

Reconstruction parcimonieuse de signal

Introduction à la reconstruction de signal et au compressed sensing

Leo Davy

supervisé par

Jean-François Crouzet

Institut de Mathématiques Alexander Grothendieck

Université de Montpellier

Faculté des sciences

Février-Mai

Contents

1	Cadre du problème	3
1.1	Problématique de la reconstruction de signal	3
1.1.1	Exemples de signaux étudiés (image, sons, tomographie, ...) et formalisation mathématique de leur description	4
1.2	Exactitude, échantillonnage et bruit	4
1.2.1	Lien entre l'exactitude et l'échantillonnage	4
1.2.2	L'importance du bruit dans les problèmes	5
1.3	Bases orthonormales et frames	5
1.3.1	Intérêt des bases orthonormales et description des outils mathématiques disponibles	5
1.3.2	Lien entre frames et base orthonormale	5
1.3.3	Exemples de frames (Fourier, ondelettes)	7
1.4	Le théorème de Shannon-Nyquist	10
1.4.1	L'échantillonnage selon Shannon d'un signal à support compact en fréquence	10
1.4.2	L'échantillonnage selon Shannon d'un signal k-sparse	10
2	Approche parcimonieuse	11
2.1	Signaux ayant une représentation parcimonieuse	11
2.1.1	Définition d'une représentation parcimonieuse	11
2.1.2	Importance de la base	11
2.2	Décroissance des coefficients et régularité	11
2.2.1	Approximation linéaire et régularité	11
2.2.2	Décroissance des coefficients de Fourier	13
2.2.3	Décroissance des coefficients d'ondelettes	13
2.3	Résolution de (P0)	16
2.3.1	Définition de (P0)	17
2.3.2	Solution optimale combinatoire	19
2.3.3	Résolution dans un dictionnaire pics/Fourier	19
2.3.4	Principe d'incertitude	19
2.4	Résolution de (P1)	19
2.4.1	Définition de (P1)	23
2.4.2	Propriétés du minimiseur	23
2.5	Lien géométrique entre (P0) et (P1)	23
2.5.1	Boules unité en grande dimension	23
2.5.2	Unicité de la solution de (P0) et (P1)	23
3	Compressed sensing et approche aléatoire	24
3.1	Axiomatisation, UUP et RIP	24
3.1.1	Définition de UUP	25
3.1.2	Exemple de familles vérifiant UUP	25
3.1.3	Définition de ERP	26

3.1.4	Exemples de familles vérifiant ERP	26
3.1.5	Lien entre RIP et ERP	26
3.2	Théorème de Candes-Tao	26
3.2.1	Enoncé du théorème	26
3.2.2	Preuve du théorème	27
3.3	Exemple de F_Ω	27
3.3.1	Ensemble de Fourier	27
3.3.2	Gaussien	27
3.4	Conséquences du théorème	27
3.4.1	Influence des paramètres	27
3.4.2	Quelques résultats numériques	27
3.5	Sur la propriété RIP	27
3.5.1	Difficulté pour un ensemble de vérifier RIP	27
3.5.2	Lien entre RIP et WERP	27
3.6	Extensions du théorème	27
3.6.1	Conditions suffisantes sur δ_K	27
3.6.2	Conditions nécessaires sur δ_K	27
3.6.3	Sur l'optimalité du résultat	27
3.7	Algorithmes	27
3.7.1	Orthogonal Matching Pursuit	27
3.7.2	Robust Orthogonal Matching Pursuit	27
3.7.3	Quelques exemples numériques	27
A	Annexe	28
A.1	Valeurs propres de $F_\Omega F_\Omega^*$ et Analyse en composante principale	28
A.2	Outils probabilistes de la preuve du théorème	28

Chapter 1

Cadre du problème

1.1 Problématique de la reconstruction de signal

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 $A : \mathcal{F} \longrightarrow \mathcal{G}$, où \mathcal{G} est une autre famille de signaux. Donnons ici quelques exemples de familles de fonctions que l'on rencontrera dans ce mémoire. Commençons avec les fonctions à temps continu (c'est à dire avec $X = \mathbb{R}^d$ pour un certain $d > 0$).

Exemple 1.1.1. Signaux à énergie finie :

1. $\mathcal{F} = L^2(\mathbb{R}^d) := \{f : \mathbb{R}^d \longrightarrow \mathbb{R} \mid \int_{\mathbb{R}^d} |f(t)|^2 dt < \infty\}$.
2. $\mathcal{F} = L^p(\mathbb{R}^d) := \{f : \mathbb{R}^d \longrightarrow \mathbb{R} \mid \int_{\mathbb{R}^d} |f(t)|^p dt < \infty\}$, pour $0 < p \leq 1$.
3. $\mathcal{F} = L^\infty(\mathbb{R}^d) := \{f : \mathbb{R}^d \longrightarrow \mathbb{R} \mid \sup_{t \in \mathbb{R}^d} |f(t)| < \infty\}$.

Signaux avec une régularité :

1. $\mathcal{F} = \mathcal{C}^0(\mathbb{R}^d) = \{f : \mathbb{R}^d \longrightarrow \mathbb{R} \text{ avec } f \text{ continue}\}$.
2. $\mathcal{F} = \mathcal{C}^r(\mathbb{R}^d) = \{f : \mathbb{R}^d \longrightarrow \mathbb{R} \text{ avec } \forall t \in \mathbb{R}^d \sup_{0 \leq k \leq r} |f^{(k)}(t)| < \infty\}$.

on pourra aussi considérer pour chacun des espaces ci-dessus, leur version à support compact, notée avec l'indice $_{\text{loc}}$ (par exemple L^p_{loc} ou $\mathcal{C}^p_{\text{loc}}$) pour indiquer que pour tout $f \in \mathcal{F}_{\text{loc}}$, il existe un $C \geq 0$ tel que $|t| \geq C \implies f(t) = 0$. On considérera également des cas où X est un ouvert ou un fermé de \mathbb{R}^d . Un autre espace qui sera peut être utilisé est l'espace de Sobolev $W^{r,p}$ qui représente des signaux à énergie finie pour $\|\cdot\|_p$, mais dont chaque dérivée (définie faiblement) d'ordre inférieur ou égal à r est elle aussi à énergie finie.

Une autre classe d'intérêt de signaux majeur est celle des signaux à temps discret (c'est à dire avec $X = \mathbb{N}^d$).

Exemple 1.1.2. Signaux à énergie finie :

1. $\mathcal{F} = l^2(\mathbb{N}^d) := \{f : \mathbb{N}^d \longrightarrow \mathbb{R} \mid \sum_{\mathbb{N}^d} |f(t)|^2 < \infty\}$.
2. $\mathcal{F} = l^p(\mathbb{N}^d) := \{f : \mathbb{N}^d \longrightarrow \mathbb{R} \mid \sum_{\mathbb{N}^d} |f(t)|^p < \infty\}$, pour $0 < p \leq 1$.
3. $\mathcal{F} = l^\infty(\mathbb{N}^d) := \{f : \mathbb{N}^d \longrightarrow \mathbb{R} \mid \sup_{t \in \mathbb{N}^d} |f(t)| < \infty\}$.

On verra plus loin que l'on peut également définir une notion de régularité intéressante pour les signaux à temps discret.

Remarque 1.1.3. Dans les exemples ci-dessus les signaux sont à valeur dans $Y = \mathbb{R}$, cependant toutes ces exemples peuvent être considérées avec \mathbb{C} comme espace d'arrivée.

1.1.1 Exemples de signaux étudiés (image, sons, tomographie, ...) et formalisation mathématique de leur description

Considérons maintenant de façon plus concrète des exemples de signaux étudiés afin d'introduire l'opérateur A .
TODO : Pour chacun des exemples ci-dessous, formaliser le problème et poser sa solution comme un problème de minimisation "argmin"

Exemple 1.1.4. L'exemple le plus simple pour introduire le sujet est celui d'un signal à une dimension, on pourra par exemple penser à un signal décrivant un son ou bien un signal électrique, d'une durée finie, et à chaque instant on peut associer une amplitude. D'un point de vue formel on pourra ainsi considérer que ce signal est $f : [0, 1] \rightarrow \mathbb{R}^+$. On pourra chercher à faire différentes opérations sur ce signal :

- *Echantillonnage*
- *Seuil*
- *Décomposition harmonique (Fourier)*
- *Filtrage*

Exemple 1.1.5. Après avoir considéré le signal à une dimension, un autre type de signal est celui des signaux en deux ou trois dimensions dont l'exemple type est celui des vidéos ou des images :

- *Débruitage*
- *Super-résolution*
- *Compression*
- *Détection / Reconnaissance*

Exemple 1.1.6. Une autre problématique essentielle que l'on considérera en profondeur dans ce mémoire est celui des problèmes inverses dans lesquels à partir d'un signal mesuré, on cherchera à reconstruire ce qui a émis ce signal.

- *Tomographie (transformée de Radon)*
- *Géologie/IRM*

Exemple 1.1.7. Récemment, des problèmes avec des signaux en grande dimension sont aussi apparus, notamment dans des problématiques de type big-data.

Cependant sur ces problèmes il y a une ambiguïté sur la définition de la dimension qui est considérée comme la dimension de l'espace de départ du signal considéré, mais cependant, chaque signal appartient à un espace qui n'a à priori aucune raison d'être fini. De plus, chacun de ces problèmes commence par une mesure qui est toujours un processus discret et le reste du traitement est réalisé sur un ordinateur qui est lui aussi un processus discret. Ainsi chacun de ces problèmes est discretisé et alors la dimension¹ du signal augmente de façon considérable, ainsi, une image photographie a typiquement une dimension $d \gg 10^6$.

1.2 Exactitude, échantillonnage et bruit

1.2.1 Lien entre l'exactitude et l'échantillonnage

Ainsi il est nécessaire d'adapter la stratégie d'échantillonnage, un échantillonnage insuffisant ou inapproprié ne permettra pas avec certitude de pouvoir récupérer l'information sous-jacente au signal. Un échantillonnage qui prendrait trop de mesures pose aussi des problèmes, premièrement car si la taille de ces mesures est trop importante, il sera difficile de faire des opérations dessus et cela compliquera la résolution du problème. Mais aussi, car augmenter le nombre

¹On considère ici la "dimension" comme étant le nombre de degrés de liberté du signal étudié.

de mesures risque de ne pas apporter davantage d'informations pour la résolution du problème ², et dans certains cas, ces mesures superflues risquent seulement de mesurer du bruit et donc de diminuer l'efficacité de la résolution. Il est donc nécessaire pour une famille de signaux donnée d'avoir des conditions nécessaires sur l'échantillonnage, pour veiller à être certain d'étudier au moins le signal, mais aussi des conditions suffisantes pour ne pas étudier trop au delà du signal.

1.2.2 L'importance du bruit dans les problèmes

Ainsi il est nécessaire de prendre en compte le fait qu'il y ait des sources de bruit dans les problèmes considérés et on cherchera donc à vérifier que les constructions qui viendront seront stables face au bruit. Une remarque importante à faire est que le bruit est généralement constitué de modifications très locales (que l'on considérera ainsi comme "hautes-fréquences").

1.3 Bases orthonormales et frames

1.3.1 Intérêt des bases orthonormales et description des outils mathématiques disponibles

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

On verra ainsi dans cette section tout d'abord des définitions et propriétés classiques des espaces de Hilbert. Ensuite on verra progressivement comment relâcher certaines des définitions initiales afin de pouvoir conserver une formule de reconstruction. Afin d'explicitier l'intérêt de ces définitions on verra deux exemples de frames. Tout d'abord le frame de Fourier, dont la compréhension sera utile pour le troisième chapitre. Finalement, nous introduirons les ondelettes par l'analyse multi-résolution, les formules que nous obtiendrons seront utilisées dans le chapitre suivant.

1.3.2 Lien entre frames et base orthonormale

Rappelons tout d'abord les définitions et propriétés d'une base orthonormale. On considère ici un espace H muni d'un produit scalaire et une famille $\{e_i\}_I$, avec I un ensemble.

Definition 1.3.1. On dira qu'une famille $\{e_i\}_I$ est :

- *libre* si pour n'importe quelle suite finie de coefficients $(\lambda_i)_I$ telle que $\sum_I \lambda_i e_i = 0$, on a $\lambda_i = 0$ pour n'importe quel $i \in I$.
- *orthogonale*³ si pour n'importe quels i et j différents on a $\langle e_i, e_j \rangle = 0$
- *génératrice* 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* si la famille est libre et génératrice.

Donc, pour tout $h \in H$, si la famille $\{e_i\}_I$ est libre et génératrice, il existe une unique suite $(\lambda_i)_{i \in I}$ de scalaires, telle que $h = \sum_{i \in I} \lambda_i e_i$. On peut alors définir un nouveau produit scalaire sur H :

$$\langle \cdot, \cdot \rangle : H \times H \longrightarrow \mathbb{R} \quad (1.1)$$

$$(h_1 = (\lambda_i)_I, h_2 = (\mu_i)_I) \longmapsto \langle h_1, h_2 \rangle = \sum_I \lambda_i \mu_i^*. \quad (1.2)$$

²On peut ici penser aux problématiques d'*overfitting* du *Machine Learning* dans lesquels un système qui est trop entraîné sur un ensemble de données devient inefficace dès qu'il est testé sur des données sur lesquelles il n'a pas été entraîné

³Dans un espace vectoriel muni d'un produit scalaire et d'une base, une famille est libre si et seulement si elle est orthogonale (cela découle des propriétés du produit scalaire, notamment qu'il est défini).

On peut remarquer que ce produit scalaire est défini de façon unique par rapport à la base $\{e_i\}_I$, cependant, si la base est normalisée ($\langle e_i, e_i \rangle = 1$), alors ce produit scalaire devient identique au produit scalaire initial de H . Ainsi, la valuation du produit scalaire est indépendante du choix de la base (tant qu'elle est normalisée)⁴. On a alors 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 :

Theoreme 1.3.2. *Soit $\{f_i\}_I$ une suite d'éléments orthonormaux dans H muni d'un produit scalaire. Alors $\overline{\text{Vect}(\{f_i\}_I)} = H$ si et seulement si*

$$\sum_I |\langle f, f_i \rangle|^2 = \|f\|_2^2, \forall f \in H.$$

Cependant, comme on le verra dans la suite, il y a des situations dans lesquelles chercher à avoir une base orthonormale est trop restrictif, on cherchera donc à relâcher les conditions sur la définition d'une base. Tout d'abord, si la famille est orthogonale, mais elle n'est pas génératrice on a

Theoreme 1.3.3. *Soit $\{f_i\}_I$ une famille orthonormale de H . Alors,*

$$\sum_I |\langle f, f_i \rangle|^2 \leq \|f\|_2^2, \forall f \in H$$

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

Theoreme 1.3.4. *Soit $\{f_i\}_I$ une famille génératrice normalisée. Alors,*

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

On peut exprimer ce théorème en disant que l'analyse par une famille génératrice capture au moins l'énergie du vecteur analysé. Au vu de ces résultats, on est amenés à considérer les définitions suivantes qui correspondent à une relaxation de la condition de normalisation ou de la condition d'orthogonalité.

Definition 1.3.5. Pour une famille d'éléments $\{f_i\}_I$ de H , alors on dit que c'est

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

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

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

$$m \|f\|^2 \leq \sum_I |\langle f, f_i \rangle|^2 \leq M \|f\|^2, \forall f \in H. \quad (1.3)$$

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

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

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

Remarque 1.3.6. • Une base orthonormale est une base de Riesz avec $m = M = 1$.

⁴Dans $L^2(\mathbb{R})$, on peut ainsi retrouver les égalités de Parseval ou de Plancherel en exprimant une fonction $f \in L^2(\mathbb{R})$ soit dans la base canonique de $L^2(\mathbb{R})$ par rapport à f donnée par le vecteur $\frac{f}{\|f\|}$ et une base de son orthogonal, soit en exprimant la fonction dans la base de Fourier.

- 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.

Ainsi, lorsque l'on dispose d'une suite $F = \{f_i\}_I$ on peut définir l'opérateur d'analyse

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

et de synthèse

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

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

$$\theta_F^* \circ \theta_F(f) = \sum_I \langle f, f_i \rangle f_i. \quad (1.4)$$

Tout d'abord on peut remarquer que, si F est une famille orthogonale, alors l'application précédente correspond presque à une projection orthogonale dans l'espace engendré par F ⁵. On va maintenant voir que si F est un frame *équilibré* (c'est à dire avec des constantes m, M égales), alors on dispose d'une formule analogue à 1.4 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 1.3,

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

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,

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

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

1.3.3 Exemples de frames (Fourier, ondelettes)

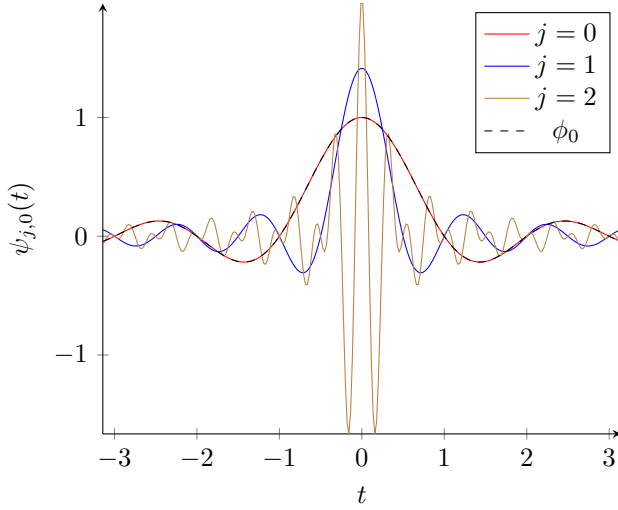
Introduisons maintenant la décomposition d'une fonction en ondelettes.

Definition 1.3.7. On dit que $\psi : \mathbb{R} \rightarrow \mathbb{R}$ est une ondelette génératrice si

$$\{\psi_{j,k} := 2^{j/2} \psi(2^j \cdot -k)\}_{j,k \in \mathbb{Z}} \quad (1.5)$$

est une famille génératrice de $L^2(\mathbb{R})$. On appellera⁶ une base engendrée par une telle fonction ψ une base d'ondelettes.

Ondelette ψ de Shannon (temporelle)



Remarque 1.3.8. Dans la définition des ondelettes $\psi_{j,k}$ engendrées par ψ , le coefficient j correspond au facteur d'échelle⁷, en raison du facteur 2^j devant la variable, au fur et à mesure que j augmente, l'ondelette est parcourue de plus en plus vite. Ainsi, augmenter j revient à augmenter la fréquence de ψ , c'est à dire d'éloigner le support de $\hat{\psi}$ de l'origine⁸. 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).

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

$$Wf = (\langle f, \psi_{j,k} \rangle)_{j,k \in \mathbb{Z}} \quad (1.6)$$

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

Supposons ainsi que l'on dispose d'un frame d'ondelettes et que ce frame est équilibré (c'est à dire que les bornes de frame m et M sont égales), alors on dispose d'une formule de reconstruction (d'après Daubechies 3.2.2)

$$f = \frac{1}{M} \sum_{j,k} \langle f, \psi_{j,k} \rangle \psi_{j,k}. \quad (1.7)$$

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 emboîtés

⁵La propriété qui n'est pas vérifiée est $(\theta_F^* \circ \theta_F)^2 = Id_{Vect(F)}$

⁶On considérera par la suite des frames d'ondelettes ou bien des bases de Riesz d'ondelettes

⁷Du point de vue des notations, on considère que j tend vers l'infini, signifie que ψ_j analyse les hautes fréquences, ce choix de notation n'est pas uniforme dans la littérature, par exemple (TODO : ajouter ref) Mallat et Daubechies, utilisent $-j$ par rapport à nos notations. Par contre les notations utilisées correspondent à celles de Jaffard et Meyer. Mais cependant tous les résultats sont bien entendu équivalents.

⁸En effet, de façon plus précise et formelle, on a $\hat{\psi}_{j,0}(\omega) = 2^{-\frac{j}{2}} \hat{\psi}(\frac{\omega}{2^j})$.

satisfaisant

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

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

$$f(\cdot) \in V_j \iff f(\cdot/2^j) \in V_0, \quad (1.8)$$

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

$$f \in V_0 \iff f(\cdot - n) \in V_0 \forall n \in \mathbb{Z}. \quad (1.9)$$

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} ,

$$f = \pi_{V_0}(f) + \sum_{i>0} \pi_{W_j}(f) \quad (1.10)$$

donc afin d'avoir une formule de reconstruction, il suffit de connaître un frame de V_0 et de même pour chaque W_j . On peut maintenant revenir aux ondelettes, on considère que l'on connaît une base $(\varphi_{0,k})_{k \in \mathbb{Z}}$ de V_0 et $(\psi_{j,k})_{j \in \mathbb{N}^*, k \in \mathbb{Z}}$ un frame de l'orthogonal de V_0 dans $L^2(\mathbb{R})$, on pose alors $W_j = W_{j-1} \oplus Vect(\{\psi_{j,k}\}_{k \in \mathbb{Z}})$ et on obtient ainsi que l'analyse multi-échelle ainsi construite fournit une formule de reconstruction⁹

$$f = \sum_{k \in \mathbb{Z}} \langle f, \varphi_{0,k} \rangle \tilde{\varphi}_{0,k} + \sum_{j=1}^{+\infty} \sum_{k \in \mathbb{Z}} \langle f, \psi_{j,k} \rangle \tilde{\psi}_{j,k}. \quad (1.11)$$

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

Proposition 1.3.9. 1. La transformée de Fourier discrète est un frame pour \mathcal{F} .

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

Preuve 1.3.10. 1. On utilisera la transformée de Fourier discrète écrite sous forme de matrice unitaire, qui est une matrice unitaire de Vandermonde avec les racines de l'unité en coefficients. Avec de l'algèbre on montre l'orthonormalité des colonnes.

2. Application du théorème suivant.

Theoreme 1.3.11. Soit $\mathcal{Q} \subset \mathbb{R}^d$ un ensemble de mesure finie, $h \in L^2(\mathbb{R}^d)$ et $\mathcal{A} = \{A_j \in GL_d(\mathbb{R})\}_J$ une famille de matrices inversibles.

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

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

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

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

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

⁹Dans la formule de reconstruction les deux sommes sur \mathbb{Z} ne sont pas problématiques car les fonctions considérées sont dans $L^2(\mathbb{R})$ donc avec une décroissance suffisamment rapide, donc seulement un nombre fini de $\langle f, \varphi_{0,k} \rangle$ sont différents de 0 si φ_0 a une décroissance suffisamment rapide (et de même pour chaque ψ_j).

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

Preuve 1.3.12. Voir [1] pour la preuve et les définitions, je les ajouterais ici et au dessus plus tard

1.4 Le théorème de Shannon-Nyquist

1.4.1 L'échantillonnage selon Shannon d'un signal à support compact en fréquence

1.4.2 L'échantillonnage selon Shannon d'un signal k-sparse

Chapter 2

Approche parcimonieuse

2.1 Signaux ayant une représentation parcimonieuse

2.1.1 Définition d'une représentation parcimonieuse

2.1.2 Importance de la base

2.2 Décroissance des coefficients et régularité

2.2.1 Approximation linéaire et régularité

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

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

$$f_N = \sum_{k=0}^{N-1} \langle f, \phi_k \rangle \tilde{\phi}_k. \quad (2.1)$$

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

$$f_N = \sum_{k=0}^{N-1} \langle f, g_k \rangle g_k \quad (2.2)$$

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

$$f = \sum_{k=0}^{+\infty} \langle f, g_k \rangle g_k. \quad (2.3)$$

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

$$f - f_N = \sum_{k=N}^{+\infty} \langle f, g_k \rangle g_k \quad (2.4)$$

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

$$\varepsilon_l(N, f) = \|f - f_N\|^2 = \sum_{k=N}^{+\infty} |\langle f, g_k \rangle|^2. \quad (2.5)$$

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.

Theoreme 2.2.1. *Soit $r > 1/2$, il existe des constantes $A, B > 0$ telles que si*

$$\sum_{k=0}^{+\infty} |k|^{2r} |\langle f, g_k \rangle|^2 < \infty, \quad (2.6)$$

alors on a

$$A \sum_{k=0}^{+\infty} k^{2r} |\langle f, g_k \rangle|^2 \leq \sum_{N=0}^{+\infty} N^{2r-1} \varepsilon_l(N, f) \leq B \sum_{k=0}^{+\infty} k^{2r} |\langle f, g_k \rangle|^2 \quad (2.7)$$

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

Preuve 2.2.2. On développe le terme au centre de l'égalité de la façon suivante

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

Puis on majore des deux côtés avec

$$\int_0^M x^{2r-1} dx \leq \sum_{N=0}^m N^{2r-1} \leq \int_1^{m+1} x^{2r-1} dx. \quad (2.8)$$

En calculant les deux intégrales on déduit

$$Am^{2r} \leq \sum_{N=0}^m N^{2r-1} \leq Bm^{2r} \quad (2.9)$$

où A et B dépendent seulement de r , ce qui nous donne 2.7. Montrons maintenant $\varepsilon_l(N, f) = o(N^{-2r})$, remarquons tout d'abord que $\varepsilon_l(N, f)$ est décroissant par rapport à la première variable, on déduit de cela

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

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

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

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

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

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

$$W_{\mathcal{B},r} = \{f : \sum_{m=0}^{+\infty} m^{2r} |\langle f, g_m \rangle|^2 < \infty\} \quad (2.10)$$

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

2.2.2 Décroissance des coefficients de Fourier

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

$$f_N(t) = \sum_{|n| \leq N/2} \langle f, e_n \rangle e_n(t) \in U_N. \quad (2.11)$$

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

$$\hat{f}'(\omega) \int_{-\infty}^{+\infty} f'(t) e^{i\omega t} dt = i\omega \int_{-\infty}^{+\infty} f(t) e^{i\omega t} dt = i\omega \hat{f}(\omega) \quad (2.12)$$

et en utilisant la formule de Plancherel on a

$$\|\hat{f}'\|_2^2 = \int_{-\infty}^{+\infty} |\omega|^2 |\hat{f}(\omega)|^2 d\omega = \int_{-\infty}^{+\infty} |f'(t)|^2 dt = \|f'\|_2^2. \quad (2.13)$$

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

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

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

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

$$W^r(\mathbb{R}) = \{f \in L^2(\mathbb{R}) : \int_{\mathbb{R}} |\omega|^{2r} |\hat{f}(\omega)|^2 d\omega < \infty\}. \quad (2.14)$$

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

2.2.3 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,

Definition 2.2.4. Soit α tel qu'il existe un entier strictement positif r tel que $r - 1 \leq \alpha < r$. On dit qu'une fonction f est α -Lipschitz en t_0 si il existe une constante $C > 0$ et un polynôme P_{t_0} de degré strictement inférieur à r , tels que pour tout t qui appartient à un voisinage T_0 de t_0 , on a

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

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

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

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

Remarque 2.2.5. 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 2.2.3, on obtient

$$Wf(j, k) = f \star \widetilde{\psi}_j(k). \quad (2.15)$$

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 commun¹.

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

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

Definition 2.2.6. Soit m un entier strictement positif. On dit que ψ a m -moments nuls si

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

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

$$\int \psi(t) p(t) dt = 0. \quad (2.16)$$

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

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

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

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

¹L'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.

Theoreme 2.2.7. Si f est α -Lipschitz en t_0 avec $0 < \alpha \leq m$ où m est un entier. Alors, il existe une constante $C > 0$ telle que

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

Preuve 2.2.8. Remarquons tout d'abord que, soit P un polynôme de degré strictement inférieur à m , et ψ une ondelette à m -moments nuls, alors, par changement de variable et d'après 2.16 on a :

$$WP(j, k) = \int 2^{j/2} \psi(2^j t - 2^{-j} k) P(t) dt = \int 2^{j/2} \psi(t') P(2^j t' - 2^{-j} k) 2^{-j} dt' = 0. \quad (2.17)$$

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

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

la dernière inégalité étant obtenue en utilisant 2.17 sur le terme de droite et en faisant entrer la valeur absolue dans la première intégrale. On utilise maintenant le fait que f est α -Lipschitz et on fait un changement de variable, on obtient ainsi

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

Pour obtenir l'inégalité suivante on utilise

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

et on a ainsi

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

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

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

Proposition 2.2.9. Soit $f :]0, 1[\rightarrow \mathbb{R}$ une fonction α -Lipschitzienne avec $\alpha > 1$, alors il existe une constante $C > 0$ et une base d'ondelette orthonormales associée à une multirésolution $\{(\psi_{j,k})_{(i,j): j \geq J, 2^j > k \geq 0}\}, \{\varphi_{J,k}\}_k$ avec $m > \alpha$ moments nuls telle que

$$\varepsilon_l(f, 2^J) = \|f - \sum_{k=0}^{2^J-1} \langle f, \varphi_{J,k} \rangle \tilde{\varphi}_{J,k}\|_2^2 \leq C 2^{-2J\alpha} = C N^{-2\alpha} \quad (2.18)$$

avec $N = 2^J$.

Preuve 2.2.10. L'existence d'une telle base d'ondelette n'est pas démontrée ici, des constructions peuvent être trouvées dans (ajouter ref) pour obtenir des bases de $L^2(\mathbb{R})$, on peut ainsi considérer une telle multirésolution donnée par une ondelette de Daubechies ou bien une coiflet à $m > \alpha$ moments nuls. Il est ensuite possible, avec quelques difficultés d'obtenir depuis ces ondelettes, une base orthonormale de $L^2(]0, 1[)$ (ajouter ref). Soit f une fonction α -Lipschitzienne sur $]0, 1[$, on a ainsi d'après la partie sur les frames et l'existence de la base d'ondelette précédente admise, une formule de reconstruction

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

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

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

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

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

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

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

on utilise maintenant le théorème 2.2.7 et on obtient²

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

ce qui prouve la proposition.

2.3 Résolution de (P0)

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

On a par exemple vu que l'on pouvait reconstruire les fonctions Lipschitziennes avec une bonne précision en utilisant une base d'ondelettes orthonormale avec un certain nombre de moments nuls. De plus, on a remarqué que si la fonction se comporte comme un polynôme d'un degré inférieur au nombre de moments nuls au voisinage d'un point, alors les coefficients d'ondelettes dans ce voisinage seront nuls. De même, les seuls coefficients d'ondelette qui seront

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

grands seront ceux au voisinage d'un point où aucune approximation par un polynôme de petit degré n'est efficace³. Ainsi, la représentation avec ces ondelettes d'une fonction ne possédant que quelques points où elle est irrégulière sera approximée avec peu de coefficients. Afin d'insister, l'intérêt de cela est que de façon naïve, afin de déterminer une fonction, il faut connaître sa valeur en chaque point, ainsi, si l'on souhaite faire un traitement par ordinateur de cette fonction, il faut stocker chacun des points de la fonction. Avec ce que l'on a fait, on sait qu'en fait on peut reconstruire la fonction avec un plus petit nombre de coefficients que la fonction n'a de points. En ce sens, la représentation en ondelettes d'une fonction Hölderienne est parcimonieuse (peu de coefficients non nuls), alors que la représentation par la valuation d'une fonction Hölderienne n'est pas parcimonieuse.

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

2.3.1 Définition de (P0)

Formalisons maintenant ce que nous avons dit ci-dessus. On considère \mathcal{F} un espace vectoriel et utilisons un dictionnaire $\Phi = \Phi_1 \cup \dots \cup \Phi_D$ de bases, où chaque Φ_d est une base de \mathcal{F} . Ainsi Φ est une concaténation de bases⁴ et on s'intéresse à la façon d'écrire un signal $f \in \mathcal{F}$ dans Φ , c'est à dire à la façon d'écrire

$$f = \sum_{\gamma} c_{\gamma} \phi_{\gamma} \quad (2.19)$$

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 2.19 sous forme matricielle en posant F_{Φ} la matrice ayant pour lignes les vecteurs ϕ_{γ} et en posant $x = (c_{\gamma})_{\gamma}$ la notation sous forme de vecteur de x , on utilisera aussi la notation $x = (x_d)_{d=1, \dots, D}$. On s'intéresse ainsi aux solutions de

$$f = F_{\Phi} x. \quad (2.20)$$

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

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

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

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

³Une analyse du théorème de Jaffard 2.2.7 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.

⁴Ainsi Φ est un frame équilibré d'après la première partie (TODO : ajouter ref)

Exemple 2.3.1. On étudie les signaux dans \mathbb{R}^N et on choisit un dictionnaire constitué de la concaténation de la base de Fourier $W = \{e_k(t) = \frac{1}{\sqrt{N}} e^{i\frac{2\pi kt}{N}}\}_{0 \leq k \leq N-1}$ et de la base canonique de Diracs⁵ $T = \{\delta_k\}_{0 \leq k \leq N-1}$. Ainsi, avec ce choix F_W est la matrice de Fourier discrète et F_T est la matrice identité de taille N . Avec le théorème suivant, on va obtenir un principe d'incertitude, qui nous garantira qu'un signal ne peut pas être parcimonieux à la fois dans la base de Dirac, et dans la base de Fourier.

Theoreme 2.3.2. Soit un signal $f \in \mathbb{R}^N$ non nul, alors

$$\|F_W f\|_0 \|F_T f\|_0 \geq N \quad (2.21)$$

et ainsi

$$\|F_W f\|_0 + \|F_T f\|_0 \geq 2\sqrt{N}. \quad (2.22)$$

Preuve 2.3.3. TODO :Ajouter Ref Tao uncertainty principle for cyclic.... Soit $0 \leq \omega \leq N-1$ un entier, alors

$$|F_W(\omega)| = |\hat{f}(\omega)| = \frac{1}{\sqrt{N}} \left| \sum_t f(t) e_{\omega}(t) \right| \quad (2.23)$$

$$\leq \frac{1}{\sqrt{N}} \sum_t |f(t)|, \quad (2.24)$$

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

$$\sup_{\omega} |F_W(\omega)| \leq \frac{1}{\sqrt{N}} \sum_t |F_T f(t)| = \frac{1}{\sqrt{N}} \langle \text{sign}(f), |f| \rangle \quad (2.25)$$

$$\leq \frac{1}{\sqrt{N}} \|F_T\|_0^{\frac{1}{2}} \langle |f|, |f| \rangle^{\frac{1}{2}} = \frac{1}{\sqrt{N}} \|F_T\|_0^{\frac{1}{2}} \langle |F_W f|, |F_W f| \rangle^{\frac{1}{2}} \quad (2.26)$$

$$\leq \frac{1}{\sqrt{N}} \|F_T\|_0^{\frac{1}{2}} \langle |f|, |f| \rangle^{\frac{1}{2}} = \frac{1}{\sqrt{N}} \|F_T\|_0^{\frac{1}{2}} \left(\sum_{\omega} |\hat{f}(\omega)|^2 \right)^{\frac{1}{2}} \leq \frac{1}{\sqrt{N}} \|F_T\|_0^{\frac{1}{2}} \|F_W f\|_0^{\frac{1}{2}} \sup_{\omega} |F_W(\omega)|. \quad (2.27)$$

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

$$\sqrt{\|F_T f\|_0 \|F_W f\|_0} \leq \frac{\|F_T f\|_0 + \|F_W f\|_0}{2} \quad (2.28)$$

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

Theoreme 2.3.4. Soit N un entier positif 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

$$\|F_W x_W\|_0 + \|F_T x_T\|_0 < \sqrt{N} \quad (2.29)$$

est l'unique solution de (P0).

⁵Chaque vecteur de cette base vérifie $\delta_{k,i} = 1$ si $i = k$ et 0 sinon.

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

$$F_\Phi(x_1 - x_2) = 0. \quad (2.30)$$

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

$$F_T \delta_T + F_W \delta_W = 0 \quad (2.31)$$

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

$$\delta_W = -F_W^t F_T \delta_T. \quad (2.32)$$

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 2.3.2, on a que δ a au moins $2\sqrt{N}$ coefficients non nuls si δ est non nul. En revenant à la situation initiale $\delta = x_1 - x_2$, on a une contradiction car à la fois x_1 et x_2 ont chacun moins de \sqrt{N} coefficients, donc $\delta = 0$. Ainsi, si une solution existe avec moins de \sqrt{N} coefficients, alors c'est la solution de (P0) et elle est unique.

Cependant, en choisissant $N = p$, où p est un nombre premier⁶, Tao 2005 (TODO: ajouter ref) a montré que l'on obtient l'inégalité

$$\|F_T f\|_0 + \|F_W f\|_0 \geq p + 1 \quad (2.33)$$

et que l'inégalité est atteinte⁷. Grâce à ce principe d'incertitude plus fort que 2.3.2, on obtient avec le même type de preuve⁸ le résultat suivant

Theoreme 2.3.5. Soit N un nombre premier et un signal $f \in \mathbb{R}^N$, alors n'importe quel x vérifiant $f = F_\Phi x = F_W x_W + F_T x_T$ et

$$\|F_T x_T\|_0 \leq \frac{N}{2} \quad (2.34)$$

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.

2.3.2 Solution optimale combinatoire

2.3.3 Résolution dans un dictionnaire pics/Fourier

2.3.4 Principe d'incertitude

2.4 Résolution de (P1)

On a ainsi vu dans la section précédente que le problème (P0) de minimisation de la solution par rapport à la parcimonie admet une solution unique dès qu'une solution existe et que cette solution vérifie une condition de la forme 2.29 ou 2.34. 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

⁶L'hypothèse p premier est essentielle, la preuve reposant sur la non-existence de sous-groupes propres du groupe cyclique $\mathbb{Z}/p\mathbb{Z}$

⁷L'inégalité est atteinte en ce sens que si $A \subset T$ et $B \subset W$ tels que $|A| + |B| \geq p + 1$ alors il existe une fonction f telle que $\text{Supp} F_T f = A$ et $\text{Supp} F_W f = B$

⁸Voir Candes-Romberg-Tao pour les détails, un lemme sur l'injectivité d'un opérateur similaire F_W est tout de même nécessaire pour conclure la preuve.

peut résoudre un autre problème pour lequel des méthodes de résolution efficaces existaient déjà. En effet nous allons voir que résoudre le problème P1,

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

permet sous certaines conditions de résoudre P0. L'intérêt de P1 est que c'est un problème de programmation linéaire et de nombreuses méthodes permettent de le résoudre. (TODO ajouter refs et détails). Précisons donc ce que nous avons affirmé, dans le même cadre que précédemment, c'est à dire dans le cas d'un dictionnaire $\Phi = (T, W)$ temps fréquence composé de la base de Fourier et de Dirac dans \mathbb{R}^N .

Theoreme 2.4.1. *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*

$$\|x_T\|_0 + \|x_W\|_0 < \frac{1}{2}\sqrt{N} \quad (2.35)$$

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

Preuve 2.4.2. Tout d'abord, remarquons que si x vérifie 2.53, alors x vérifie 2.29 et donc d'après le théorème 2.3.4 x est donc l'unique condition de P0. Il nous faut donc vérifier que cette solution est bien la solution de (P1). Supposons que x vérifie les conditions du théorème, on doit donc montrer que si x est l'unique solution de (P1), alors pour tout x_1 différent de x qui vérifie $f = Fx_1$ alors $\|x_1\|_1 > \|x\|_1$. Donc de façon équivalente, pour tout $\delta \in \mathcal{N} = \{\delta : F_\Phi \delta = 0\}$ non nul, on doit vérifier que

$$\|x + \delta\|_1 - \|x\|_1 > 0. \quad (2.36)$$

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

$$\|x + \delta\|_1 - \|x\|_1 = \sum_{\gamma \in \Gamma^c} |\delta_\gamma| + \sum_{\gamma \in \Gamma} |c_\gamma + \delta_\gamma| - |c_\gamma|. \quad (2.37)$$

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

$$|c_\gamma + \delta_\gamma| - |c_\gamma| \geq -|\delta_\gamma| \quad (2.38)$$

et en insérant cette inégalité dans la somme on obtient

$$\|x + \delta\|_1 - \|x\|_1 \geq \sum_{\gamma \in \Gamma^c} |\delta_\gamma| - \sum_{\gamma \in \Gamma} |\delta_\gamma|, \quad (2.39)$$

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

$$\sum_{\gamma \in \Gamma} |\delta_\gamma| < \sum_{\gamma \in \Gamma^c} |\delta_\gamma|. \quad (2.40)$$

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

$$\sum_{\gamma \in \Gamma} |\delta_\gamma| < \frac{1}{2} \left(\sum_{t \in T} |\delta_{T,t}| + \sum_{\omega \in W} |\delta_{W,\omega}| \right) = \frac{\|\delta_T\|_1 + \|\delta_W\|_1}{2} \quad (2.41)$$

On avait vu avec 2.32 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 2.40 devient

$$\sum_{t \in \Gamma_T} |\delta_{T,t}| + \sum_{\omega \in \Gamma_W} |\widehat{\delta_T}_\omega| < \frac{\|\delta_T\|_1 + \|\widehat{\delta_T}\|_1}{2} \quad (2.42)$$

On peut ainsi définir la quantité

$$\mu(\Gamma_T, \Gamma_W) = \sup_x \frac{\sum_{t \in \Gamma_T} |\delta_{T,t}| + \sum_{\omega \in \Gamma_W} |\widehat{\delta_{T\omega}}|}{\|\delta_T\|_1 + \|\delta_W\|_1} \quad (2.43)$$

et ainsi la condition suffisante pour que la solution de P1 soit unique devient

$$\mu(\Gamma_T, \Gamma_W) < \frac{1}{2}. \quad (2.44)$$

Avant de démontrer que cette inégalité est bien vérifiée énonçons ce que nous avons démontré qui est une généralisation du théorème 2.4.1 :

Theoreme 2.4.3. Soit N un entier positif et un signal $f \in \mathbb{R}^N$, alors n'importe quel $x = (x_T, x_W)$, où Γ_T est le support de x_T et Γ_W est le support de x_W , vérifiant $f = F_T x_T + F_W x_W$ et

$$\mu(\Gamma_T, \Gamma_W) < \frac{1}{2} \quad (2.45)$$

est l'unique solution de P1.

Montrons donc que dans le dictionnaire Fourier-Dirac cette condition est bien vérifiée, Tout d'abord minorons le dénominateur avec

$$\|\delta_T\|_1 \geq \|\delta_T\|_\infty \quad \text{et} \quad \|\widehat{\delta_T}\|_1 \geq \sqrt{N} \|\delta_T\|_\infty, \quad (2.46)$$

on a ainsi

$$\mu(\Gamma_T, \Gamma_W) \leq \frac{\sum_{t \in \Gamma_T} |\delta_{T,t}| + \sum_{\omega \in \Gamma_W} |\widehat{\delta_{T\omega}}|}{\|\delta\|_\infty (\sqrt{N} + 1)}. \quad (2.47)$$

(Je crois que ce qui suit, jusqu'à l'énoncé du prochain théorème, est faux, je donne une preuve alternative qui me semble correcte après avoir énoncé le théorème). Maintenant majorons le numérateur avec

$$\sum_{\omega \in \Gamma_W} |\widehat{\delta_{T\omega}}| = \|R_{\Gamma_W} F_W^t F_T \delta_T\|_1 \leq \|R_{\Gamma_W} F_W^t F_T\|_\infty \|\delta_T\|_\infty \quad (2.48)$$

où $\|A\|_\infty = \sup_i \|c_i\|_1$ 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 a⁹

$$\|R_{\Gamma_W} F_W^t F_T\|_\infty = \frac{|\Gamma_W|}{\sqrt{N}}. \quad (2.49)$$

On peut maintenant en combinant 2.56 avec 2.55 et en utilisant $\sum_{t \in \Gamma_T} |\delta_{T,t}| \leq |\Gamma_T| \|\delta_{T,t}\|_\infty$ écrire la majoration du numérateur de 2.54 :

$$\sum_{t \in \Gamma_T} |\delta_{T,t}| + \sum_{\omega \in \Gamma_W} |\widehat{\delta_{T\omega}}| \leq \|\delta_T\|_\infty (|\Gamma_T| + \frac{|\Gamma_W|}{\sqrt{N}}). \quad (2.50)$$

Il nous reste maintenant à appliquer la majoration que l'on vient de trouver à 2.54 et on obtient

$$\mu(\Gamma_T, \Gamma_W) \leq \frac{|\Gamma_T| + \frac{|\Gamma_W|}{\sqrt{N}}}{\sqrt{N} + 1}. \quad (2.51)$$

On utilise maintenant l'hypothèse (affaiblie) du théorème $|\Gamma_T| < \frac{\sqrt{N}}{2}$ et $|\Gamma_W| < \frac{\sqrt{N}}{2}$ et on obtient

$$\mu(\Gamma_T, \Gamma_W) < \frac{\frac{\sqrt{N}}{2} + \frac{1}{2}}{\sqrt{N} + 1} = \frac{1}{2} \frac{\sqrt{N} + 1}{\sqrt{N} + 1} = \frac{1}{2}. \quad (2.52)$$

⁹C'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 $1/\sqrt{N}$

On vient ainsi de montrer en utilisant le théorème 2.4.3 que si une solution existe vérifiant $\|x_T\|_0 < \sqrt{N}/2$ et $\|x_W\|_0 < \sqrt{N}/2$ alors c'est l'unique solution de P1. Ce qui conclut la preuve du théorème. En fait on a même montré le théorème suivant

Theoreme 2.4.4. 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

$$\|x_T\|_0 < \frac{1}{2}\sqrt{N} \quad \text{et} \quad \|x_W\|_0 < \frac{1}{2}\sqrt{N} \quad (2.53)$$

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

Preuve 2.4.5. Réécrivons une preuve correcte :

$$\mu(\Gamma_T, \Gamma_W) \leq \frac{\sum_{t \in \Gamma_T} |\delta_{T,t}| + \sum_{\omega \in \Gamma_W} |\widehat{\delta_{T\omega}}|}{\|\delta_T\|_1 + \|\delta_W\|_1}. \quad (2.54)$$

Maintenant majorons le numérateur avec

$$\sum_{\omega \in \Gamma_W} |\widehat{\delta_{T\omega}}| = \|R_{\Gamma_W} F_W^t F_T \delta_T\|_1 \leq \|R_{\Gamma_W} F_W^t F_T\|_1 \|\delta_T\|_1 \quad (2.55)$$

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 a¹⁰ :

$$\|R_{\Gamma_W} F_W^t F_T\|_1 = \frac{|\Gamma_W|}{\sqrt{N}}. \quad (2.56)$$

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

$$\sum_{t \in \Gamma_T} |\delta_{T,t}| = \|R_{\Gamma_T} F_T^t F_W \delta_W\|_1 \leq \|R_{\Gamma_T} F_T^t F_W\|_1 \|\delta_W\|_1 \quad (2.57)$$

ainsi que

$$\|R_{\Gamma_T} F_T^t F_W\|_1 = \frac{|\Gamma_T|}{\sqrt{N}}. \quad (2.58)$$

On peut maintenant rassembler les résultats:

$$\sum_{t \in \Gamma_T} |\delta_{T,t}| + \sum_{\omega \in \Gamma_W} |\widehat{\delta_{T\omega}}| \leq \|\delta_W\|_1 \frac{|\Gamma_T|}{\sqrt{N}} + \|\delta_T\|_1 \frac{|\Gamma_W|}{\sqrt{N}}. \quad (2.59)$$

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

$$\sum_{t \in \Gamma_T} |\delta_{T,t}| + \sum_{\omega \in \Gamma_W} |\widehat{\delta_{T\omega}}| < \frac{\|\delta_T\|_1 + \|\delta_W\|_1}{2}. \quad (2.60)$$

Il nous reste maintenant à appliquer la majoration que l'on vient de trouver à 2.54 et on obtient

$$\mu(\Gamma_T, \Gamma_W) < \frac{1}{2} \frac{\|\delta_T\|_1 + \|\delta_W\|_1}{\|\delta_T\|_1 + \|\delta_W\|_1} = \frac{1}{2} \quad (2.61)$$

Ce qui conclut la preuve (correcte?) du théorème.

¹⁰C'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 $1/\sqrt{N}$

2.4.1 Définition de (P1)

2.4.2 Propriétés du minimiseur

2.5 Lien géométrique entre (P0) et (P1)

2.5.1 Boules unité en grande dimension

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

Chapter 3

Compressed sensing et approche aléatoire

3.1 Axiomatisation, UUP et RIP

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

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

$$\theta : \mathcal{F} \longrightarrow \mathbb{R}^{|\Omega|} \quad (3.1)$$

$$(f_k)_{k=0, \dots, N} = f \longmapsto (y_k = \langle f, \psi_k \rangle)_{k \in \Omega} = \theta(f) \quad (3.2)$$

et une application de synthèse associée

$$\mathbb{R}^{|\Omega|} \longrightarrow \mathcal{F}^\# \subset \mathbb{R}^N \quad (3.3)$$

$$(y_k)_{k \in \Omega} \longmapsto (y_k)^\# = (f_k^\#)_{k=0, \dots, N} \quad (3.4)$$

et on cherche à obtenir une ε -reconstruction :

$$\|f - \theta(f)^\#\|_2 \leq \varepsilon, \quad \forall f \in \mathcal{F}. \quad (3.5)$$

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

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

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

$$R_\Omega : \ell^2([0, N]) \longrightarrow \ell^2(\Omega) \quad (3.6)$$

$$(g_k)_{0 \leq k \leq N} \longmapsto (g_k)_{k \in \Omega} \quad (3.7)$$

et l'inclusion prolongée par des zéros

$$R_T^* : \ell^2(T) \longrightarrow \ell^2([0, N]) \quad (3.8)$$

$$(g_k)_{k \in T} \longmapsto (g_k)_{k \in T} \oplus (0)_{k \in T^c}. \quad (3.9)$$

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

$$F_{\Omega T} = F_\Omega R_T^* : \ell^2(T) \longrightarrow \ell^2(\Omega) \quad (3.10)$$

$$(g_k)_T \longmapsto F_\Omega((g_k)_T \oplus (0)_{T^c}). \quad (3.11)$$

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

3.1.1 Définition de UUP

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

Definition 3.1.1. On dit que F_Ω vérifie λ -**UUP** si il existe ρ tel que avec probabilité $1 - \mathcal{O}(N^{-\rho/\alpha})$ on ait:
 $\forall f \subset \mathbb{R}^N$ signal tel que

$$|\text{supp}(f)| \leq \alpha K / \lambda \quad (3.12)$$

on ait l'inégalité

$$\frac{1}{2} \frac{K}{N} \|f\|_2^2 \leq \|F_\Omega f\|_2^2 \leq \frac{3}{2} \frac{K}{N} \|f\|_2^2. \quad (3.13)$$

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

Proposition 3.1.2. F_Ω vérifie λ -**UUP** si et seulement si
avec probabilité au moins $1 - \mathcal{O}(N^{-\rho/\alpha})$ on a $\forall T \subset [0, N]$ qui vérifie $|T| \leq \alpha \frac{K}{\lambda}$ alors les valeurs propres de $F_{\Omega T}$ vérifient

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

Preuve 3.1.3. A recopier.

Remarque 3.1.4. Pour expliciter le fait que cela définit bien un principe d'incertitude, considérons F_Ω comme étant la transformée de fourier discrète partielle, et un signal concentré en temps ($|\text{supp}(f)| \leq \alpha \frac{K}{\lambda}$), alors on a

$$\|F_\Omega f\|_{\ell^2} = \|\hat{f}\|_{\ell^2(\Omega)} \leq \sqrt{\frac{3K}{2N}} \|f\|_{\ell^2} \quad (3.14)$$

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

$$\frac{\|\hat{f}\|_{\ell^2(\Omega)}}{\|f\|_{\ell^2}} \longrightarrow 0 \quad (3.15)$$

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

Remarque 3.1.5. Justifions maintenant le fait que c'est un principe uniforme. Une version non uniforme (et donc plus faible) serait que pour chaque f vérifiant 3.12, alors avec probabilité au moins $1 - \mathcal{O}(N^{-\rho/\alpha})$ 3.13 est vérifié. Mais il y a beaucoup de choix possibles de f vérifiant 3.12, et parmi ceux-ci il peut y avoir un grand nombre de f ayant la propriété rare de ne pas vérifier 3.13, 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 3.13 soit vrai pour tous les f possibles vérifiant 3.12. Ce qui justifie l'appellation uniforme.

Remarque 3.1.6. ¹ Remarquons que l'on peut réécrire 3.13 peut se réécrire

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

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

3.1.2 Exemple de familles vérifiant UUP

Proposition 3.1.7. ²

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

¹A vérifier

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

3.1.3 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).

Definition 3.1.8. F_Ω vérifie **ERP** si

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

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

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

3.1.4 Exemples de familles vérifiant ERP

Proposition 3.1.9. • Les ensembles Gaussiens et binaires vérifient $\log N$ -**ERP**

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

3.1.5 Lien entre RIP et ERP

3.2 Théorème de Candes-Tao

3.2.1 Enoncé du théorème

Theoreme 3.2.1. Soit F_Ω qui vérifie λ_1 -**ERP** et λ_2 -**UUP**. On pose $\lambda = \max(\lambda_1, \lambda_2)$, soit $K \geq \lambda$. Soit f un signal dans \mathbb{R}^N tel que ses coefficients dans une base de référence décroissent comme⁵ :

$$|\theta_{(n)}| \leq C n^{-\frac{1}{p}} \quad (3.16)$$

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

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

$$\|f - f^\#\|_2 \leq C_r \left(\frac{K}{\lambda}\right)^{-r} \quad (3.17)$$

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

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

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

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

- 3.2.2 Preuve du théorème
- 3.3 Exemple de F_Ω
 - 3.3.1 Ensemble de Fourier
 - 3.3.2 Gaussien
- 3.4 Conséquences du théorème
 - 3.4.1 Influence des paramètres
 - 3.4.2 Quelques résultats numériques
- 3.5 Sur la propriété RIP
 - 3.5.1 Difficulté pour un ensemble de vérifier RIP
 - 3.5.2 Lien entre RIP et WERP
- 3.6 Extensions du théorème
 - 3.6.1 Conditions suffisantes sur δ_K
 - 3.6.2 Conditions nécessaires sur δ_K
 - 3.6.3 Sur l'optimalité du résultat
- 3.7 Algorithmes
 - 3.7.1 Orthogonal Matching Pursuit
 - 3.7.2 Robust Orthogonal Matching Pursuit
 - 3.7.3 Quelques exemples numériques

Appendix A

Annexe

A.1 Valeurs propres de $F_{\Omega}F_{\Omega}^*$ et Analyse en composante principale

A.2 Outils probabilistes de la preuve du théorème

Bibliography

- [1] Akram Aldroubi, Carlos Cabrelli, and Ursula M. Molter. “Wavelets on irregular grids with arbitrary dilation matrices and frame atoms for $L_2(\mathbb{R}^d)$ ”. In: *Applied and Computational Harmonic Analysis* 17.2 (2004). Special Issue: Frames in Harmonic Analysis, Part II, pp. 119–140. ISSN: 1063-5203. DOI: <https://doi.org/10.1016/j.acha.2004.03.005>. URL: <https://www.sciencedirect.com/science/article/pii/S1063520304000442>.