

Reconstruction parcimonieuse de signal

Introduction à la reconstruction de signal et au
compressed sensing

Leo Davy

supervisé par
Jean-François Crouzet

Institut de Mathématiques Alexander Grothendieck
Université de Montpellier
Faculté des sciences
Février-Mai

Contents

1	Cadre du problème	3
1.1	Traitement du signal, analyse et synthèse	3
1.2	Reconstructions atomiques et parcimonie	4
1.3	Exemples d'applications	7
2	Frame et reconstruction $\ \cdot\ _2$	10
2.1	Bases orthonormales et frames	10
2.1.1	Intérêt des bases orthonormales et description des outils mathématiques disponibles	10
2.1.2	Lien entre frame et base orthonormale	10
2.1.8	Exemples de frames (Fourier, ondelettes)	19
2.2	Décroissance des coefficients et régularité	25
2.2.1	Approximation linéaire et régularité	25
2.2.2	Décroissance des coefficients de Fourier	27
2.2.4	Décroissance des coefficients d'ondelettes	28
3	Reconstruction parcimonieuse et $\ \cdot\ _1$	33
3.1	Résolution de (P0)	33
3.1.1	Définition de (P0)	34
3.1.3	Solution optimale combinatoire	37
3.1.4	Résolution dans un dictionnaire pics/Fourier	37
3.1.5	Principe d'incertitude	37
3.2	Résolution de (P1)	37
3.2.5	Généralisation à des paires de bases arbitraires	42
3.2.7	Extensions du résultat	43
3.2.8	Définition de (P1)	44
3.2.9	Propriétés du minimiseur	44
3.3	Lien géométrique entre (P0) et (P1)	44
3.3.1	Boules unité en grande dimension	44
3.3.2	Unicité de la solution de (P0) et (P1)	44

4	Compressed sensing et approche aléatoire	46
4.1	Axiomatisation, UUP et RIP	46
4.1.1	Définition de UUP	47
4.1.8	Exemple de familles vérifiant UUP	48
4.1.10	Définition de ERP	48
4.1.12	Exemples de familles vérifiant ERP	49
4.1.14	Lien entre RIP et ERP	49
4.2	Théorème de Candes-Tao	49
4.2.1	Enoncé du théorème	49
4.2.2	Preuve du théorème	50
4.3	Exemple de F_Ω	50
4.3.1	Ensemble de Fourier	50
4.3.2	Gaussien	50
4.4	Conséquences du théorème	50
4.4.1	Influence des paramètres	50
4.4.2	Quelques résultats numériques	50
4.5	Sur la propriété RIP	50
4.5.1	Difficulté pour un ensemble de vérifier RIP	50
4.5.2	Lien entre RIP et WERP	50
4.6	Extensions du théorème	50
4.6.1	Conditions suffisantes sur δ_K	50
4.6.2	Conditions nécessaires sur δ_K	50
4.6.3	Sur l'optimalité du résultat	50
4.7	Algorithmes	50
4.7.1	Orthogonal Matching Pursuit	50
4.7.2	Robust Orthogonal Matching Pursuit	50
4.7.3	Quelques exemples numériques	50
A	Annexe	51
A.1	Valeurs propres de $F_\Omega F_\Omega^*$ et Analyse en composante principale	51
A.2	Outils probabilistes de la preuve du théorème	51

Chapter 1

Cadre du problème

1.1 Traitement du signal, analyse et synthèse

La reconstruction du signal est un problème que l'on considère dans le cadre du traitement du signal, c'est à dire que l'on considère qu'à un signal, on peut appliquer une transformation, et de cette transformation on obtient un nouveau signal qui aura certaines caractéristiques permettant de mieux comprendre ce signal. D'un point de vue plus formel, on considère une famille de signaux \mathcal{F} , chaque élément de cette famille étant une application $f : X \longrightarrow Y$, et on considère un opérateur $F : \mathcal{F} \longrightarrow \mathcal{G}$, où \mathcal{G} est une autre famille de signaux.

Par exemple une famille de signaux que l'on considérera est \mathbb{R}^N et on verra donc $f \in \mathbb{R}^N$ comme une application $\llbracket 0, N - 1 \rrbracket \rightarrow \mathbb{R}$. Une autre famille que l'on étudiera est $L^2(\mathbb{R}^d)$ ou bien $\ell^2(\mathbb{N}^d)$ et certains de leurs sous-espaces, par exemple on pourra se restreindre à des applications satisfaisant une forme de régularité.

L'objectif de ce mémoire est de présenter des techniques permettant la reconstruction de signaux, donc on cherchera à obtenir des opérateurs $F : \mathcal{F} \rightarrow \mathcal{F}$ qui vérifieront $f = F(f)$.

La première étape en traitement du signal est la mesure, à partir d'un signal, on lui associe des coefficients, ainsi on a une application

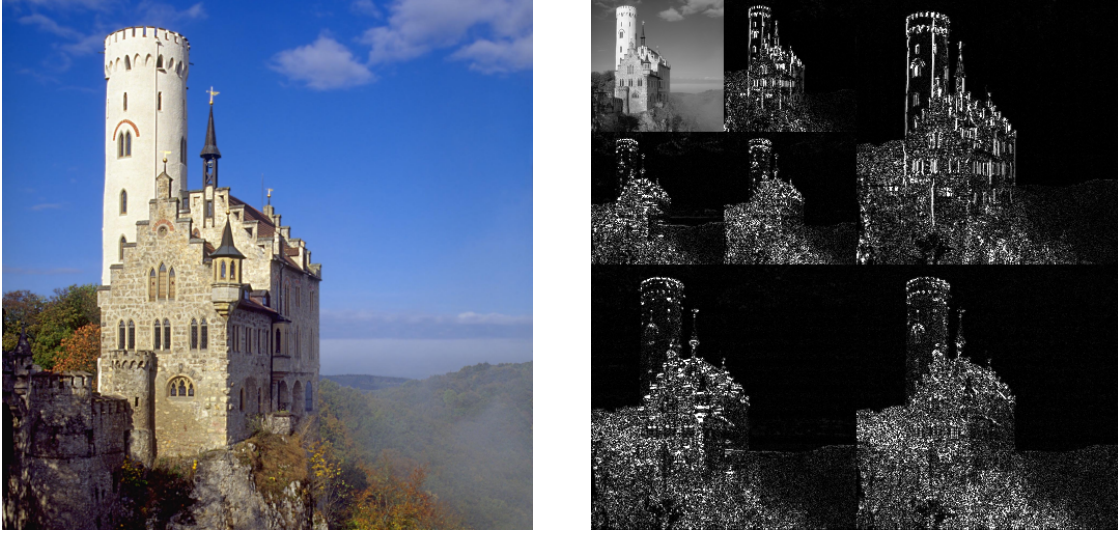
$$A : \mathcal{F} \longrightarrow \mathcal{G} \tag{1.1}$$

$$f \longmapsto A(f) = (a_g(f))_{g \in I} \tag{1.2}$$

que l'on appelle *opérateur d'analyse*.

Des opérateurs d'analyse que l'on rencontrera seront par exemples la transformée de Fourier, qui a une application associe ses coefficients de Fourier, ou bien la transformée en ondelette, qui a une application associe ses coefficients d'ondelette. Un autre exemple d'opérateur que l'on étudiera est celui qui à un élément de \mathbb{R}^N lui associe ses coefficients dans deux bases, donc un élément de \mathbb{R}^{2N} . Pour donner un dernier exemple d'opérateur que l'on considérera consistera à associer à un élément de \mathbb{R}^N la valeur de sa projection sur M sous-espaces de \mathbb{R}^N choisis au hasard, donc un élément de \mathbb{R}^M .

TODO : ajouter des figures avec un signal et sa transformée de Fourier.

Figure 1.1: Une photographie et sa transformée en ondelettes.^a

^aImages de Wikicommons

Comme nous souhaitons reconstruire des signaux, notre objectif est de pouvoir utiliser ces coefficients afin de revenir dans l'espace initial. On définit donc l'application

$$S : \mathcal{G} \longrightarrow \mathcal{F} \quad (1.3)$$

$$A(f) \longmapsto S(A(f)) \quad (1.4)$$

que l'on appelle *opérateur de synthèse*.

On aura donc une formule de reconstruction si $S \circ A = Id_{\mathcal{F}}$. Par exemple, avec \mathcal{F} bien choisi, des opérateurs de synthèse que l'on rencontrera seront par exemple, la transformée de Fourier inverse, ou bien la transformée en ondelette inverse, qui a des coefficients de Fourier, respectivement d'ondelettes, associe le signal initial. La partie ?? introduira la notion de frame qui nous permettra de caractériser des opérateurs sur lesquels on aura la garantie d'avoir une formule de reconstruction.

1.2 Reconstructions atomiques et parcimonie

Dans ce mémoire on va, suivant l'approche de Stéphane Mallat¹, approcher la problématique de la reconstruction et du traitement du signal en donnant une place centrale à la notion de parcimonie. De façon large, l'approche parcimonieuse consistera soit à minimiser le nombre de coefficients obtenus par l'analyse ou utilisés par la synthèse, soit à exploiter,

¹L'approche parcimonieuse est au coeur du programme de Stéphane Mallat, son livre qui fait référence en traitement du signal notamment concernant les ondelettes *A wavelet tour of signal processing* s'est vu dans sa dernière édition ajouter le sous-titre *The sparse way* et le cours présenté durant sa chaire au Collège de France était centré sur *Le triangle "régularité, approximation, parcimonie"*.

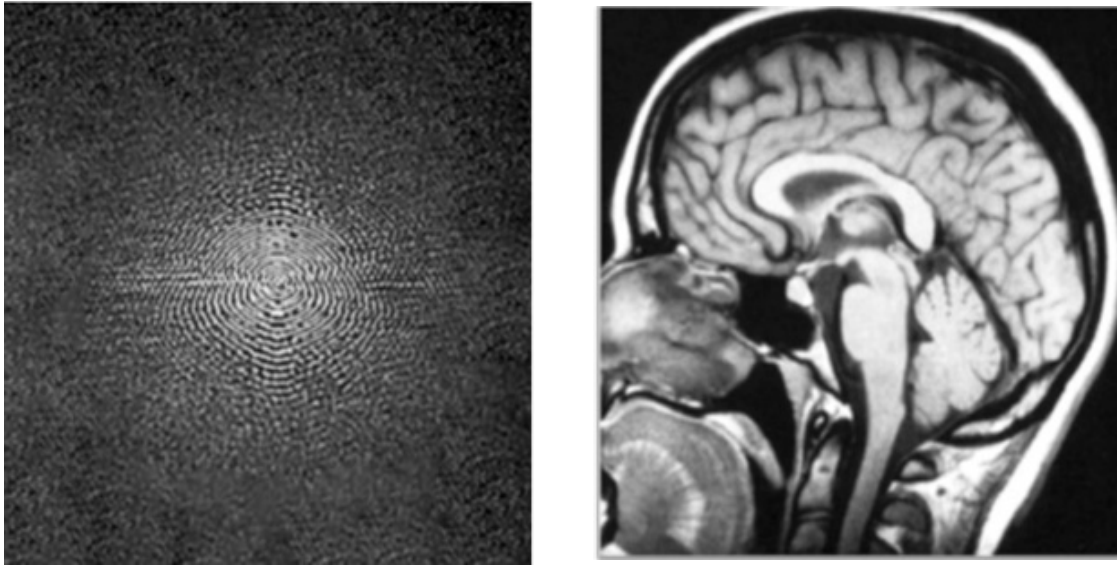


Figure 1.2: Mesure par résonance magnétique d'un cerveau (le signal), qui correspond à des coefficients de Fourier (l'analyse), et reconstruction du signal (la synthèse).^a

^aImages du site www.MRIquestions.com de Allen D. Elster.

à l'aide d'un principe d'incertitude, l'hypothèse que le signal est représentable avec peu de coefficients. Décrivons donc ces trois stratégies que l'on développera dans ce mémoire.

La première approche consiste à minimiser le nombre de coefficients que l'on obtient lors de l'analyse, ainsi, cela revient à choisir un opérateur d'analyse qui est adapté à \mathcal{F} . On verra ainsi, que si on étudie des fonctions avec certaines propriétés de régularité, il est possible de choisir une transformation en ondelettes pour laquelle les coefficients seront rapidement faibles. Donc, pour un signal donné, il peut être intéressant de regarder, pour plusieurs transformées en ondelettes, laquelle est celle qui a le minimum de coefficients non nuls. Par exemple, pour reconstruire une image en deux dimensions, il est possible de reconstruire cette image en choisissant parmi des ondelettes de Daubechies, des coiflets, des curvelets, des brushlets, ... chacune de ces ondelettes étant plus ou moins adaptée pour représenter des images avec différentes propriétés. On verra ainsi dans ce mémoire que les ondelettes ayant des moments nuls et à support compact (les ondelettes de Daubechies et les coiflets) permettent une reconstruction avec une décroissance rapide de leurs coefficients pour les fonctions Lipschitziennes ??.

Les deux autres approches reposent sur le fait que l'on dispose de différentes façons de représenter un signal. Par exemple dans le cadre de l'algèbre linéaire et d'un vecteur $f \in \mathbb{R}^N$ ou dans un espace de Hilbert, chaque choix de base revient à choisir un couple d'opérateurs analyse-synthèse (A, S) et donc à une formule de reconstruction. De la même façon, chaque transformée en ondelettes donne un couple d'opérateurs analyse-



Figure 1.3: L'ondelette de Daubechies à 2 moments nuls et la coifflet à 2 moments nuls.^a

^aImages de Wikicommons

synthèse (A, S) , on a ainsi le double bénéfice des ondelettes qui, en plus d'avoir de bonnes propriétés, existent en grand nombre.

Avec cette multitude de représentations disponibles d'un même objet, la notion de *dictionnaire* a été développée. Introduisons cette notion par l'algèbre linéaire, un dictionnaire est alors simplement une famille de vecteurs $\{f_i\}_I$, disons que ce dictionnaire engendre \mathcal{F} , alors pour chaque signal $f \in \mathcal{F}$, il existe au moins une suite de coefficients $(c_i)_J, J \subset I$, telle que

$$f = \sum_J c_i f_i. \quad (1.5)$$

Ainsi, on peut identifier la suite de coefficients $(c_i)_J$ à f , et chacune de ces suites de coefficients donne lieu à une identification avec f . Une telle suite de coefficients peut être appelée une décomposition atomique de f , chaque vecteur f_i étant alors appelé un atome. Avec une approche parcimonieuse, on souhaite alors pour un signal donné, trouver la décomposition atomique qui utilise un minimum d'atomes. Cela correspond donc à utiliser un minimum de coefficients pour la synthèse de f .

Ce problème est de nature combinatoire et très difficile à résoudre de façon générale (c'est un problème NP-difficile [23]) On verra dans ?? qu'en choisissant un dictionnaire vérifiant un principe d'incertitude alors on pourra effectivement résoudre de façon unique le problème de trouver la décomposition atomique d'un élément, sous réserve qu'une décomposition atomique suffisamment parcimonieuse existe.

La dernière approche et la plus récente est celle du compressed sensing [11], [4], [3]. Dans ce cadre, on souhaite reconstruire un signal $x_0 \in \mathbb{R}^N$ avec $k \ll N$ composantes non nulles, dans une certaine base, on se demande alors si il est possible de reconstruire ce signal en faisant $m < N$ mesures de x_0 . On verra qu'il est possible de répondre

à cette question par l'affirmative. Une partie des arguments provient de l'approche précédente, la mesure devra vérifier un principe d'incertitude avec la base dans laquelle le signal est parcimonieux, une fois que l'on a un tel opérateur d'analyse, en résolvant le problème de décomposition atomique on devrait pouvoir retrouver x_0 . Cependant, on ne connaît pas à l'avance dans quelle base le signal est parcimonieux, l'approche prend alors un tournant probabiliste et on va chercher à faire une mesure qui va presque sûrement vérifier un principe d'incertitude avec la base inconnue.

Donnons un exemple du type de résultat auquel on peut arriver à partir du compressed sensing, supposons que l'on souhaite reconstruire une image x_0 en niveaux de gris de taille 1000×1000 , donc un vecteur dans \mathbb{R}^d avec $d = 10^6$. Supposons que ce vecteur a $10000 = 10^4$ coefficients d'ondelette non nuls, alors avec 10^5 mesures, en cherchant une décomposition atomique qui coïncide avec ces mesures, on devrait pouvoir reconstruire exactement x_0 . Le mécanisme (d'après [10]) qui fait que l'on peut reconstruire ce signal est similaire au fait que si l'on tire 10000 points au hasard dans \mathbb{R}^{10^6} (l'espace de l'image) et que l'on forme un cône partant de l'origine avec ces points, alors faire coïncider le signal avec les 10^5 mesures revient à choisir comment passer d'un plan à 10^5 dimensions à ce cône. Le compressed sensing est alors rendu possible par le fait que la plupart des plans à 900000 dimensions ne coupent pas le cône, donc le signal reconstruit avec 10^5 mesures devrait effectivement être le signal original.

1.3 Exemples d'applications

Maintenant que l'on a vu l'approche que nous allons utiliser pour la reconstruction de signaux, voyons quelques exemples de problèmes de traitement du signal sur lesquels les techniques développées dans ce mémoire donnent des méthodes de résolution. Afin de formaliser ces problèmes nous aurons besoin de $\|\cdot\|_1$ et $\|\cdot\|_2$ les normes sur ℓ^1 et ℓ^2 , ainsi que la "norme" $\|\cdot\|_0$ définie pour $x := (c_i)_{i \in I}$ par $\|x\|_0 = |\{i : c_i \neq 0\}|$. Cette "norme" permet de mesurer la parcimonie de l'écriture de x , cependant, bien qu'on l'appelle ici norme (en omettant les guillemets), ce n'en est pas une car elle n'est pas homogène. L'un des objectifs de ce mémoire est de comprendre comment manipuler et obtenir des résultats avec cette norme, de nombreux résultats seront ainsi obtenus en la reliant à la norme $\|\cdot\|_1$.

Exemple 1.3.1 (Compression). Soit $y \in \mathbb{R}^N$ un signal. Normalement on doit décrire y avec ses N coefficients, avec ce qui est développé dans ce mémoire, plusieurs méthodes sont proposées pour décrire y avec $k < N$ coefficients. Est-ce possible de représenter y avec moins de N coefficients?

On répondra à cette question sous des conditions différentes avec la théorie des frames et avec la décomposition atomique.

Exemple 1.3.2 (Débruitage). Soit $y \in \mathbb{R}^N$ un signal et soit $\epsilon \in \mathbb{R}^N$ un vecteur avec des petits coefficients. Est-ce possible de retrouver y à partir de $y + \epsilon$?

On répondra à cette question avec la théorie des frames serrés et avec la décomposition atomique.

Exemple 1.3.3 (Problèmes inverses). Soit y un signal dont on fait une mesure indirecte, $\tilde{y} = Hy$ avec H un opérateur linéaire. Est-ce possible de trouver une solution x parcimonieuse à $\tilde{y} = HAx$?

Dans ce cas là Ax donne une décomposition atomique de y avec des colonnes de A comme atomes. Sous certaines conditions, la partie sur la décomposition atomique peut permettre de résoudre ce type de problèmes.

Exemple 1.3.4 (Tomographie et mesure de Radon). Dans ce contexte on cherche à mesurer un objet que l'on ne peut pas observer directement (par exemple la composition de la croûte terrestre en géologie syismique ou bien un organe en imagerie médicale par rayons X), la technique consiste alors à faire passer au travers de ce système un signal que l'on contrôle (par exemple une onde de choc ou bien un rayon X) et à en mesurer la réponse.

Soit $x = (x_i)_I$ un signal connu que l'on fait passer à travers un système $A = (a_i)_I$, on obtient donc $y_x = \sum_I a_i x_i$. On se demande alors si il est possible de trouver une décomposition atomique de A . Des techniques basées sur la recherche d'une telle décomposition ont été développées par exemple pour la géologie avec des ondelettes [15], pour l'imagerie par rayons X [5] [26] et plus généralement pour des problèmes d'inversion de la transformée de Radon [25] [24]. Dans ces articles des techniques relevant de la décomposition atomique (en minimisant la norme $\|\cdot\|_1$) ou du compressed sensing sont utilisées.

Exemple 1.3.5 (Imagerie par résonnance magnétique (IRM)). Terminons cette série d'exemples avec l'apport du compressed sensing aux techniques d'IRM, on se base sur [16] pour la discussion.

En IRM l'objectif est de reconstruire des images de l'intérieur d'un corps humain (par exemple), ainsi la mesure se doit d'être indirecte et sans danger. Cette mesure est faite en envoyant des suites de photons à travers le corps et en raison de phénomènes physiques, la mesure de l'énergie de ces photons correspond à une transformée de Fourier spatiale d'une section du corps. La mesure est limitée en vitesse par des contraintes inévitables (voir [16]) et donc il n'est pas possible de connaître l'ensemble des coefficients de Fourier.

Formalisons donc ce problème, notons F la transformée de Fourier, et F_I la restriction de F aux $|I|$ coefficients mesurés. Notons m la section du corps que l'on cherche à trouver, et on a donc $y = F_I m$ le signal obtenu. Cependant l'équation $y = F_I x$ n'admet pas de solution unique, il nous faut donc rajouter une contrainte. Grâce à la partie ?? on saura que l'on peut généralement représenter une image avec peu de coefficients dans une base d'ondelette, notons cette décomposition Wx . Grâce au compressed sensing on verra que le problème de minimiser $\|Wx\|_1$ parmi les x qui vérifient $y = F_I x$ admet une unique solution, que l'on peut calculer et c'est $x = m$.

Ainsi, avant le compressed sensing les images reconstruites demandaient une longue durée de scan (6mn avec certaines périodes d'apnée) et les résultats n'étaient pas de très bonne qualité. Après le compressed sensing, la reconstruction ne nécessitant pas de récupérer tous les coefficients, les scans sont devenus bien plus rapides et de meilleure

qualité (25s avec respiration libre).

Avant de poursuivre et de s'éloigner des applications pour passer à la théorie, notons quelques unes de leurs caractéristiques auxquelles on veillera dans les résultats obtenus. Tout d'abord, les problèmes énoncés sont résolubles avec les méthodes de ce mémoire sous certaines conditions, ainsi, sur certains problèmes appliqués, par exemple ceux dont la résolution implique l'utilisation d'une décomposition atomique, leur résolution n'est pas automatique. En effet, il est nécessaire d'avoir un dictionnaire avec des atomes permettant une décomposition atomique, et en pratique cela peut être très difficile de trouver un tel dictionnaire pour certains problèmes.

Ensuite, certains de ces problèmes sont issus de la physique, il y a donc de fortes chances pour que les mesures soient perturbées et avec une certaine imprécision. Il sera donc intéressant de vérifier que les techniques de reconstruction que l'on utilise sont stables lorsque la solution est légèrement modifiée. De la même façon, en pratique, dans les égalités du type $y = Ax$ des exemples ci-dessus, on cherchera plutôt à résoudre $\|y - Ax\|_2 < \varepsilon$.

Finalement, un dernier point auquel il faut veiller, la première étape étant la mesure, il sera intéressant de pouvoir donner pour certaines classes de signaux le nombre minimal de mesures nécessaires afin de reconstruire le signal mesuré. Le dernier exemple nous indique aussi qu'il peut être intéressant de pouvoir dire quand est-ce qu'on a suffisamment de mesures pour pouvoir reconstruire le signal.

Chapter 2

Frame et reconstruction $\|\cdot\|_2$

2.1 Bases orthonormales et frames

2.1.1 Intérêt des bases orthonormales et description des outils mathématiques disponibles

Une approche classique et pratique pour l'analyse de signaux est l'utilisation d'une base orthonormale pour représenter un signal. En effet l'intérêt est multiple: si l'on connaît une base orthonormale de décomposition d'un signal, il y a une unique façon d'écrire ce signal dans cette base, mais surtout, l'espace est alors naturellement muni d'un produit scalaire qui permettra d'utiliser tout l'outillage des espaces de Hilbert pour résoudre le problème.

On verra ainsi dans cette section tout d'abord des définitions et propriétés classiques des espaces de Hilbert. Ensuite on verra progressivement comment relâcher certaines des définitions initiales afin de pouvoir conserver une formule de reconstruction. Afin d'explicitier l'intérêt de ces définitions on verra deux exemples de frames. Tout d'abord le frame de Fourier, dont la compréhension sera utile pour le troisième chapitre. Ensuite, nous introduirons les ondelettes par l'analyse multi-résolution; les formules que nous obtiendrons seront utilisées pour montrer que certaines ondelettes permettent d'obtenir des formules de reconstruction avec une approximation rapide sur certaines classes de fonctions. Ainsi, après ce chapitre il devrait être clair que la reconstruction en norme $\|\cdot\|_2$ est faisable, de plusieurs façons et qu'il est possible que de telles reconstructions aient une représentation parcimonieuse.

2.1.2 Lien entre frame et base orthonormale

Rappelons tout d'abord les définitions et propriétés d'une base orthonormale. On considère ici H un espace de Hilbert donc un espace vectoriel muni d'une base hilbertienne et d'un produit scalaire. Pour alléger les notations, on considère que H est un \mathbb{R} -espace vectoriel, pour passer à un \mathbb{C} -espace vectoriel la seule modification est d'appliquer la conjugaison à de nombreux endroits. L'ensemble des résultats présentés restent vrais

dans le cas complexe.

Définition 2.1.3. On dira qu'une famille $\{e_i\}_I$ d'éléments de H est :

- *libre* si pour n'importe quelle suite finie de coefficients $(\lambda_i)_{J, J \subset I}$ telle que $\sum_J \lambda_i e_i = 0$, on a $\lambda_i = 0$ pour n'importe quel $i \in J$.
- *orthogonale* si pour n'importe quels i et j différents on a $\langle e_i, e_j \rangle = 0$
- *totale* si quel que soit $f \in H$ tel que pour tout $i \in I$ on a $\langle f, e_i \rangle = 0$, alors $f = 0$.
- *une base hilbertienne* si la famille est orthonormale et totale.

Donc, pour tout $h \in H$, si la famille $\{e_i\}_I$ est libre et totale, il existe une unique suite $(\lambda_i)_{i \in J}$ de scalaires, telle que $h = \sum_{i \in J} \lambda_i e_i$, on peut donc faire une identification entre h et $(\lambda_i)_J$. On peut alors expliciter le produit scalaire sur H , en notant $h_1 = (\lambda_i)_J, h_2 = (\mu_i)_J$

$$\langle \cdot, \cdot \rangle : H \times H \longrightarrow \mathbb{R} \quad (2.1)$$

$$(h_1, h_2) \longmapsto \langle h_1, h_2 \rangle = \sum_I \lambda_i \mu_i. \quad (2.2)$$

On peut remarquer que ce produit scalaire est défini de façon unique par rapport à la base hilbertienne $\{e_i\}_I$, cependant, si $U : H \rightarrow H$ est un opérateur unitaire, ce qui correspond en dimension finie à une rotation ou à un changement de base, alors on a $\langle U h_1, U h_2 \rangle = \langle h_1, h_2 \rangle$. On peut donc dire que la valuation du produit scalaire ne dépend pas du choix de la base hilbertienne. Par exemple, en supposant connu le fait que la transformée de Fourier est un opérateur unitaire, par exemple sur $L^2(\mathbb{R})$, on a l'égalité de Plancherel avec l'égalité précédente et de celle-ci découle l'égalité de Parseval en prenant $h_1 = h_2$. On a alors de façon générale le théorème suivant qui nous donne une condition nécessaire et suffisante pour que l'espace engendré par une famille $\{f_i\}$ soit dense dans H :

Théorème 2.1.3.1. Soit $\{f_i\}_I$ une suite d'éléments orthonormaux dans H muni d'un produit scalaire. Alors $\overline{\text{Vect}(\{f_i\}_I)} = H$ si et seulement si

$$\sum_I |\langle f, f_i \rangle|^2 = \|f\|^2, \quad \forall f \in H.$$

Cependant, comme on le verra dans la suite, il y a des situations dans lesquelles chercher à avoir une base orthonormale est trop restrictif, on cherchera donc à relâcher les conditions sur la définition d'une base.

Tout d'abord, si la famille est orthogonale mais n'est pas génératrice on a le résultat suivant :

Théorème 2.1.3.2. Soit $\{f_i\}_I$ une famille orthonormale de H . Alors,

$$\sum_I |\langle f, f_i \rangle|^2 \leq \|f\|^2, \quad \forall f \in H$$

.

On peut exprimer ce théorème en disant que l'analyse par une famille orthogonale n'ajoute pas d'énergie au vecteur analysé. Si la famille est génératrice,

Théorème 2.1.3.3. *Soit $\{f_i\}_I$ une famille génératrice normalisée. Alors,*

$$\|f\|^2 \leq \sum_I |\langle f, f_i \rangle|^2, \forall f \in H$$

.

On peut exprimer ce théorème en disant que l'analyse par une famille génératrice capture au moins l'énergie du vecteur analysé, cependant elle peut ne pas en capturer toute l'énergie. Au vu de ces résultats, on est amenés à considérer les définitions suivantes qui correspondent au fait de prendre des familles qui ne sont pas nécessairement orthogonales ou libres.

Définition 2.1.4. Pour une famille d'éléments $\{f_i\}_I$ de H , alors on dit que c'est

1. Une suite de *Bessel* si il existe une constante $M > 0$ telle que

$$\sum_I |\langle f, f_i \rangle|^2 \leq M \|f\|^2, \forall f \in H.$$

2. Un *frame* si il existe des constantes $M, m > 0$ telles que

$$m \|f\|^2 \leq \sum_I |\langle f, f_i \rangle|^2 \leq M \|f\|^2, \forall f \in H. \quad (2.3)$$

3. Une *base de Riesz* (ou *base inconditionnelle*) si il existe des constantes $M, m > 0$ telles que

$$m \sum |c_k|^2 \leq \left\| \sum c_k f_k \right\|^2 \leq M \sum |c_k|^2$$

pour n'importe quelle famille finie $\{c_k\}$.

En utilisant les trois théorèmes précédents on a les résultats suivants:

Proposition 2.1.5. • Une base orthonormale est une base de Riesz avec $m = M = 1$.

- Une base de Riesz est un frame dont les éléments sont linéairement indépendents.
- Un frame est une suite de Bessel dont les éléments sont générateurs.

Avec ces résultats, lorsque l'on dispose d'une famille de vecteurs $F = \{f_i\}_I$ on peut définir l'opérateur d'analyse

$$\theta_F(f) = \{\langle f, f_i \rangle\}_I$$

et de synthèse

$$\theta_F^*(\{c_i\}_I) = \sum_I c_i f_i.$$

Ainsi la composée des deux opérateurs nous donne un opérateur de projection dans l'espace vectoriel engendré par F :

$$\theta_F^* \circ \theta_F(f) = \sum_I \langle f, f_i \rangle f_i. \quad (2.4)$$

Tout d'abord on peut remarquer que, si F est une famille orthonormale, alors l'application précédente correspond à une projection orthogonale dans l'espace engendré par F . On va maintenant voir que si F est un frame *serré* (ou *équilibré*) (c'est à dire avec des constantes m, M égales), alors on dispose d'une formule analogue à 2.4 qui nous donne une projection orthogonale. L'intérêt de cela étant que, si F est génératrice de l'espace entier H , alors la projection orthogonale correspond à une formule de reconstruction. Supposons ainsi que l'on ait $m = M$, on a d'après 2.3,

$$\sum_I |\langle f_j, f_i \rangle|^2 = M \|f_j\|^2.$$

Posons $\pi = \frac{1}{M} \theta_F^* \circ \theta$ et vérifions que c'est une projection orthogonale, soit $f \in Vect(F)$, alors $f = \sum_J \lambda_j f_j$ avec $J \subset I$, d'où,

$$\langle f, f_k \rangle = \sum_J \lambda_j \langle f_j, f_k \rangle = \lambda_k \langle f_k, f_k \rangle + \sum_{j \in J - \{k\}} \lambda_j \langle f_j, f_k \rangle$$

$$\pi(f) = \frac{1}{M} \sum_I \langle f, f_i \rangle f_i = \frac{1}{M} \sum_J \lambda_j \sum_I \langle f_j, f_i \rangle f_i$$

et pour conclure, on projète $\pi(f)$, sur chaque composante f_k et on obtient

$$\begin{aligned} \langle \pi(f), f_k \rangle &= \frac{1}{M} \sum_J \lambda_j \sum_I \langle f_j, f_i \rangle \langle f_i, f_k \rangle = \frac{1}{M} \lambda_k \sum_I |\langle f_k, f_i \rangle|^2 + \frac{1}{M} \sum_{j \in J - \{k\}} \lambda_j \sum_I \langle f_j, f_i \rangle \langle f_i, f_k \rangle \\ &= \frac{1}{M} \lambda_k M \langle f_k, f_k \rangle + \frac{1}{M} \sum_{j \in J - \{k\}} \lambda_j M \langle f_j, f_k \rangle \\ &= \langle f, f_k \rangle. \end{aligned}$$

On a donc, si f est dans l'espace engendré par F , que la projection ne change pas les coordonnées de f . Sinon, si f est dans l'orthogonal de F , alors chacune de ses composantes est orthogonale à tous les f_i , donc f est dans le noyau de π . Ainsi, π est bien une projection orthogonale dans F .

On dispose donc d'une formule de reconstruction qui est valable pour tout f qui est dans l'espace engendré par F

$$f = \frac{1}{M} \sum_{f_i \in F} \langle f, f_i \rangle f_i. \quad (2.5)$$

Cependant cette formule peut sembler ajouter un inconvénient avec la constante $1/M$ par rapport à la formule de reconstruction dans une base orthonormale (cas $m = M = 1$ d'après la combinaison des théorèmes 2.1.3.2 et 2.1.3.3). Nous allons voir ci-dessous en quoi avoir un coefficient de frame serré $M > 1$ permet d'améliorer la stabilité de la formule de reconstruction, on appellera un tel frame un frame redondant.

Avant de poursuivre vers l'étude de frames plus sophistiqués dans les parties suivantes, considérons un frame élémentaire. On se place dans \mathbb{R}^2 et on considère la famille de vecteurs $\Phi = \{\varphi_1 = (0, 1), \varphi_2 = (-\frac{\sqrt{3}}{2}, -\frac{1}{2}), \varphi_3 = (\frac{\sqrt{3}}{2}, -\frac{1}{2})\}$. Cette famille n'est clairement pas orthogonale, ses éléments sont libres deux à deux, mais la famille n'est pas libre (en effet, $\varphi_1 + \varphi_2 + \varphi_3 = 0$). Cependant, cette famille forme un frame, pour le vérifier, prenons $x = (x_1, x_2)$ un élément de \mathbb{R}^2 , alors,

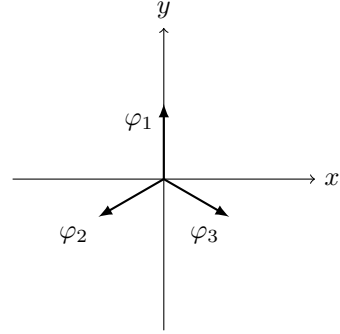


Figure 2.1: Le frame équilibré $\Phi = \{\varphi_1, \varphi_2, \varphi_3\}$ de \mathbb{R}^2

$$\begin{aligned} \sum_{i=1}^3 |\langle x, \varphi_i \rangle|^2 &= |x_2|^2 + \left| -\frac{\sqrt{3}}{2}x_1 - \frac{1}{2}x_2 \right|^2 + \left| \frac{\sqrt{3}}{2}x_2 - \frac{1}{2}x_1 \right|^2 \\ &= \frac{3}{2}(|x_1|^2 + |x_2|^2) = \frac{3}{2}\|x\|_2^2. \end{aligned}$$

Ainsi, les vecteurs $(\varphi_1, \varphi_2, \varphi_3)$ forment un frame équilibré de constante $M = \frac{3}{2}$ de \mathbb{R}^2 et on a donc la formule de reconstruction :

$$x = \frac{1}{M} \left(\sum_{i=1}^3 \langle x, \varphi_i \rangle \varphi_i \right) = \frac{2}{3} (\langle x, \varphi_1 \rangle \varphi_1 + \langle x, \varphi_2 \rangle \varphi_2 + \langle x, \varphi_3 \rangle \varphi_3) \quad (2.6)$$

Ainsi, comme aperçu par Jean Morlet dès 1986 ([7]) travailler avec des frames permet, en pratique, de pouvoir stocker des coefficients de frame avec moins de précision, donc par exemple en mettant à 0 les coefficients proches de 0, ou bien en admettant une erreur de quantification plus importante.

Remarque 2.1.6. La démonstration proposée ici poursuit l'exemple donné par Ingrid Daubechies, l'exemple présenté dans [7] est dans le cadre du frame 2.1, on prolonge ici ce raisonnement aux frames finis serrés arbitraires. En cherchant à prolonger ce résultat un résultat intermédiaire était nécessaire, le lemme 2.1.7, une preuve originale et élémentaire est proposée ici. Afin de vérifier que le résultat et la preuve étaient corrects, le résultat a été cherché dans la littérature publiée sur le sujet et trouvé dans le livre de Stéphane Mallat [22], théorème 5.2. On reproduira et discutera du théorème de Stéphane Mallat qui est plus général et dont la démonstration est plus simple après la démonstration du lemme 2.1.7.

Voyons comment formaliser l'observation de Jean Morlet, on considère un frame $F = (f_i)_{i=1,\dots,N}$ serré (c'est-à-dire $m = M$) redondant (c'est-à-dire $m > 1$). Ce frame engendre un espace vectoriel $Vect(F)$ de dimension $d \leq N$ car les éléments de F n'ont pas à être linéairement indépendants. Fixons un élément $f \in Vect(F)$. Tout d'abord, exprimons f dans une base orthonormale $(e_i)_{i=1,\dots,d}$, on a ainsi la formule standard:

$$f = \sum_{i=1}^d \langle f, e_i \rangle e_i = \sum_{i=1}^d c_i e_i.$$

Puis perturbons les coefficients de la façon suivante: fixons un $\epsilon > 0$ qui nous permettra de contrôler la taille de la perturbation introduite, et prenons d variables aléatoires réelles indépendantes $(\alpha_i)_{i=1,\dots,d}$ de moyenne nulle et de variance égale à 1. On peut alors perturber chaque coefficient c_i en y ajoutant $\epsilon \alpha_i$, on peut alors calculer l'erreur de reconstruction moyenne

$$\begin{aligned} \mathbb{E} \left(\left\| f - \sum_{i=1}^d (c_i + \epsilon \alpha_i) e_i \right\|_2^2 \right) &= \mathbb{E} \left(\sum_{i=1}^d \epsilon^2 \alpha_i^2 \right) \\ &= \epsilon^2 \sum_{i=1}^d \mathbb{E}(\alpha_i^2) = \epsilon^2 d. \end{aligned}$$

Ensuite, appliquons la même altération aux coefficients de f dans le frame F afin de calculer l'erreur de reconstruction moyenne. Considérons la formule de reconstruction

$$f = \frac{1}{M} \sum_{i=1}^N \langle f, f_i \rangle f_i$$

et prenons N variables aléatoires réelles indépendantes $(\alpha_i)_{i=1,\dots,N}$ de moyenne nulle et de variance égale à 1.

$$\begin{aligned} \mathbb{E} \left(\left\| f - \frac{1}{M} \sum_{i=1}^N (\langle f, f_i \rangle + \epsilon \alpha_i) f_i \right\|_2^2 \right) &= \mathbb{E} \left(\left\| \frac{1}{M} \sum_{i=1}^N \epsilon \alpha_i f_i \right\|_2^2 \right) \\ &= \frac{\epsilon^2}{M^2} \mathbb{E} \left(\sum_{i=1}^N \alpha_i^2 \right) = \frac{\epsilon^2}{M^2} \sum_{i=1}^N \mathbb{E}(\alpha_i^2) \\ &= \frac{\epsilon^2 N}{M^2} \end{aligned} \tag{2.7}$$

Si on souhaite comparer les majorations obtenues dans le cas orthonormal et dans le cas d'un frame serré, il nous faut donc une relation entre : le nombre d'éléments dans un frame serré (N), la constante de frame (M) et la dimension de l'espace engendré par le frame (d).

Lemme 2.1.7. *Soit $\Phi = (\varphi_i)_{i=1,\dots,N}$ un frame avec des constantes de frame $m = M$ qui engendre $Vect(\Phi)$ un espace vectoriel de dimension d . Alors*

$$\frac{N}{M} \leq d. \tag{2.8}$$

Avec ce lemme, on peut donc comparer l'erreur de reconstruction moyenne après perturbation dans le cas d'un frame 2.7 et dans le cas d'une base orthonormale:

$$\frac{N}{M^2} \leq \frac{\epsilon^2 d}{M} < \epsilon^2 d. \quad (2.9)$$

Ainsi, en revenant au cas du frame de \mathbb{R}^2 2.1, on a que si on perturbe les coefficients l'erreur de reconstruction est améliorée d'un facteur $\frac{2}{3}$ par rapport à celle que l'on obtiendrait en utilisant une base orthonormale de \mathbb{R}^2 . Maintenant passons à la preuve originale du lemme:

Preuve. Afin de prouver ce résultat rappelons que l'on a l'opérateur d'analyse

$$A : \mathbb{R}^d \longrightarrow \ell^2(N) \quad (2.10)$$

$$f \longmapsto (\langle f, \varphi_i \rangle)_{i=1, \dots, N} \quad (2.11)$$

et le frame étant équilibré, la boule $B_d(0, 1)$ dans \mathbb{R}^d de rayon 1 est envoyée sur la boule $B_N(0, M)$ de $\ell^2(N)$. Maintenant, on considère l'opérateur de synthèse,

$$S : \ell^2(N) \xrightarrow{S} \mathbb{R}^d \quad (2.12)$$

$$x = (x_i)_{i=1, \dots, N} \longmapsto \sum_{i=1}^N x_i \varphi_i \quad (2.13)$$

et on considère, sans perte de généralité que chaque φ_i est normalisé. Ainsi chaque φ_i , on peut associer une suite $(\lambda_j^i)_{j=0, \dots, d} = (\langle \varphi_i, e_j \rangle)_{j=0, \dots, d}$ de norme 1 dans \mathbb{R}^d muni de la norme euclidienne telle que $\varphi_i = \sum_{j=1}^d \lambda_j^i e_j$ où e_j est une base orthonormale de \mathbb{R}^d . L'objectif est maintenant de réécrire $S(x)$ dans la base orthonormale, on a ainsi

$$\sum_{i=1}^N x_i \varphi_i = \sum_{i=1}^N x_i \sum_{j=1}^d \lambda_j^i e_j = \sum_{j=1}^d \sum_{i=1}^N x_i \lambda_j^i e_j = \sum_{j=1}^d c_j e_j, \quad (2.14)$$

où $c_j = \sum_{i=1}^N x_i \lambda_j^i$. On va maintenant majorer les c_j de façon uniforme, chaque λ_j^i correspond à la projection de φ_i sur e_j , chaque φ_i est dans la boule unité ainsi c_j correspond à la projection de tous les φ_i sur e_j , ainsi on a avec l'inégalité de Cauchy-Schwarz la majoration

$$c_j \leq \|x\|_2 \sqrt{M}. \quad (2.15)$$

On a maintenant la majoration,

$$\|S(x)\|_2^2 = \left\| \sum_{j=1}^d c_j e_j \right\|_2^2 = \sum_{j=1}^d c_j^2 \leq dM \|x\|_2^2. \quad (2.16)$$

On va maintenant chercher une minoration de $\|S(x)\|_2^2$,

$$\|S(x)\|_2^2 = \left\| \sum_{i=1}^N x_i \sum_{j=1}^d \lambda_j^i e_j \right\|_2^2 \geq N \|x\|_2^2 \min_{i=1, \dots, N} \left\| \sum_{j=1}^d \lambda_j^i e_j \right\|_2^2 = N \|x\|_2^2 \min_{i=1, \dots, N} \sum_{j=1}^d (\lambda_j^i)^2. \quad (2.17)$$

Or,

$$\sum_{i=1}^d (\lambda_j^i)^2 = \sum_{i=1}^d |\langle \varphi_j, e_i \rangle|^2 = \|\varphi_j\|_2^2 = 1. \quad (2.18)$$

On peut donc combiner les inégalités obtenues et on a

$$N\|x\|_2^2 \leq \|S(x)\|_2^2 \leq dM\|x\|_2^2 \quad (2.19)$$

et ainsi

$$\frac{N}{M} \leq d. \quad (2.20)$$

□

Comme discuté dans la remarque 2.1.6, un résultat plus général est présenté dans le livre de Stéphane Mallat [22], le voici :

Théorème 2.1.7.1 (Mallat). *Soit $\Phi = (\varphi_i)_{i=1,\dots,N}$ un frame avec des constantes de frame m, M qui engendre $\text{Vect}(\Phi)$ un espace vectoriel de dimension d . Alors*

$$m \leq \frac{N}{d} \leq M. \quad (2.21)$$

Si le frame est serré, $m = M = \frac{N}{d}$.

Premièrement, remarquons que la dernière égalité est une conséquence de l'inégalité avec $m = M$.

Ensuite, remarquons que l'on peut réécrire 2.21 de façon équivalente :

$$d \leq \frac{N}{m} \quad \text{et} \quad \frac{N}{M} \leq d. \quad (2.22)$$

La deuxième inégalité étant celle du lemme 2.1.7, pour passer de ce lemme au théorème précédent, il faut donc montrer que cette inégalité dans le cas d'un frame qui n'est pas équilibré. En étudiant la preuve du lemme qui est donnée ici, on voit que le fait que le frame soit équilibré intervient dans la première partie, dans laquelle on cherche une majoration. Comme on cherche une majoration, en étudiant la preuve on peut voir qu'elle fonctionne aussi si $m \leq M$. Donc pour arriver au résultat de Stéphane Mallat, il reste à montrer l'inégalité

$$d \leq \frac{N}{m}. \quad (2.23)$$

On peut se convaincre, et montrer, qu'une preuve très similaire à celle du lemme est possible qui permet de prouver cette dernière inégalité et le théorème de Stéphane Mallat. En effet, il y a certaines symétries que l'on peut utiliser pour obtenir les inégalités dans l'autre sens. Il semble donc qu'un argument plus général puisse exister pour faciliter cette preuve. On présente donc la preuve de Stéphane Mallat qui utilise naturellement la trace d'un opérateur et les valeurs propres.

Preuve. Remarquons tout d'abord que la définition de l'opérateur de synthèse S par rapport à A que l'on fait ici correspond exactement à prendre l'opérateur adjoint de A , que l'on note A^* . Celui-ci vérifie pour tout $f \in Vect(F)$ et pour tout $x \in Im(A) \subset \ell^2(N)$,

$$\langle Af, x \rangle = \langle f, A^*x \rangle. \quad (2.24)$$

On peut alors remarquer

$$\|Af\|^2 = \sum_{i=1, \dots, N} \langle f, f_i \rangle^2 = \langle A^*Af, f \rangle. \quad (2.25)$$

On peut donc réécrire la condition de frame 2.3 sous la forme:

$$m\|f\|^2 \leq \langle A^*Af, f \rangle \leq M\|f\|^2. \quad (2.26)$$

On voit alors que les bornes de frame m et M correspondent à la plus petite et à la plus grande valeur propre de A^*A en dimension finie, aussi appelées valeurs singulières de A .

On peut maintenant prouver le théorème en quelques lignes : les valeurs propres de A sont donc entre m et M , ainsi

$$dm \leq Tr(A^*A) \leq dM \quad (2.27)$$

car A^*A agit sur un espace isométrique à $\ell^2(d)$. En utilisant le fait que $Tr(A^*A) = Tr(AA^*)$, on a:

$$Tr(AA^*) = \sum_{i=1, \dots, N} |\langle f_i, f_i \rangle|^2 = N \quad (2.28)$$

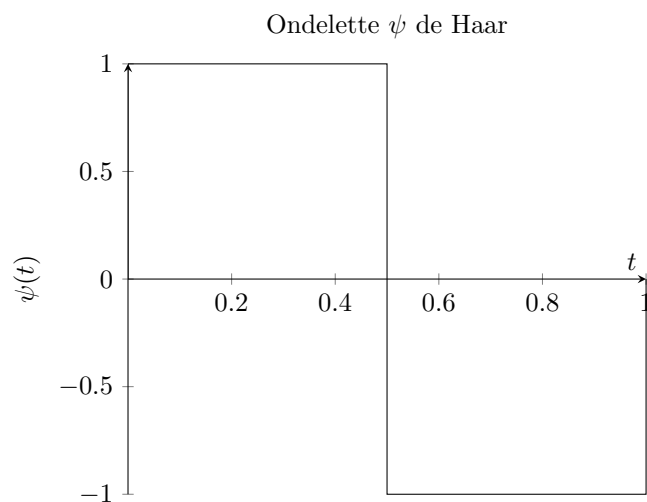
et on a alors prouvé le théorème en insérant le précédent résultat dans 2.27. \square

En comparant cette preuve et celle du lemme, on voit que la trace permet de concentrer le raisonnement du lemme en quelques lignes. On voit alors l'efficacité de la trace et l'importance des valeurs propres pour prouver ce type de résultat. Ainsi, certains des raisonnements dans la suite de ce mémoire utiliseront ce point de vue spectral.

Avant de poursuivre vers des exemples de frames, remarquons que dans cette partie un point de vue plus général aurait pu être possible. En effet, on aurait pu considérer travailler dans des espaces de Banach plutôt que de Hilbert, on aurait pu introduire d'autres notions de bases telles que les bases de Hamel, Schauder, Auerbach ou Markushevich plutôt que de Riesz ou Hilbert. Le choix de travailler dans un espace de Hilbert a été fait car c'est un cadre suffisant pour le reste du mémoire avec des propriétés très similaire en dimension finie et infinie. Il est important de remarquer que ces choix ont de l'importance seulement en dimension infinie, les notions étant équivalentes en dimension finie.

Cependant des généralisations auraient pû être possible en utilisant des notions plus restrictives, cela est notamment motivé par le fait que certaines ondelettes, par exemple, celle de Haar peuvent fournir une base raisonnable (de Schauder) de $L^p([0, 1])$, pour tout $1 \leq p < \infty$, contrairement à la base de Fourier qui n'est pas une base raisonnable pour $p = 1$. Ici on dit que la base de Fourier n'est pas raisonnable car il existe des fonctions dans $L^1([0, 1])$ dont la série de Fourier ne converge pas en norme L^1 .

Il semble qu'une façon par laquelle on aurait pu généraliser l'étude d'opérateurs ayant des propriétés similaires aux frames dans des espaces de Banach aurait été de donner une place centrale à la trace, en considérant la classe des opérateurs à trace, aussi appelés opérateurs nucléaires. A cette théorie on peut associer le nom de Grothendieck avec l'article [2]. Pour une introduction aux opérateurs à trace par les frames on peut consulter [20], par leur similarité à l'analyse complexe on peut consulter [19], pour un point de vue historique sur leur développement on peut consulter [21].



2.1.8 Exemples de frames (Fourier, ondelettes)

Historiquement, au début du XXème siècle, David Hilbert souhaitait savoir si d'autres bases de $L^2([0, 1])$ pouvaient être trouvées que la base de Fourier. C'est ainsi que sous sa direction Alfréd Haar introduit dans sa thèse [14] en 1909 la base orthogonale de $L^2([0, 1])$ donnée par ???. Jusqu'aux années 1980, aucune autre ondelette n'a été trouvée, c'est sous l'impulsion des travaux de Jean Morlet et Alexandre Grossman que le domaine des ondelettes a commencée à être exploré. Ensuite Yves Meyer a reconnu dans leurs travaux des formules analogues à des résultats d'analyse harmonique, cependant à ce moment là, tous les frames d'ondelettes qui étaient construits étaient redondants. Yves Meyer (voir [6] pour plus de détails sur l'ensemble de ces développements) a donc souhaité prouver que c'était nécessaire que les frames d'ondelettes soient redondants, et donc pas orthonormaux. En essayant de prouver cela par l'absurde, en supposant l'existence d'un frame d'ondelette orthonormal, au lieu d'obtenir une contradiction il obtint une construction d'une famille d'ondelettes orthonormales [18]. A partir de là de nombreuses constructions d'ondelettes furent faites avec différentes propriétés de régularité, de support, de décroissance,... Mentionnons tout de même la construction d'ondelettes avec des moments nuls à support compact par Ingrid Daubechies [8] ainsi que l'algorithme de transformée en ondelette discret et le concept d'analyse multi-résolution de Stéphane Mallat [17] qui ont permis aux ondelettes de devenir un outil maintenant très répandu dans les mathématiques appliquées.

Les concepts d'analyse-temps fréquence et de localisation développés dans la théorie

des ondelettes sont aujourd'hui développés en mathématiques pures dans le cadre de l'analyse micro-locale. (TODO : ajouter ref + d'autres développements des idées des ondelettes en maths pures)

Etudions donc ces ondelettes:

Définition 2.1.9. On dit que $\psi : \mathbb{R} \rightarrow \mathbb{R}$ est une ondelette génératrice si

$$\{\psi_{j,k} := 2^{j/2} \psi(2^j \cdot -k)\}_{j,k \in \mathbb{Z}} \quad (2.29)$$

est une famille génératrice de $L^2(\mathbb{R})$. On appellera¹ une base engendrée par une telle fonction ψ une base d'ondelettes.

Remarque 2.1.10. Dans la définition des ondelettes $\psi_{j,k}$ engendrées par ψ , le coefficient j correspond au facteur d'échelle², en raison du facteur 2^j devant la variable, au fur et à mesure que j augmente, l'ondelette est parcourue de plus en plus vite. Ainsi, augmenter j revient à augmenter la fréquence de ψ , c'est à dire d'éloigner le support de $\hat{\psi}$ de l'origine. Le coefficient k correspond à une translation de l'ondelette ψ_j , en ce sens, l'analyse par ondelette, permet une analyse à la fois en temps (par rapport à k) et en fréquence (par rapport à j).

En effet, de façon plus précise et formelle, on a par l'action de la transformée de Fourier sur les dilatations:

$$\widehat{\psi_{j,0}}(\omega) = 2^{-\frac{j}{2}} \hat{\psi}\left(\frac{\omega}{2^j}\right) \quad (2.30)$$

et par l'action de la transformée de Fourier sur les translations:

$$\widehat{\psi_{j,k}}(\omega) = 2^{-\frac{j}{2}} \hat{\psi}\left(\frac{\omega}{2^j}\right) e^{-2i\pi 2^{-j} k \omega}. \quad (2.31)$$

On peut voir que l'analogie temps-fréquence dans le cadre des ondelettes implique en un certain sens le point de vue temps-fréquence de l'analyse Fourier. En effet, si on prend ψ l'ondelette de Shannon (voir ??), alors sa transformée de Fourier $\hat{\psi}$ est la fonction indicatrice sur $[-2\pi, \pi] \cup [\pi, 2\pi]$. Donc, à échelle j fixée, on a que $\widehat{\psi_{j,k}}$ est supporté sur $B_j := [-2\pi 2^{j+1}, -\pi 2^j] \cup [\pi 2^j, \pi 2^{j+1}]$ et vaut:

$$\widehat{\psi_{j,k}}(\omega) = 2^{-\frac{j}{2}} e^{-2i\pi 2^{-j} k \omega}. \quad (2.32)$$

Regardons maintenant la projection de $f \in L^2(\mathbb{R})$ sur les translatés de ψ_j , on a:

$$\langle f, \psi_{j,k} \rangle = \langle \hat{f}, \widehat{\psi_{j,k}} \rangle = \int_{B_j} \hat{f}(\omega) 2^{-\frac{j}{2}} e^{2i\pi 2^{-j} k \omega} d\omega \quad (2.33)$$

¹On considérera par la suite des frames d'ondelettes ou bien des bases de Riesz d'ondelettes

²Du point de vue des notations, on considère que j tend vers l'infini, signifie que ψ_j analyse les hautes fréquences, ce choix de notation n'est pas uniforme dans la littérature, par exemple (TODO :ajouter ref) Mallat et Daubechies, utilisent $-j$ par rapport à nos notations. Par contre les notations utilisées correspondent à celles de Jaffard et Meyer. Mais cependant tous les résultats sont bien entendu équivalents.

Figure 2.2:
L'ondelette $\psi(t) = 2\text{sinc}(2t) - \text{sinc}(t)$
dite de Shannon avec
différentes dilata-
tions dyadiques ψ_j .
Lorsque j augmente,
la fonction est parcou-
rue plus rapidement
d'un facteur 2^j .

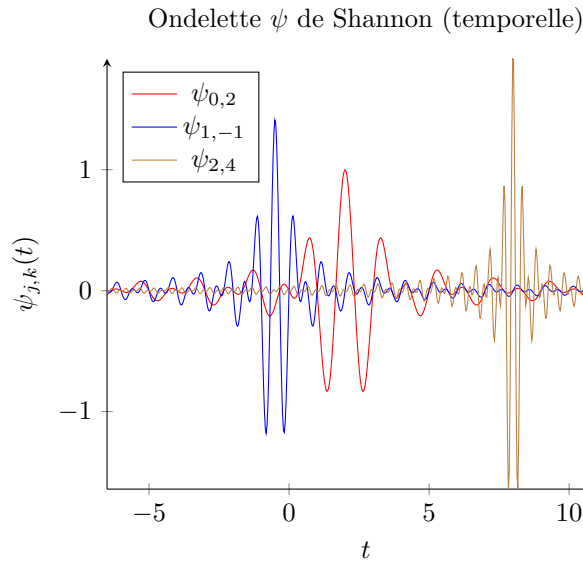
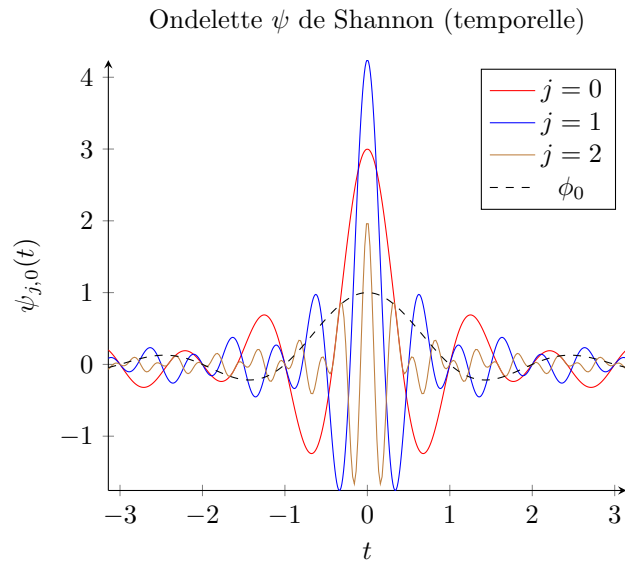


Figure 2.3: Des ver-
sions traduites $\psi_{j,k}$
des ondelettes ψ_j du
graphe précédent.
Lorsque j augmente,
la fonction est parcou-
rue plus rapidement
d'un facteur 2^j .

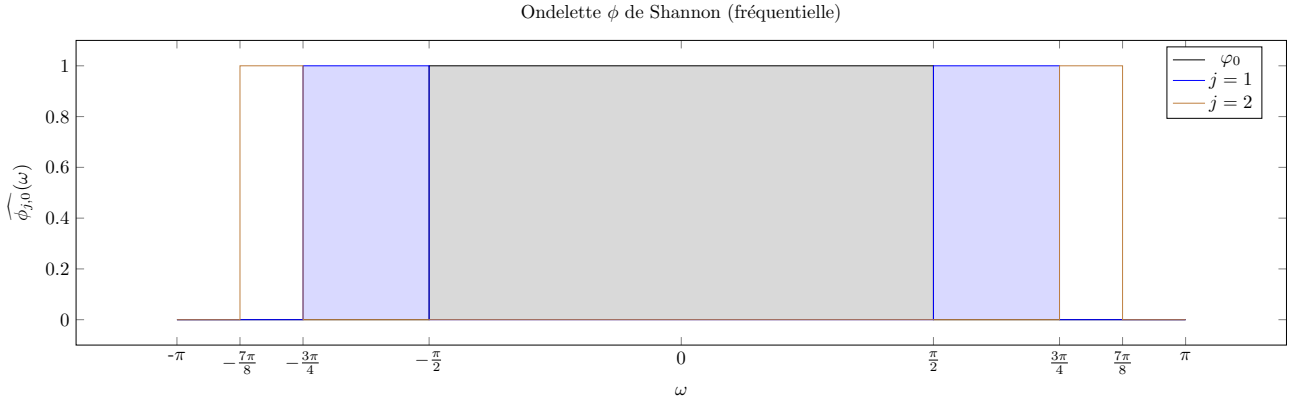


Figure 2.4: La transformée de Fourier $\hat{\psi}$ de l'ondelette de Shannon Ψ . L'ondelette d'échelle ϕ_0 est un filtre passe-bande $[-\frac{\pi}{2}, \frac{\pi}{2}]$, les ondelettes ψ_j sont des filtres passe bande qui recouvrent le reste de l'intervalle avec des supports disjoints de taille 2^{-j} . On peut ainsi vérifier toutes les propriétés de l'analyse multi-résolution pour l'ondelette de Shannon sur les fonctions à bande limitée (c'est à dire, dont le support de la transformée de Fourier est borné).

donc en sommant par rapport à k on a que la projection à j fixé correspond à un développement en série de Fourier de f avec les fréquences dans B_j . Donc l'intuition que lorsque j augmente, la fréquence augmente, coïncide exactement avec la notion de fréquence de la théorie de Fourier lorsque on considère l'ondelette de Shannon.

Donc étant donnée une famille d'ondelettes $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$, comme on l'a vu on peut associer à une fonction $f \in L^2(\mathbb{R})$, ses coefficients d'ondelettes

$$Wf = (\langle f, \psi_{j,k} \rangle)_{j,k \in \mathbb{Z}} \quad (2.34)$$

et on se demande alors si à partir de ces coefficients on peut reconstruire f . Cela revient ainsi à déterminer si la famille d'ondelettes est un frame d'ondelette. Ici on ne cherchera pas à énumérer et à vérifier des frames d'ondelettes, un très grand nombre de frames d'ondelettes existent (ajouter ref.), on admet ainsi pour l'instant l'existence des frames d'ondelettes. Un peu plus bas nous verrons et démontrerons un théorème qui permet de construire de nombreux frames d'ondelettes, ce théorème fournira également une extension des ondelettes à $L^2(\mathbb{R}^d)$ et avec une version qui montre que la présence de puissances de 2 dans la définition des ondelettes revient à un choix d'échantillonnage. Supposons ainsi que l'on dispose d'un frame d'ondelettes et que ce frame est équilibré (c'est à dire que les bornes de frame m et M sont égales), alors on dispose d'une formule de reconstruction (d'après Daubechies 3.2.2)

$$f = \frac{1}{M} \sum_{j,k} \langle f, \psi_{j,k} \rangle \psi_{j,k}. \quad (2.35)$$

Cependant bien que la formule précédente permette une reconstruction elle suppose de parcourir des indices sur \mathbb{Z} ce qui pourrait créer des complications concernant la conver-

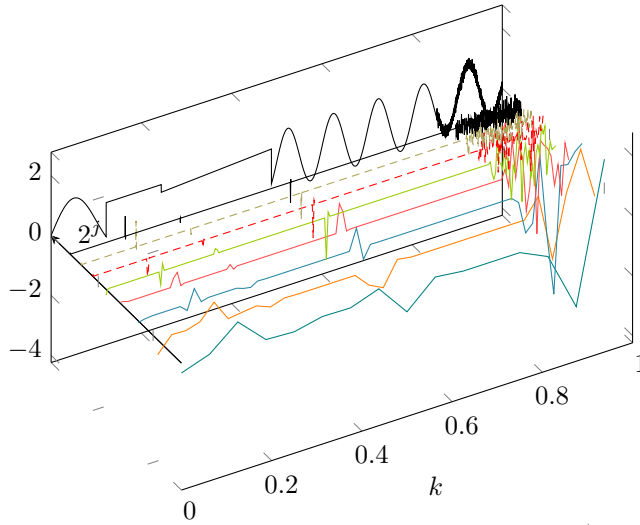


Figure 2.5: Les coefficients d'ondelette (pour l'ondelette de Daubechies D4) d'un signal présentant des zones de différentes régularité. Les coefficients les plus proches du signal correspondent aux échelles les plus fines. On remarque que ces coefficients sont très petits dans les zones sans discontinuités. On peut voir que les hautes valeurs de ces coefficients (proches des discontinuités) se propagent vers les plus grandes échelles en s'étalant.

gence (d'un point de vue théorique ou pratique). On va ici très rapidement introduire la notion d'analyse multi-échelle qui permet de simplifier la formule de reconstruction. Construisons ici une analyse multi-échelle (ici de $L^2(\mathbb{R})$), considérons tout d'abord une suite d'espaces emboîtés satisfaisant

$$\{0\} = \lim_{j \rightarrow -\infty} \bigcap_j V_i \subset \cdots \subset V_{-1} \subset V_0 \subset \cdots \subset V_i \subset V_{i+1} \subset \cdots \subset \lim_{j \rightarrow \infty} \bigcup_{-\infty}^j V_i = L^2(\mathbb{R}).$$

L'intérêt d'avoir une telle suite d'espaces emboîtés est que étant donnée une fonction $f \in L^2(\mathbb{R})$, on peut considérer sa projection orthogonale dans un V_i , on a alors une approximation f_i de f dans V_i , si on souhaite améliorer l'approximation de f il suffit alors de remonter dans ces espaces emboîtés pour avoir une reconstruction avec une précision arbitraire. Introduisons maintenant la propriété qui va permettre de voir cette suite d'espaces comme une analyse multi-échelle

$$f(\cdot) \in V_j \iff f\left(\frac{\cdot}{2^j}\right) \in V_0, \quad (2.36)$$

c'est à dire que les fonctions d'un espace V_j sont des versions dilatées d'un facteur 2^{-j} des fonctions de l'espace V_0 . Ajoutons maintenant la condition que V_0 contient toutes

les translations entières de ses éléments, c'est à dire

$$f \in V_0 \iff f(\cdot - n) \in V_0 \forall n \in \mathbb{Z}. \quad (2.37)$$

Ainsi, une fonction qui appartient à V_j s'écrit comme une combinaison linéaire de versions translatées et dilatées de fonctions appartenant à V_0 . De plus, quelque soit j on peut prendre W_j le complémentaire orthogonal de V_j dans V_{j+1} ,

$$f = \pi_{V_0}(f) + \sum_{i>0} \pi_{W_i}(f) \quad (2.38)$$

donc afin d'avoir une formule de reconstruction, il suffit de connaître un frame de V_0 et de même pour chaque W_j . On peut maintenant revenir aux ondelettes, on considère que l'on connaît $\varphi \in L^2(\mathbb{R})$ tel que $(\varphi_k)_{k \in \mathbb{Z}}$, en notant φ_k les translatés par k de φ , soit une base de V_0 et $(\psi_{j,k})_{j \in \mathbb{N}^*, k \in \mathbb{Z}}$ un frame de l'orthogonal de V_0 dans $L^2(\mathbb{R})$, on pose alors $W_j = W_{j-1} \oplus \text{Vect}(\{\psi_{j,k}\}_{k \in \mathbb{Z}})$ et on obtient ainsi que l'analyse multi-échelle ainsi construite fournit une formule de reconstruction³

$$f = \sum_{k \in \mathbb{Z}} \langle f, \varphi_k \rangle \varphi_k + \sum_{j=1}^{+\infty} \sum_{k \in \mathbb{Z}} \langle f, \psi_{j,k} \rangle \psi_{j,k}. \quad (2.39)$$

On appelle l'application φ_0 ondelette d'échelle. On a vu que les familles orthonormales forment des frames, on a donc la proposition :

Proposition 2.1.11. 1. La transformée de Fourier discrète est un frame pour \mathcal{F} .

2. Les ondelettes forment un frame pour \mathcal{F} .

Preuve 2.1.12. 1. On utilisera la transformée de Fourier discrète écrite sous forme de matrice unitaire, qui est une matrice unitaire de Vandermonde avec les racines de l'unité en coefficients. Avec de l'algèbre on montre l'orthonormalité des colonnes.

2. Application du théorème suivant.

Théorème 2.1.12.1. Soit $\mathcal{Q} \subset \mathcal{R}^d$ un ensemble de mesure finie, $h \in L^2(\mathbb{R}^d)$ et $\mathcal{A} = \{A_j \in GL_d(\mathbb{R})\}_J$ une famille de matrices inversibles.

Pour tout $j \in J$, on pose $B_j = (A_j^T)^{-1}$, $S_j = A_j^T Q$, $h_j = h(B_j \cdot)$ et soit $\mathcal{S} = \{S_j\}_J$.

On suppose que \mathcal{S} est un recouvrement de \mathbb{R}^d , \mathcal{H} est une partition de Riesz de l'unité avec des bornes p et P et que $\text{Supp}(h) \subset Q$.

Soit $X = \{x_{j,k} \in \mathbb{R}^d : j \in J, k \in K\}$ tel que quelque soit $j \in J$, l'ensemble $\{e_{x_{j,k}} \chi_Q\}_K$ forme un frame pour \mathcal{K}_Q avec des bornes m_j et M_j .

Si $m := \inf_J m_j > 0$ et $M = \sup_J M_j < \infty$, alors la collection

$$\{|\det A_j|^{1/2} \psi(A_j x - x_{j,k})\}_{j,k}$$

³Dans la formule de reconstruction les deux sommes sur \mathbb{Z} ne sont pas problématiques car les fonctions considérées sont dans $L^2(\mathbb{R})$ donc avec une décroissance suffisamment rapide, donc seulement un nombre fini de $\langle f, \varphi_{0,k} \rangle$ sont différents de 0 si φ_0 a une décroissance suffisamment rapide (et de même pour chaque ψ_j).

forme un frame d'ondelettes de $L^2(\mathbb{R}^d)$ avec des bornes mp et MP , engendré par une seule fonction ψ où ψ est la transformée de Fourier inverse de h .

Preuve 2.1.13. Voir [1] pour la preuve et les définitions, je les ajouterai ici et au dessus plus tard

2.2 Décroissance des coefficients et régularité

2.2.1 Approximation linéaire et régularité

On s'intéresse dans cette partie au lien entre une fonction $f \in L^2(]0, 1[)$ et son approximation dans une base. On verra un résultat reliant la décroissance des coefficients de la fonction dans une base fixée et la vitesse de convergence de la reconstruction. On verra ensuite à l'aide de ce résultat, que pour la base de Fourier (et resp. certaines bases d'ondelettes), on obtient des formules de reconstruction pour les fonctions dérivables (et resp. pour les fonctions Lipschitziennes) avec une erreur de reconstruction qui décroît rapidement.

On considère ainsi un espace d'approximation de fonctions $U_N \subset L^2([0, 1])$. Par construction, la meilleure approximation linéaire de f dans U_N , est la projection orthogonale f_N de f dans U_N , qui peut être obtenue à l'aide de la base biorthogonale de synthèse associée $(\tilde{\phi}_k)_{k=1, \dots, N}$ et la formule de reconstruction :

$$f_N = \sum_{k=0}^{N-1} \langle f, \phi_k \rangle \tilde{\phi}_k. \quad (2.40)$$

Afin de mesurer l'erreur d'approximation par rapport à f , on considère une base $\mathcal{B} = \{g_k\}_{k \in \mathbb{N}}$ de l'espace $L^2([0, 1])$ entier à laquelle on ajoute la condition de contenir une famille $(g_k)_{k \in I}$, avec $\#I = N$ qui forme une base de l'espace d'approximation U_N . On peut ainsi écrire, en réordonnant la famille (g_k) , $f_N \in U_N$ dans cette base :

$$f_N = \sum_{k=0}^{N-1} \langle f, g_k \rangle g_k \quad (2.41)$$

et les $\{g_k\}_{k \in \mathbb{N}}$ formant une base de $L^2([0, 1])$, on peut écrire f dans cette base:

$$f = \sum_{k=0}^{+\infty} \langle f, g_k \rangle g_k. \quad (2.42)$$

On obtient donc que la partie orthogonale à la famille $\{\phi_k\}_{k=0, \dots, N-1}$, est celle analysée par $\{g_k\}_{k \geq N}$. C'est à dire,

$$f - f_N = \sum_{k=N}^{+\infty} \langle f, g_k \rangle g_k \quad (2.43)$$

et la mesure de l'erreur d'approximation avec N coefficients est donc

$$\varepsilon_l(N, f) = \|f - f_N\|^2 = \sum_{k=N}^{+\infty} |\langle f, g_k \rangle|^2. \quad (2.44)$$

Comme on a supposé que $f \in L^2([0, 1])$ et que la famille (g_k) est génératrice, on a que l'erreur d'approximation tend vers 0 lorsque N augmente. On va maintenant s'intéresser au théorème suivant de Stéphane Mallat qui relie la décroissance des coefficients de f dans la base de $L^2([0, 1])$ à la vitesse de décroissance de l'erreur d'approximation de la fonction.

Théorème 2.2.1.1. *Soit $r > 1/2$, il existe des constantes $A, B > 0$ telles que si*

$$\sum_{k=0}^{+\infty} |k|^{2r} |\langle f, g_k \rangle|^2 < \infty, \quad (2.45)$$

alors on a

$$A \sum_{k=0}^{+\infty} k^{2r} |\langle f, g_k \rangle|^2 \leq \sum_{N=0}^{+\infty} N^{2r-1} \varepsilon_l(N, f) \leq B \sum_{k=0}^{+\infty} k^{2r} |\langle f, g_k \rangle|^2 \quad (2.46)$$

et ainsi on a $\varepsilon_l(N, f) = o(N^{-2r})$.

Preuve. On développe le terme au centre de l'égalité de la façon suivante

$$\sum_{N=0}^{\infty} N^{2r-1} \varepsilon_l(N, f) = \sum_{N=0}^{\infty} \sum_{k=N}^{\infty} N^{2r-1} |\langle f, g_k \rangle|^2 = \sum_{k=0}^{+\infty} |\langle f, g_k \rangle|^2 \sum_{N=0}^k N^{2r-1}.$$

Puis on majore des deux côtés avec

$$\int_0^M x^{2r-1} dx \leq \sum_{N=0}^m N^{2r-1} \leq \int_1^{m+1} x^{2r-1} dx. \quad (2.47)$$

En calculant les deux intégrales on déduit

$$Am^{2r} \leq \sum_{N=0}^m N^{2r-1} \leq Bm^{2r} \quad (2.48)$$

où A et B dépendent seulement de r , ce qui nous donne 2.46. Montrons maintenant $\varepsilon_l(N, f) = o(N^{-2r})$, remarquons tout d'abord que $\varepsilon_l(N, f)$ est décroissant par rapport à la première variable, on déduit de cela

$$\varepsilon_l(N, f) \sum_{m=N/2}^{N-1} m^{2r-1} \leq \sum_{m=N/2}^{+\infty} m^{2r-1} \varepsilon_l(m, f)$$

on a donc avec le calcul précédent que le terme de droite converge quel que soit le choix de N , et ainsi

$$\lim_{N \rightarrow +\infty} \sum_{m=N/2}^{+\infty} m^{2r-1} \varepsilon_l(m, f) = 0$$

donc tous les termes de l'inégalité précédente tendent vers 0. De plus, $\sum_{m=N/2}^{N-1} m^{2r-1} \geq CN^{2r}$ et donc

$$\lim_{N \rightarrow \infty} \varepsilon_l(N, f) N^{2r} = 0.$$

□

On a ainsi démontré que si f appartient à l'espace

$$W_{\mathcal{B},r} = \{f : \sum_{m=0}^{+\infty} m^{2r} |\langle f, g_m \rangle|^2 < \infty\} \quad (2.49)$$

alors l'approximation linéaire dans la base \mathcal{B} décroît au moins comme N^{-2r} . On montrera dans la prochaine sections que si \mathcal{B} est une base de Fourier alors $W_{\mathcal{B},r}$ contient les fonctions r -différentiables. On montrera ensuite que si \mathcal{B} est une base d'ondelette avec une certaine propriété alors l'espace $W_{\mathcal{B},r}$ contient les fonctions α -Lipschitziennes pour $1 < \alpha < r$. Des énoncés réciproques existent aussi et des démonstrations de ceux-ci peuvent être trouvés dans (TODO :Ajouter Ref)

2.2.2 Décroissance des coefficients de Fourier

On considère ici $\mathcal{B} = (e_n)_{n \in \mathbb{Z}}$ la base de Fourier (voir 1.3.3) de $L^2(\mathbb{R})$. On peut ainsi définir l'espace $U_N = \{f \in L^2(\mathbb{R}) : |k| > N \implies \hat{f}(k) = 0\}$ sur lequel $\mathcal{B}_N = (e_n)_{|n| \leq N}$ est une base. Avec des mots, U_N est l'espace des fonctions qui ne sont portées par aucune exponentielle complexe de fréquence supérieure ou égale à N . Le théorème de Shannon nous indique que $2N$ fréquences permettent de séparer n'importe laquelle de ces fonctions, ainsi on a la formule de reconstruction. On dispose ainsi d'une formule de projection (et de reconstruction), soit $f \in L^2(\mathbb{R})$

$$f_N(t) = \sum_{|n| \leq N/2} \langle f, e_n \rangle e_n(t) \in U_N. \quad (2.50)$$

Ainsi f_N est une approximation linéaire de f , et f sera rapidement approximée si f n'a pas trop de hautes fréquences. Montrons maintenant que la vitesse de décroissance des coefficients de Fourier est liée à la régularité de la fonction. Tout d'abord, revenons à $L^2(\mathbb{R})$ considérons que f est dérivable, alors on a en intégrant par parties

$$\hat{f}'(\omega) \int_{-\infty}^{+\infty} f'(t) e^{i\omega t} dt = i\omega \int_{-\infty}^{+\infty} f(t) e^{i\omega t} dt = i\omega \hat{f}(\omega) \quad (2.51)$$

et en utilisant la formule de Plancherel on a

$$\|\hat{f}'\|_2^2 = \int_{-\infty}^{+\infty} |\omega|^2 |\hat{f}(\omega)|^2 d\omega = \int_{-\infty}^{+\infty} |f'(t)|^2 dt = \|f'\|_2^2. \quad (2.52)$$

On est ainsi amenés à définir une régularité dans $L^2(\mathbb{R})$, distincte de la dérivabilité en un point avec la définition suivante :

Définition 2.2.3. On dit que $f \in L^2(\mathbb{R})$ est différentiable au sens de Sobolev si

$$\int_{-\infty}^{+\infty} |\omega|^2 |\hat{f}(\omega)|^2 d\omega < \infty.$$

Et avec cette définition on peut définir pour n'importe quel $r > 0$ l'espace des fonction r -différentiables de Sobolev:

$$W^r(\mathbb{R}) = \{f \in L^2(\mathbb{R}) : \int_{\mathbb{R}} |\omega|^{2r} |\hat{f}(\omega)|^2 d\omega < \infty\}. \quad (2.53)$$

Ainsi d'après le théorème sur la vitesse d'approximation de Mallat, l'approximation linéaire dans la base de Fourier d'une application r -différentiable décroît plus vite que N^{-2r} .

2.2.4 Décroissance des coefficients d'ondelettes

On va montrer dans cette section que en imposant certaines conditions sur les ondelettes, alors il est possible de démontrer que la régularité au sens de Lipschitz, implique une décroissance des coefficients d'ondelettes. On pourra ensuite relier cette décroissance aux discussions de la fin de la partie précédente.

Tout d'abord posons les définitions dont nous aurons besoin dans cette partie,

Définition 2.2.5. Soit α tel qu'il existe un entier strictement positif r tel que $r - 1 \leq \alpha < r$. On dit qu'une fonction f est α -Lipschitz en t_0 si il existe une constante $C > 0$ et un polynôme P_{t_0} de degré strictement inférieur à r , tels que pour tout t qui appartient à un voisinage T_0 de t_0 , on a

$$|f(t) - P_{t_0}(t)| \leq C|t - t_0|^\alpha$$

Avec cette définition on vérifie immédiatement que si on considère un signal r -dérivable au sens classique, alors en utilisant l'approximation avec un polynôme de Taylor du signal, on a pour tout $0 < \alpha \leq r$, que le signal est partout α -Lipschitz.

On peut facilement relier cette définition avec les ondelettes, en considérant les moments d'ondelette. On considère ainsi $\{\psi_{j,k} = \psi(2^{\frac{j}{2}}\psi(2^j \cdot -k))\}_{j,k}$ une base orthonormale d'ondelettes de $L^2([0, 1])$, et on a le coefficient d'ondelette à l'échelle j et à l'instant k donnée par

$$Wf(j, k) = \langle f, \psi_{j,k} \rangle = \int f(t) \psi_{j,k}(t) dt.$$

Remarque 2.2.6. Remarquons ici que l'on peut exprimer cette projection à partir de l'ondelette prise à l'instant 0

$$\widetilde{\psi}_j(t) = \psi_{j,0}(-t) = \psi(-2^j t),$$

et ainsi en faisant un changement de variable dans 2.2.4, on obtient

$$Wf(j, k) = f \star \widetilde{\psi}_j(k). \quad (2.54)$$

On peut ainsi interpréter le coefficient d'ondelette pris en (j, k) comme la corrélation entre le signal pris en k avec une ondelette à l'échelle j . C'est à dire qu'un coefficient avec une grande valeur indique une grande similitude entre l'ondelette et le signal (l'ondelette approxime bien le signal) alors qu'un petit coefficient indique que l'ondelette et le signal ont peu en commun⁴.

Rappelons aussi qu'une condition nécessaire pour que ψ soit une ondelette génératrice est

$$\int \psi(t) dt = 0.$$

Définition 2.2.7. Soit m un entier strictement positif. On dit que ψ a m -moments nuls si

$$\int t^k \psi(t) dt = 0 \quad , \forall k < m.$$

De cette définition on déduit que si P est un polynôme de degré strictement inférieur à m et si l'ondelette a m -moments nuls, alors

$$\int \psi(t) p(t) dt = 0. \quad (2.55)$$

Ainsi si $f(t) = P(t) + \epsilon$ où ϵ représente un bruit, on a

$$\langle f, \psi_{j,k} \rangle = o(\epsilon)$$

ainsi les coefficients de l'ondelette d'échelle suffisent à reconstruire f . C'est ainsi que si on considère une ondelette avec un certain nombre de moments nuls, alors dans les parties régulières du signal, les coefficients d'ondelettes seront petits, alors que dans les zones avec des irrégularités ou des discontinuités, les coefficients resteront grands. On peut aussi remarquer que si les irrégularités sont séparées, alors en affinant l'échelle d'analyse, le support des ondelettes diminue et alors de moins en moins de coefficients auront une valeur importante, ainsi les seuls coefficients qui resteront grand en changeant d'échelle sont ceux qui contiennent une zone irrégulière.

On va maintenant démontrer le théorème suivant de Jaffard qui permet de préciser cela,

⁴L'ondelette et le signal ont peu en commun au sens où ils sont de façon équivalente, presque orthogonaux, et donc ce coefficient a un poids faible dans la formule de reconstruction.

Theorème 2.2.7.1. *Si f est α -Lipschitz en t_0 avec $0 < \alpha \leq m$ où m est un entier. Alors, il existe une constante $C > 0$ telle que*

$$|Wf(j, k)| \leq C 2^{-j(\alpha + \frac{1}{2})} (1 + |2^{-j}t - t_0| 2^{j\alpha}).$$

Preuve. Remarquons tout d'abord que, soit P un polynôme de degré strictement inférieur à m , et ψ une ondelette à m -moments nuls, alors, par changement de variable et d'après 2.55 on a :

$$WP(j, k) = \int 2^{j/2} \psi(2^j t - 2^{-j} k) P(t) dt = \int 2^{j/2} \psi(t') P(2^j t' - 2^{-j} k) 2^{-j} dt' = 0. \quad (2.56)$$

Par hypothèse, x est α -Lipschitz en t_0 , donc il existe P_{t_0} un polynôme de degré inférieur à m tel que $|f(t) - P_{t_0}(t)| \leq C|t - t_0|^\alpha$. En utilisant la linéarité de l'intégrale et en appliquant une inégalité triangulaire on obtient

$$\begin{aligned} |Wf(j, k)| &= \left| \int (f(t) - P_{t_0}(t) + P_{t_0}(t)) \psi_{j,k}(t) dt \right| \\ &\leq \left| \int (f(t) - P_{t_0}(t)) \psi_{j,k}(t) dt \right| + \left| \int P_{t_0}(t) \psi_{j,k}(t) dt \right| \\ &\leq \int |f(t) - P_{t_0}(t)| |\psi_{j,k}(t)| dt \end{aligned}$$

la dernière inégalité étant obtenue en utilisant 2.56 sur le terme de droite et en faisant entrer la valeur absolue dans la première intégrale. On utilise maintenant le fait que f est α -Lipschitz et on fait un changement de variable, on obtient ainsi

$$\begin{aligned} |Wf(j, k)| &\leq \int |f(t) - P_{t_0}(t)| |\psi_{j,k}(t)| dt \leq C \int |t - t_0|^\alpha 2^{j/2} |\psi(2^j t - k)| dt \\ &\leq C \int |2^{-j} t' + 2^{-j} k - t_0|^\alpha 2^{-j/2} |\psi(t')| dt'. \end{aligned}$$

Pour obtenir l'inégalité suivante on utilise

$$|a + b|^\alpha \leq |2 * \max(|a|, |b|)|^\alpha \leq 2^\alpha (|a|^\alpha + |b|^\alpha)$$

et on a ainsi

$$\begin{aligned} |Wf(j, k)| &\leq 2^\alpha C \int (|2^{-j} t'|^\alpha + |2^{-j} k - t_0|^\alpha) 2^{-j/2} |\psi(t')| dt' \\ &\leq 2^\alpha C 2^{-j(\alpha + 1/2)} \left(\int |t'|^\alpha |\psi(t')| dt' + |2^{-j} k - t_0|^\alpha 2^{\alpha j} \int |\psi(t')| dt' \right) \end{aligned}$$

Ce qui donne le résultat dès que les intégrales considérées sont définies, ce qui est le cas par exemple si l'ondelette est à support compact ou bien à décroissance suffisamment rapide. \square

On peut combiner le théorème de Jaffard avec une analyse multi-échelle d'ondelettes avec la proposition suivante :

Proposition 2.2.8. *Soit $f :]0, 1[\rightarrow \mathbb{R}$ une fonction α -Lipschitzienne avec $\alpha > 1$, alors il existe une constante $C > 0$ et une base d'ondelette orthonormales associée à une multirésolution $\{(\psi_{j,k})_{(i,j):j \geq J, 2^j > k \geq 0}\}, \{\varphi_{J,k}\}_k$ avec $m > \alpha$ moments nuls telle que*

$$\varepsilon_l(f, 2^J) = \|f - \sum_{k=0}^{2^J-1} \langle f, \varphi_{J,k} \rangle \tilde{\varphi}_{J,k}\|_2^2 \leq C 2^{-2J\alpha} = C N^{-2\alpha} \quad (2.57)$$

avec $N = 2^J$.

Preuve. L'existence d'une telle base d'ondelette n'est pas démontrée ici, des constructions peuvent être trouvées dans (ajouter ref) pour obtenir des bases de $L^2(\mathbb{R})$, on peut ainsi considérer une telle multirésolution donnée par une ondelette de Daubechies ou bien une coiflet à $m > \alpha$ moments nuls. Il est ensuite possible, avec quelques difficultés d'obtenir depuis ces ondelettes, une base orthonormale de $L^2(]0, 1[)$ (ajouter ref). Soit f une fonction α -Lipschitzienne sur $]0, 1[$, on a ainsi d'après la partie sur les frames et l'existence de la base d'ondelette précédente admise, une formule de reconstruction

$$f = \sum_{k=0}^{2^J-1} \langle f, \varphi_{J,k} \rangle \tilde{\varphi}_{J,k} + \sum_{j=J+1}^{+\infty} \sum_{k=0}^{2^j-1} \langle f, \psi_{j,k} \rangle \tilde{\psi}_{j,k}.$$

On a ainsi, en réécrivant l'équation et en prenant la norme

$$\varepsilon(f, 2^J) = \|f - \sum_{k=0}^{2^J-1} \langle f, \varphi_{J,k} \rangle \tilde{\varphi}_{J,k}\|_2^2 = \left\| \sum_{j=J+1}^{+\infty} \sum_{k=0}^{2^j-1} \langle f, \psi_{j,k} \rangle \tilde{\psi}_{j,k} \right\|_2^2.$$

On peut alors majorer le terme de droite en utilisant le fait que la famille d'analyse est génératrice, on obtient

$$\varepsilon(f, 2^J) \leq \sum_{j=J+1}^{+\infty} \left\| \sum_{k=0}^{2^j-1} \langle f, \psi_{j,k} \rangle \tilde{\psi}_{j,k} \right\|_2^2$$

et en utilisant le fait que les ondelettes sont normalisées on a

$$\varepsilon(f, 2^J) \leq \sum_{j=J+1}^{+\infty} \sum_{k=0}^{2^j-1} |\langle f, \psi_{j,k} \rangle|^2$$

on utilise maintenant le théorème 2.2.7.1 et on obtient⁵

$$\begin{aligned}
 \varepsilon(f, 2^J) &\leq \sum_{j=J+1}^{+\infty} \sum_{k=0}^{2^j-1} C^2 2^{-j(2\alpha+1)} \\
 &\leq \sum_{j=J+1}^{+\infty} C^2 2^{-j(2\alpha+1)} 2^j = \sum_{j=J+1}^{+\infty} C^2 2^{-j2\alpha} \\
 &\leq \frac{C^2}{1 - 4^{-\alpha}} 2^{-2J\alpha}
 \end{aligned}$$

ce qui prouve la proposition. □

⁵Le théorème de Jaffard est pour une fonction ponctuellement Lipschitzienne, on considère ici une fonction α -Lipschitzienne en tout point, donc le terme en $(1 + \frac{|2^{-j}k - t_0|}{2^j})$ n'apparaît pas.

Chapter 3

Reconstruction parcimonieuse et $\|\cdot\|_1$

3.1 Résolution de (P0)

Dans ce qui précède, nous nous sommes intéressés aux propriétés qui font qu'une famille de vecteurs permet de reconstruire une famille de signaux. Nous avons vu différentes bases (Fourier et ondelettes) et nous avons vu que ces bases permettent de reconstruire des signaux présentant un certain type de régularité avec des coefficients qui suivent une décroissance assez rapide.

On a par exemple vu que l'on pouvait reconstruire les fonctions Lipschitziennes avec une bonne précision en utilisant une base d'ondelettes orthonormale avec un certain nombre de moments nuls. De plus, on a remarqué que si la fonction se comporte comme un polynôme d'un degré inférieur au nombre de moments nuls au voisinage d'un point, alors les coefficients d'ondelettes dans ce voisinage seront nuls. De même, les seuls coefficients d'ondelette qui seront grands seront ceux au voisinage d'un point où aucune approximation par un polynôme de petit degré n'est efficace¹. Ainsi, la représentation avec ces ondelettes d'une fonction ne possédant que quelques points où elle est irrégulière sera approximée avec peu de coefficients. Afin d'insister, l'intérêt de cela est que de façon naïve, afin de déterminer une fonction, il faut connaître sa valeur en chaque point, ainsi, si l'on souhaite faire un traitement par ordinateur de cette fonction, il faut stocker chacun des points de la fonction. Avec ce que l'on a fait, on sait qu'en fait on peut reconstruire la fonction avec un plus petit nombre de coefficients que la fonction n'a de points. En ce sens, la représentation en ondelettes d'une fonction Hölderienne est parcimonieuse (peu de coefficients non nuls), alors que la représentation par la valuation d'une fonction Hölderienne n'est pas parcimonieuse.

¹Une analyse du théorème de Jaffard 2.2.7.1 montre que les coefficients affectés par une discontinuité forment un cône dans les coefficients d'ondelette autour du point de discontinuité. Ce cône se visualise dans la représentation temps-fréquence des coefficients d'ondelette, il part du point de discontinuité et s'élargit en diminuant le coefficient d'échelle j . La largeur de ce cône dépend de la régularité α de la fonction.

Nous allons maintenant nous intéresser à l'autre direction de ce problème, c'est à dire que nous allons supposer que l'on dispose d'une famille de vecteurs et que la fonction que l'on cherche à reconstruire est une somme parcimonieuse de vecteurs de cette famille. Cependant on connaît seulement la valuation de cette fonction et pas les vecteurs sous-jacents qui permettent de représenter la fonction de façon parcimonieuse. Aussi, on n'a pas supposé que cette famille est libre donc il n'y a pas une unique façon d'obtenir cette solution, en fait il y a une infinité de solutions dès que la famille n'est pas libre. On va voir cependant que l'hypothèse de parcimonie est cruciale et qu'elle nous permettra de récupérer exactement les coefficients qui permettent l'écriture parcimonieuse de cette fonction.

3.1.1 Définition de (P0)

Formalisons maintenant ce que nous avons dit ci-dessus. On considère \mathcal{F} un espace vectoriel et utilisons un dictionnaire $\Phi = \Phi_1 \cup \dots \cup \Phi_D$ de bases, où chaque Φ_d est une base de \mathcal{F} . Ainsi Φ est une concaténation de bases² et on s'intéresse aux façon d'écrire un signal $f \in \mathcal{F}$ dans Φ , c'est à dire aux façon d'écrire

$$f = \sum_{\gamma} c_{\gamma} \phi_{\gamma} \quad (3.1)$$

où l'indice $\gamma = (d, i)$ indique le dictionnaire Φ_d correspondant ainsi que le vecteur $\phi_{d,i} \in \Phi_d$. On peut aussi écrire 3.1 sous forme matricielle en posant F_{Φ} la matrice ayant pour lignes les vecteurs ϕ_{γ} et en posant $x = (c_{\gamma})_{\gamma}$ la notation sous forme de vecteur de x , on utilisera aussi la notation $x = (x_d)_{d=1, \dots, D}$. On s'intéresse ainsi aux solutions de

$$f = F_{\Phi} x. \quad (3.2)$$

Comme discuté précédemment, le choix des coefficients c_{γ} n'est pas unique dès que $D > 1$, cependant notre objectif n'est pas simplement de reconstruire f (car n'importe quelle base Φ_i permet déjà cela), mais de trouver l'écriture de f avec le minimum de coefficients non nuls. Ainsi, le problème que l'on cherche à résoudre est

$$\min \|x\|_0 \quad \text{tel que } f = F_{\Phi} x, \quad (\text{P0})$$

où $\|x\|_0 = \#\{\gamma : c_{\gamma} \neq 0\}$ est le nombre de coefficients non nuls de x . Cependant, la résolution en toute généralité de ce problème n'est pas faisable, en effet résoudre ce problème nécessite de résoudre (P0) pour chaque combinaison de vecteurs du dictionnaire si x est dans l'image. Ainsi, le nombre de combinaisons possibles parmi tous les vecteurs croît bien trop vite pour être calculable en pratique, nous verrons donc comment résoudre ce problème en utilisant une autre méthode.

Il est important de noter qu'à ce stade il n'y a aucune raison de supposer que chercher une unique solution à (P0) a un sens. En effet, quand on a choisi le dictionnaire Φ rien

²Ainsi Φ est un frame équilibré d'après la première partie (TODO : ajouter ref)

ne nous interdisait de prendre à chaque fois la même base et on aurait ainsi D solutions identiques, ayant chacune la même parcimonie. On a ainsi D solutions, et si on prend une paire de solutions x_1, x_2 , alors $F_\Phi(x_1 - x_2) = 0$, d'où on obtient qu'à n'importe laquelle des D solutions, on peut ajouter, par exemple $x_1 - x_2$, et on obtient une nouvelle solution. Cependant, cette solution ne sera jamais moins parcimonieuse que l'une des D solutions initiales. Il est donc clair qu'il est nécessaire d'imposer des conditions sur les bases qui constituent le dictionnaire si l'on souhaite obtenir une solution unique. Afin d'étudier cela commençons par un cadre simple dans lequel résoudre P_0 a un sens.

Exemple 3.1.2. On étudie les signaux dans \mathbb{R}^N et on choisit un dictionnaire constitué de la concaténation de la base de Fourier $W = \{e_k(t) = \frac{1}{\sqrt{N}} e^{i2\pi kt/N}\}_{0 \leq k \leq N-1}$ et de la base canonique de Diracs³ $T = \{\delta_k\}_{0 \leq k \leq N-1}$. Ainsi, avec ce choix F_W est la matrice de Fourier discrète et F_T est la matrice identité de taille N . Avec le théorème suivant, on va obtenir un principe d'incertitude, qui nous garantira qu'un signal ne peut pas être parcimonieux à la fois dans la base de Dirac, et dans la base de Fourier.

Théorème 3.1.2.1. Soit un signal $f \in \mathbb{R}^N$ non nul, alors

$$\|F_W f\|_0 \|F_T f\|_0 \geq N \quad (3.3)$$

et ainsi

$$\|F_W f\|_0 + \|F_T f\|_0 \geq 2\sqrt{N}. \quad (3.4)$$

Preuve. TODO :Ajouter Ref Tao uncertainty principle for cyclic.... Soit $0 \leq \omega \leq N-1$ un entier, alors

$$|F_W(\omega)| = |\hat{f}(\omega)| = \frac{1}{\sqrt{N}} \left| \sum_t f(t) e_\omega(t) \right| \quad (3.5)$$

$$\leq \frac{1}{\sqrt{N}} \sum_t |f(t)|, \quad (3.6)$$

d'où $\sup_\omega |F_W(\omega)| \leq \frac{1}{\sqrt{N}} \sum_t |f(t)|$. On pose maintenant $\text{sign}(f) = (\frac{f(t)}{|f(t)|})_t$ pour tous les t tels que $f(t) \neq 0$ et 0 si $f(t) = 0$ et on a ainsi $\langle \text{sign}(f), \text{sign}(f) \rangle = \|F_T f\|_0$, on va ainsi pouvoir montrer le théorème en utilisant successivement l'inégalité de Cauchy-Schwarz puis l'égalité de Parseval,

$$\sup_\omega |F_W(\omega)| \leq \frac{1}{\sqrt{N}} \sum_t |F_T f(t)| = \frac{1}{\sqrt{N}} \langle \text{sign}(f), |f| \rangle \quad (3.7)$$

$$\leq \frac{1}{\sqrt{N}} \|F_T\|_0^{\frac{1}{2}} \langle |f|, |f| \rangle^{\frac{1}{2}} = \frac{1}{\sqrt{N}} \|F_T\|_0^{\frac{1}{2}} \langle |F_W f|, |F_W f| \rangle^{\frac{1}{2}} \quad (3.8)$$

$$\leq \frac{1}{\sqrt{N}} \|F_T\|_0^{\frac{1}{2}} \langle |f|, |f| \rangle^{\frac{1}{2}} = \frac{1}{\sqrt{N}} \|F_T\|_0^{\frac{1}{2}} \left(\sum_\omega |\hat{f}(\omega)|^2 \right)^{\frac{1}{2}} \leq \frac{1}{\sqrt{N}} \|F_T\|_0^{\frac{1}{2}} \|F_W f\|_0^{\frac{1}{2}} \sup_\omega |F_W(\omega)|. \quad (3.9)$$

³Chaque vecteur de cette base vérifie $\delta_{k,i} = 1$ si $i = k$ et 0 sinon.

On a ainsi montré la première partie du théorème, la deuxième partie provient directement de l'inégalité entre la moyenne arithmétique et la moyenne géométrique. En effet, on a

$$\sqrt{\|F_T f\|_0 \|F_W f\|_0} \leq \frac{\|F_T f\|_0 + \|F_W f\|_0}{2} \quad (3.10)$$

et on a déjà montré que le terme de gauche est supérieur ou égal à \sqrt{N} . \square

Observons que sans restrictions sur N , l'inégalité obtenue ne peut pas être améliorée, comme observé dans (Donoho-Stark 89 /Donoho-Huo TODO: ajouter ref), si N est un carré, alors, la fonction avec des 1 seulement aux coefficients multiples de \sqrt{N} et 0 ailleurs est sa propre transformée de Fourier et ainsi elle a $2\sqrt{N}$ coefficients dans le dictionnaire (T, W) et ainsi l'inégalité est atteinte. Une conséquence de cela est qu'une condition sur la parcimonie de la forme $\|F_T f\|_0 + \|F_W f\|_0 < K$ avec $K > \sqrt{N}$ ne pourra pas garantir l'unicité de la solution de (P0). Montrons que si $K = \sqrt{N}$ alors on a l'unicité de la solution de (P0).

Théorème 3.1.2.2. Soit N un entier positif, (T, W) le dictionnaire Fourier-Dirac et un signal $f \in \mathbb{R}^N$, alors n'importe quel x vérifiant $f = F_\Phi x = F_W x_W + F_T x_T$ et

$$\|x_W\|_0 + \|x_T\|_0 < \sqrt{N} \quad (3.11)$$

est l'unique solution de (P0).

Supposons que pour f donné non nul et supposons que l'on ait deux solutions de (P0), x_1 et x_2 , ainsi $f = F_\Phi x_1$, $f = F_\Phi x_2$ et on a aussi $\|x_1\|_0 < \sqrt{N}$, $\|x_2\|_0 < \sqrt{N}$. On a par linéarité de l'opérateur F_Φ ,

$$F_\Phi(x_1 - x_2) = 0. \quad (3.12)$$

Etudions ainsi les éléments du noyau de F_Φ , posons $\mathcal{N} = \{\delta : F_\Phi \delta = 0\}$, et pour tout $\delta \in \mathcal{N}$, écrivons $\delta = (\delta_T, \delta_W)$, on a

$$F_T \delta_T + F_W \delta_W = 0 \quad (3.13)$$

ainsi, en utilisant que les colonnes de F_W forment une base, donc F_W est une matrice orthogonale, on a

$$\delta_W = -F_W^t F_T \delta_T. \quad (3.14)$$

On a donc montré que les éléments de \mathcal{N} sont de la forme $\delta = (\delta_T, -F_W^t F_T \delta_T)$ et d'après le théorème 3.1.2.1, on a que δ a au moins $2\sqrt{N}$ coefficients non nuls si δ est non nul. En revenant à la situation initiale $\delta = x_1 - x_2$, on a une contradiction car à la fois x_1 et x_2 ont chacun moins de \sqrt{N} coefficients, donc $\delta = 0$. Ainsi, si une solution existe avec moins de \sqrt{N} coefficients, alors c'est la solution de (P0) et elle est unique.

Cependant, en choisissant $N = p$, où p est un nombre premier⁴, Tao 2005 (TODO: ajouter ref) a montré que l'on obtient l'inégalité

$$\|F_T f\|_0 + \|F_W f\|_0 \geq p + 1 \quad (3.15)$$

⁴L'hypothèse p premier est essentielle, la preuve reposant sur la non-existence de sous-groupes propres du groupe cyclique $\mathbb{Z}/p\mathbb{Z}$

et que l'inégalité est atteinte⁵. Grâce à ce principe d'incertitude plus fort que 3.1.2.1, on obtient avec le même type de preuve⁶ le résultat suivant

Theorème 3.1.2.3. Soit N un nombre premier et un signal $f \in \mathbb{R}^N$, alors n'importe quel x vérifiant $f = F_\Phi x = F_W x_W + F_T x_T$ et

$$\|x_T\|_0 \leq \frac{N}{2} \quad (3.16)$$

est l'unique solution de (P0).

On a ainsi vu qu'avec un dictionnaire constitué de Fourier et de Dirac la solution de (P0) est unique, on a également vu brièvement, qu'en renforçant le principe d'incertitude sur les deux familles, alors on peut certifier qu'on a bien obtenu *la solution de (P0) pour des signaux avec un support plus grand.*

3.1.3 Solution optimale combinatoire

3.1.4 Résolution dans un dictionnaire pics/Fourier

3.1.5 Principe d'incertitude

3.2 Résolution de (P1)

On a ainsi vu dans la section précédente que le problème (P0) de minimisation de la solution par rapport à la parcimonie admet une solution unique dès qu'une solution existe et que cette solution vérifie une condition de la forme 3.44 ou 3.16. Cependant, on a aussi vu au début de la section précédente que le problème (P0) est un problème de nature combinatoire et le nombre de combinaisons possibles augmentant très vite par rapport à N , sa résolution n'est pas faisable et ainsi il est nécessaire d'avoir une autre approche à ce problème.

La découverte qui a permis de rendre la résolution faisable, et par là permis par exemples les avancées du compressed sensing qui ont eu de nombreuses applications et dont la théorie sera étudiée dans le prochain chapitre, est que l'on peut résoudre un autre problème pour lequel des méthodes de résolution efficaces existaient déjà. En effet nous allons voir que résoudre le problème P1,

$$\min_x \|x\|_1 \text{ tel que } f = Fx \quad (\text{P1})$$

permet sous certaines conditions de résoudre P0. L'intérêt de P1 est que c'est un problème de programmation linéaire et de nombreuses méthodes permettent de le

⁵L'inégalité est atteinte en ce sens que si $A \subset T$ et $B \subset W$ tels que $|A| + |B| \geq p + 1$ alors il existe une fonction f telle que $\text{Supp} F_T f = A$ et $\text{Supp} F_W f = B$

⁶Voir Candes-Romberg-Tao pour les détails, un lemme sur l'injectivité d'un opérateur similaire F_W est tout de même nécessaire pour conclure la preuve.

résoudre. (TODO ajouter refs et détails). Précisons donc ce que nous avons affirmé, dans le même cadre que précédemment, c'est à dire dans le cas d'un dictionnaire $\Phi = (T, W)$ temps fréquence composé de la base de Fourier et de Dirac dans \mathbb{R}^N .

Théorème 3.2.0.1. *Soit N un entier positif, $\Phi = (T, W)$ est la concaténation des bases de Dirac et de Fourier et un signal $f \in \mathbb{R}^N$, alors n'importe quel $x = (x_T, x_W)$ vérifiant $f = F_T x_T + F_W x_W$ et*

$$\|x_T\|_0 < \frac{\sqrt{N}}{2} \quad \text{et} \quad \|x_W\|_0 < \frac{\sqrt{N}}{2} \quad (3.17)$$

est l'unique solution de P1, et c'est la solution de P0.

Remarque 3.2.1. Le théorème précédent a été obtenu en cherchant une preuve alternative à la preuve qui est faite par Donoho et Huo (TODO: ajouter ref), leur preuve, comme une grande partie de la section précédente, utilise le même schéma que celle qui est faite ici. Leur théorème est le suivant :

Théorème 3.2.1.1. *Soit N un entier positif, $\Phi = (T, W)$ est la concaténation des bases de Dirac et de Fourier et un signal $f \in \mathbb{R}^N$, alors n'importe quel $x = (x_T, x_W)$ vérifiant $f = F_T x_T + F_W x_W$ et*

$$\|x_T\|_0 + \|x_W\|_0 < \frac{\sqrt{N}}{2} \quad (3.18)$$

est l'unique solution de P1, et c'est la solution de P0.

La preuve qui est présentée utilise un lemme qui est une version affaiblie d'un résultat présenté dans l'article. Dans l'article, l'inégalité plus forte qui est utilisée est obtenue à l'aide d'un principe variationnel (TODO: préciser), mais comme les auteurs le remarquent, leur résultat ne semblait pas exact au sens où même lorsque l'inégalité est atteinte il n'y avait aucun contre-exemple apparent. En effet, dans le cas de la base de Fourier-Dirac, le peigne de Dirac, fournit dans certains cas un exemple de signal qui est supporté sur \sqrt{N} coefficients soit dans la base de Fourier, soit dans la base de Dirac, ainsi le problème P0 a plusieurs solutions et donc une condition nécessaire pour résoudre simultanément P1 et P0 est $\|x\|_0 < \sqrt{N}$. Or, les hypothèses du théorème ne sont plus vérifiées dès que $\|x\|_0 = \sqrt{N}$ (car au moins, soit x_T , soit x_W est supporté sur au moins $\frac{\sqrt{N}}{2}$ coefficients).

Preuve. La preuve de ce théorème se fait en plusieurs parties. Tout d'abord, remarquons que si x vérifie 3.47, alors x vérifie 3.44 et donc d'après le théorème 3.1.2.2 x est donc l'unique condition de P0. Il nous faut donc vérifier que cette solution est bien la solution de (P1). On montre ensuite un lemme qui permet de donner une condition suffisante pour qu'une paire de bases vérifie que la solution obtenue est bien celle de P1. On vérifiera ensuite que dans la paire de bases Fourier-Dirac, les conditions du lemme sont vérifiées et cela permettra de conclure la preuve du théorème. Avant d'énoncer le lemme, définissons une quantité μ qui mesure dans une paire de bases $\Phi = (T, W)$ à quel point un élément dans le noyau de F_Φ peut être supporté à la fois sur T et sur W .

Définition 3.2.2. Soit $\Phi = (T, W)$ une paire de bases, on note $\mathcal{N} = \{\delta = (\delta_T, \delta_W) : F_\Phi \delta = 0\}$, soit Γ_T (resp. Γ_W) un ensemble d'indices de T (resp. W), alors on pose

$$\mu(\Gamma_T, \Gamma_W) = \sup_{\delta \in \mathcal{N}} \frac{\sum_{t \in \Gamma_T} |\delta_{T,t}| + \sum_{\omega \in \Gamma_W} |\delta_{W,\omega}|}{\|\delta_T\|_1 + \|\delta_W\|_1} \quad (3.19)$$

Lemme 3.2.3. Soit un signal $f \in \mathbb{R}^N$ et $\Phi = (T, W)$ une paire de bases de \mathbb{R}^N , alors n'importe quel $x = (x_T, x_W)$, où Γ_T est le support de x_T et Γ_W est le support de x_W , vérifiant $f = F_T x_T + F_W x_W$ et

$$\mu(\Gamma_T, \Gamma_W) < \frac{1}{2} \quad (3.20)$$

est l'unique solution de P1.

Pour prouver le théorème on vérifiera donc dans la base de Fourier-Dirac que pour n'importe quelle paires d'indices vérifiant les conditions du théorème alors l'inégalité 3.20 sera vérifiée, et ainsi la solution de P0 sera bien la même que celle de P1 ce qui permettra de conclure la preuve du théorème. Enonçons donc cela sous la forme d'un autre lemme

Lemme 3.2.4. ⁷ Soit $\Phi = (T, W)$ la paire de bases Fourier-Dirac et soient Γ_T et Γ_W des sous ensembles d'indices de T et respectivement de W vérifiant

$$|\Gamma_T| < \frac{\sqrt{N}}{2} \quad \text{et} \quad |\Gamma_W| < \frac{\sqrt{N}}{2}, \quad (3.21)$$

alors on a,

$$\mu(\Gamma_T, \Gamma_W) < \frac{1}{2}. \quad (3.22)$$

Ainsi, une fois les lemmes démontrés, le théorème le sera aussi. \square

Commençons par la preuve du lemme 3.2.3.

Preuve. Supposons que x vérifie les conditions du lemme, c'est à dire, x est effectivement une solution de l'équation $f = F_\Phi x$ et la condition 3.20 est vérifiée sur Φ , alors on doit donc montrer que x est l'unique solution de (P1), on doit donc montrer que pour tout x_1 différent de x qui vérifie $f = F_\Phi x_1$ alors $\|x_1\|_1 > \|x\|_1$. Donc de façon équivalente, pour tout $\delta \in \mathcal{N} = \{\delta : F_\Phi \delta = 0\}$ non nul, on doit vérifier que

$$\|x + \delta\|_1 - \|x\|_1 > 0. \quad (3.23)$$

Notons $\Gamma = \{\gamma : c_\gamma \neq 0\} = \Gamma_T \cup \Gamma_W \subset [0, 2N - 1]$ l'ensemble des indices non nuls de $x = (c_\gamma)_\gamma$, on peut donc décomposer la somme

$$\|x + \delta\|_1 - \|x\|_1 = \sum_{\gamma \in \Gamma^c} |\delta_\gamma| + \sum_{\gamma \in \Gamma} |c_\gamma + \delta_\gamma| - |c_\gamma|. \quad (3.24)$$

⁷C'est ce lemme dont il est fait mention dans la remarque précédant la preuve et qui permet la généralisation du théorème de Donoho et Huo.

Par l'inégalité triangulaire on a $|c_\gamma| \leq |c_\gamma + \delta_\gamma| + |\delta_\gamma|$ quel que soit γ . On a ainsi

$$|c_\gamma + \delta_\gamma| - |c_\gamma| \geq -|\delta_\gamma| \quad (3.25)$$

et en insérant cette inégalité dans la somme on obtient

$$\|x + \delta\|_1 - \|x\|_1 \geq \sum_{\gamma \in \Gamma^c} |\delta_\gamma| - \sum_{\gamma \in \Gamma} |\delta_\gamma|, \quad (3.26)$$

ainsi une condition suffisante pour obtenir l'unicité est que pour $\delta \in \mathcal{N}$ non nul on ait

$$\sum_{\gamma \in \Gamma} |\delta_\gamma| < \sum_{\gamma \in \Gamma^c} |\delta_\gamma|. \quad (3.27)$$

Avec des mots cela revient à dire que si δ est dans \mathcal{N} et non nul, alors δ a plus de poids hors du support de x que sur le support de x . En ajoutant le terme de gauche de l'inégalité précédente des deux côtés on obtient

$$\sum_{\gamma \in \Gamma} |\delta_\gamma| < \frac{1}{2} \left(\sum_{t \in T} |\delta_{T,t}| + \sum_{\omega \in W} |\delta_{W,\omega}| \right) = \frac{\|\delta_T\|_1 + \|\delta_W\|_1}{2}. \quad (3.28)$$

Donc l'inégalité précédente est aussi une condition suffisante pour que 3.23 soit vérifiée et on peut réécrire cette inégalité sous la forme

$$\frac{\sum_{t \in \Gamma_T} |\delta_{T,t}| + \sum_{\omega \in \Gamma_W} |\delta_{W,\omega}|}{\|\delta_T\|_1 + \|\delta_W\|_1} < \frac{1}{2}. \quad (3.29)$$

On veut que l'inégalité soit vérifiée pour n'importe quel delta, donc en vérifiant la condition sur le suprémum des δ dans le noyau de F_Φ le lemme sera vrai. C'est exactement la condition 3.32 du lemme

$$\mu(\Gamma_T, \Gamma_W) := \sup_{\delta \in \mathcal{N}} \frac{\sum_{t \in \Gamma_T} |\delta_{T,t}| + \sum_{\omega \in \Gamma_W} |\delta_{W,\omega}|}{\|\delta_T\|_1 + \|\delta_W\|_1} < \frac{1}{2}. \quad (3.30)$$

Le lemme 3.2.3 est donc bien démontré.

On peut au passage remarquer qu'on peut utiliser la structure du noyau de F_Φ de la façon suivante afin d'obtenir une écriture équivalente de 3.32 mais qui utilise le fait qu'un élément du noyau de F_Φ est entièrement déterminé par ses coefficients dans l'une des deux bases. On avait vu avec 3.14 que les éléments δ de \mathcal{N} sont de la forme $(\delta_T, -F_W^t F_T \delta_T) =: (\delta_T, -\hat{\delta}_T)$, donc 3.29 devient

$$\frac{\sum_{t \in \Gamma_T} |\delta_{T,t}| + \sum_{\omega \in \Gamma_W} |\hat{\delta}_{T\omega}|}{\|\delta_T\|_1 + \|\hat{\delta}_T\|_1} < \frac{1}{2}. \quad (3.31)$$

□

On peut maintenant passer à la preuve du lemme 3.2.4

Preuve.

$$\mu(\Gamma_T, \Gamma_W) \leq \frac{\sum_{t \in \Gamma_T} |\delta_{T,t}| + \sum_{\omega \in \Gamma_W} |\hat{\delta}_{T\omega}|}{\|\delta_T\|_1 + \|\delta_W\|_1}. \quad (3.32)$$

Maintenant majorons le numérateur avec

$$\sum_{\omega \in \Gamma_W} |\hat{\delta}_{T\omega}| = \|R_{\Gamma_W} F_W^t F_T \delta_T\|_1 \leq \|R_{\Gamma_W} F_W^t F_T\|_1 \|\delta_T\|_1 \quad (3.33)$$

où $\|A\|_1 = \sup_i \|c_i\|_1$ avec c_i les colonnes de la matrice, et R_{Γ_W} est la matrice de projection dans l'espace engendré par les vecteurs indexés par Γ_W . Donc $R_{\Gamma_W} F_W^t F_T$ est une matrice à $|\Gamma_W|$ lignes et N colonnes, la norme ℓ_1 de chaque colonne est égale à $\frac{|\Gamma_W|}{\sqrt{N}}$, ainsi, on a⁸ :

$$\|R_{\Gamma_W} F_W^t F_T\|_1 = \frac{|\Gamma_W|}{\sqrt{N}}. \quad (3.34)$$

Maintenant appliquons la même chose à $\delta_T = -R_{\Gamma_T} F_T^t F_W \delta_W$:

$$\sum_{t \in \Gamma_T} |\delta_{T,t}| = \|R_{\Gamma_T} F_T^t F_W \delta_W\|_1 \leq \|R_{\Gamma_T} F_T^t F_W\|_1 \|\delta_W\|_1 \quad (3.35)$$

ainsi que

$$\|R_{\Gamma_T} F_T^t F_W\|_1 = \frac{|\Gamma_T|}{\sqrt{N}}. \quad (3.36)$$

On peut maintenant rassembler les résultats:

$$\sum_{t \in \Gamma_T} |\delta_{T,t}| + \sum_{\omega \in \Gamma_W} |\hat{\delta}_{T\omega}| \leq \|\delta_W\|_1 \frac{|\Gamma_T|}{\sqrt{N}} + \|\delta_T\|_1 \frac{|\Gamma_W|}{\sqrt{N}}. \quad (3.37)$$

On utilise maintenant les hypothèses $|\Gamma_T| < \sqrt{N}/2$ et $|\Gamma_W| < \sqrt{N}/2$, on obtient ainsi :

$$\sum_{t \in \Gamma_T} |\delta_{T,t}| + \sum_{\omega \in \Gamma_W} |\hat{\delta}_{T\omega}| < \frac{\|\delta_T\|_1 + \|\delta_W\|_1}{2}. \quad (3.38)$$

Il nous reste maintenant à appliquer la majoration que l'on vient de trouver à 3.32 et on obtient

$$\mu(\Gamma_T, \Gamma_W) < \frac{1}{2} \frac{\|\delta_T\|_1 + \|\delta_W\|_1}{\|\delta_T\|_1 + \|\delta_W\|_1} = \frac{1}{2} \quad (3.39)$$

Ce qui conclut la preuve du lemme 3.2.4 et donc du théorème 3.2.0.1. \square

⁸C'est ici que le choix de la paire de bases a une importance, la matrice $F_W^t F_T$ contient tous les produits scalaires des vecteurs de W et de T , dans le dictionnaire de Fourier-Dirac, chacun des coefficients vaut $1/\sqrt{N}$

3.2.5 Généralisation à des paires de bases arbitraires

A partir de cette preuve dans le dictionnaire Fourier-Dirac, on peut facilement obtenir une généralisation à une paire de bases orthogonales (Φ, Ψ) arbitraire. En effet, dans les preuves le choix des bases a un effet seulement sur les matrices F_Φ et F_Ψ , et plus particulièrement sur les matrices $F_\Psi^t F_\Phi$ et $F_\Phi^t F_\Psi$ qui sont transposées l'une de l'autre (et chacune de ces matrices est orthogonale). En se souvenant que les colonnes de Ψ sont les vecteurs (qui forment des bases orthonormales) (ψ_1, \dots, ψ_N) et les colonnes de Φ sont $(\varphi_1, \dots, \varphi_N)$ on peut facilement exprimer les matrices précédentes avec:

$$F_\Psi^t F_\Phi = \begin{bmatrix} \langle \psi_1, \varphi_1 \rangle & \langle \psi_1, \varphi_1 \rangle & \cdots & \langle \psi_1, \varphi_N \rangle \\ \langle \psi_2, \varphi_1 \rangle & \ddots & \vdots & \langle \psi_2, \varphi_N \rangle \\ \vdots & \dots & \ddots & \vdots \\ \langle \psi_N, \varphi_1 \rangle & \cdots & \cdots & \langle \psi_N, \varphi_N \rangle \end{bmatrix} \quad (3.40)$$

En analysant les preuves, on peut voir que la quantité qui est essentielle est :

$$M = M_{\Phi, \Psi} = M_{\Psi, \Phi} = \sup_{1 \leq i, j \leq N} |\langle \psi_i, \varphi_j \rangle|. \quad (3.41)$$

On peut voir cette quantité comme la corrélation maximale entre les vecteurs de Φ et Ψ .

On voit en particulier, que dans les cas du dictionnaire Fourier-Dirac que l'on a considéré dans les théorèmes précédents, on a $M = \frac{1}{\sqrt{N}}$ et on voit que cette quantité peut directement être insérée dans les théorèmes 3.1.2.1, 3.1.2.2 pour le problème P0 et 3.2.0.1, 3.2.4 pour le problème P1, ce n'est pas une coïncidence et on va généraliser ces résultats ci-dessous.

Théorème 3.2.5.1. *Soit un signal $f \in \mathbb{R}^N$ non nul et soit (Φ, Ψ) une paire de bases orthonormales et M la quantité définie par 3.41, alors*

$$\|F_\Phi f\|_0 \|F_\Psi f\|_0 \geq \frac{1}{M^2} \quad (3.42)$$

et ainsi

$$\|F_\Phi f\|_0 + \|F_\Psi f\|_0 \geq \frac{2}{M}. \quad (3.43)$$

Théorème 3.2.5.2. *Soit N un entier positif, (Φ, Ψ) une paire de bases orthonormales, M la quantité définie par 3.41 et un signal $f \in \mathbb{R}^N$, alors n'importe quel x vérifiant $f = F_\Phi x_\Phi + F_\Psi x_\Psi$ et*

$$\|x_\Phi\|_0 + \|x_\Psi\|_0 < \frac{1}{M} \quad (3.44)$$

est l'unique solution de (P0).

Pour ces deux théorèmes, la démonstration est immédiate en remplaçant dans les preuves les matrices F_T et F_W par F_Φ et F_Ψ et la quantité \sqrt{N} par $\frac{1}{M}$.

Lemme 3.2.6. *Soit (Φ, Ψ) une paire de bases orthonormales et M la quantité définie par 3.41 et soient Γ_T et Γ_W des sous ensembles d'indices de T et respectivement de W vérifiant*

$$|\Gamma_T| < \frac{1}{2M} \quad \text{et} \quad |\Gamma_W| < \frac{1}{2M}, \quad (3.45)$$

alors on a,

$$\mu(\Gamma_T, \Gamma_W) < \frac{1}{2}. \quad (3.46)$$

Là aussi la démonstration est immédiate en faisant les changements adaptés dans la preuve. On a donc de la même façon que précédemment, en appliquant le lemme 3.2.3:

Theorème 3.2.6.1. *Soit N un entier positif, (Φ, Ψ) une paire de bases orthonormales, M la quantité définie par 3.41 et un signal $f \in \mathbb{R}^N$, alors n'importe quel $x = (x_\Phi, x_\Psi)$ vérifiant $f = F_\Phi x_\Phi + F_\Psi x_\Psi$ et*

$$\|x_T\|_0 < \frac{1}{2M} \quad \text{et} \quad \|x_W\|_0 < \frac{1}{2M} \quad (3.47)$$

est l'unique solution de P1, et c'est la solution de P0.

3.2.7 Extensions du résultat

Comme indiqué précédemment, le théorème 3.2.0.1 obtenu est une généralisation d'un théorème de David Donoho et Xiaoming Huo dans [9], ainsi, la généralisation du théorème 3.2.0.1 qu'est le théorème 3.2.6.1, est aussi une généralisation d'un autre théorème de l'article précédent.

Cependant, d'autres généralisations des résultats de David Donoho et Xiaoming Huo ont été proposées, cela tenant notamment de l'importance de l'article par exemple pour le développement du compressed sensing ou l'étude de la cohérence de paires de bases, mais aussi car comme indiqué par les auteurs les bornes ne semblaient pas exactes.

En effet, les auteurs indiquent que leur résultat devrait pouvoir être amélioré d'un facteur 2 optimal car un contre exemple peut être montré en considérant le peigne de Dirac pour lequel $\|x_T\|_0 = \|x_W\|_0 = \frac{\sqrt{N}}{2}$. Le théorème 3.2.6.1 prouvé ici montre que ce contre exemple correspond au cas limite à partir duquel les hypothèses du théorème ne sont plus vérifiées.

Une autre généralisation des résultats de [9] a été faite par Michael Elad et Alfred Bruckstein [12]. Ceux-ci ont en effet démontré le théorème suivant:

Theorème 3.2.7.1. *Soit N un entier positif, (Φ, Ψ) une paire de bases orthonormales, M la quantité définie par 3.41 et un signal $f \in \mathbb{R}^N$, alors n'importe quel $x = (x_\Phi, x_\Psi)$ vérifiant $f = F_\Phi x_\Phi + F_\Psi x_\Psi$ et*

$$\|x_T\|_0 + \|x_W\|_0 < \frac{\sqrt{2} - 0.5}{M} = \frac{0.9142}{M} \quad (3.48)$$

est l'unique solution de P1, et c'est la solution de P0.

La preuve de ce théorème est similaire en grande partie à celle de 3.2.0.1 présentée ici et dans [9], on pourra remarquer que le point de vue matriciel adopté dans ce mémoire est également celui qui est présenté par Michael Elad et Alfred Bruckstein. La différence essentielle entre la preuve présentée ici et celle de [12] est que dans l'article un problème variationnel est résolu alors qu'ici les preuves découlent directement d'inégalités matricielles. C'est d'ailleurs la même différence entre la preuve présentée ici de 3.2.0.1 et celle de David Donoho et Xiaoming Huo [9].

On a ainsi que le théorème de Michael Elad et Alfred Bruckstein et le théorème 3.2.6.1 présenté ici sont les deux vrais sur un domaine qui couvre celui sur lequel celui de David Donoho et Xiaoming Huo est vrai, ce sont donc bien des généralisations. On peut résumer la situation avec le graphe suivant : (TODO: ajouter graphe).

Une autre généralisation des résultats de Michael Elad et Alfred Bruckstein a été faite par Arie Feuer et Arkadi Nemirovski [13] dans laquelle ils démontrent que la borne de Michael Elad et Alfred Bruckstein est atteinte, et donc optimale. Montrer que la borne est atteinte revient à montrer qu'il y a au moins un signal qui atteint l'égalité du théorème 3.2.7.1, la construction d'un tel signal est relativement élaborée. On notera seulement que dans cette construction le nombre de coordonnées dans l'une des composantes vaut $\sqrt{2} \frac{\sqrt{N}}{2}$, ce n'est donc, heureusement, pas un contre exemple aux théorèmes 3.2.0.1 et 3.2.6.1 prouvés ici.

3.2.8 Définition de (P1)

3.2.9 Propriétés du minimiseur

3.3 Lien géométrique entre (P0) et (P1)

3.3.1 Boules unité en grande dimension

3.3.2 Unicité de la solution de (P0) et (P1)

Figure 3.1: Minimisation de $y = Fx$ pour la norme ℓ^1 . La solution de P1 est généralement celle de P0 sauf si les solutions sont parallèles à l'une des faces de la boule de ℓ^1 .

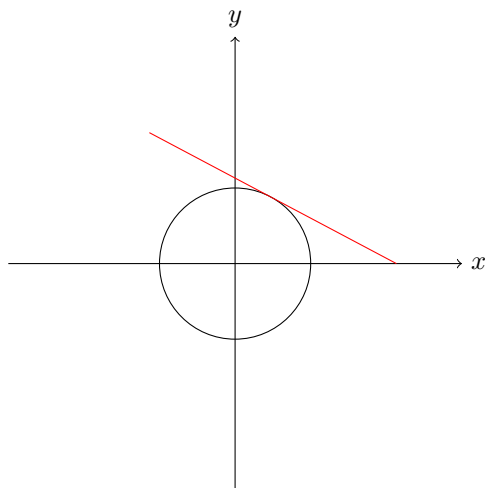
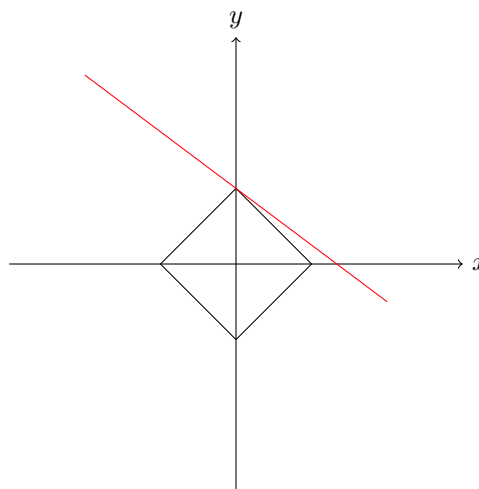


Figure 3.2: Minimisation de $y = Fx$ pour la norme ℓ^2 . La solution de ?? n'est généralement pas celle de P0 sauf si les solutions sont parallèles à l'un des axes.

Chapter 4

Compressed sensing et approche aléatoire

4.1 Axiomatisation, UUP et RIP

Dans ce chapitre on considère $\mathcal{F} \subset \mathbb{R}^N$ une classe de signaux. On cherche à pouvoir reconstruire chaque élément $f \in \mathcal{F}$ avec une précision ε en utilisant une famille de vecteurs $(\psi_k)_{k \in \Omega}$.

C'est à dire, on considère une application d'analyse,

$$\theta : \mathcal{F} \longrightarrow \mathbb{R}^{|\Omega|} \quad (4.1)$$

$$(f_k)_{k=0, \dots, N} = f \longmapsto (y_k = \langle f, \psi_k \rangle)_{k \in \Omega} = \theta(f) \quad (4.2)$$

et une application de synthèse associée

$$\mathbb{R}^{|\Omega|} \longrightarrow \mathcal{F}^\# \subset \mathbb{R}^N \quad (4.3)$$

$$(y_k)_\Omega \longmapsto (y_k)^\# = (f_k^\#)_{k=0, \dots, N} \quad (4.4)$$

et on cherche à obtenir une ε -reconstruction :

$$\|f - \theta(f)^\#\|_2 \leq \varepsilon, \quad \forall f \in \mathcal{F}. \quad (4.5)$$

Le problème est donc de choisir une famille $(\psi_k)_{k \in \Omega}$ pour qu'il soit possible d'obtenir la dernière inégalité.

On remarque aussi que $\|\theta(f)\|_0 \leq |\Omega|$, ainsi on cherchera à avoir un $K(\varepsilon) = K = |\Omega|$ dans la suite.

On considèrera F_Ω une matrice aléatoire avec $|\Omega|$ lignes et N colonnes dont les coefficients suivent une distribution de probabilités. On considère aussi que $|\Omega|$ est aussi une variable aléatoire à valeurs dans $\{0, \dots, N\}$ et on notera $K = \mathbb{E}(|\Omega|)$. On note

$$R_\Omega : \ell^2([0, N]) \longrightarrow \ell^2(\Omega) \quad (4.6)$$

$$(g_k)_{0 \leq k \leq N} \longmapsto (g_k)_{k \in \Omega} \quad (4.7)$$

et l'inclusion prolongée par des zéros

$$R_T^* : \ell^2(T) \longrightarrow \ell^2([0, N]) \quad (4.8)$$

$$(g_k)_{k \in T} \longmapsto (g_k)_{k \in T} \oplus (0)_{k \in T^c}. \quad (4.9)$$

On considèrera aussi par la suite la matrice aléatoire $F_{\Omega T}$ en conservant que les $|T|$ colonnes indexées par T de la matrice F_Ω , c'est à dire :

$$F_{\Omega T} = F_\Omega R_T^* : \ell^2(T) \longrightarrow \ell^2(\Omega) \quad (4.10)$$

$$(g_k)_T \longmapsto F_\Omega((g_k)_T \oplus (0)_T^c). \quad (4.11)$$

On remarque aussi que $F_{\Omega T}^* F_{\Omega T} : \ell^2(T) \rightarrow \ell^2(T)$ est symétrique et que l'on peut la diagonaliser sous la forme $U \Lambda U^*$ où $\Lambda = (\lambda_1 \geq \dots \geq \lambda_{|T|})$ sont les valeurs propres de $F_{\Omega T}^* F_{\Omega T}$.

4.1.1 Définition de UUP

On peut alors définir le principe uniforme d'incertitude (Uniform Uncertainty Principle),

Définition 4.1.2. On dit que F_Ω vérifie λ -**UUP** si il existe ρ tel que avec probabilité $1 - \mathcal{O}(N^{-\rho/\alpha})$ on ait:

$\forall f \subset \mathbb{R}^N$ signal tel que

$$|\text{supp}(f)| \leq \alpha K / \lambda \quad (4.12)$$

on ait l'inégalité

$$\frac{1}{2} \frac{K}{N} \|f\|_2^2 \leq \|F_\Omega f\|_2^2 \leq \frac{3}{2} \frac{K}{N} \|f\|_2^2. \quad (4.13)$$

On aurait aussi pu définir le principe uniforme d'incertitude à l'aide des valeurs propres :

Proposition 4.1.3. F_Ω vérifie λ -**UUP** si et seulement si

avec probabilité au moins $1 - \mathcal{O}(N^{-\rho/\alpha})$ on a $\forall T \subset [0, N]$ qui vérifie $|T| \leq \alpha \frac{K}{\lambda}$ alors les valeurs propres de $F_{\Omega T}$ vérifient

$$\frac{1}{2} \frac{K}{N} \leq \lambda_{\min}(\Lambda) \leq \lambda_{\max}(\Lambda) \leq \frac{3}{2} \frac{K}{N}.$$

Preuve 4.1.4. A recopier.

Remarque 4.1.5. Pour expliciter le fait que cela définit bien un principe d'incertitude, considérons F_Ω comme étant la transformée de fourier discrète partielle, et un signal concentré en temps ($|\text{supp}(f)| \leq \alpha \frac{K}{\lambda}$), alors on a

$$\|F_\Omega f\|_{\ell^2} = \|\hat{f}\|_{\ell^2(\Omega)} \leq \sqrt{\frac{3K}{2N}} \|f\|_{\ell^2} \quad (4.14)$$

en appliquant le principe d'incertitude. On déduit donc que

$$\frac{\|\hat{f}\|_{\ell^2(\Omega)}}{\|\hat{f}\|_{\ell^2}} \longrightarrow 0 \quad (4.15)$$

si $K = o(N)$, c'est à dire que si f est à support compact, il est nécessaire d'avoir un nombre de mesures K qui est au moins de l'ordre de f . Donc f ne peut pas être localisé à la fois en temps et en fréquence, ce qui justifie l'appellation "principe d'incertitude".

Remarque 4.1.6. Justifions maintenant le fait que c'est un principe uniforme. Une version non uniforme (et donc plus faible) serait que pour chaque f vérifiant 4.12, alors avec probabilité au moins $1 - \mathcal{O}(N^{-\rho/\alpha})$ 4.13 est vérifié. Mais il y a beaucoup de choix possibles de f vérifiant 4.12, et parmi ceux-ci il peut y avoir un grand nombre de f ayant la propriété rare de ne pas vérifier 4.13, et alors l'union de ces événements n'a pas nécessairement une faible probabilité de se produire.

Ainsi, le principe est uniforme car la propriété **UUP** est telle que l'on a une probabilité au moins $1 - \mathcal{O}(N^{-\rho/\alpha})$ que 4.13 soit vrai pour tous les f possibles vérifiant 4.12. Ce qui justifie l'appellation uniforme.

Remarque 4.1.7. ¹ Remarquons que l'on peut réécrire 4.13 peut se réécrire

$$(1 - \delta_K) \|f\|_2^2 \leq \|F_\Omega f\|_2^2 \leq \|f\|_2^2 (1 + \delta_K)$$

avec $\delta = 1 - \frac{K}{2N}$ ce qui rappelle la définition d'un frame avec des bornes $m = M = \frac{1}{2}$ dans le meilleur des cas. Cela justifie que certaines fois le principe uniforme d'incertitude est aussi appelé propriété d'isométrie restreinte (**RIP**) (Restricted Isometry Property).

4.1.8 Exemple de familles vérifiant UUP

Proposition 4.1.9. ²

- Les ensembles Gaussiens et binaires vérifient $\log N - UUP$
- L'ensemble de Fourier vérifie $(\log N)^6 - UUP$.

4.1.10 Définition de ERP

Un autre principe que l'on va utiliser qui nous permettra de nous assurer que l'approximation $f^\#$ obtenue est proche de f pour la norme ℓ^1 est le principe de reconstruction exacte (**ERP** - Exact Reconstruction Principle).

Définition 4.1.11. F_Ω vérifie **ERP** si

¹A vérifier

²Pour certains résultats concernant ERP et UUP : <https://www.math.ucla.edu/~tao/preprints/sparse.html>

- $\forall T \subset [0, N]$
- $\forall \sigma \in \{\pm 1\}^T$

il existe avec probabilité prépondérante, un vecteur $P \in \mathbb{R}^N$ tel que

1. $P(t) = \sigma(t), \forall t \in T$
2. P est une combinaison linéaire des lignes de F_Ω ³
3. $P(t) < \frac{1}{2}, \forall t \in T^{c4}$

4.1.12 Exemples de familles vérifiant ERP

Proposition 4.1.13. • *Les ensembles Gaussiens et binaires vérifient log N-ERP*

- *L'ensemble de Fourier vérifie log N-ERP.*

4.1.14 Lien entre RIP et ERP

4.2 Théorème de Candes-Tao

4.2.1 Enoncé du théorème

Theorème 4.2.1.1. *Soit F_Ω qui vérifie λ_1 -ERP et λ_2 -UUP. On pose $\lambda = \max(\lambda_1, \lambda_2)$, soit $K \geq \lambda$.*

Soit f un signal dans \mathbb{R}^N tel que ses coefficients dans une base de référence décroissent comme⁵ :

$$|\theta_{(n)}| \leq C n^{-\frac{1}{p}} \quad (4.16)$$

pour un certain $C > 0$ et $0 < p \leq 1$.

On pose $r = \frac{1}{p} - \frac{1}{2}$, alors n'importe quel minimiseur de (P1) vérifie :

$$\|f - f^\#\|_2 \leq C_r \left(\frac{K}{\lambda}\right)^{-r} \quad (4.17)$$

avec probabilité au moins $1 - \mathcal{O}(n^{-\frac{\rho}{\alpha}})$, pour certains ρ et α .

³C'est équivalent à P appartient au *rowspace* de F_Ω , ce qui est équivalent à : $\exists Q$ tel que $P = F_\Omega^* Q$.

⁴Le $\frac{1}{2}$ n'a pas vraiment d'importance, n'importe quelle constante $0 < \beta < 1$ permet d'obtenir les mêmes résultats

⁵les coefficient $(|\theta_{(n)}|)$ sont triés par ordre décroissant

4.2.2 Preuve du théorème

4.3 Exemple de F_Ω

4.3.1 Ensemble de Fourier

4.3.2 Gaussien

4.4 Conséquences du théorème

4.4.1 Influence des paramètres

4.4.2 Quelques résultats numériques

4.5 Sur la propriété RIP

4.5.1 Difficulté pour un ensemble de vérifier RIP

4.5.2 Lien entre RIP et WERP

4.6 Extensions du théorème

4.6.1 Conditions suffisantes sur δ_K

4.6.2 Conditions nécessaires sur δ_K

4.6.3 Sur l'optimalité du résultat

4.7 Algorithmes

4.7.1 Orthogonal Matching Pursuit

4.7.2 Robust Orthogonal Matching Pursuit

4.7.3 Quelques exemples numériques

Appendix A

Annexe

- A.1 Valeurs propres de $F_{\Omega}F_{\Omega}^*$ et Analyse en composante principale
- A.2 Outils probabilistes de la preuve du théorème

Bibliography

- [1] Aldroubi, Akram, Cabrelli, Carlos, and Molter, Ursula M. “Wavelets on irregular grids with arbitrary dilation matrices and frame atoms for $L_2(\mathbb{R}^d)$ ”. In: *Applied and Computational Harmonic Analysis* 17.2 (2004). Special Issue: Frames in Harmonic Analysis, Part II, pp. 119–140. ISSN: 1063-5203. DOI: <https://doi.org/10.1016/j.acha.2004.03.005>. URL: <https://www.sciencedirect.com/science/article/pii/S1063520304000442>.
- [2] Alexandre, Grothendieck. *Produits tensoriels topologiques et espaces nucléaires* / A. Grothendieck. fre. Memoirs of the American Mathematical Society. Providence: American Mathematical Society, 1955. ISBN: 0-8218-1216-5.
- [3] Candès, Emmanuel and Romberg, Justin. “Sparsity and incoherence in compressive sampling”. In: *Inverse Problems* 23.3 (2007), pp. 969–985. DOI: 10.1088/0266-5611/23/3/008. URL: <https://doi.org/10.1088/0266-5611/23/3/008>.
- [4] Candes, Emmanuel J. and Tao, Terence. “Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?” In: *IEEE Transactions on Information Theory* 52.12 (2006), pp. 5406–5425. DOI: 10.1109/TIT.2006.885507.
- [5] Choi, Kihwan et al. “Compressed sensing based cone-beam computed tomography reconstruction with a first-order methoda)”. In: *Medical Physics* 37.9 (2010), pp. 5113–5125. DOI: <https://doi.org/10.1118/1.3481510>. eprint: <https://aapm.onlinelibrary.wiley.com/doi/pdf/10.1118/1.3481510>. URL: <https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.3481510>.
- [6] Daubechies, I. “Where do wavelets come from? A personal point of view”. In: *Proceedings of the IEEE* 84.4 (1996), pp. 510–513. DOI: 10.1109/5.488696.
- [7] Daubechies, Ingrid. “3. Discrete Wavelet Transforms: Frames”. In: *Ten Lectures on Wavelets*, pp. 53–105. DOI: 10.1137/1.9781611970104.ch3. eprint: <https://epubs.siam.org/doi/pdf/10.1137/1.9781611970104.ch3>. URL: <https://epubs.siam.org/doi/abs/10.1137/1.9781611970104.ch3>.
- [8] Daubechies, Ingrid. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, 1992. DOI: 10.1137/1.9781611970104. eprint: <https://epubs.siam.org/doi/pdf/10.1137/1.9781611970104>. URL: <https://epubs.siam.org/doi/abs/10.1137/1.9781611970104>.

- [9] Donoho, D. L. and Huo, X. “Uncertainty Principles and Ideal Atomic Decomposition”. In: *IEEE Trans. Inf. Theor.* 47.7 (Sept. 2006), 2845–2862. ISSN: 0018-9448. DOI: 10.1109/18.959265. URL: <https://doi.org/10.1109/18.959265>.
- [10] Donoho, David L. “For most large underdetermined systems of linear equations the minimal 1-norm solution is also the sparsest solution”. In: *Communications on Pure and Applied Mathematics* 59.6 (2006), pp. 797–829. DOI: <https://doi.org/10.1002/cpa.20132>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpa.20132>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.20132>.
- [11] Donoho, D.L. “Compressed sensing”. In: *IEEE Transactions on Information Theory* 52.4 (2006), pp. 1289–1306. DOI: 10.1109/TIT.2006.871582.
- [12] Elad, M. and Bruckstein, A.M. “A generalized uncertainty principle and sparse representation in pairs of bases”. In: *IEEE Transactions on Information Theory* 48.9 (2002), pp. 2558–2567. DOI: 10.1109/TIT.2002.801410.
- [13] Feuer, Arie, Member, Senior, and Nemirovski, Arkadi. “On sparse representations in pairs of bases”. In: *IEEE Trans. Inf. Theory* (2003), pp. 1579–1581.
- [14] Haar, Alfred. “Zur Theorie der orthogonalen Funktionensysteme”. In: *Mathematische Annalen* 69.3 (1910), pp. 331–371. DOI: 10.1007/BF01456326. URL: <https://hal.archives-ouvertes.fr/hal-01333722>.
- [15] Loris, Ignace et al. “Tomographic inversion using 1-norm regularization of wavelet coefficients”. In: *Geophysical Journal International* 170.1 (July 2007), pp. 359–370. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2007.03409.x. eprint: <https://academic.oup.com/gji/article-pdf/170/1/359/5924549/170-1-359.pdf>. URL: <https://doi.org/10.1111/j.1365-246X.2007.03409.x>.
- [16] Lustig, Michael et al. “Compressed Sensing MRI”. In: *IEEE Signal Processing Magazine* 25.2 (2008), pp. 72–82. DOI: 10.1109/MSP.2007.914728.
- [17] Mallat, Stphane. *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*. 3rd. USA: Academic Press, Inc., 2008. ISBN: 0123743702.
- [18] Meyer, Yves. *Wavelets and Operators*. Ed. by Salinger, D. H. Translator. Vol. 1. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1993. DOI: 10.1017/CB09780511623820.
- [19] Queffélec Hervé et Queffélec, Martine. *Analyse complexe et applications : cours et exercices / Hervé Queffélec, Martine Queffélec*. fr. Mathématiques en devenir. Paris: Calvage et Mounet, DL 2019. ISBN: 978-2-916352-59-6.
- [20] Richman, Fred, Bridges, Douglas, and Schuster, Peter. “Trace-class operators”. In: *Houston Journal of Mathematics* 28.3 (2002), pp. 565–583.
- [21] Robert, Didier. “Sur les traces d’opérateurs (De Grothendieck à Lidskii)”. working paper or preprint. Apr. 2014. URL: <https://hal.archives-ouvertes.fr/hal-01015295>.

- [22] Stéphane, Mallat. “CHAPTER 5 - Frames”. In: *A Wavelet Tour of Signal Processing (Third Edition)*. Ed. by Stéphane, Mallat. Third Edition. Boston: Academic Press, 2009, pp. 155–204. ISBN: 978-0-12-374370-1. DOI: <https://doi.org/10.1016/B978-0-12-374370-1.00009-4>. URL: <https://www.sciencedirect.com/science/article/pii/B9780123743701000094>.
- [23] Tillmann, Andreas M. “On the Computational Intractability of Exact and Approximate Dictionary Learning”. In: *IEEE Signal Processing Letters* 22.1 (2015), pp. 45–49. DOI: 10.1109/LSP.2014.2345761.
- [24] Trad, Daniel, Ulrych, Tadeusz, and Sacchi, Mauricio. “Latest views of the sparse Radon transform”. In: *Geophysics* 68.1 (Jan. 2003), pp. 386–399. ISSN: 0016-8033. DOI: 10.1190/1.1543224. eprint: https://pubs.geoscienceworld.org/geophysics/article-pdf/68/1/386/3203946/gsgpy_68_1_386.pdf. URL: <https://doi.org/10.1190/1.1543224>.
- [25] Wang, Benfeng et al. “A Robust and Efficient Sparse Time-Invariant Radon Transform in the Mixed Time-Frequency Domain”. In: *IEEE Transactions on Geoscience and Remote Sensing* 57.10 (2019), pp. 7558–7566. DOI: 10.1109/TGRS.2019.2914086.
- [26] Zheng, Xuehang et al. *Sparse-View X-Ray CT Reconstruction Using ℓ_1 Prior with Learned Transform*. 2019. arXiv: 1711.00905 [stat.ML].