Université de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)

Présentée et soutenue par :
Laurent DALENC

Le mardi 26 août 2014

Titre :

Characterization of product BMO and iterated commutators involving Calderón-Zygmund operators

---

ED MITT : Domaine Mathématiques : Mathématiques appliquées

Unité de recherche :

UMR 5219

Directeur(s) de Thèse :

Stefanie PETERMICHL, Professeur, IMT, Université Paul Sabatier, Toulouse

Rapporteurs :

Brett WICK, Professor, School of mathematics, Georgia Institute of Technology
Jean ESTERLE, Professeur, Institut de Mathématiques de Bordeaux, Université de Bordeaux

Autre(s) membre(s) du jury :

Stanislas KUPIN, Professeur, Institut de Mathématiques de Bordeaux, Université de Bordeaux

THÈSE

En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :
Université Toulouse III Paul Sabatier (UT3 Paul Sabatier)

Discipline ou spécialité :
Mathématiques Appliquées

Présentée et soutenue par
Laurent DALENC
le : 26 Septembre 2014

Titre :
Characterization of product BMO and iterated commutators involving Calderón-Zygmund operators

---

École doctorale :
Mathématiques Informatique Télécommunications (MITT)

Unité de recherche :
UMR 5219

Directeurs de thèse :
Stefanie PETERMICHL, Professeur, IMT, Université Paul Sabatier, Toulouse

Rapporteurs :
Brett WICK, Professor, School of mathematics, Georgia Institute of Technology
Jean ESTERLE, Professeur, Institut de Mathématiques de Bordeaux, Université de Bordeaux

Membres du jury :
Stefanie PETERMICHL, Professeur, IMT, Université Paul Sabatier, Toulouse
Brett WICK, Professor, invité, School of mathematics, Georgia Institute of Technology
Jean ESTERLE, Professeur, Institut de Mathématiques de Bordeaux, Université de Bordeaux
Stanislas KUPIN, Professeur, Institut de Mathématiques de Bordeaux, Université de Bordeaux
Pascal J. THOMAS, Professeur, IMT, Université Paul Sabatier, Toulouse

Résumé

Le but de ma thèse est de décrire les familles d'opérateurs de Calderón-Zygmund qui, imbriqués au sein de commutateurs, caractérisent BMO à plusieurs paramètres. L'espace BMO à plusieurs paramètres est une généralisation de l'espace BMO classique, et a commencé à être étudié au cours des années 1980 par Chang et Fefferman. A chaque paramètre, on associe un opérateur de Calderón-Zygmund agissant sur ce paramètre, un opérateur de Calderón-Zygmund étant un opérateur à noyau. Ensuite, si b appartient à BMO, on lui associe l'opérateur M_b de multiplication par b. On considère ensuite une suite d'itérés de commutateurs ayant pour argument ces opérateurs de Calderón-Zygmund et M_b.
Le but est alors d'étudier le rapport entre la norme BMO de b et celle de ces itérés de commutateurs agissant sur L^2.
Le premier résultat concernant cette théorie est du à Coifman, Rochberg et Weiss qui ont démontré dans le cas du paramètre un que les transformées de Riesz, qui sont des opérateurs de Calderón-Zygmund, caractérisent BMO.
Le résultat suivant est du à Uchiyama, qui, lui, a proposé un critère portant sur une famille d'opérateurs de Calderón-Zygmund, pour savoir s'ils généralisent la décomposition de Stein-Fefferman, puis Li a fourni un critère englobant celui de Uchiyama pour savoir si un commutateur caractérise BMO à un paramètre.
Le premier théorème dans le cas du multiparamètre est du à Ferguson-Lacey qui ont montré dans le cas du paramètre t=2 que les transformées de Hilbert caractérisent BMO, puis Lacey-Ferguson l'on étendu à un nombre quelconque d'itérations.
Enfin, Lacey-Petermichl-Wick-Pipher ont étendu ce résultat au transformées de Riesz dans le cas du multiparamètre.
C'est, dans un premier temps, ce résultat que l'on a généralisé, fournissant un critère permettant de savoir si une famille d'opérateurs de Calderón-Zygmund caractérisent BMO à plusieurs paramètres.
Enfin, nous avons montré que la norme du commutateur est, à une constante multiplicative près, majorée par la norme BMO de b pour n'importe quel type d'opérateurs de Calderón-Zygmund, en utilisant le théorème de représentation de Hytonen qui permet de réduire le problème au cas des shifts dyadiques.

Abstract

The aim of my thesis was to find criteria on families of Calderón-Zygmund operators to know if, with iterated commutators, they characterize product BMO space. Multiparameter BMO space is a generalization of classical BMO space, and began to be studied during the eighties by Chang and Fefferman.
To each parameter is associated a Calderón-Zygmund operator acting on this parameter. We define also, associated to b in BMO, the operator M_b of multiplication by b. Then we define iterated commutators with those Calderón-Zygmund operators and the operator M_b. Then the aim is to study the relation between the BMO norm of b and the norm of the commutator acting on L^2.
The first result in one parameter case is due to Coifman, Rochberg and Weiss, who proved that Riesz transforms characterize BMO.
The next result is due to Uchiyama, who gave a criterion on families of Calderon-Zygmund operators to show if they generalize Stein-Fefferman decomposition. Then Li gave another criterion on those families of Calderon-Zygmund operators, to show if commutators characterize BMO.
The first result in multiparameter case is due to Ferguson-Lacey, who proved in the case of parameter t=2 that Hilbert transform characterize BMO. Then Lacey and Terwilleger extended this result to arbitraly number of iterations.
Finally, Lacey-Petermichl-Wick-Pipher extended this result to Riesz transform in product BMO space.
So, first, I found a criterion on families of Calderón-Zygmund operators to know if they characterize product BMO space.
Finally, I proved that commutators norms are majorized, up to a multiplicative constant, by BMO norm of b in multiparameter case for any kind of Calderón-Zygmund commutators, using the representation theorem of Hytonen, which reduces the problem to dyadic shifts.

Chapitre 0

Introduction

0.1 Opérateurs de Calderón-Zygmund

L'histoire des opérateurs de Calderón-Zygmund est longue : Yves Meyer distingue dans son traité ([18] et [19]) sur la théorie des ondelettes trois générations d'opérateurs de Calderón-Zygmund, les deux plus importantes étant la première et la troisième.

0.1.1 Première génération d'opérateurs de Calderón-Zygmund

Un opérateur de Calderón-Zygmund de première génération est défini à partir de ce qu'on appelle un noyau. Parmi ces opérateurs, les deux plus connus sont la transformée de Hilbert en dimension 1 et sa généralisation à la dimension quelconque, la transformée de Riesz. Les opérateurs de ce type sont d'abord définis sur un ensemble de fonctions tests, en général soit sur l'ensemble des fonctions de Schwarz, soit sur les fonctions indéfiniment différentiables à support compact. La technique par la suite consiste à procéder par densité pour les définir par exemple sur des espaces de Lebesgue ou de Hardy.
Ces opérateurs ont été étudié en détail par Elias M. Stein dans son fameux traité "Singular integrals and differential properties of functions" ([25]).

définition 1. *: Opérateurs de Calderón-Zygmund de première génération*

Soit $\Omega : \mathbb{R}^n \setminus \{0\} \to \mathbb{C}$ *une fonction homogène de degré 0 vérifiant les propriétés :*
(i) $\int_{\mathbb{S}^{n-1}} \Omega(x)\, d\sigma = 0$
(ii) soit w *défini par :*

$$\forall \delta \in [0,1], w(\delta) = \underset{|x-x'|\leq\delta, |x|=|x'|=1}{sup} |\Omega(x) - \Omega(x')|$$

alors w *doit vérifier :*

$$\int_0^1 \frac{w(\delta)}{\delta}\, d\delta < +\infty$$

Le noyau K est défini par :

$$\forall x \in \mathbb{R}^n \setminus \{0\}, K(x) = \frac{\Omega(x)}{|x|^n}$$

Nous posons alors

$$\forall f \in \mathcal{C}_0^\infty(\mathbb{R}^n), \forall x \notin supp(f), Tf(x) = \int_{\mathbb{R}^n} \frac{\Omega(y)}{|y|^n} f(x-y)\, dy$$

et T est l'opérateur de Calderón-Zygmund associé au noyau K.

Nous voyons dans cette définition que T est un opérateur de convolution. On dit que ce type d'opérateur est à intégrale singulière car le noyau K admet une singularité en 0. Ce qui fait converger l'intégrale en dépit de cette singularité est la propriété (i) de Ω que l'on appelle en anglais "cancellative".
L'opérateur T défini sur un ensemble de fonctions tests peut alors s'étendre à des espaces de Lebesgue.
En effet, T vérifie le théorème suivant :

théorème 1. *Soit T l'opérateur défini précédemment à partir du noyau K sur un ensemble de fonctions tests. Alors T s'étend de manière unique en un opérateur de $L^p(\mathbb{R}^n)$ pour $1 < p < +\infty$, soit*

$$\forall 1 < p < +\infty, T : L^p(\mathbb{R}^n) \longrightarrow L^p(\mathbb{R}^n)$$

Pour le cas p=1, T est dit de type faible (1,1), soit

$$T : L^1(\mathbb{R}^n) \longrightarrow L^{1,\infty}(\mathbb{R}^n)$$

ce qui signifie :

$$\exists A > 0, \forall f \in L^1(\mathbb{R}^n), \forall \alpha > 0$$

$$|\{x \in \mathbb{R}^n, |Tf(x)| > \alpha\}| \leq \frac{A}{\alpha} ||f||_1$$

Nous pouvons désormais donner les définitions des transformées de Hilbert et de Riesz qui sont des opérateurs de Calderón-Zygmund.

définition 2. *: Transformée de Hilbert*
C'est l'opérateur de Calderón-Zygmund associé au noyau $K(x) = \frac{1}{\pi x}$

définition 3. *: Transformée de Riesz*
La transformée de Riesz, notée $R_j, 1 \leq j \leq n$ est l'opérateur de Calderón-Zygmund associé au noyau $K(x) = c_n \frac{x_j}{|x|^{n+1}}, c_n = \frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}}$

0.1.2 La définition actuelle des opérateurs de Calderón-Zygmund

Une définition moderne des opérateurs de Calderón-Zygmund inclus la définition précédente. Voici la définition la plus générale :

définition 4. *Opérateur de Calderón-Zygmund*
Soit $K : \mathbb{R}^d \times \mathbb{R}^d \setminus \Delta \longrightarrow \mathbb{C}$ *avec* $\Delta = \{(x,x), x \in \mathbb{R}^d\}$
On suppose en outre que K vérifie :
(i) $\exists C_0 > 0, \forall (x,y) \in \mathbb{R}^d \times \mathbb{R}^d \setminus \Delta, |K(x,y)| \leq \frac{C_0}{|x-y|^d}$
(ii) $\exists C_\psi > 0, \forall x, x', y \in \mathbb{R}^d, x \neq y, x' \neq y, |x-y| > 2|x-x'|,$

$$|K(x,y) - K(x',y)| + |K(y,x) - K(y,x')| \leq \frac{C_\psi}{|x-y|^d}\psi(\frac{|x-x'|}{|x-y|})$$

On définit alors l'opérateur de Calderón-Zygmund T associé au noyau K sur l'ensemble des fonctions tests $\mathcal{C}_0^\infty(\mathbb{R}^d)$ *de la façon suivante :*

$$\forall f \in \mathcal{C}_0^\infty(\mathbb{R}^d), \forall x \notin supp(f), Tf(x) = \int_{\mathbb{R}^d} K(x,y)f(y)\,dy$$

Si T s'étend en un opérateur borné sur $L^2(\mathbb{R}^d)$, on dit que T est un opérateur borné de Calderón-Zygmund.
Les plus petites constantes C_0 et C_ψ telles que (i) et (ii) soient vérifiées sont notées respectivement $||K||_{CZ_0}$ et $||K||_{CZ_\psi}$.
La fonction ψ la plus utilisée est $\psi(t) = t^\alpha, \alpha \in (0,1]$

0.2 Les espaces de Hardy réels et BMO à un paramètre

Les espaces de Hardy et BMO, notés respectivement $H^1(\mathbb{R}^d)$ et $BMO(\mathbb{R}^d)$, sont intimement liés à la théorie des opérateurs de Calderón-Zygmund. La définition de l'espace BMO est beaucoup plus récente que celle des espaces de Hardy, et Charles Fefferman a prouvé que $BMO(\mathbb{R}^d)$ était en réalité le dual topologique de $H^1(\mathbb{R}^d)$.
L'intérêt des espaces de Hardy et BMO dans la théorie des opérateurs de Calderón-Zygmund est dû au fait que les espaces de Lebesgue $L^1(\mathbb{R}^d)$ et $L^\infty(\mathbb{R}^d)$ ne sont pas stables par ceux-ci, alors que si T est par exemple un opérateur de Calderón-Zygmund de première génération, celui-ci applique $H^1(\mathbb{R}^d)$ dans $L^1(\mathbb{R}^d)$ et $L^\infty(\mathbb{R}^d)$ dans $BMO(\mathbb{R}^d)$
En outre, nous avons les injections continues suivantes :

$$H^1(\mathbb{R}^d) \circlearrowleft L^1(\mathbb{R}^d)$$

$$L^\infty(\mathbb{R}^d) \circlearrowleft BMO(\mathbb{R}^d)$$

et comme $L^1(\mathbb{R}^d)$ et $L^\infty(\mathbb{R}^d)$, les espaces de Hardy et BMO ne sont pas réflexifs.
Définissons à présent l'espace de Hardy $H^1(\mathbb{R}^d)$

définition 5. *: Espace de Hardy* $H^1(\mathbb{R}^d)$

Soit $\Phi \in \mathcal{S}(\mathbb{R}^d)$ *avec* $\int_{\mathbb{R}^d} \Phi(x)\,dx \neq 0$. *Posons* $\forall t > 0, \Phi_t(x) = \frac{1}{t^d}\Phi(\frac{x}{t})$.
Soit $f \in \mathcal{S}'(\mathbb{R}^d)$, *définissons*

$$\forall x \in \mathbb{R}^d, M_\Phi f(x) = \sup_{t>0} |(f \star \Phi_t)(x)|$$

Alors $f \in H^1(\mathbb{R}^d)$ *ssi* $M_\Phi f \in L^1(\mathbb{R}^d)$

Il existe également une caractérisation de l'espace de Hardy $H^1(\mathbb{R}^d)$ à l'aide des transformées de Riesz :

théorème 2. *Soit* $f \in L^1(\mathbb{R}^d)$, *alors* $f \in H^1(\mathbb{R}^d)$ *ssi*

$$\forall 1 \leq j \leq d, R_j f \in L^1(\mathbb{R}^d)$$

définition 6. *Structure d'espace vectoriel normé sur l'espace de Hardy* $H^1(\mathbb{R}^d)$

Nous pouvons définir une structure d'espace vectoriel normé sur $H^1(\mathbb{R}^d)$ *par les normes équivalentes suivantes :*

$$||f||_{H^1} = ||M_\Phi f||_{L^1}$$

ou

$$||f||_{H^1} = ||f||_{L^1} + \sum_{j=1}^{d} ||R_j f||_{L^1}$$

Définissons à présent l'espace $BMO(\mathbb{R}^d)$.

définition 7. *Espace* $BMO(\mathbb{R}^d)$

Soit $f \in L^1_{loc}(\mathbb{R}^d)$ *et soit* $\mathcal{B}$ *l'ensemble des boules de* $\mathbb{R}^d$, *alors*

$$f \in BMO(\mathbb{R}^d) \text{ ssi } \sup_{B \in \mathcal{B}} \frac{1}{|B|} \int_B |f - \langle f \rangle_B|\,dx < +\infty$$

où $|B|$ *est la mesure de la boule* B *et où*

$$\langle f \rangle_B = \frac{1}{|B|} \int_B f\,dx$$

est la moyenne de f *sur* B
Nous notons alors

$$||f||_{BMO} = \sup_{B \in \mathcal{B}} \frac{1}{|B|} \int_B |f - \langle f \rangle_B|\,dx$$

Malheureusement, la dernière quantité n'est pas une norme. Pour cela, il faut quotienter cet espace par l'ensemble des fonctions constantes. Nous obtenons alors un espace vectoriel normé que nous noterons toujours $BMO(\mathbb{R}^d)$

0.3 Interprétation en termes d'ondelettes

La théorie des ondelettes a révolutionné la manière d'appréhender la théorie des opérateurs de Calderón-Zygmund ainsi que celle des espaces de Hardy et BMO. Nous allons en évoquer les éléments de base.
Une base d'ondelettes est engendrée par une ondelette-mère et un ondelette-père et est indexée par l'ensemble des cubes dyadiques appartenant à la grille dyadique dans $\mathbb{R}^d$. Définissons tout d'abord celle-ci.

définition 8. *: Grille dyadique dans* $\mathbb{R}^d$

Nous noterons $\mathcal{D}_d$ *la grille dyadique dans* $\mathbb{R}^d$*, ensemble des cubes dyadiques dans* $\mathbb{R}^d$*, définie par :*

$$\mathcal{D}_d = \{2^k j + [0, 2^k)^d, j \in \mathbb{Z}^d, k \in \mathbb{Z}\}$$

Nous pouvons à présent définir les bases d'ondelettes dans $\mathbb{R}^d$

définition 9. *: base d'ondelette dans* $\mathbb{R}^d$

Soit $w^0 = w$ *l'ondelette-mère et* $w^1 = W$ *l'ondelette-père.*
Posons pour $\epsilon \in \{0, 1\}$ *et pour* $I \in \mathcal{D}_1$

$$w_I^\epsilon(x) = \frac{1}{\sqrt{|I|}} w^\epsilon(\frac{x - c(I)}{|I|})$$

où $c(I)$ *est le centre du segment* I
Posons ensuite $Sig_d = \{0, 1\}^d \setminus \{\vec{1}\}$ *où* $\vec{1} = (1, 1, ..., 1)$
On définit alors :

$$\forall Q \in \mathcal{D}_d, \forall \epsilon \in Sig_d, w_Q^\epsilon(x_1, ..., x_d) = \prod_{j=1}^{d} w_{I_j}^{\epsilon_j}(x_j)$$

avec $Q = I_1 \times ... \times I_d$
On appelle alors $\{w_Q^\epsilon, Q \in \mathcal{D}_d, \epsilon \in Sig_d\}$ *la base d'ondelette engendrée par l'ondelette-mère* w *et l'ondelette-père* W.

Nous avons alors le théorème suivant

théorème 3. *La base d'ondelette définie précedemment forme une base hilbertienne de* $L^2(\mathbb{R}^d)$ *et une base inconditionnelle de* $L^p(\mathbb{R}^d), 1 < p < +\infty$

Sûrement, la base d'ondelette la plus connue et la plus utilisée est la base de Haar. Nous allons en rappeler la définition, ainsi que celle de la base d'ondelette de Meyer.

définition 10. *: base de Haar*

la base d'ondelette de Haar est la base d'ondelette engendrée par l'ondelette-mère $w^0 = 1_{[0,\frac{1}{2}]} - 1_{[\frac{1}{2},1]}$ *et l'ondelette-père* $w^1 = 1_{[0,1]}$

définition 11. *: ondelette de Meyer*
la base d'ondelette de Meyer est la base d'ondelette engendrée par une ondelette-mère $w^0 \in \mathcal{S}(\mathbb{R}^d)$ *telle que* $supp(\widehat{w^0}) \subset [-\frac{8}{3}, -\frac{2}{3}] \cup [\frac{2}{3}, \frac{8}{3}]$ *et* $\widehat{w^0} \equiv 1$ *sur* $[-2, -1] \cup [1, 2]$

Nous sommes désormais en mesure de donner une caractérisation de l'espace de Hardy et de l'espace BMO par l'intermédiaire de la théorie des ondelettes.

théorème 4. *Soit* $\{w_Q^\epsilon, Q \in \mathcal{D}_d, \epsilon \in Sig_d\}$ *une base d'ondelette. Soit* $f \in L^1(\mathbb{R}^d)$
Alors $f \in H^1(\mathbb{R}^d)$ *ssi*

$$(\sum_{Q,\epsilon} |\langle f, w_Q^\epsilon \rangle|^2 |w_Q^\epsilon(x)|^2)^{\frac{1}{2}} \in L^1(\mathbb{R}^d)$$

théorème 5. *Soit* $\{w_Q^\epsilon, Q \in \mathcal{D}_d, \epsilon \in Sig_d\}$ *une base d'ondelette.*
Alors $f \in BMO(\mathbb{R}^d)$ *ssi*

$$\sup_{Q \in \mathcal{D}_d} (\frac{1}{|Q|} \sum_{Q' \subset Q, \epsilon} |\langle f, w_{Q'}^\epsilon \rangle|^2)^{\frac{1}{2}} < +\infty$$

et si nous prenons cette quantité comme norme, nous obtenons toujours la topologie BMO

Ces deux derniers théorèmes sont démontrés dans [18]

0.4 Espaces de Hardy, BMO, L^p à plusieurs paramètres

Jusqu'à présent nous n'avons considéré que les espaces de Hardy, BMO et de Lebesgue classiques, c'est-à-dire à un paramètre. Maintenant nous allons définir une généralisation de ces espaces, les espaces de Hardy, BMO et de Lebesgue à plusieurs paramètres. Ceux-ci ont commencé à être étudiés aux cours des années 1980 par Sun-Yung A. Chang et Robert Fefferman ([3] et [4]).
Afin de définir ces espaces, nous avons besoin d'étendre la théorie des ondelettes au cas du multiparamètre. Pour cela, au lieu de raisonner dans $\mathbb{R}^d$ nous le ferons dans $\mathbb{R}^{\vec{d}} = \bigotimes_{s=1}^t \mathbb{R}^{d_s}$ avec t le nombre de paramètres et $\vec{d} = (d_1, ..., d_t)$.
Dans un premier temps nous allons définir la grille dyadique à t paramètres, c'est-à-dire l'ensemble des rectangles dyadiques dans $\mathbb{R}^{\vec{d}}$

définition 12. *: grille dyadique à t paramètres*
Soit $\vec{d} = (d_1, ...d_t) \in \mathbb{N}^t$*, la grille dyadique à t paramètres* $\mathcal{D}_{\vec{d}}$ *est définie par :*

$$\mathcal{D}_{\vec{d}} = \bigotimes_{s=1}^t \mathcal{D}_{d_s} = \{R \in \mathbb{R}^{\vec{d}}, R = \prod_{s=1}^t Q_s, Q_s \in \mathcal{D}_{d_s}\}$$

Ensuite nous définissons l'ensemble des signatures à t paramètres :

définition 13. *Ensemble des signatures à t paramètres*
C'est l'ensemble

$$Sig_{\vec{d}} = \{\vec{\epsilon} = (\epsilon_1, ..., \epsilon_t), \epsilon_s \in Sig_{d_s}, 1 \leq s \leq t\}$$

Nous pouvons à présent définir la base d'ondelette à t paramètres :

définition 14. *Soit w^0 une ondelette-mère et w^1 l'ondelette-père, la base d'ondelette à t paramètres est l'ensemble $\{w_R^{\vec{\epsilon}}, R \in \mathcal{D}_{\vec{d}}, \vec{\epsilon} \in Sig_{\vec{d}}\}$ telle que*

$$\forall R \in \mathcal{D}_{\vec{d}}, R = \prod_{s=1}^{t} Q_s, \forall \vec{\epsilon} \in Sig_{\vec{d}}, \forall (x_1, ...x_t) \in \mathbb{R}^{\vec{d}}, w_R^{\vec{\epsilon}} = \prod_{s=1}^{t} w_{Q_s}^{\epsilon_s}(x_s)$$

Nous pouvons à présent faire le lien avec les espaces de Lebesgue à t paramètres

théorème 6. *La base d'ondelette à t paramètres est une base hilbertienne de l'espace de Lebesgue à t paramètres $L^2(\mathbb{R}^{\vec{d}}) = \bigotimes_{s=1}^{t} L^2(\mathbb{R}^{d_s})$ et une base inconditionnelle de $L^p(\mathbb{R}^{\vec{d}}) = \bigotimes_{s=1}^{t} L^p(\mathbb{R}^{d_s})$ pour $1 < p < +\infty$*

Nous sommes désormais en mesure de définir les espaces de Hardy et BMO à t paramètres.
Soit $(w_R^{\vec{\epsilon}})$ une base d'ondelettes à t paramètres.
Soit $\varphi_1 \in \mathcal{C}_0^{\infty}(\mathbb{R}^{\vec{d}})$ telle que $\varphi_1 \geq 0$ et $\int_{\mathbb{R}^{\vec{d}}} \varphi_1 \, dx = 1$
Définissons $Dil_R^2 \varphi_1(x_1, ..., x_t) = \frac{1}{\sqrt{|R|}} \varphi_1(\frac{x_1 - c(Q_1)}{l(Q_1)}, ... \frac{x_t - c(Q_t)}{l(Q_t)})$

Définissons à partir de cela les quantités :

$$Mf(x) = \sup_{R \in \mathcal{D}_{\vec{d}}, \vec{\epsilon} \in Sig_{\vec{d}}} Dil_R^2 \varphi_1(x) |\langle f, Dil_R^2 \varphi_1 \rangle|$$

$$Sf(x) = (\sum_{R \in \mathcal{D}_{\vec{d}}, \vec{\epsilon} \in Sig_{\vec{d}}} |w_R^{\vec{\epsilon}}(x)|^2 |\langle f, w_R^{\vec{\epsilon}} \rangle|^2)^{\frac{1}{2}}$$

Soit $1 \leq s \leq t$ et $\vec{0} \leq \vec{j} \leq \vec{d}$ c'est-à-dire $0 \leq j_s \leq d_s$ pour $1 \leq s \leq t$. Définissons l'opérateur sur $L^2(\mathbb{R}^{\vec{d}})$ par

$$R_{s,j_s} = (\bigotimes_{i=1, i \neq s}^{t} Id_{L^2(\mathbb{R}^{d_i})}) \otimes R_{j_s}$$

où R_{j_s} est la transformée de Riesz dans $L^2(\mathbb{R}^{d_s})$ avec la convention $R_0 = id_{L^2(\mathbb{R}^{d_s})}$

définition 15. *Espace de Hardy à t paramètres*
Les quantités $||Mf||_1$, $||f||_1 + ||Sf||_1$ et $\sum_{\vec{0} \leq \vec{j} \leq \vec{d}} || \prod_{s=1}^{t} R_{s,j_s} f||_1$ définissent la même topologie d'espace vectoriel normé. Cet espace est appelé espace de Hardy à t paramètres et est noté $H^1(\mathbb{R}^{\vec{d}})$

Il ne nous reste plus qu'à définir l'espace BMO à t paramètres. Pour cela, nous avons besoin d'une base d'ondelette à t paramètres. Soit $(w_R^{\vec{\epsilon}})$ une telle base.

définition 16. *Espace BMO à t paramètres*
Soit $b \in L^2(\mathbb{R}^{\vec{d}})$. Définissons la quantité

$$||b||_{BMO} = \sup_{U \subset \mathbb{R}^{\vec{d}} ouvert, |U| < +\infty} \left(\frac{1}{|U|} \sum_{R \subset U, R \in \mathcal{D}_{\vec{d}}, \vec{\epsilon} \in Sig_{\vec{d}}} |\langle b, w_R^{\vec{\epsilon}} \rangle|^2\right)^{\frac{1}{2}}$$

Soit

$$V = \{b \in L^2(\mathbb{R}^{\vec{d}}), ||b||_{BMO} < +\infty\rangle$$

$BMO(\mathbb{R}^{\vec{d}})$ est l'espace vectoriel normé défini comme le complété de V pour la norme $||.||_{BMO}$

Comme dans le cas du paramètre 1, L'espace BMO à t paramètres est en réalité le dual de l'espace de Hardy à t paramètres.

théorème 7. *Chang-Fefferman*
L'espace $BMO(\mathbb{R}^{\vec{d}})$ est le dual topologique de l'espace de Hardy $H^1(\mathbb{R}^{\vec{d}})$:

$$BMO(\mathbb{R}^{\vec{d}}) = H^1(\mathbb{R}^{\vec{d}})^*$$

0.5 Espace BMO, opérateurs de Calderón-Zygmund et commutateurs

Nous abordons à présent le but de ma thèse. Celle-ci consiste à déterminer les familles finies d'opérateurs de Calderón-Zygmund qui, par le biais de commutateurs, caractérisent BMO à t paramètres.
Soit $(\mathcal{T}_s)_{1 \leq s \leq t}$ une famille finie d'opérateurs de Calderón-Zygmund tels que si $T_s \in \mathcal{T}_s, 1 \leq s \leq t$, T_s agit sur le paramètres s de $L^2(\mathbb{R}^{\vec{d}})$. Le but est d'étudier l'inégalité dans le cas du multiparamètre :

$$C_1||b||_{BMO} \leq \sup_{T_s \in \mathcal{T}_s, 1 \leq s \leq t} ||[T_1, [T_2, ..., [T_t, M_b]...]]||_2 \leq C_2||b||_{BMO}$$

avec M_b l'opérateur de multiplication par b, et le commutateur $[A, B]$ défini par $[A, B] = AB - BA$.
Je vais à présent rappeler les théorèmes qui ont auparavant tenté de répondre à ces questions. Les premiers théorèmes concernent les espaces BMO classiques à 1 paramètre, le premier étant celui de Coifman, Rochberg et Weiss ([5]) :

théorème 8. *Coifman, Rochberg, Weiss*
Soit $b \in BMO(\mathbb{R}^d)$, T un opérateur de Calderón-Zygmund de première génération, alors $[T, M_b]$ est un opérateur borné de $L^2(\mathbb{R}^d)$ et :

$$||[T, M_b]||_2 \leq C_T||b||_{BMO}$$

Réciproquement, si pour $1 \leq j \leq d$, $[R_j, M_b]$ *est un opérateur borné de* $L^2(\mathbb{R}^d)$ *alors* $b \in BMO(\mathbb{R}^d)$ *et*

$$||b||_{BMO} \leq A \sum_{1 \leq j \leq d} ||[R_j, M_b]||_2$$

Par conséquent les transformées de Riesz caractérisent $BMO(\mathbb{R}^d)$ *et nous avons l'inégalité :*

$$C_1||b||_{BMO} \leq \sup_{1 \leq j \leq d} ||[R_j, M_b]||_2 \leq C_2||b||_{BMO}$$

avec $C_1, C_2 > 0$

Le problème à présent est de déterminer si d'autres opérateurs de Calderón-Zygmund que les transformées de Riesz caractérisent BMO.
Uchiyama a en partie répondu à cette question en fournissant un critère permettant de savoir si une famille d'opérateurs de Calderón-Zygmund caractérise BMO. En réalité il a étendu le théorème de décomposition de Stein-Fefferman à une plus large classe d'opérateurs de Calderón-Zygmund ([27]).

théorème 9. *Uchiyama*
Soit $(\theta_j)_{1 \leq j \leq n} \in (\mathcal{C}^\infty(\mathbb{R}^d \setminus \{0\}))^n$ *des fonctions homogènes de degré 0.*
Définissons pour $1 \leq j \leq n$ *l'opérateur* T_j *par*

$$\forall \xi \in \mathbb{R}^d p.p, \widehat{T_j f}(\xi) = \theta_j(\xi)\hat{f}(\xi)$$

Supposons en outre que la famille $(\theta_j)_{1 \leq j \leq n}$ *vérifie le critère de Uchiyama, c'est-à-dire :*

$$\forall \xi \in \mathbb{S}^{d-1}, rang(\begin{pmatrix} \theta_1(\xi) & ... & \theta_n(\xi) \\ \theta_1(-\xi) & ... & \theta_n(-\xi) \end{pmatrix}) = 2$$

Alors la famille $(T_j)_{1 \leq j \leq n}$ *caratérise* $BMO(\mathbb{R}^d)$, *c'est-à-dire :*

$$\forall b \in BMO(\mathbb{R}^d), b \text{ à support compact}, \exists g_1, ...g_n \in L^\infty(\mathbb{R}^d), b = \sum_{j=1}^n T_j g_j$$

Les transformées de Riesz auxquelles on rajoute l'identité vérifie le critère de Uchiyama. On a donc bien une généralisation du théorème de décomposition de Stein-Fefferman.

Le dernier théorème qui concerne la caractérisation de l'espace BMO à un paramètre est le théorème de Li ([17]).

théorème 10. *Li*
Soit $(\theta_j)_{1 \leq j \leq n} \in (\mathcal{C}^\infty(\mathbb{R}^d \setminus \{0\}))^n$ *des fonctions homogènes de degré 0.*
Définissons pour $1 \leq j \leq n$ *l'opérateur* T_j *par*

$$\forall \xi \in \mathbb{R}^d a.e, \widehat{T_j f}(\xi) = \theta_j(\xi)\hat{f}(\xi)$$

Supposons en outre que la famille $(\theta_j)_{1\leq j\leq n}$ *vérifie le critère de Li, c'est-à-dire :* $\exists\psi : \mathbb{S}^{d-1} \to \mathbb{S}^{d-1}$ *continue,* $\exists\delta_0 > 0$,

$$\forall\xi \in \mathbb{S}^{d-1}, \sum_{j=1}^{n} |\theta_j(\xi) - \theta_j(\psi(\xi))|^2 \geq \delta_0$$

Alors, si $b \in L^2(\mathbb{R}^d)$ *et si* $1 < p < \infty$, *nous avons les équivalences suivantes :*
(i) $b \in BMO(\mathbb{R}^d)$
(ii) $[T_j, M_b]$ *borné sur* $L^p(\mathbb{R}^d)$, $1 \leq j \leq n$
(iii) $[T_j, M_b]$ *borné sur* $L^q(\mathbb{R}^d)$, $1 \leq j \leq n, 1 < q < \infty$
et alors la famille $(T_j)_{1\leq j\leq n}$ *caractérise BMO*

L'hypothèse de Li généralise celle de Uchiyama, car, si $(\theta_j)_{1\leq j\leq n}$ vérifie l'hypothèse de Uchiyama, cette famille de fonctions vérifie celle de Li avec $\psi(\xi) = -\xi$.

Le cas du multiparamètre est plus récent : le théorème sur lequel nous nous baserons, noté LPPW ([14]), est dû à M. Lacey, S. Petermichl, J. Pipher, B. Wick. Ceux-ci ont démontré que les transformées de Riesz caractérisaient $BMO(\mathbb{R}^{\vec{d}})$ ([14])

théorème 11. *Lacey, Petermichl, Pipher, Wick*
Les transformées de Riesz caractérisent l'espace $BMO(\mathbb{R}^{\vec{d}})$ *à t paramètres et il existe* $C_1, C_2 > 0$ *tels que nous ayons l'inégalité :*

$$C_1||b||_{BMO} \leq \sup_{\vec{0}\leq\vec{j}\leq\vec{d}} ||[R_{1,j_1}[...[R_{t,j_t}, M_b]...]||_2 \leq C_2||b||_{BMO}$$

A présent nous allons scinder le problème en deux : d'une part considerer l'inégalité

$$C_1||b||_{BMO} \leq \max_{T_s\in\mathcal{T}_s, 1\leq s\leq t} ||[T_1, [...[T_t, M_b]...]||_2$$

appelée borne inférieure, et d'autre part l'inégalité

$$||[T_1, [...[T_t, M_b]...]||_2 \leq C_2||b||_{BMO}$$

appelée borne supérieure.

En ce qui concerne la borne inférieure, Stefanie Petermichl et moi-même avons tenté d'étendre le critère de Li au multiparamètre. Malheureusement nous avons échoué, car nous voulions utiliser le théorème de Stone-Weierstrass et utiliser l'idée de [14], mais malheureusement nous n'avions pas un assez bon contrôle des dérivées sur les multipliers de Fourier pour arriver à nos fins. En revanche, par le théorème de Nachbin, qui est un dérivé du théorème de Stone-Weiestrass, nous avions alors ce contrôle des dérivées désiré, et nous en avons déduit un nouveau critère autre que celui de Li.

Voici ce que nous avons démontré :

Pour chaque $1 \leq s \leq t$, nous avons une collection d'opérateurs de Calderón-Zygmund $\mathcal{T}_s = \{T_{s,1}, ..., T_{s,n_s}\}$, dont les noyaux sont homogènes de degré $-d_s$, de symboles $\Theta_s = \{\theta_{s,k_s} \in \mathcal{C}^\infty(\mathcal{S}^{d_s-1}) : 1 \leq k_s \leq n_s\}$ homogènes de degré 0. A une fonction f et un symbole b est associé la suite de commutateurs itérés

$$C_{\vec{k}}(b, \cdot) = [T_{1,k_1}[...[T_{t,k_t}, M_b]...]].$$

Ici $1 \leq s \leq t, \vec{k} = (k_1, ..., k_t), 0 \leq k_s \leq n_s$ et T_{s,k_s} est un opérateur de Calderón-Zygmund appartenant à la famille $\mathcal{T}_s$ agissant sur le paramètre s.

Nous imposons les restrictions suivantes sur les classes $\mathcal{T}_s$ pour chaque paramètre s séparément, plus facilement exprimés à l'aide de leur symbole par :

– $\forall x \neq y \in \mathbb{S}^{d_s-1} \ \exists \ \theta_{s,i}$ tel que $\theta_{s,i}(x) \neq \theta_{s,i}(y)$
  (séparation des points sur la sphère)
– $\forall \ x \in \mathbb{S}^{d_s-1} \ \forall t$ tangent à $\mathbb{S}^{d_s-1}$ en $x \ \exists \ i$ tel que $\frac{\partial \theta_{s,i}}{\partial t}(x) \neq 0$
  (existence de dérivées tangentielles non triviales)

Si K n'est pas à valeurs réelles, nous imposons la condition supplémentaire suivante :

– Θ_s est fermé par conjugaison complexe.

Les ensembles $\mathcal{T}_s$ infinis vérifient également notre théorème .

Alors, nous avons les théorèmes :

théorème 12. *Sous les conditions précédentes sur les classes $\mathcal{T}_s$, il existe des constantes $C_1, C_2 > 0$ telles que* $\forall b \in \mathrm{BMO}(\mathbb{R}^{\vec{d}})$

$$C_1 ||b||_{BMO} \leq \sup_{0 \leq k_s \leq n_s} ||[T_{1,k_1}[...[T_{t,k_t}, M_b]...]]||_2 \leq C_2 ||b||_{BMO}$$

où la norme BMO doit être comprise au sens de Chang et Fefferman. T_{s,k_s} est le $k_s^{ème}$ choix d'opérateur de Calderón-Zygmund dans la famille $\mathcal{T}_s$ agissant sur le paramètre s.

Comme nous le savons, ce type de théorème admet une formulation équivalente en terme de factorisation faible d'espace de Hardy. Si $\vec{k}$ est un vecteur avec $1 \leq k_s \leq d_s$ et $1 \leq s \leq t$, définissons $\Pi_{\vec{k}}$ l'opérateur bilinéaire à partir des commutateurs itérés par :

$$\langle C_{\vec{k}}(b, f), g \rangle_{L^2} = \langle b, \Pi_{\vec{k}}(f, g) \rangle_{L^2}.$$

L'opérateur $\Pi_{\vec{k}}$ peut être exprimé comme la combinaison linéaire d'itérés d'opérateurs de Calderón-Zygmund T_{s,k_s} (et de leurs adjoints), appliqués à f et g.

Utilisons la notation

$$\|f\|_{L^2 * L^2} = \inf \left\{ \sum_{\vec{k}} \sum_j \|\phi_j^{\vec{k}}\|_2 \|\psi_j^{\vec{k}}\|_2 \right\}$$

où l'infimum décrit toutes les décompositions possibles $f = \sum_{\vec{k}} \sum_j \Pi_{\vec{k}}(\phi_j^{\vec{k}}, \psi_j^{\vec{k}})$. En raisonnant par dualité, nous obtenons le théorème :

théorème 13. *Nous avons $H^1(\mathbb{R}^{\vec{d}}) = L^2 * L^2$. Pour tout $f \in H^1(\mathbb{R}^{\vec{d}})$ il existe deux suites $\phi_j^{\vec{k}}, \psi_j^{\vec{k}} \in L^2$ telle que $f = \sum_{\vec{k}} \sum_j \Pi_{\vec{k}}(\phi_j^{\vec{j}}, \psi_j^{\vec{k}})$ avec $\|f\|_{H^1} \sim \sum_{\vec{k}} \sum_j \|\phi_j^{\vec{k}}\|_2 \|\psi_j^{\vec{k}}\|_2$.*

Quant au problème de la borne supérieure, Yumeng Ou et moi-même avons obtenu un résultat optimal, car nous avons montré que l'inégalité était vrai pour n'importe quel type d'opérateur de Calderón-Zygmund vérifiant la deuxième définition que j'ai donnée. Pour cela, nous avons utilisé le théorème de représentation de Hytonen qui permet de décomposer n'importe quel opérateur de Calderon-Zygmund comme une moyenne de Shifts dyadiques. L'étape suivante consiste à décomposer le commutateur en somme de paraproduits, et à utiliser les propriété de ceux-ci.

Voici le théorème et le corollaire que nous avons démontré :

théorème 14. *Soit $b \in BMO(\mathbb{R}^{\vec{d}})$ et $(T_i)_{1 \leq i \leq t}$ une collection d'opérateurs de Calderón-Zygmund, avec chaque T_i agissant sur le paramètre i de $\mathbb{R}^{\vec{d}} = \mathbb{R}^{d_1} \otimes \cdots \otimes \mathbb{R}^{d_t}$. Alors,*

$$\|[\ldots[[M_b, T_1], T_2] \ldots, T_t]\|_{L^2 \to L^2} \leq C \|b\|_{BMO}$$

où C ne dépend que de $\vec{d}$ et de $\prod_{i=1}^t \|T_i\|_{CZ}$.

corollary 1. *Soit $(T_{i,s_i})_{1 \leq i \leq t, 1 \leq s_i \leq n_i}$ une famille d'opérateurs de Calderón-Zygmund caractérisant l'espace $BMO_{prod}(\mathbb{R}^{\vec{d}})$, c'est-à-dire, $\exists C_1, C_2 > 0$, tel que*

$$C_1 \|b\|_{BMO} \leq \sup_{1 \leq i \leq t, 1 \leq s_i \leq n_i} \|[\ldots[[M_b, T_{1,s_1}], T_{2,s_2}] \ldots, T_{t,s_t}]\|_{L^2 \to L^2} \leq C_2 \|b\|_{BMO}.$$

Alors, $\exists \epsilon > 0$ telle que pour toute famille d'opérateurs de Calderón-Zygmund $(T'_{i,s_i})_{1 \leq i \leq t, 1 \leq s_i \leq n_i}$ vérifiant $\|T'_{i,s_i}\|_{CZ} \leq \epsilon$, la famille $(T_{i,s_i} + T'_{i,s_i})_{1 \leq i \leq t, 1 \leq s_i \leq n_i}$ caractérise encore $BMO(\mathbb{R}^{\vec{d}})$.
En particulier, pour toute famille d'opérateurs de Calderón-Zygmund $(T'_{i,s_i})_{1 \leq i \leq t, 1 \leq s_i \leq n_i}$, il existe $\epsilon_1, \ldots, \epsilon_t > 0$ telle que pour tout $0 < c_i < \epsilon_i$, $1 \leq i \leq t$, la famille $(T_{i,s_i} + c_i T'_{i,s_i})_{1 \leq i \leq t, 1 \leq s_i \leq n_i}$ caractérise $BMO(\mathbb{R}^{\vec{d}})$.

Table des matières

Chapitre 1

La borne inférieure

1.1 Introduction

A classical result of Nehari [23] shows that a Hankel operator with antianalytic symbol b is bounded if and only if the symbol belongs to BMO. This theorem has an equivalent formulation by means of commutators of a symbol function b and the Hilbert transform, as the latter are a combination of orthogonal Hankel operators. Nehari's result leans on analytic structure in several crucial ways : the classical factorization result for H^1 functions on the disk and the fact that the Hilbert transform is a Fourier projection operator.

The classical text of Coifman, Rochberg and Weiss [5] extended the one-parameter theory to real analysis in the sense that the Hilbert transforms were replaced by Riesz transforms. In their text, they obtained sufficiency, i.e. that a BMO symbol b yields an L^2 bounded commutator for certain more general, convolution type singular integral operators. For necessity, they showed that the collection of Riesz transforms was representative enough. This is quite natural, in the view of the definition of H^1 requiring Riesz transforms being back in L^1 as well as the Fefferman-Stein decomposition of BMO using Riesz kernels.

Uchiyama [27] revisited said decomposition, with a very technical but constructive proof. It remarkably replaced the class of Riesz transforms by more general classes of kernel operators obeying a certain point separation criterion for their Fourier multiplier symbols. See also [28] and [26] for natural questions in this direction. Li [17] used a criterion similar to Uchiyama's, to show that it was also a sufficiently representative class to characterize BMO by means of commutators.

All of these results date back to the 70s, 80s and 90s and consider H^1 spaces in one parameter and simple, i.e. non-iterated commutators.

It is well known that the product theory and with it the product BMO space, as identified by Chang and Fefferman [3], [4] have more complicated structure. We remind of Carleson's interesting example [2] illustrating this difference. The techniques to tackle the analogs of the above questions in several parameters are very different and have brought, with the works of Lacey and his collaborators, valuable new insight and use to existing theories, for example in the interpretation of Journé's lemma in combination with Carleson's example.

Ferguson and Lacey proved in [6] that the iterated commutator of the Hilbert transform and multiplication by a symbol b characterize BMO, and with it, they proved the equivalent weak factorization result for H^1 on the bidisk. Lacey and Terwilliger extended this result to an arbitrary number of iterates in [16], requiring thus, among others, a refinement of Pipher's iterated multi-parameter version of Journé's lemma. The real variable analog, the result of Coifman, Rochberg and Weiss [5] using Riesz transforms instead of Hilbert transforms, was extended to the multi parameter setting in [14]. In this current paper, we extend in part, the direction of Uchiyama and Li to several parameters. We formulate a sufficient condition on a family of Calderón-Zygmund operators, so that their iterated commutators characterize BMO :

For vectors $\vec{d} = (d_1, ..., d_t) \in \mathbb{N}^t$, we consider product spaces

$$\mathbb{R}^{\vec{d}} = \mathbb{R}^{d_1} \times ... \times \mathbb{R}^{d_t}.$$

For each $1 \leq s \leq t$, we have a collection of Calderón-Zygmund operators $\mathcal{T}_s = \{T_{s,1}, ..., T_{s,n_s}\}$, whose kernels are homogeneous of degree $-d_s$, with Fourier multiplier symbols $\Theta_s = \{\theta_{s,k_s} \in \mathcal{C}^\infty(\mathcal{S}^{d_s-1}) : 1 \leq k_s \leq n_s\}$ that are in turn homogeneous of degree 0. For appropriate functions f and symbols b we consider the family of iterated commutators

$$C_{\vec{k}}(b, \cdot) = [T_{1,k_1}[...[T_{t,k_t}, M_b]...]].$$

Here $1 \leq s \leq t, \vec{k} = (k_1, ..., k_t), 0 \leq k_s \leq n_s$ and T_{s,k_s} denotes the k_sth choice of Calderón-Zygmund operator in the family $\mathcal{T}_s$ acting in the sth variable.

We impose the following restrictions on the classes $\mathcal{T}_s$ for each parameter s separately, easiest formulated in terms of their symbols :

– $\forall x \neq y \in \mathbb{S}^{d_s-1}\ \exists\ \theta_{s,i}$ so that $\theta_{s,i}(x) \neq \theta_{s,i}(y)$
  (full point separation on the sphere)
– $\forall\ x \in \mathbb{S}^{d_s-1}\ \forall t$ tangent to $\mathbb{S}^{d_s-1}$ at $x\ \exists\ i$ so that $\frac{\partial \theta_{s,i}}{\partial t}(x) \neq 0$
  (existence of non-trivial tangential derivatives)

In the case that the kernels K are not real valued, it appears that a last condition is needed :

– Θ_s is closed under complex conjugation.

Infinite sets $\mathcal{T}_s$ are also included in our theorem at no additional cost.

example 1. *It is easy to check that the family of Riesz transforms in $\mathbb{R}^{d_s}$ satisfies these properties.*

example 2. *It is also not hard to check that the family of all rotations of any one smooth, dilation and translation invariant Calderón-Zygmund operator T with a discontinuity in 0 of its symbol in any given direction has these properties. Precisely* *we mean an operator T that has a smooth symbol m that is homogeneous of degree zero with the property that there exists $\xi \in \mathcal{S}^{d_s-1}$ such that $m(\xi) \neq m(-\xi)$. Notice that in many cases, such as when we choose T to be the first Riesz transform, a small number of rotations are sufficient to make up a family with the required properties.*

theorem 1. *Under the conditions above on the classes $\mathcal{T}_s$, there exist constants $C_1, C_2 > 0$ so that $\forall b \in \mathrm{BMO}(\mathbb{R}^{\vec{d}})$*

$$C_1||b||_{BMO} \leq \sup_{0 \leq k_s \leq n_s} ||[T_{1,k_1}[...[T_{t,k_t}, M_b]...]]||_2 \leq C_2||b||_{BMO}$$

where we mean the product BMO *norm according to Chang and Fefferman. T_{s,k_s} denotes the k_sth choice of Calderón-Zygmund operator in the family $\mathcal{T}_s$ acting in the sth variable.*

It is well known, that theorems of this form have an equivalent formulation in the language of weak factorization of Hardy spaces. For $\vec{k}$ a vector with $1 \leq k_s \leq d_s$ and $1 \leq s \leq t$, let us denote by $\Pi_{\vec{k}}$ the bilinear operator obtained by unwinding the commutator :

$$\langle C_{\vec{k}}(b, f), g\rangle_{L^2} = \langle b, \Pi_{\vec{k}}(f, g)\rangle_{L^2}.$$

The operator $\Pi_{\vec{k}}$ can be expressed as linear combination of iterates of Calderón-Zygmund operators T_{s,k_s} (and their adjoints), applied to f, g.

Using the notation

$$\|f\|_{L^2 * L^2} = \inf \left\{ \sum_{\vec{k}} \sum_{j} \|\phi_j^{\vec{k}}\|_2 \|\psi_j^{\vec{k}}\|_2 \right\}$$

where the infimum runs over all possible decompositions of $f = \sum_{\vec{k}} \sum_j \Pi_{\vec{k}}(\phi_j^{\vec{k}}, \psi_j^{\vec{k}})$. With the help of the relevant commutator theorem, it is an exercise in duality to see the following :

theorem 2. *We have $H^1(\mathbb{R}^{\vec{d}}) = L^2 * L^2$. For any $f \in H^1(\mathbb{R}^{\vec{d}})$ there exist sequences $\phi_j^{\vec{k}}, \psi_j^{\vec{k}} \in L^2$ such that $f = \sum_{\vec{k}} \sum_j \Pi_{\vec{k}}(\phi_j^{\vec{j}}, \psi_j^{\vec{k}})$ with $\|f\|_{H^1} \sim \sum_{\vec{k}} \sum_j \|\phi_j^{\vec{k}}\|_2 \|\psi_j^{\vec{k}}\|_2$.*

In this text we prefer the language of commutators in terms of upper (sufficiency) and lower (necessity) bounds.

Our proof follows the machinery developed by Lacey and collaborators in [6], [16], [14]. In particular, we refine a strategy from [14], to pass from the complex variable case and the Hilbert transform to the real variable and Riesz transform case. The Fourier multipliers of the Riesz transforms are very special - monomials on the sphere. We establish such a passage for much more general multiplier operators.

It seems not possible to use any previously proved characterization theorems directly. We can however reuse some of the general strategy and in particular, we manage to 'black box' the very technical wavelet support and paraproduct estimates found in different versions in previous works. In [14], this part appears to be the most streamlined and is general enough to apply to our situation.

1.2 A Brief Review of Multi-Parameter Theory

1.2.1 Wavelets in Higher Dimensions and Several Parameters

We will use the following dilation and translation operators on $\mathbb{R}^d$

$$\mathrm{Tr_y} f(x) := f(x-y)\,, \quad y \in \mathbb{R}^{\mathrm{d}}, \tag{1.2.1}$$

$$\mathrm{Dil_a}^{(\mathrm{p})} f(x) := a^{\text{-d/p}} f(x/a)\,, \quad a > 0\,, 0 < p \leq \infty\,. \tag{1.2.2}$$

These will also be applied to sets, in an obvious fashion, in the case of $p=\infty$.

By the *(d-dimensional) dyadic grid* in $\mathbb{R}^d$ we mean the collection of cubes

$$\mathcal{D}_d := \Big\{ j2^k + [0,2^k)^d \mid j \in \mathbb{Z}^d\,, k \in \mathbb{Z} \Big\}\,.$$

An elementary example of a wavelet system is the Haar system generated by $h = -\mathbf{1}_{(0,1/2)} + \mathbf{1}_{(1/2,1)}$ and $W = \mathbf{1}_{(0,1)}$. The principle requirement is that the functions $\{\mathrm{Tr}_{c(I)} \mathrm{Dil}_I^{(2)} w \mid I \in \mathcal{D}_1\}$ form an orthonormal basis for $L^2(\mathbb{R})$.

The wavelet in this text should be thought of *Meyer wavelet* though, due to its extraordinary Fourier support properties. Although not explicit in this text, we borrow certain technical estimates that make decisive use of this feature of the Meyer wavelet.

For $\varepsilon \in \{0,1\}$, set $w^0 = w$ and $w^1 = W$, the superscript ^{0} denoting that 'the function has mean 0', while a superscript ^{1} denotes that 'the function is an L^2 normalized indicator function'. In one dimension, for an interval I, set

$$w_I^{\varepsilon} := \mathrm{Tr}_{c(I)} \mathrm{Dil}_{|I|}^{(2)} w^{\varepsilon}\,.$$

Multiresolution wavelets, such as the Haar or the Meyer wavelet have the useful identity

$$\sum_{I \supsetneq J} \langle f, w_I \rangle w_I = \langle f, w_J^1 \rangle w_J^1\,. \tag{1.2.3}$$

The passage from $\mathbb{R}$ to $\mathbb{R}^d$ consists of a product of d wavelets associated to intervals of the same size, so that the resulting wavelet is associated to a cube.

Let $\sigma_d := \{0,1\}^d - \{\vec{1}\}$, which we refer to as *signatures*. In d dimensions, for a cube Q with side $|I|$, i.e., $Q = I_1 \times \cdots \times I_d$, and a choice of $\varepsilon \in \sigma_d$, set

$$w_Q^{\varepsilon}(x_1,\ldots,x_d) := \prod_{j=1}^d w_{I_j}^{\varepsilon_j}(x_j).$$

It is then the case that the collection of functions

$$\mathrm{Wavelet}_{\mathcal{D}_d} := \{ w_Q^{\varepsilon} \mid Q \in \mathcal{D}_d\,, \varepsilon \in \sigma_d \}$$

form a wavelet basis for $L^p(\mathbb{R}^d)$ for any choice of d dimensional dyadic grid $\mathcal{D}_d$. Here, we are using the notation $\vec{1} = (1, \ldots, 1)$.

The passage to the tensor product setting, $\mathbb{R}^{\vec{d}} = \mathbb{R}^{d_1} \times ... \times \mathbb{R}^{d_t}$ consists of a product of t wavelets associated to cubes of possibly different size, so that the resulting wavelet is associated to a rectangle.

For a vector $\vec{d} = (d_1, \ldots, d_t)$, and $1 \le s \le t$, let $\mathcal{D}_{d_s}$ be a choice of d_s dimensional dyadic grid, and let

$$\mathcal{D}_{\vec{d}} = \otimes_{s=1}^t \mathcal{D}_{d_s} \ .$$

Also, let $\sigma_{\vec{d}} := \{\vec{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_t) : \varepsilon_s \in \sigma_{d_s}\}$. Note that each ε_s is a vector, and so $\vec{\varepsilon}$ is a 'vector of vectors'. For a rectangle $R = Q_1 \times \cdots \times Q_t$, being a product of cubes of possibly different dimensions, and a choice of vectors $\vec{\varepsilon} \in \sigma_{\vec{d}}$ set

$$w_R^{\vec{\varepsilon}}(x_1, \ldots, x_t) = \prod_{s=1}^t w_{Q_s}^{\varepsilon_s}(x_s).$$

These are the appropriate functions and bases to analyze multi-parameter paraproducts and commutators.

So the collection of wavelets associated to a dyadic grid in the product setting $\mathcal{D}_{\vec{d}}$ is

$$\left\{ w_R^{\vec{\varepsilon}} \mid R \in \mathcal{D}_{\vec{d}}\,, \vec{\varepsilon} \in \sigma_{\vec{d}} \right\} \ .$$

This is a basis in $L^p(\mathbb{R}^{\vec{d}})$.

1.2.2 Chang–Fefferman BMO

Let us describe product Hardy space theory. By this, we mean the Hardy spaces associated with domains like $\otimes_{s=1}^t \mathbb{R}^{d_s}$.

The Hardy space $H^1(\mathbb{R}^d)$ denotes the class of functions with the norm

$$\sum_{j=0}^d \|R_j f\|_1$$

where R_j denotes the jth Riesz transform. We adopt the convention that R_0, the 0th Riesz transform, is the identity. This space is invariant under the one parameter family of isotropic dilations, while $H^1(\mathbb{R}^{\vec{d}})$ is invariant under dilations of each coordinate separately. This invariance under a t parameter family of dilations gave rise to the term 'multi-parameter' theory.

The product space $H^1(\mathbb{R}^{\vec{d}})$ has a variety of equivalent norms, in terms of square functions, (strong) maximal functions and Riesz transforms.

The dual of the real Hardy space is

$$H^1(\mathbb{R}^{\vec{d}})^* = \mathrm{BMO}(\mathbb{R}^{\vec{d}}),$$

the t-fold product BMO space. It is a Theorem of Chang and Fefferman [4] that this space has a characterization in terms of a product Carleson measure.

Define

$$\|b\|_{\mathrm{BMO}(\mathbb{R}^{\vec{d}})} := \sup_{U\subset\mathbb{R}^{\vec{d}}} \left[|U|^{-1} \sum_{R\subset U} \sum_{\vec{\varepsilon}\in\sigma_{\vec{d}}} |\langle b, w_R^{\vec{\varepsilon}}\rangle|^2 \right]^{1/2}. \tag{1.2.4}$$

Here the supremum is taken over all open subsets $U \subset \mathbb{R}^{\vec{d}}$ with finite measure, and we use a wavelet basis $w_R^{\vec{\varepsilon}}$.

theorem 3. *(Chang, Fefferman) We have the equivalence of norms*

$$\|b\|_{(H^1(\mathbb{R}^{\vec{d}}))^*} \approx \|b\|_{BMO(\mathbb{R}^{\vec{d}})}$$

That is, $BMO(\mathbb{R}^{\vec{d}})$ is the dual to $H^1(\mathbb{R}^{\vec{d}})$.

Notice that this space BMO is invariant under a t-parameter family of dilations. Here the dilations are isotropic in each parameter separately. This fact is also represented by the choice of our wavelet system.

1.2.3 Journé's Lemma

Notice that the supremum in the wavelet definition of BMO runs over open sets of finite measure. This supremum restricted just to rectangles gives the definition of the larger rectangular BMO. There is a substantial geometric difference : the maximal dyadic sub-rectangles of any arbitrary rectangle are disjoint while those maximal dyadic sub-rectangles in open sets are not necessarily comparable by inclusion. It is in part due to this difference that, in the same way as in [6], a geometric lemma by Journé [12] involving rectangles in the plane, particularly useful in handling collections of rectangles not comparable by inclusion, comes into play. It was first observed by Ferguson and Lacey that Journé's lemma could be improved to partially compare rectangular BMO and product BMO of two parameters.

An n-dimensional version of Journé's original lemma is due to Pipher [24] and makes use of iterations. This is the reason why we are going to have to replace the rectangular BMO space by another version of BMO that allows us to induct on the number of parameters in our commutator and therefore make use of the iterated nature of Journé's lemma in more than two parameters. This idea was first used in [16].

Say that a collection of rectangles $\mathcal{U} \subset \mathcal{D}_{\vec{d}}$ has $t-1$ parameters if and only if there is a choice of coordinate s so that for all $R, R' \in \mathcal{U}$ we have $Q_s = Q'_s$, that is the sth coordinate of the rectangles are all one fixed d_s dimensional cube.

We then define

$$\|f\|_{\mathrm{BMO}_{-1}(\mathbb{R}^{\vec{d}})} = \sup_{\mathcal{U}\text{ has }t-1\text{ parameters}} \left(|\operatorname{sh}(\mathcal{U})|^{-1} \sum_{\vec{\varepsilon}} \sum_{R\in\mathcal{U}} |\langle f, w_R^{\vec{\varepsilon}}\rangle|^2 \right)^{1/2}$$

In this notation, a collection of rectangles has a shadow given by $\mathrm{sh}(\mathcal{U}) = \cup\{R : R \in \mathcal{U}\}$. The -1 subscript is used to indicate that we have 'reduced by one parameter' in the definition. The reader may be more familiar with the rectangular BMO space mentioned above. In two parameters, the space BMO_{-1} is larger than rectangular BMO.

Carleson produced examples of functions which acted as linear functionals on $H^1(\mathbb{R}^{\vec{d}})$ with norm one, yet had arbitrarily small rectangular BMO norm (and hence arbitrarily small BMO_{-1} norm).

Here is the precise version of the above mentioned refinement of Journé's lemma. It permits us, with certain restrictions and by inducing a damping factor, to control the BMO norm by the BMO_{-1}norm.

lemma 1. *Let $\mathcal{U}$ be a collection of rectangles of finite shadow. For any $a > 0$, we can construct $V \supset \mathrm{sh}(\mathcal{U})$ together with a function $E : \mathcal{U} \to [1, \infty]$ so that $E(R) \cdot R \subset V$ for all $R \in \mathcal{U}$, $|V| < (1 + a)|\,\mathrm{sh}(\mathcal{U})|$, and last that*

$$\left\| \sum_{\vec{\varepsilon}} \sum_{R \in \mathcal{U}} E(R)^{-C} \langle b, w_R^{\vec{\varepsilon}} \rangle w_R^{\vec{\varepsilon}} \right\|_{\mathrm{BMO}} \leq K_a \|b\|_{\mathrm{BMO}_{-1}}.$$

Here C depends only on $\vec{d}$ and K_a on a and $\vec{d}$.

A good and more complete reference on the subject is [?].

1.2.4 Remarks on the Upper Bound

We are going to assume that K is a smooth Calderón–Zygmund convolution kernel on $\mathbb{R}^d \times \mathbb{R}^d$. This means that the kernel is a distribution that satisfies the estimates below for $x \neq y$

$$|\nabla^j K(y)| \leq N|y|^{-d-j}\,, \quad j = 0, 1, 2, \ldots, d + 1\,.$$
$$||\widehat{K}||_{L^\infty} \leq N\,.$$

The first estimate combines the standard size and smoothness estimate. The last assumption is equivalent to assuming that the operator defined on Schwartz functions by

$$T_K f(x) = \int_{\mathbb{R}^d} K(x - y) f(y)\, dy$$

extends to a bounded operator on $L^2(\mathbb{R}^d)$.

If $K_1, \ldots, K_t$ is a sequence of Calderón–Zygmund kernels, with K_s defined on $\mathbb{R}^{d_s} \times \mathbb{R}^{d_s}$. It is not obvious that the corresponding tensor product operator

$$T_{K_1} \otimes \cdots \otimes T_{K_t}$$

is a bounded operator on $L^p(\mathbb{R}^{\vec{d}})$. This is a consequence of multi-parameter Calderón–Zygmund theory.

By [14], theorem (5.3), multi-parameter commutators are bounded operators if the symbol belongs to BMO :

theorem 4. *For* $1 < p < \infty$,

$$||[T_{K_1}, ..., [T_{K_t}, M_b]...]||_p \lesssim ||b||_{BMO}$$

By BMO*, we mean Chang–Fefferman* BMO*. The implied constant depends upon the vector* $\vec{d}$*, and the* T_{K_s}*.*

The Calderón-Zygmund operators we are concerned about in this text are assumed to have 'infinite' smoothness in the sense of the estimates on the kernel in the beginning of (1.2.4) and are therefore included in the result above.

The rest of the paper is dedicated to establishing a lower estimate of our commutators by means of product BMO. We are going to follow the iteration strategy in [6], [16], [14]. The one-dimensional case is very special : the Hilbert transform is both Calderón-Zygmund as well as half space Fourier projection operator. We have lost this feature in higher dimensions, but it motivates the use of Calderón-Zygmund operators close to projection operators, such as in [14].

1.3 Cone operators

In dimension $d \geqslant 2$, a cone $C \subset \mathbb{R}^d$ is given by the data (ξ, Q) where $\xi \in \mathbb{R}^d$ is the direction of the cone and the cube $Q \subset \xi^{\perp}$ centered at the origin is its aperture. The cone consists of all vectors θ that take the form $(\theta_\xi \xi, \theta^{\perp})$ where $\theta_\xi = \theta.\xi$ and $\theta^{\perp} \in \theta_\xi Q$. By λC we mean the dilated cone with data $(\xi, \lambda Q)$.

Given a cone C, we consider its Fourier projection operator defined via $\widehat{P_C f} = \mathbf{1}_C \hat{f}$. Due to the fact that the apertures are cubes, such operators are combinations of Fourier projections onto half spaces and as such admit uniform L^p bounds. For a given cone C we consider a smooth Calderón-Zygmund operator T_C with a kernel K_C whose Fourier symbol $\widehat{K_C} \in C^{\infty}$ and satisfies the estimate $\mathbf{1}_C \leqslant \widehat{K_C} \leqslant \mathbf{1}_{(1+\tau)C}$.

remark 1. *The derivatives of the symbols* $\widehat{K_C}$ *increase with the aperture of the cones. In the course of the proof it will be important that the* L^p *bounds of operators* T_C *do not grow with the aperture of the cones. We thank the special nature of the cone operators and their closeness to half plane projections for this fact. By a rotation argument, we may assume that the cone* C *has direction* x_1*. There exists a smoothed symbol* m *of the sort described, so that higher derivatives in consecutive directions* $x_2, \ldots, x_n$ *are controlled independently of the aperture. In the remaining variable,* x_1*, the derivatives grow with the aperture, but we control total variation of the derivatives in* $x_2, \ldots, x_n$*. In doing so and carefully reading the Marcinkiewicz multiplier theorem, it provides us with* L^p *bounds independent of the aperture. The details are left to the interested reader. We refer to [8] page 363 for a detailed statement of the Marcinkiewicz multiplier theorem.*

1.3.1 Selection of a Representative Class of Cones

Following the idea in [14], we select classes of cones that are going to give us a certain auxiliary lower bound. We felt the need to refine this process, which is necessary due to the fact that we consider more general classes of Calderón-Zygmund operators instead of just the class of Riesz transforms.

Let b be our BMO function that we normalize to have norm 1. Let U be the open set that gives us the supremum in the BMO norm of b and denote by $\mathcal{U}$ the collection of rectangles $R \subset U$. Let us renormalize, by an appropriate dilation, the size of the set $\mathrm{sh}(\mathcal{U})$ to be comparable to 1. Let $\beta = P_{\mathcal{U}} b$, the wavelet projection onto those wavelets adapted to rectangles in the class $\mathcal{U}$.

Given a cone C, with data (ξ, Q), we denote by H_C the half plane projection that corresponds to the direction ξ, the convolution operator whose symbol is $\chi_{(0,\infty)}(\xi \cdot \theta)$. Recall that T_C denotes the Calderón-Zygmund operator adapted to the cone and P_C the Fourier projection associated to the cone. Given a vector of cones $\vec{C} = (C_s)_{1\leq s\leq t}$ we denote by $H_{\vec{C}}, T_{\vec{C}}, P_{\vec{C}}$ their tensor products.

lemma 2. *Let b be the set of all* BMO *functions normalized as above. For all such b, let $U, \mathcal{U}, \beta$ be as above. For any $\kappa > 0$ we can select a finite set of pairs $(\vec{D}, \vec{C})$of vectors of cones $\vec{D} = (D_s)_{1\leq s\leq t}$ where $D_s \subset \mathbb{R}^{d_s}$ with data (ξ_s, Q_s) and $C_s \subset \mathbb{R}^{d_s}, 1 \leq s \leq t$ with data (ξ'_s, Q'_s) so that for each β there is a pair $(\vec{D}, \vec{C})$ with the following properties.*

1. $D_s \subset C_s$
2. $\|T_{\vec{D}}\beta\|_2 \geq 4^{-t}$
3. $\|(H_{\vec{D}} - T_{\vec{D}})\beta\|_4 \leq \kappa$
4. $\|(H_{\vec{C}} - P_{\vec{C}})|T_{\vec{D}}\beta|^2\|_2 \leq \kappa$

Proof. We first select a finite collection of cones D_s. Let us for the moment fix b. Let η be a small positive number to be determined later. It will be in relation with the aperture of the cones : given η, the aperture Q_s is chosen large enough so that

$$\mathbb{P}(D_s \cap \mathbb{S}^{d_s-1} | \mathbb{S}^{d_s-1}) \geq \frac{1}{2} - \eta.$$

We consider random rotations $D_s^{\phi_s}$ of D_s and write $\vec{D}^{\phi}$ for component-wise independent rotation.

Averaging the L^2 norms gives us

$$\mathbb{E}(\|P_{\vec{D}^{\phi}}\beta\|_2^2) = \mathbb{E}(\int_{\vec{D}^{\phi}} |\hat{\beta}(\xi)|^2 d\xi) \geq (\frac{1}{2} - \eta)^t$$

as well as

$$\mathbb{E}(\|(H_{\vec{D}^{\phi}} - P_{\vec{D}^{\phi}})\beta\|_2^2) \leq \eta^t.$$

Notice that for all choices of ϕ, we have

$$0 \leq \|T_{\vec{D}^{\phi}}\beta\|_2 \leq 1$$

as well as

$$0 \leq \|(H_{\vec{D}\phi} - T_{\vec{D}\phi})\beta\|_2 \leq 1.$$

Together, this provides us with the estimates

$$\mathbb{P}(\|T_{\vec{D}\phi}\beta\|_2 \geq 4^{-t}) \geq \frac{4^t(\frac{1}{2}-\eta)^t - 1}{4^t - 1}$$

and

$$\mathbb{P}(\|(H_{\vec{D}\phi} - T_{\vec{D}\phi})\beta\|_2 \geq \eta^{-t}) \leq \eta^{\frac{t}{2}}.$$

Since $\lim_{\eta\to 0} \frac{4^t(\frac{1}{2}-\eta)^t-1}{4^t-1} = \frac{1}{2^t-1}$ and $\lim_{\eta\to 0} 1 - \eta^{\frac{t}{2}} = 1$, the sum of the above probabilities exceeds 1 for small enough η. In this case we are sure to be able to select directions so that

$$\|T_{\vec{D}\phi}\beta\|_2 \geq 4^{-t}$$

and

$$\|(H_{\vec{D}\phi} - T_{\vec{D}\phi})\beta\|_2 \leq \eta^{-t/2}.$$

We have to preserve the smallness of the latter estimate when passing to the L^4 norm. We have half plane projection operators H_D and Calderón-Zygmund operators T_D that have, according to remark 1 above, uniform L^p bounds. It is essential that the L^p norms do not grow when $\eta \to 0$. Recall that small η induce large aperture for the cones D. Also remember that β is normalized in L^2 as well as in BMO. We therefore have uniform L^8 bounds of the following : $\|(H_{\vec{D}\phi} - T_{\vec{D}\phi})\beta\|_8 \leq K$ where the constant K neither depends on the aperture nor the direction of the cones.

By interpolation we get $\|(H_{\vec{D}\phi} - T_{\vec{D}\phi})\beta\|_4 \lesssim \eta^{-t/6}$. We choose η small enough so that both the above inequalities hold as well as $\eta^{-t/6} < \kappa$.

We have seen that there exists a fixed η so that for each b the set $b(\eta) \subset \mathbb{S}^{d-1}$ of admissible directions ξ is not empty. Notice that $b(\eta) \subset b(\eta/2)$. Furthermore, there exists $r(\eta)$ so that the ball $B(\xi, r(\eta)) \cap \mathbb{S}^{d-1} \subset b(\eta/2)$ for all $\xi \in b(\eta)$. So by increasing the aperture, a dense enough finite sample set of directions will therefore provide an admissible direction for all appropriately normalized BMO functions b.

We turn to the selection of cones C_s, keeping in mind that cones D_s have already been chosen. Due to uniform L^4 estimates of $T_{\vec{D}\phi}$ we see that $\||T_{\vec{D}}\beta|^2\|_2 \leq K$ for some universal K. So, in particular, for any vector of cones $\vec{C}$, we have $\|(H_{\vec{C}} - P_{\vec{C}})|T_{\vec{D}}\beta|^2\|_2 \leq K$.

Take $\varsigma < \eta/2$ a small positive number. Choosing the aperture of the cones C_s large enough so that

$$\mathbb{P}(C_s \cap \mathbb{S}^{d_s-1}|\mathbb{S}^{d_s-1}) \geq \frac{1}{2} - \varsigma$$

gives us the estimate

$$\mathbb{E}\|(H_{\vec{C}\phi} - P_{\vec{C}\phi})|T_{\vec{D}}\beta|^2\|_2 \leq K\varsigma^t$$

Similarly to above,

$$\mathbb{P}(\|(H_{\vec{C}\phi} - P_{\vec{C}\phi})|T_{\vec{D}}\beta|^2\|_2 \geq K\varsigma^{t/2}) \leq \varsigma^{t/2}$$

If $D_s = (\xi_s, Q_s)$ let E_{ξ_s} be the hyperplane perpendicular to ξ_s and H_{ξ_s} the corresponding half space that contains D_s. Let $\alpha = \min\{\angle(\xi_1, \xi_2) : \xi_1 \in D_s, \xi_2 \in E_{\xi_s}\}$ where $\angle$ denotes the angle between vectors. Notice that α only depends upon η. Consider now the circular cone $A_{\xi_s} = \{\xi : \angle(\xi, \xi_s) < \alpha/4\}$. There exists a fixed larger aperture Q'_s, only depending on α so that $(\xi, Q'_s) \supset (\xi_s, Q_s)$ whenever $\xi \in D_{\xi_s}$. We are free to choose ς small enough so that

$$\mathbb{P}(A_{\xi_s} \cap \mathbb{S}^{d_s-1} | \mathbb{S}^{d_s-1}) \geq \varsigma^{1/2}$$

as well as $K\varsigma^{t/2} < \kappa$. Since

$$\mathbb{P}(\|(H_{\vec{C}\phi} - P_{\vec{C}\phi})|T_{\vec{D}}\beta|^2\|_2 \geq K\varsigma^{t/2}) \leq \varsigma^{t/2}$$

we are sure to find $C_s = (\xi'_s, Q'_s)$ with the required properties.
□

By slightly enlarging the aperture of cones C_s and an argument similar to the one above, we obtain a finite collection of cones C_s with the required properties.

We form commutators using arbitrary cones $C_s = (\xi_s, Q_s)$. Let us define

$$\|b\|_{\vec{Q}} = \sup \|[T_{C_1}, ...[T_{C_t}, M_b]...]\|_{2\to 2}$$

where the supremum is taken over all choices of cone transforms $T_{C_s} = T_{(\xi_s, Q_s)}$ in which the direction ξ_s varies and the aperture of the cone is fixed to be Q_s for each parameter s separately. Here T_{C_s} acts in the sth variable. In [14] the following theorem was proven :

theorem 5. $\|b\|_{\vec{Q}} \sim \|b\|_{\mathrm{BMO}}$ *with constants depending upon the aperture of the cones.*

We are going to need information that is somewhat more specific. It is valuable to us to know for which test function, depending on the symbol b, the commutator becomes large.

lemma 3. *If* $\gamma = T_{\vec{D}}\beta$ *with cones* $\vec{D}, \vec{C}$ *chosen as in the lemma, then*

$$\|[T_{C_1}, ...[T_{C_t}, M_b]...]\bar{\gamma}\|_2 \gtrsim 1.$$

The proof of a similar estimate is implicit in [14], section 7. Although the cones in our text have somewhat different properties (D_s and C_s do not necessarily share the same direction), the pairs (D_s, C_s) were chosen to enable the use of the proof in [14]. We sketch the part of the proof that illustrates the special use of the cone operators.

Proof For a fixed, small δ_{-1} to be chosen, we start with a BMO function b so that $\|b\|_{\mathrm{BMO}} < \delta_{-1}$ is small. Let U be the supremal set in the definition of BMO and

$\mathcal{U}$ the corresponding collection of dyadic rectangles with its shadow $sh(\mathcal{U})$. Journé's lemma provides us with a slightly larger set V. Let $\mathcal{V} = \{R : R \subset V, R \not\subset sh(\mathcal{U})\}$. Let $\mathcal{W}$ denote the rest of the dyadic rectangles. We use the collections $\mathcal{U}, \mathcal{V}$ and $\mathcal{W}$ to split our symbol $b = P_{\mathcal{U}}b + P_{\mathcal{V}}b + P_{\mathcal{W}}b$. By linearity we obtain three commutators tested on $\bar{\gamma}$. We will see that only the commutator with symbol $P_{\mathcal{U}}b$ is large and the other two are negligible error terms.

We first observe that with $\beta = P_{\mathcal{U}}b$ and $\gamma = T_{\vec{D}}\beta$, we have $\|[T_{C_1}, ...[T_{C_t}, M_\beta]...]\bar{\gamma}\|_2 \gtrsim 1$. Observe that the only non-zero term in this commutator is $T_{C_1}...T_{C_t}(P_{\mathcal{U}}b)\bar{\gamma}$ since any cone operator falling on $\bar{\gamma}$ is zero. Consider now the splitting

$$T_{\vec{C}}(\gamma + (H_{\vec{C}} - T_{\vec{D}})\beta + (I - H_{\vec{C}})\beta)\bar{\gamma}.$$

The last term is zero since $(I - H_{\vec{C}})\beta$ and $\bar{\gamma}$ are supported on the same half space away from the cones $\vec{C}$. The second term is small due to the choice of the cone in lemma (2). The first term is large and explains the motivation using cone transforms :

$$||T_{\vec{C}}(\gamma.\bar{\gamma})||_2 + \kappa \geq ||H_{\vec{C}}(\gamma.\bar{\gamma})||_2 \gtrsim ||\gamma.\bar{\gamma}||_2 = ||\gamma||_4^2 \gtrsim 1$$

This follows as the Fourier transform of $\overline{\gamma} \cdot \gamma$ is symmetric with respect to the half planes determined by the cones ; the last inequality uses the Littlewood-Paley inequalities.

Next, we will see that $\|[T_{C_1}, ...[T_{C_t}, M_{P_{\mathcal{V}}b}]...]\bar{\gamma}\|_2 \lesssim \delta_J^{1/4}$. It is easy to see that

$$||[T_{C_1}, ..., [T_{C_t}, M_{\mathcal{P}_{\mathcal{V}}b}]...]\bar{\gamma}||_2 \lesssim ||\mathcal{P}_{\mathcal{V}}b||_4 ||\gamma||_4 \lesssim ||\mathcal{P}_{\mathcal{V}}b||_4$$

where the implied constant depends upon the L^4 norms of the Cone transforms. But, by Journé's lemma, we have that

$$||\mathcal{P}_{\mathcal{V}}b||_2 \leq \delta_J^{\frac{1}{2}}, \qquad ||\mathcal{P}_{\mathcal{V}}b||_{BMO} \leq 1 .$$

Together they imply

$$||\mathcal{P}_{\mathcal{V}}b||_4 \leq \delta_J^{\frac{1}{4}}.$$

For the technical estimate of the last term $\|[T_{C_1}, ...[T_{C_t}, M_{P_{\mathcal{W}}b}]...]\bar{\gamma}\|_2 \lesssim K_J\delta_{-1}$ we refer to [14] section 7, proof of (7.9). Here K_J depends upon the constant δ_J. □

We gather the information and are left with the following :

theorem 6. *For each parameter s there exists a finite collection $\mathcal{C}_s$ of cones $C_{s,k_s} = (\xi_{k_s}, Q_s)$ with $1 \leqslant k_s \leqslant n_s$ of fixed aperture Q_s so that*

$$\|b\|_{BMO} \lesssim \sup \|[T_{C_{1,k_1}}, ...[T_{C_{t,k_t}}, M_b]...]\|_{2\to 2} \lesssim \|b\|_{BMO}$$

for all BMO functions b. Here the supremum runs over all $C_{s,k_s} \in \mathcal{C}_s$.

It will be essential for us to approximate symbols of cone operators using polynomials in members of our given collections of symbols.

1.3.2 Approximation of Cones via the Family Θ

For a fixed parameter, given our family Θ, we wish to approximate the symbol of cone projection operators by means of polynomials in θ_i. For technical reasons, we need a very good approximation that controls also the supremum norm of derivatives of the symbols, say of order d. On one hand, we require the resulting approximations to be Calderón-Zygmund operators with enough additional smoothness on the kernel. This is a necessary requirement to control error terms that arise from the commutator with symbol $\mathrm{P}_{\mathcal{W}}\, b$. This error estimate is not carried out in this text and can be found in [14], section 7, where the added smoothness is crucial. On the other hand, we need the L^p estimates of the approximations to stay controlled when $\epsilon \to 0$.

Nachbin's beautiful theorem [22] allows us, under certain conditions on the family, to do so. We state it in the form we are going to need.

theorem 7. *Let $\mathfrak{M}$ be a compact smooth manifold. Let B be a closed real subalgebra of $A = (C^m(\mathfrak{M}), \tau_m)$ where τ_m is the topology induced by the norm of uniform convergence in C^m. Then $B = A$ if and only if B contains the function 1, $\forall x \neq y \in \mathfrak{M} \exists f \in B$ such that $f(x) \neq f(y)$ and for every $x \in \mathfrak{M}$ and $0 \neq v \in T_x(\mathfrak{M})$ there exists $f \in B$ such that $\mathrm{df}(x)(v) \neq 0$.*

It is not hard to check that under the additional assumtion that B be closed under complex conjugation, there is a complex version.

lemma 4. *For a given d-dimensional pair of cones D and C as in lemma (2), let $H_{-\xi_C}, H_{-\xi_D}$ denote the opposing half spaces, respectively. Choose a function $h_{C,D} \in C^d(\mathbb{S}^{d-1})$ with values between 0 and 1 such that*

- $h_{C,D}(\xi) = 1 \forall \xi \in C$
- $h_{C,D}(\xi) = 0 \forall \xi \in H_{-\xi_C} \cup H_{-\xi_D}$.

Given any small $\epsilon > 0$, there exists an operator $F_{C,D}$ with symbol $v_{C,D}$, that is a polynomial in $\theta \in \Theta$ so that $\|v_{C,D} - h_{C,D}\|_{\tau_d} < \epsilon$, where $\|.\|_{\tau_d}$ is the norm of uniform convergence in C^d. We have universal L^p estimates for the associated kernel operators $F_{C,D} : \|F_{C,D}\|_p \lesssim K_p$ where this constant is independent of the choice of the cone and universal for small ϵ.

Proof. Thanks to our assumptions, the part concerning the approximations is almost clear. Just observe that we may add the identity operator I with multiplier 1 to our collection. That is the collection Θ characterizes BMO if and only if $\Theta \cup \{1\}$ does. In the case that the kernels are real valued, we did not assume that Θ be closed under complex conjugation. In this case, consider $\Theta \cup \bar{\Theta}$ characterizes BMO if and only if Θ does. Observe that if T_θ denotes the Calderón-Zygmund operator associated to the symbol θ, then $T^*_\theta = T_{\bar{\theta}}$. Observe also that $[T, b] = [T^*, \bar{b}]^*$. If the kernel $K(x)$ of T is real, then $K(-x)$ is the kernel of T^*. It is easy to verify that

$$[T_1, [T_2^*, b]]f = [T_1[T_2, b^{(\cdot,-)}]]f^{(\cdot,-)}.$$

Here $f^{(\cdot,-)}(x,y) = f(x,-y)$ so f has a sign change in the second set of variables. Its obvious generalization holds when more iterates and adjoints are present. The BMO and L^2 norms are preserved under these reflections.

It remains the important point of universal L^p estimates . Thanks to the control on the derivatives granted to us by Nachbin's theorem, we may apply a standard multiplier theorem to obtain uniform L^p bounds.

□

1.4 Lower bound, Calderón-Zygmund operators

We induct on the number t of parameters, that is the number of coordinates in $\vec{d} = (d_1, \ldots, d_t)$.We assume that $d_s \geqslant 2$ for all s. The case when $d_s = 1$ for some s reduces our choices of admissible operators to the Hilbert transform. This case is easier and merely complicates notation for us.

The base case $t = 1$ of our induction argument is stronger than what we need and a theorem by Li :

theorem 8. *Let $\mathcal{T}$ be a collection of Calderón-Zygmund operators, where the following restriction is imposed : the symbols θ_i of the $T_i \in \mathcal{T}$ are infinitely smooth and satisfy $\sum |\theta_i(x) - \theta_i(-x)| \neq 0$ for all $x \in \mathbb{S}^{d-1}$.*

In the case of $t = 1$ for all $d \geqslant 2$ and symbols b on $\mathbb{R}^d$ we have

$$\|b\|_{\mathrm{BMO}} \lesssim \sup_{1 \leqslant k \leqslant n} \|[M_b, T_k]\|_{2\to 2}.$$

Here T_k denotes the kth choice of operator in the family $\mathcal{T}$.

We are also going to need the following weaker lower bound in terms of the BMO_{-1} norm in terms of iterated commutators using our families of Calderón-Zygmund operators.

lemma 5. *Let $t \geqslant 2$. Given classes $\mathcal{T}_s$ of Calderón-Zygmund operators with the class of their symbols Θ_s. Assume that for each parameter $1 \leqslant s \leqslant t$ separately we have*

1. *$\forall x \neq y \in \mathbb{S}^{d_s-1}\ \exists\ \theta_{s,i}$ so that $\theta_{s,i}(x) \neq \theta_{s,i}(y)$*
2. *$\forall\ x \in \mathbb{S}^{d_s-1}\ \forall\ t$ tangent to $\mathbb{S}^{d_s-1}$ at $x\ \exists\ i$ so that $\frac{\partial \theta_{s,i}}{\partial t}(x) \neq 0$*

and assume that under these same conditions the lower bound holds in the case of $t-1$ parameters in terms of product BMO*. Then we have the estimate*

$$\|b\|_{\mathrm{BMO}_{-1}} \lesssim \sup_{\vec{k}} \|C_{\vec{k}}(b,\cdot)\|_{2\to 2},$$

where $C_{\vec{k}}(b,\cdot) = [T_{1,k_1}[\ldots[T_{t,k_t}, M_b]\ldots]]$. Here $1 \leq s \leq t, \vec{k} = (k_1, \ldots, k_t), 0 \leq k_s \leq n_s$ and T_{s,k_s} denotes the k_sth choice of operator in the family $\mathcal{T}_s$ acting in the sth variable.

The proof uses a well established equivalent formulation of commutator estimates and weak factorization. This argument goes back to Ferguson and Sadosky [7]. Our case is closest to the proof of lemma (6.3) in [14], replacing the collection of Riesz transforms by our families $\mathcal{T}_s$. We include a sketch for the sake of completeness.

We assume that $t \geq 2$ and use the induction hypothesis to establish a lower bound in terms of our BMO norm with $t-1$ parameters.

Proof. It is sufficient to demonstrate that the following inequality holds,

$$||b||_{(L^2 * L^2)^*} \gtrsim ||b||_{BMO_{-1}}$$

and this will be established, inducting on the number of parameters. Assume the truth of the Theorem in $t-1$ parameters.

Given a smooth symbol $b(x_1, \dots x_t) = b(x_1, x')$ of t parameters, we assume that $||b||_{BMO_{-1}} = 1$. Assume the supremum is achieved by the collection $\mathcal{U}$ of $\mathcal{D}_{\vec{d}}$ of $t-1$ parameters. Say that the rectangles in $\mathcal{U}$ agree in the first coordinate, to a fixed cube $Q \subset \mathbb{R}^{d_1}$. After normalization, assume that $|Q| = 1$ and $|\mathrm{sh}(\mathcal{U})| \approx 1$.

Then define

$$\psi = \sum_{R \in \mathcal{U}} \sum_{\vec{\varepsilon} \in \mathrm{Sig}_{\vec{d}}} \langle b, w_R^{\vec{\varepsilon}} \rangle w_R^{\vec{\varepsilon}}.$$

Note that $\langle b, \psi \rangle = 1$. To prove the claim, it is then enough to prove that $||\psi||_{L^2(\mathbb{R}^d) * L^2(\mathbb{R}^d)} \lesssim 1$. Observe that $\psi(x) = \psi_1(x_1)\psi'(x')$ and $\psi_1 \in H^1(\mathbb{R}^{d_1})$ with

$$||\psi_1||_{H^1(\mathbb{R}^d)} = 1.$$

To ψ_1, apply the one-parameter weak factorization of $H^1(\mathbb{R}^{d_1})$ resulting from the one-parameter characterization result of Li. There exists functions $f_n^j, g_n^j \in L^2(\mathbb{R}^{d_1})$, $n \in \mathbb{N}$, $1 \leq j_1 \leq d_1$, such that

$$\psi_1 = \sum_{n=1}^{\infty} \sum_{j_1=1}^{d_1} \Pi_{1,j_1}(f_n^{j_1}, g_n^{j_1})$$

where $\Pi_{1,j_1}(p,q) := T_{1,j_1}(p)q + pT_{1,j_1}(q)$. One next sees that $\psi' \in H^1(\otimes_{l=2}^t \mathbb{R}^{d_l})$ with norm controlled by a constant. By the induction hypothesis in $t-1$ parameters, in particular that $H^1(\otimes_{l=2}^t \mathbb{R}^{d_l}) = L^2(\otimes_{s=2}^t \mathbb{R}^{d_s}) * L^2(\otimes_{s=2}^t \mathbb{R}^{d_s})$, we have $f_m^{\vec{j}}, g_m^{\vec{j}} \in L^2(\otimes_{s=2}^t \mathbb{R}^{n_s})$ with $m \in \mathbb{N}$ and $\vec{j}$ a vector with $1 \leq j_s \leq d_s$ for $s = 2, \dots, t$ such that

$$\psi' = \sum_{m=1}^{\infty} \sum_{\vec{j}} \Pi_{\vec{j}}(f_m^{\vec{j}}, g_m^{\vec{j}}), \qquad \sum_{m=1}^{\infty} \sum_{\vec{j}} ||f_m^{\vec{j}}||_2 ||g||_m^{\vec{j}} \lesssim 1.$$

This immediately implies $||b||_{(L^2 * L^2)^*} \gtrsim ||b||_{BMO_{-1}}$ since $\psi = \psi_1\psi'$, and we have a weak factorization of ψ with $||\psi||_{L^2(\mathbb{R}^d) * L^2(\mathbb{R}^d)} \lesssim 1$.

□

We now turn to the induction step in the main theorem, to finish the proof of the lower estimate in terms of BMO in t parameters.

Proof. We start with any BMO function b so that $\|b\|_{\mathrm{BMO}-1} < \delta_{-1}$ is small. Notice that we have no loss of generality here : due to lemma (5), we already have a lower bound for such b where $\|b\|_{\mathrm{BMO}_{-1}} \geqslant \delta_{-1}$.

We normalize the function b as before, find the function β and obtain cones D_s, the function $\gamma := T_{\vec{D}}\beta$ and cones C_s according to lemma (2). For a small positive number ϵ to be chosen, that determines the precision with which we approximate the cone transforms T_{C_s}, obtain operators T_s, polynomials in $\Theta \cup \bar{\Theta} \cup \{1\}$.

We are going to see that, indeed, the following estimate holds :

$$\|[T_1, ...[T_t, M_b]...]\bar{\gamma}\|_2 \gtrsim 1$$

Similar to before in the proof of lemma (3), we split the estimate into one large term and two error terms with the help of Journé's lemma. To be precise, we split the symbol function b into its parts $b = P_{\mathcal{U}}b + P_{\mathcal{V}}b + P_{\mathcal{W}}b$.

The commutator $\|[T_1, ...[T_t, M_\beta]...]\bar{\gamma}\|_2$ consists of terms of the form $T\beta T'\bar{\gamma}$ where T, T' are combinations of T_s and the identity. In the case where T' is not the identity, it follows from lemma (4) that the symbol of T' is at most ϵ on the Fourier support of $\bar{\gamma}$. Such components are small :

$$\|T\beta T'\bar{\gamma}\|_2 \lesssim \|\beta T'\bar{\gamma}\|_2 \lesssim \|\beta\|_4 \|T'\bar{\gamma}\|_4 \lesssim \epsilon^{1/3}.$$

To obtain the last inequality, we have to preserve the trivially small L^2 norm $\|T'\bar{\gamma}\|_2$ when passing to L^4. To do so, recall that β is normalized both in BMO and L^2. Observe also that T' is at most ϵ on the Fourier support of $\bar{\gamma}$, which gives us $\|T'\bar{\gamma}\|_2 \leq \epsilon$. In addition, T' has universal L^8 norms independent of ϵ by lemma 4. It remains to interpolate to obtain the estimate above.

Now we are left with term $T\beta\bar{\gamma} = T_1 \ldots . T_t \beta\bar{\gamma}$ which we estimate as follows. Remember that $\gamma = T_{\vec{D}}\beta$ and write

$$\beta = \gamma + (H_{\vec{D}} - T_{\vec{D}})\beta + (I - H_{\vec{D}})\beta,$$

thus obtaining three terms. We will see that only one of them is large.

The functions $(I - H_{\vec{D}})\beta$ and $\bar{\gamma}$ are supported on the same product of half spaces complementary to cones D_s. We know that the symbol $h_{C,D}$ vanishes and therefore the T_s are at most ϵ, so

$$\|T((I - H_{\vec{D}})\beta \cdot \bar{\gamma})\|_2 \leqslant \epsilon \|(I - H_{\vec{D}})\beta \cdot \bar{\gamma}\|_2 \leqslant \epsilon \|(I - H_{\vec{D}})\beta\|_4 \|\bar{\gamma}\|_4.$$

Recall the compositions of half plane projection operators have uniform L^p bounds and that L^4 norms of both β and γ are controlled.

For the part $T((H_{\vec{D}} - T_{\vec{D}})\beta \cdot \bar{\gamma})$ we rely on the estimate from lemma 2 of the L^4 norm

$$\|T((H_{\vec{D}} - T_{\vec{D}})\beta \cdot \bar{\gamma})\|_2 \lesssim \|(H_{\vec{D}} - T_{\vec{D}})\beta \cdot \bar{\gamma}\|_2 \leqslant \kappa \|\bar{\gamma}\|_4 \lesssim \kappa.$$

For the term $T(\gamma\bar\gamma)$we consider

$$|\|T\gamma\bar\gamma\|_2 - \|H_{\vec C}\gamma\bar\gamma\|_2| \leqslant \|(T-H_{\vec C})\gamma\bar\gamma\|_2 \leqslant \|(T-T_{\vec C})\gamma\bar\gamma\|_2 + \|(T_{\vec C}-H_{\vec C})\gamma\bar\gamma\|_2 \lesssim \epsilon+\kappa.$$

Since $\gamma\bar\gamma$ is real with symmetric Fourier transform, we have $\|H_{\vec C}\gamma\bar\gamma\|_2 \gtrsim \|\gamma\bar\gamma\|_2 = \|\gamma\|_4^2$. Furthermore

$$\|\gamma\|_4^2 \gtrsim \left\| \left(\sum_{\varepsilon} \sum_{R\in\mathcal{U}} \frac{|\langle\gamma, w_R\rangle|^2}{|R|} \mathbf{1}_R \right)^{1/2} \right\|_4^2 \gtrsim \left\| \left(\sum_{\varepsilon} \sum_{R\in\mathcal{U}} \frac{|\langle\gamma, w_R\rangle|^2}{|R|} \mathbf{1}_R \right)^{1/2} \right\|_2^2 \gtrsim 1.$$

The first inequality uses a Littlewood Paley inequality and to see the second inequality, note that the rectangles in $\mathcal{U}$ are contained in a set of measure bounded by 1. We have therefore proved that $\|T(\gamma\bar\gamma)\|_2 \gtrsim 1$.

We wish to prove that commutators that arise with our Calder'on-Zygmund operators themselves are large, not just specific polynomials in those operators. To do so, observe the following elementary fact. Let T, T' be Calderón-Zygmund operators. Then

$$[TT', M_b] = T[T', M_b] + [T, M_b]T'.$$

If the symbols of T_s and T'_s are polynomials in the θ_s, it follows that for some choice of operators associated to θ_s,

$$\|[T_{1,k_1}[\dots[T_{t,k_t},\beta]]]\bar\gamma'\|_2 \gtrsim 1$$

where $\bar\gamma'$ is of the form $T\bar\gamma$ and where T is a composition of operators T_{s,l_s}. Notice here that it is essential that we only approximate a finite set of cone operators so that we control degrees and coefficients of the arising polynomials. This point is imperative, since we do not control degree or coefficients with Nachbin's approximation.

Recall that $\beta = P_{\mathcal{U}} b$ and that all dyadic rectangles are split into three groups $\mathcal{U}\dot\cup\mathcal{V}\dot\cup\mathcal{W}$. In order to see that the norm of the commutator satisfies $\|[T_{1,k_1}[\dots[T_{t,k_t},b]]]\|_{2\to 2} \gtrsim 1$, we use test function $\bar\gamma'$ and split the estimate according to partial sums of the symbol b of only those rectangles belonging to classes $\mathcal{U},\mathcal{V},\mathcal{W}$ respectively. We have already seen that

$$\|[T_{1,k_1}[\dots[T_{t,k_t},P_{\mathcal{U}}b]]]\bar\gamma'\|_2 \gtrsim 1.$$

It remains to see that the remaining parts are small. We are going to see that

$$\|[T_{1,k_1}[\dots[T_{t,k_t},P_{\mathcal{V}}b]]]\bar\gamma'\|_2 \lesssim \delta_J^{1/4},$$

the part of the estimate responsive to Journé's lemma and also that

$$\|[T_{1,k_1}[\dots[T_{t,k_t},P_{\mathcal{W}}b]]]\bar\gamma'\|_2 \lesssim K_J\delta_{-1}.$$

For these two estimates, we can follow directly the arguments in [14].

The first estimate illustrates the use of Journé's lemma in this context. We do not need to use any cancellation of the commutator :

$$\|[T_{1,k_1}[\dots[T_{t,k_t},P_{\mathcal{V}}b]]]\bar\gamma'\|_2 \lesssim \|P_{\mathcal{V}}b\|_4\|\gamma\|_4$$

where the implied constant depends upon L^2 and L^4 operator norms of the T_{s,k_s}. The L^4 norm of γ is uniformly controlled and by construction we have $\|P_{\mathcal{V}} b\|_{\mathrm{BMO}} \leqslant 1$. Last, Journé's lemma provides us with the estimate $\|P_{\mathcal{V}} b\|_2^2 \leqslant \delta_J$. Interpolation then gives $\|P_{\mathcal{V}} b\|_4 \lesssim \delta_J^{1/4}$.

The last estimate requires a very careful analysis, but does not use the specifics of our operators, except the control on a large number of derivatives of the kernel. We therefore appeal to the version in [14], section 7, where the estimate was stated for Riesz transforms but in fact carried out for more general Calderón-Zygmund operators with control on a large number of derivatives, such as the ones we have here.

□

1.5 Concluding Remarks

Our theorem is a generalization of the Riesz transform case, but it falls short of recovering the full Uchiyama-Li criterion in several parameters. Li's criterion only requires point separation of all pairs ξ and $-\xi$ on the sphere. This criterion is quite natural as it makes sure there is an operator in the family that has a singularity in a given direction, for all directions. Due to the method of proof, we felt the need to require point separation for all pairs of points as well as a derivative condition. The strategy to obtain lower bounds in this multi-parameter setting remains analytic in nature - while we are not able to use Fourier projections directly as in one dimension, we build operators that are close enough to still pretend we are in the one-dimensional setting. Families that have Li's criterion are not enough to approximate the operators we need in the norm of uniform convergence in $\mathcal{C}(\mathbb{S}^{d-1})$ much less in $\mathcal{C}^n(\mathbb{S}^{d-1})$. We require the latter because we need excellent convergence of multiplier symbols on the Fourier transform side in order to draw meaningful conclusions. It is interesting to remark that, in cases like ours, one easily proves a version of Stone Weierstrass theorem that can handle defects in the sense that it is clear which algebra is generated by a family of functions with defects, such as a lack of point separation for a given pair of ξ and ζ in $\mathbb{S}^{d-1}$. One uses factor spaces to see that the generated algebra will have the exact same set of defects : the algebra generated by a family that lacks point separation for a set of pairs (ξ, ζ) will be the subalgebra with that same property. The situation is not so simple if one needs uniform approximation in $\mathcal{C}^n(\mathbb{S}^{d-1})$. Due to the necessary conditions on the tangential derivatives, the situation becomes very complex when the family has defects, such as a lack of point separation in just one point or the lack of non-zero tangential derivatives. The corresponding subalgebras are unknown since the 1950s.

Chapitre 2

La borne supérieure

2.1 Introduction

In [14] the product BMO space on $\mathbb{R}^{d_1} \otimes \cdots \otimes \mathbb{R}^{d_t}$ was characterized by the multi-parameter iterated commutators of Riesz transforms. This extended to the product setting the classical results of R. Coifman, R. Rochberg and G. Weiss [5], a characterization of classical BMO in terms of boundedness on $L^2(\mathbb{R}^d)$ of the commutator of a singular integral operator with a multiplication operator, which by duality also implies a weak factorization result of $H^1(\mathbb{R}^d)$.

In the multi-parameter setting, let M_b be the operator of pointwise multiplication by $b \in BMO(\mathbb{R}^{\vec{d}})$. Let T_i be the Calderón-Zygmund operators on $\mathbb{R}^{d_i}$. One seeks to characterize product BMO in terms of commutators in the sense that

$$\|b\|_{BMO} \lesssim \|[\ldots[[M_b, T_1], T_2] \ldots, T_t]\|_{L^2 \to L^2} \lesssim \|b\|_{BMO}$$

where the first and second inequality will be referred to as lower bound and upper bound, respectively.

In the case of Hilbert transform, the above result in bi-parameter setting was proved by M. Lacey and S. Ferguson in [13], where the upper bound was first shown by S. Ferguson and C. Sadosky [7]. M. Lacey and E. Tervilleger [16] then extended the result to the multi-parameter setting. The Riesz transform result was proved by M. Lacey, S. Petermichl, J. Pipher and B. Wick in [14], where they obtained a more general upper bound result for any Calderón-Zygmund operators of convolution type with high degree of smoothness. Later on in [15] they simplified the proof of upper bound for Riesz transforms by means of dyadic shifts.

In this paper, we prove the upper bound for any given collection of Calderón-Zygmund operators. As a corollary, we prove new characterizations of product BMO in terms of commutators of Calderón-Zygmund operators.

The main theorem of the paper is the following.

theorem 9. *Let $b \in BMO(\mathbb{R}^{\vec{d}})$ and $(T_i)_{1\leq i\leq t}$ be a collection of Calderón-Zygmund operators, with each T_i acting on parameter i of $\mathbb{R}^{\vec{d}} = \mathbb{R}^{d_1} \otimes \cdots \otimes \mathbb{R}^{d_t}$. Then,*

$$\|[\ldots[[M_b, T_1], T_2]\ldots, T_t]\|_{L^2\to L^2} \leq C\|b\|_{BMO}$$

where C depends only on $\vec{d}$ and $\prod_{i=1}^t \|T_i\|_{CZ}$.

One of the interesting results implied by the theorem is that a perturbation of a collection of operators characterizing product BMO still characterizes product BMO. Since Calderón-Zygmund operators form a linear space, whose norm can be made arbitrarily small by multiplying a small constant, it means that once we have a collection of operators characterizing BMO, say, Riesz transforms, we automatically obtain infinitely many collections of operators which also characterize BMO. We organize this observation into the following corollary, whose proof is given at the end of the paper.

corollary 2. *Let $(T_{i,s_i})_{1\leq i\leq t, 1\leq s_i\leq n_i}$ be a family of Calderón-Zygmund operators characterizing the space $BMO_{prod}(\mathbb{R}^{\vec{d}})$, that is, $\exists C_1, C_2 > 0$, such that*

$$C_1\|b\|_{BMO} \leq \sup_{1\leq i\leq t, 1\leq s_i\leq n_i} \|[\ldots[[M_b, T_{1,s_1}], T_{2,s_2}]\ldots, T_{t,s_t}]\|_{L^2\to L^2} \leq C_2\|b\|_{BMO}.$$

Then, $\exists\epsilon > 0$ such that for any family of Calderón-Zygmund operators $(T'_{i,s_i})_{1\leq i\leq t,1\leq s_i\leq n_i}$ satisfying $\|T'_{i,s_i}\|_{CZ} \leq \epsilon$, the family $(T_{i,s_i}+T'_{i,s_i})_{1\leq i\leq t,1\leq s_i\leq n_i}$ still characterizes $BMO(\mathbb{R}^{\vec{d}})$. In particular, for an arbitrary family of Calderón-Zygmund operators $(T'_{i,s_i})_{1\leq i\leq t,1\leq s_i\leq n_i}$, there exist $\epsilon_1, \ldots, \epsilon_t > 0$ such that for any $0 < c_i < \epsilon_i$, $1 \leq i \leq t$, the family $(T_{i,s_i} + c_iT'_{i,s_i})_{1\leq i\leq t,1\leq s_i\leq n_i}$ characterizes $BMO(\mathbb{R}^{\vec{d}})$.

The main tool in the proof of the main theorem is the representation theorem by T. Hytönen [9], which states that any Calderón-Zygmund operator can be represented as a probabilistic average of simple dyadic shift operators. While the earliest version of this theorem appeared in [10], here we choose to apply a slightly different one given in [9]. In our proof, we will reduce the problem to the upper bound for commutators with dyadic shifts. This is the first use of Hytönen's representation theorem to commutator theory. The novelty of this approach to the upper bound is two fold. First, in contrast to typical methods dealing with multi-parameter theory, the main estimates for the dyadic shifts can be iterated. Second, new paraproducts are introduced, and this is where the delicate estimates in product theory are required.

The paper is organized as follows. In Section 2, we recall several preliminary results on dyadic shifts, representation theorem and multi-parameter paraproducts. In Section 3, a full proof of the main theorem in its one-parameter case is introduced. The proof of the main theorem in arbitrarily many parameters is presented in Section 4, while the last section is devoted to the proof of the corollary mentioned above.

2.1.1 Dyadic shifts and representation theorem

Recall that while the standard dyadic grid is defined as

$$\mathcal{D}^0 := \{2^{-k}([0,1)^d + m) : k \in \mathbb{Z}, m \in \mathbb{Z}^d\},$$

for any parameter $\omega = (\omega_j)_{j\in\mathbb{Z}} \in (\{0,1\}^d)^{\mathbb{Z}}$, one can define an associated shifted dyadic grid as

$$\mathcal{D}^\omega := \{I \dot{+} \omega : I \in \mathcal{D}^0\}$$

where

$$I \dot{+} \omega := I + \sum_{j:2^{-j}<\ell(I)} 2^{-j}\omega_j.$$

For a fixed shifted grid $\mathcal{D}^\omega$ and $i, j \in \mathbb{Z}_+$, a dyadic shift operator S^{ij}_ω is defined to be bounded on L^2 with operator norm less than 1. Specifically,

$$S^{ij}_\omega f := \sum_{K\in\mathcal{D}^\omega} \sum_{\substack{I\in\mathcal{D}^\omega, I\subset K \\ \ell(I)=2^{-i}\ell(K)}} \sum_{\substack{J\in\mathcal{D}^\omega, J\subset K \\ \ell(J)=2^{-j}\ell(K)}} a_{IJK}\langle f, h_I\rangle h_J := \sum_K \sum_{I,J\subset K}^{(i,j)} a_{IJK}\langle f, h_I\rangle h_J,$$

with $|a_{IJK}| \leq |I|^{1/2}|J|^{1/2}/|K|$. S^{ij}_ω is called cancellative if all the Haar functions in the definition are cancellative, otherwise, it is called noncancellative.

Recall that in one dimension, any dyadic interval I is associated with a cancellative Haar function $h^0_I = |I|^{-1/2}(\chi_{I_l} - \chi_{I_r})$ and a noncancellative one $h^1_I = |I|^{-1/2}\chi_I$. While in d dimensions, each cube $I = I_1 \times \cdots \times I_d$ is associated with 2^d Haar functions :

$$h^\epsilon_I(x) = h^{(\epsilon_1,\ldots,\epsilon_d)}_{I_1\times\cdots\times I_d}(x_1,\ldots,x_d) = \prod_{i=1}^d h^{\epsilon_i}_{I_i}(x_i),\ \epsilon \in \{0,1\}^d,$$

where h^1_I is called noncancellative, while all the other $2^d - 1$ Haar functions h^ϵ_I for $\epsilon \in \{0,1\}^d \setminus \{1\}$ are cancellative. Note that all the cancellative Haar functions for a fixed grid form an orthonormal basis of $L^2(\mathbb{R}^d)$. And in this paper, we usually suppress the parameter ϵ to abbreviate the notation.

We now introduce T. Hytönen's representation theorem, a key tool in our proof. Interested readers can find its proof and more detailed discussion in [9] and [10]. The operator T mentioned in the following will denote a Calderón-Zygmund operator associated with a δ-standard kernel K. T. Hytönen [9] proved the following theorem :

theorem 10. *Let T be a Calderón-Zygmund operator, then it has an expansion, say for $f, g \in C^\infty_0(\mathbb{R}^d)$,*

$$\langle g, Tf\rangle = c \cdot \|T\|_{CZ} \cdot \mathbb{E}_\omega \sum_{i,j=0}^\infty 2^{-\max(i,j)\delta/2}\langle g, S^{ij}_\omega f\rangle,$$

where c is a dimensional constant and S^{ij}_ω is a dyadic shift of parameter (i,j) on the dyadic grid $\mathcal{D}^\omega$; all of them except possibly S^{00}_ω are cancellative.

According to the proof of Theorem 10, in the representation of any T, only S^{00}_ω may be noncancellative, and if this is the case, only one of $\{h_I\}$, $\{h_J\}$ in its definition is noncancellative, i.e. S^{00}_ω is a paraproduct.

2.1.2 Multi-parameter paraproducts

Recall that a multi-parameter paraproduct associated with function b can be viewed as a bilinear operator which is defined as

$$B(b,f) = \sum_{R\in\mathcal{D}_{\vec{d}}} \pm\langle b, h_R^{\epsilon_1}\rangle\langle f, h_R^{\epsilon_2}\rangle h_R^{\epsilon_3}|R|^{-1/2},$$

where $\epsilon_j \in \{0,1\}^{\vec{d}}$ and $\mathcal{D}_{\vec{d}}$ denotes the tensor product of dyadic grids. Note that $h_R^{\epsilon_j}$ is cancellative if and only if $\epsilon_j \neq \vec{1}$. According to Journé [11] and later on improved by C. Muscalu, J. Pipher, T. Tao and C. Thiele [20] [21], one has the following boundedness result.

theorem 11. *Let $\vec{d} = (d_1,\dots,d_t)$ and $\epsilon_j = (\epsilon_{j,1},\dots,\epsilon_{j,t})$. If $\epsilon_1 \neq \vec{1}$ and $\forall 1\le s\le t$, there is at most one of $j=2,3$ such that $\epsilon_{j,s}=\vec{1}$, then the operator B satisfies*

$$B:\, BMO_{prod}(\mathbb{R}^{\vec{d}})\times L^p(\mathbb{R}^{\vec{d}}) \to L^p(\mathbb{R}^{\vec{d}}), \quad 1<p<\infty.$$

2.2 Proof of the one-parameter case

In this section, we present a detailed proof of the main theorem in the one-parameter setting, which will later on be generalized to work in the multi-parameter setting. As an essential part of the proof, delicate estimates of several new paraproducts will be introduced.

Given a BMO function b and a Calderón-Zygmund operator T, one could represent the commutator $[b,T]$ as an average of $[b,S^{ij}_\omega]$ due to Theorem 10. Then, in order to prove the upper bound inequality, it suffices to prove that for any $f\in C_0^\infty(\mathbb{R}^d)$,

$$\|\sum_{i,j=0}^{\infty} 2^{-\max(i,j)\delta/2}[b,S^{ij}_\omega]f\|_{L^2} \lesssim \|b\|_{BMO}\|f\|_{L^2}$$

uniformly in ω. In the following we will write S^{ij} for short as the argument doesn't depend on ω explicitly.

The strategy of the proof is the following. First, we decompose b and f using Haar basis. Second, we split the sum into several parts and rewrite each of them as a paraproduct or its variant. (Note that in some of the cases one may end up with a paraproduct composed with a dyadic shift instead). Finally, we apply Theorem 11 and its variants to obtain sufficiently good decay estimates.

One can decompose $[b, S^{ij}]f$ as

$$\begin{aligned}[b, S^{ij}]f &= \sum_{I,J} \langle b, h_I\rangle \langle f, h_J\rangle [h_I, S^{ij}]h_J \\ &= \sum_{I,J} \langle b, h_I\rangle \langle f, h_J\rangle \left(h_I S^{ij} h_J - S^{ij}(h_I h_J)\right) := I + II,\end{aligned}$$

where in the following I and II will be referred to as first term and second term, respectively. In order to further organize the sum and extract the correct paraproducts structure, even in the simplest one-parameter case, one needs to divide up the sum into many different parts to analyze, depending on the relative sizes of i, j and I, J.

2.2.1 Case $(i, j) \neq (0, 0)$

Let's first look at the case when $(i, j) \neq (0, 0)$, meaning that all the Haar functions appearing are cancellative. Hence,

$$[b, S^{ij}]f = \sum_{I,J} \langle b, h_I\rangle \langle f, h_J\rangle \left(h_I \sum_{J' \subset J^{(i)}}^{(j)} a_{JJ'J^{(i)}} h_{J'} - \sum_K \sum_{I'', J'' \subset K}^{(i,j)} a_{I''J''K} \langle h_I h_J, h_{I''}\rangle h_{J''} \right),$$

where $J^{(i)}$ denotes the i^{th} ancestor of J.

First, we claim that it suffices to estimate the part $I \subset J^{(i)}$. Indeed, it is obvious that when $I \cap J^{(i)} = \emptyset$, both terms in the parenthesis are zero. Furthermore, by the cancellation structure of the commutator, when $I \supsetneq J^{(i)}$, the term $[h_I, S^{ij}]h_J$ is also zero. To see this, as h_I is constant on $J^{(i)}$, fixing an arbitrary $x_0 \in J^{(i)}$ implies

$$h_I S^{ij} h_J - S^{ij}(h_I h_J) = h_I(x_0) S^{ij} h_J - S^{ij}(h_I(x_0) h_J) = 0.$$

Note that this is the only part of the proof where one needs the particular commutator structure.

Next, we discuss the case $j > i$ and $j \leq i$ separately. In both cases, we split the sum into two parts ($\ell(I) \leq 2^{i-j}\ell(J)$ and $2^{i-j}\ell(J) < \ell(I) \leq 2^i \ell(J)$) and show that the L^2 norms of both of the first and second terms are bounded.

Case $j > i$

We first estimate the part $\ell(I) \leq 2^{i-j}\ell(J)$ of the first term $h_I S^{ij} h_J$. According to the size of I, there exists only one $J' \subset J^{(i)}$ such that $\ell(J') = 2^{i-j}\ell(J)$ and $I \subset J'$.

Hence, the first term in $[b, S^{ij}]f$ becomes

$$\begin{aligned}
&\sum_J \sum_{\substack{I\subset J^{(i)}\\ \ell(I)\leq 2^{i-j}\ell(J)}} \langle b, h_I\rangle\langle f, h_J\rangle h_I a_{JJ'J^{(i)}} h_{J'}\\
&= \sum_K \sum_{J'\subset K}^{(j)} \sum_{J\subset K}^{(i)} \sum_{\substack{I\subset J'\\ \ell(I)\leq 2^{-j}\ell(K)}} \langle b, h_I\rangle h_I a_{JJ'K}\langle f, h_J\rangle h_{J'}\\
&= \sum_K \sum_{J'\subset K}^{(j)} \sum_{\substack{I\subset J'\\ \ell(I)\leq 2^{-j}\ell(K)}} \langle b, h_I\rangle h_I\langle S^{ij}f, h_{J'}\rangle h_{J'}\\
&= \sum_I \sum_{J'\supset I} \langle b, h_I\rangle h_I\langle S^{ij}f, h_{J'}\rangle h_{J'}\\
&= \sum_I \langle b, h_I\rangle h_I\langle S^{ij}f, h_I\rangle h_I + \sum_I \langle b, h_I\rangle h_I\langle S^{ij}f, h_I^1\rangle h_I^1\\
&= \sum_I \langle b, h_I\rangle\langle S^{ij}f, h_I\rangle h_I^{\epsilon}|I|^{-1/2} + \sum_I \langle b, h_I\rangle\langle S^{ij}f, h_I^1\rangle h_I|I|^{-1/2}.
\end{aligned}$$

Both of the terms above are classical one-parameter paraproducts of type $B(b, S^{ij}f)$, whose L^2 norm is bounded by $\|b\|_{BMO}\|f\|_{L^2}$ due to Theorem 11, which is good enough since the decaying factor $2^{-j\delta/2}$ in front would ensure the summability of the sum over i, j.

Now we turn to estimate the second term $S^{ij}(h_I h_J)$. Due to the supports of Haar functions, this term is nontrivial only when $I \cap J \neq \emptyset$. Because of $j > i$, one has $I \subsetneq J$, which means that h_J is a constant on I. Hence, the second term is

$$\begin{aligned}
&S^{ij}\Big(\sum_I \langle b, h_I\rangle \sum_{J\supset I^{(j-i)}} \langle f, h_J\rangle h_I h_J\Big)\\
&= S^{ij}\Big(\sum_I \langle b, h_I\rangle \sum_{J\supsetneq I^{(j-i)}} \langle f, h_J\rangle h_I h_J\Big) + S^{ij}\Big(\sum_I \langle b, h_I\rangle\langle f, h_{I^{(j-i)}}\rangle h_I h_{I^{(j-i)}}\Big)\\
&= S^{ij}\Big(\sum_I \langle b, h_I\rangle\langle f, h^1_{I^{(j-i)}}\rangle h_I h^1_{I^{(j-i)}}\Big) + S^{ij}\Big(\sum_I \langle b, h_I\rangle\langle f, h_{I^{(j-i)}}\rangle h_I h_{I^{(j-i)}}\Big)\\
&= S^{ij}\Big(\sum_I \pm\langle b, h_I\rangle\langle f, h^1_{I^{(j-i)}}\rangle h_I|I^{(j-i)}|^{-1/2}\Big) + S^{ij}\Big(\sum_I \pm\langle b, h_I\rangle\langle f, h_{I^{(j-i)}}\rangle h_I|I^{(j-i)}|^{-1/2}\Big),
\end{aligned}$$

where on the last line the two terms inside the dyadic shift are both variants of paraproduct. The desired estimate would follow easily if we could show that both of them are bounded : $BMO \times L^2 \to L^2$ with norm controlled by some dimensional constant, for which we need the following lemma.

lemma 6. *Given $b \in BMO(\mathbb{R}^d)$ and $k \geq 0$, define an operator*

$$B_1(b, f) = \sum_I \pm\langle b, h_I\rangle\langle f, h^{\epsilon}_{I^{(k)}}\rangle h_I^{\epsilon'}|I^{(k)}|^{-1/2},$$

where at least one of ϵ, ϵ' is not $\vec{1}$. Then $\|B_1(b, f)\|_{L^2} \lesssim \|b\|_{BMO}\|f\|_{L^2}$ with a constant independent of k.

Proof. We only prove the case when $\epsilon \neq \vec{1}$, the other one is similar. And let's omit ϵ to keep the notation concise. For any $g \in L^2(\mathbb{R}^d)$ with $\|g\|_{L^2} \leq 1$,

$$\langle B_1(b,f), g\rangle = \langle b, \sum_I \pm\langle f, h_{I^{(k)}}\rangle\langle g, h_I^{\epsilon'}\rangle h_I |I^{(k)}|^{-1/2}\rangle,$$

it suffices to show that

$$\|\sum_I \pm\langle f, h_{I^{(k)}}\rangle\langle g, h_I^{\epsilon'}\rangle h_I |I^{(k)}|^{-1/2}\|_{H^1} \lesssim \|f\|_{L^2}\|g\|_{L^2}.$$

To see this, write

$$\begin{aligned} &S(\sum_I \pm\langle f, h_{I^{(k)}}\rangle\langle g, h_I^{\epsilon'}\rangle h_I |I^{(k)}|^{-1/2})^2 = \sum_I |\langle f, h_{I^{(k)}}\rangle\langle g, h_I^{\epsilon'}\rangle|^2 \frac{\chi_I(x)}{|I||I^{(k)}|} \\ &\leq \sup_{x\in I} \frac{\langle g, h_I^{\epsilon'}\rangle^2}{|I|} \sum_I \langle f, h_{I^{(k)}}\rangle^2 \frac{\chi_I(x)}{|I^{(k)}|} = \sup_{x\in I} \frac{\langle g, h_I^{\epsilon'}\rangle^2}{|I|} 2^{-kd} \sum_I \langle f, h_{I^{(k)}}\rangle^2 \frac{\chi_I(x)}{|I|} \\ &\leq 2^{-kd}(Mg(x))^2(S_{(k)}f(x))^2, \end{aligned}$$

where $S_{(k)}f := (\sum_I |\langle f, h_{I^{(k)}}\rangle|^2|I|^{-1}\chi_I)^{1/2}$. Since $\|S_{(k)}f\|_{L^2}^2 \leq \sum_I |\langle f, h_{I^{(k)}}\rangle|^2 = \sum_J \sum_{I\subset J}^{(k)} |\langle f, h_J\rangle|^2 = 2^{kd}\|f\|_{L^2}^2$,

$$\begin{aligned} \|\sum_I \pm\langle f, h_{I^{(k)}}\rangle\langle g, h_I^{\epsilon'}\rangle h_I |I^{(k)}|^{-1/2}\|_{H^1} &\lesssim \|S(\sum_I \pm\langle f, h_{I^{(k)}}\rangle\langle g, h_I^{\epsilon'}\rangle h_I |I^{(k)}|^{-1/2})\|_{L^1} \\ &\leq 2^{-kd/2}\|(Mg)(S_{(k)}f)\|_{L^1} \leq 2^{-kd/2}2^{kd/2}\|f\|_{L^2}\|g\|_{L^2} = \|f\|_{L^2}\|g\|_{L^2}. \end{aligned}$$

□

It is easy to see that this lemma yields the desired estimate of the second term due to the decaying factor $2^{-j\delta/2}$, which completes the discussion of the part $\ell(I) \leq 2^{i-j}\ell(J)$.

Now we turn to consider the part $2^{i-j}\ell(J) < \ell(I) \leq 2^i\ell(J)$ and again start with the first term, which can be written as

$$\begin{aligned} &\sum_J \sum_{\substack{I\subset J^{(i)} \\ \ell(I)>2^{i-j}\ell(J)}} \langle b, h_I\rangle\langle f, h_J\rangle h_I \sum_{\substack{J'\subset J^{(i)}, J'\subsetneq I \\ \ell(J')=2^{i-j}\ell(J)}} a_{JJ'J^{(i)}} h_{J'} \\ &= \sum_K \sum_{J'\subset K}^{(j)} \sum_{J\subset K}^{(i)} \sum_{J'\subsetneq I\subset K} \langle b, h_I\rangle\langle f, h_J\rangle h_I a_{JJ'K} h_{J'} \\ &= \sum_K \sum_{J'\subset K}^{(j)} \sum_{J'\subsetneq I\subset K} \langle b, h_I\rangle h_I \langle S^{ij}f, h_{J'}\rangle h_{J'} \\ &= \sum_J \sum_{J\subsetneq I\subset J^{(j)}} \langle b, h_I\rangle h_I \langle S^{ij}f, h_J\rangle h_J \\ &= \sum_{k=1}^j \sum_J \pm\langle b, h_{J^{(k)}}\rangle\langle S^{ij}f, h_J\rangle h_J |J^{(k)}|^{-1/2}. \end{aligned}$$

We will prove the following lemma, which shows the estimate for a variant of the classical one-parameter paraproduct appeared above.

lemma 7. *Given $b \in BMO(\mathbb{R}^d)$ and $k \geq 0$, define an operator*

$$B_2(b,f) = \sum_I \pm \langle b, h_{I^{(k)}} \rangle \langle f, h_I \rangle h_I |I^{(k)}|^{-1/2},$$

where all the Haar functions are cancellative. Then $\|B_2(b,f)\|_{L^2} \lesssim \|b\|_{BMO}\|f\|_{L^2}$ with a constant independent of k.

Before we proceed to its proof, note that for the application to our problem, there is no need to generalize the lemma to include cases when some of the Haar functions are noncancellative. To see this, observe that only when $i = j = 0$ are there possibly noncancellative Haar functions appearing in the dyadic shift. However, the entire part $2^{i-j}\ell(J) < \ell(I) \leq 2^i \ell(J)$ would then vanish since in that case $\ell(I) \leq \ell(J)$.

Proof. The proof of this result is in the same fashion as the one of Lemma 6. It suffices to show that for any $g \in L^2(\mathbb{R}^d)$ with $\|g\|_{L^2} \leq 1$,

$$\|S(\sum_I \pm \langle f, h_I \rangle \langle g, h_I \rangle h_{I^{(k)}} |I^{(k)}|^{-1/2})\|_{L^1} \lesssim \|f\|_{L^2}\|g\|_{L^2}.$$

To see this, write

$$S(\sum_I \pm \langle f, h_I \rangle \langle g, h_I \rangle h_{I^{(k)}} |I^{(k)}|^{-1/2})^2 = \sum_J \left(\sum_{I:I^{(k)}=J} \langle f, h_I \rangle \langle g, h_I \rangle |I^{(k)}|^{-1/2} \right)^2 \frac{\chi_J}{|J|}$$

which together with Cauchy-Schwartz inequality implies

$$\begin{aligned}
S(\sum_I \pm \langle f, h_I \rangle \langle g, h_I \rangle h_{I^{(k)}} |I^{(k)}|^{-1/2}) &\leq \sum_J \left(\sum_{I:I^{(k)}=J} \langle f, h_I \rangle \langle g, h_I \rangle | \frac{\chi_J}{|J|} \right) \\
&\leq \sum_J \left(\sum_{I:I^{(k)}=J} |\langle f, h_I \rangle|^2 \right)^{1/2} \left(\sum_{I:I^{(k)}=J} |\langle g, h_I \rangle|^2 \right)^{1/2} \frac{\chi_J}{|J|} \\
&\leq \left(\sum_J \sum_{I:I^{(k)}=J} |\langle f, h_I \rangle|^2 \frac{\chi_J}{|J|} \right)^{1/2} \left(\sum_J \sum_{I:I^{(k)}=J} |\langle g, h_I \rangle|^2 \frac{\chi_J}{|J|} \right)^{1/2} \\
&:= (S^{(k)} f)(S^{(k)} g).
\end{aligned}$$

where the operator $S^{(k)} f := (\sum_J \sum_{I:I^{(k)}=J} |\langle f, h_I \rangle|^2 |J|^{-1} \chi_J)^{1/2}$. We claim that $S^{(k)} : L^2 \to L^2$ with norm bounded by a dimensional constant, which does not depend on k. Combining this with another Cauchy-Schwartz will complete the proof.

To show the claim, denote $\alpha_J = (\sum_{I:I^{(k)}=J} |\langle f, h_I \rangle|^2)^{1/2}$ for any J and define $F(x) = \sum_J \alpha_J h_J(x)$. Then

$$\begin{aligned}
\|S^{(k)} f\|_{L^2}^2 &= \|(\sum_J \alpha_J^2 \frac{\chi_J}{|J|})^{1/2}\|_{L^2}^2 = \|SF\|_{L^2}^2 \\
&\lesssim \|F\|_{L^2}^2 = \sum_J \alpha_J^2 = \sum_J \sum_{I:I^{(k)}=J} |\langle f, h_I \rangle|^2 = \sum_I |\langle f, h_I \rangle|^2 = \|f\|_{L^2}^2.
\end{aligned}$$

□

Applying this lemma to our previous calculation of the first term implies that the L^2 norm of it is bounded by $\|\sum_{k=1}^{j} B_2(b,f)\|_{L^2} \lesssim j\|b\|_{BMO}\|f\|_{L^2}$. When summing over i,j finally, the extra j here won't matter as the decaying factor $2^{-j\delta/2}$ is much smaller. This completes the discussion of the first term.

Now the only part left is the second term, which can be further split into two parts as

$$I := S^{ij}(\sum_{J}\sum_{\substack{I\subset J\\ \ell(I)>2^{i-j}\ell(J)}}\langle b,h_I\rangle\langle f,h_J\rangle h_I h_J)$$

$$II := S^{ij}(\sum_{J}\sum_{J\subsetneq I\subset J^{(i)}}\langle b,h_I\rangle\langle f,h_J\rangle h_I h_J),$$

due to the supports of Haar functions.

Note that II is of exactly the same form as the sum appeared in the estimate of the first term except that j has been changed to i. Hence, the same reasoning implies that $\|II\|_{L^2} \lesssim i\|b\|_{BMO}\|f\|_{L^2}$. Again, because of the existence of the decaying factor in front, summing over i,j won't blow this up. It thus suffices to estimate part I. And this can also be achieved through a similar technique by observing that

$$\begin{aligned}
I &= S^{ij}(\sum_{I}\sum_{I\subset J\subsetneq I^{(j-i)}}\langle b,h_I\rangle\langle f,h_J\rangle h_I h_J)\\
&= S^{ij}(\sum_{I}\langle b,h_I\rangle\langle f,h_I\rangle h_I^{\epsilon}|I|^{-1/2}) + S^{ij}(\sum_{I}\sum_{I\subsetneq J\subsetneq I^{(j-i)}}\langle b,h_I\rangle\langle f,h_J\rangle h_I h_J)\\
&= S^{ij}(\sum_{I}\langle b,h_I\rangle\langle f,h_I\rangle h_I^{\epsilon}|I|^{-1/2}) + \sum_{k=1}^{j-i-1} S^{ij}(\sum_{I}\langle b,h_I\rangle\langle f,h_{I^{(k)}}\rangle h_I|I^{(k)}|^{-1/2}),
\end{aligned}$$

where the first term is S^{ij} acting on a classical paraproduct of type $B(b,f)$, and the second term can be estimated using Lemma 6. This completes the discussion of case $j>i$.

Case $j \le i$

Similarly as before, we start with the part $\ell(I) \le 2^{i-j}\ell(J)$. The estimate for the first term can be obtained exactly the same as in the case $j>i$. It thus suffices to estimate the second term, which is nonzero only when $I\cap J \neq \emptyset$. We then have to consider two different possibilities : $I\subset J$ or $J\subsetneq I\subset J^{(i-j)}$.

If $I \subset J$, we have

$$\begin{aligned}
&S^{ij}(\sum_I \langle b, h_I\rangle h_I \sum_{J\supset I}\langle f, h_J\rangle h_J)\\
&= S^{ij}(\sum_I \langle b, h_I\rangle h_I \sum_{J\supsetneq I}\langle f, h_J\rangle h_J) + S^{ij}(\sum_I \langle b, h_I\rangle\langle f, h_I\rangle h_I^{\epsilon}|I|^{-1/2})\\
&= S^{ij}(\sum_I \langle b, h_I\rangle h_I \langle f, h_I^1\rangle h_I^1) + S^{ij}(\sum_I \langle b, h_I\rangle\langle f, h_I\rangle h_I^{\epsilon}|I|^{-1/2})\\
&= S^{ij}(\sum_I \langle b, h_I\rangle \langle f, h_I^1\rangle h_I|I|^{-1/2}) + S^{ij}(\sum_I \langle b, h_I\rangle\langle f, h_I\rangle h_I^{\epsilon}|I|^{-1/2}),
\end{aligned}$$

which is S^{ij} acting on two classical paraproducts, implying the desired estimate.

If $J \subsetneq I \subset J^{(i-j)}$ instead, we have

$$S^{ij}(\sum_J \sum_{J\subsetneq I\subset J^{(i-j)}} \langle b, h_I\rangle\langle f, h_J\rangle h_I h_J),$$

which is of exactly the same form as the sum appeared in the estimate of the first term in part $2^{i-j}\ell(J) < \ell(I) \leq 2^i\ell(J)$ of the case $j > i$, except that here we have $i-j$ in place of j. Hence, the argument above implies that the L^2 norm of it is bounded by $(i-j)\|b\|_{BMO}\|f\|_{L^2}$. Taking into account of the decaying factor $2^{-i\delta/2}$, this completes the estimate of part $\ell(I) \leq 2^{i-j}\ell(J)$.

Next, let's consider the part $2^{i-j}\ell(J) < \ell(I) \leq 2^i\ell(J)$, where again the first term can be estimated as same as in the case $j > i$, and it thus remains to study the second term. Since $i \leq j$ and $I \cap J \neq \emptyset$, one has $I \supsetneq J$, corresponding to which the sum is

$$S^{ij}(\sum_J \sum_{J\subsetneq I\subset J^{(i)}} \langle b, h_I\rangle\langle f, h_J\rangle h_I h_J),$$

which is the same as II appeared above at the end of Case $j > i$, the desired estimate then follows.

2.2.2 Case $(i, j) = (0, 0)$

The only different case that may occur here is when S^{00} is noncancellative. And it suffices to assume that it is of the type $S^{00}f = \sum_I a_I\langle f, h_I^1\rangle h_I$, since if we switch the positions of cancellative and noncancellative Haar functions, what we get is none other than its adjoint. Furthermore, following from our discussion at the beginning, it suffices to consider the case $I \subset J$.

The estimate for the second term is the same as the one appeared in part $\ell(I) \leq 2^{i-j}\ell(J)$ of the case $j \leq i$, which we omit. To study the first term, one observes that for any h_J,

$$S^{00}h_J = \sum_{I\subsetneq J} a_I\langle h_J, h_I^1\rangle h_I = \sum_{I\subsetneq J} a_I|I|^{1/2}h_I h_J.$$

Hence, the first term becomes

$$\sum_{I\subset J}\sum_{I'\subsetneq J}\langle b, h_I\rangle h_I\langle f, h_J\rangle a_{I'}|I'|^{1/2}h_{I'}h_J = \sum_{I\subset I'\subsetneq J} + \sum_{I'\subsetneq I\subset J} := I + II.$$

One write

$$
\begin{aligned}
I &= \sum_I \langle b, h_I\rangle h_I \left(\sum_{I\subset I'\subsetneq J} a_{I'}\langle f, h_J\rangle h_J |I'|^{1/2} h_{I'} \right) \\
&= \sum_I \langle b, h_I\rangle h_I \left(\sum_{I\subset I'} a_{I'} |I'|^{1/2} h_{I'} \langle f, h^1_{I'}\rangle h^1_{I'} \right) \\
&= \sum_I \langle b, h_I\rangle h_I \left(\sum_{I\subset I'} a_{I'} \langle f, h^1_{I'}\rangle h_{I'} \right) \\
&= \sum_I \langle b, h_I\rangle h_I \left(\sum_{I\subset I'} \langle S^{00} f, h_{I'}\rangle h_{I'} \right) \\
&= \sum_I \langle b, h_I\rangle h_I \langle S^{00} f, h_I\rangle h_I + \sum_I \langle b, h_I\rangle h_I \langle S^{00} f, h^1_I\rangle h^1_I \\
&= \sum_I \langle b, h_I\rangle \langle S^{00} f, h_I\rangle h^\epsilon_I |I|^{-1/2} + \sum_I \langle b, h_I\rangle \langle S^{00} f, h^1_I\rangle h_I |I|^{-1/2},
\end{aligned}
$$

which is the sum of two classical paraproducts of type $B(b, S^{00} f)$.

For part II, we would like to rewrite it as paraproducts composed with S^{00}. Observe that

$$
II = \sum_{I'\subsetneq I} \langle b, h_I\rangle h_I a_{I'} |I'|^{1/2} h_{I'} (\langle f, h^1_I\rangle h^1_I + \langle f, h_I\rangle h_I) := II' + II''.
$$

Since the two terms above are similar, we only estimate the first one.

$$
\begin{aligned}
II' &= \sum_{I'} a_{I'} |I'|^{1/2} h_{I'} \left(\sum_{I\supsetneq I'} \langle b, h_I\rangle |I|^{-1/2} \langle f, h^1_I\rangle h_I \right) \\
&:= \sum_{I'} a_{I'} |I'|^{1/2} h_{I'} \sum_{I\supsetneq I'} \langle S_b f, h_I\rangle h_I \\
&= \sum_{I'} a_{I'} \langle S_b f, h^1_{I'}\rangle h_{I'} = S^{00}(S_b f),
\end{aligned}
$$

where the operator $S_b f := \sum_I \langle b, h_I\rangle |I|^{-1/2} \langle f, h^1_I\rangle h_I$ is a paraproduct, and the boundedness of L^2 norm follows. This then completes the proof of the main theorem in the one-parameter setting.

2.3 Proof of the main theorem

In this section, we articulate the proof of the main theorem in the general setting by presenting estimates of several selective cases. We will show that the multi-parameter theorem can be obtained by an iteration of the one-parameter argument, so we can focus on the bi-parameter case as an example. The main idea is to show that the commutator can be written as a finite sum of either t-parameter paraproducts or such paraproducts composed with some one-parameter dyadic shifts in several variables. We present three cases to illustrate the strategy required in any of the cases that arise.

Use Theorem 10 twice for both variables we have

$$[[b,T_1],T_2]f = c\|T_1\|_{CZ}\|T_2\|_{CZ}\cdot\mathbb{E}_{\omega_1}\mathbb{E}_{\omega_2}\sum_{i_1,j_1=0}^{\infty}\sum_{i_2,j_2=0}^{\infty}2^{-\max(i_1,j_1)\delta/2}2^{-\max(i_2,j_2)\delta/2}[[b,S_{\omega_1}^{i_1j_1}],S_{\omega_2}^{i_2j_2}]f.$$

Since our estimate in the following doesn't depend on the parameters ω_1,ω_2 explicitly, we will omit them in the notation. And just like how we treat the one-parameter commutator, one split

$$\begin{aligned}
&[[b,S_1^{i_1j_1}],S_2^{i_2j_2}]f\\
&=\sum_{I_1,J_1}\sum_{I_2,J_2}\langle b,h_{I_1}\otimes u_{I_2}\rangle\langle f,h_{J_1}\otimes u_{J_2}\rangle[h_{I_1},S_1^{i_1j_1}]h_{J_1}\otimes[u_{I_2},S_2^{i_2j_2}]u_{J_2}\\
&=\sum_{I_1,J_1}\sum_{I_2,J_2}\langle b,h_{I_1}\otimes u_{I_2}\rangle\langle f,h_{J_1}\otimes u_{J_2}\rangle[h_{I_1}S_1^{i_1j_1}h_{J_1}\otimes u_{I_2}S_2^{i_2j_2}u_{J_2}\\
&\quad -h_{I_1}S_1^{i_1j_1}h_{J_1}\otimes S_2^{i_2j_2}(u_{I_2}u_{J_2})-S_1^{i_1j_1}(h_{I_1}h_{J_1})\otimes u_{I_2}S_2^{i_2j_2}u_{J_2}+S_1^{i_1j_1}(h_{I_1}h_{J_1})\otimes S_2^{i_2j_2}(u_{I_2}u_{J_2})],
\end{aligned}$$

where we use u_I to denote Haar function as well, but for the second variable. According to the cancellation of the commutator structure, the summand in the above is nonzero only when $I_1\subset J_1^{(i_1)}$ and $I_2\subset J_2^{(i_2)}$. Again, we discuss the cancellative and noncancellative cases separately.

2.3.1 Case $(i_1,j_1)\neq(0,0)$ and $(i_2,j_2)\neq(0,0)$

The goal in this case is to rewrite each of the four sums above into a bi-parameter paraproduct whose norm decays fast enough so that it can be summed. Since the steps involved are iterations of the one-parameter argument, we only present the details for one of the typical mixed terms. Before we start, we first state the following lemma which handles the boundedness of several variants of classical bi-parameter paraproducts.

lemma 8. *Given $b\in BMO(\mathbb{R}^n\times\mathbb{R}^m)$ and $k,l\geq 0$, define the following operators*

$$B_1(b,f)=\sum_{I,J}\pm\langle b,h_I\otimes u_J\rangle\langle f,h_{I^{(k)}}^{\epsilon_1}\otimes u_{J^{(l)}}^{\epsilon_2}\rangle h_I^{\epsilon_1'}\otimes u_J^{\epsilon_2'}|I^{(k)}|^{-1/2}|J^{(l)}|^{-1/2},$$

$$B_2(b,f)=\sum_{I,J}\pm\langle b,h_{I^{(k)}}\otimes u_J\rangle\langle f,h_I\otimes u_{J^{(l)}}^{\epsilon_2}\rangle h_I\otimes u_J^{\epsilon_2'}|I^{(k)}|^{-1/2}|J^{(l)}|^{-1/2},$$

$$B_3(b,f)=\sum_{I,J}\pm\langle b,h_I\otimes u_{J^{(l)}}\rangle\langle f,h_{I^{(k)}}^{\epsilon_1}\otimes u_J\rangle h_I^{\epsilon_1'}\otimes u_J|I^{(k)}|^{-1/2}|J^{(l)}|^{-1/2},$$

$$B_4(b,f)=\sum_{I,J}\pm\langle b,h_{I^{(k)}}\otimes u_{J^{(l)}}\rangle\langle f,h_I\otimes u_J\rangle h_I\otimes u_J|I^{(k)}|^{-1/2}|J^{(l)}|^{-1/2},$$

where at least one of ϵ_i, ϵ_i' is not $\vec{1}$, for $i = 1, 2$. Then, $\forall i = 1, 2, 3, 4$, $\|B_i(b, f)\|_{L^2} \lesssim \|b\|_{BMO}\|f\|_{L^2}$ with a constant independent of k, l.

The proof of the lemma is exactly the same as its one-parameter counterpart, except that for B_2, B_3 one uses hybrid square-maximal functions as majorization. We omit it here. Note that in multi-parameter setting, parallel results of the lemma still hold.

Now we begin to study the part of the sum corresponding to $S_1^{i_1j_1}(h_{I_1}h_{J_1}) \otimes u_{I_2}S_2^{i_2j_2}u_{J_2}$ when $i_1 \geq j_1, i_2 < j_2$ and $\ell(I_1) \leq 2^{i_1-j_1}\ell(J_1), 2^{i_2-j_2}\ell(J_2) < \ell(I_2) \leq 2^{i_2}\ell(J_2)$. The other parts can be handled similarly. First, in order to reorganize the second variable, one applies the one-parameter argument for the first term of the part $2^{i-j}\ell(J) < \ell(I) \leq 2^i\ell(J)$ in Section 2.2.1 to obtain

$$\begin{aligned}
&\sum_{J_1}\sum_{J_2}\sum_{\substack{I_1\subset J_1^{(i_1)}\\ \ell(I_1)\leq 2^{i_1-j_1}\ell(J_1)}}\sum_{\substack{I_2\subset J_2^{(i_2)}\\ \ell(I_2)>2^{i_2-j_2}\ell(J_2)}} \langle b, h_{I_1}\otimes u_{I_2}\rangle\langle f, h_{J_1}\otimes u_{J_2}\rangle S_1^{i_1j_1}(h_{I_1}h_{J_1})\otimes u_{I_2}S_2^{i_2j_2}u_{J_2}\\
&=\sum_{J_1}\sum_{\substack{I_1\subset J_1^{(i_1)}\\ \ell(I_1)\leq 2^{i_1-j_1}\ell(J_1)}} S_1^{i_1j_1}(h_{I_1}h_{J_1})\sum_{J_2}\sum_{\substack{I_2\subset J_2^{(i_2)}\\ \ell(I_2)>2^{i_2-j_2}\ell(J_2)}} \langle\langle b, h_{I_1}\rangle_1, u_{I_2}\rangle_2\langle\langle f, h_{J_1}\rangle_1, u_{J_2}\rangle_2 u_{I_2}S_2^{i_2j_2}u_{J_2}\\
&=\sum_{J_1}\sum_{\substack{I_1\subset J_1^{(i_1)}\\ \ell(I_1)\leq 2^{i_1-j_1}\ell(J_1)}} S_1^{i_1j_1}(h_{I_1}h_{J_1})\sum_{l=1}^{j_2}\sum_{J}\pm\langle\langle b, h_{I_1}\rangle_1, u_{J^{(l)}}\rangle_2\langle S_2^{i_2j_2}(\langle f, h_{J_1}\rangle_1), u_J\rangle_2 u_J|J^{(l)}|^{-1/2}\\
&=\sum_{l=1}^{j_2}\sum_{J} u_J|J^{(l)}|^{-1/2}\sum_{J_1}\sum_{\substack{I_1\subset J_1^{(i_1)}\\ \ell(I_1)\leq 2^{i_1-j_1}\ell(J_1)}} \pm\langle\langle b, u_{J^{(l)}}\rangle_2, h_{I_1}\rangle_1\langle\langle S_2^{i_2j_2}f, u_J\rangle_2, h_{J_1}\rangle_1 S_1^{i_1j_1}(h_{I_1}h_{J_1}).
\end{aligned}$$

Next, to deal with the first variable, one applies the one-parameter argument for the second term of the part $\ell(I) \leq 2^{i-j}\ell(J)$ in Section 2.2.1, which means we now need to discuss two different cases : $I_1 \subset J_1$ or $J_1 \subsetneq I_1 \subset J_1^{(i_1-j_1)}$. As they are very similar, we only study the first one as an example. The corresponding one-parameter technique gives us

$$\begin{aligned}
&\sum_{l=1}^{j_2}\sum_{J} u_J|J^{(l)}|^{-1/2}S_1^{i_1j_1}(\sum_{I}\pm\langle\langle b, u_{J^{(l)}}\rangle_2, h_I\rangle_1\langle\langle S_2^{i_2j_2}f, u_J\rangle_2, h_I^1\rangle_1 h_I|I|^{-1/2}+\\
&\qquad\sum_{I}\pm\langle\langle b, u_{J^{(l)}}\rangle_2, h_I\rangle_1\langle\langle S_2^{i_2j_2}f, u_J\rangle_2, h_I\rangle_1 h_I^{\epsilon_1}|I|^{-1/2})\\
&=\sum_{l=1}^{j_2} S_1^{i_1j_1}(\sum_{I,J}\pm\langle b, h_I\otimes u_{J^{(l)}}\rangle\langle S_2^{i_2j_2}f, h_I^1\otimes u_J\rangle h_I\otimes u_J|I|^{-1/2}|J^{(l)}|^{-1/2}+\\
&\qquad\sum_{I,J}\pm\langle b, h_I\otimes u_{J^{(l)}}\rangle\langle S_2^{i_2j_2}f, h_I\otimes u_J\rangle h_I^{\epsilon_1}\otimes u_J|I|^{-1/2}|J^{(l)}|^{-1/2}),
\end{aligned}$$

which is dyadic shift $S_1^{i_1j_1}$ acting on two paraproducts of type $B_3(b, S_2^{i_2j_2}f)$, whose boundedness can be derived from Lemma 8. And the constant j_2 in front won't matter thanks to the decaying factor $2^{-\max(i_2,j_2)\delta/2}$.

2.3.2 Case $(i_1, j_1) = (0, 0)$ or $(i_2, j_2) = (0, 0)$

In this section, we take two noncancellative shifts S_1^{00} and S_2^{00} as an example, as the mixed cancellative-noncancellative cases are even simpler. More specifically, we will study the term corresponding to $h_{I_1} S_1^{00} h_{J_1} \otimes u_{I_2} S_2^{00} u_{J_2}$, while the other ones can be dealt with using the techniques from this section together with the previous one. From the proof of the representation theorem, one may end up with dyadic shifts which are paraproducts. Moreover, recall that the summands are nonzero only if $I_1 \subset J_1, I_2 \subset J_2$ due to cancellation. We will now discuss two different cases depending on whether the two dyadic shifts are chosen to be of the same or different types of paraproducts.

Type I

In this case, we assume that

$$S_1^{00} f = \sum_I a_I^1 \langle f, h_I^1 \rangle h_I, \quad S_2^{00} f = \sum_J a_J^2 \langle f, u_J^1 \rangle u_J.$$

Since both of the dyadic shifts are of the same type, we can iterate the one-parameter argument to derive the desired result. Write

$$\begin{aligned}
&\sum_{I_1 \subset J_1} \sum_{I_2 \subset J_2} \langle b, h_{I_1} \otimes u_{I_2} \rangle \langle f, h_{J_1} \otimes u_{J_2} \rangle h_{I_1} S_1^{00} h_{J_1} \otimes u_{I_2} S_2^{00} u_{J_2} \\
&= \sum_{I_1 \subset J_1} \sum_{I' \subsetneq J_1} \sum_{I_2 \subset J_2} \sum_{J' \subsetneq J_2} \langle b, h_{I_1} \otimes u_{I_2} \rangle \langle f, h_{J_1} \otimes u_{J_2} \rangle a_{I'}^1 |I'|^{1/2} a_{J'}^2 |J'|^{1/2} h_{I_1} h_{I'} h_{J_1} \otimes u_{I_2} u_{J'} u_{J_2} \\
&= \left(\sum_{I_1 = I' \subsetneq J_1} + \sum_{I_1 \subsetneq I' \subsetneq J_1} + \sum_{I' \subsetneq I_1 \subsetneq J_1} + \sum_{I' \subsetneq I_1 = J_1} \right) \left(\sum_{I_2 = J' \subsetneq J_2} + \sum_{I_2 \subsetneq J' \subsetneq J_2} + \sum_{J' \subsetneq I_2 \subsetneq J_2} + \sum_{J' \subsetneq I_2 = J_2} \right).
\end{aligned}$$

We only estimate the mixed case $\sum_{I_1 \subsetneq I' \subsetneq J_1} \sum_{J' \subsetneq I_2 \subsetneq J_2}$, as the other parts can be handled similarly. The strategy here is to first reorganize the first variable to move the shift S_1^{00} into the pairing to act on f, then for the second variable to rewrite the full sum as a bi-parameter paraproduct composed with S_2^{00}. Note that this technique can be applied to handle the multi-parameter commutators with ease. To be specific,

$$\begin{aligned}
&\sum_{I_1 \subsetneq I' \subsetneq J_1} \sum_{J' \subsetneq I_2 \subsetneq J_2} \langle b, h_{I_1} \otimes u_{I_2} \rangle \langle f, h_{J_1} \otimes u_{J_2} \rangle a_{I'}^1 |I'|^{1/2} a_{J'}^2 |J'|^{1/2} h_{I_1} h_{I'} h_{J_1} \otimes u_{I_2} u_{J'} u_{J_2} \\
&= \sum_{J' \subsetneq I_2 \subsetneq J_2} a_{J'}^2 |J'|^{1/2} u_{I_2} u_{J'} u_{J_2} \sum_{I_1} \langle \langle b, u_{I_2} \rangle_2, h_{I_1} \rangle_1 h_{I_1} \left(\sum_{I_1 \subsetneq I' \subsetneq J_1} a_{I'}^1 |I'|^{1/2} \langle \langle f, u_{J_2} \rangle_2, h_{J_1} \rangle_1 h_{J_1} h_{I'} \right) \\
&= \sum_{J' \subsetneq I_2 \subsetneq J_2} a_{J'}^2 |J'|^{1/2} u_{I_2} u_{J'} u_{J_2} \sum_I \langle \langle b, u_{I_2} \rangle_2, h_I \rangle_1 \langle S_1^{00}(\langle f, u_{J_2} \rangle_2), h_I^1 \rangle_1 h_I |I|^{-1/2},
\end{aligned}$$

where the last step follows from the one-parameter argument for part I in Section 2.2.2. Next, rewrite the above as

$$\begin{aligned}
&\sum_I h_I |I|^{-1/2} \sum_{J' \subsetneq I_2} a^2_{J'} |J'|^{1/2} u_{I_2} u_{J'} \langle b, h_I \otimes u_{I_2}\rangle \langle\langle S^{00}_1 f, h^1_I\rangle_1, u^1_{I_2}\rangle_2 u^1_{I_2} \\
&= \sum_{J'} a^2_{J'} |J'|^{1/2} u_{J'} \sum_I h_I |I|^{-1/2} \sum_{I_2 \supsetneq J'} (\langle b, h_I \otimes u_{I_2}\rangle |I_2|^{-1/2}) \langle\langle S^{00}_1 f, h^1_I\rangle_1, u^1_{I_2}\rangle_2 u_{I_2} \\
&:= \sum_{J'} a^2_{J'} |J'|^{1/2} u_{J'} \sum_I h_I |I|^{-1/2} \sum_{I_2 \supsetneq J'} \langle S^I(\langle S^{00}_1 f, h^1_I\rangle_1), u_{I_2}\rangle_2 u_{I_2},
\end{aligned}$$

where for any I, $S^I f := \sum_J b^I_J \langle f, u^1_J\rangle_2 u_J$ with $b^I_J := \langle b, h_I \otimes u_J\rangle |J|^{-1/2}$. The operator S^I here can be thought of as a dyadic shift in the second variable associated with a fixed cube I in the first variable. Then, the above equals

$$\begin{aligned}
&\sum_{J'} a^2_{J'} |J'|^{1/2} u_{J'} \sum_I h_I |I|^{-1/2} \langle S^I(\langle S^{00}_1 f, h^1_I\rangle_1), u^1_{J'}\rangle_2 |J'|^{-1/2} \\
&= S^{00}_2 (\sum_I h_I |I|^{-1/2} S^I(\langle S^{00}_1 f, h^1_I\rangle_1)) \\
&= S^{00}_2 (\sum_I h_I |I|^{-1/2} \sum_J b^I_J \langle\langle S^{00}_1 f, h^1_I\rangle_1, u_J\rangle_2 u_J) \\
&= S^{00}_2 (\sum_{I,J} \langle b, h_I \otimes u_J\rangle \langle S^{00}_1 f, h^1_I \otimes u_J\rangle h_I \otimes u_J |I|^{-1/2} |J|^{-1/2}),
\end{aligned}$$

where the last item is S^{00}_2 acting on a classical bi-parameter paraproduct of type $B(b, S^{00}_1 f)$.

Type II

Now we discuss a mixed case where

$$S^{00}_1 f = \sum_I a^1_I \langle f, h^1_I\rangle h_I, \quad S^{00}_2 f = \sum_J a^2_J \langle f, u_J\rangle u^1_J.$$

The first half of the argument is devoted to move S^{00}_1 into the pairing to act on f, exactly the same as in Type I. But for the second variable, one needs to argue by duality instead. Write

$$\begin{aligned}
&\sum_{I_1 \subset J_1} \sum_{I_2 \subset J_2} \langle b, h_{I_1} \otimes u_{I_2}\rangle \langle f, h_{J_1} \otimes u_{J_2}\rangle h_{I_1} S^{00}_1 h_{J_1} \otimes u_{I_2} S^{00}_2 u_{J_2} \\
&= \sum_{I_1 \subset J_1} \sum_{I' \subsetneq J_1} \sum_{I_2 \subset J_2} \langle b, h_{I_1} \otimes u_{I_2}\rangle \langle f, h_{J_1} \otimes u_{J_2}\rangle a^1_{I'} |I'|^{1/2} h_{I_1} h_{I'} h_{J_1} \otimes u_{I_2} S^{00}_2 u_{J_2} \\
&= \left(\sum_{I_1 = I' \subsetneq J_1} + \sum_{I_1 \subsetneq I' \subsetneq J_1} + \sum_{I' \subsetneq I_1 \subsetneq J_1} + \sum_{I' \subsetneq I_1 = J_1} \right) \left(\sum_{I_2 \subsetneq J_2} + \sum_{I_2 = J_2} \right).
\end{aligned}$$

We only estimate the case $\sum_{I_1\subsetneq I'\subsetneq J_1}\sum_{I_2\subsetneq J_2}$, as the other parts are similar. The argument in the previous case implies

$$\begin{aligned}&\sum_{I_1\subsetneq I'\subsetneq J_1}\sum_{I_2\subsetneq J_2}\langle b, h_{I_1}\otimes u_{I_2}\rangle\langle f, h_{J_1}\otimes u_{J_2}\rangle a^1_{I'}|I'|^{1/2}h_{I_1}h_{I'}h_{J_1}\otimes u_{I_2}S_2^{00}u_{J_2}\\ &=\sum_{I_2\subsetneq J_2}u_{I_2}S_2^{00}u_{J_2}\sum_I\langle\langle b, u_{I_2}\rangle_2, h_I\rangle_1\langle S_1^{00}(\langle f, u_{J_2}\rangle_2), h_I^1\rangle_1 h_I|I|^{-1/2}.\end{aligned}$$

Next, let $g\in L^2(\mathbb{R}^n\times\mathbb{R}^m)$ with norm 1 be the function pairing with which the above achieves its L^2 norm. We have

$$\begin{aligned}&\langle\sum_{I_2\subsetneq J_2}u_{I_2}S_2^{00}u_{J_2}\sum_I\langle\langle b, u_{I_2}\rangle_2, h_I\rangle_1\langle S_1^{00}(\langle f, u_{J_2}\rangle_2), h_I^1\rangle_1 h_I|I|^{-1/2}, g\rangle\\ &=\langle\sum_{I_2\subsetneq J_2}\sum_I u_{I_2}S_2^{00}u_{J_2}\langle\langle b, u_{I_2}\rangle_2, h_I\rangle_1\langle S_1^{00}(\langle f, u_{J_2}\rangle_2), h_I^1\rangle_1 h_I|I|^{-1/2}, \sum_{K_1,K_2}\langle g, h_{K_1}\otimes u_{K_2}\rangle h_{K_1}\otimes u_{K_2}\rangle\\ &=\sum_I\sum_{I_2\subsetneq J_2}\sum_{K_1,K_2}\langle b, h_I\otimes u_{I_2}\rangle\langle S_1^{00}f, h_I^1\otimes u_{J_2}\rangle\langle g, h_{K_1}\otimes u_{K_2}\rangle\langle h_I|I|^{-1/2}\otimes u_{I_2}S_2^{00}u_{J_2}, h_{K_1}\otimes u_{K_2}\rangle\\ &=\sum_I\sum_{I_2\subsetneq J_2}|I|^{-1/2}\langle b, h_I\otimes u_{I_2}\rangle\langle S_1^{00}f, h_I^1\otimes u_{J_2}\rangle\langle g, h_I\otimes u_{I_2}\rangle\langle u_{I_2}S_2^{00}u_{J_2}, u_{I_2}\rangle_2,\end{aligned}$$

where the last step is because $u_{I_2}S_2^{00}u_{J_2}=a^2_{J_2}|J_2|^{-1/2}u_{I_2}$. Then, one can rewrite the above as

$$\begin{aligned}&\langle S_1^{00}f,\sum_I\sum_{I_2\subsetneq J_2}|I|^{-1/2}\langle u_{J_2}, S_2^{00*}(u^2_{I_2})\rangle_2\langle b, h_I\otimes u_{I_2}\rangle\langle g, h_I\otimes u_{I_2}\rangle h_I^1\otimes u_{J_2}\rangle\\ &=\langle S_1^{00}f,\sum_{I,J}\langle b, h_I\otimes u_J\rangle\langle g, h_I\otimes u_J\rangle|I|^{-1/2}\langle u_J^1, S_2^{00*}(u_J^2)\rangle_2 h_I^1u_J^1\rangle.\end{aligned}$$

Since the boundedness of dyadic shifts implies $|\langle u_J^1, S_2^{00*}(u_J^2)\rangle_2|\le|J|^{-1/2}$, the above pairing which equals the L^2 norm of the sum can be estimated by a bi-parameter paraproduct argument. And this finishes the proof of the main theorem.

2.4 Proof of the Corollary

We end the paper with a proof of the perturbation result of Corollary 2.

Proof. Given $(T_{i,s_i})_{1\le i\le t,1\le s_i\le n_i}$ and an arbitrary family of operators $(T'_{i,s_i})_{1\le i\le t,1\le s_i\le n_i}$, for any $1\le i\le t$, let s_i, $1\le s_i\le n_i$ be fixed. Then

$$\begin{aligned}&[\dots[[M_b, T_{1,s_1}+T'_{1,s_1}], T_{2,s_2}+T'_{2,s_2}]\dots, T_{t,s_t}+T'_{t,s_t}]\\ &=[\dots[[M_b, T_{1,s_1}], T_{2,s_2}]\dots, T_{t,s_t}]+\sum_{j\in\Lambda}[\dots[[M_b, T^j_{1,s_1}], T^j_{2,s_2}]\dots, T^j_{t,s_t}],\end{aligned}$$

where Λ is a finite index set, $T^j_{i,s_i}=T_{i,s_i}$ or T'_{i,s_i}, and $\forall j\in\Lambda$, $\exists 1\le i\le t$ s.t. $T^j_{i,s_i}=T'_{i,s_i}$.

By assumption,

$$C_1\|b\|_{BMO} \leq \sup_{1\leq i\leq t, 1\leq s_i\leq n_i} \|[\dots[[M_b, T_{1,s_1}], T_{2,s_2}]\dots, T_{t,s_t}]\|_{L^2\to L^2} \leq C_2\|b\|_{BMO},$$

and Theorem 9 implies

$$\|[\dots[[M_b, T^j_{1,s_1}], T^j_{2,s_2}]\dots, T^j_{t,s_t}]\|_{L^2\to L^2} \leq C\|b\|_{BMO} \prod_{i=1}^{t} \|T^j_{i,s_i}\|_{CZ}.$$

Hence, there exists sufficiently small $\epsilon > 0$, such that $\|T'_{i,s_i}\|_{CZ} \leq \epsilon$ implies

$$\sum_{j\in\Lambda} \|[\dots[[M_b, T^j_{1,s_1}], T^j_{2,s_2}]\dots, T^j_{t,s_t}]\|_{L^2\to L^2} \leq \frac{C_1}{2}\|b\|_{BMO}, \forall 1\leq i\leq t, 1\leq s_i\leq n_i,$$

Then, by triangular inequality and taking the supremum we have

$$\begin{aligned}\frac{C_1}{2}\|b\|_{BMO} &\leq \sup_{1\leq i\leq t, 1\leq s_i\leq n_i} \|[\dots[[M_b, T_{1,s_1}+T'_{1,s_1}], T_{2,s_2}+T'_{2,s_2}]\dots, T_{t,s_t}+T'_{t,s_t}]\|_{L^2\to L^2}\\ &\leq \left(C_2+\frac{C_1}{2}\right)\|b\|_{BMO},\end{aligned}$$

which completes the proof of the first assertion. While the second assertion follows easily because Calderón-Zygmund operators form a Banach space under the CZ norm.
□

Bibliographie

[1] Carlos Cabrelli, Michael Lacey, Ursula Molter, Jill Pipher, *Variations on the theme of Journe's lemma*, Houston J. Math **32** (3) (2006) 833-861.

[2] L. Carleson, *A counterexample for measures bounded on H^p spaces for the bidisk*, Mittag-Leffler Rep. No. 7, Inst. Mittag-Leffler, 1974.

[3] Sun-Yung A. Chang, Robert Fefferman, *A continuous version of duality of H^1 with BMO on the bidisk*, Ann. of Math. (2) **112** (1) (1980) 179-201.

[4] Sun-Yung A. Chang, Robert Fefferman, *Some recent developments in Fourier analysis and H^p-theory on product domains*, Bull. Amer. Math. Soc. (N.S.) **12** (1) (1985) 1-43.

[5] R. Coifman, R. Rochberg and G. Weiss, *Factorization theorems for Hardy spaces in several variables*, Ann. of Math. (2) **103** (1976), no. 3, 611-635.

[6] Sarah H. Ferguson, Michael T. Lacey, *A characterization of product BMO by commutators*, Acta Math. **189** (2) (2002) 143-160.

[7] S. Ferguson and C. Sadosky, *Characterizations of bounded mean oscillation on the polydisk in terms of Hankel operators and Carleson measures*, J. Anal. Math. **81** (2000), 239-267.

[8] Loukas Grafakos, *Classical Fourier Analysis*, second edition, Grad. Texts in Math., vol. 249, Springer, New York, ISBN 978-0-387-09431-1, 2008,xvi+489pp.

[9] T. Hytönen, *Representation of singular integrals by dyadic operators, and the A_2 theorem*, preprint (2011).

[10] T. Hytönen, *The sharp weighted bound for general Calderón-Zygmund operators*, Ann. of Math. (2) **175** (2012), no. 3, 1473-1506.

[11] J. Journé, *Calderón-Zygmund operators on product spaces*, Rev. Mat. Iberoamericana **1** (1985), no. 3, 55-91.

[12] J. Journé *A covering lemma for product spaces*, Proc. Amer. Math. Soc. **96** (4) (1986) 593-598.

[13] S. Ferguson and M. Lacey, *A characterization of product BMO by commutators*, Acta Math. **189** (2002), no. 2, 143-160.

[14] M. Lacey, S. Petermichl, J. Pipher and B. Wick, *Multiparameter Riesz commutators*, Amer. J. Math. **35** (2009), no. 1, 159-183.

[15] M. Lacey, S. Petermichl, J. Pipher and B. Wick, *Iterated Riesz commutators : a simple proof of boundedness*, Contemp. Math., 505, Amer. Math. Soc., Providence, RI (2010), 171-178.

[16] M. Lacey and E. Terwilleger, *Hankel operators in several complex variables and product BMO*, Houston J. Math. **35** (2009), no. 1, 159-183.

[17] Song-Ying Li, *Characterization of the boundedness for a family of commutators on L^p*,Colloq. Math. **70** (1) (1996) 59-71.

[18] Yves Meyer *Wavelets and operators*, Cambridge University Press , Cambridge studies in advanced mathematics 37, (1992).

[19] Yves Meyer *Ondelettes et opérateurs 2 : Opérateurs de Calderon-Zygmund* Hermann, (1990) ISBN 2 7005661267.

[20] C. Muscalu, J. Pipher, T. Tao and C. Thiele, *Bi-parameter paraproducts*, Acta Math. **193** (2004), no. 2, 269-296.

[21] C. Muscalu, J. Pipher, T. Tao and C. Thiele, *Multi-parameter paraproducts*, Rev. Mat. Iberoamericana **22** (2006), no. 3, 963-976.

[22] Leopoldo Nachbin, *Sur les algèbres denses de fonctions différentiables sur une variété*, C.R. Acad. Sci. Paris **228** (1949) 1549-1551.

[23] Zeev Nehari, *On bounded bilinear forms*, Ann. of Math. (2) **65** (1957) 153-162.

[24] Jill Pipher, *Journé's covering lemma and its extension to higher dimension*, Duke Math. J. **53** (3) (1986) 683-690.

[25] Elias M. Stein, *Singular Integrals and Differential Properties of Functions*, Princeton University Press, (1970).

[26] Akihito Uchiyama, *On the compactness of operators of Hankel type* Tôhoku Math. J. (2) **30** (1) (1978) 163-171.

[27] Akihito Uchiyama, *A constructive proof of the Fefferman-Stein decomposition of $BMO(\mathbb{R}^n)$*, Acta. Math. **148** (1982) 215-241.

[28] Akihito Uchiyama, *On the characterization of $H^p(\mathbb{R}^n)$ in terms of Fourier multipliers*, Proc. Amer. Soc. **109** (1) (1990) 117-123.

Abstract

The aim of my thesis was to find criteria on families of Calderón-Zygmund operators to know if, with iterated commutators, they characterize product BMO space. Multiparameter BMO space is a generalization of classical BMO space, and began to be studied during the eighties by Chang and Fefferman.
To each parameter is associated a Calderón-Zygmund operator acting on this parameter. We define also, associated to b in BMO, the operator M_b of multiplication by b. Then we define iterated commutators with those Calderón-Zygmund operators and the operator M_b. Then the aim is to study the relation between the BMO norm of b and the norm of the commutator acting on L^2.
The first result in one parameter case is due to Coifman, Rochberg and Weiss, who proved that Riesz transforms characterize BMO.
The next result is due to Uchiyama, who gave a criterion on families of Calderon-Zygmund operators to show if they generalize Stein-Fefferman decomposition. Then Li gave another criterion on those families of Calderon-Zygmund operators, to show if commutators characterize BMO.
The first result in multiparameter case is due to Ferguson-Lacey, who proved in the case of parameter t=2 that Hilbert transform characterize BMO. Then Lacey and Terwilleger extended this result to arbitraly number of iterations.
Finally, Lacey-Petermichl-Wick-Pipher extended this result to Riesz transform in product BMO space.
So, first, I found a criterion on families of Calderón-Zygmund operators to know if they characterize product BMO space.
Finally, I proved that commutators norms are majorized, up to a multiplicative constant, by BMO norm of b in multiparameter case for any kind of Calderón-Zygmund commutators, using the representation theorem of Hytonen, which reduces the problem to dyadic shifts.

Résumé

Le but de ma thèse est de décrire les familles d'opérateurs de Calderón-Zygmund qui, imbriqués au sein de commutateurs, caractérisent BMO à plusieurs paramètres. L'espace BMO à plusieurs paramètres est une généralisation de l'espace BMO classique, et a commencé à être étudié au cours des années 1980 par Chang et Fefferman. A chaque paramètre, on associe un opérateur de Calderón-Zygmund agissant sur ce paramètre, un opérateur de Calderón-Zygmund étant un opérateur à noyau. Ensuite, si b appartient à BMO, on lui associe l'opérateur M_b de multiplication par b. On considère ensuite une suite d'itérés de commutateurs ayant pour argument ces opérateurs de Calderón-Zygmund et M_b.
Le but est alors d'étudier le rapport entre la norme BMO de b et celle de ces itérés de commutateurs agissant sur L^2.
Le premier résultat concernant cette théorie est du à Coifman, Rochberg et Weiss qui ont démontré dans le cas du paramètre un que les transformées de Riesz, qui sont des opérateurs de Calderón-Zygmund, caractérisent BMO.
Le résultat suivant est du à Uchiyama, qui, lui, a proposé un critère portant sur une famille d'opérateurs de Calderón-Zygmund, pour savoir s'ils généralisent la décomposition de Stein-Fefferman, puis Li a fourni un critère englobant celui de Uchiyama pour savoir si un commutateur caractérise BMO à un paramètre.
Le premier théorème dans le cas du multiparamètre est du à Ferguson-Lacey qui ont montré dans le cas du paramètre t=2 que les transformées de Hilbert caractérisent BMO, puis Lacey-Ferguson l'on étendu à un nombre quelconque d'itérations.
Enfin, Lacey-Petermichl-Wick-Pipher ont étendu ce résultat au transformées de Riesz dans le cas du multiparamètre.
C'est, dans un premier temps, ce résultat que l'on a généralisé, fournissant un critère permettant de savoir si une famille d'opérateurs de Calderón-Zygmund caractérisent BMO à plusieurs paramètres.
Enfin, nous avons montré que la norme du commutateur est, à une constante multiplicative près, majorée par la norme BMO de b pour n'importe quel type d'opérateurs de Calderón-Zygmund, en utilisant le théorème de représentation de Hytonen qui permet de réduire le problème au cas des shifts dyadiques