Weighted LP estimates on Riemannian manifolds

Kamilia Dahmani

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le *7 Décembre, 2018* **par :**

KAMILIA DAHMANI

Weighted L^p estimates on Riemannian manifolds

JURY

OLIVIER GUÉDON	Université Paris-Est Marne La Vallée	Rapporteur
STEFANIE PETERMICHL	Université Toulouse 3	Directrice de thèse
BERNHARD HAAK	Université de Bordeaux	Examinateur
KOMLA DOMELEVO	Université Toulouse 3	Co-Directeur de thèse
EMMANUEL RUSS	Université Grenoble Alpes	Rapporteur

École doctorale et spécialité :
MITT : Domaine Mathématiques : Mathématiques appliquées
Unité de Recherche :
Institut de Mathématiques de Toulouse (UMR 5219)
Directeur(s) de Thèse :
Stefanie PETERMICHL et *Komla DOMELEVO*
Rapporteurs :
Olivier Guédon et *Emmanuel RUSS*

Contents

Remerciements

J'aimerais tout d'abord remercier Stefanie Petermichl pour m'avoir pris sous son aile pendant ces trois années, m'avoir proposé un sujet sur lequel j'ai adoré travailler avec elle et pour m'avoir fait voyager au sens figuré dans le monde de la recherche mais également au sens littéral à travers le monde afin de me former auprès d'experts en analyse harmonique.

Je tiens à remercier également Komla Domelevo pour m'avoir fait découvrir l'analyse stochastique, pour avoir répondu à toutes mes questions avec beaucoup de pédagogie, pour avoir enrichi ma culture en géométrie différentielle ainsi que pour tous ses conseils lors de la rédaction de ce manuscrit.

J'aimerais remercier mes amis/collègues du bureau principal 1R3-202 qui ont rendu cette aventure inoubliable : à Victor pour ses nombreux conseils, à Sourav pour toutes nos discussions de fin de journée, à Susely pour avoir volé à mon secours à chaque fois que Latex faisait des siennes et enfin à Anthony pour sa touche orientale. Sachez que vous me manquerez tous énormément (mais ne t'inquiète pas Anthony, toi un peu plus ;-)).

Je remercie également tous les autres thésards qui m'ont côtoyés pour leur accueil, les aides dans les tâches administratives, la préparation des enseignements, et j'en passe. Je souhaite à ceux qui soutiendront après moi de s'épanouir dans leurs projets et de réussir.

Durant ces trois années j'ai eu la chance de renconter beaucoup de doctorants et post doctorants à Toulouse mais aussi ailleurs et avec lesquels je me suis beaucoup amusée. Je souhaite donc remercier Pierre et Kevin pour les débats sur les formations à suivre à la mécanique des fluides, mais aussi Alex pour son invitation au laboratoire DIAM de Delft, Queen Lashi pour nos nombreuses conversations philosophiques, Clément et Marine pour leur accueil au laboratoire de mathématiques à Besançon et enfin Valentin pour les bons moments partagés en conférences. Je remercie également Kristina avec qui j'ai adoré parler de mathématiques et faire des to-do list.

Ma gratitude s'adresse aussi à Catherine Stasiulis pour toute son aide, pour m'avoir reçue avec beaucoup de gentillesse et toujours avec le sourire. Je remercie également Mihai Maris, dont le bureau m'était toujours ouvert, pour ses nombreux conseils.

Un grand merci à El Maati Ouhabaz pour avoir suivi mon parcours depuis le début du Master et à Frédéric Bernicot pour m'avoir invité au laboratoire Jean Leray afin de partager mon intérêt pour les estimations à poids.

Emmanuel Russ et Olivier Guédon m'ont fait l'honneur de bien vouloir consacrer du temps à la lecture de ce manuscrit. Je tiens à les remercier pour l'intérêt qu'ils ont porté à mes travaux. Je remercie vivement Bernhard Haak pour avoir accepté de faire partie de mon jury, mais aussi pour avoir rendu les Travaux Dirigés d'analyse fonctionnelle auxquels j'ai assisté aussi intéressants !

J'ai une profonde pensée pour Maëva qui a été une jum exceptionnelle malgré tout le stress qu'elle a dû subir elle même, mais aussi à mes amies de longue date (Mouna, Diane, Sasha) qui ont toujours trouvé le fin mot lorsque le moral n'était pas au plus haut. Merci à Marie pour avoir été à mes côtés ces derniers mois et pour tout son encouragement.

Une mention particulière également à Jimmy : supporter un(e) thésard(e) n'est jamais chose simple. Merci pour ton soutien sans faille et ta bonne humeur au quotidien. Merci également de me rappeler que la distance n'est qu'un prétexte de plus pour se voir plus souvent.

Enfin, je remercie de tout coeur mes parents pour avoir toujours cru en moi, pour leur soutien infini et pour avoir formé le meilleur fan club au monde. Ma babushka également qui, malgré la distance, m'a envoyé toutes les ondes positives du monde afin que je réussisse au mieux mon parcours scolaire et ce, depuis mes 5 ans.

Abstract

The topics addressed in this thesis lie in the field of harmonic analysis and more precisely, weighted inequalities. Our main interests are the weighted L^p-bounds of the Riesz transforms on complete Riemannian manifolds and the sharpness of the bounds in terms of the power of the characteristic of the weights. We first obtain a linear and dimensionless result on non necessarily homogeneous spaces, when $p = 2$ and the Bakry-Emery curvature is non-negative. We use here an analytical approach by exhibiting a concrete Bellman function. Next, using stochastic techniques and sparse domination, we prove that the Riesz transforms are L^p-bounded for $p \in (1, +\infty)$ and obtain the previous result for free. Finally, we use an elegant change in the precedent proof to weaken the condition on the curvature and assume it is bounded from below.

Keywords: Riesz transforms, weighted inequalities, Bellman functions, Bakry-Emery curvature, sparse operators, stochastic representation of the Riesz transform.

Résumé

Cette thèse s'inscrit dans le domaine de l'analyse harmonique et plus exactement, des estimations à poids. Un intérêt particulier est porté aux estimations L^p à poids des transformées de Riesz sur des variétés Riemanniennes complètes ainsi qu'à l'optimalité des résultats en terme de la puissance de la caractéristique des poids. On obtient un premier résultat (en terme de la linéarité et de la non dépendance de la dimension) sur des espaces pas nécessairement de type homogène, lorsque $p = 2$ et la courbure de Bakry-Emery est positive. On utilise pour cela une approche analytique en exhibant une fonction de Bellman concrète. Puis, en utilisant des techniques stochastiques et une domination éparse, on démontre que les transformées de Riesz sont bornées sur L^p, pour $p \in (1, +\infty)$ et on déduit également le résultat précèdent. Enfin, on utilise un changement élégant dans la preuve précèdente pour affaiblir l'hypothèse sur la courbure et la supposer minorée.

Mots-clés: Transformées de Riesz, inegalités à poids, fonctions de Bellman, courbure de Bakry-Emery, opérateurs épars, représentation stochastique des transformées de Riesz.

Chapter 1

Introduction

1.1 Motivation (version française)

1.1.1 Contexte historique

On s'intéresse dans cette thèse à la continuité dans des espaces L^p pondérés des transformées de Riesz dans des variétés Riemanniennes.

L'étude de l'estimation L^p de la norme de la transformée de Hilbert sur la droite réelle a commencé à partir des travaux de Riesz [74] et Pichorides [71]. En effet, en 1972 Pichorides a montré que la constante optimale pour la norme de la transformée de Hilbert H est donnée par

$$C_p = \begin{cases} 2\tan\dfrac{\pi}{2p} & \text{si } 1 < p \leq 2; \\ \cot\dfrac{\pi}{2p} & \text{si } 2 < p < \infty. \end{cases}$$

Dans le cadre Euclidien, la i-ème transformée de Riesz sur $\mathbb{R}^n$ est définie par

$$R_i = \frac{\partial}{\partial x_i}(-\Delta)^{-1/2},$$

où $\Delta = \sum_{i=1}^n \partial^2/\partial^2_{x_i}$ est la Laplacien usuel sur $\mathbb{R}^n$. Le vecteur de la transformée de Riesz R est défini comme étant la collection $R = (R_1, R_2, \ldots, R_n)$. Notons que dans le cas où la dimension vaut 1, la transformée de Riesz est la transformée de Hilbert.

Concernant les estimations L^p de la norme du vecteur de Riesz sur $\mathbb{R}^n$, T. Iwaniec et G. Martin ont prouvé dans [46] que pour tout $1 < p < \infty$, il existe une constante $C_p > 0$ indépendante de n telle que

$$\|R_j f\|_p \leq C_p \|f\|_p,\ \forall j = 1, \cdots, n.$$

Cette constante est égale à la constante de Pichorides mentionnée plus haut. Quant au vecteur de Riesz, le meilleur résultat connu à ce jour a été présenté par Bañuelos et

Wang dans [9]

$$\|Rf\|_p \leq 2(p^* - 1)\|f\|_p, \ \forall \ p \in (1, \infty), \text{ où } p^* = \max(p, \frac{p}{p-1}).$$

On réfère également le lecteur aux travaux apparus dans [77, 60, 72, 25].

En 1960, utilisant des outils d'analyse complexe, Helson et Szego ont démontré dans [38] que la transformée de Hilbert est bornée sur $L^2(\omega)$ si, et seulement si le poids ω pouvait s'écrire sous la forme $\omega = \exp(\varphi + H\psi)$, où $\varphi, \psi \in L^\infty$ et $\|\psi\|_{L^\infty} < \pi/2$.
C'est en 1973 que Hunt, Muckenhoupt et Wheeden ont démontré dans [41] que la condition A_p caractérise également la continuité de la transformée de Hilbert sur $L^p(\omega)$. Plus précisément, ils ont démontré que H est bornée sur $L^p(\omega), p \in (1, \infty)$ si, et seulement si ω appartient à la classe A_p des poids. C'est-à-dire que la caractéristique du poids, notée $Q_p(\omega)$, est finie, où

$$Q_p(\omega) := \sup_Q \left(\frac{1}{|Q|}\int_Q \omega(x)\,dx\right)\left(\frac{1}{|Q|}\int_Q \omega^{-\frac{1}{p-1}}(x)\,dx\right)^{p-1} < \infty,$$

et où le sup est pris sur toutes les boules $Q \subset \mathbb{R}$.

Les inégalités pondérées (ou à poids) sont intervenues naturellement en analyse avec l'apparition de la théorie des intégrales singulières. Dans la théorie des EDP par exemple, les poids apparaissent dans l'étude des EDP à coefficients dégénérés, celles à domaines de géométrie non lisse ou encore dans les équations avec données initiales rugueuses (rough initial data). Sur $\mathbb{R}^n$, le prototype des poids A_p sont les fonctions puissances. En effet, en utilisant les intégrales de Riemann impropres, il est aisé de voir que pour tout $a \in \mathbb{R}$ et pour $p > 1$, $|x|^a \in A_p$ si, et seulement si $-n < a < (p-1)n$.

Durant les 15 dernières années, un grand intérêt a été apporté aux estimations optimales des normes des opérateurs d'intégrale singulière T, et ce en fonction de la caractéristique A_p du poids, $Q_p(\omega)$. Le but est de démontrer des inégalités de la forme

$$\|Tf\|_{L^p(\omega)} \leq C Q_p(\omega)^r \|f\|_{L^p(\omega)}$$

pour un certain r, et où la constante C est indépendante de f et de ω. Puisque $Q_p(\omega) > 1$, l'objectif est de trouver des estimations où r est le plus petit possible. Ce type de questions concernant l'optimalité de la puissance de la caractéristique sont aujourd'hui connues sous le nom de *la conjecture* A_p.

Durant plusieurs années, de nombreux résultats ont été présentés. On peut notamment citer un problème de longue date (voir Fefferman-Kenig-Pipher [31] et Astala-Iwaniec-Saksman [3]) qui a été résolu grâce à la norme optimale en terme de la caractéristique du poids de l'opérateur de Beurling-Ahlfors [69]. En 2000, J. Wittwer a démontré dans [82] une estimation optimale pour les transformées de martingales, en utilisant les résultats bi pondérés de [61]. Pour ce qui est de la transformée de Hilbert, l'estimation de sa norme a été améliorée au fil du temps : S. Buckley a démontré dans [15] que la transformée de Hilbert est bornée par le carré de la caractéristique A_2 du poids. Dans [68], S. Petermichl et S. Pott ont amélioré ce résultat passant de la puissance 2 à la puissance 3/2. Ce problème a finalement été résolu par Stefanie Petermichl dans [66] montrant ainsi un résultat optimal en terme de la caractéristique du poids dans $L^p(\omega)$. Un an plus tard, elle a également résolu ce problème pour la transformée de Riesz dans [67]. Plus généralement, la conjecture A_2 a été complètement résolue en 2012 par T. Hytonen, pour tout opérateur de Calderón-Zygmund.

Résolution de la conjecture A_p, [42]

Theorem 1.1. *Soit $T \in \mathcal{L}(L^2(\mathbb{R}^N))$ un opérateur de Calderón-Zygmund. Alors pour tout $\omega \in A_p$,*

$$\|Tf\|_{L^p(\omega)} \leq C_p(T) Q_p(\omega)^{\max\{1,1/(p-1)\}} \|f\|_{L^p(\omega)},\ p \in (1,\infty)$$

et le résultat est optimal en terme de la puissance de la caractéristique $Q_p(\omega)$.

Pour prouver ce théorème, T. Hytonen s'est d'abord placé dans le cas où $p = 2$. Ensuite, il a représenté l'opérateur de Calderón-Zygmund comme étant une moyenne de "dyadic shifts". Enfin, il a conclu en utilisant le résultat d'extrapolation de Rubio de Francia suivant (cf. [75],[17],[24])

Theorem 1.2. *Soit T un opérateur sous linéaire. Supposons que pour un certain $r \in [1,\infty)$ et tout $\omega \in A_r$, l'opérateur T satisfait pour tout $B > 1$ ce qui suit*

$$\|T\|_{L^r(u)} \leq N_r(B),\ \forall u \in A_r,\ Q_r(u) \leq B.$$

Alors pour tout $1 < p < \infty$ et tout $B > 1$, il existe une constante $N_p(B) > 0$ telle que

$$\|T\|_{L^p(\omega)} \leq N_p(B),\ \forall \omega \in A_p,\ Q_p(\omega) \leq B.$$

De plus,

$$N_p(B) \leq \begin{cases} 2^{1/r} N_r(2C(p')^{\frac{p-r}{p-1}} B), & si\ p > r \\ 2^{\frac{r-1}{r}} N_r(2^{r-1}(C(p)^{p-r}B)^{\frac{r-1}{p-1}}), & si\ p < r \end{cases}$$

où la constante $C(p)$ dépend uniquement de p et apparaît dans l'inégalité

$$\|Mf\|_{L^p(\omega)} \leq C(p) Q_p(\omega)^{\max\{1,1/(p-1)\}} \|f\|_{L^p(\omega)},\ p \in (1,\infty)$$

où M est la fonction maximale de Hardy-Littlewood.

1.1.2 Cadre du travail

On se place dans le cadre de cette thèse sur une variété Riemannienne (X, g, μ_φ) munie d'une mesure de type $d\mu_\varphi = e^{-\varphi} d\mu$, $\varphi \in C^2(X)$. On munit de plus l'espace de la courbure de Bakry-Emery $\mathrm{Ric}_\varphi = \mathrm{Ric} + \nabla^2\varphi$. Dans ce cas, le vecteur de Riesz est défini par

$$\mathcal{R}_\varphi = \nabla \circ (-\Delta_\varphi)^{-1/2},$$

où $\Delta_\varphi = \Delta - \nabla\varphi \cdot \nabla$.

La question de la continuité des transformées de Riesz sur des variétés Riemanniennes a fait l'objet de nombreux travaux. Nous pouvons notammenter citer [78, 60, 10, 36, 12, 72, 2]. Dans [29], Bakry stipule que, sous des hypothèses de courbure de Ricci minorée, les transformées de Riesz sont bornées sur des variétés Riemanniennes complètes. De plus, en utilisant des techniques d'analyse stochastique, des résultats indépendants de la dimension ont été obtenus dans [54] et [58]. Les mêmes résultats ont été obtenus également dans [16] en utilisant cette fois des techniques déterministes.

Concernant les poids, les classes de poids considérées durant les dernières décennies étaient uniquement définies en fonction des volumes des boules. Ceci a fortement contribué à l'étude des espaces de type homogène. On rappelle qu'une mesure sur un espace métrique X est dite doublante si la mesure de toute boule est approximativement la mesure de son double. Plus précisément, s'il existe une constante $C > 0$ telle que pour tout $x \in X$ et $r > 0$, on a

$$\mu(B(x, 2r)) \le C\mu(B(x, r)),$$

où $\mu(B(x, r))$ est le volume de la boule $B(x, r)$. Dans ce cas, on dit que la mesure μ est C-doublante et que (X, μ) est un espace de type homogène.

L'un des principaux obstacles de cette thèse est que les mesures considérées ne sont pas nécessairement doublantes. C'est pour cette raison que nous nous intéressons à une nouvelle classe de poids qui serait plus adéquate pour l'étude des transformées de Riesz sur X. Ainsi, nous utilisons des poids qui appartiennent à la classe Poisson-A_p, de caractéristique $\widetilde{Q}_p(\omega)$, où on considère des moyennes de Poisson plutôt que des moyennes sur des boules comme c'est le cas pour la classe A_p de Muckenhoupt. Ceci nous permet non seulement de nous intéresser à des mesures qui ne sont pas forcément doublante mais aussi d'obtenir des résultats optimaux et indépendants de la dimension pour le vecteur de Riesz:

$$\|\mathcal{R}_\varphi\|_{L^2(\omega)\to L^2(\omega)} \lesssim \widetilde{Q}_2(\omega).$$

Nous soulignons que dans ce contexte, le résultat est optimale en terme de la puissance de la caractéristique $\widetilde{Q}_2(\omega)$ puisque la puissance vaut 1. Ce type d'estimation linéaire est très récent dans le cas d'un espace X de mesure finie et pour $\varphi = 0$ [14]. Cette thèse a donc les nouveautés suivantes :

1. Une estimation pondérée même dans le cas où $\varphi \not\equiv 0$.

2. L'estimation est optimale en terme de la puissance de la caractéristique $\widetilde{Q}_2(\omega)$.

3. L'estimation est indépendante de la dimension.

Même dans le cas $\varphi = 0$, ces opérateurs ne sont pas nécessairement de type Calderón-Zygmund. En effet, ces transformées de Riesz appartiennent à une classe générale d'opérateurs sans noyau et les premières estimations pondérées furent établies par Auscher et Martell dans [4]. L'estimation optimale a été récemment étendue par Bernicot, Frey et Petermichl dans [14]. Toutefois, leur approche ne concerne que le cas où $\varphi = 0$. Le caractère doublant de la mesure $\mu_0 = e^{-0}\mu$ est utilisé et la constante dépend de la dimension de la variété.

On montre tout d'abord dans cette thèse que sur une variété Riemannienne complète (X, g, μ) munie d'une mesure $e^{-\varphi}d\mu$ et d'une courbure de Bakry-Emery positive, on obtient une estimation pondérée et linéaire de la norme des transformées de Riesz. De plus, l'estimation ne dépend pas de la dimension de la variété. La preuve de ce résultat repose notamment sur une représentation du vecteur de Riesz au moyen du semi groupe de Poisson sur les fonctions (noté P_t) et sur les 1-formes différentielles (noté $\vec{P}_t$). Ceci ramène le problème à une estimation bilinéaire faisant intervenir $\overline{\nabla} P_t f$ et $\overline{\nabla} \vec{P}_t \vec{g}$ où f est une fonction, $\vec{g}$ est une 1-forme différentielle et $\overline{\nabla}$ le gradient en temps et en espace. La démonstration de cette estimation bilinéaire fait appel à une construction d'une fonction de Bellman explicite. Cette stratégie ressemble à celle de Carbonaro-Dragičević dans [16] dans le cas non pondéré ($w = 1$). Une différence notable dans ce travail est la complexité de la fonction de Bellman de six variables. Cette dernière est issue d'une analyse de [61] ainsi que [70].

La procédure grâce à laquelle on transfère l'étude de certains opérateurs d'analyse harmonique vers celle de l'analyse stochastique est standard. Pour une fonction f dans $L^p(X)$, on considère $\tilde{f}$ son extension harmonique sur $X \times \mathbb{R}^+$. On compose cette fonction avec le processus $Z_t = (B_t^X, B_t)$ nommé bruit de fond ou "background radiation process" défini par Gundy et Varopoulos dans [35] sur $X \times \mathbb{R}^+$. Ce processus représente des trajectoires du mouvement Brownien sur le demi espace supérieur qui débutent leurs trajets à l'infini et qui s'arrêtent au moment où elles touchent un bord. On produit ainsi une nouvelle transformée de martingale grâce à une matrice A de taille $(n+1) \times (n+1)$. On note $(A * M^f)_t$ la transformée de martingale, où $(M_t^f)_t$ est la martingale associée à f (en utilisant la formule d'Itô). Enfin, on projette cette martingale par l'espérance conditionnelle et on obtient ainsi l'opérateur recherché. On déduit ensuite des propriétés de l'opérateur en question à l'aide de propriétés basiques d'analyse stochastique. Parmi les exemples classiques on a

- Les transformées de Riesz sur $\mathbb{R}^n$ [35],
- L'opérateur de Beurling-Ahlfors sur $\mathbb{C}$ [7],
- La transformée de Hilbert sur $\mathbb{R}$ [23],

- Les transformées de Riesz et l'opérateur de Beurling-Ahlfors sur des variétés Riemanniennes [55].

Dans le second résultat de cette thèse, on présente une estimation L^p de la norme du vecteur de Riesz sur une variété Riemannienne à géométrie bornée, indépendante de la dimension et pour $p \in (1, \infty)$. Il découlera par la suite le résultat énoncé précédemment sur $L^2(w)$, avec une meilleure constante numérique. On se placera encore une fois dans le cas où la variété est munie d'une courbure positive. On se fondera sur une représentation stochastique de X.-D. Li des transformées de Riesz sur une variété [54, 58, 55]. Afin de construire des mouvements Browniens sur une variété, on étudiera en détail la construction de Eells-Elworthy-Malliavin définie comme étant la projection sur une variété d'une solution d'une équation stochastique différentielle étant elle même définie sur le fibré principal des cadres (orthonormal frame bundle) $\mathcal{O}(X)$. Cette construction permet de décrire les processus de diffusion sur une variété et donne naissance à la notion de transport parallèle le long de courbes [39, 28].

Notre preuve est différente de celle du premier résultat car elle ne repose pas sur une fonction de Bellman. En effet, on utilise une domination éparse du processus stochastique de Li. L'idée de domination éparse est particulièrement bien adaptée aux bornes pondérées. Elle joue d'ailleurs un rôle central (dans une version discrète où l'opérateur épars est défini sur des cubes dyadiques) dans une preuve de la conjecture A_2 (voir [52] et [49].) Une originalité de l'approche proposée dans ce travail est de construite un opérateur épars associé à un temps d'arrêt continu ([21]) Cela permet d'obtenir des bornes indépendantes de la dimension de la variété.

Le processus stochastique introduit par Li est une certaine semi-martingale, construite à partir de B_t, B_t^X, deux mouvements Browniens respectivement définis sur $\mathbb{R}_+$ et X et de deux martingales auxiliaires (X_t) et (Y_t), subordonnée à (X_t) de telle sorte qu'il (le processus stochastique de Li) vérifie une certaine équation différentielle stochastique. De fait, l'argument que l'on utilise requiert plusieurs outils, notamment une estimation de type faible de la fonction maximale du processus en question. Cette estimation de type faible est apparue dans [8]. Elle nous permet par la suite de dominer le processus en utilisant la méthode éparse introduite dans [21].

Enfin, on présente une troisième preuve de l'estimation de la norme L^p du vecteur de Riesz sur une variété à géométrie bornée. L'avantage de cette preuve est qu'elle permet d'étendre le résultat aux courbures de Bakry-Emery négative. Notons que cette estimation ne dépend pas de la borne inférieure de la courbure. La preuve de ce troisième résultat diffère légèrement de celle du second dans le sens où on considère la courbure minorée. Le semi groupe de Poisson et les transformées de Riesz doivent donc être définis en conséquence. On introduit également une sous martingale qui est la somme d'une martingale apparue précédemment et d'un processus de variation fini croissant. Cela nous permet de contrôler des termes dépendant de la borne inférieure de la courbure qui apparaissent dans ce nouveau cas. Le reste de la preuve ressemble fortement à ce qui précède.

1.1.3 La transformée de Riesz sur $\mathbb{R}^n$

Sur la droite réelle, on définit la transformée de Hilbert par

$$Hf(x) = \frac{1}{\pi} v.p. \int_{\mathbb{R}} \frac{f(t)}{x-t} dt.$$

De même, on a un analogue n-dimensionnel de cet opérateur. En effet, il existe n opérateurs sur $\mathbb{R}^n$ que l'on appelle les transformées de Riesz. Ces opérateurs sont définis pour tout $1 \leq j \leq n$ par

$$\begin{aligned} R_j f(x) &:= c_n \lim_{\epsilon \to 0} \int_{\mathbb{R}^n \setminus B_\epsilon(x)} \frac{(t_j - x_j) f(t)}{|x-t|^{n+1}} \, dt \\ &= c_n \ v.p. \int_{\mathbb{R}^n} \frac{(t_j - x_j) f(t)}{|x-t|^{n+1}} \, dt, \end{aligned}$$

avec $c_n = \dfrac{\Gamma[(n+1)/2]}{\pi^{(n+1)/2}}$.
De plus, ces opérateurs peuvent être vus comme une convolution avec le noyau

$$K_j(x) = c_n \ v.p. \frac{x_j}{|x|^{n+1}},$$

faisant d'eux des opérateurs de type Calderón-Zygmund.
On peut également définir ces opérateurs grâce à une relation faisant intervenir le Laplacien et les premières dérivées partielles de la façon suivante

$$R_j \circ \sqrt{-\Delta} = \partial_j$$

ou plus formellement, on peut écrire

$$R = \nabla \circ (-\Delta)^{-1/2}.$$

(Dans cette dernière formule, R est un vecteur de n composantes.)

Les transformées de Riesz apparaissent dans la théorie du potentiel ainsi que l'analyse harmonique. En particulier, elles permettent de déduire des informations sur la totalité du hessien d'une fonction en ne connaissant que son Laplacien. Nous citons parmi les propriétés les plus remarquables des transformées de Riesz les suivantes

- $\mathcal{F}(R_j f)(\xi) = -i \dfrac{\xi_j}{|\xi|} \mathcal{F} f(\xi) \ \forall j = 1, \cdots, n.$
- Une conséquence immédiate est que les transformées de Riesz sont bornées sur L^2.
- Pour tout $1 < p < \infty$, T. Iwaniec et G. Martin ont montré qu'il existe une constante $C_p > 0$ indépendante de n telle que

  $$\|R_j f\|_p \leq C_p \|f\|_p, \ \forall j = 1, \cdots, n.$$

  Cette constante s'appelle la constante de Pichorides et vaut $C_p = \cot(\frac{\pi}{2p^*})$, où $p^* = \max(p, \frac{p}{p-1})$.

- En définissant $\|Rf\|_p := \|(\sum_{i=1}^n |R_i f|^2)^{1/2}\|_p$ la norme dans L^p du vecteur de Riesz, Bañuelos et Wang ont démontré que

$$\|Rf\|_p \leq 2(p^* - 1)\|f\|_p, \ \forall\, p \in (1, \infty).$$

Cette constante est la meilleure connue à ce jour. Il est conjecturé dans [5] que la constante optimale devrait être $C_p = \cot(\frac{\pi}{2p^*})$. Notons que la borne de Bañuelos et Wang n'est pas optimale même dans le cas où $p = 2$, qui par la transformée de Fourier vaut 1.
La connaissance de la valeur exacte (ou du moins une bonne estimation) de la norme p-ième des transformées de Riesz sur $\mathbb{R}^n$ est un sujet récent dont l'importance apparaît notamment dans la théorie des opérateurs quasi conformes et les EDP qui y sont associées, ainsi que dans la théorie L^p Hodge.

1.1.4 Représentation analytique de la transformée de Riesz sur $\mathbb{R}^n$

Dans cette partie, nous mettrons en évidence le lien fort entre les transformées de Riesz et les fonctions harmoniques sur $\mathbb{R}^n$. En effet, on présentera une formule de type Littlewood-Paley appliquée aux transformées de Riesz en utilisant les semi groupes de Poisson.
Ainsi, soit $R = (R_1, \cdots, R_n)$ le vecteur de Riesz sur $\mathbb{R}^n$ et P_t le semi groupe de Poisson agissant sur des fonctions. Alors on a

$$\langle Rf, g\rangle_{L^2(\mathbb{R}^n)} = -4\int_0^\infty \langle \nabla P_t f, \frac{d}{dt}P_t g\rangle_{L^2(\mathbb{R}^n)} t dt.$$

Pour démontrer ce résultat, on utilise le fait que pour des fonctions F suffisamment décroissantes on a la formule suivante

$$F(0) = \int_0^\infty F''(t) t dt,$$

grâce à une double intégration par parties.
En particulier, si on prend $F(t) = \langle P_t Rf, P_t g\rangle_{L^2(\mathbb{R}^n)}$, on obtient

$$\begin{aligned} F(0) &= \langle Rf, g\rangle_{L^2(\mathbb{R}^n)} \\ &= \int_0^\infty \frac{d^2}{dt^2}\langle P_t Rf, P_t g\rangle_{L^2(\mathbb{R}^n)} t dt \\ &= \int_0^\infty (\langle \frac{d^2}{dt^2}P_t Rf, P_t g\rangle_{L^2(\mathbb{R}^n)} + \langle P_t Rf, \frac{d^2}{dt^2}P_t g\rangle_{L^2(\mathbb{R}^n)} \\ &\quad + 2\langle \frac{d}{dt}P_t Rf, \frac{d}{dt}P_t g\rangle_{L^2(\mathbb{R}^n)}) t dt \\ &= 4\int_0^\infty \langle \sqrt{-\Delta}P_t Rf, \sqrt{-\Delta}P_t g\rangle_{L^2(\mathbb{R}^n)} t dt, \end{aligned}$$

où les deux dernières égalités sont une conséquence du fait que $\partial_t P_t f = -\sqrt{-\Delta}P_t f$, $\partial^2_{tt} P_t f = -\Delta P_t f$ et que Δ est symétrique. Sachant que $R = \nabla \circ (-\Delta)^{-1/2}$ et que $(-\Delta)^{1/2}$ commute avec P_t et ∇, on conclut en utilisant une fois de plus que $\partial_t P_t g = -\sqrt{-\Delta}P_t g$.

1.1.5 Représentation probabiliste des transformées de Riesz sur $\mathbb{R}^n$

On a vu précédemment une représentation déterministe des transformées de Riesz en utilisant des semi groupes. On se propose de présenter maintenant une approche probabiliste. L'idée d'une telle représentation de la transformée de Riesz provient des travaux de Gundy et Varopoulos dans [35] puis de ceux de Gundy et Silverstein dans [37] via une espérance conditionnelle d'une transformée de martingale. Le cœur de cette représentation se trouve dans la définition du processus appelé bruit de fond ou "background radiation" et le fait qu'une fonction $f \in L^p(\mathbb{R}^n)$ peut être exprimée comme une espérance conditionnelle d'une transformation simple d'une intégrale stochastique associée à f, faisant apparaître son extension harmonique sur le plan $\mathbb{R}^n \times \mathbb{R}_+$ ainsi qu'un $(n+1) \times (n+1)$-dimensionnel mouvement Brownien.

En effet, suivant la démarche de Gundy et Varopoulos, on définit le processus $Z_t = (X_t, B_t)$, où X_t est un mouvement Brownien sur $\mathbb{R}^n$ et B_t un mouvement Brownien sur $\mathbb{R}$ commençant à un certain niveau $y > 0$. Soit $\tau = \inf\{t \geq 0 : B_t = 0\}$ un temps d'arrêt. Le processus $(Z_t)_{0 \leq t \leq \tau}$ est un processus de diffusion sur $\mathbb{R}^n \times \mathbb{R}_+$ et s'arrête au temps τ touchant ainsi le bord $\mathbb{R}^n \times \{0\}$.
Soit g une fonction et $Qg(x, y) = e^{-y\sqrt{-\Delta}}g(x)$ son semi groupe de Poisson i.e. l'extension harmonique de g sur $\mathbb{R}^n \times \mathbb{R}_+$. On a par la formule d'Itô

$$g(X_\tau) = Qg(Z_\tau) = Qg(Z_0) + \int_0^\tau \nabla Qg(Z_s) dZ_s.$$

On affirme que pour une matrice A de taille $(n+1) \times (n+1)$ on a

$$\langle \lim_{y \to \infty} \mathbb{E}\left(\int_0^\tau A \nabla Qf(Z_s) dZ_s | X_\tau = x \right), g \rangle = 2 \int_0^\infty \int_{\mathbb{R}^n} \langle A \nabla Qf(x, z), \nabla Qg(x, z) \rangle z dx dz. \tag{1.1}$$

En effet, soit $T_A f(x) = \mathbb{E}\left(\int_0^\tau A \nabla Qf(Z_s) dZ_s | X_\tau = x \right)$. En dualisant on obtient alors

$$\begin{aligned}
\langle T_A f, g \rangle \quad &= \quad \mathbb{E}\left(\langle \mathbb{E}\left(\int_0^\tau A \nabla Qf(Z_s) dZ_s | X_\tau = x \right), g(X_\tau) \rangle \right) \\
&= \quad \mathbb{E}\left(\mathbb{E}\left(\langle \int_0^\tau A \nabla Qf(Z_s) dZ_s, g(X_\tau) \rangle | X_\tau = x \right) \right) \\
&= \quad \mathbb{E}\left(\langle \int_0^\tau A \nabla Qf(Z_s) dZ_s, g(X_\tau) \rangle \right) \\
&= \quad \mathbb{E}\left(\langle \int_0^\tau A \nabla Qf(Z_s) dZ_s, Qg(Z_0) \rangle \right) \\
&\qquad + \mathbb{E}\left(\langle \int_0^\tau A \nabla Qf(Z_s) dZ_s, \int_0^\tau \nabla Qg(Z_s) dZ_s \rangle \right) \\
&= \quad \mathbb{E}\left(\int_0^\tau \langle A \nabla Qf(Z_s), \nabla Qg(Z_s) \rangle ds \right) \\
&\underset{y \to +\infty}{\longrightarrow} \quad 2 \int_0^\infty \int_{\mathbb{R}^n} \langle A \nabla Qf(x, z), \nabla Qg(x, z) \rangle z dx dz.
\end{aligned}$$

Alors, pour une matrice A qui permute y et x_k (et des zéros ailleurs), on obtient en utilisant la représentation analytique des transformées de Riesz l'identité suivante

$$\begin{aligned}\int_{\mathbb{R}^n}\langle \lim_{y\to\infty}\mathbb{E}\left(\int_0^\tau \frac{\partial Qf}{\partial x_k}dy_s|X_\tau = x\right), g(X_\tau)\rangle dx &= 2\int_0^\infty\int_{\mathbb{R}^n}\langle\frac{\partial Qf}{\partial x_k}(x,y), \frac{\partial Qg}{\partial y}(x,y)\rangle ydxdy\\ &= -\frac{1}{2}\int_{\mathbb{R}^n}\langle R_k f, g\rangle dx.\end{aligned}$$

On obtient dans ce cas le résultat suivant [35]

Theorem 1.3. *Soit $A_j = (a_{ik})$ la matrice de taille $(n+1)\times(n+1)$ où $a_{ik} = 0$ sauf pour $i = n+1$, $k = j$ et $a_{(n+1)j} = 1$. On a la représentation de Gundy Varopoulos suivante*

$$-\frac{1}{2}R_j f = \lim_{y\to\infty}\mathbb{E}\left(\int_0^\tau A_j\nabla Qf(Z_s)dZ_s|X_\tau = x\right).$$

1.1.6 Exemple : l'espace Gaussien

Maintenant que l'on a une bonne compréhension de ce qui se passe sur $\mathbb{R}^n$ muni de la mesure de Lebesgue, une question naturelle est de se demander ce qu'il se passe lorsque l'on change la mesure. En effet, le but est de remplacer la mesure de Lebesgue dx par la mesure Gaussienne γ sur $\mathbb{R}^n$ définie par

$$d\gamma(x) = e^{-\frac{\|x\|^2}{2}}.$$

L'espace $(\mathbb{R}^n, d\gamma)$ s'appelle dans ce cas l'espace Gaussien. On définit également un nouveau Laplacien

$$\Delta_{OU}f(x) = \Delta f(x) - x\cdot\nabla f(x),\ f\in C_c^\infty$$

appelé l'opérateur d'Ornstein-Uhlenbeck.

L'opérateur Δ_{OU} défini sur $C_c^\infty(\mathbb{R}^n)$ admet une extension auto adjointe sur $L^2(\mathbb{R}^n, d\gamma)$ que l'on notera encore Δ_{OU}. Cet opérateur est négatif et symétrique par rapport à la mesure γ puisque l'on a pour tous $f, g \in C_c^\infty$

$$\int_{\mathbb{R}^n}(\nabla f,\nabla g)d\gamma(x) = -\int_{\mathbb{R}^n}f(\Delta_{OU}g)d\gamma(x) = -\int_{\mathbb{R}^n}g(\Delta_{OU}f)d\gamma(x)$$

Il génère donc un semi groupe de diffusion P_t défini par la formule de Mehler

$$P_t f(x) := \int_{\mathbb{R}^n}f(e^{-t}x + \sqrt{1-e^{-2t}}y)d\gamma(y).$$

Ce semi groupe peut être exprimé à l'aide du noyau de Mehler M_t

$$P_t f(x) = \int_{\mathbb{R}^n}M_t(x,y)f(y)d\gamma(y),$$

où

$$M_t(x,y) = \frac{1}{\pi^{n/2}(1-e^{-2t})^{n/2}}\exp\left(\frac{\|e^{-tx}-y\|^2}{1-e^{-2t}}\right).$$

Notons que ce semi groupe est solution du problème de Cauchy suivant

$$\left\{\begin{array}{rcl} \partial_t u(t,x) & = & \Delta u(t,x) - x\cdot\nabla u(t,x), \\ u_0 & = & f \in L^2(\mathbb{R}^n), \end{array}\right.$$

et on note formellement le semi groupe d'Ornstein Uhlenbeck $e^{t\Delta_{OU}}f$. Pour en savoir plus sur ce semigroupe, on réfère le lecteur à [76].

Meyer a introduit dans [60] le vecteur de Riesz associé à l'opérateur d'Ornstein Uhlenbeck Δ_{OU} comme étant l'opérateur

$$R(\Delta_{OU}) := \nabla(-\Delta_{OU})^{-1/2}.$$

Représentation analytique : On obtient une représentation analytique du vecteur de Riesz associé à l'opérateur d'Ornstein Uhlenbeck en ajustant la preuve déjà énoncée précédemment, pour $R = \nabla(-\Delta_{OU})^{-1/2}$, $T_t f(x) = e^{-t\sqrt{-\Delta_{OU}}}f(x)$ le semi groupe de Poisson et où $(\mathbb{R}^n, dx)$ est remplacé par $(\mathbb{R}^n, d\gamma(x))$. On obtient alors ce qui suit

$$\langle\nabla(-\Delta_{OU})^{-1/2}f, g\rangle_{L^2(\mathbb{R}^n, d\gamma(x))} = -4\int_0^\infty \langle\nabla T_t f, \frac{d}{dt}T_t g\rangle_{L^2(\mathbb{R}^n, d\gamma(x))} t dt.$$

Représentation probabiliste : On peut associer à l'opérateur d'Ornstein Uhlenbeck Δ_{OU} un processus de diffusion (X_t) sur $\mathbb{R}^n$ qui satisfait

$$dX_t = dW_t - X_t dt,$$

où (W_t) est le mouvement Brownien sur $\mathbb{R}^n$. Cette équation différentielle stochastique est résolue par une variation de paramètres. En effet, soit $f(X_t, t) = X_t e^t$. On obtient par la formule d'Itô

$$\begin{aligned} df(X_t, t) &= X_t\, e^t\, dt + e^t\, dX_t \\ &= X_t e^t\, dt + e^t\, (dW_t - X_t dt) \\ &= e^t\, dW_t. \end{aligned}$$

En intégrant entre 0 et t on obtient

$$X_t = X_0 e^{-t} + e^{-t}\int_0^t e^s\, dW_s.$$

On affirme que l'on a la représentation probabiliste suivante des transformées de Riesz associées à l'opérateur d'Ornstein Uhlenbeck

$$-\frac{1}{2}\nabla(-\Delta_{OU})^{-1/2}f(x) = \lim_{y\to\infty}\mathbb{E}_y\left[e^{-\tau}\int_0^\tau e^s \nabla e^{-B_s\sqrt{-\Delta_{OU}}}f(X_s)dB_s | X_\tau = x\right].$$

On voit que cette représentation diffère légèrement de la formule donnée par Gundy et Varopoulos de par la présence de deux nouveaux termes. Pour expliquer ce phénomène, nous avons besoin de nouveaux arguments, principalement tirés de la géométrie différentielle. C'est d'ailleurs l'une des raisons qui a motivé les travaux de cette thèse sur des variétés Riemanniennes.

1.1.7 Méthodes et résultats

Premier résultat

La méthode de la fonction de Bellman a été utilisée à l'origine dans la théorie du contrôle par Richard E. Bellman. Burkholder l'a par la suite introduite en analyse harmonique en 1984 pour obtenir certaines inégalités sur les transformées de martingale, avec une variante de la méthode. Cette méthode est réapparue dans les années 90 grâce à Nazarov, Treil et Volberg pour montrer/redémontrer de nombreux résultats en analyse harmonique. Cette méthode s'avère être un outil extrêmement puissant et un moyen très naturel de traiter les inégalités pondérées et de trouver une dépendance de la norme de certains opérateurs classiques en analyse harmonique sur des espaces L^p pondérés avec la caractéristique A_p du poids.
Le principal défi de cette méthode consiste à trouver une fonction appropriée satisfaisant toutes les propriétés souhaitées, puis à utiliser des arguments de convexité. Bien que l'unicité ne soit pas requise, la connaissance de ces fonctions nécessite beaucoup de pratique. Les auteurs eux-mêmes décrivent cette méthode comme un savoir-faire artisanal. Chaque problème a sa propre fonction de Bellman dépendant d'un certain nombre de variables qui changent d'un problème à l'autre.

Le premier résultat de cette thèse concerne le vecteur de Riesz $\mathcal{R}_\varphi$ sur une variété Riemannienne (X, g, μ) munie d'une mesure de type $e^{-\varphi}d\mu$ où un poids supplémentaire est présent : nous étudions donc la norme du vecteur de Riesz sur l'espace pondéré $L^2(\omega) = L^2(\omega e^{-\varphi}d\mu)$.

En exhibant une fonction de Bellman appropriée dont l'origine provient d'une analyse profonde du papier [61], nous prouvons que sur une variété Riemannienne complète (X, g, μ) munie de la mesure $d\mu_\varphi = e^{-\varphi}d\mu$ tel que la courbure de Bakry-Emery soit positive et que $\mu_\varphi(X) < \infty$, nous disposons d'une estimation linéaire et indépendante de la dimension de la norme pondérée du vecteur de Riesz en terme de la caractéristique du poids dans la classe Poisson A_2:

$$\|\mathcal{R}_\varphi\|_{L^2(\omega)\to L^2(\omega)} \lesssim \widetilde{Q}_2(\omega).$$

Plus précisément, on définit pour tout $Q \geq 1$, la fonction $B_Q = B_1 + B_2 + B_3 + B_4$ sur le domaine

$$D_Q = \{\mathcal{X} := (X, Y, \mathrm{x}, \mathrm{y}, r, s) : \mathrm{x}^2 \leq Xr, \langle \mathrm{y}, \mathrm{y}\rangle \leq Ys, 1 \leq rs \leq Q\}$$

par

- $B_1(Z, H, \mathrm{x}, \mathrm{y}, r, s) = X - \dfrac{\mathrm{x}^2}{r} + Y - \dfrac{\langle \mathrm{y}, \mathrm{y}\rangle}{s}$,
- $B_2(X, Y, \mathrm{x}, \mathrm{y}, r, s) = X - \dfrac{\mathrm{x}^2}{r} + Y - \dfrac{\langle \mathrm{y}, \mathrm{y}\rangle}{s + \dfrac{M(r,s)}{Q^2}}$,

- $B_3(X,Y,\mathrm{x},\mathrm{y},r,s) = X - \dfrac{\mathrm{x}^2}{r + \dfrac{N(r,s)}{Q^2}} + Y - \dfrac{\langle \mathrm{y},\mathrm{y}\rangle}{s},$

  où
  $$M(r,s) = -\frac{4Q^2}{r} - rs^2 + (4Q^2+1)s$$
  et
  $$N(r,s) = -\frac{4Q^2}{s} - sr^2 + (4Q^2+1)r.$$

- $B_4 = B_{41} + B_{42} + B_{43}$ avec

  - $B_{41}(X,Y,\mathrm{x},\mathrm{y},r,s) = X - \dfrac{\mathrm{x}^2}{r + \dfrac{\widetilde{M}(r,s)}{Q}} + Y - \dfrac{\langle \mathrm{y},\mathrm{y}\rangle}{s}$

  - $B_{42}(X,Y,\mathrm{x},\mathrm{y},r,s) = X - \dfrac{\mathrm{x}^2}{r} + Y - \dfrac{\langle \mathrm{y},\mathrm{y}\rangle}{s + \dfrac{\widetilde{N}(r,s)}{Q}}$

  - $B_{43}(X,Y,\mathrm{x},\mathrm{y},r,s) = \sup_{a>0}\left(X - \dfrac{\mathrm{x}^2}{r + a\dfrac{K(r,s)}{Q}} + Y - \dfrac{\langle \mathrm{y},\mathrm{y}\rangle}{s + a^{-1}\dfrac{K(r,s)}{Q}} \right)$

    où
    $K(r,s) = \sqrt{Q}\sqrt{rs} - \dfrac{rs}{4},$
    $\widetilde{M}(r,s) = -\dfrac{4Q}{s} - \dfrac{r^2 s}{4Q} + (4Q+1)r$
    et
    $\widetilde{N}(r,s) = -\dfrac{4Q}{r} - \dfrac{s^2 r}{4Q} + (4Q+1)s.$

Cette fonction satisfait les propriétés suivantes

1. $0 \leq B_Q \leq 884(X+Y)$;
2. $-d^2 B_Q \geq \frac{4}{Q}|d\mathrm{x}||d\mathrm{y}|$, où B_Q est C^2;
3. $\partial_\nu B_Q \leq 0$.

A l'aide de cette fonction et de ses propriétés, nous sommes en mesure de démontrer l'inégalité bilinéaire suivante

$$\int_0^\infty \int_X |\overline{\nabla} P_t f(x)||\overline{\nabla} \vec{P}_t \vec{g}(x)| t\, d\mu_\varphi(x) dt \leq 221\ \widetilde{Q}_2(\omega)\|f\|_{L^2(X,\omega\mu_\varphi)}\|\vec{g}\|_{L^2(T^*X,\omega^{-1}\mu_\varphi)},$$

où $P_t f$ est le semi groupe de poisson associé à la fonction f, $\vec{P}_t\vec{g}$ est le semi groupe de Poisson associé à la 1-forme différentielle $\vec{g}$ et $\overline{\nabla}$ est le gradient en temps et en espace. On définit le vecteur de Riesz sur L^2 par

$$\mathcal{R}_\varphi : \overline{R(-\Delta_\varphi)} \to L^2(T^*X).$$

En utilisant la formule

$$\int_X \langle \mathcal{R}_\varphi f(x), \vec{g}(x)\rangle \, d\mu_\varphi(x) = 4\int_0^\infty \int_X \left\langle dP_t f(x), \frac{d}{dt}\vec{P}_t\vec{g}(x)\right\rangle d\mu_\varphi(x)\, tdt,$$

il découle que

$$\|\mathcal{R}_\varphi f\|_{L^2(T^*X,\omega\mu_\varphi)} \leq 884\widetilde{Q}_2(\omega)\|f\|_{L^2(X,\omega\mu_\varphi)}.$$

Ces estimations sont connues comme étant optimales en terme de la caractéristique du poids même lorsque $X = \mathbb{T}$ et $\varphi = 0$ [70].

Pour plus d'informations sur ce travail, nous référons le lecteur à [18].

Second résultat

Notre approche pour le deuxième résultat est un peu plus probabiliste et utilise la représentation martingale suivante de la transformation de Riesz sur des variétés Riemanniennes complètes, présentée pour la première fois dans [54] :

$$-\frac{1}{2}\mathcal{R}_\varphi f(x) = \lim_{y\to\infty} \mathbb{E}_y\left[M_\tau \int_0^\tau M_s^{-1} dQ(f)(B_s^X, B_s)dB_s | B_\tau^X = x\right]. \qquad (1.2)$$

Il a été très profitable d'utiliser des notions d'analyse stochastique, car elle est étroitement liée à l'analyse harmonique et permet d'obtenir des résultats optimaux pour les normes L^p de divers opérateurs importants.

Notre résultat est puissant car il estime Z^*, la fonction maximale d'un certain processus Z que nous définirons ultérieurement. La technique utilisée dans la preuve s'appelle *la domination éparse*. Cette technique récente est dûe à Nazarov et Lerner dans [52] et à Lacey dans [49] en 2015.

Pour un opérateur T et une fonction f convenable, le but est d'établir un contrôle ponctuel de Tf par un opérateur épars S i.e $|Tf| \lesssim S|f|$ puis d'utiliser le fait que la propriété éparse permet d'insérer des poids et d'en déduire la puissance optimale pour la constante A_p. Bien que la domination éparse de Lacey implique immédiatement des inégalités pondérées avec dépendance optimale vis-à-vis de la caractéristique A_p du poids, elle est définie sur des cubes et ne peut pas fournir d'estimations indépendantes de la dimension, ni de résultats satisfaisants sur des espaces non homogènes. Nous contournerons ce problème en utilisant un opérateur épars avec des temps d'arrêt continus, comme dans [21].

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité et $X = (X_t)_t$ un processus stochastique. L'opérateur $X \mapsto S(X)$ est dit épars s'il existe une suite croissante de temps d'arrêt adaptés $0 = T^{-1} \leq T^0 \leq \cdots$ et des ensembles emboîtés $E_j = \{T^j < \infty\}$, $E_j \subset E_{j-1}$ tels que

$$S(X) = \sum_{j=-1}^{\infty} X_{T^j}\chi_{E_j} \text{ où } X_{T^j} = \mathbb{E}(X|\mathcal{F}_{T^j});$$

$$\forall A_j \subset E_j,\ A_j \in \mathcal{F}_{T_j} \text{ tel que } \mathbb{P}(A_j \cap E_{j+1}) \leq \frac{1}{2}\mathbb{P}(A_j).$$

De plus, l'équation (1.2) peut être ainsi réécrite

$$-\frac{1}{2}(R_\varphi f)(x) = \lim_{y\to\infty} \mathbb{E}_y \left[Z_\tau | B_\tau^X = x \right],$$

où Z_t est une semi-martingale définie à l'aide des martingales auxiliaires X_t et Y_t définies comme suit

$$X_t = Qf(B_t^X, B_t) - Qf(B_0^X, y) = \int_0^t (\nabla, \partial_y) Qf(B_s^X, B_s)(U_s dW_s, dB_s),$$

$$Y_t = \int_0^t \nabla Qf(B_s^X, B_s) dB_s,$$

$$Z_t = M_t \int_0^t M_s^{-1} dY_s,$$

et Y_t est différentiellement subordonnée à X_t.

On peut alors montrer que pour tout $p \in (1, \infty)$:

$$\|\mathcal{R}_\varphi\|_{L^p \to L^p} \leq 16 \frac{p^2}{p-1}$$

et

$$\|\mathcal{R}_\varphi\|_{L^2(\omega) \to L^2(\omega)} \lesssim \widetilde{Q}_2(\omega).$$

Pour plus d'informations sur ce travail, nous renvoyons le lecteur à [19].

Troisième résultat

En utilisant la même approche probabiliste que dans le deuxième résultat, nous nous concentrons cette fois sur la courbure de Bakry-Emery, suivant [16]. Les marches aléatoires et les semi groupes de Poisson sur les variétés sont un sujet délicat et nous renvoyons le lecteur au texte de Emery [29].

Supposons que $\mathrm{Ric}_\varphi \geq -a,\ a \geq 0$. On définit le semi groupe de Poisson et le vecteur de Riesz en conséquence

$$Q^a(f)(x, y) = e^{-y\sqrt{aId - \Delta_\varphi}} f(x)$$

et

$$\mathcal{R}_\varphi^a = d(aId - \Delta_\varphi)^{-1/2}.$$

La pierre angulaire de ce résultat est un remplacement élégant de la martingale (X_t) par une sous-martingale (X_t^a), somme de (X_t) et d'un processus de variation finie croissant. Par la formule d'Itô, (X_t^a) n'est rien d'autre que $Q^a f$, l'extension de Poisson de f. Cette astuce nous permet de contrôler les termes issus de la courbure négative.

L'introduction de la sous-martingale implique un autre changement, à savoir la définition de l'opérateur épars S. En effet, nous définissons

$$S(X^a) = \sum_{j=-1}^{\infty} \mathbb{E}(X^a|\mathcal{F}_{T^j})\chi_{E_j} \quad \text{au lieu de} \quad S(X^a) = \sum_{j=-1}^{\infty} X^a_{T^j}\chi_{E_j}$$

car contrairement au cas des martingales, ici on a une simple inégalité

$$\mathbb{E}(X^a|\mathcal{F}_{T^j}) \geq X^a_{T^j}.$$

1.1.8 Applications et perspectives

Nous explorons brièvement dans cette partie quelques pistes qui s'inscrivent dans le même cadre que cette thèse. Nous suggérons au lecteur intéressé ce qui suit

1. L'opérateur de Beurling-Ahlfors;
2. Les intégrales fractionnaires;
3. Les bornes $L^p(w)$ des transformées de Riesz sur une variété Riemannienne.

L'opérateur de Beurling-Ahlfors

On peut obtenir une estimation pondérée de la norme de l'opérateur de Beurling-Ahlfors agissant sur des variétés. L'intérêt de cet opérateur provient d'un célèbre problème de régularité dans [3]. L'idée est d'utiliser à nouveau une formule de représentation de transformation de martingale pour l'opérateur de Beurling-Ahlfors étendue aux 1-formes sur des variétés riemanniennes complètes. Cette idée est tirée de [56, 57].
Soit (X, g, μ_φ) et B_t^X définis comme précédemment. L'opérateur de Beurling Ahlfors est défini sur les variétés comme suit

$$B = (d_\varphi^* d - d d_\varphi^*)(\vec{\Delta}_\varphi)^{-1},$$

où d désigne la dérivée extérieure, d_φ^* son opérateur adjoint et $\vec{\Delta}_\varphi = dd_\varphi^* + d_\varphi^* d$ est le Laplacien (pondéré) de Hodge-de Rham agissant sur les $1-$formes différentielles.
On définit les matrices $A_1 = (a_i a_j^*)$ et $A_2 = (a_i^* a_j)$ comme dans [56, Section 3] et

$$B = A_2 - A_1.$$

Contrairement au vecteur de Riesz, on définit le semi groupe de chaleur rétrograde généré par le Laplacien de Hodge de Rham Laplacien par

$$P\vec{g}(x, T-s) = e^{-(T-s)\vec{\Delta}_\varphi}\vec{g}(x),\ \forall x \in X,\ s \in [0,T],\ \vec{g} \in C_0^\infty(\Lambda T^* X),$$

pour tout $T > 0$.
La représentation probabiliste de l'opérateur de Beurling-Ahlfors sur des variétés Riemanniennes complètes où $\mathrm{Ric}_\varphi \geq 0$ est la suivante

$$S_{A_i}^T \vec{g}(x) = \mathbb{E}\left(M_T \int_0^T M_t^{-1} A_i \nabla P\vec{g}(B_t^X, T-t) dX_t | B_T^X = x \right),\ i = 1, 2$$

et

$$S_B^T\vec{g}(x) = 2\lim_{T\to\infty}\mathbb{E}\left(M_T\int_0^T M_t^{-1}B\nabla P\vec{g}(B_t^X, T-t)dX_t|B_T^X = x\right).$$

Soit

$$Z_t = M_t\int_0^t M_s^{-1}B\nabla P\vec{g}(B_s^X, T-s)dB_s^X$$

et

$$X_t = M_t\int_0^t M_s^{-1}\nabla P\vec{g}(B_s^X, T-s)dB_s^X.$$

Soit $Y_t = \dfrac{1}{\|B\|_{op}}Z_t$, où $\|\cdot\|_{op}$ désigne la norme d'opérateur. Y_t satisfait

$$dY_t = V_tY_t + \frac{B}{\|B\|_{op}}dX_t$$

et est différentiellement subordonnée à X_t. On obtient alors

$$\|Z^*\|_{L^p(T^*X,w)} \leq 16\frac{p^2}{p-1}\|B\|_{op}^2\widetilde{Q}_p(w)^{\max(1,\frac{1}{p-1})}\|X\|_{L^p(X,w)}.$$

Intégrales fractionnaires

Une autre famille d'opérateurs intéressante est celle des intégrales fractionnaires associées à un semi groupe de Feller $(T_t)_t$ dont la dimension Varopoulos est d. Nous définissons les intégrales fractionnaires d'ordre $\alpha \in (0,d)$ comme suit

$$I_\alpha f(x) = \frac{1}{\Gamma(\frac{\alpha}{2})}\int_0^\infty t^{\alpha/2-1}T_tf(x)dt.$$

Encore une fois, nous pouvons considérer la représentation probabiliste des intégrales fractionnaires étudiées sur $\mathbb{R}^d$ (voir [1]) et sur des espaces localement compacts (voir [47]) puis étendre cette représentation aux variétés Riemanniennes complètes en utilisant l'approche de Li dans [58]. On obtient ainsi

$$S_\alpha^a f(x) = \mathbb{E}^a\left[M_\tau\int_0^\tau M_s^{-1}B_s^\alpha\frac{\partial Q(f)}{\partial y}(B_s^X, B_s)dB_s|B_\tau^X = x\right].$$

Il existe alors une constante $C_{\alpha,d} > 0$ telle que

$$S_\alpha^a f \underset{a\to+\infty}{\longrightarrow} C_{\alpha,d}I_\alpha f,$$

au sens des distributions.

Si l'on peut trouver une fonction similaire à celle utilisée par Wang pour prouver la continuité des intégrales fractionnaires, on peut espérer obtenir des résultats intéressants concernant les inégalités de Hardy-Littlewood-Sobolev.

Bornes $L^p(w)$ des transformées de Riesz sur une variété Riemannienne

On aimerait étendre les résultats de nos travaux précédents aux espaces $L^p(w)$ pour tout $p \in (1, +\infty)$. Malheureusement, la preuve ne marche plus dans le cas $p \neq 2$ si on introduit des poids u et w tels que $u^p w = u$. Cependant, nous pouvons espérer obtenir des résultats positifs en utilisant le théorème d'extrapolation dans sa version probabiliste, utilisé par exemple dans [23].

1.2 Motivation

1.2.1 Historical context

In this thesis, we are interested in L^p and weighted L^p bounds of the Hilbert and the Riesz transforms.
The L^p estimate of the Hilbert transform on the real line dates back to the work of Riesz [74] and Pichorides [71]. Indeed, in 1972 Pichorides proved that the best constant for the norm of the Hilbert transform H is given by

$$C_p = \begin{cases} 2\tan\dfrac{\pi}{2p} & \text{for } 1 < p \leq 2; \\ \cot\dfrac{\pi}{2p} & \text{for } 2 < p < \infty. \end{cases}$$

In the Euclidean setting, the i–th Riesz transform in $\mathbb{R}^n$ is defined as

$$R_i = \frac{\partial}{\partial x_i}(-\Delta)^{-1/2},$$

where $\Delta = \sum_{i=1}^n \partial^2/\partial^2_{x_i}$ is the usual Laplacian in $\mathbb{R}^n$. The vector Riesz transform R is defined as the collection $R = (R_1, R_2, \ldots, R_n)$. Note that in the one-dimensional setting, the Riesz transform is nothing but the Hilbert transform. Regarding the L^p estimate of the Riesz vector in $\mathbb{R}^n$, T. Iwaniec and G. Martin proved in [46] that for all $1 < p < \infty$, there exists a constant $C_p > 0$ independent of n such that

$$\|R_j f\|_p \leq C_p \|f\|_p, \ \forall j = 1, \cdots, n.$$

This constant is equal to Pichorides constant. For the Riesz vector, the best result known so far is given by Bañuelos and Wang in [9]

$$\|Rf\|_p \leq 2(p^* - 1)\|f\|_p, \ \forall \ p \in (1, \infty), \text{ where } p^* = \max(p, \frac{p}{p-1}).$$

We also refer to [77, 60, 72, 25]. The knowledge of the exact value (or at least a good estimate) of the p-norm of the Riesz transforms on $\mathbb{R}^n$ is a recent matter whose importance appears in the theory of quasiconformal mappings and related PDEs.

In 1960 and using complex analysis tools, Helson and Szego proved in [38] that the Hilbert transform is bounded on $L^2(\omega)$ if and only if the weight ω can be represented as

$\omega = \exp(\varphi + H\psi)$, where $\varphi, \psi \in L^\infty$, $\|\psi\|_{L^\infty} < \pi/2$. Later in 1973, Hunt, Muckenhoupt and Wheeden proved in [41] that the A_p condition also characterizes the boundedness of the Hilbert transform on $L^p(\omega)$. In fact, they proved that H is bounded on $L^p(\omega), p \in (1,\infty)$ if, and only if ω belongs to the A_p-class of weights, that is

$$Q_p(\omega) := \sup_Q \left(\frac{1}{|Q|}\int_Q \omega(x)\,dx\right)\left(\frac{1}{|Q|}\int_Q \omega^{-\frac{1}{p-1}}(x)\,dx\right)^{p-1} < \infty,$$

where the supremum is taken over all cubes $Q \subset \mathbb{R}$.

Weighted inequalities arose naturally in analysis with the birth of singular integral theory. In the theory of PDEs for example, weights appear to treat PDEs with degenerate coefficients, domains with non smooth geometry or equations with rough initial data. On $\mathbb{R}^n$, the prototypical A_p weights are the power weights: by using for example improper Riemannian integrals, it is easy to see that for all $a \in \mathbb{R}$ and for $p > 1$, $|x|^a \in A_p$ if and only if $-n < a < (p-1)n$.

In the last 15 years, it has been of interest to find sharp bounds for the norm of a singular integral operator T in terms of the A_p characteristic $Q_p(\omega)$ of the weight. The aim is to prove an estimate of the form

$$\|Tf\|_{L^p(\omega)} \leq CQ_p(\omega)^r\|f\|_{L^p(\omega)}$$

for a suitable r where the constant C is independent of f or ω. Since $Q_p(\omega) > 1$, the focus is on finding estimates with r as small as possible. Questions of such optimal norm estimates have become known as A_p conjectures.

Over the years, many results were presented. For instance, a long standing regularity problem (see for example Fefferman-Kenig-Pipher [31] and Astala-Iwaniec-Saksman [3]) has been solved through the optimal weighted norm estimate of the Beurling-Ahlfors operator, a classical Calderón-Zygmund operator, using the heat flow characteristic of the weight. See Petermichl-Volberg [69]. In 2000, J. Wittwer proved in [82] a sharp estimate for the martingale transform, using important developments on a corresponding two weight question in [61]. As for the Hilbert transform, its bound has been improved several times: S. Buckley proved in [15] that the Hilbert transform is bounded by the square of the classical A_2 characteristic of the weight. In [68] S. Petermichl and S. Pott improved the power in this estimate from 2 to 3/2. This problem has finally been solved by Stefanie Petermichl in [66] who proved a sharp bound for the operator norm of the Hilbert transform in $L^p(\omega)$. A year later, she also solved it for the Riesz transforms in [67]. It was finally in 2012 that the so-called A_2-conjecture has been completely solved by T. Hytonen, for any Calderón-Zygmund operator.

Theorem 1.4 (Resolution of the A_p conjecture, [42]). *Let $T \in \mathcal{L}(L^2(\mathbb{R}^N))$ be a fixed Calderón-Zygmund operator. Then for all $\omega \in A_p$,*

$$\|Tf\|_{L^p(\omega)} \leq C_p(T)Q_p(\omega)^{\max\{1,1/(p-1)\}}\|f\|_{L^p(\omega)},\ p \in (1,\infty)$$

and the result is sharp in the power of $Q_p(\omega)$.

The proof was reduced to the case when $p = 2$, then the Calderón-Zygmund operator was represented as an average of "dyadic shifts" and finally, the following extrapolation theorem due to Rubio de Francia was used.

Theorem 1.5 ([75],[17],[24])**.** *Let T be a sub-linear operator. Suppose that for some r in $[1,\infty)$ and every ω in A_r, the operator T satisfies for each $B > 1$ the following*

$$\|T\|_{L^r(u)} \leq N_r(B),\ \forall u \in A_r,\ Q_r(u) \leq B.$$

Then for any $1 < p < \infty$ and all $B > 1$, there exists a constant $N_p(B) > 0$ such that

$$\|T\|_{L^p(\omega)} \leq N_p(B),\ \forall \omega \in A_p,\ Q_p(\omega) \leq B$$

Moreover,

$$N_p(B) \leq \begin{cases} 2^{1/r} N_r(2C(p')^{\frac{p-r}{p-1}} B), & \text{if } p > r \\ 2^{\frac{r-1}{r}} N_r(2^{r-1}(C(p)^{p-r}B)^{\frac{r-1}{p-1}}), & \text{if } p < r \end{cases}$$

where the constant $C(p)$ depends only on p and that appears in

$$\|Mf\|_{L^p(\omega)} \leq C(p) Q_p(\omega)^{\max\{1,1/(p-1)\}} \|f\|_{L^p(\omega)},\ p \in (1,\infty)$$

where M is the Hardy-Littlewood maximal function.

Several other proofs of the A_2-conjecture were presented, each of them simplifying Hytonen's proof and contributing to a better understanding of the field. See for example [53] whose proof exploited a local mean oscillation inequality and [49] who used sparse operators to obtain a pointwise control of the operator T.

1.2.2 Context of the work in this thesis

The focus in this thesis is on the Riesz vector on Riemannian manifolds (X, g, μ_φ) endowed with measures of the type $d\mu_\varphi = e^{-\varphi} d\mu,\ \varphi \in C^2(X)$. If in addition we endow the space with the Bakry-Emery curvature $\mathrm{Ric}_\varphi = \mathrm{Ric} + \nabla^2\varphi$, then the Riesz vector is defined as $\mathcal{R}_\varphi = \nabla \circ (-\Delta_\varphi)^{-1/2}$ with $\Delta_\varphi = \Delta - \nabla\varphi \cdot \nabla$.

For early considerations of L^p boundedness of Riesz transforms on manifolds, we refer to [78]. We mention also the works [60, 10, 36, 12, 72, 2] among which the papers of Bakry provide estimates of Riesz transforms for complete Riemannian manifolds under the general condition that the Bakry–Emery curvature is bounded below (see [29]). Using stochastic techniques, linear dimensionless estimates of the Bakry–Riesz vector on manifolds were announced in [54] and [58]. Using deterministic techniques, such estimates were proved in [16]. See also [6] for second order Riesz transforms on manifolds and [8] for Riesz transforms on manifolds, correcting a previous gap in the probabilistic proof.

Regarding the weighted theory, the classes of weights considered in the last decades were merely defined in terms of volume of balls, so the entire weighted theory has been extended to the doubling framework. We recall that a measure on a metric space X is said to be doubling if the measure of any ball is approximately the measure of its double, or more precisely, if there is a constant $C > 0$ such that for all $x \in X$ and $r > 0$, we have

$$\mu(B(x, 2r)) \leq C\mu(B(x, r)),$$

where $\mu(B(x, r))$ is the volume of the ball $B(x, r)$. In this case, we say that μ is C-doubling and that (X, μ) is a homogeneous space.

One challenge in this thesis is that the measures considered are **not necessarily doubling**. For this reason, we are interested in a version of the A_p class with characteristic $Q_p(\omega)$ which is particularly well-suited for working with the Riesz transforms on X. Namely, we use the Poisson-A_p class with characteristic $\widetilde{Q}_p(\omega)$, which considers Poisson averages instead of box averages in the definition of A_p. This allows us to tackle some measures that may have mild non-homogeneity as well as obtain a sharp bound free of dimension for the Riesz vector:

$$\|\mathcal{R}_\varphi\|_{L^2(\omega)\to L^2(\omega)} \lesssim \widetilde{Q}_2(\omega).$$

We stress both the continuity of the operator in this setting, its rate of continuity, i.e. the first power of the characteristic $\widetilde{Q}_2(\omega)$ as well as the fact that implied constants do not depend upon the dimension. The linear estimate (in terms of the classical characteristic that induces a dimensional growth) is very recent in the case of general X with bounded geometry and $\varphi = 0$ [14]. This thesis has therefore the following novelties:

1. A weighted estimate holds even in the case $\varphi \not\equiv 0$.

2. The estimate is sharp in terms of dependence on the power of $\widetilde{Q}_2(\omega)$.

3. The estimate is free of dimension.

Even in the case $\varphi = 0$, these operators are not necessarily of Calderón-Zygmund type. These Riesz transforms on manifolds fit into the class of non-kernel operators, whose weighted theory was established in Auscher-Martell [4]. The optimal weighted norm estimates for these types of operators, even without the extra $e^{-\varphi}$, have only recently been found in [14] (in terms of the classical characteristic). In all these proofs, the doubling feature of the measure $\mu_0 = e^{-0}\mu$ is heavily used and dimensional growth occurs.

We first show in this thesis (see Chapter 4) that on a complete Riemannian manifold (X, g, μ) endowed with measure $e^{-\varphi}d\mu$ and non-negative Bakry-Emery curvature, we have a dimension-free linear weighted norm estimates for the arising Riesz vector in terms of the Poisson flow A_2 characteristic of the weight. The proof is by a simple but beautiful fact that follows from a Littlewood-Paley type formula for the Riesz vector using the Poisson flow [1], followed by a Bellman function, resembling the strategy used by Carbonaro-Dragičević [16] for the unweighted case in L^p. Their proof relies on Bakry's idea of adapted Poisson flow on one-forms [12] in combination with the concise but powerful Bellman function of Nazarov-Treil [63] that is adapted to L^p estimates. A key difference here is the complication of the weighted Bellman function, that has to be known in an explicit manner. This Bellman function of six variables is derived through an analysis of [61] as well as [70]. A similar function was constructed in [22]. Properties that go beyond those needed for martingale multipliers are required to obtain the desired Riesz transform estimates on manifolds, which is in a sharp contrast to the Euclidean case [23], where existence of this Bellman function suffices.

The search for optimal estimates in weighted spaces has greatly improved our understanding of central operators in harmonic analysis and has developed numerous tools with a probabilistic flavor. Notably, the first solution of an A_2 problem [69] uses an underlying estimate for predictable martingale multipliers of dyadic martingales under a change of law by Wittwer [82]. Regarding the Riesz transforms, it has been known at least since Gundy-Varopoulos [35] that the Riesz transforms of a function can be written as conditional expectation of a simple transformation of a martingale associated to the function.

The procedure by which one transfers the study of certain operators in harmonic analysis to stochastic analysis is standard. For a function f in $L^p(X)$, we let $\tilde{f}$ be its harmonic extension on $X \times \mathbb{R}^+$. We compose this function with the background radiation process $Z_t = (B_t^X, B_t)$ defined by Gundy and Varopoulos in [35] on $X \times \mathbb{R}^+$ which is Brownian trajectories in the upper half space started at infinity and stopped when hitting the boundary. We produce a new martingale, the so called martingale transform by using a $(n+1) \times (n+1)$ matrix A. We denote the martingale transform by $(A * M^f)_t$, where $(M_t^f)_t$ is the martingale associated to f (using Itô formula). Finally, we project it by conditional expectation to obtain the desired operator. We deduce properties of this operator from basic properties of stochastic analysis. The classical examples are

- The Riesz transforms on $\mathbb{R}^n$ [35],
- The Beurling-Ahlfors operator on $\mathbb{C}$ [7],
- The Hilbert transform on $\mathbb{R}$ [23],
- The Riesz transform and the Beurling-Ahlfors operator on Riemannian manifolds [55].

[1] This transference approach has also been used for the Hilbert transform on the disk [70] in the early days of the sharp weighted theory and for the Riesz vector in Euclidean space [23] to obtain a sharp dimensionless estimate with respect to the well adapted Poisson A_2 characteristic.

In the second result of this thesis (see Chapter 6), we present the dimensionless L^p boundedness of the Riesz vector on manifolds with bounded geometry, for all $p \in (1, \infty)$. As a corollary, we obtain for free the result obtained in [18] with a better numerical constant. We will only consider manifolds with non-negative curvature and use stochastic tools relying on the stochastic representation of Riesz transforms on manifolds by X.-D. Li [54, 58, 55]. To this end, we use Eells-Elworthy-Malliavin construction of a Brownian motion on manifolds (see Chapter 5). It is defined as the projection on the manifold of the solution of a stochastic differential equation, or SDE, (called horizontal Brownian motion), defined on the orthonormal frame bundle $\mathcal{O}(M)$. This construction allows to describe diffusion processes on manifolds and gives rise to the notion of parallel transport along the paths of the diffusion [39, 28].

Our proof is very different from previous ones in that it does not rely on a Bellman function for the problem. Rather, it develops a sparse domination of the stochastic process of Li. See [52] for a first domination pointwise, the elegant and short argument in [49] for the first probabilistic object, namely a discrete time martingale transform and also [21] for the continuous time case. One can deduce, from such domination a dimensionless bound. The sparse operators are particularly well suited for working with weights, which is why this so obtained dimensionless estimate also holds in the weighted setting.

The stochastic process by Li is a specific semi-martingale, built using a pair of martingales that have differential subordination and solving a certain stochastic differential equation. As such, our argument required several tools. One of them is a weak type estimate of the maximal operator of this process, which first appeared in [8]. This is the only part of our proof that uses a (simple) Bellman function. The explicit form of the function is essential and not just its convexity and size properties. The first derivative of said Bellman function is used to control a drift term that arises because the process we consider is not a martingale. Further, we then show that this process has a sparse domination, according to the definition of sparse operator in [21]. The specific form of the defining stochastic equation is used.

Finally, we present a third proof of the dimensionless L^p boundedness of the Riesz vector on manifolds with bounded geometry (See Chapter 6, Section 6.4). Our proof has the significant advantage that it allows negative Bakry-Emery curvatures, yielding to a strong conclusion, namely that of a new dimensionless weighted L^p estimate with optimal exponent and independent of the lower bound of the curvature. Other than previous arguments, only a small part of our proof differs from that of the second result in that we consider the lower bound of the curvature. The Poisson flow and the Riesz transforms are thus defined accordingly. We also introduce a sub-martingale which is the sum of a previously defined martingale and an increasing finite variation process. This allows us to to control terms depending on the lower bound of the curvature, appearing in this case. The rest of the proof resembles that of the second result.

1.3 Notations

We give here a few notations that will be used throughout this thesis. Notations that are used later but not mentioned here are either standard or will be defined when needed.

- $\mathbb{D} = \{z \in \mathbb{C} : |z| \leq 1\}$ will denote the unit disk of the complex plane.
- $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\} = \partial\mathbb{D}$ will denote the unit circle of the complex plane.
- $\mathcal{L}(E; F)$ is the space of bounded linear operators from E into F. $\mathcal{L}(E) := \mathcal{L}(E, E)$.
- Let A be a subset of X. The characteristic function χ_A is the function $\chi_A \colon X \to \{0, 1\}$ defined as
$$\chi_A(x) := \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$
- Let $1 \leq p < \infty$ and (X, μ) be a measure space. Define
$$L^p(X) = \{f \text{ real or complex-valued measurable function on } X \text{ such that } \int_X |f|^p \, d\mu < \infty\}$$
and
$$\|f\|_p = \left(\int_X |f|^p \, d\mu \right)^{1/p}.$$
For a function ω that is positive almost everywhere, let
$$\|f\|_{L^p(\omega)} = \left(\int_X |f|^p \omega \, d\mu \right)^{1/p}$$
- $L^1_{loc}(X)$ is the space of locally integrable functions
$$L^1_{loc}(X) = \{f \text{ measurable s.t. } f\chi_K \in L^1(K), \ \forall K \subset X, \ K \text{ compact}\}.$$
- $C_c^\infty(X)$ stands for the set of smooth (in the sense of having continuous derivatives of all orders) functions that take values in X and are compactly supported. It is also called the set of bump functions.
- On (X, μ) and for $A \subseteq X$, we define $|A| = \mu(A)$.
- $p.v. \left(\frac{1}{x}\right) : C_c^\infty(\mathbb{R}) \to \mathbb{C}$ is defined via the Cauchy principal value as
$$\left[p.v. \left(\frac{1}{x} \right) \right](u) = \lim_{\varepsilon \to 0^+} \int_{\mathbb{R} \setminus [-\varepsilon; \varepsilon]} \frac{u(x)}{x} \, \mathrm{d}x = \int_0^{+\infty} \frac{u(x) - u(-x)}{x} \, \mathrm{d}x \,,$$
for $u \in C_c^\infty(\mathbb{R})$.

- Let $f \in L^1_{loc}(\mathbb{R}^n)$. The Hardy-Littlewood maximal function associated to f is the function $Mf : \mathbb{R}^n \to [0, +\infty]$ defined by

$$Mf(x) = \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)|\, dy.$$

- $T^{r,s}(\mathbb{R}^n)$ is the tensor product

$$T^{r,s}(\mathbb{R}^n) = \underbrace{\mathbb{R}^n \otimes \cdots \mathbb{R}^n}_{r} \otimes \underbrace{(\mathbb{R}^n)^* \times \cdots (\mathbb{R}^n)^*}_{s}.$$

- $x \wedge y = \min(x,y)$ and $x \vee y = \max(x,y)$.

Chapter 2

Preliminaries

2.1 The Hilbert and Riesz transforms

The topic of the Hilbert transform is motivated by its close connection with some of the most important problems in analysis. We can cite for instance the Riemann Hilbert problem for holomorphic functions, BMO spaces, convergence on $L^p(\mathbb{T})$ of partial Fourier sums... etc. But also in more applied mathematics like signal processing [34, 26].
Historically, the Hilbert transform was named after David Hilbert. Its first use dates back to 1905 in Hilbert's work concerning periodic functions, or equivalently for functions on the circle and the theory of the Hilbert transform depended on techniques of complex analysis. With the development of the Calderón-Zygmund school and the extension of one-dimensional theory to higher dimensions, real-variable methods replaced complex analysis. These new methods led to the application of singular integrals (on the real line) in other domains.

2.1.1 The Hilbert transform on $\mathbb{T}$

It is well-known that there is an intimate connection between the Hilbert transform and conjugate harmonic functions in the context of complex analysis.
More specifically, given a real-valued function $f \in L^2(\mathbb{T})$ where $\mathbb{T}$ is the boundary of the disk $\mathbb{D}$, we can produce a function u by the mean of the Poisson integral formula on $\mathbb{D}$ such that $u = f$ (almost everywhere) on $\mathbb{T}$. We may find a harmonic conjugate of u, say $u^\dagger$, such that $u^\dagger(0) = 0$ and $u + iu^\dagger$ is holomorphic on $\mathbb{D}$. What we hope to do is to produce a boundary function $f^\dagger$ for $u^\dagger$. It turns out that the Hilbert transform $H : f \longmapsto f^\dagger$ is such that $P_y * f + iP_y * f^\dagger$ is holomorphic on $\mathbb{D}$ and it is defined on the disk by

$$Hf(e^{i\theta}) = \frac{1}{2\pi} p.v \int_0^{2\pi} f(e^{i\theta}) cot \left(\frac{\theta - t}{2} \right) dt.$$

More explicitly, we have the following scheme

$$\begin{tikzcd}
f \in L^2(\mathbb{T}) \arrow[r, "\text{Boundary Value of } u"] \arrow[d, "H"] & u = P_r * f \in Harm(\mathbb{D}) \arrow[d, "u^\dagger \in Harm(\mathbb{D})"] \\
f^\dagger & u + iu^\dagger \in Hol(\mathbb{D}) \arrow[l, "\text{Boundary Value of } u^\dagger"]
\end{tikzcd}$$

2.1.2 The Hilbert transform on $\mathbb{R}$

Next, we can define its analogue[1] on the real line $\mathbb{R}$, where you can think of $\mathbb{R}$ as the boundary of the upper-half space $\mathbb{C}_+$. This means that for a function $f \in L^2(\mathbb{R})$ there exists a unique harmonic function $F(x,y) = P_y * f(x)$ in $\mathbb{C}_+$, called the Poisson transform of f, such that, in L^2-sense,

$$\lim_{y \to 0^+} F(x,y) = f(x).$$

The function P_y is the Poisson kernel and it is defined by

$$P_y(x) = \frac{1}{\pi} \frac{y}{x^2 + y^2}.$$

Moreover, this function F admits a unique harmonic function G in $\mathbb{C}_+$ vanishing at infinity and such that $F + iG$ is holomorphic in the upper half plane. This function G is known as the conjugate harmonic function to F in $\mathbb{C}_+$ and it is often denoted as $Q_y * f(x)$, where

$$Q_y(x) = \frac{1}{\pi} \frac{x}{x^2 + y^2}$$

is the conjugate Poisson kernel. The Hilbert transform H may then be defined as the boundary value $\lim_{y \to 0+} G(x,y)$, taken in L^2 -sense. In other words:

$$Q_y * f(x) = P_y * (Hf)(x).$$

In this context, H is given by

$$Hf(x) = \frac{1}{\pi} p.v \int_{\mathbb{R}} \frac{f(t)}{x - t} dt.$$

2.1.3 Properties

In what follows, we focus on the Hilbert transform defined on the real line. The following properties are very classical and their proofs as well as more details can be found in any book on harmonic analysis. See for example [33].

[1]The Hilbert transform on the real line and on the disk are denoted the same. But there is no danger of confusion since the class of functions under consideration specifies the context.

One of the most important property of the Hilbert transform is its Fourier transform, given by the following formula

$$\mathcal{F}(Hf)(\xi) = -i\text{sgn}(\xi)\mathcal{F}f(\xi)$$

where sgn is the usual sign function. This formula can be seen as an equivalent definition of the Hilbert transform, since we know that multiplication on Fourier side corresponds to convolution on space:

$$Hf(x) = K_H * f(x),$$

where $K_H(x) := p.v.\dfrac{1}{\pi x}$. In other words, the Hilbert transform H can be considered as a singular integral operator of convolution type whose kernel is $\dfrac{1}{\pi x}$.

Among the numerous properties of the Hilbert transform, due to its great importance in several domains, we are going to list those that will be useful later in this work.

- It is easy to see that H is an isometry on L^2 i.e. $\|Hf\|_2 = \|f\|_2$ by Plancherel's theorem.
- Although the kernel of H is not integrable (and hence we cannot use Young's inequality to prove boundedness on L^p), Marcel Riesz generalized in 1927 the previous result to $L^p(\mathbb{R})$ for all $1 < p < \infty$. In 1972, Pichorides proved that the best constant for the norm of the operator is given by
  $$C_p = \begin{cases} \tan\frac{\pi}{2p} & \text{for } 1 < p \leq 2; \\ \cot\frac{\pi}{2p} & \text{for } 2 < p < \infty. \end{cases}$$
  The same best constants hold for the periodic Hilbert transform.

One may wonder what happens at the endpoints 1 and infinity. By computing the Hilbert transform of the interval [0,1], which is a function in $L^1 \cap L^\infty$, one sees that

$$H(\chi_{[0,1]}) = \frac{1}{\pi} log\left(\frac{|x|}{|x-1|}\right)$$

which is not bounded on L^1, nor on L^∞. As a matter of fact, we have the following properties at the endpoints

- H is of weak-type (1,1). This result is due to Kolmogorov, in 1927 and it means that
  $$\mu\left(\{x \in \mathbb{R} : |Hf(x)| > \lambda\}\right) \leq C\frac{\|f\|_1}{\lambda}, C \geq 1.$$
- Fefferman proved[2] in 1971 that H maps bounded functions into a larger space called bounded mean oscillations and denoted by BMO.

[2]The definition of H on L^∞ slightly differs from the original one on L^p.

- The Hilbert transform relates the square root of the Laplacian to the derivative in this way
$$H \circ \sqrt{-\Delta} = \partial_x.$$
This formula is often used when the studied spaces are groups or manifolds.

2.1.4 The Riesz transform on $\mathbb{R}^n$

We can now study in the same way an n-dimensional analogue of the Hilbert transform. It turns out that there exist n operators in $\mathbb{R}^n$, called the Riesz transforms, which have analogous properties to those of the Hilbert transform on $\mathbb{R}$. These operators are defined for all $1 \leq j \leq n$ by

$$\begin{aligned} R_j f(x) &:= c_n \lim_{\epsilon \to 0} \int_{\mathbb{R}^n \setminus B_\epsilon(x)} \frac{(t_j - x_j) f(t)}{|x-t|^{n+1}} \, dt \\ &= c_n \; p.v. \int_{\mathbb{R}^n} \frac{(t_j - x_j) f(t)}{|x-t|^{n+1}} \, dt, \end{aligned}$$

with $c_n = \dfrac{\Gamma[(n+1)/2]}{\pi^{(n+1)/2}}$, a dimensional normalization. Again, these operators are given by a convolution with the kernel

$$K_j(x) = c_n \; p.v. \frac{x_j}{|x|^{n+1}}.$$

Equivalently, we can define these operators by the means of the Laplacian and first order derivatives as follows
$$R_j \circ \sqrt{-\Delta} = \partial_j$$
or more formally
$$R = \nabla \circ (-\Delta)^{-1/2}.$$
(In this formula, R is understood to be the vector of n components.)
The Riesz transforms arises in the study of differentiability properties of harmonic potentials in potential theory and harmonic analysis. In particular, the Riesz transforms allows to recover information about the entire hessian of a function from knowledge of only its Laplacian.
It is natural to believe that the Riesz transforms' properties are similar to the Hilbert transform ones. We may cite for instance

- $\mathcal{F}(R_j f)(\xi) = -i \dfrac{\xi_j}{|\xi|} \mathcal{F} f(\xi) \; \forall j = 1, \cdots, n.$
- An immediate consequence is that the Riesz transforms are bounded on L^2.
- For all $1 < p < \infty$, T. Iwaniec and G. Martin proved that there exists a constant $C_p > 0$ independent of n such that
$$\|R_j f\|_p \leq C_p \|f\|_p, \; \forall j = 1, \cdots, n.$$
This constant is equal to Pichorides constant $C_p = \cot(\frac{\pi}{2p^*})$, where $p^* = \max(p, \frac{p}{p-1})$.

- By defining $\|Rf\|_p := \|(\sum_{i=1}^n |R_i f|^2)^{1/2}\|_p$ the L^p norm of the vectorial Riesz transform, Bañuelos and Wang proved that

$$\|Rf\|_p \leq 2(p^* - 1)\|f\|_p, \ \forall \ p \in (1, \infty).$$

  This is so far the best known result and the sharp bound remains open. It is conjectured in Problem 6 in [5] that the sharp bound should be $C_p = \cot(\frac{\pi}{2p^*})$. We note that Bañuelos and Wang's bound does not give the sharp bound even when $p = 2$, which by the Fourier transform is 1.
  The knowledge of the exact value (or even a better estimate) of the L^p norm of the Riesz vector leads to important applications in the study of quasi-conformal mappings and related non-linear geometric PDEs as well as in the L^p Hodge decomposition theory.

2.2 Semigroups

This section is devoted to semigroups and their properties. They mostly appear in the theory of linear evolution equations, but they also have a rich interplay with other subjects in functional analysis, stochastic analysis and mathematical physics.

Strongly continuous semigroups (which will be defined below) arise when we want to find a solution for a Cauchy problem given by

$$\begin{cases} u'(t) &= Au(t), \quad t \in [0, T] \\ u(0) &= u_0, \quad u_0 \in X, \end{cases} \tag{2.1}$$

where A is a linear (non necessarily bounded) operator defined on a domain $D(A)$ of a Banach space X. The most common problem is the heat equation, where $(A, D(A))$ is (Δ, D), for some open domain $D \subset \mathbb{R}^n$.
A naive approach would be to suggest a solution given by $u(t) = e^{t\Delta}u_0$. The problem is that if we work on a very general space, this exponential cannot be defined via matrices anymore. This is why we turn ourselves to semigroups.

Definition 2.1. *Let X be a Banach space over $\mathbb{R}$ or $\mathbb{C}$.*

- *A semigroup on X is a map $T : \mathbb{R}_+ \to \mathcal{L}(X)$ such that*
  - $T(0) = Id$
  - $\forall t, s \geq 0 : \ T(t+s) = T(t)T(s).$
- *We say that $(T(t))_{t\geq 0}$ is strongly continuous on X if*

$$\forall x_0 \in X : \ \|T(t)x_0 - x_0\| \to 0, \ \text{as } t \downarrow 0$$

The first two axioms are algebraic while the last one is topological, and states that the map T is continuous in the strong operator topology. We denote by C_0-semigroups all strongly continuous semigroups.

Remark 2.1. *In what follows, we are only considering C_0-semigroups.*

As mentioned earlier, the theory of strongly continuous semigroups was developed in order to study existence and uniqueness of solutions to the evolution equations. To that purpose, we need to define the *generator* of a semigroup.

Definition 2.2. *Let $(T(t))_{t\geq 0}$ be a strongly continuous semigroup on X. The generator of $(T(t))_{t\geq 0}$ is the operator A defined by*

$$D(A) = \{x \in X \text{ such that } \lim_{t\downarrow 0} \frac{1}{t}\,(T(t) - I)\,x \text{ exists in } X\}$$

$$A\,x = \lim_{t\downarrow 0} \frac{1}{t}\,(T(t) - I)\,x,\ \forall\ x \in D(A).$$

The strongly continuous semigroup T with generator A is often denoted by the symbol e^{At}. This notation is compatible with the notation for matrix exponentials, and for functions of an operator defined via functional calculus (for example, via the spectral theorem). Another common notation for semigroups is P_t.
The next proposition will suggest that strongly continuous semigroups generated by an operator A are indeed the objects to use in order to construct a solution of the Problem (2.1). For more details and proofs of the classical results given below see, e.g., [64], [30] and references therein.

Proposition 2.1. *Let $(T(t))_{t\geq 0}$ be a strongly continuous semigroup on X whose generator is A.*

- *If $f \in D(A)$, then $T(t)f \in D(A)$, $\forall t \geq 0$.*
- *If $f \in D(A)$, then $t \mapsto T(t)f$ is differentiable on $\mathbb{R}_+$ and*

$$\frac{d}{dt}T(t)f = AT(t)f = T(t)Af.$$

Among C_o-semigroups, there are particular semigroups that enjoy specific properties. We list now those that shall frequently appear in this thesis.

Definition 2.3. *Let $(T(t))_{t\geq 0}$ be a strongly continuous semigroup on X. We say that*

- *$(T(t))_{t\geq 0}$ is positive if $(T(t))f \geq 0$, $\forall f \geq 0$, $\forall t \geq 0$, for almost every $x \in X$.*
- *$(T(t))_{t\geq 0}$ is contractive if $\|T(t)\|_{\mathcal{L}(E)} \leq 1$, $\forall t \geq 0$.*
- *$(T(t))_{t\geq 0}$ is L^∞-contractive if for every $t \geq 0$ and $f \in L^2(X) \cap L^\infty(X)$ we have*

$$\|T(t)f\|_{L^\infty} \leq \|f\|_{L^\infty}$$

- *If the semigroup is both contractive and positive, we say that it is sub-Markovian.*

- *A Markov semigroup is a sub-Markov semigroup which is conservative, i.e.* $T(t)1 = 1$*. This property is referred to as mass conservation.*

Example 2.1. *As mentioned earlier, the following example is one of the main purposes to develop the semigroup theory. It is often called the heat semigroup, Gaussian semigroup or n-diffusion semigroup. We consider it in this example on* $L^2(\mathbb{R}^n)$*, where it is defined explicitly by*

$$P_t f(x) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-|x-y|^2/4t} f(y)dy,$$

for $t > 0$*,* $x \in \mathbb{R}^n$ *and* $f \in L^2(\mathbb{R}^n)$*. It is the unique solution of the heat equation*

$$\begin{cases} \partial_t u(t,x) & = & \Delta u(t,x), \\ u_0 & = & f \in L^2(\mathbb{R}^n), \end{cases}$$

and we usually denote the heat semigroup by $e^{t\Delta} f$*.*
This semigroup P_t *forms a Markov (and hence strongly continuous) semigroup on* $L^2(\mathbb{R}^n)$ *for* $t > 0$ *with* $P_0 = Id$ *and its generator coincides with the closure of the Laplace operator*

$$\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$$

for every function in the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ *[30].*
By putting

$$p_t(x) = \frac{1}{(4\pi t)^{n/2}} e^{-|x|^2/4t},$$

we rewrite the semigroup as

$$P_t f(x) = p_t * f(x),$$

and p_t *is called the kernel of the semigroup. The kernel of the semigroup is* C^∞ *in* $\mathbb{R}_+ \times \mathbb{R}^n$*, positive, and satisfies*

$$\partial_t p_t = \Delta p_t \quad \text{and} \quad \int_{\mathbb{R}^n} p_t(x)dx = 1.$$

We will see further that we can associate to the semigroup P_t *(and more generally to any Markov semigroup) a Markov process.*

Example 2.2. *Another important example on* $\mathbb{R}^n$ *is the Poisson semigroup, that we denote in this example by* P_t*. In fact, if we denote by* H_t *the heat semigroup, then these semigroups are related by Bochner's subordination formula as follows*

$$\begin{aligned} P_t f(x) &= \frac{1}{\sqrt{\pi}} \int_0^\infty \frac{e^{-s}}{\sqrt{s}} H_{\frac{t^2}{4s}} f(x) ds \\ &= \int_0^\infty H_s f(x) d\lambda_s, \end{aligned}$$

where $d\lambda_s = \dfrac{t}{2\sqrt{\pi}} e^{-t^2/4s} s^{-3/2} ds.$

The Poisson semigroup P_t *satisfies*

$$\begin{cases} \partial^2_{tt} u(t,x) &= -\Delta u(t,x), \\ u_0 &= f \in L^2(\mathbb{R}^n). \end{cases}$$

Another common notation of the Poisson semigroup is $e^{-t\sqrt{-\Delta}} f$.

2.3 Analytical representation of the Riesz transform

We highlight in this section the intimate connection of Riesz transforms and harmonic functions on $\mathbb{R}^n$. Indeed, we present a Littlewood-Paley type formula for the Riesz transforms using the Poisson flow.

Lemma 2.1. *Let* $R = (R_1, \cdots, R_n)$ *be the Riesz vector on* $\mathbb{R}^n$ *and* P_t *be the Poisson semigroup. Then we have*

$$\langle Rf, g\rangle_{L^2(\mathbb{R}^n)} = -4 \int_0^\infty \langle \nabla P_t f, \frac{d}{dt} P_t g\rangle_{L^2(\mathbb{R}^n)} t dt.$$

Proof. To prove this lemma, we use the fact that for sufficiently decaying function F, we have

$$F(0) = \int_0^\infty F''(t) t dt,$$

by integrating by parts twice.
Hence if we take $F(t) = \langle P_t Rf, P_t g\rangle_{L^2(\mathbb{R}^n)}$, we obtain

$$\begin{aligned} F(0) &= \langle Rf, g\rangle_{L^2(\mathbb{R}^n)} \\ &= \int_0^\infty \frac{d^2}{dt^2} \langle P_t Rf, P_t g\rangle_{L^2(\mathbb{R}^n)} t dt \\ &= \int_0^\infty (\langle \frac{d^2}{dt^2} P_t Rf, P_t g\rangle_{L^2(\mathbb{R}^n)} + \langle P_t Rf, \frac{d^2}{dt^2} P_t g\rangle_{L^2(\mathbb{R}^n)} \\ &\quad + 2\langle \frac{d}{dt} P_t Rf, \frac{d}{dt} P_t g\rangle_{L^2(\mathbb{R}^n)}) t dt \\ &= 4 \int_0^\infty \langle \sqrt{-\Delta} P_t Rf, \sqrt{-\Delta} P_t g\rangle_{L^2(\mathbb{R}^n)} t dt, \end{aligned}$$

where the two last equalities come from the fact that $\partial_t P_t f = -\sqrt{-\Delta} P_t f$, $\partial^2_{tt} P_t f = -\Delta P_t f$ and Δ is symmetric. Next, recall that $R = \nabla \circ (-\Delta)^{-1/2}$ and $(-\Delta)^{1/2}$ commutes with P_t and ∇. Using one more time the fact that $\partial_t P_t g = -\sqrt{-\Delta} P_t g$ we obtain

$$\langle Rf, g\rangle_{L^2(\mathbb{R}^n)} = -4 \int_0^\infty \langle \nabla P_t f, \frac{d}{dt} P_t g\rangle_{L^2(\mathbb{R}^n)} t dt,$$

as claimed. $\square$

2.4 Stochastic calculus

Recent years have witnessed a considerable effort of using probability theory of stochastic processes as a powerful tool in the study of problems from harmonic analysis. The close connection of stochastic differential theory and harmonic analysis allows to obtain sharp results. We recall here some well-known notions and facts in probability theory. Most of definitions are taken from the book [73].

2.4.1 Bases of stochastic calculus

In what follows, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathcal{F}_t)_{t\geq 0} \subseteq \mathcal{F}$ be a filtration of the underlying probability space such that $\mathcal{F}_s \subseteq \mathcal{F}_t$ for all $s \leq t$.
We always assume that the filtered probability space satisfies the usual hypothesis i.e.

1. $\mathcal{F}_0$ contains all the $\mathbb{P}$-null sets of $\mathcal{F}$;

2. $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$, meaning that the filtration is right continuous.

Conditional expectation. The conditional expectation of a random variable is its expected value given that a certain set of condition is known to occur.

Definition 2.4. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathcal{F}_t)_{t\geq 0}$ be a filtration of the underlying probability space, $X : \Omega \to \mathbb{R}^n$ be a random variable on that probability space with finite expectation and $\mathcal{H} \subseteq \mathcal{F}$ a sub-sigma-algebra of $\mathcal{F}$. Then there exists a random variable Z that is $\mathcal{H}$-measurable and integrable such that for every bounded and $\mathcal{H}$-measurable random variable U we have*

$$\mathbb{E}[XU] \;=\; \mathbb{E}[ZU].$$

We write then

$$Z \;=\; \mathbb{E}[X|\mathcal{H}]$$

and call Z the conditional expectation of X given $\mathcal{H}$. This random variable is well defined because if Y is another random variable satisfying the same property, then $Y = Z$ almost surely.

The following properties are considered standard and their proofs can be found in the literature.

Proposition 2.2. *Let X, Y be random variables and $\mathcal{F}$ a filtration. Then*

1. *The conditional expectation is linear;*

2. *If X is $\mathcal{F}$-measurable then $\mathbb{E}(XY|\mathcal{F}) = X\mathbb{E}(Y|\mathcal{F})$;*

3. *$\mathbb{E}(\mathbb{E}(X|\mathcal{F})) = \mathbb{E}(X)$;*

4. *If $f : \mathbb{R} \to \mathbb{R}$ is a convex function, then $f(E(X|\mathcal{F})) \leq E(f(X)|\mathcal{F})$;*

5. *The conditional expectation is $L^p(\Omega, \mathcal{F}, \mathbb{P})$ contractive for $p \geq 1$;*

6. *If $\mathcal{H}_1 \subset \mathcal{H}_2 \subset \mathcal{F}$ then $\mathbb{E}(\mathbb{E}(X|\mathcal{H}_2)|\mathcal{H}_1) = \mathbb{E}(X|\mathcal{H}_1)$.*

More generally for a fixed a random variable W such that $W > 0$ a.s, we define a weighted probability space $(\Omega, \mathcal{F}, \hat{\mathbb{P}})$ -where $\hat{\mathbb{P}} = \dfrac{W d\mathbb{P}}{\mathbb{E}(W)}$ is the weighted measure- as

$$\hat{\mathbb{P}}(A) = \frac{\mathbb{E}(\chi_A W)}{\mathbb{E}(W)},$$

where A is a subset of Ω. We denote as well by $\mathbb{E}_W(\cdot)$ the expectation with respect to $\hat{\mathbb{P}}$.
By definition of the conditional expectation, we also have

$$\mathbb{E}_W(\chi_A|\mathcal{F}_t) = \frac{\mathbb{E}(\chi_A W|\mathcal{F}_t)}{\mathbb{E}(W|\mathcal{F}_t)}.$$

Indeed, let X be a $\mathcal{F}$-measurable random variable and W a weight. Let $Y = \dfrac{\mathbb{E}(XW|\mathcal{F})}{\mathbb{E}(W|\mathcal{F})}$.
We want to prove that for every $G \in \mathcal{F}$ we have $\mathbb{E}_W(X\chi_G) = \mathbb{E}_W(Y\chi_G)$. By uniqueness of the conditional expectation we would have $Y = \mathbb{E}_W(X|\mathcal{F})$ a.s. By using the properties of the conditional expectation listed above we have

$$\begin{aligned}
\mathbb{E}_W(Y\chi_G) = \frac{\mathbb{E}(Y\chi_G W)}{\mathbb{E}(W)} &= \frac{\mathbb{E}\left(\frac{\mathbb{E}(XW|\mathcal{F})}{\mathbb{E}(W|\mathcal{F})}\chi_G W\right)}{\mathbb{E}(W)} \\
&= \frac{\mathbb{E}\left(\mathbb{E}\left(\frac{\mathbb{E}(XW|\mathcal{F})}{\mathbb{E}(W|\mathcal{F})}\chi_G W|\mathcal{F}\right)\right)}{\mathbb{E}(W)} \\
&= \frac{\mathbb{E}\left(\frac{\mathbb{E}(XW|\mathcal{F})}{\mathbb{E}(W|\mathcal{F})}\chi_G \mathbb{E}\left(W|\mathcal{F}\right)\right)}{\mathbb{E}(W)} \\
&= \frac{\mathbb{E}\left(\mathbb{E}(X\chi_G W|\mathcal{F})\right)}{\mathbb{E}(W)} \\
&= \frac{\mathbb{E}(X\chi_G W)}{\mathbb{E}(W)} \\
&= \mathbb{E}_W(X\chi_G),
\end{aligned}$$

as claimed.

Definition 2.5. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A stochastic process is a collection $(X_t)_{t\geq 0}$ of random variables that are measurable with respect to $\mathcal{F}$.*
A process X such that for each t it is $\mathcal{F}_t$-measurable is called adapted.
A process X is called càdlàg if it has right continuous sample paths, with left limits.

A stopping time is a random variable whose value is interpreted as the time at which a given stochastic process (or sequence of random variables) exhibits a certain behaviour of interest. More precisely

Definition 2.6. *A random variable $\tau : \Omega \to \mathbb{R}_+ \cup \{+\infty\}$ is called a stopping time with respect to $(\mathcal{F}_t)_{t\geq 0}$ if*

$$\forall t \geq 0, \quad \{\omega \in \Omega : \ \tau(\omega) \leq t\} \in \mathcal{F}_t.$$

A martingale is a sequence of random variables for which, at a particular time in the realized sequence, the expectation of the next value in the sequence is equal to the present observed value even given knowledge of all prior observed values.

Definition 2.7. *Let $M = (M_t)_{t\geq 0}$ be a stochastic process. We say that $(M_t)_{t\geq 0}$ is a martingale with respect to $(\mathcal{F}_t)_{t\geq 0}$ if*

1) $(M_t)_{t\geq 0}$ is adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$.

2) $(M_t)_{t\geq 0}$ is integrable for every t.

3) $\mathbb{E}(M_t|\mathcal{F}_s) = M_s, \ \forall t \geq s$.

If we replace in 3) the equality sign with $\leq$ or $\geq$ then M is called super-martingale or sub-martingale, respectively.

Proposition 2.3. *Let $(M_t)_{t\geq 0}$ be a martingale. We have for all $t \geq s \geq 0$*

$$\mathbb{E}(M_t) = \mathbb{E}(\mathbb{E}(M_t|\mathcal{F}_s)) = \mathbb{E}(M_s) = \ldots = \mathbb{E}(M_0)$$

In other words, the sequence $(\mathbb{E}(M_t))_{t\geq 0}$ is constant.

Brownian motions are named after the botanist Robert Brown and are originated as a model of the phenomenon that pollen grains suspended in water have a continual swarm behaviour. They are nowadays a fundamental example in the theory of continuous in time stochastic processes.

Definition 2.8. *A stochastic process $B = (B_t)_{t\geq 0}$ adapted to a filtration $\mathcal{F}$ is called a Brownian motion if*

1. $B_0 = 0$ almost surely;

2. B_t is almost surely continuous;

3. B_t has independent increments;

4. $B_t - B_s \sim \mathcal{N}(0, t-s)$ for $0 \leq s \leq t$.

Definition 2.9. *Let X_t, $t \geq 0$ be a real-valued stochastic process defined on $(\Omega, \mathcal{F}, \mathbb{P})$. We denote by $\langle X\rangle_t$ or $\langle X, X\rangle_t$ its quadratic variation and define it by*

$$\langle X, X\rangle_t = \lim_{\|P\|\to 0} \sum_{k=1}^{n} (X_{t_k} - X_{t_{k-1}})^2,$$

where P ranges over partitions of the interval $[0,t]$ and the norm of the partition P is the mesh. This limit, if it exists, is defined using convergence in probability.
More generally, the covariation of two processes X and Y is defined as

$$\begin{aligned}\langle X, Y\rangle_t &= \lim_{\|P\|\to 0} \sum_{k=1}^{n} \left(X_{t_k} - X_{t_{k-1}}\right)\left(Y_{t_k} - Y_{t_{k-1}}\right) \\ &= \tfrac{1}{4}(\langle X+Y\rangle_t - \langle X-Y\rangle_t),\end{aligned}$$

where the second equality is given by the polarization identity.

Example 2.3. *If $B = (B_t)_{t\geq 0}$ is a Brownian motion then $\langle B, B\rangle_t = t$.*

Definition 2.10. *Let X and Y be two martingales. We say that the martingale Y is said differentially subordinate to the martingale X if the process $(\langle X, X\rangle_t - \langle Y, Y\rangle_t)_{t\geq 0}$ is non-negative and non-decreasing in t.*

Itô calculus. The Itô formula serves as the stochastic calculus counterpart of the chain rule. It can be heuristically derived by forming the Taylor series expansion of the function up to its second derivatives.

Theorem 2.1. *Let $(M_t)_{t\geq 0}$ be a continuous martingale and $F : \mathbb{R} \to \mathbb{R}$ a twice differentiable function. Then $(F(M_t))_{t\geq 0}$ is a semi-martingale and*

$$F(M_t) = F(M_0) + \underbrace{\int_0^t F'(M_s)dM_s}_{\textit{martingale part}} + \underbrace{\frac{1}{2}\int_0^t F''(M_s)d\langle M, M\rangle_s}_{\textit{finite variation part}}, \ \forall t \geq 0, \ \mathbb{P} - a.s.$$

Equivalently for all $t \geq 0$ and $\mathbb{P} - a.s$,

$$dF(M_t) = F'(M_t)dM_t + \frac{1}{2}F''(M_t)d\langle M, M\rangle_t.$$

Note that semi-martingales are real valued process defined on the filtered probability space that can be decomposed as sum of a martingale and a cádlág adapted process of locally bounded variation.

We also present another form of Itô formula to find the differential of a time-dependent function of a stochastic process.

Theorem 2.2 (second form of Itô formula)**.** *Let $(M_t)_{t\geq 0}$ be a continuous martingale and $F : [0,T]\times\mathbb{R}\to\mathbb{R}$ a function that is differentiable in the first variable (time) and twice differentiable in the second one. Then*

$$\begin{aligned} F(t,M_t) = \quad & F(0,M_0) + \int_0^t \frac{\partial F}{\partial s}(s,M_s)ds + \int_0^t \frac{\partial F}{\partial x}(s,M_s)dM_s \\ & + \frac{1}{2}\int_0^t \frac{\partial^2 F}{\partial x^2}(s,M_s)d\langle M,M\rangle_s, \ \forall t\geq 0, \ \mathbb{P}-a.s. \end{aligned}$$

Equivalently for all $t\geq 0$ and $\mathbb{P}-a.s$,

$$dF(t,M_t) = \frac{\partial F}{\partial t}(t,M_t)dt + \frac{\partial F}{\partial x}(t,M_t)dM_t + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}(t,M_t)d\langle M,M\rangle_t.$$

Remark 2.2. *1) Another version of the previous theorem exists for $\mathbb{R}^d$-valued martingales where we sum the integrals in d variables.*

2) There is another alternative to the Itô formula which is the Stratonovich formula. Unlike the Itô calculus, Stratonovich integrals are defined such that the chain rule of ordinary calculus holds, which makes them easier to be manipulated. It is possible to convert between the two formulas whenever one definition is more convenient by the formula

$$\int_0^t H_s(\omega)\circ dZ_s(\omega) = \int_0^t H_s(\omega)dZ_s(\omega) + \frac{1}{2}\langle H,Z\rangle_t.$$

The symbol $\circ$ is called Itô's circle.

We end our review of martingales with Doob's inequality.

Theorem 2.3. *Let $X_t = \mathbb{E}(X|\mathcal{F}_t)$ be the martingale generated by the $\mathbb{P}$-integrable random variable X (more generally we can assume that X_t is a non-negative sub-martingale) and $X^* := \sup_{s\geq 0}|X_s|$ is the maximal function of $(X_s)_{0\leq s}$. Then for every $p>1$ we have*

$$\|X\|_p \leq \|X^*\|_p \leq \frac{p}{p-1}\|X\|_p,$$

where $\|X\|_p^p := \sup_{t\geq 0}\|X_t\|_p^p = \int_\Omega |X_t(w)|^p d\mathbb{P}(w)$.
Moreover, these constants are the best possible.

We present a more general weighted version of Doob's inequality. Let us first recall the general setting of A_p martingales. Fix a random variable W such that $W>0$ a.s. For $p>1$ we say that W is an A_p weight if

$$Q_p(W) = \sup_t \|W_t\left(\mathbb{E}[(W^{\frac{-1}{p-1}})|\mathcal{F}_t]\right)^{p-1}\|_{L^\infty} < \infty,$$

where $W_t = \mathbb{E}(W|\mathcal{F}_t)$.

The following was proved in [21]

Theorem 2.4. *Let $X_t = \mathbb{E}_W(X|\mathcal{F}_t)$ be the martingale generated by the $Wd\mathbb{P}$-integrable random variable X. We denote by $X^* = \sup_{t\geq 0} X_t$ the maximal function of X. Then for every $p > 1$ we have*

$$\|X^*\|_{L^p(W)} \leq Q_p(W)^{\frac{1}{p-1}} \frac{p^{p'}}{p-1} \|X\|_{L^p(W)},$$

where p' is the conjugate of p. Moreover, this result is sharp in terms of the dependence on the characteristic.

2.4.2 Feynman-Kac formula

Diffusion processes and more specifically Brownian motions originated in physics as mathematical models of motion of molecules which are subject to collisions with other molecules in a gas or fluid.
In the 40's, Richard Feynman discovered that the Schrodinger equation can be solved by averaging over paths. Based on this discovery, Mark Kac observed that a similar representation works for solutions of the heat equation with an external term. This representation is now called the Feynman-Kac formula.
It comes today as no surprise that the formula has been generalized for other diffusion processes. Indeed, let $\mathbb{P}^x$ be a family of probability measures on some probability space, one for each possible initial point x under which the stochastic process $X : [0, \infty) \times \Omega \to \mathbb{R}$ is a diffusion process that starts at x. That is, under each $\mathbb{P}^x$ the process X_t obeys the following differential equation

$$\begin{cases} dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t \\ X_0 = x \end{cases}$$

where W is a Brownian motion, μ and σ are respectively the drift and the diffusion fields.

Definition 2.11. *Define the infinitesimal generator of the process X_t to be the (generally unbounded) differential operator acting on suitable functions f by*

$$Af(t, x) = \lim_{t \downarrow 0} \frac{\mathbb{E}[f(t, X_t)] - f(t, x)}{t}.$$

For a process X defined as above and any function f compactly supported and C^2, we have

$$Lf(t, x) = \mu(x)f'(t, x) + \frac{1}{2}\sigma^2(t, x)f''(t, x).$$

Then we have the following Feynman-Kac formula

Theorem 2.5. *Assume that μ and σ are globally Lipschitz and of at most a polynomial growth in the variable x. Let f and K be continuous functions such that $K \geq 0$ and $f(x) = O(|x|)$ as $|x| \to \infty$. Then the function u defined by*

$$u(t, x) = \mathbb{E}^x \left(exp \left(- \int_0^t K(s, X_s)ds \right) f(X_t) \right)$$

satisfies the diffusion equation

$$\frac{\partial u}{\partial t} = Lu - Ku,$$

with the initial condition $u(0,x) = f(x)$.

Remark 2.3. *The Feynman-Kac theorem for 1-dimensional diffusion processes extends naturally to multidimensional diffusion processes.*

Remark 2.4. 1. *We can also to a Markov semigroup* $T(t)$ *a Markov process* $(X_t)_{t\geq 0}$. *Indeed, let* $(\Omega, \mathcal{F})$ *be a measurable space. For* $x \in \Omega$, *let* $\mathbb{P}^x$ *be the law of* X_t^x, *the process starting at* x *a.s and* $\mathbb{E}^x$ *be the expectation under* $\mathbb{P}^x$. *Then for any* $f \in L^2(d\mu)$ *we have*

$$T(t)f(x) = \mathbb{E}^x(f(X_t)) = \mathbb{E}(f(X_t)|X_0 = x).$$

2. *Back to Example 2.1, we can associate a Markov process* $(X_t)_{t\geq 0}$ *to* P_t, *for any* $f \in L^2(\mathbb{R}^n)$ *we have*

$$P_t f(x) = \mathbb{E}^x(f(X_t)).$$

2.5 Probabilistic representation of the Riesz transform

We previously saw in Section 2.3 a deterministic representation of the Riesz transforms involving semigroups. There is also a probabilistic approach due to Gundy and Varopoulos in [35] and later to Gundy and Silverstein in [37] which expresses the Riesz transform as the conditional expectation of a martingale transform. The core of this representation is the definition of the background radiation process and the fact that a function $f \in L^p(\mathbb{R}^n)$ can be expressed as a stochastic integral involving its harmonic extension to the upper half space $\mathbb{R}^n \times \mathbb{R}_+$ and a $(n+1)\times(n+1)$-dimensional Brownian motion.

Indeed, following Gundy and Varopoulos, let $Z_t = (X_t, B_t)$, where X_t is the Brownian motion on $\mathbb{R}^n$ and B_t a Brownian motion on $\mathbb{R}$ starting at $y > 0$. Let $\tau = \inf\{t \geq 0 : B_t = 0\}$. The process $(Z_t)_{0\leq t\leq \tau}$ is a diffusion process on $\mathbb{R}^n \times \mathbb{R}_+$ and terminates at time τ upon hitting the boundary $\mathbb{R}^n \times \{0\}$.
Define $Qg(x,y) = e^{-y\sqrt{-\Delta}}g(x)$ to be the Poisson integral of g i.e. the harmonic extension of g on $\mathbb{R}^n \times \mathbb{R}_+$. By Itô formula we have

$$g(X_\tau) = Qg(Z_\tau) = Qg(Z_0) + \int_0^\tau \nabla Qg(Z_s)dZ_s.$$

First, we claim that for any $(n+1)\times(n+1)$ matrix A, we have

$$\langle \lim_{y\to\infty} \mathbb{E}\left(\int_0^\tau A\nabla Qf(Z_s)dZ_s|X_\tau = x\right), g\rangle = 2\int_0^\infty \int_{\mathbb{R}^n} \langle A\nabla Qf(x,z), \nabla Qg(x,z)\rangle z dx dz. \tag{2.2}$$

Indeed, let $T_A f(x) = \mathbb{E}\left(\int_0^\tau A\nabla Q f(Z_s) dZ_s | X_\tau = x\right)$. By dualizing we have

$$\begin{aligned}
\langle T_A f, g\rangle &= \mathbb{E}\left(\langle \mathbb{E}\left(\int_0^\tau A\nabla Q f(Z_s) dZ_s | X_\tau = x\right), g(X_\tau)\rangle\right) \\
&= \mathbb{E}\left(\mathbb{E}\left(\langle \int_0^\tau A\nabla Q f(Z_s) dZ_s, g(X_\tau)\rangle | X_\tau = x\right)\right) \\
&= \mathbb{E}\left(\langle \int_0^\tau A\nabla Q f(Z_s) dZ_s, g(X_\tau)\rangle\right) \\
&= \mathbb{E}\left(\langle \int_0^\tau A\nabla Q f(Z_s) dZ_s, Qg(Z_0)\rangle\right) \\
&\quad + \mathbb{E}\left(\langle \int_0^\tau A\nabla Q f(Z_s) dZ_s, \int_0^\tau \nabla Q g(Z_s) dZ_s\rangle\right) \\
&= \mathbb{E}\left(\int_0^\tau \langle A\nabla Q f(Z_s), \nabla Q g(Z_s)\rangle ds\right) \\
&\underset{y\to+\infty}{\longrightarrow} 2\int_0^\infty \int_{\mathbb{R}^n} \langle A\nabla Q f(x,z), \nabla Q g(x,z)\rangle z dx dz.
\end{aligned}$$

In the last equality, we used Lemma 5.2. More rigorous calculations will be detailed in Chapter 5 in a more general context.

Next, for a matrix A that permutes y and x_k (with zeros everywhere else), we obtain the following

$$\begin{aligned}
\int_{\mathbb{R}^n} \langle \lim_{y\to\infty} \mathbb{E}\left(\int_0^\tau \frac{\partial Q f}{\partial x_k} dy_s | X_\tau = x\right), g(X_\tau)\rangle dx &= 2\int_0^\infty \int_{\mathbb{R}^n} \langle \frac{\partial Q f}{\partial x_k}(x,y), \frac{\partial Q g}{\partial y}(x,y)\rangle y dx dy \\
&= -\frac{1}{2}\int_{\mathbb{R}^n} \langle R_k f, g\rangle dx,
\end{aligned}$$

where in the last equality we used the analytical representation of the Riesz transform. We can now state the following result [35]

Theorem 2.6. *Let $A_j = (a_{ik})$ be the $(n+1)\times(n+1)$ matrix with $a_{ik} = 0$ unless $i = n+1$, $k = j$ and $a_{(n+1)j} = 1$. Then we have the Gundy Varopoulos formula*

$$-\frac{1}{2} R_j f = \lim_{y\to\infty} \mathbb{E}\left(\int_0^\tau A_j \nabla Q f(Z_s) dZ_s | X_\tau = x\right).$$

2.6 Example: the Gaussian space

Now that we have a pretty good understanding at what is happening on $\mathbb{R}^n$ endowed with the Lebesgue measure, we want to investigate further. Instead of working on $(\mathbb{R}^n, dx)$, we replace the Lebesgue measure by the Gaussian measure γ on $\mathbb{R}^n$ such that

$$d\gamma(x) = e^{-\frac{\|x\|^2}{2}}.$$

The pair $(\mathbb{R}^n, d\gamma)$ is called the Gaussian space. We define a new Laplacian

$$\Delta_{OU} f(x) = \Delta f(x) - x \cdot \nabla f(x), \;\; f \in C_c^\infty$$

called the Ornstein-Uhlenbeck operator. This definition allows to obtain the same results known in the classical Calderón-Zygmund theory with the Laplacian on the classical Lebesgue space $(\mathbb{R}^n, dx)$.
The operator Δ_{OU} defined on $C_c^\infty(\mathbb{R}^n)$ admits a self-adjoint extension to $L^2(\mathbb{R}^n, d\gamma)$, also denoted Δ_{OU}. This operator is negative and symmetric with respect to the measure γ since for all $f, g \in C_c^\infty$ we have

$$\int_{\mathbb{R}^n} (\nabla f, \nabla g) d\gamma(x) = -\int_{\mathbb{R}^n} f(\Delta_{OU} g) d\gamma(x) = -\int_{\mathbb{R}^n} g(\Delta_{OU} f) d\gamma(x)$$

Hence, it generates a diffusion semigroup P_t defined by the Mehler formula

$$P_t f(x) := \int_{\mathbb{R}^n} f(e^{-t}x + \sqrt{1 - e^{-2t}}y) d\gamma(y).$$

It can also be expressed using the so-called Mehler kernel M_t

$$P_t f(x) = \int_{\mathbb{R}^n} M_t(x, y) f(y) d\gamma(y),$$

where

$$M_t(x, y) = \frac{1}{\pi^{n/2}(1 - e^{-2t})^{n/2}} \exp\left(\frac{\|e^{-tx} - y\|^2}{1 - e^{-2t}}\right).$$

This semigroup is solution of

$$\begin{cases} \partial_t u(t, x) & = & \Delta u(t, x) - x \cdot \nabla u(t, x), \\ u_0 & = & f \in L^2(\mathbb{R}^n), \end{cases}$$

and we usually denote the Ornstein Uhlenbeck semigroup by $e^{t\Delta_{OU}} f$. We refer the reader to [76] for more details and list below some properties of this semigroup

Proposition 2.4. *For all $p \in [1, \infty)$ we have*

1. *$\{P_t, t \geq 0\}$ is a contraction semigroup on $L^p(\mathbb{R}^n, d\gamma)$;*
2. *$\{P_t, t \geq 0\}$ is strongly continuous on $L^p(\mathbb{R}^n, d\gamma)$;*
3. *γ is an invariant measure for $\{P_t, t \geq 0\}$.*

Finally, in [60], Meyer introduced the Riesz vector associated with the Ornstein Uhlenbeck operator Δ_{OU} by

$$R(\Delta_{OU}) = \nabla(-\Delta_{OU})^{-1/2}.$$

Analytical representation We obtain the analytical representation of the Riesz vector associated with the Ornstein Uhlenbeck operator by repeating the same proof as in Section 2.3, where $R = \nabla(-\Delta_{OU})^{-1/2}$, $T_t f(x) = e^{-t\sqrt{-\Delta_{OU}}}f(x)$ is the Poisson semigroup and $(\mathbb{R}^n, dx)$ is replaced by $(\mathbb{R}^n, d\gamma(x))$. Since the properties on the semigroup remain the same with respect to the Gaussian space, we obtain the following weak type representation

$$\langle \nabla(-\Delta_{OU})^{-1/2}f, g\rangle_{L^2(\mathbb{R}^n, d\gamma(x))} = -4\int_0^\infty \langle \nabla T_t f, \frac{d}{dt}T_t g\rangle_{L^2(\mathbb{R}^n, d\gamma(x))} t dt.$$

Probabilistic representation We may associate to the Ornstein Uhlenbeck operator Δ_{OU} a diffusion process (X_t) on $\mathbb{R}^n$ that satisfies

$$dX_t = dW_t - X_t dt,$$

where (W_t) is the Brownian motion on $\mathbb{R}^n$. This stochastic differential equation is solved by variation of parameters. Indeed, let $f(X_t, t) = X_t e^t$. Using Itô formula we get

$$\begin{aligned} df(X_t, t) &= X_t\, e^t\, dt + e^t\, dX_t \\ &= X_t e^t\, dt + e^t\,(dW_t - X_t dt) \\ &= e^t\, dW_t. \end{aligned}$$

Integrating from 0 to t we obtain

$$X_t = X_0 e^{-t} + e^{-t}\int_0^t e^s\, dW_s.$$

We claim that we have the following probabilistic representation (which we will prove in Chapter 5)

$$-\frac{1}{2}\nabla(-\Delta_{OU})^{-1/2}f(x) = \lim_{y\to\infty}\mathbb{E}_y\left[e^{-\tau}\int_0^\tau e^s \nabla e^{-B_s\sqrt{-\Delta_{OU}}}f(X_s)dB_s | X_\tau = x\right].$$

We see that the probabilistic representation of the Riesz vector associated with the Ornstein Uhlenbeck operator slightly differs from the Gundy Varopoulos representation in that new terms appeared. To explain this phenomenon, we do need some new ideas and arguments, mainly from differential geometry. For this reason, the next section will be devoted to define more general notions and as a consequence, define the Riesz transforms on Riemannian manifolds.

2.7 Riemannian geometry

In this section we will introduce some aspects of differential geometry. We aim to be as intuitive as possible in the understanding of the defined objects.

The class of spaces studied in what follows will be that of Riemannian manifolds. Intuitively speaking, a manifold is nothing more than a metric space (X, d) which locally looks like $\mathbb{R}^n$ with its usual Euclidean metric, for some integer n. Riemannian manifolds are differentiable manifolds (manifolds that allow differentiation and integration) with an extra bit of structure, a Riemannian metric, that allows to measure lengths and angles of tangent vectors.

Since Riemannian manifolds are differentiable manifolds, one can attach to every point $x \in X$ a tangent space denoted by T_xX, which is the set of all tangent vectors. Intuitively, tangent vectors are those vectors that are tangent to the surface of the manifold. The most common definition used for computations is to think of tangent vectors as directions in which we want to differentiate functions. For each $x \in X$, the derivation $D : C^\infty(X) \to \mathbb{R}$ is a linear operator that satisfies the Leibniz identity

$$D(fg) = f(x)D(g) + g(x)D(f). \tag{2.3}$$

Hence, T_xX is the set of all linear derivations at the point x.

We can assemble all these tangent spaces together and this will form the tangent bundle $TX = \bigcup_{x\in X} T_xX$. It is a smooth manifold. Smooth functions on TX are called vector fields on X. This means that we can get a vector field by attaching to each point of the manifold a tangent vector from the corresponding tangent space.
This construction of the tangent bundle can be done in a different manner. Instead of using tangent spaces, we consider their duals T_x^*X. Putting together all these spaces will form a bundle T^*X.
For a smooth function $f : X \to \mathbb{R}$, we define the smooth section df of T^*X by

$$df(x)(Y) := Y(f) \in \mathbb{R},$$

for all $Y \in T_xX$ thought of as a derivation in the sense (2.3). We refer the interested reader to [32, Theorem 1.51] for the definition of $Y(f)$. We call df the differential of f.

The reasons why we focus on Riemannian manifolds is that they are interesting for

- **Metric geometry.**
  Riemannian manifolds are metric spaces. This means that there exists a Riemannian metric which is a 2-tensor field g that is symmetric and positive semi-definite. On each tangent space T_xX, it determines an inner product $\langle Y, Z\rangle := g(Y, Z)$, for $Y, Z \in T_xX$.
  Metric geometry has another important feature which is geodesics. Intuitively geodesics are a generalization of the notion of straight lines realizing the shortest distance between two points on "non flat" manifolds. We refer to [50] for a more rigorous definition of geodesics. In this thesis, we use the fact that the geodesic distance from a point $x \in X$ is smooth, expect on the cut locus of the point x and the point x itself. In particular, we will use the Laplacian local comparison theorem away from x and $\mathrm{cut}(x)$ [81].

- **Calculus.**
  Riemannian manifolds are by definition differentiable manifolds. It means that the

usual notions of multivariable calculus on differentiable manifolds apply (derivatives, vector and tensor fields, integration of differential forms).
Following [51] and [50], let (X, g) be a complete Riemannian manifold. For each $x \in X$ we denote by T_xX and T^*_xX the tangent and the cotangent spaces at x, respectively. There is a canonical way of converting tangent vectors into cotangent vectors and vice versa. For instance, we define a map

$$\begin{array}{rcl} T^*_xX & \to & T_xX \\ w & \mapsto & \sharp w \end{array}$$

by requiring that $\langle \sharp w, Y \rangle = w(Y)$ where Y is an arbitrary vector and $\langle \cdot, \cdot \rangle$ is the inner product defined by g.
For every $j, k \in \mathbb{N}$, we set

$$T^{j,k}_xX := \underbrace{T_xX \otimes \cdots \otimes T_xX}_{j \text{ times}} \otimes \underbrace{T^*_xX \otimes \cdots \otimes T^*_xX}_{k \text{ times}}$$

and we denote by $T^{j,k}X$ the fibre bundle over X whose fibre at x is $T^{j,k}_xX$ i.e. $T^{j,k}X = \bigcup_{x\in X} T^{j,k}_xX$.
A tensor of type (j, k) is just a section of $T^{j,k}X$ i.e. a continuous map $\sigma : X \to T^{j,k}X$ such that $\pi(\sigma(x)) = x, \forall x \in X$, where $\pi : T^{j,k}X \to X$ is the projection map.
We denote the space of smooth tensors of type (j, k) by $C^\infty(T^{j,k}X)$. Functions on X are identified with tensors of type $(0, 0)$.
The Riemannian scalar product on T_xX induces a scalar product $\langle \cdot, \cdot \rangle_{T^{j,k}_xX}$ on $T^{j,k}_xX$. We set $| \cdot |^2_{T^{j,k}_xX} = \langle \cdot, \cdot \rangle_{T^{j,k}_xX}$.
Finally, Λ^kT^*X denotes the bundle of k-forms.

We call

$$d : C^\infty(\Lambda^kT^*X) \to C^\infty(\Lambda^{k+1}T^*X)$$

the exterior derivative. It satisfies the following properties

1. For $k = 0$, $d : C^\infty(X) \to \Lambda T^*X$ is the differential on functions defined above.
2. For $f \in C^\infty(X)$, $d(df) = 0$.
3. For α a p-form and β a q-form, we have

$$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^p \alpha \wedge d\beta,$$

   where $\wedge$ is the exterior product.

To sum up, the exterior derivative extends the notion of differential of a function to differential forms of higher degree.

We also define ∇ to be the covariant derivative, i.e. a family of linear connections

$$\begin{array}{rcccc} \nabla & : & C^\infty(TX) \times C^\infty(TX) & \to & C^\infty(TX) \\ & & (Y, Z) & \mapsto & \nabla_Y Z \end{array}$$

that satisfy

1. ∇ is linear over C^∞ in Y i.e., for all $f, g \in C^\infty$, we have
$$\nabla_{fY_1+gY_2} Z = f\nabla_{Y_1} Z + g\nabla_{Y_2} Z.$$
2. ∇ is linear over $\mathbb{R}$ in Z i.e., for all $a, a \in \mathbb{R}$, we have
$$\nabla_Y(aZ_1 + bZ_2) = a\nabla_Y Z_1 + b\nabla_Y Z_2.$$
3. For all $f \in C^\infty$, ∇ satisfies the Leibniz rule
$$\nabla_Y(fZ) = f\nabla_Y Z + (Yf)Z.$$

$\nabla_Y Z$ is then called the covariant derivative of Z in the direction Y. To sum up, the covariant derivative generalizes the directional derivative from vector calculus to tensor fields.

Remark 2.5. – *Connections are not uniquely defined on a manifold, since comparing tangent vectors attached to different points is a priori impossible. However, in the case of a Riemannian manifold, we can choose a unique connection with certain properties that we will call the Levi-Civita connection.*

– *We recall that on functions, d and ∇ coincide with the differential.*

- **Measure theory.**
Any oriented Riemannian manifold has a canonical measure given by the volume form μ defined such that $d\mu(x) = \sqrt{\det g(x)}dx$. It allows to integrate functions and to define L^p spaces on Riemannian manifolds.
For each $p \in [1, \infty]$ and $j, k \in \mathbb{N}$, let $L^p(T^{j,k}X, \mu)$ be the Banach space of all measurable tensors u of type (j, k) with
$$\|u\|_{L^p(T^{j,k}X,\mu)} = \begin{cases} \left(\int_X |u|^p_{T^{j,k}_x X} d\mu(x)\right)^{1/p}, & \text{if } p \in [1, \infty); \\ \operatorname{ess\,sup}_{x \in X} |u|_{T^{j,k}_x X}, & \text{if } p = \infty. \end{cases}$$
We drop the subscripts when there is no ambiguity.

- **Curvature.**
Riemannian manifolds are the most natural setting for studying the notion of curvature.
Let us first give an intuitive idea of how to visualize the curvature. Take $X = S^2$,

the 2-dimensional sphere in $\mathbb{R}^3$. It is obviously a manifold, and as discussed earlier, we can consider tangent vectors in TX as vectors in $\mathbb{R}^3$ which are attached to points in S^2 along some tangent lines. One way to see that the sphere is curved is to consider a tangent vector to the sphere at a point A and then transport it to another point B along a circle so that it keeps pointing in the same direction on the sphere. This is done by keeping constant the angle between the tangent vector and the direction in which it moves. We call this a parallel transport. In the same way, we can transport this tangent vector from the point B to another point C. Carrying it back to the point A, we end up with a different tangent vector. This phenomenon reflects the curvature of the sphere.
The figure below illustrates the curvature of the sphere.

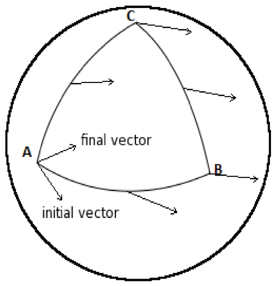

Figure 2.1: The parallel transport of a vector on a 2D sphere.

Somehow, the curvature provides one way of measuring the degree to which the geometry determined by a given Riemannian metric might differ from that of ordinary Euclidean n-space.
One of the most frequent object that appears when studying estimates of operators acting on manifolds is the Ricci curvature.
The Ricci curvature, denoted by Ric, plays an important role in the analysis of Markov semigroups through the Bochner-Lichnerowicz-Weitzenbock formula by connecting the Laplace operator to the Ricci curvature. See Section 2.7.1. It also appears very frequently when studying estimates of operators acting on manifolds.

For reasons coming from the study of hyper contractive diffusions processes, Bakry and Emery defined in [13] a generalization of the Ricci curvature by studying *carré du champ* and *carré du champ itéré*. The model example is the weighted Riemannian manifold $(X, g, e^{-\varphi} d\mu(x))$, where $\varphi : X \to \mathbb{R}$ is a smooth potential. In this case,

$$\mathrm{Ric}_\varphi = \mathrm{Ric} + \nabla^2 \varphi,$$

where Ric denotes the Ricci curvature tensor on X and $\nabla^2\varphi$ is the Hessian of φ. Moreover, the self-adjoint ($\varphi-$)Laplacian with respect to the weighted measure is

$$\Delta_\varphi f = \Delta f - \nabla f \cdot \nabla \varphi.$$

When φ is a constant function, we recover of course the Ricci curvature, making the Bakry Emery curvature an extension of the Ricci curvature.

Example 2.4. *Taking $X = \mathbb{R}^n$ and $\varphi(x) = \frac{\|x\|^2}{2}$, we have $d\mu(x) = d\gamma(x)$, the Gaussian measure. In this case, we have*

$$\mathrm{Ric}_{OU} = 1.$$

One can of course consider even more variants of heat semigroup (by considering Riemannian manifolds for example) in which case we need to adapt some structures. This will be the aim of Chapters 4 and 6.

In what follows, we are going to focus on manifolds with (Bakry-Emery) Ricci curvature bounded from below. This assumption on the curvature allows for example to obtain results on the behaviour of balls, by the Bishop-Gromov comparison inequality [81].

2.7.1 Laplacians

We conclude this section by defining different Laplace operators acting on manifolds.

We define the Laplace-Beltrami operator acting on functions as follows

$$\Delta f = -d^* df,$$

where d^* is the adjoint of d on $L^2(X)$.

Remark 2.6. *We use a sign convention that differs from the one in the literature on differential geometry. The main reason is to coincide with the Laplacian on $\mathbb{R}^n$.*

Alternatively, the operator can be generalized to operate on 1-forms, using the exterior derivatives and their adjoints d^* on L^2. The Hodge-de Rham Laplacian acting on 1-forms $\vec{g}$ is then defined by

$$\vec{\Delta}\vec{g} = dd^*\vec{g} + d^*d\vec{g}.$$

Note that the Laplace Beltrami operator is non positive while the Hodge-de Rham operator is positive. Moreover, these two operators are related as follows

$$\begin{aligned} \vec{\Delta} d &= (dd^* + d^*d)d \\ &= d(d^*d) + 0 \\ &= d(-\Delta). \end{aligned}$$

More generally, if we consider a weighted manifold $(X, g, e^{-\varphi}d\mu(x))$, it is natural to change the exterior derivative to make sure it is still adjoint. To this end we define

$$d^*_\varphi = d^* + i_{\nabla\varphi},$$

where $i_{\nabla\varphi}$ denotes the inner multiplication by $\nabla\varphi$ on Λ^1. Indeed, we calculate

$$\begin{aligned}\int_X \langle df, g\rangle e^{-\varphi} d\mu(x) &= \int_X \langle df, e^{-\varphi} g\rangle d\mu(x)\\ &= \int_X \langle f, d^*(e^{-\varphi} g)\rangle d\mu(x)\\ &= \int_X \langle f, e^{\varphi} d^*(e^{-\varphi} g)\rangle e^{-\varphi} d\mu(x)\\ &= \int_X \langle f, d^*_\varphi g\rangle e^{-\varphi} d\mu(x),\end{aligned}$$

then by identification,

$$\begin{aligned}d^*_\varphi g &= e^{\varphi} d^*(e^{-\varphi} g))\rangle\\ &\overset{(1)}{=} e^{\varphi}(e^{-\varphi} d^* g - \langle g, de^{-\varphi}\rangle)\\ &= d^* g + \langle g, \nabla\varphi\rangle,\end{aligned}$$

yielding to the result. In (1), we have used the following equality for a function k and a 1-form df

$$d^*(kdf) = kd^*df - \langle df, dk\rangle.$$

Indeed,

$$\begin{aligned}\langle kd^*df, h\rangle &= \langle d^*df, kh\rangle\\ &= \langle df, d(kh)\rangle\\ &= \langle df, kdh + hdk\rangle\\ &= \langle\langle df, dk\rangle, h\rangle + \langle d^*(kdf), h\rangle.\end{aligned}$$

We also have the following relation

$$\begin{aligned}\vec{\Delta}_\varphi d &= (dd^*_\varphi + d^*_\varphi d)d\\ &= d(d^*_\varphi d)\\ &= d(d^* + i_{\nabla\varphi})d\\ &= d(d^*d + \langle\nabla\varphi, d\rangle)\\ &= d(-\Delta + \langle\nabla\varphi, d\rangle)\\ &= d(-\Delta_\varphi).\end{aligned} \tag{2.4}$$

Finally, let ∇ be the Levi-Civita connection on X. Define

$$\underline{\Delta} = \mathrm{Trace}\nabla^2 \tag{2.5}$$

to be the rough Laplacian acting on 1-forms as well where

$$\nabla^2_{X,Y} T = \nabla_X \nabla_Y T - \nabla_{\nabla_X Y} T.$$

When there is no confusion, we will denote by Δ either the Laplace-Beltrami (acting on functions), or the rough Laplacian (acting on 1-forms). By the Bochner-Weitzenböck formula we have

$$\vec{\Delta}\vec{g} = -\underline{\Delta}\vec{g} + \mathrm{Ric}(\cdot, \sharp\vec{g}),$$

where Ric is the curvature of the manifold.
More generally, the Hodge-de Rham operator is a second order differential operator acting on k-forms defined by

$$\vec{\Delta} = d^{k-1}d^{*,k-1} + d^{*,k}d^k.$$

When there is no ambiguity about the order of the forms we are working with, we simply write $\vec{\Delta} = dd^* + d^*d$. The Bochner-Weitzenböck formula remains valid as well for higher degree forms by replacing the Ricci curvature by the Weitzenböck curvature.

2.8 Methods and results

2.8.1 First result

The Bellman function method was originally used in control theory by Richard E. Bellman. It was later introduced to harmonic analysis by Burkholder in 1984 to obtain sharp inequalities for martingale transforms, with a variation of the method. It reappeared in the nineties with the help of Nazarov, Treil and Volberg to prove/reprove many results in harmonic analysis. This method turns out to be an extremely powerful tool and a very natural way to deal with weighted inequalities and to find sharp dependence of the norm of some classical operators in harmonic analysis on weighted L^p spaces on the A_p characteristic of the weight.
The biggest challenge of this method is to find a suitable function satisfying all desired properties and then use some convexity arguments. Although uniqueness is not required, finding these functions demands a lot of practice. The authors themselves describe this method as a craftsmanship. Each problem has its own Bellman function, depending on a number of variables that changes from one case to another.

The first result of this thesis concerns the Riesz vector $\mathcal{R}_\varphi$ on Riemannian manifolds (X, g, μ) defined respecting the measures of the type $e^{-\varphi}d\mu$ where an additional weight is present: in weighted spaces $L^2(\omega) = L^2(\omega e^{-\varphi}d\mu)$ we study the operator norm of the Riesz vector.

By exhibiting a suitable Bellman function whose origin come from an analysis of the paper [61], we prove that on a complete Riemannian manifold (X, g, μ) endowed with measure $d\mu_\varphi = e^{-\varphi}d\mu$ such that the Bakry-Emery curvature is non-negative and $\mu_\varphi(X) < \infty$, we have a dimension-free linear weighted norm estimates for the arising Riesz vector in terms of the Poisson flow A_2 characteristic of the weight:

$$\|\mathcal{R}_\varphi\|_{L^2(\omega)\to L^2(\omega)} \lesssim \widetilde{Q}_2(\omega).$$

Such estimates are known to be in sharp dependence on the power of the Poisson characteristic even when $X = \mathbb{T}$ and $\varphi = 0$ [70].

For more on this work, we refer to [18].

2.8.2 Second result

Our approach for the second result is a bit more probabilistic and uses a martingale representation of the Riesz transform on complete Riemannian manifolds, first presented in [54].
It has been very profitable to deal with the stochastic differential theory, since it is in close connection with harmonic analysis and allows to obtain sharp results of L^p bounds for various important operators.

Our result is of strong nature because it estimates Z^*, the maximal function of Z. The technique used in the proof is called the sparse domination. It is quite recent and it was first due to Nazarov and Lerner in [52] and Lacey in [49]. It is a very powerful proof of the A_2 conjecture [49]. The topic has been very active lately. See for example [14] where sparse domination is extended beyond Calderón-Zygmund theory, for non integral operators[3] and [21] for a more probabilistic approach.

The sparse domination is a recent technique developed by Nazarov and Lerner in [52] and Lacey in 2015 [49]. For an operator T and suitable f, the purpose is to establish pointwise control of Tf by a sparse operator S i.e $|Tf| \lesssim S|f|$ then use the fact that sparseness property allows to insert weights and recover the best power for the A_p constant. Although Lacey's sparse domination immediately implies weighted inequalities with sharp dependence upon the A_p characteristic of the weight, it is defined on cubes and cannot provide dimensionless estimates, nore satisfactory results on non-doubling spaces. We bypass this problem by using a sparse operator with continuous stopping times, as in [21].
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $X = (X_t)_t$ a stochastic process. We say that the operator $X \mapsto S(X)$ is called sparse if there exists an increasing sequence of adapted stopping times $0 = T^{-1} \leq T^0 \leq \cdots$ with nested sets $E_j = \{T^j < \infty\}$, $E_j \subset E_{j-1}$ so that

$$S(X) = \sum_{j=-1}^{\infty} X_{T^j} \chi_{E_j} \text{ where } X_{T^j} = \mathbb{E}(X|\mathcal{F}_{T^j});$$

$$\forall A_j \subset E_j,\ A_j \in \mathcal{F}_{T_j} \text{ there holds } \mathbb{P}(A_j \cap E_{j+1}) \leq \frac{1}{2}\mathbb{P}(A_j).$$

The great advantages of the recursive proof is that it only relies on some weak-L^1 estimate and that homogeneity is not needed.

[3]For more on this theory, we recommend Auscher and Martell's paper [4].

By writing the Riesz transform as the conditional expectation of a stochastic process and using Wang's function [80], we show that for any $p \in (1, \infty)$:

$$\|\mathcal{R}_\varphi\|_{L^p \to L^p} \leq 16 \frac{p^2}{p-1}$$

and

$$\|\mathcal{R}_\varphi\|_{L^2(\omega) \to L^2(\omega)} \lesssim \widetilde{Q}_2(\omega).$$

For more on this work, we refer to [19].

2.8.3 Third result

Using the same probabilistic approach as in the second result, we focus this time on Bakry-Emery curvature, following [16] and considering a probabilistic approach. Random walks and Poisson flows on manifolds are a delicate matter and we refer the reader to the excellent text by Emery [29].

Assuming that $\mathrm{Ric}_\varphi \geq -a,\ a \geq 0$, we define the Poisson flow and Riesz transforms accordingly. The corner stone proof of this result is an elegant replacement of the martingale (X_t) by a sub-martingale (X_t^a), which is the sum of (X_t) and an increasing finite variation process. By Itô formula, (X_t^a) is nothing but $Q^a f$, the Poisson extension of f. This trick allows us to tackle the less forgiving negative curvature part.

Another change that comes with the introduction of the sub-martingale is the definition of the sparse operator S. Indeed, we define

$$S(X^a) = \sum_{j=-1}^{\infty} \mathbb{E}(X^a | \mathcal{F}_{T^j}) \chi_{E_j} \quad \text{instead of} \quad S(X^a) = \sum_{j=-1}^{\infty} X^a_{T^j} \chi_{E_j}$$

because unlike martingales, we only have the following inequality

$$\mathbb{E}(X^a | \mathcal{F}_{T^j}) \geq X^a_{T^j}.$$

Chapter 3

Bellman functions and their applications on martingale transform

In this chapter, we will construct Bellman functions following Wittwer's work in [82] to prove that the weighted L^2 bound of the dyadic martingale transforms is linear in terms of the characteristic of the weight. To this end, we first introduce the Haar basis in L^2 and its equivalent in $L^2(w)$, then we will present some intermediate results which in turn will allow us to prove Wittwer's result. Finally, an analogy will be made with the Bakry-Riesz vector which will be studied in the next chapter.

3.1 Haar basis

Let $\mathcal{D}$ denote the collection of all dyadic intervals in $\mathbb{R}$, that is the collection

$$\mathcal{D} = \{[n2^k, (n+1)2^k) : n, k \in \mathbb{Z}\}.$$

Each dyadic grid gives rise to an orthonormal system in L^2 called the Haar system $\{h_I, I \in \mathcal{D}\}$ defined by

$$h_I = \frac{\chi_{I_l} - \chi_{I_r}}{\sqrt{|I|}},$$

where I_l denotes the left half of I and I_r the right one.
Note that a function $f \in L^2$ can be expanded as follows

$$f = \sum_{I \in \mathcal{D}} (f, h_I) h_I,$$

where $(\cdot \, , \, \cdot)$ is the standard inner product in L^2 and we have

$$\|f\|_2^2 = \sum_{I \in \mathcal{D}} |(f, h_I)|^2.$$

It is often more convenient to work with a basis in $L^2(w)$ that is orthonormal. For this reason, instead of working with the classical Haar basis, we define the disbalanced Haar basis.

Lemma 3.1 (Disbalanced Haar functions)**.** *There exist constants x_I and A_I such that*

$$x_I h_I = h_I^w + A_I \chi_I$$

and (h_I^w) satisfies

1) h_I^w vanishes outside I is constant on I_l and I_r,

2) $\int h_I^w w dx = 0$,

3) $\|h_I^w\|_{L^2(w)} = 1$.

This means that $(h_I^w)_{I\in\mathcal{D}}$ is orthonormal in $L^2(w)$ and we may calculate explicitly the constants. Indeed from 2) we obtain

$$\begin{aligned}\int (x_I h_I - A_I \chi_I) w dx = 0 \quad &\Leftrightarrow \quad x_I \int h_I w dx = A_I \int_I w \\ &\Leftrightarrow \quad x_I \frac{\sqrt{|I|}}{2} (\langle w \rangle_{I_l} - \langle w \rangle_{I_r}) = A_I |I| \langle w \rangle_I .\end{aligned}$$

Hence,

$$A_I = \frac{x_I}{2\sqrt{|I|}} \frac{\langle w \rangle_{I_l} - \langle w \rangle_{I_r}}{\langle w \rangle_I}.$$

From 3) we obtain

$$\begin{aligned}\int (h_I^w)^2 w dx = 1 \quad &\Leftrightarrow \quad \int \Big((x_I h_I)^2 + A_I^2 \chi_I - 2 x_I A_I h_I \chi_I \Big) w dx = 1 \\ &\Leftrightarrow \quad x_I^2 \left(\frac{\langle w \rangle_{I_l} + \langle w \rangle_{I_r}}{2} - \frac{(\langle w \rangle_{I_l} - \langle w \rangle_{I_r})^2}{4 \langle w \rangle_I} \right) = 1 \\ &\Leftrightarrow \quad x_I^2 \left(\frac{4\langle w \rangle_I^2 - (\langle w \rangle_{I_l} - \langle w \rangle_{I_r})^2}{4 \langle w \rangle_I} \right) = 1 \\ &\Leftrightarrow \quad x_I^2 \left(\frac{(2\langle w \rangle_I - (\langle w \rangle_{I_l} - \langle w \rangle_{I_r}))(2\langle w \rangle_I + (\langle w \rangle_{I_l} - \langle w \rangle_{I_r}))}{4 \langle w \rangle_I} \right) = 1.\end{aligned}$$

Using the fact that $2\langle w \rangle_I = \langle w \rangle_{I_l} + \langle w \rangle_{I_r}$, we obtain

$$x_I^2 \left(\frac{\langle w \rangle_{I_l} \langle w \rangle_{I_r}}{\langle w \rangle_I} \right) = 1.$$

Hence,

$$x_I = \sqrt{\frac{\langle w \rangle_I}{\langle w \rangle_{I_l} \langle w \rangle_{I_r}}},$$

where $\langle f \rangle_I$ denotes the average value of f on $I \in \mathcal{D}$.

3.2 Useful results

In this section, we present several results that will be later used to prove Wittwer's result [82] regarding dyadic martingale transforms. All of the proofs will rely on the Bellman functions technique. We first start by presenting Theorem 3.1, which is the weighted version of the Carleson embedding theorem. Theorems 3.2 and 3.3 are respectively the first version and the "*easier to apply*" version of the weighted bilinear Carleson embedding theorem. Finally, Lemmata 3.2, 3.3 and 3.4 are different estimates concerning weights. We recall that the main goal of this section is to reduce the study of weighted L^2 norm of the dyadic martingale transforms to the study of four sums, using the results below. We hope that the use of the Bellman functions technique in every proof will convince the readers of its importance and elegance.

Theorem 3.1 (The weighted Carleson embedding theorem). *Let $\{\alpha_I\}_{I\in\mathcal{D}}$ be a sequence of non-negative numbers. If for all $J \in \mathcal{D}$*

$$\frac{1}{|J|}\sum_{I\subseteq J}\langle w\rangle_I^2\alpha_I \le C\langle w\rangle_J,$$

then for every $f \in L^2$

$$\sum_{I\in\mathcal{D}}\langle fw^{1/2}\rangle_I^2\alpha_I \le 4C\|f\|_2^2.$$

Moreover, the constant 4 is sharp.

Proof. We may assume that f is non-negative, since otherwise we can split it into the positive and the negative part. We also assume without loss of generality that $C = 1$. Observe that it is enough to prove that for every J fixed in $\mathcal{D}$,

$$\frac{1}{|J|}\sum_{I\subseteq J}\langle fw^{1/2}\rangle_I^2\alpha_I \le 4\langle f^2\rangle_J \tag{3.1}$$

because the original inequality can be obtained in the limit by using the monotone convergence theorem.
The left hand side of the inequality can be rewritten as

$$\frac{1}{|J|}\langle fw^{1/2}\rangle_J^2\alpha_J + \frac{1}{2|J_l|}\sum_{I\subseteq J_l}\langle fw^{1/2}\rangle_I^2\alpha_I + \frac{1}{2|J_r|}\sum_{I\subseteq J_r}\langle fw^{1/2}\rangle_I^2\alpha_I, \tag{3.2}$$

where J_l and J_r denote the left half and the right half of the interval J.
We see from (3.2) and the Carleson condition that we need to define the following variables

$$\langle fw^{1/2}\rangle_J = x,\quad \langle f^2\rangle_J = X,\quad \langle w\rangle_J = w,\quad \frac{1}{|J|}\sum_{I\subseteq J}\langle w\rangle_I^2\alpha_I = M$$

which naturally restricts us to the following domain

$$D = \{(X, x, w, M) : x^2 \le Xw; 0 \le M \le w; X, x, w \ge 0\}.$$

We also consider the abstract Bellman function

$$B(X,x,w,M)=\sup_{f,w,\alpha}\frac{1}{|J|}\sum_{I\subseteq J}\langle fw^{1/2}\rangle_I^2\alpha_I$$

where the supremum is taken over all f, w and α satisfying

$$\langle fw^{1/2}\rangle_J=x,\quad \langle f^2\rangle_J=X,\quad \langle w\rangle_J=w,\quad \frac{1}{|J|}\sum_{I\subseteq J}\langle w\rangle_I^2\alpha_I=M.$$

Note that the function B does not depend on the choice of the interval J.
The function B is such that $0\le B\le 4X$, which comes from the fact that we assumed (3.1) to be true. Moreover, by (3.2) we have

$$B(X,x,w,M)-\frac{1}{2}\left(B(X_l,x_l,w_l,M_l)+B(X_r,x_r,w_r,M_r)\right)\ge\frac{1}{|J|}x^2\alpha_J, \tag{3.3}$$

whenever (X,x,w,M), (X_l,x_l,w_l,M_l) and (X_r,x_r,w_r,M_r) belong to the domain D and satisfy $X=\dfrac{X_l+X_r}{2}$, $x=\dfrac{x_l+x_r}{2}$, $w=\dfrac{w_l+w_r}{2}$ and $M-\dfrac{M_l+M_r}{2}=h=\dfrac{1}{|J|}w_J^2\alpha_J$.
The inequality (rather than equality) comes from the fact that fixing averages separately on J_l and J_r leads to a smaller set of functions than if fixing the average on J.
Condition (3.3) means in particular that B is concave. Moreover, if we take $X=X_l=X_r$, $x=x_l=x_r$, $w=w_l=w_r$ and $M-h=M_l=M_r$, then we obtain

$$\begin{aligned}B(X,x,w,M)-B(X,x,w,M-h)&\ge\frac{1}{|J|}x^2\alpha_J\\&=h\frac{x^2}{w_J^2},\end{aligned}$$

meaning that $\dfrac{\partial B}{\partial M}\ge\dfrac{x^2}{w^2}$. It turns out that these two infinitesimal conditions also imply (3.3). Indeed,

$$\begin{aligned}h\frac{x^2}{w_J^2}&\le B(X,x,w,M)-B(X,x,w,M-h)\\&\le B(X,x,w,M)-\frac{1}{2}\left(B(X_l,x_l,w_l,M_l)+B(X_r,x_r,w_r,M_r)\right),\end{aligned}$$

where the first inequality comes from $\dfrac{\partial B}{\partial M}\ge\dfrac{x^2}{w^2}$ and the second from the concavity of B.
Therefore, we conclude that B satisfies on D

1. $0\le B\le 4X$;
2. $\dfrac{\partial B}{\partial M}\ge\dfrac{x^2}{w^2}$;

3. $-d^2B \geq 0$.

Conversely, and this is the most important part of the proof, assuming we have a function B satisfying the previous properties, we may prove (3.1) by applying (3.3) n times. We obtain

$$\begin{aligned}
4\int_J f^2 &= 4\langle f^2\rangle_J |J| \\
&\geq |J| B\left(\langle f^2\rangle_J, \langle f w^{1/2}\rangle_J, \langle w\rangle_J, \frac{1}{|J|}\sum_{I\subseteq J}\langle w\rangle_I^2 \alpha_I\right) \\
&\geq \sum_{I\subset J, |I|=2^{-n}|J|} |I| B\left(\langle f^2\rangle_I, \langle f w^{1/2}\rangle_I, \langle w\rangle_I, \frac{1}{|I|}\sum_{K\subset I}\langle w\rangle_K^2 \alpha_K\right) \\
&\quad + \sum_{I\subseteq J, |I|>2^{-n}|J|} \langle f w^{1/2}\rangle_I^2 \alpha_I \\
&\geq \sum_{I\subseteq J, |I|>2^{-n}|J|} \langle f w^{1/2}\rangle_I^2 \alpha_I.
\end{aligned}$$

We obtain the final result by applying the above estimates on intervals $[-2^n, 0]$ and $[0, 2^n]$ and passing to the limit when $n \to +\infty$.

Let us now try to find a function B that satisfies the estimates above.
We may suppose that B has the form

$$B(X, x, w, M) = 4\left(X - \alpha(x, w, M)\right),$$

where

1. $|\alpha(x, w, M)| \leq x^2 w^{-1}$;

2. α is convex;

3. $\dfrac{\partial \alpha}{\partial M} \leq -\dfrac{x^2}{4w^2}$.

The first property comes from the fact that $x^2 \leq Xw$, the second one because B is concave and the last one because $\dfrac{\partial B}{\partial M} \geq \dfrac{x^2}{w^2}$ and we have a factor 4 in the construction of B. The first and third properties suggest to try α of the form

$$\alpha(x, w, M) = \gamma(w, M)x^2,$$

where

1. $|\gamma(w, M)| \leq w^{-1}$;

2. $\dfrac{\partial \gamma}{\partial M} \leq -\dfrac{1}{4w^2}$.

We also recall that $M \leq w$. Observe that $\gamma(w, M) = \frac{1}{w+M}$ satisfies the hypotheses. In conclusion, one possible choice of B is

$$B(X, x, w, M) = 4\left(X - \frac{x^2}{w+M}\right).$$

$\square$

Remark 3.1.
- *It is reasonable to think of M as a concave function of w in the sense*

$$M(w_J) - \frac{M(w_{J_l}) + M(w_{J_r})}{2} \geq \frac{1}{|J|}\langle w\rangle_J^2 \alpha_J.$$

Then we can consider γ as a function of w and $M(w)$. This observation will be of great help in our future investigations on the origin of the function used to prove the weighted boundedness of the Riesz transform on manifolds.

- *In the unweighted Carleson embedding theorem, the function B is*

$$B(X, x, M) = 4\left(X - \frac{x^2}{1+M}\right),$$

where M refers to the unweighted Carleson condition

$$M = \frac{1}{|J|}\sum_{I \subseteq J} \alpha_I \leq 1.$$

- *The sharpness of the constant 4 was proved in [62, Theorem 3.3]*

Theorem 3.2 (The bilinear weighted Carleson embedding theorem, [82, 61])**.** *Let $\{\alpha_I\}_{I\in\mathcal{D}}$ be a sequence of non-negative numbers. If for all $J \in \mathcal{D}$*

$$\frac{1}{|J|}\int_J \left(\sum_{I\subseteq J} \alpha_I \langle w\rangle_I \chi_I(x)\right)^2 v dx \leq \langle w\rangle_J$$

and

$$\frac{1}{|J|}\int_J \left(\sum_{I\subseteq J} \alpha_I \langle v\rangle_I \chi_I(x)\right)^2 w dx \leq \langle v\rangle_J$$

then for all $f, g \in L^2$

$$\sum_{I\in\mathcal{D}} \langle f w^{1/2}\rangle_I \langle g v^{1/2}\rangle_I \alpha_I |I| \leq 24\|f\|_2\|g\|_2. \tag{3.4}$$

Proof. We prove the bilinear weighted Carleson embedding in the same spirit as we did it for Theorem 3.1.

First step: we define the abstract Bellman function. Indeed, for a fixed interval $J \in \mathcal{D}$, let us consider the variables

$$X = \langle f^2 \rangle_J, \; Y = \langle g^2 \rangle_J, \; x = \langle f w^{1/2} \rangle_J, \; y = \langle g v^{1/2} \rangle_J, \; r = \langle w \rangle_J, \; s = \langle v \rangle_J,$$

$$M = \frac{1}{|J|} \int_J \left(\sum_{I \subseteq J} \alpha_I \langle w \rangle_I \chi_I(x) \right)^2 v dx, \; N = \frac{1}{|J|} \int_J \left(\sum_{I \subseteq J} \alpha_I \langle v \rangle_I \chi_I(x) \right)^2 w dx.$$

We also define the abstract Bellman function

$$B(X, Y, x, y, r, s, M, N) = \sup \frac{1}{|J|} \sum_{I \subseteq J} \langle f w^{1/2} \rangle_I \langle g v^{1/2} \rangle_I \alpha_I |I|,$$

where the supremum is taken over all f, g, w, v, α non negative with averages fixed as above and such that the Carleson conditions are satisfied.
Observe that we have

$$M_J - \frac{M_{J_l} + M_{J_r}}{2} = \left(\langle w \rangle_J \alpha_J \right)^2 \langle v \rangle_J = \langle w \rangle_J \alpha_J K_J$$

and

$$N_J - \frac{N_{J_l} + N_{J_r}}{2} = \left(\langle v \rangle_J \alpha_J \right)^2 \langle w \rangle_J = \langle v \rangle_J \alpha_J K_J,$$

where

$$K := \frac{1}{|J|} \sum_{I \subseteq J} \langle w \rangle_I \langle v \rangle_I \alpha_I |I|.$$

This naturally leads us to consider a new variable K which satisfies

$$K_J - \frac{K_{J_l} + K_{J_r}}{2} = \langle w \rangle_J \langle v \rangle_J \alpha_J.$$

Considering the nine variables, this restricts us to the domain

$$D = \{(X, Y, x, y, r, s, M, N, K) : x^2 \leq X r, y^2 \leq Y s, 1 \leq rs \leq Q, \\ 0 \leq M \leq r, 0 \leq N \leq s, 0 \leq K \leq \sqrt{rs}\}.$$

The only non-trivial inequality to prove is $K \leq \sqrt{rs}$. In fact, by Cauchy Schwarz inequality, $K^2 \leq Ms$ (and likewise $K^2 \leq Nr$). Use the fact that $M \leq r$ by assumption on M (or $N \leq s$) to conclude.
Rather than proving the inequality (3.4) as it is stated, we will prove that

$$\sum_{I \in \mathcal{D}} \langle f w^{1/2} \rangle_I \langle g v^{1/2} \rangle_I \alpha_I |I| \leq 24(\|f\|_2^2 + \|g\|_2^2).$$

We get back to the original inequality by replacing f by λf and g by $\frac{g}{\lambda}$ and then minimize in λ.
If we denote the 9-tuple $(X, Y, x, y, r, s, M, N, K)$ by a, the function B satisfies

$$0 \leq B(a) \leq C(X + Y),$$

$$B(a)-\frac{1}{2}\left(B(a_l)+B(a_r)\right)\geq xy\alpha_J. \tag{3.5}$$

Conversely, if we have a function B that satisfies the previous properties, then

$$\begin{aligned}
|J|\left(\langle f^2\rangle_J+\langle g^2\rangle_J\right) &\geq |J|B\left(\langle f^2\rangle_J,\cdots,\frac{1}{|J|}\sum_{I\subseteq J}\langle w\rangle_I\langle v\rangle_I\alpha_I|I|\right)\\
&\geq \sum_{I\subset J,|I|=2^{-n}|J|}|I|B\left(\langle f^2\rangle_I,\cdots,\frac{1}{|I|}\sum_{K\subseteq I}\langle w\rangle_K\langle v\rangle_K\alpha_K|K|\right)\\
&\quad+\sum_{I\subset J,|I|>2^{-n}|J|}\langle fw^{1/2}\rangle_I\langle gv^{1/2}\rangle_I\alpha_I|I|\\
&\geq \sum_{I\subset J,|I|>2^{-n}|J|}\langle fw^{1/2}\rangle_I\langle gv^{1/2}\rangle_I\alpha_I|I|
\end{aligned}$$

and the result is proved by using the size property of B and letting $n\to\infty$.

Second step: we construct a concrete Bellman function B that satisfies our problem. To have (3.5), it is sufficient to have

$$B(a)-\frac{1}{2}\left(B(a_l)+B(a_r)\right)\gtrsim\frac{xy}{rs}\left(K_J-\frac{K_{J_l}+K_{J_r}}{2}\right), \tag{3.6}$$

for $K\lesssim\dfrac{yr}{x}$ and $K\lesssim\dfrac{xs}{y}$,

$$B(a)-\frac{1}{2}\left(B(a_l)+B(a_r)\right)\gtrsim\frac{x^2}{r^2}\left(M_J-\frac{M_{J_l}+M_{J_r}}{2}\right), \tag{3.7}$$

for $K\gtrsim\dfrac{yr}{x}$, and

$$B(a)-\frac{1}{2}\left(B(a_l)+B(a_r)\right)\gtrsim\frac{y^2}{s^2}\left(N_J-\frac{N_{J_l}+N_{J_r}}{2}\right), \tag{3.8}$$

for $K\gtrsim\dfrac{xs}{y}$.

Indeed, by using the fact that $K_J-\frac{K_{J_l}+K_{J_r}}{2}\geq rs\alpha_J$, the result is immediate. We do the same work with the remaining inequalities, and use the size restrictions on K.

We have now reduced our problem to finding a function B that satisfies (3.6)-(3.8). We will be looking for a function B of the form

$$B(a)=B_1(b)+B_2(c)+B_3(d),$$

where $b=(X,Y,x,y,r,s,M)$, $c=(X,Y,x,y,r,s,N)$ and $d=(X,Y,x,y,r,s,K)$. We previously saw that we need to divide the domain D into 3 sub-domains $R_1=\{yr\geq$

$xK\}\bigcap\{xs \geq yK\}$, $R_2 = \{xs \leq yK\}$ and $R_3 = \{yr \leq xK\}$. We define on the whole domain D the following function

$$B_3(d) = X + Y - \sup_{t>0}\left(\frac{x^2}{r + tK(r,s)} + \frac{y^2}{s + t^{-1}K(r,s)}\right).$$

Let us show that this function satisfies (3.6). Indeed, B_3 is concave as the supremum of concave functions. Moreover, by differentiating B_3 with respect to t, we obtain

$$\frac{\partial B_3}{\partial t} = -\frac{x^2 K}{(r+tK)^2} + \frac{y^2 K}{(ts+K)^2},$$

which yields to

$$\frac{\partial B_3}{\partial t} = 0 \Leftrightarrow t = \frac{ry - Kx}{sx - Ky}.$$

Thus the supremum is attained at $t_m := \dfrac{yr - xK}{xs - yK}$. When both the denominator and numerator of t_m are positive, K is small and we are in R_1. Now because we would like to prove (3.6), one can apply chain rule and show that $\frac{\partial B_3}{\partial K} \geq \frac{xy}{2rs}$ on a smaller region, namely $\{yr \geq 4xK\}\bigcap\{xs \geq 4yK\}$, at the expense of enlarging R_2 and R_3. Another method was presented in [22] where the region was $R_1' = \{yr \geq 2xK\}\bigcap\{xs \geq 2yK\}$, allowing a better numerical constant. We add it here for the sake of completeness.
By omitting the variables X and Y and when $t = t_m$,

$$\begin{aligned} B_3(d) &= -\frac{x^2(sx - Ky)}{r(sx-Ky) + (ry - Kx)K} \\ &\quad -\frac{y^2(ry - Kx)}{s(ry - Kx) + (sx - Ky)K} \\ &= -\frac{sx^2 - 2Kxy + ry^2}{sr - K^2}. \end{aligned}$$

Recall that we are only interested in K's that are in R_1'. We have

$$\frac{\partial B_3}{\partial K} = 2\frac{(xs - Ky)(yr - Kx)}{(sr - K^2)^2}.$$

We claim that on R_1'

$$\frac{\partial B_3}{\partial K} \geq \frac{xy}{2rs}.$$

Indeed, by multiplying both numerator and denominator of $\dfrac{\partial B_3}{\partial K}$ by xy and then expand we obtain

$$\frac{\partial B_3}{\partial K} = 2\frac{(xs - Ky)(yr - Kx)xy}{[r(xs - Ky) + K(yr - Kx)]\,[s(yr - Kx) + K(xs - Ky)]}.$$

We need to verify that

$$\begin{aligned} 4rs &\geq \frac{[r(xs-Ky)+K(yr-Kx)]\,[s(yr-Kx)+K(xs-Ky)]}{(xs-Ky)(yr-Kx)} \\ &= K^2+sr+Ks\frac{yr-Kx}{xs-Ky}+Kr\frac{xs-Ky}{yr-Kx}. \end{aligned}$$

We know from previous calculations that $K^2 \leq rs$ (by fixing the constant $c_1 = 1$ in the hypothesis of the bilinear embedding theorem). Moreover on R_1', $xs - yK \geq yK$ and $yr - xK \geq xK$, which leads to the result. In fact, the constant 2 appearing in the definition of R_1' can be replaced by any $1+\epsilon$, $\epsilon > 0$. In this case we would have had $\dfrac{\partial B_3}{\partial K} \geq \dfrac{xy}{(1+\frac{1}{\epsilon})rs}$.
Therefore,

$$B_3(d) - \frac{1}{2}\left(B_3(d_l) + B_3(d_r)\right) \geq \frac{xy}{2rs}\left(K_J - \frac{K_{J_l}+K_{J_r}}{2}\right)$$

Now for the functions B_1 and B_2. Note that they are defined to deal with the "super concavity" for K's that are not in R_1'. We use the same idea as in the proof of Theorem 3.1. We define

$$B_1(b) = X - \frac{x^2}{r+M} + Y - \frac{y^2}{s}$$

and claim that it satisfies (3.7). Indeed, we have

$$-d^2 B_1 \geq 0$$

everywhere on D and

$$\frac{\partial B_1}{\partial M} \geq \frac{x^2}{4r^2}.$$

Therefore

$$\begin{aligned} B_1(b) - \frac{1}{2}\left(B_1(b_l) + B_1(b_r)\right) &\geq \frac{\partial B_1}{\partial M}\left(M_J - \frac{M_{J_l}+M_{J_r}}{2}\right) \\ &\geq \frac{x^2}{4r^2}\left(M_J - \frac{M_{J_l}+M_{J_r}}{2}\right). \end{aligned}$$

Recall that $M_J - \frac{M_{J_l}+M_{J_r}}{2} = r\alpha_J K_J$ and for $K \geq \dfrac{yr}{2x}$, the right hand side is bigger than $\dfrac{xy\alpha_J}{8}$.
Analogously, let

$$B_2(c) = X - \frac{x^2}{r} + Y - \frac{y^2}{s+N}.$$

We prove likewise that it satisfies (3.8) and hence for $K \geq \dfrac{xs}{2y}$ we have again

$$-d^2 B_2 \geq \frac{xy\alpha_J}{8}.$$

This means that whether K is in R'_1 or not and thanks to the global concavity of B_1, B_2 and B_3, we always have

$$-d^2B \geq \frac{1}{8}xy\alpha_J.$$

Consequently, the Bilinear weighted Carleson embedding theorem is proved. □

Remark 3.2. *The concrete function B is actually C^1 and piecewise C^2 [22]. Hence the last inequality in the proof actually means that that on each domain where B is C^2 we have*

$$-d^2B \geq \frac{1}{8}xy\alpha_J.$$

This theorem is historically the first version of the Bilinear Carleson embedding theorem. None of the conditions above are necessary and the theorem is valid for two-weighted problems. We present now a version with simpler assumptions.

Theorem 3.3 ([65])**.** *Let $\{\alpha_I\}_{I\in\mathcal{D}}$ be a sequence of non-negative numbers, w and w^{-1} weights such that $1 \leq \langle w\rangle_I\langle w^{-1}\rangle_I \leq Q$. If for all $J \in \mathcal{D}$*

1. $\frac{1}{|J|}\sum_{I\subseteq J}\frac{\alpha_I}{\langle w\rangle_I} \leq Q\langle w^{-1}\rangle_J,$
2. $\frac{1}{|J|}\sum_{I\subseteq J}\frac{\alpha_I}{\langle w^{-1}\rangle_I} \leq Q\langle w\rangle_J,$
3. $\frac{1}{|J|}\sum_{I\subseteq J}\alpha_I \leq Q,$

then for all $f, g \in L^2$

$$\sum_{I\in\mathcal{D}}\frac{\alpha_I}{\langle w\rangle_I\langle w^{-1}\rangle_I}\langle fw^{1/2}\rangle_I\langle gw^{-1/2}\rangle_I \leq 24Q\|f\|_2\|g\|_2. \tag{3.9}$$

Proof. The proof of this theorem is similar to that of Theorem 3.2 so we only present a sketch of it.
Instead of proving (3.9), we prove

$$\frac{1}{|J|}\sum_{I\subseteq J}\frac{\alpha_I}{\langle w\rangle_I\langle w^{-1}\rangle_I}\langle fw^{1/2}\rangle_I\langle gw^{-1/2}\rangle_I \leq 24Q(\langle f^2\rangle_J + \langle g^2\rangle_J)$$

and get back to the original inequality by using homogeneity.
We define a Bellman function B on

$$\begin{aligned}D = \{a = (X, Y, x, y, r, s, M, N, K) : x^2 \leq Xr, y^2 \leq Ys, 1 \leq rs \leq Q,\\ 0 \leq M \leq Q^2r, 0 \leq N \leq Q^2s, 0 \leq K \leq Q\}.\end{aligned}$$

by

$$B(a) = B_1(b) + B_2(c) + B_3(d),$$

where $b = (X, Y, x, y, r, s, M)$, $c = (X, Y, x, y, r, s, N)$, $d = (X, Y, x, y, r, s, K)$ and

$$B_1(b) = X - \frac{x^2}{r + \frac{M}{Q^2}} + Y - \frac{y^2}{s},$$

$$B_2(c) = X - \frac{x^2}{r} + Y - \frac{y^2}{s + \frac{N}{Q^2}},$$

$$B_3(d) = X + Y - \sup_{t>0} \left(\frac{x^2}{r + t\frac{K}{Q}} + \frac{y^2}{s + t^{-1}\frac{K}{Q}} \right).$$

The function B is such that

1. $0 \leq B \leq 3(X + Y)$;
2. $-d^2 B \geq C \dfrac{xy}{Qrs} |drds|$.

Indeed, the size property is immediate. For the Hessian, note that we have by the same computations as in the proof of Theorem 3.2 the following first derivatives on D

1. $\dfrac{\partial B_1}{\partial M} \geq \dfrac{x^2}{4Q^2 r^2}$;
2. $\dfrac{\partial B_2}{\partial N} \geq \dfrac{y^2}{4Q^2 s^2}$;
3. $\dfrac{\partial B_3}{\partial K} \geq \begin{cases} \dfrac{xy}{2Qrs} & \text{if } \dfrac{K}{Q} \leq \dfrac{yr}{2x} \text{ and } \dfrac{K}{Q} \leq \dfrac{xs}{2y}, \\ 0 & \text{elsewhere on } D. \end{cases}$

Using these estimates and the same reasoning as in Theorem 3.1 we obtain

$$\begin{aligned} B(a) \quad &\geq \quad B_1(X, Y, x, y, r, s, M - \Delta M) + \Delta M \frac{\partial B_1}{\partial M} + B_2(X, Y, x, y, r, s, N - \Delta N) \\ &\qquad + \Delta N \frac{\partial B_2}{\partial N} + B_3(X, Y, x, y, r, s, K - \Delta K) + \Delta K \frac{\partial B_3}{\partial K} \\ &\geq \quad B(X, Y, x, y, r, s, M - \Delta M, N - \Delta N, K - \Delta K) \\ &\qquad + (\frac{K}{s}\frac{\partial B_1}{\partial M} + \frac{K}{r}\frac{\partial B_2}{\partial N} + \frac{\partial B_3}{\partial K})\Delta K \\ &\geq \quad B(X, Y, x, y, r, s, M - \Delta M, N - \Delta N, K - \Delta K) + \Delta K \frac{xy}{8Qrs} \\ &\geq \quad \frac{1}{2}\left(B(a_l) + B(a_r)\right) + \Delta K \frac{xy}{8Qrs}. \end{aligned} \tag{3.10}$$

The second inequality is true if we assume that $M = \dfrac{M_l + M_r}{2} + \Delta M$, $\Delta M = \dfrac{K}{s}\Delta K$, $N = \dfrac{N_l + N_r}{2} + \Delta N$, $\Delta N = \dfrac{K}{r}\Delta K$ and $K = \dfrac{K_l + K_r}{2} + \Delta K$. In the third inequality we

used the previous derivatives and the different size restrictions on K. The last inequality is given by the concavity of B.
To finish the proof, fix

$$X = \langle f^2\rangle_J,\ Y = \langle g^2\rangle_J,\ x = \langle fw^{1/2}\rangle_J,\ y = \langle gw^{-1/2}\rangle_J,\ r = \langle w\rangle_J,\ s = \langle w^{-1}\rangle_J,$$

$$M_J = \frac{1}{|J|}\sum_{I\subseteq J}\frac{\alpha_I}{\langle w^{-1}\rangle_I}K_I,\ N_J = \frac{1}{|J|}\sum_{I\subseteq J}\frac{\alpha_I}{\langle w\rangle_I}K_I \text{ and } \ K_J = \frac{1}{|J|}\sum_{I\subseteq J}\alpha_I.$$

Moreover, the assumption made previously on M, N and K are true since $M = \dfrac{M_l + M_r}{2} + \dfrac{\alpha_J K_J}{\langle w^{-1}\rangle_J|J|}$, $N = \dfrac{N_l + N_r}{2} + \dfrac{\alpha_J}{\langle w\rangle_J|J|}$ and $K = \dfrac{K_l + K_r}{2} + \dfrac{\alpha_J}{|J|}$. Thus we put $\Delta K = \dfrac{\alpha_J}{|J|}$, $\Delta M = \dfrac{\alpha_J K_J}{\langle w^{-1}\rangle_J|J|} = \dfrac{K_J\Delta K}{s}$ and $\Delta N = \dfrac{\alpha_J K_J}{\langle w\rangle_J|J|} = \dfrac{K_J\Delta K}{r}$.
All of the variables are in D hence by applying the size property and (3.10) n times we obtain

$$3|J|(\langle f^2\rangle_J + \langle g^2\rangle_J) \geq \sum_{I\subseteq J, |I|>2^{-n}|J|}\frac{\alpha_I}{8Q\langle w\rangle_I\langle w^{-1}\rangle_I}\langle fw^{1/2}\rangle_I\langle gw^{-1/2}\rangle_I.$$

It suffices then to let $n\to\infty$ to get the result. □

Lemma 3.2 ([40]). *Let w be a weight in A_2 and Q its characteristic. Then for all $J\in\mathcal{D}$,*

$$\frac{1}{|J|}\sum_{I\subseteq J}\langle w^{-1}\rangle_I|\langle w\rangle_{I_l} - \langle w\rangle_{I_r}|^2|I| \leq \frac{80}{3}Q^2\langle w\rangle_J.$$

Proof. We prove this lemma by the Bellman method. We define a Bellman function B by

$$B(r,s) = (4Q^2+1)r - \frac{4Q^2}{s} - sr^2.$$

The function B is such that

1. $0 \leq B \leq 5Q^2 r$ on $D = \{(r,s)\in(\mathbb{R}_+)^2, 1\leq rs\leq Q\}$;

2. $-d^2B \geq \frac{3}{2}s(dr)^2$ on $D = \{(r,s)\in(\mathbb{R}_+)^2, 1\leq rs\leq Q\}$;

3. $-d^2B \geq 0$ on $D_2 = \{(r,s)\in(\mathbb{R}_+)^2, 1\leq rs\leq 2Q\}$.

Indeed, the majorization is immediate. Moreover,

$$\begin{aligned}(4Q^2+1)r - \frac{4Q^2}{s} - r^2 s &= \frac{4Q^2}{s}(rs-1) - r(rs-1)\\ &= (rs-1)(\frac{4Q^2}{s} - r)\\ &\geq 0.\end{aligned}$$

Hence the size property is proved. We also have

$$-d^2B - \frac{3}{2}s(dr)^2 = \frac{1}{2}s(dr)^2 - 4r(dsdr) + \frac{8Q^2}{s^3}(ds)^2.$$

This form is positive semidefinite on D because all the principal minors of its corresponding matrix (the diagonal elements and the determinant) are non negative. The same calculations can be made to prove the third point.
The properties of B being stated and checked, we pass to the proof of the lemma. First, we prove that the infinitesimal inequality $-d^2B \geq \frac{3}{2}s(dr)^2$ is equivalent to

$$B(r,s) - \frac{B(r_r,s_r)+B(r_l,s_l)}{2} \geq \frac{3}{16}s|r_r - r_l|^2, \tag{3.11}$$

for any three points (r,s), (r_r,s_r) and (r_l,s_l) in D such that $r = \frac{r_l+r_r}{2}$, $s = \frac{s_l+s_r}{2}$ and $\dfrac{(1-t)a_l+(1+t)a_r}{2} \in D$, where $t \in [0,1], a = (r,s)$.
First, define $\Delta r = \frac{r_l-r_r}{2}$ and $\Delta s = \frac{s_l-s_r}{2}$. By writing Taylor's formula at (r,s) for $B(r \pm \Delta r, s \pm \Delta s)$ we obtain

$$B(r\pm\Delta r, s\pm\Delta s) = B(r,s) \pm \nabla B(r,s)\begin{pmatrix}\Delta r\\ \Delta s\end{pmatrix} + \frac{1}{2}\langle d^2B(r,s)\begin{pmatrix}\Delta r\\ \Delta s\end{pmatrix}, \begin{pmatrix}\Delta r\\ \Delta s\end{pmatrix}\rangle.$$

Then,

$$\begin{aligned} B(r,s) - \frac{1}{2}(B(r_r,s_r)+B(r_l,s_l)) &= -\frac{1}{2}\langle d^2B(r,s)\begin{pmatrix}\Delta r\\ \Delta s\end{pmatrix}, \begin{pmatrix}\Delta r\\ \Delta s\end{pmatrix}\rangle \\ &\geq \frac{3}{16}s|r_r-r_l|^2. \end{aligned}$$

Since $|r_r - r_l|^2 = 4(\Delta r)^2$

$$-d^2B \geq \frac{3}{2}s(dr)^2.$$

In the other direction, define a function $\gamma : [-1,1] \to \mathbb{R}$, of class C^2. By a double integration by parts we have the following

$$\frac{1}{2}\int_{-1}^{1}(1-|t|)\gamma''(t)dt = \frac{\gamma(1)+\gamma(-1)-2\gamma(0)}{2}.$$

We apply this formula to $\gamma(t) = B(a+t\Delta a)$, where $a = (r,s)$. Notice that when the parameter t runs over $[-1,1]$, $a + t\Delta \in D_2$. Hence $t \mapsto \gamma(t)$ is concave. Moreover, when $t \in [0,1]$ or $t \in [-1,0]$, $a + t\Delta a$ is in D and we have a better estimate on $-\gamma''$. We obtain then

$$\begin{aligned}
B(a) - \frac{1}{2}(B(a+\Delta a) + B(a-\Delta a)) &= \gamma(0) - \frac{1}{2}\gamma(1) - \frac{1}{2}\gamma(-1) \\
&= -\frac{1}{2}\int_{-1}^{1}(1-|t|)\gamma''(t)dt \\
&= -\frac{1}{2}\int_{-1}^{1}(1-|t|)\langle d^2B(a+t\Delta a)\Delta a, \Delta a\rangle dt \\
&\geq \frac{1}{2}\int_{-1}^{1}(1-|t|)\frac{3}{2}(s+t\Delta s)(\Delta r)^2 dt \\
&= \frac{3}{2}s(\Delta r)^2\frac{1}{2}\int_{-1}^{1}(1-|t|)dt \\
&\geq \frac{3}{16}s|r_r - r_l|^2.
\end{aligned}$$

To finish the proof of the lemma, fix $r = \langle w\rangle_J$, $s = \langle w^{-1}\rangle_J$, $r_r = \langle w\rangle_{J_r}$, $r_l = \langle w\rangle_{J_l}$, $s_r = \langle w^{-1}\rangle_{J_r}$ and $s_l = \langle w^{-1}\rangle_{J_l}$. Next, apply (3.11) to get

$$\begin{aligned}
B(\langle w\rangle_J, \langle w^{-1}\rangle_J) \geq\ & \frac{1}{2}(B(\langle w\rangle_{J_l}, \langle w^{-1}\rangle_{J_l}) + B(\langle w\rangle_{J_r}, \langle w^{-1}\rangle_{J_r}) \\
& + \frac{3}{16}|\langle w\rangle_{J_r} - \langle w\rangle_{J_l}|^2|\langle w^{-1}\rangle_J|.
\end{aligned}$$

By applying this inequality on sub-intervals of J_l and J_r then use size property of the function B, we obtain

$$\frac{1}{|J|}\sum_{I\subseteq J}\langle w^{-1}\rangle_I|\langle w\rangle_{I_l} - \langle w\rangle_{I_r}|^2|I| \leq \frac{80}{3}Q^2\langle w\rangle_J.$$

Hence the lemma is proved. $\square$

Lemma 3.3 ([82]). *Let w be a weight in A_2 and Q its characteristic. Then for all $J \in \mathcal{D}$,*

$$\frac{1}{|J|}\sum_{I\subseteq J}\frac{\langle w\rangle_{I_l} - \langle w\rangle_{I_r}}{\langle w\rangle_I}\frac{\langle w^{-1}\rangle_{I_l} - \langle w^{-1}\rangle_{I_r}}{\langle w^{-1}\rangle_I}\langle w\rangle_I|I| \leq 40Q\langle w\rangle_J.$$

Proof. We prove this lemma by the Bellman method again. Let

$$B(r,s) = \frac{-4Q}{s} - \frac{r^2 s}{4Q} + (4Q+1)r$$

defined on the domain

$$D = \{(r,s) \in (\mathbb{R}_+)^2, 1 \leq rs \leq Q\}.$$

Then B is such that

1. $0 \leq B \leq 5Qr$;

2. $-d^2B \geq \dfrac{|dsdr|}{s}$.

Indeed the majorization is immediate and

$$\begin{aligned} \frac{-4Q}{s} - \frac{r^2 s}{4Q} + (4Q+1)r &\geq r - \frac{r^2 s}{4Q} \\ &\geq 0, \end{aligned}$$

because $4Qr \geq r^2 s$, which proves the size property. Moreover the matrices

$$-d^2B - \frac{drds}{s} = \begin{pmatrix} \frac{s}{2Q} & \frac{r}{2Q} - \frac{1}{s} \\ \frac{r}{2Q} - \frac{1}{s} & \frac{8Q}{s^3} \end{pmatrix}$$

and

$$-d^2B + \frac{drds}{s} = \begin{pmatrix} \frac{s}{2Q} & \frac{r}{2Q} + \frac{1}{s} \\ \frac{r}{2Q} + \frac{1}{s} & \frac{8Q}{s^3} \end{pmatrix}$$

are positive and semi definite on D because the diagonal elements are non negative and

$$\begin{aligned} \det(-d^2B - \frac{drds}{s}) &= \frac{4}{s^2} - \left(\frac{r}{2Q} - \frac{1}{s}\right)^2 \\ &= \frac{12Q^2 - r^2s^2 + 4rs}{4Q^2s^2} \geq 0, \end{aligned}$$

$$\begin{aligned} \det(-d^2B + \frac{drds}{s}) &= \frac{2}{s^2} - \left(\frac{r}{2Q} + \frac{1}{s}\right)^2 \\ &= \frac{12Q^2 - r^2s^2 - 4rs}{4Q^2s^2} \geq 0. \end{aligned}$$

As in the proof of Lemma 3.2, one can show that this is equivalent to

$$B(r,s) - \frac{B(r_l,s_l) + B(r_r,s_r)}{2} \geq \frac{1}{8}\frac{|s_l - s_r||r_l - r_r|}{s},$$

for any three points (r,s), (r_l,s_l) and (r_r,s_r) in D such that $r = \frac{r_r+r_l}{2}$, $s = \frac{s_r+s_l}{2}$ and $\dfrac{(1-t)a_l + (1+t)a_r}{2} \in D$, where $t \in [0,1], a = (r,s)$. Hence, by choosing $r = \langle w \rangle_J$, $s = \langle w^{-1} \rangle_J$, $r_r = \langle w \rangle_{J_r}$, $r_l = \langle w \rangle_{J_l}$, $s_r = \langle w^{-1} \rangle_{J_r}$ and $s_l = \langle w^{-1} \rangle_{J_l}$, we prove by iterating the inequality above and using the non negativity of B that

$$\frac{1}{|J|}\sum_{I \subseteq J} \frac{\langle w \rangle_{I_l} - \langle w \rangle_{I_r}}{\langle w \rangle_I} \frac{\langle w^{-1} \rangle_{I_l} - \langle w^{-1} \rangle_{I_r}}{\langle w^{-1} \rangle_I} \langle w \rangle_I |I| \leq 8B(\langle w \rangle_J, \langle w^{-1} \rangle_J).$$

The proof of the lemma is finished by using the majorization on B. □

Lemma 3.4 ([61]). *Let w be a weight in A_2 and Q its characteristic. Then for all $J \in \mathcal{D}$,*

$$\frac{1}{|J|}\sum_{I\subseteq J}\frac{\langle w\rangle_{I_l}-\langle w\rangle_{I_r}}{\langle w\rangle_I}\frac{\langle w^{-1}\rangle_{I_l}-\langle w^{-1}\rangle_{I_r}}{\langle w^{-1}\rangle_I}\langle w\rangle_I\langle w^{-1}\rangle_I|I| \leq 32\sqrt{Q}\sqrt{\langle w\rangle_J\langle w^{-1}\rangle_J}.$$

Proof. We prove this lemma by the Bellman method again. Let

$$B(r,s)=\sqrt{Qrs}-\frac{rs}{4},$$

defined on the domain

$$D=\{(r,s)\in(\mathbb{R}_+)^2, 1\leq rs\leq Q\}.$$

Then B is such that

1. $0 \leq B \leq \sqrt{Qrs}$;
2. $-d^2B \geq \frac{1}{4}|dsdr|$.

Indeed, size property is immediate. Moreover,

$$-d^2B \pm \frac{drds}{4} = \begin{pmatrix} \dfrac{\sqrt{Qs}}{4\sqrt{r^3}} & \dfrac{1}{4}-\dfrac{\sqrt{Q}}{4\sqrt{rs}}\pm\dfrac{1}{4} \\ \dfrac{1}{4}-\dfrac{\sqrt{Q}}{4\sqrt{rs}}\pm\dfrac{1}{4} & \dfrac{\sqrt{Qr}}{4\sqrt{s^3}} \end{pmatrix}$$

is positive semi definite on D because its diagonal elements are non negative and

$$\det(-d^2B-\frac{1}{4}dsdr)=0,$$

$$\det(-d^2B+\frac{1}{4}dsdr)=\frac{\sqrt{Q}}{4\sqrt{rs}}-\frac{1}{4}\geq 0.$$

Thus, one can show as in proof of Lemma 3.2 that the infinitesimal inequality $-d^2B \geq \frac{1}{4}|dsdr|$ is equivalent to

$$B(r,s)-\frac{B(r_l,s_l)+B(r_r,s_r)}{2}\geq\frac{1}{32}|s_l-s_r||r_l-r_r|,$$

for any three points (r,s), (r_r,s_r) and (r_l,s_l) in D such that $r=\frac{r_l+r_r}{2}$, $s=\frac{s_l+s_r}{2}$ and $\dfrac{(1-t)a_l+(1+t)a_r}{2}\in D$, where $t\in[0,1]$, $a=(r,s)$. The end of the proof is similar to those of Lemmata 3.2 and 3.3 and the details are left to the reader. $\square$

3.3 The martingale transform on $L^2(w)$

Define the dyadic martingale transform T_σ as

$$T_\sigma f = \sum_{I\in\mathcal{D}} \sigma_I (f, h_I) h_I,$$

where σ_I takes the values +1 and -1 only. The martingale transform is the dyadic analogue of singular integral operators in the continuous setting. For a weight $\omega \in A_2$, we want to estimate $\|T_\sigma f\|_{L^2(\omega)}$ and minimize the power of the characteristic $Q_2(\omega)$

$$Q_2(\omega) := \sup_I \frac{1}{|I|}\int_I \omega \frac{1}{|I|}\int_I \omega^{-1} < \infty$$

where the supremum runs over intervals I. By dualizing we obtain

$$\begin{aligned}
\|T_\sigma f\|_{L^2(\omega)} &= \sup_{\|g\|_{L^2(\omega^{-1})}=1} \left| \int (T_\sigma f)\cdot g dx \right| \\
&= \sup_{\|g\|_{L^2(\omega^{-1})}=1} \left| \int \sum_{I,J\in\mathcal{D}} \sigma_I (f, h_I)(g, h_J) h_I(x) h_J(x) dx \right| \\
&= \sup_{\|g\|_{L^2(\omega^{-1})}=1} \left| \sum_{I\in\mathcal{D}} \sigma_I (f, h_I)(g, h_I) \right|
\end{aligned}$$

Replacing f by $f\omega^{1/2}$ and g by $g\omega^{-1/2}$ we obtain

$$\|T_\sigma f\|_{L^2(\omega)} = \sup_{\|g\|_{L^2}=1} \left| \sum_{I\in\mathcal{D}} \sigma_I (f\omega^{-1/2}, h_I)(g\omega^{1/2}, h_I) \right|. \tag{3.12}$$

Instead of using the classical Haar system, we need to switch to a more appropriate system of disbalanced Haar functions (h_I^ω) that is orthonormal in $L^2(\omega)$. We use Lemma 3.1 to substitute the first and second h_I in (3.12) by $h_I^{\omega^{-1}}$ and h_I^ω respectively and pass the absolute value inside to obtain a sum of four summands $\Sigma_1 + \Sigma_2 + \Sigma_3 + \Sigma_4$ which are

$$\Sigma_1 = \sum_{I\in\mathcal{D}} |(f\omega^{-1/2}, h_I^{\omega^{-1}})(g\omega^{1/2}, h_I^\omega)\frac{1}{x_I^\omega x_I^{\omega^{-1}}}|,$$

$$\Sigma_2 = \sum_{I\in\mathcal{D}} |\langle f\omega^{-1/2}, \chi_I\rangle\langle g\omega^{1/2}, h_I^\omega\rangle \frac{A_I^{\omega^{-1}}}{x_I^{\omega^{-1}}}\frac{1}{x_I^\omega}|,$$

$$\Sigma_3 = \sum_{I\in\mathcal{D}} |\langle g\omega^{1/2}, \chi_I\rangle\langle f\omega^{-1/2}, h_I^\omega\rangle \frac{A_I^{\omega}}{x_I^{\omega}}\frac{1}{x_I^{\omega^{-1}}}|,$$

and

$$\Sigma_4 = \sum_{I\in\mathcal{D}} |\langle f\omega^{-1/2}\rangle_I \langle g\omega^{1/2}\rangle_I \frac{A_I^{\omega^{-1}}}{x_I^{\omega^{-1}}}\frac{A_I^\omega}{x_I^\omega}|I|^2|.$$

We are going to bound each of these sums by $cQ_2\|f\|_2\|g\|_2$ using the Bellman function method. This method is particularly adapted to bound the dyadic martingale transform. Our approach will consist of reducing the boundedness of each sum to that of bounding a quantity independent of the choice of the dyadic interval. Next, we write an inequality for the above quantity by splitting the dyadic interval into its left and right parts. Assuming we have a solution to the inequality, we bound the quantity and then conversely, we find a solution to the inequality.
This result is originally due to Janine Wittwer in [82]. In what follows, we detail these steps on each sum and explicitly present all the calculations. A parallel analysis will be conducted in order to understand the origin of the function used to prove the weighted boundedness of the Bakry-Riesz transform on manifolds.

More specifically, we have
Estimate of Σ_1

$$\Sigma_1 = \sum_{I\in\mathcal{D}} |(f\omega^{-1/2}, h_I^{\omega^{-1}})(g\omega^{1/2}, h_I^{\omega})\frac{1}{x_I^{\omega}x_I^{\omega^{-1}}}|$$

First method: We first use the fact that by polarization identity we have $w_{I_l}w_{I_r} = w_I^2 - (\Delta w)^2 \leq w_I^2$ (and the same goes for w^{-1}) to bound $\frac{1}{x_I^{\omega}x_I^{\omega^{-1}}} \leq Q_2^{1/2}$ (by definition of x_I^{ω}) then use Cauchy-Schwarz and Bessel inequalities to conclude.
Second method: We proceed as in the proof of the weighted Carleson embedding theorem (Theorem 3.1. Instead of proving that

$$\sum_{I\in\mathcal{D}} |(f\omega^{-1/2}, h_I^{\omega^{-1}})(g\omega^{1/2}, h_I^{\omega})\frac{1}{x_I^{\omega}x_I^{\omega^{-1}}}| \leq Q_2^{1/2}\|f\|_2\|g\|_2,$$

we prove the following larger inequality

$$\sum_{I\in\mathcal{D}} \frac{1}{Q_2^{1/2}}|(f\omega^{-1/2}, h_I^{\omega^{-1}})(g\omega^{1/2}, h_I^{\omega})\frac{1}{x_I^{\omega}x_I^{\omega^{-1}}}| \leq (\|f\|_2^2 + \|g\|_2^2). \tag{3.13}$$

We define a function B_1 of six variables on the domain

$$D = \{(X, Y, x, y, r, s) \in \mathbb{R}^4 \times \mathbb{R}_+^2 : x^2 \leq Xr, y^2 \leq Ys, 1 \leq rs \leq Q_2\}.$$

as follows

$$B_1(X, Y, x, y, r, s) = \frac{1}{|J|}\sup_{f,g,\omega}\sum_{I\subseteq J}\frac{1}{Q_2^{1/2}}|(f\omega^{-1/2}, h_I^{\omega^{-1}})(g\omega^{1/2}, h_I^{\omega})\frac{1}{x_I^{\omega}x_I^{\omega^{-1}}}|$$

for a fixed $J \in \mathcal{D}$ such that $I \subseteq J$, where the supremum is taken over all f, g and w that satisfy

$$\langle fw^{-1/2}\rangle_J = x, \quad \langle f^2\rangle_J = X, \quad \langle w^{-1}\rangle_J = r, \quad \langle gw^{1/2}\rangle_J = y, \quad \langle g^2\rangle_J = Y, \quad \langle w\rangle_J = s.$$

The function B_1 satisfies the size property $0 \le B_1 \le (X+Y)$ because we want (3.13) to be true and for $a=(X,Y,x,y,r,s)$, the main inequality

$$B_1(a) \ge \frac{1}{2}(B_1(a_l)+B_1(a_r)) + \frac{1}{Q_2^{1/2}|J|}|(f\omega^{-1/2}, h_J^{\omega^{-1}})(g\omega^{1/2}, h_J^{\omega})\frac{1}{x_J^{\omega}x_J^{\omega^{-1}}}|.$$

The second part of the right hand side of the inequality becomes (after using that $h_J^{\omega} = x_J h_J - A_J\chi_J$)

$$\frac{1}{4Q_2^{1/2}}\left((x_l - x_r) - \frac{r_l - r_r}{r}x\right)\left((y_l - y_r) - \frac{s_l - s_r}{s}y\right).$$

Using Taylor's formula, it is easy to see (cf. Lemma 3.2) that the main inequality is equivalent to its infinitesimal version

$$-d^2B_1 \ge \frac{2}{Q_2^{1/2}}|dx - \frac{dr}{r}x||dy - \frac{ds}{s}y|.$$

Conversely, if we have a function satisfying the previous properties, then by applying n times the main inequality to the right half and the left half of J_l and J_r we obtain

$$\begin{aligned} |J|B_1(a) &\ge \frac{1}{2}\sum_{I\in\mathcal{D},\ |I|=2^{-n}|J|}(B_1(a_I)+B_1(a_I)) \\ &\quad + \sum_{I\in\mathcal{D},|I|>2^{-n}|J|}\frac{1}{Q_2^{1/2}}|(f\omega^{-1/2}, h_I^{\omega^{-1}})(g\omega^{1/2}, h_I^{\omega})\frac{1}{x_I^{\omega}x_I^{\omega^{-1}}}|. \end{aligned}$$

Hence by size property the estimate is obtained after letting $n \to +\infty$.

Finally, the function

$$B_1(X,Y,x,y,r,s) = X - \frac{x^2}{r} + Y - \frac{y^2}{s}$$

satisfies the desired properties. Indeed, one can easily check that it satisfies on D the size property. Moreover,

$$\begin{aligned} -d^2B_1 &= \frac{2}{r}(dx)^2 - \frac{4x}{r^2}(dxdr) + \frac{2x^2}{r^3}(dr)^2 + \frac{2}{s}(dy)^2 - \frac{4y}{s^2}(dyds) + \frac{2y^2}{s^3}(ds)^2 \\ &= \frac{2}{r}|dx - x\frac{dr}{r}|^2 + \frac{2}{s}|dy - y\frac{ds}{s}|^2 \\ &\ge \frac{4}{\sqrt{Q_2}}|dx - x\frac{dr}{r}||dy - y\frac{ds}{s}|. \end{aligned}$$

Consequently, $|\sum_1| \le 4Q_2\|f\|_2\|g\|_2$

Next, Σ_2 and Σ_3 are symmetric so we only focus on Σ_2.

Estimate of Σ_2

$$\begin{aligned}\Sigma_2 &= \sum_{I\in\mathcal{D}} |\langle f\omega^{-1/2}, \chi_I\rangle\langle g\omega^{1/2}, h_I^{\omega}\rangle \frac{A_I^{\omega^{-1}}}{x_I^{\omega^{-1}}}\frac{1}{x_I^{\omega}}| \\ &= \sum_{I\in\mathcal{D}} |\langle f\omega^{-1/2}\rangle_I\langle g\omega^{1/2}, h_I^{\omega}\rangle \frac{A_I^{\omega^{-1}}}{x_I^{\omega^{-1}}}\frac{1}{x_I^{\omega}}|I||.\end{aligned}$$

First method: We use the fact that

$$|\frac{A_I^{\omega^{-1}}}{x_I^{\omega^{-1}}}\frac{1}{x_I^{\omega}}|I|| \le \sqrt{\langle w\rangle_I}\frac{|\langle w^{-1}\rangle_{I_l} - \langle w^{-1}\rangle_{I_r}|}{\langle w^{-1}\rangle_I}\sqrt{|I|}$$

and split the sum by using Cauchy Schwarz inequality to obtain

$$|\Sigma_2| \le \left(\sum_{I\in\mathcal{D}} |\langle g\omega^{1/2}, h_I^{\omega}\rangle|^2\right)^{1/2}\left(\sum_{I\in\mathcal{D}} |\langle f\omega^{-1/2}\rangle_I|^2\langle w\rangle_I\frac{|\langle w^{-1}\rangle_{I_l} - \langle w^{-1}\rangle_{I_r}|^2}{\langle w^{-1}\rangle_I^2}|I|\right)^{1/2}.$$

Again, for the first part, we can use Bessel inequalities and bound the term by $\|g\|_2$. For the second part, we apply the weighted Carleson theorem for $\alpha_I = \langle w\rangle_I\frac{|\langle w^{-1}\rangle_{I_l} - \langle w^{-1}\rangle_{I-r}|^2}{\langle w^{-1}\rangle_I^2}|I|$ since it satisfies the Carleson condition

$$\frac{1}{|J|}\sum_{I\subseteq J}\langle w\rangle_I|\langle w^{-1}\rangle_{I_l} - \langle w^{-1}\rangle_{I_r}|^2|I| \le \frac{80}{3}Q_2^2\langle w^{-1}\rangle_J$$

by Lemma 3.2. Hence the second part is bounded by $11Q_2\|f\|_2$.

Second method: Another method consists of exhibiting a Bellman function. Let

$$\langle fw^{-1/2}\rangle_J = x,\quad \langle f^2\rangle_J = X,\quad \langle w^{-1}\rangle_J = r,\quad \langle gw^{1/2}\rangle_J = y,\quad \langle g^2\rangle_J = Y,\quad \langle w\rangle_J = s.$$

Since $h_J^{\omega} = x_J h_J - A_J\chi_J$, we have

$$\begin{aligned}&\textstyle\sum_{I\in\mathcal{D}} |\langle f\omega^{-1/2}\rangle_I\langle g\omega^{1/2}, h_I^{\omega}\rangle \frac{A_I^{\omega^{-1}}}{x_I^{\omega^{-1}}}\frac{1}{x_I^{\omega}}|I|| \\ &= \textstyle\sum_{I\in\mathcal{D}} \frac{|I|}{4}|\frac{\langle w^{-1}\rangle_{I_l} - \langle w^{-1}\rangle_{I_r}}{\langle w^{-1}\rangle_I}x_I\left((\langle y\rangle_{I_l} - \langle y\rangle_{I_r}) - \frac{\langle w\rangle_{I_l} - \langle w\rangle_{I_r}}{\langle w\rangle_I}\right)|.\end{aligned}$$

Let

$$B_2(X,Y,x,y,r,s) = \frac{1}{|J|}\sum_{I\in\mathcal{D}}\frac{|I|}{4Q}|\frac{\langle w^{-1}\rangle_{I_l} - \langle w^{-1}\rangle_{I_r}}{\langle w^{-1}\rangle_I}x_I\left((\langle y\rangle_{I_l} - \langle y\rangle_{I_r}) - \frac{\langle w\rangle_{I_l} - \langle w\rangle_{I_r}}{\langle w\rangle_I}\right)|.$$

Then we have on

$$D = \{(X,Y,x,y,r,s) \in \mathbb{R}^4\times\mathbb{R}_+^2 : x^2 \le Xr, y^2 \le Ys, 1 \le rs \le Q\},$$

1. $0 \leq B_2 \lesssim X + Y$;
2. For $a = (X, Y, x, y, r, s)$,
$$B_2(a) - \frac{1}{2}(B_2(a_l) + B_2(a_r)) \geq \frac{1}{4Q} \left| \frac{\langle w^{-1}\rangle_{J_l} - \langle w^{-1}\rangle_{J_r}}{\langle w^{-1}\rangle_J} x_J \left((\langle y\rangle_{J_l} - \langle y\rangle_{J_r}) - \frac{\langle w\rangle_{J_l} - \langle w\rangle_{J_r}}{\langle w\rangle_J} \right) \right|,$$
which is equivalent to its infinitesimal version
$$-d^2 B_2 \geq \frac{2}{Q} \left|\frac{x}{r} dr\right| \left|dy - \frac{ds}{s} y\right|.$$

The function
$$B_2(X, Y, x, y, r, s) = 2\sqrt{3} \left(X - \frac{x^2}{r + \frac{M(r,s)}{Q^2}} + Y - \frac{y^2}{s} \right)$$
is suitable with
$$M(r, s) = (4Q^2 + 1)r - \frac{4Q^2}{s} - r^2 s,$$
taken from Lemma 3.2. Indeed, size property is immediate. Moreover, the function
$$F(X, x, r, z) = X - \frac{x^2}{r + z}$$
is concave for $y, z \geq 0$ because
$$-d^2 F = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{2}{y+z} & \frac{-2x}{(y+z)^2} & \frac{-2x}{(y+z)^2} \\ 0 & \frac{-2x}{(y+z)^2} & \frac{2x^2}{(y+z)^3} & \frac{2x^2}{(y+z)^3} \\ 0 & \frac{-2x}{(y+z)^2} & \frac{2x^2}{(y+z)^3} & \frac{2x^2}{(y+z)^3} \end{pmatrix}$$
is positive semi definite. Hence B_2 is a sum of 2 concave functions. Then if we write $B_{22}(X, x, r, s) = 2\sqrt{3}\left(X - \frac{x^2}{r + \frac{M(r,s)}{Q^2}} \right)$ and consider $H = B_{22} \circ M$, we have by chain rule
$$\begin{aligned} \frac{-d^2 H}{2\sqrt{3}} &= -d^2 B_{22} + \frac{\partial B_{22}}{\partial (M/Q^2)} \frac{1}{Q^2}(-d^2 M) + \frac{2}{s}\left|dy - y\frac{ds}{s}\right|^2 \\ &\geq \frac{\partial B_{22}}{\partial (M/Q^2)} \frac{1}{Q^2}(-d^2 M) + \frac{2}{s}\left|dy - y\frac{ds}{s}\right|^2 \\ &\geq \frac{x^2}{36 r^2} \frac{3}{2Q^2} s (dr)^2 + \frac{2}{s}\left|dy - y\frac{ds}{s}\right|^2 \\ &\geq \frac{x}{\sqrt{3} Q r} dr \left|dy - y\frac{ds}{s}\right|. \end{aligned}$$

Consequently, $|\sum_2| \leq 14Q_2\|f\|_2\|g\|_2$

Estimate of Σ_3 By symmetry,

$$B_3(X,Y,x,y,r,s) = 2\sqrt{3}\left(X - \frac{x^2}{r} + Y - \frac{y^2}{s + \frac{N(r,s)}{Q_2^2}}\right)$$

where

$$N(r,s) = (4Q^2+1)s - \frac{4Q^2}{r} - rs^2.$$

Again, $|\Sigma_3| \leq 14Q_2\|f\|_2\|g\|_2$.

Estimate of Σ_4 We turn now to the last sum Σ_4.

$$\Sigma_4 = \sum_{I\in\mathcal{D}} |\langle f\omega^{-1/2}\rangle_I \langle g\omega^{1/2}\rangle_I \frac{A_I^{\omega^{-1}}}{x_I^{\omega^{-1}}}\frac{A_I^{\omega}}{x_I^{\omega}}|I|^2|.$$

We estimate this sum by using the weighted bilinear Carleson embedding theorem for $\alpha_I = \frac{A_I^{\omega^{-1}}}{x_I^{\omega^{-1}}}\frac{A_I^{\omega}}{x_I^{\omega}}|I|$. We claim that for such α_I the conditions of the weighted bilinear Carleson embedding are satisfied. Indeed, we only prove the first inequality since the second one is proved analogously.

$$\begin{aligned}
\frac{1}{|J|}\int_J \left(\sum_{I\subseteq J} \alpha_I \langle w\rangle_I \chi_I(x)\right)^2 w^{-1}dx &= \frac{1}{|J|}\int_J \left(\sum_{I\subseteq J, K\subseteq J} \alpha_I \langle w\rangle_I \chi_I(x)\alpha_K\langle w\rangle_K \chi_K(x)\right) w^{-1}dx \\
&= \frac{1}{|J|}\sum_{I\subseteq J, K\subseteq J} \alpha_I\langle w\rangle_I \alpha_K \langle w\rangle_K \int_J \chi_I(x)\chi_K(x) w^{-1}dx \\
&= \frac{1}{|J|}\Big(\sum_{I=K\subseteq J} \alpha_I^2 \langle w\rangle_I^2 \langle w^{-1}\rangle_I |I| \\
&\quad +2\sum_{I,K\subseteq J, K\subsetneq I} \alpha_I\langle w\rangle_I \alpha_K\langle w\rangle_K \langle w^{-1}\rangle_K |K|\Big).
\end{aligned}$$

The first sum is (for $\alpha_I = \frac{A_I^{\omega^{-1}}}{x_I^{\omega^{-1}}}\frac{A_I^{\omega}}{x_I^{\omega}}|I|$)

$$\frac{1}{16|J|}\sum_{I\subseteq J} |\langle w\rangle_{I_l} - \langle w\rangle_{I_r}|^2 \frac{|\langle w^{-1}\rangle_{I_l} - \langle w^{-1}\rangle_{I_r}|^2}{\langle w^{-1}\rangle_I^2}\langle w^{-1}\rangle_I |I|.$$

However, a simple calculation shows that $\frac{|\langle w^{-1}\rangle_{I_l} - \langle w^{-1}\rangle_{I_r}|^2}{\langle w^{-1}\rangle_I^2} \leq 4$ and we estimate the above by

$$\frac{1}{4|J|}\sum_{I\subseteq J} |\langle w\rangle_{I_l} - \langle w\rangle_{I_r}|^2 \langle w^{-1}\rangle_I |I| \leq \frac{20}{3} Q_2^2 \langle w\rangle_J,$$

by Lemma 3.2. The second sum is

$$\frac{1}{2|J|}\sum_{I\subseteq J}\frac{\langle w\rangle_{I_l}-\langle w\rangle_{I_r}}{\langle w\rangle_I}\frac{\langle w^{-1}\rangle_{I_l}-\langle w^{-1}\rangle_{I_r}}{\langle w^{-1}\rangle_I}\langle w\rangle_I\sum_{K\subsetneq I}\frac{\langle w\rangle_{K_l}-\langle w\rangle_{K_r}}{\langle w\rangle_K}\frac{\langle w^{-1}\rangle_{K_l}-\langle w^{-1}\rangle_{K_r}}{\langle w^{-1}\rangle_K}\langle w\rangle_K\langle w^{-1}\rangle_K|K|.$$

By Lemma 3.4, the inner sum is bounded by $32Q_2|I|$. Hence the above estimate is bounded by

$$\frac{16Q_2}{|J|}\sum_{I\subseteq J}\frac{\langle w\rangle_{I_l}-\langle w\rangle_{I_r}}{\langle w\rangle_I}\frac{\langle w^{-1}\rangle_{I_l}-\langle w^{-1}\rangle_{I_r}}{\langle w^{-1}\rangle_I}\langle w\rangle_I|I|,$$

which is in turn bounded by $640Q^2\langle w\rangle_J$, by Lemma 3.3. Consequently,

$$\frac{1}{|J|}\int_J\left(\sum_{I\subseteq J}\alpha_I\langle w\rangle_I\chi_I(x)\right)^2 w^{-1}dx\le\frac{1940}{3}Q^2\langle w\rangle_J,$$

and the conditions of Theorem 3.2 are satisfied so we can use the weighted bilinear Carleson embedding where $v=\frac{w^{-1}}{\frac{1940}{3}Q^2}$. Therefore,

$$\sum_{I\in\mathcal{D}}\langle fw^{1/2}\rangle_I\langle gw^{-1/2}\rangle_I\alpha_I|I|\le 24\cdot\sqrt{\frac{1940}{3}}Q_2\|f\|_{L^2(w)}\|g\|_{L^2(w^{-1})}$$

and

$$|\Sigma_4|\le 624Q_2\|f\|_2\|g\|_2.$$

In conclusion, we obtain the following weighted estimate on martingale transforms

$$\|T_\sigma f\|_{L^2(w)}\le 656Q_2\|f\|_{L^2(w)}.$$

Remark 3.3. *In the case of the Bakry-Riesz vector (see Chapter 4), we will prove that*

$$\int_X\langle\mathcal{R}_\varphi f(x),\vec{g}(x)\rangle\,d\mu_\varphi(x)=4\int_0^\infty\int_X\left\langle dP_tf(x),\frac{d}{dt}\vec{P}_t\vec{g}(x)\right\rangle d\mu_\varphi(x)\,tdt.$$

and

$$\|\mathcal{R}_\varphi f\|_{L^2(w)}\le 4\sup_{\|\vec{g}\|_{L^2(\omega^{-1})}\le 1}\int_0^\infty\int_X|\nabla P_tf(x)||\vec{\nabla}\vec{P}_t\vec{g}(x)|d\mu_\varphi(x)tdt$$

which we want to be less than $C\widetilde{Q}_2(w)\|f\|_{L^2(w)}\|\vec{g}\|_{L^2(w^{-1})}$.
This means that $B=\sum_{i=1}^4 B_i$ has to satisfy

$$-d^2B\gtrsim\frac{1}{Q_2}|dx||dy|$$

However we know that

$$\begin{aligned} -d^2B \gtrsim & \ \frac{1}{Q_2}(|dx - \tfrac{dr}{r}x||dy - \tfrac{ds}{s}y| \\ & +|\frac{x}{r}dr| \times |dy - \tfrac{ds}{s}y| \\ & +|dx - \tfrac{dr}{r}x| \times |\frac{y}{s}ds| \\ & +A) \\ \gtrsim & \ \frac{1}{Q_2}|dx||dy|. \end{aligned}$$

A careful reader may notice that by triangle inequality it suffices for A to be equal to $|\frac{xy}{rs}drds|$.

Chapter 4

Sharp dimension-free weighted bounds for the Bakry-Riesz vector.

There are no L^p spaces, only weighted L^2.

Antonio Cordoba

We prove in this chapter a sharp dimensionless weighted L^2 estimate of the Riesz vector on a Riemannian manifold with non-negative Bakry-Emery curvature. The proof is by the method of Bellman functions, where the explicit expression of a Bellman function of six variables is essential.
An important fact is that our estimate in terms of the Poisson characteristic of the weight includes the case of the Gauss space as well as other spaces that are not necessarily of homogeneous type.

4.1 Development

Before stating the main result of this chapter and its proof, we define the context of this work, as well as some notations. We refer the reader to Chapter 2 for more detailed definitions and properties of the analytical and geometrical objects used throughout this chapter.

4.1.1 Setting and notations

Let (X, g, μ) be a complete Riemannian manifold of dimension N without boundary. Let Δ be the negative Laplace-Beltrami operator on X. Given $\varphi \in C^2(X)$, consider the weighted measure on X defined by

$$d\mu_\varphi(x) = e^{-\varphi(x)} d\mu(x),$$

and assume that constants are integrable with respect to this weighted measure. The importance of this assumption will appear in Section 4.5. Next, denote by Δ_φ the associated weighted Laplacian defined on $C_c^\infty(X)$ by

$$\Delta_\varphi f = \Delta f - \nabla f \cdot \nabla \varphi.$$

Notice that for all $f, g \in C_c^\infty(X)$, we have

$$\int_X (\nabla f, \nabla g) d\mu_\varphi(x) = -\int_X f \Delta_\varphi g d\mu_\varphi(x) = -\int_X g \Delta_\varphi f d\mu_\varphi(x). \tag{4.1}$$

It was proved in [12] and [79] that on complete Riemannian manifolds, the operator Δ_φ is essentially self-adjoint on $L^2(X, \mu_\varphi)$. We will still note Δ_φ its unique self-adjoint extension.
The Bakry-Emery curvature tensor associated with Δ_φ is defined by

$$\mathrm{Ric}_\varphi = \mathrm{Ric} + \nabla^2 \varphi,$$

where Ric denotes the Ricci curvature tensor on X and $\nabla^2\varphi$ is the Hessian of φ (i.e. a $2-$tensor). All over this paper, we will consider that $\mathrm{Ric}_\varphi \geq 0$.

For each x in X, we denote the tangent space and its dual, the cotangent space at x respectively by T_xX and T_x^*X so that

$$TX = \bigcup_{x\in X} T_xX \text{ and } T^*X = \bigcup_{x \in X} T_x^*X.$$

We denote by $< \cdot, \cdot >$ either the inner product in TX and T^*X, or in the Lebesgue space $L^2(X, \mu_\varphi)$ with a subscript to avoid ambiguities.
By $\| \cdot \|_{L^2(X, \omega\mu_\varphi)}$ and $\| \cdot \|_{L^2(T^*X, \omega^{-1}\mu_\varphi)}$, we denote respectively the norm in $L^2(X, \omega\mu_\varphi)$ and $L^2(T^*X, \omega^{-1}\mu_\varphi)$, where ω and ω^{-1} are weights that belong to $L^1_{loc}(X, \mu_\varphi)$.

We denote by d and ∇ respectively the exterior and the covariant derivative and d_φ^* and ∇_φ^* their $L^2(\mu_\varphi)$-adjoint operators. An easy computation shows that

$$d_\varphi^* = -d^* + i_{\nabla\varphi},$$

where $i_{\nabla\varphi}$ denotes the inner multiplication by $\nabla\varphi$ on Λ^1.
We also define $\overline{\nabla}$ as the total covariant derivative on $X \times \mathbb{R}_+$ that satisfies $|\overline{\nabla}\eta| = \sqrt{|\nabla\eta|^2 + |\partial_t\eta|^2}$, for all η in $C^\infty(T^*X(X \times \mathbb{R}_+))$.

We consider $\vec{\Delta}_\varphi = dd_\varphi^* + d_\varphi^* d$ to be the weighted Hodge-De Rham Laplacian acting on $1-$forms. As for the Laplace Beltrami operator, $\vec{\Delta}_\varphi$ initially defined on smooth $1-$forms with compact support is essentially self-adjoint on $L^2(T^*X, \mu_\varphi)$ and again, we will denote $\vec{\Delta}_\varphi$ its unique self-adjoint extension.

Finally, self-adjointness of the considered weighted Laplacians allows to define by spectral theory the following semigroups that respectively act on functions and 1-forms

$$P_t = \exp(-t(-\Delta_\varphi)^{1/2}) \text{ and } \vec{P}_t = \exp(-t(\vec{\Delta}_\varphi)^{1/2}).$$

Note that the semigroup P_t acting on functions is an integral operator with positive kernel that we will note p_t [79].

We are concerned in this chapter with a special class of weights, called Poisson$-A_2$ and noted $\widetilde{A}_2$. We say that $\omega \in \widetilde{A}_2$ if and only if

$$\widetilde{Q}_2(\omega) := \sup_{(x,t)\in X\times\mathbb{R}_+} P_t(\omega)(x)P_t(\omega^{-1})(x) < \infty.$$

The weights involved in this definition are a priori in $L^2(X,\mu_\varphi)$.

Remark 4.1. *For $\omega \in L^1_{loc}(X,\mu_\varphi)$, we define its two-sided truncation*

$$\omega_n = n^{-1}\chi_{\omega\leq n^{-1}} + \omega\chi_{n^{-1}\leq\omega\leq n} + n\chi_{\omega\geq n},$$

where χ is the characteristic function and $n \in \mathbb{N}^*$. The truncated weight ω_n is clearly in $L^2(X,\mu_\varphi)$ and satisfies some interesting properties that we are going to see later. For the moment, we are going to work with ω_n and then extend our results to ω, including a definition for $\widetilde{Q}_2(\omega)$ when ω is only locally integrable. We are also going to suppose that $P_t\omega_n$ and $P_t\omega_n^{-1}$ are finite almost everywhere so that $\widetilde{Q}_2(\omega_n)$ makes sense.

Remark 4.2. *Throughout this chapter, C will denote constants whose values may change even in a chain of inequalities. These constants are independent of the dimension of the manifold and other important quantities.*

4.1.2 Preliminaries

The following lemma slightly differs from the one appearing in [12] since it involves weights. The stated results will be of great utility in the next sections.

Lemma 4.1. *For every $f \in C_c^\infty(X)$, $\vec{g} \in C_c^\infty(T^*X)$ and $\omega \in L^2(X,\mu_\varphi)$,*

a) $|P_t f(x)|^2 \leq P_t\left(|f|^2\omega\right)(x)\, P_t\omega^{-1}(x)$.

b) $dP_t = \vec{P}_t d$ *and consequently* $\vec{\Delta}_\varphi d = d(-\Delta_\varphi)$.

c) $d_\varphi^* P_t = \vec{P}_t d_\varphi^*$ *and consequently* $\vec{\Delta}_\varphi d_\varphi^* = d_\varphi^*(-\Delta_\varphi)$.
If we also have $\mathrm{Ric}_\varphi \geq 0$ *then*

d) $\left|e^{-t\vec{\Delta}_\varphi}\vec{g}(x)\right|_{T_x^*X} \leq e^{t\Delta_\varphi}|\vec{g}(x)|_{T_x^*X}$.

e) $\left|\vec{P}_t\vec{g}(x)\right|^2_{T_x^*X} \leq P_t\left(|\vec{g}|^2_{T_x^*X}\,\omega^{-1}\right)(x)\, P_t\omega(x)$.

Proof. Items (b), (c) and (d) in the lemma have been proved in [12].
For item (a), we use the integral expression of P_t and Hölder's inequality.

$$\begin{aligned} P_t f(x) &= \int_X p_t^{1/2}(x,y) f(y) \omega^{1/2}(y) p_t^{1/2}(x,y) \omega^{-1/2}(y) d\mu_\varphi(y) \\ &\leq \left(\int_X p_t(x,y) |f(y)|^2 \omega(y) d\mu_\varphi(y) \right)^{1/2} \times \left(\int_X p_t(x,y) \omega^{-1}(y) d\mu_\varphi(y) \right)^{1/2}. \end{aligned}$$

To conclude, simply raise to the power 2 the above inequality.
To prove item (e), note that by [12, Inequality (1.4)], one can write

$$\left| \vec{P}_t \vec{g}\,(x) \right|_{T_x^* X} \leq P_t |\vec{g}|_{T_x^* X}(x).$$

The proof is then analogous to the one of item (a). □

4.2 Bilinear embedding and its corollary

In this section, we state the main result of this paper and its corollary for the boundedness of the Riesz transform. There have been considerable efforts in bounding the Riesz transform as well as finding its exact norm. The reasons behind these interests come from [46] where it was pointed out that sharp estimates of the norm of the Riesz transform imply important results for non linear geometric PDEs and in the Hodge decomposition theory. The proof of Theorem 4.1 will be given in Section 4.4.

Theorem 4.1. *Suppose that (X, μ_φ) is a complete Riemannian manifold without boundary and of finite measure. Suppose also that $\mathrm{Ric}_\varphi \geq 0$ and that ω_n and ω_n^{-1} are a.e positive weights defined as in Remark 4.1. Then for all $f \in C_c^\infty(X)$ and $\vec{g} \in C_c^\infty(T^*X)$, we have the following dimension-free estimate*

$$\int_0^\infty \int_X |\nabla P_t f(x)| |\vec{\nabla} \vec{P}_t \vec{g}(x)| t \, d\mu_\varphi(x) dt \leq 221 \; \widetilde{Q}_2(\omega_n) \|f\|_{L^2(X, \omega_n \mu_\varphi)} \|\vec{g}\|_{L^2(T^*X, \omega_n^{-1} \mu_\varphi)}. \tag{4.2}$$

Let $\mathcal{R}_\varphi$ denote the Riesz transform initially defined on the range space of $-\Delta_\varphi$, $R(-\Delta_\varphi)$ by

$$\mathcal{R}_\varphi = d \circ (-\Delta_\varphi)^{-1/2}$$

and that extends to a contraction

$$\mathcal{R}_\varphi : \overline{R(-\Delta_\varphi)} \to L^2(T^*X).$$

We have then the following corollary :

Corollary 4.1. *Under the above conditions,*

$$\|\mathcal{R}_\varphi f\|_{L^2(T^*X, \omega_n \mu_\varphi)} \leq 884 \widetilde{Q}_2(\omega_n) \|f\|_{L^2(X, \omega \mu_\varphi)},$$

for all $f \in \overline{L^2(X, \omega_n \mu_\varphi) \cap R(-\Delta_\varphi)}^{L^2}$.

Proof. The idea of the proof is to represent the Riesz transform by using Poisson semigroups on functions and differential forms. Indeed, for every $f \in L^2(X, \omega_n\mu_\varphi) \cap R(-\Delta_\varphi)$ and $\vec{g} \in C_c^\infty\left(T^*X, \omega_n^{-1}\mu_\varphi\right)$ we have the well known fact (see below for the proof)

$$\int_X \langle \mathcal{R}_\varphi f\left(x\right), \vec{g}\left(x\right)\rangle \, d\mu_\varphi\left(x\right) = 4\int_0^\infty \int_X \left\langle dP_t f\left(x\right), \frac{d}{dt}\vec{P}_t\vec{g}\left(x\right)\right\rangle d\mu_\varphi\left(x\right) tdt. \tag{4.3}$$

We assume for the moment that claim (4.3) is true. Corollary 4.1 follows from the well-known formula

$$\begin{aligned}
\|\mathcal{R}_\varphi f\|_{L^2(\omega_n\mu_\varphi)} &= \sup_{\|\vec{g}\|_{L^2(\omega_n^{-1}\mu_\varphi)}\le 1} |\langle \mathcal{R}_\varphi f, \vec{g}\rangle_{L^2(\mu_\varphi)}| \\
&= 4 \sup_{\|\vec{g}\|_{L^2(\omega_n^{-1}\mu_\varphi)}\le 1} \left| \int_0^\infty \left\langle dP_t f, \frac{d}{dt}\vec{P}_t\vec{g}\right\rangle_{L^2(\mu_\varphi)} tdt \right| \\
&= 4 \sup_{\|\vec{g}\|_{L^2(\omega_n^{-1}\mu_\varphi)}\le 1} \left| \int_0^\infty \int_X \left\langle dP_t f(x), \frac{d}{dt}\vec{P}_t\vec{g}(x)\right\rangle d\mu_\varphi(x) tdt \right| \\
&\le 4 \sup_{\|\vec{g}\|_{L^2(\omega_n^{-1}\mu_\varphi)}\le 1} \int_0^\infty \int_X |\nabla P_t f(x)||\vec{\nabla}\vec{P}_t\vec{g}(x)| d\mu_\varphi(x) tdt \\
&\le 884\widetilde{Q}_2(\omega_n)\|f\|_{L^2(X,\omega_n\mu_\varphi)},
\end{aligned}$$

by Theorem 4.1 and because $\|\vec{g}\|_{L^2(\omega_n^{-1}\mu_\varphi)} \le 1$.
To prove the claim (4.3), consider the function

$$h\left(t\right) = \left\langle \vec{P}_t\mathcal{R}_\varphi f, \vec{P}_t\vec{g}\right\rangle_{L^2(\mu_\varphi)}.$$

Since $\langle \mathcal{R}_\varphi f, \vec{g}\rangle_{L^2(\mu_\varphi)} = h\left(0\right)$, it suffices to show that

$$h\left(0\right) = \int_0^\infty h''(t) tdt = -4\int_0^\infty \left\langle dP_t f, \frac{d}{dt}\vec{P}_t\vec{g}\right\rangle_{L^2(\mu_\varphi)} tdt \tag{4.4}$$

In order to prove the first equality in (4.4), it is enough to show that both $h\left(t\right)$ and $th'\left(t\right)$ tend to zero as $t \to \infty$. First, note that by Lemma 4.1, $\vec{P}_t\mathcal{R}_\varphi f = \mathcal{R}_\varphi P_t f$. Therefore, by the L^2 contractivity of both $\mathcal{R}_\varphi$ and $\vec{P}_t$,

$$|h\left(t\right)| \le \|P_t f\|_{L^2(\omega_n\mu_\varphi)} \|\vec{g}\|_{L^2\left(\omega_n^{-1}\mu_\varphi\right)}. \tag{4.5}$$

Since $f \in R\left(-\Delta_\varphi\right)$, the spectral theorem gives that $P_t f \to 0$ in $L^2\left(\omega_n\mu_\varphi\right)$ as $t \to \infty$. Similarly, Lemma 4.1 gives

$$\begin{aligned}
h'(t) &= 2\langle -(\vec{\Delta}_\varphi)^{1/2}\vec{P}_t d(-\Delta_\varphi)^{-1/2} f, \vec{P}_t\vec{g}\rangle_{L^2(\mu_\varphi)} \\
&= -2\left\langle P_t f, P_t d_\varphi^*\vec{g}\right\rangle_{L^2(\mu_\varphi)},
\end{aligned}$$

therefore $\lim_{t\to+\infty} t\,|h'(t)| = 0$ as before.
The second equality in (4.4) can be verified by a straightforward calculation, again with the help of Lemma 4.1. Indeed,

$$h''(t) = -2\left(\langle \frac{d}{dt}P_tf, P_td^*_\varphi\vec{g}\rangle_{L^2(\mu_\varphi)} + \langle P_tf, \frac{d}{dt}P_td^*_\varphi\vec{g}\rangle_{L^2(\mu_\varphi)}\right).$$

By Lemma 4.1,

$$\langle \frac{d}{dt}P_tf, P_td^*_\varphi\vec{g}\rangle_{L^2(\mu_\varphi)} = \langle dP_tf, \frac{d}{dt}\vec{P}_t\vec{g}\rangle_{L^2(\mu_\varphi)},$$

and

$$\begin{aligned}\langle P_tf, \frac{d}{dt}P_td^*_\varphi\vec{g}\rangle_{L^2(\mu_\varphi)} &= \langle P_tf, d^*_\varphi\frac{d}{dt}\vec{P}_t\vec{g}\rangle_{L^2(\mu_\varphi)}\\ &= \langle dP_tf, \frac{d}{dt}\vec{P}_t\vec{g}\rangle_{L^2(\mu_\varphi)}.\end{aligned}$$

Thus, we get the desired result. □

4.3 The Bellman function

As mentioned before, the main tool used to prove Theorem 4.1 is a particular Bellman function that is constructed explicitly. A substantial part of its origin lies in the seminal paper by Nazarov, Treil and Volberg [61]. It has been developed in [82], [70] and [22], with the first explicit expression in [22]. Our construction differs from the one in [22] in that this construction is slightly shorter and and gives an explicit numerical constants. Another difference is that we restrict ourselves to infinitesimal convexity estimates.
In fact, for any $Q \geq 1$, we show that we can exhibit a function B_Q in domain

$$D_Q = \{\mathcal{X} := (X, Y, \mathrm{x}, \mathrm{y}, r, s) : \mathrm{x}^2 \leq Xr, \langle \mathrm{y}, \mathrm{y}\rangle \leq Ys, 1 \leq rs \leq Q\},$$

which is a subset of $\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}_+ \times \mathbb{R}_+$. The function B_Q is globally in C^1 and piecewise in C^2 such that

$$0 \leq B_Q \leq 884(X+Y); \tag{4.6}$$

$$-d^2B_Q \geq \frac{4}{Q}|d\mathrm{x}||d\mathrm{y}|;\ \text{where } B_Q \text{ is in } C^2. \tag{4.7}$$

Furthermore, B_Q is radial in x and y in the sense that

$$B_Q(X, Y, \mathrm{x}, \mathrm{y}, r, s) = \overline{B}_Q(X, Y, |\mathrm{x}|, |\mathrm{y}|, r, s.)$$

Consequently, the domain of $\overline{B}_Q$ is defined accordingly in $\mathbb{R}^6$. Writing $\nu = |\mathrm{y}| \in \mathbb{R}_+$ we have in addition

$$\partial_\nu \overline{B}_Q \leq 0. \tag{4.8}$$

Remark 4.3. *We use a Bellman function involving real variables. As a consequence, Theorem 4.1 and Corollary 4.1 hold for real-valued functions and forms. One may find the corresponding estimates for complex-valued functions and forms by separating the real and imaginary parts.*

Remark 4.4. *The property (4.7) means that for all 6-tuple $\mathcal{X} = (X, Y, \mathrm{x}, \mathrm{y}, r, s)$ in D_Q, we have*

$$\left\langle -d^2 B_Q d\mathcal{X}, d\mathcal{X} \right\rangle \geq \frac{4}{Q} |d\mathrm{x}||d\mathrm{y}|.$$

The strategy used to build this Bellman function relies on a careful analysis of the previous subsection. Hence, we consider the following Bellman functions $B_Q = B_1 + B_2 + B_3 + B_4$ and $\overline{B}_Q = \overline{B}_1 + \overline{B}_2 + \overline{B}_3 + \overline{B}_4$ of six variables such that

- $B_1(Z, H, \mathrm{x}, \mathrm{y}, r, s) = X - \dfrac{\mathrm{x}^2}{r} + Y - \dfrac{\langle \mathrm{y}, \mathrm{y} \rangle}{s}$,
- $B_2(X, Y, \mathrm{x}, \mathrm{y}, r, s) = X - \dfrac{\mathrm{x}^2}{r} + Y - \dfrac{\langle \mathrm{y}, \mathrm{y} \rangle}{s + \dfrac{M(r,s)}{Q^2}}$,
- $B_3(X, Y, \mathrm{x}, \mathrm{y}, r, s) = X - \dfrac{\mathrm{x}^2}{r + \dfrac{N(r,s)}{Q^2}} + Y - \dfrac{\langle \mathrm{y}, \mathrm{y} \rangle}{s}$,

  where
  $$M(r, s) = -\frac{4Q^2}{r} - rs^2 + (4Q^2 + 1)s$$
  and
  $$N(r, s) = -\frac{4Q^2}{s} - sr^2 + (4Q^2 + 1)r.$$

- $B_4 = B_{41} + B_{42} + B_{43}$ with
  - $B_{41}(X, Y, \mathrm{x}, \mathrm{y}, r, s) = X - \dfrac{\mathrm{x}^2}{r + \dfrac{\widetilde{M}(r,s)}{Q}} + Y - \dfrac{\langle \mathrm{y}, \mathrm{y} \rangle}{s}$
  - $B_{42}(X, Y, \mathrm{x}, \mathrm{y}, r, s) = X - \dfrac{\mathrm{x}^2}{r} + Y - \dfrac{\langle \mathrm{y}, \mathrm{y} \rangle}{s + \dfrac{\widetilde{N}(r,s)}{Q}}$
  - $B_{43}(X, Y, \mathrm{x}, \mathrm{y}, r, s) = \sup_{a>0} \left(X - \dfrac{\mathrm{x}^2}{r + a\dfrac{K(r,s)}{Q}} + Y - \dfrac{\langle \mathrm{y}, \mathrm{y} \rangle}{s + a^{-1}\dfrac{K(r,s)}{Q}} \right)$

    where
    $K(r, s) = \sqrt{Q}\sqrt{rs} - \dfrac{rs}{4}$,

$$\widetilde{M}(r,s) = -\frac{4Q}{s} - \frac{r^2 s}{4Q} + (4Q+1)r$$
and
$$\widetilde{N}(r,s) = -\frac{4Q}{r} - \frac{s^2 r}{4Q} + (4Q+1)s.$$

Explaining the construction of $K, \widetilde{M}$ and $\widetilde{N}$: We know from Remark 3.3 that $-d^2 B_{4i} \gtrsim |\frac{xy}{Q_2 rs} drds|$, for $i = 1, \cdots, 3$.
We start by B_{43}. The appropriate function K should satisfy by Theorem 3.3 the following size property $0 \leq K \leq Q_2$ and $-d^2 K \gtrsim |dsdr|$. The second restriction comes from $\Delta K = \frac{\alpha_J}{|J|}$, where $\alpha_J = (\langle w \rangle_{J_l} - \langle w \rangle_{J_r})(\langle w^{-1} \rangle_{J_l} - \langle w^{-1} \rangle_{J_r})\frac{|J|}{4}$. By Lemma 3.4, we know that the function
$$K(r,s) = \sqrt{Qrs} - \frac{rs}{4}$$
is appropriate.
Concerning B_{41}, we know it should be of the form
$$X - \frac{x^2}{r + \frac{\widetilde{M}(r,s)}{Q_2}} + Y - \frac{y^2}{s}.$$

It is now clear that it is concave. We have by chain rule
$$-d^2 B_{41} \geq \frac{\partial S}{\partial \widetilde{M}} (-d^2 \widetilde{M}),$$
where S is a function of X, Y, x, y, r, s and $\widetilde{M}(r,s)$.
However,
$$\frac{\partial S}{\partial \widetilde{M}} = \frac{x^2}{Q_2 \left(r + \frac{\widetilde{M}(r,s)}{Q_2} \right)^2} \geq \frac{x^2}{C Q_2 r^2},$$
for $0 \leq \widetilde{M} \leq C Q_2 r$. We use the fact that $1 \geq \frac{K}{Q_2} \geq \frac{yr}{2x}$ as follows
$$\begin{aligned} -d^2 B_{41} &\gtrsim \frac{x^2}{r^2}(-d^2 \widetilde{M}) \\ &\geq \frac{x^2}{Q_2 r^2}(-d^2 \widetilde{M}) \frac{yr}{2x} \\ &\gtrsim \frac{xy}{Q_2 r}(-d^2 \widetilde{M}). \end{aligned}$$

Hence, $-d^2 \widetilde{M} \gtrsim \frac{|dsdr|}{s}$. The function
$$\widetilde{M}(r,s) = \frac{-4Q}{s} - \frac{r^2 s}{4Q} + (4Q+1)r,$$

defined in Lemma 3.3 satisfies the hypothesis.
By symmetry,

$$B_{42}(X,Y,x,y,r,s) = X - \frac{x^2}{r} + Y - \frac{y^2}{s + \dfrac{\widetilde{N}(r,s)}{Q_2}},$$

where

$$\widetilde{N}(r,s) = \frac{-4Q}{r} - \frac{s^2 r}{4Q} + (4Q+1)s.$$

Remark 4.5. *1. In [22], instead of defining functions $\widetilde{M}$ and $\widetilde{N}$, the authors only need to define function K. B_{41} is then of the form*

$$X - \frac{x^2}{r + \left(r - \dfrac{1}{s\left(\dfrac{K(r,s)}{Q_2}+1\right)}\right)} + Y - \frac{y^2}{s}.$$

B_{42} is defined analogously and B_{43} remains the same. The calculation of the Hessian of B_{41} is done in two steps. Define

$$M(r,s) = r - \frac{1}{s\left(\frac{K(r,s)}{Q_2}+1\right)}, \quad F(x,y,r,s,M) = \frac{-x^2}{r+M} - \frac{y^2}{s}$$

$$\text{and } G(r,s,K) = r - \frac{1}{s(\frac{K}{Q_2}+1)}.$$

Notice that F and G are concave. Indeed,

$$-d^2G = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \dfrac{2}{s^3\left(\dfrac{K}{Q_2}+1\right)} & \dfrac{1}{Q_2 s^2\left(\dfrac{K}{Q_2}+1\right)^2} \\ 0 & \dfrac{1}{Q_2 s^2\left(\dfrac{K}{Q_2}+1\right)^2} & \dfrac{2}{Q_2^2 s\left(\dfrac{K}{Q_2}+1\right)^3} \end{pmatrix}$$

and

$$-d^2F = \begin{pmatrix} \dfrac{2}{r+M} & \dfrac{-2x}{(r+M)^2} & \dfrac{-2x}{(r+M)^2} & 0 & 0 \\ \dfrac{-2x}{(r+M)^2} & \dfrac{2x^2}{(r+M)^3} & \dfrac{2x^2}{(r+M)^3} & 0 & 0 \\ \dfrac{-2x}{(r+M)^2} & \dfrac{2x^2}{(r+M)^3} & \dfrac{2x^2}{(r+M)^3} & 0 & 0 \\ 0 & 0 & 0 & \dfrac{2}{s} & \dfrac{2y}{s^2} \\ 0 & 0 & 0 & \dfrac{2y}{s^2} & \dfrac{2y^2}{s^3} \end{pmatrix}$$

have non-negative principal minors. Hence they are positive semi-definite. We have then by chain rules

$$\begin{aligned}-d^2B_{41} &= -d^2F + \partial_M F(-d^2M) \\ &\geq \partial_M F(-d^2M)\end{aligned}$$

and

$$\begin{aligned}-d^2M &= -d^2G + \partial_K G(-d^2K) \\ &\geq \partial_K G(-d^2K).\end{aligned}$$

Thus

$$\begin{aligned}-d^2B_{41} &\geq \partial_M F \partial_K G(-d^2K) \\ &\geq \frac{x^2}{4r^2}\frac{1}{4Q_2 s}\frac{1}{4}|dsdr| \\ &\geq \frac{1}{128Q_2}\frac{xy}{rs}|dsdr|,\end{aligned}$$

where in the last inequality we used the fact that $1 \geq \dfrac{K}{Q_2} \geq \dfrac{yr}{2x}$. *The advantage of this method is that the assumptions are now easier to verify, compared to Theorem 3.3.*

We refer the reader to the previous chapter for the proofs of properties on size and derivative estimates of the functions $M, N, K, \widetilde{M}$ and $\widetilde{N}$. We recall the main results for the sake of completeness.

Functions M and N:

$$0 \leq M \leq 5Q^2 s \text{ and } -d^2M \geq r(ds)^2,$$

$$0 \leq N \leq 5Q^2 r \text{ and } -d^2N \geq s(dr)^2.$$

Function K:

$$0 \leq K \leq Q \text{ and } -d^2K \geq \frac{1}{4}|dsdr|.$$

Functions $\widetilde{M}$ and $\widetilde{N}$:

$$0 \leq \widetilde{M} \leq 5Qr \text{ and } -d^2\widetilde{M} \geq \frac{|dsdr|}{s},$$

$$0 \leq \widetilde{N} \leq 5Qs \text{ and } -d^2\widetilde{N} \geq \frac{|dsdr|}{r}.$$

Now, define $\Pi = \{\frac{K}{Q} = \frac{\mathrm{x}s}{|\mathrm{y}|}\} \cup \{\frac{K}{Q} = \frac{|\mathrm{y}|r}{\mathrm{x}}\}$. The function B_Q satisfies the following :

Lemma 4.2. *For every $(X, Y, \mathrm{x}, \mathrm{y}, r, s)$ in D_Q,*

1) $0 \leq B_Q \leq 6(X+Y)$;

2) $\partial_\nu \overline{B}_{Q,} \leq 0$;

3) *Function B_Q is globally C^1. Moreover, if $(X, Y, \mathrm{x}, \mathrm{y}, r, s) \in D_Q \setminus \Pi$, then it is C^2 everywhere except on the set Π and we have $-d^2 B_Q \geq \frac{4}{Q}|d\mathrm{x}||d\mathrm{y}|$.*

The proof of this lemma was more or less already proved in the case of martingale transforms but we recall for the sake of completeness, by calculation explicitly the numerical constants.

Proof. 1. The result follows directly due to the construction of B_Q as well as the hypothesis on D_Q.

2. For more convenience, we will deal with each function $\overline{B}_i, i = 1, \ldots, 4$ separately. It is clear that the derivatives in the variable ν of $\overline{B}_1$, $\overline{B}_2$, $\overline{B}_3$, $\overline{B}_{41}$ and $\overline{B}_{42}$ are negative by straightforward computations. It remains to study $\overline{B}_{43}$. Let us rewrite it in the form

$$\overline{B}_{43}(X, Y, |\mathrm{x}|, |\mathrm{y}|, r, s) = X + Y - \sup_{a>0} \beta(a, X, Y, |\mathrm{x}|, |\mathrm{y}|, r, s),$$

with

$$\beta(a, X, Y, |\mathrm{x}|, |\mathrm{y}|, r, s) = \frac{\mathrm{x}^2}{r + aK(r,s)/Q} + \frac{\nu^2}{s + a^{-1}K(r,s)/Q}.$$

The function β is continuously differentiable in $a > 0$ and

$$\frac{\partial \beta}{\partial a} = -\frac{\mathrm{x}^2 K/Q}{(r + aK/Q)^2} + \frac{\nu^2 K/Q}{(as + K/Q)^2},$$

which yields to

$$\frac{\partial \beta}{\partial a} = 0 \Leftrightarrow a = \frac{Qr\nu - K\mathrm{x}}{Qs\mathrm{x} - K\nu},$$

provided this fraction is finite and non-null.

Let $a_m := \frac{Qr\nu - K\mathrm{x}}{Qs\mathrm{x} - K\nu}$. If both numerator and denominator are positive, $\partial\beta/\partial a$ changes sign from positive to negative. Which means that the extremum is a maximum and it is attained at a_m. In this case,

$$\overline{B}_{43} = X - \frac{\mathrm{x}^2}{r + \dfrac{K(r,s)}{Q} \dfrac{Qr\nu - \mathrm{x}K(r,s)}{Qs\mathrm{x} - \nu K(r,s)}} + Y - \frac{\nu^2}{s + \dfrac{K(r,s)}{Q} \dfrac{Qs\mathrm{x} - \nu K(r,s)}{Qr\nu - \mathrm{x}K(r,s)}}.$$

If a_m is null or infinite, then $\overline{B}_{43}$ is $X + Y - \dfrac{\mathrm{x}^2}{r}$ and $X + Y - \dfrac{\nu^2}{s}$, respectively.

To compute $\partial_\nu \overline{B}_{43}$, consider a one-parameter family of functions

$$\overline{B}^a_{43}(X, Y, |\mathrm{x}|, |\mathrm{y}|, r, s) = X + Y - \beta(a, X, Y, |\mathrm{x}|, |\mathrm{y}|, r, s).$$

When a_m is strictly positive and finite, it is clear that $\overline{B}_{43} = \overline{B}_{43}^{a_m}$.
By chain rule

$$\frac{\partial \overline{B}_{43}}{\partial \nu} = \left.\frac{\partial \overline{B}_{43}^{a}}{\partial a}\right|_{a=a_m} \cdot \frac{\partial a}{\partial \nu} + \left.\frac{\partial \overline{B}_{43}^{a}}{\partial \nu}\right|_{a=a_m}.$$

But $\left.\frac{\partial B_{43}^{a}}{\partial a}\right|_{a=a_m} = 0$, since β attains its maximum at a_m. Consequently

$$\begin{aligned}
\frac{\partial \overline{B}_{43}}{\partial \nu} &= \left.\frac{\partial \overline{B}_{43}^{a}}{\partial \nu}\right|_{a=a_m} \\
&= -\frac{2\nu}{s + a_m^{-1}K/Q} \\
&\leq 0.
\end{aligned}$$

When a_m is null or infinite, then $\frac{\partial \overline{B}_{43}}{\partial \nu}$ is null or equal to $-\frac{2\nu}{s} \leq 0$, respectively (see [22] for more details on the behavior of B_{43}), which finishes the proof of (4.8).

3. To prove this assertion, we refer the reader to the paper of Petermichl and Domelevo [22, Lemma 3], where the authors have checked that the first partial derivatives of B_{43} are continuous throughout three regions, namely R_1 where $|\mathrm{y}|r - \mathrm{x}\frac{K}{Q} > 0$ and $\mathrm{x}s - |\mathrm{y}|\frac{K}{Q} > 0$, R_2 where $|\mathrm{y}|r - \mathrm{x}\frac{K}{Q} > 0$ and $\mathrm{x}s - |\mathrm{y}|\frac{K}{Q} \leq 0$ and R_3 where $\mathrm{x}s - |\mathrm{y}|\frac{K}{Q} > 0$ and $|\mathrm{y}|r - \mathrm{x}\frac{K}{Q} \leq 0$.
Now, to prove the concavity property, we will study as before the Hessian of each of B_1, B_2, B_3 and B_4 separately and then sum the results to obtain the desired estimate.
Case of B_1: A direct computation of the Hessian yields

$$\begin{aligned}
-d^2 B_1 &= \frac{2\mathrm{x}^2}{r}\left|\frac{d\mathrm{x}}{\mathrm{x}} - \frac{dr}{r}\right|^2 + \frac{2}{s}\langle d\mathrm{y} - \frac{\mathrm{y}}{s}ds, d\mathrm{y} - \frac{\mathrm{y}}{s}ds\rangle \\
&\geq \frac{4\mathrm{x}}{\sqrt{rs}}\left|\frac{d\mathrm{x}}{\mathrm{x}} - \frac{dr}{r}\right|\left|d\mathrm{y} - \frac{\mathrm{y}}{s}ds\right| \\
&\geq \frac{4}{Q}\left|d\mathrm{x} - \frac{\mathrm{x}}{r}dr\right|\left|d\mathrm{y} - \frac{\mathrm{y}}{s}ds\right|.
\end{aligned}$$

Case of B_2 and B_3: We can deduce from the Hessian of B_2 that of B_3 simply by replacing the variables x by y and r by s. As in the previous case, the Hessian of the first part of B_2 is bounded from below by $\frac{2\mathrm{x}^2}{r}\left|\frac{d\mathrm{x}}{\mathrm{x}} - \frac{dr}{r}\right|^2$. As for the second part, we use the following lemma:

Lemma 4.3. *If a function B has the form*

$$B(F,f,w,M) = F - \frac{f^2}{w+M}$$

and M is a function depending on the variables w and v, then if we write $H = B \circ M$, we have

$$-d^2 H \geq -\frac{\partial B}{\partial M} d^2 M.$$

Proof. By the chain rule, we compute the Hessian of H and get

$$\begin{aligned} \left\langle -d^2 H(df,dw,dv), \begin{pmatrix} df \\ dw \\ dv \end{pmatrix} \right\rangle &= \left\langle -d^2 B\left(df, d_w M + d_v M\right), \begin{pmatrix} df \\ d_w M + d_v M \end{pmatrix} \right\rangle \\ &\quad + \frac{\partial B}{\partial M} \left\langle -d^2 M(dw,dv), \begin{pmatrix} dw \\ dv \end{pmatrix} \right\rangle, \end{aligned}$$

where $d_w M + d_v M = \frac{\partial M}{\partial w} dw + \frac{\partial M}{\partial v} dv$. But as seen previously, B is concave, that is $-d^2 B \geq 0$. Thus, we can drop the first term on the right side of the equality and the lemma is proved. $\square$

Consequently, the Hessian of the second part of B_2 is bounded from below by $\dfrac{\langle \mathrm{y},\mathrm{y}\rangle}{36Q^2 s^2} r(ds)^2$.
Finally,

$$\begin{aligned} -d^2 B_2 &\geq \frac{2\mathrm{x}^2}{r} \left| \frac{d\mathrm{x}}{\mathrm{x}} - \frac{dr}{r} \right|^2 + \frac{\langle \mathrm{y},\mathrm{y}\rangle}{36Q^2 s^2} r(ds)^2 \\ &\geq \frac{\sqrt{2}}{3} \frac{|\mathrm{y}| r}{Qrs} |ds| \left| d\mathrm{x} - \frac{\mathrm{x}}{r} dr \right|. \end{aligned}$$

Analogously,

$$-d^2 B_3 \geq \frac{\sqrt{2}}{3} \frac{\mathrm{x}s}{Qrs} |dr| \left| d\mathrm{y} - \frac{\mathrm{y}}{s} ds \right|.$$

Case of B_4: Once again, we will study separately B_{41}, B_{42} and B_{43}.

It can be easily shown by using Lemma 4.3 that functions B_{41} and B_{42} are concave everywhere on the domain since $\dfrac{\partial B_{41}}{\partial \widetilde{M}}, \dfrac{\partial B_{42}}{\partial \widetilde{N}} \geq 0$ and $\widetilde{M}$, $\widetilde{N}$ are both concave. Function B_{43} is concave as well as the infimum of a family of concave functions, since β is convex for all $a > 0$. B_4 is hence concave as the sum of these functions. Moreover, these functions have an additional "super concavity" property on more restricted domains. These domains, which we already defined in the proof of

Lemma 4.2, appear naturally by looking at different values of $a_m = \frac{Qr\nu - K\mathrm{x}}{Qs\mathrm{x} - K\nu}$. We first study B_{43}. Recall that

$$B_{43}(X,Y,\mathrm{x},\mathrm{y},r,s) = \sup_{a>0}\left(X - \frac{\mathrm{x}^2}{r + a\dfrac{K(r,s)}{Q}} + Y - \frac{\langle \mathrm{y},\mathrm{y}\rangle}{s + a^{-1}\dfrac{K(r,s)}{Q}}\right),$$

and that the supremum is attained at $a = a_m > 0$, when both its numerator and denominator are positive. In this case, K is relatively small and we have (omitting variables X and Y)

$$\begin{aligned}
B_{43}(X,Y,\mathrm{x},\mathrm{y},r,s) &= -\frac{\mathrm{x}^2}{r + a_m\dfrac{K}{Q}} - \frac{\langle \mathrm{y},\mathrm{y}\rangle}{s + a_m^{-1}\dfrac{K}{Q}} \\
&= -\frac{\mathrm{x}^2(Qs\mathrm{x} - K|\mathrm{y}|)}{r(Qs\mathrm{x} - K|\mathrm{y}|) + (Qr|\mathrm{y}| - K\mathrm{x})\dfrac{K}{Q}} \\
&\quad -\frac{\langle \mathrm{y},\mathrm{y}\rangle(Qr|\mathrm{y}| - K\mathrm{x})}{s(Qr|\mathrm{y}| - K\mathrm{x}) + (Qs\mathrm{x} - K|\mathrm{y}|)\dfrac{K}{Q}} \\
&= -\frac{sQ^2\mathrm{x}^2 - 2QK\mathrm{x}|\mathrm{y}| + rQ^2\langle \mathrm{y},\mathrm{y}\rangle}{Q^2sr - K^2}. \qquad (4.9)
\end{aligned}$$

When a_m is null or infinite, then K is big, meaning that $K \geq \frac{Qr|\mathrm{y}|}{\mathrm{x}}$ or $K \geq \frac{Qs\mathrm{x}}{|\mathrm{y}|}$, and the supremum is attained at the boundary. Note that we can't have these two inequalities at once, unless $\mathrm{x}, |\mathrm{y}| = 0$. When K is big, B_{43} is respectively $-\frac{\mathrm{x}^2}{r}$ or $-\frac{\langle \mathrm{y},\mathrm{y}\rangle}{s}$.

Now if we restrict ourselves to the set R_1' where $K \leq \frac{Qr|\mathrm{y}|}{2\mathrm{x}}$ and $K \leq \frac{Qs\mathrm{x}}{2|\mathrm{y}|}$, which is contained in the set R_1 where $K \leq \frac{Qr|\mathrm{y}|}{\mathrm{x}}$ and $K \leq \frac{Qs\mathrm{x}}{|\mathrm{y}|}$, then it was shown by [22] that $\dfrac{\partial B_{43}}{\partial K} \gtrsim \dfrac{\mathrm{x}|\mathrm{y}|}{rs}$. Indeed by (4.9),

$$\frac{\partial B_{43}}{\partial K} = 2Q\frac{(Q\mathrm{x}s - K|\mathrm{y}|)(Q|\mathrm{y}|r - K\mathrm{x})}{(Q^2sr - K^2)^2}$$

We multiply by $\mathrm{x}|\mathrm{y}|$ both the numerator and denominator and then expand the denominator to obtain

$$\frac{\partial B_{43}}{\partial K} = 2Q\frac{(Q\mathrm{x}s - K|\mathrm{y}|)(Q|\mathrm{y}|r - K\mathrm{x})\mathrm{x}|\mathrm{y}|}{\left[Qr(Q\mathrm{x}s - K|\mathrm{y}|) + K(Q|\mathrm{y}|r - K\mathrm{x})\right]\left[Qs(Q|\mathrm{y}|r - K\mathrm{x}) + K(Q\mathrm{x}s - K|\mathrm{y}|)\right]}.$$

We need to find a constant c such that $\dfrac{\partial B_{43}}{\partial K} \geq \dfrac{c}{Q}\dfrac{\mathrm{x}|\mathrm{y}|}{rs}$ on R_1'. Meaning that

$$\begin{aligned}
\frac{2Q^2rs}{c} &\geq \frac{Qr(Q\mathrm{x}s - K|\mathrm{y}|) + K(Q|\mathrm{y}|r - K\mathrm{x})}{Q|\mathrm{y}|r - K\mathrm{x}} \times \frac{Qs(Q|\mathrm{y}|r - K\mathrm{x}) + K(Q\mathrm{x}s - K|\mathrm{y}|)}{Q\mathrm{x}s - K|\mathrm{y}|} \\
&= Q^2rs + K^2 + QKr\frac{Q\mathrm{x}s - K|\mathrm{y}|}{Q|\mathrm{y}|r - K\mathrm{x}} + QKs\frac{Q|\mathrm{y}|r - K\mathrm{x}}{Q\mathrm{x}s - K|\mathrm{y}|}.
\end{aligned}$$

We know that $K^2 \leq Q^2 \leq Q^2rs$ and on R_1', $Q|\mathrm{y}|r - \mathrm{x}K \geq \mathrm{x}K$ and $Q\mathrm{x}s - |\mathrm{y}|K \geq |\mathrm{y}|K$. Thus

$$\begin{aligned}
QKr\frac{Q\mathrm{x}s - K|\mathrm{y}|}{Q|\mathrm{y}|r - K\mathrm{x}} &\leq Qr\frac{Q\mathrm{x}s - K|\mathrm{y}|}{\mathrm{x}}, \\
QKs\frac{Q|\mathrm{y}|r - K\mathrm{x}}{Q\mathrm{x}s - K|\mathrm{y}|} &\leq Qs\frac{Q|\mathrm{y}|r - K\mathrm{x}}{|\mathrm{y}|}
\end{aligned}$$

and the sum of the two terms on the left is then less than $2Q^2rs$. By combining all the results we get

$$4Q^2rs \leq \frac{2}{c}Q^2rs,$$

so $c = 1/2$ works. Since $-d^2K \geq \dfrac{1}{4}|dsdr|$, this means that in R_1' we have

$$-d^2B_{43} \geq \frac{\mathrm{x}|\mathrm{y}|}{8Qrs}|dsdr|.$$

Functions B_{41} and B_{42} are introduced to deal with the concavity for other $K's$. In fact, we only need to study one of them, say B_{41}, as the result is the same for the other function by symmetry of the variables. The idea is to apply chain rule and Lemma 4.3, knowing that $K \geq \frac{Qr|\mathrm{y}|}{2\mathrm{x}}$.
More precisely, let

$$H(\mathrm{x},\mathrm{y},r,s) = S(\mathrm{x},\mathrm{y},r,s,\widetilde{M}(r,s)) = \frac{-\mathrm{x}^2}{r + \dfrac{\widetilde{M}(r,s)}{Q}} - \frac{\langle \mathrm{y},\mathrm{y}\rangle}{s}.$$

Notice that again, we omit the variables Z and H because they do not play a role for the Hessian.
One checks by calculating the Hessian of S that $-S$ is convex, which means that $-d^2S \geq 0$. By introducing $d_r\widetilde{M} = \dfrac{\partial \widetilde{M}}{\partial r}dr$, $d_s\widetilde{M} = \dfrac{\partial \widetilde{M}}{\partial s}ds$ and applying chain rule

we obtain

$$\begin{aligned}
&\langle -d^2 H(\mathrm{dx}, \mathrm{dy}, dr, ds), (\mathrm{dx}, \mathrm{dy}, dr, ds)\rangle = \\
&\langle -d^2 S(\mathrm{dx}, \mathrm{dy}, dr, ds, d_r\widetilde{M} + d_s\widetilde{M}), (\mathrm{dx}, \mathrm{dy}, dr, ds, d_r\widetilde{M} + d_s\widetilde{M})\rangle \\
&-\frac{\partial S}{\partial \widetilde{M}}\langle d^2\widetilde{M}(dr, ds), (dr, ds)\rangle \\
&\geq -\frac{\partial S}{\partial \widetilde{M}}\langle d^2\widetilde{M}(dr, ds), (dr, ds)\rangle \\
&\geq \frac{1}{72Q}\frac{\mathrm{x}|\mathrm{y}|}{rs}|dsdr|,
\end{aligned}$$

because $\dfrac{\partial S}{\partial \widetilde{M}} = \dfrac{\mathrm{x}^2}{Q\left(r + \frac{\widetilde{M}}{Q}\right)^2} \geq \dfrac{\mathrm{x}^2}{36Qr^2}$ and $-d^2\widetilde{M} \geq \dfrac{|dsdr|}{s}$. Moreover, we use the fact that $1 \geq \dfrac{K}{Q} \geq \dfrac{|\mathrm{y}|r}{2\mathrm{x}}$.

Finally, we obtain the following estimates

$$-d^2 B_{41} \geq \frac{1}{72}\frac{\mathrm{x}|\mathrm{y}|}{Qrs}|dsdr|, \quad \text{ when } K \geq \frac{Qr|\mathrm{y}|}{2\mathrm{x}} \quad \text{and}$$

$$-d^2 B_{42} \geq \frac{1}{72}\frac{\mathrm{x}|\mathrm{y}|}{Qrs}|dsdr|, \quad \text{ when } K \geq \frac{Qs\mathrm{x}}{2|\mathrm{y}|}.$$

Thanks to the global concavity of B_4 and to the more refined estimates of each part of it on three complementary sub-domains we have

$$-d^2 B_4 \geq \frac{1}{72}\frac{\mathrm{x}|\mathrm{y}|}{Qrs}|dsdr|,$$

on the regions where B_{43} is C^2.

Conclusion: In order to finish the proof, it suffices to choose constants $C_1 = 1$, $C_2 = C_3 = 6\sqrt{2}$ and $C_4 = 288$. The constant 884 appearing in Theorem 4.1 is due to the fact that we want

$$-d^2 B = -\sum_{i=1}^{4} c_i d^2 B_i \geq \frac{4}{Q_2}|dxdy|.$$

Recall that we have

$$-d^2 B_1 \geq \frac{4}{Q_2}|dx - \frac{x}{r}dr||dy - \frac{y}{s}ds|, \quad -d^2 B_2 \geq \frac{\sqrt{2}}{3Q_2}|dx - \frac{x}{r}dr|\frac{y}{s}|ds|,$$

$$-d^2 B_3 \geq \frac{\sqrt{2}}{3Q_2}|dy - \frac{y}{s}ds|\frac{x}{r}|dr| \text{ and}$$

$$-d^2B_4 \geq \begin{cases} \dfrac{1}{8Q_2}\dfrac{xy}{rs}|drds|, \text{ when } K \in R_1' \\ \dfrac{1}{72Q_2}\dfrac{xy}{rs}|drds|, \text{ otherwise.} \end{cases}$$

It means that whether K is in R_1' or not, we have

$$-d^2B_4 \geq \frac{1}{72Q_2}\frac{xy}{rs}|drds|.$$

Hence if we take $C_1 = 1$, $C_2 = C_3 = 6\sqrt{2}$ and $C_4 = 288$, we obtain

$$-\sum_{i=1}^{4} C_i d^2 B_i \geq \frac{4}{Q_2}|dxdy|,$$

which yields to the following size property

$$0 \leq B = \sum_{i=1}^{4} C_i B_i \leq 884(X+Y),$$

as claimed.

□

As mentioned earlier, B_Q fails to be C^2 everywhere. We can add smoothness by taking convolutions with mollifiers: for a fixed compact $\mathcal{K}$ in the interior of $\overline{D}_Q$, choose $\varepsilon > 0$ such that $\varepsilon < dist(\mathcal{K}, \partial\overline{D}_Q)$. Consider $\overline{B}_{\varepsilon,Q}(\overline{\mathcal{X}}) = \overline{B}_Q * \frac{1}{\varepsilon^6}\psi(\frac{\overline{\mathcal{X}}}{\varepsilon})$, where ψ is a bell-shaped infinitely differentiable function with compact support in the unit ball of $\mathbb{R}^6$.
The resulting functions $B_{\varepsilon,Q}$ and $\overline{B}_{\varepsilon,Q}$ are clearly smooth in a small neighborhoord of $\mathcal{K}$ and satisfy the following properties

1') $0 \leq B_{\varepsilon,Q} \leq 884(X+Y+2\varepsilon)$;

2') $\partial_\nu \overline{B}_{\varepsilon,Q} \leq 0$;

3') $-d^2 B_{\varepsilon,Q} \geq \frac{4}{Q}|d\mathrm{x}||d\mathrm{y}|$.

Proof. 1') Recall that the original function satisfies $0 \leq B_Q \leq 6(X+Y)$. Thus, it is easy to see that size property of the new weighted and mollified function $B_{\varepsilon,Q}$ changes only by a factor depending on the distance between the compact $\mathcal{K}$ and ∂D_Q, as well as the sum of weights C_i, $i \in \{1, \cdots, 4\}$.

2') The non-positivity of $\partial_\nu \overline{B}_{\varepsilon,Q}$ is preserved because the function $\overline{B}_Q$ is globally C^1 and satisfies

$$\partial_\nu \overline{B}_{\varepsilon,Q}(\overline{\mathcal{X}}) = \partial_\nu \overline{B}_Q * \psi_\varepsilon(\overline{\mathcal{X}}),$$

for $\overline{\mathcal{X}} = (X, Y, |\mathrm{x}|, \nu, r, s)$.
Since $\partial_\nu \widetilde{B}_Q$ is negative and ψ_ε is positive, we obtain the result.

3') The statement is true because B_Q is C^1 and we integrate over a compact set. Moreover, the second order derivatives exist almost everywhere (because Π is of measure zero) and are locally integrable, which means they coincide with the second order distributional derivatives. One can find more about this procedure in several previous texts on Bellman functions.

$\square$

Remark 4.6. *Concerning the domain of $B_{\varepsilon,Q}$, its construction is standard (see for example [22]) and requires some technicalities which we summarize as follows: we truncate our variables from below at level $\varepsilon > 0$ and from above at level ε^{-1}. cf. for instance Lemma 4.4. Next we define the mollified function $B_{\varepsilon,Q}$ as above. All the inequalities coming from size property are still verified, even if that means expanding the domain by a constant depending on ε.*

4.4 Proof of the main result

This section is devoted to the proof of Theorem 4.1.
First of all, let's fix $(x,t) \in X \times \mathbb{R}_+$ and define $\widetilde{B}_{\varepsilon,Q} : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \times T^*X \times \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\widetilde{B}_{\varepsilon,Q}(X, Y, \mathrm{x}, \mathrm{y}, r, s) = B_{\varepsilon,Q}(X, Y, |\mathrm{x}|, |\mathrm{y}|_{T_x^*X}, r, s).$$

Let Q be in what follows the Poisson characteristic i. e.

$$Q = \widetilde{Q}_2(\omega) := \sup_{(x,t)\in X\times\mathbb{R}_+} P_t(\omega)(x)P_t(\omega^{-1})(x) < \infty.$$

Let us also define for a certain $t_0 > 0$ small enough the vector

$$\begin{aligned} &v(x, t + t_0) \\ &= \Big(P_{t+t_0}|f|^2\omega_n(x), P_{t+t_0}|\vec{g}|^2_{T_x^*X}\omega_n^{-1}(x), P_{t+t_0}f(x), \vec{P}_{t+t_0}\vec{g}(x), P_{t+t_0}\omega_n^{-1}(x), P_{t+t_0}\omega_n(x)\Big)\end{aligned}$$

and in parallel

$$\begin{aligned} &\overline{v}(x, t + t_0) \\ &= \Big(P_{t+t_0}\left(|f|^2\omega_n\right)(x), P_{t+t_0}\left(|\vec{g}|^2_{T_x^*X}\omega_n^{-1}\right)(x), P_{t+t_0}f(x), |\vec{P}_{t+t_0}\vec{g}(x)|_{T_x^*X}, P_{t+t_0}\omega_n^{-1}(x), P_{t+t_0}\omega_n(x)\Big)\end{aligned}$$

where we recall that P_t and $\vec{P}_t$ stand for the weighted Poisson extensions. It is important to mention that $v(x, t + t_0) \in D_Q$ and $(x,t) \mapsto v(x,t)$ maps compacts in $X \times \mathbb{R}_+$ to compacts in D_Q. Indeed, the following inequalities

$$|P_{t+t_0}f|^2 \le P_{t+t_0}\left(|f|^2\omega_n\right)P_{t+t_0}\omega_n^{-1} \quad \text{and} \quad |\vec{P}_{t+t_0}\vec{g}|^2_{T_x^*X} \le P_{t+t_0}\left(|\vec{g}|^2_{T_x^*X}\omega_n^{-1}\right)P_{t+t_0}\omega_n$$

are true by Lemma 4.1. It is also clear that $P_{t+t_0}\omega_n(x)P_{t+t_0}\omega_n^{-1}(x) \le Q$ by the very definition of Q. It remains to show that it is greater than 1.

Since the semigroup P_t is Markovian, $P_t 1 = 1$ by [11]. Thus

$$\begin{aligned} 1 &= \int_X p_{t+t_0}^{1/2}(x,y)\omega_n^{1/2}(y)p_{t+t_0}^{1/2}(x,y)\omega_n^{-1/2}(y)d\mu_\varphi(y) \\ &\leq (P_{t+t_0}\omega_n(x))^{1/2} \times \left(P_{t+t_0}\omega_n^{-1}(x)\right)^{1/2}. \end{aligned}$$

Besides, the mapping property holds because v is a continuous function.
Next, define

$$b_\varepsilon(x,t+t_0) = \widetilde{B}_{\varepsilon,Q}(v(x,t+t_0))$$

and consider the operators

$$\Delta_{\varphi,t} = \partial_{tt}^2 + \Delta_\varphi \text{ and } \vec{\Delta}_{\varphi,t} = \partial_{tt}^2 + \vec{\Delta_\varphi}.$$

The fact that $\widetilde{B}_{\varepsilon,Q}$ is radial allows us to define the Bellman function on manifolds. Our goal is to find a link between $\Delta_{\varphi,t}b_\varepsilon$ and $d^2\widetilde{B}_{\varepsilon,Q}$ and then estimate the integral

$$-\int\int \Delta_{\varphi,t}b_\varepsilon(x,t+t_0)d\mu_\varphi(x)tdt$$

from below and above.

4.4.1 Estimate from below

Proposition 4.1. *Suppose that* $\mathrm{Ric}_\varphi \geq 0$. *Then for all* $(x,t) \in X \times \mathbb{R}_+$,

$$-\Delta_{\varphi,t}b_\varepsilon(x,t+t_0) \geq \frac{4}{Q}|\overline{\nabla}P_{t+t_0}f(x)||\overline{\nabla}\vec{P}_{t+t_0}\vec{g}(x)|.$$

Proof. Following [16, Lemma 12] and [16, Proposition 13] we use the function $B_{\varepsilon,Q}$ and the corresponding function $\overline{B}_{\varepsilon,Q}$ to define the quantity

$$\begin{aligned} F(x,t+t_0) &= -\Delta_{\varphi,t}\widetilde{B}_{\varepsilon,Q}(v(x,t+t_0)) \\ &\quad -\frac{\partial_\nu \overline{B}_{\varepsilon,Q}(\overline{v}(x,t+t_0))}{|\vec{P}_t\vec{g}|_{T_x^*X}} \times (-\mathrm{Ric}_\varphi(\sharp\vec{P}_t\vec{g},\sharp\vec{P}_t\vec{g})), \end{aligned}$$

where $\sharp : T_x^*X \to T_xX$ is the sharp musical isomorphism.
Note that we have a different sign convention for $\Delta_{\varphi,t}$ and use concavity instead of convexity. The calculations used to compute F are omitted since they follow exactly the same steps as in [16], that is, computing $\Delta_{\varphi,t}\widetilde{B}_{\varepsilon,Q}$ and writing F in terms of its variables and their different derivatives by using the Bochner formula in [12, eq. (0.3)]. We obtain at the end

$$F(x,t+t_0) \geq \frac{4}{Q}|\overline{\nabla}P_{t+t_0}f(x)||\overline{\nabla}\vec{P}_{t+t_0}\vec{g}(x)|.$$

To verify this inequality, one can use exponential local coordinates and inequality (4.7), since the expression of F holds pointwise.
Furthermore, since $\partial_\nu\overline{\widetilde{B}}_{\varepsilon,Q} \leq 0$ and $\mathrm{Ric}_\varphi \geq 0$ we have

$$-\Delta_{\varphi,t}\widetilde{B}_{\varepsilon,Q}(v(x,t+t_0)) \geq F(x,t+t_0),$$

thus yielding the result. $\square$

4.4.2 Estimate from above

In order to prove Theorem 4.1, it suffices to show that

$$-\int\int \Delta_{\varphi,t} b_\varepsilon(x,t+t_0)d\mu_\varphi(x)tdt \leq C\|f\|_{L^2(X,\omega_n\mu_\varphi)}\|\vec{g}\|_{L^2(T^*X,\omega_n^{-1}\mu_\varphi)}.$$

In fact, for a fixed point $o \in X$, $l > 1$ and $s > 0$ such that $t_0 \in (0, 1/s)$, we have the following result:

Proposition 4.2. *Suppose that* $\mathrm{Ric}_\varphi \geq 0$. *Then for every* (x,t) *in the compact set* $K_{s,l} := \overline{B(o,l)} \times [1/s, s]$ *we have*

$$\overline{\lim_{s\to\infty}}\,\overline{\lim_{l\to\infty}}\int_{1/s}^{s}\int_{B(o,l)} -\Delta_{\varphi,t} b_\varepsilon(x,t+t_0)d\mu_\varphi(x)tdt \leq 884(1+\varepsilon)(\|f\|^2_{L^2(X,\omega_n\mu_\varphi)}+\|\vec{g}\|^2_{L^2(T^*X,\omega_n^{-1}\mu_\varphi)}).$$

Proof. Let $\rho(x,o)$ be the geodesic distance on X between o and x. Define $\Lambda \in C_c^\infty([0,\infty))$ be a decreasing function such that $0 \leq \Lambda \leq 1$, $\Lambda = 1$ in $[0,1]$ and $\Lambda = 0$ in $[2,\infty)$. We are interested in the following composite function :

$$\Lambda\left(\frac{\rho(x,o)^2}{l^2}\right), l > 1.$$

Observe that this composite function is always positive and equals to 1 when $\rho(x,o) < l$. Recall also that by Proposition 4.1, $-\Delta_{\varphi,t} b_\varepsilon \geq 0$ and so

$$\int_{1/s}^{s}\int_{B(o,l)} -\Delta_{\varphi,t} b_\varepsilon(x,t+t_0)d\mu_\varphi(x)tdt \leq \int_{1/s}^{s}\int_X -\Delta_{\varphi,t} b_\varepsilon(x,t+t_0)\Lambda\left(\frac{\rho(x,o)^2}{l^2}\right)d\mu_\varphi(x)tdt.$$

To prove the lemma, we shall show that

$$\begin{aligned} &\overline{\lim_{s\to\infty}}\,\overline{\lim_{l\to\infty}}\int_{1/s}^{s}\int_X -\partial^2_{tt} b_\varepsilon(x,t+t_0)\Lambda\left(\frac{\rho(x,o)^2}{l^2}\right)d\mu_\varphi(x)tdt \\ \leq\ & 884(1+\varepsilon)(\|f\|^2_{L^2(X,\omega_n\mu_\varphi)} + \|\vec{g}\|^2_{L^2(T^*X,\omega_n^{-1}\mu_\varphi)}), \end{aligned} \tag{4.10}$$

and

$$\lim_{l\to\infty}\int_{1/s}^{s}\int_X -\Delta_\varphi b_\varepsilon(x,t+t_0)\Lambda\left(\frac{\rho(x,o)^2}{l^2}\right)d\mu_\varphi(x)tdt = 0. \tag{4.11}$$

We first prove (4.10). An integration by parts in the variable t gives

$$\int_{1/s}^{s} -\partial^2_{tt} b_\varepsilon\left(x,t+t_0\right)tdt = \frac{1}{s}\partial_t b_\varepsilon\left(x,\frac{1}{s}+t_0\right) - s\partial_t b_\varepsilon\left(x,s+t_0\right) + b_\varepsilon\left(x,s+t_0\right) - b_\varepsilon\left(x,\frac{1}{s}+t_0\right).$$

The size property (4.6) implies

$$\begin{aligned} b_\varepsilon\left(x,s+t_0\right) - b_\varepsilon\left(x,\frac{1}{s}+t_0\right) &\leq b_\varepsilon\left(x,s+t_0\right) \\ &\leq 884(1+\varepsilon)\left(P_{s+t_0}(|f|^2\omega_n)(x) + P_{s+t_0}(|\vec{g}|^2_{T^*_xX}\,\omega_n^{-1})(x)\right). \end{aligned}$$

It follows by contractivity of the semigroup P_t on $L^r(\mu_\varphi)$ for every $r \in [1,+\infty]$ (read [79] as a reference) that

$$\begin{aligned}\int_{\frac{1}{s}}^{s}\int_X -\partial_{tt}^2 b_\varepsilon(x,t+t_0)\,\Lambda\left(\frac{\rho(x,o)^2}{l^2}\right)d\mu_\varphi(x)\,tdt \quad &\leq \quad 884(1+\varepsilon)\left(\|f\|^2_{L^2(X,\omega_n\mu_\varphi)}+\|\vec{g}\|^2_{L^2(T^*X,\omega_n^{-1}\mu_\varphi)}\right)\\ &\quad + \left\|s\partial_t b_\varepsilon(\cdot,s+t_0)-\frac{1}{s}\partial_t b_\varepsilon\left(\cdot,\frac{1}{s}+t_0\right)\right\|_{L^1(\mu_\varphi)},\end{aligned}$$

Therefore, in order to prove (4.10) it is enough to show that

$$\lim_{s\to+\infty}\|s\partial_t b_\varepsilon(\cdot,s+t_0)\|_{L^1(\mu_\varphi)}=0, \tag{4.12}$$

and

$$\lim_{s\to+\infty}\left\|\frac{1}{s}\partial_t b_\varepsilon\left(\cdot,\frac{1}{s}+t_0\right)\right\|_{L^1(\mu_\varphi)}=0, \tag{4.13}$$

By applying chain rule we obtain

$$\begin{aligned}\partial_t b_\varepsilon(x,s+t_0) \quad &= \quad \frac{\partial\widetilde{B}_{\varepsilon,Q}}{\partial X}(v)\partial_t P_{s+t_0}\left(|f|^2\omega_n\right)(x)+\frac{\partial\widetilde{B}_{\varepsilon,Q}}{\partial Y}(v)\partial_t P_{s+t_0}\left(|\vec{g}|^2_{T_x^*X}\omega_n^{-1}\right)(x)\\ &\quad +\frac{\partial\widetilde{B}_{\varepsilon,Q}}{\partial \mathrm{x}}(v)\partial_t P_{s+t_0}f(x)+\langle\frac{\partial\widetilde{B}_{\varepsilon,Q}}{\partial \mathrm{y}}(v),\partial_t\vec{P}_{s+t_0}\vec{g}(x)\rangle_{T_x^*X}\\ &\quad +\frac{\partial\widetilde{B}_{\varepsilon,Q}}{\partial r}(v)\partial_t P_{s+t_0}\omega_n(x)+\frac{\partial\widetilde{B}_{\varepsilon,Q}}{\partial s}(v)\partial_t P_{s+t_0}\omega_n^{-1}(x).\end{aligned}$$

Thus, using Hölder's inequality with some α and its conjugate exponent α' we obtain

$$\begin{aligned}\|s\partial_t b_\varepsilon(\cdot,s+t_0)\|_{L^1(\mu_\varphi)} \quad &\leq \quad \underbrace{\left\||\frac{\partial\widetilde{B}_{\varepsilon,Q}}{\partial X}(v)|+\cdots+|\frac{\partial\widetilde{B}_{\varepsilon,Q}}{\partial s}(v)|\right\|_{L^\alpha(\mu_\varphi)}}_{\Sigma_1}\\ &\quad \times\underbrace{\left\|s\left(|\partial_t P_{s+t_0}\left(|f|^2\omega_n\right)(\cdot)|+\cdots+|\partial_t P_{s+t_0}\omega_n^{-1}(\cdot)|\right)\right\|_{L^{\alpha'}(\mu_\varphi)}}_{\Sigma_2}.\end{aligned} \tag{4.14}$$

Let's study Σ_1. First of all, by triangle inequality, one sees that Σ_1 can be majorized by the sum of norms of each partial derivative of $\widetilde{B}_{\varepsilon,Q}$. These partial derivatives show terms in P_t and $\vec{P}_t$. We will then need to estimate the obtained norms uniformly in $t>0$ so that we can send s to $+\infty$ in (4.12). This can be done by using Hölder's inequality and contractivity of both P_t and $\vec{P}_t$.
For instance, we already know from Section 3 (proof of Lemma 4.2, 2)) that when a_m is strictly finite and positive,

$$\langle\frac{\partial B_{Q,\mathcal{K}}}{\partial \mathrm{y}},d\mathrm{y}\rangle=-\frac{6\langle \mathrm{y},d\mathrm{y}\rangle}{s}-\frac{2\langle \mathrm{y},d\mathrm{y}\rangle}{s+\frac{M}{Q^2}}-\frac{2\langle \mathrm{y},d\mathrm{y}\rangle}{s+\frac{\widetilde{N}}{Q}}-\frac{2\langle \mathrm{y},d\mathrm{y}\rangle}{s+a_m^{-1}\frac{K}{Q}},\ \text{for all } d\mathrm{y},$$

which implies that

$$|\frac{\partial B_{Q,\mathcal{K}}}{\partial \mathrm{y}}| \leq |\frac{6\mathrm{y}}{s}| + |\frac{2\mathrm{y}}{s+\frac{M}{Q^2}}| + |\frac{2\mathrm{y}}{s+\frac{\widetilde{N}}{Q}}| + |\frac{2\mathrm{y}}{s+a_m^{-1}\frac{K}{Q}}|.$$

Recall that $1 \leq rs$ and that M, $\widetilde{N}$ and $a_m^{-1}\frac{K}{Q}$ are positive. Thus, we have $|\frac{\partial B_{Q,\mathcal{K}}}{\partial \mathrm{y}}| \leq 12|\mathrm{y}|r$. Since $B_{Q,\mathcal{K}}$ is C^1, we can also dominate $|\frac{\partial B_{\varepsilon,Q}}{\partial \mathrm{y}}|$, but with a constant depending on ε. Meaning that there exists a constant $C_\varepsilon > 0$ such that

$$|\frac{\partial B_{\varepsilon,Q}}{\partial \mathrm{y}}(\mathcal{X})| \leq C_\varepsilon |\mathrm{y}| r.$$

Finally, by replacing the variable $\mathcal{X}$ by $v(x, t+t_0)$ we obtain

$$\left\| \frac{\partial \widetilde{B}_{\varepsilon,Q}}{\partial \mathrm{y}} \right\|_{L^\alpha(\mu_\varphi)} \leq C_\varepsilon \left\| P_{t+t_0}\omega_n^{-1} P_{t+t_0} |\vec{g}|_{T_x^* X} \right\|_{L^\alpha(\mu_\varphi)},$$

and by symmetry,

$$\left\| \frac{\partial \widetilde{B}_{\varepsilon,Q}}{\partial \mathrm{x}} \right\|_{L^\alpha(\mu_\varphi)} \leq C_\varepsilon \left\| P_{t+t_0}\omega_n P_{t+t_0} f \right\|_{L^\alpha(\mu_\varphi)}.$$

The result is the same up to a constant when a_m is null or infinite. Using the same arguments as above, we can dominate the first partial derivatives in the other variables. Indeed,

$$\left\| \frac{\partial \widetilde{B}_{\varepsilon,Q}}{\partial X} \right\|_{L^\alpha(\mu_\varphi)} = \left\| \frac{\partial \widetilde{B}_{\varepsilon,Q}}{\partial Y} \right\|_{L^\alpha(\mu_\varphi)} \leq C_\varepsilon.$$

Moreover, if a_m is finite and positive, we have

$$\begin{aligned}
\frac{\partial B_{Q,\mathcal{K}}}{\partial r} \quad = \quad & \frac{3\mathrm{x}^2}{r^2} + \frac{Q^2\langle \mathrm{y},\mathrm{y}\rangle \dfrac{\partial M(r,s)}{\partial r}}{(Q^2 s + M(r,s))^2} + \frac{Q^2\mathrm{x}^2(Q^2 + \dfrac{\partial N(r,s)}{\partial r})}{(Q^2 r + N(r,s))^2} \\
& + \frac{Q\mathrm{x}^2(Q + \dfrac{\partial \widetilde{M}(r,s)}{\partial r})}{(Qr + \widetilde{M}(r,s))^2} + \frac{Q\langle \mathrm{y},\mathrm{y}\rangle \dfrac{\partial \widetilde{N}(r,s)}{\partial r}}{(Qs + \widetilde{N}(r,s))^2} \\
& + \left(\frac{\mathrm{x}^2(1 + \dfrac{a_m}{Q}\dfrac{\partial K(r,s)}{\partial r})}{(r + a_m \dfrac{K(r,s)}{Q})^2} + \frac{\langle \mathrm{y},\mathrm{y}\rangle \dfrac{a_m^{-1}}{Q}\dfrac{\partial K(r,s)}{\partial r}}{(s + a_m^{-1}\dfrac{K(r,s)}{Q})^2} \right),
\end{aligned}$$

where the last derivative between brackets represents $\partial_r B_{43}$ and has been calculated as $\partial_{\mathrm{y}} B_{43}$ in Section 3.

The partial derivatives in r of M, N, K, $\widetilde{M}$ and $\widetilde{N}$ are

$$\begin{aligned}
\frac{\partial M(r,s)}{\partial r} &= \frac{4Q^2}{r^2} - s^2,\ \frac{\partial N(r,s)}{\partial r} = -2sr + (4Q^2+1),\\
\frac{\partial K(r,s)}{\partial r} &= \frac{\sqrt{Q}}{2}\sqrt{\frac{s}{r}} - \frac{s}{4},\\
\frac{\partial \widetilde{M}(r,s)}{\partial r} &= -\frac{rs}{2Q} + (4Q+1) \text{ and } \frac{\partial \widetilde{N}(r,s)}{\partial r} = \frac{4Q}{r^2} - \frac{s^2}{4Q}.
\end{aligned}$$

Thus, using that $rs \geq 1$, we obtain

$$\begin{aligned}
\left|\frac{\partial B_{Q,\mathcal{K}}}{\partial r}\right| &\leq 3\mathrm{x}^2 s^2 + 4\langle \mathrm{y}, \mathrm{y}\rangle + 8\mathrm{x}^2 s^2 + 7\mathrm{x}^2 s^2 + 5\langle \mathrm{y}, \mathrm{y}\rangle\\
&\quad + \left(\underbrace{\frac{\mathrm{x}^2(1 + \frac{a_m}{Q}\frac{\partial K(r,s)}{\partial r})}{(r + a_m\frac{K(r,s)}{Q})^2}}_{A} + \underbrace{\frac{\langle \mathrm{y}, \mathrm{y}\rangle \frac{a_m^{-1}}{Q}\frac{\partial K(r,s)}{\partial r}}{(s + a_m^{-1}\frac{K(r,s)}{Q})^2}}_{B} \right).
\end{aligned}$$

The next step is to bound A and B from above. In fact, since $\frac{\partial K(r,s)}{\partial r} = \frac{\sqrt{Q}}{2}\sqrt{\frac{s}{r}} - \frac{s}{4}$, one can observe that $r\frac{\partial K(r,s)}{\partial r} \leq K(r,s)$. Using this observation, we bound A as follows

$$\begin{aligned}
A &\leq \frac{\mathrm{x}^2\left(1 + \frac{a_m}{Q}\frac{K(r,s)}{r}\right)}{r^2\left(1 + a_m\frac{K(r,s)}{Qr}\right)^2}\\
&\leq \frac{\mathrm{x}^2}{r^2\left(1 + a_m\frac{K(r,s)}{Qr}\right)}\\
&\leq \frac{\mathrm{x}^2}{r^2} \leq \mathrm{x}^2 s^2.
\end{aligned}$$

By the same strategy we bound B

$$\begin{aligned}
B &\leq \frac{\langle \mathrm{y}, \mathrm{y}\rangle \dfrac{a_m^{-1}}{Q} \dfrac{K(r,s)}{r}}{r^2 \left(\dfrac{s}{r} + a_m^{-1} \dfrac{K(r,s)}{Qr}\right)^2} \\
&\leq \frac{\langle \mathrm{y}, \mathrm{y}\rangle \left(\dfrac{a_m^{-1}}{Q} \dfrac{K(r,s)}{r} + \dfrac{s}{r}\right)}{r^2 \left(\dfrac{s}{r} + a_m^{-1} \dfrac{K(r,s)}{Qr}\right)^2} \\
&\leq \frac{\langle \mathrm{y}, \mathrm{y}\rangle}{r^2 \left(\dfrac{s}{r} + a_m^{-1} \dfrac{K(r,s)}{Qr}\right)} \\
&\leq \frac{\langle \mathrm{y}, \mathrm{y}\rangle}{s^2} \leq \langle \mathrm{y}, \mathrm{y}\rangle r^2,
\end{aligned}$$

and finally we obtain

$$\left|\frac{\partial B_{Q,\mathcal{K}}}{\partial r}\right| \leq C\left(\mathrm{x}^2 s^2 + \langle \mathrm{y}, \mathrm{y}\rangle r^2 + \langle \mathrm{y}, \mathrm{y}\rangle\right).$$

Therefore,

$$\begin{aligned}
\left\|\frac{\partial \widetilde{B}_{\varepsilon,Q}}{\partial r}\right\|_{L^\alpha(\mu_\varphi)} \leq\ & C_\varepsilon \Big(\|P_{t+t_0} f P_{t+t_0} \omega_n\|^2_{L^\alpha(\mu_\varphi)} + \\
& \left\|P_{t+t_0} |\vec{g}|_{T_x^* X} P_{t+t_0} \omega_n^{-1}\right\|^2_{L^\alpha(\mu_\varphi)} + \left\|\vec{P}_{t+t_0} \vec{g}\right\|^2_{L^\alpha(\mu_\varphi)} \Big)
\end{aligned}$$

and again by symmetry,

$$\begin{aligned}
\left\|\frac{\partial \widetilde{B}_{\varepsilon,Q}}{\partial s}\right\|_{L^\alpha(\mu_\varphi)} \leq\ & C_\varepsilon \Big(\left\|P_{t+t_0} |\vec{g}|_{T_x^* X} P_{t+t_0} \omega_n^{-1}\right\|^2_{L^\alpha(\mu_\varphi)} + \\
& \|P_{t+t_0} f P_{t+t_0} \omega_n\|^2_{L^\alpha(\mu_\varphi)} + \|P_{t+t_0} f\|^2_{L^\alpha(\mu_\varphi)} \Big).
\end{aligned}$$

Now, if a_m is null or infinite, then $\dfrac{\partial B_{43}}{\partial r}$ is either $-\dfrac{|\mathrm{x}|^2}{r^2}$ or 0. We repeat the previous calculations and obtain the same results up to a constant.
As said before, we now need to estimate from above these norms for each $i = 1, \ldots, 6$ uniformly in $t > 0$. To do so, we use Hölder's inequality and contractivity of both P_t and $\vec{P}_t$ in $L^r(\mu_\varphi)$ for all $r \in [1, +\infty]$. In other terms, we have shown that Σ_1 appearing in (4.14) can be majorized in the following way

$$\Sigma_1 \leq C\left(\varepsilon, f, \vec{g}, \omega_n, \omega_n^{-1}\right)$$

uniformly in $t > 0$.

As for part Σ_2, all we have to do is to show that the quantity

$$\left\| s \left(\left| \partial_t P_{s+t_0} \left(|f|^2 \omega_n \right) \right| + \left| \partial_t P_{s+t_0} \left(|\vec{g}|^2_{T^*_x X} \omega_n^{-1} \right) \right| + \cdots + \left| \partial_t P_{s+t_0} \omega_n^{-1} \right| \right) \right\|_{L^{\alpha'}(\mu_\varphi)}$$

tends to 0 as s tends to infinity. In fact, simply observe that by the Hilbert space spectral representation theory, each of $s\partial_t P_{s+t_0} f$, $s\partial_t \vec{P}_{s+t_0} \vec{g}$, $s\partial_t P_{s+t_0} \left(|f|^2 \omega_n \right)$, $s\partial_t P_{s+t_0} \left(|\vec{g}|^2_{T^*_x X} \omega_n^{-1} \right)$, $s\partial_t P_{s+t_0} \omega_n$ and $s\partial_t P_{s+t_0} \omega_n^{-1}$ converge to 0 in $L^2(\mu_\varphi)$ as s tends to infinity, because functions $P_{s+t_0} f$, $\vec{P}_{s+t_0} \vec{g}$, $P_{s+t_0} \omega_n$ and $P_{s+t_0} \omega_n^{-1}$ are square integrable.
To conclude, notice that $\| t\partial_t P_t \|_{L^r(X,\mu_\varphi)} + \left\| t\partial_t \vec{P}_t \right\|_{L^r(X,\mu_\varphi)}$ is uniformly bounded in t for all r in $(1, \infty)$ [30, Theorem 4.6 (c)], because P_t and $\vec{P}_t$ are symmetric Markov semigroups respectively on $L^2 (X, \mu_\varphi)$ and $L^2 (T^*X, \mu_\varphi)$ and thus extend to bounded holomorphic semigroups respectively on $L^r (X, \mu_\varphi)$ and $L^r (T^*X, \mu_\varphi)$, for all r in $(1, \infty)$ [20, Theorem 1.4.2].
We follow the same procedure to prove (4.13).

We now prove (4.11). By integrating by parts twice, we have

$$- \int_{\frac{1}{s}}^{s} \int_X \Delta_\varphi b_\varepsilon (x, t + t_0) \Lambda \left(\frac{\rho(x, o)^2}{l^2} \right) d\mu_\varphi (x) \, t dt$$

$$= - \int_{\frac{1}{s}}^{s} \int_X b_\varepsilon (x, t + t_0) \Delta_\varphi \Lambda \left(\frac{\rho(x, o)^2}{l^2} \right) d\mu_\varphi (x) \, t dt.$$

A simple computation based on [12, p 140] gives

$$\begin{aligned} -\Delta_\varphi \Lambda \left(\frac{\rho(x, o)^2}{l^2} \right) &= -\frac{2}{l^2} \left(|d\rho(x, o)|^2 + \rho(x, o) \Delta_\varphi \rho(x, o) \right) \Lambda' \left(\frac{\rho^2(x, o)}{l^2} \right) \\ &\quad - \frac{4\rho^2(x, o)}{l^4} |d\rho(x, o)|^2 \Lambda'' \left(\frac{\rho^2(x, o)}{l^2} \right), \end{aligned}$$

for all $x \in X \backslash (\mathrm{cut}(o) \cup \{o\})$, where $\mathrm{cut}(o)$ denotes the cut locus of the point o. In addition, since $\mathrm{Ric}_\varphi \geq 0$, by [81, Theorem 1.1] we have the local comparison result

$$\Delta_\varphi \rho(x, o) \leq C \frac{1}{\rho(x, o)}, \tag{4.15}$$

for all $x \in X \backslash (\mathrm{cut}(o) \cup \{o\})$.
Since $\|d\rho\|_\infty \leq 1$, and $\mathrm{supp}\, \Lambda$ is in $[0, 2]$, by (4.15) there exists $C > 0$ such that

$$-\Delta_\varphi \Lambda \left(\frac{\rho(x, o)^2}{l^2} \right) \geq -C \left(\|\Lambda'\|_\infty + \|\Lambda''\|_\infty \right) \chi_{B(o,2l) \backslash B(o,l)}, \tag{4.16}$$

for all $x \in X \backslash (\mathrm{cut}(o) \cup \{o\})$ and provided $l \geq 1$. Moreover, (4.16) holds weakly on X and in particular, we have

$$\int_X -\Delta_\varphi b_\varepsilon (x, t+t_0) \Lambda \left(\frac{\rho(x,o)^2}{l^2} \right) d\mu_\varphi (x) \geq -C \int_{B(o,2l) \backslash B(o,l)} b_\varepsilon (x, t+t_0) \, d\mu_\varphi (x) .$$

Hence, the size property (4.6) implies that

$$\begin{aligned} &\int_X -\Delta_\varphi b_\varepsilon (x, t+t_0) \Lambda \left(\frac{\rho(x,o)^2}{l^2} \right) d\mu_\varphi (x) \\ &\geq -C \int_{B(o,2l) \backslash B(o,l)} \left(P_{t+t_0} \left(|f|^2 \omega_n \right) (x) + P_{t+t_0} \left(|\vec{g}|^2_{T_x^* X} \, \omega_n^{-1} \right) (x) \right) d\mu_\varphi (x). \end{aligned}$$

Denote the integrand on the right-hand side by $\Psi_l (t)$. Since $\lim_{l \to +\infty} \Psi_l = 0$ pointwise on $\mathbb{R}_+$ and $0 \leq \Psi_l (t) \leq \|f\|^2_{L^2(X, \omega_n \mu_\varphi)} + \|\vec{g}\|^2_{L^2(T^* X, \omega_n^{-1} \mu_\varphi)}$, the Lebesgue dominated convergence theorem implies

$$\liminf_{l \to +\infty} \int_{\frac{1}{s}}^s \int_X -\Delta_\varphi b_\varepsilon (x, t+t_0) \Lambda \left(\frac{\rho(x,o)^2}{l^2} \right) d\mu_\varphi (x) \geq 0. \tag{4.17}$$

It remains to prove that

$$\limsup_{l \to +\infty} \int_{\frac{1}{s}}^s \int_X -\Delta_\varphi b_\varepsilon (x, t+t_0) \Lambda \left(\frac{\rho(x,o)^2}{l^2} \right) d\mu_\varphi (x) \, t dt \leq 0.$$

Consider the function

$$R(x, t+t_0) = 884 \, (1+\varepsilon) \left(P_{t+t_0} \left(|f|^2 \omega_n \right) (x) + P_{t+t_0} \left(|\vec{g}|^2_{T_x^* X} \, \omega_n^{-1} \right) (x) \right).$$

We have $b_\varepsilon - R \leq 0$ on $K_{s,l}$ and by an argument similar to the one we used to prove (4.17) one shows that

$$\limsup_{l \to +\infty} \int_{\frac{1}{s}}^s \int_X -\Delta_\varphi \left(b_\varepsilon (x, t+t_0) - R(x, t+t_0) \right) \Lambda \left(\frac{\rho(x,o)^2}{l^2} \right) d\mu_\varphi (x) \, t dt \leq 0.$$

It suffices then to prove that

$$\limsup_{l \to +\infty} \int_{1/s}^s \int_X \Delta_\varphi R \, (x, t+t_0) \Lambda \left(\frac{\rho(x,o)^2}{l^2} \right) d\mu_\varphi (x) \, t dt = 0, \tag{4.18}$$

using an integration by parts. To this purpose, notice first that the composite function $\Lambda \left(\frac{\rho^2}{l^2} \right)$ is equal to zero for $\rho \geq 2l$ and hence we have

$$\| d\Lambda \left(\frac{\rho^2}{l^2} \right) \|_\infty \leq \frac{4 \|\Lambda'\|_\infty}{l}.$$

Then,

$$\left|\int_{\frac{1}{s}}^{s}\int_X \Delta_\varphi R\left(x,t+t_0\right)\Lambda\left(\frac{\rho(x,o)^2}{l^2}\right)d\mu_\varphi\left(x\right)tdt\right| \leq \frac{4\left\|\Lambda'\right\|_\infty}{l}\int_{\frac{1}{s}}^{s}\int_X \left|dR\left(x,t+t_0\right)\right|d\mu_\varphi\left(x\right)tdt \tag{4.19}$$

Now, notice that by Lemma 4.1 and semigroup contractivity we have

$$\begin{aligned}
\int_X |dR(x,t+t_0)|d\mu_\varphi(x) &\leq C(\varepsilon)\Big(\int_X [|\vec{P}_t\left(dP_{t_0}(|f|^2\,\omega_n)\right)|(x) \\
&\quad + |\vec{P}_t\left(dP_{t_0}(|\vec{g}|^2_{T_x^*X}\,\omega_n^{-1})\right)|(x)d\mu_\varphi(x)]\Big) \\
&\leq C(\varepsilon)\int_X [|dP_{t_0}(|f|^2\,\omega_n)|(x)+|dP_{t_0}(|\vec{g}|^2_{T_x^*X}\,\omega_n^{-1})|(x)]\,d\mu_\varphi(x) \\
&\leq C(\varepsilon)\Bigg[\left(\int_X |dP_{t_0}(|f|^2\omega_n)|^2 d\mu_\varphi(x)\right)^{1/2}\times\left(\int_X d\mu_\varphi(x)\right)^{1/2} \\
&\quad + \left(\int_X |dP_{t_0}(|\vec{g}|^2_{T_x^*X}\omega_n^{-1})|^2 d\mu_\varphi(x)\right)^{1/2}\times\left(\int_X d\mu_\varphi(x)\right)^{1/2}\Bigg]
\end{aligned}$$

Now observe that by using (4.1) and because X is without a boundary we have

$$\begin{aligned}
\int_X |dP_{t_0}(|f|^2\omega_n)(x)|^2 d\mu_\varphi(x) &\leq \left(\int_X |P_{t_0}(|f|^2\omega_n)(x)|^2 d\mu_\varphi(x)\right)^{1/2} \\
&\quad \times\left(\int_X |\Delta_\varphi P_{t_0}(|f|^2\omega_n)(x)|^2 d\mu_\varphi(x)\right)^{1/2} \\
&= \|P_{t_0}(|f|^2\omega_n)\|_{L^2(X,\mu_\varphi)}\times\|\Delta_\varphi P_{t_0}(|f|^2\omega_n)\|_{L^2(X,\mu_\varphi)} \\
&\leq \|f\|_\infty^4\|\omega_n\|^2_{L^2(X,\mu_\varphi)}.
\end{aligned}$$

The last inequality holds since Δ_φ is self-adjoint on L^2 and hence admits bounded functional calculus. We obtain similar results fo the operator $\vec{\Delta}_\varphi$ and so

$$\int_X |dR(x,t)|d\mu_\varphi(x) \leq C(\varepsilon)\left(\|f\|_\infty^4\|\omega_n\|^2_{L^2(X,\mu_\varphi)}+\|\vec{g}\|_\infty^4\|\omega_n^{-1}\|^2_{L^2(X,\mu_\varphi)}\right)$$

As a consequence, the right-hand side integral of (4.19) is finite. Letting l tend to infinity implies (4.18) and concludes the proof of the proposition. □

4.4.3 Conclusion

Proof of Theorem 4.1. To finish the proof of the theorem, we use a standard trick. Indeed, by combining the reverse Fatou lemma as well as Propositions 4.1 and 4.2 and passing to the limit as ε tends to 0 we get

$$\int_0^\infty\int_X |\nabla P_t f(x)||\vec{\nabla}\vec{P}_t\vec{g}(x)|t\,d\mu_\varphi(x)dt \leq 221\widetilde{Q}_2(\omega_n)(\|f\|^2_{L^2(X,\omega_n\mu_\varphi)}+\|\vec{g}\|^2_{L^2(T^*X,\omega_n^{-1}\mu_\varphi)}).$$

We now apply the above inequality to λf and $\lambda^{-1}\vec{g}$ instead of f and $\vec{g}$ and then minimize the result in $\lambda > 0$. □

4.5 Enlarging the set of weights

Now that we have boundedness results for a certain type of weights in $L^2(X, \mu_\varphi)$, we will enlarge the set of weights ω satisfying Theorem 4.1 to all $L^1_{loc}(X, \mu_\varphi)$, provided that $\widetilde{Q}_2(\omega)$ is well defined and finite. We heavily use in the following proof the fact that constants are in $L^2(X, \mu_\varphi)$[1]. The main problem would be defining $P_t\omega$ when $\omega \in L^1_{loc}(X, \mu_\varphi)$. We will proceed as follows:
As stated in Remark 4.1, take $\omega \in L^1_{loc}(X, \mu_\varphi)$ and define its two-sided truncation

$$\omega_n = n^{-1}\chi_{\omega \leq n^{-1}} + \omega\chi_{n^{-1} \leq \omega \leq n} + n\chi_{\omega \geq n},$$

where χ is the characteristic function. Then we have the following properties

$$\omega_n \in L^2(X, \mu_\varphi); \tag{4.20}$$

$$\widetilde{Q}_2(\omega_n) \leq \widetilde{Q}_2(\omega); \tag{4.21}$$

$$\widetilde{Q}_2(\omega_n) \underset{n\to+\infty}{\longrightarrow} \widetilde{Q}_2(\omega). \tag{4.22}$$

This means that we can approximate a function in the class $\widetilde{A}_2$ by bounded functions from the same class, with control of their $\widetilde{A}_2$ constants.
Property (4.20) is immediate because constant functions are integrable with respect to our measure. Besides, $\omega_n \underset{n\to+\infty}{\longrightarrow} \omega$ and consequently, the definition of $P_t\omega$ arises naturally by posing $P_t\omega = \lim_{n\to\infty} P_t\omega_n$. This limit exists and makes sense because $P_t\omega \leq \underline{\lim}_{n\to\infty} P_t\omega_n$ by Fatou's lemma. Furthermore,

$$P_t\omega_n \leq \frac{1}{n} + P_t(\omega\chi_{n^{-1}\leq\omega\leq n} + n\chi_{\omega\geq n}).$$

The first term tends to zero as n tends to infinity and the term between brackets is increasing in n, which means that we can use the monotone convergence theorem to obtain

$$\begin{aligned} P_t\omega &\leq \underline{\lim}_{n\to\infty} P_t\omega_n \\ &\leq \overline{\lim}_{n\to\infty} P_t\omega_n \\ &\leq \overline{\lim}_{n\to\infty} \left(\frac{1}{n} + P_t(\omega\chi_{n^{-1}\leq\omega\leq n} + n\chi_{\omega\geq n})\right) \\ &= P_t\omega. \end{aligned}$$

We need the following preliminary lemma, where the weight is only cut from above, to prove properties (4.21) and (4.22):

Lemma 4.4. *Let $\omega^n = \omega\chi_{\omega\leq n} + n\chi_{\omega\geq n}$. Then we have $\widetilde{Q}_2(\omega^n) \leq \widetilde{Q}_2(\omega)$ and $\widetilde{Q}_2(\omega^n) \underset{n\to+\infty}{\longrightarrow} \widetilde{Q}_2(\omega)$.*

[1] This condition implies that $\mu_\varphi(X)$ is finite and therefore the kernel of $-\Delta_\varphi$ is the set of constant functions on X.

Proof. If $\omega^n = \omega\chi_{\omega\leq n} + n\chi_{\omega\geq n}$, then $(\omega^n)^{-1} = \omega^{-1}\chi_{\omega^{-1}\geq n^{-1}} + n^{-1}\chi_{\omega^{-1}\leq n^{-1}}$.
Thus, we need to prove that

$$P_t\omega P_t\omega^{-1} - P_t\omega^n P_t(\omega^n)^{-1} \geq 0.$$

Write $P_t\omega = P_t\left(\omega\chi_{\omega\leq n}\right) + P_t\left(\omega\chi_{\omega>n}\right)$ and denote

$$\begin{aligned}
A &= P_t\left(\omega\chi_{\omega\leq n}\right);\ A^{-1} = P_t\left(\omega^{-1}\chi_{\omega\leq n}\right);\\
B &= P_t\left(\omega\chi_{\omega>n}\right);\ B^{-1} = P_t\left(\omega^{-1}\chi_{\omega>n}\right);\\
B_n &= nP_t\left(\chi_{\omega>n}\right);\ B_n^{-1} = n^{-1}P_t\left(\chi_{\omega>n}\right);\\
\alpha_1 &= \int p_t(x,y)\chi_{\omega\leq n}(y)d\mu_\varphi(y);\\
\alpha_2 &= \int p_t(x,y)\chi_{\omega>n}(y)d\mu_\varphi(y).
\end{aligned}$$

so that

$$\begin{aligned}
P_t\omega P_t\omega^{-1} - P_t\omega^n P_t(\omega^n)^{-1} &= (A+B)(A^{-1}+B^{-1}) - (A+B_n)(A^{-1}+B_n^{-1})\\
&= A(B^{-1}-B_n^{-1}) + A^{-1}(B-B_n)\\
&\quad +(BB^{-1}-B_nB_n^{-1}).
\end{aligned}$$

Remark 4.7. *We stress that in the previous notations, A^{-1} and B^{-1} are not the inverses of A and B. Rather, it means that the weights we are considering are ω^{-1}.*

The last term between brackets is positive because $B_nB_n^{-1} = \alpha_2^2$ and

$$\alpha_2^2 \leq 1\times\alpha_2 \leq BB^{-1},$$

by Jensen's inequality. For the other terms, notice that

$$\begin{aligned}
A\left(B^{-1}-B_n^{-1}\right) &= \int A\left(p_t(x,y)\omega^{-1}(y)\chi_{\omega>n}(y)\right)d\mu_\varphi(y) - \int A\left(n^{-1}p_t(x,y)\chi_{\omega>n}(y)\right)d\mu_\varphi(y)\\
&= \int A\left(p_t(x,y)(n\omega^{-1}(y)n^{-1} - \omega(y)\omega^{-1}(y)n^{-1})\chi_{\omega>n}(y)\right)d\mu_\varphi(y)\\
&= \int A\left(p_t(x,y)(n-\omega(y))n^{-1}\omega^{-1}(y)\chi_{\omega>n}(y)\right)d\mu_\varphi(y)
\end{aligned}$$

and analogously,

$$A^{-1}\left(B-B_n\right) = \int A^{-1}\left(p_t(x,y)(\omega(y)-n)\chi_{\omega>n}(y)\right)d\mu_\varphi(y)$$

Hence,

$$P_t\omega P_t\omega^{-1} - P_t\omega^n P_t(\omega^n)^{-1} \geq \int p_t(x,y)\left(\frac{\omega(y)-n}{n\omega(y)}\left(n\omega(y)A^{-1}-A\right)\right)\chi_{\omega>n}(y)d\mu_\varphi(y)$$

The kernel p_t being positive, the integral on the right side is positive too, since $\omega > n$, $A \leq n\alpha_1$ and $A^{-1} \geq n^{-1}\alpha_1$.
Taking supremum over $(x,t) \in X\times\mathbb{R}_+$ on the left-hand side finishes the proof. □

The same results hold for the cutting from below $w_{(n)} = n^{-1}\chi_{\omega\leq n^{-1}} + \omega\chi_{n^{-1}<\omega}$ and (4.21) and (4.22) follow immediately by writing $w_n = (w^n)_{(1/n)}$.

As a consequence, if we let n tend to infinity in (4.2), it only remains to prove that

$$\|f\|_{L^2(X,\omega_n\mu_\varphi)} \underset{n\to+\infty}{\longrightarrow} \|f\|_{L^2(X,\omega\mu_\varphi)},$$

and

$$\|\vec{g}\|_{L^2(T_x^*X,\omega_n^{-1}\mu_\varphi)} \underset{n\to+\infty}{\longrightarrow} \|\vec{g}\|_{L^2(T_x^*X,\omega^{-1}\mu_\varphi)}.$$

To do that, we are going to use the dominated convergence theorem. Indeed, by construction of ω_n and ω_n^{-1}, we know that

$$\omega_n \leq \omega + 1 \text{ and } \omega_n^{-1} \leq \omega^{-1} + 1.$$

We can pass to the limit in n since $f \in L^2(X,\mu_\varphi)\cap L^2(X,\omega\mu_\varphi)$ and $\vec{g} \in L^2(T_x^*X,\omega^{-1}\mu_\varphi)\cap L^2(T_x^*X,\mu_\varphi)$.
We can also recover Corollary 4.1 by using Formula (4.3) and pass to the limit in n in (4.5), again by the dominated convergence theorem.

4.6 Case of the Gauss space

We now present a concrete example for the previous weighted estimates, namely the Gauss space, which is obtained when $X = \mathbb{R}^n$ and $\varphi(x) = \dfrac{\|x\|^2}{2}$. We then have the Gaussian measure

$$d\gamma(x) = \exp(-\frac{\|x\|^2}{2})dx$$

on $\mathbb{R}^n$ and the Ornstein-Uhlenbeck operator [2]

$$Lf(x) = \Delta f(x) - x\cdot\nabla f(x)$$

on $L^2(\mathbb{R}^n, d\gamma)$. This operator generates a diffusion semigroup P_t, which has been the object of many investigations during the last decades.
Note also that $\mathrm{Ric}_L \geq 0$.

If we define by $\mathcal{R}_L = d\circ(-L)^{-1/2}$ the Riesz transform associated to the Ornstein-Uhlenbeck semigroup, then

Corollary 4.2. *For all $f \in \overline{L^2(\mathbb{R}^n,\omega\gamma)\cap R(-L)}^{L^2}$ and $\omega,\omega^{-1} \in L^1_{loc}(\mathbb{R}^n,\gamma)$ such that $\omega > 0$ γ-a.e we have*

$$\|\mathcal{R}_L f\|_{L^2(\mathbb{R}^n,\omega\gamma)} \leq 884\widetilde{Q}_2(\omega)\|f\|_{L^2(\mathbb{R}^n,\omega\gamma)}.$$

[2] The Ornstein-Uhlenbeck operator is an integral operator that admits a kernel called the Mehler kernel, which can be given by an explicit representation.

4.7 Remark on the sharpness of the result

The Hilbert transform on the unit disk $\mathbb{T}$ with $\varphi = 0$ is a particular case of our main theorem, where 1 is integrable. Sharpness (in terms of the power of the characteristic) in this context, using the Poisson characteristic, was already stated in [70]. However, the argument was based on a reference citing the linear comparability of the classical and Poisson A_2 characteristics for power weights on the real line, which does not hold. Sharpness using Poisson characteristic was proved only recently in [23] on the real line, using probabilistic methods. The passage to the unit disk that is explained in [70] gives the sharpness of our result.

Chapter 5

Stochastic calculus on manifolds

The construction of Eells-Elworthy-Malliavin [27] of a Brownian motion on a manifold is perhaps the most elegant and satisfactory construction. Roughly speaking, it realizes the Brownian motion as a projection of the solution of the SDE on the orthonormal frame bundle $\mathcal{O}(X)$ over X. The advantage of this construction of Brownian motion is that it is intrinsic and it provides a path-wise construction obtained by solving a globally defined stochastic differential equation.
In this chapter, we are going to define the necessary background in differential geometry and then present the construction of the Brownian motion on manifolds [39]. Next, we move on to the heat equation on differential forms [43]. Finally, we introduce the Itô formula on 1-forms and as an application, we deduce a probabilistic representation of the Riesz transform [54] and as an application, the probabilistic representation of the Riesz transform associated to the Ornstein Uhlenbeck operator.

5.1 Orthonormal frame bundle and parallel transport

5.1.1 Parallel transport

Comparing different vectors at different positions on a manifold is quite challenging since tangent vectors point in different directions. Parallel transport provides a way to compare a vector in one tangent plane to a vector in another, by moving the vector along a curve without changing it. We refer to Section 2.7 for the definitions of geometric objects.

Definition 5.1. *Let X be equipped with the Levi-Civita connection. A vector field V[1] along a curve x_t on X is said to be parallel along $x_t : I \subset \mathbb{R} \longrightarrow X$ if $\nabla_{\dot{x}_t} V(x_t) = 0$ for every t in I. The vector V_{x_t} is then called the parallel transport of V_{x_0} along the curve and it is locally uniquely determined by V_{x_0}.*

One can show that [48, Proposition 3.3, page 71]

[1]It is possible to generalize this definition to a section V of E where $E \to X$ is a vector bundle.

Proposition 5.1. *Let x_t be a smooth curve on X. Then*

- *The parallel transport is independent of any specific parametrization of the curve x_t.*
- *The parallel transport along the reversed curve $y_t = x_{1-t}$ is the inverse of the parallel transport along x_t.*
- *The parallel transport along the composition of two curves x_t and y_t such that $x_1 = y_0$ is the composition of the corresponding parallel transports along x_t and y_t.*

5.1.2 Orthonormal frame bundle

The manifold X being Riemannian, the tangent space T_xX is endowed with a Euclidean structure given by the Riemannian metric g_x. Therefore, it is an inner product space. This means that there exist orthonormal frames on T_xX which are nothing more than orthonormal ordered basis for T_xX. Let $\mathcal{O}_x(X)$ be the set of orthonormal frames of the tangent space T_xX i.e. an ordered basis $u = (u_1, \cdots, u_n)$ of T_xX consisting of vectors which are orthonormal with respect to the bi-linear form g_x. We define the orthonormal frame bundle by

$$\mathcal{O}(X) = \bigcup_{x\in X} \mathcal{O}_x(X).$$

$\mathcal{O}(X)$ is a principal fibre bundle over X with the orthogonal group $O(n)$ and it is called the bundle of orthonormal frames on X. The group $O(n)$ acts on $\mathcal{O}(X)$ by the following right-action

$$\begin{array}{rlcc} \lhd g & : O(n) \times \mathcal{O}(X) & \longrightarrow & \mathcal{O}(X) \\ & (g,p) & \longmapsto & p \lhd g = p \cdot g \end{array}$$

We can think of this action as pushing frames along fibres. The bundle $\mathcal{O}(X)$ has a natural structure of a smooth manifold of dimension $\frac{n(n+1)}{2}$ since it is isomorphic to $\mathbb{R}^n \times O(n)$.

Let $\pi : \mathcal{O}(X) \to X$ be the canonical projection. Each $u \in \mathcal{O}(X)$ is an ordered orthonormal basis for T_xX, or, equivalently, a linear isometry

$$u : \mathbb{R}^n \to T_{\pi u}X,$$

such that $u(e_i) = u_i$, where $(e_i)_{1\leq i\leq n}$ denotes the standard basis on $\mathbb{R}^n$.
A tangent vector $V \in T_u\mathcal{O}(X)$ is called *vertical* if it is tangent to the fibre $\mathcal{O}_{\pi u}(X)$. The space of vertical vectors at u is denoted by $V_u\mathcal{O}(X) := T_u\mathcal{O}_{\pi u}(X)$. It is a sub-manifold of dimension $\frac{n(n-1)}{2}$ because it is of same dimension as the fibre $\mathcal{O}_{\pi u}(X)$, which is isomorphic to $\{\pi u\} \times O(n)$.
A curve u_t in $\mathcal{O}(X)$ is said to be *horizontal* if for each $e \in \mathbb{R}^n$, the vector field u_te is parallel along the projection curve πu_t.
A tangent vector $H \in T_u\mathcal{O}(X)$ is called *horizontal* if it is the tangent vector of a horizontal curve from u. The space of horizontal vectors at u is denoted by $H_u\mathcal{O}(X)$[2]. By

[2] Another equivalent definition of a curve u_t in $\mathcal{O}(X)$ to be horizontal is that $\dot{u}_t \in H\mathcal{O}(X)$.

the next decomposition (5.1), it is clear that $H_u\mathcal{O}(X)$ is a sub-manifold of dimension $\frac{n(n+1)}{2} - \frac{n(n-1)}{2} = n$.
We have the following direct sum decomposition

$$T_u\mathcal{O}(X) = H_u\mathcal{O}(X) \oplus V_u\mathcal{O}(X). \tag{5.1}$$

We see from this decomposition that the vertical bundle $V\mathcal{O}(X) = \bigcup_{u\in\mathcal{O}(X)} V_u\mathcal{O}(X)$ is uniquely defined while a horizontal bundle $H\mathcal{O}(X) = \bigcup_{u\in\mathcal{O}(X)} H_u\mathcal{O}(X)$ is a choice of a subspace of $T\mathcal{O}(X)$ such that we have (5.1). We emphasize that the use of the words "the" and "a" is crucial: while the vertical subspace is exclusively determined by fibration, there is an infinite number of horizontal subspaces to form the direct sum. The assignment of such horizontal spaces is called a *connection* on $\mathcal{O}(X)$:

Definition 5.2. *A connection in $\mathcal{O}(X)$ is a smoothly varying assignment to each point u in $\mathcal{O}(X)$ of a subspace $H_u\mathcal{O}(X)$ of $T_u\mathcal{O}(X)$ such that*

1. $T_u\mathcal{O}(X) = H_u\mathcal{O}(X) \oplus V_u\mathcal{O}(X), \quad \forall u \in \mathcal{O}(X)$.
2. $(\lhd g)_* H_u\mathcal{O}(X) = H_{u\lhd g}\mathcal{O}(X), \quad \forall u \in \mathcal{O}(X), \forall g \in O(n)$, *where $(\lhd g)_*$ is the push-forward of $\lhd g$.*

Remark 5.1. 1. *By "smoothly varying" we mean that if a vector field T is smooth, then so are its horizontal and vertical parts.*

2. *The second point of the definition implies that the decomposition (5.1) is compatible with the right action of $O(n)$ on $\mathcal{O}(X)$.*
3. *A connection can be associated with a certain one-form ω on $\mathcal{O}(X)$. This one-form allows to define alternatively a connection. We refer the reader to [48, Chapter2].*

The projection π induces an isomorphism $\pi_* : H_u\mathcal{O}(X) \to T_xX$[3]. Since π_* is an isomorphism, for each $T \in T_xX$ and a frame u at x, there is a unique horizontal vector T^* such that $\pi_*T^* = T$. It is called the horizontal lift of T from u. In particular for each $e \in \mathbb{R}^n$, we define a horizontal vector field H_e such that for every $u \in \mathcal{O}(X)$, $H_e(u)$ is the horizontal lift of $u(e)$ from u. More specifically, we define *the fundamental horizontal vector fields* by $H_i = H_{e_i}$, where $(e_1, \cdots, e_n)$ is the canonical basis of $\mathbb{R}^n$. We have

$$\pi_* H_i(u) = u(e_i), \quad H_i(u) \in H_u\mathcal{O}(X)$$

and

$$\begin{array}{rccc} H_i & : & \mathcal{O}(X) & \to & T\mathcal{O}(X) \\ & & u & \mapsto & H_i(u) \in T_u\mathcal{O}(X) \end{array}$$

which means that for each $i = 1, \cdots, n$, H_i is a vector field on $\mathcal{O}(X)$. We will sometimes use the equivalent definition of a vector field which is a linear map $H_i : C^\infty(\mathcal{O}(X)) \to C^\infty(\mathcal{O}(X))$ such that:

$$H_i(fg) = fH_i(g) + H_i(f)g,$$

[3]Another definition of the space of vertical vectors is that it is the kernel of π_*.

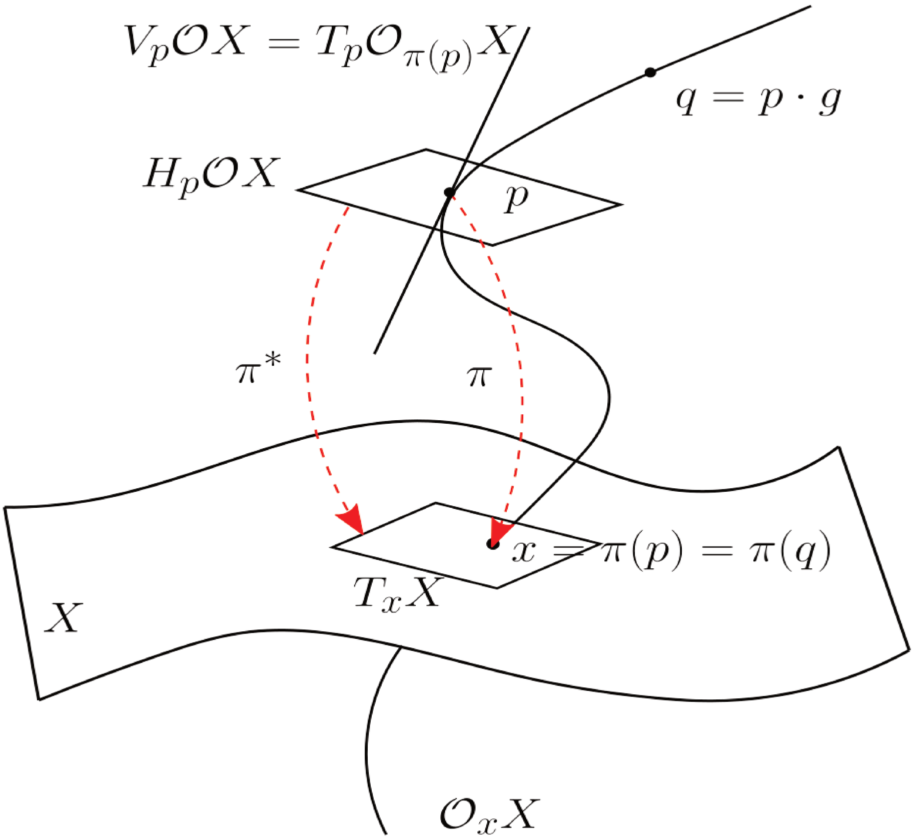

Figure 5.1: The projection, vertical and horizontal spaces on a manifold.

for all $f, g \in C^\infty(\mathcal{O}(X))$.
The following diagram enlightens the relation between the manifold, the frame bundle, its tangent space and its subspaces.

$$\begin{tikzcd}
V\mathcal{O}(X) & T\mathcal{O}(X) \arrow[l, "H+V \mapsto V"'] \arrow[r, "H+V \mapsto H"] & H\mathcal{O}(X) \arrow[r] \arrow[d, "\pi_*"'] & \mathcal{O}(X) \arrow[d, "\pi"] \\
 & & TX \arrow[r] & X
\end{tikzcd}$$

The idea of a parallel transport of fibres depends on that of a horizontal lift of a curve which lies on the base manifold. Intuitively, the projection of a horizontal lift of a curve to the base manifold gives the same curve we started with. We would like to add other conditions on this lifting in order to "connect" neighbouring fibres. More precisely, we have the following definition:

Definition 5.3. *Let $x : [0,1] \to X$ be a curve on X. The horizontal lift of x_t through $u_0 \in \mathcal{O}(X)$ is the unique curve*

$$u : [0,1] \to \mathcal{O}(X)$$

that starts at $u(0) = u_0 \in \mathcal{O}_{x_0}(X)$ such that for all $t \in [0,1]$

1. *$\pi \circ u_t = x_t$;*
2. *$\mathrm{ver}(H_u(u_t)) = 0$;*
3. *$\pi_*(H_u(u_t)) = H_x(x_t)$,*

where $H_u(u_t)$ is the tangent vector to the lifted curve u_t, at each point along the curve. The same is true for $H_x(x_t)$.

Remark 5.2.
1. *The second condition means that the tangent vectors to the curve u_t lie entirely in the horizontal space at each point.*
2. *A curve on X has several horizontal lifts on $\mathcal{O}(X)$. The uniqueness in the definition above comes from the choice of $u_0 \in \mathcal{O}_{x_0}(X)$.*

The method for writing down an explicit expression of a horizontal lift of a curve x_t through $u_0 \in \mathcal{O}_{x_0}(X)$ will be done in three steps:
First step: we produce an arbitrary curve $\delta : [0,1] \to \mathcal{O}(X)$ such that $\pi \circ \delta = x$. This curve is produced by the means of a local section $\sigma : [0,1] \to \mathcal{O}(X)$ such that $\pi \circ \sigma = \mathrm{Id_X}$, where (m, U) is a local chart on the base manifold X. We let $\delta_t = \sigma \circ x_t$.
Second step: next, we generate the horizontal lift curve u_t through $u_0 \in \mathcal{O}_{x_0}(X)$ by action of a curve $g : [0,1] \to O(n)$ so that

$$u_t = \delta_t \lhd g_t.$$

One can think of g_t as an adjustment variable which shifts δ_t around fibres and hence enables to describe how much δ_t deviates from being horizontal.
This curve g_t will be the unique solution of a first order ODE with initial condition $g(0) = g_0 \in O(n)$ such that

$$\delta_0 \lhd g_0 = u_0 \in \mathcal{O}(X).$$

Detailed derivations and calculations to obtain the ODE can be found in [48, page 69] and [44, page 265].
Third step: we obtain the following ODE that governs the necessary shifts in the fibre:

$$\dot{g}(t) = -(\omega_{\delta_t}(H_\delta(\delta_t)))g(t),$$

where ω_{δ_t} is the one-form introduced in Remark 5.1. This ODE can explicitly be solved in local charts[4] on the base manifold X. The solution can be found in [44, pages 265-266].

We summarize in Figure 5.2 the construction of a horizontal lift of a curve.

Now that the horizontal lift of a curve through u_0 has been constructed, we can define the parallel transport (or displacement) of fibres along a curve. This map is obtained by varying u_0 in the fibre $\pi^{-1}(x(0))$ into the fibre $\pi^{-1}(x(1))$ by mapping u_0 into u_1:

Definition 5.4. *Let $u : [0,1] \to \mathcal{O}(X)$ be the horizontal lift through $u_0 \in \mathcal{O}_{x_0}(X)$ of a curve $x : [0,1] \to X$. The parallel transport map along x_t is the map*

$$\begin{array}{cccc} T_x & : & \mathcal{O}_{x_0}(X) & \to & \mathcal{O}_{x_1}(X) \\ & & u_0 & \mapsto & u_1 \end{array}$$

Note that this map is actually an isomorphism of fibres.

We can recover the definition of a parallel vector field V on TX along a curve x_t introduced in Definition 5.1 by using the horizontal lift of x_t through $u_0 \in \pi^{-1}(x(0))$ as follows:
For a piecewise C^1 curve $x : [0,1] \to X$ and a vector $V_0 \in T_{x_0}X$, we define *the parallel translate of V_0 along x_t*

$$//_t(V_0) = u_t(u_0^{-1}(V_0)),$$

where u_t is the horizontal lift of x_t from u_0. Note that this definition is independent of the choice of u_0.
The mapping $//_t : T_{x_0}X \to T_{x_1}X$ is a linear isomorphism. The parallel translation preserves inner products

$$\langle //_t V, //_t V' \rangle_{T_{x_t}X} = \langle V, V' \rangle_{T_{x_0}X}.$$

This is due to the fact that orthonormal frames preserve inner products as well.
Now if we have a a vector field V along x_t i.e. $V : X \to TX$ such that $V(x_t) \in T_{x_t}X$ for each t, we define its *covariant derivative along x_t* by

$$\frac{DV}{\partial t} := //_t \frac{d}{dt} //_t^{-1} V.$$

[4] The Picard-Lindelöf theorem guarantees the existence and uniqueness of the solution.

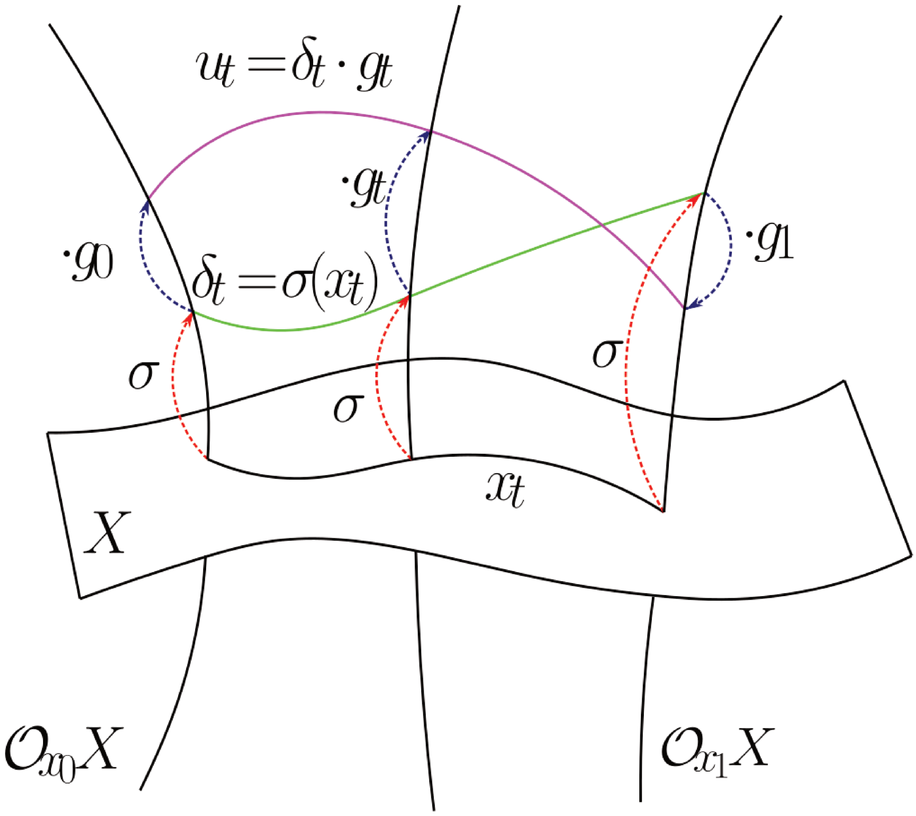

Figure 5.2: Horizontal lift of a curve.

Thus, V is parallel along x_t if, and only if $\frac{DV}{\partial t} = 0$. One can show that the two definitions are equivalent by using local charts [39, Equation (2.1.1)] and [28, Page 300].

5.2 Construction of Brownian motion

We define the operator

$$\Delta_{\mathcal{O}(X)} := \sum_{i=1}^{n} H_i^2$$

as the Bochner's horizontal Laplacian on $\mathcal{O}(X)$. The Eells-Elworthy-Malliavin construction is based on the following relation

$$\Delta f(x) = \Delta_{\mathcal{O}(X)}(f \circ \pi)(u), \tag{5.2}$$

for any smooth function $f : X \longrightarrow \mathbb{R}$ and any $u \in \mathcal{O}(X)$ such that $\pi u = x$.

Consider the following SDE on $\mathcal{O}(X)$ in Stratonovich form

$$dU_t = \sum_{i=1}^{n} H_i(U_t) \circ dW_t^i. \tag{5.3}$$

It is driven by an n-dimensional Brownian motion W. Once an initial frame U_0 is given, the unique solution of this SDE is called a horizontal Brownian motion on $\mathcal{O}(X)$. We recall the result on the generator of this type of SDE's:

Proposition 5.2. *An SDE in the following Stratonovich form*

$$dB_t^X = V_\alpha(B_t^X) \circ dW_t^\alpha + V_0(B_t^X)dt,$$

where Einstein summation convention is used, generates a diffusion process with generator

$$L = \frac{1}{2}\sum_\alpha V_\alpha^2 + V_0.$$

Therefore, the solution of (5.3) is a $\Delta_{\mathcal{O}(X)}/2$-diffusion process. For a smooth function $F : \mathcal{O}(X) \longrightarrow \mathbb{R}$ we can write Itô formula

$$\begin{aligned} F(U_t) &= F(U_0) + \sum_{i=1}^{n}\int_0^t H_iF(U_s)dW_s^i + \frac{1}{2}\sum_{i,j=1}^{n}\int_0^t H_iH_jF(U_s)d\langle W^i, W^j\rangle_s \\ &= F(U_0) + \sum_{i=1}^{n}\int_0^t H_iF(U_s)dW_s^i + \frac{1}{2}\int_0^t \Delta_{\mathcal{O}(X)}F(U_s)ds. \end{aligned}$$

Now if we apply this formula to particular function $F = f \circ \pi$, which is the lift of f on X, the by Formula (5.2) we obtain

$$f(B_t^X) = f(B_0^X) + \sum_{i=1}^{n}\int_0^t H_i(f \circ \pi)(U_s)dW_s^i + \frac{1}{2}\int_0^t \Delta f(B_s^X)ds,$$

where $B^X = \pi U$ is the projection of the horizontal Brownian motion $U = (U_t)_t$ on X. Hence, $(B_t^X)_t$ is a Brownian motion on the manifold X starting from $B_0^X = \pi U_0$.
Running the machinery backwards, if we want to construct a Brownian motion starting from x, we fix a frame $u \in \mathcal{O}(X)$ over x (i.e. $\pi u = x$). There exists then a unique horizontal Brownian motion $(U_t)_t$ starting from the frame u and its projection $B^X = \pi U$ is a Brownian motion starting from x.

Once we have constructed a Brownian motion $B^X = (B_t^X)_t$ on a manifold, it is not hard to write down the anti-development of B^X on $\mathbb{R}^n$:

$$W_t = \int_0^t U_s^{-1} \circ dB_s^X.$$

This equation drives (5.3). We notice that the correspondences

$$W \leftrightarrow B^X \leftrightarrow U$$

are very useful since one can convert a manifold-valued process B^X into Euclidean space valued process W, which is easier to handle. It is important to keep in mind that these correspondences depend on the connection used in order to defined horizontal lift for vectors. In our case we use the Riemannian connection (or Levi Civita connection) but the whole construction can be made with another affine connection. The only difference is that the orthonormal frame bundle should be replaced by the general linear frame bundle and the orthogonal group $O(n)$ by $GL(n, \mathbb{R})$.
It is also worth to mention that the parallel transport associated with the Levi Civita connection preserves the orthogonality of frames. This is due to the fact that the Levi Civita connection is compatible with the Riemannian metric.

5.3 Heat equation and 1-forms

We previously saw that the solution of the heat equation

$$\begin{cases} \partial_t v(t,x) &= \frac{\Delta}{2} v(t,x) \quad \text{on } \mathbb{R}_*^+ \times X \\ v_0(x) &= f(x), \quad \forall x \in X \end{cases} \tag{5.4}$$

can be uniquely solved as

$$v(t,x) = \mathbb{E}(f(B_t^X) | B_0^X = x),$$

where $B^X = (B_t^X)_t$ is a Brownian motion starting at x and f is a function. In order to generalize this fact to the case of the heat equation for 1-forms, Itô considered in [45] the previous problem, where v and f are now 1-forms on X. The problem is that $f(B_t^X)$ is attached to B_t^X which varies with t, while $v(t,x)$ should be attached to $x = B_0^X$. Therefore we should shift $f(B_t^X)$ back to B_0^X along the path B_t^X by parallel displacement. This is how the notion of stochastic parallel displacement appeared. Malliavin presented

in [59] a new approach: instead of translating a 1-form at B_t^X to a 1-form at B_0^X, we translate an orthonormal frame U_0 at B_0^X to an orthonormal frame U_t at B_t^X along the Brownian curve. We follow [39], where this approach is explained.
First, we need to lift the handled objects in (5.4) to their scalarization on the orthonormal frame bundle $\mathcal{O}(X)$.

Let $(x^1, \cdots, x^n)$ be a local coordinate system. It induces a basis $\left(\frac{\partial}{\partial x^i}\right)_i$ of T_xX and a dual basis $(dx^i)_i$ of T_x^*X. The system $\left(\frac{\partial}{\partial x^I} \otimes dx^J\right)$ is a basis of $T_x^{p,q}X$, where $I = (i_1, \cdots, i_p)$, $J = (j_1, \cdots, j_q)$ are multi indices of degrees respectively p and q, $\frac{\partial}{\partial x^I} = \frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_p}}$ and $dx^J = dx^{j_1} \otimes \cdots \otimes dx^{j_q}$.
A tensor field θ of type (p, q) is a mapping

$$\theta : x \in X \mapsto \theta(x) \in T_x^{p,q}X,$$

whose components are $\theta^I_J(x)$ with respect to the previous basis.

Following [39], we assume that a frame u is canonically extended to an isometry $u : T^{p,q}\mathbb{R}^n \to T_x^{p,q}X$. The scalarization[5] of θ at u is the map

$$\tilde{\theta} : \mathcal{O}(X) \to T^{p,q}\mathbb{R}^n$$

defined by

$$\tilde{\theta}(u) = u^{-1}(\theta)(\pi u).$$

Equivalently, we define it by

$$\tilde{\theta}(u) = \theta^I_J(\pi u) e_I \otimes e_*^J,$$

assuming that $u(e_I \otimes e_*^J) = \frac{\partial}{\partial x^I} \otimes dx^J$.
The scalarization gives the coordinates of θ in the frame u at x. This map is $O(n)$-equivariant in the sense that $\tilde{\theta}(u \cdot g) = g^{-1} \cdot \tilde{\theta}(u)$, where g on the right side means the usual extension of the action of $O(n)$ from $\mathbb{R}^n$ to $T^{p,q}\mathbb{R}^n$.
Likewise,

$$\Delta_{\mathcal{O}(X)} \tilde{v}(u) = u^{-1}(\underline{\Delta} v(\pi u))$$

where v is a 1-form and $\underline{\Delta} = \mathrm{Trace}\nabla^2$ is the rough Laplacian acting on tensors (see (2.5)). This comes from the fact that $H_i \tilde{v}(u) = \widetilde{\nabla v}(u)$ and $\Delta_{\mathcal{O}(X)} = \sum_{i=1}^n H_i^2$ [39, Proposition 3.1.2]. Moreover, we define

$$\widetilde{\mathrm{Ric}}_u = u^{-1} \, \mathrm{Ric}_{\pi u} \, u,$$

[5] Malliavin stated in [59, Proposition 2.3.1] that there is an isomorphic correspondence between 1-forms and the space of C^2 functions defined on $\mathcal{O}(X)$, $\mathbb{R}^n$-valued and $O(n)$-equivariant. This isomorphism can actually be generalized to any type of tensor fields on X.

where $\pi u = x$ and $\mathrm{Ric}_{\pi u} : T_xX \to T_xX$ is the Ricci transform defined by

$$\mathrm{Ric}_x(V) = \sharp(\mathrm{Ric}(V,\cdot)).$$

Note that these 3 lifts are $\mathbb{R}^n$-valued.
Finally, (5.4) is equivalent to

$$\begin{cases} \partial_t \tilde{v} &= \frac{\Delta_{\mathcal{O}(X)}}{2}\tilde{v} \quad \text{on } \mathbb{R}_*^+ \times \mathcal{O}(X) \\ \tilde{v}_0 &= \tilde{f}, \quad \text{on } \mathcal{O}(X) \end{cases} \tag{5.5}$$

Although the rough Laplacian is the natural operator to choose for the heat equation, it is preferable to use Hodge-de Rham operator, which is geometrically more significant. We have seen that the difference $\vec{\Delta} - \underline{\Delta}$ is a fibre-wise linear operator by the generalized Bochner-Weitzenböck formula and more specifically it is equal to $\mathrm{Ric}(\cdot, \sharp\cdot)$. The new problem that we consider then is

$$\begin{cases} \partial_t \tilde{v} &= \frac{\vec{\Delta}_{\mathcal{O}(X)}}{2}\tilde{v}, \quad \text{on } \mathbb{R}_*^+ \times \mathcal{O}(X) \\ \tilde{v}_0 &= \tilde{f}, \quad \text{on } \mathcal{O}(X) \end{cases} \tag{5.6}$$

where

$$\begin{aligned} \vec{\Delta}_{\mathcal{O}(X)}\tilde{v}(u) &= \Delta_{\mathcal{O}(X)}\tilde{v}(u) - \widetilde{\mathrm{Ric}}_u\tilde{v}(u) \quad \text{by Bochner-Weitzenböck formula} \\ &= u^{-1}(\vec{\Delta}v(\pi u)) \quad \text{by scalarization.} \end{aligned}$$

The solution of (5.6) can be obtained by using a matrix version of the weighted Feynmac Kac formula discussed in Subsection 2.4.2. For this purpose, let M_t be a $\mathrm{End}(\mathbb{R}^n)$-valued multiplicative functional determined by

$$\frac{dM_t}{dt} = -\frac{\widetilde{\mathrm{Ric}}_{U_t}}{2}M_t, \qquad M_0 = I_n.$$

The solution of (5.6) is then

$$\tilde{v}(t,u) = \mathbb{E}_u(M_t\tilde{v}_0(U_t)).$$

Correspondingly, the solution of the manifold version of (5.6) is

$$v(t,x) = \mathbb{E}_x(U_0M_tU_t^{-1}\left(v_0(B_t^X)\right)),$$

where $U = (U_t)_t$ is the horizontal lift of the Brownian motion B^X.

Proof. Suppose that $\tilde{v}$ is a solution. By differentiating $M_s\tilde{v}(t-s,U_s)$ and using Itô formula we obtain

$$\begin{aligned} d(M_s\tilde{v}(t-s,U_s)) &= M_s d\tilde{v}(t-s,U_s) - \frac{\widetilde{\mathrm{Ric}}_{U_t}}{2}M_s\tilde{v}(t-s,U_s)ds \\ &= \sum_{i=1}^n M_sH_i\tilde{v}(t-s,U_s)dW_s^i \\ &\quad + M_s\left(-\partial_s + \frac{\Delta_{\mathcal{O}(X)}}{2} - \frac{\widetilde{\mathrm{Ric}}_{U_t}}{2}\right)\tilde{v}(t-s,U_s)ds \end{aligned}$$

The last term vanishes because $\tilde{v}$ is a solution of (5.6), meaning that $(M_s \tilde{v}(t-s, U_s)), s \in [0,t]$ is a martingale. The proof is completed by equating the expected valued at $s = 0$ and $s = t$.
The second formula is just a rewriting of the first one. □

Remark 5.3. *We often use a fixed frame U_0 to identify the tangent space T_xX with $\mathbb{R}^n$. Under this identification U_0 becomes the identity map and it is often omitted in the solution. The corresponding writing of the second formula becomes then*

$$v(t,x) = \mathbb{E}_x(M_t U_t^{-1} v_0(B_t^X)).$$

Remark 5.4. *The solution $v(t,x)$ is commonly denoted by $\vec{P}_t v_0(x)$ or $e^{-t\vec{\Delta}} v_0(x)$, the heat semigroup. As in the Euclidean case (cf. Example 2.2), we also have a probabilistic representation for the Poisson semigroup acting on 1-forms by using the Bochner subordination formula [54]. Indeed, let B_t be the standard Brownian motion on $\mathbb{R}$ starting from $B_0 = y > 0$ and define*

$$\tau_y = \inf\{t > 0 : B_t = 0\}.$$

It is known by subordination formula that

$$\mathbb{E}_y(e^{-\lambda \tau_y}) = e^{-y\sqrt{\lambda}}.$$

Then by spectral decomposition and the previous result, we have for every $\vec{g} \in C_c^\infty(X, \Lambda^1(T^*X))$,

$$\begin{aligned} e^{-y\sqrt{\vec{\Delta}}} \vec{g}(x) &= \mathbb{E}_y(e^{-\vec{\Delta}\tau_y} \vec{g}(x)) = \mathbb{E}_y(\mathbb{E}_x(M_{\tau_y} U_{\tau_y}^{-1} \vec{g}(B_{\tau_y}^X))) \\ &= \mathbb{E}_{(x,y)}(M_{\tau_y} U_{\tau_y}^{-1} \vec{g}(B_{\tau_y}^X)). \end{aligned}$$

Similarly, we have a probabilistic representation for the Poisson semigroup acting on functions. Indeed,

$$\begin{aligned} e^{-y\sqrt{-\Delta}} f(x) &= \mathbb{E}_y(e^{\Delta \tau_y} f(x)) = \mathbb{E}_y(\mathbb{E}_x(f(B_{\tau_y}))) \\ &= \mathbb{E}_{(x,y)}(f(B_{\tau_y})). \end{aligned}$$

5.4 Probabilistic representation of the Riesz transform on manifolds

In this section, we define a function $\varphi \in C_c^\infty(X)$ as in Section 2.7.1. The reason is, as stated before, that in next chapters we will consider weighted Laplacians with some extra terms involving φ. The following results can be found in [54].

Let $(B_t^X)_t$ be the Δ_φ-diffusion process such that $B_0^X = x$. By Itô's theory of diffusion process on Riemannian manifolds, there exists a Brownian motion (W_t) on $\mathbb{R}^n$ such that

$$dB_t^X = U_t \circ dW_t - \nabla\varphi(B_t^X) dt$$

where $U_t : T^*_{X_0}X \longrightarrow T^*_{X_t}X$ is the stochastic parallel transport along $\{B^X_s,\ 0 \leq s \leq t\}$. Let $(B_t)_t$ be a 1-dimensional Brownian motion starting at $y > 0$ whose generator is $\frac{d^2}{dy^2}$ instead of $\frac{d^2}{2dy^2}$. Assume that B_t is independent of B^X_t. We introduce the background radiation process on $X \times \mathbb{R}^+$ by $Z_t = (B^X_t, B_t)$ following [35] and whose generator is $\Delta_\varphi + \frac{d^2}{dy^2}$.
Before presenting a probabilistic representation of the Riesz transform, we present a lemma which aim is to write a version of Itô formula that holds for 1-forms.

Lemma 5.1. *Let $\vec{g} \in C^\infty_c(X, \Lambda^1(T^*X))$ and $\vec{Q}\vec{g}(x,y) = e^{-y\sqrt{-\vec{\Delta}_\varphi}}\vec{g}(x)$. Define $\tau = \inf\{t > 0 : B_t = 0\}$. We have*

$$\vec{g}(B^X_\tau) = M^{*,-1}_\tau \vec{Q}\vec{g}(B^X_0, B_0) + M^{*,-1}_\tau \int_0^\tau M^*_s(\nabla, \partial_y)\vec{Q}\vec{g}(B^X_s, B_s)(U_s dW_s, dB_s). \quad (5.7)$$

Proof. By using the covariant Itô formula on Riemannian manifolds, the product rule as well as the fact that M^*_t is the solution of the SDE

$$dM^*_t = -\operatorname{Ric}_\varphi(B^X_t)M^*_t dt,$$

we have

$$\begin{aligned}
d(M^*_t\vec{Q}\vec{g}(B^X_t, B_t)) &= -\operatorname{Ric}_\varphi(B^X_t)M^*_t\vec{Q}\vec{g}(B^X_t, B_t)dt \\
&\quad + M^*_t(\nabla, \partial_y)\vec{Q}\vec{g}(B^X_t, B_t)(dB^X_t, dB_t) \\
&\quad + M^*_t\nabla^2\vec{Q}\vec{g}(B^X_t, B_t)d\langle B^X, B^X\rangle_t + M^*_t\partial^2_{y^2}\vec{Q}\vec{g}(B^X_t, B_t)dt \\
&\overset{(1)}{=} M^*_t(\Delta - \operatorname{Ric}_\varphi(B^X_t))\vec{Q}\vec{g}(B^X_t, B_t)dt \\
&\quad + M^*_t(\nabla, \partial_y)\vec{Q}\vec{g}(B^X_t, B_t)(U_t dW_t, dB_t) \\
&\quad - M^*_t(\nabla_{\nabla\varphi(B^X_t)})\vec{Q}\vec{g}(B^X_t, B_t)dt + M^*_t\partial^2_{y^2}\vec{Q}\vec{g}(B^X_t, B_t)dt \\
&\overset{(2)}{=} M^*_t(\nabla, \partial_y)\vec{Q}\vec{g}(B^X_t, B_t)(U_t dW_t, dB_t).
\end{aligned}$$

In (1), we have used the fact that

$$dB^X_t = U_t \circ dW_t - \nabla\varphi(B^X_t)dt.$$

Hence, after integration and computing the square brackets we obtain

$$d\langle B^X, B^X\rangle_t = \sum_{i,j=1}^n (U^i_t, U^j_t)\delta_{i,j}dt$$

which in turn implies that

$$\begin{aligned}
\nabla^2\vec{Q}\vec{g}(B^X_t, B_t)d\langle B^X, B^X\rangle_t &= \sum_{i=1}^n \nabla^2\vec{Q}\vec{g}(B^X_t, B_t)(U^i_t, U^i_t)dt \\
&= \operatorname{Trace}\nabla^2\vec{Q}\vec{g}(B^X_t, B_t)dt \\
&= \Delta\vec{Q}\vec{g}(B^X_t, B_t)dt.
\end{aligned}$$

In (2), we have used the generalized Bochner Weitzenböck formula

$$-\vec{\Delta}_\varphi = \Delta - \nabla_{\nabla\varphi} - \mathrm{Ric}_\varphi \tag{5.8}$$

and the harmonicity of $\vec{Q}$. Recall that $\vec{\Delta}_\varphi$ is the Hodge Laplacian and that Δ is the Bochner Laplacian.
The proof is completed by integrating from $t = 0$ to $t = \tau$. □

Remark 5.5. *Recall that although Δ initially acts on functions (in which case it is called the Laplace Beltrami operator), its definition can be extended to tensor fields. Indeed, for any tensor T we have*

$$\Delta T = \mathrm{Trace}\nabla^2 T,$$

where

$$\nabla^2_{X,Y} T = \nabla_X \nabla_Y T - \nabla_{\nabla_X Y} T.$$

The next lemma is a probabilistic result due to P-A. Meyer [60]. Recall that $Z_t = (X_t^X, X_t)$ and it starts at $(x,y) \in X \times \mathbb{R}_+$. We denote by $\mathbb{E}_y$ and $\mathbb{P}_y$ the expectation and probability of Z_t, respectively. This means that

$$\mathbb{E}_y = \int_X \mathbb{E}_{(x,y)} dx$$

and

$$\mathbb{P}_y = \int_X \mathbb{P}_{(x,y)} dx.$$

Lemma 5.2. *Suppose that there exists a non-negative constant such that*

$$\mathrm{Ric}_\varphi = \mathrm{Ric} + \nabla\varphi \geq -a.$$

Then, for all non-negative measurable functions f on X, we have

$$\mathbb{E}_y(f(B^X_\tau)) = \int_X f(x) d\mu(x).$$

Moreover, for ll non-negative measurable functions F or for all measurable F such that $F(x,\eta)\eta \in L^1(\lambda(dx) \otimes d\eta)$, we have the Green function formula

$$\mathbb{E}_y\left[\int_0^\tau F(Z_t)dt\right] = 2\int_0^\infty \int_X F(x,\eta)(y \wedge \eta) d\mu(x) d\eta.$$

Proof. Let P_y be the heat semigroup. By Remark 5.4 and using the fact that P_y is Markovian and symmetric, one can write

$$\begin{aligned}
\mathbb{E}_y(f(B^X_\tau)) &= \int_X \mathbb{E}_{(x,y)}(f(B^X_\tau)) d\mu(x) \\
&= \int_X P_y f(x) d\mu(x) \\
&= \langle P_y f, 1\rangle \\
&= \langle f, P_y 1\rangle \\
&= \int_X f(x) d\mu(x).
\end{aligned}$$

The second part of the Lemma is detailed in [54, Proposition 3.1]. □

The following formula gives a natural extension of the probabilistic representation of 1-forms, which will be used to deduce the representation of the Riesz transform à la Gundy-Varopoulos.

Theorem 5.1. *Let $\vec{g} \in C_c^\infty(X, \Lambda^1(T^*X))$ and $\vec{Q}\vec{g}(x,y) = e^{-y\sqrt{-\vec{\Delta}_\varphi}}\vec{g}(x)$. Then*

$$\frac{1}{2}\vec{g}(x) = \lim_{y\to+\infty} \mathbb{E}_y \left[M_\tau \int_0^\tau M_s^{-1}\partial_y \vec{Q}\vec{g}(B_s^X, B_s)dB_s | B_\tau^X = x \right]. \tag{5.9}$$

Proof. Take $\vec{h} \in C_c^\infty(X, \Lambda^1(T^*X))$. We want to show that

$$\langle \vec{g}, \vec{h}\rangle_{L^2(\mu_\varphi)} = 2\lim_{y\to+\infty}\int_X \left(\mathbb{E}_y \left[M_\tau \int_0^\tau M_s^{-1}\partial_y \vec{Q}\vec{g}(B_s^X, B_s)dB_s | B_\tau^X = x \right], \vec{h}(x) \right)_{T_x^*X} d\mu_\varphi(x). \tag{5.10}$$

From now on, we will drop the subscripts of of the inner products to facilitate the notations. By (5.7) and Lemma 5.2 we have

$$\begin{aligned}
&\int_X \left(\mathbb{E}_y \left[M_\tau \int_0^\tau M_s^{-1}\partial_y \vec{Q}\vec{g}(B_s^X, B_s)dB_s | B_\tau^X = x \right], \vec{h}(x) \right) d\mu_\varphi(x) \\
&\qquad = \mathbb{E}_y \left[\left(M_\tau \textstyle\int_0^\tau M_s^{-1}\partial_y \vec{Q}\vec{g}(B_s^X, B_s)dB_s, \vec{h}(B_\tau^X) \right) \right] \\
&\qquad\qquad = I_1(y) + I_2(y),
\end{aligned}$$

where

$$\begin{aligned}
I_1(y) &= \mathbb{E}_y \left[\left(M_\tau \int_0^\tau M_s^{-1}\partial_y \vec{Q}\vec{g}(B_s^X, B_s)dB_s, M_\tau^{*,-1}\vec{h}(B_0^X, B_0) \right) \right] \\
I_2(y) &= \mathbb{E}_y \left[\left(M_\tau \int_0^\tau M_s^{-1}\partial_y \vec{Q}\vec{g}(B_s^X, B_s)dB_s, M_\tau^{*,-1} \int_0^\tau M_s^*(\nabla, \partial_y)\vec{Q}\vec{h}(B_s^X, B_s)(U_s dW_s, dB_s) \right) \right].
\end{aligned}$$

Since $\int_0^\tau M_s^{-1}\partial_y \vec{Q}\vec{g}(B_s^X, B_s)dB_s$ is a martingale, one can write

$$\begin{aligned}
I_1(y) &= \mathbb{E}_y \left[\left(\int_0^\tau M_s^{-1}\partial_y \vec{Q}\vec{g}(B_s^X, B_s)dB_s, \vec{h}(B_0^X, B_0) \right) \right] \\
&= \mathbb{E}_y \left[\left(\mathbb{E}\left[\int_0^\tau M_s^{-1}\partial_y \vec{Q}\vec{g}(B_s^X, B_s)dB_s | (B_0^X, B_0) \right], \vec{h}(B_0^X, B_0) \right) \right] \\
&= 0.
\end{aligned}$$

As for I_2, we use the Itô L^2-isometry identity to obtain

$$\begin{aligned}
I_2(y) &= \mathbb{E}_y \left[\left(\int_0^\tau M_s^{-1}\partial_y \vec{Q}\vec{g}(B_s^X, B_s)dB_s, \int_0^\tau M_s^*(\nabla, \partial_y)\vec{Q}\vec{h}(B_s^X, B_s)(U_s dW_s, dB_s) \right) \right] \\
&= \mathbb{E}_y \left[\int_0^\tau \left(M_s^{-1}\partial_y \vec{Q}\vec{g}(B_s^X, B_s), M_s^*\partial_y \vec{Q}\vec{h}(B_s^X, B_s) \right) ds \right] \\
&= \mathbb{E}_y \left[\int_0^\tau \left(\partial_y \vec{Q}\vec{g}(B_s^X, B_s), \partial_y \vec{Q}\vec{h}(B_s^X, B_s) \right) ds \right] \\
&= 2\int_X \int_0^\infty (y\wedge z) \left(\partial_z \vec{Q}\vec{g}(x,z), \partial_z \vec{Q}\vec{h}(x,z) \right) dz d\mu_\varphi(x) \text{ by Lemma 5.2, } 2^{nd} \text{ part.}
\end{aligned}$$

Since we are interested in letting y tend to infinity, the term $y \wedge z$ becomes simply z and we can use Littlewood-Paley identity to finish the proof. Indeed, define $\{E_\lambda, \lambda \geq 0\}$ be the spectral resolution associated to the infinitesimal generator of the heat semigroup. Then the Poisson semigroup is defined by

$$\vec{Q}\vec{g}(t,x) = \int_0^\infty e^{-\lambda^{1/2}t} dE_\lambda \vec{g}(x).$$

Then by Fubini theorem we have

$$\int_X \int_0^\infty z \left(\partial_z \vec{Q}\vec{g}(x,z), \partial_z \vec{Q}\vec{h}(x,z)\right) dz d\mu_\varphi(x) = \int_0^\infty z \langle \partial_z \vec{Q}\vec{g}(\cdot,z), \partial_z \vec{Q}\vec{h}(\cdot,z)\rangle_{L^2(\mu_\varphi)} dz.$$

By using the spectral resolution the RHS is equal to

$$\int_0^\infty z \langle \int_0^\infty \lambda^{1/2} e^{-\lambda^{1/2}z} dE_\lambda \vec{g}(x), \int_0^\infty \lambda^{1/2} e^{-\lambda^{1/2}z} dE_\lambda \vec{h}(x)\rangle dz,$$

which in turn can be computed as follows

$$\begin{aligned}
\int_0^\infty z \int_0^\infty \lambda e^{-2\lambda^{1/2}z} d\langle E_\lambda \vec{g}(x), E_\lambda \vec{h}(x)\rangle &= \int_0^\infty \lambda \int_0^\infty z e^{-2\lambda^{1/2}z} d\langle E_\lambda \vec{g}(x), E_\lambda \vec{h}(x)\rangle \\
&= \frac{\Gamma(2)}{4} \int_0^\infty d\langle E_\lambda \vec{g}(x), E_\lambda \vec{h}(x)\rangle \\
&= \frac{1}{4} \int_X \left(\vec{g}(x), \vec{h}(x)\right) d\mu_\varphi(x).
\end{aligned}$$

Here, Γ is the Gamma function. Hence

$$2 \lim_{y\to+\infty} \int_X \int_0^\infty (y \wedge z) \left(\partial_z \vec{Q}\vec{g}(x,z), \partial_z \vec{Q}\vec{h}(x,z)\right) dz d\mu_\varphi(x) = \frac{1}{2} \int_X (\vec{g}(x), \vec{h}(x)) d\mu_\varphi(x).$$

Equation (5.10) is obtained by multiplying by 2 both sides of the previous equality. $\square$

We deduce from this formula the particular probabilistic representation of the Riesz transform associated to Δ_φ by applying (5.9) to $\vec{g} = d \circ (-\Delta_\varphi)^{-1/2} f$ and using the commutation formula $d \circ (-\Delta_\varphi) = \vec{\Delta}_\varphi \circ d$ in (2.4) as follows

$$\begin{aligned}
\partial_y e^{-y\sqrt{\vec{\Delta}_\varphi}} (d \circ \Delta_\varphi^{-1/2} f) &= -\sqrt{\vec{\Delta}_\varphi} e^{-y\sqrt{\vec{\Delta}_\varphi}} (d \circ (-\Delta_\varphi)^{-1/2} f) \\
&= -e^{-y\sqrt{\vec{\Delta}_\varphi}} \sqrt{\vec{\Delta}_\varphi} d \circ (-\Delta_\varphi)^{-1/2} f \\
&= -e^{-y\sqrt{\vec{\Delta}_\varphi}} d(-\Delta_\varphi)^{1/2} (-\Delta_\varphi)^{-1/2} f \\
&= -e^{-y\sqrt{\vec{\Delta}_\varphi}} df \\
&= -de^{-y\sqrt{-\Delta_\varphi}} f \\
&= -dQf(\,\cdot\,, y).
\end{aligned}$$

Finally, we obtain that

$$-\frac{1}{2}\mathcal{R}_\varphi f(x) = \lim_{y\to\infty} \mathbb{E}_y \left[M_\tau \int_0^\tau M_s^{-1} dQ(f)(B_s^X, B_s) dB_s | B_\tau^X = x \right]. \tag{5.11}$$

We are now ready to present the probabilistic representation of the Riesz transform associated with the Ornstein Uhlenbeck operator.

Example 5.1. *Taking $X = \mathbb{R}^n$ and $\varphi(x) = \frac{\|x\|^2}{2}$, we have $d\mu(x) = d\gamma(x)$, the Gaussian measure. In this case, we have $M_t = e^{-t} Id$, $\forall t \geq 0$ and we recover the probabilistic representation of the Riesz transform associated with Δ_{OU} obtained in Section 2.6*

$$-\frac{1}{2}\nabla(-\Delta_{OU})^{-1/2} f(x) = \lim_{y\to\infty} \mathbb{E}_y \left[e^{-\tau} \int_0^\tau e^s \nabla e^{-B_s\sqrt{-\Delta_{OU}}} f(X_s) dB_s | X_\tau = x \right].$$

Note that the above formula does not depend on the dimension of $\mathbb{R}^n$.

Remark 5.6. *The martingale representation of the Riesz transform acting on k-forms for $k \geq 1$ remains the same [55]. The only change occurs in the Bakry-Emery curvature which becomes the (weighted) Weitzenböck curvature.*

Chapter 6

Sharp L^p estimate of the Bakry Riesz transform

> In this work the assumption of quadratic integrability will be replaced by the integrability of $|f(x)|^p$. The analysis of these function classes will shed a particular light on the real and apparent advantages of the exponent 2; one can also expect that it will provide essential material for an axiomatic study of function spaces.
>
> F. Riesz, 1910

We present in this chapter a new proof of the dimensionless L^p boundedness of the Riesz vector on manifolds with bounded geometry. Our proof has the significant advantage that it allows for a much stronger conclusion, namely that of a new dimensionless L^p estimate and weighted L^2 estimate with optimal exponent. Other than previous arguments, only a small part of our proof is based on special auxiliary functions, the core of the argument is a weak type estimate and a sparse decomposition of the stochastic process by X.D. Li, whose projection is the Riesz vector.

Probabilistic representation of the Bakry-Riesz vector on manifolds. Using a martingale approach, we previously saw that one can represent the Riesz vector R_φ (associated to the Laplacian Δ_φ) via a probabilistic representation. In the literature, it first appeared in [35], where the Riesz transform was defined on $\mathbb{R}^n$. In [2] Arcozzi extended this formula to compact Lie groups and spheres. In [54] and [58], Li presented a new formula adapted to complete Riemannian manifolds. In reference to Chapter 5, the representation formula of the Riesz vector in this setting for a complete manifold with $\mathrm{Ric}_\varphi \geq 0$ is as follows

$$-\frac{1}{2}(R_\varphi f)(x) = \lim_{y\to\infty} \mathbb{E}_y \left[M_\tau \int_0^\tau M_s^{-1} dQ(f)(B_s^X, B_s) dB_s | B_\tau^X = x \right], \tag{6.1}$$

where

- $Q(f)(x,y) = e^{-y\sqrt{-\Delta_\varphi}} f(x)$ is the Poisson semigroup;
- $\tau = \inf\{t > 0 : B_t = 0\}$ is the stopping time upon hitting the boundary of the upper half space;
- M_t is the solution to the matrix-valued stochastic differential equation

  $$dM_t = V_t M_t dt, \quad M_0 = Id,$$

  for some adapted and continuous process $(V_t)_{t\geq 0}$ taking values in the set of symmetric and non-positive $n \times n$ matrices.

Equivalently, one can rewrite this formula as

$$-\frac{1}{2}(R_\varphi f)(x) = \lim_{y\to\infty} \mathbb{E}_y \left[Z_\tau | B_\tau^X = x \right], \tag{6.2}$$

where Z_t is a semi-martingale defined thanks to the auxiliary martingales X_t and Y_t (adapted to the filtration $\mathcal{F}_t = \sigma(B_s^X, B_s, s \leq t)$) as follows

$$X_t = Qf(B_t^X, B_t) - Qf(B_0^X, y) = \int_0^t (\nabla, \partial_y) Qf(B_s^X, B_s)(U_s dW_s, dB_s),$$

$$Y_t = \int_0^t \nabla Qf(B_s^X, B_s) dB_s,$$

$$Z_t = M_t \int_0^t M_s^{-1} dY_s,$$

where Y_t is by construction differentially subordinate to X_t since

$$\langle X, X\rangle_t = \int_0^t \sum_{i=1}^n |\nabla_i Qf(B_s^X, B_s)|^2 ds + \int_0^t |\partial_y Qf(B_s^X, B_s)|^2 ds$$

and

$$\langle Y, Y\rangle_t = \int_0^t \sum_{i=1}^n |\nabla_i Qf(B_s^X, B_s)|^2 ds.$$

6.1 Main results

We prove in Theorem 6.1 a dimensionless estimate in L^p spaces for the Riesz vector on manifold with non-negative curvature. The first proofs of this result are recent [16], [58], [8] and all based on a form of a Bellman function. Our proof is via a sparse domination with continuous index. All these cited Bellman proofs give a better numeric estimate than our proof, but as mentioned earlier, our proof extends (for free) to the weighted L^2 case, which the previous ones do not. Our estimate is linear in p, which means proportional to $(p-1)^{-1}$ when $p<2$ and to $p-1$ when $p>2$. We note that Bañuelos and Osekowski have in [8] the best numeric constant in this case. We note also that the proof in [16] gives the linear estimate with p also in the case where the curvature is merely bounded below (and possibly negative) with an appropriately defined Riesz vector involving a Laplacian with a modified spectrum.

Theorem 6.1 (L^p estimate)**.** *Suppose that X is a complete Riemannian manifold without boundary and $\mathrm{Ric}_\varphi \geq 0$. Then for all $f \in C_c^\infty(X)$ and $p \in (1,\infty)$, we have the following dimension-free estimate*

$$\|R_\varphi f\|_{L^p(T^*X)} \leq 16\frac{p^2}{p-1}\|f\|_{L^p(X)}. \tag{6.3}$$

We prove also a dimensionless weighted estimate in L^2 spaces for the Riesz vector on manifold with non-negative curvature. In the Euclidean setting, see [23]. For the case of manifolds, such an estimate was already known in the case $p=2$ see [18]. A priori the weight has to be globally in L^2 so as to be able to define the flow characteristic.

$$\tilde{Q}_2(w) = \sup_{x,y}(Q(w))(x,y)Q(w^{-1}(x,y)).$$

The collection of weights for which this characteristic is finite is denoted $\tilde{A}_2$. There is also a natural way to extend the class of the weights to resemble more the classical case allowing local L^1 weights. In this case we require that constants are integrable in X with the measure $d\mu_\varphi$ so as to prove the theorem for cut weights, such as in [18], that are in $L^1 \cap L^\infty \cap L^2$ and then define the characteristic by a limiting procedure and deduce the theorem. See Chapter 4 for detailed exposition in the case $p=2$.

Theorem 6.2 (weighted L^2 estimate)**.** *Suppose that X is a complete Riemannian manifold without boundary and $\mathrm{Ric}_\varphi \geq 0$. Then for all $f \in C_c^\infty(X)$ and $w \in \tilde{A}_2$, we have the following dimension-free estimate*

$$\|R_\varphi f\|_{L^2(T^*X,w)} \leq 64\tilde{Q}_2(w)\|f\|_{L^2(X,w)}. \tag{6.4}$$

The technique used in this paper resembles the sparse domination principle for discrete time martingale transforms which originally appeared in [49]. This technique has witnessed considerable efforts in the last several years and has been used to prove numerous new results in harmonic analysis, using sparse operators defined on cubes.

These cannot give dimensionless estimates, nor are satisfactory results known in the non-doubling case. As in [21] we use a sparse operator with continuous stopping times, dominating Li's process Z_t whose projection is the Riesz vector. This is what enables us to use the flow itself without cutting it into cubes, thus resulting in clean dimensionless estimates.

Following [21], we say that the operator $X \mapsto S(X)$ is called sparse if there exists an increasing sequence of adapted stopping times $0 = T^{-1} \leq T^0 \leq \cdots$ with nested sets $E_j = \{T^j < \infty\}$, $E_j \subset E_{j-1}$ so that

$$S(X) = \sum_{j=-1}^{\infty} X_{T^j} \chi_{E_j} \text{ where } X_{T^j} = \mathbb{E}(X|\mathcal{F}_{T^j}); \tag{6.5}$$

$$\forall A_j \subset E_j,\ A_j \in \mathcal{F}_{T_j} \text{ there holds } \mathbb{P}(A_j \cap E_{j+1}) \leq \frac{1}{2}\mathbb{P}(A_j). \tag{6.6}$$

The estimate we aim to show will be a consequence of a sparse domination of the stochastic process Z_t (see [52], [49] and [21]). Other than in [21] the object is not a martingale, so the sparse domination is different and the key of the proof relies on the weak-L^1 estimate for the maximal function of the studied stochastic operator. We do not aim at the fullest generality here, keeping our goal in mind, an estimate of the Riesz vector. Certain assumptions can certainly be weakened, as the attentive reader will observe.

Lemma 6.1 (Weak-type estimate)**.** *Let $X = (X)_t$ be a real valued continuous path martingale and $Y = (Y)_t$ a vector valued continuous path martingale so that Y is differentially subordinate with respect to X. Let further $Z = (Z)_t$ be a continuous path semi-martingale whose increments satisfy $dZ_t = V_t Z_t dt + dY_t$ with (V_t) continuous adapted process with values in non-positive, symmetric $n \times n$ matrices. Let $\lambda > 0$. We have*

$$\mathbb{P}\left((|Z_t| + |X_t|)^* \geq \lambda\right) \leq 2\lambda^{-1}\|X\|_1.$$

Theorem 6.3 (Sparse domination)**.** *Let $X = (X)_t$ be a real valued non-negative continuous path martingale and $Y = (Y)_t$ a vector valued continuous path martingale so that Y is differentially subordinate with respect to X. Let further $Z = (Z)_t$ be a continuous path semi-martingale whose increments satisfy $dZ_t = V_t Z_t dt + dY_t$ with (V_t) continuous adapted process with values in non-positive, symmetric $n \times n$ matrices. Then there exists a sparse domination such that*

$$Z^* \leq 4S(X).$$

We recall that we denote by $Z^* = \sup_{t \geq 0} |Z_t|$ the maximal function associated with Z.

Theorem 6.4. *Let $X = (X_t)_t$ be a real valued non-negative continuous path martingale and $Y = (Y_t)_t$ a vector valued continuous path martingale so that Y is differentially subordinate with respect to X. Let further $Z = (Z_t)_t$ be a continuous path semi-martingale* *whose increments satisfy $dZ_t = V_t Z_t dt + dY_t$ with (V_t) continuous adapted process with values in non-positive, symmetric $d \times d$ matrices. Then there holds the weighted estimate*

$$\|Z^*\|_{L^2(w)} \leq 32 Q_2^{\mathcal{F}}(w) \|X\|_{L^2(w)}$$

and

$$\|Z^*\|_{L^p} \leq 8 \frac{p^2}{p-1} \|X\|_{L^p}.$$

In general for filtered spaces, the A_p characteristic of w (identified with its closure) is

$$Q_p^{\mathcal{F}}(w) = \sup_{\tau} \|\mathbb{E}((\frac{w_\tau}{w})^{\frac{1}{p-1}} \mid \mathcal{F}_\tau)^{p-1}\|_\infty = \sup_{\tau} \|\mathbb{E}(w \mid \mathcal{F}_\tau) \mathbb{E}(w^{\frac{-1}{p-1}} \mid \mathcal{F}_\tau)^{p-1}\|_\infty.$$

In the case of interest to us, the characteristic that appears is the one that corresponds to the filtration used by Li at height y, denoted $\mathcal{F}^{(y)}$. It can be seen, similarly as is known to the Euclidean case, that this characteristic, is equal to the Poisson flow characteristic when $y \to +\infty$.

6.2 The stochastic process Z

In this section, we prove Lemma 6.1 and Theorem 6.3.

Proof. (of Lemma 6.1).

This proof is modelled after the exposition in Wang [80]. We aim to show

$$\mathbb{P}((|Z_t| + |X_t|)^* \geq \lambda) \leq 2\lambda^{-1} \|X\|_1. \tag{6.7}$$

Indeed, it suffices to show the inequality for $\lambda = 1$. To do this, define functions $V, U : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ by

$$V(x,y) = \begin{cases} -2|x| & \text{when } |x| + |y| < 1, \\ 1 - 2|x| & \text{when } |x| + |y| \geq 1. \end{cases}$$

$$U(x,y) = \begin{cases} U_1(x,y) = |y|^2 - |x|^2 & \text{when } |x| + |y| < 1, \\ U_2(x,y) = 1 - 2|x| & \text{when } |x| + |y| \geq 1. \end{cases}$$

Let us first observe that everywhere $V \leq U$ and that $U_1 \leq U_2$.
Define the stopping time

$$T = \inf\{t \geq 0 : |X_t| + |Z_t| \geq 1\}.$$

Then $|X_T| + |Z_T| \geq 1$ and $|X_t| + |Z_t| < 1$ for $t < T$.
We aim to prove that $\mathbb{E}U(X_T, Z_T) \leq 0$, since $V \leq U$ the result will follow (see the end of the argument, where we detail the step). We split

$$\mathbb{E}U(X_T, Z_T) = \mathbb{E}(U(X_T, Z_T)\chi_{\{T>0\}}) + \mathbb{E}(U(X_T, Z_T)\chi_{\{T=0\}})$$

and we show that these contributions are both non-positive.

Part 1: $\{T = 0\}$.
For such ω where $T = 0$ then by definition of T we have $|X_0|+|Z_0| \geq 1$ and $U(X_0, Z_0) = 1 - 2|X_0|$. Assuming that $|Z_0| \leq |X_0|$, then

$$1 \leq |X_0| + |Z_0| \leq 2|X_0|,$$

i.e. $1 - 2|X_0| \leq 0$ and hence

$$\mathbb{E}(U(X_T, Z_T)\chi_{\{T=0\}}) = \mathbb{E}(U(X_0, Z_0)\chi_{\{T=0\}}) \leq 0.$$

Part 2: $\{T > 0\}$.
By simple calculations on the derivatives of U we check that

$$\begin{aligned} \partial_{y_i} U(x, y) &= 2y_i && (6.8)\\ \partial^2_{xx} U(x, y) &= -2, && (6.9)\\ \partial^2_{xy_j} U(x, y) &= 0, && (6.10)\\ \partial^2_{y_i y_j} U(x, y) &= 2\delta_{ij}, && (6.11) \end{aligned}$$

for $|x| + |y| < 1$ and where δ_{ij} is the Kronecker delta.
On $\{T > 0\}$, the process evolves in the set $\{(x, y) : \ |x| + |y| < 1\}$, in the interior of which the function U is twice differentiable, which means that we have the following Itô formula

$$U(X_T, Z_T) = U(X_0, Z_0) + I_1 + \frac{1}{2} I_2,$$

with

$$\begin{aligned} I_1 &= \int_0^T \partial_x U(X_s, Z_s) dX_s + \sum_i \int_0^T \partial_{y_i} U(X_s, Z_s) dZ_s^i \\ I_2 &= \int_0^T \partial^2_{xx} U(X_s, Z_s) d\langle X, X\rangle_s + 2\sum_i \int_0^T \partial^2_{xy_i} U(X_s, Z_s) d\langle X, Z^i\rangle_s \\ &\quad + \sum_i \sum_j \int_0^T \partial^2_{y_i y_j} U(X_s, Z_s) d\langle Z^i, Z^j\rangle_s. \end{aligned}$$

Let's first study I_1:
Recall that Z_t satisfies the following stochastic differential equation

$$dZ_t = V_t Z_t dt + dY_t. \tag{6.12}$$

Now if we replace this formula in the expression of I_1, we will obtain a local martingale part which is

$$\int_0^T \partial_x U(X_s, Z_s) dX_s + \int_0^T \langle \partial_y U(X_s, Z_s), dY_s\rangle$$

and a process

$$A_T = \int_0^T \langle \partial_y U(X_s, Z_s), V_s Z_s\rangle ds.$$

We may assume that the local martingale is a true martingale without loss of generality and hence its expectation is null. As for the process A_T, by (6.8) we have

$$A_T = 2\int_0^T \langle Z_s, V_s Z_s\rangle ds \leq 0.$$

The non-positivity holds because the integrand is non-positive as well, since V takes values in the class of non-positive matrices. Notice that just like in [18], the form of the partial derivative of U in the variable y is crucial.

Now we deal with I_2:
By the formulas (6.9)-(6.11), we obtain that

$$\frac{1}{2}I_2 = (\langle Z, Z\rangle_T - |Z_0|^2 - \langle X, X\rangle_T + |X_0|^2)\chi_{\{T>0\}},$$

and hence it suffices to prove

$$(\langle Z, Z\rangle_T - |Z_0|^2 - \langle X, X\rangle_T + |X_0|^2)\chi_{\{T>0\}} \leq 0, \tag{6.13}$$

for any stopping time T. Recall that for all t we have $dZ_t = V_t Z_t dt + dY_t$. Thus by integrating we have,

$$Z_t - Z_0 = \int_0^t V_s Z_s ds + Y_t - Y_0.$$

Taking the quadratic covariance on both sides we obtain

$$\begin{aligned} \langle Z, Z\rangle_t - |Z_0|^2 &= \langle Y, Y\rangle_t - |Y_0|^2, \quad \forall t \geq 0 \\ &\leq \langle X, X\rangle_t - |X_0|^2 \;\text{ by differential subordination} \end{aligned}$$

which in turn implies that $\mathbb{E}(I_2) \leq 0$.
Finally, $U(X_0, Z_0) = |Z_0|^2 - |X_0|^2 \leq 0$.

It remains to show the weak estimate (6.7):
We have $V \leq U$ everywhere and $\mathbb{E}U(X_T, Z_T) \leq 0$. Applying this result to the stopped processes $X_{(T\wedge t)}$ and $Z_{(T\wedge t)}$ we obtain $\mathbb{E}(U(X_{(T\wedge t)}, Z_{(T\wedge t)}) \leq 0$. Observing that the function $U_2(x, y) = 1 - 2|x|$ is concave and larger than $U_1(x, y) = y^2 - x^2$ on the set $\{|x| + |y| \leq 1\}$, one obtains that $\mathbb{E}(U(X_t, Z_t)) \leq \mathbb{E}(U(X_{(T\wedge t)}, Z_{(T\wedge t)}) <= 0$.
Therefore,

$$\begin{aligned} 0 &\geq \mathbb{E}U(X_t, Z_t) \\ &\geq \mathbb{E}V(X_t, Z_t) \\ &= \mathbb{E}(V(X_t, Z_t)\chi_{\{|X_t|+|Z_t|\geq 1\}}) + \mathbb{E}(V(X_t, Z_t)\chi_{\{|X_t|+|Z_t|<1\}}) \\ &= \mathbb{E}(1 - 2|X_t|\chi_{\{|X_t|+|Z_t|\geq 1\}}) + \mathbb{E}((-2|X_t|)\chi_{\{|X_t|+|Z_t|<1\}}) \\ &= \mathbb{P}(|X_t| + |Z_t| \geq 1) - 2\mathbb{E}|X_t|, \end{aligned}$$

from which we deduce

$$\mathbb{P}((|X_t| + |Z_t|)^* \geq 1) \leq 2\|X\|_1$$

and so the lemma is proved.

□

Proof. (of Theorem 6.3). Now that we have a weak type result by Lemma 6.1, we are able to use a sparse argument as in [21]. Recall for convenience we assumed X non-negative.
To prove that $Z^*(w) \leq 4S(X)(w)$, we will prove that for all $t \geq 0$, we have

$$|Z_t(\omega)| \leq 4S(X)(\omega). \tag{6.14}$$

To do so we will successively construct filtrations and an increasing sequence of stopping times $(T^k)_{k=-1}^{\infty}$ then use the decomposition

$$Z_t = \sum_{k=0}^{\infty} Z_t \chi_{t \in [T^{k-1}, T^k)}. \tag{6.15}$$

First let $\mathcal{F}_t^0 = \mathcal{F}_t$ and consider the processes $Z^0 = \dfrac{Z}{X_0}$ and $Y^0 = \dfrac{Y}{X_0}$ and $X^0 = \dfrac{X}{X_0}$. These fractions are well defined because we have assumed the process X non-negative. Moreover, $X_0 > 0$ because otherwise we would have $Z = X = 0$ everywhere.
Further, we define the set

$$E_0 = \{\omega \in \Omega : Z^{0*}(\omega) \vee X^{0*}(\omega) > 4\}.$$

Obviously, X^0, Y^0 and Z^0 satisfy the assumptions of Lemma 6.1 so we can apply it with $\lambda = 4$ to estimate $\mathbb{P}(E_0)$. Indeed,

$$\begin{aligned}
\mathbb{P}(E_0) &\leq \frac{2}{4}\|X^0\|_1 \\
&= \frac{1}{2}\mathbb{E}(\frac{X}{X_0}) \\
&= \frac{1}{2}\mathbb{E}\left(\mathbb{E}(\frac{X}{X_0}|\mathcal{F}_0)\right) \\
&= \frac{1}{2}\mathbb{E}\left(\frac{\mathbb{E}(X|\mathcal{F}_0)}{X_0}\right) \\
&= \frac{1}{2},
\end{aligned}$$

where we used some properties of the conditional expectation.
We can associate $T^{-1} = 0$ and a stopping time

$$T^0(\omega) = \inf\{t > 0 : |Z_t^0(\omega)| \vee X_t^0(\omega)\} > 4\}$$

as the hitting time of the set $L = (4, \infty)$, which is finite in E_0, almost surely, by definition.
The key of the proof, besides the weak type estimate, relies on recursivity in order to construct a sparse operator. The construction of the sparse decomposition differs from the one in [21] because the operator we want to estimate is a perturbation of Y (which is differentially subordinate to X).
If $t \in [0, T^0)$ then $|Z_t^0| \leq 4$ by definition of T^0 and so

$$|Z_t| \leq 4\mathbb{E}(X|\mathcal{F}_0) \leq 4S(X).$$

If $T^0 = \infty$ then we are done. Otherwise, we work on E_0. Now let the filtration $(\mathcal{F}_t^1)_{t\geqslant 0} = (\mathcal{F}_{t\vee T^0})_{t\geqslant 0}$ and define the following processes

$$Y_t^{(1)} = Y_{t\vee T^0} - Y_{T^0} \quad \text{and} \quad Y_t^1 = \frac{Y_t^{(1)}}{\mathbb{E}(X|\mathcal{F}_{T^0})}\chi_{E_0},$$

$$X_t^{(1)} = X_{t\vee T^0} \quad \text{and} \quad X_t^1 = \frac{X_t^{(1)}}{\mathbb{E}(X|\mathcal{F}_{T^0})}\chi_{E_0},$$

Moreover, we define the process $Z_t^{(1)}$ to satisfy $Z_t^{(1)} = 0$ for $t \leq T^0$ and for $t > T^0$ the stochastic differential equation

$$dZ_t^{(1)} = V_t Z_t^{(1)} dt + dY_t.$$

Notice that for all $t \geq 0$ this process satisfies

$$dZ_t^{(1)} = V_t Z_t^{(1)} dt + dY_t^{(1)}.$$

We can also define an auxiliary process $W_t^{(1)} = Z_t - Z_t^{(1)}$ and notice that for $t \geq T^0$ it solves the homogeneous equation

$$dW_t^{(1)} = V_t W_t^{(1)} dt.$$

Moreover, we have

$$d\langle W_t^{(1)}, W_t^{(1)}\rangle = 2\langle dW_t^{(1)}, W_t^{(1)}\rangle = 2\langle V_t W_t^{(1)} dt, W_t^{(1)}\rangle \leq 0,$$

because V_t takes values in the class of non-positive matrices. So we have shown that $W_t^{(1)}$ is decreasing and for $t \geq T^0$ we have

$$|Z_t - Z_t^{(1)}| = |W_t^{(1)}| \leq |W_{T^0}^{(1)}| = |Z_{T^0}|.$$

Finally, define

$$Z_t^1 = \frac{Z_t^{(1)}}{\mathbb{E}(X|\mathcal{F}_{T^0})}\chi_{E_0}$$

and as before

$$E_1 = \{\omega \in E_0 : Z^{1*}(\omega) \vee X^{1*}(\omega) > 4\}$$

with its corresponding stopping time

$$T^1(\omega) = \inf\{t > 0 : |Z_t^1(\omega)| \vee X_t^1(\omega) > 4\}.$$

The newly defined processes X^1, Y^1 and Z^1 again satisfy the assumptions of Lemma 6.1 (with respect to the new filtration $(\mathcal{F}_t^1)_{t\geqslant 0}$). So we can apply the weak type estimate to obtain

$$\mathbb{P}(E_1) \leq \frac{1}{2}\mathbb{P}(E_0),$$

and moreover for every $A^0 \subseteq E_0$, $A^0 \in \mathcal{F}_{T^0}$

$$\mathbb{P}(A^0 \cap E_1) \leqslant \frac{1}{2}\mathbb{P}(A^0).$$

Now if $t \in [T^0, T^1)$ we can write

$$Z_t = Z_t^{(1)} + (Z_t - Z_t^{(1)}).$$

By definition of T^1 we have $Z_t^1 \leq 4$ and hence

$$Z_t^{(1)} \leq 4\mathbb{E}(X|\mathcal{F}_{T^0})\chi_{E_0}.$$

On the other hand, we know that $|Z_t - Z_t^{(1)}| \leq |Z_{T^0}|$. Since $(Z_t)_t$ is continuous, we have

$$|Z^0_{T^0}| = |Z^0_{T^0-}| \leq 4,$$

which implies

$$|Z_{T^0-}| \leq 4\mathbb{E}(X|\mathcal{F}_0).$$

Finally,

$$|Z_t| \leq |Z_t^{(1)}| + |(Z_t - Z_t^{(1)})| \leq 4\mathbb{E}(X|\mathcal{F}_0) + 4\mathbb{E}(X|\mathcal{F}_{T^0})\chi_{E_0} \leq 4S(X).$$

If $T^1 = \infty$ then we are done. Otherwise, we work on E_1. We keep repeating the procedure by defining successive processes

$$\text{For } k > 0: \; X_t^{(k)} = X_{T^{k-1}} + \int_{T^{k-1}}^{\max\{T^{k-1},t\}} dX_s.$$

and

$$Y_t^{(k)} = \int_{T^{k-1}}^{\max\{T^{k-1},t\}} dY_s.$$

Observe that these are martingales in $\mathcal{F}$ for all k and that $Y^{(k)}$ is differentially subordinate to $X^{(k)}$. Notice that the processes

$$X_t^k = \frac{X_t^{(k)}}{X_{T^{k-1}}} \text{ and } Y_t^k = \frac{Y_t^{(k)}}{X_{T^{k-1}}},$$

are also adapted in $\mathcal{F} = (\mathcal{F}_t)_{t\geq 0}$ since at times $t < T^{k-1}$ these processes are constant and hence adapted and at later times the denominator is measurable. Notice that the event $\{T^{k-1} < t\} \in \mathcal{F}_t$ since T^{k-1} is a stopping time.
Now set $Z_t^{(k)}$ be the process satisfying $Z_t^{(k)} = 0$ if $t \leq T^{k-1}$ and evolving for $t > T^{k-1}$ according to

$$dZ_t^{(k)} = V_t Z_t^{(k)} dt + dY_t$$

with initial condition at time T^{k-1} be set 0. Notice that the so defined process $Z_t^{(k)}$ is adapted to $(\mathcal{F}_t)_{t\geq 0}$ and solves $dZ_t^{(k)} = V_t Z_t^{(k)} dt + dY_t^{(k)}$ for all times with zero increments for $t < T^{k-1}$. For times $t \geq T^{k-1}$ we know that $W_t^{(k)} = (X - Z^{(k)})_t$ solves the homogeneous equation

$$dW_t^{(k)} = V_t W_t^{(k)} dt$$

with initial condition $W_{T^{k-1}}^{(k)} = Z_{T^{k-1}}$. Now observe that

$$d\langle W_t^{(k)}, W_t^{(k)} \rangle = 2\langle dW_t^{(k)}, W_t^{(k)} \rangle = 2\langle V_t W_t^{(k)} dt, W_t^{(k)} \rangle \leq 0$$

again because V_t takes values in the non-positive matrices. So we have for $t \geq T^{k-1}$ that

$$|Z_t - Z_t^{(k)}| = |W_t^{(k)}| \leq |W_{T^{k-1}}^{(k)}| = |Z_{T^{k-1}}|.$$

Using similar arguments as above, we can consider $Z_t^k = \frac{Z_t^{(k)}}{X_{T^{k-1}}}$ and retain these properties, now with respect to martingales X^k and Y^k.
We also define the sets

$$E_k = \{\omega \in E_{k-1} : Z^{k*}(\omega) \vee X^{k*}(\omega) > 4\}$$

and their associated stopping times

$$T^k = \inf\{t > 0 : |Z_t^k(\omega)| \vee X_t^k(\omega) > 4\}.$$

By the above, we know that processes X^k, Y^k and Z^k satisfy the assumptions of the weak type estimate and we thus control $|E_k| \leq \frac{1}{2}\|X^k \chi_{E_{k-1}}\|_1 \leq \frac{1}{2}|E_{k-1}|$.
Consequently for $t \in [T^{k-1}, T^k)$ we get

$$|Z_t| \leq |Z_t^{(k)}| + |Z_t - Z_t^{(k)}| \leqslant \sum_{j=-1}^{k-1} 4\mathbb{E}(X|\mathcal{F}_{T^j})\chi_{E_j} \leq 4S(X),$$

by considering $E_{-1} = \Omega$. Now the decomposition (6.15) gives us the desired estimate (6.14). $\square$

Proof. (of Theorem 6.4). This follows from the sparse domination and the corresponding estimate for the sparse operator, see [21]. One of the reasons why sparse domination is convenient for weighted estimates, is that it allows the characteristic to appear one time, making the estimate optimal in terms of the characteristic of the weight for $p \geq 2$. Another important fact comes from the maximal function that is bounded on $L^p(w)$, independently of the characteristic.
Let U be random variable in $L^2(u)$, where $u = w^{-1}$. By dualizing, we want to prove that

$$\mathbb{E}(S(X)|U|) \leq c_2 Q_2(w)\|X\|_{L^2(w)}\|U\|_{L^2(u)}. \tag{6.16}$$

The left hand side of (6.16) is, by definition of $S(X)$, equal to

$$\mathbb{E}\left(\sum_{j=-1}^{\infty}\mathbb{E}(X|\mathcal{F}_{T^j})(\omega)\chi_{E_j}(\omega)|U|\right) = \mathbb{E}\left(\sum_{j=-1}^{\infty}\mathbb{E}\left[\mathbb{E}(X|\mathcal{F}_{T^j})(\omega)\chi_{E_j}(\omega)|U||\mathcal{F}_{T^j}\right]\right)$$
$$= \mathbb{E}\left(\sum_{j=-1}^{\infty}\mathbb{E}(X|\mathcal{F}_{T^j})(\omega)\chi_{E_j}(\omega)\mathbb{E}(|U||\mathcal{F}_{T^j})(\omega)\right).$$

We can expand the above sum by inserting the weights and so we get

$$\mathbb{E}\Big(\sum_{j=-1}^{\infty} w_{T^j}(w_{T^j})^{-1}u_{T^j}(u_{T^j})^{-1}\mathbb{E}(X|\mathcal{F}_{T^j})(\omega)\chi_{E_j}(\omega)\mathbb{E}(|U||\mathcal{F}_{T^j})(\omega)\Big)$$
$$\leq Q_2(w)\mathbb{E}\big(\sum_{j=-1}^{\infty}(w_{T^j})^{-1}(u_{T^j})^{-1}\mathbb{E}(X|\mathcal{F}_{T^j})(\omega)\chi_{E_j}(\omega)\mathbb{E}(|U||\mathcal{F}_{T^j})(\omega)\big),$$

where the last inequality follows from definition of $Q_2(w)$. Now let

$$X_t^{\text{new}} := \frac{\mathbb{E}(X|\mathcal{F}_t)}{u_t} \quad \text{and} \quad U_t^{\text{new}} := \frac{\mathbb{E}(|U||\mathcal{F}_t)}{w_t}.$$

Then $(X_t^{\text{new}})_t$ is a $\mathbb{E}_u$ martingale and $(U_t^{\text{new}})_t$ is a $\mathbb{E}_w$ martingale where $\mathbb{E}_u(X) := \frac{\mathbb{E}(Xu)}{\mathbb{E}(u)}$ and the same goes for $\mathbb{E}_w$. This means that the left hand side of (6.16) is less that

$$Q_2(w)\mathbb{E}\Big(\sum_{j=-1}^{\infty} X_{T^j}^{\text{new}}U_{T^j}^{\text{new}}\chi_{E_j}(\omega)\Big).$$

For every fixed j, $X_{T^j}^{\text{new}}U_{T^j}^{\text{new}}\chi_{E_j}$ is $\mathcal{F}_{T^j}$ measurable. Hence we can approximate it from below by step functions

$$\sum_k \alpha_k^j \chi_{A_k^j} \nearrow X_{T^j}^{\text{new}}U_{T^j}^{\text{new}}\chi_{E_j},$$

where $(\alpha_k^j)_k$ are constants, $A_k^j \in \mathcal{F}_{T^j}$ are disjoint and $\bigcup_k A_k^j = E_j$. The reason why we define these step functions is because we want to create *plateaux* on which we will use sparse arguments.
We obviously have on A_k^j

$$\alpha_k^j \leq (X^{\text{new}})^*(U^{\text{new}})^*.$$

Moreover, if we denote by $S_k^j = A_k^j \backslash (A_k^j \cap E_{j+1})$, we know by sparsity that

$$\mathbb{P}(A_k^j) \leq 2\mathbb{P}(S_k^j)$$

and that the sets S_k^j are disjoint in both parameters. Hence, one can write for finite sums

$$\begin{aligned}
\mathbb{E}\left(\sum_{j=-1}^{J}\sum_{k}\alpha_k^j\chi_{A_k^j}\right) &= \sum_{j=-1}^{J}\sum_{k}\alpha_k^j\mathbb{P}(A_k^j)\\
&\leq 2\sum_{j=-1}^{J}\sum_{k}\alpha_k^j\mathbb{P}(S_k^j)\\
&= 2\mathbb{E}\left(\sum_{j=-1}^{J}\sum_{k}\alpha_k^j\chi_{S_k^j}\right)\\
&\leq 2\mathbb{E}\left(\sum_{j=-1}^{J}\sum_{k}(X^{\text{new}})^*(U^{\text{new}})^*u^{1/2}w^{1/2}\chi_{S_k^j}\right)\\
&\leq 2\left(\mathbb{E}\left(\sum_{j=-1}^{J}\sum_{k}(X^{\text{new}})^{*2}u\chi_{S_k^j}\right)\right)^{1/2}\\
&\qquad\times\left(\mathbb{E}\left(\sum_{j=-1}^{J}\sum_{k}(U^{\text{new}})^{*2}w\chi_{S_k^j}\right)\right)^{1/2}\\
&\leq 2\left(\mathbb{E}\left((X^{\text{new}})^{*2}u\right)\right)^{1/2}\left(\mathbb{E}\left((U^{\text{new}})^{*2}w\right)\right)^{1/2}\\
&= (\mathbb{E}(u))^{1/2}(\mathbb{E}_u(X^{\text{new}})^{*2}))^{1/2}(\mathbb{E}(w))^{1/2}(\mathbb{E}_w(U^{\text{new}})^{*2}))^{1/2}\\
&\leq 8(\mathbb{E}(u))^{1/2}(\mathbb{E}_u(X^{\text{new}})^{2}))^{1/2}(\mathbb{E}(w))^{1/2}(\mathbb{E}_w(U^{\text{new}})^{2}))^{1/2}\\
&= 8\|X\|_{L^2(w)}\|U\|_{L^2(u)}.
\end{aligned}$$

Letting $J \to \infty$ allows to conclude that

$$\mathbb{E}(S(X)|U|) \leq 8Q_2(w)\|X\|_{L^2(w)}\|U\|_{L^2(u)}.$$

We explain the changes in the unweighted L^p case, $p \in (1, +\infty)$. Let U be a random variable in L^q where q is the conjugate of p. As before we set up by duality

$$\mathbb{E}(S(X)|U|) = \mathbb{E}\left(\sum_{j=-1}^{\infty}\mathbb{E}(X|\mathcal{F}_{T^j})(\omega)\chi_{E_j}(\omega)\mathbb{E}(|U||\mathcal{F}_{T^j})(\omega)\right).$$

For a fixed j, we approximate as before the non negative and measurable function $\mathbb{E}(X|\mathcal{F}_{T^j})(\omega)\chi_{E_j}(\omega)\mathbb{E}(|U||\mathcal{F}_{T^j})(\omega)$ from below by step functions, use the sparse condition, then Holder's inequality and Doob's inequality in L^p and L^q to obtain the desired estimate. The final norm is a product of the constant 4 appearing in Lemma 6.3, the constant 2 arising in the sparse condition and the product of the constants in Doob's inequality $\dfrac{p}{p-1}\dfrac{q}{q-1} = \dfrac{p^2}{p-1}$. □

6.3 The Riesz vector

The proof of the main result now follows from standard arguments.
Following Li [54], recall that

$$X_t = Qf(B_t^X, B_t) - Qf(B_0^X, y).$$

By taking the probabilistic representation of the Riesz transform (6.1) and using the fact that the conditional expectation is L^p-contractive, one can write

$$\begin{aligned}
\|R_\varphi f\|_{L^p}^p &\leq \lim_{y\to\infty} 2^p \|Z_\tau\|_{L^p}^p \\
&\leq \lim_{y\to\infty} 2^p \|Z^*\|_{L^p}^p \\
&\leq \lim_{y\to\infty} (16\frac{p^2}{p-1})^p \|X\|_{L^p}^p \\
&\leq \lim_{y\to\infty} (16\frac{p^2}{p-1})^p \left(\|Qf(B_\tau^X, B_\tau)\|_{L^p}^p + \|Qf(B_0^X, y)\|_{L^p}^p\right) \\
&\leq (16\frac{p^2}{p-1})^p \|f(B_\tau^X)\|_{L^p}^p \\
&\leq (16\frac{p^2}{p-1})^p \|f\|_{L^p}^p,
\end{aligned}$$

Notice that sparse domination itself depends upon the used filtration (and hence y). Here the norm $\|X\|_{L^p}$ is at $t=\infty$, which is τ in our stopped processes. We use that $\|Qf(B_0^X, y)\|_\infty \to 0$ as $y \to \infty$. On $L^2(w)$, we also obtain the announced result since $Q_2^{\mathcal{F}(y)}(w)$ is the A_2 characteristic that corresponds to the filtration when $B_0 = y$ and $\tilde{Q}_2(w)$ is the Poisson flow characteristic.

6.4 Negative curvature case

A natural question is to know whether we can extend these results to a manifold whose Ricci curvature is bounded from below by a negative constant, as opposed to the previous section where it was non negative. It would be interesting to obtain a result independent of the curvature as in [16].
In this section, we assume that the Bakry-Emery curvature is bounded from below i.e

$$\mathrm{Ric}_\varphi \geq -a, \quad a \geq 0.$$

The probabilistic representation formula of the Riesz vector in this setting for a complete manifold with $\mathrm{Ric}_\varphi \geq -a$ is as follows

$$-\frac{1}{2}(R_\varphi f)(x) = \lim_{y\to\infty} \mathbb{E}_y \left[e^{-a\tau} M_\tau \int_0^\tau e^{as} M_s^{-1} dQ^a(f)(B_s^M, B_s) dB_s | B_\tau^M = x \right], \quad (6.17)$$

where

- $Q^a(f)(x,y) = e^{-y\sqrt{aId-\Delta_\varphi}}f(x)$ is the Poisson semigroup associated to the operator $aId - \Delta_\varphi$;

- $\tau = \inf\{t > 0 : B_t = 0\}$ is the stopping time upon hitting the boundary of the upper half space $X \times \mathbb{R}_+$;

- M_t is the solution to the matrix-valued stochastic differential equation

$$dM_t = V_t M_t dt, \quad M_0 = Id,$$

  for some adapted and continuous process $(V_t)_{t\geq 0}$ taking values in the set of symmetric and non-positive $n \times n$ matrices.

Equivalently, one can rewrite this formula as

$$-\frac{1}{2}(R_\varphi f)(x) = \lim_{y\to\infty} \mathbb{E}_y \left[Z^a_\tau | B^M_\tau = x \right],$$

where $(Z^a_t)_{t\geq 0}$ is a semi-martingale defined thanks to the auxiliary martingales $(X^a_t)_{t\geq 0}$ and $(Y^a_t)_{t\geq 0}$ as follows

$$X^a_t = \int_0^t (\nabla, \partial_y) Q^a f(B^M_s, B_s)(U_s dW_s, dB_s),$$

$$Y^a_t = \int_0^t \nabla Q^a f(B^M_s, B_s) dB_s,$$

$$Z^a_t = e^{-at} M_t \int_0^t e^{as} M_s^{-1} dY^a_s,$$

A first approach would be to apply the method in the previous sections to the processes $(Z^a_t)_t$ and $(X^a_t)_t$ and notice that by replacing V_t by $aI_n - V_t$ the sparse decomposition remains the same. The problem appears in the final steps when applying Itô formula on X^a_t since $X^a_t = Q^a f(B^X_t, B_t) - Q^a f(B^X_0, y) - a\int_0^t Q^a f(B^X_s, B_s)ds$. When $t = \tau$, we need to control the term $a\int_0^\tau Q^a f(B^X_s, B_s)ds$.
In the unweighted case, we bypass this problem by using Lenglart-Lépingle-Pratelli inequality and obtain a result that does not depend on a (with a slightly different constant in p). In the weighted case, we define the process

$$\begin{aligned}\widetilde{X}_t &= X^a_t + a\int_0^t Q^a f(B^M_s, B_s)ds + Q^a f(B^M_0, B_0) \\ &= Q^a f(B^M_t, B_t),\end{aligned}$$

by Itô formula. If we denote

$$A_t = a\int_0^t Q^a f(B^M_s, B_s)ds + Q^a f(B^M_0, B_0),$$

we can see that A_t is an increasing finite version process. Now we can write

$$\widetilde{X}_t = X_t^a + A_t$$

and conclude that $(\widetilde{X}_t)_{t\geq 0}$ is a submartingale. Indeed,

$$\begin{aligned}\mathbb{E}(\widetilde{X}_t|\mathcal{F}_s) &= \mathbb{E}(X_t^a + A_t|\mathcal{F}_s) \geq \mathbb{E}(X_t^a + A_s|\mathcal{F}_s) \\ &= \mathbb{E}(X_t^a|F_s) + \mathbb{E}(A_s|\mathcal{F}_s) = X_s^a + A_s = \widetilde{X}_s.\end{aligned}$$

Instead of proving a weak-type estimate for the martingale $X^a = (X_t^a)$ as we suggested earlier, we will prove a weak-type estimate for the submartingale $(\widetilde{X}_t)_{t\geq 0}$.

Lemma 6.2 (Weak-type estimate)**.** *Let X and Y be continuous path martingales so that Y is differentially subordinate with respect to X. Let further $\widetilde{X}_t = X_t + A_t$ be a non-negative* uniformly integrable *submartingale where A_t is a FV increasing process. Finally, let Z be a continuous path semi-martingale such that $|Z_0| \leq |\widetilde{X}_0|$ and whose increments satisfy $dZ_t = (V_t - aI_n)Z_t dt + dY_t$, where V_t is a continuous adapted process with values in non-positive, symmetric $n \times n$ matrices and $a \geq 0$. Then for all $\lambda > 0$ we have*

$$\mathbb{P}\left((|Z_t| + |\widetilde{X}_t|)^* \geq \lambda\right) \leq 2\lambda^{-1}\|\widetilde{X}\|_1. \tag{6.18}$$

Remark 6.1. *In the statement of the above lemma, by $\widetilde{X}$ we actually mean $\widetilde{X}_\infty$, i.e. the limit of the submartingale $(\widetilde{X}_t)$.*

Sketch of proof for the weak type estimate. This proof is following the same outline as the proof of the weak-type estimate in the previous section (which is modelled after the exposition in Wang).

Indeed, it suffices to show the inequality (6.18) for $\lambda = 1$. To do this, we define functions $U, V : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ by

$$V(x,y) = \begin{cases} -2|x| & \text{when } |x| + |y| < 1, \\ 1 - 2|x| & \text{when } |x| + |y| \geq 1, \end{cases}$$

and

$$U(x,y) = \begin{cases} |y|^2 - |x|^2 & \text{when } |x| + |y| < 1, \\ 1 - 2|x| & \text{when } |x| + |y| \geq 1. \end{cases}$$

First we can observe that everywhere $V \leq U$. We define the stopping time

$$T = \inf\{t \geq 0 : |\widetilde{X}_t| + |Z_t| \geq 1\}$$

and notice that $|\widetilde{X}_T| + |Z_T| \geq 1$ and $|\widetilde{X}_t| + |Z_t| < 1$ for $t < T$.
We want to prove

$$\mathbb{E}U(\widetilde{X}_T, Z_T) \leq 0.$$

Since $V \le U$, the above inequality will give us exactly what we need. We split

$$\mathbb{E}U(\widetilde{X}_T, Z_T) = \mathbb{E}(U(\widetilde{X}_T, Z_T)\chi_{\{T=0\}}) + \mathbb{E}(U(\widetilde{X}_T, Z_T)\chi_{\{T>0\}})$$

and we show that these contributions both satisfy some desired inequalities.

Part 1: $T = 0$.
For such ω where $T = 0$ by definition of T we have $|\widetilde{X}_0| + |Z_0| \ge 1$ and $U(\widetilde{X}_0, Z_0) = 1 - 2|\widetilde{X}_0| \le 0$. Since $|Z_0| \le |\widetilde{X}_0|$, we have

$$1 \le |\widetilde{X}_0| + |Z_0| \le 2|\widetilde{X}_0|,$$

i.e. $1 - 2|\widetilde{X}_0| \le 0$ and hence

$$\mathbb{E}(U(\widetilde{X}_T, Z_T)\chi_{\{T=0\}}) = \mathbb{E}(U(\widetilde{X}_0, Z_0)\chi_{\{T=0\}}) \le 0.$$

Part 2: $T > 0$.
We use the same calculations on the derivatives of U on $|x| + |y| < 1$ and apply Itô formula to obtain

$$U(\widetilde{X}_T, Z_T) = U(\widetilde{X}_0, Z_0) + I_1 + \frac{1}{2}I_2,$$

with

$$\begin{aligned} I_1 &= \int_0^T \partial_x U(\widetilde{X}_s, Z_s) d\widetilde{X}_s + \int_0^T \sum_i \partial_{y_i} U(\widetilde{X}_s, Z_s) dZ_s^i \\ I_2 &= \int_0^T \partial^2_{xx} U(\widetilde{X}_s, Z_s) d\langle \widetilde{X}, \widetilde{X}\rangle_s + 2\sum_i \int_0^T \partial^2_{xy_i} U(\widetilde{X}_s, Z_s) d\langle \widetilde{X}, Z^i\rangle_s \\ &\quad + \sum_i \sum_j \int_0^T \partial^2_{y_i y_j} U(\widetilde{X}_s, Z_s) d\langle Z^i, Z^j\rangle_s. \end{aligned}$$

Let's first study I_1:
Recall that Z_t satisfies the following stochastic differential equation

$$dZ_t = (V_t - aI_n)Z_t dt + dY_t$$

and that

$$\widetilde{X}_t = X_t + A_t \qquad \text{which implies} \qquad d\widetilde{X}_t = dX_t + dA_t,$$

where (X_t) is a martingale, and (A_t) is an increasing FV process.

Now if we replace $d\widetilde{X}_s$ and dZ_s in the expression of I_1 by the above formulas, we will obtain a local martingale part which is

$$\int_0^T \partial_x U(\widetilde{X}_s, Z_s) dX_s + \int_0^T \langle \partial_y U(\widetilde{X}_s, Z_s), dY_s\rangle$$

and a process

$$N_T = \int_0^T \partial_x U(\widetilde{X}_s, Z_s) dA_s + \int_0^T \langle \partial_y U(\widetilde{X}_s, Z_s), (V_s - aI_n) Z_s \rangle ds.$$

We may assume that the local martingale is a true martingale without loss of generality and hence its expectation is null. As for the process N_T, we have

$$N_T = -2 \int_0^T \widetilde{X}_s dA_s + 2 \int_0^T \langle Z_s, (V_s - aI_n) Z_s \rangle ds \leq 0.$$

Since (A_t) in an increasing FV proces we have $dA_s \geq 0$ which together with the non-negativity of $(\widetilde{X}_t)$ gives us

$$-2 \int_0^T \widetilde{X}_s dA_s \leq 0.$$

The non-positivity of the second integral also holds since $V - aI_n$ takes values in the class of non-positive matrices, so we have $\langle Z_s, (V_s - aI_n) Z_s \rangle \leq 0$.
In conclusion, taking the expectation of I_1 gives us

$$\mathbb{E}(I_1) \leq 0.$$

Now we deal with I_2:
Using the calculations on the derivatives of U we obtain that

$$\frac{1}{2} I_2 = (\langle Z, Z \rangle_T - |Z_0|^2 - \langle \widetilde{X}, \widetilde{X} \rangle_T + |\widetilde{X}_0|^2) \chi_{\{T>0\}}.$$

Recall that for all t we have $dZ_t = (V_t - aI_n) Z_t dt + dY_t$. Thus by integrating we have,

$$Z_t - Z_0 = \int_0^t (V_s - aI_n) Z_s ds + Y_t - Y_0.$$

Taking the quadratic covariance on both sides we obtain

$$\begin{aligned} \langle Z, Z \rangle_t - |Z_0|^2 &= \langle Y, Y \rangle_t - |Y_0|^2, \quad \forall t \geq 0 \\ &\leq \langle X, X \rangle_t - |X_0|^2 \ \text{ by differential subordination.} \end{aligned}$$

Since $\widetilde{X}_t = X_t + A_t$, where (A_t) is a FV process we have for all $t > 0$

$$\begin{aligned} \langle \widetilde{X}, \widetilde{X} \rangle_t &= |\widetilde{X}_0|^2 + \langle \widetilde{X}, \widetilde{X} \rangle_t^c \\ &= |\widetilde{X}_0|^2 + \langle X, X \rangle_t^c \\ &= |\widetilde{X}_0|^2 + \langle X, X \rangle_t - |X_0|^2 \end{aligned}$$

and so

$$\langle Z, Z \rangle_t - |Z_0|^2 \leq \langle X, X \rangle_t - |X_0|^2 = \langle \widetilde{X}, \widetilde{X} \rangle_t - |\widetilde{X}_0|^2,$$

which gives us

$$(\langle Z, Z\rangle_T - |Z_0|^2 - \langle \widetilde{X}, \widetilde{X}\rangle_T + |\widetilde{X}_0|^2)\chi_{\{T>0\}} \le 0.$$

Taking the expectation of the above inequality implies that

$$\mathbb{E}(\frac{I_2}{2}) \le 0.$$

Finally, since $T > 0$ we have $U(\widetilde{X}_0, Z_0) = |Z_0|^2 - |\widetilde{X}_0|^2 \le 0$ and taking the expectation in the previous Itô formula implies

$$\mathbb{E}(U(\widetilde{X}_T, Z_T)\chi_{\{T>0\}}) \le \mathbb{E}(|\widetilde{X}_0|^2) - \mathbb{E}(|X_0|^2).$$

It remains to show the weak estimate. We have $V \le U$ everywhere and $\mathbb{E}U(\widetilde{X}_T, Z_T) \le 0$. Therefore we may show as in the previous section that

$$0 \ge \mathbb{P}(|\widetilde{X}_t| + |Z_t| \ge 1) - 2\mathbb{E}|\widetilde{X}_t|,$$

from which we deduce

$$\mathbb{P}((|\widetilde{X}_t| + |Z_t|)^* \ge 1) \le 2\|\widetilde{X}\|_1$$

Finally, we can replace $(\widetilde{X}_t)$ and (Z_t) in the above inequality with $(\lambda^{-1}\widetilde{X}_t)$ and $(\lambda^{-1}Z_t)$ to get (6.18)

$$\mathbb{P}((|\widetilde{X}_t| + |Z_t|)^* \ge \lambda) \le 2\lambda^{-1}\|\widetilde{X}\|_1$$

and so the lemma is proved. □

The sparse domination also follows the outline of the previous section, but we have to change the definition of the sparse operator, because $(\widetilde{X}_t)$ is not a martingale.

Theorem 6.5 (Sparse domination)**.** *Let X and Y be continuous path martingales so that Y is differentially subordinate with respect to X. Let further $\widetilde{X}_t = X_t + A_t$ be a non-negative* uniformly integrable *submartingale where A_t is a FV increasing process. Finally, let Z be a continuous path semi-martingale such that $|Z_0| \le |\widetilde{X}_0|$ and whose increments satisfy $dZ_t = (V_t - aI_n)Z_t dt + dY_t$, where V_t is a continuous adapted process with values in non-positive, symmetric $n \times n$ matrices and $a \ge 0$. Then there exists a sparse domination such that*

$$Z^*(\omega) \le 4S(\widetilde{X})(\omega), \tag{6.19}$$

where

$$S(\widetilde{X})(\omega) = \sum_{j=-1}^{\infty} \mathbb{E}(\widetilde{X}|\mathcal{F}_{T^j})(\omega)\chi_{E_j}(\omega)$$

and E_j and T^j are defined as in the previous section.

Remark 6.2. *We needed to replace $\widetilde{X}_{T^j}$ by $\mathbb{E}(\widetilde{X}|\mathcal{F}_{T^j})$ because we are dealing with submartingales and unlike in the martingale case, we don't have an equality but*

$$\mathbb{E}(\widetilde{X}|\mathcal{F}_{T^j}) \geq \widetilde{X}_{T^j}. \tag{6.20}$$

We also need the sets E_j to satisfy the sparse condition

$$\forall A_j \subset E_j,\ A_j \in \mathcal{F}_{T_j} \text{ there holds } \mathbb{P}(A_j \cap E_{j+1}) \leq \frac{1}{2}\mathbb{P}(A_j).$$

To prove the above sparse condition we have to use WTE, but unlike in the martingale case we cannot control the L^1 norm of

$$\frac{\widetilde{X}}{\widetilde{X}_0} \qquad or \qquad \frac{\widetilde{X}}{\widetilde{X}_{T^j}}$$

because of (6.20). On the other hand, we can control the L^1 norm of

$$\frac{\widetilde{X}}{\mathbb{E}(\widetilde{X}|\mathcal{F}_0)} \qquad and \qquad \frac{\widetilde{X}}{\mathbb{E}(\widetilde{X}|\mathcal{F}_{T^j})}.$$

Now that we changed the definition of the sparse operator, the proof of this lemma is exactly the same as in the case $a = 0$. The proof of the L^2 estimate is also the same and to save the length of this chapter, we refer the interested reader to the previous section ($a = 0$).

Chapter 7

Applications and Perspectives

We briefly explore in this chapter a few tracks that fit into the same framework as this thesis. We suggest the following leads to the interested reader

1. The Beurling-Ahlfors operator;
2. Fractional integrals;
3. $L^p(w)$ boundedness of the Riesz transforms on Riemannian manifolds.

7.1 The Beurling-Ahlfors operator

We present a weighted estimate for the Beurling-Ahlfors operator acting on manifolds. The interest of this operator comes from a famous borderline regularity problem in [3]. The idea is to use again a martingale transform representation formula for the Beurling-Ahlfors transform extended to 1-forms over complete Riemannian manifolds by [56, 57]. Throughout this section, let (X, g, μ_φ) and B_t^X be defined as in the previous chapter. On manifolds, the Beurling-Ahlfors operator is given by

$$B = (d_\varphi^* d - dd_\varphi^*)(\vec{\Delta}_\varphi)^{-1},$$

where d denotes the exterior derivative, d_φ^* its adjoint operator and $\vec{\Delta}_\varphi = dd_\varphi^* + d_\varphi^* d$ is the (weighted) Hodge-de Rham Laplacian acting on $1-$forms.
Recall that by the Weitzenbock formula, the Hodge de Rham Laplacian is related to the rough Laplacian by

$$\vec{\Delta}_\varphi = -\mathrm{Tr}\nabla^2 + \mathrm{Ric}_\varphi \,.$$

Define the matrices $A_1 = (a_i a_j^*)$ and $A_2 = (a_i^* a_j)$ as in [56, Section 3] and

$$B = A_2 - A_1.$$

Unlike the Riesz transform, we define the background heat semigroup generated by the Hodge de Rham Laplacian by

$$P\vec{g}(x, T-s) = e^{-(T-s)\vec{\Delta}_\varphi}\vec{g}(x), \ \forall x \in X, \ s \in [0, T], \ \vec{g} \in C_0^\infty(\Lambda T^* X),$$

for any fixed $T > 0$.
The probabilistic representation of the Beurling-Ahlfors operator for complete Riemannian manifolds with $\mathrm{Ric}_\varphi \geq 0$ is as follows

$$S^T_{A_i}\vec{g}(x) = \mathbb{E}\left(M_T \int_0^T M_t^{-1} A_i \nabla P\vec{g}(B^X_t, T-t) dX_t | B^X_T = x\right), \ i = 1, 2$$

and

$$S^T_B\vec{g}(x) = 2 \lim_{T\to\infty} \mathbb{E}\left(M_T \int_0^T M_t^{-1} B \nabla P\vec{g}(B^X_t, T-t) dX_t | B^X_T = x\right).$$

Let

$$Z_t = M_t \int_0^t M_s^{-1} B \nabla P\vec{g}(B^X_s, T-s) dB^X_s$$

and

$$X_t = M_t \int_0^t M_s^{-1} \nabla P\vec{g}(B^X_s, T-s) dB^X_s.$$

Let $Y_t = \frac{1}{\|B\|_{op}} Z_t$, where $\|\cdot\|_{op}$ denotes the operator norm. Y_t satisfies

$$dY_t = V_t Y_t + \frac{B}{\|B\|_{op}} dX_t$$

and it is differentially subordinate to X_t. We repeat the previous proof in Chapter 6 verbatim and obtain

$$\|Z^*\|_{L^p(T^*X,w)} \leq 16 \frac{p^2}{p-1} \|B\|^2_{op} \widetilde{Q}_p(w)^{\max(1, \frac{1}{p-1})} \|X\|_{L^p(X,w)}. \tag{7.1}$$

7.2 Fractional integrals

Another interesting family of operators is the fractional integrals associated to a Feller semigroup $(T_t)_t$ whose Varopoulos dimension is d. We define the fractional integrals of order $\alpha \in (0, d)$ as follows

$$I_\alpha f(x) = \frac{1}{\Gamma(\frac{\alpha}{2})} \int_0^\infty t^{\alpha/2-1} T_t f(x) dt.$$

Again, we may consider the probabilistic representation of the fractional integrals studied on $\mathbb{R}^d$ in [1] and on locally compact spaces in [47] and extend it to complete Riemannian manifolds using Li's approach in [58]. Using the same notations as in Chapter 6, we obtain

$$S^a_\alpha f(x) = \mathbb{E}^a\left[M_\tau \int_0^\tau M_s^{-1} B^\alpha_s \frac{\partial Q(f)}{\partial y}(B^X_s, B_s) dB_s | B^X_\tau = x\right].$$

Then there exists a constant $C_{\alpha,d} > 0$ such that

$$S^a_\alpha f \underset{a\to+\infty}{\longrightarrow} C_{\alpha,d} I_\alpha f,$$

in the distributional sense.

If one can think of a function similar to Wang's one [80] to prove the boundedness of the fractional integrals, we may hope to obtain some interesting results concerning the Hardy Littlewood Sobolev inequalities.

7.3 $L^p(w)$ boundedness of the Riesz transforms on Riemannian manifolds

We refer the reader to Theorem 6.4. We would like to extend this result to $L^p(w)$ for any $p \in (1, +\infty)$. Unfortunately, the proof in the case $p \neq 2$ fails if we rewrite the proof of Theorem 6.4 and introduce weights u and w such that $u^p w = u$. However, we may hope to obtain some positive results using the probabilistic extrapolation theorem, used for instance in [23].

Bibliography

[1] David Applebaum and Rodrigo Bañuelos. Probabilistic approach to fractional integrals and the Hardy-Littlewood-Sobolev inequality. In *Analytic methods in interdisciplinary applications*, volume 116 of *Springer Proc. Math. Stat.*, pages 17–40. Springer, Cham, 2015.

[2] Nicola Arcozzi. Riesz transforms on compact Lie groups, spheres and Gauss space. *Ark. Mat.*, 36(2):201–231, 1998.

[3] K. Astala, T. Iwaniec, and E. Saksman. Beltrami operators in the plane. *Duke Math. J.*, 107(1):27–56, 2001.

[4] P. Auscher and J.M. Martell. Weighted norm inequalities, off-diagonal estimates and elliptic operators. *Contemporary Mathematics*, pages 61–83, 2010.

[5] R. Ba nuelos. The foundational inequalities of D. L. Burkholder and some of their ramifications. *Illinois J. Math.*, 54:789–868, 2010.

[6] Rodrigo Bañuelos and Fabrice Baudoin. Martingale transforms and their projection operators on manifolds. *Potential Anal.*, 38(4):1071–1089, 2013.

[7] Rodrigo Bañuelos and Prabhu Janakiraman. L^p-bounds for the Beurling-Ahlfors transform. *Trans. Amer. Math. Soc.*, 360(7):3603–3612, 2008.

[8] Rodrigo Bañuelos and Adam Osekowski. Sharp martingale inequalities and applications to Riesz transforms on manifolds, Lie groups and Gauss space. *J. Funct. Anal.*, 269(6):1652–1713, 2015.

[9] Rodrigo Bañuelos and Gang Wang. Sharp inequalities for martingales with applications to the Beurling-Ahlfors and Riesz transforms. *Duke Math. J.*, 80(3):575–600, 1995.

[10] D. Bakry. Transformations de Riesz pour les semi-groupes symétriques. II. étude sous la condition $\Gamma_2 \geq 0$. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 145–174. Springer, Berlin, 1985.

[11] D. Bakry. Un critère de non-explosion pour certaines diffusions sur une variété riemannienne complète. *Comptes rendus de l'Académie des Sciences Paris*, 303(1):23–26, 1986.

[12] D. Bakry. Etude des transformations de Riesz dans les variétés riemanniennes à courbure de Ricci minorée. *Séminaire de probabilités de Strasbourg*, 1247(1), 1987.

[13] D. Bakry and Michel Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

[14] F. Bernicot, D. Frey, and S. Petermichl. Sharp weighted norm estimates beyond Calderon-Zygmund theory. *Anal. PDE*, 9:1079–1113, 2016.

[15] S. Buckley. Summation condition on weights. *Michigan Math. J.*, 40:153–170, 1993.

[16] A. Carbonaro and O. Dragicevic. Bellman function and linear dimension-free estimates in a theorem of Bakry. *J. Funct. Anal.*, 265(7):1085–1104, 2013.

[17] D.V. Cruz Uribe, J.M. Martell, and C. Pérez. Weights, extrapolation and the theory of Rubio de Francia. *Operator Theory: Advances and Applications*, 215, 2011.

[18] K Dahmani. Sharp dimension free bound for the Bakry-Riesz vector. arXiv:1611.07696, 2016.

[19] K Dahmani, K Domelevo, and S Petermichl. Dimensionless l^p estimates for the Riesz vector on manifolds. arXiv:1802.00366, 2018.

[20] E.B. Davies. *Heat kernel and spectral theory.* Cambridge university press, 1989.

[21] K. Domelevo and S. Petermichl. A sharp maximal inequality for differentially subordinate martingales under a change of law. *arXiv:1607.06319*, 2016.

[22] K. Domelevo and S. Petermichl. Differential subordination under change of law. *arXiv:1604.01606*, 2016.

[23] K. Domelevo, S. Petermichl, and J. Wittwer. A linear dimensionless bound for the weighted Riesz vector. *hal-01097113*, 2014.

[24] O. Dragicevic, L. Grafakos, M.C. Pereyra, and S. Petermichl. Extrapolation and sharp norm estimates for classical operators on weighted Lebesgue spaces. *Publ. Mat.*, • 49:73–91, 2005.

[25] Oliver Dragicević and Alexander Volberg. Bellman function for the estimates of Littlewood-Paley type and asymptotic estimates in the $p-1$ problem. *C. R. Math. Acad. Sci. Paris*, 340(10):731–734, 2005.

[26] Javier Duoandikoetxea. *Fourier analysis*, volume 29 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001. Translated and revised from the 1995 Spanish original by David Cruz-Uribe.

[27] J. Eells and K.D. Elworthy. Stochastic dynamical systems. *Control theory and topics in functional analysis*, pages 179–185, 1976.

[28] David Elworthy. Geometric aspects of diffusions on manifolds. In *École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87*, volume 1362 of *Lecture Notes in Math.*, pages 277–425. Springer, Berlin, 1988.

[29] M. Emery. *Stochastic Calculus in Manifolds.* Universitext, Springer,, 1989.

[30] K.J. Engel and R. Nagel. One-parameter semigroups for linear evolution equations. *Graduate Texts in Mathematics, Springer-Verlag*, 194, 2000.

[31] R. Fefferman, C. Kenig, and J. Pipher. The theory of weights and the Dirichlet problem for elliptic equations. *Ann. of Math.*, 134(1):65–124, 1991.

[32] Sylvestre Gallot, Dominique Hulin, and Jacques Lafontaine. *Riemannian geometry.* Universitext. Springer-Verlag, Berlin, third edition, 2004.

[33] L. Grafakos. *Classical Fourier Analysis* $\bullet$, volume 249. Graduate Texts in Mathematics, 2008.

[34] Loukas Grafakos. *Classical and modern Fourier analysis.* Pearson Education, Inc., Upper Saddle River, NJ, 2004.

[35] R.F. Gundy and N. Varopoulos. Les transformations de Riesz et les intégrales stochastiques. *C.R. Acad. Sci. Paris*, 289:13–16, 1979.

[36] Richard F. Gundy. Sur les transformations de Riesz pour le semi-groupe d'Ornstein-Uhlenbeck. *C. R. Acad. Sci. Paris Sér. I Math.*, 303(19):967–970, 1986.

[37] Richard F. Gundy and Martin L. Silverstein. On a probabilistic interpretation for the Riesz transforms. In *Functional analysis in Markov processes (Katata/Kyoto, 1981)*, volume 923 of *Lecture Notes in Math.*, pages 199–203. Springer, Berlin-New York, 1982.

[38] H. Helson and G. Szego. A problem in prediction theory. *Ann. Mat. Pura Appl.*, pages 107–138, 1960.

[39] E.P. Hsu. *Stochastic Analysis on Manifolds*, volume 38. Graduate Studies in Mathematics, 2001.

[40] S. Hukovic, S. Treil, and A. Volberg. The Bellman functions and sharp weighted inequalities for square functions. In *Complex analysis, operators, and related topics*, volume 113 of *Oper. Theory Adv. Appl.*, pages 97–113. Birkhäuser, Basel, 2000.

[41] R. Hunt, B. Muckenhoupt, and R. Wheeden. Weighted norm inequalities for the conjugate function and Hilbert transform. *Trans. Amer. Math. Soc.*, 176:227–251, 1973.

[42] T. Hytonen. The sharp weighted bound for general Calderón-Zygmund operators. *Ann. of Math.*, 175(3):1473–1506, 2012.

[43] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*. North-Holland Mathematical Library, 1981.

[44] Chris J. Isham. *Modern differential geometry for physicists*, volume 61 of *World Scientific Lecture Notes in Physics*. World Scientific Publishing Co., Inc., River Edge, NJ, second edition, 1999.

[45] K. Itô. The Brownian motion and tensor fields on Riemannian manifolds. In *Proceedings of the International Congress of Mathematics*, 1962.

[46] Tadeusz Iwaniec and Gaven Martin. Riesz transforms and related singular integrals. *J. Reine Angew. Math.*, 473:25–57, 1996.

[47] Daesung Kim. Martingale transforms and the Hardy-Littlewood-Sobolev inequality for semigroups. *Potential Anal.*, 45(4):795–807, 2016.

[48] N. Kobayashi and K. Nomizu. *Foundations of differential geometry*, volume 1. Interscience Publishers, 1963.

[49] M.T. Lacey. An elementary proof of the A_2 Bound. *Israel J. Math.*, 2015. to appear.

[50] J.M. Lee. *Riemmannian Manifolds an introduction to curvature*. Graduate texts in mathematics, 1997.

[51] J.M. Lee. *Introduction to smooth manifolds*. Graduate texts in mathematics, 2000.

[52] A Lerner and Nazarov F. Intuitive dyadic calculus: the basics. arXiv:1508.05639, 2015.

[53] A.K. Lerner. A simple proof of the A_2 conjecture. *Int. Math. Res. Not.*, 2013.

[54] X.-D. Li. Martingale transforms and L^p-norm estimates of Riesz transforms on complete Riemannian manifolds. *Probab. Theory Relat. Fields*, 2008.

[55] X.-D Li. On the l^p-estimates for Beurling-Ahlfors and Riesz transforms on Riemannian manifolds. arXiv:1304.1168, 2013.

[56] Xiang-Dong Li. On the weak L^p-Hodge decomposition and Beurling-Ahlfors transforms on complete Riemannian manifolds. *Probab. Theory Related Fields*, 150(1-2):111–144, 2011.

[57] Xiang-Dong Li. Erratum to: On the weak L^p Hodge decomposition and Beurling-Ahlfors transforms on complete Riemannian manifolds [mr2800906]. *Probab. Theory Related Fields*, 159(1-2):409–411, 2014.

[58] Xiang-Dong Li. Erratum to “Riesz transforms on forms and L^p-Hodge decomposition on complete Riemanian manifolds” [mr2677005]. *Rev. Mat. Iberoam.*, 30(1):369–370, 2014.

[59] P. Malliavin. Formule de la moyenne, calcul de perturbations et théorèmes d'annulation pour les formes harmoniques. *J. Funct. Analy.*, pages 274–291, 1974.

[60] P.-A. Meyer. Transformations de Riesz pour les lois gaussiennes. In *Seminar on probability, XVIII*, volume 1059 of *Lecture Notes in Math.*, pages 179–193. Springer, Berlin, 1984.

[61] F. Nazarov, S. Treil, and A. Volberg. The Bellman functions and two-weight inequalities for Haar multipliers. *J. Amer. Math. Soc.*, 12, 1999.

[62] F. Nazarov, S. Treil, and A. Volberg. Bellman function in stochastic control and harmonic analysis. In *Systems, approximation, singular integral operators, and related topics (Bordeaux, 2000)*, volume 129 of *Oper. Theory Adv. Appl.*, pages 393–423. Birkhäuser, Basel, 2001.

[63] F. L. Nazarov and S. R. Treil. The hunt for a Bellman function: applications to estimates for singular integral operators and to other classical problems of harmonic analysis. *Algebra i Analiz*, 8(5), 1996.

[64] E.M. Ouhabaz. *Analysis of Heat Equations on Domains.* LMS-31, 2004.

[65] S. Petermichl. The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p characteristic. *Amer. J. Math.*, 129(5):1355–1375, 2007.

[66] S. Petermichl. The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical $A(p)$ characteristic. *American Journal of Mathematics*, 129(5):1355–1375, 2007.

[67] S. Petermichl. The sharp weighted bound for the Riesz transforms. *American Mathematical Society*, 136(4):1237–1249, 2008.

[68] S. Petermichl and S. Pott. An estimate for weighted Hilbert transform via square functions. *Trans. Amer. Math. Soc.*, 354(4):1699–1703, • .

[69] S. Petermichl and A. Volberg. Heating of the Ahlfors-Beurling operator: weakly quasiregular maps on the plane are quasiregular. *Duke Math. J.*, 112(2):281–305, 2002.

[70] S. Petermichl and J. Wittwer. A sharp estimate for the weighted Hilbert transform via Bellman functions. *Michigan Math. J.*, 2002.

[71] S. K. Pichorides. On the best values of the constants in the theorems of M. Riesz, Zygmund and Kolmogorov. *Studia Math.*, 44:165–179. (errata insert), 1972. Collection of articles honoring the completion by Antoni Zygmund of 50 years of scientific activity, II.

[72] Gilles Pisier. Riesz transforms: a simpler analytic proof of P.-A. Meyer's inequality. In *Séminaire de Probabilités, XXII*, volume 1321 of *Lecture Notes in Math.*, pages 485–501. Springer, Berlin, 1988.

[73] Philip E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2005. Second edition. Version 2.1, Corrected third printing.

[74] Marcel Riesz. Sur les fonctions conjuguées. *Math. Z.*, 27(1):218–244, 1928.

[75] J.L. Rubio De Francia. Factorization theory and A_p weights. *Amer. J. Math.*, 106(3):533–547, 1984.

[76] Peter Sjögren. Operators associated with the Hermite semigroup—a survey. In *Proceedings of the conference dedicated to Professor Miguel de Guzmán (El Escorial, 1996)*, volume 3, pages 813–823, 1997.

[77] Elias M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

[78] Elias M. Stein. *Topics in harmonic analysis related to the Littlewood-Paley theory*. Annals of Mathematics Studies, No. 63. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1970.

[79] R. Strichartz. Analysis of the Laplacian on the complete Riemannian manifold. *J. Funct. Anal*, 1983.

[80] Gang Wang. Differential subordination and strong differential subordination for continuous-time martingales and related sharp inequalities. *Ann. Probab.*, 23(2):522–551, 1995.

[81] G. Wei and W. Wylie. Comparison geometry for the Bakry-Emery Ricci tensor. *J. Differential Geom.*, 2009.

[82] J. Wittwer. A sharp estimate on the norm of the martingale transform. *Mathematical Research Letters*, 7(1):1–12, January 2000.