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

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 $[\![0, N-1]\!] \to \mathbb{R}$. Une autre famille que l'on étudiera est $L^2(\mathbb{R})$ et de façon plus générale les espaces de Hilbert.
L'objectif de ce mémoire est de présenter des techniques permettant la reconstruction de signaux, donc on cherchera à obtenir des opérateurs $F : \mathcal{F} \to \mathcal{F}$ 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.

Comme nous souhaitons reconstruire des signaux, notre objectif est de pouvoir

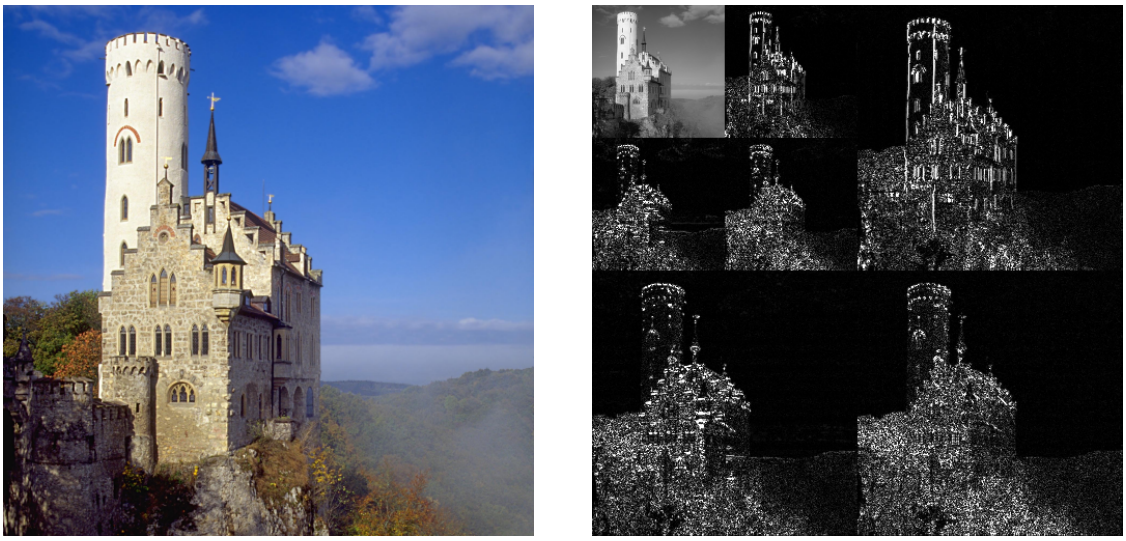


Figure 1.1: Une photographie et sa tranformée en ondelettes.[a]

[a]Images de Wikicommons

utiliser ces coefficients afin de revenir dans l'espace initial. On définit donc l'application

$$S : \mathcal{G} \longrightarrow \mathcal{F} \tag{1.3}$$
$$A(f) \longmapsto S(A(f)) \tag{1.4}$$

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 Mallat[1], approcher la problématique de la recontruction 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, à l'aide d'un principe d'incertitude, l'hypothèse que le signal est représentable avec

[1]L'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"*.

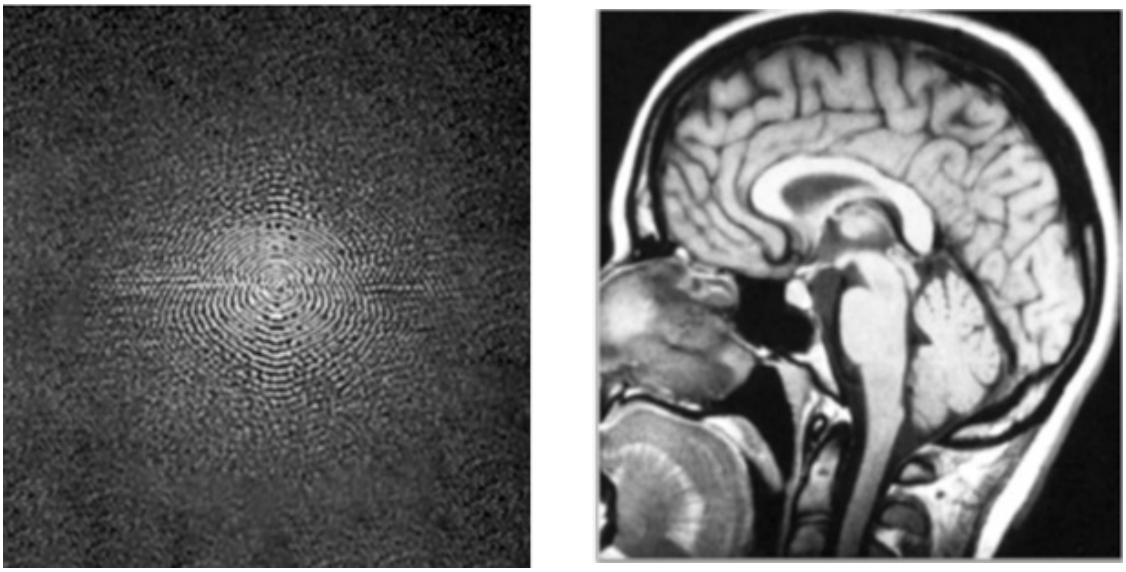


Figure 1.2: Mesure par résonnance magnétique d'un cerveau (le signal), qui correspond à des coefficients de Fourier (l'analyse), et reconstruction du signal (la synthèse).[a]

[a]Images du site `www.MRIquestions.com` de Allen D. Elster.

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 (c'est le principe de l'algorithme Best Ortho Basis de Coifman et Meyer [11]). 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 coiflet à 2 moments nuls.[a]

[a]Images 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

$$f = \sum_J c_i f_i. \tag{1.5}$$

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 appellée 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, on appelera ce problème P0.
Ce problème est de nature combinatoire et très difficile à résoudre de façon générale (c'est un problème NP-difficile [35]) On verra dans le chapitre 3 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 [18], [5], [4]. Dans ce cadre, on souhaite reconstruire un signal $x_0 \in \mathbb{R}^N$ avec $k << N$ 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 connait 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 niveux de gris de taille 1000×1000, donc un vecteur dans $\mathbb{R}^d$ avec $d = 10^6$. Supposons que ce vecteur a 20000 coefficients d'ondelette non nuls, alors avec 10^5 mesures, en cherchant une décomposition atomique qui coincide avec ces mesures, on devrait pouvoir reconstruire exactement x_0. Le mécanisme (d'après [16]) 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 coincider 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" 0 $||\cdot||_0$ 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 sysmique 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 [24], pour l'imagerie par rayons X [10] [38] et plus généralement pour des problèmes d'inversion de la transformée de Radon [37] [36]. 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 [25] 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 [25]) et donc il n'est pas possible de connaitre 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 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 résultat 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 résolvables 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 suffisament 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 connait 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 relacher certaines des définitions initiales afin de pouvoir conserver une formule de reconstruction. Afin d'expliciter 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[1].

[1]En fait la représentation obtenue avec un frame n'est généralement pas exactement parcimonieuse, de nombreux coefficients peuvent avoir une petite valeur différente de 0, cependant il est possible de conserver une bonne formule de reconstruction avec moins de coefficients, soit par seuillage, soit par les techniques plus élaborées du chapitre 4

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 $h_1 = (\lambda_i)_J, h_2 = (\mu_i)_2$

$$\langle\cdot,\cdot\rangle : H \times H \longrightarrow \mathbb{R} \tag{2.1}$$

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

On peut remarquer que ce produit scalaire est défini de façon unique par rapport à la base hilbertienne $\{e_i\}_I$, cependant, si $U : H \to H$ est un opérateur unitaire, ce qui correspond en dimension finie à une rotation ou à un changement de base, alors on a $\langle Uh_1, Uh_2\ \rangle = \langle h_1, h_2\rangle$. 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:

Theorème 2.1.3.1. *Soit* $\{f_i\}_I$ *une suite d'éléments orthonormaux dans* H *muni d'un produit scalaire. Alors* $\overline{Vect(\{f_i\}_I)} = H$ *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 relacher 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 :

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

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

.

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,

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

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

.

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*

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

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

$$m||f||^2 \leqslant \sum_I |\langle f, f_i\rangle|^2 \leqslant M||f||^2, \forall f \in H. \tag{2.3}$$

3. *Une base de Riesz (ou base inconditionnelle) si il existe des constantes $M, m > 0$ telles que*
$$m \sum |c_k|^2 \leqslant || \sum c_k f_k||^2 \leqslant M \sum |c_k|^2$$
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 :

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

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 coordonées 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

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

Cependant cette formule peut sembler ajouter un inconvenient 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 appelera un tel frame un frame redondant. Par ailleurs, une démonstration est donnée en annexe A.1.2 qui montre que la solution obtenue par les coefficients de frame est la solution minimale pour la norme ℓ^2.
En annexe A.1.1 une généralisation de la formule de reconstruction est donnée pour des frames non serrés, celle ci est obtenue avec un algorithme ayant une vitesse de convergence expontielle.

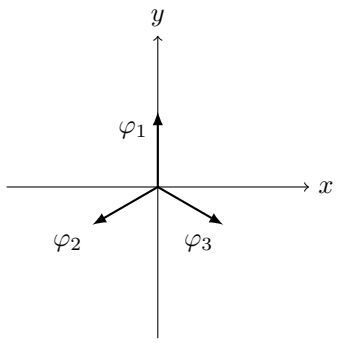


Figure 2.1: Le frame équilibré $\Phi = \{\varphi_1, \varphi_2, \varphi_3\}$ de $\mathbb{R}^2$

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,

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

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 :

$$x = \frac{1}{M}(\sum_{i=1}^{3} \langle x, \varphi_i \rangle \varphi_i) = \frac{2}{3}(\langle x, \varphi_1 \rangle \varphi_1 + \langle x, \varphi_1 \rangle \varphi_1 + \langle x, \varphi_1 \rangle \varphi_1) \tag{2.6}$$

Ainsi, comme aperçu par Jean Morlet dès 1986 ([13]) 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 [13] 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 [33], 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 $F = (f_i)_{i=1,\cdots,N}$ serré (c'est-à-dire $m = M$) redondant (c'est-à-dire $m > 1$). Ce frame engendre un espace vectoriel $Vect(F)$ de dimension $d \leqslant N$ 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 $(e_i)_{i=1,\cdots,d}$, 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 $(\alpha_i)_{i=1,\cdots,d}$ 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

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

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 $(\alpha_i)_{i=1,\cdots,N}$ de moyenne nulle et

de variance égale à 1.

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

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

$$\frac{N}{M} \leqslant d. \qquad (2.8)$$

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:

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

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

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

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

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,

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

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

et on considère, sans perte de généralité que chaque φ_i est normalisé. Ainsi chaque à chaque φ_i, on peut associer une suite $(\lambda_j^i)_{j=0,\cdots,d} = (\langle\varphi_i, e_j\rangle)_{j=0,\cdots,d}$ 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

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

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

$$c_j \leqslant ||x||_2 \sqrt{M}. \tag{2.15}$$

On a maintenant la majoration,

$$||S(x)||_2^2 = || \sum_{j=1}^{d} c_j e_j ||_2^2 = \sum_{j=1}^{d} c_j^2 \leqslant dM ||x||_2^2. \tag{2.16}$$

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

$$||S(x)||_2^2 = || \sum_{i=1}^{N} x_i \sum_{j=1}^{d} \lambda_j^i e_j ||_2^2 \geqslant N ||x||_2^2 \min_{i=1,\cdots,N} || \sum_{j=1}^{d} \lambda_j^i e_j ||_2^2 = N ||x||_2^2 \min_{i=1,\cdots,N} \sum_{j=1}^{d} (\lambda_j^i)^2. \tag{2.17}$$

Or,

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

On peut donc combiner les inégalités obtenues et on a

$$N ||x||_2^2 \leqslant ||S(x)||_2^2 \leqslant dM ||x||_2^2 \tag{2.19}$$

et ainsi

$$\frac{N}{M} \leqslant d. \tag{2.20}$$

□

Comme discuté dans la remarque 2.1.6, un résultat plus général est présenté dans le livre de Stéphane Mallat [33], le voici:

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

$$m \leqslant \frac{N}{d} \leqslant M. \tag{2.21}$$

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:

$$d \leqslant \frac{N}{m} \quad \text{et} \quad \frac{N}{M} \leqslant d. \tag{2.22}$$

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 $m \leqslant M$. Donc pour arriver au résultat de Stéphane Mallat, il reste à montrer l'inégalité

$$d \leqslant \frac{N}{m}. \tag{2.23}$$

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 $x \in Im(A) \subset \ell^2(N)$,

$$\langle Af, x \rangle = \langle f, A^* x \rangle. \tag{2.24}$$

On peut alors remarquer

$$||Af||^2 = \sum_{i=1,\cdots,N} \langle f, f_i \rangle^2 = \langle A^* Af, f \rangle. \tag{2.25}$$

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

$$m||f||^2 \leqslant \langle A^* Af, f \rangle \leqslant M||f||^2. \tag{2.26}$$

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

$$dm \leqslant Tr(A^* A) \leqslant dM \tag{2.27}$$

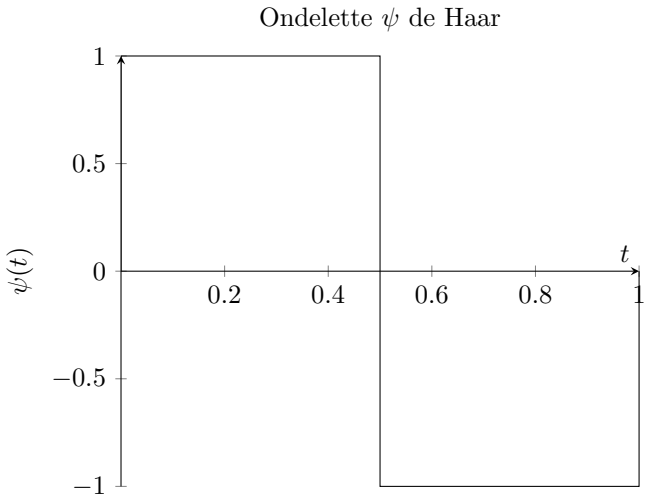
car A^*A agit sur un espace isométrique à $\ell^2(d)$. En utilisant le fait que $Tr(A^*A) = Tr(AA^*)$, on a:

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

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

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 plûtot que de Riesz ou Hilbert. On pourra consulter [32] notamment pour le lien qui est fait entre les définitions de ces bases et leurs liens avec les ondelettes. 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 $1 \leqslant p < \infty$, 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 [30], par leur similarité à l'analyse complexe on peut consulter [29], pour un point de vue historique sur leur développement on peut consulter [31].

2.1.8 Frames d'ondelettes et analyse multi-résolution

Historiquement, au début du XXème siècle, le problème de la non-convergence de la série de Fourier de certaines fonctions continues amena David Hilbert à se demander si ce phénomène était inhérent aux bases orthogonales de $L^2([0,1])$. C'est ainsi que sous sa direction Alfréd Haar introduit dans sa thèse [22] 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 commencé 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 [12] 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 [28]. A partir de là de nombreuses constructions d'ondelettes fûrent 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 [14] ainsi que l'algorithme de transformée en ondelette discret et le concept d'analyse multi-résolution de Stéphane Mallat [27] qui ont permis aux ondelettes de devenir un outil maintenant très répandu dans les mathématiques appliquées et que l'on utilisera dans la suite du mémoire. 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. On pourra à nouveau consulter [32] pour un exposé général sur le développement et le rôle des ondelettes dans les mathématiques actuelles. Etudions donc ces ondelettes:

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

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

est une famille génératrice de $L^2(\mathbb{R})$. On appelera[2] 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'échelle[3], 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

[2]On considérera par la suite des frames d'ondelettes ou bien des bases de Riesz d'ondelettes

[3]Du 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 Stéphane Mallat et Ingrid Daubechies, utilisent $-j$ par rapport à nos notations. Par contre les notations utilisées correspondent à celles de Stéphane Jaffard et Yves Meyer. Mais cependant tous les résultats sont bien entendu équivalents.

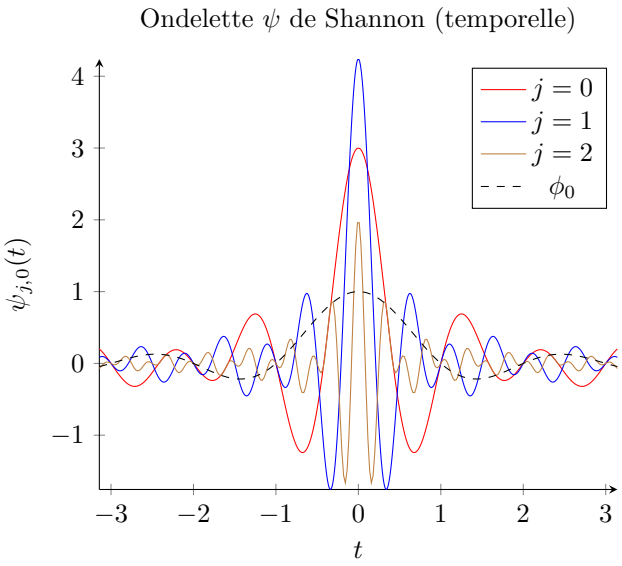


Figure 2.2: L'ondelette $\psi(t) = 2sinc(2t) - sinc(t)$ dite de Shannon avec différentes dilatations dyadiques ψ_j. Lorsque j augmente, la fonction est parcourue plus rapidement d'un facteur 2^j.

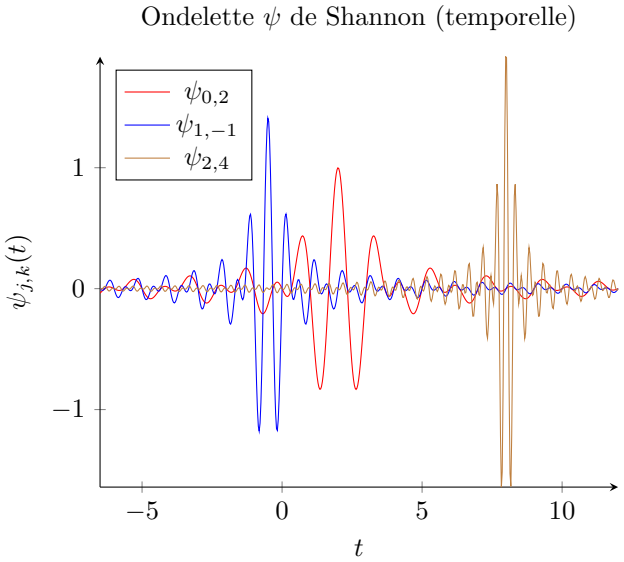


Figure 2.3: Des versions translatées $\psi_{j,k}$ des ondelettes ψ_j du graphe précédent. Lorsque j augmente, la fonction est parcourue plus rapidement d'un facteur 2^j.

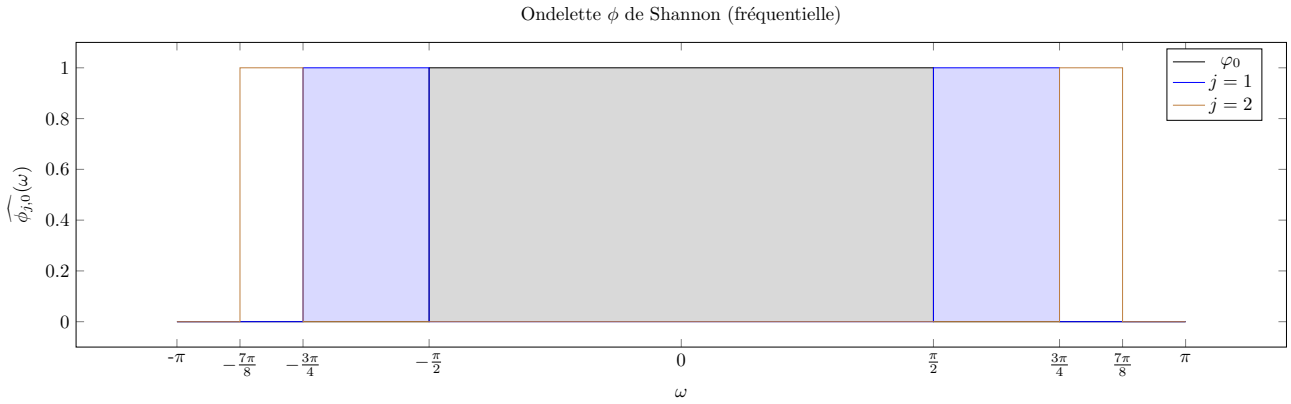


Figure 2.4: La transformée de Fourier $\hat{\psi}$ de l'ondelette de Shannon Ψ. L'ondelette d'échelle ϕ_0 est un filtre passe-bande $[-\frac{\pi}{2}, \frac{\pi}{2}]$, les ondelettes ψ_j sont des filtres passe bande qui recouvrent le reste de l'intervalle avec des supports disjoints de taille 2^{-j}. 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).

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:

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

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

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

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:

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

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

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

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, coincide 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

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

et on se demande alors si à partir de ces coefficients on peut reconstruire f. Cela revient

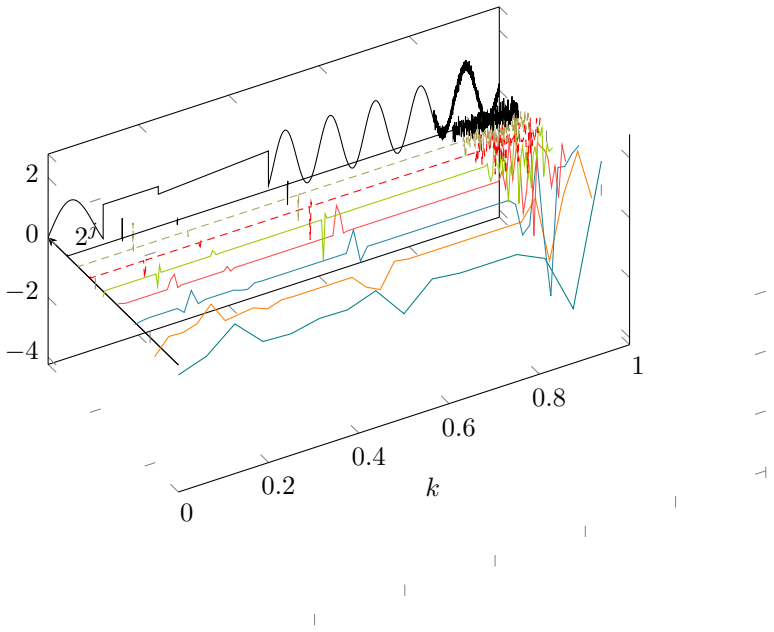


Figure 2.5: Les coefficients d'ondelette (pour l'ondelette de Daubechies D4) d'un signal présentant des zones de différentes régularité. Les coefficients les plus proches du signaux correspondent aux échelles les plus fines. On remarque que ces coefficients sont très petits dans les zones sans discontinuités.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

ainsi à déterminer si la famille d'ondelettes est un frame d'ondelette. Ici on ne cherchera pas a énumérer et a vérifier des frames d'ondelettes, un très grand nombre de frames d'ondelettes existent[4], on admet ainsi l'existence des frames d'ondelettes.
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

[4]Pour une introduction au sujet des ondelettes par les frames on peut recommander le livre d'Ingrid Daubechies [14], pour un théorème donnant une construction simple d'ondelettes en dimension $d \geqslant 1$ avec un échantillonage arbitraire (au lieu de l'échantillonage dyadique) on peut consulter [1] et pour une liste d'ondelettes `https://en.wikipedia.org/wiki/Wavelet#List_of_wavelets`

de reconstruction (d'après Daubechies 3.2.2)

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

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 convergence (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 emboités satisfaisant

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

L'intérêt d'avoir une telle suite d'espaces emmboités 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 emboités 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

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

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 W_0. Ajoutons maintenant la condition que V_0 contient toutes les translations entières de ses éléments, c'est à dire

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

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},

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

donc afin d'avoir une formule de reconstruction, il suffit de connaitre un frame de V_0 et de même pour chaque W_j. On peut maintenant revenir aux ondelettes, on considère que l'on connait $\varphi \in L^2(\mathbb{R})$ 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 $W_j = W_{j-1}\bigoplus Vect(\{\psi_{j,k}\}_{k\in\mathbb{Z}})$ et on obtient ainsi que l'analyse multi-échelle ainsi

construite fournit une formule de reconstruction[5]

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

On appelle l'application φ_0 ondelette d'échelle.

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 décroit 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 $(\tilde{\phi}_k)_{k=1,\cdots,N}$ et la formule de reconstruction :

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

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 :

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

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

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

[5]Dans 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 suffisament rapide, donc seulement un nombre fini de $\langle f, \varphi_{0,k} \rangle$ sont différents de 0 si φ_0 a une décroissance suffisament rapide (et de même pour chaque ψ_j).

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

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

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

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

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.

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

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

alors on a

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

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

Preuve. On développe le terme au centre de l'égalité dela 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

$$\int_0^M x^{2r-1} dx \leqslant \sum_{N=0}^{m} N^{2r-1} \leqslant \int_1^{m+1} x^{2r-1} dx. \tag{2.47}$$

En calculant les deux intégrales on déduit

$$Am^{2r} \leqslant \sum_{N=0}^{m} N^{2r-1} \leqslant Bm^{2r} \tag{2.48}$$

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

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

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

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

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

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

□

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

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

alors l'approximation linéaire dans la base $\mathcal{B}$ décroit 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 l'article de Stéphane Jaffard [23] ou bien dans le livre de Stéphane Mallat [27].

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 $\mathcal{B}_N = (e_n)_{|n|\leqslant N}$ 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 $f \in L^2(\mathbb{R})$

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

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

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

et en utilisant la formule de Plancherel on a

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

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:

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

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 décroit 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 $r - 1 \leqslant \alpha < r$. 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

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

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 $0 < \alpha \leqslant r$, 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

$$Wf(j,k) = f \star \widetilde{\psi_j}(k). \tag{2.54}$$

On peut ainsi interpreter 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 commun[6].

Rappelons aussi qu'un 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

$$\int \psi(t)p(t)dt = 0. \tag{2.55}$$

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 coefficient 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,

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

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

$$|Wf(j,k)| \leqslant C2^{-j(\alpha+\frac{1}{2})}(1 + |2^{-j}t - t_0|2^{j\alpha}).$$

Preuve. Remarquons tout d'abord que, soit P un polynôme de degré strictement inférieur à m, et ψ une ondelette a 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 $|f(t) - P_{t_0}(t)| \leqslant C|t - t_0|^\alpha$. En utilisant la linéarité de l'intégrale et en appliquant une inégalité triangulaire on obtient

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

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

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

Pour obtenir l'inégalité suivante on utilise

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

et on a ainsi

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

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. □

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

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

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

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

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

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

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

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

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

on utilise maintenant le théorème 2.2.7.1 et on obtient[7]

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

ce qui prouve la proposition. □

[7]Le 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'apparait pas.

Chapter 3

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

3.1 Introduction à (P0)

Dans ce qui précède, nous nous sommes intéressés aux propriétés qui font qu'une famille de vecteurs permet de reconstruire une famillle 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 efficace[1]. 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 naive, afin de déterminer une fonction, il faut connaitre 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.

[1] Une 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 connait 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.
Avant de poursuivre, il est important de remarquer que résoudre un tel problème requiert à priori de choisir la bonne décomposition parmi toutes les colonnes possibles qui est un problème combinatoire très difficile. On va voir que pour exploiter la parcimonie on utilisera de façon répétitive la norme $||\cdot||_1$ pour trouver la solution parcimonieuse, contrairement au chapitre précédent dans lequel la norme $||\cdot||_2$ était omniprésente. On peut donner une intuition géométrique à la préférence de la norme 1 par rapport à la norme 2 pour résoudre un problème ayant un lien avec la parcimonie. On verra que ce que l'on cherche à résoudre est une équation du type $y = Fx$, donc pour trouver la solution minimale pour une certaine norme, il suffit de prendre la boule centrée en 0 de plus petit rayon avec la norme correspondante et la solution minimale revient à trouver l'intersection entre cette boule et le plan des solutions de $y = Fx$.

On peut donc représenter cela avec la figure **??**, la boule unité pour la norme $||\cdot||_0$ correspondant aux axes (privés de l'origine), il est clair que la plupart des équation du type $y = Fx$ auront leur solution en norme 1 qui coincide avec celle en norme 0. Par contre, un tel phénomène n'est pas vérifié pour la norme euclidienne.

3.2 Résolution de (P0)

Formalisons maintenant ce que nous avons dit ci-dessus. On considère $\mathcal{F}$ un espace vectoriel et utilisons un dictionnaire $\Phi = \Phi_1 \cup \cdots \cup \Phi_D$ de bases, où chaque Φ_d est une base de $\mathcal{F}$. Ainsi Φ est une concaténation de bases[2] et on s'intéresse aux façon d'écrire un signal $f \in \mathcal{F}$ dans Φ, c'est à dire aux façon d'écrire

$$f = \sum_{\gamma} c_\gamma \phi_\gamma \tag{3.1}$$

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 $x = (x_d)_{d=1,\cdots,D}$ On s'intéresse ainsi aux solutions de

$$f = F_\Phi x. \tag{3.2}$$

[2]Ainsi Φ définit un frame serré.

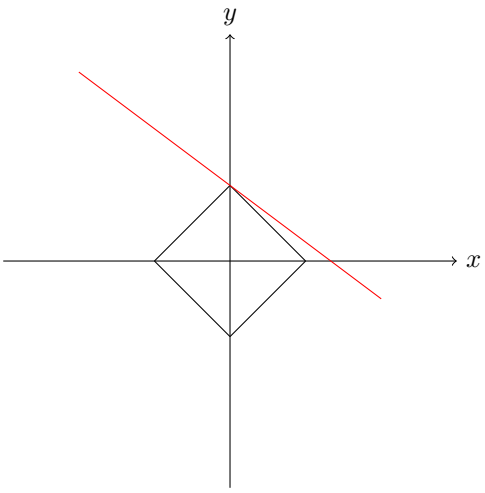


Figure 3.1: Minimisation de $y = Fx$ pour la norme ℓ^1. La solution de P1 est généralement celle de P0 sauf si les solutions sont parallèles à l'une des faces de la boule de ℓ^1.

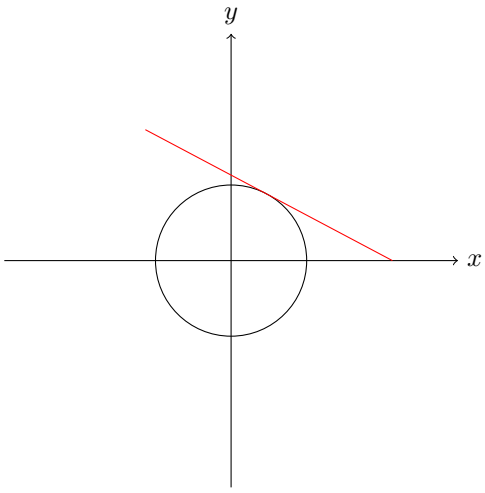


Figure 3.2: Minimisation de $y = Fx$ pour la norme ℓ^2. La solution de **??** n'est généralement pas celle de P0 sauf si les solutions sont parallèles à l'un des axes.

Comme discuté précedemment, 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

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

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 croit bien trop vite pour permettre la résolution de P0 de cette façon, 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 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 3.2.1. On étudie les signaux dans $\mathbb{R}^N$ et on choisi un dictionnaire constitué de la concaténation de la base de Fourier $W = \{e_k(t) = \frac{1}{\sqrt{N}} e^{\frac{i2\pi kt}{N}}\}_{0 \leqslant k \leqslant N-1}$ et de la base canonique de Diracs[3] $T = \{\delta_k\}_{0 \leqslant k \leqslant N-1}$. Ainsi, avec ce choix F_W est la matrice de Fourier discrète normalisée 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.

Theorème 3.2.1.1. *Soit un signal $f \in \mathbb{R}^{\mathbb{N}}$ non nul, alors*

$$||F_W f||_0 ||F_T f||_0 \geqslant N \tag{3.3}$$

et ainsi

$$||F_W f||_0 + ||F_T f||_0 \geqslant 2\sqrt{N}. \tag{3.4}$$

Preuve. La preuve qui est faite ici est standard, on peut par exemple la trouver dans l'article [34], on précise cependant la matrice F_T qui permettra d'obtenir des généralisations directes par la suite. Soit $0 \leqslant \omega \leqslant N-1$ un entier, alors

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

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

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

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

$$\leqslant \frac{1}{\sqrt{N}} ||F_T||_0^{\frac{1}{2}} \langle |f|, |f| \rangle^{\frac{1}{2}} = \frac{1}{\sqrt{N}} ||F_T||_0^{\frac{1}{2}} \langle |F_W f|, |F_W f| \rangle^{\frac{1}{2}} \tag{3.8}$$

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

$$\leqslant \frac{1}{\sqrt{N}} ||F_T||_0^{\frac{1}{2}} ||F_W f||_0^{\frac{1}{2}} \sup_\omega |F_W(\omega)|. \tag{3.10}$$

[3]Chaque 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

$$\sqrt{||F_T f||_0 \; ||F_W f||_0} \leqslant \frac{||F_T f||_0 + ||F_W f||_0}{2} \tag{3.11}$$

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

Observons que sans restrictions sur N, l'inégalité obtenue ne peut pas être améliorée, comme observé dans [17] [15], 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 $||F_T f||_0 + ||F_W||_0 < K$ 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).

Theorème 3.2.1.2. *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*

$$||x_W||_0 + ||x_T||_0 < \sqrt{N} \tag{3.12}$$

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,

$$F_\Phi(x_1 - x_2) = 0. \tag{3.13}$$

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

$$F_T \delta_T + F_W \delta_W = 0 \tag{3.14}$$

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

$$\delta_W = -F_W^t F_T \delta_T. \tag{3.15}$$

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 3.2.1.1, on a que δ a au moins $2\sqrt{N}$ coefficients non nuls si δ est non nul. En revenant a 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 premier[4], Terence Tao [34] a montré que l'on obtient l'inégalité

$$||F_T f||_0 + ||F_W f||_0 \geqslant p + 1 \tag{3.16}$$

[4]L'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 atteinte[5]. Grâce à ce principe d'incertitude plus fort que 3.2.1.1, on obtient avec le même type de preuve[6] le résultat suivant

Theorème 3.2.1.3. *Soit p 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*

$$||x_T||_0 \leqslant \frac{p}{2} \tag{3.17}$$

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.*

3.3 Résolution de (P1)

On a ainsi vu dans la section précedente 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 3.45 ou 3.17. 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,

$$min_x ||x||_1 \text{tel que } f = Fx \tag{P1}$$

permet sous certaines conditions de résoudre P0. L'intéret de P1 est que c'est un problème de programmation linéaire et de nombreuses méthodes permettent de le résoudre. Dans l'annexe A.1.3 on pourra trouver la définition de l'algorithme Matching Pursuit ainsi que quelques références vers d'autres algorithmes similaires. Précisons donc ce que nous avons affirmé, dans le même cadre que précedemment, 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 $\mathbb{R}^N$.

[5]L'inégalité est atteinte en ce sens qu'il existe $A \subset T$ et $B \subset W$ tels que $|A| + |B| = p + 1$ et il existe une fonction f telle que $SuppF_T f = A$ et $SuppF_W f = B$

[6]Voir [3] pour les détails, un lemme sur l'injectivité d'un opérateur similaire F_W est tout de même nécessaire pour conclure la preuve.

Theorème 3.3.0.1. *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*

$$||x_T||_0 < \frac{\sqrt{N}}{2} \quad et \quad ||x_W||_0 < \frac{\sqrt{N}}{2} \tag{3.18}$$

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

Remarque 3.3.1. Le théorème précédent a été obtenu en cherchant une preuve alternative à la preuve qui est faite par David Donoho et Xiaoming Huo [15], leur preuve utilise le même schéma que celle qui est faite ici. Leur théorème est le suivant :

Theorème 3.3.1.1. *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*

$$||x_T||_0 + ||x_W||_0 < \frac{\sqrt{N}}{2} \tag{3.19}$$

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, 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 4.49, alors x vérifie 3.45 et donc d'après le théorème 3.2.1.2 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 3.3.2. 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

$$\mu(\Gamma_T, \Gamma_W) = \sup_{\delta \in \mathcal{N}} \frac{\sum_{t \in \Gamma_T} |\delta_{T,t}| + \sum_{\omega \in \Gamma_W} |\delta_{W,\omega}|}{||\delta_T||_1 + ||\delta_W||_1} \tag{3.20}$$

Lemme 3.3.3. *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*

$$\mu(\Gamma_T, \Gamma_W) < \frac{1}{2} \tag{3.21}$$

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 3.21 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 3.3.4. [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*

$$|\Gamma_T| < \frac{\sqrt{N}}{2} \quad et \quad |\Gamma_W| \ < \frac{\sqrt{N}}{2}, \tag{3.22}$$

alors on a,

$$\mu(\Gamma_T, \Gamma_W) < \frac{1}{2}. \tag{3.23}$$

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

Commençons par la preuve du lemme 3.3.3.

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 3.21 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

$$||x + \delta||_1 - ||x|| > 0. \tag{3.24}$$

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

$$||x + \delta||_1 - ||x||_1 = \sum_{\gamma \in \Gamma^c} |\delta_\gamma| + \sum_{\gamma \in \Gamma} |c_\gamma + \delta_\gamma| - |c_\gamma|. \tag{3.25}$$

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

$$|c_\gamma + \delta_\gamma| - |c_\gamma| \ \geqslant -|\delta_\gamma| \tag{3.26}$$

[7] C'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.

et en insérant cette inégalité dans la somme on obtient

$$||x+\delta||_1 - ||x||_1 \geqslant \sum_{\gamma\in\Gamma^c} |\delta_\gamma| - \sum_{\gamma\in\Gamma} |\delta_\gamma|, \tag{3.27}$$

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

$$\sum_{\gamma\in\Gamma} |\delta_\gamma| < \sum_{\gamma\in\Gamma^c} |\delta_\gamma|. \tag{3.28}$$

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

$$\sum_{\gamma\in\Gamma} |\delta_\gamma| < \frac{1}{2}\left(\sum_{t\in T} |\delta_{T,t}| + \sum_{\omega\in W} |\delta_{W,\omega}|\right) = \frac{||\delta_T||_1 + ||\delta_W||_1}{2}. \tag{3.29}$$

Donc l'inégalité précédente est aussi une condition suffisante pour que 3.24 soit vérifiée et on peut réécrire cette inégalité sous la forme

$$\frac{\sum_{t\in\Gamma_T} |\delta_{T,t}| + \sum_{\omega\in\Gamma_W} |\delta_{W,\omega}|}{||\delta_T||_1 + ||\delta_W||_1} < \frac{1}{2}. \tag{3.30}$$

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 3.33 du lemme

$$\mu(\Gamma_T, \Gamma_W) := \sup_{\delta\in\mathcal{N}} \frac{\sum_{t\in\Gamma_T} |\delta_{T,t}| + \sum_{\omega\in\Gamma_W} |\delta_{W,\omega}|}{||\delta_T||_1 + ||\delta_W||_1} < \frac{1}{2}. \tag{3.31}$$

Le lemme 3.3.3 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 3.33 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 3.15 que les éléments δ de $\mathcal{N}$ sont de la forme $(\delta_T, -F_W^t F_T \delta_T) =: (\delta_T, -\widehat{\delta_T})$, donc 3.30 devient

$$\frac{\sum_{t\in\Gamma_T} |\delta_{T,t}| + \sum_{\omega\in\Gamma_W} |\widehat{\delta_{T\omega}}|}{||\delta_T||_1 + ||\widehat{\delta_T}||_1} < \frac{1}{2}. \tag{3.32}$$

□

On peut maintenant passer à la preuve du lemme 3.3.4

Preuve.

$$\mu(\Gamma_T, \Gamma_W) \leqslant \frac{\sum_{t\in\Gamma_T} |\delta_{T,t}| + \sum_{\omega\in\Gamma_W} |\widehat{\delta_{T\omega}}|}{||\delta_T||_1 + ||\delta_W||_1}. \tag{3.33}$$

Maintenant majorons le numérateur avec

$$\sum_{\omega\in\Gamma_W} |\widehat{\delta_{T\omega}}| = ||R_{\Gamma_W} F_W^t F_T \delta_T||_1 \leqslant ||R_{\Gamma_W} F_W^t F_T||_1 ||\delta_T||_1 \tag{3.34}$$

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 a[8] :

$$||R_{\Gamma_W} F_W^t F_T||_1 = \frac{|\Gamma_W|}{\sqrt{N}}. \tag{3.35}$$

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

$$\sum_{t\in\Gamma_T} |\delta_{T,t}| = ||R_{\Gamma_T} F_T^t F_W \delta_W||_1 \leqslant ||R_{\Gamma_T} F_T^t F_W||_1 ||\delta_T||_1 \tag{3.36}$$

ainsi que

$$||R_{\Gamma_T} F_T^t F_W||_1 = \frac{|\Gamma_T|}{\sqrt{N}}. \tag{3.37}$$

On peut maintenant rassembler les résultats:

$$\sum_{t\in\Gamma_T} |\delta_{T,t}| + \sum_{\omega\in\Gamma_W} |\widehat{\delta_{T\omega}}| \leqslant ||\delta_W||_1 \frac{|\Gamma_T|}{\sqrt{N}} + ||\delta_T||_1 \frac{|\Gamma_W|}{\sqrt{N}}. \tag{3.38}$$

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

$$\sum_{t\in\Gamma_T} |\delta_{T,t}| + \sum_{\omega\in\Gamma_W} |\widehat{\delta_{T\omega}}| < \frac{||\delta_T||_1 + ||\delta_W||_1}{2}. \tag{3.39}$$

Il nous reste maintenant à appliquer la majoration que l'on vient de trouver à 3.33 et on obtient

$$\mu(\Gamma_T, \Gamma_W) < \frac{1}{2} \frac{||\delta_T||_1 + ||\delta_W||_1}{||\delta_T||_1 + ||\delta_W||_1} = \frac{1}{2} \tag{3.40}$$

Ce qui conclut la preuve du lemme 3.3.4 et donc du théorème 3.3.0.1. □

3.4 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

[8] C'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 $1/\sqrt{N}$

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) $(\psi_1, \cdots, \psi_N)$ et les colonnes de Φ sont $(\varphi_1, \cdots, \varphi_N)$ on peut facilement exprimer les matrices précédentes avec:

$$F_\Psi^t F_\Phi = \begin{bmatrix} \langle \psi_1, \varphi_1 \rangle & \langle \psi_1, \varphi_1 \rangle & \cdots & \langle \psi_1, \varphi_N \rangle \\ \langle \psi_2, \varphi_1 \rangle & \ddots & \vdots & \langle \psi_2, \varphi_N \rangle \\ \vdots & \cdots & \ddots & \vdots \\ \langle \psi_N, \varphi_1 \rangle & \cdots & \cdots & \langle \psi_N, \varphi_N \rangle. \end{bmatrix} \tag{3.41}$$

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

$$M = M_{\Phi,\Psi} = M_{\Psi,\Phi} = \sup_{1 \leqslant i,j \leqslant N} |\langle \psi_i, \varphi_j \rangle|. \tag{3.42}$$

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 théorèmeset 3.2.1.1, 3.2.1.2 pour le problème P0 et 3.3.0.1, 3.3.4 pour le problème P1, ce n'est pas une coincidence et on va généraliser ces résultats ci-dessous.

Theorème 3.4.0.1. *Soit un signal $f \in \mathbb{R}^{\mathbb{N}}$ non nul et soit (Φ, Ψ) une paire de bases orthonormales et M la quantité définie par 3.42, alors*

$$||F_\Phi f||_0 ||F_\Psi f||_0 \geqslant \frac{1}{M^2} \tag{3.43}$$

et ainsi

$$||F_\Phi f||_0 + ||F_\Psi f||_0 \geqslant \frac{2}{M}. \tag{3.44}$$

Theorème 3.4.0.2. *Soit N un entier positif, (Φ, Ψ) une paire de bases orthonormales, M la quantité définie par 3.42 et un signal $f \in \mathbb{R}^N$, alors n'importe quel x vérifiant $f = F_\Phi x_\Phi + F_\Psi x_\Psi$ et*

$$||x_\Phi||_0 + ||x_\Psi||_0 < \frac{1}{M} \tag{3.45}$$

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 3.4.1. *Soit (Φ, Ψ) une paire de bases orthonormales et M la quantité définie par 3.42 et soient Γ_T et Γ_W des sous ensembles d'indices de T et respectivement de W vérifiant*

$$|\Gamma_T| < \frac{1}{2M} \quad et \quad |\Gamma_W| < \frac{1}{2M}, \tag{3.46}$$

alors on a,

$$\mu(\Gamma_T, \Gamma_W) < \frac{1}{2}. \tag{3.47}$$

Là aussi la démonstration est immédiate en faisant les changements adaptés dans la preuve. On a donc de la même façon que précedemment, en appliquant le lemme 3.3.3:

Theorème 3.4.1.1. *Soit N un entier positif, (Φ, Ψ) une paire de bases orthonormales, M la quantité définie par 3.42 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*

$$||x_T||_0 < \frac{1}{2M} \quad et \quad ||x_W||_0 < \frac{1}{2M} \tag{3.48}$$

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

3.5 Extensions du résultat

Comme indiqué précedemment, le théorème 3.3.0.1 obtenu est une généralisation d'un théorème de David Donoho et Xiaoming Huo dans [15], ainsi, la généralisation du théorème 3.3.0.1 qu'est le théorème 4.4.0.3, 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 4.4.0.3 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 [15] a été faite par Michael Elad et Alfred Bruckstein [19]. Ceux-ci ont en effet démontré le théorème suivant:

Theorème 3.5.0.1. *Soit N un entier positif, (Φ, Ψ) une paire de bases orthonormales, M la quantité définie par 3.42 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*

$$||x_T||_0 + ||x_W||_0 < \frac{\sqrt{2} - 0.5}{M} = \frac{0.9142}{M} \tag{3.49}$$

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 3.3.0.1 présentée ici et dans [15], 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 [19] 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 3.3.0.1 et celle de David Donoho et Xiaoming Huo [15].
On a ainsi que le théorème de Michael Elad et Alfred Bruckstein et le théorème 4.4.0.3 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 [20] 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 3.5.0.1, la construction d'un tel signal est rélativement 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 3.3.0.1 et 4.4.0.3 prouvés ici.

Chapter 4

Compressed sensing et approche aléatoire

4.1 Introduction au Compressed Sensing

Le chapitre précédent correspond à des résultats publiés entre 1998 et 2003, grâce à eux une classe de problème en apparence impossibles à résoudre (le problème P0) devenaient finalement accessibles sous des hypothèses de parcimonie. Cependant, les résultats prouvés donnent des informations quantitatives sur les plus petits signaux qui feront que la résolution de P0 en résolvant P1 ne fonctionne pas. Mais en fait, il y a très peu de tels contre-exemples et dans la pratique il était déjà observé que la résolution de P0 par P1 fonctionnait pour des signaux moins parcimonieux que ceux étudiés. C'est ainsi que plutôt que de chercher des résultats toujours vrais comme dans le chapitre précédent la recherche de nouveaux théorèmes s'est tournée vers des questions, demandant une approche probabiliste, du type : Quel s peut on choisir pour reconstruire *presque* tous les signaux s-parcimonieux en résolvant P1 ?
La deuxième question qui s'est posée, plus subtile, est la suivante, supposons que l'on sache que le signal est parcimonieux dans une base quelconque : Quel est le nombre minimal m de mesures que l'on peut faire pour être sûr[1] de reconstruire tous les signaux s-parcimonieux de cette base quelconque ?
A cette deuxième question, la réponse à apporter est plus claire qu'à la première question. On cherche à reconstruire un signal à N coefficients qui n'a que s coefficients non nuls, cependant on veut pouvoir le reconstruire avec $m < N$ mesure, ce signal est arbitraire donc on ne sait pas où sont les s-coefficients non nuls, il faut donc que la mesure se fasse sur un grand nombre de coefficients à la fois. Donc, dans la base dans laquelle on mesure le signal, il ne doit pas être parcimonieux. Il nous faut donc une base qui vérifie un principe d'incertitude avec la base dans laquelle le signal est parcimonieux. C'est le contenu de la condition **Uniform Uncertainty Principle** (**UUP**).
A partir de là il semble qu'il y ait une possibiité pour reconstruire le signal car n'importe

[1] On aurait bien sûr pu mettre *presque sûr* au lieu de *sûr*

quel coefficient peut être mesuré avec un nombre assez faible de mesures. Cependant une difficulté apparait directement, chaque mesure va mesurer plusieurs coefficients à la fois et il nous faut donc suffisament de mesures pour être certain de distinguer chaque coefficient non nul. Il nous faudra donc des garanties sur la famille utilisée pour la mesure pour qu'elle puisse nous permettre de distinguer les coefficients non nuls des coefficients nuls des signaux s-parcimonieux. C'est le contenu de la condition **Exact reconstruction principle** (**ERP**).
La deuxième difficulté est que l'on ne connait pas à l'avance la base dans laquelle le signal est parcimonieux, on sait aussi que l'on fera un nombre limité m de mesures, donc le principe d'incertitude devra être vérifié entre la base des mesures et n'importe quelle restriction à m coordonées de la base inconnue. Cela revient donc à la projection dans un espace à m dimensions arbitraires. C'est ainsi qu'ont été publiés par Emmanuel Candes, Justin Romberg et Terence Tao [7], [6], [3], et indépendamment par David Donoho [18], les articles fondateurs du Compressed Sensing, dont les théorèmes montrent qu'un tel raisonnement peut-être démontré.
On suivra dans un premier temps l'article de Emmanuel Candes et Terence Tao pour prouver un théorème de compressed sensing, ensuite, on verra sans preuve des théorèmes de Emmanuel Candes et Justin Romberg dont l'énoncé sera plus simple à comprendre et qui mettra en valeur l'incohérence, quantité clé du chapitre précédent. Dans ce chapitre, on verra **UUP** et **ERP** comme des axiomes, on ne démontrera pas que de telles familles existent, de la même façon que dans le chapitre 2 nous n'avons pas démontré que les ondelettes de Daubechies existent. On verra cependant des exemples de familles qui vérifient ces conditions et des preuves de ces résultats, avec des bornes plus ou moins optimales peuvent être trouvées dns les articles précedemment cités ou bien dans [21].

4.2 Axiomatisation, UUP et RIP

4.2.1 Notations

Tout d'abord réintroduisons certaines choses déjà vues dans ce mémoire et les notations qui seront utilisées. Dans ce chapitre on considère $\mathcal{F} \subset \mathbb{R}^N$ un classe de signaux. On cherche à pouvoir reconstruire chaque élément $f \in \mathcal{F}$ avec une précision ε en utilisant une famille de vecteurs $\Psi = (\psi_k)_{k\in\Omega}$.
C'est à dire, on considère une application d'analyse,

$$A : \mathcal{F} \longrightarrow \mathbb{R}^{|\Omega|} \tag{4.1}$$

$$(f_k)_{k=0,\cdots,N} = f \longmapsto (\langle f, \psi_k \rangle)_{k\in\Omega} \tag{4.2}$$

qui a chaque signal associe la projection sur chaque élément de Ψ. Pour l'instant cela est similaire à la situation dans l'étude des frames, la différence essentielle concerne Ω, dans la situation des frames Ω est fixé et généralement $|\Omega| \geqslant N$. Comme discuté, on s'intéresse ici à la projection dans un sous espace arbitraire, c'est ainsi le rôle que va jouer Ω en étant une variable aléatoire, et comme on souhaite faire un nombre minimal

de mesures, on va chercher à avoir que le nombre de mesures moyen $K = \mathbb{E}(|\Omega|)$ sera inférieur à N. La difficulté vient alors dans la construction de l'application de synthèse associée, pour l'instant définissons la pour fixer les notations :

$$S : \mathbb{R}^{|\Omega|} \longrightarrow \mathbb{R}^N \tag{4.3}$$

$$(y_k)_\Omega \longmapsto (y_k)^\# = (f_k^\#)_{k=0,\cdots,N} =: f^\# \tag{4.4}$$

et on cherche à obtenir une ε-reconstruction :

$$||f - f^\#||_2 \leqslant \varepsilon \quad , \forall f \in \mathcal{F}. \tag{4.5}$$

Le problème est donc de choisir une famille $(\psi_k)_{k\in\Omega}$ pour qu'il soit possible d'obtenir la dernière inégalité. Il est clair que le problème est mal posé si on prend $\mathcal{F} = \mathbb{R}^N$, on verra plus bas sur quelles familles de signaux on pourra démontrer des résultats.
Notons F_Ω la matrice, aléatoire, avec $|\Omega|$ lignes et N colonnes qui représente A. Pour aider à se fixer les idées sur le type de matrice que l'on considérera, F_Ω peut être la restriction de $|\Omega|$ lignes d'une matrice de Fourier, ou bien une matrice dont les coefficientss suivent une loi normale de moyenne nulle et de variance égale à 1.
Notons maintenant R_Ω la matrice de restriction aux indices de Ω.

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

$$(g_k)_{0\leqslant k\leqslant N} \longmapsto (g_k)_{k\in\Omega} \tag{4.7}$$

et l'inclusion prolongée par des zéros

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

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

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 :

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

$$(g_k)_T \longmapsto F_\Omega((g_k)_T \oplus (0)_T^c). \tag{4.11}$$

On remarque aussi que $F_{\Omega T}^* F_{\Omega T} : \ell^2(T) \to \ell^2(T)$ est symétrique et que l'on peut la diagonaliser sous la forme $U\Lambda U^*$ où $\Lambda = (\lambda_1 \geqslant \cdots \geqslant \lambda_{|T|})$ sont les valeurs propres de $F_{\Omega T}^* F_{\Omega T}$ qui sont aussi appelées valeurs singulières de $F_{\Omega T}$.
Dans ce chapitre on considérera trois modèles de matrices aléatoires, on a déjà vu le modèle aléatoire de Fourier en échantillonant les lignes indéxées par Ω de la matrice de Fourier. On verra aussi la matrice gaussienne dont les coefficients suivent une loi normale centrée réduite normalisée et le modèle de Bernoulli où les coefficients sont égaux à $+1$ ou -1.
Deux modèles d'échantillonage sont utilisés dans la suite, si on fixe $|\Omega|$ et que Ω peut être n'importe quel ensemble de taille Ω on peut parler de modèle uniforme. Si Ω est obtenu en faisant un échantillonage qui suit une loi de Bernoulli, on peut parler de modèle de Bernoulli.

4.2.2 Définition de UUP

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

Définition 4.2.3. 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

$$|supp(f)| \leqslant \alpha K/\lambda \tag{4.12}$$

on ait l'inégalité

$$\frac{1}{2}\frac{K}{N}||f||_2^2 \leqslant ||F_\Omega f||_2^2 \leqslant \frac{3}{2}\frac{K}{N}||f||_2^2. \tag{4.13}$$

Quelques mots s'imposent sur cette définition. Tout d'abord, comme mentionné en début de chapitre, on ne cherche pas à avoir un résultat toujours vrai, on veut seulement qu'il soit presque toujours vrai, d'où le fait que le résultat soit vrai avec une probabilité $1 - \mathcal{O}(N^{-\rho/\alpha})$. Remarquons aussi que le résultat est exprimé en terme de $\mathcal{O}$ et on en déduit donc que le résultat peut devenir vrai avec probabilité égale à 1 si on peut choisir N arbitrairement grand.
Ensuite, on voit que les signaux sur lesquels le résultat est vrai sont ceux dont le support est inférieur $\alpha K/\lambda$, pour couvrir un maximum de signaux, on cherchera donc à avoir F_Ω qui vérifie λ-**UUP** avec λ aussi petit que possible.
Concernant 4.13 avec l'étude des frames et de leurs liens avec les bases orthonormales, il devrait être clair que 4.13 est entre une condition de frame et de base orthonormale.

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

Proposition 4.2.4. *F_Ω vérifie λ-**UUP** si et seulement si*
avec probabilité au moins $1 - \mathcal{O}(N^{-\rho/\alpha})$ pour un certain $\rho > 0$ on a $\forall T \subset [0, N]$ qui vérifie $|T| \leqslant \alpha\frac{K}{\lambda}$ alors les valeurs propres de $F_{\Omega T}$ vérifient

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

où λ_{max} (resp. λ_{min}) est la valeur propre maximale (resp. minimale) de $F_{\Omega T}F_{\Omega T}^*$.

Preuve 4.2.5. Pour montrer que la définition implique la proposition, on prend un vecteur propre f de $F_{\Omega T}F_{\Omega T}^*$ de valeur propre maximale (resp. minimale), d'où:

$$||F_{\Omega T}f||_2^2 = f^t F_{\Omega T}^* F_{\Omega T} f = \lambda_{max}||f||_2^2 \tag{4.14}$$

et le dernier terme est plus petit (resp. plus grand pour λ_{min}) que $\frac{3K}{2N}||f||_2^2$ (resp. $\frac{K}{2N}||f||_2^2$)par hypothèse.
Dans l'autre sens, on prend un vecteur f tel que T est le support de f, alors $F_{\Omega T}f = F_\Omega f$ et donc:

$$\lambda_{min}\ ||f||_2^2 \leqslant ||F_\Omega f||_2^2 \leqslant \lambda_{max}\ ||f||_2^2 \tag{4.15}$$

et il suffit de remplacer λ_{min} et λ_{max} par leurs valeurs dans la proposition.

Remarque 4.2.6. 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 ($|supp(f)| \leqslant \alpha\frac{K}{\lambda}$), alors on a

$$||F_\Omega f||_{\ell^2} = ||\hat{f}||_{\ell^2(\Omega)} \leqslant \sqrt{\frac{3K}{2N}}||f||_{\ell^2} \tag{4.16}$$

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

$$\frac{||\hat{f}||_{\ell^2(\Omega)}}{||\hat{f}||_{\ell^2}} \longrightarrow 0 \tag{4.17}$$

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'appélation "principe d'incertitude".

Remarque 4.2.7. 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'appélation uniforme.

Remarque 4.2.8. Remarquons que l'on peut réécrire 4.13 peut se réécrire

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

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).

Proposition 4.2.9. [2]

- *Les ensembles Gaussiens et binaires vérifient* $\log(N) - \boldsymbol{UUP}$
- *L'ensemble de Fourier vérifie* $\log(N) - \boldsymbol{UUP}$.

On peut trouver des démonstration dans [6] et [21].

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

4.2.10 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 4.2.11. F_Ω vérifie λ-**ERP** si

- $\forall T \subset [0, N]$ vérifiant $|T| \leqslant \alpha \frac{K}{\lambda}$
- $\forall \sigma \in \{\pm 1\}^T$

il existe avec probabilité au moins $1 - \mathcal{O}(N^{-\rho/\alpha})$ pour un certain $\rho > 0$, 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. $P(t) < \frac{1}{2}, \forall t \in T^c$[4]

Comme discuté dans l'introduction de ce chapitre, cela revient à pouvoir reconstruire (et surtout distinguer) n'importe quelle suite de signes supportée sur T

Proposition 4.2.12.
- *Les ensembles Gaussiens et binaires vérifient* $\log N$-***ERP***
- *L'ensemble de Fourier vérifie* $\log N$-***ERP***.

4.3 Théorème de Candes-Tao

Avec les notations et les définitions ci-dessus on peut maintenant énoncer puis prouver le théorème de Emmanuel Candes et Terence Tao.

Theorème 4.3.0.1. *Soit F_Ω qui vérifie λ_1-**ERP** et λ_2-**UUP**. On pose $\lambda = \max(\lambda_1, \lambda_2)$, soit $K \geqslant \lambda$.*
Soit f un signal dans $\mathbb{R}^N$ tel que ses coefficients dans une base de référence décroissent comme[5] *:*

$$|\theta_{(n)}| \leqslant C n^{-\frac{1}{p}} \tag{4.18}$$

pour un certain $C > 0$ et $0 < p \leqslant 1$.
On pose $r = \frac{1}{p} - \frac{1}{2}$, alors n'importe quel minimiseur de (P1) vérifie :

$$||f - f^{\#}||_2 \leqslant C_r (\frac{K}{\lambda})^{-r} \tag{4.19}$$

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

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

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

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

Tout d'abord on peut préciser la famille de signaux que l'on considère, ce sont ceux dont les coefficients décroissent comme une loi en puissance dans une certaine base. C'est une telle décroissance que l'on a par exemple identifié pour les coefficients d'ondelettes d'une fonction lipschitzienne par rapport à l'échelle dans le théorème de Stéphane Jaffard 2.2.7.1. On voit donc que le théorème peut s'appliquer de façon générale. De plus, contrairement aux résultats du chapitre 3, on ne demande pas à ce que le signal soit exactement parcimonieux pour avoir une bonne reconstruction ℓ^2. L'avantage de ce théorème est donc qu'en résolvant P1, même si une solution vraiment parcimonieuse n'existe pas, on a quand même une reconstruction du signal pour la norme euclidienne.
Passons à la preuve du théorème, on discutera ensuite de certaines conséquences ainsi que de résultats analogues.

Preuve. Soit F_Ω comme dans le théorème. Soit $f \in \mathbb{R}^N$ dont les coefficients vérifient 4.18, on note alors $f^\#$ une solution $y = F_\Omega f$. L'objectif est de déterminer

$$||f - f^\#||_2. \tag{4.20}$$

On note T l'ensemble des T plus grandes valeurs de $|f| + |f^\#|$. Comme vu dans les preuves du chapitre 3, on a:

$$F_\Omega(f - f^\#) = 0. \tag{4.21}$$

Cependant, les hypothèses que l'on peut utiliser (**ERP** et **UUP**) ne peuvent être utilisées que dans sur des ensemble de taille T où $T \leqslant \alpha K/\lambda$. Donc, plutôt que de considérer $f - f^\#$, on va considérer $h = (f - f^\#)1_T$, où on note 1_T la restriction aux indices de T. Il est alors possible de démontrer en appliquant **UUP** (voir le lemme A.2.4) que l'on peut trouver $g \in \mathbb{R}^N$ qui s'écrit $g = F_\Omega^* V$ pour un certain $V \in \mathbb{R}^\Omega$ et qui vérifie à la fois de prendre les mêmes valeurs que h sur T et

$$\sum_{t\in E} |g(t)|^2 \leqslant C \sum_{t\in T} |f(t) - f^\#(t)|^2 \tag{4.22}$$

pour n'importe que ensemble E disjoint de T et de taille $|E| = \mathcal{O}(K/\lambda)$. Avec des mots, l'image de g par F_Ω vaut 0 sur T, et hors de T les valeurs de g sont majorées par la distance euclidienne de f et $f^\#$. On peut alors utiliser le fait que g s'écrive $g = F_\Omega^* V$:

$$\langle f - f^\#, g\rangle = \langle f - f^\#, F_\Omega^* V\rangle = \langle F_\Omega(f - f^\#), V\langle = 0 \tag{4.23}$$

d'après 4.21. Ainsi, on a

$$\sum_{t=0,\cdots,N} (f - f^\#)(t)g(t) = 0 \tag{4.24}$$

que l'on peut réécrire

$$\sum_{t\in T}(f - f^\#)(t)g(t) = ||f - f^\#||^2_{\ell^2(T)} = -\sum_{t\in T^c}(f - f^\#)(t)g(t). \tag{4.25}$$

Par ailleurs, on a toujours:

$$||f - f^{\#}||_{\ell^1(T^c)} \leqslant ||f||_{\ell^1(T^c)} + ||f||_{\ell^1(T^c)} \tag{4.26}$$

D'après un lemme qui utilise **ERP** que l'on peut trouver en annexe A.2.1

$$||f + f^{\#}||_{\ell^1(T^c)} \leqslant C|T|^{1-\frac{1}{p}}. \tag{4.27}$$

et d'après un autre lemme A.2.2 démontré en annexe on a

$$||f - f^{\#}||_{\ell^\infty(T^c)} \leqslant C|T|^{-1/p} \tag{4.28}$$

Avec ces deux inégalités on peut appliquer l'inégalité de Hölder pour obtenir :

$$||f - f^{\#}||_{\ell^2(T^c)} = \sqrt{|||f - f^{\#}|^2||_{\ell^1(T^c)}} \leqslant \sqrt{||f + f^{\#}||_{\ell^1(T^c)}||f - f^{\#}||_{\ell^\infty(T^c)}} \leqslant C|T|^{\frac{1}{2}-\frac{1}{2p}}. \tag{4.29}$$

Pour prouver le théorème, il reste donc à montrer qu'une telle majoration est encore vraie sur T. Pour obtenir cela on va réordonner les indices, tout d'abord on énumère $T^c = (n_1, \cdots, n_{N-|T|})$ tels que les coefficients correspondants dans $|f - f^{\#}|$ soient classés par ordre décroissant. Maintenant on regroupe ces coefficients en blocs de taille $|T|$, à part peut-être pour le dernier qui peut être plus petit. On les notes $B_J = \{n_j, J|T| < j \leqslant (J+1)|T|\}$ ces blocs pour $J = 0, \cdots, \lfloor N/|T| \rfloor$. On a alors, d'après Cauchy-Schwarz, à J fixé:

$$\sum_{j \in B_J} (f - f^{\#})(n_j) g(n_j) \leqslant ||f - f^{\#}||_{\ell^2(B_J)} ||g||_{\ell^2(B_J)}. \tag{4.30}$$

Or d'après 4.22, $||g||_{\ell^2(B_J)} \leqslant C||f - f^{\#}||_{\ell^2(T)}$, donc l'inégalité précédente devient:

$$\sum_{j \in B_J} (f - f^{\#})(n_j) g(n_j) \leqslant C||f - f^{\#}||_{\ell^2(T)}||f - f^{\#}||_{\ell^2(B_J)} \leqslant C||f - f^{\#}||_{\ell^2(T)} I_J \tag{4.31}$$

en notant

$$I_J := ||f - f^{\#}||_{\ell^2(B_J)} = \sqrt{\sum_{j=J|T|+1}^{(J+1)|T|} |(f - f^{\#})(n_j)|^2}. \tag{4.32}$$

Comme les coefficients sont par ordre décroissants, on va pouvoir obtenir des inégalités successives entre les I_J. Tout d'abord, pour $J = 0$, on a clairement :

$$I_0 \leqslant \sqrt{|T|}||f - f^{\#}||_{\ell^\infty(B_J)} \leqslant \sqrt{|T|}|(f - f^{\#})(n_0)|. \tag{4.33}$$

D'après 4.28 on a donc:

$$I_0 \leqslant C|T|^{\frac{1}{2}-\frac{1}{p}} = C|T|^{-r}. \tag{4.34}$$

Pour $J \geqslant 1$ on a ainsi:

$$I_J = ||f - f^{\#}||_{\ell^2(B_J)} \leqslant |T|^{\frac{1}{2}}|f - f^{\#}|(n_{J|T|+1}) \leqslant |T|^{\frac{1}{2}}|T|^{-1}\ ||f - f^{\#}||_{\ell^1(B_{J-1})}. \tag{4.35}$$

Afin de conclure il nous faut donc évaluer la somme sur J des I_J,

$$\sum_{J\geqslant 0} I_J \leqslant I_0 + \sum_{J\geqslant 1} I_J \leqslant C|T|^{-r} + \frac{1}{|T|^{\frac{1}{2}}} \sum_{J\geqslant 0} I_J. \tag{4.36}$$

On déduit de la précédente inégalité:

$$\sum_{J\geqslant 0} I_J \leqslant C|T|^{-r} + |T|^{-\frac{1}{2}} \, ||f - f^{\#}||_{\ell^1(T^c)} \tag{4.37}$$

et d'après 4.27 on a ainsi:

$$\sum_{J\geqslant 0} I_J \leqslant C|T|^{-r} + C|T|^{\frac{1}{2}-\frac{1}{p}} = 2C|T|^{-r}. \tag{4.38}$$

Les B_J formant une partition de $\{0, \cdots, N\}\backslash T$ on a, d'après **??**, puis 4.36:

$$\sum_{t\in T^c} (f - f^{\#})(t)g(t) \leqslant \sum_J C||f - f^{\#}||_{\ell^2(T)}||f - f^{\#}||_{\ell^2(B_J)} \tag{4.39}$$

$$\leqslant C||f - f^{\#}||_{\ell^2(T)} \sum_J I_J \leqslant 2C||f - f^{\#}||_{\ell^2(T)}|T|^{-r} \tag{4.40}$$

Et donc finalement avec 4.25 on a:

$$||f - f^{\#}||^2_{\ell^2(T)} \leqslant 2C||f - f^{\#}||_{\ell^2(T)}|T|^{-r} \tag{4.41}$$

que l'on peut réécrire :

$$||f - f^{\#}||_{\ell^2(T)} \leqslant 2C|T|^{-r}. \tag{4.42}$$

On a donc bien démontré le théorème. □

Dans la preuve, deux résultats intermédiaires ont été admis. Le premier utilise *UUP* et est une conséquence d'un théorème d'extension, avec des mots, on a construit un vecteur dans l'espace engendré par F^*_Ω qui coincide avec un autre vecteur arbitraire sur un support fini et qui vérifie une forme de stabilité ℓ^2. On pourra trouver une démonstration de ce résultat en annexe A.2.4.
Le second résultat utilise *ERP* et permet de démontrer un autre résultat sur la concentration en dehors d'un support fixé, mais pour la norme ℓ^1. On pourra en trouver un résultat formel en annexe également A.2.1.
Cependant il aurait été possible d'utiliser d'autres conditions, la condition **UUP** n'étant pas très difficile à vérifier contrairement à la condition **ERP** qui doit être vérifiée pour n'importe quelle suite de signes. Il est ainsi possible d'affaiblir cette condition en **Weak Exact Reconstruction Principle** (**WERP**) dans laquelle plutôt que d'avoir **ERP** pour n'importe quelle suite de signe, on souhaite pour presque n'importe quelle suite de signes **ERP**. En fait, Terence Tao et Emmanuel Candes prouvent dans [6] que **WERP** et **UUP** impliquent **ERP** pour presque n'importe quel signal.

4.4 Théorème de Donoho

Etudions maintenant un autre théorème fondamental du Compressed Sensing, et en fait très similaire à celui de Emmanuel Candes et Terence Tao, avec le théorème de David Donoho. Ce théorème est présenté ici car il est plus facile à exprimer que le théorème précédent, il n'y a pas besoin d'autant de notations et définitions, aussi, on va voir que ce résultat est assez proche de l'esprit des résultats du chapitre 3. Cependant on ne fera pas la preuve de ce résultat pour plusieurs raisons, premièrement, après le théorème précédent et le chapitre 3 il devrait être clair qu'un tel résultat soit démontrables, deuxièmement, les techniques de preuves utilisées par David Donoho sont assez élaborées et reposent sur l'usage d'inégalités sur des lois de probabilités, finalement, il aurait été difficile d'introduire de façon raisonnable ces résultats sans ajouter trop de pages supplémentaires à ce mémoire (qui est peut-être déjà trop long). Cela étant dit, le résultat de David Donoho n'a besoin que de d'un rappel de définitions utilisée dans les chapitres précédents. Soit (Φ, Ψ) une paires de matrices et $(\psi_1, \cdots, \psi_N)$ les colonnes de Ψ et $(\varphi_1, \cdots, \varphi_N)$ les colonnes de Φ. On note la cohérence entre Φ et Ψ:

$$M_{\Phi,\Psi} = \sup_{i,j} |\langle \psi_i, \ \varphi_j \rangle| \tag{4.43}$$

On notera en particulier $M_\Phi = \sup_{i,j} \ |\varphi_{i,j}|$.

On peut alors énoncer le théorème de David Donoho:

Theorème 4.4.0.1. *Soit Φ une base orthonormale de $\mathbb{R}^N$. Soit $T \subset \{0, \cdots, N\}$ un sous-ensemble fixé et soit $z \in \ \{\pm 1\}^T$ une suite de signes tirée uniformément au hasard ($\mathbb{P}(z(t) = 1) = \mathbb{P}(z(t) = -1) = \frac{1}{2}$ pour tout $t \in T$). Si le nombre de mesures m vérifie:*

$$m \geqslant C_0 |T| N M_\Phi^2 \log(\frac{N}{\delta}) \tag{4.44}$$

et

$$m \geqslant C_1 \ \log^2(\frac{N}{\delta}) \tag{4.45}$$

pour des constantes fixées C_0 et C_1. Alors, avec probabilité au moins $1 - \delta$, on peut reconstruire n'importe quel signal x_0 supporté sur T et ayant la même suite de signes que z à partir de m mesures

$$y = F_{\Phi\Omega} x_0 \tag{4.46}$$

en résolvant P1.

On ne démontre pas ce théorème mais discutons de son lien avec le théorème de Emmanuel Candes et Terence Tao, dans celui ci on choisit d'abord une suite de signes, qui n'est pas arbitraire et qui a été obtenue par un processus aléatoire et on peut ensuite conclure sur la reconstruction, on n'est donc pas dans le cas de l'**Exact Reconstruction Principle** mais plutôt du **Weak Exact Reconstruction Principle**.
Ensuite, ici, le lien entre la cohérence de la matrice utilisée pour mesurer et le nombre

de mesures à réaliser est explicite, en fait ce nombre de mesures est choisi car il permet de vérifier (théorème 1.2 [7]), que les valeurs singulières de $F_{\Phi,\Omega T}$ sont toutes proches de $\frac{m}{N}$, c'est bien le contenu de l'**Uniform Uncertainty Principle**.
Aussi, il est possible de relier les λ_1 et λ_2 du théorème de Emmanuel Candes et Terence Tao aux deux inégalités du théorème de Donoho. On remarquera juste que dans le cas où F_Φ est la matrice de Fourier, alors $M_\Phi = \frac{1}{\sqrt{N}}$, on utilise donc la seconde inégalité si $|T|$ est petit. De même pour une matrice dont les coefficients suivent une loi normale centrée de variance égale à 1, alors[6] $M_\Phi \leqslant C\frac{1}{N}\sqrt{\log(N)}$, on est donc également dans le cas de la seconde inégalité du théorème, si $|T|$ est petit. Cependant, on peut revenir au cas de la première inégalité en supposant $C_0 > C_1 \log(\frac{N}{\delta})$.
On a donc le corollaire suivant :

Theorème 4.4.0.2. *Soit F_Ω la restriction de la matrice de Fourier ou d'une matrice gaussienne à un sous ensemble $\Omega \subset \{0, \cdots, N-1\}$ qui vérifie:*

$$|\Omega| \geqslant CS\log(N) \tag{4.47}$$

pour une certaine constante C. Soit f un signal s-parcimonieux, donc tel qu'il existe x_0 avec s-composantes vérifiant $f = F_\Omega x_0$ alors avec probabilité tendant vers 1 la solution x de P1 est $x = x_0$.

On peut aussi déduire du théorème de Donoho la version asymptotique des théorèmes 3.3.0.1, 4.4.0.3 et 3.5.0.1 en considérant que l'on fait toutes les mesures ($|\Omega| = N$):

Theorème 4.4.0.3. *Soit N un entier positif, (Φ, Ψ) une paire de bases orthonormales, M la quantité définie*

$$M = \sup_{i,j} |\langle \varphi_i, \psi_j \rangle| \tag{4.48}$$

et un signal $f \in \mathbb{R}^N$, alors avec probabilité tendant vers 1, n'importe quel $x = (x_\Phi, x_\Psi)$ vérifiant $f = F_\Phi x_\Phi + F_\Psi x_\Psi$ et

$$||x_\Phi||_0 + ||x_\Psi||_0 < C\frac{N}{M^2\log(N)} \tag{4.49}$$

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

[6]Pour montrer cela il faut montrer que l'espérance de la valeur maximale de n tirages de loi normale centrée réduite est majoré par $C\sqrt{\log(n)}$, la démonstration étant hors du sujet du mémoire n'est pas faite ici, on pourr consulter [21] pour de nombreux résultats et démonstrations sur les matrices aléatoires dans le cadre du compressed sensing.

Chapter 5

Conclusion

5.1 Conclusion

Dans ce mémoire de nombreuses méthodes pour reconstruire un signal ont été proposées. Tout d'abord on a vu que pour la norme ℓ^2 la théorie des frames était particulièrement adaptée et englobait notamment le cadre des bases orthonormales. On a aussi vu que les formules de reconstruction obtenues étaient robustes face au bruit. On a aussi vu qu'il était possible de construire des frames de nombreuses façon, notamment à partir de bases orthonormales. On a donc poursuivi en étudiant des bases orthonormales qui vérifient une décroissance rapide des coefficients avec la régularité du signal analysé. Ainsi, en faisant un choix de base adapté, il semble qu'une forme de parcimonie de la représentation du signal soit possible.
Cependant les représentations données par les frames ne favorisent pas la parcimonie exacte du résultat, de nombreux petits coefficients sont généralement présents. On a ainsi étudié la faisabilité d'avoir une représentation exactement parcimonieuse, on a vu qu'il était possible d'obtenir une telle représentation de façon unique sans résoudre de problème combinatoireP0 mais en résolvant un problème de programmation linéaire P1. On a même vu que la possibilité d'une telle reconstruction dans une paire de bases orthonormales reposait sur la mesure de la cohérence des deux bases qui s'exprime de façon très simple. Cependant, le domaine de résolution équivalente entre P0 et P1 était plus restreint que nécessaire pour une grande majorité de signaux.
C'est ainsi que le compressed sensing a été introduit permettant d'exploiter la parcimonie de façon double. Le dernier résultat du chapitre précédent nous a donné une condition suffisante pour avoir l'équivalence de la reconstruction entre P0 et P1 dans un dictionnaire pour une majorité de signaux. Les théorèmes de Emmanuel Candes, Justin Romberg, Terence Tao et David Donoho ont eux permis le tour de force de permettre la reconstruction d'un signal de façon exacte avec un système en apparence clairement sous-déterminé mais avec lesquels l'hypothèse de parcimonie permet une reconstruction exacte. Un tel tour de force a été possible en exploitant de façon fine les résultats et surtout les idées des chapitres 2 et 3.
Les derniers résultats permettant de dire que les signaux parcimonieux peuvent (presque

tous) être représenté dans un espace (aléatoire) de plus petite dimension, on peut conclure que l'on a bien fait le tour du triangle "régularité, parcimonie, approximation basse dimension".
On peut donc résumer la situation de la façon suivante (sous réserve de quelques abus de langage mais dont la version formelle a été le sujet du mémoire): Si le signal est régulier, on peut le représenter de façon parcimonieuse. Si le signal est parcimonieux, on peut le représenter en basse dimension. Si on a une représentation en basse dimension aléatoire, on peut retrouver le signal régulier original.

Appendix A

Annexe

A.1 Algorithmes

Présentons ici quelques algorithmes qui peuvent être utilisés pour mettre en oeuvre les résultats de ce mémoire. On peut mentionner l'article [9] pour une courte introduction à plusieurs technique en norme 2 et en norme 1, et on pourra consulter **??** pour des algorithmes plus récents et avancés qui tirent notamment parti du compressed sensing

A.1.1 Frames

Tout d'abord montrons que la reconstruction avec les coefficients de frame minimise la norme ℓ^2. Considérons qu'un dictionnaire $\Phi = (\varphi_i)_{i\in I}$ est un frame, on note F_Φ l'opérateur d'analyse avec pour lignes les colonnes de Φ et F_Φ^* l'opérateur de synthèse vérifiant $F_\Phi^* \circ F_\Phi = Id$ sur $Vect(\Phi)$. Alors on a la proposition suivante, d'après [14]

Proposition A.1.2. *Si $f = \sum_{i\in I} c_i\varphi_i$ pour une suite de coefficients $(c_i)_I \in \ell^2(I)$, alors la décomposition est minimale en norme ℓ^2 seulement si $c_i = \langle f, \varphi_i\rangle$ pour tout $i \in I$. De façon équivalente, si il existe un $i \in I$ tel que $c_i \neq \langle f, \varphi_i\rangle$, alors*

$$\sum_I |c_i|^2 > \sum_I |\langle f, \varphi_j\rangle|^2. \tag{A.1}$$

Preuve. On a donc $f = F_\Phi^* c$ pour un certain $c \in \ell^2(I)$. On décompose maintenant c entre sa partie a dans l'espace vectoriel engendré par F_Φ, et celle b dans l'orthogonal de cet espace, on a donc $c = a + b$ et par le théorème de Pythagore $||c||^2 = ||a||^2 + ||b||^2$. Donc il existe aussi $x \in \mathcal{H}$ (où $\mathcal{H}$ est l'espace de Hilbert sur lequel on suppose que le frame est défini) tel que $a = F_\Phi x$, donc $c = F_\Phi x + b$. Ainsi $f = F_\Phi^* c = F_\Phi^* F_\Phi x + F_\Phi^* b$, le dernier terme valant 0 par construction on a $x = f$, donc $c = F_\Phi f + b$. On obtient donc:

$$\sum_I |c_j|^2 = ||c||^2 = ||F_\Phi f||^2 + ||b||^2 = \sum_I |\langle f, \varphi_i\rangle|^2 + ||b||^2 \tag{A.2}$$

qui est strictement supérieur à $\sum_I |\langle f, \varphi_i\rangle|^2$, sauf si $b = 0$ et donc $c = F_\Phi f = (\langle f, \varphi_i\rangle)_I$. □

Voyons maintenant comment obtenir une formule de reconstruction lorsque le frame n'est pas serré. Dans le cas où le frame est serré on a vu que l'on a la formule de reconstruction

$$f = \frac{1}{M}\sum_I \langle f, \varphi_i\rangle \varphi_i = \frac{1}{M} Id \tag{A.3}$$

cependant si le frame n'est pas serré on a $m \neq M$ et donc la formule de reconstruction n'est plus valide. Il nous faudrait donc plutôt une formule de reconstruction du type:

$$f = \sum_I \langle f, \tilde{\varphi}_i\rangle \varphi_i \tag{A.4}$$

où les $\tilde{\varphi}_i$ dépendent du frame. On peut montrer que ces vecteurs correspondent au frame dual de F et qu'ils sont définis par

$$\tilde{\varphi}_i = (F_\Phi^* F_\Phi)^{-1}\varphi_i. \tag{A.5}$$

Maintenant on considère que les constantes m, M sont proches l'unes de l'autre, c'est à dire $r = M/m - 1 << 1$, alors la formule de reconstruction obtenue avec $\frac{1}{M}Id$ dans le cas d'un frame serré devrait, dans l'autre cas être que $F_\Phi^* F_\Phi$ est proche de $\frac{m+M}{2}Id$. Ainsi $(F_\Phi^* F_\Phi)^{-1}$ devrait être proche de $\frac{2}{m+M}Id$ et de même on devrait avoir que $\tilde{\varphi}_i$ doit être proche de $\frac{2}{m+M}Id$. On peut préciser cela avec

$$f = \frac{2}{m+M} Id \sum_I \langle f, \varphi_i\rangle \varphi_i + Rf \tag{A.6}$$

avec $R = Id - \frac{2}{m+M}F^*F$ et il suffit de remplacer R dans la précédente équation pour vérifier qu'elle est correcte. Or, en notant $||\cdot||$ la norme d'opérateur, on a par contruction $m \leqslant ||F_\Phi^* F_\Phi|| \leqslant M$ et donc:

$$-(Id - \frac{2}{m+M}MId) \leqslant R \leqslant Id - \frac{2}{m+M}mId \tag{A.7}$$

ce qui est identique à $-\frac{M-m}{M+m}Id \leqslant R \leqslant \frac{M-m}{M+m}Id$, d'où:

$$||R|| \leqslant \frac{M-m}{M+m} = \frac{r}{2+r}. \tag{A.8}$$

Donc l'erreur de reconstruction en norme euclidienne de f est donc $\frac{r}{2+r}||f||$. A partir de là il n'est pas difficile d'obtenir de meilleures formules de reconstruction, il suffit de répéter le processus. Avec la définition que l'on a de R on a

$$F_\Phi^* F_\Phi = \frac{m+M}{2}(Id - R) \tag{A.9}$$

et on a aussi montré $||R|| < 1$, on cherche $(F_\Phi^* F_\Phi)^{-1}$ pour connaitre les $\tilde{\varphi}_i$, donc il suffit d'inverser le terme de droite. Pour cela il suffit de développer la série de $(Id - R)^{-1}$ qui converge en norme. On a donc:

$$\tilde{\varphi}_i = (F_\Phi^* F_\Phi)^{-1}\varphi_i = \frac{2}{m+M}\sum_{k=0}^{\infty} R^k \varphi_i. \tag{A.10}$$

On peut donc construire une approximation à n itérations:

$$\tilde{\varphi}_i^n = \frac{2}{m+M} \sum_{k=0}^{n} R^k \varphi_i. \tag{A.11}$$

En utilisant le fait que l'on peut réécrire l'égalité précédente sous la forme

$$\tilde{\varphi}_i^n = Id - \frac{2}{m+M} \sum_{k=n+1}^{\infty} R^k \varphi_i = (Id - R^{n+1})\tilde{\varphi}_i. \tag{A.12}$$

on peut montrer que l'erreur d'une approximation de f avec N itérations est inférieure ou égale à $(\frac{r}{2+r})^{n+1}||f||$, on a ainsi une décroissance exponentielle en N.

A.1.3 Matching Pursuit

Présentons maintenant l'algorithme Matching Pursuit introduit par Stéphane Mallat [26]. Cet algorithme de la plus grande simplicité permet de faire une minimisation en norme 1 et peut notamment être utilisé pour la résolution de P1.
On dispose d'un dictionnaire $\Phi = (\varphi_1, \cdots, \varphi_N)$ et on cherche à minimiser la reconstruction en norme 1 d'un signal f. L'idée est très simple, à l'étape k on utilise deux vecteurs, le premier, $f^{(k)}$ représente l'approximation de f, le second, $r^{(k)}$ représente le résidu de l'approximation.
Initialement, $f^{(0)} = 0$ et $r^{(0)} = f$.
Pour passer de l'étape k à l'étape $k+1$, on ajoute à $f^{(k)}$ le vecteur de Φ qui est maximalement corrélé avec $r^{(k)}$ (multiplié par le coefficient de corrélation), et on définit $s^{(k+1)} = f - f^{(k+1)}$.
De nombreuses extensions de cet algorithme sont possible telles que l'Orthogonal Matching Pursuit ou le Robust Orthogonal Matching Pursuit. Un autre algorithme similaire est l'algorithme Basis Pursuit de Scott Chen, David Donoho et Michael Saunders [8] qui exploite notamment les résultats des chapitres 3 et 4.

A.2 Lemmes du théorème de Candes-Tao

Dans la preuve du théorème d'Emmanuel Candes et Terence Tao deux résultats ont été nécessaires, démontrons les ici. Le premier lemme peut être vu comme une stabilité en norme ℓ^1.

Lemme A.2.1. *Soit F_Ω une matrice d'analyse qui vérifie **ERP**. Soit f un signal fixé de la forme $f = f_0 + h$ où f_0 est un signal supporté sur un ensemble T qui vérifie*

$$|T| \leqslant \alpha \frac{K}{\lambda}. \tag{A.13}$$

Alors avec probabilité au moins $1 - \mathcal{O}(N^{-\frac{\rho}{\alpha}})$ alors n'importe quelle solution de P1

$$\min ||f^\#||_{\ell^1} \text{ tel que } F_\Omega f = F_\Omega f^\# \tag{A.14}$$

vérifie

$$||f^{\#}||_{\ell^1(T^c)} \leqslant 4||h||_{\ell^1}. \tag{A.15}$$

Preuve. Tout d'abord, $f^{\#}$ étant un minimiseur de P1, on a:

$$||f^{\#}||_{\ell^1} \leqslant ||f||_{\ell^1} \leqslant ||f_0||_{\ell^1} + ||h||_{\ell^1}. \tag{A.16}$$

On utilise maintenant le fait que F_Ω vérifie **ERP**, on a donc, avec probabilité au moins $1-\mathcal{O}(N^{-\frac{\rho}{\alpha}})$, il existe un vecteur P qui s'écrit $P = F_\Omega^* V$ (P est une combinaison linéaire des lignes de F_Ω) pour un certain $V \in \ell^2(\Omega)$ tel que pour tout $t \in T$, $P(t) = signe(f_0(t))$ et pour tout $t \in T^c$, $|P(t)| \leqslant \frac{1}{2}$.
Tout d'abord, utilisons $P = F_\Omega^* V$:

$$\langle f^{\#}, P\rangle = \langle f^{\#}, F_\Omega^* V\rangle = \langle F_\Omega f^{\#}, V\rangle = \langle F_\Omega(f_0 + h), V\rangle = \langle f_0 + h, F_\Omega^* V\rangle = \langle f_0 + h, P\rangle. \tag{A.17}$$

Utilisons maintenant le fait que P ait les mêmes signes que f_0 sur son support et que $|P(t)| < 1$ hors de T:

$$\langle f^{\#}, P\rangle = \langle f_0, P\rangle + \langle h, P\ \rangle \geqslant ||f_0||_{\ell^1} - ||h||_{\ell^1}. \tag{A.18}$$

et dans l'autre sens, avec $|P(t)| \leqslant \frac{1}{2}$, on a

$$|\langle f^{\#}, P\rangle| \leqslant \sum_{t\in T} |f^{\#}(t)P(t)| + \sum_{t\in T^c} |f^{\#}(t)P(t)| \tag{A.19}$$

$$\leqslant \sum_{t\in T} |f^{\#}(t)| + \frac{1}{2}\sum_{t\in T^c} |f^{\#}(t)| \tag{A.20}$$

$$= ||f^{\#}||_{\ell^1} - \frac{1}{2}||f^{\#}||_{\ell^1(T^c)}. \tag{A.21}$$

On a donc obtenu

$$||f_0||_{\ell^1} - ||h||_{\ell^1} \leqslant ||f^{\#}||_{\ell^1} - \frac{1}{2}||f^{\#}||_{\ell^1(T^c)}. \tag{A.22}$$

En remplaçant $||f^{\#}||_{\ell^1}$ par ce qui a été obtenu à l'inégalite A.16, on obtient finalement

$$-||h||_{\ell^1} \leqslant ||h||_{\ell^1} - \frac{1}{2}||f^{\#}||_{\ell^1(T^c)} \tag{A.23}$$

et en réorganisant les termes on a bien prouvé le lemme. □

On a donc démontré ce lemme qui nous indique que n'importe quelle solution de P1 est concentrée sur le support du signal parcimonieux sous-jacent. On peut maintenant montrer que si f vérifie la propriété de décroissance de ses coefficients dans une base fixée

$$|\theta|_{(n)} \leqslant Cn^{-\frac{1}{p}} \tag{A.24}$$

comme dans le théorème. Si on suppose que l'ensemble T contient les $|T|$ plus grands coefficients de f dans la base fixée, alors, en notant $f = f_0 + h$ avec $f_0 = f_T$ et $h = f_{T^c}$ alors

$$||h||_{\ell^1} = ||f||_{\ell^1(T^c)} \leqslant C \sum_{t=T+1}^{N} t^{-\frac{1}{p}} \tag{A.25}$$

On peut alors majorer facilement le terme de droite en utilisant d'abord le fait que l'on peut majorer le terme de droite par la série correspondante, celle-ci étant convergente car $0 < p < 1$, et ensuite on peut majorer la série par l'intégrale, ce qui donne

$$||f||_{\ell^1(T^c)} \leqslant C \sum_{t=T+1}^{N} t^{-\frac{1}{p}} \leqslant C \sum_{t=T+1}^{\infty} t^{-\frac{1}{p}} \leqslant C \int_T^{\infty} t^{-\frac{1}{p}}\, dt \leqslant C \frac{T^{1-\frac{1}{p}}}{\frac{1}{p}-1} = C_p T^{1-\frac{1}{p}} \tag{A.26}$$

et ainsi en appliquant le lemme

$$||f^{\#}||_{\ell^1(T^c)} \ \leqslant 4C_p|T|^{1-\frac{1}{p}}. \tag{A.27}$$

A partir de cela on peut obtenir une majoration sur les coefficients de $f^{\#}$ avec le lemme suivant:

Lemme A.2.2. *Soit $f^{\#}$ et T comme dans le lemme A.2.1, écrivons $|f^{\#}|_{(0)} \geqslant |f^{\#}|_{(1)} \geqslant \cdots \geqslant |f^{\#}|_{(N)}$ les coefficients de $f^{\#}$ par ordre décroissant. Alors ces coefficients vérifient pour tout $m > |T|$:*

$$|f^{\#}|_{(m)} \ \leqslant C_p \frac{|T|^{1-\frac{1}{p}}}{m-|T|}. \tag{A.28}$$

Preuve A.2.3. T est comme dans le lemme, donc c'est l'ensemble des $|T|$ plus grandes valeurs de f. Notons E_m les m plus grandes valeurs de $f^{\#}$. On a clairement $|E_m \cap T^c| \ \geqslant m - |T|$ et donc

$$||f^{\#}||_{\ell^1(E_m \cap T^c)} \geqslant (m-|T|)|f^{\#}|_{(m)}. \tag{A.29}$$

On a aussi

$$||f^{\#}||_{\ell^1(E_m \cap T^c)} \ \leqslant ||f^{\#}||_{\ell^1(T^c)} \ \leqslant C|T|^{1-\frac{1}{p}} \tag{A.30}$$

la dernière inégalité étant obtenue avec A.27. En combinant les deux inégalités précédentes on a alors le résultat souhaité.

Maintenant que l'on a montré le résultat utilisé dans la preuve qui utilise **ERP**, passons au résultat qui utilise **UUP**.
Les lemmes précédents peuvent être vus comme des résultats sur la norme ℓ^1. Le prochain résultat est lui par rapport à la norme ℓ^2.

Lemme A.2.4. *Soit F_Ω une matrice qui vérifie **UUP**. Alors, avec probabilité au moins $1 - \mathcal{O}(N^{-\frac{\rho}{\alpha}})$ alors, pour tout ensemble $T \subset \{0, \cdots, N-1\}$ vérifiant $|T| \ \leqslant \alpha \frac{K}{\lambda}$ et pour n'importe quel $f \in \ell^2(T)$, il existe $f^{ext} \in \ell^2(N)$ tel que:*

- *pour tout $t \in T$, $f(t) = f^{ext}(t)$ (c'est à dire $R_T f^{ext} = f$)*

- *f^{ext} est une combinaison linéaires des lignes de F_Ω (c'est à dire $f^{ext} = F_\Omega V$ pour un certain $V \in \ell^2(\Omega)$)*
- *De plus f^{ext} vérifie pour tout $E \subset \{0, \cdots, N-1\}$*

$$||f^{ext}||_{\ell^2(E)} \leqslant C\sqrt{(1 + \frac{E}{\alpha K/\lambda})}||f||_{\ell^2(T)} \tag{A.31}$$

Preuve. On peut donc supposer que **UUP** est vérifié, on a donc d'après le chapitre sur les frames que $F^*_{\Omega T}F_{\Omega T}$ est inversible, et d'après la proposition 4.2.4, en notant $||\cdot||$ la norme d'opérateur, on a

$$||(F^*_{\Omega T}F_\Omega)^{-1}|| \leqslant \frac{2N}{K}. \tag{A.32}$$

Posons

$$f^{ext} \ = F^*_\Omega F_{\Omega T}(F^*_{\Omega T}F_{\Omega T})^{-1}f. \tag{A.33}$$

On vérifie ainsi directement les deux premiers résultats du lemme. Dans un premier temps supposons que $|E| \leqslant \frac{\alpha K}{\lambda}$, on a alors avec la proposition 4.2.4 que $||F_{\Omega T}|| \leqslant \sqrt{\frac{3K}{2N}}$. On peut maintenant remarquer en notant $V = F_{\Omega T}(F^*_{\Omega T}F_{\Omega T})^{-1}f$:

$$||f^{ext}||_{\ell^2(E)} = ||R_E f_{ext}||_{\ell^2} = ||F^*_{\Omega E}V|| \leqslant \sqrt{\frac{3K}{2N}} \tag{A.34}$$

et on peut alors calculer la norme ℓ^2 de f^{ext} avec A.33:

$$||f^{ext}||_{\ell^2(E)} \leqslant ||F^*_{\Omega E}||\ ||F_{\Omega T}||\ ||(F^*_{\Omega T}F_{\Omega T})^{-1}||||f||_{\ell^2(T)} \tag{A.35}$$

et en remplaçant avec les majorations des normes d'opérateurs que l'on a obtenues on a:

$$||f^{ext}||_{\ell^2(E)} \leqslant \frac{3K}{2N}\frac{2N}{K}||f||_{\ell^2(T)} = 3||f||_{\ell^2(T)}. \tag{A.36}$$

Maintenant débarassons nous de l'hypothèse $|E| \ \leqslant \alpha K/\lambda$ en écrivant E comme une union disjointe d'ensembles chacun de taille inférieure ou égale à $\alpha K/\lambda$, c'est à dire $E = E_1 \bigcup E_2 \bigcup \cdots E_n$ avec $n = \lceil \frac{|E|}{\alpha K/\lambda} \rceil$ et $E_i \leqslant \alpha K/\lambda$, ainsi,

$$||f^{ext}||^2_{\ell^2(E)} \ = \sum_{i=1}^{n} \ ||f^{ext}||^2_{\ell^2(E_i)} \leqslant (1 + \frac{|E|}{\alpha K/\lambda})3||f||^2_{\ell^2(T)} \tag{A.37}$$

et on a donc bien le résultat souhaité. □

Bibliography

[1] Aldroubi, Akram, Cabrelli, Carlos, and Molter, Ursula M. "Wavelets on irregular grids with arbitrary dilation matrices and frame atoms for L2(Rd)". 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] Candes, E. J., Romberg, J., and Tao, T. "Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information". In: *IEEE Trans. Inf. Theor.* 52.2 (Feb. 2006), 489–509. ISSN: 0018-9448. DOI: 10.1109/TIT.2005.862083. URL: https://doi.org/10.1109/TIT.2005.862083.

[4] 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.

[5] 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.

[6] 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.

[7] Candès, Emmanuel and Romberg, Justin. "Sparsity and incoherence in compressive sampling". In: *Inverse Problems* 23.3 (2007), 969–985. ISSN: 1361-6420. DOI: 10.1088/0266-5611/23/3/008. URL: http://dx.doi.org/10.1088/0266-5611/23/3/008.

[8] Chen, Scott Shaobing, Donoho, David L., and Saunders, Michael A. "Atomic Decomposition by Basis Pursuit". In: *SIAM Rev.* 43.1 (Jan. 2001), 129–159. ISSN: 0036-1445. DOI: 10.1137/S003614450037906X. URL: https://doi.org/10.1137/S003614450037906X.

[9] Chen, Shaobing and Donoho, D. "Basis pursuit". In: *Proceedings of 1994 28th Asilomar Conference on Signals, Systems and Computers*. Vol. 1. 1994, 41–44 vol.1. DOI: 10.1109/ACSSC.1994.471413.

[10] 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.

[11] Coifman, R.R. and Wickerhauser, M.V. "Entropy-based algorithms for best basis selection". In: *IEEE Transactions on Information Theory* 38.2 (1992), pp. 713–718. DOI: 10.1109/18.119732.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] Donoho, David L. and Stark, Philip B. "Uncertainty Principles and Signal Recovery". In: *SIAM Journal on Applied Mathematics* 49.3 (1989), pp. 906–931. ISSN: 00361399. URL: http://www.jstor.org/stable/2101993.

[18] Donoho, D.L. "Compressed sensing". In: *IEEE Transactions on Information Theory* 52.4 (2006), pp. 1289–1306. DOI: 10.1109/TIT.2006.871582.

[19] 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.

[20] Feuer, Arie, Member, Senior, and Nemirovski, Arkadi. "On sparse representations in pairs of bases". In: *IEEE Trans. Inf. Theory* (2003), pp. 1579–1581.

[21] Foucart, Simon and Rauhut, Holger. *A Mathematical Introduction to Compressive Sensing*. Birkhäuser Basel, 2013. ISBN: 0817649476.

[22] 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.

[23] Jaffard, S. "POINTWISE SMOOTHNESS, TWO-MICROLOCALIZATION AND WAVELET COEFFICIENTS". In: *Publicacions Matemàtiques* 35.1 (1991), pp. 155–168. ISSN: 02141493, 20144350. URL: http://www.jstor.org/stable/43736311.

[24] 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.

[25] Lustig, Michael et al. "Compressed Sensing MRI". In: *IEEE Signal Processing Magazine* 25.2 (2008), pp. 72–82. DOI: 10.1109/MSP.2007.914728.

[26] Mallat, S.G. and Zhang, Zhifeng. "Matching pursuits with time-frequency dictionaries". In: *IEEE Transactions on Signal Processing* 41.12 (1993), pp. 3397–3415. DOI: 10.1109/78.258082.

[27] Mallat, Stphane. *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*. 3rd. USA: Academic Press, Inc., 2008. ISBN: 0123743702.

[28] 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/CBO9780511623820.

[29] Queffélec Hervé et Queffélec, Martine. *Analyse complexe et applications : cours et exercices / Hervé Queffélec, Martine Queffélec*. fre. Mathématiques en devenir. Paris: Calvage et Mounet, DL 2019. ISBN: 978-2-916352-59-6.

[30] Richman, Fred, Bridges, Douglas, and Schuster, Peter. "Trace-class operators". In: *Houston Journal of Mathematics* 28.3 (2002), pp. 565–583.

[31] 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.

[32] S.Jaffard. "Décompositions en ondelettes". In: (). URL: https://perso.math.u-pem.fr/jaffard.stephane/pdf/decompositions_en_ondelettes.pdf.

[33] 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.

[34] Tao, Terence. "An uncertainty principle for cyclic groups of prime order". In: *arXiv Mathematics e-prints*, math/0308286 (Aug. 2003), math/0308286. arXiv: math/0308286 [math.CA].

[35] 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.

[36] 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.

[37] 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.

[38] Zheng, Xuehang et al. *Sparse-View X-Ray CT Reconstruction Using ℓ_1 Prior with Learned Transform*. 2019. arXiv: 1711.00905 [stat.ML].