

Tomographie (MHT921), Partie I, Automne 2009

Alain Yger

10 mars 2011

Table des matières

1	Introduction	1
2	Les modèles mathématiques des mécanismes tomographiques	3
2.1	La transformation de Radon	3
2.1.1	La définition	3
2.1.2	La formule « des tranches »	5
2.1.3	Une formulation (un peu) différente : les formules d'inversion	7
2.1.4	La formule de rétro-projection revisitée : le choix du filtre	9
2.1.5	La transformation de Radon à l'épreuve de la méthode de Kaczmarz	11
2.1.6	La transformation de Radon à l'épreuve de la SVD	14
2.2	S.P.E.C.T et Transformation de Radon atténuée	23
2.2.1	Le problème physique : présentation succincte de la S.P.E.C.T	23
2.2.2	Radon atténuée à l'épreuve de la méthode ART	26
2.2.3	Le mécanisme d'inversion analytique	28
3	Processus stochastiques ; filtrage et déconvolution	31
3.1	Une version réaliste pour les données de la SPECT : le modèle stochastique	31
3.1.1	Processus Poissonniens $1D$	31
3.1.2	Notion de processus Poissonnien $2D$	32
3.1.3	Le modèle stochastique de la SPECT	33
3.2	Processus stationnaires : le filtrage de Wiener	33
3.2.1	Retour sur la notion de processus ; stationnarité	33
3.2.2	Déconvolution linéaire ; « pseudo-filtrage » de Wiener	40
3.2.3	Une méthode ancienne : la gammagraphie par ouverture de codage	44
3.3	Algorithmes de « Matching Pursuit » ; P.O.D	46
3.3.1	Dictionnaires et « <i>Matching Pursuit</i> »	47
3.3.2	La décomposition orthogonale propre	50
3.3.3	P.O.D et non stationnarité	51
4	Une esquisse de présentation des outils mathématiques de l'IRM	53
4.1	Introduction	53
4.2	Champ statique et champ Radio-Fréquence	53
4.3	La phase de « précession »	55
4.4	La phase de « relaxation »	58

Chapitre 1

Introduction

Au travers de la présentation de diverses méthodes classiques en tomographie médicale en relation avec des transformations mathématiques (tomodensitométrie¹, tomographie d'émission photonique², tomographie par émission de positrons³, Gammapgraphie par ouverture de codage, échographie) ou avec des mécanismes physiques relevant de l'électromagnétisme (IRMN⁴), nous présenterons dans ce cours tant les outils mathématiques soutendant l'instrumentation elle-même (transformations du type Radon ou *X-Ray*, analyse de Fourier, déconvolution) que l'analyse et le débruitage des images produites.

Nous ferons tout d'abord un tour d'horizon des transformations mathématiques sous-jacentes, à commencer par les transformées intégrales que sont la transformation de Radon et la *X-Ray* transform, ainsi que les transformations correspondant à la convolution avec un masque se présentant soit comme une fonction caractéristique de code, soit comme le spectre d'une telle fonction caractéristique. Chaque fois, nous situerons ces transformations en relation avec leur application pratique à l'un des procédés tomographiques cités plus haut.

Nous tenterons également une approche de présentation schématique de l'imagerie par résonance magnétique nucléaire (IRMN), pour laquelle un modèle mathématique rigoureux est loin d'être à ce jour encore compris ni même acquis (une bonne dose d'empirisme subsistant encore ici).

Sera envisagée plus tard, en relation avec des modèles empruntés soit à l'imagerie médicale elle-même, soit à la sismique, le traitement des images *via* des méthodes relevant de l'analyse hilbertienne (*matching pursuit*, *Proper Orthogonal Decomposition*), de l'analyse harmonique (ondelettes, filtrage de Wiener ou pseudo-Wiener) ou

¹Ou encore CAT-*scanner*, pour *Computerized Axial Tomography*, ou simplement *scanner* : c'est le procédé le plus connu, où le patient est soumis à un balayage de rayons X émanant d'un dispositif pivotant autour du corps, doublé d'un système de réception enregistrant l'intensité des rayons en sortie (après atténuation au travers des tissus traversés).

²*Single Photon Emission Computed Tomography*, ou encore en abrégé SPECT : le radiotracer absorbé par l'organe émet un rayonnement photonique (modélisable, on le verra, par un processus poissonnien), rayonnement capté par la caméra (un ensemble de collimateurs pivotant autour du corps).

³*Positron Emission Tomography*, ou encore en abrégé PET : le radiotracer émet des positrons (ou positons, c'est la même chose), qui suite à leur collision avec des électrons de la matière du tissu, génèrent deux photons en coïncidence.

⁴Imagerie par Résonance Magnétique Nucléaire : un signal est émis lors de la démagnétisation des photons préalablement soumis à un champ magnétique, et l'onde ainsi produite est enregistrée.

de l'analyse non-linéaire (minimisation L^1 , méthodes parcimonieuses,...). Cette seconde partie du cours (traitée par J.F. Giovannelli) sera dévolue plus généralement aux méthodes d'obédience déterministe ou statistique utilisées face à de tels problèmes dits "problèmes inverses" en mathématiques.

Le logiciel MATLAB et ses deux *toolboxes* *Signal Processing* et surtout *Image Processing* seront pour nous un outil d'illustration constant. Les TP sous ce logiciel illustreront le cours et les routines émaillant ces notes (élaborées sur des modèles test en dimension deux⁵) ont été implémentées sous ce logiciel.

Une référence bibliographique actualisée présentant un panorama des méthodes décrites dans la première partie de ce cours est par exemple l'ouvrage de Charles L. Epstein, *Introduction to the Mathematics of Medical Imaging*, Second edition, Siambooks, 2007⁶. Nous utiliserons aussi, pour les aspects plus proprement mathématiques, l'ouvrage de F. Natterer, *The mathematics of computerized tomography*, Teubner-Wiley, New York, 1986. Pour les ouvrages relevant plus généralement de l'analyse des signaux et des images, on renvoie à la liste des ouvrages de référence proposée sur le site du master⁷, en particulier l'ouvrage de Gonzalez, Woods et Ed-dins, *Digital Image Processing using MATLAB*. Nous donnerons d'autres références bibliographiques plus spécifiques au fil de ces notes.

Concernant les références bibliographiques en relation avec la seconde partie du cours, se reporter à la fiche actualisée de l'UE MHT921, en ligne sur le site. On y trouvera une présentation du contenu de l'UE sous sa nouvelle forme.



FIG. 1.1 – Le dispositif de CAT-scanner

⁵Evidemment, la réalité pratique voudrait que l'on travaille en trois dimensions; cependant il arrive fréquemment (comme c'est le cas en tomographie axiale) que l'on se ramène à travailler sur des coupes 2D du domaine 3D que l'on prétend explorer.

⁶Voir le site <http://www.ec-securehost.com/SIAM/OT102.html>

⁷<http://www.math.u-bordeaux1.fr/~yger/referencesbase.pdf>

Chapitre 2

Les modèles mathématiques des mécanismes tomographiques

2.1 La transformation de Radon

2.1.1 La définition

Notre espace de référence sera \mathbb{R}^n , mais nos modèles seront le plus souvent \mathbb{R}^2 et \mathbb{R}^3 . L'ensemble des hyperplans affines de \mathbb{R}^n (les droites en dimension 2, les plans en dimension 3) est en bijection avec le produit cartésien de la sphère unité \mathbb{S}^{n-1} avec \mathbb{R} , à condition d'identifier le couple $(\vec{\theta}, s)$ avec le couple $(-\vec{\theta}, -s)$: un hyperplan Π de \mathbb{R}^n est en effet repéré par le vecteur $\pm\vec{\theta}$ dirigeant Π^\perp et l'abscisse algébrique $\pm p$ du pied de la perpendiculaire à ce plan menée depuis l'origine, cette normale à Π^\perp étant supposée orientée par le vecteur $\vec{\theta}$, les couples $(\vec{\theta}, p)$ et $(-\vec{\theta}, -s)$ repérant bien sûr le même hyperplan.

Définition 2.1 *Si f est une fonction source définie dans \mathbb{R}^n , mesurable, à valeurs complexes, intégrable sur chaque hyperplan (contre la mesure de Lebesgue $(n-1)$ -dimensionnelle sur cet hyperplan), la transformée de Radon¹ de f (on dit aussi la transformée de Radon par les hyperplans) est par définition l'application de $\mathbb{S}^{n-1} \times \mathbb{R}$ qui à $(\vec{\theta}, s)$ associe*

$$R[f](\vec{\theta}, s) := \int \dots \int_{y \in \vec{\theta}^\perp} f(s\vec{\theta} + y) dm_{\vec{\theta}^\perp}(y) \quad (2.1)$$

(c'est-à-dire l'intégrale de la fonction source le long de la « coupe » correspondant à l'hyperplan repéré par $(\vec{\theta}, p)$ ou $(-\vec{\theta}, -p)$).

Remarque 2.1. En dimension 2, on peut se contenter de balayer avec θ (si $\vec{\theta} := (\cos \theta, \sin \theta)$) le secteur angulaire $[0, \pi[$ puisque l'on sait que $R[f](\vec{\theta}, s) = R[f](\vec{\theta}, -s)$; on obtiendra bien en se contentant du balayage $(\theta, s) \in [0, \pi[\times \mathbb{R}$ toute l'information concernant la transformée de Radon de f .

¹Pour cette section, ma référence première pour rédiger ces notes est l'excellent livre de Frank Natterer, *The mathematics of computerized tomography*, Teubner-Wiley, New York, 1986. Cette transformation (et les formules d'inversion qui l'accompagnent) sont apparues dans les travaux du mathématicien autrichien Johannes Radon dès 1917; on trouvera d'ailleurs le texte original de Radon en appendice dans l'ouvrage de F. Natterer.

Exemple 2.1. En dimension 2 (ce sera notre dimension « jouet »), la transformée de Radon de f est donc, suivant la remarque précédente, la fonction

$$(\vec{\theta}, s) \longmapsto \int_{\mathbb{R}} f(s \cos \theta - t \sin \theta, s \sin \theta + t \cos \theta) dt.$$

En dimension deux, si la source f est à support borné dans $[-R, R]^2$ et se présente sous une forme discrétisée (donc comme une source modélisée par exemple par une brillance sur une grille de N pixels sur N pixels), il n'est nullement évident de calculer la transformée de Radon car, étant donnée une ligne intersectant la grille de pixels, on est confronté au délicat problème de « recenser » les pixels suffisamment voisins de la ligne pour être supposés du point de vue de la discrétisation du problème en faire partie. Sous MATLAB, la routine

```
>> frot=imrotate(f,-theta,'crop');
```

permettant de faire pivoter la grille de pixels et l'image source f qu'elle porte de l'angle $-\theta$ est une routine certes intéressante, mais à manier avec précaution, comme cela sera vu en TP. Un exemple-test bien classique et que nous croiserons dans ce cours est le fantôme f de Shepp-Logan que nous avons figuré sur la figure de gauche dans la figure 2.1 ci dessous, tandis qu'à droite, nous avons affiché le module de l'image de sa transformée de Radon (cette seconde image est habituellement appelée le *sinogramme* de f). Sur la figure 2.2 plus loin, l'image par la transformation

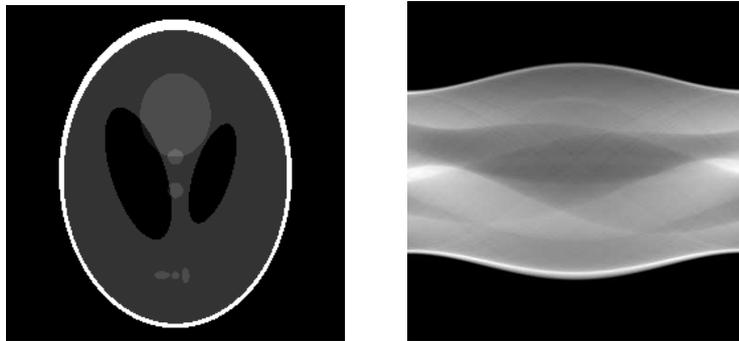


FIG. 2.1 – Le fantôme de Shepp-Logan (à gauche) et son sinogramme (à droite)

de Radon d'une combinaison de sources assez bien localisées fait apparaître sur le sinogramme des tracés sinusoidaux (déphasés les uns par rapport aux autres), ce que l'on retrouve naturellement si l'on suppose (cas extrême) que $f = \delta_{x_0, y_0}$ et que l'on élimine le paramètre t entre les deux relations

$$\begin{aligned} s \cos \theta - t \sin \theta &= x_0 \\ s \sin \theta + t \cos \theta &= y_0, \end{aligned}$$

ce qui fournit

$$s = x_0 \cos \theta + y_0 \sin \theta.$$

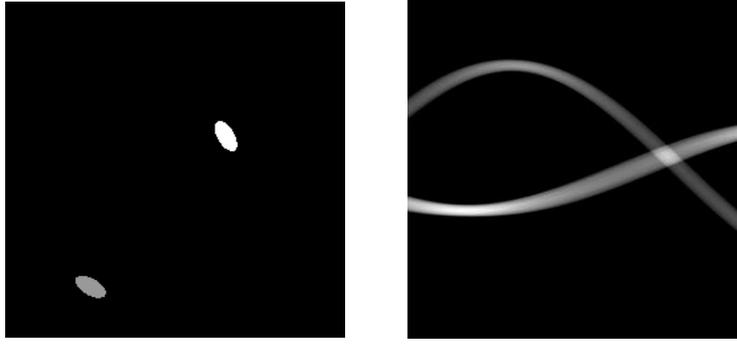


FIG. 2.2 – Deux sources localisées (à gauche) et le sinogramme (à droite)

2.1.2 La formule « des tranches »

Si f est une source supposée continue et à support borné, le théorème de Fubini assure que, pour tout $\omega \in \mathbb{R}$, pour $\vec{\theta} \in \mathbb{S}^{n-1}$,

$$\begin{aligned} \int_{\mathbb{R}} e^{-i\omega s} R[f](\vec{\theta}, s) ds &= \int_{\mathbb{R}} e^{-i\omega s} \left(\int \dots \int_{\vec{\theta}^\perp} f(s\vec{\theta} + y) dm_{\vec{\theta}^\perp}(y) \right) \\ &= \int \dots \int_{\mathbb{R}^n} f(u) e^{-i\omega \langle \vec{\theta}, u \rangle} du_1 \dots du_n = \widehat{f}(\omega \vec{\theta}). \end{aligned} \quad (2.2)$$

La formule (2.2), dite « des tranches » permet, lorsque la transformée de Radon est connue complètement (pour tout $\vec{\theta}$ et pour tout $s \in \mathbb{R}$) d'en déduire, pour chaque $\vec{\theta} \in \mathbb{R}$, le signal

$$\omega \in \mathbb{R} \longmapsto \widehat{f}(\omega \vec{\theta}).$$

Or, la formule d'inversion de Fourier assure que si f est dans $L^2(\mathbb{R}^n)$, \widehat{f} l'est aussi et que l'on a

$$\begin{aligned} f &:= \frac{1}{(2\pi)^n} \lim_{R \rightarrow +\infty} \int \dots \int_{\|x\| < R} \widehat{f}(\xi) e^{i\langle \xi, (\cdot) \rangle} d\xi_1 \dots d\xi_n \\ &= \frac{1}{(2\pi)^n} \lim_{R \rightarrow +\infty} \int_0^R \left(\int \dots \int_{\mathbb{S}^{n-1}} \widehat{f}(r\vec{\theta}) e^{ir\langle \vec{\theta}, (\cdot) \rangle} d\sigma_{\mathbb{S}^{n-1}}(\vec{\theta}) \right) r^{n-1} dr, \end{aligned} \quad (2.3)$$

la limite étant prise dans $L^2(\mathbb{R}^n, dx)$.

L'écueil auquel se heurte l'utilisation de ce mécanisme d'inversion tient au principe d'incertitude inhérent à l'analyse de Fourier : plus on localise une particule, plus son spectre se trouve diffus (les gaussiennes réalisant le meilleur compromis). Pour restituer f (et en particulier les « accidents » de f , ce qui souvent est ce qui nous intéresse), il est nécessaire de connaître le spectre de f partout, ou, ce qui revient au même, pour chaque $\vec{\theta} \in \mathbb{S}^{n-1}$, le spectre de

$$s \longmapsto R[f](\vec{\theta}, s)$$

pour tout $r > 0$, en particulier pour les r très grand, ce qui est impossible. Il est par conséquent nécessaire d'utiliser un filtre passe-bas (on dit aussi un « filtre rampe ») de réponse impulsionnelle $s \mapsto h(s)$ (de spectre \widehat{h} ayant son support localisé dans la

fenêtre fréquentielle $[-R_0, R_0]$), de remplacer, pour chaque $\vec{\theta} \in \mathbb{S}^{n-1}$, $s \mapsto R[f](\vec{\theta}, s)$ par sa version filtrée

$$s \mapsto [h * R[f](\vec{\theta}, \cdot)](s)$$

et de restituer enfin la source f *via* la formule dite de « rétro-projection filtrée » :

$$\begin{aligned} f &\simeq \frac{1}{(2\pi)^n} \int_0^{R_0} \left(\int \dots \int_{\mathbb{S}^{n-1}} \mathcal{F} \left[h * R[f](\vec{\theta}, \cdot) \right] (r) e^{ir\langle \vec{\theta}, \cdot \rangle} d\sigma_{\mathbb{S}^{n-1}} \right) r^{n-1} dr \\ &\simeq \frac{1}{(2\pi)^n} \int \dots \int_{\|\xi\| \leq R_0} \mathcal{F} \left[h * R[f](\xi/\|\xi\|, \cdot) \right] (\|\xi\|) e^{i\langle \xi, \cdot \rangle} d\xi_1 \dots d\xi_n \end{aligned} \quad (2.4)$$

(ici \mathcal{F} désigne la prise de transformée de Fourier).

Autre défaut de cette formule de rétro-projection, la présence du facteur r^{n-1} amplifiant les hautes fréquences, ce malgré l'utilisation du filtre rampe.

Cette formule ne laisse d'autre part envisager aucune possibilité de restitution locale de la source à partir de données partielles (par exemple la simple connaissance de $R[f]$ sur tous les hyperplans traversant une certaine zone pour espérer restituer la source dans la zone en question); le recours à la transformation de Fourier dans la formulation du mécanisme de rétroprojection nous prive de cette possibilité, si tant est qu'elle pourrait exister². On pourra se rendre compte de ces phénomènes en utilisant la routine MATLAB

```
>> f= iradon(R,theta);
```

basée sur l'implémentation de la méthode de rétro-projection filtrée. Nous pouvons comparer par exemple le fantôme de Shepp-Logan avec la source obtenue après bruitage significatif (ici le bruit additionnel est un bruit blanc gaussien) *via* le synopsis

```
>> f= phantom (256);
>> theta = 0:.5:180;
>> R = radon (f,theta);
>> RR = R + 10*randn(size(R));
>> fbruite = iradon(RR,theta);
```

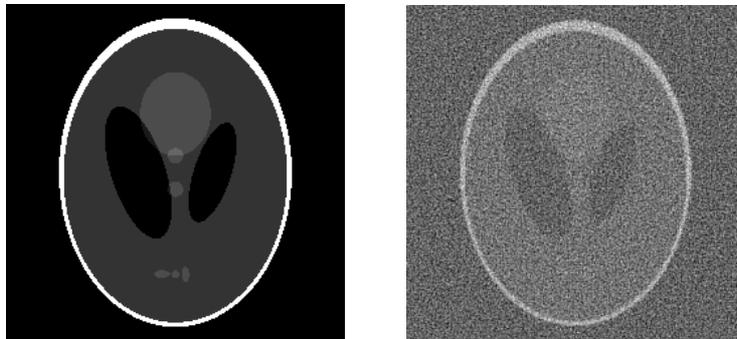


FIG. 2.3 – Le fantôme de Shepp-Logan (à gauche) et la version après bruitage en sortie, puis rétroprojection filtrée.

²C'est le cas, nous y reviendrons, en dimension impaire.

Il faut souligner aussi le délicat problème (de nature plus informatique celui là) inhérent au fait que si le domaine source est repéré en coordonnées cartésiennes, le spectre de la fonction source f se calcule certes à partir de la connaissance de la transformée de Radon $R[f]$ de f , mais dans un domaine fréquentiel où le repérage des points se fait cette fois en coordonnées polaires ! Des commandes telles que

```
>> frot imrotate (f,theta,'crop')
```

(ou des techniques de traitement d'image implémentées sous d'autres langages de programmation telles celles vues dans l'UE INF472) devront encore certainement venir à notre secours.

2.1.3 Une formulation (un peu) différente : les formules d'inversion

Avant de définir un mécanisme inversant la transformation de Radon, nous introduirons une transformation (notée R^\sharp) « duale » de celle opérée par Radon au sens suivant : pour obtenir la valeur ponctuelle de $R^\sharp[g]$ (où g est une fonction continue sur $\mathbb{S}^{n-1} \times \mathbb{R}$ telle que $g(\vec{\theta}, s) = g(-\vec{\theta}, -s)$), on intègre cette fonction g , pensée comme une fonction définie sur l'ensemble des hyperplans de \mathbb{R}^n , sur l'ensemble des hyperplans passant par le point x , ce qui donne

$$R^\sharp[g](x) := \int_{\mathbb{S}^{n-1}} g(\vec{\theta}, \langle x, \vec{\theta} \rangle) d\sigma_{\mathbb{S}^{n-1}}(\vec{\theta}). \quad (2.5)$$

La formule (2.3) s'écrit encore (formellement) sous la forme

$$\begin{aligned} f &:= \frac{1}{(2\pi)^n} \lim_{R \rightarrow +\infty} \int \dots \int_{\|x\| < R} \widehat{f}(\xi) e^{i\langle \xi, (\cdot) \rangle} d\xi_1 \dots d\xi_n \\ &= \frac{1}{(2\pi)^n} \lim_{R \rightarrow +\infty} \int_0^R \left(\int \dots \int_{\mathbb{S}^{n-1}} [R[f](\vec{\theta}, \cdot)](r) e^{ir\langle \vec{\theta}, (\cdot) \rangle} d\sigma_{\mathbb{S}^{n-1}}(\vec{\theta}) \right) r^{n-1} dr \\ &= \frac{1}{2} \frac{1}{(2\pi)^n} \lim_{R \rightarrow +\infty} \int_{-R}^R \left(\int \dots \int_{\mathbb{S}^{n-1}} [R[f](\vec{\theta}, \cdot)](r) e^{ir\langle \vec{\theta}, (\cdot) \rangle} d\sigma_{\mathbb{S}^{n-1}}(\vec{\theta}) \right) |r|^{n-1} dr \\ &= \frac{1}{2} \frac{1}{(2\pi)^{n-1}} R^\sharp[I^{1-n}[R[f](\vec{\theta}, \cdot)]], \end{aligned} \quad (2.6)$$

où I^α désigne, pour $\alpha < -p$, l'opérateur (dit de Riesz) qui à un signal $\Phi(\cdot)$ sur \mathbb{R}^p , associe le signal dont la transformée de Fourier est

$$\omega \mapsto |\omega|^{-\alpha} \widehat{f}(\omega)$$

(dans la formule ci-dessus, on utilise I^{1-n} dans le cas $p = 1$, car l'opérateur agit sur $s \mapsto R[f](\vec{\theta}, s)$, fonction d'une variable, ce pour chaque $\vec{\theta}$ fixé).

Si n est impair, $n - 1$ est pair et la formule (2.6) s'écrit simplement sous la forme

$$f = \frac{1}{2} \frac{(-1)^{\frac{n-1}{2}}}{(2\pi)^{n-1}} R^\sharp \left[\frac{\partial^{n-1}}{\partial s^{n-1}} [R(f)(\cdot, \cdot)] \right] \quad (2.7)$$

(on a convenu de souligner la variable par rapport à laquelle se fait la différentiation). Si n est pair, il faut introduire la transformée de Hilbert \mathbb{H} sur $L^2(\mathbb{R}, dt)$, qui est

par définition l'opérateur qui à un signal analogique S associe le signal $\mathbb{H}[S]$ dont le spectre est

$$\omega \mapsto -i \operatorname{signe}(\omega) \widehat{S}(\omega)$$

qui correspond à la convolution avec le noyau singulier $t \mapsto 1/t$, *i.e*

$$\mathbb{H}[S] = \frac{1}{\pi} \times \text{valeur principale} \left(\int_{\mathbb{R}} \frac{S(\tau)}{(\cdot) - \tau} d\tau \right).$$

Dans ce second cas (n pair, ce qui englobe en particulier le cas $n = 2$), la formule d'inversion devient

$$f = \frac{1}{2} \frac{(-1)^{\frac{n-2}{2}}}{(2\pi)^{n-1}} R^\sharp \left[\mathbb{H} \left[\frac{\partial^{n-1}}{\partial s^{n-1}} [R(f)(\cdot, \cdot)] \right] \right] \quad (2.8)$$

(même remarque que précédemment concernant le soulignement).

Remarque 2.2. On peut aussi remarquer que si n est impair, l'opérateur I^{1-n} agissant sur les signaux fonctions de n variables est en fait l'opérateur

$$I^{1-n} = (-1)^{\frac{n-1}{2}} \Delta^{\frac{n-1}{2}},$$

où

$$\Delta := \sum_{k=1}^n \frac{\partial^2}{\partial x_k^2}$$

est l'opérateur Laplacien. Dans ce cas, une autre manière de formuler la formule d'inversion consiste à écrire

$$f = \frac{1}{2} \frac{(-1)^{\frac{n-1}{2}}}{(2\pi)^{n-1}} \Delta^{\frac{n-1}{2}} [R^\sharp[R[f]]]. \quad (2.9)$$

Malheureusement, cette formule n'est valable que si n est impair et dans le cas n pair, il faut se contenter de la formule

$$f = I^{1-n} [R^\sharp[R[f]]], \quad (2.10)$$

où l'action de I^{1-n} consiste à multiplier le spectre par $\|\omega\|^{n-1}$, ce que l'on fait en filtrant au préalable $x \mapsto R^\sharp[R[f]](x)$ avec un filtre rampe. Il s'agit simplement ici une autre manière de voir la formule de rétro-projection filtrée. Nous en verrons encore une autre plus loin (section 2.1.4).

Remarque 2.3. Dans le cas $n = 2$ (et plus généralement dans le cas n pair), la formule (2.8) peut être écrite explicitement si l'on fait l'hypothèse que la source f est une fonction C^∞ à support compact ; dans ce cas en effet, il en est de même, pour chaque $\vec{\theta} \in \mathbb{S}^1$, pour la fonction $\vec{\theta} \mapsto R[f](\vec{\theta}, s)$ et la transformée de Hilbert, pour $\vec{\theta}$ fixé, de la fonction

$$s \mapsto \frac{\partial}{\partial s} (R[f](\vec{\theta}, s))$$

est le signal

$$\begin{aligned} s \mapsto & \left(\frac{1}{\pi} \lim_{\epsilon \rightarrow 0} \int_{|\sigma| \geq \epsilon} \frac{\partial}{\partial \sigma} (R[f](\vec{\theta}, \sigma)) \frac{d\sigma}{s - \sigma} \right) \\ & = -\frac{1}{2\pi} \int_{\mathbb{R}} \left(\frac{\partial}{\partial \sigma} (R[f](\vec{\theta}, s + \sigma)) - \frac{\partial}{\partial \sigma} (R[f](\vec{\theta}, s - \sigma)) \right) \frac{d\sigma}{\sigma}. \end{aligned}$$

La formule d'inversion (2.8) se lit alors dans le cas $n = 2$ (en substituant l'action de la transformation de Hilbert comme ci-dessus)

$$f(x, y) = -\frac{1}{\pi} \int_0^\infty \frac{d}{dt} \left[\frac{1}{2\pi} \int_0^{2\pi} R[f](\vec{\theta}, x \cos \theta + y \sin \theta + t) d\theta \right] \frac{dt}{t}. \quad (2.11)$$

C'est cette formule d'inversion que Radon prouva dans son article fondateur de 1917.

2.1.4 La formule de rétro-projection revisitée : le choix du filtre

Une formule majeure couplant le rôle de la transformation de Radon R et de sa transformation duale R^\sharp est la suivante :

Proposition 2.1 *Si f est une fonction source (supposée dans l'espace $\mathcal{S}(\mathbb{R}^n)$ des fonctions C^∞ à décroissance rapide ainsi que toutes leurs dérivées) et*

$$w : (\vec{\theta}, s) \in \mathbb{S}^{n-1} \times \mathbb{R} \mapsto w(\vec{\theta}, s)$$

une fonction (dite fonction « fenêtre » définie, de classe C^∞ , et de plus à décroissance rapide, sur le cylindre $\mathbb{S}^{n-1} \times \mathbb{R}$), on a la relation

$$R^\sharp[w] *_{\mathbb{R}^n} f = R^\sharp \left[(\vec{\theta}, s) \mapsto \left(w(\vec{\theta}, \cdot) *_{\mathbb{R}} R[f](\vec{\theta}, \cdot) \right) (s) \right]. \quad (2.12)$$

De plus, l'action de l'opérateur R^\sharp sur la fenêtre w se traduit par

$$\widehat{R^\sharp[w]}(\omega) = (2\pi)^{n-1} \|\omega\|^{1-n} \left(\mathcal{F} \left[w \left(\frac{\omega}{\|\omega\|}, \cdot \right) \right] (\|\omega\|) + \mathcal{F} \left[w \left(-\frac{\omega}{\|\omega\|}, \cdot \right) \right] (-\|\omega\|) \right). \quad (2.13)$$

Preuve. Il s'agit encore d'une application du théorème de Fubini (pour ce qui est de la formule (2.12)); en effet, le membre de gauche de la formule s'écrit

$$\begin{aligned} & \int_{\mathbb{S}^{n-1}} \int_{\mathbb{R}^n} w(\vec{\theta}, \langle x - y, \vec{\theta} \rangle) f(y) dy d\sigma_{\mathbb{S}^{n-1}}(\theta) \\ &= \int_{\mathbb{S}^{n-1}} \int_{\mathbb{R}} \int_{\vec{\theta}^\perp} g(\langle x, \vec{\theta} \rangle - s) f(s\vec{\theta} + z) dm_{\vec{\theta}^\perp}(z) dy d\sigma_{\mathbb{S}^{n-1}}(\theta) \\ &= \int_{\mathbb{S}^{n-1}} \left[\int_{\mathbb{R}} w(\vec{\theta}, \langle x, \vec{\theta} \rangle - s) R[f](\vec{\theta}, s) ds \right] d\sigma_{\mathbb{S}^{n-1}}(\theta), \end{aligned}$$

ce qui donne la formule voulue. La seconde formule (2.13) tient au fait que si φ est une fonction-test dans \mathbb{R}^n (disons C^∞ et à support compact), on a, d'après la formule des tranches et la formule d'inversion de Fourier,

$$\begin{aligned} \int_{\mathbb{R}^n} R^\sharp[w](x) \widehat{\varphi}(x) dx &= \int_{\mathbb{S}^{n-1}} \left[\int_{\mathbb{R}} w(\vec{\theta}, s) R[\widehat{\varphi}](\vec{\theta}, s) ds \right] d\sigma_{\mathbb{S}^{n-1}}(\vec{\theta}) \\ &= \int_{\mathbb{S}^{n-1}} \left[\frac{1}{2\pi} \int_{\mathbb{R}} \mathcal{F}[w(\vec{\theta}, \cdot)](s) \overline{\mathcal{F}[R[\widehat{\varphi}](\vec{\theta}, \cdot)]}(s) ds \right] d\sigma_{\mathbb{S}^{n-1}}(\vec{\theta}) \\ &= (2\pi)^{n-1} \int_{\mathbb{S}^{n-1}} \left[\int_{\mathbb{R}} \mathcal{F}[w(\vec{\theta}, \cdot)](s) \varphi(s\vec{\theta}) ds \right] d\sigma_{\mathbb{S}^{n-1}}(\vec{\theta}) \\ &= (2\pi)^{n-1} \left[\int_{\mathbb{S}^{n-1}} \int_0^\infty \mathcal{F}[w(\vec{\theta}, \cdot)](r) \varphi(r\vec{\theta}) dr d\sigma_{\mathbb{S}^{n-1}}(\vec{\theta}) \right. \\ &\quad \left. - \int_{\mathbb{S}^{n-1}} \int_0^\infty \mathcal{F}[w(\vec{\theta}, \cdot)](r) \varphi(-r\vec{\theta}) dr d\sigma_{\mathbb{S}^{n-1}}(\vec{\theta}) \right] \\ &= (2\pi)^{n-1} \left[\int_{\mathbb{S}^{n-1}} \int_0^\infty \mathcal{F}[w(\vec{\theta}, \cdot)](r) \varphi(r\vec{\theta}) dr d\sigma_{\mathbb{S}^{n-1}}(\vec{\theta}) \right. \\ &\quad \left. - \int_{\mathbb{S}^{n-1}} \int_0^\infty \mathcal{F}[w(-\vec{\theta}, \cdot)](-r) \varphi(r\vec{\theta}) dr d\sigma_{\mathbb{S}^{n-1}}(\vec{\theta}) \right] \\ &= (2\pi)^{n-1} \int_{\mathbb{R}^n} \|\omega\|^{1-n} G_w(\omega) \varphi(\omega) d\omega, \end{aligned}$$

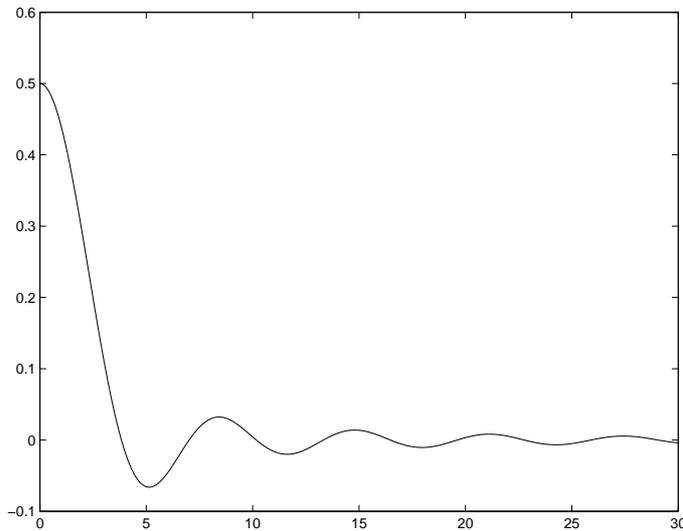


FIG. 2.4 – Le graphe de $\xi \mapsto \frac{J_1(\xi)}{\xi}$ sur $[0, 30]$

avec

$$G_w(\omega) = \mathcal{F}\left[w\left(\frac{\omega}{\|\omega\|}, \cdot\right)\right](\|\omega\|) + \mathcal{F}\left[w\left(-\frac{\omega}{\|\omega\|}, \cdot\right)\right](\|\omega\|),$$

l'égalité entre les deux dernières lignes s'obtenant en exprimant l'intégrale figurant à la dernière ligne en coordonnées sphériques $\omega = r\vec{\theta}$. Comme la fonction test φ est arbitraire, on a bien l'identité (2.13) voulue. \diamond

Le choix naturel de fonction « fenêtre » w s'imposant dans le mécanisme de rétro-projection filtrée revue sous l'angle de la formule (2.12) est celui pour lequel $R^\sharp[w]$ a un spectre approchant au mieux celui de l'impulsion de Dirac à l'origine, soit celui pour lequel $\widehat{R^\sharp[w]}$ est la fonction caractéristique χ_Ω d'une boule $B(0, \Omega)$ dans le domaine des fréquences (la structure radiale du spectre étant ici à privilégier); d'après la formule d'inversion de Fourier, ce choix de fenêtre w correspond à ce que

$$R^\sharp[w](x) = \frac{\Omega^n}{(2\pi)^{n/2}} \frac{J_{n/2}(\Omega\|x\|)}{(\Omega\|x\|)^{n/2}}, \quad (2.14)$$

où $J_{n/2}$ est la fonction de Bessel d'indice $n/2$. Ce choix correspond (puisque $R^\sharp[w]$ est une fonction radiale) à une fenêtre w de dépendant en fait que de la variable s (et non de $\vec{\theta}$) avec

$$\widehat{w}(\xi) = \frac{1}{2}(2\pi)^{1-n} \widehat{R^\sharp[w]}((\xi, 0, \dots, 0)) |\xi|^{n-1} = \frac{1}{2}(2\pi)^{1-n} |\xi|^{n-1} \chi_{[0, \Omega]}(|\xi|). \quad (2.15)$$

Le paramètre essentiel ici est le choix de Ω . Si la fonction source présente un spectre de répartition d'énergie essentiellement contenue dans la boule de rayon Ω_0 dans le domaine des fréquences, le choix de $w = w_\Omega$ comme dans (2.14) avec $\Omega \geq \Omega_0$ fournira une bonne restitution de f ; il pourra d'ailleurs s'avérer judicieux d'éviter la coupure brutale de $\chi_{[0, \Omega]}$ en Ω en multipliant préalablement $\chi_{[0, \Omega]}$ par la fenêtre de Hamming sur $[-\Omega, \Omega]$ (ce pour pallier aux effets d'*aliasing*). Si le spectre de la fonction source n'est pas essentiellement à support dans la boule de centre Ω_0 , il est important d'avoir en tête le graphe de la fonction

$$\xi \in [0, \infty] \mapsto \frac{J_{n/2}(\xi)}{\xi}$$

puisque le résultat obtenu par rétro-projection filtrée (selon la formule (2.12) est précisément la convolution de f avec la fonction $R^\sharp[w_\Omega]$, fonction s'exprimant (voir (2.14)) en termes précisément de cette fonction (on a figuré sur la figure 2.4 le graphe de cette fonction pour $n = 2$).

2.1.5 La transformation de Radon à l'épreuve de la méthode de Kaczmarz

Rappels généraux sur la méthode itérative de Kaczmarz (point de vue géométrique)

Nous allons rappeler ici une méthode cruciale inspirée de l'algorithmique hilbertienne pour résoudre de manière approchée un système d'équations affines posé dans un espace de Hilbert H , puis nous transposerons cette méthode (suivant divers points de vue, certains numériques, certains plus conceptuels) au cadre de l'inversion « approchée » de la transformation de Radon, suivant la technique qualifiée aujourd'hui de ART³. L'intérêt majeur de cette nouvelle approche (comparée à celle fournie par la rétro-projection filtrée) est que nous pourrons ensuite l'adapter au cadre de transformations intégrales moins élémentaires que la transformation de Radon, telle la transformation de Radon atténuée impliquée, on le verra, dans la modélisation de la SPECT.

On suppose que l'on dispose d'un \mathbb{C} -espace de Hilbert H , de p espaces de Hilbert H_1, \dots, H_p (réels ou complexes), et de la donnée, chaque fois, d'un opérateur linéaire continu surjectif

$$R_j : H \longmapsto H_j .$$

Étant donnés p éléments $g_1 \in H_1, \dots, g_p \in H_p$, on souhaite résoudre de manière approchée le système d'équations affines :

$$R_j[f] = g_j, \quad j = 1, \dots, p . \quad (2.16)$$

Exemple 2.1. On prend $H := L^2(\mathbb{B}^n)$, où \mathbb{B}^n est la boule unité dans \mathbb{R}^n ,

$$H_j = h := L^2_{\mathbb{C}}([-1, 1]^n, (1 - s^2)^{(1-n)/2} ds)$$

et R_j l'opérateur qui à une fonction source f dans H associe l'élément de h défini comme la fonction

$$s \mapsto R[f](\vec{\theta}_j, s),$$

où $\vec{\theta}_j \in \mathbb{S}^{n-1}$ et la fonction source f a été prolongée par 0 hors de \mathbb{B}^n . Ces opérateurs sont évidemment surjectifs et continus puisque l'inégalité de Cauchy-Schwarz nous assure de ce que, pour tout $s \in [-1, 1]$,

$$|R[f](\vec{\theta}_j, s)|^2 \leq (1 - s^2)^{(n-1)/2} \times \text{volume}(\mathbb{B}^{n-1}) \times \int_{\vec{\theta}_j^\perp} |f(s\vec{\theta}_j + y)|^2 dm_{\vec{\theta}_j^\perp}(y),$$

ce qui implique

$$\|R_j f\|_h \leq \sqrt{\text{vol}(\mathbb{B}^{n-1})} \|f\|_H .$$

On se pose le problème de trouver une version approchée \tilde{f} de f en disposant des diverses données $s \mapsto R[f](\vec{\theta}_j, s)$ pour une collection (finie) d'angles $\vec{\theta}_1, \dots, \vec{\theta}_p$.

Exemple 2.2 (un autre regard sur Radon, cette fois totalement discrétisé). On plaque sur le disque $[-1, 1]^2$ une grille discrète de pas τ (dans les deux directions) et on considère une

³Algebraic Reconstruction Technique.

fonction source f définie dans ce carré et constante sur chaque maille. Une telle image f se trouve ainsi identifiée à un vecteur (f_1, \dots, f_N) , où N désigne le nombre de mailles de la grille. Si l'on se donne une collection de lignes L_1, \dots, L_p et que l'on note, pour chaque ligne L_j , $j = 1, \dots, p$, pour chaque maille S_k , $k = 1, \dots, N$,

$$a_{j,k} := \text{longueur}(L_j \cap S_k)$$

et $a_j = (a_{j,1}, \dots, a_{j,N})$, $j = 1, \dots, p$, chercher (dans \mathbb{R}^N ou \mathbb{C}^N) une solution approchée du système linéaire

$$a_j \cdot X = b_j, \quad j = 1, \dots, p$$

(b_1, \dots, b_p désignant des scalaires) revient à poser dans le cadre discret le problème continu consistant à chercher une source approchée f solution de

$$\int_{L_j} f(x) dm_{L_j}(x) = b_j, \quad j = 1, \dots, p.$$

La résolution approchée de ce système linéaire peut être attaquée par la méthode que l'on se propose de décrire.

Si F_j désigne, pour chaque $j = 1, \dots, p$, le noyau de R_j et si T la projection orthogonale sur $F_1 \cap \dots \cap F_p$, alors, pour toute solution f du système (2.16), $(I - T)(f)$ représente la solution de ce même système de norme minimale. C'est sur ce principe qu'est fondée la méthode de Kaczmarz, ainsi que sur un schéma algorithmique très simple qu'il est aisé de visualiser, comme on le verra sur la figure 2.5 plus loin.

On introduit, pour décrire la procédure de Kaczmarz, un paramètre dit « de relaxation » ϖ dont on verra ultérieurement l'intérêt dans nos modèles pratiques, pour l'instant seulement assujetti (c'est important) à appartenir à l'intervalle ouvert $]0, 2[$.

Pour chaque $j = 1, \dots, p$, on note Q_j l'opérateur de projection orthogonale sur le sous-espace fermé F_j et l'on introduit les versions « relaxées » des Q_j que sont les opérateurs

$$Q_j^\varpi := (1 - \varpi)\text{Id}_H + \varpi Q_j, \quad j = 1, \dots, p.$$

On note que

$$\|Q_j^\varpi\|^2 - \|f\|^2 = (2 - \varpi)\varpi(\|Q_j\|^2 - \|f\|^2), \quad (2.17)$$

et c'est dans le fait que $\varpi(2 - \varpi) > 0$ qu'intervient fondamentalement la contrainte sur le paramètre de relaxation. On voit facilement que si

$$Q^\varpi := Q_p^\varpi \circ \dots \circ Q_1^\varpi,$$

alors l'inégalité (2.17) implique

$$\text{Ker}(\text{Id}_H - Q^\varpi) = \bigcap_{j=1}^p \text{Ker}(\text{Id}_H - Q_j).$$

La propriété (assez facile à établir et dont on pourra se convaincre graphiquement en examinant la figure 2.5) suivant laquelle la suite d'opérateurs $(Q_p \circ \dots \circ Q_1)^k$, $k = 1, 2, \dots$, converge fortement vers la projection orthogonale sur l'intersection des sous-espaces $F_j = \text{Ker}(\text{Id}_H - Q_j)$, $j = 1, \dots, p$, c'est-à-dire que pour tout f dans H ,

$$\lim_{k \rightarrow +\infty} (Q_p \circ \dots \circ Q_1)^k(f) = \text{Pr}_{F_1 \cap \dots \cap F_p}[f]. \quad (2.18)$$

Cette propriété subsiste lorsque l'on introduit le paramètre de relaxation et que l'on remplace chaque Q_j par Q_j^ϖ . C'est de cette propriété que découle la proposition sur laquelle s'appuie la méthode itérative de Kaczmarz.

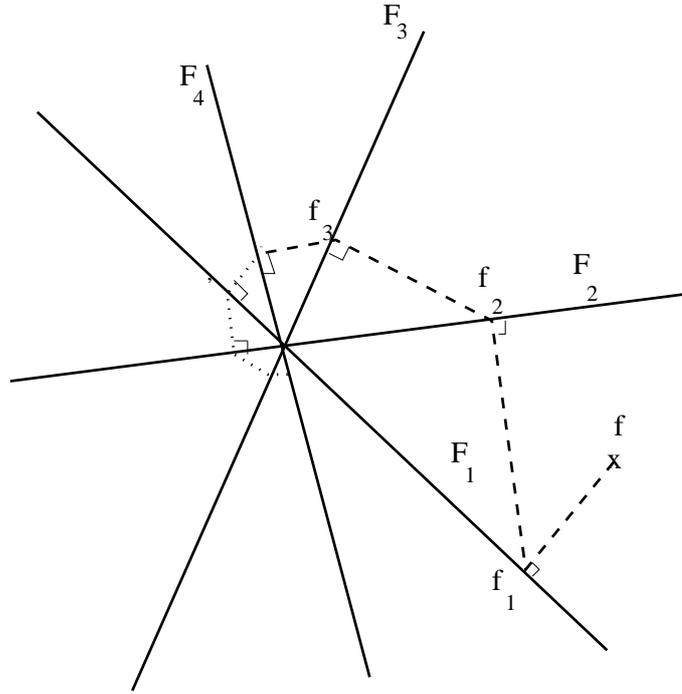


FIG. 2.5 – Une vision heuristique de la méthode de Kaczmarz

Proposition 2.2 Soit $f^0 \in \bigoplus_{j=1}^p F_j^\perp$, où $F_j = \text{Ker } R_j$, $j = 1, \dots, p$. Étant donnés $g_1 \in H_1, \dots, g_p \in H_p$, on note $P_{g,j}$ l'opérateur (affine) de projection orthogonale sur le sous-espace affine fermé des solutions de $R_j[f] = g_j$. Soit $\varpi \in]0, 2[$. La suite d'éléments de H initiée à f^0 et définie par la relation inductive

$$f^{k+1} = \left[\left((1 - \varpi)\text{Id}_H + \varpi P_{g,p} \right) \circ \dots \circ \left((1 - \varpi)\text{Id}_H + \varpi P_{g,1} \right) \right]^k [f^k] \quad (2.19)$$

converge vers la solution de norme minimale du système (2.16).

Preuve. Si l'on note

$$\begin{aligned} P_g^\varpi &= \left((1 - \varpi)\text{Id}_H + \varpi P_{g,p} \right) \circ \dots \circ \left((1 - \varpi)\text{Id}_H + \varpi P_{g,1} \right) \\ Q^\varpi &= \left((1 - \varpi)\text{Id}_H + \varpi \text{Pr}_{F_p} \right) \circ \dots \circ \left((1 - \varpi)\text{Id}_H + \varpi \text{Pr}_{F_1} \right), \end{aligned}$$

on voit par récurrence sur k que, pour toute solution f du système (2.16), on a

$$f^k = (P_g^\varpi)^k [f^0] = f + (Q^\varpi)^k [f^0 - f].$$

Si k tend vers l'infini, il résulte de (2.18) que la suite f^k converge vers $f + T(f^0 - f)$, où T désigne la projection orthogonale sur l'intersection des F_j . Comme $T(f^0) = 0$ du fait de l'hypothèse sur f^0 , la suite f^k converge vers $(I - T)(f)$ qui est bien la solution de norme minimale du système (2.16) voulue (cette solution ne dépend pas de f). \diamond

L'algorithme de Kaczmarz et la transformation de Radon

On reprend la situation de l'exemple 2.1. Il est facile de vérifier que l'adjoint de l'opérateur R_j est l'opérateur

$$R_j^* : g \in L^2([-1, 1], (1 - s^2)^{(1-n)/2} ds) \longmapsto \left[x \mapsto (1 - \langle x, \vec{\theta} \rangle^2)^{(1-n)/2} g(\langle x, \vec{\theta} \rangle) \right].$$

En recomposant avec R_j , on trouve, si $g \in L^2([-1, 1], (1 - s^2)^{(1-n)/2} ds)$,

$$R_j \circ R_j^*(g) : s \mapsto \int_{\theta_j^\perp} (1 - s^2)^{(1-n)/2} g(s) dm_{\theta_j^\perp}(y) = \text{vol}(\mathbb{B}^{n-1}) g(s). \quad (2.20)$$

Fixons ici g_1, \dots, g_p dans $h = L^2([-1, 1], (1 - s^2)^{(1-n)/2} ds)$ et calculons explicitement la projection orthogonale P_j sur le sous-espace des solutions affines du système

$$R[f](\vec{\theta}_j, \cdot) = g_j, \quad j = 1, \dots, p. \quad (2.21)$$

Soit $h \in H$. Comme $P_{g,j}h - h \in (\text{Ker } R_j)^\perp = \overline{\text{Im } R_j^*}$, on peut trouver, pour chaque $j = 1, \dots, p$, une suite $(u_{j,l})_l$ d'éléments de h telle que

$$P_{g,j}[h] - h = \lim_{l \rightarrow +\infty} R_j^*[u_{j,l}]. \quad (2.22)$$

En prenant l'image par R_j , on obtient (d'après (2.20))

$$g_j - R_j[h] = \text{vol}(\mathbb{B}^{n-1}) \times \lim_{l \rightarrow +\infty} u_{j,l}. \quad (2.23)$$

En combinant (2.22) et (2.23), on obtient

$$P_{g,j}[h] = h + \frac{R_j^*[g_j - R_j[h]]}{\text{vol}(\mathbb{B}^{n-1})}, \quad (2.24)$$

ce qui permet de mettre immédiatement en route l'algorithme itératif de Kaczmarz en partant, par exemple, de $f^0 = 0$.

Reste à élucider (ce n'est pas une question facile !) la manière de profiter la plus intelligemment du choix du paramètre de relaxation ϖ . Pour cela, il faut nous préoccuper non plus seulement de l'aspect géométrique du problème, mais de son aspect quantitatif (comment contrôler la convergence de la suite $(f^k)_k$?). On y reviendra après une digression du côté de la « décomposition en valeurs singulières », outil important dans ce type de problème inverse.

2.1.6 La transformation de Radon à l'épreuve de la SVD

Quelques rappels d'algèbre (en dimension finie) : la décomposition en valeurs singulières

Soit R une application linéaire (supposée pour l'instant surjective) de \mathbb{R}^N dans \mathbb{R}^p , représentée par une matrice A lorsque \mathbb{R}^N et \mathbb{R}^p sont rapportés à leurs bases canoniques respectives (l'hypothèse de surjectivité implique évidemment $N \geq p$). D'après la formule du rang, le noyau de R est un sous-espace vectoriel F de \mathbb{R}^N de dimension $N - p$ de \mathbb{R}^N , que l'on peut rapporter à une base orthonormée (e_{p+1}, \dots, e_N) (pour le produit scalaire canonique). Si l'on complète (suivant le procédé de Gram-Schmidt) cette base en une base orthonormée (e_1, \dots, e_N) de \mathbb{R}^N , on constate que (e_1, \dots, e_p) constitue une base du sous-espace F^\perp constitué des vecteurs de \mathbb{R}^N orthogonaux au noyau de R . La restriction de R à F^\perp est une application linéaire bijective entre F^\perp et \mathbb{R}^p . Si \mathbb{R}^p est rapporté à la base (e_1, \dots, e_p) et \mathbb{R}^N à sa base canonique, la matrice de R dans ces bases s'écrit sous la forme :

$$\begin{pmatrix} 0 & \cdots & \cdots & 0 \\ B & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix}$$

où B est une matrice réelle $p \times p$ de déterminant non nul ; il existe donc une matrice orthogonale réelle V_1 de taille (N, N) telle que

$$A = \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ B & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix} \cdot V_1^* .$$

La matrice $B \cdot B^*$ (ici $B^* = {}^t B$ car B est réelle) est une matrice symétrique réelle que l'on peut donc écrire

$$B \cdot B^* = U \cdot \text{diag}(\lambda_1, \dots, \lambda_p) \cdot U^* ,$$

où U est une matrice orthogonale réelle de taille (p, p) ; d'autre part, pour tout $v \in \mathbb{R}^p$,

$$\langle B \cdot B^*(v), v \rangle = \langle B^*(v), B^*(v) \rangle \geq 0 ,$$

ce qui implique que pour $j = 1, \dots, p$, $\lambda_j = \sigma_j^2 \geq 0$ (on appelle σ_j la racine positive du réel positif λ_j) ; on peut d'ailleurs faire en sorte que $\sigma_1^2 \geq \sigma_2^2 \geq \dots \geq \sigma_p^2$. Comme $\det B \neq 0$, ces nombres réels positifs σ_j^2 , $j = 1, \dots, p$ sont tous non nuls. Les nombres réels positifs $\sigma_1, \dots, \sigma_p$ correspondants sont dits *valeurs singulières* de l'application linéaire R (il est clair que ces nombres ne dépendent que de R et non du choix des bases dans lesquelles l'action de R est exprimée).

Comme

$$B \cdot B^* = \left(U \cdot \text{diag}(\sigma_1, \dots, \sigma_p) \right) \cdot \left(U \cdot \text{diag}(\sigma_1, \dots, \sigma_p) \right)^* ,$$

la matrice

$$B^* \cdot \left(\left(U \cdot \text{diag}(\sigma_1, \dots, \sigma_p) \right)^* \right)^{-1}$$

est une matrice orthogonale réelle W de taille (p, p) et l'on peut donc écrire

$$B^* = W \cdot \text{diag}(\sigma_1, \dots, \sigma_p) \cdot U^* ,$$

soit

$$B = U \cdot \text{diag}(\sigma_1, \dots, \sigma_p) \cdot W^* .$$

On peut donc ainsi écrire

$$A = U \cdot \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots \\ 0 & \cdots & \cdots & \sigma_p & 0 & \cdots & \cdots & 0 \end{pmatrix} \cdot V^*$$

où U et V sont respectivement des matrices orthogonales réelles de tailles (p, p) et (N, N) .

Une telle décomposition est dite *décomposition en valeurs singulières* de la matrice A (ou de l'opérateur R) et on la retrouve sous un environnement de calcul scientifique (comme MATLAB ou SCILAB) sous la commande :

```
>> [U,D,V] = svd(A);
```

Cette transformation permet (quitte à effectuer des transformations orthogonales à la source et au but) de ramener l'étude d'une application linéaire surjective de \mathbb{R}^N dans \mathbb{R}^p à l'étude d'une application linéaire dont la matrice est une matrice diagonale

de taille (p, p) , complétée à droite par une matrice de zéros de taille $(p, N - p)$. C'est un outil très intéressant du point de vue pratique. On se doute en particulier du rôle important joué par les vecteurs colonnes de V , vecteurs « modèle » de l'espace source \mathbb{R}^N lorsque'il est soumis à la transformation linéaire R de matrice A .

Si R n'est pas surjective, cette décomposition reste valide (mais $p - \text{rang}(R)$ valeurs singulières μ_j sont alors nulles).

Tout ceci peut se transposer au cadre complexe (R étant une application \mathbb{C} -linéaire de \mathbb{C}^N dans \mathbb{C}^p); il suffit juste de remplacer « orthogonale » par « unitaire ». Les valeurs singulières restent des nombres strictement positifs.

La construction que nous venons de faire nous a permis de trouver un système orthonormé de \mathbb{R}^p (constitué en fait de $r = \text{rang}(R)$ éléments u_1, \dots, u_r), un système orthonormé v_1, \dots, v_r dans \mathbb{R}^N , tels que l'on puisse écrire, pour tout $X \in \mathbb{R}^N$,

$$A[X] = \sum_{k=1}^r \sigma_k \langle X, v_k \rangle u_k, \quad (2.25)$$

où $\sigma_1, \dots, \sigma_r$ sont des nombres strictement positifs. Les nombres strictement positifs σ_k^2 , $k = 1, \dots, r$ sont les valeurs propres non nulles de la matrice symétrique (ou hermitienne) AA^* , les vecteurs u_1, \dots, u_r sont les vecteurs propres (normalisés) correspondants et les v_k sont définis par

$$v_k = \frac{1}{\sigma_k} A^*[u_k], \quad k = 1, \dots, r.$$

On dit que (2.25) est une *décomposition en valeurs singulières pour l'opérateur A* . Si $Y = Y_0$ est un vecteur de \mathbb{R}^p donné, on remarque que

$$\sum_{k=1}^r \frac{1}{\sigma_k} \langle Y_0, u_k \rangle v_k$$

est le vecteur de norme minimale parmi tous les vecteurs $X \in \mathbb{R}^N$ minimisant

$$\varphi \longmapsto \|A[X] - Y_0\|^2,$$

donc un candidat envisageable comme antécédent de Y_0 *via* l'application A (application qui bien sûr n'est en général ni injective ni surjective).

Décomposition en valeurs singulières de la transformation de Radon

Un système de polynômes orthogonaux $(G_{n,k})_{k \in \mathbb{N}}$ joue un rôle particulièrement important dans l'étude de la transformation de Radon; ce sont les polynômes de Gegenbauer d'indice $\lambda = n/2$. Lui est liée la collection $(G_{n,k})_{k \in \mathbb{N}}$ de fonctions polynomiales ($G_{n,k}$ étant une fonction polynomiale de degré exactement k) formant un système orthogonal (et total) dans l'espace de Hilbert $L^2([-1, 1], (1 - s^2)^{(n-1)/2} dt)$, en général normalisées de manière à ce que

$$G_{n,k}(1) = \sup_{[-1,1]} |G_{n,k}(s)| = 1.$$

Si l'on se réfère aux tables, on trouve, pour tout $k \in \mathbb{N}$,

$$\int_{[-1,1]} |G_{n,k}(s)|^2 (1 - s^2)^{(n-1)/2} ds = \frac{2^{n-1} (\Gamma((n+1)/2))^2 k!}{(k + n/2) \Gamma(n+k)}.$$

Les fonctions polynomiales $G_{n,2k}$ sont paires, les fonctions polynomiales $G_{n,2k+1}$ sont impaires. Chacune de ces fonctions polynomiales est une fonction oscillante, la fréquence (tout comme l'amplitude) des oscillations s'amplifiant au fur et à mesure que l'on approche des extrémités de l'intervalle $[-1, 1]$ (voir la figure 2.6). On

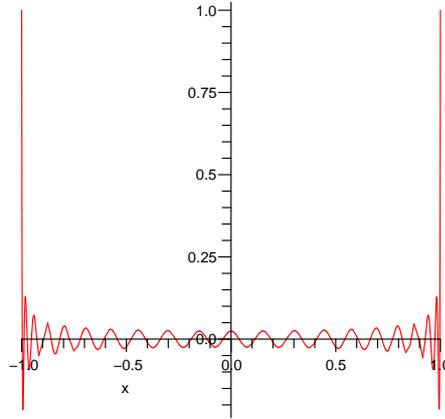


FIG. 2.6 – Le graphe de $G_{2,40}$

trouvera sous MAPLE comment générer (et tracer les graphes) des fonctions polynomiales de Gegenbauer. Les relations de récurrence que l'on utilise habituellement pour les calculer (ainsi normalisées) sont les relations

$$\begin{aligned} k\tilde{G}_{n,k}(s) &= 2(k+n/2-1)s\tilde{G}_{n,k-1}(s) - (k+n-2)\tilde{G}_{n,k-2}(s), \quad k \geq 2 \\ G_{n,k}(s) &= \frac{\tilde{G}_{n,k}(s)}{\tilde{G}_{n,k}(1)} \end{aligned}$$

initiées avec $\tilde{G}_{n,0}(s) = 1$, $\tilde{G}_{n,1}(s) = ns$ (on pourra ainsi aisément les générer sous MATLAB).

Le fait que les polynômes de Gegenbauer jouent un rôle important dans la description de la transformation de Radon tient à un lemme calculatoire que nous établirons ici (on pourra l'admettre), intimement lié aux propriétés d'orthogonalité des polynômes de Gegenbauer.

Lemme 2.1 Si $\vec{\theta}_1$ et $\vec{\theta}_2$ sont deux directions de \mathbb{S}^{n-1} ,

$$R[x \mapsto G_{n,k}(\langle x, \vec{\theta}_2 \rangle)](\vec{\theta}_1, s) = c(n) G_{n,k}(\langle \vec{\theta}_1, \vec{\theta}_2 \rangle) \times (1-s^2)^{(n-1)/2} G_{n,k}(s) \quad (2.26)$$

avec, compte tenu des normalisations

$$c(n) = \frac{\pi^{(n-1)/2}}{\Gamma\left(\frac{n+1}{2}\right)}.$$

Preuve. On fera la preuve dans le cas $n = 2$ pour simplifier. Le membre de gauche de la formule (2.26) est (si l'on suppose la fonction polynomiale $G_{2,k}$ tronquée à $[-1, 1]$)

$$\begin{aligned} \int_{\mathbb{R}} G_{2,k}(\langle s\vec{\theta}_1 + t\theta_1^\perp, \theta_2 \rangle) dt &= \int_{|t| \leq \sqrt{1-s^2}} G_{2,k}(s \cos \varphi + t \sin \varphi) dt \\ &= \int_{-1}^1 G_{2,k}(s(\cos \varphi + \sqrt{1-s^2}u)) du, \end{aligned}$$

où $\langle \vec{\theta}_1, \vec{\theta}_2 \rangle = \cos \varphi$. En développant $G_{2,k}$ comme un polynôme de degré k , on constate que le membre de gauche de (2.26) s'écrit

$$\sqrt{1-s^2} \times P_k(\vec{\theta}_1, s),$$

où $P_k(\vec{\theta}_1, \cdot)$ est une fonction polynomiale de degré k . Du fait de l'orthogonalité des polynômes de Gegenbauer $G_{2,k}$ dans $L^2([-1, 1], \sqrt{1-s^2})$, on constate que

$$\int_{[-1,1]} \sqrt{1-s^2} P_k(s) G_{2,l}(s) ds = 0 \quad \forall l > k;$$

par symétrie de l'opérateur

$$g \mapsto \left[s \mapsto R[x \mapsto g(\langle x, \vec{\theta}_2 \rangle)](\vec{\theta}_1, s) \right],$$

on constate aussi que

$$\int_{[-1,1]} \sqrt{1-s^2} P_k(s) G_{2,l}(s) ds = 0 \quad \forall l < k.$$

En utilisant le fait que $G_{2,k}(1) = 1$, on constate donc ici que

$$P_k = G_{2,k}(\langle \vec{\theta}_1, \vec{\theta}_2 \rangle),$$

ce qui est la formule (2.26) voulue dans le cas $n = 2$; le cas général se traite de manière similaire. \diamond

La formule (2.26) permet d'exhiber une base orthonormée de l'espace de Hilbert

$$L^2(\mathbb{S}^{n-1} \times [-1, 1], d\sigma_{\mathbb{S}^{n-1}} \otimes (1-s^2)^{(1-n)/2} ds)$$

constituée de fonctions propres pour l'opérateur $R \circ R^*$, la transformation de Radon étant ici considérée comme un opérateur continu

$$R : L^2(\mathbb{B}^n) \longrightarrow L^2(\mathbb{S}^{n-1} \times [-1, 1], d\sigma_{\mathbb{S}^{n-1}} \otimes (1-s^2)^{(1-n)/2} ds).$$

Comme on le vérifie en effet grâce au théorème de Fubini, l'adjoint R^* de R transforme une fonction $g : (\vec{\theta}, s) \mapsto g(\vec{\theta}, s)$ représentant un élément de l'espace de Hilbert $L^2(\mathbb{S}^{n-1} \times [-1, 1], d\sigma_{\mathbb{S}^{n-1}} \otimes (1-s^2)^{(1-n)/2} ds)$ en l'élément de $L^2(\mathbb{B}^n)$ de représentant

$$x \longmapsto \int_{\mathbb{S}^{n-1}} (1 - \langle \vec{\theta}, x \rangle)^{(1-n)/2} g(\vec{\theta}, \langle \vec{\theta}, x \rangle) d\sigma_{\mathbb{S}^{n-1}}(\vec{\theta}).$$

En composant à gauche avec la transformation de Radon, on trouve

$$R[R^*[g]] = R \circ R^*[g] : (\vec{\theta}, s) \longmapsto \int_{\mathbb{S}^{n-1}} R_{\vec{\theta}}[R_{\vec{\tau}}^*[g]](s) d\sigma_{\mathbb{S}^{n-1}}(\vec{\tau}).$$

Si l'on prend pour g une fonction de la forme

$$h \otimes u_{n,k} : (\vec{\theta}, s) \longmapsto h(\vec{\theta}) \times (1-s^2)^{(n-1)/2} G_{n,k}(s), \quad k \in \mathbb{N},$$

où h est une fonction C^∞ sur \mathbb{S}^{n-1} , on trouve donc, en utilisant la relation (2.26)

$$\begin{aligned} (RR^*[h \otimes u_{n,k}])(\vec{\theta}, s) &= c(n) \left(\int_{\mathbb{S}^{n-1}} G_{n,k}(\langle \vec{\tau}, \vec{\theta} \rangle) h(\vec{\tau}) d\sigma_{\mathbb{S}^{n-1}}(\vec{\tau}) \right) \\ &\quad \times (1-s^2)^{(n-1)/2} G_{n,k}(s). \end{aligned} \quad (2.27)$$

SVD de la transformée de Radon en dimension 2.

Dans le cas particulier où $n = 2$, on remarque que si

$$h_l : \theta \longmapsto e^{il\theta}, \quad l \in \mathbb{Z},$$

le système $(h_l)_{l \in \mathbb{N}}$ forme un système orthogonal total de $L^2(\mathbb{S}^1, d\sigma_{\mathbb{S}^1})$ tandis que le système $(u_{2,k})_{k \in \mathbb{N}}$, où

$$u_{2,k} : s \in [-1, 1] \longmapsto \sqrt{1-s^2} \times G_{2,k}(s), \quad k \in \mathbb{N},$$

forme un système orthogonal total de $L^2([-1, 1], ds/\sqrt{1-s^2})$, ce qui implique que le système des fonctions « doublement oscillantes » (en θ et en s) $(h_l \otimes u_{2,k})_{l \in \mathbb{Z}, k \in \mathbb{N}}$ constitue un système orthogonal total de l'espace de Hilbert d'arrivée

$$L^2(\mathbb{S}^1 \times [-1, 1], d\sigma_{\mathbb{S}^1} \otimes ds/\sqrt{1-s^2})$$

de l'opérateur $R \circ R^*$. La formule (2.27) se lit dans le cas $n = 2$ et avec $h = h_l$, $l \in \mathbb{Z}$,

$$RR^*[h_l \otimes u_{2,k}] = \left(\int_0^{2\pi} G_{2,k}(\cos u) e^{ilu} du \right) h_l \otimes u_{2,k} \quad (2.28)$$

$$= \frac{1}{2} \left(\int_{-\pi}^{\pi} G_{2,k}(\cos u) \cos(lu) du \right) h_l \otimes u_{2,k}. \quad (2.29)$$

Les fonctions $h_l \otimes u_{2,k}$, où $k+l$ est pair, constituent un système orthogonal engendrant dans $L^2(\mathbb{S}^1 \times [-1, 1], d\sigma_{\mathbb{S}^1} \otimes ds/\sqrt{1-s^2})$ le sous-espace fermé \mathbf{G}_2 constitué des classes de fonctions telles que $g(\vec{\theta}, s) \equiv g(-\vec{\theta}, -s)$; pour chaque tel couple (k, l) , on notera

$$\sigma_{2,l,k}^2 = \frac{1}{2} \int_{-\pi}^{\pi} G_{2,k}(\cos u) \cos(lu) du.$$

(ces nombres sont positifs et nuls si $l > k$ puisque $G_{2,k}$ est un polynôme de degré k). En fait, on peut les calculer explicitement en remarquant que

$$G_{2,k}(t) = \frac{1}{k+1} \frac{\sin((k+1)\arccos t)}{\sin(\arccos t)}, \quad k \in \mathbb{N}.$$

On constate alors que, si $0 \leq |l| \leq k$,

$$\sigma_{2,l,k} = 2\sqrt{\frac{\pi}{k+1}}, \quad k \in \mathbb{N}, \quad l \in \mathbb{Z}, \quad k+l \text{ pair}.$$

Si l'on définit, pour tous les couples $(l, k) \in \mathbb{Z} \times \mathbb{N}$ tels que $k+l$ soit pair et $|l| \leq k$,

$$g_{2,l,k} := \frac{h_l \otimes u_{2,k}}{\|h_l \otimes u_{2,k}\|_2},$$

la norme L^2 étant prise dans $L^2(\mathbb{S}^1 \times [-1, 1], d\sigma_{\mathbb{S}^1} \otimes ds/\sqrt{1-s^2})$, et que l'on pose

$$f_{2,l,k} = \frac{1}{\sigma_{2,l,k}} R^*[g_{2,l,k}],$$

on dispose avec le système $(f_{2,l,k})_{l,k}$ d'une base hilbertienne de $L^2(\mathbb{B}^2)$, avec le système $(g_{2,l,k})_{l,k}$ d'une base hilbertienne du sous-espace \mathbf{G}_2 . Les valeurs singulières de l'opérateur R sont les $\sigma_{2,l,k}$, $0 \leq |l| \leq k$, avec $k \in \mathbb{N}$, $l \in \mathbb{Z}$ et $k+l$ pair.

Au final, nous avons construit un système orthonormé $(g_{2,l,k})_{l,k}$ de

$$H' := L^2(\mathbb{S}^1 \times [-1, 1], d\sigma_{\mathbb{S}^1} \otimes ds / \sqrt{1-s^2})$$

et un système orthonormé $(f_{2,l,k})_{k,l}$ de $H := L^2(\mathbb{B}^2)$, avec $k \in \mathbb{N}$, $l \in \mathbb{Z}$, $|l| \leq k$, $k+l$ pair, ce de manière à ce que, pour tout $f \in H$,

$$R[f] = 2 \sum_{k=0}^{\infty} \sqrt{\frac{\pi}{k+1}} \left(\sum_{\substack{|l| \leq k \\ k+l \text{ pair}}} \langle f, f_{2,l,k} \rangle g_{2,l,k} \right), \quad (2.30)$$

ce qui constitue bien, sur le modèle de (2.25), une décomposition en valeurs singulières pour R .

SVD de la transformée de Radon en dimension $n > 2$.

À la place des h_l , $l \in \mathbb{Z}$, on utilise une base orthonormée de $L^2(\mathbb{S}^{n-1})$ constituée d'harmoniques sphériques (h_l étant la restriction à la sphère d'un polynôme harmonique homogène de degré l). Pour tous les couples (l, k) de $\mathbb{N} \times \mathbb{N}$ tels que $l+k$ soit pair, on note

$$g_{n,l,k} := \frac{h_l \otimes u_{n,k}}{\|h_l \otimes u_{n,k}\|_2},$$

où cette fois

$$u_{n,k}(s) = (1-s^2)^{(n-1)/2} G_{n,k}(s), \quad s \in [-1, 1].$$

On observe que l'on a encore (grâce aux identités (2.27) appliquées avec $h = h_l$) des relations du type

$$RR^*[h_l \otimes u_{n,k}] = \sigma_{n,l,k}^2 h_l \otimes u_{n,k}, \quad k, l \in \mathbb{N}, \quad k+l \text{ pair},$$

avec des nombres positifs $\sigma_{n,l,k}$, d'ailleurs calculables suivant les formules intégrales :

$$\sigma_{n,l,k} = \sqrt{\frac{\pi^{(n-1)/2}}{\Gamma\left(\frac{n+1}{2}\right)} \int_{-1}^1 G_{n,k}(t) G_{n-2,l}(t) (1-t^2)^{(n-3)/2} dt}$$

(et nuls si $l > k$). En posant cette fois, pour tous les couples (l, k) d'entiers positifs tels que $k+l$ est pair et $l \leq k$,

$$f_{n,l,k} = \frac{1}{\sigma_{n,l,k}} R^*[g_{n,l,k}],$$

on dispose avec le système $(f_{n,l,k})_{l,k}$ ⁴ d'une base hilbertienne de $L^2(\mathbb{B}^n)$, avec le système $(g_{n,l,k})_{l,k}$ d'une base hilbertienne du sous-espace \mathbf{G}_n de

$$L^2(\mathbb{S}^{n-1} \times [-1, 1], d\sigma_{\mathbb{S}^{n-1}} \otimes (1-s^2)^{(1-n)/2} ds)$$

⁴Pour ne pas alourdir les notations, nous confondrons sous un même indice l l'indexation de toutes les harmoniques sphériques de degré l (il y en a $N(n, l)$, nous y reviendrons).

constitué des classes de fonctions g telles que $g(\vec{\theta}, s) = g(-\vec{\theta}, -s)$. Les valeurs singulières correspondantes de l'opérateur R sont les $\sigma_{n,l,k}$, $0 \leq l \leq k$, avec $k, l \in \mathbb{N}$ et $k+l$ pair.

Comme dans le cas de la dimension 2, nous avons, sur le modèle de (2.25) et de (2.30) une décomposition en valeurs singulières de la transformation de Radon de la forme

$$R[f] = \sum_{k=0}^{\infty} \left(\sum_{\substack{|l| \leq k \\ k+l \text{ pair}}} \sigma_{n,l,k} \langle f, f_{n,l,k} \rangle g_{n,l,k} \right).$$

Notons simplement ici que pour chaque l , il y a $N(n, l)$ harmoniques sphériques⁵ de degré l et il serait donc plus judicieux d'écrire la formule ci-dessus

$$R[f] = \sum_{k=0}^{\infty} \left(\sum_{\substack{0 \leq l \leq k \\ k+l \text{ pair}}} \sigma_{n,l,k} \left(\sum_{\lambda=1}^{N(n,l)} \langle f, f_{n,l,\lambda,k} \rangle g_{n,l,\lambda,k} \right) \right).$$

Comment exploiter la SVD pour « inverser » ?

Supposons que l'on ait un opérateur continu (par exemple R) d'un espace de Hilbert H dans un autre espace de Hilbert H' et que l'on veuille, étant donné $g \in H'$, trouver l'antécédent (ou du moins un antécédent jugé convenable) de g via R , c'est-à-dire un élément de H tel que $R[f] = g$. L'idée que l'on proposera ici est celle qui consiste à proposer comme choix dde f , parmi tous les éléments minimisant sur H l'application

$$\varphi \longmapsto \|R[\varphi] - g\|_{H'},$$

celui de norme minimale. En fait cet élément f , dit inverse de Moore-Penrose de g via R est l'unique solution de l'équation

$$R^*(R[f]) = R^*[g]$$

si $R^* : H' \rightarrow H$ désigne l'adjoint de R .

Si l'on dispose d'un système orthogonal $(g_k)_k$ de H' , d'un système orthogonal $(f_k)_k$ de H et d'une représentation de R de la forme

$$R[f] = \sum_{k=1}^{\infty} \sigma_k \langle f, f_k \rangle g_k,$$

où les scalaires σ_k sont des nombres strictement positifs (c'est précisément ce que l'on appelle une décomposition en valeurs singulières de R), l'élément de H de norme minimale parmi les éléments h minimisant

$$h \longmapsto \|R[h] - g\|_{H'}$$

est précisément donné par

$$f = \sum_{k=1}^{\infty} \frac{1}{\sigma_k} \langle g, g_k \rangle f_k$$

(c'est l'unique solution de $R^*[R[f]] = R^*[g]$).

⁵En fait, $N(n, 0) = 1$ et $N(n, l) = \frac{(2l+n-2)!(n+l-3)!}{l!(n-2)!}$, l'une des meilleures références sur les harmoniques sphériques étant l'article de R.T. Seeley, *Spherical Harmonics*, Amer. Math. Monthly 73,1966, 115-121.

Quelques précisions sur l'estimation d'erreur dans l'algorithme de Kaczmarz

Revenons dans cette section sur la démarche algorithmique de Kaczmarz telle que nous l'avons mise en pratique pour la transformation de Radon avec la proposition 2.2 transcrite dans ce cas avec les formules (2.24).

Soient $\vec{\theta}_1, \dots, \vec{\theta}_p$ p éléments de \mathbb{S}^{n-1} tels que $\vec{\theta}_{j_1} \neq \pm \vec{\theta}_{j_2}$ pour deux indices quelconques j_1 et j_2 distincts. Pour $k \in \mathbb{N}$ et $j = 1, \dots, p$, notons

$$f_{n,k}^{\vec{\theta}_j} : x \in \mathbb{B}^n \longmapsto G_{n,k}(\langle x, \vec{\theta}_j \rangle).$$

Ces fonctions source particulières engendrent un sous-espace de dimension finie de $L^2(\mathbb{B}^n)$ (on vérifie que les fonctions $f_{n,k}^{\vec{\theta}_1}, \dots, f_{n,k}^{\vec{\theta}_{k+1}}$ constituent d'ailleurs, si $k < p$, une base de ce sous-espace, que l'on note \mathcal{H}_k). Ce sous-espace est d'ailleurs stable par chacun des opérateurs Q_1, \dots, Q_p correspondant aux projections orthogonales respectives sur les noyaux des opérateurs R_1, \dots, R_p puisque la relation (2.26) implique, pour j_1, j_2 dans $\{1, \dots, p\}$ et pour $k \in \mathbb{N}$,

$$Q_{j_2}[f_{n,k}^{\vec{\theta}_{j_1}}] = f_{n,k}^{\vec{\theta}_{j_1}} - \frac{c(n)}{\text{vol}(\mathbb{B}^{n-1})} \times G_{2,k}(\langle \vec{\theta}_{j_1}, \vec{\theta}_{j_2} \rangle) \times f_{n,k}^{\vec{\theta}_{j_2}}. \quad (2.31)$$

On peut d'ailleurs définir aussi les espaces \mathcal{H}_k pour $k \geq p$ et remarquer que

$$\bigoplus_{j=1}^p \overline{\text{Im } R_j^*} = \bigoplus_{j=1}^p (\text{Ker } [R_j])^\perp = \bigoplus_{k=0}^{\infty} \mathcal{H}_k.$$

Supposons que l'algorithme de Kaczmarz soit initié à $f^0 = 0$ et soit d_g la distance de l'origine à l'intersection des sous-espaces affines $\{f; R_j[f] = g_j\}$, $j = 1, \dots, p$.

Si $\varpi \in]0, 2[$ est un paramètre de relaxation et si $\rho_k(\varpi)$ désigne, pour $k = 0, \dots, p-1$, le rayon spectral de la restriction de $Q_p^\varpi \circ \dots \circ Q_1^\varpi$ au sous-espace de dimension finie \mathcal{H}_k ⁶, la composante $e_{k,M}$ sur \mathcal{H}_k de l'erreur après M itérations de l'algorithme de Kaczmarz est estimée en

$$\|e_{k,M}\|_{L^2(\mathbb{B}^n)} \leq (\rho_k(\varpi))^M d_g.$$

Comme la composante de f sur les espaces \mathcal{H}_k pour $k \geq p$ ne saurait être déterminée par les données $R_j f = g_j$ (puisque les éléments de \mathcal{H}_k pour $k \geq p$ ne peuvent être reconstitués qu'à partir d'au moins $k+1 \geq p+1$ coupes du fait de leur « fréquence » d'oscillation k), il n'a de sens (dans l'estimation d'erreur au niveau de la méthode de Kaczmarz) que de considérer le comportement de $\|e_{k,M}\|_{L^2(\mathbb{B}^n)}$ pour $k < p$.

On constate en dimension $n = 2$, si les angles sont pris dans l'ordre croissant entre 0 et π , qu'un choix de ϖ petit⁷ privilégie (du point de vue de la vitesse de convergence de l'erreur) les valeurs de k petites (donc l'accès aux structures cohérentes de la

⁶Ce rayon spectral peut se calculer soit de manière directe, soit de manière approchée *via* la formule du rayon spectral, en explicitant la matrice de l'opérateur $Q_p^\varpi \circ \dots \circ Q_1^\varpi$ exprimée dans la base $f_{n,k}^{\vec{\theta}_1}, \dots, f_{n,k}^{\vec{\theta}_{k+1}}$ de \mathcal{H}_k ; on trouvera de très intéressants résultats obtenus au terme de ce calcul (suivant les valeurs de ϖ) concernant le comportement de $k \mapsto \rho_k(\varpi)$ en fin de section V.4 du livre de F. Natterer, *The Mathematics of Computerized Tomography*, Wiley & Sons, 1986.

⁷Le choix de .1, .2 est souvent fait en pratique.

fonction source inconnue) par rapport aux valeurs de k plus grandes (donc l'accès aux détails ou accidents de la fonction source). Au contraire, un choix de ϖ grand (par exemple le choix classique $\varpi = 1$) renverse cette situation.

Si les angles sont pris dans un ordre aléatoire (ce qui semble de fait à privilégier), les $\rho_k(\varpi)$ sont d'autant plus petits (pour les valeurs de k petit) que ϖ est grand ; choisir le paramètre de relaxation grand est dans ce cas plus intéressant pour récupérer vite les structures cohérentes ; pour k grand, les valeurs maximales de $\varpi \mapsto \rho_k(\varpi)$ lorsque $\varpi > .5$ dans ce cas restent en dessous des valeurs observées lorsque l'on impose la clause de croissance au choix de l'ordre des angles. On pourra en TP vérifier en pratique ces constats faits à la lumière d'une analyse théorique de la décomposition en valeurs singulières de la transformation de Radon et de l'algorithme de Kaczmarz.

Le choix du paramètre de relaxation dans la méthode de Kaczmarz n'est donc pas anodin et la connaissance explicite de la décomposition en valeurs singulières de la transformée de Radon permet de se faire une idée de la raison pour laquelle il en est ainsi. On comprend aussi au travers de la connaissance de cette décomposition en valeurs singulières (et du caractère oscillant en $\vec{\theta}$ et s des fonctions propres de $R \circ R^*$) pourquoi les bases d'ondelettes (en $\vec{\theta}$ et s) ne sont pas des outils adéquats à l'analyse d'un sinogramme, nous y reviendrons lorsque nous étudierons le lien entre la transformation de Radon et la transformée bidimensionnelle continue en ondelettes.

2.2 S.P.E.C.T et Transformation de Radon atténuée

2.2.1 Le problème physique : présentation succincte de la S.P.E.C.T

On rappelle⁸ que l'on entend sous la terminologie S.P.E.C.T (*Single Photon Emission Computed Tomography* la *tomographie d'émission monophotonique* (T.E.M.P en français). Il s'agit de visualiser la localisation tridimensionnelle dans l'organisme d'un radio-traceur (c'est-à-dire une substance radio-active émettrice de photons). Grâce à l'analyse des coupes $2D$ suivant un axe, on déduit, coupe par coupe, la « carte » de concentration du radio-traceur dans les coupes reconstruites.

Le choix du radio-traceur n'est pas anodin : il dépend de sa capacité à suivre un métabolisme (ou à permettre de réaliser un diagnostic) ; ce peut être un atome seul (iode 123), une molécule « marquée » (diphosphate marqué au technetium 99m) ; c'est ce dernier isotope, de part sa courte demi-vie physique (6 heures) et la facilité de son obtention par générateur portable, qui s'avère d'emploi le plus courant (on le note ^{99m}Tc), dans 90% des cas (détection de métastases, scintigraphies ventriculaires *via* le marquage des globules rouges).

L'organisme « chargé » (du fait qu'on lui a fait absorber préalablement l'isotope radio-traceur) émet alors des photons Gamma, qui seront captés dans les collimateurs d'un dispositif de Gamma caméra (en rotation autour de l'axe orthogonal aux plans des

⁸Pour rédiger cette section, je me suis appuyé sur le mémoire de la toute récente thèse d'Elie Nasr : « Méthodes hilbertiennes pour la correction d'atténuation en tomographie d'émission monophotonique », thèse soutenue à Bordeaux 1 en Juillet 2008.

coupes). L'étude des dispositifs enregistreurs (c'est-à-dire des Gamma caméras) et en particulier de l'architecture des grilles de collimateurs (faut-il privilégier des collimateurs « fins » de manière à privilégier la précision de la détermination de l'azimut dont les photons proviennent ou au contraire des grilles de collimateurs à « gros trous » pour espérer capter le plus de photons possible ?) fera l'objet d'une discussion ultérieure, lorsque nous évoquerons le problème sous-jacent de déconvolution. Dans cette section, nous nous focaliserons uniquement sur la modélisation du rayonnement lui-même.

Le modèle physique de propagation des photons au travers du milieu est fourni par l'équation de transport. Plaçons nous pour simplifier en dimension 2 (c'est-à-dire dans un plan de coupe perpendiculaire à l'axe du dispositif de Gamma-caméra en rotation).

Pour l'instant, nous envisagerons un modèle déterministe pour ce qui concerne le « terme source »

$$(x, y, t) \mapsto f(x, y, t) dx dy dt$$

correspondant au rayonnement (mesure du nombre de particules émises pendant l'intervalle temporel infinitésimal $[t, t + dt]$ depuis la zone $[x, x + dx] \times [y, y + dy]$). On désigne par

$$(x, y, t) \mapsto \Phi(x, y, t)$$

la densité de particules émises traversant dans l'intervalle temporel $[t, t + dt]$ la zone $[x, x + dx] \times [y, y + dy]$ et par

$$(x, y, t) \mapsto a(x, y, t) dx dy dt$$

la mesure à densité correspondant à la « carte d'atténuation » linéique du tissu dans lequel sont appelées à se propager ces photons (on suppose que cette densité calculée au point (x, y) et à l'instant t ne dépend pas de la direction de la droite de propagation issue de (x, y) dans le plan de coupe étudié). L'atténuation est due à plusieurs phénomènes ; c'est la somme d'une mesure d'absorption a_σ photoélectrique, d'une mesure d'absorption a_{Compton} due à la déflexion par effet Compton de certains photons suite à des diffusions, d'une mesure a_{Rayleigh} due à la diffusion par effet Rayleigh, et enfin d'une mesure d'atténuation a_{paires} générée par le phénomène de création de paires. Le coefficient instantané d'atténuation linéique (c'est-à-dire dans la propagation des photons suivant une droite depuis le point (x, y, z_0) , en particulier suivant une droite située dans le plan de coupe $\{z = z_0\}$) s'exprime en

$$a(x, y, t)_{z=z_0} = \rho(x, y, z_0) \times \frac{\text{Nombre d'Avogadro}}{\text{Masse atomique}(x, y, z_0)} \times \sigma(x, y, t)_{z=z_0},$$

où $(x, y, t) \mapsto \sigma(x, y, t)_{z=z_0}$ est la « section efficace totale » (exprimée en cm^2 par atome au point (x, y, z_0) et à l'instant t , $\rho(x, y, z_0)$ la densité du tissu exprimée en grammes par cm^3 au point (x, y, z_0) , la masse atomique du milieu au point (x, y, z_0) étant exprimée, elle, en grammes par molécule. Le modèle mathématique décrit par l'équation de transport est

$$\frac{\partial \Phi}{\partial t}(x, y, t) + \langle v(x, y, t), \nabla_{x,y} \Phi(x, y, t) \rangle = -|v(x, y, t)| a(x, y, t) \Phi(x, y, t) + f(x, y, t),$$

où $v(x, y, z, t)$ désigne le vecteur vitesse instantané. Étudions la propagation le long de la droite

$$L_{s,\vec{\theta}} := \{(x, y) ; (x, y) = s\vec{\theta} + \xi\vec{\theta}^\perp, \xi \in \mathbb{R}\}.$$

Si l'on suppose $|v|$ constant (dans la propagation le long de cette droite, on a donc $v = |v|\vec{\theta}^\perp$) et Φ stationnaire (indépendant de t), on obtient donc

$$\frac{\partial \Phi(s\vec{\theta} + \xi\vec{\theta}^\perp)}{\partial \xi} = -a(s\vec{\theta} + \xi\vec{\theta}^\perp)\Phi(s\vec{\theta} + \xi\vec{\theta}^\perp) + f(s\vec{\theta} + \xi\vec{\theta}^\perp).$$

Cette équation différentielle du premier ordre s'intègre, si l'on suppose a indépendant du temps et f nulle hors d'un disque, en

$$\Phi(s\vec{\theta} + t\vec{\theta}^\perp) = \int_{-\infty}^t \exp \left[- \int_{\xi}^{\infty} a(s\vec{\theta} + \eta\vec{\theta}^\perp) d\eta \right] f(s\vec{\theta} + \xi\vec{\theta}^\perp) d\xi.$$

Suivant ce modèle (évidemment très simplifié), on constate donc que la densité de photons « captée » par le collimateur (supposé infiniment fin et profond) placé à l'extrémité de $L_{s,\vec{\theta}}$ dans la direction vers laquelle pointe $\vec{\theta}^\perp$ se trouve être

$$I(s, \vec{\theta}) = \int_{-\infty}^{\infty} \exp \left[- \int_{\xi}^{\infty} a(s\vec{\theta} + \eta\vec{\theta}^\perp) d\eta \right] f(s\vec{\theta} + \xi\vec{\theta}^\perp) d\xi.$$

Si $a \simeq 0$, on retrouve ici $R[f](\vec{\theta}, s)$ et ainsi la transformation de Radon. Si a n'est plus supposée négligeable, on obtient une version « atténuée » de la transformation de Radon, dite précisément *transformation de Radon atténuée*.

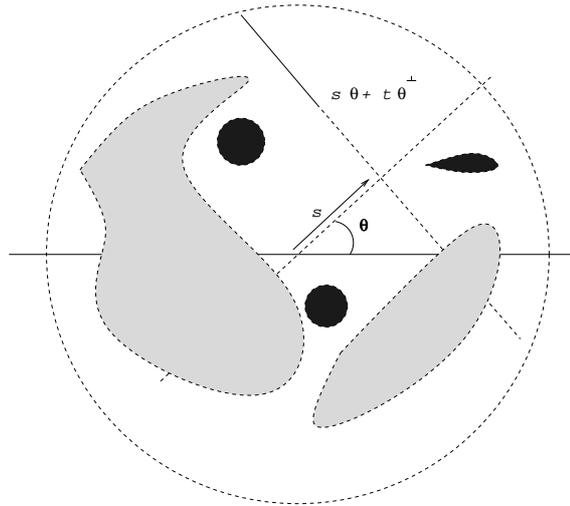


FIG. 2.7 – La transformée de Radon atténuée

En pratique, on verra ultérieurement que la carte d'atténuation ne saurait être supposée nulle, ni même de norme uniforme petite! Les modèles réalistes supposent, face à une source inconnue f de norme uniforme égale à 1 (par exemple, comme vous le verrez en TP, une combinaison de fonctions caractéristiques de disques de rayon petit), une atténuation telle que la *fan-beam fonction*

$$D[a] : (x, y, \theta) \mapsto \int_0^{\infty} a((x, y) + \eta\theta) d\eta \quad (2.32)$$

prenne ses valeurs dans l'intervalle $[3, 4]$, configuration qui exclut la possibilité de se libérer de la prise d'exponentielle en utilisant par exemple l'approximation au

premier ordre $e^{-u} \sim 1 - u$ (ce qui aurait permis d'envisager la correction entre la transformée de Radon et sa version atténuée comme une fonction bilinéaire de a et de f)⁹.

2.2.2 Radon atténuée à l'épreuve de la méthode ART

On utilisera dans cette partie la méthode de Kaczmarz pour la résolution algébrique du problème d'inversion de la transformation de Radon atténuée en présence d'une atténuation connue. On puisera donc dans les outils introduits dans la section 2.1.5.

On se limitera ici au cadre de la dimension $n = 2$; l'espace H sera ici l'espace

$$H = L^2_{\mathbb{C}}(\overline{D(0,1)}, dx dy)$$

des fonctions-source, mesurables d'énergie finie (à valeurs réelles), dans la boule unité fermée $\overline{D(0,1)}$, équipée de la mesure de Lebesgue $dx dy$.

On se donne une application continue

$$a : \overline{D(0,1)} \longmapsto [0, \infty[$$

correspondant à une répartition d'atténuation dans le milieu ambiant $\overline{D(0,1)}$. Pour $\vec{\theta} \in \mathbb{S}^1$, on définit l'opérateur

$$R_{a,\vec{\theta}} : H \longrightarrow h := L^2([-1,1], ds/\sqrt{1-s^2})$$

où $R_{a,\vec{\theta}}$ admet pour représentant la fonction définie presque partout dans $[-1,1]$ (et d'énergie finie)

$$s \longmapsto \int_{\langle (x,y), \vec{\theta} \rangle = s} \exp\left(-\int_0^{+\infty} a((x,y) + \eta \vec{\theta}^\perp) d\eta\right) I(x,y) d\lambda_{s,\vec{\theta}}(x,y),$$

$d\lambda_{\vec{\theta}}$ désignant la mesure de Lebesgue linéique sur la droite

$$L_{s,\vec{\theta}} := \{(x,y) ; \langle (x,y), \vec{\theta} \rangle = s\}.$$

L'opérateur $R_{a,\vec{\theta}}$ s'écrit comme la composée de l'opérateur continu $M_{a,\vec{\theta}}$ (de H dans lui-même) de multiplication par la fonction continue

$$A_{a,\vec{\theta}} : (x,y) \longmapsto \exp\left(-\int_0^{+\infty} a((x,y) + \eta \vec{\theta}^\perp) d\eta\right)$$

et de l'opérateur $R_{\vec{\theta}} : H \longrightarrow L^2([-1,1], \frac{ds}{\sqrt{1-s^2}})$ qui à f associe la fonction définie presque partout dans $[-1,1]$ (et d'énergie finie)

$$s \longmapsto \int_{\{\langle (x,y), \vec{\theta} \rangle = s\}} f(x,y) d\lambda_{s,\vec{\theta}}(x,y) = \int_{\mathbb{R}} f(s\vec{\theta} + \xi \vec{\theta}^\perp) d\xi.$$

L'opérateur $R_{a,\vec{\theta}} = R_{\vec{\theta}} \circ M_{a,\vec{\theta}}$ est un opérateur surjectif puisque $R_{\vec{\theta}}$ l'est; l'adjoint de $R_{a,\vec{\theta}}$ est

$$R_{a,\vec{\theta}}^* = M_{a,\vec{\theta}}^* \circ R_{\vec{\theta}}^* = M_{a,\vec{\theta}} \circ R_{\vec{\theta}}^*$$

⁹Notons d'ailleurs que si l'on s'autorisait une telle approximation, la connaissance de $R_a[f](\vec{\theta}, s)$ et de $R_a[f](\vec{\theta} + \pi, -s)$ nous ramène par addition à l'expression de $2(R[a] \times R[f])(\vec{\theta}, s)$.

puisque $M_{a,\vec{\theta}}$ est évidemment auto-adjoint. Or on a

$$R_{\vec{\theta}}^* : g \longmapsto \left((x, y) \longmapsto \left[\frac{g}{\sqrt{1-s^2}} \right]_{s=\langle (x,y), \vec{\theta} \rangle} \right).$$

On sait aussi que $R_{\vec{\theta}} \circ R_{\vec{\theta}}^* = 2 \text{Id}_h$. L'opérateur

$$R_{a,\vec{\theta}} \circ R_{a,\vec{\theta}}^* = R_{\vec{\theta}} \circ M_{a,\vec{\theta}}^2 \circ R_{\vec{\theta}}^*$$

est donc l'opérateur qui à g dans l'espace de Hilbert h associe la fonction

$$\begin{aligned} s &\longmapsto \int_{\{(x,y), \vec{\theta}\}=s} \frac{g(s)}{\sqrt{1-s^2}} (A_{a,\vec{\theta}}(x, y))^2 d\lambda_{s,\vec{\theta}}(x, y) \\ &= \frac{g(s)}{\sqrt{1-s^2}} \int_{\{(x,y), \theta\}=s} (A_{a,\vec{\theta}}(x, y))^2 d\lambda_{s,\vec{\theta}}(x, y) \\ &= \frac{g(s)}{\sqrt{1-s^2}} \int_{-\sqrt{1-s^2}}^{\sqrt{1-s^2}} \exp \left[-2 \left(\int_{\{0 \leq \eta \leq \sqrt{1-s^2}-\xi\}} a(s\vec{\theta} + (\xi + \eta)\vec{\theta}^\perp) d\eta \right) \right] d\xi. \end{aligned}$$

Cet opérateur $R_{a,\vec{\theta}} \circ R_{a,\vec{\theta}}^*$ est donc l'opérateur de multiplication par la fonction bornée positive (en fait strictement bornée inférieurement aussi)

$$\begin{aligned} B_{a,\vec{\theta}} : \\ s \in [-1, 1] &\longmapsto \frac{\int_{-\sqrt{1-s^2}}^{\sqrt{1-s^2}} \exp \left[-\text{Re} \left(\int_{\{0 \leq \eta \leq \sqrt{1-s^2}-\xi\}} a(s\theta + (\xi + \eta)\theta^\perp) d\eta \right) \right] d\xi}{\sqrt{1-s^2}} \\ &= \frac{R_{\vec{\theta}}([A_{a,\vec{\theta}}]^2)}{\sqrt{1-s^2}} \end{aligned}$$

C'est donc un opérateur inversible continu de l'espace de Hilbert h dans lui-même. Si l'on considère différents angles $\vec{\theta}_j \in [0, 2\pi[$, $j = 1, \dots, p$ et p éléments g_1, \dots, g_p de l'espace de Hilbert h , le calcul explicite des projections

$$f \in H \longmapsto P_{a,j} f \in H$$

sur les sous-espaces $\{f; R_{a,\vec{\theta}_j} f = g_j\}$, $j = 1, \dots, p$, est maintenant tout à fait explicite grâce aux formules de Kaczmarz avec relaxation (il s'agit ici de la version modifiée de la formule (2.24)) :

$$\begin{aligned} P_{a,j}^\varpi f &= f + \varpi R_{a,\vec{\theta}_j}^* \circ [R_{a,\vec{\theta}_j} \circ R_{a,\vec{\theta}_j}^*]^{-1} (g_j - R_{a,\vec{\theta}_j} [f]) \\ &= f + \varpi R_{a,\vec{\theta}_j}^* \left(\frac{g_j - R_{a,\vec{\theta}_j} [f]}{B_{a,\vec{\theta}_j}} \right) \\ &= f + \varpi A_{a,\vec{\theta}_j} \times \left(\frac{g_j - R_{a,\vec{\theta}_j} [f]}{\sqrt{1-s^2} B_{a,\vec{\theta}_j}} \right)_{s=\langle (x,y), \vec{\theta}_j \rangle} \\ &= f + \varpi A_{a,\vec{\theta}_j} \times \left(\frac{g_j - R_{a,\vec{\theta}_j} [f]}{R_{\vec{\theta}_j}([A_{a,\vec{\theta}_j}]^2)} \right)_{s=\langle (x,y), \vec{\theta}_j \rangle} \end{aligned}$$

En injectant l'expression de $P_{a,j}^\varpi$, l'algorithme itératif de Kaczmarz avec relaxation pour passer de f^k à f^{k+1} s'écrit :

$$\begin{aligned} f_0^k &= f^k \\ f_j^k &= P_{a,j}^\varpi f_{j-1}^k = f_{j-1}^k + \varpi A_{a,\vec{\theta}_j} \times \left(\frac{g_j - R_{a,\vec{\theta}_j}[f]}{R_{\theta_j}([A_{a,\vec{\theta}_j}]^2)} \right)_{s=\langle(x,y),\vec{\theta}_j\rangle}, j = 1, \dots, p \\ f^{k+1} &= f_p^k \end{aligned}$$

2.2.3 Le mécanisme d'inversion analytique

On doit à R.G. Novikov¹⁰ la première formule d'inversion pour la transformation de Radon atténuée. Un autre mécanisme d'inversion analytique (d'inspiration très voisine) a été ensuite proposé par F. Natterer¹¹ et c'est la démarche proposée par Natterer que nous allons brièvement présenter ici.

On rappelle que la transformation de Hilbert \mathbb{H} est définie (sur $L^2(\mathbb{R})$) par

$$g \in L^2(\mathbb{R}) \longmapsto \mathbb{H}[g] \in L^2(\mathbb{R}) : s \longmapsto \mathbb{H}[g](s) := \frac{1}{\pi} \int_{\mathbb{R}} \frac{g(\xi)}{s - \xi} d\xi,$$

l'intégrale de droite étant définie presque partout¹².

Décrivons le processus d'inversion proposé par Natterer, à savoir comment retrouver une fonction source $(x, y) \mapsto f(x, y)$ dont on connaît la transformée de Radon atténuée

$$R_a[f] : (\vec{\theta}, s) \in \mathbb{S}^1 \times \mathbb{R} \longmapsto \int_{\mathbb{R}} \exp \left[- \int_{\xi}^{\infty} a(s\vec{\theta} + \eta\vec{\theta}^\perp) d\eta \right] f(s\vec{\theta} + \xi\vec{\theta}^\perp) d\xi.$$

On suppose ici la carte d'atténuation connue.

1. On calcule dans un premier temps la fonction

$$h_a : (\vec{\theta}, s) \in \mathbb{S}^1 \times \mathbb{R} \mapsto \frac{(\text{Id} + i\mathbb{H})}{2} \left[\xi \mapsto R[a](\vec{\theta}, \xi) \right] (s).$$

Notons au passage que l'opérateur $\text{Id} + i\mathbb{H}$ correspond précisément à la transformation de Hilbert telle qu'elle est implémentée sous MATLAB (routine « hilbert »); ce calcul est donc un calcul qu'il est possible d'implémenter sous MATLAB en utilisant précisément la routine « hilbert ».

2. On calcule ensuite la fonction

$$\begin{aligned} F &= F_{R_a[f]} : (\vec{\theta}, s) \longmapsto e^{-h_a(\vec{\theta}, s)} \mathbb{H} \left[\xi \mapsto e^{h_a(\vec{\theta}, \xi)} R_a[f](\vec{\theta}, \xi) \right] (s) \\ &= \frac{e^{-h_a(\vec{\theta}, s)}}{\pi} \int_{\mathbb{R}} \frac{e^{h_a(\vec{\theta}, \xi)}}{s - \xi} \left[\int_{\mathbb{R}} e^{-Da(\xi\vec{\theta} + \eta\vec{\theta}^\perp, \theta^\perp)} f(\xi\vec{\theta} + \eta\vec{\theta}^\perp) d\eta \right] d\xi, \end{aligned}$$

où $D[a]$ est la *fan-beam fonction* associée à la carte d'atténuation (supposée connue) $(x, y) \mapsto a(x, y)$ (telle qu'elle est définie en (2.32)).

¹⁰An inversion formula for the attenuated Radon transform, Ark. Mat. 40, 2002, pp. 145-167.

¹¹Inversion of the attenuated Radon transform, Inverse Problems 17, 2001, pp. 113-119.

¹²Il faut prendre garde au fait que ce n'est pas tout à fait la transformée de Hilbert telle qu'elle est implémentée par exemple sous MATLAB : ici \mathbb{H} correspond, au niveau du spectre, à la multiplication par $\omega \mapsto -i\text{signe}(\omega)$ tandis que la transformée implémentée sous MATLAB correspond, au niveau du spectre, à la multiplication par $\omega \mapsto 1 + \text{signe}(\omega)$.

3. On calcule ensuite le champ de vecteurs dans \mathbb{R}^2 défini par

$$X = (x, y) \longmapsto \vec{\Phi}_{R_a[f]}(X) = \left(\int_0^{2\pi} \cos \theta e^{D[a](X, \vec{\theta}^\perp)} F(\vec{\theta}, \langle X, \vec{\theta} \rangle) d\theta, \int_0^{2\pi} \sin \theta e^{D[a](X, \vec{\theta}^\perp)} F(\vec{\theta}, \langle X, \vec{\theta} \rangle) d\theta \right).$$

4. On calcule enfin, au sens des distributions,

$$(x, y) \longmapsto \frac{1}{4\pi} \operatorname{Re} \left[\operatorname{div} \vec{\Phi}_{R_a[f]} \right] (x, y) = f(x, y)$$

pour retrouver la fonction source f (div désignant la prise de divergence).

L'implémentation sous MATLAB de ce mécanisme d'inversion se fait en profitant au maximum de la routine « inrotate » .

Chapitre 3

Processus stochastiques ; filtrage et déconvolution

3.1 Une version réaliste pour les données de la SPECT : le modèle stochastique

Dans la sous-section 2.2.1 du chapitre précédent, nous avons proposé un modèle déterministe pour décrire le nombre moyen $f(x, y) dx dy$ de photons émis depuis une zone infinitésimale $[x, x + dx] \times [y, y + dy]$ du disque unité (la dépendance en temps est ici encore oubliée) ou encore le nombre moyen $\Phi(x, y) dx dy$ de photons traversant cette même zone $[x, x + dx] \times [y, y + dy]$. Or il se trouve qu'une modélisation mathématique réaliste nous oblige à considérer que pour chaque (x, y) , nous avons en fait des « processus stochastiques continus » $(f_{x', y'})_{(x', y') \sim (x, y)}$ et $(\Phi_{x', y'})_{(x', y') \sim (x, y)}$ (tous les deux au voisinage de (x, y)) qui se trouvent être des processus Poissonniens de paramètres respectifs $f(x, y)$ et $\Phi(x, y)$ ¹.

3.1.1 Processus Poissonniens 1D

Supposons que l'on dispose d'une liste infinie d'instants « marqués » $\{\tau_\iota ; \iota \in \mathcal{I}\}$ et que l'on « jette » N de ces instants de manière aléatoire sur l'intervalle temporel $[0, T]$, les jets étant mutuellement indépendants et le résultat de chaque jet étant régi par la loi uniforme sur $[-T/2, T/2]$. La probabilité que k d'entre eux ($0 \leq k \leq N$) exactement « tombent » dans un sous-intervalle précisé $[t_1, t_2]$, avec $0 \leq t_1 \leq t_2 \leq T$, est égale à

$$p_{[t_1, t_2], N} = \binom{N}{k} \times \left(\frac{t_2 - t_1}{T}\right)^k \times \left(1 - \frac{t_2 - t_1}{T}\right)^{N-k}$$

(on retrouve une loi binomiale de $\mathcal{B}(N, (t_2 - t_1)/T)$) ; si l'on suppose N très grand (devant k) et $t_2 - t_1$ petit devant T , on peut utiliser pour $p_{[t_1, t_2], N}$ (qui, on le remarque, ne dépend que de la longueur $t_2 - t_1$ de $[t_1, t_2]$) l'approximation suivante :

$$\begin{aligned} p_{[t_1, t_2], N}(k) &\simeq e^{-N(t_2 - t_1)/T} \times \frac{1}{k!} \times \frac{N!}{(N - k)!} \times \left(\frac{N(t_2 - t_1)}{T}\right)^k \\ &\simeq e^{-N(t_2 - t_1)/T} \times \frac{1}{k!} \times \left(\frac{N(t_2 - t_1)}{T}\right)^k ; \end{aligned}$$

¹Un processus stochastique continu (ici plan) sera pour l'instant pour nous juste une collection de variables aléatoires $(X_{x', y'})_{x', y'}$ indexée par deux paramètres (x', y') parcourant un ouvert de \mathbb{R}^2 ; nous reviendrons plus en détails plus loin sur cette notion pour la préciser.

si maintenant, on suppose que N et T tendent tous les deux vers $+\infty$, mais que le nombre N/T reste constant et égal à un paramètre λ , l'approximation devient

$$p_{[t_1, t_2]}(k) = e^{-\lambda(t_2 - t_1)} \frac{(\lambda(t_2 - t_1))^k}{k!}.$$

À chaque intervalle $[t_1, t_2]$, on a ainsi associé un aléa N_{t_1, t_2} à valeurs dans \mathbb{N} dont le résultat fournit le nombre de points « tombant » sur le segment $[t_1, t_2]$. On remarque que N_{t_1, t_2} suit la même loi que $N_{0, t_2 - t_1}$. On introduit un processus stochastique sur $[0, \infty[$, c'est-à-dire une collection de variables aléatoires définies sur le même espace probabilisé et indexée par $t \in [0, \infty[$, notée $(X_t)_{t > 0}$, en posant : $X_t = N_{0, t}$.

Ce processus (les variables sont à valeurs dans \mathbb{N}) se plie à un certain nombre de règles :

1. Si l'on a des instants successifs $0 \leq t_1 \leq t_2 \leq \dots \leq t_k$, les variables

$$X_{t_1}, X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \dots$$

sont mutuellement indépendantes ;

2. $\mathbb{P}(X_{t+dt} - X_t = 1) = \lambda dt + o(dt)$;
3. $\mathbb{P}(X_{t+dt} - X_t > 1) = o(dt)$.

On caractérise ainsi (avec ces trois conditions) la notion de *processus de Poisson* ou *Processus Poissonnien* et λ est le paramètre du processus (notons que λt est la moyenne de X_t pour tout $t \geq 0$) ; en effet, on montre que X_t suit une de Poisson de paramètre λt , ceci étant en accord avec la construction du processus à partir des lois $p_{[t_1, t_2]}$.

Remarque 3.1. Un processus de Poisson n'est nullement stationnaire car on voit facilement que si $t_1 < t_2$,

$$\begin{aligned} \mathbb{E}[X_{t_1} X_{t_2}] &= \mathbb{E}[X_{t_1}^2] + \mathbb{E}[X_{t_1}(X_{t_2} - X_{t_1})] \\ &= \lambda t_1 + (\lambda t_1)^2 + (\lambda t_1) \times (\lambda(t_2 - t_1)) \\ &= \lambda t_1 + \lambda^2 t_1 t_2 \end{aligned}$$

(les rôles de t_1 et t_2 étant échangés lorsque $t_2 > t_1$) ; on ne trouve pas ici une fonction de $t_2 - t_1$!

3.1.2 Notion de processus Poissonnien 2D

Nous pouvons répéter le même type de construction en dimension 2 en supposant cette fois que nous avons un stock infini de pixels $\{(x_\iota, y_\iota) ; \iota \in \mathcal{I}\}$ que nous jettons sur un carré $[-T/2, T/2]^2$. Nous construisons une collection N_K (indexée par les pavés $[a_1, a_2] \times [b_1, b_2]$ de \mathbb{R}^2) en considérant que le résultat de l'aléa N_K est le nombre de pixels marqués « tombant » dans le pavé K . La loi de cet aléa N_K ne dépend que de la surface aire (K) du pavé et suit une loi de Poisson de paramètre $\lambda \times \text{aire } K$, où λ est la valeur (supposée constante) de N/T^2 lorsque N (le nombre de pixels jetés) et T^2 (l'aire du carré sur lequel on les jette) deviennent tous les deux grands. On définit un processus $X_{x,y}$ sur \mathbb{R}^2 en posant

$$N_{x,y} := N_{[0,|x|] \times [0,|y|]} \quad \forall x, y \in \mathbb{R}^2.$$

Ce processus vérifie alors les règles suivantes :

1. Si $(x_1, y_1), \dots, (x_n, y_n)$ sont des points tels que $0 \leq |x_1| \leq |x_2| \leq \dots \leq |x_n|$ et $0 \leq |y_1| \leq |y_2| \leq \dots \leq |y_n|$, les variables

$$X_{x_1, y_1}, X_{x_2, y_2} - X_{x_1, y_1}, X_{x_3, y_3} - X_{x_2, y_2}, \dots$$

sont mutuellement indépendantes ;

2. $\mathbb{P}(X_{x+dx, y+dy} - X_{x, y} = 1) = \lambda dx dy + o(dx dy)$;
3. $\mathbb{P}(X_{x+dx, y+dy} - X_{x, y} > 1) = o(dx dy)$.

On caractérise ainsi (avec ces trois conditions) la notion de *processus de Poisson* ou *Processus Poissonnien 2D* et λ est le paramètre du processus (notons que $\lambda|x||y|$ est la moyenne de $X_{x, y}$ pour tout $(x, y) \in \mathbb{R}$; ainsi $\lambda dx dy$ est la moyenne de l'aléa $X_{dx, dy}$ et λ peut donc être considéré comme une densité « moyenne » de pixels marqués.

Comme en 1D, il n'y a pas stationnarité d'un tel processus.

3.1.3 Le modèle stochastique de la SPECT

Dans la description du mécanisme physique de la SPECT, il faut considérer en fait que la valeur de la fonction source $(x, y) \mapsto f(x, y)$ au point (x_0, y_0) est en fait un processus Poissonnien 2D $(f_{x', y'})_{(x', y') \sim (x_0, y_0)}$ au voisinage de (x_0, y_0) , de paramètre λ précisément égal à $f(x_0, y_0)$. La fonction densité f représentant la fonction source représente le paramètre d'un bruit poissonnien d'émission. Le processus stochastique $(X_{x, y})_{(x, y) \in D(0, 1)}$ n'est plus un processus Poissonnien de paramètre constant, mais on a, pour tout sous-ensemble mesurable A du disque :

$$\int_A \mathbb{E}[f_{x, y}] dx dy = \int \int_A f(x, y) dx dy,$$

ce qui assure $f(x, y) = \mathbb{E}[f_{x, y}]$ pour presque tout (x, y) dans le disque. Ce que l'on reconstitue *via* le procédé tomographique et la fonction espérance (ou moyenne) du processus source d'émission

$$(x, y) \mapsto f_{x, y}$$

qui se présente localement comme un processus Poissonnien. La même remarque vaut pour le sens à donner à

$$(x, y) \mapsto \Phi(x, y) \stackrel{\text{P-P}}{=} \mathbb{E}[\Phi_{x, y}],$$

où $(x, y) \mapsto \Phi_{x, y}$ est un processus (localement) Poissonnien.

Remarque 3.2. Si l'on imagine le disque $D(0, 1)$ comme un support discrétisé, il faut voir $f(x_i, y_i)$ (lorsque (x_i, y_i) est un pixel) comme le paramètre $\lambda(x_i, y_i)$ (*i.e* la valeur moyenne) d'une variable aléatoire f_{x_i, y_i} à valeurs dans \mathbb{N} et suivant une loi de Poisson :

$$\mathbb{P}(f_{x_i, y_i} = k) = e^{-f(x_i, y_i)} \frac{(f(x_i, y_i))^k}{k!}.$$

3.2 Processus stationnaires : le filtrage de Wiener

3.2.1 Retour sur la notion de processus ; stationnarité

La mesure de tout phénomène physique, qu'il s'agisse d'un phénomène suivi au cours d'une évolution continue du temps (*resp.* de l'espace) ou suivant un maillage

temporel (*resp.* spatial), est entachée d'erreurs de calculs ou de bruit inhérents à l'instrumentation ou à l'espace ambiant dans lequel est effectué l'enregistrement ; ces avatars échappent à une modélisation déterministe, d'où la nécessité d'introduire parallèlement (ou en contrepoint) au point de vue déterministe, dans les problèmes inhérents à la tomographie comme ailleurs, un point de vue stochastique. Cette nécessité se trouve confortée par le fait que bien souvent les signaux que l'on souhaite enregistrer ou mesurer correspondent eux même à des modèles probabilistes (tels des bruits poissonniens d'émission en imagerie médicale ou échographie, voir la section précédente). La notion de « signal stochastique » est modélisée en mathématiques sous le concept de processus. On parle de « processus discret » (ou encore de « série chronologique ») lorsque le temps ou l'espace (par exemple \mathbb{R}^2 ou \mathbb{R}^3 en traitement d'images ou en tomographie) sont discrétisés, de « processus continu » sinon. Dans ce cours, nous nous placerons dans la mesure du possible dans le cadre des processus discrets, ce afin de ne pas avoir à mettre en œuvre la théorie de l'intégration stochastique élaborée par I. Ito ².

Un processus aléatoire discret est une suite $(X_k)_{k \in \mathbb{Z}^n}$ indexée par \mathbb{Z}^n (les cas les plus importants étant $n = 1$ et $n = 2$ correspondant respectivement au cadre des signaux et des images) de variables aléatoires réelles ou complexes toutes définies sur le même espace probabilisé $(\mathcal{O}, \mathcal{T}, \mathbb{P})$. Si toutes les variables aléatoires X_k , $k \in \mathbb{Z}^n$ ont des moments d'ordre deux finis ($\mathbb{E}[|X_k|^2] < \infty$ pour tout $k \in \mathbb{Z}^n$), le processus $(X_k)_{k \in \mathbb{Z}^n}$ est dit processus L^2 . En théorie, un processus aléatoire continu est une collection $(X_{x_1, \dots, x_n})_{(x_1, \dots, x_n) \in \mathbb{R}^n}$ (indexée donc cette fois par \mathbb{R}^n) de variables aléatoires toutes définies sur le même espace probabilisé ; le processus sera dit L^2 si X_{x_1, \dots, x_n} est dans L^2 tout $x \in \mathbb{R}^n$.

La notion de stationnarité pour un processus est une notion essentielle, traduisant le fait que la « loi » du processus n'est pas affectée par une translation dans le temps ou dans l'espace : un processus aléatoire discret $(X_k)_{k \in \mathbb{Z}}$ sur un espace probabilisé $(\mathcal{O}, \mathcal{T}, \mathbb{P})$ est dit stationnaire (au sens fort) si et seulement si pour tout $p \in \mathbb{N}$, pour tout k_1, \dots, k_p dans \mathbb{Z} , la loi conjointe du vecteur de variables aléatoires $(X_{k_1+l}, \dots, X_{k_p+l})$ est la même que celle de $(X_{k_1}, \dots, X_{k_p})$ pour tout $l \in \mathbb{Z}$. Trop contraignante à manier, cette notion de stationnarité est souvent remplacée par une notion plus faible, ce lorsque l'on a affaire à des processus discrets L^2 . Pour les processus continus, la notion forte de stationnarité est identique et, ici encore, elle se doit d'être remplacée par une notion plus faible dans le cas des processus continus L^2 .

Définition 3.1 *Un processus discret L^2 , $X = (X_k)_{k \in \mathbb{Z}^n}$, est dit stationnaire (au second ordre) si et seulement si*

- d'une part l'espérance $\mathbb{E}[X_k] = \int_{\mathcal{O}} X_k d\mathbb{P}$ est indépendante de $k \in \mathbb{Z}^n$;
- d'autre part, la fonction

$$r_{X,X} : (k_1, k_2) \in \mathbb{Z}^n \times \mathbb{Z}^n \longmapsto \mathbb{E}[X_{k_1} \overline{X_{k_2}}] = \int_{\Omega} X_{k_1}(\omega) \overline{X_{k_2}(\omega)} d\mathbb{P}$$

s'exprime sous la forme $(k_1, k_2) \mapsto s_{X,X}(k_1 - k_2)$, la fonction $s_{X,X} : \mathbb{Z}^n \rightarrow \mathbb{C}$ étant dite dans ce cas fonction d'autocorrélation du processus discret $(X_k)_{k \in \mathbb{Z}}$ stationnaire au second ordre.

²Cependant, le modèle de bruit poissonnien que nous avons introduit dans la section précédente est un modèle de processus continu, et on comprendra donc qu'il soit difficile dans un cours de tomographie de contourner l'obstacle que constitue le maniement des processus continus et, par là même de l'intégrale stochastique.

Un processus continu $(X_{x_1, \dots, x_n})_{(x_1, \dots, x_n) \in \mathbb{R}^n}$ supposé L^2 est stationnaire au second ordre si d'une part $x \mapsto \mathbb{E}[X_x]$ est égale μ -presque partout à une constante et si la fonction d'autocorrélation $R_{X,X} : (x, y) \mapsto \mathbb{E}[X_x \overline{X_y}]$ est une fonction localement intégrable égale (au sens des mesures) à $(x, y) \mapsto S_{X,X}(x - y)$, où $S_{X,X}$ est une mesure sur \mathbb{R}^n dite mesure d'autocorrélation du processus continu L^2 stationnaire au second ordre.

Remarque 3.3. Si X et Y sont deux processus discrets L^2 stationnaires au second ordre, la fonction $(k_1, k_2) \mapsto \mathbb{E}[X_{k_1} \overline{Y_{k_2}}]$ s'écrit sous la forme $(k_1, k_2) \mapsto s_{X,Y}(k_1 - k_2)$, la fonction $s_{X,Y}$ étant appelée fonction d'intercorrélacion des processus X et Y .

Définition 3.2 Si $(X_k)_{k \in \mathbb{Z}^n}$ est un processus discret L^2 stationnaire au second ordre, de fonction d'autocorrélation $s_{X,X}$ telle que $(s_{X,X}(k))_{k \in \mathbb{Z}^n} \in l^2_{\mathbb{C}}(\mathbb{Z}^n)$, on appelle densité spectrale $\text{psd}[X]$ de puissance de X la transformée de Fourier de la suite $(s_{X,X}(k))_k$, c'est-à-dire la fonction 2π -périodique en toutes les variables obtenue comme limite dans $L^2_{\mathbb{C}}(\mathbb{R}/(2\pi\mathbb{Z}))$ de la suite de polynômes trigonométriques

$$\left(\omega \mapsto \sum_{k \in [-N, N]^n} s_{X,X}(k) e^{-i\langle k, \omega \rangle} \right)_{N \in \mathbb{N}}.$$

Si $(X_x)_{x \in \mathbb{R}^n}$ est un processus continu L^2 stationnaire au second ordre, de mesure d'autocorrélation $S_{X,X}$ telle que $\int_{\mathbb{R}^n} |S_{X,X}| < +\infty$, on appelle densité spectrale de puissance $\text{PSD}[X]$ la fonction bornée

$$\omega \in \mathbb{R}^n \mapsto \langle S_{X,X}, e^{-i\langle \omega, \cdot \rangle} \rangle,$$

c'est-à-dire le spectre de la mesure d'autocorrélation du processus.

Proposition 3.1 Soit $X = (X_k)_{k \in \mathbb{Z}^n}$ un processus L^2 stationnaire au second ordre, comme dans la définition 3.2. La densité spectrale de puissance de X admet un représentant positif ou nul sur \mathbb{R}^n tout entier; ceci reste vrai pour un processus continu stationnaire au second ordre.

Preuve. Pour simplifier, on se limitera pour la preuve au cas $n = 1$ (et aux processus discrets). Il suffit juste de constater en invoquant la formule de Plancherel que pour tout polynôme trigonométrique partout positif $Q : \omega \mapsto |a_0 + a_1 e^{i\omega} + \dots + a_M e^{iM\omega}|^2$, on a

$$\begin{aligned} \int_{[0, 2\pi]} \text{psd}[X](\omega) Q(\omega) d\omega &= 2\pi \sum_{k \in \mathbb{Z}} s_{X,X}(k) \left(\sum_{l_1 - l_2 = k} \overline{a_{l_1}} a_{l_2} \right) \\ &= 2\pi \sum_{k \in \mathbb{Z}} \sum_{l_1 - l_2 = k} \overline{a_{l_1}} a_{l_2} \mathbb{E}[X_{l_1} \overline{X_{l_2}}] = \mathbb{E} \left[\left| \sum_{l=0}^M \overline{a_l} X_l \right|^2 \right] \geq 0. \end{aligned}$$

La proposition est démontrée. \diamond

On dit qu'un processus stochastique L^2 -discret stationnaire au second ordre est *ergodique au sens du calcul de moyenne* si et seulement si la suite de variables aléatoires

$$\eta_N := \frac{1}{2N-1} \sum_{l=-(N-1)}^{N-1} X_l, \quad N = 1, 2, \dots$$

converge presque partout vers l'aléa constant $\eta = \mathbb{E}[X_0]$. Le processus est dit *ergodique au sens du calcul de corrélation* si et seulement si pour chaque entier $k \in \mathbb{Z}$, le processus $\text{cov}_k X$ défini par

$$[\text{cov}_k X]_l := X_{l+k} \overline{X}_l, \quad l \in \mathbb{Z},$$

est ergodique au sens de la moyenne. Lorsqu'un processus L^2 -discret et stationnaire au second ordre X est ergodique au sens de sa moyenne, on peut approcher cette moyenne $\mathbb{E}[X] = \eta$ en calculant l'aléa η_N en un point particulier de l'espace d'évènements, à savoir le point qui correspond à la version précisée du processus dont on dispose, ce lorsque N est assez grand ; on dispose ainsi d'une estimation de l'espérance mathématique de X . Si le processus est en même temps ergodique au sens de la corrélation, on peut calculer aussi sa fonction d'autocorrélation en calculant, pour tout k dans \mathbb{Z} , et pour N assez grand, la valeur de l'aléa

$$\frac{1}{2N-1} \sum_{l=-(N-1)}^{N-1} X_{l+k} \overline{X}_l$$

au point de l'espace d'évènements qui correspond à la version précisée du processus dont on dispose et que l'on traite. C'est par le biais de ces estimateurs que seront calculés l'espérance d'un processus L^2 -discret stationnaire à l'ordre 2 et sa fonction d'autocorrélation, donc sa densité spectrale de puissance. Il est par conséquent intéressant de disposer d'un critère permettant d'assurer l'ergodicité en moyenne. Nous avons pour cela la proposition suivante.

Proposition 3.2 *Si X est un processus L^2 -discret stationnaire au second ordre tel que $\lim_{|k| \rightarrow +\infty} s_{X,X}(k) = 0$, alors X est ergodique au sens de la moyenne.*

Preuve. D'après l'inégalité de Markov, on a, pour tout $\epsilon > 0$ et pour tout $N \in \mathbb{N}^*$, si $\eta := \mathbb{E}[X_0]$,

$$\begin{aligned} \mathbb{P}\left(\left|\frac{1}{2N-1} \sum_{k=-(N-1)}^{N-1} (X_k - \eta)\right| \geq \epsilon\right) &\leq \frac{\sum_{k=-(N-1)}^{N-1} \sum_{l=-(N-1)}^{N-1} (E[X_k \overline{X}_l] - |\eta|^2)}{(2N-1)^2 \epsilon^2} \\ &\leq \frac{\sum_{k=-2(N-1)}^{2(N-1)} s_{X,X}(k) \left(1 - \frac{|k|}{2N-1}\right)}{(2N-1)^2 \epsilon^2} \end{aligned}$$

En utilisant la convergence vers 0 au sens de Cesaro de la suite $(S_{X,X}(k))_k$, on voit que le second membre tend vers 0 lorsque N tend vers l'infini et que par conséquent la suite $(\eta_N)_N$ converge stochastiquement vers l'aléa presque partout égal à $\eta = \mathbb{E}[X_0]$; cette convergence stochastique implique l'ergodicité au sens de la moyenne. \diamond

Remarque 3.4. Les diverses notions introduites ici (ergodicité au sens de la moyenne, au sens de la corrélation), seraient transposables au cas des processus 1D continus L^2 stationnaires au second ordre, mais il faudrait introduire l'intégrale stochastique, ce que nous ne ferons pas ici. Notons que c'est l'hypothèse d'ergodicité au sens de la corrélation que l'on fait dans la pratique pour calculer la densité spectrale de puissance d'un processus discret supposé stationnaire au second ordre.

Le calcul de la psd *via* la méthode de Welch

Supposons que $X = (X_k)_{k \in \mathbb{Z}}$ soit un processus discret L^2 -stationnaire au second ordre au sens de la définition 3.1. Deux problèmes apparaissent lorsque l'on prétend estimer la densité spectrale de puissance de ce signal stochastique : d'une part, l'observateur n'a en général à sa disposition qu'un seul échantillon de ce processus, à savoir le signal observé (nous en reparlerons plus loin), d'autre part, il n'est envisageable que d'estimer des versions tronquées

$$\omega \longmapsto \text{psd}[X; N](\omega) := \sum_0^{N-1} s_{X,X}(k) e^{-ik\omega}, \quad N \gg 1,$$

de la densité spectrale de puissance $\text{psd}[X]$. Si l'on suppose X ergodique au sens de la corrélation (voir la sous-section précédente), on peut, à partir d'un échantillon $(x_k)_k$ du processus, une fois N choisi très grand, « estimer » $s_{X,X}(k)$ pour $k = 0, \dots, N-1$ comme suit :

$$s_{X,X}(k) \simeq s_{X,X}^{(b)}(k) := \frac{1}{N-k} \sum_{l=0}^{N-k-1} x_{l+k} \bar{x}_l. \quad (3.1)$$

ce premier calcul approché (dit estimation « biaisée », d'où le (b) en exposant), s'il est intéressant lorsque $|k|$ est petit comparé à N , perd une part de son sens lorsque $|k|$ se rapproche de N (car alors la corrélation n'est pas calculée à partir d'un nombre d'échantillons suffisant pour que l'on puisse justifier d'après l'hypothèse d'ergodicité que l'on a affaire ici à un estimateur approximation raisonnable de $s_{X,X}(k)$). On peut lui préférer, tenant compte de la remarque ci-dessus, l'approximation (dite « non biaisée », d'où le (nb) en exposant) de $s_{X,X}(k)$ pour $k = 0, \dots, N-1$, par

$$s_{X,X}(k) \simeq s_{X,X}^{(nb)}(k) := \frac{1}{N} \sum_{l=0}^{N-k-1} x_{l+k} \bar{x}_l. \quad (3.2)$$

Cette dernière approximation est dite estimation non biaisée. On peut maintenant, reprenant l'une ou l'autre de nos estimations (3.1) ou (3.2), approcher la densité spectrale de puissance $\text{psd}[X]$ échantillonnée aux points $\omega_j = 2\pi j/N$, pour $j = 0, \dots, N-1$ de deux manières, en prenant les valeurs en ces points de l'une ou l'autre des fonctions :

$$\begin{aligned} \text{psd}[X] \simeq \text{psd}[X; N]^{(b)} : \omega &\longmapsto \sum_{l=0}^{N-1} \bar{x}_l e^{il\omega} \left(\sum_{k=0}^{N-l-1} \frac{x_{l+k} e^{-i(l+k)\omega}}{N-k} \right) \\ \text{psd}[X] \simeq \text{psd}[X; N]^{(nb)} : \omega &\longmapsto \frac{1}{N} \sum_{l=0}^{N-1} \bar{x}_l e^{il\omega} \left(\sum_{k=0}^{N-1-l} x_{l+k} e^{-i(l+k)\omega} \right) \\ &= \frac{\left| \sum_{k=0}^{N-1} x_k e^{-ik\omega} \right|^2}{N}. \end{aligned} \quad (3.3)$$

La méthode proposée par P.D. Welch dans les années 1970 est aujourd'hui la méthode la plus couramment utilisée pour fournir un estimateur la densité spectrale

d'un signal stochastique stationnaire (à l'ordre 2), étant entendu que l'on ne dispose que d'une seule observation $(x_k)_k$ sur un très long pas de temps. C'est en segmentant ce signal sur des plages temporelles distinctes (mais de même longueur) que l'on générera divers échantillons de ce signal stochastique. Voici ci-dessous le synopsis de cette méthode, inspirée de la formule approchée (3.3).

1. Le champ temporel d'étude $\{0, \dots, N - 1\}$ est subdivisé en q fenêtres (qui peuvent éventuellement se chevaucher avec un certain ratio de recouvrement) de longueur M .
2. On se donne un signal digital positif $(h(0), \dots, h(M - 1))$ « pondérateur » en privilégiant la zone centrale de la fenêtre aux extrémités.
3. Pour chaque fenêtre temporelle $\{k_0, \dots, k_0 + M - 1\}$ (initiée ici à k_0), on utilise l'algorithme de \widetilde{M} -DFT pour calculer en $\widetilde{M} \geq M$ points régulièrement espacés sur $[-\pi, \pi]$ la valeur prise par la fonction

$$\omega \mapsto \frac{\left| \sum_{k=0}^{M-1} h(k)x_{k_0+k}e^{-ik\omega} \right|^2}{|h(0)|^2 + \dots + |h(M-1)|^2}.$$

4. On moyennise enfin tous les résultats obtenus avec les diverses fenêtres.

C'est cet algorithme que l'on retrouve sous l'environnement MATLAB avec la routine `spectrum.welch(Window, Segment, Overlap)` combinée avec la routine `psd` :

```
Fs=1000;
t=0:1/Fs:.3;
x=cos(2*pi*t*200)+randn(size(t));
Hs=spectrum.welch;
psd(Hs,x,'Fs',Fs)
```

Bruits et filtrage de processus discrets

Dans cette section, nous transposons l'action des filtres digitaux du cadre déterministe au cadre stochastique. On rappelle que le filtre digital stable de réponse impulsionnelle $(h(k))_k$, où $\sum_k |h(k)| < \infty$ est l'opérateur $l^2(\mathbb{Z}) \rightarrow l^2(\mathbb{Z})$ de convolution avec la suite $(h(k))_k$, i.e

$$s(k) = \sum_{l \in \mathbb{Z}} h(l)e(k-l) = \sum_l h(k-l)e_l.$$

Proposition 3.3 *Soit X un processus discret L^2 sur un espace probabilisé $(\mathcal{O}, \mathcal{T}, \mathbb{P})$, stationnaire au second ordre, tel que la suite $(s_{X,X}(k))_{k \in \mathbb{Z}}$ soit dans $l^2_{\mathbb{C}}(\mathbb{Z})$. Si \mathcal{L} est un filtre digital stable de réponse impulsionnelle $(h(k))_{k \in \mathbb{Z}}$, on peut définir un autre processus discret $Y = \mathcal{L}[X]$ (lui aussi L^2 et stationnaire au second ordre) en posant*

$$Y_k : \stackrel{L^2(\mathcal{O}, \mathcal{T}, \mathbb{P})}{=} \lim_{N \rightarrow +\infty} \sum_{-N}^N h(l)X_{k-l}, \quad k \in \mathbb{Z}.$$

De plus, on a dans $L^2_{\mathbb{C}}(\mathbb{R}/(2\pi\mathbb{Z}))$ la relation suivante entre les densités spectrales de puissance des processus X et $Y = \mathcal{L}[X]$.

$$\text{psd}[Y](\omega) = \left| \sum_{k=-\infty}^{+\infty} h(k)e^{-i\omega k} \right|^2 \times \text{psd}[X](\omega) \quad (3.4)$$

Preuve. Si F est un sous-ensemble fini de \mathbb{Z} , on a immédiatement, pour tout $k \in \mathbb{Z}$,

$$\mathbb{E} \left[\left| \sum_{l \in F} h(l) X_{k-l} \right|^2 \right] = \left| \sum_{l_1 \in F} \sum_{l_2 \in F} h(l_1) \overline{h(l_2)} s_{X,X}(l_1 - l_2) \right| \leq \|h\|_1^2 \times \|s_{X,X}\|_\infty.$$

Le critère de Cauchy s'applique donc et implique la convergence dans $L^2_{\mathbb{C}}(\mathcal{O}, \mathcal{T}, \mathbb{P})$ de la série bilatère $\sum_{l \in \mathbb{Z}} h(l) X_{k-l}$ qui était censée définir formellement Y_k . Le processus $(Y_k)_k$ ainsi défini est donc bien un processus discret L^2 . On constate que $\mathbb{E}[Y_k] = \mathbb{E}[X_0] \times \sum_l h(l)$ ne dépend pas de k et que, pour $k_1, k_2 \in \mathbb{Z}$, $\mathbb{E}[Y_{k_1} \overline{Y_{k_2}}]$ vaut

$$\begin{aligned} & \lim_{N \rightarrow \infty} \left(\sum_{l_1 = -N}^N \sum_{l_2 = -N}^N h(l_1) \overline{h(l_2)} s_{X,X}((k_1 - k_2) - (l_1 - l_2)) \right) \\ &= \sum_{l \in \mathbb{Z}} H(k_1 - k_2 - l) s_{X,X}(l) \end{aligned}$$

(où $H(k) := \sum_{l \in \mathbb{Z}} h(k+l) \overline{h(l)}$ pour $k \in \mathbb{Z}$) ne dépend que de $k_1 - k_2$. Le processus $(Y_k)_k$ est donc bien lui aussi stationnaire à l'ordre 2 et sa fonction d'autocorrélation $s_{Y,Y}$ s'obtient comme la réponse à l'entrée $s_{X,X}$ du filtre digital (lui aussi stable) de réponse impulsionnelle $(H(k))_k$. En prenant les transformées de Fourier des signaux digitaux $s_{X,X}$ et $s_{Y,Y}$, on en déduit bien la formule (3.4). \diamond

L'observation d'un processus discret L^2 (supposé stationnaire à l'ordre 2) X se trouve dans la pratique entaché d'un « bruit » ; ce bruit additif (dont peuvent être responsables soit le phénomène observé lui-même, soit le dispositif de mesure du phénomène susceptible d'impliquer une erreur) peut être modélisé par un autre processus discret L^2 noté B . Les modèles « idéaux » les plus simples sont les bruits blancs.

Définition 3.3 *On appelle bruit blanc discret tout processus discret L^2 stationnaire au second ordre B tel que $\mathbb{E}[B_0] = 0$, que $s_{B,B}(0) = \sigma^2 > 0$ et $s_{B,B}(k) = 0$ sinon, donc ayant pour densité spectrale de puissance la fonction identiquement égale à σ^2 ; la quantité σ^2 est dite variance du bruit blanc.*

C'est à Norbert Wiener ³ que l'on doit la proposition suivante, d'un intérêt pratique majeur, soutendant le principe de ce que l'on appelle le filtrage de Wiener.

Proposition 3.4 *Soient X et Y deux processus discrets L^2 sur un même espace probabilisé $(\mathcal{O}, \mathcal{T}, \mathbb{P})$, tous les deux stationnaires au second ordre, tels que les trois suites $(s_{X,X}(k))_k$, $(s_{Y,Y}(k))_k$, $(s_{X,Y}(k))_k$ soient dans $l^2_{\mathbb{C}}(\mathbb{Z})$ et qu'il existe un filtre digital stable \mathcal{L} de manière à ce que l'on ait dans $L^2_{\mathbb{C}}(\mathbb{R}/(2\pi\mathbb{Z}))$ la relation*

$$\widehat{s_{X,Y}} = \widehat{\mathcal{L}} \times \text{psd}[X + Y]. \quad (3.5)$$

Le filtre \mathcal{L} est, de tous les filtres stables digitaux, celui pour lequel l'erreur quadratique $\mathbb{E}[|(\mathcal{L}[X + Y])_k - X_k|^2]$ (indépendante de k du fait de la stationnarité à l'ordre 2 de X et Y) est minimale ; c'est donc le meilleur filtre dont on puisse espérer disposer pour identifier la présence du processus X lorsque l'on dispose de l'observation du processus « perturbé » $\tilde{X} = X + Y$.

³Norbert Wiener, 1864-1964, mathématicien appliqué américain, l'un des pionniers de l'analyse harmonique appliquée moderne.

Preuve. Ce résultat facile repose sur le principe des moindres carrés. Si \mathcal{L} est un filtre stable quelconque, on sait (proposition 3.3) que le processus $\mathcal{L}[X+B]$ est encore un processus discret L^2 , stationnaire à l'ordre 2, ce qui implique que les nombres positifs $\mathbb{E}[|(\mathcal{L}[X+Y])_k - X_k|^2]$ ne dépendent pas de k et sont égaux à une « erreur » $\epsilon(\mathcal{L})$. Comme $(\mathcal{L}[X+Y])_k = \sum_l h(l)X_{k-l}$, un filtre stable \mathcal{L}_{opt} tel que cette erreur $\epsilon_{\mathcal{L}}$ soit minimale serait le filtre dont la réponse impulsionnelle des coefficients $(h_{\text{opt}}(k))_k$ est telle que pour un entier k arbitraire dans \mathbb{Z} (c'est vrai alors pour tous), l'aléa $\sum_l h_{\text{opt}}(l)(X_{k-l} + Y_{k-l})$ représente (dans $L^2_{\mathbb{C}}(\mathcal{O}, \mathcal{T}, \mathbb{P})$) la projection orthogonale de X_k sur le sous espace engendré par les $X_{\nu} + B_{\nu}$, $\nu \in \mathbb{Z}$. Ceci revient donc à écrire les conditions d'orthogonalité

$$\mathbb{E}\left[\left(X_k - \sum_{l=-\infty}^{\infty} h_{\text{opt}}(l)X_{k-l}\right)\left(\overline{X_{\nu}} + \overline{Y_{\nu}}\right)\right] = 0, \quad \nu \in \mathbb{Z}, \quad k \in \mathbb{Z}.$$

Une fois éliminée la redondance de cette liste de conditions (résultant de la stationnarité au sens faible des processus), il reste que ce système se réduit à la liste de conditions

$$s_{X, X+Y}(k) := \sum_{l=-\infty}^{\infty} h_{\text{opt}}(l)s_{X+Y, X+Y}(k-l), \quad k \in \mathbb{Z}.$$

En prenant les transformées de Fourier des suites (indexées par k) figurant aux deux membres, on trouve bien que les transformées de Fourier des deux filtres \mathcal{L} et \mathcal{L}_{opt} sont égales, d'où le résultat. \diamond

Exemple. Si X est un processus discret L^2 stationnaire à l'ordre 2 (d'autocorrélation dans $l^2(\mathbb{Z})$) et B un bruit blanc de variance σ^2 indépendant de X , l'autocorrélation de $X+B$ est aussi dans $l^2(\mathbb{Z})$ et l'on a $\widehat{s_{X,B}} = \text{psd}[X]$ ainsi que $\text{psd}[X+B] = \text{psd}[X] + \sigma^2$. Si la fonction $\text{psd}[X]/(\text{psd}[X] + \sigma^2)$ est le spectre d'un filtre stable \mathcal{L} , ce filtre \mathcal{L} est le filtre optimal de la proposition 3.4.

Remarque 3.5. Le problème que pose le *design* du filtre de Wiener, construit pour filtrer un processus donné $(X_k)_{k \in \mathbb{Z}}$ donné et entaché d'un bruit blanc (comme dans l'exemple précédent) est qu'il est nécessaire de connaître le processus $(X_k)_{k \in \mathbb{Z}}$ au travers de sa psd pour construire le filtre ! L'autre défaut est évidemment l'hypothèse préalable de stationnarité qu'il est nécessaire de faire sur le processus $(X_k)_{k \in \mathbb{Z}}$. En ce sens, le filtrage de Wiener est plus 'destiné' à l'identification du processus $(X_k)_{k \in \mathbb{Z}}$ au sein d'un environnement bruité (le processus est-il réellement présent ou non ?) qu'à sa reconstruction. D'autres méthodes, inspirées de l'idée de Wiener, mais transposées cette fois au problème de la *déconvolution* joueront, elles, un rôle au niveau de la reconstruction proprement dite. Nous allons les voir dans la sous section suivante.

3.2.2 Déconvolution linéaire ; « pseudo-filtrage » de Wiener

Dans cette section, nous nous donnons un espace probabilisé $L^2(\mathcal{O}, \mathcal{T}, \mathbb{P})$ et un processus continu mesurable $X = (X_t)_t$, L^2 et stationnaire au second ordre au sens de la définition 3.1 (volet continu), tel que la mesure d'autocorrélation $S_{X,X}$ soit de masse totale finie et que l'on puisse donc en définir la densité spectrale de puissance $\text{PSD}[X]$

$$\text{PSD}[X] : \omega \longmapsto \langle R_{X,X}(t), e^{-i\omega t} \rangle = \int_{\mathbb{R}} e^{-i\omega t} dR_{X,X}(t).$$

On pourrait d'ailleurs tout aussi bien envisager des processus continus stationnaires au second ordre $2D$; pour simplifier ici, nous nous limiterons au cas $1D$.

Nous supposons que notre processus (supposé *a priori inconnu*) passe *simultanément* à l'intérieur de plusieurs filtres analogiques $\mathcal{L}_1, \dots, \mathcal{L}_N$, dits *filtres convoluteurs*, dont nous supposerons ici que les réponses impulsionnelles h_1, \dots, h_N , sont des distributions à support compact (nous avons en tête ici de tels modèles en rapport avec les exemples de filtres-masques que nous avons en vue en tomographie). Les réponses de ces systèmes à l'entrée X sont les N processus continus mesurables $Y^{(1)}, \dots, Y^{(N)}$, où

$$Y_t^{(k)} = [h_k * X]_t = [\mathcal{L}_k[X]]_t, \quad k = 1, \dots, N, \quad t \in \mathbb{R}.$$

Il ressort des hypothèses faites tant sur le processus X que sur les filtres \mathcal{L}_j que ces processus sont encore des processus continus mesurables L^2 stationnaires au second ordre, tels que

$$\text{PSD}[Y^{(k)}](\omega) = |\hat{h}_k(\omega)|^2 \text{PSD}[X](\omega) \quad \forall k = 1, \dots, N.$$

Nous noterons $\vec{\mathcal{L}}$ la matrice colonne d'opérateurs dont les composantes sont les opérateurs $\mathcal{L}_1, \dots, \mathcal{L}_N$. On suppose enfin que le vecteur colonne \vec{Y} constitué des processus $Y^{(k)}$ est enregistré entaché d'un bruit \vec{B} (environnement, appareillage de mesure), vecteur colonne de processus $B^{(k)}$, $k = 1, \dots, N$, tous indépendants de X , tous de moyenne nulle, et tels que la matrice de mesures

$$(t_1, t_2) \mapsto R_{B,B}(t_1, t_2) = \left(\mathbb{E} \left[B_{t_1}^{(k)} \overline{B_{t_2}^{(l)}} \right] \right)_{1 \leq k, l \leq N}$$

soit la matrice $\delta(t_1 - t_2) \sigma^2 I_N$, où I_N est la matrice identité $N \times N$.

Le problème de la déconvolution (à instrumentation $\mathcal{L}_1, \dots, \mathcal{L}_N$ connue) consiste en la conception de N appareils $D^{(1)}, \dots, D^{(N)}$ (dits filtres *deconvoluteurs*) que l'on supposera correspondre à N filtres analogiques stationnaires, tels que l'erreur

$$\mathbb{E} \left[\left| X_t - \sum_{l=1}^N [D^{(l)} * (Y^{(l)} + B^{(l)})]_t \right|^2 \right]$$

(erreur qui est égale presque partout à une constante indépendante de t du fait de la stationnarité de X , des $Y^{(l)}$ et des $B^{(l)}$) soit minimale parmi toutes les erreurs générées par tous les choix de systèmes de déconvoluteurs stationnaires donnés.

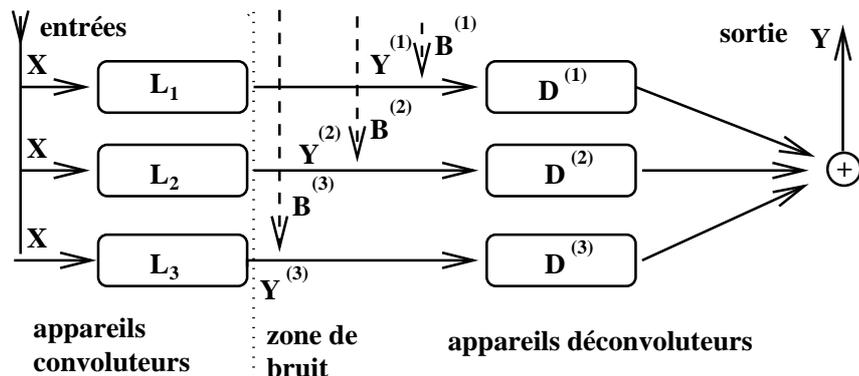


FIG. 3.1 – Appareillage convolution / déconvolution

Ce mécanisme, que nous avons illustré sur la figure 3.1, est tout à fait inspiré des développements de la sous-section précédente et l'on a la

Proposition 3.5 *Sous toutes les hypothèses faites ci-dessus, le choix optimum de filtres analogiques stationnaires déconvoluteurs pour le problème de déconvolution correspondant aux appareils $\mathcal{L}_1, \dots, \mathcal{L}_N$, est donné par*

$$\widehat{D}^{(k)} = \frac{\overline{\widehat{h}_k} \text{PSD}[X]}{\left(\sum_{l=1}^N |\widehat{h}_l|^2 \right) \text{PSD}[X] + \sigma^2}, \quad k = 1, \dots, N. \quad (3.6)$$

Preuve. On note \vec{D}^* la matrice ligne $\vec{D}^* = (D^{(1)}, \dots, D^{(N)})$ constituée d'un choix de déconvoluteurs analogiques stationnaires potentiels. La valeur de l'erreur (dans $L^2(\mathcal{O}, \mathcal{T}, \mathbb{P})$) entre les variables X_t (valeur du processus d'entrée à l'instant t) et Y_t (valeur du processus de sortie à l'instant t), se trouve être, d'après les hypothèses de stationnarité au sens faible portant sur X et \vec{B} , presque partout égale à une constante $e_{\vec{D}}$. Du fait de l'indépendance des processus X et \vec{B} , on a

$$e_{\vec{D}} = \mathbb{E} \left[\left| \vec{D}^* \vec{\mathcal{L}}[X]_t - X_t \right|^2 \right] + \mathbb{E} \left[\left| \vec{D}^* \vec{B}_t \right|^2 \right] \quad \text{presque partout.} \quad (3.7)$$

En utilisant la formule de Plancherel et la règle de calcul selon laquelle la densité spectrale de puissance d'un processus se trouve transformée par le filtrage, nous pouvons transformer l'identité (3.7) en l'identité (où t n'apparaît plus)

$$e_{\vec{D}} = \frac{1}{2\pi} \left(\int_{\mathbb{R}} \left| \widehat{\vec{D}^*}(\omega) \widehat{\vec{\mathcal{L}}}(\omega) - 1 \right|^2 \text{PSD}[X](\omega) d\omega + \int_{\mathbb{R}} \widehat{\vec{D}^*}(\omega) \text{PSD}[\vec{B}](\omega) \overline{\widehat{\vec{D}^*}(\omega)} d\omega \right) \quad (3.8)$$

où $\text{PSD}[\vec{B}](\omega)$ est la matrice $N \times N$ des fonctions 2π périodiques $\text{PSD}[B^{(k)}]$, $1 \leq k, l \leq N$. Sous les hypothèses faites ici, cette fonction à valeurs matricielles est en fait constante et l'on a

$$\text{PSD}[\vec{B}](\omega) \equiv \sigma^2 I_N.$$

On a donc, par conséquent

$$e_{\vec{D}} = \frac{1}{2\pi} \int_{\mathbb{R}} \left[\left| \widehat{\vec{D}^*}(\omega) \widehat{\vec{\mathcal{L}}}(\omega) - 1 \right|^2 \text{PSD}[X](\omega) + \sigma^2 \sum_{k=1}^N |\widehat{D}^{(k)}(\omega)|^2 \right] d\omega.$$

Si l'on note $\|\widehat{h}\|^2 = |\widehat{h}_1|^2 + \dots + |\widehat{h}_N|^2$ et $\|\widehat{D}\|^2 = |\widehat{D}^{(1)}|^2 + \dots + |\widehat{D}^{(N)}|^2$, on a

$$\begin{aligned} & \text{PSD}[X] \times \left| \sum_{k=1}^N \widehat{D}^{(k)} \widehat{h}_k - 1 \right|^2 + \sigma^2 \|\widehat{D}\|^2 \\ &= \sum_{k=1}^N |\widehat{D}^{(k)}|^2 (\sigma^2 + |\widehat{h}_k|^2) - 2 \text{Re} \left[\sum_{k=1}^N \widehat{h}_k \widehat{D}^{(k)} \right] + \text{terme ind. de } \widehat{\vec{D}} \\ &= \sum_{k=1}^N (\sigma^2 + S_{X,X} |\widehat{h}_k|^2) \left| \widehat{D}^{(k)} - \frac{\overline{\widehat{h}_k} \text{PSD}[X]}{\text{PSD}[X] \|\widehat{h}\|^2 + \sigma^2} \right|^2 + \text{terme ind. de } \widehat{\vec{D}}. \end{aligned}$$

Le choix de \vec{D} de manière à ce que cette expression soit minimale est bien celui qui est donné par les relations

$$\widehat{D}^{(k)} = \frac{\overline{\widehat{h}_k} \text{PSD}[X]}{\left(\sum_{l=1}^N |\widehat{h}_l|^2 \right) \text{PSD}[X] + \sigma^2}, \quad k = 1, \dots, N.$$

La proposition est ainsi démontrée. \diamond

Remarque 3.5. Si l'on fait l'hypothèse selon laquelle les variables X_t sont peu corrélées, on prendra comme filtres approximant les filtres convoluteurs optimaux les filtres $D^{(k)}$, $k = 1, \dots, N$, dont le spectre est donné par

$$\widehat{D}^{(k)} = \frac{\overline{\widehat{h}_k}}{\left(\sum_{l=1}^N |\widehat{h}_l|^2\right) + \epsilon^2}, \quad k = 1, \dots, N, \quad (3.9)$$

où $\epsilon^2 = \sigma^2/\text{PSD}[X](0)$ désigne ce que l'on peut qualifier de *rapport bruit sur signal*. De tels filtres déconvoluteurs sont appelés *pseudo-filtres de Wiener* et couramment utilisés par exemple dans les techniques d'imagerie médicale (il est en effet évident que tout ce que nous venons de faire en 1D se transporte mot pour mot à la situation 2D).

Nous pouvons également énoncer une version déterministe de la proposition 3.5.

Proposition 3.6 *Supposons que $\mathcal{L}_1, \dots, \mathcal{L}_N$ soient N filtres analogiques stationnaires stables, de réponses impulsionnelles $h_1, \dots, h_N \in L^1(\mathbb{R})$ et que s_1, \dots, s_N soient N signaux d'énergie finie sur l'axe des temps. Si ϵ est un paramètre strictement positif donné, le minimum, lorsque e décrit $L^2(\mathbb{R})$, de l'expression*

$$\epsilon^2 \|e\|_2^2 + \sum_{k=1}^N \|\mathcal{L}_k[e] - s_k\|_2^2$$

est réalisé pour le signal dont le spectre est donné par

$$\widehat{e}^{\text{opt}} = \frac{1}{\|\widehat{h}\|^2 + \epsilon^2} \left(\sum_{k=1}^N \overline{\widehat{h}_k} \widehat{s}_k \right),$$

soit encore

$$e^{\text{opt}} = \sum_{k=1}^N D_{\mathcal{L}, \epsilon}^{(k)}[s_k],$$

où $D_{\mathcal{L}, \epsilon}^{(k)}$ est le filtre analogique stationnaire de transformée de Fourier

$$\omega \mapsto \frac{\overline{\widehat{h}_k}}{\|\widehat{h}\|^2 + \epsilon^2},$$

avec

$$\|\widehat{h}\|^2 := |\widehat{h}_1|^2 + \dots + |\widehat{h}_N|^2.$$

Les filtres analogiques $D_{\mathcal{L}, \epsilon}^{(k)}$ ainsi définis sont appelés (de part l'analogie présentée avec ceux de la remarque 3.5 (en (3.9)) *pseudo-filtres de Wiener déconvoluteurs associés à la convolution multi-canaux \mathcal{L} et à la tolérance bruit/signal ϵ .*

Preuve. Elle est en tout point identique à la preuve de la proposition 3.5 ; le terme à minimiser s'écrit aussi grâce à la formule de Plancherel

$$\begin{aligned} & \frac{1}{2\pi} \int_{\mathbb{R}} \left[\epsilon^2 |\widehat{e}(\omega)|^2 + \sum_{k=1}^N |\widehat{h}_k(\omega) \widehat{e}(\omega) - \widehat{s}_k(\omega)|^2 \right] d\omega \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \left(\|\widehat{h}\|^2 + \epsilon^2 \right) \left[\left| \widehat{e} - \frac{\sum_{k=1}^N \overline{\widehat{h}_k} \widehat{s}_k}{\|\widehat{h}\|^2 + \epsilon^2} \right|^2 + \text{terme ind. de } \widehat{e} \right] d\omega, \end{aligned}$$

d'où le résultat. \diamond

3.2.3 Une méthode ancienne : la gammagraphie par ouverture de codage

Pour voir l'intérêt du pseudo-filtrage de Wiener (tel qu'il est envisagé dans la sous-section précédente), nous allons évoquer ici une méthode beaucoup plus simple à implémenter (mais bien sûr en général bien moins performante) que les méthodes du type CAT-scanner, SPECT ou PET déjà discutées, à savoir la *gammagraphie par ouverture codée*.

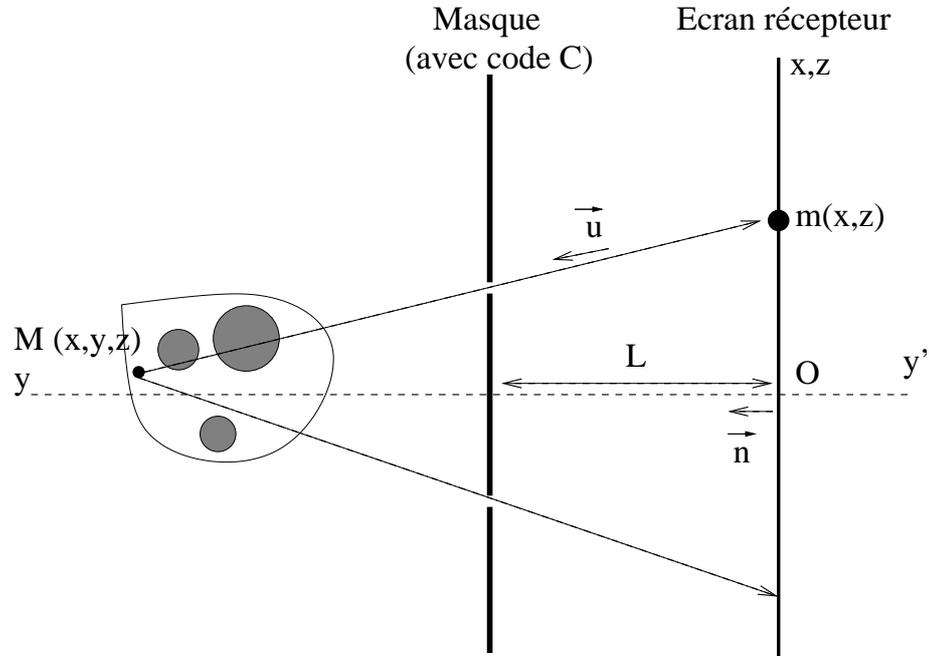


FIG. 3.2 – Dispositif gammagraphique par ouverture codée

Le mécanisme en est très simple et s'appuie sur l'optique géométrique ; il est d'usage fréquent en astronomie et astrophysique par exemple⁴.

Supposons que l'on place un écran récepteur (dans un plan vertical xOz) à une distance au moins égale à L d'un corps émettant un rayonnement (bruit Poissonien d'émission) de densité $f(x, y, z)$ au point courant $M(x, y, z)$ (ici $f(x, y, z)$ représente, on l'a vu le paramètre local $\lambda(x, y, z)$ du processus Poissonien d'émission (voir la figure 3.2, où l'on a pour l'instant omis de faire figurer le « masque » représenté comme le plan $x0'z$, $0' = (0, L, 0)$). Au point $m = (x_0, 0, z_0)$, le signal capté à l'issue du rayonnement est le signal

$$I(x_0, z_0) = \frac{1}{4\pi} \iiint_{y>L} f(x, y, z) \frac{\langle \vec{n}(x, y, z), \vec{u}(x, y, z; x_0, z_0) \rangle}{(x - x_0)^2 + (z - z_0)^2 + y^2} dx dy dz$$

⁴En médecine nucléaire, il a été utilisé dans les années 1980-1990, sous l'impulsion d'André Maréchal à l'Institut d'Optique d'Orsay ; par exemple en scintigraphie cardiaque (CHU d'Amiens, équipe du Professeur J. Fonroget) ; voir à ce sujet les travaux théoriques de J. Brunol, N. de Beaucoeuray, S. Lowenthal, Déconvolution analogique en imagerie par ouverture codée appliquée à la médecine nucléaire, *Optica Acta*, 25, 2 (1978), pp. 113-124, ou aussi les travaux de N. Ohyama, T. Honda, J. Tsujiuchi, Tomogram reconstruction using advanced coded aperture imaging, *Optics communications*, 36, 6 (1981), pp. 434-438.

$$= \frac{1}{4\pi} \iiint_{y>L} f(x, y, z) \frac{\cos^3(\vec{n}(x, y, z), \vec{u}(x, y, z; x_0, z_0))}{y^2} dx dy dz,$$

où $\vec{n} := (0, 1, 0)$ et $\vec{u}(x, y, z; x_0, z_0)$ désigne le vecteur directeur de la demi-droite mM ; ceci résulte de l'optique géométrique, avec les habituelles approximations dites « des petits angles ». Pour sérier mieux l'arrivée du rayonnement sur l'écran, il est naturel d'incorporer au dispositif une plaque de plomb (matérialisée par le plan $x0'z$) dans laquelle on a évidé un « code » C . Dans ce cas, le signal capté au point $(x_0, 0, z_0)$ de l'écran devient (grâce au théorème de Thalès)

$$\begin{aligned} I_C(x_0, z_0) &= \frac{1}{4\pi} \iiint_{y>L} f(x, y, z) C\left(\frac{y-L}{y}x_0 + \frac{L}{y}x, \frac{y-L}{y}z_0 + \frac{L}{y}z\right) \times \\ &\quad \times \frac{\cos^3(\vec{n}(x, y, z), \vec{u}(x, y, z; x_0, z_0))}{(y+L)^2} dx dy dz \\ &\simeq \frac{1}{4\pi} \iiint_{y>L} f(x, y, z) C\left(\frac{y-L}{y}x_0 + \frac{L}{y}x, \frac{y-L}{y}z_0 + \frac{L}{y}z\right) \frac{dx dy dz}{y^2} \end{aligned}$$

si l'on suppose le corps rayonnant placé près de l'axe $y'0y$ et l'angle (\vec{n}, \vec{u}) assez petit sur le support de la fonction

$$(x, y, z) \mapsto f(x, y, z) C\left(\frac{y-L}{y}x_0 + \frac{L}{y}x, \frac{y-L}{y}z_0 + \frac{L}{y}z\right)$$

figurant sous l'intégrale. En effet, seuls les rayons susceptibles de traverser la plaque de plomb, donc de passer au travers du code, sont dans ce cas captés (voir la figure 3.2). Si le code est symétrique $C(u, v) = C(-u, -v)$, on peut représenter $I_C(x_0, z_0)$ (après changement de variables sous l'intégrale) sous la forme

$$\begin{aligned} I_C(x_0, z_0) &= \\ &= \frac{1}{4\pi} \int_{y>L} \left[\iint f\left(\frac{y-L}{L}u, y, \frac{y-L}{L}w\right) C_{\frac{y-L}{L}}(x_0 - u, z_0 - w) dudw \right] \frac{(y-L)^2 dy}{L^2 y^2} \end{aligned}$$

où, pour $\lambda > 0$, C_λ désigne l'homothétique $(u, v) \mapsto C(\lambda u, \lambda v)$ du code C . On voit ainsi que la contribution $f(\cdot, y, \cdot) dy$ de la densité d'émission dans la « tranche verticale infinitésimale » $\mathbb{R} \times [y, y + dy] \times \mathbb{R}$, après homothétie (dépendant de y suivant les axes des x et des z), une convolution avec le code modifié

$$(u, v) \mapsto \left(\frac{y-L}{yL}\right)^2 C_{\frac{y-L}{L}}(u, v),$$

toutes ces contributions (pour les tranches verticales infinitésimales $\mathbb{R} \times [y, y + dy] \times \mathbb{R}$) étant ensuite ajoutées pour réaliser l'« empilement » que constitue $(x, z) \mapsto I_C(x, z)$.

Si l'on discrétise le problème en supposant que l'organe rayonnant se trouve découpé en N tranches verticales régulièrement espacées aux cotes $L < y_1 < y_2 < \dots < y_N$ (chacune d'épaisseur $e = (y_N - y_1)/N$), le résultat capté sur l'écran est discrétisé en

$$I_C(x_0, z_0) = e \sum_{k=1}^N \mu_k (C_{\lambda_k} * f_k)(x_0, z_0),$$

où $\lambda_k := (y_k - L)/L$, $\mu_k := (\lambda_k/y_k)^2$ et

$$f_k(x, z) = f(\lambda_k x, y_k, \lambda_k z), \quad k = 1, \dots, N.$$

Du fait que, pour $j \neq k$, la fonction

$$(\omega_1, \omega_2) \mapsto \widehat{C}_{\lambda_j}(\omega_1, \omega_2) \overline{\widehat{C}_{\lambda_k}(\omega_1, \omega_2)}$$

se présente comme une fonction oscillante (au contraire de ce qui se passe si $j = k$), la méthode (inspirée du pseudo-filtrage de Wiener) qui consiste, pour espérer récupérer f_k par « déconvolution », à filtrer I_C/e avec le filtre analogique de transformée de Fourier

$$\widehat{\mathcal{D}}_k = \frac{\mu_k \overline{\widehat{C}_{\lambda_k}}}{\sum_{j=1}^N \mu_j^2 |\widehat{C}_{\lambda_j}|^2 + \epsilon^2}$$

(où, ici encore, ϵ^2 peut être interprété comme un rapport bruit/signal) donne (après filtrage à *posteriori* dans le plan des fréquences) d'assez bons résultats lorsque la fonction source f à restituer (*via* la restitution des coupes f_k) est constituée d'entités bien identifiées, séparées et localisées (par exemple des nécroses dans la paroi du ventricule en scintigraphie cardiaque). On relèvera la similarité entre la méthode utilisée ici et celle utilisée pour la réalisation de filtres déconvoluteurs « optimaux » en termes de minimisation au sens des moindres carrés.

3.3 Algorithmes de « Matching Pursuit » ; P.O.D

Cette section, si elle s'éloigne du fil directeur de ce chapitre, n'en a pas moins sa place ici dans la mesure où elle présente des algorithmes d'intérêt constant en traitement du signal ou analyse d'images : le premier lorsqu'il s'agit de « traquer » une information contre un « dictionnaire » (par exemple une image médicale contre un dictionnaire d'images pathologiques), ou de restituer une source d'émission a priori inconnue (mais que l'on sait être d'un certain type) à partir de sa réponse à un dispositif gammagraphique comme dans la section précédente, *etc.* ; le second lorsque l'on espère isoler dans un signal non stationnaire les structures cohérentes (elles stationnaires!).

Si l'orthogonalité est une notion clef dans tout processus d'analyse et de synthèse en théorie de l'information, c'est malheureusement une notion « fragile » : par exemple le procédé d'orthonormalisation de Gram-Schmidt dans \mathbb{R}^n , initié à partir d'un système libre au sein duquel duquel deux vecteurs sont « presque » colinéaires, peut générer des combinaisons de ces vecteurs de base affectés d'énormes coefficients! Pour prendre un exemple plus concret, il est souvent utile aujourd'hui, face aux questions de sécurité ou de *copyright*, d'authentifier une image en y sur-imprimant un code (une « marque »⁵) que seul son propriétaire est à même de déceler ; si la réalisation préalable de décompositions orthogonales permet de « signer » l'image en codant intelligemment les divers composants de l'une de ces décompositions, il n'est aucunement évident que l'orthogonalité entre composants soit un tant soit peu préservée lorsque l'image subit un quelconque traitement (compression, modification géométrique, filtrage, *etc.*) ; dès lors, le processus d'authentification de l'image perturbée par son propriétaire devient impossible ! Voici encore un second exemple : en imagerie médicale, il s'avère fréquemment utile, plutôt que de tenter d'inverser le mécanisme tomographique \mathcal{R} qui a permis d'obtenir une image I , de chercher à

⁵C'est ce que l'on appelle le « *watermarking* ».

« pister » cette image contre un dictionnaire (éventuellement redondant, mais on veillera à limiter cela) d'images $\mathcal{R}[f_j]$, où f_1, \dots, f_N sont des états source pathologiques auxquels on aimerait confronter l'état f dont on ne connaît que $\mathcal{R}[f]$; on pressent en effet que f s'apparente à une combinaison de ces états pathologiques f_j , $j = 1, \dots, N$, combinaison qu'il serait judicieux de retrouver sous forme « organisée » : d'abord la pathologie (parmi les f_j) la mieux « corrélée » à f , puis celle qui, bien qu'aussi corrélée significativement à f , l'est un peu moins, *etc.* Ce sont ces idées que nous allons esquisser dans cette section.

3.3.1 Dictionnaires et « Matching Pursuit »

Étant donnée, dans un espace de Hilbert, un élément dont on souhaite extraire une information (aux fins d'analyse, de classification, ou de synthèse ultérieure), l'un des algorithmes les plus naïfs (mais de fait aussi les plus robustes) que l'on puisse imaginer est celui qui consiste à « pister » h contre un « dictionnaire » que l'on aura au préalable composé d'éléments « test » dont on s'attend *a priori* que l'élément h proposé soit une combinaison linéaire. Il peut s'avérer utile (on le verra plus loin) d'élaborer un dictionnaire d'éléments « test » à partir de l'élément h que l'on prétend analyser. On donnera dans cette section deux versions (l'une élémentaire, l'autre un peu plus élaborée) de l'algorithme mathématique soutendant ce scénario, dit algorithme de « Matching Pursuit »⁶.

Le cadre proposé est le cadre hilbertien des espaces $H = L_{\mathbb{K}}^2(\mathbb{R})$ ou $l_{\mathbb{K}}^2(\mathbb{Z})$ avec $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} ; on pourrait tout aussi bien envisager $L_{\mathbb{K}}^2(\mathbb{R}^n)$ ou $l_{\mathbb{K}}^2(\mathbb{Z}^n)$; il faut aussi ajouter que, même dans le cadre de la dimension finie, $H = \mathbb{R}^N$ ou $H = \mathbb{C}^N$, le recours à pareille démarche se justifie.

On convient d'appeler *dictionnaire* une liste finie ou dénombrable d'éléments de H engendrant un sous-espace dense; pour simplifier, on suppose tous les éléments du dictionnaire normalisés et de norme égale à 1. Étant donné un tel dictionnaire \mathcal{D} et un élément h de H , l'algorithme le plus simple proposé (MP pour « Matching Pursuit ») repose sur la proposition suivante.

Proposition 3.7 *Soit \mathcal{D} un dictionnaire d'éléments de H et h un élément de H ; on suppose que l'on peut construire de manière itérative une suite d'éléments $(d_k)_{k \in \mathbb{N}}$ du dictionnaire \mathcal{D} (les d_j dépendant de h), une suite de scalaires $\lambda_1, \dots, \lambda_k, \dots$ (dépendant aussi de h) tels que $|\langle h, d_1 \rangle| = \max_{d \in \mathcal{D}} |\langle h, d \rangle|$, $\lambda_1 d_1 = \text{Proj}_{\mathbb{K}d_1}(h)$ et, pour tout $k \geq 1$,*

$$\left| \left\langle h - \sum_{l=1}^k \lambda_l d_l, d_{k+1} \right\rangle \right| = \max_{d \in \mathcal{D}} \left| \left\langle h - \sum_{l=1}^k \lambda_l d_l, d \right\rangle \right| \quad \text{et}$$

$$\lambda_{k+1} d_{k+1} = \text{Proj}_{\mathbb{K}d_{k+1}} \left(h - \sum_{l=1}^k \lambda_l d_l \right).$$

Alors, la suite

$$(h_k)_{k \geq 1} := \left(\sum_{l=1}^k \lambda_l d_l \right)_{k \geq 1}$$

⁶La terminologie anglo-saxonne résume l'objectif : « traquer » (ou « poursuivre ») en essayant d'« ajuster » aux combinaisons d'éléments du dictionnaire (c'est la phase de « matching »).

converge vers h dans H ; on dit qu'elle réalise le « matching » de h suivant le dictionnaire \mathcal{D} .

Remarque 3.6. Comme le dictionnaire \mathcal{D} peut fort bien être redondant (on a d'ailleurs tout intérêt à ce qu'il en soit ainsi), la décomposition de h suivant \mathcal{D} n'a rien d'unique ; la méthode proposée dans ce scénario algorithmique fournit simplement une décomposition efficace en « traquant » (de manière « hiérarchique ») les éléments tests du dictionnaire qui comptent le plus dans la représentation de h .

Posons $v_l := \lambda_l d_l$ pour $l \in \mathbb{N}^*$ et notons r_k le reste de la décomposition de h une fois l'étape k de l'algorithme effectuée, soit $r_k := h - \sum_{l=1}^k v_l$, partant de $r_0 = h$. Si $1 \leq k < p$, on a $r_k = r_p + \sum_{l=k+1}^p v_l$. Comme on effectue à chaque cran de l'algorithme une projection orthogonale unidimensionnelle sur la droite vectorielle engendrée par le nouvel élément du dictionnaire trouvé, on a, pour tout $l \in \mathbb{N}^*$, $\|r_k\|^2 = \|r_{k-1}\|^2 - |\langle r_{k-1}, d_k \rangle|^2$; la suite $(\|r_k\|^2)_{k \geq 0}$ est donc une suite décroissante minorée, donc convergente ; on en déduit aussi

$$\sum_{l=1}^{\infty} \|v_l\|^2 < \infty. \quad (3.10)$$

Pour montrer que la suite $(r_k)_{k \geq 0}$ converge, nous allons prouver qu'elle est de Cauchy. Fixons donc $\epsilon > 0$; on a, si $1 \leq k < p$, $\|r_k - r_p\|^2 = \|r_k\|^2 + \|r_p\|^2 - 2\|r_p\|^2 - 2 \sum_{l=k+1}^p \operatorname{Re} \langle r_p, v_l \rangle$; mais on a, pour tout l entre $k+1$ et p ,

$$|\langle r_p, v_l \rangle| \leq \|v_l\| \times \|v_{p+1}\| ; \quad (3.11)$$

en effet, toujours pour l entre $k+1$ et p ,

$$\begin{aligned} |\langle r_p, v_l \rangle| &= |\langle r_p, \langle r_{l-1}, d_l \rangle d_l \rangle| = |\langle r_p, d_l \rangle| \times |\langle r_{l-1}, d_l \rangle| \\ &= \|v_l\| |\langle r_p, d_l \rangle| \leq \|v_l\| \times \|v_{p+1}\| \end{aligned}$$

du fait même du principe de l'algorithme, puisque $|\langle r_p, d_{p+1} \rangle|$ (qui est par définition la norme de v_{p+1}) maximise tous les $|\langle r_p, d \rangle|$ avec $d \in \mathcal{D}$. On a donc

$$\|r_k - r_p\|^2 \leq \|r_k\|^2 - \|r_p\|^2 + 2\|v_{p+1}\| \sum_{l=k+1}^p \|v_l\|. \quad (3.12)$$

Or $\|v_{p+1}\| \sum_{l=k+1}^p \|v_l\| \leq \|v_{p+1}\| \sum_{l=1}^{p+1} \|v_l\|$; mais la clause (3.10) implique

$$\lim_{p \rightarrow \infty} \left(\inf_{q \geq p} \left(\|v_q\| \sum_{l=1}^q \|v_l\| \right) \right) = 0. \quad (3.13)$$

Si k est assez grand, on est donc certain que $\|r_k\|^2 - \|r_p\|^2 \leq \epsilon$; on choisira k assez grand pour qu'il en soit ainsi. Il résulte de (3.13) qu'il existe $q > p$ tel que $\|v_q\| \sum_{l=1}^q \|v_l\| \leq \epsilon$; mais on a $\|r_k - r_p\| \leq \|r_k - r_q\| + \|r_p - r_q\|$. En appliquant (3.12) (mais cette fois avec les couples (k, q) et (p, q) au lieu de (p, k)), on voit que $\max(\|r_k - r_q\|^2, \|r_p - r_q\|^2) \leq 3\epsilon$. En mettant tout ensemble, on conclut à ce que $\|r_k - r_p\| \leq 2\sqrt{3}\epsilon$ pour ce choix de k , ce qui implique bien que la suite $(r_k)_{k \geq 0}$ est de Cauchy, donc convergente. Mais on a aussi la convergence vers 0 de $\|v_k\|$ lorsque k tend vers l'infini ; on a donc $\lim_{k \rightarrow \infty} |\langle r_k, d_{k+1} \rangle| = 0$. Le principe de l'algorithme

(appliqué une fois de plus) nous assure que pour tout élément d du dictionnaire, $\lim_{k \rightarrow \infty} |\langle r_k, d \rangle| = 0$. Comme le dictionnaire forme une partie totale, la limite de la suite $(r_k)_{k \geq 0}$ doit être orthogonale à tous les atomes du dictionnaire, est par conséquent nulle, et la proposition est démontrée. \diamond

Cet algorithme est entaché d'un défaut : il faut à chaque étape travailler avec le dictionnaire complet et l'on ne peut se permettre d'éliminer les éléments « tests » dès lors qu'ils sont apparus ; pour corriger cet état de fait et transformer l'algorithme de *Matching Pursuit* en un algorithme « glouton », on peut en introduire une version orthogonale en le couplant avec le procédé d'orthonormalisation de Gram-Schmidt : c'est l'algorithme *Matching Pursuit Orthogonal* (M.P.O) dont on présente ici la description. Il s'agit, outre la démarche décrite dans la proposition 3.7, d'imposer à chaque itération (à savoir la détection de d_{k+1}) l'orthogonalité du reste avec les éléments du dictionnaire précédemment sélectionnés ; pareil algorithme nécessite bien sûr, dès cet élément test d_{k+1} déterminé, un ré-ajustement des coefficients du « résumé » $\sum_{l=1}^k \alpha_l^{(k)} d_l$ dont on disposait avant que la recherche de d_{k+1} ne soit mise en route.

Voici en quelques points la démarche algorithmique que génèrent les diverses opérations impliquées dans cette variante par la détection du $(k+1)$ -ème atome. Supposons que, précédant cette étape, on dispose d'une liste d_1, \dots, d_k d'éléments du dictionnaire sélectionnés (aux k -crans précédents), d'un « résumé » $R_k := \sum_{l=1}^k \alpha_l^{(k)} d_l$, et de l'inverse de la matrice de Gram G_k des atomes d_1, \dots, d_k ⁷. On effectue alors la détection de l'élément test d_{k+1} suivant la règle :

$$\left| \left\langle h - \sum_{l=1}^k \alpha_l^{(k)} d_l, d_{k+1} \right\rangle \right| = \max_{d \in \mathcal{D}, d \neq d_1, \dots, d_k} \left| \left\langle h - \sum_{l=1}^k \alpha_l^{(k)} d_l, d \right\rangle \right|;$$

on calcule ensuite le vecteur ligne ${}^t C_k$ des corrélations (*i.e* produits scalaires) de ce nouvel élément test d_{k+1} avec les k éléments d_1, \dots, d_k du dictionnaire précédemment sélectionnés :

$${}^t C_k := (\langle d_{k+1}, d_1 \rangle, \dots, \langle d_{k+1}, d_k \rangle)$$

puis le vecteur colonne $B_k = G_k^{-1} C_k$ de coordonnées $B_k(l)$, $l = 1, \dots, k$; notons que B_k, C_k seront stockés en mémoire ; on calcule enfin le coefficient $\alpha_{k+1}^{(k+1)}$ appelé à affecter le nouvel atome sélectionné d_{k+1} dans la nouvelle version R_{k+1} du résumé, selon la formule

$$\alpha_{k+1}^{(k+1)} = \frac{\langle h - R_k, d_{k+1} \rangle}{1 - \sum_{l=1}^k B_k(l) \langle d_l, d_{k+1} \rangle},$$

puis on effectue les ré-ajustements nécessaires sur les coefficients $\alpha_1^{(k)}, \dots, \alpha_k^{(k)}$ pour obtenir l'orthogonalité du nouveau reste $h - R_{k+1}$ avec les éléments test d_1, \dots, d_{k+1} ; ces ré-ajustements conduisent à l'expression du nouveau résumé :

$$R_{k+1} = \sum_{l=1}^k (\alpha_l^{(k)} - \alpha_{k+1}^{(k+1)} B_k(l)) d_l + \alpha_{k+1}^{(k+1)} d_{k+1}.$$

Avant de procéder à l'étape suivante, on injecte les données jusque là stockées B_k, C_k, G_k^{-1} pour préparer l'itération ultérieure, en l'occurrence calculer l'inverse de

⁷Ce calcul, on le verra plus loin, aura été effectué de manière récursive (*via* une information mémorisée lors de l'itération précédente).

la matrice de Gram G_{k+1} des éléments test d_1, \dots, d_{k+1} ; ce calcul s'effectue suivant la formule :

$$G_{k+1}^{-1} = \begin{pmatrix} G_k^{-1} + \rho_k B_k B_k^* & -\rho_k B_k \\ -\rho_k B_k^* & \rho_k \end{pmatrix} \quad \text{avec} \quad \rho_k := \frac{1}{1 - C_k^* B_k}.$$

3.3.2 La décomposition orthogonale propre

Comme nous l'avons mentionné dans la section précédente, la réalisation préalable d'un dictionnaire à partir précisément de l'élément que l'on prétend analyser peut s'avérer un outil utile ; c'est ce dictionnaire qui servira à la mise en œuvre des algorithmes décrits dans la section précédente.

Exemple. Si s est un signal digital $(s(1), \dots, s(N))$ de longueur N , on peut, s'inspirant de la démarche conduisant par exemple au calcul de la fonction d'autocorrélation et de la densité spectrale de puissance (méthode de Welch décrite dans la section 3.2.1), considérer comme dictionnaire la liste des signaux digitaux $(0, \dots, 0, s(k), \dots, s(k+M-1), 0, \dots, 0)$ obtenus en tronquant s hors d'une fenêtre de longueur $M \ll N$ que l'on translate (avec un pas suffisamment petit, la valeur idéale – mais coûteuse – en étant 1) dans l'intervalle $\{1, \dots, N\}$; de même, pour une image, on peut, s'inspirant du classique algorithme JPEG, envisager des troncatures de l'image par un masque de 8 pixels sur 8 pixels que l'on translate chaque fois d'au plus d'un cran suivant l'une ou l'autre direction ; on parle à propos de dictionnaires obtenus de cette manière de dictionnaires de « *snapshots* » (on pourrait dire des « instantanés » du signal ou de l'image digitale).

L'un des premiers objectifs à réaliser lorsque l'on construit un dictionnaire à partir d'un élément h donné est que les éléments du dictionnaire soient beaucoup plus faciles à stocker que l'élément h lui-même (dans l'exemple des dictionnaires de « *snapshots* », stocker un élément du dictionnaire revient à stocker $M \ll N$ entrées non nulles ainsi que l'indice initial de la fenêtre).

Si $H = H_N = l_{\mathbb{K}}^2(\{1, \dots, N\})$ ($\mathbb{K} = \mathbb{R}$ ou \mathbb{C}) et que le dictionnaire $\mathcal{D}[h]$ (de cardinal $D = N - M$) construit pour analyser h est composé d'instantanés de h de longueur $M \ll N$ pris sur une fenêtre de longueur M que l'on décale chaque fois d'une unité (chacun de ces instantanés étant assimilé à un élément de $H_M := l_{\mathbb{K}}^2(\{1, \dots, M\})$) on exploite les idées inspirées par le principe de la décomposition en valeurs singulières pour réaliser à partir de $\mathcal{D}[h]$ un $\mathcal{D}[h]$ originel si on l'utilise pour implémenter la « traque » de h via l'algorithme dictionnaire orthogonal cette fois, donc non redondant, plus efficace et moins coûteux que le dictionnaire M.P décrit dans la section précédente. La redondance de $\mathcal{D}[h]$ n'est pas cependant ici ignorée : elle est fondamentalement prise en compte pour la construction précisément de ce nouveau dictionnaire. L'idée est en effet de considérer la matrice de Gram $D \times D$ définie par

$$G := \left(\langle d_j, d_k \rangle_{H_M} \right)_{1 \leq j, k \leq D}$$

et d'en calculer la liste des valeurs propres $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_M \geq \lambda_{M+1} = 0 \geq \dots \geq \lambda_D = 0$. On notera U_1, \dots, U_M le système orthonormé de \mathbb{K}^D constituée de vecteurs propres de G (respectivement associés aux valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_M$). Les premiers de ces vecteurs propres U_1, U_2, \dots , indiquent en quelque sorte les « directions principales » statistiquement les plus probables pour le sous-espace de H_N engendré par les instantanés d_1, \dots, d_D , considérés cette fois comme éléments de H_N et tronqués par des zéros hors d'une fenêtre de longueur M fixée, prise dans la liste des fenêtres choisies pour la réalisation des « *snapshots* ». Pour chaque $j = 1, \dots, D$, on posera

donc⁸

$$D_j := \sum_{k=1}^D u_{jk} d_k, \quad j = 1, \dots, M$$

et on réalisera, après avoir choisi un indice r entre 1 et D , un nouveau dictionnaire $\mathcal{D}_r[h]$ en concaténant les r dictionnaires $\mathcal{D}_{r,j}[h]$, $j = 1, \dots, r$, obtenus en positionnant D_j dans $[N/M]$ fenêtres consécutives (sans chevauchement cette fois) de l'intervalle $\{1, \dots, N\}$; ce dictionnaire $\mathcal{D}_r[h]$ est de cardinal $[N/M] \times r$, significativement inférieur à $N - M$ si r est petit devant M . Les éléments de ce nouveau dictionnaire $\mathcal{D}_r[h]$ (qui sont cette fois deux-à-deux orthogonaux dans H_N) sont dits *modes propres* de h et le *matching* de h contre ce dictionnaire (d'autant plus intéressant que l'on peut choisir $r \ll M$, ce qui est d'autant plus possible que la suite (λ_k) décroît rapidement) est un exemple de *décomposition orthogonale propre*⁹. Si les r modes propres D_1, \dots, D_r de h sont ainsi stockés, on espère pouvoir proposer dans chacune des $[N/M]$ fenêtres consécutives une version « résumé » de h sous la forme d'une combinaison linéaire de D_1, \dots, D_r ; le nombre d'entrées nécessaires pour stocker cette version comprimée sera donc $r \times [N/M]$ (il y a $[N/M]$ fenêtres consécutives) et la compression sera d'autant plus intéressante que le taux r/M pourra être pris petit.

3.3.3 P.O.D et non stationnarité

La construction d'un dictionnaire $\mathcal{D}_r[h]$, $1 \leq r \leq M$ à partir d'un dictionnaire exhaustif d'instantanés (de longueur M) de h , du fait que son principe se fonde sur des idées statistiques, ne se révèle réellement efficace que si h répond à des critères de stationnarité (sous l'angle déterministe ou stochastique). Toutefois, en l'absence de tels critères, la détermination préalable des modes propres D_1, \dots, D_M (à partir d'un dictionnaire exhaustif d'instantanés), couplée avec l'algorithme M.P.O, ouvre une voie pour isoler les structures cohérentes au sein de h : il est naturel de lancer l'exploration de h avec l'algorithme M.P.O en utilisant, pour $1 \leq r \leq M$, le dictionnaire $\tilde{\mathcal{D}}_r[h]$ (de cardinal toujours $\simeq D$) dont les éléments sont obtenus en positionnant l'un des D_j , $j = 1, \dots, r$ dans l'une des $[D/r]$ fenêtres de longueur M de l'intervalle $\{1, \dots, M\}$, prises de r en r .

Exemple. On peut analyser suivant de telles méthodes les signaux temporels générés par les équations de Navier-Stokes qui régissent par exemple les écoulements turbulents (inhérents aux phénomènes physiologiques), signaux constituant typiquement des signaux non stationnaires auxquels on peut envisager d'utiliser à titre d'exercice ces techniques. On peut par exemple considérer les signaux correspondant respectivement aux enregistrements (en fonction du temps) de la vorticité (module du rotationnel du champ de vitesse) et de la pression en un point pour l'écoulement d'un fluide (en présence d'obstacles induisant la turbulence) dans un canal $2D$.

⁸Notons que les D_j , $j = 1, \dots, M$, construits ainsi peuvent aussi s'obtenir comme les colonnes de la matrice U dans la décomposition en valeurs singulières $[U, \text{Diag}, V] = \text{svd}(D)$, où D est la matrice $M \times D$ dont les colonnes sont les éléments d_1, \dots, d_D de H_M .

⁹P.O.D pour « Proper Orthogonal Decomposition » dans la terminologie anglo-saxonne; on parle aussi d'*analyse en composantes principales*.

Chapitre 4

Une esquisse de présentation des outils mathématiques de l'IRM

4.1 Introduction

Le principe de l'Imagerie par Résonance Magnétique Nucléaire (IRMN) se fonde sur les propriétés magnétiques des noyaux des atomes. Nous y retrouverons cependant (mais à la fin seulement) la transformation de Radon (3D cette fois, c'est-à-dire par les plans). Ce que nous prétendons juste faire ici est juste d'isoler les outils mathématiques sous-jacents et relativement « tractables » (équation de Bloch, repère mobile, transformation de Fourier, inversion par rétroprojection de la transformation de Radon par les plans en trois dimensions) en pointant du doigt également les incontournables problèmes de nature physique¹.

4.2 Champ statique et champ Radio-Fréquence

Le corps à étudier est placé dans un environnement où il se trouve soumis à un champ magnétique constant $B_{00}\vec{k}$, avec, on le verra, B_{00} constante positive très grande (ce champ est créé par un puissant aimant, l'aimant de l'IRM).

Les particules (les protons en sont un exemple) possèdent un moment magnétique de spin \vec{M} que l'on peut visualiser comme un vecteur unitaire de \mathbb{R}^3 . En fait le moment de spin \vec{M} d'un proton est en principe un vecteur d'opérateurs ($\hat{M}_x, \hat{M}_y, \hat{M}_z$) déduit du spin 1/2 de ce proton par multiplication par la constante $gq/2m$, où q désigne la charge du proton, m sa masse et g la constante de Landé qui vaut dans le cas particulier d'un proton $g \simeq 5.586$. Le spin 1/2 d'un proton s'exprime sous la forme générale de la superposition de deux états propres, soit comme le quotient de deux nombres complexes α/β , ou encore comme un point de la sphère de Riemann, c'est-à-dire un point de la sphère \mathbb{S}^2 . C'est cette remarque qui nous permet d'assimiler une fois pour toutes le moment magnétique de spin \vec{M} d'un proton à un vecteur

¹Nous nous sommes ici inspirés de diverses présentations de Charles Epstein, dont le texte d'un exposé « A lecture on Selective RF-pulses in MRI » disponible en ligne sur son site web : <http://www.math.upenn.edu/~cle/> Un site intéressant illustrant ce chapitre avec des codes MATLAB simples est par exemple <http://mrsrl.stanford.edu/~brian/bloch/> L'article de C. Epstein, Introduction to Magnetic Resonance Imaging for mathematicians, Annales de l'Institut Fourier, 54 (5), 2004, pp. 1697-1716, propose une excellente introduction au sujet et a fortement inspiré ces notes.

unitaire de \mathbb{R}^3 .

Sous l'effet d'un champ magnétique (dépendant des coordonnées d'espace x, y, z et du temps) l'équation différentielle reliant moment magnétique de spin et champ magnétique ambiant pour un proton positionné au point (x, y, z) est

$$\frac{d\vec{M}}{dt} = \gamma(1 + \sigma)\vec{B}(x, y, z, t) \wedge \vec{M}, \quad (4.1)$$

introduite par Félix Bloch. Ici, γ , exprimé en MHz par Tesla, désigne le *rapport gyromagnétique* du noyau (pour nous, en l'occurrence ici le proton). Parce que le noyau de l'atome d'hydrogène ^1H est constitué d'un seul proton et possède un rapport gyromagnétique fort ($\gamma \simeq 42.6$ MHz par Tesla, soit *grosso modo* la fréquence des ondes radio), et que l'atome d'hydrogène est présent dans quasiment toutes les structures organiques (ne serait-ce que sous forme de la molécule H_2O), c'est sur lui que l'on se base pour travailler en IRMN; c'est donc la population des noyaux de l'atome d'hydrogène que l'on prétend pister.

La *fréquence angulaire de Larmor* du champ constant $B_{00}\vec{k}$ est obtenue comme le produit γB_{00} . La constante σ traduit des variations de la fréquence de Larmor liée à des variations de la distribution d'électrons (dans la terminologie anglo-saxonne, c'est le *chemical shift*) et on la négligera ici en supposant $\sigma = 0$.

Lorsque $\vec{B} = B_{00}\vec{k}$, la résolution de l'équation (4.1) avec $\sigma = 0$ est immédiate et conduit à

$$\vec{M}(t) = U^{\text{ref}}(t)\vec{M}(0), \quad t \in [0, T],$$

où

$$U^{\text{ref}}(t) = \begin{pmatrix} \cos(\omega_0 t) & -\sin(\omega_0 t) & 0 \\ \sin(\omega_0 t) & \cos(\omega_0 t) & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

on comprend l'interprétation de ω_0 comme fréquence angulaire et on voit que ω_0 s'interprète ici comme la vitesse de rotation angulaire du moment magnétique de spin autour de l'axe vertical \vec{k} .

On introduit maintenant un vecteur $\vec{\theta}$ de \mathbb{S}^3 , une constante positive $c \ll B_{00}$ et on introduit le champ magnétique (lui variable dans l'espace)

$$\vec{b}_{0,\vec{\theta}} : X = (x, y, z) \mapsto c\langle X, \vec{\theta} \rangle \vec{k};$$

la superposition du champ constant $B_{00}\vec{k}$ et du champ $\vec{b}_{0,\vec{\theta}}$ sera notée $\vec{B}_{0,\vec{\theta}}$ et appelée *champ magnétique statique* dans la direction $\vec{\theta}$. On note que les « lignes de niveau » de ce champ statique (c'est-à-dire les surfaces sur lesquelles il est constant) sont les plans

$$\langle X, \vec{\theta} \rangle = s, \quad s \in \mathbb{R}.$$

On retrouve la collection des plans de coupe perpendiculaires à la direction $\vec{\theta}$ qui apparaissent dans l'expression de la transformation de Radon $(\vec{\theta}, s) \mapsto R[\cdot](\vec{\theta}, s)$. Cette remarque nous servira ultérieurement.

Lorsque la population de protons est soumise à ce champ statique $\vec{B}_{0,\vec{\theta}}$, les protons que constituent les noyaux des atomes d'hydrogène ^1H (population que l'on a donc décidé de cibler ici) s'alignent (tels des toupies) soit dans la direction du champ, soit dans la direction opposée (on parle de protons parallèles et anti-parallèles); si

l'on examine le bilan de la magnétisation sur tous les protons, de fait 99.999% des contributions à la magnétisation s'annulent deux-à-deux ², mais le bilan final est une magnétisation positive (infime puisqu'elle ne touche que la population pour laquelle il n'y a pas annulation) fournit un bilan de magnétisation dirigé suivant la direction du champ. C'est de ce bilan positif que l'on va essayer de capter un signal.

4.3 La phase de « précession »

Pour ce faire, nous allons dans un premier temps (phase dite de *précession*) ajouter au champ statique $\vec{B}_{0,\vec{\theta}}$ un champ « excitant » tournant

$$(x, y, z, t) \mapsto \vec{B}_{\text{RF}}(x, y, z, t) := e^{i\omega_0 t}(a(t)\vec{i} + b(t)\vec{j})$$

(la vitesse de rotation angulaire étant la fréquence angulaire de Larmor du champ statique $B_{00}\vec{k}$, soit $\omega_0 = \gamma B_{00}$), ce pendant un laps de temps de longueur T à partir de l'instant $t = 0$. Le rapport d'intensité entre champ statique et champ « excitant » tournant est très grand (de l'ordre de 10^6); ce champ est par contre orthogonal en tout point au champ statique; notons qu'il est, lui, indépendant de (x, y, z) . La notation RF vaut ici pour « Radio-Fréquence ».

Quel est l'effet de ce champ excitant additionnel sur le comportement du moment magnétique face à cette fois la nouvelle équation de Bloch

$$\frac{d\vec{M}}{dt} = \gamma(\vec{B}_{0,\vec{\theta}} + \vec{B}_{\text{RF}}) \wedge \vec{M}, \quad (4.2)$$

dans le laps de temps (de fait très court $[0, T]$) où il est appliqué? Pour simplifier les choses, on préfère poser

$$\vec{M}(t) = U_0^{\text{ref}}(t)\vec{\mu}(t)$$

et chercher le système différentiel vérifié par $\vec{\mu}$, mais en le posant dans le repère mobile obtenu en tournant le repère $(\vec{i}, \vec{j}, \vec{k})$ de $\omega_0 t$ autour de l'axe des z . L'équation (4.2) devient, si $\vec{B}_{\vec{\theta}} := \vec{B}_{0,\vec{\theta}} + \vec{B}_{\text{RF}}$,

$$\begin{aligned} (U_0^{\text{ref}}(t))^{-1} \frac{d\vec{M}}{dt} &= (U_0^{\text{ref}}(t))^{-1} \frac{dU_0^{\text{ref}}}{dt} \vec{\mu}(t) + \frac{d\vec{\mu}}{dt} \\ &= (U_0^{\text{ref}})^{-1} [\gamma \vec{B}_{\vec{\theta}} \wedge (U_0^{\text{ref}} \vec{\mu}(t))] + \frac{d\vec{\mu}}{dt} \\ &= \left[(U_0^{\text{ref}})^{-1} \cdot \left([\gamma \vec{B}_{\vec{\theta}} \cdot U_0^{\text{ref}}] \right) \right] \wedge \vec{\mu}(t) + \frac{d\vec{\mu}}{dt} \\ &= \gamma \left[(U_0^{\text{ref}})^{-1} \vec{B}_{\vec{\theta}} \right] \wedge [U_0^{\text{ref}} \vec{\mu}(t)] + \frac{d\vec{\mu}}{dt}. \end{aligned}$$

Ceci implique que dans le repère mobile tournant (de $\omega_0 t$ autour de l'axe $z'Oz$), l'équation de Bloch se lit

$$\frac{d\vec{\mu}}{dt} = \vec{B}_{\text{effective},\vec{\theta}} \wedge \vec{\mu}(t), \quad (4.3)$$

²Un autre point de vue serait de supposer que 99.999% des protons de fait ne s'alignent pas mais restent dans un état non organisé (chaotique) tandis que 0.001% d'entre eux s'alignent dans la direction du champ, ce qui revient de fait au final au même si l'on examine le bilan global; le point important est que l'on ne va de fait profiter que du comportement « organisé » d'une toute petite minorité de protons!

où

$$\vec{B}_{\text{effective},\vec{\theta}} = (U_0^{\text{res}})^{-1} \vec{B}_{\vec{\theta}} - \frac{\omega_0}{\gamma} \vec{k}.$$

En travaillant dans le repère tournant, on peut donc se ramener à supposer l'équation de Bloch de la forme (4.3), avec

$$\vec{B}_{\text{effective},\vec{\theta}} = \gamma^{-1}(\alpha(t), \beta(t), s),$$

s étant indépendant de t (mais dépendant de la position X du proton) et prenant ses valeurs dans $[-s_0, s_0]$ ($s = \gamma(B_{00} + c\langle X, \vec{\theta} \rangle) - \omega_0 = c\langle X, \vec{\theta} \rangle$). Une valeur donnée de s , couplée avec le choix d'une direction $\vec{\theta}$, est à interpréter comme le repérage d'un plan de coupe dans \mathbb{R}^3 . Le choix du vecteur $\vec{\theta}$ conditionne le champ de gradient additionnel linéaire suivant précisément la direction $\vec{\theta}$. Le champ $\vec{B}_{\text{effective},\vec{\theta}}$ dépend de la position X du proton ; lorsque les protons considérés sont dans un plan de coupe $\langle X, \vec{\theta} \rangle = \xi$, avec $\xi \in \mathbb{R}$, ce champ est fonction de ξ , au sens où $s = \gamma c \xi$. On peut donc considérer que fixer une valeur de s revient à fixer précisément un de ces plans de coupe. On interprète s comme le « décalage de résonance » (*resonance offset*) ou encore « décalage de fréquence ». L'explication de cette terminologie réside dans le fait que l'équation de Bloch dans le repère mobile s'écrit très simplement :

$$\begin{pmatrix} \frac{d\mu_{s,1}}{dt} \\ \frac{d\mu_{s,2}}{dt} \\ \frac{d\mu_{s,3}}{dt} \end{pmatrix} = \begin{pmatrix} 0 & s & -\beta(t) \\ -s & 0 & \alpha(t) \\ \beta(t) & -\alpha(t) & 0 \end{pmatrix} \begin{pmatrix} \mu_{s,1}(t) \\ \mu_{s,2}(t) \\ \mu_{s,3}(t) \end{pmatrix}.$$

Ce système différentiel du premier ordre se lit *via* la transformation de Fourier (et les notations complexes)

$$\begin{aligned} e^{-ist} \frac{d}{dt} [e^{ist} (\mu_{s,1}(t) + i\mu_{s,2}(t))] &= i(\alpha(t) + i\beta(t)) \mu_{s,3}(t) \\ \frac{d\mu_{s,3}}{dt} &= \beta(t) \mu_{s,1}(t) - \alpha(t) \mu_{s,2}(t). \end{aligned} \quad (4.4)$$

On a donc, pour $t \in [0, T]$,

$$(\mu_{s,1} + i\mu_{s,2})(t) = ie^{-ist} \int_0^t e^{is\tau} \mu_{s,3}(\tau) (\alpha(\tau) + i\beta(\tau)) d\tau.$$

L'interprétation de s comme un « décalage de fréquence » est ainsi claire. À l'instant final T (ou va s'achever la phase de précession), la composante transverse $\mu_{s,1} + i\mu_{s,2}$ du moment magnétique de spin (exprimé dans le repère mobile) s'écrira

$$(\mu_{s,1} + i\mu_{s,2})(T) = ie^{-isT} \int_0^T e^{is\tau} \mu_{s,3}(\tau) (\alpha(\tau) + i\beta(\tau)) d\tau.$$

Si l'on suppose que μ_3 reste voisin de 1 pendant le laps de temps $[0, T]$ où est subi le champ excitant Radio-Fréquence, on peut supposer $\mu_3 \equiv 1$ sur $[0, T]$ et l'on obtient donc

$$(\mu_{s,1} + i\mu_{s,2})(T) = e^{-isT} \int_0^T e^{is\tau} (\alpha + i\beta)(\tau) d\tau ; \quad (4.5)$$

on voit dans ce cas que la composante transverse du moment magnétique de spin $\vec{\mu}_s$ à l'instant t , considérée comme fonction de s , est le spectre de la fonction $\alpha + i\beta$ (qui est appelée « enveloppe RF ») à un déphasage près.

Si l'on spécifie une fonction $\vec{\mu}_{\text{profil}} : s \in [-s_0, s_0] \longrightarrow \mathbb{S}^2$ (on l'appellera le *profil de magnétisation*), on peut se poser le problème de *design* suivant : déterminer, T désignant un laps de temps fixé, les fonctions $t \mapsto \alpha(t)$, $t \mapsto \beta(t)$ de manière à ce que

$$\vec{\mu}_s(T) = \vec{\mu}_{\text{profil}}(s) \quad \forall s \in [-s_0, s_0].$$

Un exemple typique de profil est celui qui consiste à choisir un (petit) intervalle I de $[-s_0, s_0]$ (il est intéressant de se souvenir ici que s étant en correspondance avec un plan orthogonal à la direction $\vec{\theta}$, un tel petit intervalle est en correspondance avec une « tranche » perpendiculaire à cette direction), à se donner un angle ψ (on l'appellera un « flip ») entre 0 et 180 degrés et à choisir le profil $\mu_{\text{profil}}^{I,\psi}$ comme :

$$\mu_{\text{profil}}^{I,\psi}(s) := \begin{cases} (\sin \psi, 0, \cos \psi) = \sin \psi \vec{i} + \cos \psi \vec{k} & \text{si } s \in I \\ (0, 0, 1) = \vec{k} & \text{si } s \notin I. \end{cases}$$

La surimposition du champ RF adapté à ce profil pendant le laps de temps $[0, T]$ aura donc pour effet de faire dévier la magnétisation d'un proton de sa position originelle \vec{k} (lorsque $t = 0$) à sa position finale $\sin \psi \vec{i} + \cos \psi \vec{k}$, ce lorsque $s \in I$ (donc pour des protons appartenant à une certaine tranche fine perpendiculaire à la direction $\vec{\theta}$).

Malheureusement dans la pratique, les « flips » que l'on cherche à réaliser (pour que la phase de relaxation dont on parlera plus loin se révèle la plus efficace possible) sont des « flips » de 90, voire 180 degrés. Il n'est donc plus question de faire l'approximation $\mu_{s,3} \simeq 1$ qui nous permettait d'ajuster les fonctions $t \in [0, T] \mapsto \alpha(t)$ et $t \in [0, T] \mapsto \beta(t)$ pour qu'il en soit ainsi. Si ψ est petit, il est par contre assez facile d'obtenir la RF-enveloppe $t \mapsto (\alpha + i\beta)(t)$ permettant de réaliser cet objectif : si l'on oublie le déphasage, on voit que $\alpha + i\beta$ s'obtient comme l'inverse (par Fourier) de la fonction caractéristique de I , multipliée par $\sin \psi$; par exemple, si $I = [-\sigma, \sigma]$ et si l'on choisit

$$\alpha + i\beta = \sin \psi \times \chi_{[0,T]}(t) \times \frac{\sin(\sigma(t - T/2))}{t - T/2}$$

on réalise, compte tenu de la formule (4.5) un champ RF de manière à ce que

$$\mu_{s,1}(T) + i\mu_{s,2}(T) \simeq ie^{-isT/2} \sin \psi \times \chi_{[-\sigma,\sigma]}(s).$$

Si l'on renverse le champ statique $\vec{B}_{0,\vec{\theta}}$ et que l'on poursuit la phase de précession pendant un laps de temps égal à $T/2$, on parviendra à corriger le déphasage.

Lorsque l'on a affaire à de grands angles (90, 180 degrés, comme c'est le cas sur les machines couramment utilisées), le *design* de l'enveloppe RF est plus complexe du point de vue mathématique et fait appel par exemple soit à la méthode SLR³, soit aux outils mathématiques du scattering inverse (voir l'article déjà mentionné de C. Epstein aux Annales de l'Institut Fourier, ce ne sont pas des mathématiques trop difficiles).

³J. Pauly, P. Le Roux, D. Nishimura, A. Macovski, Parameter relations for the Shinnar-Le Roux selective excitation pulse design algorithm, IEEE Trans. on Med. Imaging. 10, 1991,53-65.

4.4 La phase de « relaxation »

Nous en donnons ici une présentation extrêmement simplifiée. Il faut pour commencer penser qu'à l'issue de la phase de précession, le champ magnétique de spin des protons a été dévié (tel un parapluie que l'on aurait ouvert) de son équilibre, puis que l'on laisse automatiquement le parapluie se refermer (en cliquant sur le presseur). Ce « retour à l'équilibre » va induire un signal dont on va s'intéresser aux composantes longitudinales (orthogonales à l'axe du parapluie).

Raisonnons à partir de maintenant de manière non plus locale (au niveau du comportement de chaque proton comme nous l'avons fait dans la section précédente) mais de manière globale. En présence du champ statique $\vec{B}_{0,0}$, la densité de magnétisation de spin s'exprime sous la forme

$$\frac{C}{\text{Temp}} \rho(x, y, z) \vec{B}_{00} \vec{k}$$

où C est une constante absolue, Temp la température, ρ la densité des molécules d'eau. Au terme donc de la phase de précession ainsi précédemment décrite (on a pris pour simplifier une enveloppe RF de manière à réaliser au final le profil de fréquence $s \mapsto (1, 0, 0) = \vec{i}$, donc un « flip » de $\psi = 90$ degrés), on obtient ainsi une densité de magnétisation de spin (M_x, M_y, M_z) avec

$$M_x + iM_y = \frac{CB_{00}\rho(x, y, z)}{\text{Temp}} e^{i(\omega_0 + c\langle X, \vec{\theta} \rangle)}.$$

Dès l'instant où se relâche l'action du champ RF, se produit un phénomène de relaxation durant lequel la population de protons revient à son équilibre (en ce qui concerne les moments de magnétisation de spin), c'est-à-dire à l'état précédent l'initiation du champ excitant RF. En fait, cette relaxation se traduit (au niveau de la composante $M_x + iM_y$) par une décroissance exponentielle, la densité de magnétisation $M_x + iM_y$ s'exprimant alors comme

$$(x, y, z, t) \mapsto \frac{CB_{00}\rho(x, y, z)}{\text{Temp}} e^{i(\omega_0 + c\langle X, \vec{\theta} \rangle)t} \times e^{-t/T_2}.$$

On a ici un champ magnétique décroissant qui induit donc (d'après la loi de Faraday) dans des bobines placées ici de manière à leur axe soit l'axe des x et l'axe des y) un signal complexe $S_x + iS_y$ (on compose ainsi le résultat des deux courants induits) qui précisément sera le signal mesuré ici (dit signal FID pour « Free Induction Decay »). Si l'on calcule la contribution totale de toute la population de protons et que l'on intègre donc sur tout le corps étudié, le signal FID mesuré sera en fait proportionnel (le facteur dépend des instruments de mesure, de la sensibilité du détecteur, *etc.*) à

$$\begin{aligned} \text{FID} &= \frac{CB_{00}e^{-t/T_2}e^{i\omega_0 t}}{\text{Temp}} \iiint \rho(x, y, z)(\omega_0 + c\langle X, \vec{\theta} \rangle) e^{itc\gamma\langle X, \vec{\theta} \rangle} dx dy dz \\ &\simeq \frac{C\omega_0 B_{00}e^{-t/T_2}e^{i\omega_0 t}}{\text{Temp}} \iiint \rho(x, y, z) e^{itc\gamma\langle X, \vec{\theta} \rangle} dx dy dz. \end{aligned}$$

À ce stade, en échantillonnant en temps et en jouant sur la possibilité de faire varier $\vec{\theta}$, on peut récupérer la transformée de Fourier (3D) de ρ au voisinage de l'origine, donc au final la densité ρ .

Ce n'est toutefois pas en général ainsi que l'on procède. Il s'avère de fait beaucoup plus intéressant de choisir le profil de fréquence de manière à réaliser un « flip » (par exemple de 90 degrés) correspondant à un petit intervalle I de $[-s_0, s_0]$, donc, on l'a vu, à ne prendre en compte que les protons appartenant à une tranche fine (d'épaisseur $|I|$) perpendiculaire à la direction $\vec{\theta}$. Ceci nous conduit donc à ce que les signaux captés soit les signaux

$$\frac{C\omega_0 B_{00} e^{-t/T_2} e^{i(\omega_0+s)t}}{\text{Temp}} \iint_{\vec{\theta}^\perp} \rho(s\vec{\theta} + \vec{u}) \, dudv, \quad s \in \mathbb{R}.$$

En prenant le spectre (en t), puis en faisant varier $\vec{\theta}$, on retrouve la transformée de Radon de ρ par les plans,

$$(\vec{\theta}, s) \longmapsto \iint_{\vec{\theta}^\perp} \rho(s\vec{\theta} + \vec{u}) \, dudv$$

que l'on sait inverser par rétroprojection. On retrouve à ce stade ici la tomographie classique.

Encore une fois, il s'agit ici d'une présentation piétonnière du mécanisme de l'IRMN, qu'il nous paraissait cependant important de faire figurer dans ce cours de tomographie.