y^{-1}(dZ)Y^{-1}(d\bar{Z})\right) \text{ avec } Z = X + iY$$

$$\begin{cases} X = 0 \\ Y = i.R_n \end{cases} \longrightarrow ds^2 = Tr\left(\left(R_n^{-1}(dR_n)\right)^2\right)$$

- La métrique est définie : $ds^2 = Tr\left(\left(R^{-1}(dR)\right)^2\right)$
- Distance associée :
 - Par intégration, on obtient :

$$d^2(R_1, R_2) = \left\| \log\left(R_1^{-1/2} \cdot R_2 \cdot R_1^{-1/2}\right) \right\|^2 = \sum_{k=1}^n \log^2(\lambda_k)$$

Avec $\det(R_2 - \lambda R_1) = 0$

- Dans le cas général, pour l'espace de Siegel

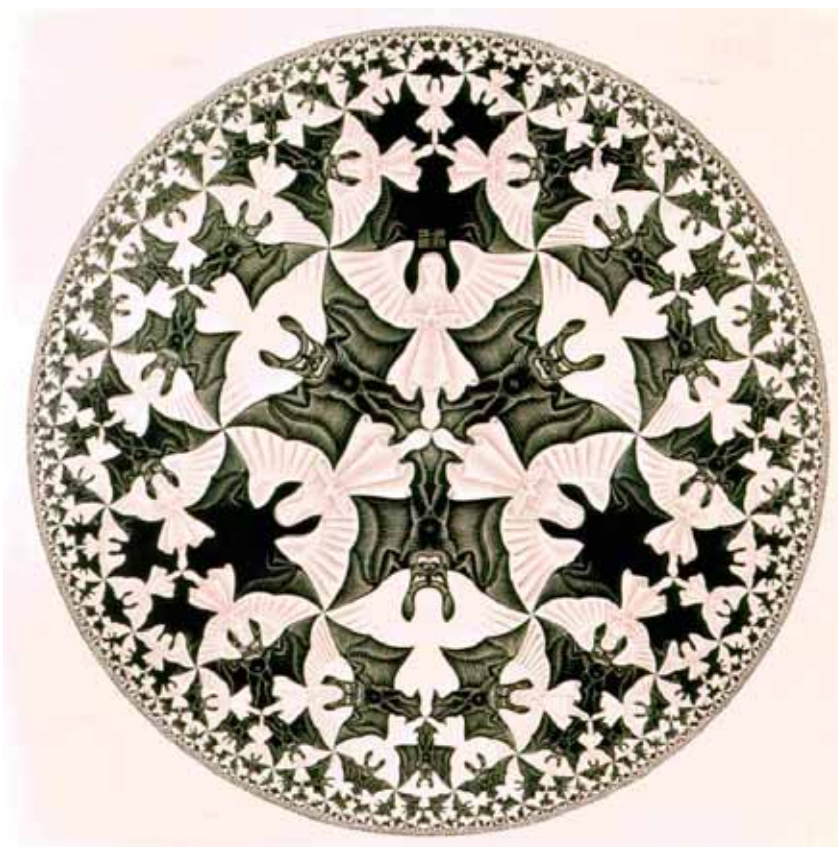
$$Z = X + iY \in SH_n \text{ avec } X \neq 0$$

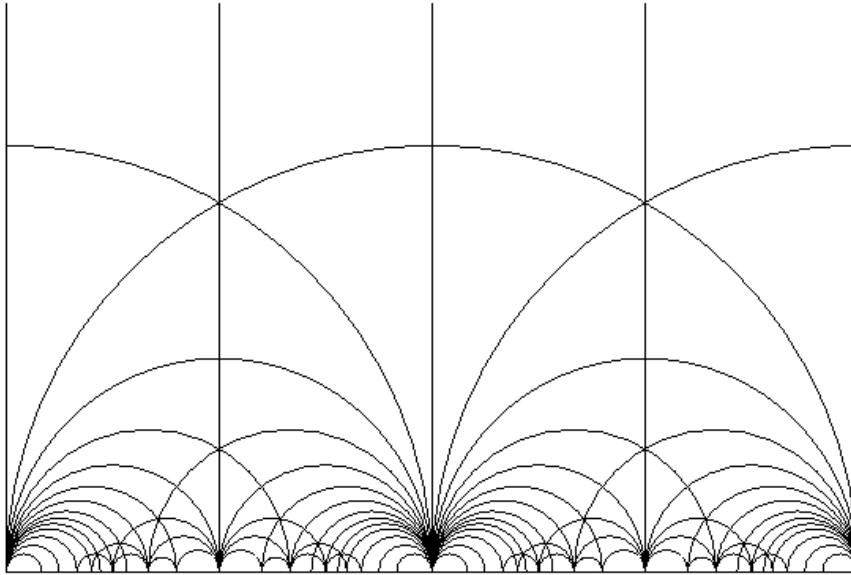
$$d_{Siegel}^2(Z_1, Z_2) = \left(\sum_{k=1}^n \log^2\left(\frac{1 + \sqrt{\lambda_k}}{1 - \sqrt{\lambda_k}}\right) \right) \text{ avec } Z_1, Z_2 \in SH_n$$

Avec $\det(R(Z_1, Z_2) - \lambda \cdot I) = 0$

$$R(Z_1, Z_2) = (Z_1 - Z_2)(Z_1 - \bar{Z}_2)^{-1}(\bar{Z}_1 - \bar{Z}_2)(\bar{Z}_1 - Z_2)^{-1}$$

Utilisation dans l'art de la géométrie hyperbolique (Escher)





Demi-Plan supérieur de Poincare

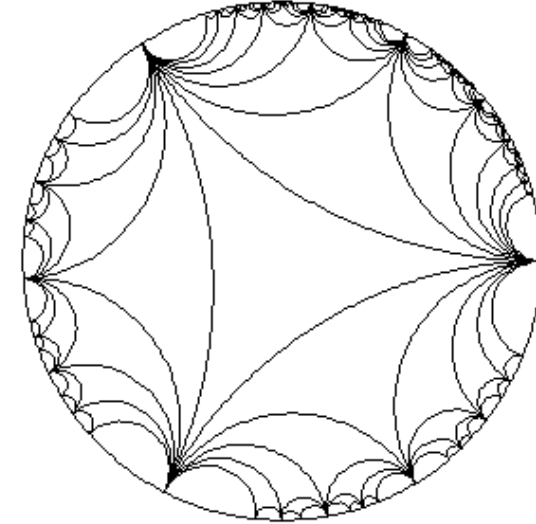
$$ds^2 = \frac{dx^2 + dy^2}{y^2} = \frac{|dz|^2}{y^2} = y^{-1} dzy^{-1} dz^*$$

avec $z = x + iy$

Demi-Plan supérieur de Siegel

$$ds^2 = Tr(Y^{-1} dZY^{-1} dZ^*)$$

avec $Z = X + iY$



Disque de Poincaré

$$ds^2 = \frac{|dz|^2}{(1 - |z|^2)^2}$$

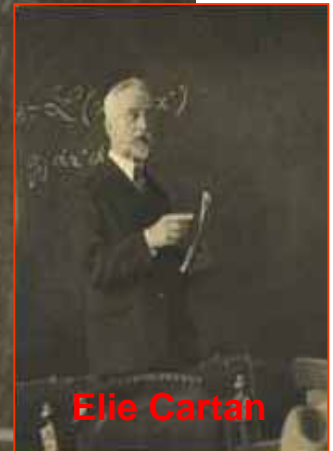
$$ds^2 = (1 - zz^*)^{-1} dz (1 - zz^*)^{-1} dz^*$$

Disque de Siegel

$$ds^2 = Tr \left[(1 - ZZ^*)^{-1} dZ (1 - ZZ^*)^{-1} dZ^* \right]$$

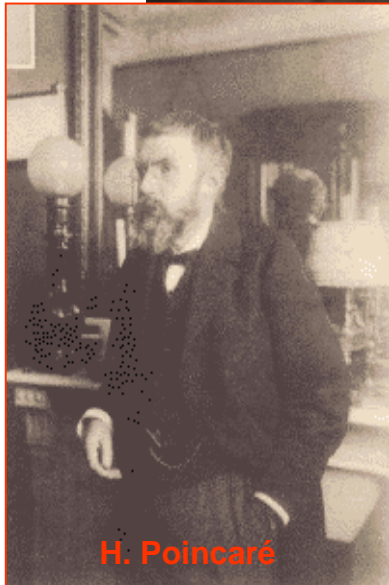


Carl Ludwig Siegel
avec George Polya



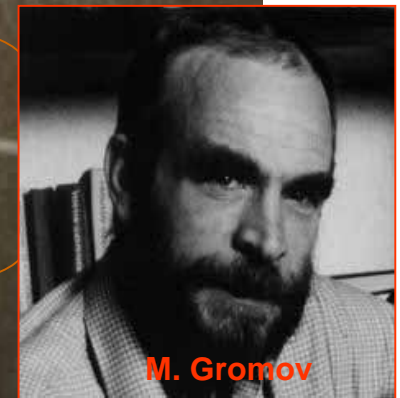
Elie Cartan

$$ds_{Siegel}^2 = Tr\left(Y^{-1}(dZ)Y^{-1}(d\bar{Z})\right) \text{ avec } Z = X + iY$$



H. Poincaré

« At one point Siegel thought that too many unnecessary things were being published, so he decided not to publish anything at all »
George Polya
The Polya Picture Album, Encounters of a Mathematician, Birkäuser



M. Gromov

La géométrie du demi-plan de Siegel et le cas particulier

Demi-Plan de Siegel

$SH_n = \{Z = X + iY \in \text{Sym}(n, \mathbb{C}) / \text{Im}(Z) = Y > 0\}$

$\det(R_1^{-1/2} \cdot R_2 \cdot R_1^{-1/2} - \lambda \cdot I) = 0$

$d^2(R_1, R_2) = \sum_{k=1}^n \log^2(\lambda_k)$

$d^2(Z_1, Z_2) = \sum_{i=1}^n \log^2\left(\frac{1 + \sqrt{\lambda_i}}{1 - \sqrt{\lambda_i}}\right)$

$Z_k = X_k + iY_k$

$Z_k = i \cdot R_k$ si $W_k \equiv N(0, R_k)$

$ds^2 = \text{Trace}[(R^{-1}dR)^2]$

$ds^2 = \text{Tr}(Y^{-1}(dZ)Y^{-1}(d\bar{Z}))$

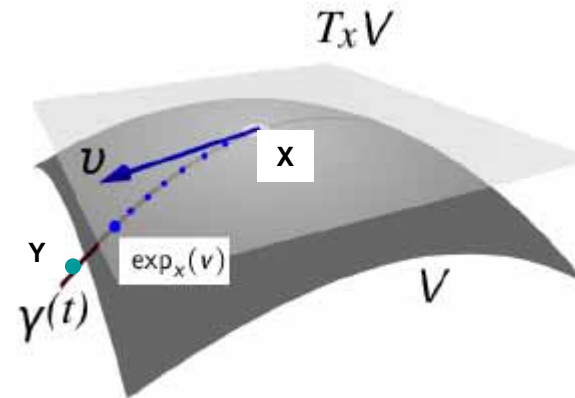
$R(Z_1, Z_2) = (Z_1 - Z_2)(Z_1 - \bar{Z}_2)^{-1}(\bar{Z}_1 - \bar{Z}_2)(\bar{Z}_1 - Z_2)^{-1}$

$\det(R(Z_1, Z_2) - \lambda \cdot I) = 0$

Géodésique et moyenne géométrique symétrisée

$$\begin{cases} v_X^Y = \text{grad}_X^Y(V) = -\exp_X^{-1}(V) = -X^{1/2} \log(X^{-1/2} Y X^{-1/2}) X^{1/2} \\ \exp_X(v_X^Y, t) = X^{1/2} e^{-t(X^{-1/2} \cdot v_X^Y \cdot X^{-1/2})} X^{1/2} \end{cases}$$

$$\Rightarrow \exp_X(v_X^Y, t) = X^{1/2} e^{t \log(X^{-1/2} Y X^{-1/2})} X^{1/2}$$



- La géodésique entre 2 matrices \$X\$ et \$Y\$ est donnée par :

$$\gamma(t) = X^{1/2} e^{t \log(X^{-1/2} Y X^{-1/2})} X^{1/2} = X^{1/2} (X^{-1/2} Y X^{-1/2})^t X^{1/2}$$

$$\text{avec } t \in [0,1] \quad \gamma(0) = X \quad \text{et} \quad \gamma(1) = Y$$

- La moyenne est définie comme barycentre au sens de Fréchet :

$$X \circ Y = \gamma(1/2) = X^{1/2} (X^{-1/2} Y X^{-1/2})^{1/2} X^{1/2}$$

moyenne géométrique symétrisée

- La moyenne qui apparaît est bien une matrice symétrique ou hermitienne définie positive :

$$A \circ B = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^{1/2} A^{1/2}$$

- Il s'agit bien du barycentre de Fréchet, qui minimise :

$$\text{Min}_X \left(\left\| \log \left(A^{-1/2} X A^{-1/2} \right) \right\|^2 + \left\| \log \left(B^{-1/2} X B^{-1/2} \right) \right\|^2 \right)$$

- On peut remarquer que quand les matrices commutent, il s'agit de la moyenne géométrique classique :

$$A \circ B = (AB)^{1/2} \text{ qui n'est pas symétrique définie positive}$$

- Une autre façon de retrouver la moyenne géométrique « symétrisée » consiste à la trouver comme solution de l'équation de Riccati :

$$X A^{-1} X = B \Rightarrow A^{-1/2} (X A^{-1/2} A^{-1/2} X) A^{-1/2} = A^{-1/2} B A^{-1/2}$$

$$\Rightarrow (A^{-1/2} X A^{-1/2}) (A^{-1/2} X A^{-1/2}) = A^{-1/2} B A^{-1/2}$$

$$\Rightarrow A^{-1/2} X A^{-1/2} = \left(A^{-1/2} B A^{-1/2} \right)^{1/2} \Rightarrow X = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^{1/2} A^{1/2}$$

Il s'agit d'un espace Riemannien symétrique



- En géométrie riemannienne, un espace localement symétrique est une variété riemannienne (M,g) telle que, (localement) autour de chaque point x , il existe une symétrie σ_x qui inverse les géodésiques issues de x et qui est une isométrie (locale).
- Un espace localement symétrique est dit symétrique si les symétries peuvent se prolonger à tout l'espace. De manière équivalente, une variété riemannienne (M,g) est dite symétrique lorsque, pour tout point x de M , il existe une isométrie $\sigma_x:M\rightarrow M$ vérifiant :
 - $\sigma_x(x) = x$;
 - $d\sigma_x(x) = -\text{Id}$.
- Cette isométrie σ_x est appelée l'involution en x .
- Propriétés :
 - Tout espace symétrique est une variété riemannienne géodésiquement complète, donc complète en vertu du théorème de Hopf-Rinow.
 - Il existe une et une unique involution en x .
 - Le tenseur de courbure d'une variété riemannienne est parallèle.

Il s'agit d'un espace Riemannien symétrique

- Pour l'espace considéré, la géodésique entre deux matrices A et B est donnée par :

$$\gamma(t) = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^t A^{1/2} = A^{1/2} e^{t \cdot \log(A^{-1/2} B A^{-1/2})} A^{1/2} \text{ avec } t \in [0,1]$$

- Pour l'espace symétrique des matrices symétriques ou hermitiennes définies positives, pour chaque paire (A,B), il existe une isométrie bijective $G_{(A,B)}$ qui vérifie :

$$G_{(A,B)} A = B \quad \text{et} \quad G_{(A,B)} B = A$$

- Cette isométrie a un unique point fixe Z qui est donné par le milieu de (A,B), c'est à dire la moyenne géométrique précédemment définie $d(G_{(A,B)} X, X) = 2d(X, Z)$:

$$G_{(A,B)} X = (A \circ B) X^{-1} (A \circ B) \text{ avec } A \circ B = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^{1/2} A^{1/2}$$

Et les propriétés : $G_{(A,B)}(A \circ B) = (A \circ B)$ et $dG_{(A,B)}(A \circ B) = -I$

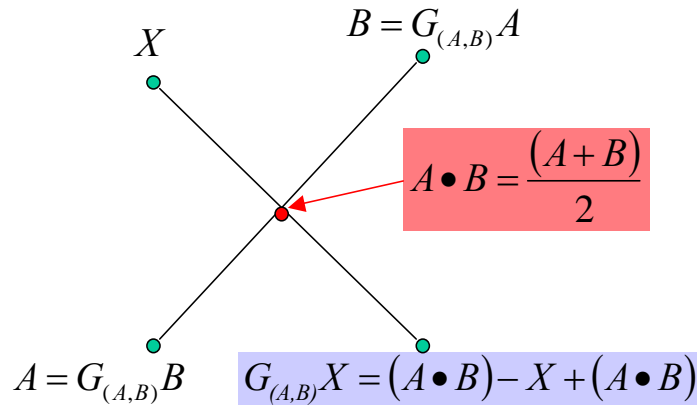
- Ceci est différent de l'approche classique en traitement du signal qui suppose un espace normé euclidien :

$$G_{(A,B)} X = (A \bullet B) - X + (A \bullet B)$$

avec $A \bullet B = \frac{A + B}{2}$ et $\|A - B\|_F$

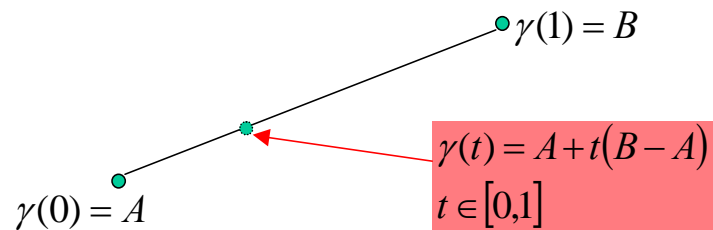
Géodésique et moyenne géométrique symétrisée

Espace Normé : isométrie
(approche traitement du signal Classique)

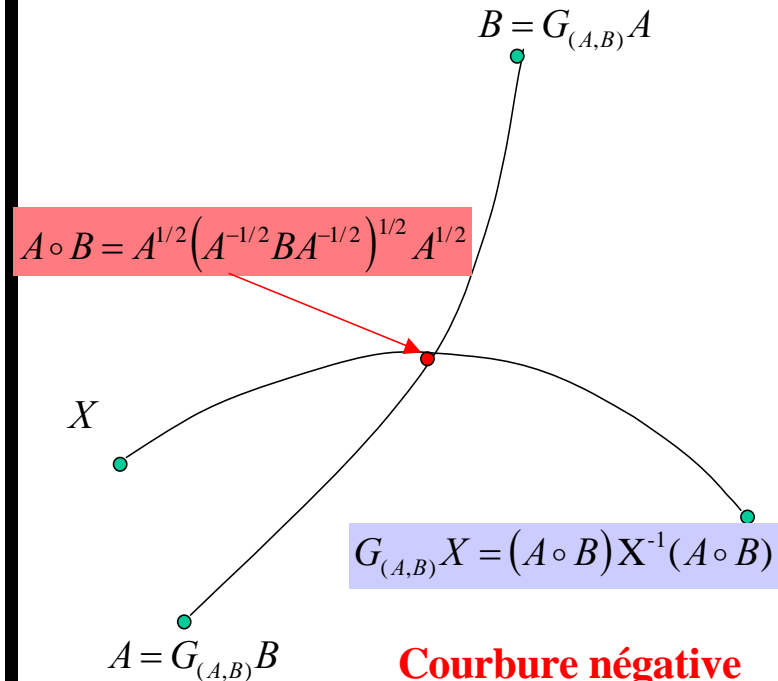


Courbure nulle

Espace Normé : Géodésique

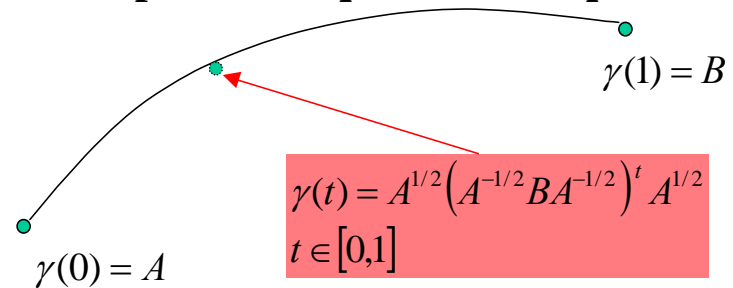


Espace métrique : isométrie



Courbure négative

Espace métrique : Géodésique



« La Messe est dite » par Marcel Berger

Marcel Berger (IHES), « 150 ans de Géométrie Riemannienne »,
Géométrie au 20^{ième} siècle, Histoire et horizons, Hermann Éditeur, 2005



• Premier Miracle :

La théorie des espaces symétriques peut être considérée comme le premier miracle de la géométrie riemannienne, en fait comme un nœud de forte densité dans l'arbre de toutes les mathématiques ... On doit à Elie Cartan dans les années 1926 d'avoir découvert que ces géométries sont, dans une dimension donnée, en nombre fini, et en outre toutes classées.

• Second Miracle :

Entre les variétés localement symétriques et les variétés riemanniennes générales, il existe une catégorie intermédiaire, celle des variétés kählériennes. ... On a alors affaire pour décrire le panorama des métriques kählériennes sur notre variété, non pas à un espace de formes différentielles quadratiques, très lourd, mais à un espace vectoriel de fonctions numériques. ... La richesse Kählerienne fait dire à certains que la géométrie kählérienne est plus importante que la géométrie riemannienne.

• Pas d'espoir d'autre miracle :

Ne cherchez pas d'autres miracles du genre des espaces (localement) symétriques et des variétés kählériennes. En effet, c'est un fait depuis 1953 que les seules variétés riemanniennes irréductibles qui admettent un invariant par transport parallèle autre que g elle-même (et sa forme volume) sont les espaces localement symétriques, les variétés kählériennes, les variétés kählérienne de Calabi-Yau, et les variétés hyperkähleriennes.

Plus de détail : Marcel Berger, « A Panoramic View of Riemannian Geometry », Springer 2003

Algorithmes de la moyenne/Médiane : critère à optimiser

- Problème à résoudre « Barycentre de N matrices » $\{B_k\}_{k=1}^N$:

$$X = \arg \min_X f_\alpha(X) = \arg \min_X \sum_{k=1}^N d^\alpha(X, B_k)$$

- Moyenne barycentrique

$$\alpha = 2 \quad \longrightarrow \quad X = \arg \min_X f_2(X) = \arg \min_X \sum_{k=1}^N d^2(X, B_k)$$

$$\text{grad}[f_2(X)] = -\sum_{k=1}^N \exp_X^{-1}(B_k) \in T_X V \quad \longrightarrow \quad X_{n+1} = \exp_{X_n}(-\varepsilon \cdot \text{grad}[f_2(X_n)])$$

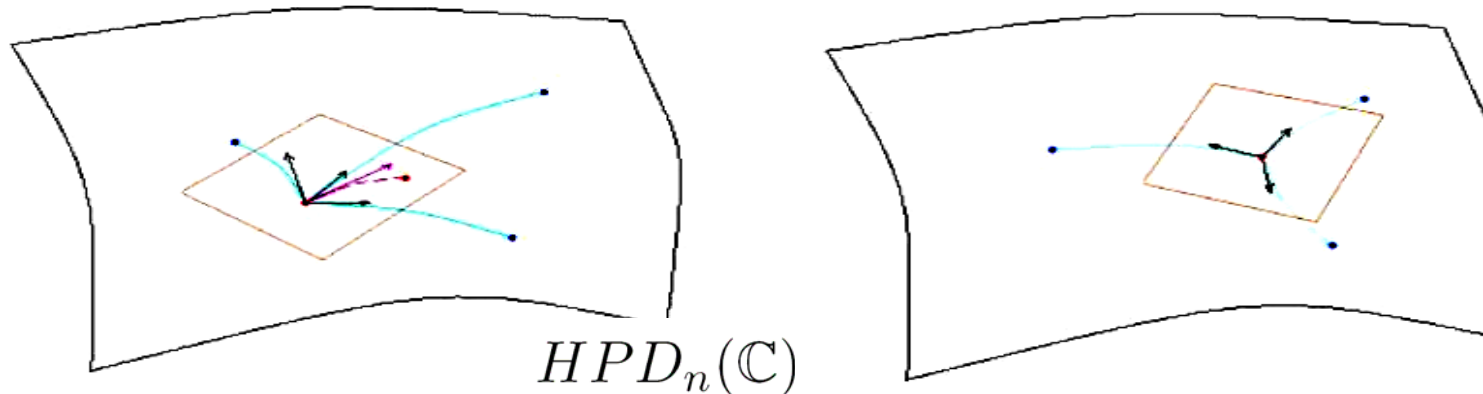
- Mediane (Point de Fermat-Weber)

$$\alpha = 1 \quad \longrightarrow \quad X = \arg \min_X f_1(X) = \arg \min_A \sum_{k=1}^N d(X, B_k)$$

$$\text{grad}[f_1(X)] = -\sum_{k=1}^N \frac{\exp_X^{-1}(B_k)}{d(X, B_k)} \in T_X V \quad \longrightarrow \quad X_{n+1} = \exp_{X_n}(-\varepsilon \cdot \text{grad}[f_1(X_n)])$$

Il s'agit d'un flot de gradient sur la variété

- Algorithme par descente de gradient



- La tangente en X à la géodésique rejoignant X à B_k

$$\gamma_k(t) = X^{1/2} \left(X^{-1/2} B_k X^{-1/2} \right)^{t_k} X^{1/2} = X^{1/2} e^{t_k \log(X^{-1/2} B_k X^{-1/2})} X^{1/2}$$

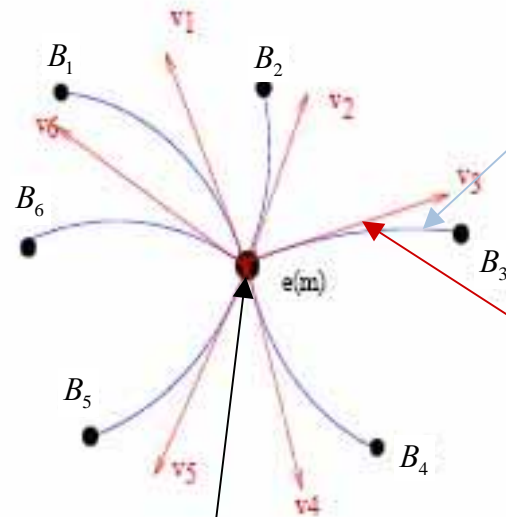
$$\Rightarrow \left. \frac{d\gamma_k(t)}{dt_k} \right|_{t_k=0} = X^{1/2} \log(X^{-1/2} B_k X^{-1/2}) X^{1/2}$$

- La somme des tangentes en X des géodésiques de X aux $\{B_k\}_{k=1}^N$

$$G_X = \sum_{k=1}^N \left. \frac{d\gamma_k(t)}{dt_k} \right|_{t=0} \Rightarrow G_X = X^{1/2} \left(\sum_{k=1}^N \log(X^{-1/2} B_k X^{-1/2}) \right) X^{1/2}$$

Karcher Barycenter & Jacobi Field

$$\gamma_k(t) = X^{1/2} \left(X^{-1/2} B_k X^{-1/2} \right)^t X^{1/2} = X^{1/2} e^{t \log \left(X^{-1/2} B_k X^{-1/2} \right)} X^{1/2}$$



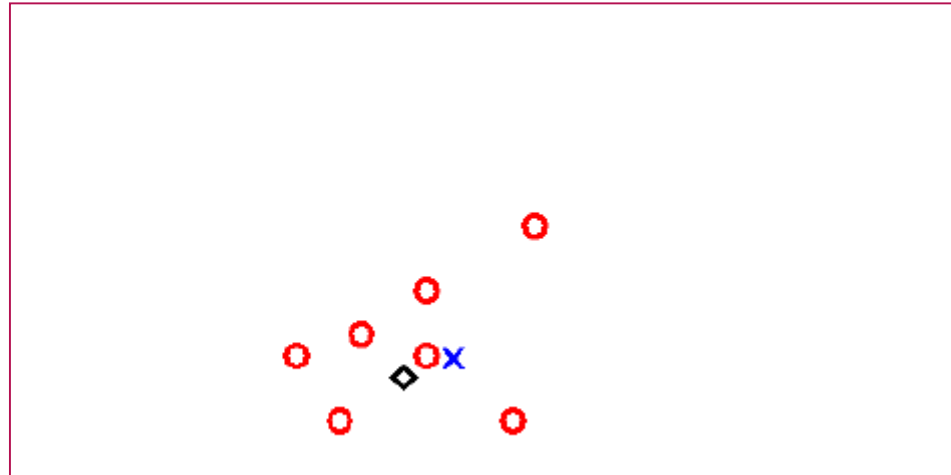
$$\left. \frac{d\gamma_k(t)}{dt} \right|_{t=0} = X^{1/2} \log \left(X^{-1/2} B_k X^{-1/2} \right) X^{1/2}$$

$$\sum_{k=1}^N \left. \frac{d\gamma_k(t)}{dt} \right|_{t=0} = X^{1/2} \left(\sum_{k=1}^N \log \left(X^{-1/2} B_k X^{-1/2} \right) \right) X^{1/2} = 0$$

Médiane plus robuste que moyenne

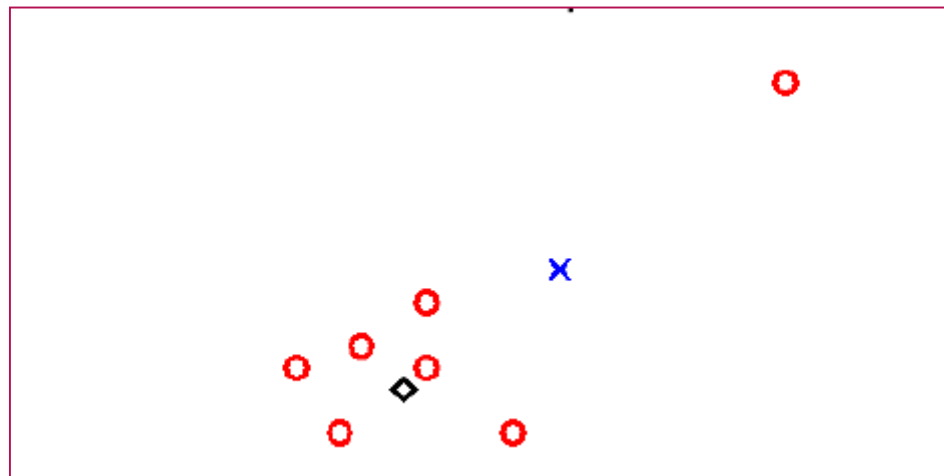
X : moyenne

◇ : médiane



La médiane est insensible à la présence de points aberrants tant qu'ils restent minoritaires

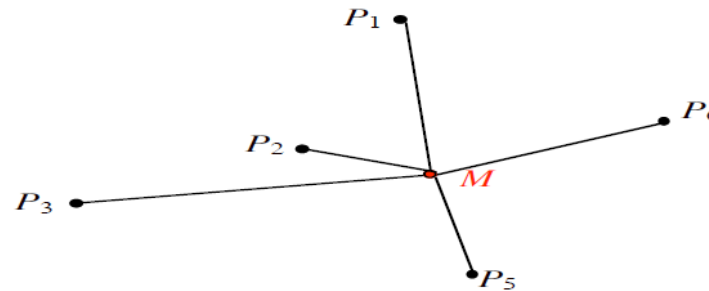
La moyenne est immédiatement perturbée par la présence de points aberrants



Mediane et point de Fermat-Weber



Let $P_1, P_2 \dots P_n$ be n points of a metric space. A point M is a *median* of these n points if M is a point of the space minimizing the sum $d(M, P_1) + d(M, P_2) + \dots + d(M, P_n)$ of its distances to the n points.



The median of six points



$$\int_0^{\text{median}} P_X(x).dx = 0.5 \Leftrightarrow \text{median} = \underset{m}{\text{Min}} E(|x - m|)$$

Laplace (1774) :

Laplace a appelé cette valeur « **le milieu de probabilité** » ou « **la valeur probable** ». Le terme médiane a été introduit par Cournot dans *l'Exposition de la théorie des chances* en 1883.

$$\alpha = 1 \quad \longrightarrow \quad X = \arg \min_X f_1(X) = \arg \min_X \sum_{k=1}^N d(X, B_k)$$

$$\text{grad}[f_1(X)] = - \sum_{k=1}^N \frac{\exp_X^{-1}(B_k)}{d(X, B_k)} \in T_X V \quad \longrightarrow \quad X_{n+1} = \exp_{X_n}(-\varepsilon \cdot \text{grad}[f_1(X_n)])$$

$$\exp_X^{-1}(U) = X^{1/2} \log(X^{-1/2} U X^{-1/2}) X^{1/2}$$

$$\exp_X(V) = X^{1/2} e^{X^{-1/2} V X^{-1/2}} X^{1/2} \quad \longrightarrow$$

$$X_{n+1} = X_n^{1/2} e^{\varepsilon \sum_{k=1}^N \frac{\log(X_n^{-1/2} B_k X_n^{-1/2})}{\|\log(X_n^{-1/2} B_k X_n^{-1/2})\|_F}} X_n^{1/2}$$

avec $\|\log(X_n^{-1/2} B_k X_n^{-1/2})\|_F = \sqrt{\sum_{i=1}^M \log^2(\lambda_i)}$

et $\det(X_n - \lambda B_k) = 0$

Algorithme de gradient non convergent

- Algorithme médian par descente de gradient

$$X_{n+1} = X_n^{1/2} e^{\varepsilon \sum_{k=1}^N \frac{\log(X_n^{-1/2} B_k X_n^{-1/2})}{\left\| \log(X_n^{-1/2} B_k X_n^{-1/2}) \right\|_F}} X_n^{1/2}$$

avec $\left\| \log(X_n^{-1/2} B_k X_n^{-1/2}) \right\|_F = \sqrt{\sum_{i=1}^M \log^2(\lambda_i)}$

et $\det(X_n - \lambda B_k) = 0$

- Non différentiabilité aux points de données : problème de convergence
 - En théorie : régularisation au voisinage des points de données et extraction diagonale (algorithme général adapté de Huling Le, mais long d'exécution)
 - En pratique :
 - Assez souvent, les points donnés « entourent » la médiane : convergence
 - Prendre un pas voisin de la précision voulue (50 itérations suffisent ici)

Mediane : autre algorithme de gradient régularisé

- Autre Algorithme Régularisé proposé

- Critère de minimisation régularisé

$$X_{m+1} = \arg \min_X \sum_{k=1}^N \frac{d^2(X, B_k)}{d(X_m, B_k)}$$

- Résolution itérative de problèmes d'optimisation convexes régularisés, dont la suite des solutions converge vers la médiane.

- Étape m :

$$\begin{cases} X_{m+1,0} = X_m = X_{m,n_{convergence}} \\ X_{m+1,n+1} = X_{m+1,n}^{1/2} e^{\varepsilon \sum_{k=1}^N \frac{\log(X_{m+1,n}^{-1/2} B_k X_{m+1,n}^{-1/2})}{\|\log(X_m^{-1/2} B_k X_m^{-1/2})\|_F}} X_{m+1,n}^{1/2} \end{cases}$$

$$\text{arrêt : } \left\| \log \left(X_{m+1,n_{converge}+1}^{-1/2} X_{m,n_{converge}} X_{m+1,n_{converge}+1}^{-1/2} \right) \right\|_F < \zeta$$

$$\text{alors } X_{m+1} = X_{m+1,n_{converge}}$$

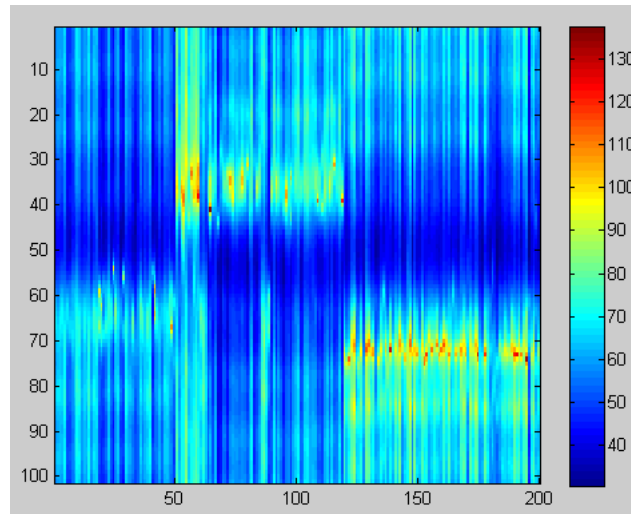
- Critère d'arrêt :

$$\left\| \log \left(X_{m+1}^{-1/2} X_m X_{m+1}^{-1/2} \right) \right\|_F < \zeta$$

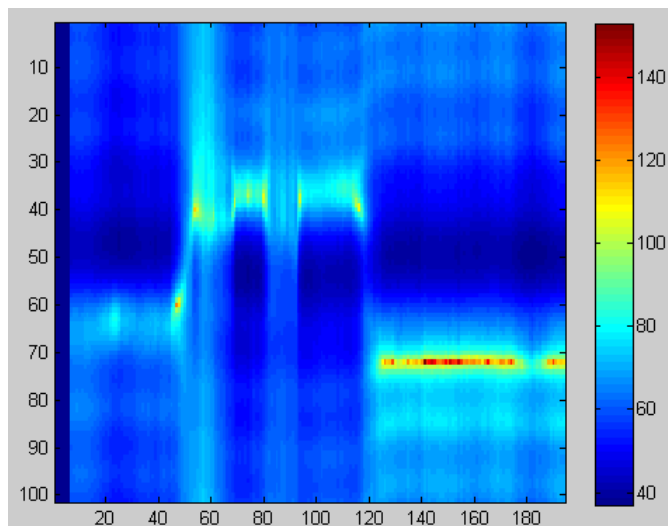
Tests sur données simulées



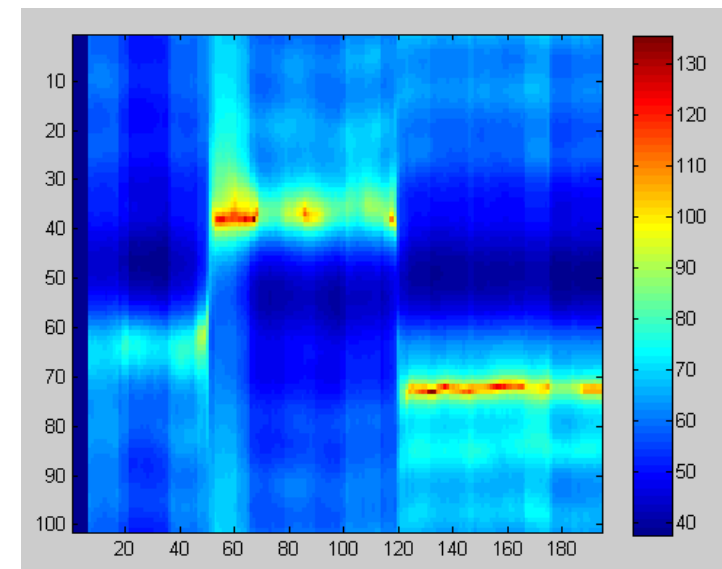
Spectre initial



Spectre moyen



Spectre médian





La géométrie d'Erich Kähler étend la géométrie Riemannienne au domaine complexe :

- La forme riemannienne définie positive définissant la métrique de Kähler est donnée par :

$$ds^2 = 2 \sum_{i,j=1}^n g_{i\bar{j}} \cdot dz^i dz^{\bar{j}}$$

- Condition de Kähler condition : Il existe localement une fonction potentielle de Kähler, Φ (et les équivalentes Pluri-harmoniques) telle que :

$$g_{i\bar{j}} = \frac{\partial^2 \Phi}{\partial z^i \partial \bar{z}^j}$$

- Le tenseur de R est alors donné par :

$$R_{i\bar{j}} = - \frac{\partial^2 \log(\det g_{k\bar{l}})}{\partial z_i \partial \bar{z}_j}$$

- Et la courbure scalaire :

$$R = \sum_{k,l=1}^n g^{k\bar{l}} \cdot R_{k\bar{l}}$$

■ Modèle du signal :

- Modèle multivarié complexe circulaire :

$$p(Z_n / R_n) = (\pi)^{-n} |R_n|^{-1} \cdot e^{-Tr[\hat{R}_n \cdot R_n^{-1}]}$$

$$\text{avec } \hat{R}_n = (Z_n - m_n)(Z_n - m_n)^+ \text{ et } E[\hat{R}_n] = R_n$$

- Modèle Radar :

$m_n = 0$ processus de moyenne nulle

$R_n = E[Z_n Z_n^+]$ Toeplitz Hermitienne Définie Positive

$Z_n = [z_1 \cdots z_n]^T$ avec $z_k = x_k + iy_k = \rho_k e^{i\phi_k}$

- Modèle autorégressif complexe :

$$z_n = -\sum_{k=1}^N a_k^{(N)} z_{n-k} + b_n \text{ avec } E[b_n b_{n-k}^*] = \delta_{k,0} \sigma^2 \text{ et } A_N = [a_1^{(N)} \cdots a_N^{(N)}]^T$$

- Lien avec l'algorithme de Issai Schur's algorithm (1875-1941)

- **[Alpay] D. Alpay, « Algorithme de Schur, espaces à noyau reproduisant et théorie des systèmes », Panoramas et synthèse, n°6, Société Mathématique de France**

Estimation des paramètres autorégressifs régularisés

- Nous utilisons l'algorithme de Burg régularisé (brevet THALES : F. Barbaresco, « Procédé et dispositif de détermination du spectre de fréquence d'un signal », brevet n° 95 06983, Juin 1995)

.Initialisation :

$$f_0(k) = b_0(k) = z(k) \quad , \quad k=1, \dots, N \quad (N: \text{nb. ech.})$$

$$P_0 = \frac{1}{N} \cdot \sum_{k=1}^N |z(k)|^2$$

$$a_0^{(0)} = 1$$

.Itération (n): Pour $n = 1$ à M

$$\mu_n = - \frac{\frac{2}{N-n} \sum_{k=n+1}^N f_{n-1}(k) \cdot b_{n-1}^*(k-1) + 2 \cdot \sum_{k=1}^{n-1} \beta_k^{(n)} \cdot a_k^{(n-1)} \cdot a_{n-k}^{(n-1)}}{\frac{1}{N-n} \sum_{k=n+1}^N |f_{n-1}(k)|^2 + |b_{n-1}(k-1)|^2 + 2 \cdot \sum_{k=0}^{n-1} \beta_k^{(n)} \cdot |a_k^{(n-1)}|^2} \quad \text{avec} \quad \beta_k^{(n)} = \gamma_1 \cdot (2\pi)^2 \cdot (k-n)^2$$

$$\begin{cases} a_0^{(n)} = 1 \\ a_k^{(n)} = a_k^{(n-1)} + \mu_n \cdot a_{n-k}^{(n-1)*} \quad , \quad k=1, \dots, n-1 \\ a_n^{(n)} = \mu_n \end{cases}$$

$$\begin{cases} f_n(k) = f_{n-1}(k) + \mu_n \cdot b_{n-1}(k-1) \\ b_n(k) = b_{n-1}(k-1) + \mu_n^* \cdot f_{n-1}(k) \end{cases}$$

modèle autorégressif et matrice de covariance

- La matrice de covariance et son inverse peuvent être paramétrée à partir des paramètres du modèle autorégressif :
 - Structure Blocs de la matrice de covariance :

$$R_n^{-1} = \begin{bmatrix} \alpha_{n-1} & \alpha_{n-1} \cdot A_{n-1}^+ \\ \alpha_{n-1} \cdot A_{n-1} & R_{n-1}^{-1} + \alpha_{n-1} \cdot A_{n-1} \cdot A_{n-1}^+ \end{bmatrix} \quad R_n = \begin{bmatrix} \alpha_{n-1}^{-1} + A_{n-1}^+ \cdot R_{n-1} \cdot A_{n-1} & -A_{n-1}^+ \cdot R_{n-1} \\ -R_{n-1} \cdot A_{n-1} & R_{n-1} \end{bmatrix}$$

$$\text{avec } \alpha_n^{-1} = [1 - |\mu_n|^2] \alpha_{n-1}^{-1} \quad \text{et} \quad A_n = \begin{bmatrix} A_{n-1} \\ 0 \end{bmatrix} + \mu_n \cdot \begin{bmatrix} A_{n-1}^{(-)} \\ 1 \end{bmatrix} \quad \text{où } V^{(-)} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & \ddots & 0 \\ 1 & 0 & 0 \end{bmatrix} \cdot V^*$$

Récurtivité sur l'ordre de la métrique de Siegel



■ Si on utilise la structure blocks des matrices, on peut calculer la métrique de Siegel à l'ordre n par rapport à l'ordre n-1

■ Equation réursive sur la métrique de Siegel faisait apparaître la matrice de Fisher sur les paramètres autorégressifs:

$$ds_n^2 = Tr\left[\left(R_n \cdot dR_n^{-1}\right)^2\right] = ds_{n-1}^2 + \left(\frac{d\alpha_{n-1}}{\alpha_{n-1}}\right)^2 + \alpha_{n-1} \cdot dA_{n-1}^+ \cdot R_{n-1} \cdot dA_{n-1}$$

■ On peut remarquer que le second terme s'écrit :

$$\left(\frac{d\alpha_{n-1}}{\alpha_{n-1}}\right)^2 + \alpha_{n-1} \cdot dA_{n-1}^+ \cdot R_{n-1} \cdot dA_{n-1} = d \begin{bmatrix} \log \alpha_{n-1} \\ A_{n-1} \end{bmatrix}^+ \cdot \begin{bmatrix} 1 & 0 \\ 0 & \alpha_{n-1} R_{n-1} \end{bmatrix} \cdot d \begin{bmatrix} \log \alpha_{n-1} \\ A_{n-1} \end{bmatrix}$$

■ Ceci s'interprète en terme de géométrie de l'information

$$\Sigma_{\hat{A}_n} = I[A_{n-1}] = \alpha_{n-1} R_{n-1} \quad \xrightarrow{\text{Extension du cas}} \quad Z = \begin{bmatrix} \log \alpha_{n-1} \\ A_{n-1} \end{bmatrix} \begin{bmatrix} \log \alpha_{n-1} \\ A_{n-1} \end{bmatrix}^+ + i \begin{bmatrix} 1 & 0 \\ 0 & \alpha_{n-1} R_{n-1} \end{bmatrix}$$

$\log \alpha_{n-1} \perp A_{n-1}$

**Extension du cas
Gaussien scalaire
(Géométrie de l'information)**



Métrie de Kähler d'un modèle autorégressif complexe



■ Dans le cadre de la géométrie Affine de l'information, la métrique est donnée par le Hessien de l'entropie qui joue ici le rôle du potentiel de Kähler :

- L'entropie d'un modèle multivarié gaussien de moyenne nulle :

$$p(Z_n / R_n) = (\pi)^{-n} \cdot |R_n|^{-1} \cdot e^{-Tr[\hat{R}_n \cdot R_n^{-1}]} \implies \tilde{\Phi}(R) = -\log(\det R) - n \log(\pi e)$$

avec $\hat{R}_n = (Z_n - m_n) \cdot (Z_n - m_n)^+$

et $E[\hat{R}_n] = R_n$

$$g_{ij} \equiv \frac{\partial^2 \tilde{\Phi}}{\partial H_i \partial H_j} \text{ et } H = -R$$

- On peut choisir comme modèle de paramétrisation les coefficients de réflexion

$$\tilde{\Phi}(R_n) = \sum_{k=1}^{n-1} (n-k) \cdot \ln[1 - |\mu_k|^2] + n \cdot \ln[\pi \cdot e \cdot P_0]$$

$$\alpha_n^{-1} = [1 - |\mu_n|^2] \alpha_{n-1}^{-1}$$

$$\det(R_n) = \prod_{k=0}^{n-1} \alpha_k^{-1} = \alpha_0^{-n} \prod_{k=1}^{n-1} [1 - |\mu_k|^2]^{n-k}$$

avec $\alpha_0^{-1} = P_0 = \frac{1}{n} \sum_{k=1}^n |z_k|^2$

Papier initial d'Erich Kähler, 1932, Hambourg

« Kähler Erich, Mathematical Works », Edited by R. Berndt and O. Riemenschneider, Berlin, Walter de Gruyter, ix, 2003

Erich Kähler

$$ds^2 = \sum \frac{\partial^2 U}{\partial x_i \partial \bar{x}_k} dx_i d\bar{x}_k$$

invariant gegenüber der Gruppe der „hyperfuchschen“ Transformationen⁽⁴⁾. Analog ist für die „hyperabelschen“ Transformationen

$$x_i = \frac{\alpha_i x_i + \beta_i}{\gamma_i x_i + \delta_i} \quad (i = 1, 2, \dots, n),$$

die die Einheitskreise

$$1 - x_i \bar{x}_i = 0 \quad (i = 1, 2, \dots, n)$$

festlassen, die aus dem Potential

$$U = \sum_{i=1}^n k_i \log(1 - x_i \bar{x}_i) \quad (k_i \text{ Konstanten})$$

ableitbare Metrik invariant, und es ist klar, daß auch alle Zwischenfälle, etwa die aus hyperfuchschen Transformationen in r und s ($r+s=n$) Variablen komponierten Gruppen, zu Metriken von jenem Typus führen.

Über eine bemerkenswerte Hermitesche Metrik.

Von ERICH KÄHLER in Hamburg.

1.

Bei der Untersuchung der Invarianten einer reell $2n$ -dimensionalen HERMITESCHEN Metrik¹⁾

$$(1) \quad ds^2 = \sum g_{i\bar{k}} dx_i d\bar{x}_k$$

...

Auf die Verwendung der vorliegenden formalen Entwicklungen für die Theorie der automorphen Funktionen und auf die Analogie der Gleichungen

$$D(U) = 0, \quad \Delta(U) = e^{kU}$$

mit den klassischen Differentialgleichungen

$$\Delta U = 0, \quad \Delta U = e^{kU}$$

auf die bereits G. GIRAUD¹⁾ und A. BLOCH²⁾ hingewiesen haben, gedenke ich in einer späteren Arbeit einzugehen.

¹⁾ G. GIRAUD, Sur une équation aux dérivées partielles, non linéaire etc. Comptes Rendus 166, I (1918), p. 893.

²⁾ A. BLOCH, Sur une nouvelle et importante généralisation de l'équation de Laplace. L'Enseignement Mathém. 26 (1927), p. 52.

Hamburg, den 22. Oktober 1932.

L'entropie $U = -\log \det[R]$, pour les modèles autorégressifs complexes, peut être considéré comme un potentiel de Kähler qui est paramétrisé à partir des coefficients de réflexion. Cette expression est alors la même que celle proposée par E. Kähler

Métrie de Kähler : cas hyper-Abelian

- On retrouve la métrie proposée par E. Kähler dans son papier initial : :

- Kähler a nommé ce cas « hyper-Abelian » :

$$\Phi = \sum_{k=1}^{n-1} \alpha_k \cdot \ln \left[1 - |z_k|^2 \right] = \ln K_D(z, z)$$

avec

$$\text{Noyau de Bregman : } K_D(z, z) = \prod_{k=1}^{n-1} \left(1 - |z_k|^2 \right)^{\alpha_k}$$

$$\text{et } \left\{ z / |z_k| < 1 \quad \forall k = 1, \dots, n \right\}$$

- Autre métrie proposée par Kähler, appelé cas « hyper-fuchien » :

$$\Phi = \alpha \cdot \ln \left(1 - \sum_{k=1}^n |z_k|^2 \right) \quad \text{with} \quad \left\{ z / \sum_{k=1}^n |z_k|^2 < 1 \right\}$$

Métrie de Kähler d'un modèle autorégressif complexe

On définit une métrique « Doppler » metric dans le cas de modèles autorégressifs complexes comme le Hessien d'un potentiel de Kähler, où le potentiel est donné comme dans le cas de la géométrie affine de l'information par l'Entropie:

- Le potentiel de Kähler paramétré par $[P_0 \quad \mu_1 \quad \cdots \quad \mu_{n-1}]^T$:

$$\tilde{\Phi}(R_n) = \sum_{k=1}^{n-1} (n-k) \cdot \ln[1 - |\mu_k|^2] + n \cdot \ln[\pi \cdot e \cdot \alpha_0^{-1}]$$

- La métrique associée est alors calculée :

$$\theta^{(n)} = [P_0 \quad \mu_1 \quad \cdots \quad \mu_{n-1}]^T = [\theta_1^{(n)} \quad \cdots \quad \theta_n^{(n)}]^T$$

$$g_{11} = n\alpha_0^2 = nP_0^{-2} \quad g_{ij} = \frac{(n-i) \cdot \delta_{ij}}{(1 - |\mu_i|^2)^2}$$

$$ds_n^2 = n \cdot \left(\frac{dP_0}{P_0} \right)^2 + \sum_{i=1}^{n-1} (n-i) \frac{|d\mu_i|^2}{(1 - |\mu_i|^2)^2}$$

Courbure scalaire du modèle autorégressif complexe

- On utilise l'expression du tenseur de Ricci dans le cas de la géométrie de Kähler :

- En géométrie de Kähler, le tenseur de Ricci est donné par :

$$R_{i\bar{j}} = -\frac{\partial^2 \log(\det g_{k\bar{l}})}{\partial z_i \partial \bar{z}_j}$$

- Pour le cas du modèle autorégressif complexe, on obtient alors :

$$\begin{cases} R_{11} = -2 \frac{1}{P_0^2} \\ R_{k\bar{l}} = -2 \frac{\delta_{kl}}{(1 - |\mu_k|^2)^2} \text{ for } k = 2, \dots, n-1 \end{cases}$$

- On en déduit le calcul de la courbure scalaire qui est négative :

$$R = \sum_{k,l} g^{k\bar{l}} \cdot R_{k\bar{l}}$$

$$R = -2 \cdot \left[\sum_{j=0}^{n-1} \frac{1}{(n-j)} \right]_{n \rightarrow \infty} \rightarrow -\infty$$

Modèle autorégressif # métrique de Kähler-Einstein



La métrique précédente n'est pas une métrique de type Kähler-Einstein, mais une structure « matricielle » proche

- Une métrique est dite de Kähler-Einstein metric si son tenseur de ricci est proportionnel à la métrique :

$$R_{i\bar{j}} = k_0 \cdot g_{k\bar{l}} \quad \text{avec } k_0 : \text{constant} \Rightarrow -\frac{\partial^2 \log(\det g_{k\bar{l}})}{\partial z_i \partial \bar{z}_j} = k_0 \cdot \frac{\partial^2 \Phi}{\partial z^i \partial \bar{z}^j}$$

- Dans le cas de la métrique de Kähler-Einstein, le potentiel de Kähler est solution de l'équation de Monge-Ampère :

$$\det(g_{k\bar{l}}) = |\psi|^2 e^{-k_0 \Phi} \quad \text{avec} \quad \begin{cases} \Phi : \text{Potentiel de Kähler} \\ \psi : \text{fonction holomorphe} \end{cases}$$

- Pour un modèle autorégressif complexe, on a :

$$[R_{ij}] = B^{(n)} [g_{ij}] \quad \text{avec} \quad R = \text{Tr}[B^{(n)}] = -2 \cdot \left[\sum_{j=0}^{n-1} \frac{1}{(n-j)} \right]$$

$$\text{et où } B^{(n)} = -2 \text{diag} \{ \dots, (n-i)^{-1}, \dots \}$$

Algorithme médian sur les coefficients de réflexions 1/2



■ On se donne $(\theta_1, \dots, \theta_N)$, où $\theta_k = (P_0^{(k)}, \mu_1^{(k)}, \dots, \mu_{n-1}^{(k)}) = (P_0^{(k)}, \mu^{(k)}) \in R_+ \times D^{n-1}$ est un n uplet des coefficients de réflexion $\mu^{(k)} = (\mu_1^{(k)}, \dots, \mu_{n-1}^{(k)})$ à tous les ordres précédés du coefficient de puissance $P_0^{(k)}$.

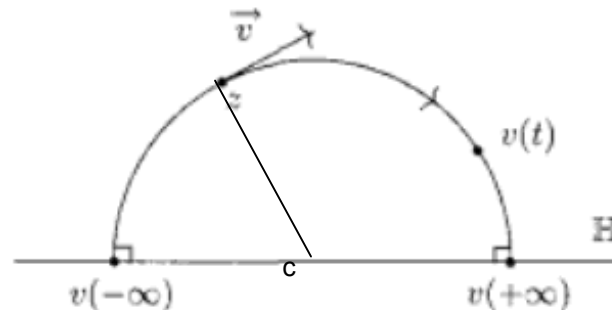
■ On calcule d'abord $P_0^{median} = \exp(\text{mediane}(\log(P_0^{(1)}), \log(P_0^{(2)}), \dots, \log(P_0^{(N)})))$, dans R (méthode classique de calcul du médian pour une valeur scalaire).

■ On estime ensuite la médiane des coefficients de réflexions complexes algorithmiquement : on ramène ces coefficients du disque unité dans le demi-plan de Poincaré $C^{-1} : D \rightarrow H$ avec :

$$\mu_j^{(k)} \mapsto z_j^{(k)} = i \left(\frac{1 + \mu_j^{(k)}}{1 - \mu_j^{(k)}} \right)$$

On note $z^{(k)} = C^{-1}(\mu^{(k)})$

■ Les géodésiques sont alors des demi-cercles centrés sur l'axe des abscisses



Algorithme médian sur les CR

- On initialise l'algorithme à un point quelconque Z dans H :

$$Z^{(0)} = [Z_1^{(0)} \dots Z_{n-1}^{(0)}]^T$$

- A l'étape p , on se déplace selon le gradient:

$$t_{Z,l}^{(p)} = \sum_{k=1}^N i \times \text{signe}(Z_l^{(p)} - z_l^{(k)}) \frac{Z_l^{(p)} - c_l^{(k)}}{|Z_l^{(p)} - c_l^{(k)}| + 1/p} \quad \text{avec} \quad c_l^{(k)} = \frac{|z_l^{(k)}|^2 - |Z_l^{(p)}|^2}{2\text{Re}(z_l^{(k)} - Z_l^{(p)})}$$

où Z est l'estimé courant et c_k désigne le centre de la géodésique rejoignant Z à z_k .

$$c_{Z_l^{(p)}} = \text{Re}(Z_l^{(p)}) + \frac{\text{Im}(t_{Z,l}^{(p)})}{\text{Re}(t_{Z,l}^{(p)})} \text{Im}(Z_l^{(p)})$$

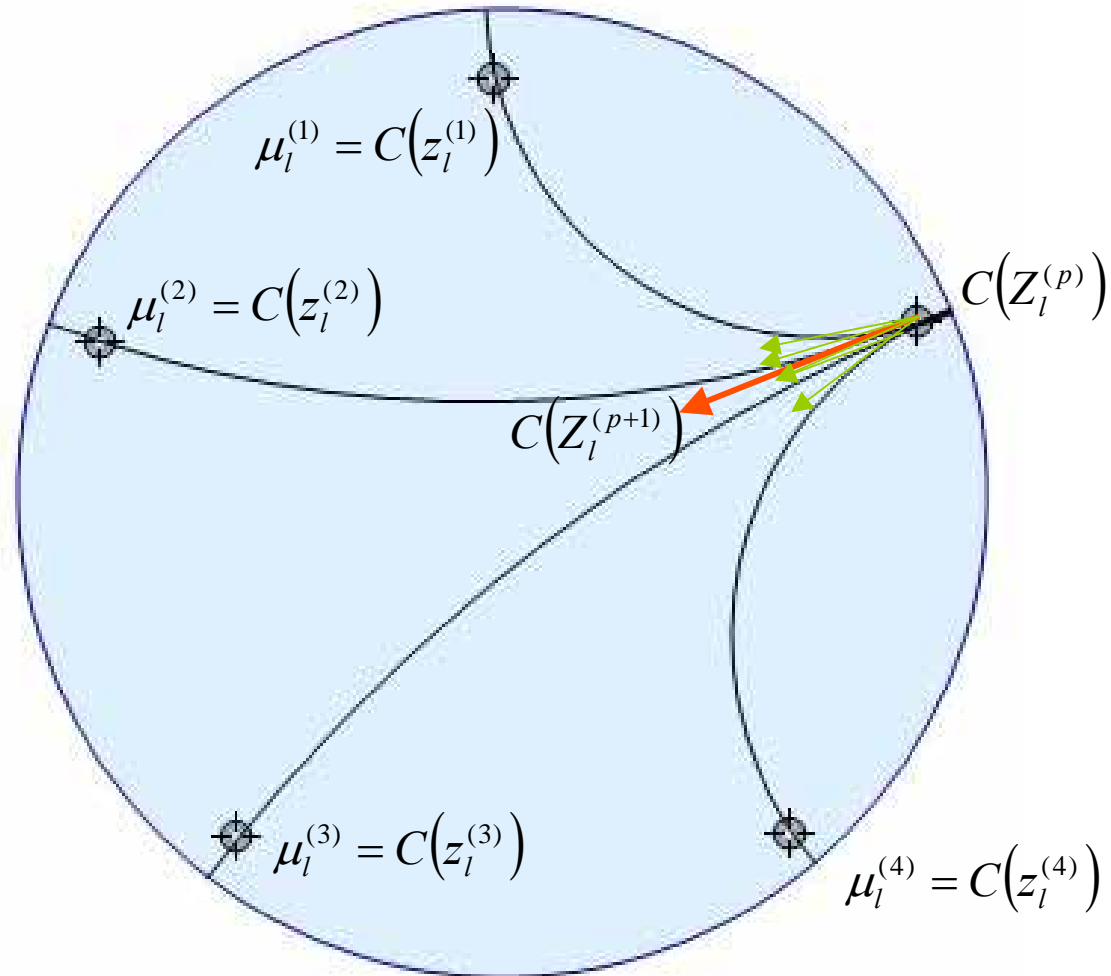
$$\phi_l^{(p)} = \arg(Z_l^{(p)} - c_{Z_l^{(p)}})$$

$$Z_l^{(p+1)} = c_{Z_l^{(p)}} + |Z_l^{(p)} - c_{Z_l^{(p)}}| \exp \left[i \left(\phi_l^{(p)} - \frac{1}{2Mk} \text{signe}(\text{Re}(Z_l^{(p)} - c_{Z_l^{(p)}})) \frac{Z_l^{(p)} - c_{Z_l^{(p)}}}{|Z_l^{(p)} - c_{Z_l^{(p)}}| + 1/p} \right) \right]$$

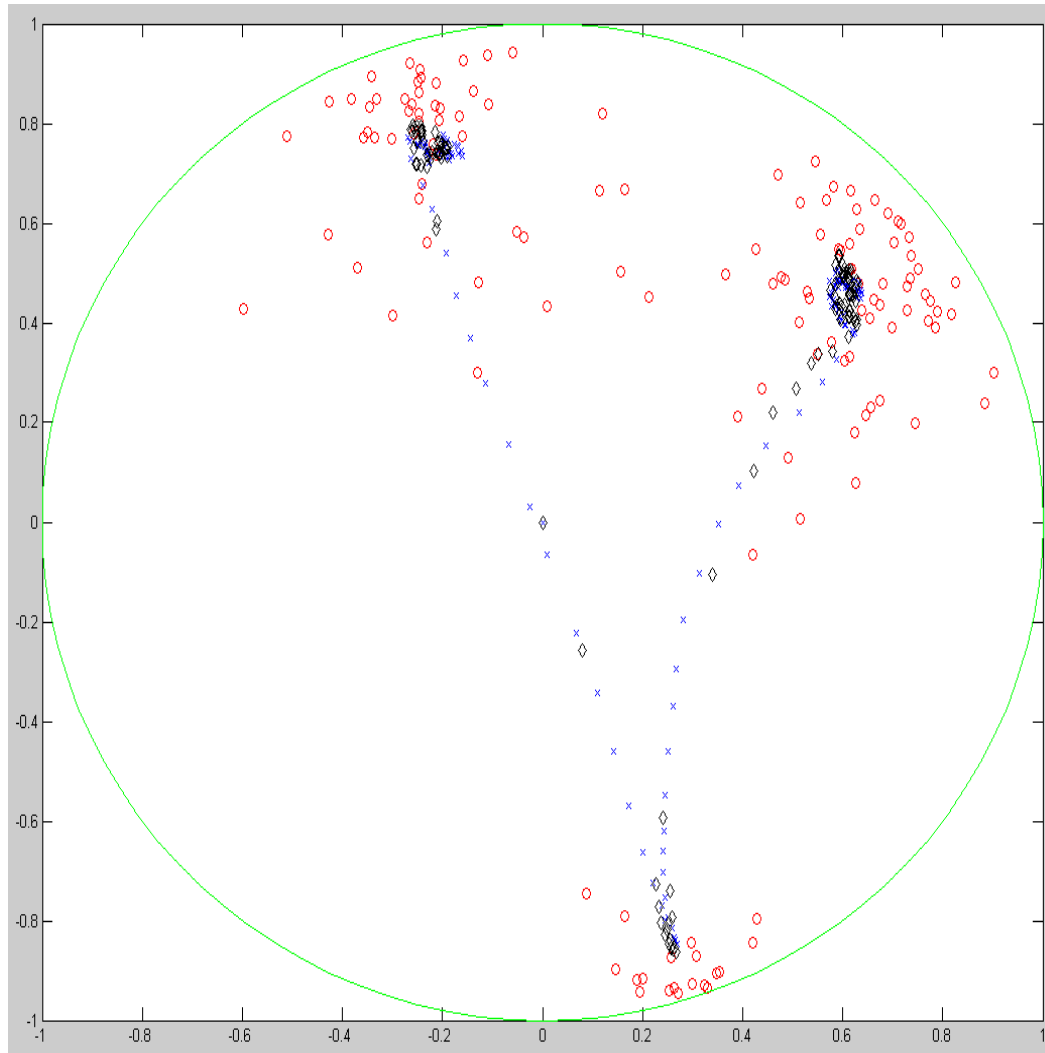
- L'entier p , initialisé à 1, décroît si l'écart entre deux itérés est moindre que $1/p$.
- Après convergence, on retourne à :

$$\Theta = (P, C(Z)) \quad \text{avec} \quad \begin{cases} P_0^{median} = \exp(\text{mediane}(\log(P_0^{(1)}), \log(P_0^{(2)}), \dots, \log(P_0^{(N)}))) \\ \mu_l^{median} = C(Z_l^{(conv)}) = \frac{Z_l^{(conv)} - i}{Z_l^{(conv)} + i} \end{cases}$$

Flot de gradient dans le disque unité de Poincaré



Résultats sur le premier coefficient de réflexion



Superposition pour toutes les cases distances des moyennes et médianes locale, avec les points de données

Equation de la chaleur de Joseph Fourier



- « Les équations différentielles de la propagation de la chaleur expriment les conditions les plus générales, et ramènent les questions physiques à des problèmes d'analyse pure, ce qui est proprement l'objet de la théorie ... Les formes des corps sont variées à l'infini, la distribution de la chaleur qui les pénètre peut être arbitraire et confuse; mais toutes les inégalités s'effacent rapidement et disparaissent à mesure que le temps s'écoule. La marche du phénomène devenue plus régulière et plus simple, demeure enfin assujettie à une loi déterminée qui est la même pour tous les cas, et qui ne porte plus aucune empreinte sensible de la disposition initialeLes théories nouvelles, expliquées dans notre ouvrage sont réunies pour toujours aux sciences mathématiques et reposent comme elles sur des fondements invariables; elles conserveront tous les éléments qu'elles possèdent aujourd'hui, et elles acquerront, continuellement plus d'étendue. »

Joseph Fourier (1768-1830), « Discours préliminaire à la théorie analytique de la chaleur »

Équation de diffusion & géodésique



- Équation de diffusion sur un graphe 1D de données scalaires



- Dans un espace vectoriel normé dans le cas unidimensionnel, l'équation de diffusion s'écrit via le Laplacien discret:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \Rightarrow \frac{\partial u}{\partial t} = \frac{1}{\nabla x} \left(\frac{u_{n+1} - u_n}{\nabla x} - \frac{u_n - u_{n-1}}{\nabla x} \right) = \frac{2}{\nabla x^2} (\hat{u}_n - u_n)$$

avec la moyenne arithmétique : $\hat{u}_n = (u_{n+1} + u_{n-1}) / 2$

- L'équation de Fourier discrétisée peut également s'écrire :

$$u_{n,t+1} = u_{n,t} + \frac{2\nabla t}{\nabla x^2} [\hat{u}_{n,t} - u_{n,t}] = (1 - \rho) u_{n,t} + \rho \hat{u}_{n,t} \text{ avec } \rho = \frac{2\nabla t}{\nabla x^2}$$

- Par analogie sur la base du flot de Karcher, il est possible d'écrire une équation de diffusion sur un graphe 1D de matrices, via la moyenne géométrique $\hat{A}_{n,t}$ des voisins $(A_{n-1,t}, A_{n+1,t})$:

$$X_{n,t+1} = X_{n,t}^{1/2} e^{\frac{2\nabla t}{\nabla x^2} \log(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2})} X_{n,t}^{1/2} = X_{n,t}^{1/2} \left(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2} \right)^\rho X_{n,t}^{1/2}$$

$$\text{avec } \rho = \frac{2\nabla t}{\nabla x^2} \text{ et } \hat{X}_{n,t} = X_{n+1,t}^{1/2} \left(X_{n+1,t}^{-1/2} X_{n-1,t} X_{n+1,t}^{-1/2} \right)^{1/2} X_{n+1,t}^{1/2}$$

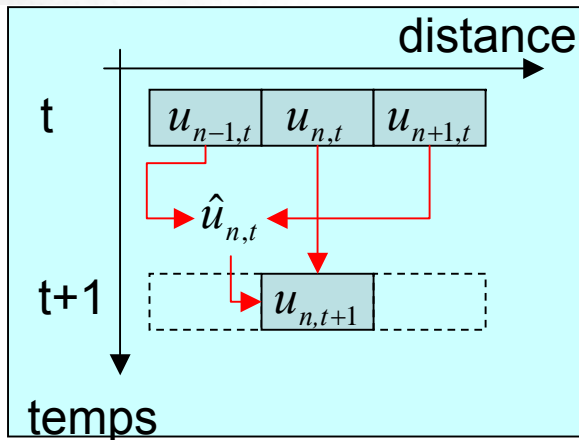
Equation de diffusion sur un graphe 1D scalaire

- Equation de diffusion sur un graphe 1D de données scalaires



$$u_{n,t+1} = (1 - \rho) \cdot u_{n,t} + \rho \cdot \hat{u}_{n,t}$$

avec $\rho = \frac{2\nabla t}{\nabla x^2}$ et $\hat{u}_{n,t} = \frac{u_{n+1,t} + u_{n-1,t}}{2}$



$$u_{n,t+1} = u_{n,t} \bullet_{\rho} \hat{u}_{n,t}$$

$\rho \mapsto 1$

$$\hat{u}_{n,t} = u_{n-1,t} \bullet_{1/2} u_{n+1,t}$$

$u_{n-1,t}$

$u_{n+1,t}$

$\rho \mapsto 0$

$u_{n,t}$

$$u \bullet_{\rho} v = (1 - \rho)u + \rho \cdot v$$

$$\rho \in [0, 1]$$

THALES

Equation de diffusion sur un graphe 1D de matrices

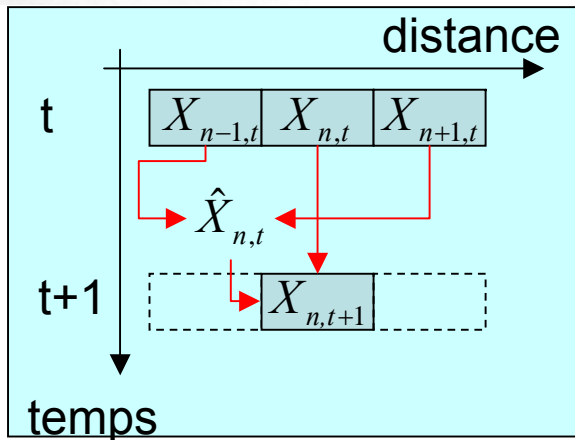


- Equation de diffusion sur graphe 1D de matrices SPD



$$X_{n,t+1} = X_{n,t}^{1/2} e^{\frac{2\nabla t}{\nabla x^2} \log(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2})} X_{n,t}^{1/2} = X_{n,t}^{1/2} \left(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2} \right)^\rho X_{n,t}^{1/2}$$

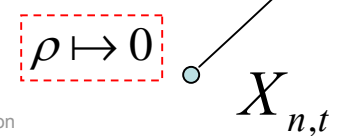
avec $\rho = \frac{2\nabla t}{\nabla x^2}$ et $\hat{X}_{n,t} = X_{n+1,t}^{-1/2} \left(X_{n+1,t} X_{n-1,t} X_{n+1,t} \right)^{1/2} X_{n+1,t}^{-1/2}$



$$X_{n,t+1} = X_{n,t} \bullet_\rho \hat{X}_{n,t}$$

$$\hat{X}_{n,t} = X_{n-1,t} \bullet_{1/2} X_{n+1,t}$$

$$\rho \mapsto 0$$



$$X_{n-1,t}$$

$$X_{n+1,t}$$

$$A \bullet_\rho B = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^\rho A^{1/2}$$

$$\rho \in [0,1]$$

Équation de diffusion : contrainte de discrétisation

- Une première remarque concernant une contrainte qui apparaît :

$$u_{n,t+1} = u_{n,t} + \frac{2\nabla t}{\nabla x^2} [\hat{u}_{n,t} - u_{n,t}] = (1 - \rho) \cdot u_{n,t} + \rho \cdot \hat{u}_{n,t} \text{ avec } \rho = \frac{2\nabla t}{\nabla x^2}$$

$$u_{n,t+1} \in \begin{cases} [u_{n,t}, \hat{u}_{n,t}] & \text{si } u_{n,t} < \hat{u}_{n,t} \\ [\hat{u}_{n,t}, u_{n,t}] & \text{si } \hat{u}_{n,t} < u_{n,t} \end{cases} \Leftrightarrow \rho < 1 \text{ soit } \frac{2}{\nabla x} < \nu \text{ en posant } \nu = \frac{\nabla x}{\nabla t}$$

Si on vérifie Shannon $\nabla x = T_E = \frac{1}{f_E}$ avec $f_{\max} < \frac{f_E}{2} \Rightarrow 4f_{\max} < 2f_E < \nu$

Équation de diffusion : Du discret au continu

- L'équation de diffusion discrète sur un graphe 1D de matrices symétriques ou hermitiennes définies positives est donnée par :

$$X_{n,t+1} = X_{n,t}^{1/2} e^{\frac{2\nabla t}{\nabla x^2} \log(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2})} X_{n,t}^{1/2} = X_{n,t}^{1/2} \left(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2} \right)^\rho X_{n,t}^{1/2}$$

$$\text{avec } \rho = \frac{2\nabla t}{\nabla x^2} \text{ et } \hat{X}_{n,t} = X_{n+1,t}^{1/2} \left(X_{n+1,t}^{-1/2} X_{n-1,t} X_{n+1,t}^{-1/2} \right)^{1/2} X_{n+1,t}^{1/2}$$

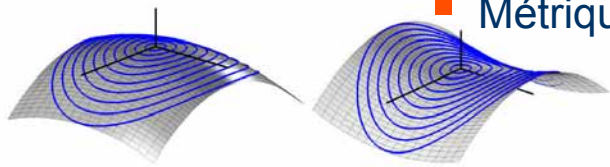
- Pour remonter à l'équation continue de diffusion, il faut utiliser l'équation de Campbell-Hausdorff :

$$\log(e^X e^Y) = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}[X, [X, Y]] + \frac{1}{12}[Y, [Y, X]] \\ + \text{termes de degré } \geq 4$$

$$\text{avec } [X, Y] = XY - YX$$

Équation de diffusion anisotrope

- Regardons maintenant le cas anisotrope classique :
 - Dans le cas général diffusion sur une variété, l'équation de diffusion de Laplace-Beltrami s'écrit :



■ Métrique :

$$ds^2 = \sum_{k=1}^N g_{i,j} dx_i dx_j \quad \text{avec} \quad g = [g_{i,j}]_{i,j=1}^N$$

■ Équation de diffusion :

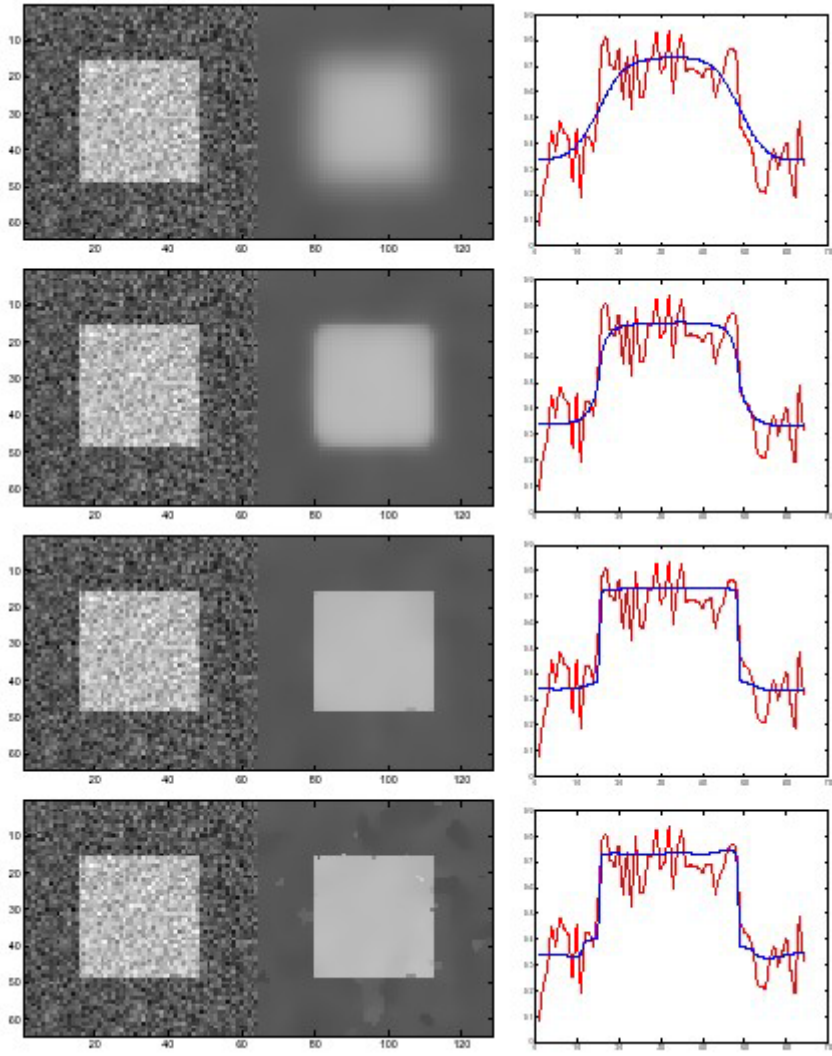
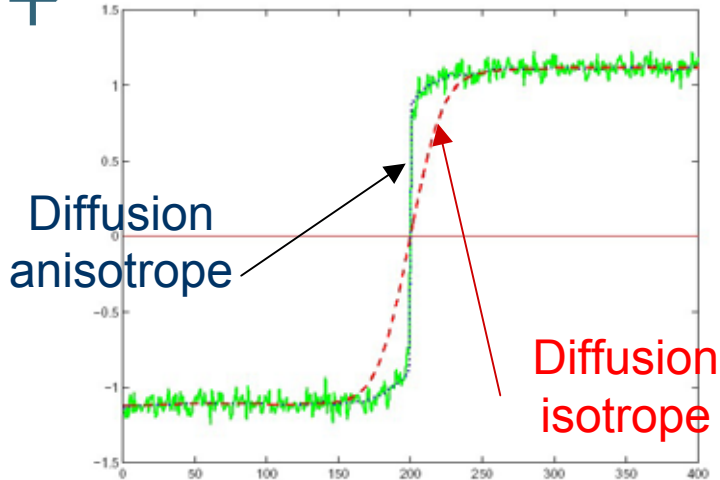
$$\frac{\partial u}{\partial t} = \frac{1}{\sqrt{\det(g)}} \sum_{i,j} \frac{\partial}{\partial x_i} \left(\sqrt{\det(g)} g^{ij} \frac{\partial u}{\partial x_j} \right) \quad \text{avec} \quad g^{-1} = [g^{i,j}]_{i,j=1}^N$$

■ Dans le cas 1D isotrope : $ds^2 = dx^2 \Rightarrow \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$

■ Dans le cas 1D anisotrope : $ds^2 = dx^2 + du^2 = \left(1 + \left(\frac{\partial u}{\partial x} \right)^2 \right) dx^2$
soit :

$$g = \left(1 + \left(\frac{\partial u}{\partial x} \right)^2 \right) \Rightarrow \frac{\partial u}{\partial t} = \left(1 + \left(\frac{\partial u}{\partial x} \right)^2 \right)^{-1/2} \frac{\partial}{\partial x} \left(\left(1 + \left(\frac{\partial u}{\partial x} \right)^2 \right)^{-1/2} \frac{\partial u}{\partial x} \right)$$

Diffusion anisotrope ↩



diffusion anisotrope sur un graphe de matrices

- Dans le cas 1D anisotrope, par analogie sur un graphe de matrices :

$$X_{n,t+1} = X_{n,t}^{1/2} e^{\gamma \log(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2})} X_{n,t}^{1/2} = X_{n,t}^{1/2} \left(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2} \right)^\gamma X_{n,t}^{1/2}$$

$$\text{avec } \hat{X}_{n,t} = X_{n+1,t}^{1/2} e^{\beta \log(X_{n+1,t}^{-1/2} X_{n-1,t} X_{n+1,t}^{-1/2})} X_{n+1,t}^{1/2} = X_{n+1,t}^{1/2} \left(X_{n+1,t}^{-1/2} X_{n-1,t} X_{n+1,t}^{-1/2} \right)^\beta X_{n+1,t}^{1/2}$$

$$\beta = \frac{\alpha_{n,t}^+}{\left(\alpha_{n,t}^+ + \alpha_{n,t}^- \right)}, \quad \gamma = \frac{\nabla t \cdot \alpha_{n,t} \left(\alpha_{n,t}^+ + \alpha_{n,t}^- \right)}{\nabla x^2}, \quad \alpha_{n,t} = \left(1 + \left(\frac{d(X_{n+1,t}, X_{n-1,t})}{\nabla x} \right)^2 \right)^{-1/2}$$

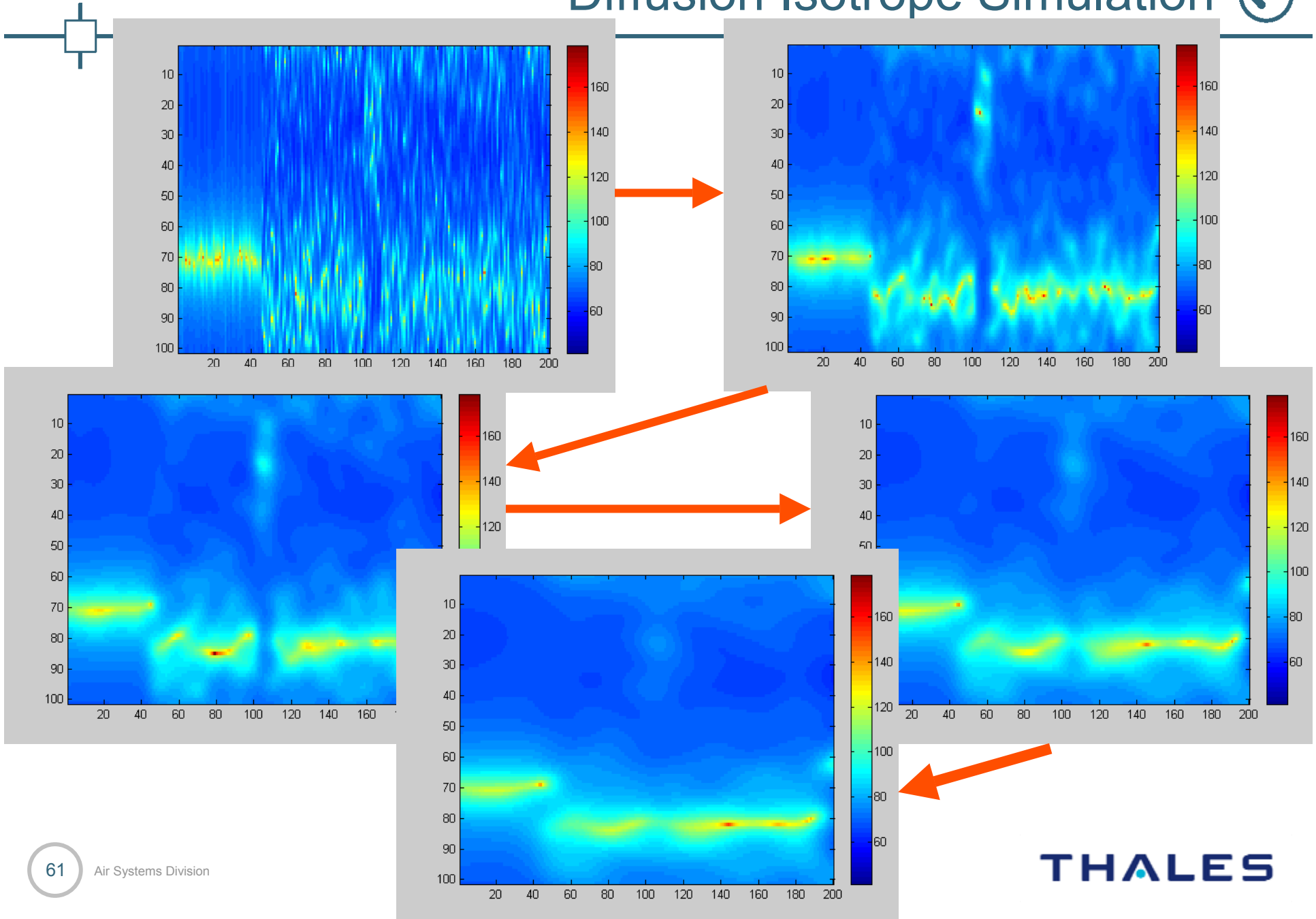
$$\alpha_{n,t}^+ = \left(1 + \left(\frac{d(X_{n+1,t}, X_{n,t})}{\nabla x} \right)^2 \right)^{-1/2} \quad \text{et} \quad \alpha_{n,t}^- = \left(1 + \left(\frac{d(X_{n,t}, X_{n-1,t})}{\nabla x} \right)^2 \right)^{-1/2}$$

- Pour la distance , on prend :

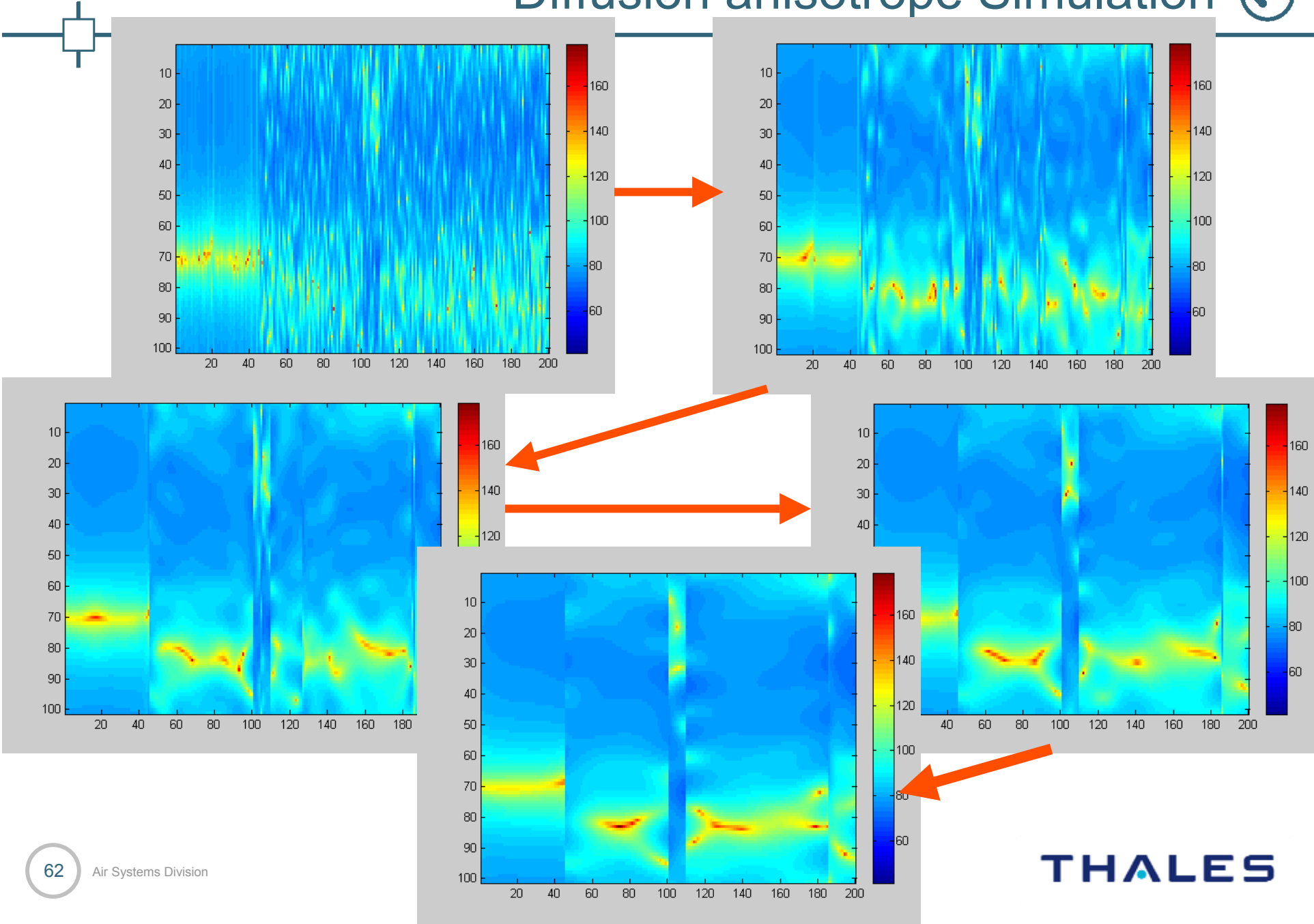
$$d(X_1, X_2) = \left\| \log(X_1^{-1/2} X_2 X_1^{-1/2}) \right\|_F = \sqrt{\sum_{i=1}^M \log^2(\lambda_i)}$$

$$\text{et } \det(X_1 - \lambda X_2) = 0$$

Diffusion Isotrope Simulation

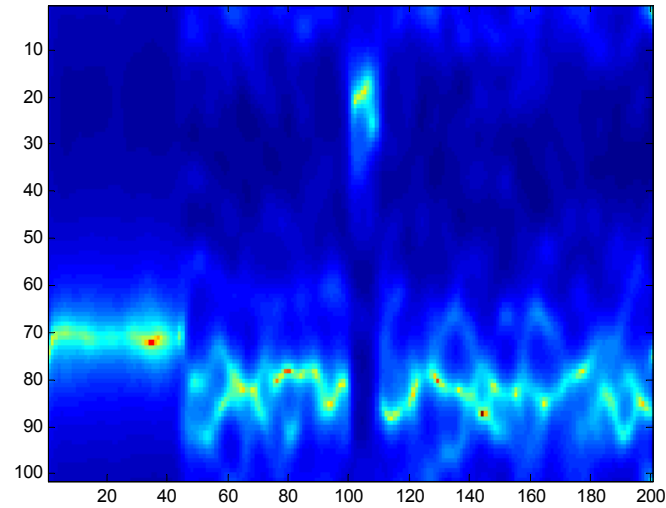


Diffusion anisotrope Simulation

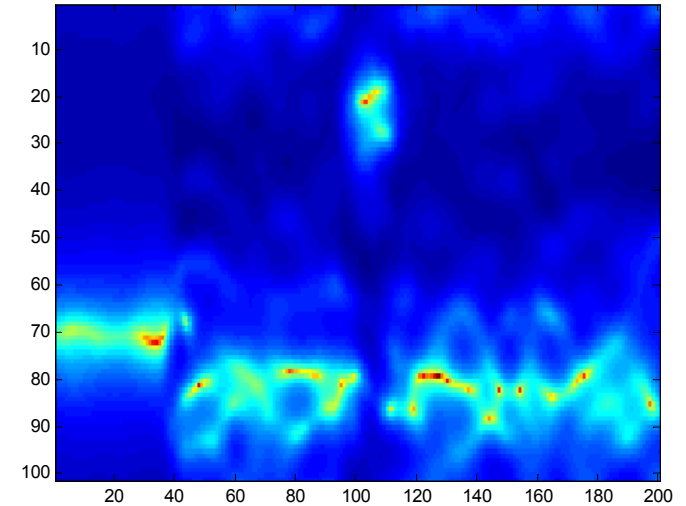


Les différentes diffusion à 50 itérations

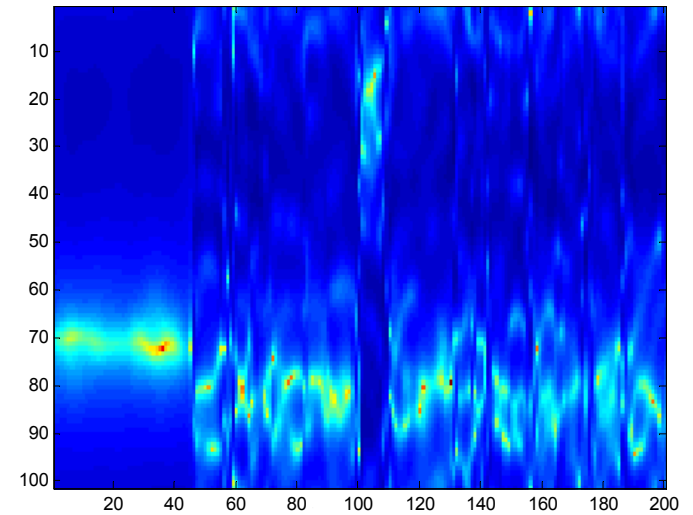
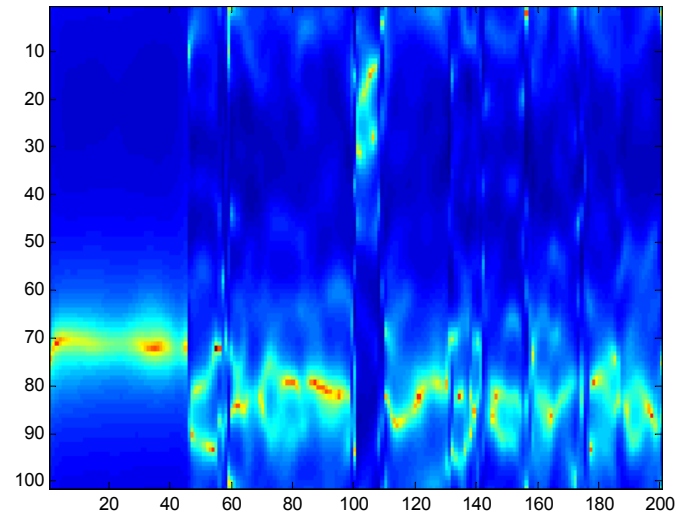
Diffusion
Isotrope



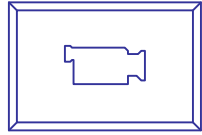
Diffusion médianage



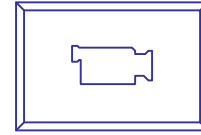
Diffusion
Anisotrope



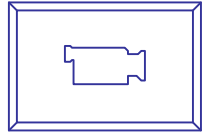
Les différentes diffusions



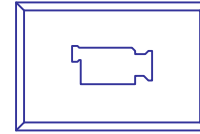
Diffusion Isotrope



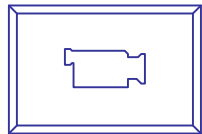
Diffusion Anisotrope



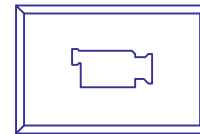
Medianage Isotrope



Medianage Anisotrope



Moyennage Isotrope

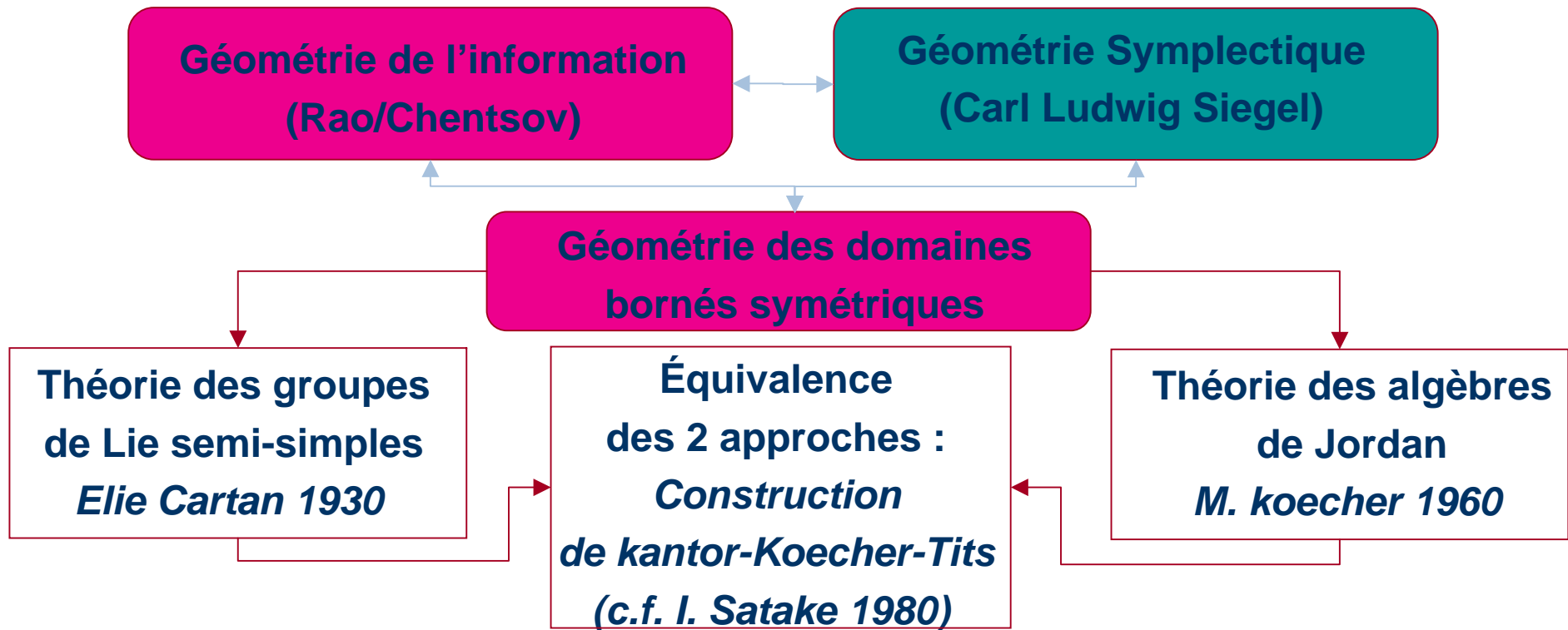


Moyennage Anisotrope

Les 2 sources géométriques

- Géométrie de l'information & Géométrie des cônes symétriques

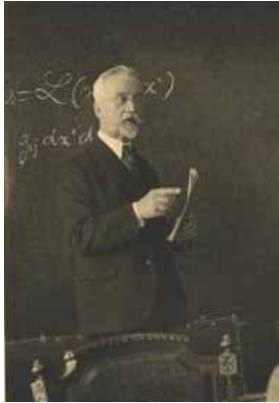
- Ces 2 géométries ont de multiples liens avec les travaux de Carl Ludwig Siegel (1896-1981) en Géométrie symplectique :



Références Bibliographiques

- [1] F. Barbaresco & G. Bouyt, "Espace Riemannien symétrique et géométrie des espaces de matrices de covariance : équations de diffusion et calculs de médianes", colloque GRETSI'09, Dijon, Sept. 2009
- [2] F. Barbaresco, "New Foundation of Radar Doppler and Array Processing based on German Mathematical Works: Geometry of Metric Spaces with negative curvature and Von-Mangoldt-Cartan-Hadamard Manifolds", IRS'09, Hamburg, Sept. 2009
- [3] F. Barbaresco, "Interactions between Symmetric Cone and Information Geometries", LIX Fall Colloquium ETCV'08, Ecole Polytechnique, Springer Lecture Notes in Computer Science, vol.5416, pp. 124-163, 2009, <http://www.springerlink.com/content/978-3-642-00825-2>
- [4] F. Barbaresco, "Applications of Information Geometry to Radar Signal Processing", Video Lectures, http://videolectures.net/etvc08_barbaresco_aoigt/
- [5] M. Arnaudon and X. Li, « Barycentres of measures transported by stochastic flows », Ann. Probab. 33, no. 4, pp.1509-1543, 2005
- [6] C.L. Siegel, «Symplectic Geometry», Acad. Press, 1964
- [7] M. Émery, G. Mokobodzki, «Sur le barycentre d'une probabilité dans une variété » . Séminaire de probabilités de Strasbourg, n° 25, p. 220-233, 1991
- [8] H. Karcher, «Riemannian center of mass and mollifier smoothing», Comm. Pure Applied Math., °30, pp.509-541, 1977
- [9] A. Terras, «Harmonic Analysis on Symmetric Spaces and Applications II», Springer-Verlag, 1988
- [10] V. Arsigny «Log-Euclidean metrics for fast and simple calculus on diffusion tensors», in Magnetic Resonance in Medicine, Volume 56 Issue 2, Pages 411 - 421, Jun 2006.
- [11] M. Calvo, J. Oller, «A distance between multivariate normal distributions based in an embedding into the Siegel Group», Journal of Multivariate Analysis, vol.35, pp. 223-242, 1990
- [12] M. Calvo, J. Oller, «An explicit solution of information geodesic equations for the multivariate normal model», Statistics and Decisions., vol.9, pp. 119-138, 1991
- [13] Huiling Le, "Estimation of Riemannian Barycenters", Proc. London Math. Soc., pp. 193-200, 2004
- [14] Harris, W. F., "The average eye", Ophthalmic Physiol. Opt., n°24, pp. 580-585, 2004
- [15] H.C.F. von Mangoldt, "Über diejenigen Punkte auf positiv gekrümmten flächen, welche die eigenschaft haben, dass die von ihnen ausgehenden geodätischen Linien nie aufhören, kürzeste Linien zu sein", J. Reine Angew.Math., n°91, pp.23-52, 1881
- [16] E. Kähler, "Über eine bemerkenswerte Hermitesche Metrik", Abh., Math. Sem. Hamburg Univ., n°9, pp.173-186, 1933
- [17] K. T. Sturm, "Probability measures on metric spaces of nonpositive curvature". Contemp. Math. 338 , 357-390, 2003
- [18] E. Cartan, "Sur les domaines bornés homogènes de l'espace de n variables complexes", Abh. Math. Semin. hamb. Univ., n°11, pp.116-162, 1935, www.numdam.org
- [19] C.R. Rao, "Information and Accuracy attainable in the estimation of statistical parameters", Bull. Calcutta Math. Soc., n°37, pp.81-91, 1945
- [20] S. Yoshizawa and K. Tanabe, "Dual Differential Geometry associated with the Kullback-Leibler Information on the Gaussian Distributions and its 2-parameter Deformations", SUT Journal of Mathematics, vol.35,n°1, pp.113-137, 1999
- [21] T.Ando and al., "Geometric Means", Linear Algebra Appl., vol.385, pp.305-334, 2004
- [22] F. Bruhat and J. Tits, « Groupes réductifs sur un corps local », IHES, n°41, pp.5-251, 1972
- [23] H. Cartan, "Ouverts fondamentaux pour le groupe modulaire", Séminaire Henri Cartan, tome n°10, n°1, exp.n°3, p.1-12, 1957
- [24] N.N. Chentsov, "Statistical Decision Rules and Optimal Inferences", Trans. of Math. Monog., n°53, Amer. Math. Society, Providence, 1982
- [25] J. Faraut and A. Koranyi, "Analysis on Symmetric Cones", Oxford University Press, 1994
- [26] K. Koufany, "Analyse et Géométrie des domaines bornés symétriques", HDR, Institut de Mathématiques Elie Cartan, Nancy, Nov. 2006
- [27] C. P. Niculescu, "An Extension of the Mazur-Ulam Theorem", American Institute of Physics Proc., vol. 729, n°248-256, 2004
- [28] M. Berger, « Panoramic View of Riemannian Geometry », Springer, 2004

Contributions des mathématiciens français



E. Cartan, « Leçons sur la géométrie des espaces de Riemann », 2nd édition, revue et augmentée. Paris, Gauthier-Villars, 1946



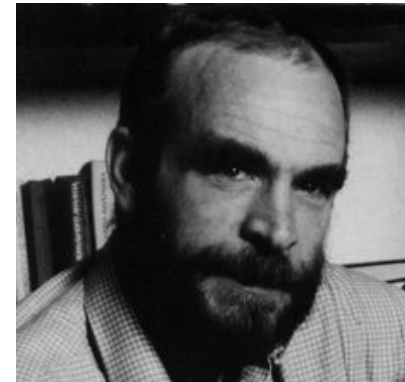
F. Bruhat and J. Tits, « Groupes réductifs sur un corps local », IHES, n°41, pp.5-251, 1972



M. Fréchet, "L'intégrale abstraite d'une fonction abstraite d'une variable abstraite et son application à la moyenne d'un élément aléatoire de nature quelconque", Revue Scientifique, pp. 483-512, 1944



J. Hadamard, « Les surfaces à courbures opposées et leurs lignes géodésiques », J. Math, sér. 5, 4, pp.27-73, 1898



M. Gromov, "Hyperbolic groups". Essays in group theory. Math. Sci. Res. Inst. Publ. 8. New York: Springer. pp. 75-263, 1987

Contributions des mathématiciens allemands



H.C.F. von Mangoldt, "Über diejenigen Punkte auf positiv gekrümmten flächen, welche die eigenschaft haben, dass die von ihnen ausgehenden geodätischen Linien nie aufhören, kürzeste Linien zu sein", J. Reine Angew.Math., n°91, pp.23-52, 1881

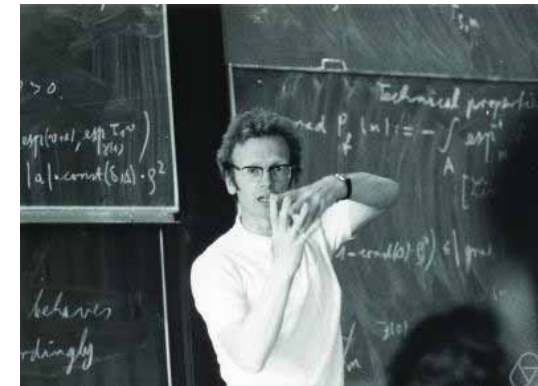


E. Kähler, "Über eine bemerkenswerte Hermitesche Metrik", Abh., Math. Sem. Hamburg Univ., n°9, pp.173-186, 1933

C.L. Siegel, "Symplectic Geometry", Academic Press, New York, 1964



K. T. Sturm, "Probability measures on metric spaces of nonpositive curvature". In vol.: Heat kernels and analysis on manifolds, graphs, and metric spaces, Contemp. Math. 338 , 357-390, 2003



H. Karcher, "Riemannian center of mass and mollifier smoothing", Comm. Pure Applied Math., °30, pp.509-541, 1977

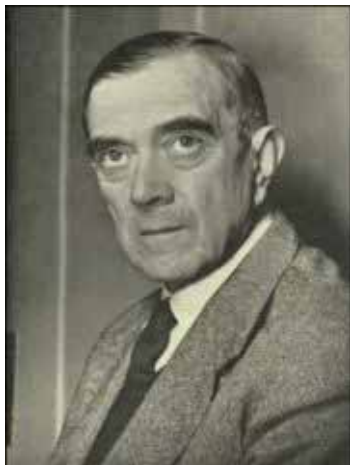
Contributions des mathématiciens italiens



E. Torricelli, Opere di Evangelista Torricelli, G. Loria and G. Vassura (Eds), Vol.I, 2ème partie, pp.90-97, Vol. III, pp.426-431, Faenza, 1919

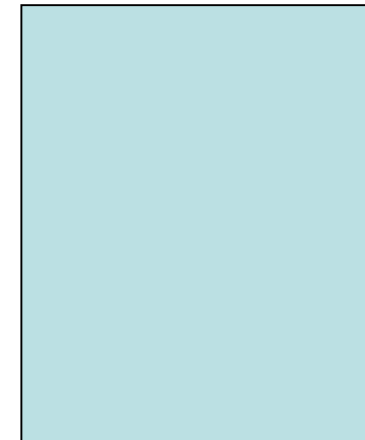


B.F. Cavalieri, « Exercitationes geometricae », Bologna, 1647



C. Gini, «L'uomo medio », Giornali degli economiste e rivista de statica, n°48, pp.1-24, 1914

C. GINI, « Di una formula comprensiva delle medie », Metron 13 , 3-22, 1938



L. Galvani and C. Gini « Di talune estensioni dei concetti di media ai caratteri qualitativi », Metron, 8, 3-209, 1929

Contribution des mathématiciens anglais



M.G. Kendall, « Note on the estimation of a ranking »,
Journal of the Royal Statistical
Society, 105, 119, 1942

M.G. Kendall, « Rank Correlation Methods », Londres,
Griffin, 1955



Henri Cartan

André Weil

René de Possel

Charles Ehresmann

Laurent Schwartz



Jean Dieudonné

Claude Chevalley

Pierre Samuel

Jean-Pierre Serre

Adrien Douady

BOURBAKI Group : 60 years ago

Metric Measure Space

Séminaire BOURBAKI
60ème année, 2007-2008, n° 990

Mars 2008

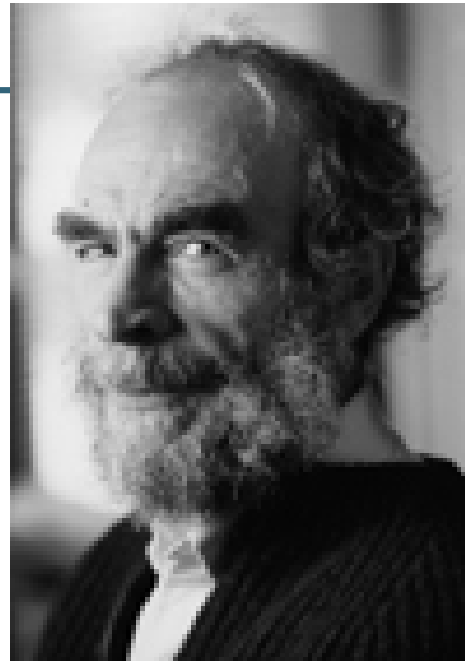
GÉOMÉTRIE DES ESPACES MÉTRIQUES MESURÉS : LES TRAVAUX DE LOTT, VILLANI, STURM

par Michel LEDOUX

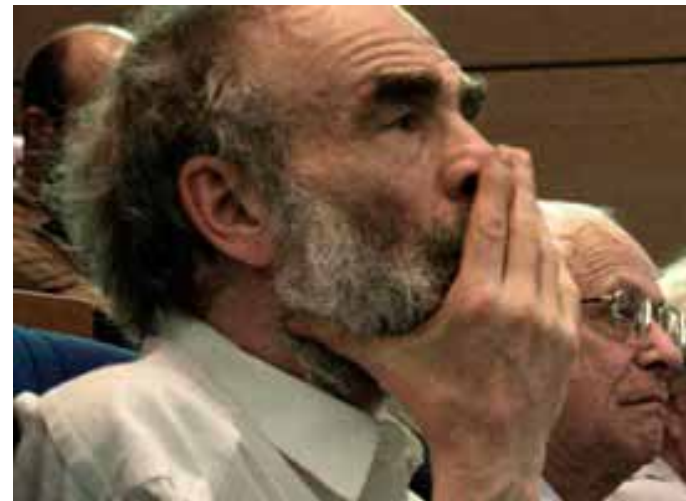
L'étude géométrique des espaces métriques mesurés s'est récemment dotée d'une définition synthétique de borne inférieure de courbure de Ricci issue de la théorie du transport optimal de mesures grâce aux travaux parallèles et complémentaires de J. Lott et C. Villani [L-V1-2] et K.-T. Sturm [Stu2-3].

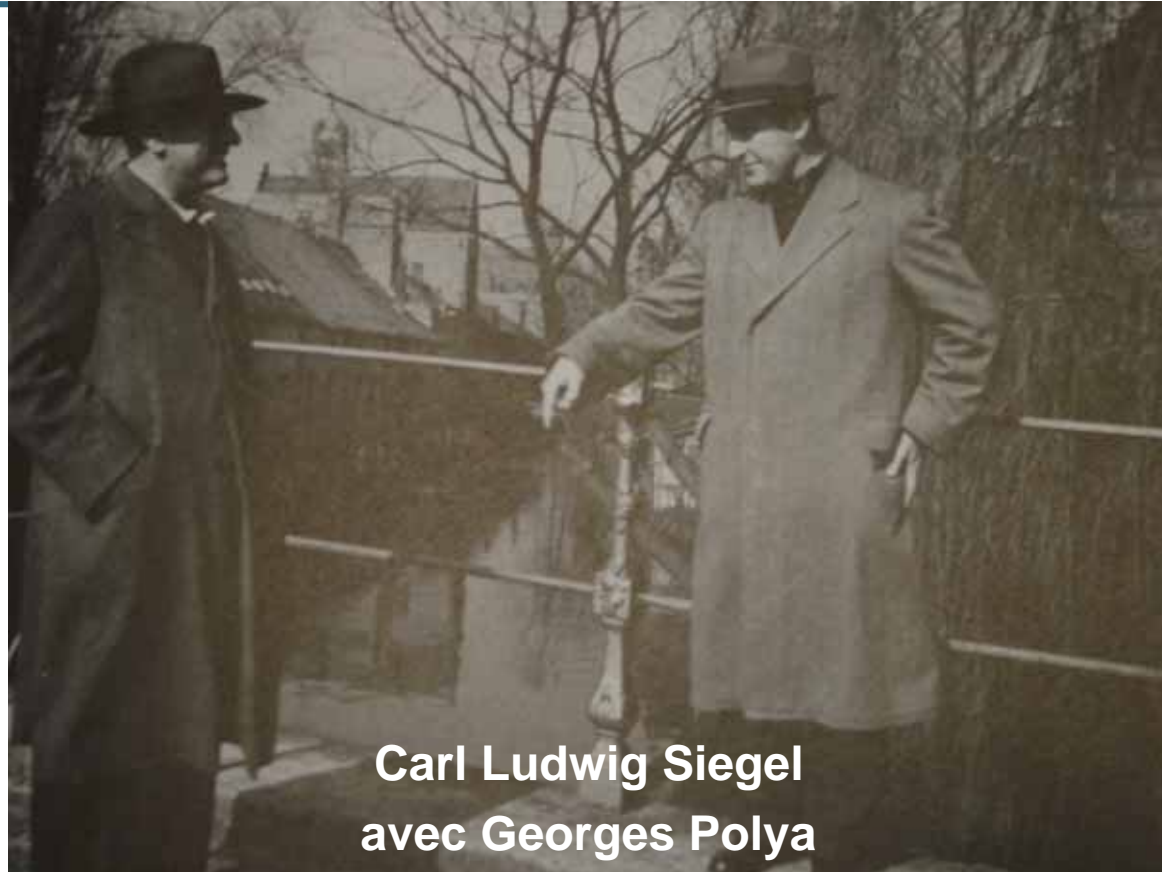


M. Gromov, “Hyperbolic Groups”,
Essays in Group Theory, Math. Sci.
Res. Inst. Publ. 8, New York, pp.75-
263, 1987



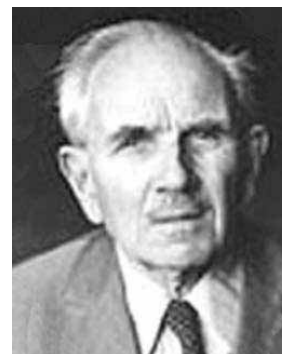
IHES
Institut des Hautes Etudes
Scientifiques
Bures sur Yvettes
France





**Carl Ludwig Siegel
avec Georges Polya**

■ *“ Constaté que les théories les plus parfaites sont les guides les plus sûrs pour résoudre les problèmes concrets~; avoir assez confiance en sa science pour prendre des responsabilités techniques. Puissent beaucoup de mathématiciens connaître un jour ces joies très saines, quelques humbles qu'ils les jugent !”*



■ **Jean LERAY**

THALES



Planches supplémentaires
Version Longue
Détails des calculs

THALES



Introduction de la métrique la plus naturelle aux espaces de matrices de covariance

Métrieque et géométrie pour les matrices de covariance

- La géométrie de l'information pour les lois multivariées gaussiennes de moyenne nulle et la géométrie différentielle intrinsèque des matrices hermitiennes définies positives (cas particulier de la géométrie des espaces de Siegel) conduit à la même métrieque :

- Géométrie de l'information :

$$p(Z_n / R_n) = (\pi)^{-n} \cdot |R_n|^{-1} \cdot e^{-Tr[\hat{R}_n \cdot R_n^{-1}]}$$

avec $g_{ij}(\theta) = -E \left[\frac{\partial^2 \ln p(Z_n / \theta_n)}{\partial \theta_i \cdot \partial \theta_j^*} \right]$

avec $\hat{R}_n = (Z_n - m_n) \cdot (Z_n - m_n)^+$

et $E[\hat{R}_n] = R_n$ $m_n = 0 \implies ds^2 = Tr \left((R_n^{-1} (dR_n))^2 \right)$

- Géométrie du demi-plan supérieur de C.L. Siegel :

$$SH_n = \{ Z = X + iY \in Sym(n, C) / \text{Im}(Z) = Y > 0 \}$$

$$ds_{Siegel}^2 = Tr \left(Y^{-1} (dZ) Y^{-1} (d\bar{Z}) \right) \text{ avec } Z = X + iY$$

$$\begin{cases} X = 0 \\ Y = i \cdot R_n \end{cases} \implies ds^2 = Tr \left((R_n^{-1} (dR_n))^2 \right)$$

Bonne métrique au sens de la géométrie de l'information



- La métrique est définie : $ds^2 = Tr\left(\left(R_n^{-1}(dR_n)\right)^2\right)$
 - Avantage vu de la Géométrie de l'information :
 - métrique invariante par changement non singulier de paramétrage

$$w = W(\theta) \Rightarrow ds^2(w) = ds^2(\theta)$$

- Métrique tenant compte de la statistique des paramètres

$$R_\theta = E\left[\left(\theta - \hat{\theta}\right)\left(\theta - \hat{\theta}\right)^+\right] \geq I[\theta]^{-1} : \text{borne de Cramer - Rao}$$

$$\text{et } ds^2 = d\theta^+ . I(\theta) . d\theta = d\theta^+ . R_\theta^{-1} . d\theta$$

Bonne métrique au sens de la géométrie des matrices

- La métrique est définie : $ds^2 = Tr\left(\left(R_n^{-1}(dR_n)\right)^2\right)$
- Avantage vu de la Géométrie de Siegel :
 - Les isométries de l'espace de Siegel SH_n sont données par le groupe quotient $PSp(n, R) \equiv Sp(n, R) / \{\pm I_{2n}\}$ avec $Sp(n, F)$ le groupe symplectique :

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \Rightarrow M(Z) = (AZ + B)(CZ + D)^{-1}$$

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in Sp(n, F) \Leftrightarrow \begin{cases} A^T C \text{ et } B^T D \text{ symétrique} \\ A^T D - C^T B = I_n \end{cases}$$

$$Sp(n, F) \equiv \left\{ M \in GL(2n, F) / M^T J M = J \right\}, J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \in SL(2n, R)$$

- La seule métrique invariante par $M(Z)$ est la métrique de Siegel :

$$ds_{Siegel}^2 = Tr\left(Y^{-1}(dZ)Y^{-1}(d\bar{Z})\right) \text{ avec } Z = X + iY$$

$$\begin{cases} X = 0 \\ Y = i.R_n \end{cases} \longrightarrow ds^2 = Tr\left(\left(R_n^{-1}(dR_n)\right)^2\right)$$

- La métrique est définie : $ds^2 = Tr\left(\left(R^{-1}(dR)\right)^2\right)$
- Distance associée :
 - Par intégration, on obtient :

$$d^2(R_1, R_2) = \left\| \log\left(R_1^{-1/2} \cdot R_2 \cdot R_1^{-1/2}\right) \right\|^2 = \sum_{k=1}^n \log^2(\lambda_k)$$

Avec $\det(R_2 - \lambda R_1) = 0$

- Dans le cas général, pour l'espace de Siegel

$$Z = X + iY \in SH_n \text{ avec } X \neq 0$$

$$d_{Siegel}^2(Z_1, Z_2) = \left(\sum_{k=1}^n \log^2\left(\frac{1 + \sqrt{\lambda_k}}{1 - \sqrt{\lambda_k}}\right) \right) \text{ avec } Z_1, Z_2 \in SH_n$$

Avec $\det(R(Z_1, Z_2) - \lambda \cdot I) = 0$

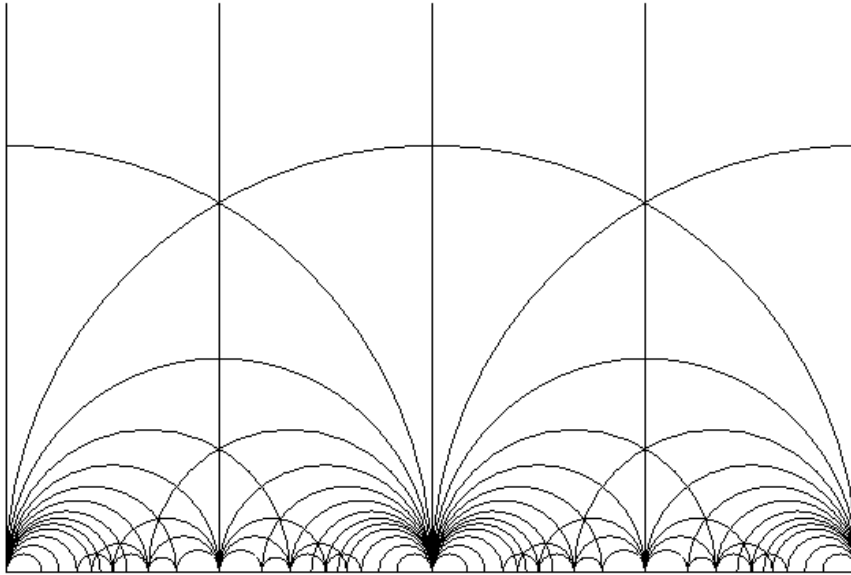
$$R(Z_1, Z_2) = (Z_1 - Z_2)(Z_1 - \bar{Z}_2)^{-1}(\bar{Z}_1 - \bar{Z}_2)(\bar{Z}_1 - Z_2)^{-1}$$

- Dans le cas de l'espace de Siegel :
 - Il suffit de remarquer que
 - la différentielle 2nd de $Z \rightarrow R(Z_1, Z)$ en $Z_1 = Z$ est donnée par l'expression :

$$R(Z_1, Z) = (Z_1 - Z)(Z_1 - \bar{Z})^{-1}(\bar{Z}_1 - \bar{Z})(\bar{Z}_1 - Z)^{-1}$$

$$\Rightarrow D^2R = 2dZ(Z - \bar{Z})^{-1}d\bar{Z}(\bar{Z} - Z)^{-1} = (1/2).dZY^{-1}d\bar{Z}Y^{-1}$$

$$\Rightarrow ds^2 = Tr(Y^{-1}dZY^{-1}d\bar{Z}) = 2.Tr(D^2R)$$



Demi-Plan supérieur de Poincare

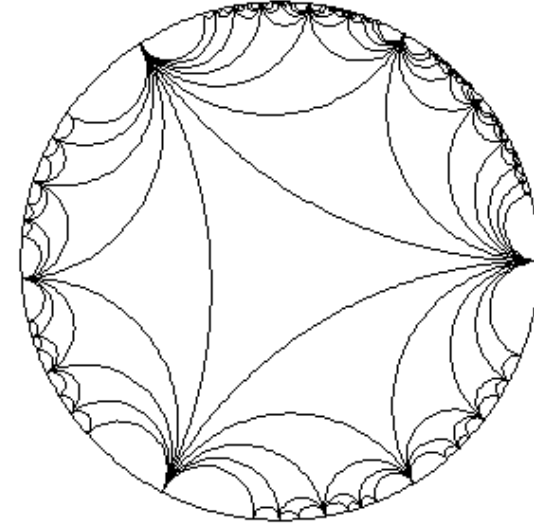
$$ds^2 = \frac{dx^2 + dy^2}{y^2} = \frac{|dz|^2}{y^2} = y^{-1} dzy^{-1} dz^*$$

avec $z = x + iy$

Demi-Plan supérieur de Siegel

$$ds^2 = Tr(Y^{-1} dZY^{-1} dZ^*)$$

avec $Z = X + iY$



Disque de Poincaré

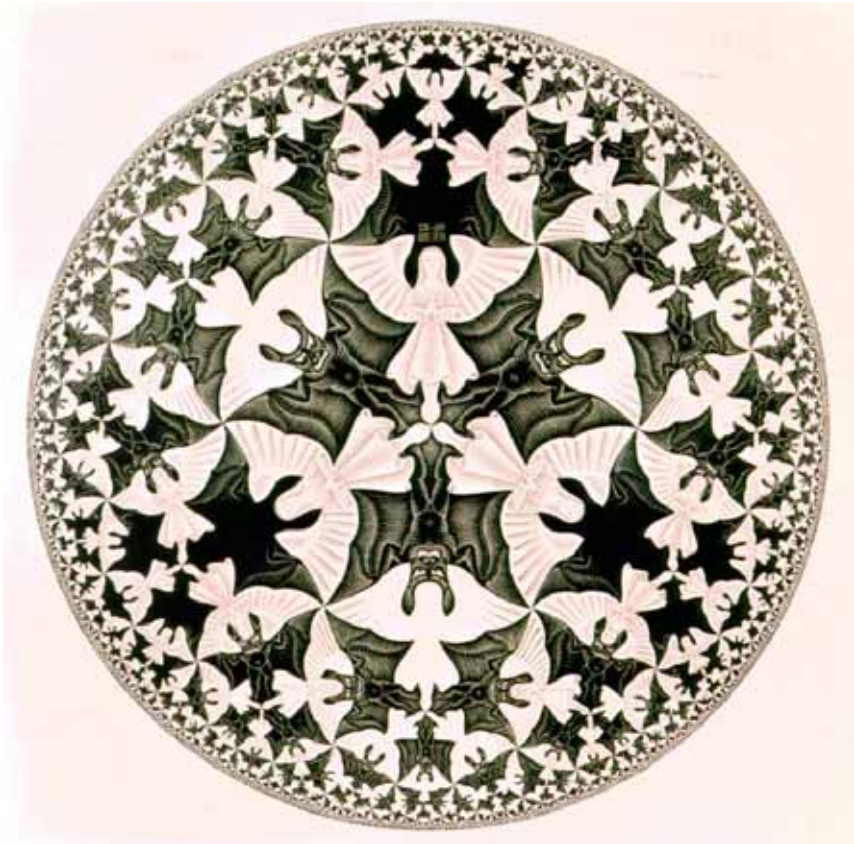
$$ds^2 = \frac{|dz|^2}{(1 - |z|^2)^2}$$

$$ds^2 = (1 - zz^*)^{-1} dz (1 - zz^*)^{-1} dz^*$$

Disque de Siegel

$$ds^2 = Tr \left[(1 - ZZ^*)^{-1} dZ (1 - ZZ^*)^{-1} dZ^* \right]$$

Utilisation dans l'art de la géométrie hyperbolique (Escher)



La géométrie du demi-plan de Siegel



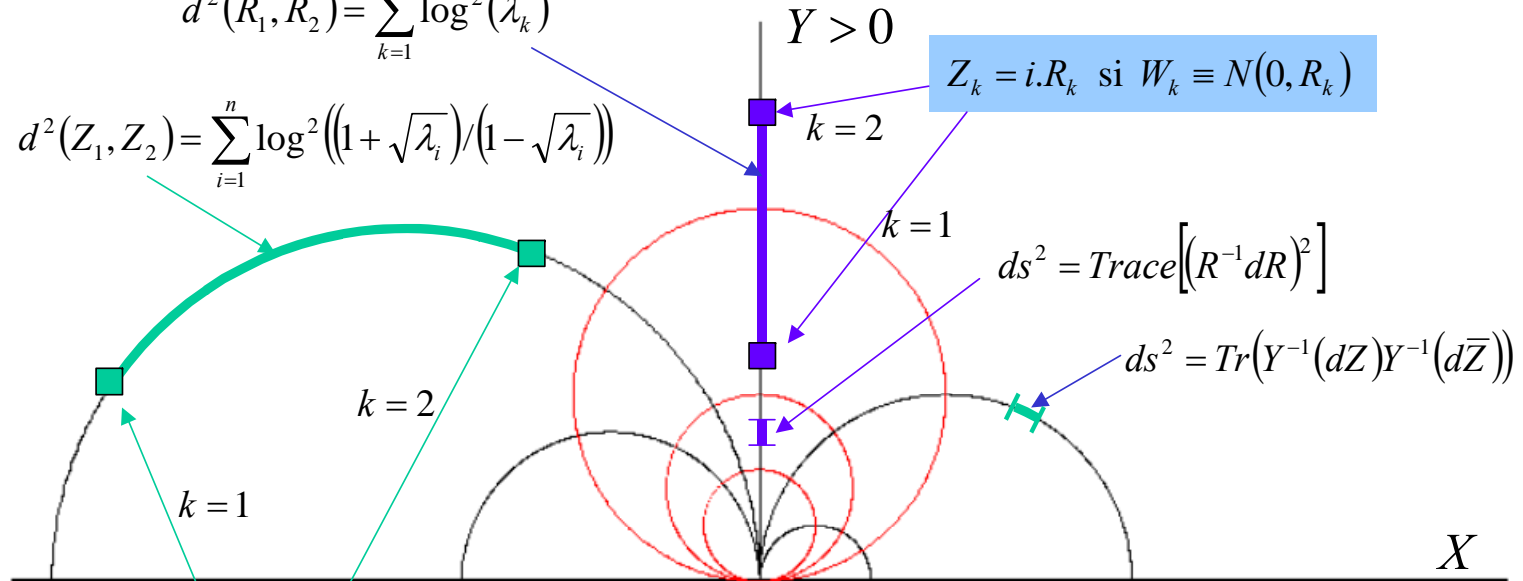
Demi-Plan de Siegel

$$SH_n = \{Z = X + iY \in \text{Sym}(n, \mathbb{C}) / \text{Im}(Z) = Y > 0\}$$

$$\det(R_1^{-1/2} \cdot R_2 \cdot R_1^{-1/2} - \lambda \cdot I) = 0$$

$$d^2(R_1, R_2) = \sum_{k=1}^n \log^2(\lambda_k)$$

$$d^2(Z_1, Z_2) = \sum_{i=1}^n \log^2\left(\frac{1 + \sqrt{\lambda_i}}{1 - \sqrt{\lambda_i}}\right)$$



$$R(Z_1, Z_2) = (Z_1 - Z_2)(Z_1 - \bar{Z}_2)^{-1}(\bar{Z}_1 - \bar{Z}_2)(\bar{Z}_1 - Z_2)^{-1}$$

$$\det(R(Z_1, Z_2) - \lambda \cdot I) = 0$$

THALES



La source de la géométrie de l'information



■ Fondements combinatoire de la divergence de Kullback

- La divergence de Kullback peut apparaître naturellement sur des bases combinatoires en utilisant la formule de Stirling :

Soit la loi multinomiale de N éléments répartis sur M niveaux $\{n_i\}$

$$P_M(n_1, n_2, \dots, n_M / q_1, \dots, q_M) = N! \prod_{i=1}^M \frac{q_i^{n_i}}{n_i!}$$

avec q_i les a priori , $\sum_{i=1}^M n_i = N$ et $p_i = \frac{n_i}{N}$

En utilisant la formule de Stirling : $n! \approx n^n \cdot e^{-n} \cdot \sqrt{2 \cdot \pi \cdot n}$ quand $n \rightarrow +\infty$

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \log[P_M] = \sum_{i=1}^M p_i \cdot \log \left[\frac{p_i}{q_i} \right] = K(p, q)$$

■ Fondements variationnels de la divergence de Kullback

- Donsker et Varadhan ont proposé une définition variationnelle de la divergence de Kullback :

$$K(p, q) = \sup_{\phi} \left[E_p(\phi) - \log E_q(e^{\phi}) \right]$$

Géométrie de Chentsov & métrique de Siegel



■ Géométrie de l'information et Divergence de Kullback (Rao, 1945)

- La métrique Riemannienne de l'information peut être obtenue en faisant un développement de Taylor à l'ordre 2 de la divergence de Kullback :

$$K[p(x/\theta), p(x/\theta + d\theta)] = \frac{1}{2!} \sum_{i,j} g_{ij}(\theta) \cdot d\theta_i \cdot d\theta_j + O(|d\theta|^3)$$

$$g_{ij}(\theta) = - \int p(x/\theta) \frac{\partial^2 \log p(x/\theta)}{\partial \theta_i \partial \theta_j} \cdot dx = -E \left[\frac{\partial^2 \log p(x/\theta)}{\partial \theta_i \partial \theta_j} \right]$$

■ Géométrie de l'information et matrice de Fisher

- La métrique Riemannienne de l'information s'identifie à la matrice d'information de Fisher $I(\theta)$:

$$I(\theta) = [g_{ij}(\theta)]_{i,j} \quad \text{et} \quad g_{ij}(\theta) = E \left[\frac{\partial \log p(x/\theta)}{\partial \theta_i} \cdot \frac{\partial \log p(x/\theta)}{\partial \theta_j} \right]$$

$$ds^2 = d\theta^T \cdot I(\theta) \cdot d\theta = \text{Var}[d \log p(x/\theta)] = E[(d \log p(x/\theta))^2]$$

- La matrice de Fisher intervient dans la borne de Cramer-Rao :

$$E \left[(\theta - \hat{\theta})(\theta - \hat{\theta})^T \right] \geq I(\theta)^{-1}$$

$$w = W(\theta) \Rightarrow ds^2(w) = ds^2(\theta)$$

Exemple du cas scalaire : distribution gaussienne

■ Distribution exponentielle : cas loi Gaussienne

- La matrice de l'information de Fisher est donnée par :

$$I(\theta) = \sigma^{-2} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \text{ avec } E[\theta\theta^T] \geq I(\theta)^{-1} \text{ et } \theta = \begin{pmatrix} m \\ \sigma \end{pmatrix}$$

- Cette matrice définit la métrique dans l'espace des paramètres :

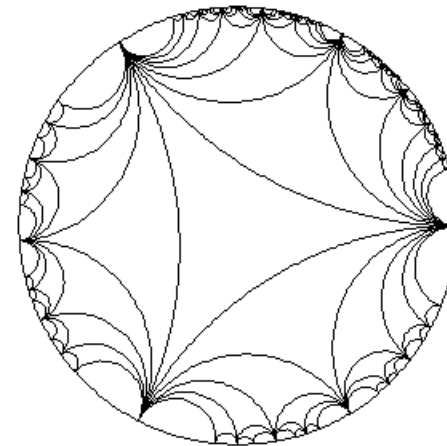
$$ds^2 = d\theta^T \cdot I(\theta) \cdot d\theta = \frac{dm^2}{\sigma^2} + 2 \cdot \frac{d\sigma^2}{\sigma^2} = 2 \cdot \sigma^{-2} \left[\left(\frac{dm}{\sqrt{2}} \right)^2 + (d\sigma)^2 \right]$$

- Il s'agit du modèle de Géométrie hyperbolique de Poincaré :

$$z = \frac{m}{\sqrt{2}} + i \cdot \sigma \quad \omega = \frac{z - i}{z + i} \quad (|\omega| < 1)$$

$$\Rightarrow ds^2 = 8 \cdot \frac{|d\omega|^2}{(1 - |\omega|^2)^2}$$

La géométrie des distributions Gaussiennes est celle du disque de Poincaré



Exemple du cas scalaire : distribution gaussienne

■ Distribution exponentielle : cas loi Gaussienne

- La métrique est donnée comme suit avec $r = |\omega|$, et en intégrant sur une radiale :

$$ds^2 = 8 \cdot \left(\frac{dr}{1-r^2} \right)^2 = 2 \cdot d \ln \frac{1+r}{1-r}$$

- On utilise ensuite la transformation suivante pour exprimer la métrique entre 2 points du disque unité :

$$v = \phi_\tau(\omega) = \frac{\omega - \tau}{\bar{\tau}\omega - 1} \cdot e^{j \cdot \varphi} \quad \text{et} \quad 0 = \phi_\tau(\tau)$$

- La distance est finalement donnée par :

$$D^2(\{m_1, \sigma_1\}, \{m_2, \sigma_2\}) = 2 \cdot \left(\ln \frac{1 + \delta(\omega, \tau)}{1 - \delta(\omega, \tau)} \right)^2 \quad \text{avec} \quad \delta(\omega, \tau) = \left| \frac{\omega - \tau}{1 - \omega\bar{\tau}} \right|$$

$$z = \frac{m}{\sqrt{2}} + i \cdot \sigma \quad \text{et} \quad \omega = \frac{z - i}{z + i}$$

On considère le signal suivre une loi multivariée gaussienne circulaire :

$$p(X_n / R_n) = (\pi)^{-n} \cdot |R_n|^{-1} \cdot e^{-Tr[\hat{R}_n \cdot R_n^{-1}]}$$

avec $\hat{R}_n = (X_n - m_n) \cdot (X_n - m_n)^+$ et $E[\hat{R}_n] = R_n$

La métrieque de Fisher est donnée par :

$$g_{ij}(\theta) = -E \left[\frac{\partial \ln p(X_n / \theta_n)}{\partial \theta_i \cdot \partial \theta_j} \right] = -Tr[\partial_i R_n \cdot \partial_j R_n^{-1}] + \partial_i m_n^+ \cdot R_n^{-1} \cdot \partial_j m_n$$

On considère le processus de moyenne nulle :

$$R_n \cdot R_n^{-1} = I_n \Rightarrow \mathcal{R}_n = -R_n \cdot \mathcal{R}_n^{-1} \cdot R_n \quad \text{then} \quad g_{ij}(\theta) = Tr \left[\left(R_n \cdot \partial_i R_n^{-1} \right) \cdot \left(R_n \cdot \partial_j R_n^{-1} \right) \right]$$

$$\Rightarrow ds^2(\theta) = Tr \left[R_n \cdot \left(\sum_i \partial_i R_n^{-1} \cdot d\theta_i \right) \cdot R_n \cdot \left(\sum_j \partial_j R_n^{-1} \cdot d\theta_j \right) \right]$$

On en déduit la métrieque de Rao :

comme $dR_n^{-1} = \sum_i \partial_i R_n^{-1} d\theta_i \Rightarrow$

$$ds^2 = Tr \left[\left(R_n dR_n^{-1} \right)^2 \right] = Tr \left[\left(R_n^{-1} dR_n \right)^2 \right]$$



■ Gaussiennes multivariées de moyenne nulle

- Ces distributions sont uniquement paramétrées par la matrice de covariance et la métrique est celle de Carl Ludwig Siegel

$$ds^2 = Tr\left(\left(R^{-1}dR\right)^2\right) = Tr\left(\left(d \log R\right)^2\right) = \left\|R^{-1/2}dRR^{-1/2}\right\|^2$$

**Métrique de
Carl Ludwig Siegel**

avec $\|A\|^2 = \langle A, A \rangle$ et $\langle A, B \rangle = Tr(AB^T)$

Comme $R^{-1}R = I \Rightarrow dR^{-1}R = -R^{-1}dR$

Alors $ds_R^2 = ds_{R^{-1}}^2$

**Invariance par rapport
à l'inversion**

$$D(R_1, R_2) = D(R_1^{-1}, R_2^{-1})$$

■ Gaussiennes multivariées de moyenne non nulle

- Les isométries dans le cas de moyenne non nulle sont les homéomorphismes suivants :

$$(m, R) \rightarrow (m', R') = (A^T m + a, A^T R A) \text{ isométrie}$$

$$ds^2 \mapsto ds'^2 = ds^2 \quad \text{avec } (a, A) \in R^n \times GL(n, R)$$

$$ds^2 = dm^T R^{-1} dm + Tr\left(\left(R^{-1}dR\right)^2\right)$$



De la métrique de Siegel à la distance de Siegel

- la distance de Siegel déduite de sa métrique est obtenue à partir des valeurs propres étendues dans le cas $\{m_1 = m_2\}$:

$$ds^2 = \left\| R^{-1/2} dR R^{-1/2} \right\|^2$$

Métrique de Siegel

$$\log(A) = \sum_{n=1}^{\infty} (-1)^{n+1} \cdot \frac{(A-I)^n}{n}$$

$$D^2(R_1, R_2) = \left\| \log(R_1^{-1/2} \cdot R_2 \cdot R_1^{-1/2}) \right\|^2 = \sum_{k=1}^n \log^2(\lambda_k)$$

avec $\det(R_1^{-1/2} \cdot R_2 \cdot R_1^{-1/2} - \lambda \cdot I) = \det(R_2 - \lambda R_1) = 0$

$$\|\Sigma\|^2 = \langle \Sigma, \Sigma \rangle \quad \text{avec} \quad \langle A, B \rangle = \text{Tr}[AB^T]$$

Théorème de James

- Lorsque $\Sigma_1 = \Sigma_2$, alors $\frac{n}{2} \sum_{k=1}^n \log^2(\lambda_k)$ suit une loi du Khi-2 à $n \cdot (n+1) / 2$ degrés de liberté.

Généralisation de Swain (fonction de contraste)

- Expression de la distance à partir d'une fonction de contraste C^3 sur $[0, \infty[$:

$$D^2(R_1, R_2) = \sum_{k=1}^n v(\lambda_k)$$

$$v(1) = v'(1) = 0, \quad v''(\lambda) = 1/2$$

et $v(\lambda) > 0$ si $\lambda > 0, \lambda \neq 1$

$$D_v^2(\Sigma_1, \Sigma_2) = D_w^2(\Sigma_1, \Sigma_2) \quad \text{si} \quad w(\lambda) = v(\lambda^{-1})$$

THALES



**La source de la géométrie
intrinsèque de l'espace de
Siegel**

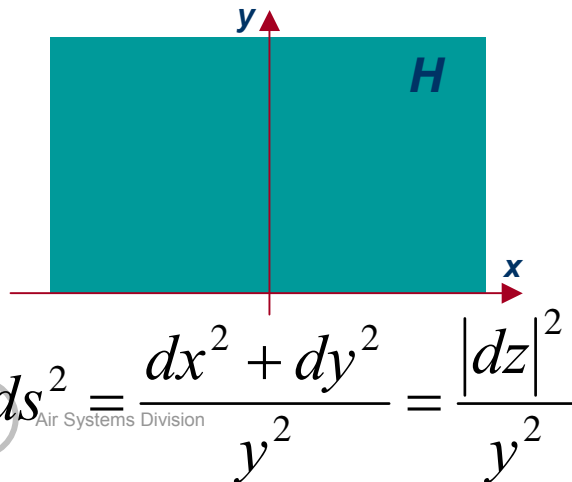


Opérations du groupe SL_2R sur le plan hyperbolique

- Les travaux de Carl Ludwig Siegel exposé dans son ouvrage « Symplectic Geometry » sont une généralisation de SL_2R sur le demi-plan hyperbolique de Poincaré
- La transformation de Möbius est une action transitive qui transforme le demi-plan complexe en lui-même (espace homogène) :

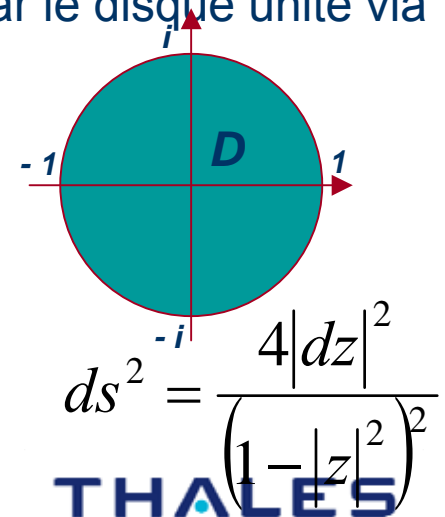
$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2R \quad , \quad M(z) = \frac{az + b}{cz + d} \quad \text{et} \quad z \in H = \{z \in C : \text{Im}(z) > 0\}$$

- M et $-M$ ont la même action et on considère l'action du groupe quotient : $PSL_2R = SL_2R / \langle \pm I_2 \rangle$
- Cet espace a également un autre modèle donné par le disque unité via la transformation de Cayley :



$$D = \{z \in C : |z| < 1\}$$

$$\left| \begin{array}{l} H \rightarrow D \\ z \mapsto \frac{z-i}{z+i} \end{array} \right. \quad \text{et} \quad \left| \begin{array}{l} D \rightarrow H \\ z \mapsto i \frac{1+z}{1-z} \end{array} \right.$$





Opérations du groupe $SL_2\mathbb{R}$ sur le plan hyperbolique

- Ces modèles peuvent être compactifié (tel que chaque transformation ait un point fixe) :

$$Cl(H) = \{z \in \mathbb{C} : \text{im}(z) > 0\} \cup \{\infty\} \text{ et } Cl(D) = \{z \in \mathbb{C} : |z| \leq 1\}$$

- Toute matrice de $SL_2\mathbb{R}$, non égale à $\pm I_2$ est conjuguée dans $GL_2\mathbb{R}$ à l'une des 3 suivantes (transformations hyperbolique/elliptique/parabolique) :

| | | Point(s) fixe(s) |
|---|------------------------------------|----------------------------|
| $\sigma(M) \neq \pm 1 \in \mathbb{R} \Rightarrow g_\lambda = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}$ | $z \mapsto \lambda^2 z$ | $0, \infty \in \partial H$ |
| $\sigma(M) = \pm 1 \Rightarrow \pm g_1 = \pm \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ | $z \mapsto z + 1$ | $\infty \in \partial H$ |
| $\sigma(M) = a + ib$ $a^2 + b^2 = 1, b \neq 0 \Rightarrow g_{a+ib} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ | $z \mapsto \frac{az + b}{-bz + a}$ | $i \in H$ |

Géométrie Symplectique et Espace de Siegel



■ Définition du groupe Symplectique :

- Le groupe Symplectique est défini comme suit :

$Sym(n, F) \subset M(n, F)$: espace des matrices $n \times n$ symétriques ($F = R, C$)

$$Sp(n, F) \equiv \left\{ M \in GL(2n, F) / M^T J M = J \right\}, J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \in SL(2n, R)$$

de façon équivalente : $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in Sp(n, F) \Leftrightarrow M^{-1} = \begin{pmatrix} D^T & -B^T \\ -C^T & A^T \end{pmatrix}$

$\Leftrightarrow A^T C$ et $B^T D$ symétriques et $A^T D - C^T B = I_n$

- Siegel a introduit l'espace SH_n « $n^{\text{ième}}$ demi-plan supérieur de Siegel ». SH_1 est le demi-plan supérieur hyperbolique H .

$$SH_n = \{ Z = X + iY \in Sym(n, C) / \text{Im}(Z) = Y > 0 \}$$

- L'espace de Siegel est l'espace homogène correspondant au groupe symplectique $Sp(n, R)$ quotient par le sous-groupe compact maximal $K_n \equiv Sp(n, R) \cap SO(2n, R)$. Le groupe $Sp(n, R) / K_n$ est une sous variété de $X_{2n} = GL(2n, C) / U_{2n}$



■ Espace homogène

- Le groupe $PSp(n, R) \equiv Sp(n, R) / \{\pm I_{2n}\}$ est le groupe des biholomorphismes de SH_n via les transformations de Möbius généralisées :

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \Rightarrow M(Z) = (AZ + B)(CZ + D)^{-1}$$

- On montre que $M(Z)$ appartient à SH_n : $M(Z) \in SH_n$

$$M \begin{bmatrix} Z \\ I \end{bmatrix} = \begin{bmatrix} E \\ F \end{bmatrix} = \begin{bmatrix} AZ + B \\ CZ + D \end{bmatrix}$$

$$\begin{cases} M^T J M = J \\ Z \in SH_n \quad (Z^T = Z, \text{Im}(Z) > 0) \end{cases} \Rightarrow \begin{cases} F \text{ invertible} \Rightarrow M(Z) = EF^{-1} \\ EF^{-1} \text{ symétrique} \\ \text{Im}(EF^{-1}) > 0 \end{cases}$$

- L'action du groupe symplectique sur SH_n est transitive :

$$\phi_1(Z) \equiv \left. \begin{aligned} & \begin{pmatrix} I_n & X \\ 0 & I_n \end{pmatrix} \begin{pmatrix} Y^{1/2} & 0 \\ 0 & Y^{-1/2} \end{pmatrix} = \begin{pmatrix} Y^{1/2} & XY^{-1/2} \\ 0 & Y^{-1/2} \end{pmatrix} \\ & Z \mapsto Y^{1/2} Z Y^{1/2} \quad \text{et} \quad Z \mapsto Z + X \end{aligned} \right\} \Rightarrow \phi_1(iI) = X + iY$$

Géométrie Symplectique et Espace de Siegel

■ Action sur le demi-plan de siegel

- Si on utilise que $Z \in SH_n$ alors $Z^T = Z, \text{Im}(Z) = \frac{1}{2i}(Z - \bar{Z}) > 0$

$$M \begin{bmatrix} Z \\ I \end{bmatrix} = \begin{bmatrix} E \\ F \end{bmatrix} = \begin{bmatrix} AZ + B \\ CZ + D \end{bmatrix} \quad \text{et} \quad M^T J M = J$$

$$\left\{ \begin{array}{l} \begin{bmatrix} E \\ F \end{bmatrix}^T J \begin{bmatrix} E \\ F \end{bmatrix} = J \Rightarrow \begin{bmatrix} Z^T & I \end{bmatrix} J \begin{bmatrix} Z \\ I \end{bmatrix} = 0 \quad \text{car } Z^T = Z \end{array} \right.$$

$$\left\{ \begin{array}{l} -\frac{1}{2i} \begin{bmatrix} \bar{E} & \bar{F} \end{bmatrix} J \begin{bmatrix} E \\ F \end{bmatrix} = -\frac{1}{2i} \begin{bmatrix} \bar{Z} & I \end{bmatrix} J \begin{bmatrix} Z \\ I \end{bmatrix} > 0 \quad \text{car } \text{Im}(Z) > 0 \end{array} \right.$$

$$\left\{ \begin{array}{l} E^T F = F^T E \quad (1) \end{array} \right.$$

$$\left\{ \begin{array}{l} -\frac{1}{2i} (\bar{E}F - \bar{F}E) > 0 \quad (2) \end{array} \right.$$

si $Fv = 0 \Rightarrow \bar{v}\bar{F} = 0 \Rightarrow \bar{v}(\bar{E}F - \bar{F}E)v = 0 \Rightarrow v = 0 \Rightarrow F$ invertible

(1) $\Rightarrow EF^{-1}$ symétrique et (2) $\Rightarrow \text{Im}(EF^{-1}) > 0$



■ Algèbre Linéaire Symplectique

- Une matrice symplectique est de déterminant 1 : $\det(M) = 1$
- Soit $M \in Sp_{2n}R$, on appelle (et la transformation associée) hyperbolique si M n'a aucune valeur propre sur le cercle unité complexe : $\sigma(M) = \{\lambda_i\}$ et $|\lambda_i| \neq 1, \forall i = 1, \dots, n$
- Soit M une matrice Symplectique sans valeur propre de module unité, alors M est symplectiquement similaire à une matrice de la forme :

$$\begin{pmatrix} A & 0 \\ 0 & A^{-T} \end{pmatrix}$$

- Le sous-groupe $Sp_{2n}R \cap O_{2n}$ (matrices symplectiques orthogonales) est caractérisé par des matrices de la forme :

$$\begin{pmatrix} A & B \\ -B & A \end{pmatrix} \text{ avec } A^T A + B^T B = I_n \text{ et } A^T B \text{ symétrique}$$



■ Métrique de Siegel

- Le groupe $PSp(n, R) \equiv Sp(n, R) / \{\pm I_{2n}\}$ est le groupe des biholomorphismes de SH^n via les transformations de Möbius généralisées :

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \Rightarrow M(Z) = (AZ + B)(CZ + D)^{-1}$$

- $PSp(n, R)$ agit comme un sous-groupe d'isométries :

$$ds_{Siegel}^2 = Tr\left(Y^{-1}(dZ)Y^{-1}(d\bar{Z})\right) \text{ avec } Z = X + iY$$

$$Z' = (AZ + B)(CZ + D)^{-1} \Rightarrow Tr\left(Y'^{-1}(dZ')Y'^{-1}(d\bar{Z}')\right) = Tr\left(Y^{-1}(dZ)Y^{-1}(d\bar{Z})\right)$$

- La distance associée est :

$$d_{Siegel}(Z_1, Z_2) = \left(\sum_{k=1}^n \log^2 \left(\frac{1 + \sqrt{r_k}}{1 - \sqrt{r_k}} \right) \right)^{1/2}$$

- Avec les r_k les valeurs propres du bi-rapport :

$$R(Z_1, Z_2) = (Z_1 - Z_2)(Z_1 - \bar{Z}_2)^{-1}(\bar{Z}_1 - \bar{Z}_2)(\bar{Z}_1 - Z_2)^{-1}$$

- Soit (A,B) et (C,D) deux paires de points dans $Sp_{2n}R / K$, alors il existe une transformation symplectique qui transforme A en C et B en D si et seulement si : $\sigma_j(A^{-1}B) = \sigma_j(C^{-1}D) \quad \forall j$

Géométrie Symplectique et Espace de Siegel



Invariants différentiels dans l'espace de Siegel

- On cherche des expressions différentielles invariantes par le groupe $G = \text{Sp}(n, \mathbb{R})$: $Z' = (AZ + B)(CZ + D)^{-1}$

$$\left. \begin{aligned} dZ' &= (CZ + D)^{-1t} (dZ)(CZ + D)^{-1} \\ \text{Im}(Z') &= (C\bar{Z} + D)^{-1t} \text{Im}(Z)(CZ + D)^{-1} \end{aligned} \right\} \text{ car } A^t D - C^t B = I_n$$

- Etant donné $Z = X + iY$, les matrices les plus générales qui transforment Z dans le point iI_n sont celles de la forme :

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \begin{bmatrix} Y^{-1/2} & -Y^{-1/2} X \\ 0 & Y^{1/2} \end{bmatrix} = \begin{bmatrix} \alpha Y^{-1/2} & \beta Y^{1/2} - \alpha Y^{-1/2} X \\ -\beta Y^{1/2} & \beta Y^{-1/2} X + \alpha Y^{1/2} \end{bmatrix}$$

$(\alpha - i\beta)$ unitaire , $CZ + D = (\alpha - i\beta)Y^{1/2}$ et $C\bar{Z} + D = (\alpha + i\beta)Y^{1/2}$

- On en déduit :

$$dZ' = (\alpha - i\beta)^{-1t} Y^{-1/2} (dZ) Y^{-1/2} (\alpha - i\beta)^{-1}$$

$$\text{Im}(Z') = Y' = (\alpha + i\beta)^{-1t} Y^{-1/2} \text{Im}(Z) Y^{-1/2} (\alpha - i\beta)^{-1} = Y^{-1/2} \text{Im}(Z) Y^{-1/2} = I_n$$

- ainsi que l'identité :

$$\text{Tr}(Y'^{-1} dZ' Y'^{-1} d\bar{Z}') = \text{Tr}(dZ' d\bar{Z}')$$

$$= \text{Tr}(Y^{1/2} (dZ) Y^{-1} (d\bar{Z}) Y^{-1/2}) = \text{Tr}(Y^{-1} dZ Y^{-1} d\bar{Z})$$

Géométrie Symplectique et Espace de Siegel



■ Invariants différentiels dans l'espace de Siegel

- Nous venons de montrer que la formule différentielle (quadratique hermitienne) suivante est invariante par le groupe $G=Sp(n,R)$:

$$ds^2 = Tr\left(Y^{-1}(dZ)Y^{-1}(d\bar{Z})\right)$$

- Cette expression définit un ds^2 hermitien défini positif invariant par G .
- Cette métrique ds^2 est une métrique Kählérienne

$$\Omega = \frac{1}{2i} Tr\left(Y^{-1}dZ \wedge Y^{-1}d\bar{Z}\right) = Tr\left(Y^{-1}dYY^{-1} \wedge dX\right)$$

$$Y^{-1}(dY)Y^{-1} = -d(Y^{-1}) \Rightarrow \Omega = Tr\left(dX \wedge d(Y^{-1})\right) \Rightarrow d\Omega = 0$$

- L'élément de volume dans l'espace de Siegel invariant par G est donné par:

$$\left. \begin{array}{l} X = (x_{ij})_{i,j} \\ Y^{-1} = (y'_{ij})_{i,j} \end{array} \right\} \Rightarrow \Omega = \sum_i \delta x_{ii} \wedge \delta y'_{ii} + 2 \sum_{i < j} \delta x_{ij} \wedge \delta y'_{ij}$$

- En élevant à la puissance extérieure d'ordre $n(n+1)/2$, le volume invariant par G est donné par :

$$\left(\prod_{i \leq j} \delta x_{ij} \right) \wedge \left(\prod_{i \leq j} \delta y'_{ij} \right) \text{IALES}$$

Géométrie Symplectique et Espace de Siegel



■ Métrique de Siegel

- $PSp(n, R)$ agit comme un sous-groupe d'isométries :

$$ds_{Siegel}^2 = Tr\left(Y^{-1}(dZ)Y^{-1}(d\bar{Z})\right) \text{ avec } Z = X + iY$$

- Lien avec la métrique $d_p^{i/p}$:

$$d_p(A, B) = \left(\sum_{i=1}^n \left| \log \sigma_j(A^{-1}B) \right|^p \right)^{1/p} \xrightarrow{p \rightarrow \infty} \text{Max} \left\{ \left| \log \sigma_1(A^{-1}B) \right|, \left| \log \sigma_n(A^{-1}B) \right| \right\}$$

- La bijection suivante relie la métrique de Siegel à la métrique d_2 :

$$\begin{aligned} \Phi_1: SH_n &\rightarrow Sp(n, R)/K_n \\ Z &\mapsto \phi_1(Z)K_n \end{aligned} \quad O \in K_n = \text{Stab}(iI_n) \text{ et } O = \begin{pmatrix} A & B \\ -B & A \end{pmatrix}$$

avec $\phi_1: SH_n \rightarrow Sp(n, R)$ avec $-A^T = A$ et $B^T = B$

$$\phi_1(Z) \equiv \begin{pmatrix} I_n & X \\ 0 & I_n \end{pmatrix} \begin{pmatrix} Y^{1/2} & 0 \\ 0 & Y^{-1/2} \end{pmatrix} = \begin{pmatrix} Y^{1/2} & XY^{-1/2} \\ 0 & Y^{-1/2} \end{pmatrix}$$

$$ds_{Siegel}(Z, W) = \sqrt{2}d_2(\phi_1(Z), \phi_1(W)) \quad \forall Z, W \in SH_n$$



Autres modèles de l'Espace de Siegel

- La compactification naturelle de SH_n est donnée par SD_n le domaine bornée qui est une extension du disque unité de Poincaré, SD_n biholomorphe à SH_n .

$$SD_n = \{Z \in \text{Sym}(n, \mathbb{C}) / \|Z\|_2 < I\} \text{ avec } \|Z\|_2 = Z\bar{Z}$$

- La frontière de SD_n est appelée frontière de Shilov, c'est l'ensemble des matrices nxn unitaires symétriques $USym(n)$.
- Les deux espaces sont reliés par la transformée de Cayley :

$$\Phi_2: SH_n \rightarrow SD_n$$

$$Z \mapsto (Z - iI_n)(Z + iI_n)^{-1}$$

$$\Phi_2^{-1}: SD_n \rightarrow SH_n$$

$$Z \mapsto i(I_n + Z)(I_n - Z)^{-1}$$

- Les biholomorphismes de SD_n sont données par les transformations de Möbius généralisées induites par le sous-groupe $Sp(n, R)'$:

$$Sp(n, R)' \equiv Sp(n, \mathbb{C}) \cap SU(n, n)$$

$$SU(n, n) \equiv \{M \in SL(2n, \mathbb{C}) / M^* \text{diag}(I_n, -I_n)M = \text{diag}(I_n, -I_n)\}$$

$$M \in \text{Stab}(0) \equiv \{M \in Sp(n, R)' / M(0) = 0\}$$

$$\Leftrightarrow M(Z) = UZU^T, U \in U_n, Z \in SD_n$$

■ Autre modèle de l'espace de Siegel

- Modèle projectif de SH_n :

$$\Phi_3: SH_n \rightarrow SPH_n$$

$$Z \mapsto \begin{bmatrix} Z \\ I_n \end{bmatrix} \quad \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{bmatrix} Z \\ I_n \end{bmatrix} = \begin{bmatrix} AZ + B \\ CZ + D \end{bmatrix} = \begin{bmatrix} (AZ + B)(CZ + D)^{-1} \\ I_n \end{bmatrix}$$

■ Remarque : matrice triangulaire supérieure

- Soit G un sous-groupe de matrices triangulaires supérieures de type blocs 2x2 dans $Sp(n, R)$ alors G est généré par les translations et les congruences :

$$Z \mapsto T(Z) = Z + B, \quad B \in \text{Sym}(n, R), \quad T = \begin{pmatrix} I_n & B \\ 0 & I_n \end{pmatrix}$$

$$Z \mapsto Q(Z) = AZA^T, \quad A \in GL(n, R), \quad Q = \begin{pmatrix} A & 0 \\ 0 & (A^T)^{-1} \end{pmatrix}$$

Géométrie Symplectique et Espace de Siegel



■ Lien avec le cône symétrique

- Pour nous ramener à notre cas d'étude, prenons $x=0$:

$$Z = iY \quad (X = 0) \quad \text{avec} \quad \text{Im}(Z) = Y > 0$$

- La correspondance entre les points de l'espace de Siegel et les matrices P symplectiques symétriques et positives :

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \quad (PJ)^2 = -I_{2n}$$

$$PJL = -iL \quad \text{avec} \quad L = \begin{bmatrix} iY \\ I_n \end{bmatrix} \quad \text{et} \quad P = \begin{pmatrix} Y & 0 \\ 0 & Y^{-1} \end{pmatrix}$$

- Autre représentation de l'espace de Siegel :

$$W = (I_n - Y)(I_n + Y)^{-1} \quad \text{avec} \quad \overline{W}W < 1$$

- ds^2 hermitien défini positif, définit une métrique Kählérienne invariant par $G = \text{Sp}(n, \mathbb{R})$:

$$ds^2 = \text{Tr}(Y^{-1}(dY)Y^{-1}(dY))$$

THALES



Moyenne (géométrique symétrisée) & Géodésique

Moyenne Géométrique de matrices définies positives



Rappel sur la moyenne géométrique de nombres réels positifs

■ Définitions de la moyenne géométrique de nombres réels positifs :

1) $x = \sqrt{ab}$ *moyenne géométrique*

2) solution de $xa^{-1}x = b$

3) $\sqrt{ab} = \lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n$ avec $a_0 = a$, $b_0 = b$ et $a_{n+1} = \frac{2a_n b_n}{a_n + b_n}$ et $b_{n+1} = \frac{a_n + b_n}{2}$

$\min\{a, b\} \leq \frac{2ab}{a+b} \leq \sqrt{ab} \leq \frac{a+b}{2} \leq \max\{a, b\}$ (*harmonique, géométrique et arithmétique*)

4) $\text{Max}_{x \in \mathbb{R}}(x)$ tel que $\begin{pmatrix} a & x \\ x & b \end{pmatrix} > 0$

5) $\delta(a, b) = 2\delta(a, \sqrt{ab}) = 2\delta(b, \sqrt{ab})$ avec $\delta(x, y) = |\log(x/y)|$

6) $e^{\frac{x+y}{2}} = \sqrt{ab}$ avec $a = e^x, b = e^y$

7) point fixe positif de la transformée de Möbius : $X = \begin{pmatrix} 0 & b \\ a^{-1} & 0 \end{pmatrix}$

8) $\sqrt{ab} = \lim_{t \rightarrow \infty} x(t)$ avec $\frac{dx(t)}{dt} = b - (1/a)x(t)$ et $x(0) = x_0 > 0$

9) $x > 0$ tel que $F''(x)(a) = 1/b$ avec $F(x) = -\log x$

Moyenne arithmetico-géométrique : C.F. Gauss

- Gauss (à l'âge de 20 ans) & Lagrange indépendamment ont prouvés :

- Moyenne arithmético-géométrique de a et b :

$$\forall a \geq b > 0, (a_0, b_0) = (a, b) \text{ et } \begin{cases} a_{n+1} = \frac{a_n + b_n}{2} \\ b_{n+1} = \sqrt{a_n b_n} \end{cases}$$

$$a_{n+1}^2 - b_{n+1}^2 = \left(\frac{a_n + b_n}{2} \right)^2 - a_n b_n = \left(\frac{a_n - b_n}{2} \right)^2 \geq 0$$

$$\Rightarrow \forall n \geq 0 \quad b = b_0 \leq \dots \leq b_n \leq b_{n+1} \leq a_{n+1} \leq a_n \leq \dots \leq a_0 = a$$

- Est reliée à l'intégrale :

$$I(a, b) = \frac{\pi/2}{M(a, b)} \text{ avec } \begin{cases} M(a, b) = \lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n \\ I(a, b) = \int_0^{\pi/2} \frac{dt}{\sqrt{a^2 \cos^2 t + b^2 \sin^2 t}} \end{cases}$$

- La transformation de Landen donne :

$$\begin{cases} I(a_1, b_1) = I\left(\frac{a+b}{2}, \sqrt{ab}\right) = I(a, b) \\ 2J(a_1, b_1) = J(a, b) + abI(a, b) \end{cases} \text{ avec } J(a, b) = \int_0^{\pi/2} \sqrt{a^2 \cos^2 t + b^2 \sin^2 t} dt$$

$$\text{Landen transf.: } u = t + \text{Arc tan}\left(\frac{b}{a} \tan t\right)$$

4.J(a,b) : Périmètre de l'ellipse de demi-axes a et b

Moyenne Géométrique de matrices définies positives

■ Extension de la moyenne géométrique à des matrices définies positives :

- Le procédé itératif présenté pour les nombres réels positifs s'étend aux matrices symétriques définies positives :

$$A \circ B = \lim_{n \rightarrow \infty} A_n = \lim_{n \rightarrow \infty} B_n \text{ avec } A_0 = A, B_0 = B$$

$$A_{n+1} = 2(A_n^{-1} + B_n^{-1})^{-1} \text{ et } B_{n+1} = (A_n + B_n)/2$$

- La moyenne géométrique est solution de l'équation de l'EDP :

$$\forall A, B \in \text{Sym}^{++}, A \circ B = \lim_{t \rightarrow \infty} X(t) \text{ avec } \begin{cases} \frac{dX}{dt} = B - XA^{-1}X \\ X(0) \in \text{Sym}^{++} \end{cases}$$

- La moyenne géométrique est solution de :

$$\forall A, B \in \text{Sym}^{++}, A \circ B \text{ plus grande matrice (Loewner) t.q. } \begin{pmatrix} A & X \\ X & B \end{pmatrix} > 0$$

- La moyenne géométrique est également donnée par :

$$F''(X)(A) = X^{-1}AX = B^{-1} \text{ avec } F(X) = -\log \det(X)$$

moyenne géométrique symétrisée

- La moyenne qui apparaît est bien une matrice symétrique ou hermitienne définie positive :

$$A \circ B = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^{1/2} A^{1/2}$$

- Il s'agit bien du barycentre de Fréchet, qui minimise :

$$\text{Min}_X \left(\left\| \log \left(A^{-1/2} X A^{-1/2} \right) \right\|^2 + \left\| \log \left(B^{-1/2} X B^{-1/2} \right) \right\|^2 \right)$$

- On peut remarquer que quand les matrices commutent, il s'agit de la moyenne géométrique classique :

$$A \circ B = (AB)^{1/2} \text{ qui n'est pas symétrique définie positive}$$

- Une autre façon de retrouver la moyenne géométrique « symétrisée » consiste à la trouver comme solution de l'équation de Riccati :

$$X A^{-1} X = B \Rightarrow A^{-1/2} (X A^{-1/2} A^{-1/2} X) A^{-1/2} = A^{-1/2} B A^{-1/2}$$

$$\Rightarrow (A^{-1/2} X A^{-1/2}) (A^{-1/2} X A^{-1/2}) = A^{-1/2} B A^{-1/2}$$

$$\Rightarrow A^{-1/2} X A^{-1/2} = \left(A^{-1/2} B A^{-1/2} \right)^{1/2} \Rightarrow X = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^{1/2} A^{1/2}$$

Moyenne de 2 matrices définies positives



■ Pourquoi la moyenne géométrique plutôt que la moyenne arithmétique :

- La Moyenne arithmétique ne tient pas en compte du fait que les matrices sont des estimés statistiques :

$$\hat{R}_{arithmétique} = \frac{A + B}{2}$$

- Pour un modèle de loi multivariée gaussienne, la variance de la matrice R est proportionnelle à R. Pour tenir compte de cette variance pour la moyenne arithmétique, il faut considérer l'inverse de cette matrice qui fait intervenir le déterminant suivant qui n'a aucune raison d'avoir un bon comportement : $\det[(A + B) / 2]^{-1}$

- En revanche pour la moyenne géométrique :

$$A \circ B = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^{1/2} A^{1/2} \Rightarrow A \circ B = (AB)^{1/2}$$

- Elle tient compte de la variance de la matrice A et B commutent

$$\left. \begin{aligned} ds^2 &= Tr \left(\left(R^{-1} (dR) \right)^2 \right) \\ w &= \Theta(\theta) \end{aligned} \right\} \Rightarrow ds_w^2 = ds_\theta^2$$



- On en déduit l'expression de la géodésique

$$D(A, \gamma(t))^2 = \left\| \log \left(A^{-1/2} \gamma(t) A^{-1/2} \right) \right\|^2$$

$$t \in [0,1] \quad t \rightarrow \gamma(t)$$

$$D(A, \gamma(t)) = t.D(A, B)$$

$$A \circ B = \gamma(1/2)$$

$$\gamma(t) = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^t A^{1/2} \quad \text{avec } 0 \leq t \leq 1$$

$$\gamma(0) = A \quad \text{et} \quad \gamma(1) = B$$

$$\gamma(1/2) = A \circ B$$

Racine Carrée d'une matrices définies positives

- Méthodes récursives du calcul de la racine carrée
 - Itération de Denman-Beabers :

$$X_{k+1} = \begin{bmatrix} 0 & Y_{k+1} \\ Z_{k+1} & 0 \end{bmatrix} = \frac{1}{2} \left(\begin{bmatrix} 0 & Y_k \\ Z_k & 0 \end{bmatrix} + \begin{bmatrix} 0 & Z_k^{-1} \\ Y_k^{-1} & 0 \end{bmatrix} \right)$$

avec $X_0 = \begin{bmatrix} 0 & A \\ I & 0 \end{bmatrix}$ alors $\lim_{k \rightarrow \infty} X_k = \begin{bmatrix} 0 & A^{1/2} \\ A^{-1/2} & 0 \end{bmatrix}$

- Itération de Schultz (sans inversion de matrice) :

$$X_{k+1} = \begin{bmatrix} 0 & Y_{k+1} \\ Z_{k+1} & 0 \end{bmatrix} = \frac{1}{2} X_k (3I - X_k^2)$$

- Méthode de Björk via la décomposition de Schur :

$$Q^+ A Q = T \quad \text{Avec } T \text{ triangulaire supérieure}$$

Par récurrence, on calcule : $U^2 = T \Rightarrow A^{1/2} = Q U Q^+$



■ Moyenne de N matrices symétriques définies positives :

- Comment étendre les approches précédentes pour trouver la définition d'une moyenne sur N matrices symétriques définies positives. Une approche naïve consisterait à faire :

1) $(A_1 A_2 \dots A_N)^{1/N}$

ne sont pas symétriques définies positives
si les A_i ne commutent pas

2) $e^{(\log A_1 + \log A_2)/2} \neq A_1^{1/2} (A_1^{-1/2} A_2 A_1^{-1/2})^{1/2} A_1^{1/2}$

et $A \geq B \Rightarrow e^A \geq e^B$ non valide

L'extension de la définition de la moyenne de N matrices Symétriques définies positives est un problème ouvert

Moyenne Géométrique de matrices définies positives

■ Moyenne de 3 matrices symétriques définies positives :

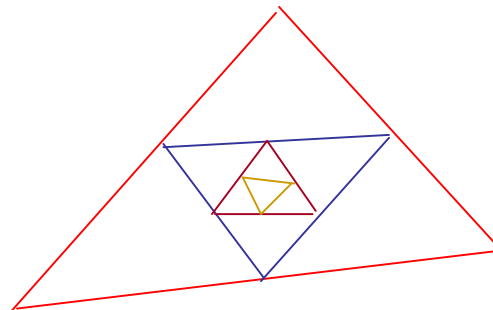
- Une première idée pour 3 matrices proposé par Denes Petz en utilisant une récursion interprétée comme une procédure de symétrisation :

$$A \circ B = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^{1/2} A^{1/2}$$

$$\left\{ \begin{array}{l} A_1 = A , B_1 = B , C_1 = C \end{array} \right.$$

$$\left\{ \begin{array}{l} A_{n+1} = A_n \circ B_n , B_{n+1} = A_n \circ C_n , C_{n+1} = B_n \circ C_n \end{array} \right.$$

$$G(A, B, C) = \lim_{n \rightarrow \infty} A_n = \lim_{n \rightarrow \infty} B_n = \lim_{n \rightarrow \infty} C_n$$





■ Moyenne de N matrices symétriques définies positives :

- Une extension de la moyenne géométrique à N matrices définies positives a été proposée récemment par T. Ando sur la base des propriétés suivantes :

P1: Si A, B et C commutent alors $G(A, B, C) = (ABC)^{1/3}$
et $G(A, A, A) = A$

P2: $G(\alpha A, \beta B, \gamma C) = (\alpha\beta\gamma)^{1/3} G(A, B, C)$ avec $\alpha, \beta, \gamma > 0$

P3: $\forall \pi(A, B, C)$ permutation de (A, B, C)
alors $G(A, B, C) = G(\pi(A, B, C))$

P4: Si $A \geq A_0, B \geq B_0, C \geq C_0 \Rightarrow G(A, B, C) \geq G(A_0, B_0, C_0)$

P5: $\{A_n\}, \{B_n\}, \{C_n\}$ séquences décroissantes monotones
 $\Rightarrow \{G(A_n, B_n, C_n)\} \xrightarrow{n \rightarrow \infty} G(A, B, C)$

P6: $\forall S$ inversible, $G(S^* A S, S^* B S, S^* C S) = S^* G(A, B, C) S$

P7: $G(\lambda A_1 + (1 - \lambda) A_2, \lambda B_1 + (1 - \lambda) B_2, \lambda C_1 + (1 - \lambda) C_2) \geq$
 $\lambda G(A_1, B_1, C_1) + (1 - \lambda) G(A_2, B_2, C_2)$ avec $0 < \lambda < 1$

P8: $G(A, B, C) = G(A^{-1}, B^{-1}, C^{-1})^{-1}$

P9: $\det G(A, B, C) = (\det A \cdot \det B \cdot \det C)^{1/3}$

Moyenne Géométrique de matrices définies positives

■ Moyenne de N matrices symétriques définies positives :

- T. Ando, C.K. Li et R. Mathias ont proposé la procédure suivantes pour N matrices :

$$G(A_1, A_2) = A_1^{1/2} \left(A_1^{-1/2} A_2 A_1^{-1/2} \right)^{1/2} A_1^{1/2}$$

Pour $k = 3, \dots, N$ faire

Connaissant la moyenne pour k matrices : $G(A_1, \dots, A_k)$

$\left\{ G((A_i)_{i \neq l}) \right\}_{l=1}^{k+1}$ connu

$$\left[A^{(1)} = (A_1, \dots, A_{k+1}) \right.$$

$$\left. A^{(r+1)} = T(A^{(r)}) = \left(G((A_i^{(r)})_{i \neq 1}), \dots, G((A_i^{(r)})_{i \neq k+1}) \right), r = 1, 2, \dots \right.$$

$$G(A_1, \dots, A_k, A_{k+1}) = \tilde{A} \quad \text{avec} \quad \lim_{r \rightarrow \infty} A^{(r)} = (\tilde{A}, \dots, \tilde{A})$$

THALES



Du Barycentre de Fréchet au Barycentre de Karcher

Algorithmes de la moyenne/Médiane : critère à optimiser

- Problème à résoudre « Barycentre de N matrices » $\{B_k\}_{k=1}^N$:

$$X = \arg \min_X f_\alpha(X) = \arg \min_X \sum_{k=1}^N d^\alpha(X, B_k)$$

- Moyenne barycentrique

$$\alpha = 2 \quad \longrightarrow \quad X = \arg \min_X f_2(X) = \arg \min_X \sum_{k=1}^N d^2(X, B_k)$$

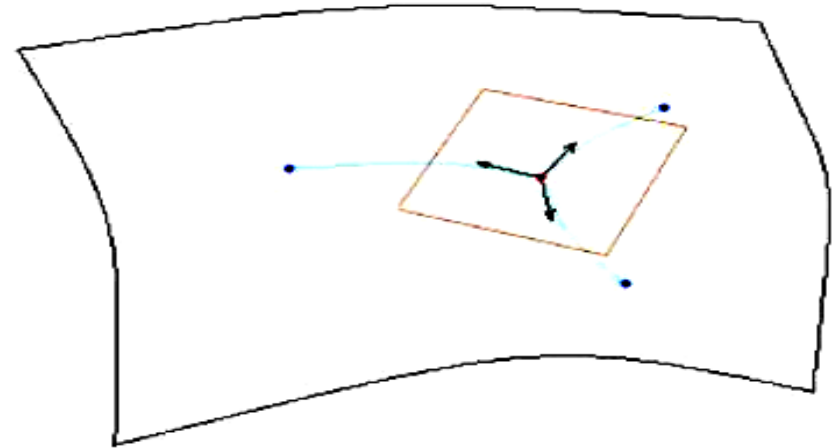
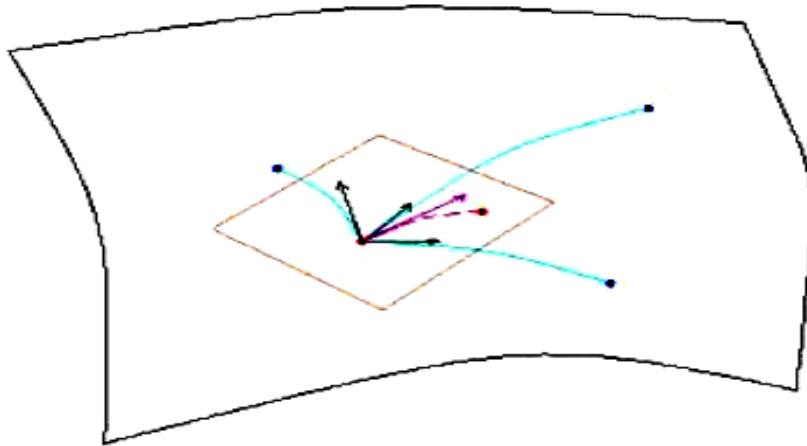
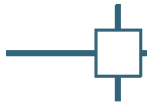
$$\text{grad}[f_2(X)] = -\sum_{k=1}^N \exp_X^{-1}(B_k) \in T_X V \quad \longrightarrow \quad X_{n+1} = \exp_{X_n}(-\varepsilon \cdot \text{grad}[f_2(X_n)])$$

- Mediane (Point de Fermat-Weber)

$$\alpha = 1 \quad \longrightarrow \quad X = \arg \min_X f_1(X) = \arg \min_A \sum_{k=1}^N d(X, B_k)$$

$$\text{grad}[f_1(X)] = -\sum_{k=1}^N \frac{\exp_X^{-1}(B_k)}{d(X, B_k)} \in T_X V \quad \longrightarrow \quad X_{n+1} = \exp_{X_n}(-\varepsilon \cdot \text{grad}[f_1(X_n)])$$

Flot de gradient sur une variété



THALES



Moyennage de matrices



$$\alpha = 2 \implies X = \arg \min_X f_2(X) = \arg \min_X \sum_{k=1}^N d^2(X, B_k)$$

$$\text{avec } d^2(X, B_k) = \langle \log(X^{-1/2} B_k X^{-1/2}), \log(X^{-1/2} B_k X^{-1/2}) \rangle$$

$$X_{n+1} = \exp_{X_n}(-\varepsilon \cdot \text{grad}[f_2(X_n)]) \text{ avec } \text{grad}[f_2(X)] = -\sum_{k=1}^N \exp_X^{-1}(B_k) \in T_X V$$

$$\exp_X^{-1}(U) = X^{1/2} \log(X^{-1/2} U X^{-1/2}) X^{1/2}$$

$$\exp_X(V) = X^{1/2} e^{X^{-1/2} V X^{-1/2}} X^{1/2}$$

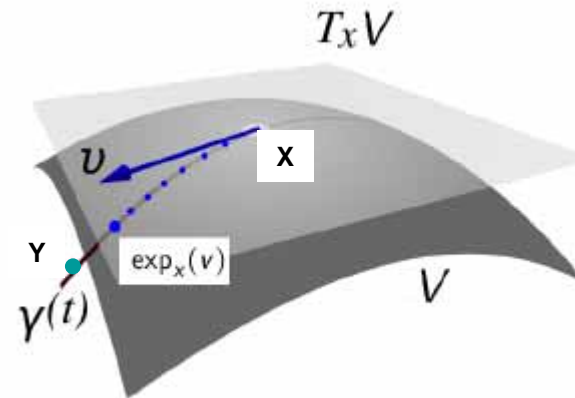


$$X_{n+1} = X_n^{1/2} e^{\varepsilon \sum_{k=1}^N \log(X_n^{-1/2} B_k X_n^{-1/2})} X_n^{1/2}$$

Géodésique et moyenne géométrique symétrisée

$$\begin{cases} v_X^Y = \text{grad}_X^Y(V) = -\exp_X^{-1}(V) = -X^{1/2} \log(X^{-1/2} Y X^{-1/2}) X^{1/2} \\ \exp_X(v_X^Y, t) = X^{1/2} e^{-t(X^{-1/2} \cdot v_X^Y \cdot X^{-1/2})} X^{1/2} \end{cases}$$

$$\Rightarrow \exp_X(v_X^Y, t) = X^{1/2} e^{t \log(X^{-1/2} Y X^{-1/2})} X^{1/2}$$



- La géodésique entre 2 matrices X et Y est donnée par :

$$\gamma(t) = X^{1/2} e^{t \log(X^{-1/2} Y X^{-1/2})} X^{1/2} = X^{1/2} (X^{-1/2} Y X^{-1/2})^t X^{1/2}$$

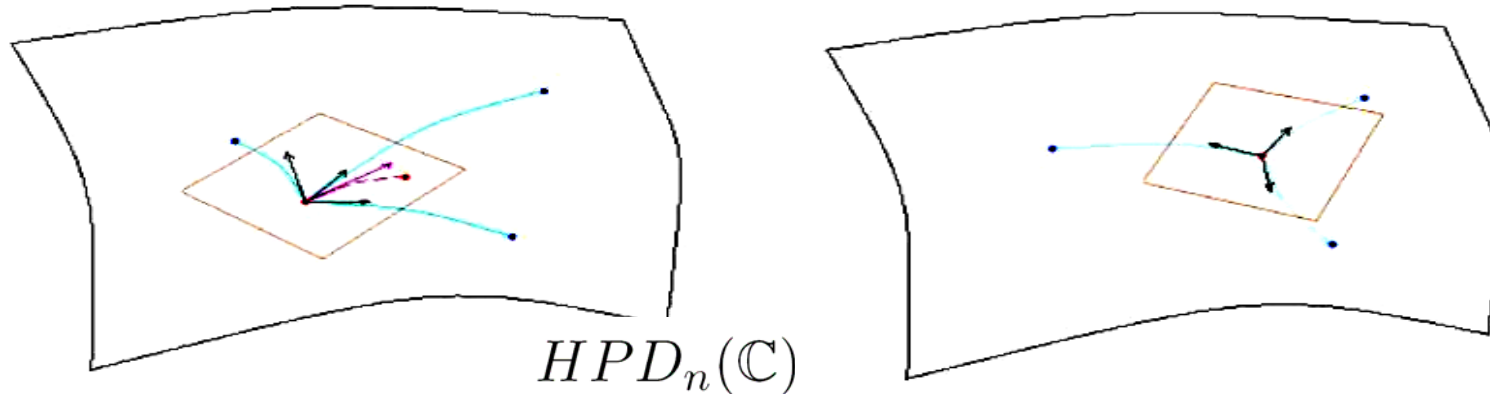
$$\text{avec } t \in [0,1] \quad \gamma(0) = X \quad \text{et} \quad \gamma(1) = Y$$

- La moyenne est définie comme barycentre au sens de Fréchet :

$$X \circ Y = \gamma(1/2) = X^{1/2} (X^{-1/2} Y X^{-1/2})^{1/2} X^{1/2}$$

Il s'agit d'un flot de gradient sur la variété

- Algorithme par descente de gradient



- La tangente en X à la géodésique rejoignant X à B_k

$$\gamma_k(t) = X^{1/2} \left(X^{-1/2} B_k X^{-1/2} \right)^{t_k} X^{1/2} = X^{1/2} e^{t_k \log(X^{-1/2} B_k X^{-1/2})} X^{1/2}$$

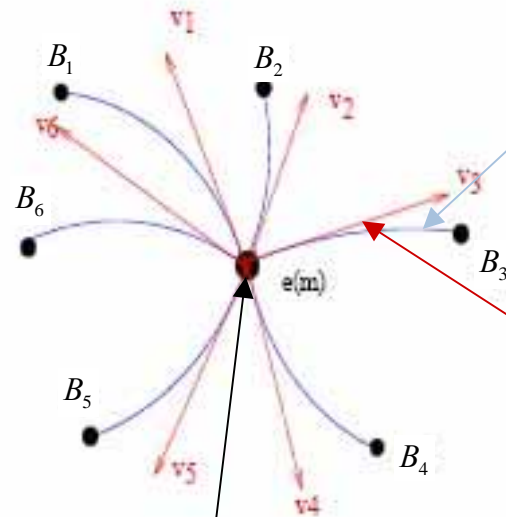
$$\Rightarrow \left. \frac{d\gamma_k(t)}{dt_k} \right|_{t_k=0} = X^{1/2} \log(X^{-1/2} B_k X^{-1/2}) X^{1/2}$$

- La somme des tangentes en X des géodésiques de X aux $\{B_k\}_{k=1}^N$

$$G_X = \sum_{k=1}^N \left. \frac{d\gamma_k(t)}{dt_k} \right|_{t=0} \Rightarrow G_X = X^{1/2} \left(\sum_{k=1}^N \log(X^{-1/2} B_k X^{-1/2}) \right) X^{1/2}$$

Karcher Barycenter & Jacobi Field

$$\gamma_k(t) = X^{1/2} \left(X^{-1/2} B_k X^{-1/2} \right)^t X^{1/2} = X^{1/2} e^{t \log \left(X^{-1/2} B_k X^{-1/2} \right)} X^{1/2}$$



$$\left. \frac{d\gamma_k(t)}{dt} \right|_{t=0} = X^{1/2} \log \left(X^{-1/2} B_k X^{-1/2} \right) X^{1/2}$$

$$\sum_{k=1}^N \left. \frac{d\gamma_k(t)}{dt} \right|_{t=0} = X^{1/2} \left(\sum_{k=1}^N \log \left(X^{-1/2} B_k X^{-1/2} \right) \right) X^{1/2} = 0$$

- Soit γ une géodésique et on introduit une famille de geodesiques $c(s,t) = \exp_y(s \exp_y^{-1} \gamma(t))$, où y est fixé.
- soit $c' = \frac{d}{ds} c(s,t)$ et $\dot{c} = \frac{d}{dt} c(s,t)$. Alors $s \mapsto \dot{c}(s,t)$ est une famille de champ de Jacobi.
- De plus, $d(y, \gamma(t)) = \|c'(s,t)\|$.
- Ceci permet de calculer

$$\frac{d}{dt} \langle c'(s,t), c'(s,t) \rangle$$

- $\frac{D}{dt}$ et $\frac{D}{ds}$ correspondent aux dérivées cocvariantes

- Par bilinéarité, on trouve

$$\frac{d}{dt} \langle c'(s, t), c'(s, t) \rangle = 2 \left\langle \frac{D}{dt} c'(s, t), c'(s, t) \right\rangle = 2 \int_0^1 \left\langle \frac{D}{ds} \dot{c}(s, t), c'(s, t) \right\rangle ds$$

En utilisant le fait que $\frac{D}{dt} c' = \frac{D}{ds} \dot{c}$, et que $\frac{D}{ds} c' = 0$ (équation géodésique),

du au fait que $\|c'\|$ ne dépend pas de s .

- Alors,

$$2 \int_0^1 \left\langle \frac{D}{ds} \dot{c}(s, t), c'(s, t) \right\rangle ds = 2 \int_0^1 \frac{d}{ds} \langle \dot{c}(s, t), c'(s, t) \rangle ds = \langle \dot{c}(1, t), c'(1, t) \rangle$$

comme $c'(0, t) = 0$

- Par définition de c , $\dot{c}(1, t) = \dot{\gamma}(t)$ et $c'(1, t) = -\exp_{\gamma(t)}^{-1}(y)$
- On conclue que

$$\frac{d}{dt} \langle c'(s, t), c'(s, t) \rangle = -2 \langle \dot{\gamma}(t), \exp_{\gamma(t)}^{-1}(y) \rangle$$

- Ceci est valide pour tout chemin géodésique γ , aussi avec $t = 0$ et $\gamma(0) = x$, on obtient le résultat final

$$\nabla_x \text{dist}(x, y)^2 = -2 \exp_x^{-1}(y)$$

THALES



Notion d'espace symétrique

Il s'agit d'un espace Riemannien symétrique



- En géométrie riemannienne, un espace localement symétrique est une variété riemannienne (M,g) telle que, (localement) autour de chaque point x , il existe une symétrie σ_x qui inverse les géodésiques issues de x et qui est une isométrie (locale).
- Un espace localement symétrique est dit symétrique si les symétries peuvent se prolonger à tout l'espace. De manière équivalente, une variété riemannienne (M,g) est dite symétrique lorsque, pour tout point x de M , il existe une isométrie $\sigma_x: M \rightarrow M$ vérifiant :
 - $\sigma_x(x) = x$;
 - $d\sigma_x(x) = -\text{Id}$.
- Cette isométrie σ_x est appelée l'involution en x .
- Propriétés :
 - Tout espace symétrique est une variété riemannienne géodésiquement complète, donc complète en vertu du théorème de Hopf-Rinow.
 - Il existe une et une unique involution en x .
 - Le tenseur de courbure d'une variété riemannienne est parallèle.

Il s'agit d'un espace Riemannien symétrique

- Pour l'espace considéré, la géodésique entre deux matrices A et B est donnée par :

$$\gamma(t) = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^t A^{1/2} = A^{1/2} e^{t \cdot \log(A^{-1/2} B A^{-1/2})} A^{1/2} \text{ avec } t \in [0,1]$$

- Pour l'espace symétrique des matrices symétriques ou hermitiennes définies positives, pour chaque paire (A,B), il existe une isométrie bijective $G_{(A,B)}$ qui vérifie :

$$G_{(A,B)} A = B \quad \text{et} \quad G_{(A,B)} B = A$$

- Cette isométrie a un unique point fixe Z qui est donné par le milieu de (A,B), c'est à dire la moyenne géométrique précédemment définie $d(G_{(A,B)} X, X) = 2d(X, Z)$:

$$G_{(A,B)} X = (A \circ B) X^{-1} (A \circ B) \text{ avec } A \circ B = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^{1/2} A^{1/2}$$

Et les propriétés : $G_{(A,B)}(A \circ B) = (A \circ B)$ et $dG_{(A,B)}(A \circ B) = -I$

- Ceci est différent de l'approche classique en traitement du signal qui suppose un espace normé euclidien :

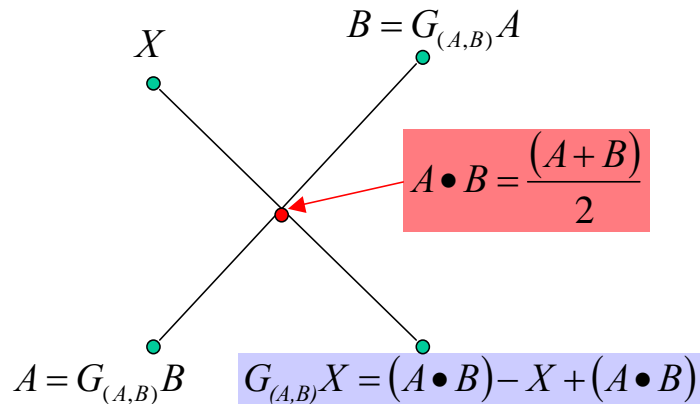
$$G_{(A,B)} X = (A \bullet B) - X + (A \bullet B)$$

avec $A \bullet B = \frac{A + B}{2}$ et $\|A - B\|_F$

Géodésique et moyenne géométrique symétrisée

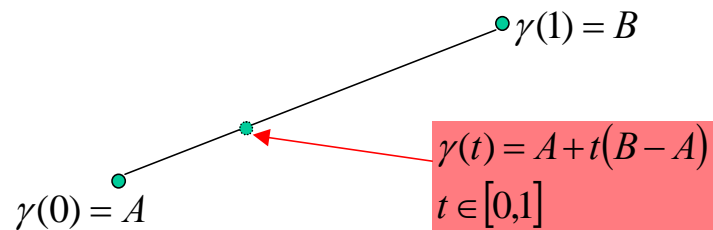


Espace Normé : isométrie
(approche traitement du signal Classique)

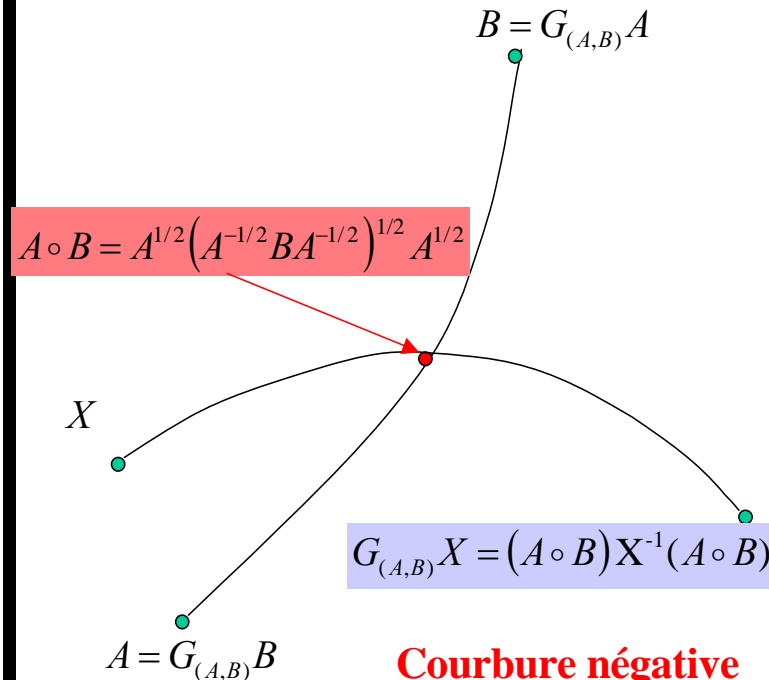


Courbure nulle

Espace Normé : Géodésique

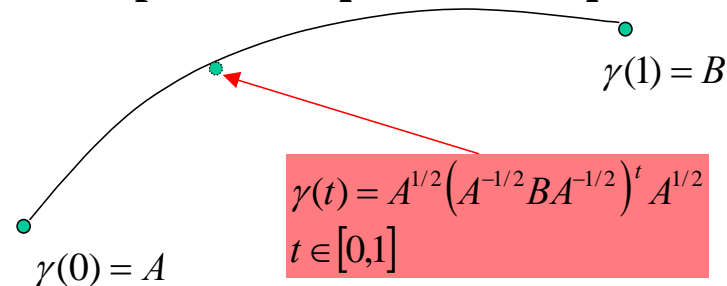


Espace métrique : isométrie



Courbure négative

Espace métrique : Géodésique



« La Messe est dite » par Marcel Berger

Marcel Berger (IHES), « 150 ans de Géométrie Riemannienne »,
Géométrie au 20^{ième} siècle, Histoire et horizons, Hermann Éditeur, 2005



M. Berger

• Premier Miracle :

La théorie des espaces symétriques peut être considérée comme le premier miracle de la géométrie riemannienne, en fait comme un nœud de forte densité dans l'arbre de toutes les mathématiques. ... On doit à Elie Cartan dans les années 1926 d'avoir découvert que ces géométries sont, dans une dimension donnée, en nombre fini, et en outre toutes classées.

• Second Miracle :

Entre les variétés localement symétriques et les variétés riemanniennes générales, il existe une catégorie intermédiaire, celle des variétés kählériennes. ... On a alors affaire pour décrire le panorama des métriques kählériennes sur notre variété, non pas à un espace de formes différentielles quadratiques, très lourd, mais à un espace vectoriel de fonctions numériques (le potentiel de Kähler). ... La richesse Kählerienne fait dire à certains que la géométrie kählérienne est plus importante que la géométrie riemannienne.

• Pas d'espoir d'autre miracle :

Ne cherchez pas d'autres miracles du genre des espaces (localement) symétriques et des variétés kählériennes. En effet, c'est un fait depuis 1953 que les seules variétés riemanniennes irréductibles qui admettent un invariant par transport parallèle autre que g elle-même (et sa forme volume) sont les espaces localement symétriques, les variétés kählériennes, les variétés kählérienne de Calabi-Yau, et les variétés hyperkähleriennes.

THALES



Médianage de matrices

$$\alpha = 1 \quad \longrightarrow \quad X = \arg \min_X f_1(X) = \arg \min_X \sum_{k=1}^N d(X, B_k)$$

$$\text{grad}[f_1(X)] = - \sum_{k=1}^N \frac{\exp_X^{-1}(B_k)}{d(X, B_k)} \in T_X V \quad \longrightarrow \quad X_{n+1} = \exp_{X_n}(-\varepsilon \cdot \text{grad}[f_1(X_n)])$$

$$\exp_X^{-1}(U) = X^{1/2} \log(X^{-1/2} U X^{-1/2}) X^{1/2}$$

$$\exp_X(V) = X^{1/2} e^{X^{-1/2} V X^{-1/2}} X^{1/2} \quad \longrightarrow$$

$$X_{n+1} = X_n^{1/2} e^{\varepsilon \sum_{k=1}^N \frac{\log(X_n^{-1/2} B_k X_n^{-1/2})}{\|\log(X_n^{-1/2} B_k X_n^{-1/2})\|_F}} X_n^{1/2}$$

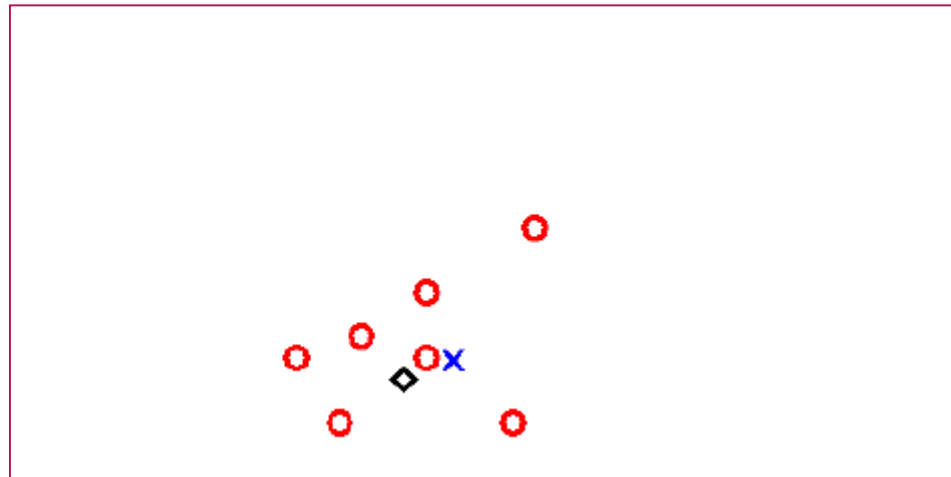
avec $\|\log(X_n^{-1/2} B_k X_n^{-1/2})\|_F = \sqrt{\sum_{i=1}^M \log^2(\lambda_i)}$

et $\det(X_n - \lambda B_k) = 0$

Médiane plus robuste que moyenne

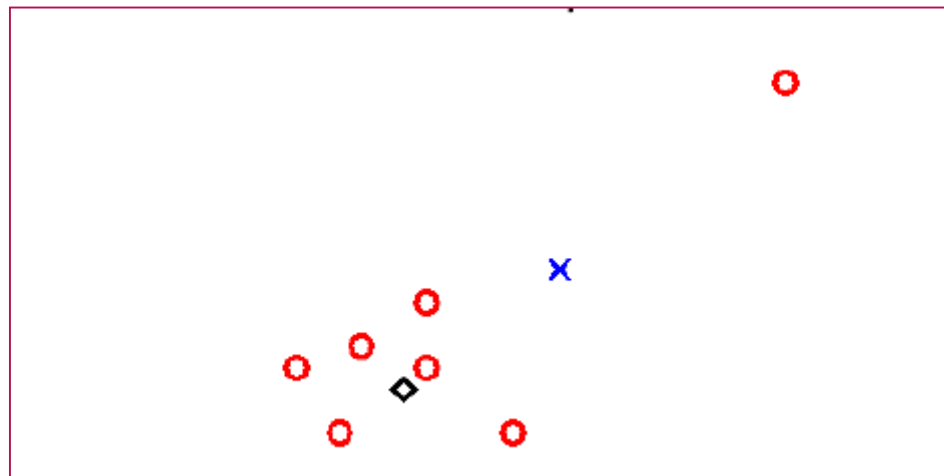
X : moyenne

◇ : médiane



La médiane est insensible à la présence de points aberrants tant qu'ils restent minoritaires

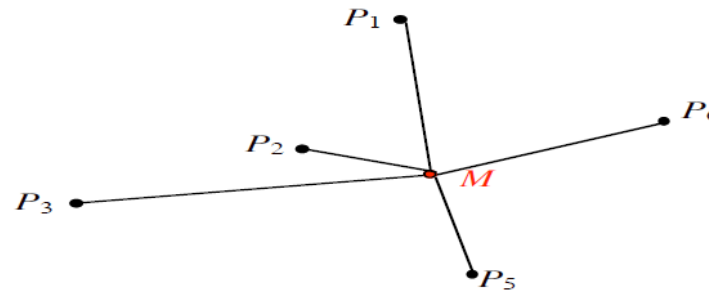
La moyenne est immédiatement perturbée par la présence de points aberrants



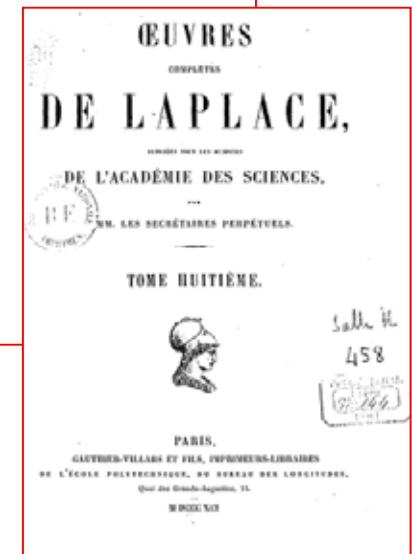
Mediane et point de Fermat-Weber



Let $P_1, P_2 \dots P_n$ be n points of a metric space. A point M is a *median* of these n points if M is a point of the space minimizing the sum $d(M, P_1) + d(M, P_2) + \dots + d(M, P_n)$ of its distances to the n points.



The median of six points



$$\int_0^{\text{median}} P_X(x).dx = 0.5 \Leftrightarrow \text{median} = \underset{m}{\text{Min}} E(|x - m|)$$

Laplace (1774) :

Laplace a appelée cette valeur « **le milieu de probabilité** » ou « **la valeur probable** ». Le terme mediane a été introduit par Cournot dans *l'Exposition de la théorie des chances* en 1883.

Algorithme de gradient non convergent

- Algorithme médian par descente de gradient

$$X_{n+1} = X_n^{1/2} e^{\varepsilon \sum_{k=1}^N \frac{\log(X_n^{-1/2} B_k X_n^{-1/2})}{\left\| \log(X_n^{-1/2} B_k X_n^{-1/2}) \right\|_F}} X_n^{1/2}$$

avec $\left\| \log(X_n^{-1/2} B_k X_n^{-1/2}) \right\|_F = \sqrt{\sum_{i=1}^M \log^2(\lambda_i)}$

et $\det(X_n - \lambda B_k) = 0$

- Non différentiabilité aux points de données : problème de convergence
 - En théorie : régularisation au voisinage des points de données et extraction diagonale (algorithme général adapté de Huling Le, mais long d'exécution)
 - En pratique :
 - Assez souvent, les points donnés « entourent » la médiane : convergence
 - Prendre un pas voisin de la précision voulue (50 itérations suffisent ici)

Mediane : autre algorithme de gradient régularisé

- Autre Algorithme Régularisé proposé

- Critère de minimisation régularisé

$$X_{m+1} = \arg \min_X \sum_{k=1}^N \frac{d^2(X, B_k)}{d(X_m, B_k)}$$

- Résolution itérative de problèmes d'optimisation convexes régularisés, dont la suite des solutions converge vers la médiane.

- Étape m :

$$\begin{cases} X_{m+1,0} = X_m = X_{m,n_{convergence}} \\ X_{m+1,n+1} = X_{m+1,n}^{1/2} e^{\varepsilon \sum_{k=1}^N \frac{\log(X_{m+1,n}^{-1/2} B_k X_{m+1,n}^{-1/2})}{\|\log(X_m^{-1/2} B_k X_m^{-1/2})\|_F}} X_{m+1,n}^{1/2} \end{cases}$$

$$\text{arrêt : } \left\| \log \left(X_{m+1,n_{convergence}+1}^{-1/2} X_{m,n_{convergence}} X_{m+1,n_{convergence}+1}^{-1/2} \right) \right\|_F < \zeta$$

$$\text{alors } X_{m+1} = X_{m+1,n_{convergence}}$$

- Critère d'arrêt :

$$\left\| \log \left(X_{m+1}^{-1/2} X_m X_{m+1}^{-1/2} \right) \right\|_F < \zeta$$

THALES



Equation de diffusion sur un
graphe de matrices

Equation de la chaleur de Joseph Fourier



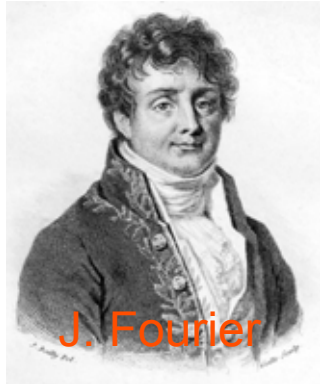
- « Les équations différentielles de la propagation de la chaleur expriment les conditions les plus générales, et ramènent les questions physiques à des problèmes d'analyse pure, ce qui est proprement l'objet de la théorie ... Les formes des corps sont variées à l'infini, la distribution de la chaleur qui les pénètre peut être arbitraire et confuse; mais toutes les inégalités s'effacent rapidement et disparaissent à mesure que le temps s'écoule. La marche du phénomène devenue plus régulière et plus simple, demeure enfin assujettie à une loi déterminée qui est la même pour tous les cas, et qui ne porte plus aucune empreinte sensible de la disposition initialeLes théories nouvelles, expliquées dans notre ouvrage sont réunies pour toujours aux sciences mathématiques et reposent comme elles sur des fondements invariables; elles conserveront tous les éléments qu'elles possèdent aujourd'hui, et elles acquerront, continuellement plus d'étendue. »

Joseph Fourier (1768-1830), « Discours préliminaire à la théorie analytique de la chaleur »

Équation de diffusion & géodésique



- Équation de diffusion sur un graphe 1D de données scalaires



- Dans un espace vectoriel normé dans le cas unidimensionnel, l'équation de diffusion s'écrit via le Laplacien discret:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \Rightarrow \frac{\partial u}{\partial t} = \frac{1}{\nabla x} \left(\frac{u_{n+1} - u_n}{\nabla x} - \frac{u_n - u_{n-1}}{\nabla x} \right) = \frac{2}{\nabla x^2} (\hat{u}_n - u_n)$$

avec la moyenne arithmétique : $\hat{u}_n = (u_{n+1} + u_{n-1}) / 2$

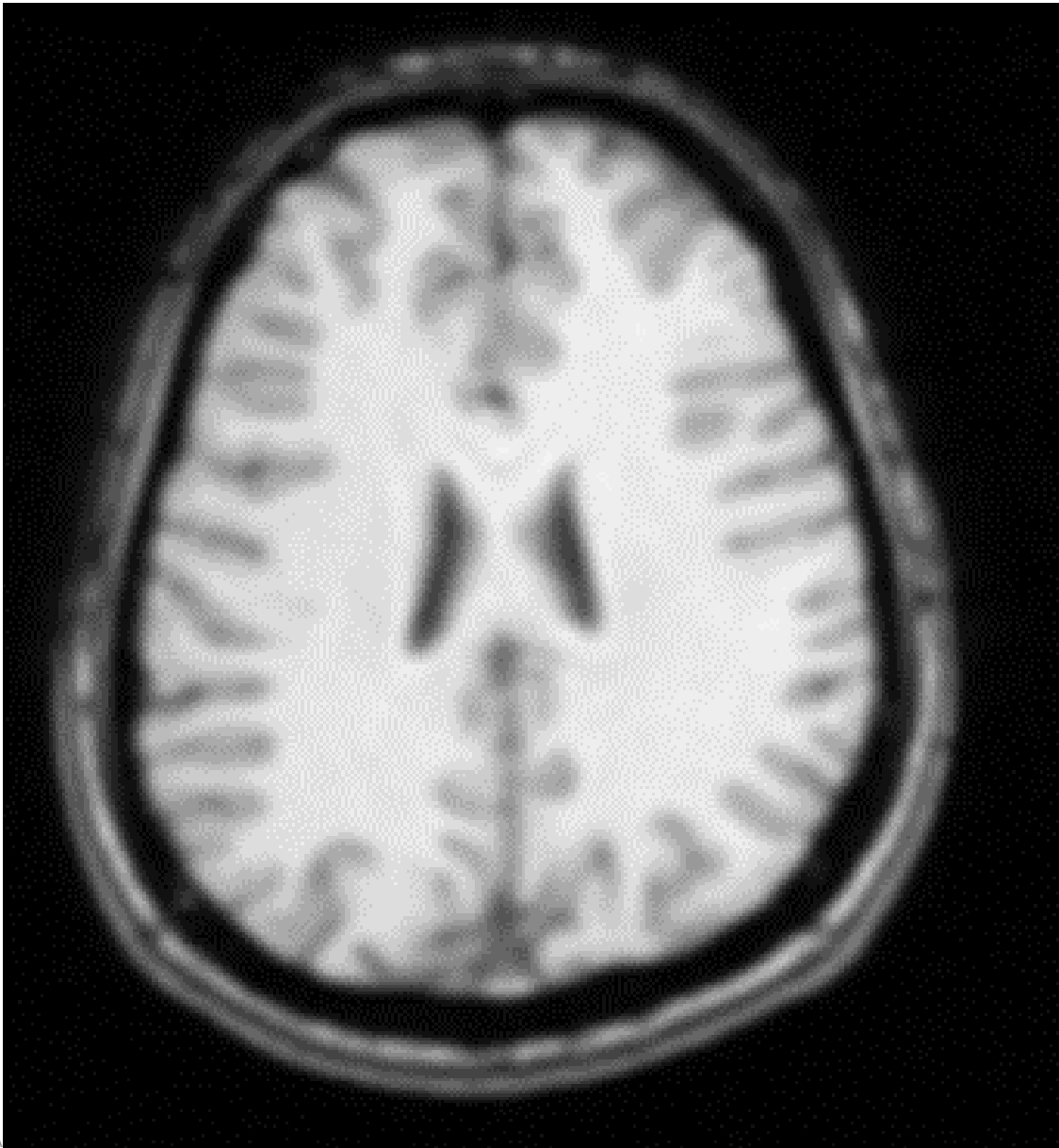
- L'équation de Fourier discrétisée peut également s'écrire :

$$u_{n,t+1} = u_{n,t} + \frac{2\nabla t}{\nabla x^2} [\hat{u}_{n,t} - u_{n,t}] = (1 - \rho) u_{n,t} + \rho \hat{u}_{n,t} \text{ avec } \rho = \frac{2\nabla t}{\nabla x^2}$$

- Par analogie sur la base du flot de Karcher, il est possible d'écrire une équation de diffusion sur un graphe 1D de matrices, via la moyenne géométrique $\hat{A}_{n,t}$ des voisins $(A_{n-1,t}, A_{n+1,t})$:

$$X_{n,t+1} = X_{n,t}^{1/2} e^{\frac{2\nabla t}{\nabla x^2} \log(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2})} X_{n,t}^{1/2} = X_{n,t}^{1/2} \left(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2} \right)^\rho X_{n,t}^{1/2}$$

$$\text{avec } \rho = \frac{2\nabla t}{\nabla x^2} \text{ et } \hat{X}_{n,t} = X_{n+1,t}^{1/2} \left(X_{n+1,t}^{-1/2} X_{n-1,t} X_{n+1,t}^{-1/2} \right)^{1/2} X_{n+1,t}^{1/2}$$



$$\frac{\partial u}{\partial t} = \Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$$

Équation de diffusion & géodésique

- Équation de diffusion sur un graphe 1D de matrices



$$X_{n,t+1} = X_{n,t}^{1/2} e^{\frac{2\nabla t}{\nabla x^2} \log(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2})} X_{n,t}^{1/2} = X_{n,t}^{1/2} \left(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2} \right)^\rho X_{n,t}^{1/2}$$

avec $\rho = \frac{2\nabla t}{\nabla x^2}$ et $\hat{X}_{n,t} = X_{n+1,t}^{1/2} \left(X_{n+1,t}^{-1/2} X_{n-1,t} X_{n+1,t}^{-1/2} \right)^{1/2} X_{n+1,t}^{1/2}$

- Interprétation géométrique : A partir des géodésiques $\gamma_{X_{n,t}}^{X_{n,t+1}}(\lambda)$ et $\gamma_{X_{n,t}}^{\hat{X}_{n,t}}(\lambda)$ entre $X_{n,t}$, $X_{n,t+1}$ et $\hat{X}_{n,t}$:

$$\log(X_{n,t}^{-1/2} X_{n,t+1} X_{n,t}^{-1/2}) = \frac{2\nabla t}{\nabla x^2} \log(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2})$$

$$\Rightarrow \left. \frac{d\gamma_{X_{n,t}}^{X_{n,t+1}}(\lambda)}{d\lambda} \right|_{\lambda=0} = \frac{2\nabla t}{\nabla x^2} \left. \frac{d\gamma_{X_{n,t}}^{\hat{X}_{n,t}}(\lambda)}{d\lambda} \right|_{\lambda=0} \Rightarrow T_{\gamma_{X_{n,t}}^{X_{n,t+1}}} // T_{\gamma_{X_{n,t}}^{\hat{X}_{n,t}}}$$

- Comme dans le cas euclidien $x \bullet_\beta y = \beta \cdot x + (1 - \beta) y$ on remplace la moyenne pondérée par un barycentre pondéré géodésique :

$$A \bullet_\beta B = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^\beta A^{1/2}$$

Équation de diffusion : contrainte de discrétisation

- Une première remarque concernant une contrainte qui apparaît :

$$u_{n,t+1} = u_{n,t} + \frac{2\nabla t}{\nabla x^2} [\hat{u}_{n,t} - u_{n,t}] = (1 - \rho) \cdot u_{n,t} + \rho \cdot \hat{u}_{n,t} \text{ avec } \rho = \frac{2\nabla t}{\nabla x^2}$$

$$u_{n,t+1} \in \begin{cases} [u_{n,t}, \hat{u}_{n,t}] & \text{si } u_{n,t} < \hat{u}_{n,t} \\ [\hat{u}_{n,t}, u_{n,t}] & \text{si } \hat{u}_{n,t} < u_{n,t} \end{cases} \Leftrightarrow \rho < 1 \text{ soit } \frac{2}{\nabla x} < \nu \text{ en posant } \nu = \frac{\nabla x}{\nabla t}$$

Si on vérifie Shannon $\nabla x = T_E = \frac{1}{f_E}$ avec $f_{\max} < \frac{f_E}{2} \Rightarrow 4f_{\max} < 2f_E < \nu$

Équation de diffusion : Du discret au continu

- L'équation de diffusion discrète sur un graphe 1D de matrices symétriques ou hermitiennes définies positives est donnée par :

$$X_{n,t+1} = X_{n,t}^{1/2} e^{\frac{2\nabla t}{\nabla x^2} \log(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2})} X_{n,t}^{1/2} = X_{n,t}^{1/2} \left(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2} \right)^\rho X_{n,t}^{1/2}$$

$$\text{avec } \rho = \frac{2\nabla t}{\nabla x^2} \text{ et } \hat{X}_{n,t} = X_{n+1,t}^{1/2} \left(X_{n+1,t}^{-1/2} X_{n-1,t} X_{n+1,t}^{-1/2} \right)^{1/2} X_{n+1,t}^{1/2}$$

- Pour remonter à l'équation continue de diffusion, il faut utiliser l'équation de Campbell-Hausdorff :

$$\log(e^X e^Y) = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}[X, [X, Y]] + \frac{1}{12}[Y, [Y, X]] \\ + \text{termes de degré } \geq 4$$

$$\text{avec } [X, Y] = XY - YX$$

Équation de diffusion : Du discret au continu

- L'équation de diffusion discrète ré-écrite via l'équation de Campbell-Hausdorff :

$$e^{\log(X_{n,t}^{-1/2} X_{n,t+1} X_{n,t}^{-1/2})} e^{-\rho \log(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2})} = I \Rightarrow \log \left(e^{\log(X_{n,t}^{-1/2} X_{n,t+1} X_{n,t}^{-1/2})} e^{-\rho \log(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2})} \right) = 0$$

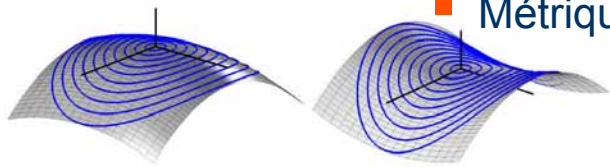
$$\Rightarrow \log(X_{n,t}^{-1/2} X_{n,t+1} X_{n,t}^{-1/2}) - \rho \log(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2}) + \frac{1}{2} \left[\log(X_{n,t}^{-1/2} X_{n,t+1} X_{n,t}^{-1/2}), -\rho \log(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2}) \right] = 0$$

$$\Rightarrow \frac{\log(X_{n,t}^{-1/2} X_{n,t+1} X_{n,t}^{-1/2})}{\Delta t} = \frac{2}{\Delta x^2} \log(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2}) + \frac{1}{\Delta x^2} \left[\log(X_{n,t}^{-1/2} X_{n,t+1} X_{n,t}^{-1/2}), \log(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2}) \right]$$

$$\frac{\partial \log X}{\partial t} = \frac{\partial^2 \log X}{\partial x^2} + \frac{dt}{2} \left[\frac{\partial \log X}{\partial t}, \frac{\partial^2 \log X}{\partial x^2} \right]$$

Équation de diffusion anisotrope

- Regardons maintenant le cas anisotrope classique :
 - Dans le cas général diffusion sur une variété, l'équation de diffusion de Laplace-Beltrami s'écrit :



■ Métrique :

$$ds^2 = \sum_{k=1}^N g_{i,j} dx_i dx_j \quad \text{avec} \quad g = [g_{i,j}]_{i,j=1}^N$$

■ Équation de diffusion :

$$\frac{\partial u}{\partial t} = \frac{1}{\sqrt{\det(g)}} \sum_{i,j} \frac{\partial}{\partial x_i} \left(\sqrt{\det(g)} g^{ij} \frac{\partial u}{\partial x_j} \right) \quad \text{avec} \quad g^{-1} = [g^{i,j}]_{i,j=1}^N$$

- Dans le cas 1D isotrope : $ds^2 = dx^2 \Rightarrow \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$
- Dans le cas 1D anisotrope : $ds^2 = dx^2 + du^2 = \left(1 + \left(\frac{\partial u}{\partial x} \right)^2 \right) dx^2$
soit :

$$g = \left(1 + \left(\frac{\partial u}{\partial x} \right)^2 \right) \Rightarrow \frac{\partial u}{\partial t} = \left(1 + \left(\frac{\partial u}{\partial x} \right)^2 \right)^{-1/2} \frac{\partial}{\partial x} \left(\left(1 + \left(\frac{\partial u}{\partial x} \right)^2 \right)^{-1/2} \frac{\partial u}{\partial x} \right)$$

Équation de diffusion anisotrope : cas scalaire 1D

- Dans le cas 1D anistrophe :

$$\frac{\partial u}{\partial t} = \left(1 + \left(\frac{\partial u}{\partial x} \right)^2 \right)^{-1/2} \frac{\partial}{\partial x} \left(\left(1 + \left(\frac{\partial u}{\partial x} \right)^2 \right)^{-1/2} \frac{\partial u}{\partial x} \right)$$

- Si on discrétise :

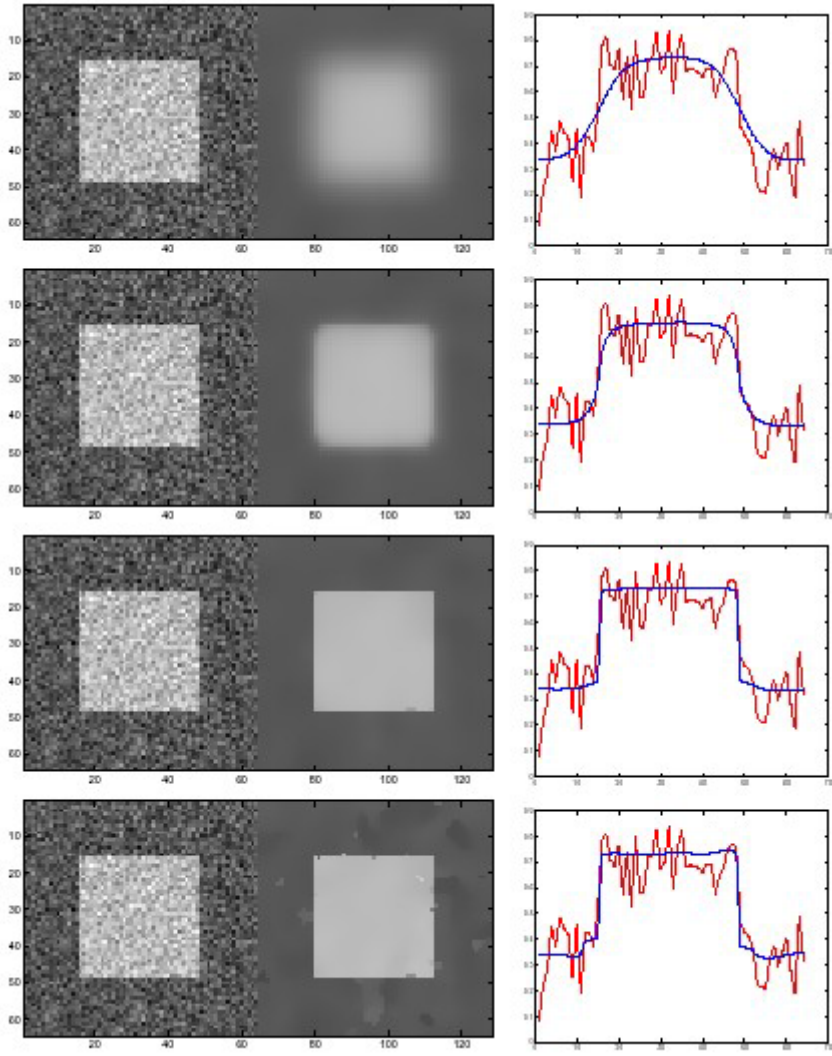
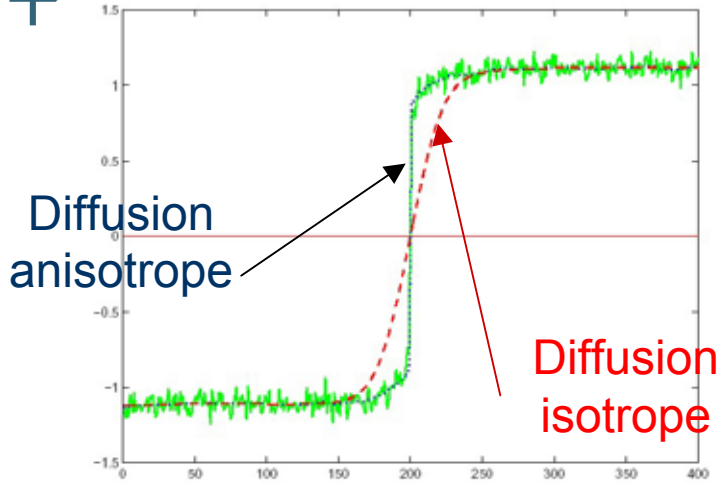
$$u_{n,t+1} = u_{n,t} + \frac{\nabla t \cdot \alpha_{n,t}}{\nabla x} \left[\alpha_{n,t}^+ \left(\frac{u_{n+1,t} - u_{n,t}}{\nabla x} \right) - \alpha_{n,t}^- \left(\frac{u_{n,t} - u_{n-1,t}}{\nabla x} \right) \right]$$

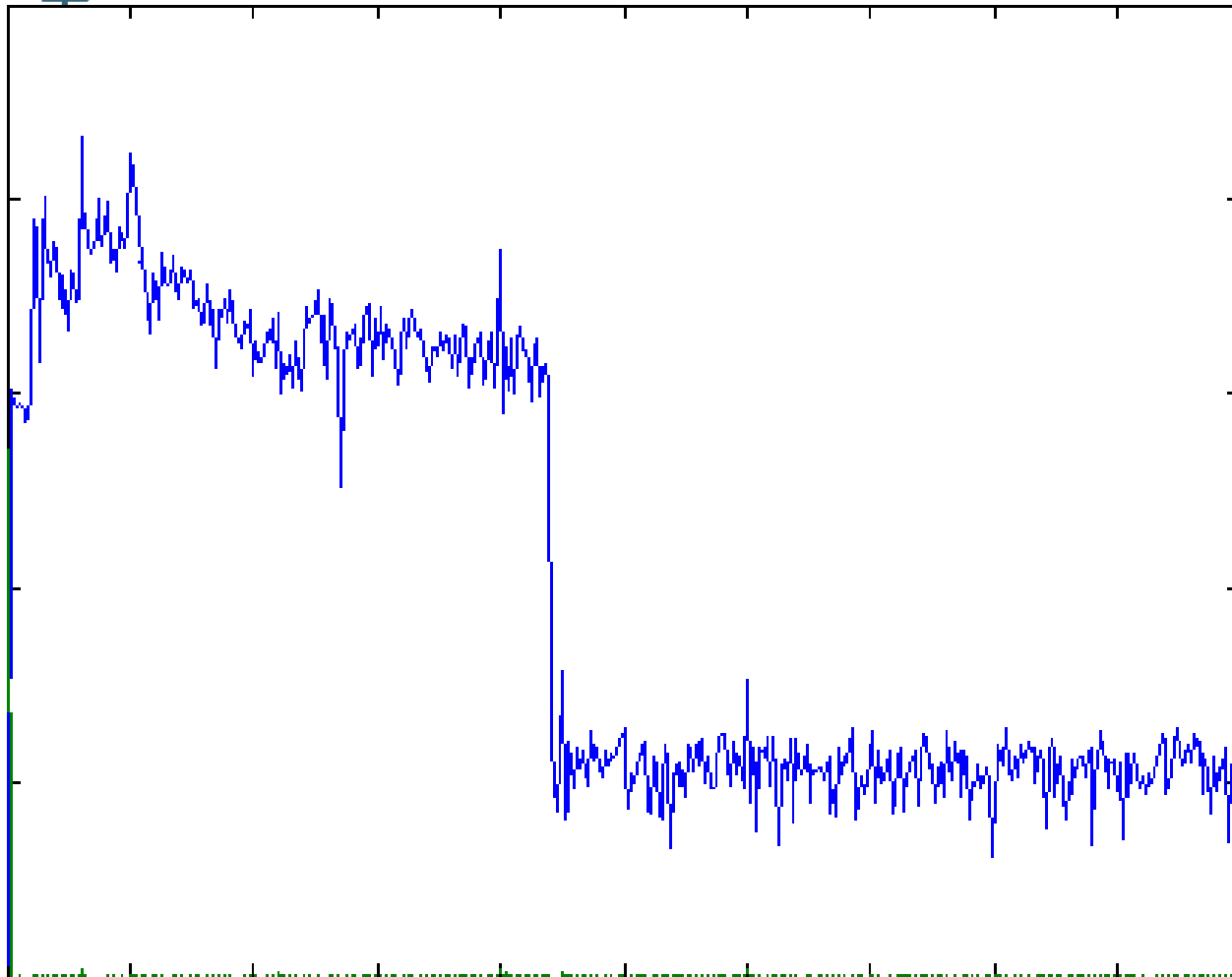
avec

$$\alpha_{n,t} = \left(1 + \left(\frac{u_{n+1,t} - u_{n-1,t}}{\nabla x} \right)^2 \right)$$

$$\alpha_{n,t}^+ = \left(1 + \left(\frac{u_{n+1,t} - u_{n,t}}{\nabla x} \right)^2 \right) \quad \alpha_{n,t}^- = \left(1 + \left(\frac{u_{n,t} - u_{n-1,t}}{\nabla x} \right)^2 \right)$$

Diffusion anisotrope ↩





En bleu, signal soumis à la diffusion anisotrope

En vert, différence entre le signal original auquel est soustrait le signal diffusé anisotropiquement au cours de la diffusion (apparition des cibles)

diffusion anisotrope sur un graphe de matrices

- Dans le cas 1D anisotrope, par analogie sur un graphe de matrices :

$$X_{n,t+1} = X_{n,t}^{1/2} e^{\gamma \log(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2})} X_{n,t}^{1/2} = X_{n,t}^{1/2} \left(X_{n,t}^{-1/2} \hat{X}_{n,t} X_{n,t}^{-1/2} \right)^\gamma X_{n,t}^{1/2}$$

$$\text{avec } \hat{X}_{n,t} = X_{n+1,t}^{1/2} e^{\beta \log(X_{n+1,t}^{-1/2} X_{n-1,t} X_{n+1,t}^{-1/2})} X_{n+1,t}^{1/2} = X_{n+1,t}^{1/2} \left(X_{n+1,t}^{-1/2} X_{n-1,t} X_{n+1,t}^{-1/2} \right)^\beta X_{n+1,t}^{1/2}$$

$$\beta = \frac{\alpha_{n,t}^+}{\left(\alpha_{n,t}^+ + \alpha_{n,t}^- \right)}, \quad \gamma = \frac{\nabla t \cdot \alpha_{n,t} \left(\alpha_{n,t}^+ + \alpha_{n,t}^- \right)}{\nabla x^2}, \quad \alpha_{n,t} = \left(1 + \left(\frac{d(X_{n+1,t}, X_{n-1,t})}{\nabla x} \right)^2 \right)^{-1/2}$$

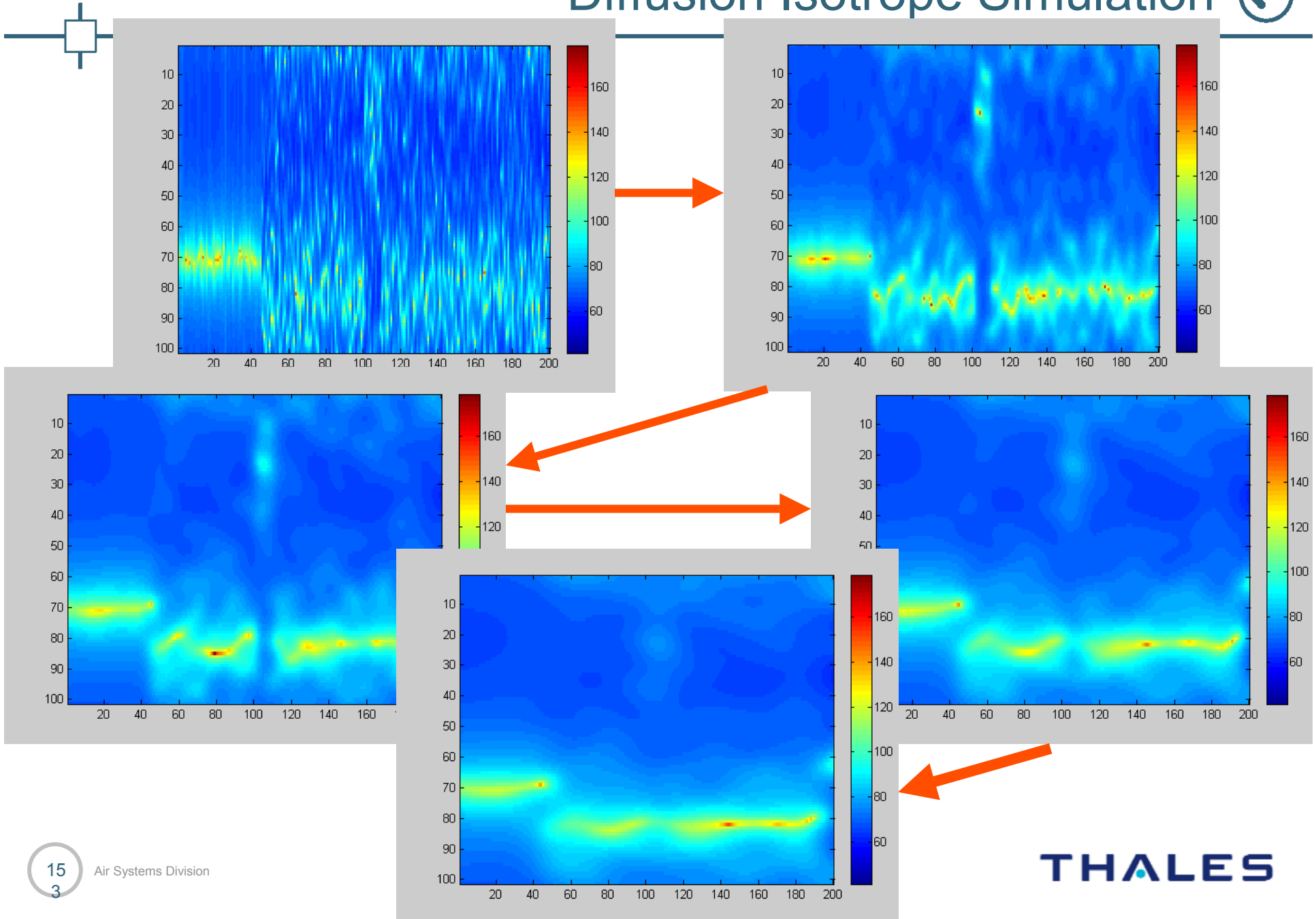
$$\alpha_{n,t}^+ = \left(1 + \left(\frac{d(X_{n+1,t}, X_{n,t})}{\nabla x} \right)^2 \right)^{-1/2} \quad \text{et} \quad \alpha_{n,t}^- = \left(1 + \left(\frac{d(X_{n,t}, X_{n-1,t})}{\nabla x} \right)^2 \right)^{-1/2}$$

- Pour la distance , on prend :

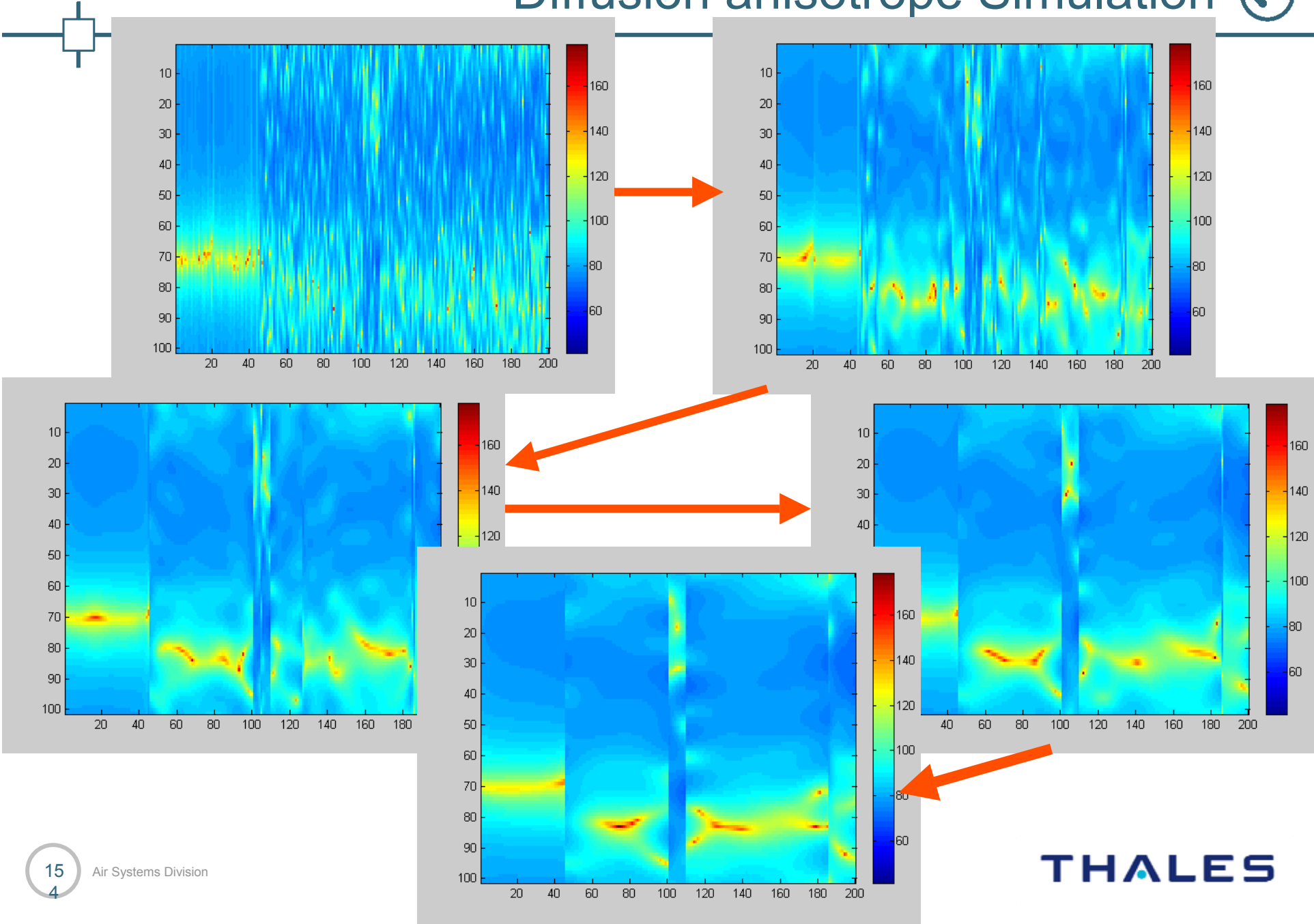
$$d(X_1, X_2) = \left\| \log(X_1^{-1/2} X_2 X_1^{-1/2}) \right\|_F = \sqrt{\sum_{i=1}^M \log^2(\lambda_i)}$$

$$\text{et } \det(X_1 - \lambda X_2) = 0$$

Diffusion Isotrope Simulation



Diffusion anisotrope Simulation



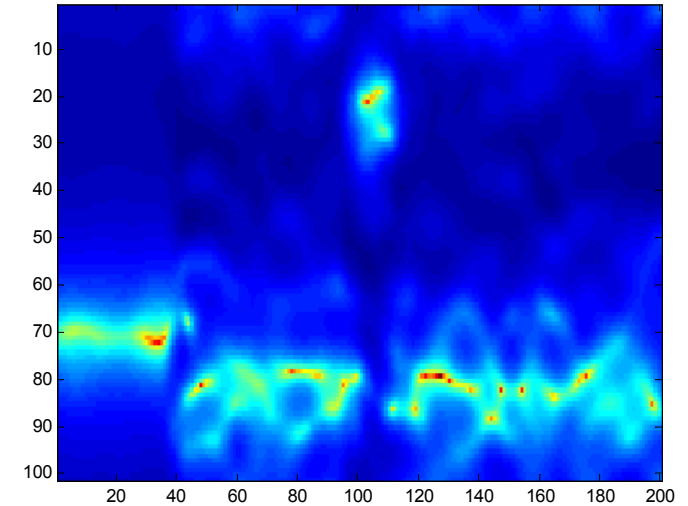
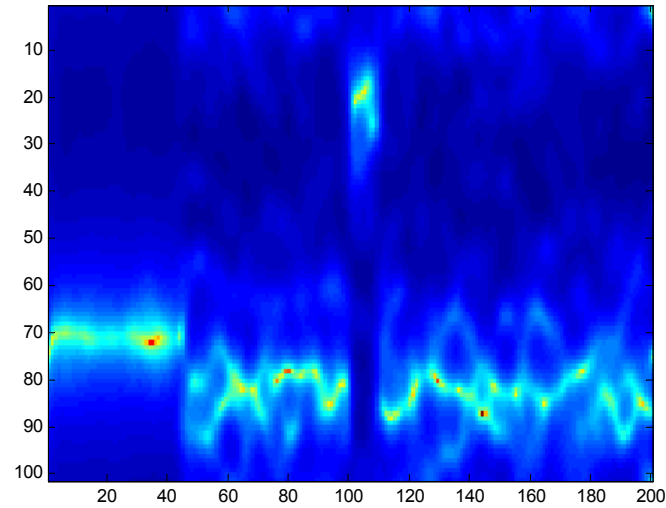
Les différentes diffusion à 50 itérations



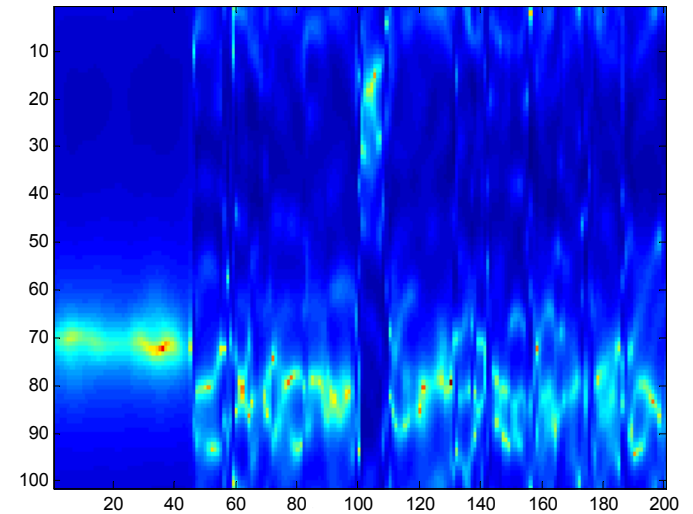
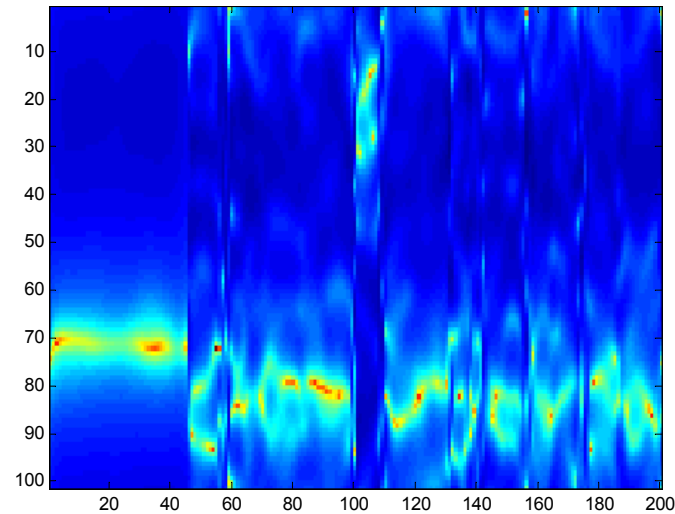
Diffusion classique

Diffusion médianage

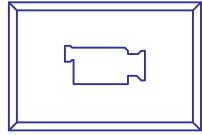
Diffusion
Isotrope



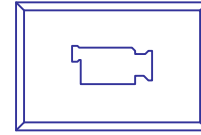
Diffusion
Anisotrope



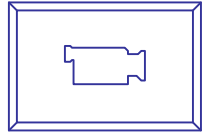
Les différentes diffusions



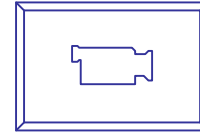
Diffusion Isotrope



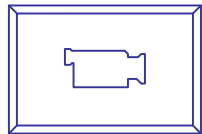
Diffusion Anisotrope



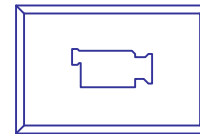
Medianage Isotrope



Medianage Anisotrope



Moyennage Isotrope



Moyennage Anisotrope

THALES



Métrique de Siegel pour les modèles autorégressifs complexes

Modèle Multivarié Gaussien Complexe Circulaire

■ Modèle du signal Radar :

- Modèle Multivarié Gaussien Complexe Circulaire :

$$p(Y_n / R_n) = (\pi)^{-n} |R_n|^{-1} \cdot e^{-Tr[\hat{R}_n \cdot R_n^{-1}]}$$

- Modèle Radar : avec $\hat{R}_n = (Y_n - m_n) \cdot (Y_n - m_n)^+$ et $E[\hat{R}_n] = R_n$

$m_n = 0$ processus de moyenne nulle

$R_n = E[Y_n Y_n^+]$ hermitienne Toeplitz définie positive

$$Y_n = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \text{ avec } y_k = a_k + ib_k = \rho_k e^{i\phi_k}$$

- Utilisation d'un modèle autorégressif complexe :

$$y_n = -\sum_{k=1}^N a_k^{(N)} y_{n-k} + b_n \text{ avec } E[b_n b_{n-k}^*] = \delta_{k,0} \sigma^2 \text{ et } A_N = [a_1^{(N)} \dots a_N^{(N)}]^T$$

- Littérature mathématique + lien avec l'algorithme d'Issai Schur (1875-1941)

- **[Alpay] D. Alpay, « Algorithme de Schur, espaces à noyau reproduisant et théorie des systèmes », Panoramas et synthèse, n°6, Société Mathématique de France, 1998**

Structures blocks des modèles CAR



- Les modèles autorégressifs complexe modélise le signal radar multivarié gaussien complexe circulaire de moyenne nulle.

- Rappel : Métrique de Rao-Siegel dans le cas Hermitien défini positif :

$$ds^2 = Tr\left(\left(R^{-1}dR\right)^2\right) = \left\|R^{-1/2}dRR^{-1/2}\right\|^2 \text{ avec } \|A\|^2 = \langle A, A \rangle \text{ et } \langle A, B \rangle = Tr(AB^T)$$

$$D^2(R_1, R_2) = \left\|\log\left(R_1^{-1/2}R_2R_1^{-1/2}\right)\right\|^2 = \sum_{k=1}^n \log^2(\lambda_k) \text{ et } \det\left(R_1^{-1/2}R_2R_1^{-1/2} - \lambda I\right) = 0$$

- La matrice de covariance a une structure blocks :

$$R_n^{-1} = \begin{bmatrix} \alpha_{n-1} & \alpha_{n-1} \cdot A_{n-1}^+ \\ \alpha_{n-1} \cdot A_{n-1} & R_{n-1}^{-1} + \alpha_{n-1} \cdot A_{n-1} \cdot A_{n-1}^+ \end{bmatrix} \quad R_n = \begin{bmatrix} \alpha_{n-1}^{-1} + A_{n-1}^+ \cdot R_{n-1} \cdot A_{n-1} & -A_{n-1}^+ \cdot R_{n-1} \\ -R_{n-1} \cdot A_{n-1} & R_{n-1} \end{bmatrix}$$

avec $\alpha_n^{-1} = \left[1 - |\mu_n|^2\right] \alpha_{n-1}^{-1}$ et $A_n = \begin{bmatrix} A_{n-1} \\ 0 \end{bmatrix} + \mu_n \cdot \begin{bmatrix} A_{n-1}^{(-)} \\ 1 \end{bmatrix}$ où $V^{(-)} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & \ddots & 0 \\ 1 & 0 & 0 \end{bmatrix} \cdot V^*$

Géométrie du demi-plan/disque de siegel

- Etude du domaine fondamental de Minkowski :
 - Domaine fondamental étudié par Minkowski (utilisation du système de coordonnées partielles d'Iwasawa) (lien avec l'algorithme de Goldberg : inversion d'une matrice de covariance avec modèle autorégressif)

For $Y \in P_n$ what is sometimes called the Iwasawa decomposition is given by

$$Y[g] = \bar{T}gY\bar{g} \quad Y = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix} \begin{bmatrix} 1 & & x_{ij} \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

with $d_i > 0$ and $x_{ij} \in \mathbf{R}$. This is the co-ordinate system Minkowski employed to find his fundamental domain.

$$Y = \begin{pmatrix} v & 0 \\ 0 & W \end{pmatrix} \begin{bmatrix} 1 & T_x \\ 0 & I_{n-1} \end{bmatrix}$$

where $v > 0$, $W \in P_{n-1}$, $x \in \mathbf{R}^{n-1}$, i.e., $T_x = (x_1, \dots, x_{n-1})$, $x_j \in \mathbf{R}$.

Structures blocks des modèles CAR



■ La matrice $R_n^{(1)-1/2+} \cdot R_n^{(2)} \cdot R_n^{(1)-1/2}$ ou $R_n^{(1)1/2+} \cdot R_n^{(2)-1} \cdot R_n^{(1)1/2}$ qui intervient dans la métrique de Siegel conserve la structure Blocks

■ Conservation de la structure Blocks :

$$\Omega_n = R_n^{(1)1/2+} \cdot R_n^{(2)-1} \cdot R_n^{(1)1/2} = \begin{bmatrix} \beta_{n-1} & \beta_{n-1} \cdot W_{n-1}^+ \\ \beta_{n-1} \cdot W_{n-1} & \Omega_{n-1} + \beta_{n-1} \cdot W_{n-1} \cdot W_{n-1}^+ \end{bmatrix}$$

avec $W_{n-1} = \sqrt{\alpha_{n-1}^{(1)}} \cdot R_{n-1}^{(1)1/2+} \cdot [A_{n-1}^{(2)} - A_{n-1}^{(1)}]$ et $\beta_{n-1} = \frac{\alpha_{n-1}^{(2)}}{\alpha_{n-1}^{(1)}}$

■ On en déduit que les valeurs propres étendues sont zéros de la fonction :

$$\begin{cases} F^{(n)}(\lambda_k^{(n)}) = \lambda_k^{(n)} - \beta_{n-1} + \beta_{n-1} \cdot \lambda_k^{(n)} \cdot \sum_{i=1}^{n-1} \frac{|W_{n-1}^+ \cdot X_i^{(n-1)}|^2}{(\lambda_i^{(n-1)} - \lambda_k^{(n)})} = 0 \\ \frac{X_k^{(n)}}{X_{k,1}^{(n)}} = \begin{bmatrix} 1 \\ -\lambda_k^{(n)} \cdot U_{n-1} \cdot (\Lambda_{n-1} - \lambda_k^{(n)} \cdot I_{n-1})^{-1} \cdot U_{n-1}^+ \cdot W_{n-1} \end{bmatrix} \end{cases}$$

Structures blocks des modèles CAR



■ On en déduit un algorithme récursif sur l'ordre et parallélisable pour chaque valeur propre étendue :

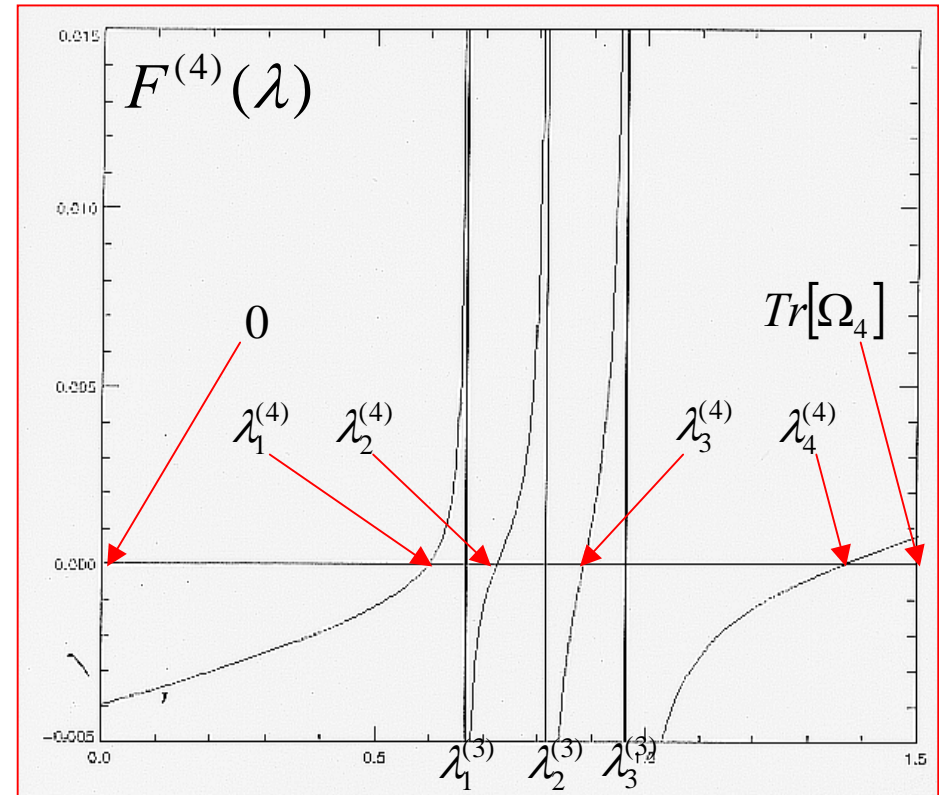
$$\begin{cases} F^{(n)}(\lambda_k^{(n)}) = \lambda_k^{(n)} - \beta_{n-1} + \beta_{n-1} \cdot \lambda_k^{(n)} \cdot \sum_{i=1}^{n-1} \frac{|W_{n-1}^+ \cdot X_i^{(n-1)}|^2}{(\lambda_i^{(n-1)} - \lambda_k^{(n)})^2} = 0 \\ \frac{X_k^{(n)}}{X_{k,1}^{(n)}} = \begin{bmatrix} 1 \\ -\lambda_k^{(n)} \cdot U_{n-1} \cdot (\Lambda_{n-1} - \lambda_k^{(n)} \cdot I_{n-1})^{-1} \cdot U_{n-1}^+ \cdot W_{n-1} \end{bmatrix} \end{cases}$$

Renormalisation à chaque itération :

$$\frac{X_k^{(n)}}{X_{k,1}^{(n)}} \rightarrow \left(1 + \lambda_k^{(n)2} \cdot \sum_{i=1}^{n-1} \frac{|W_{n-1}^+ \cdot X_i^{(n-1)}|^2}{(\lambda_i^{(n-1)} - \lambda_k^{(n)})^2} \right)^{-1/2} \cdot \frac{X_k^{(n)}}{X_{k,1}^{(n)}}$$

Interprétation en terme de projection

$$\begin{cases} \sum_{k=1}^n \left(\frac{\partial F^{(n)}(\lambda_k^{(n)})}{\partial \eta} \right)^{-1} = 1 \\ \begin{bmatrix} 1 \\ A_{n-1} \end{bmatrix} = \sum_{k=1}^n \left(\frac{\partial F^{(n)}(\lambda_k^{(n)})}{\partial \eta} \right)^{-1} \cdot \frac{X_k^{(n)}}{X_{k,1}^{(n)}} \end{cases}$$



$$0 < \lambda_n^{(n)} < \lambda_{n-1}^{(n-1)} < \lambda_{n-1}^{(n)} < \dots < \lambda_2^{(n)} < \lambda_1^{(n-1)} < \lambda_1^{(n)} < Tr[\Omega_n]$$

Courbe strictement croissante sur chaque intervalle :

$$\frac{\partial F^{(n)}(\eta)}{\partial \eta} = \frac{\alpha_{n-1}}{\eta_k^{(n)} \cdot |X_{k,1}^{(n)}|^2} = 1 + \alpha_{n-1} \cdot \sum_{k=1}^{n-1} \frac{\eta_k^{(n-1)} \cdot |A_{n-1}^+ \cdot X_k^{(n-1)}|^2}{(\eta_k^{(n-1)} - \eta)^2} > 1$$

avec $Tr[\Omega_n] = Tr[\Omega_{n-1}] + \alpha_{n-1} \cdot \begin{bmatrix} 1 \\ A_{n-1} \end{bmatrix}^+ \begin{bmatrix} 1 \\ A_{n-1} \end{bmatrix}$

Récurtivité sur l'ordre de la métrique de Siegel



■ Si on utilise la structure blocks des matrices, on peut calculer la métrique de Siegel à l'ordre n par rapport à l'ordre n-1

■ Equation réursive sur la métrique de Siegel faisait apparaître la matrice de Fisher sur les paramètres autorégressifs:

$$ds_n^2 = Tr\left[\left(R_n \cdot dR_n^{-1}\right)^2\right] = ds_{n-1}^2 + \left(\frac{d\alpha_{n-1}}{\alpha_{n-1}}\right)^2 + \alpha_{n-1} \cdot dA_{n-1}^+ \cdot R_{n-1} \cdot dA_{n-1}$$

■ On peut remarquer que le second terme s'écrit :

$$\left(\frac{d\alpha_{n-1}}{\alpha_{n-1}}\right)^2 + \alpha_{n-1} \cdot dA_{n-1}^+ \cdot R_{n-1} \cdot dA_{n-1} = d \begin{bmatrix} \log \alpha_{n-1} \\ A_{n-1} \end{bmatrix}^+ \cdot \begin{bmatrix} 1 & 0 \\ 0 & \alpha_{n-1} R_{n-1} \end{bmatrix} d \begin{bmatrix} \log \alpha_{n-1} \\ A_{n-1} \end{bmatrix}$$

■ Ceci s'interprète en terme de géométrie de l'information

$$\Sigma_{\hat{A}_n} = I[A_{n-1}] = \alpha_{n-1} R_{n-1} \xrightarrow{\text{red arrow}} Z = \begin{bmatrix} \log \alpha_{n-1} \\ A_{n-1} \end{bmatrix} \begin{bmatrix} \log \alpha_{n-1} \\ A_{n-1} \end{bmatrix}^+ + i \begin{bmatrix} 1 & 0 \\ 0 & \alpha_{n-1} R_{n-1} \end{bmatrix}$$

$\log \alpha_{n-1} \perp A_{n-1}$

**Extension du cas
Gaussien scalaire
(Géométrie de l'information)**



Racine carré non-symétrique du groupe de Siegel

On s'intéresse à la décomposition de Cholesky de la matrice :

En faisant intervenir les racine carrée des matrices définies positives :

$$\Omega_n = (\alpha_n \cdot R_n)^{-1} = W_n \cdot W_n^+ = \left(1 - |\mu_n|^2\right) \begin{bmatrix} 1 & A_{n-1}^+ \\ A_{n-1} & \Omega_{n-1} + A_{n-1} \cdot A_{n-1}^+ \end{bmatrix}$$

avec $W_n = \sqrt{1 - |\mu_n|^2} \begin{bmatrix} 1 & 0 \\ A_{n-1} & \Omega_{n-1}^{1/2} \end{bmatrix}$ et $\Omega_{n-1} = \Omega_{n-1}^{1/2} \cdot \Omega_{n-1}^{1/2+}$

Toute distribution de variable n-dimensionnel est associée $Z \sim N_n(0, I_n)$ naturellement au groupe Affine : c'est l'élément tel que son action sur le vecteur $X \sim N_n(A_n, \Omega_n)$ le transforme en vecteur aléatoire :

$$\begin{bmatrix} 1 & 0 \\ A_{n-1} & \Omega_{n-1}^{1/2} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ Z \end{bmatrix} = \begin{bmatrix} 1 \\ \Omega_{n-1}^{1/2} \cdot Z + A_{n-1} \end{bmatrix} = \begin{bmatrix} 1 \\ X \end{bmatrix}$$

On peut considérer cette représentation des éléments du groupe affine comme la racine carrée non-symétrique d'un élément du groupe de Siegel :

$$\begin{bmatrix} 1 & A_{n-1}^+ \\ A_{n-1} & \Omega_{n-1} + A_{n-1} \cdot A_{n-1}^+ \end{bmatrix}$$

Expression récursive de la divergence de Kullback

On s'intéresse à l'expression récursive de la divergence de Kullback :

- La structure de la divergence de Kullback est donnée par :

$$K(R_n^{(1)}, R_n^{(2)}) = \frac{1}{2} \left[\text{Tr}[\Sigma_n] - \ln |\Sigma_n| - n \right]$$

$$\Sigma_n = R_n^{(1)1/2+} \cdot R_n^{(2)-1} \cdot R_n^{(1)1/2} = \begin{bmatrix} 2\beta_{n-1} & \beta_{n-1} \cdot V_{n-1}^+ \\ \beta_{n-1} \cdot V_{n-1} & \Sigma_{n-1} + \beta_{n-1} \cdot V_{n-1} \cdot V_{n-1}^+ \end{bmatrix}$$

où $V_{n-1} = F_{n-1}^{(1)+} \cdot [A_{n-1}^{(2)} - A_{n-1}^{(1)}]$ et $\beta_{n-1} = \frac{\alpha_{n-1}^{(2)}}{\alpha_{n-1}^{(1)}}$ avec $\alpha_n^{(1)} \cdot R_n^{(1)} = F_n^{(1)} \cdot F_n^{(1)+}$

$$F_n^{(1)} = \frac{1}{\sqrt{1 - |\mu_n^{(1)}|^2}} \begin{bmatrix} 1 & -A_{n-1}^{(1)+} \\ 0 & I_{n-1} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & F_{n-1}^{(1)} \end{bmatrix}$$

- Si on utilise de plus les relations suivantes :

$$|\Sigma_n| = \beta_{n-1} \cdot |\Sigma_{n-1}| \quad \text{Tr}[\Sigma_n] = \text{Tr}[\Sigma_{n-1}] + \beta_{n-1} \cdot [1 + V_{n-1}^+ \cdot V_{n-1}]$$

- On obtient l'expression récursive sur l'ordre de la divergence de Kullback :

$$K(R_n^{(1)}, R_n^{(2)}) = K(R_{n-1}^{(1)}, R_{n-1}^{(2)}) + \frac{1}{2} \left[(\beta_{n-1} - 1) + \beta_{n-1} \cdot V_{n-1}^+ \cdot V_{n-1} - \ln(\beta_{n-1}) \right]$$

THALES



**Modélisation variationnelle de
l'analyse autorégressive
complexe et le flot de la chaleur
sur ces modèles**

Formalisation variationnelle de l'analyse AR

Le théorème de Parseval permet de modéliser l'analyse autorégressive complexe sous forme variationnelle :

Le théorème de Parseval nous dit que :

$$\int_0^1 |A^{(n)}(f)|^2 \hat{S}_z(f) df = E[|e_n(k)|^2] \quad e_n(k) = z_k - \hat{z}_k = \sum_{k=0}^n a_k^{(n)} z_{n-k}$$

$$A^{(n)}(f) = \sum_{k=0}^n a_k^{(n)} e^{-i2\pi f k} \quad (\text{et } a_0^{(n)} = 1), \quad \hat{S}_z(f) = \sum_{k=-\infty}^{+\infty} \hat{r}_k e^{-j2\pi f k} \quad \text{et } \hat{r}_k = E[z_m z_{m-k}^*]$$

Si on rajoute un terme de régularisation (a priori de douceur spectrale)

⋮

$$S[A^{(n)}(f)] = \int_0^1 L(f, A^{(n)}, \nabla_f A^{(n)}) df \quad \text{avec} \quad \text{cas quadratique}$$

$$L(f, A^{(n)}, \nabla_f A^{(n)}) = |A^{(n)}(f)|^2 \hat{S}_z(f) + \lambda \Phi(\nabla_f A^{(n)}(f)) \quad \Phi(X) = |X|^2$$

Grâce à l'isomorphisme canonique entre le plan Euclidien et le plan complexe, nous formalisons le problème de régularisation de la prédiction linéaire sous l'angle du calcul des variations.

Formalisation variationnelle de l'analyse AR

La solution variationnelle de l'analyse autorégressive, donnée par l'équation d'Euler-Lagrange Complexe, est solution de l'EDP de type « Schrödinger » :

- Equation d'Euler-Lagrange « complexe » (régularisation quadratique) :

$$L(f, A^{(n)}, \nabla_f A^{(n)}) = |A^{(n)}(f)|^2 \hat{S}_z(f) + \lambda |\nabla_f A^{(n)}(f)|^2$$

$$\frac{d}{df} \left[\frac{\partial L}{\partial \nabla_f A^{(n)*}} \right] = \frac{\partial L}{\partial A^{(n)*}}$$

- Solution de l'EDP :

$$\frac{\partial A^{(n)}(f)}{\partial t} = \lambda \cdot \frac{\partial^2 A^{(n)}(f)}{\partial f^2} - A^{(n)}(f) \cdot \hat{S}_z(f)$$

- Dans le cas non régularisé, le Cepstre évolue comme suit :

$$C^{(n)}(f) = -\ln A^{(n)}(f) \Rightarrow \frac{\partial C^{(n)}(f)}{\partial t} = S_z(f)$$

Formalisation variationnelle de l'analyse AR



Le théorème de Kolmogorov-Szegö-Krein devient alors une conséquence du théorème d'Emmy Noether :

- La fonctionnelle est invariante selon :

$$S(\phi_s \bullet A^{(n)}(\psi_n(f))) = S(A^{(n)}(f))$$

avec $\phi_s \bullet A^{(n)}(f) = \phi_s(A^{(n)}(f)) = e^{i \cdot s} \cdot A^{(n)}(f)$ et $\psi_n(f) = f + n$

où $s \in \mathfrak{R}$ et $n \in N$

- Le théorème de Noether nous indique alors que :

$$\frac{\partial \left[|A^{(n)}(f)|^2 \cdot S_z(f) \right]}{\partial f} = 0 \Rightarrow |A^{(n)}(f)|^2 \cdot S_z(f) = cste \quad \forall f$$

- Ce qui est l'objet du théorème de Kolmogorov-Szegö-Krein :

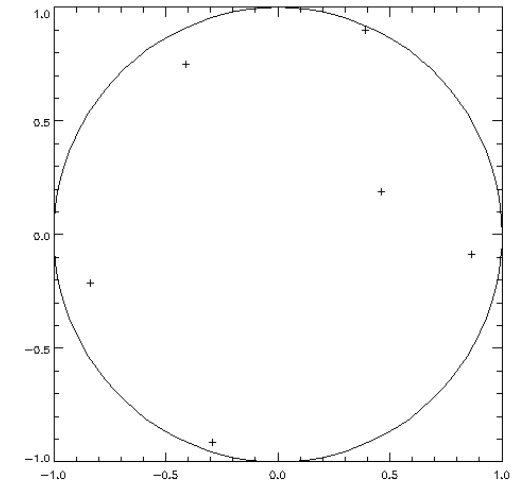
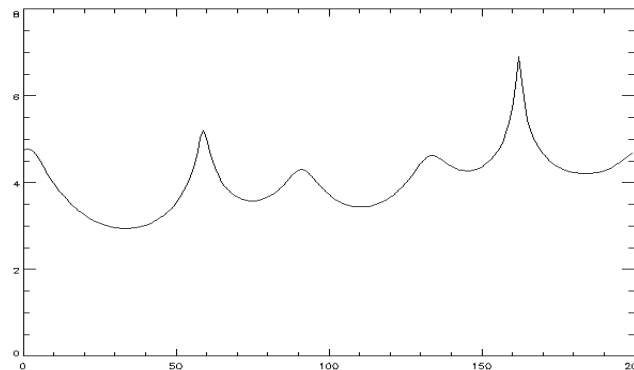
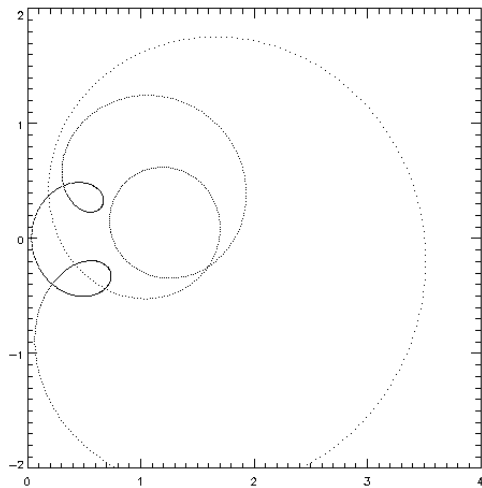
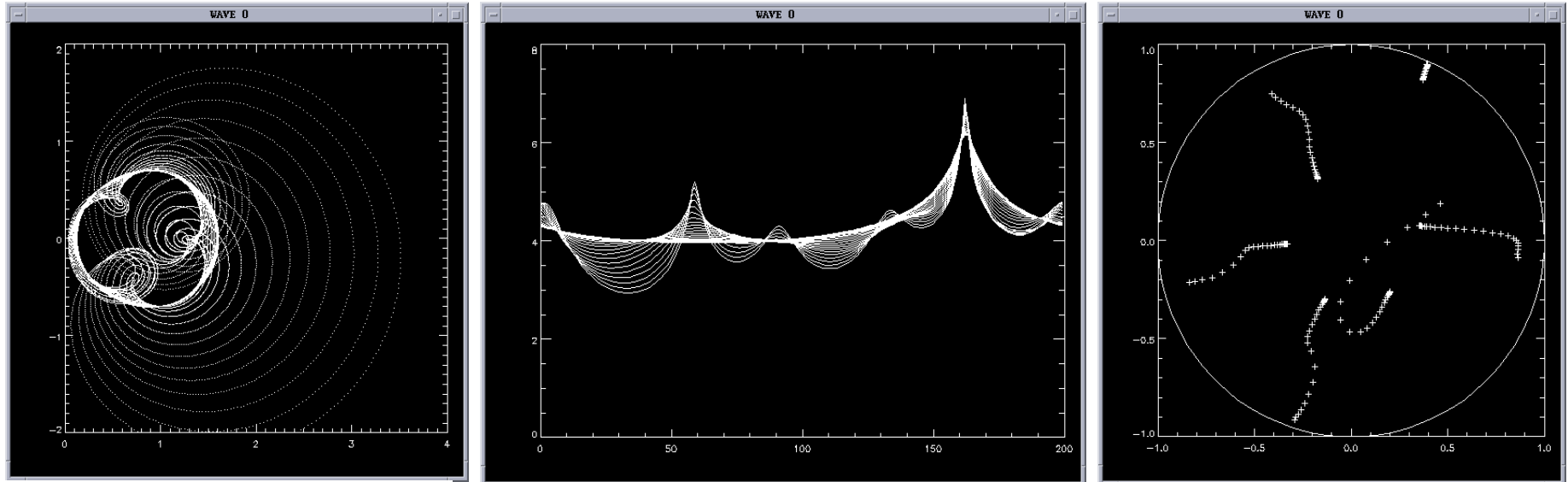
$$\text{Inf}_{A^{(n)}} \int_{-1/2}^{1/2} |A^{(n)}(f)|^2 \cdot S_z(f) \cdot df = e^{-\int_{-1/2}^{1/2} \log S_z(f) \cdot df}$$

$$S_z(f) = \frac{\sigma^2}{|A^{(n)}(f)|^2}$$

Régularisation variationnelle d'un modèle AR

Exemple de régularisation :

- On observe la disparition des pics parasites sur une analyse Autorégressive à très faible nombre d'échantillon :



THALES



Procédé par
Médiane de modèles
autorégressifs complexes



La géométrie d'Erich Kähler étend la géométrie Riemannienne au domaine complexe :

- La forme riemannienne définie positive définissant la métrique de Kähler est donnée par :

$$ds^2 = 2 \sum_{i,j=1}^n g_{i\bar{j}} \cdot dz^i dz^{\bar{j}}$$

- Condition de Kähler condition : Il existe localement une fonction potentielle de Kähler, Φ (et les équivalentes Pluri-harmoniques) telle que :

$$g_{i\bar{j}} = \frac{\partial^2 \Phi}{\partial z^i \partial \bar{z}^j}$$

- Le tenseur de courbure est alors donné par :

$$R_{i\bar{j}} = - \frac{\partial^2 \log(\det g_{k\bar{l}})}{\partial z_i \partial \bar{z}_j}$$

- Et la courbure scalaire :

$$R = \sum_{k,l=1}^n g^{k\bar{l}} \cdot R_{k\bar{l}}$$

Modèle Autorégressif Complexe

■ Modèle du signal radar :

- Modèle multivarié complexe circulaire :

$$p(Z_n / R_n) = (\pi)^{-n} |R_n|^{-1} \cdot e^{-Tr[\hat{R}_n \cdot R_n^{-1}]}$$

$$\text{avec } \hat{R}_n = (Z_n - m_n)(Z_n - m_n)^+ \text{ et } E[\hat{R}_n] = R_n$$

- Modèle Radar :

$m_n = 0$ processus de moyenne nulle

$R_n = E[Z_n Z_n^+]$ Toeplitz Hermitienne Définie Positive

$Z_n = [z_1 \cdots z_n]^T$ avec $z_k = x_k + iy_k = \rho_k e^{i\phi_k}$

- Modèle autorégressif complexe :

$$z_n = -\sum_{k=1}^N a_k^{(N)} z_{n-k} + b_n \text{ avec } E[b_n b_{n-k}^*] = \delta_{k,0} \sigma^2 \text{ et } A_N = [a_1^{(N)} \cdots a_N^{(N)}]^T$$

- Lien avec l'algorithme de Issai Schur's algorithm (1875-1941)

- **[Alpay] D. Alpay, « Algorithme de Schur, espaces à noyau reproduisant et théorie des systèmes », Panoramas et synthèse, n°6, Société Mathématique de France**

Estimation des paramètres autorégressifs régularisés

- Nous utilisons l'algorithme de Burg régularisé (brevet THALES : F. Barbaresco, « Procédé et dispositif de détermination du spectre de fréquence d'un signal », brevet n° 95 06983, Juin 1995)

.Initialisation :

$$f_0(k) = b_0(k) = z(k) \quad , \quad k=1, \dots, N \quad (N: \text{nb. ech.})$$

$$P_0 = \frac{1}{N} \cdot \sum_{k=1}^N |z(k)|^2$$

$$a_0^{(0)} = 1$$

.Itération (n): Pour $n = 1$ à M

$$\mu_n = - \frac{\frac{2}{N-n} \sum_{k=n+1}^N f_{n-1}(k) \cdot b_{n-1}^*(k-1) + 2 \cdot \sum_{k=1}^{n-1} \beta_k^{(n)} \cdot a_k^{(n-1)} \cdot a_{n-k}^{(n-1)}}{\frac{1}{N-n} \sum_{k=n+1}^N |f_{n-1}(k)|^2 + |b_{n-1}(k-1)|^2 + 2 \cdot \sum_{k=0}^{n-1} \beta_k^{(n)} \cdot |a_k^{(n-1)}|^2} \quad \text{avec} \quad \beta_k^{(n)} = \gamma_1 \cdot (2\pi)^2 \cdot (k-n)^2$$

$$\begin{cases} a_0^{(n)} = 1 \\ a_k^{(n)} = a_k^{(n-1)} + \mu_n \cdot a_{n-k}^{(n-1)*} \quad , \quad k=1, \dots, n-1 \\ a_n^{(n)} = \mu_n \end{cases}$$

$$\begin{cases} f_n(k) = f_{n-1}(k) + \mu_n \cdot b_{n-1}(k-1) \\ b_n(k) = b_{n-1}(k-1) + \mu_n^* \cdot f_{n-1}(k) \end{cases}$$

modèle autorégressif et matrice de covariance

- La matrice de covariance et son inverse peuvent être paramétrée à partir des paramètres du modèle autorégressif :
 - Structure Blocs de la matrice de covariance :

$$R_n^{-1} = \begin{bmatrix} \alpha_{n-1} & \alpha_{n-1} \cdot A_{n-1}^+ \\ \alpha_{n-1} \cdot A_{n-1} & R_{n-1}^{-1} + \alpha_{n-1} \cdot A_{n-1} \cdot A_{n-1}^+ \end{bmatrix} \quad R_n = \begin{bmatrix} \alpha_{n-1}^{-1} + A_{n-1}^+ \cdot R_{n-1} \cdot A_{n-1} & -A_{n-1}^+ \cdot R_{n-1} \\ -R_{n-1} \cdot A_{n-1} & R_{n-1} \end{bmatrix}$$

avec $\alpha_n^{-1} = [1 - |\mu_n|^2] \alpha_{n-1}^{-1}$ et $A_n = \begin{bmatrix} A_{n-1} \\ 0 \end{bmatrix} + \mu_n \cdot \begin{bmatrix} A_{n-1}^{(-)} \\ 1 \end{bmatrix}$ où $V^{(-)} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & \ddots & 0 \\ 1 & 0 & 0 \end{bmatrix} \cdot V^*$

Métrie de Kähler d'un modèle autorégressif complexe



■ Dans le cadre de la géométrie Affine de l'information, la métrique est donnée par le Hessien de l'entropie qui joue ici le rôle du potentiel de Kähler :

- L'entropie d'un modèle multivarié gaussien de moyenne nulle :

$$p(Z_n / R_n) = (\pi)^{-n} \cdot |R_n|^{-1} \cdot e^{-Tr[\hat{R}_n \cdot R_n^{-1}]} \implies \tilde{\Phi}(R) = -\log(\det R) - n \log(\pi e)$$

avec $\hat{R}_n = (Z_n - m_n) \cdot (Z_n - m_n)^+$

et $E[\hat{R}_n] = R_n$

$$g_{ij} \equiv \frac{\partial^2 \tilde{\Phi}}{\partial H_i \partial H_j} \text{ et } H = -R$$

- On peut choisir comme modèle de paramétrisation les coefficients de réflexion

$$\tilde{\Phi}(R_n) = \sum_{k=1}^{n-1} (n-k) \cdot \ln[1 - |\mu_k|^2] + n \cdot \ln[\pi \cdot e \cdot P_0]$$

$$\alpha_n^{-1} = [1 - |\mu_n|^2] \alpha_{n-1}^{-1}$$

$$\det(R_n) = \prod_{k=0}^{n-1} \alpha_k^{-1} = \alpha_0^{-n} \prod_{k=1}^{n-1} [1 - |\mu_k|^2]^{n-k}$$

avec $\alpha_0^{-1} = P_0 = \frac{1}{n} \sum_{k=1}^n |z_k|^2$

Papier initial d'Erich Kähler, 1932, Hambourg

« Kähler Erich, Mathematical Works », Edited by R. Berndt and O. Riemenschneider, Berlin, Walter de Gruyter, ix, 2003

Erich Kähler

$$ds^2 = \sum \frac{\partial^2 U}{\partial x_i \partial \bar{x}_k} dx_i d\bar{x}_k$$

invariant gegenüber der Gruppe der „hyperfuchschen“ Transformationen⁽⁴⁾. Analog ist für die „hyperabelschen“ Transformationen

$$x_i = \frac{\alpha_i x_i + \beta_i}{\gamma_i x_i + \delta_i} \quad (i = 1, 2, \dots, n),$$

die die Einheitskreise

$$1 - x_i \bar{x}_i = 0 \quad (i = 1, 2, \dots, n)$$

festlassen, die aus dem Potential

$$U = \sum_{i=1}^n k_i \log(1 - x_i \bar{x}_i) \quad (k_i \text{ Konstanten})$$

ableitbare Metrik invariant, und es ist klar, daß auch alle Zwischenfälle, etwa die aus hyperfuchschen Transformationen in r und s ($r+s = n$) Variablen komponierten Gruppen, zu Metriken von jenem Typus führen.

Über eine bemerkenswerte Hermitesche Metrik.

Von ERICH KÄHLER in Hamburg.

1.

Bei der Untersuchung der Invarianten einer reell $2n$ -dimensionalen HERMITESCHEN Metrik¹⁾

$$(1) \quad ds^2 = \sum g_{i\bar{k}} dx_i d\bar{x}_k$$

...

Auf die Verwendung der vorliegenden formalen Entwicklungen für die Theorie der automorphen Funktionen und auf die Analogie der Gleichungen

$$D(U) = 0, \quad \Delta(U) = e^{kU}$$

mit den klassischen Differentialgleichungen

$$\Delta U = 0, \quad \Delta U = e^{kU}$$

auf die bereits G. GIRAUD¹⁾ und A. BLOCH²⁾ hingewiesen haben, gedenke ich in einer späteren Arbeit einzugehen.

¹⁾ G. GIRAUD, Sur une équation aux dérivées partielles, non linéaire etc. Comptes Rendus 166, I (1918), p. 893.

²⁾ A. BLOCH, Sur une nouvelle et importante généralisation de l'équation de Laplace. L'Enseignement Mathém. 26 (1927), p. 52.

Hamburg, den 22. Oktober 1932.

L'entropie $U = -\log \det[R]$, pour les modèles autorégressifs complexes, peut être considéré comme un potentiel de Kähler qui est paramétrisé à partir des coefficients de réflexion. Cette expression est alors la même que celle proposée par E. Kähler

Métrie de Kähler : cas hyper-Abelian

- On retrouve la métrie proposée par E. Kähler dans son papier initial : :

- Kähler a nommé ce cas « hyper-Abelian » :

$$\Phi = \sum_{k=1}^{n-1} \alpha_k \cdot \ln \left[1 - |z_k|^2 \right] = \ln K_D(z, z)$$

avec

$$\text{Noyau de Bregman : } K_D(z, z) = \prod_{k=1}^{n-1} \left(1 - |z_k|^2 \right)^{\alpha_k}$$

$$\text{et } \left\{ z / |z_k| < 1 \quad \forall k = 1, \dots, n \right\}$$

- Autre métrie proposée par Kähler, appelé cas « hyper-fuchien » :

$$\Phi = \alpha \cdot \ln \left(1 - \sum_{k=1}^n |z_k|^2 \right) \quad \text{with} \quad \left\{ z / \sum_{k=1}^n |z_k|^2 < 1 \right\}$$

Métrie de Kähler d'un modèle autorégressif complexe

On définit une métrique « Doppler » metric dans le cas de modèles autorégressifs complexes comme le Hessien d'un potentiel de Kähler, où le potentiel est donné comme dans le cas de la géométrie affine de l'information par l'Entropie:

- Le potentiel de Kähler paramétré par $[P_0 \quad \mu_1 \quad \cdots \quad \mu_{n-1}]^T$:

$$\tilde{\Phi}(R_n) = \sum_{k=1}^{n-1} (n-k) \cdot \ln[1 - |\mu_k|^2] + n \cdot \ln[\pi \cdot e \cdot \alpha_0^{-1}]$$

- La métrique associée est alors calculée :

$$\theta^{(n)} = [P_0 \quad \mu_1 \quad \cdots \quad \mu_{n-1}]^T = [\theta_1^{(n)} \quad \cdots \quad \theta_n^{(n)}]^T$$

$$g_{11} = n\alpha_0^2 = nP_0^{-2} \quad g_{ij} = \frac{(n-i) \cdot \delta_{ij}}{(1 - |\mu_i|^2)^2}$$

$$ds_n^2 = n \cdot \left(\frac{dP_0}{P_0} \right)^2 + \sum_{i=1}^{n-1} (n-i) \frac{|d\mu_i|^2}{(1 - |\mu_i|^2)^2}$$

Courbure scalaire du modèle autorégressif complexe

- On utilise l'expression du tenseur de Ricci dans le cas de la géométrie de Kähler :

- En géométrie de Kähler, le tenseur de Ricci est donné par :

$$R_{i\bar{j}} = -\frac{\partial^2 \log(\det g_{k\bar{l}})}{\partial z_i \partial \bar{z}_j}$$

- Pour le cas du modèle autorégressif complexe, on obtient alors :

$$\begin{cases} R_{11} = -2 \frac{1}{P_0^2} \\ R_{k\bar{l}} = -2 \frac{\delta_{kl}}{(1 - |\mu_k|^2)^2} \text{ for } k = 2, \dots, n-1 \end{cases}$$

- On en déduit le calcul de la courbure scalaire qui est négative :

$$R = \sum_{k,l} g^{k\bar{l}} \cdot R_{k\bar{l}}$$

$$R = -2 \cdot \left[\sum_{j=0}^{n-1} \frac{1}{(n-j)} \right]_{n \rightarrow \infty} \rightarrow -\infty$$

Modèle autorégressif # métrique de Kähler-Einstein

La métrique précédente n'est pas une métrique de type Kähler-Einstein, mais une structure « matricielle » proche

- Une métrique est dite de Kähler-Einstein metric si son tenseur de ricci est proportionnel à la métrique :

$$R_{i\bar{j}} = k_0 \cdot g_{k\bar{l}} \quad \text{avec } k_0 : \text{constant} \Rightarrow -\frac{\partial^2 \log(\det g_{k\bar{l}})}{\partial z_i \partial \bar{z}_j} = k_0 \cdot \frac{\partial^2 \Phi}{\partial z^i \partial \bar{z}^j}$$

- Dans le cas de la métrique de Kähler-Einstein, le potentiel de Kähler est solution de l'équation de Monge-Ampère :

$$\det(g_{k\bar{l}}) = |\psi|^2 e^{-k_0 \Phi} \quad \text{avec} \quad \begin{cases} \Phi : \text{Potentiel de Kähler} \\ \psi : \text{fonction holomorphe} \end{cases}$$

- Pour un modèle autorégressif complexe, on a :

$$[R_{ij}] = B^{(n)} [g_{ij}] \quad \text{avec} \quad R = \text{Tr}[B^{(n)}] = -2 \cdot \left[\sum_{j=0}^{n-1} \frac{1}{(n-j)} \right]$$

$$\text{et où } B^{(n)} = -2 \text{diag} \{ \dots, (n-i)^{-1}, \dots \}$$

Flot de Calabi & Kähler-Einstein

Il existe 2 flots particuliers qui sont déduits de fonctionnelles dépendant uniquement de la courbure scalaire intégrée sur la variété

- Le flot de Kähler-Einstein est donné par :

Action de Hilbert

$$\int_M R(g) dV(g)$$



Flot de Ricci

$$\frac{\partial g_{\bar{i}\bar{j}}}{\partial t} = -R_{\bar{i}\bar{j}} + \frac{\bar{R}}{n} g_{\bar{i}\bar{j}}$$

$$\bar{R} = \frac{\int_M R(g) dV(g)}{\int_M dV(g)}$$

- Le flot de Calabi est donné par (il agit directement sur le potentiel) :

Action de Calabi

$$S(g) = \int_M R^2(g) dV(g)$$



Flot de Calabi

$$\frac{\partial g_{\bar{i}\bar{j}}}{\partial t} = \frac{\partial^2 R}{\partial z_i \partial z_j^*}$$

$$g_{\bar{i}\bar{j}} = \frac{\partial^2 \phi(z, z^*)}{\partial z_i \partial z_j^*}$$

Flot de Calabi

$$\frac{\partial \phi}{\partial t} = R - \bar{R}$$

- L'inégalité entre les deux fonctionnelles est donnée par :

$$S(g) \geq \left(\int_M R(g) dV(g) \right)^2 / \int_M dV(g)$$

Ionnas Bakas : Relation entre les 2 flots

- En dimension 2 de la variable complexe, les deux flots sont reliés l'un à l'autre :

- Ionnas Bakas a montré le lien entre les 2 flots en remarquant que :

$$\frac{\partial g_{i\bar{j}}}{\partial t} = -R_{i\bar{j}} \Rightarrow \begin{cases} \frac{\partial^2 g_{i\bar{j}}}{\partial t^2} = -\frac{\partial R_{i\bar{j}}}{\partial t} \\ R_{i\bar{j}} = -\frac{\partial^2 \log(\det(g))}{\partial z_i \partial z_j^*} \end{cases} \Rightarrow \begin{cases} \frac{\partial^2 g_{i\bar{j}}}{\partial t^2} = \frac{\partial^2}{\partial z_i \partial z_j^*} \left(\frac{\partial \log(\det(g))}{\partial t} \right) \\ \frac{\partial \log(\det(g))}{\partial t} = -\sum_{i,j} g^{i\bar{j}} R_{i\bar{j}} = -R \end{cases}$$

$$\Rightarrow \frac{\partial^2 g_{i\bar{j}}}{\partial t^2} = -\frac{\partial^2 R}{\partial z_i \partial z_j^*}$$

- Si la dérivée seconde par rapport au temps du flot de Kähler-Ricci est identifiée à l'opposée de la dérivée première du temps du flot de Calabi, alors les 2 flots sont équivalents :

$$\frac{\partial^2}{\partial t_R^2} = -\frac{\partial}{\partial t_C}$$

Ionnas Bakas : flots agissant sur le potentiel

■ Bakas a établie les relation suivante $ds^2 = 2e^{\Phi(z, z^*, t)} dz \cdot dz^*$

$$R_{zz^*} = -\frac{\partial^2 \Phi}{\partial z \partial z^*} \quad R_{zz^*} = -\frac{\partial^2 \log(\det(g))}{\partial z \partial z^*} \quad \text{et} \quad g_{zz^*} = e^{\Phi}$$

Flot de Kähler-Ricci

$$\frac{\partial g_{i\bar{j}}}{\partial t} = -R_{i\bar{j}} \Rightarrow \frac{\partial e^{\Phi}}{\partial t} = \frac{\partial^2 \Phi}{\partial z \partial z^*} \Rightarrow e^{\Phi} \frac{\partial \Phi}{\partial t} = \frac{\partial^2 \Phi}{\partial z \partial z^*}$$

$$\Rightarrow \frac{\partial \Phi}{\partial t} = \Delta \Phi \quad \text{avec} \quad \Delta. = e^{-\Phi} \frac{\partial^2}{\partial z \partial z^*}$$

Flot de Calabi

$$\frac{\partial g_{i\bar{j}}}{\partial t} = \frac{\partial^2 R}{\partial z \partial z^*} \quad \text{et} \quad R = \sum_{i,j} g^{i\bar{j}} R_{i\bar{j}} = -e^{-\Phi} \frac{\partial^2 \Phi}{\partial z \partial z^*} = -\Delta \Phi$$

$$\Rightarrow \frac{\partial e^{\Phi}}{\partial t} = -\frac{\partial^2 \Delta \Phi}{\partial z \partial z^*}$$

$$\Rightarrow \frac{\partial \Phi}{\partial t} = -\Delta \Delta \Phi$$

flots de kähler-Ricci et Calabi sur les modèle AR

- Si on étudie sur la base de la métrique de Siegel (donné par le Hessien de l'Entropie du processus AR), la métrique est déformée comme suit :

- Flot de Kähler-Ricci sur les coefficients de réflexion :

$$\frac{\partial g_{ij}}{\partial t} = -R_{ij}$$

$$\begin{cases} |\mu_i(t)|^2 = 1 - \left(1 - |\mu_i(0)|^2\right) e^{-\frac{t}{(n-i)}} \xrightarrow{t \rightarrow \infty} |\mu_i(\infty)|^2 = 1 \\ P_0(t) = P_0(0) e^{\frac{t}{n}} \xrightarrow{t \rightarrow \infty} P_0(\infty) = +\infty \end{cases}$$

- Flot de Calabi sur les coefficients de réflexion :

$$\begin{cases} \frac{\partial \phi}{\partial t} = R(\phi) - \bar{R} \\ \phi(0, t) = \psi(t) \end{cases}$$

$$\begin{cases} |\mu_i(t)|^2 = 1 - \left(1 - |\mu_i(0)|^2\right) e^{-\frac{2t}{(n-i)^2}} \xrightarrow{t \rightarrow \infty} |\mu_i(\infty)|^2 = 1 \\ P_0(t) = P_0(0) e^{\frac{2t}{n^2}} \xrightarrow{t \rightarrow \infty} P_0(\infty) = +\infty \end{cases}$$

Algorithme médian sur les coefficients de réflexions 1/2

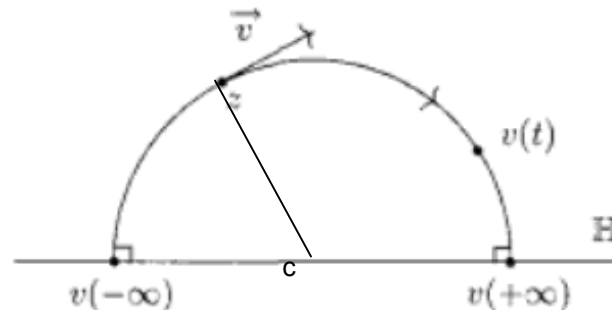


- On se donne $(\theta_1, \dots, \theta_N)$, où $\theta_k = (P_0^{(k)}, \mu_1^{(k)}, \dots, \mu_{n-1}^{(k)}) = (P_0^{(k)}, \mu^{(k)}) \in R_+ \times D^{n-1}$ est un n uplet des coefficients de réflexion $\mu^{(k)} = (\mu_1^{(k)}, \dots, \mu_{n-1}^{(k)})$ à tous les ordres précédés du coefficient de puissance $P_0^{(k)}$.
- On calcule d'abord $P_0^{median} = \exp(\text{mediane}(\log(P_0^{(1)}), \log(P_0^{(2)}), \dots, \log(P_0^{(N)})))$, dans R (méthode classique de calcul du médian pour une valeur scalaire).
- On estime ensuite la médiane des coefficients de réflexions complexes algorithmiquement : on ramène ces coefficients du disque unité dans le demi-plan de Poincaré $C^{-1} : D \rightarrow H$ avec :

$$\mu_j^{(k)} \mapsto z_j^{(k)} = i \left(\frac{1 + \mu_j^{(k)}}{1 - \mu_j^{(k)}} \right)$$

On note $z^{(k)} = C^{-1}(\mu^{(k)})$

- Les géodésiques sont alors des demi-cercles centrés sur l'axe des abscisses



Algorithme médian sur les CR

- On initialise l'algorithme à un point quelconque Z dans H :

$$Z^{(0)} = [Z_1^{(0)} \dots Z_{n-1}^{(0)}]^T$$

- A l'étape p , on se déplace selon le gradient:

$$t_{Z,l}^{(p)} = \sum_{k=1}^N i \times \text{signe}(Z_l^{(p)} - z_l^{(k)}) \frac{Z_l^{(p)} - c_l^{(k)}}{|Z_l^{(p)} - c_l^{(k)}| + 1/p} \quad \text{avec} \quad c_l^{(k)} = \frac{|z_l^{(k)}|^2 - |Z_l^{(p)}|^2}{2\text{Re}(z_l^{(k)} - Z_l^{(p)})}$$

où Z est l'estimé courant et c_k désigne le centre de la géodésique rejoignant Z à z_k .

$$c_{Z_l^{(p)}} = \text{Re}(Z_l^{(p)}) + \frac{\text{Im}(t_{Z,l}^{(p)})}{\text{Re}(t_{Z,l}^{(p)})} \text{Im}(Z_l^{(p)})$$

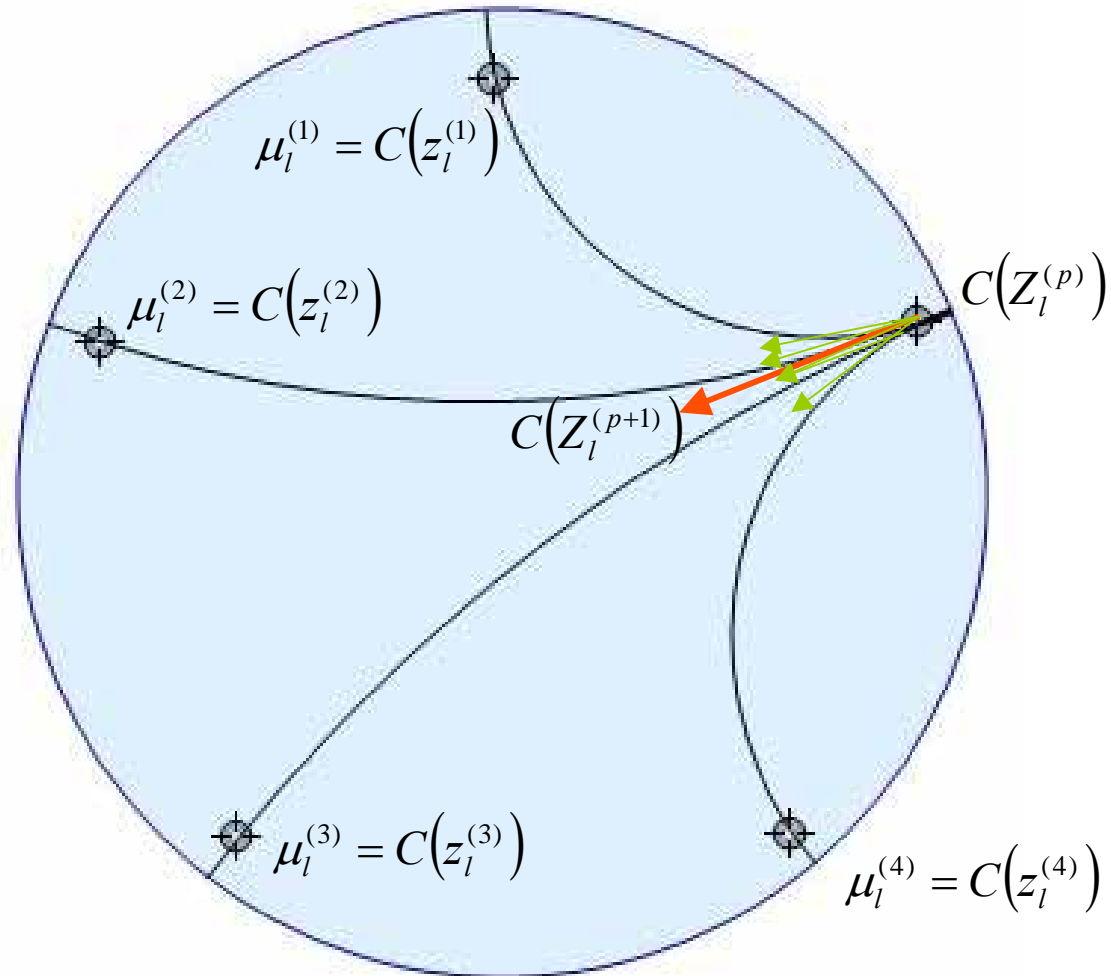
$$\phi_l^{(p)} = \arg(Z_l^{(p)} - c_{Z_l^{(p)}})$$

$$Z_l^{(p+1)} = c_{Z_l^{(p)}} + |Z_l^{(p)} - c_{Z_l^{(p)}}| \exp \left[i \left(\phi_l^{(p)} - \frac{1}{2Mk} \text{signe}(\text{Re}(Z_l^{(p)} - c_{Z_l^{(p)}})) \frac{Z_l^{(p)} - c_{Z_l^{(p)}}}{|Z_l^{(p)} - c_{Z_l^{(p)}}| + 1/p} \right) \right]$$

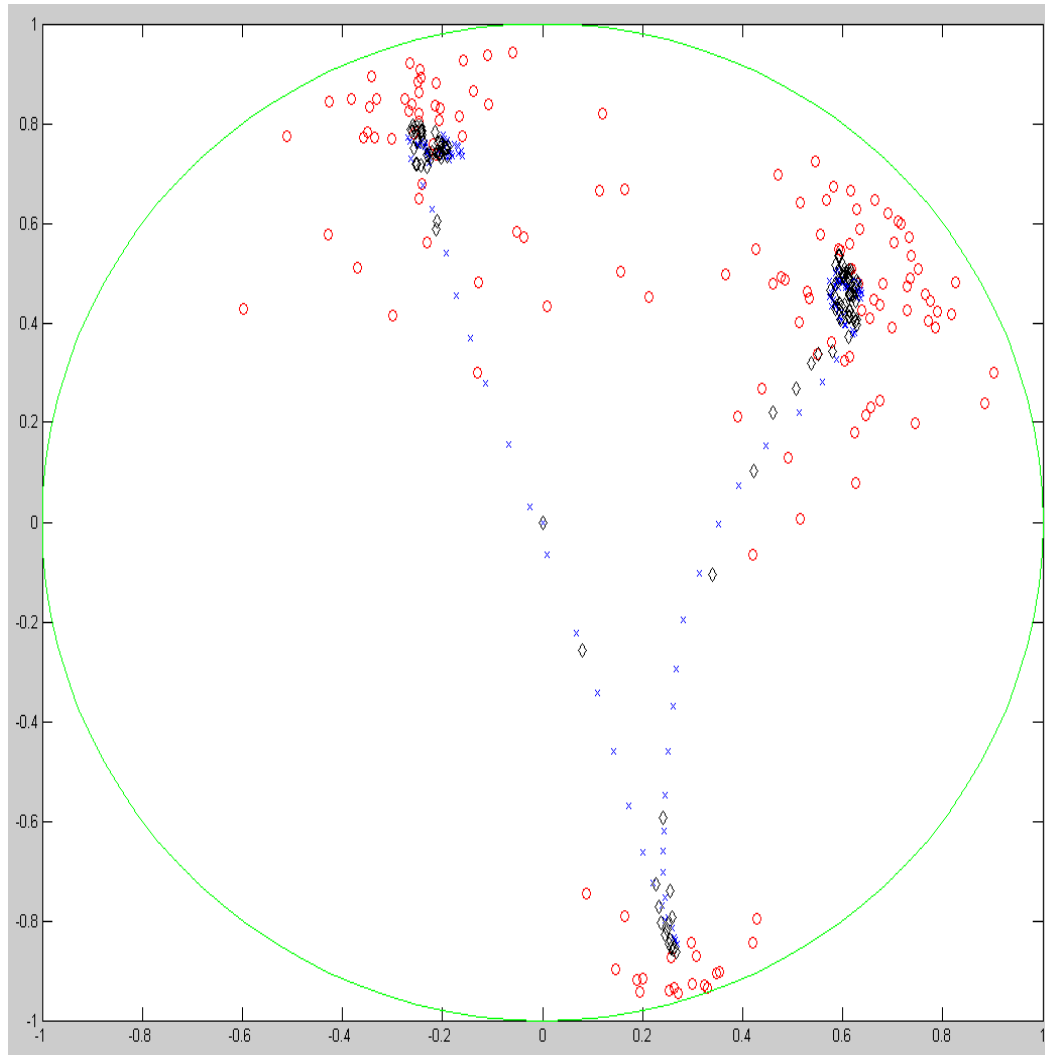
- L'entier p , initialisé à 1, décroît si l'écart entre deux itérés est moindre que $1/p$.
- Après convergence, on retourne à :

$$\Theta = (P, C(Z)) \quad \text{avec} \quad \begin{cases} P_0^{median} = \exp(\text{mediane}(\log(P_0^{(1)}), \log(P_0^{(2)}), \dots, \log(P_0^{(N)}))) \\ \mu_l^{median} = C(Z_l^{(conv)}) = \frac{Z_l^{(conv)} - i}{Z_l^{(conv)} + i} \end{cases}$$

Flot de gradient dans le disque unité de Poincaré



Résultats sur le premier coefficient de réflexion



Superposition pour toutes les cases distances des moyennes et médianes locale, avec les points de données

THALES



Flot de la chaleur & Géométrie de l'information

■ L'équation de la chaleur fait intervenir une moyenne :

■ Opérateur de Laplace intervenant dans le Flot de Fourier de la chaleur:

$$\frac{\partial^2 \theta}{\partial x^2} = \lim_{\nabla x \rightarrow 0} \frac{[\theta(x + \nabla x) - \theta(x)] - [\theta(x) - \theta(x - \nabla x)]}{(\nabla x)^2}$$

$$\frac{\partial^2 \theta}{\partial x^2} \approx \frac{2}{(\nabla x)^2} [\hat{\theta}(x) - \theta(x)] \quad \text{avec} \quad \hat{\theta}(x) = \frac{\theta(x + \nabla x) + \theta(x - \nabla x)}{2}$$

■ Ceci reste vrai avec le Laplacien en dimension supérieure :

$$\Delta \theta = \operatorname{div}(\nabla \theta) = \frac{\partial^2 \theta}{\partial x_1^2} + \dots + \frac{\partial^2 \theta}{\partial x_n^2} \approx \frac{2^n}{(\nabla x)^2} [\hat{\theta}(x) - \theta(x)]$$

$$\text{avec} \quad \hat{\theta}(x) = \frac{\theta(x + \nabla x_1) + \theta(x - \nabla x_1) + \dots + \theta(x + \nabla x_n) + \theta(x - \nabla x_n)}{2^n}$$

■ L'équation de la chaleur fait intervenir une moyenne :

- Opérateur de Laplace-Beltrami intervenant dans le Flot de Beltrami sur une variété :

$$\Delta_g \theta = \sum_{i,j} \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x_i} \left(\sqrt{|g|} g^{ij} \frac{\partial \theta}{\partial x_j} \right)$$

$$\Delta_g \theta = \sum_{\lambda,\mu} g^{\lambda\mu} \partial_{\lambda\mu}^2 \theta = \frac{2^{n^2} \sum_{\lambda,\mu} g^{\lambda\mu}}{(\nabla x)^2} [\hat{\theta}(x) - \theta(x)]$$

$$\text{avec } \hat{\theta}(x) = \sum_{\lambda,\mu} \alpha_{\lambda\mu} \left(\frac{\theta(x + \nabla x_\lambda) + \theta(x - \nabla x_\mu)}{2} \right)$$

$$\text{où } \alpha_{\lambda\mu} = \frac{g^{\lambda\mu}}{\sum_{\lambda,\mu} g^{\lambda\mu}} \quad \text{et } \nabla x = \nabla x_1 = \nabla x_2 = \dots = \nabla x_n$$

- L'équation de la chaleur se met donc sous la forme générale :

$$\frac{\partial \theta}{\partial t} = \Delta \theta = \rho [\hat{\theta}(x) - \theta(x)]$$

Équation de la chaleur et Géométrie de l'information

Equation de la chaleur

$$\frac{\partial \theta}{\partial t} = \Delta \theta = \rho [\hat{\theta}(x) - \theta(x)]$$

$$\frac{d\theta}{dt} \frac{d\theta^+}{dt} = \rho^2 (\theta - \hat{\theta})(\theta - \hat{\theta})^+$$

$$R_\theta \geq I(\theta)^{-1}$$

$$E[d\theta \cdot d\theta^+] = \rho^2 R_\theta dt^2 \geq \rho^2 I(\theta)^{-1} dt^2$$

$$E[d(\theta - \hat{\theta})d(\theta - \hat{\theta})^+] = E(d\theta \cdot d\theta^+) - d\hat{\theta} \cdot d\hat{\theta}^+$$

$$\text{Tr} \left[I(\theta) \cdot E \left[d(\theta - \hat{\theta})^+ d(\theta - \hat{\theta}) \right] \right] + d\hat{\theta}^+ \cdot I(\theta) \cdot d\hat{\theta} \geq n \rho^2 dt^2$$

$$\xrightarrow{\hat{\theta} \mapsto \theta} d\theta^+ \cdot I(\theta) \cdot d\theta = ds_I^2 = n \rho^2 dt^2$$

$$\frac{d\theta^+}{dt} \cdot I(\theta) \cdot \frac{d\theta}{dt} = n \rho^2$$

■ Équation de la chaleur et géodésique :

- La géodésique est donnée par :

$$\frac{\partial \theta}{\partial t} = \Delta \theta = \rho [\hat{\theta}(x) - \theta(x)] \Rightarrow \frac{d\theta}{dt} \cdot I(\theta) \cdot \frac{d\theta}{dt} = n\rho^2$$

- Pour une courbe géodésique $\theta = \theta(t)$ son vecteur tangent $\dot{\theta} = \frac{d\theta}{dt}(t)$ est de longueur constante par rapport à la métrique ds_I , soit :

$$\sum_{i,j} g_{ij} \frac{d\theta_i}{dt} \frac{d\theta_j}{dt} = n\rho^2$$

- La constante peut être choisie pour être de valeur unité quand le paramètre t est choisi comme étant le paramètre d'arc curviligne « s ».

THALES



Conclusion

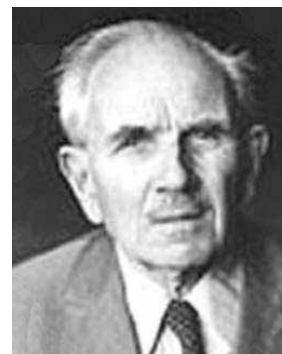
Références Bibliographiques

- [1] F. Barbaresco & G. Bouyt, "Espace Riemannien symétrique et géométrie des espaces de matrices de covariance : équations de diffusion et calculs de médianes", colloque GRETSI'09, Dijon, Sept. 2009
- [2] F. Barbaresco, "New Foundation of Radar Doppler and Array Processing based on German Mathematical Works: Geometry of Metric Spaces with negative curvature and Von-Mangoldt-Cartan-Hadamard Manifolds", IRS'09, Hamburg, Sept. 2009
- [3] F. Barbaresco, "Interactions between Symmetric Cone and Information Geometries", LIX Fall Colloquium ETCV'08, Ecole Polytechnique, Springer Lecture Notes in Computer Science, vol.5416, pp. 124-163, 2009, <http://www.springerlink.com/content/978-3-642-00825-2>
- [4] F. Barbaresco, "Applications of Information Geometry to Radar Signal Processing", Video Lectures, http://videolectures.net/etvc08_barbaresco_aoigt/
- [5] M. Arnaudon and X. Li, « Barycentres of measures transported by stochastic flows », Ann. Probab. 33, no. 4, pp.1509-1543, 2005
- [6] C.L. Siegel, «Symplectic Geometry», Acad. Press, 1964
- [7] M. Émery, G. Mokobodzki, «Sur le barycentre d'une probabilité dans une variété » . Séminaire de probabilités de Strasbourg, n° 25, p. 220-233, 1991
- [8] H. Karcher, «Riemannian center of mass and mollifier smoothing», Comm. Pure Applied Math., °30,pp.509-541,1977
- [9] A. Terras, «Harmonic Analysis on Symmetric Spaces and Applications II», Springer-Verlag, 1988
- [10] V. Arsigny «Log-Euclidean metrics for fast and simple calculus on diffusion tensors», in Magnetic Resonance in Medicine, Volume 56 Issue 2, Pages 411 - 421, Jun 2006.
- [11] M. Calvo, J. Oller, «A distance between multivariate normal distributions based in an embedding into the Siegel Group», Journal of Multivariate Analysis, vol.35, pp. 223-242, 1990
- [12] M. Calvo, J. Oller, «An explicit solution of information geodesic equations for the multivariate normal model», Statistics and Decisions., vol.9, pp. 119-138, 1991
- [13] Huiling Le, "Estimation of Riemannian Barycenters", Proc. London Math. Soc., pp. 193-200, 2004
- [14] Harris, W. F., "The average eye", Ophthalmic Physiol. Opt., n°24, pp. 580-585, 2004
- [15] H.C.F. von Mangoldt, "Über diejenigen Punkte auf positiv gekrümmten flächen, welche die eigenschaft haben, dass die von ihnen ausgehenden geodätischen Linien nie aufhören, kürzeste Linien zu sein", J. Reine Angew.Math., n°91, pp.23-52, 1881
- [16] E. Kähler, "Über eine bemerkenswerte Hermitesche Metrik", Abh., Math. Sem. Hamburg Univ., n°9, pp.173-186, 1933
- [17] K. T. Sturm, "Probability measures on metric spaces of nonpositive curvature". Contemp. Math. 338 , 357-390, 2003
- [18] E. Cartan, "Sur les domaines bornés homogènes de l'espace de n variables complexes", Abh. Math. Semin. hamb. Univ., n°11, pp.116-162, 1935, www.numdam.org
- [19] C.R. Rao, "Information and Accuracy attainable in the estimation of statistical parameters", Bull. Calcutta Math. Soc., n°37, pp.81-91, 1945
- [20] S. Yoshizawa and K. Tanabe, "Dual Differential Geometry associated with the Kullback-Leibler Information on the Gaussian Distributions and its 2-parameter Deformations", SUT Journal of Mathematics, vol.35,n°1, pp.113-137, 1999
- [21] T.Ando and al., "Geometric Means", Linear Algebra Appl., vol.385, pp.305-334, 2004
- [22] F. Bruhat and J. Tits, « Groupes réductifs sur un corps local », IHES, n°41, pp.5-251, 1972
- [23] H. Cartan, "Ouverts fondamentaux pour le groupe modulaire", Séminaire Henri Cartan, tome n°10, n°1, exp.n°3, p.1-12, 1957
- [24] N.N. Chentsov, "Statistical Decision Rules and Optimal Inferences", Trans. of Math. Monog., n°53, Amer. Math. Society, Providence, 1982
- [25] J. Faraut and A. Koranyi, "Analysis on Symmetric Cones", Oxford University Press, 1994
- [26] K. Koufany, "Analyse et Géométrie des domaines bornés symétriques", HDR, Institut de Mathématiques Elie Cartan, Nancy, Nov. 2006
- [27] C. P. Niculescu, "An Extension of the Mazur-Ulam Theorem", American Institute of Physics Proc., vol. 729, n°248-256, 2004
- [28] M. Berger, « Panoramic View of Riemannian Geometry », Springer, 2004



**Carl Ludwig Siegel
avec Georges Polya**

■ *“ Constaté que les théories les plus parfaites sont les guides les plus sûrs pour résoudre les problèmes concrets~; avoir assez confiance en sa science pour prendre des responsabilités techniques. Puissent beaucoup de mathématiciens connaître un jour ces joies très saines, quelques humbles qu'ils les jugent !”*



■ **Jean LERAY**