WI5005

Generation of wavelets by semigroups

Timothy de Moor

Delft Institute of Applied Mathematics

to obtain the degree of Master of Science

at the Delft University of Technology,

to be defended publicly on Wednesday September 12, 2018 at 14:00.

Student number:	4272536	
Project duration:	January 10, 2018 – September 12, 2018	
Thesis committee:	Prof. dr. J.M.A.M. van Neerven,	TU Delft, Responsible professor
	Dr. D. Frey,	TU Delft, Supervisor
	Dr. ir. R. Heusdens,	TU Delft

October 22, 2018

TU Delft Delft University of Technology

Challenge the future

Abstract

Wavelets are a recent development in signal processing. These kind of functions are both well-localized in time and in frequency, and so using these to transform the signal gives insight where certain frequencies are needed. The classical way of constructing wavelets, as described by Daubechies and Meyer [5, 12] is only well-suited for $\mathbb{R}^n$, so new methods are developed for a broader range of spaces. In this paper, we describe the algorithm developed by Coifman and Maggioni [2], and the algorithm developed by Coulhon et al. [3]. Lastly, we modify the last algorithm using the finite speed of propagation property, and so we obtain a new way of developing wavelets.

Preface

This paper is written as part of my Master Thesis WI5005. It has been written between Januari 2018 and September 2018, to fulfill the graduation requirements of the master Applied Mathematics at the Delft University of Technology.

The project started with the question whether it was possible to generate wavelets from noncompact semigroups, given the algorithm by Coifman and Maggioni. After working out their algorithm, it turned out that a crucial step would fail. My supervisor Dorothee Frey then came up with a sampling method and gave me the paper by Coulhon et al. The remainder of the time was spent working out their algorithm and modifying it using the finite speed of propagation property.

I would like to thank Dorothee for the support she gave me. I also would like to thank Cindy Bosman, whose coffee pas I could borrow in order to keep be concentrated.

Timothy de Moor

Hoofddorp, September 5th, 2018

Contents

1 Introduction **5**
1.1 Applications of multiresolution analyses and wavelets 8
1.2 Organization of the thesis 13

2 Preliminaries **14**
2.1 Functional calculus and the Bochner integral. 14
2.2 Interpolation estimates for holomorphic functions 15
2.3 The Riesz-Thorin Interpolation Theorem 17
2.4 Finite speed of propagation 21

3 Generalising operator properties from one L^p space to the other **29**

4 Symmetric Diffusion Semigroups **34**
4.1 Basic definitions 34
4.2 Multiresolution analysis 35

5 Wavelets **37**

6 Construction of an orthogonal basis **38**

7 Construction using partitions of unity **46**
7.1 Assumptions 46
7.2 Integral operators and kernels 47
7.3 The sampling theorem 59
7.4 Construction of the wavelets 67

8 Wavelets and the finite speed of propagation **82**
8.1 Partition of unity without compact support 82
8.2 Gradient estimates 85

9 Conclusion **92**
9.1 Future work 92

1 Introduction

In the age of big data, new solutions have to be found to analyze this data. To eliminate the noise out of this data, several algorithms have been constructed. One of the ways is using a lowpass filter: taking the Fourier transform and throwing away the higher frequencies. A newer approach uses wavelets, which have the advantage of localizing features on different scales, so it is easy to keep the important information, while throwing away the noise at the other scales and locations. It turns out that using some smart techniques, the construction of wavelets in these cases is of the same order of the Fast Fourier Transform.

Moreover, if one wants to process signals or images, usually the Fourier transform is used to go to the frequency scale, and then the signal is filtered. This has a few drawbacks, as the Fourier transform loses information about the position, and at places where the signal changes heavily, it can be hard to properly reconstruct the signal back out of the Fourier transform. Wavelets can solve these problems: they have fast decay on the space side, and compact support on the frequency side (or the other way around), therefore they are well-localized on both sides of the Fourier transform. By convoluting the signal with these functions, shifted and scaled, only the places where the signal changes heavily will have a large contribution. Therefore, both the position of these points in the signal, and the behaviour can be reconstructed much better.

In this paper, we focus us on differential equations, and the aid of wavelets therein. If the differential equation has nonhomogeneous terms, the Fourier transform may have issues with the stability of these terms. We work from the semigroups generated by those differential operators, to build wavelets. Those wavelets can then be helpful for solving the equation. We consider a few algorithms for making wavelets, based on different assumptions on this semigroup we base the construction on. We give a short explanation of the different approaches.

For compact semigroups, Coifman and Maggioni [2] use that the eigenvalues only have an accumulation point at 0. So we can consider the eigenfunctions for which the eigenvalues are bound away from zero. Then their span is finite-dimensional, and we can use a modified version of the classical approach of Meyer [12]. As the eigenfunctions are in general not localized, but the scaling functions and wavelets are, they invented a trick: They approximate the spaces constructed by the eigenfunctions by spaces which are spanned by localized functions. After finding a way of orthogonalizing these functions in such a way that both the localization and the approximation are preserved, they end up with the following theorem:

Theorem 1.1. *Fix $\varepsilon > 0$. Let (X, d) be a (quasi-)metric space, and let μ be a doubling Borel measure, such that $\mu(X) < \infty$. Let $(T_t)_{t\geq 0}$ be a compact selfadjoint contraction semigroup on $L^2(x, \mu)$ generated by a nonnegative selfadjoint operator, such that $d(\operatorname{supp}(T_t f)) < \delta + a$ for every f such that $d(\operatorname{supp}(f)) < a$ for some $a > 0$, where $d(A)$ is the diameter of A. Here, δ only depends on the space X. Suppose that Φ_0 is a set of functions such that $d(\operatorname{supp}\varphi) < a\delta$ for* *each $\varphi \in \Phi_0$ ($a < c$, where c is a parameter depending on the space) and such that $\|P_{\Phi_0}\xi - \xi\| < \varepsilon$ for each eigenfunction ξ of T_1 such that the corresponding eigenvalue is greater than ε. (P_{Φ_0} is the orthogonal projection on the space Φ_0.)*

Then there exists a sequence of subspaces W_j, such that $L^2(X,\mu) = \bigoplus_{j=-1}^{\infty} W_j$, and each W_j is spanned by an orthonormal basis of wavelet functions $\{\psi_{j,i}\}_{i\in\mathcal{I}(j)}$ such that $d(\operatorname{supp}\psi_{j,i}) \leq c\delta^j$.

They consider the function spaces $\Phi_j := T^{2^j}\Phi_{j-1}$. These spaces turn out to approximate the spaces V_j spanned by the eigenfunctions ξ_λ, where the eigenvalue $\lambda > \varepsilon^{1-2^{j+1}}$ quite well. By orthogonalizing the spaces using a multiscale approach, a base of these function spaces is found. These are projected back on the orthogonal complements of the spaces V_{j+1} in V_j, and then orthogonalized again. This way, an orthogonal wavelet basis is constructed

So although it is not very apparent from the theorem, the construction is dependent on T. This means that if there are two semigroups satisfying the assumptions for the same space, in general different wavelets will be constructed. The construction will in general fail for noncompact semigroups, as the orthogonalization step, which uses a modified Gram-Schmidt procedure, requires a finite basis.

For these noncompact semigroups, a different approach has to be found. For instance, the approach of Coulhon, Kerkyacharian and Petrushev [3], is an alternative. To obtain localized wavelets, they use a partition of unity. Then they apply the square root of the nonnegative selfadjoint operator L to it, using the selfadjoint functional calculus. Using some technical kernel estimates, they manage to prove that each of the operators $\Psi_j(\sqrt{L})$, where Ψ_j are the functions in the partition of unity, is an integral operator. The kernels of these operators turn out to be wavelets. Eventually, the following properties are proven:

Theorem 1.2. *Let (X,d,μ) be a metric space with a doubling measure, and let L be a nonnegative selfadjoint operator. It generates a semigroup $P_t := e^{-tL}$. Assume that each P_t is an integral operator with kernel $p_t(x,y)$ that satisfies the following properties:*

- *For $t > 0$ and $x, y \in X$*

$$p_t(x,y) \leq \frac{Ce^{-cd^2(x,y)/t}}{\sqrt{\mu(B(x,\sqrt{t}))\mu(B(y,\sqrt{t}))}}.$$

- *There exists an $\alpha > 0$ such that for all $t > 0$ and $x, y, y' \in X$ with $d(y,y') < \sqrt{t}$*

$$|p_t(x,y) - p_t(x,y')| \leq C\left(\frac{d(y,y')}{\sqrt{t}}\right)^{\alpha} \frac{e^{-td^2(x,y)/t}}{\sqrt{\mu(B(x,\sqrt{t}))\mu(B(y,\sqrt{t}))}}.$$

- *For $t > 0$ and $x \in X$*

$$\int_X p_t(x,y)\,\mathrm{d}\mu(y) = 1$$

Let $(\Psi_j)_{j\geq 0}$ be a partition of unity on $\mathbb{R}_+$, such that $\operatorname{supp}\Psi_0 \subseteq [0,b]$ *and* $\operatorname{supp}\Psi_j \subseteq [b^{j-1}, b^{j+1}]$ *for some $b > 1$. Then there exists a system of wavelets $\{\psi_{j\xi}\}$ such that $\xi \in X$ and*

(a) each $\psi_{j\xi}$ has exponential decay, and is Hölder continuous of order α.

(b) $\|\psi_{j\xi}\|_p \simeq \mu(B(\xi, b^{-j}))^{1/p-1/2}$, where the constants in the equivalence depend only on p.

(c) $\phi(\sqrt{L})\psi_{0\xi} = \psi_{0\xi}$ for $\phi \in C_c^\infty([0,\infty))$, $\phi \equiv 1$ on $[0,b]$. Moreover, $\varphi(b^{-j}\sqrt{L})\psi_{j\xi} = \psi_{j\xi}$ for $\varphi \in C_c^\infty([0,\infty))$, $\varphi \equiv 1$ on $[b^{-j}, b^j]$.

(d) For each $f \in L^2(X,\mu)$, we have that

$$\frac{1}{4}\|f\|_2^2 \leq \sum_{j\geq 0}\sum_{\xi} |\langle f, \psi_{j\xi}\rangle| \leq 2\|f\|_2^2$$

The third approach we have in the paper is new. It is a modification of the one from Coulhon, Kerkyacharian and Petrushev [3]. We change the assumptions for the semigroup P_t, and obtain wavelets using the same construction as above, but with different properties. For example, we use the Finite Speed of Propagation [4, (3.5)], and obtain that the wavelets have compact support instead. The theorem above changes into the following theorem:

Theorem 1.3. *Let (X,d) be either $\mathbb{R}^d$ with the Euclidean metric, or a Riemannian manifold, and let μ a doubling Borel measure on X. Moreover, assume that the p-Poincaré inequality 8.1 holds for the (Riemannian) gradient. Then assume that L is a nonnegative selfadjoint operator. It generates a semigroup $P_t := e^{-tL}$. Assume that each P_t is an integral operator and its kernel p_t satisfies the bound*

$$p_t(x,y) \leq \frac{Ce^{-cd^2(x,y)/t}}{\sqrt{\mu(B(x,\sqrt{t}))\mu(B(y,\sqrt{t}))}}.$$

and assume that

$$\sup_{t\geq 0}\|t^{1/2}\nabla P_t\|_{p\to p} < \infty$$

Then there exists a $b > 1$, and a countable system $\{\psi_{j\xi}\}_{j\geq 0, \xi\in X}$ of wavelets which obey the following conditions:

a) Every ψ_j has compact support. Moreover, $\psi_{j\xi}$ can be written as the kernel of an operator Ψ_j: $\psi_{j\xi} = \Psi_j(\sqrt{L})(\cdot,\xi)$. This operator satisfies the condition $\|b^{-j}\nabla\Psi_j\|_{p\to p} < \infty$, uniformly in j.

b)

$$\|\psi_{j\xi}\|_p \simeq \mu(B(\xi, b^{-j}))^{1/p-1/2}$$

c) *For every* $\phi, \varphi \in C_c^\infty(\mathbb{R})$*, such that* $\phi \equiv 1$ *on* $[0,1]$ *and* $\varphi \equiv 1$ *on* $[b^{-1}, b]$*, we can find an* $\varepsilon > 0$ *such that*

$$\|\phi(b^{-1}\sqrt{L})\psi_{0\xi} - \psi_{0\xi}\|_p \leq \varepsilon$$

and

$$\|\varphi(b^{-j}\sqrt{L})\psi_{j\xi} - \psi_{j\xi}\|_p \leq \varepsilon$$

d) *For every* $f \in L^2(X, \mu)$*, we have the equivalence*

$$\frac{1}{4}\|f\|_2^2 \leq \sum_{j\geq 0}\sum_{\xi} |\langle f, \psi_{j\xi}\rangle|^2 \leq 2\|f\|_2^2$$

The theorem can be proven in a more general setting, but for the sake of the argument, the space X is chosen such that the term "gradient" is well-defined. It is beyond the scope of this thesis to explain what we mean by a "gradient" in more general spaces.

Property (c) differs from the corresponding property of the second approach. This follows because we want to use the finite speed of propagation. For this, we need to have a partition of unity consisting of functions with compact Fourier transform, which at best can have exponential decay. This directly transmits over to property (c).

1.1 Applications of multiresolution analyses and wavelets

We have not actually defined what a wavelets is. A wavelets is a band-limited function with exponential decay. This way, it is localized in both space domain and frequency domain. A few examples of wavelets can be found in figure 1. These wavelets are scaled down by powers of two, and translated, such that they form some sort of basis for L^2. The projection of an L^2-function onto the various wavelets yield coefficients. This is called the "wavelet transform", and the sequence of coefficients obtained this way is the transformed function.

As an example why wavelets are useful, consider the function $f(x) = e^{-\frac{x^2}{2}} + \frac{1}{2\pi}(||x-1|-1| - |x-1|+1)$. To compute the Fourier and wavelet transforms, it is sampled on the interval $[-4, 4]$ at 512 points, and put into the computer. The discrete Fourier transform is shown in figure 6, and the wavelet transform, using Daubechies "db20" wavelets [5], is shown in figure 8. In the zoomed pictures, the effect of the absolute value disturbance is shown clearly: The Fourier transform has some wiggles at the high frequencies, as indifferentiabilities have high frequency contributions. In the case of the wavelet transform, we see three wiggles repeat at a logarithmic scale. These are the wavelets in the neighborhood of $x = 0, 1, 2$, and as the rest of the function is smooth, the other places do not give any significant contribution.

To see the difference these localized contributions make, we cut off the top half of the frequencies/wavelets. As the amplitude of the cut off frequencies are quite small, the reversed Fourier/wavelet transform gives a signal which

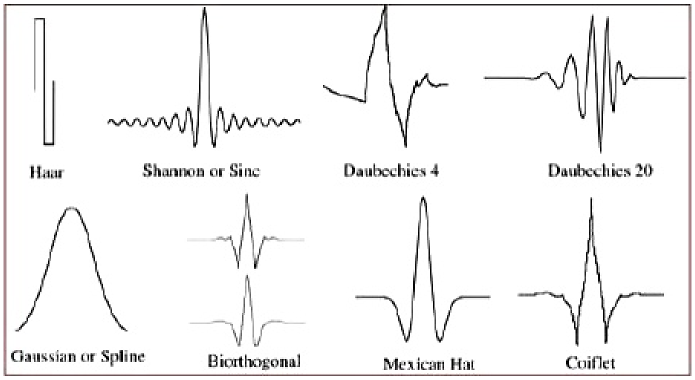

Figure 1: A few examples of wavelets. Image retrieved from http://www.continuummechanics.org/wavelets.html

resembles f quite well, as can be seen in figure 2. The biggest differences are at the points $x = 0, 1, 2$, as can be seen in the zoomed in pictures. At these places, the reverse Fourier transform is a smoothened out version of the original signal, while the reverse wavelet transform models the indifferentiability better. This is, because only the top three peaks in figure 8 are cut off, while there are a lot of wiggles in figure 6 thrown away.

Multiresolution analyses and wavelets can solve more problems than just the drawbacks noted above: The following examples are due to Coifman and Maggioni, as application of their approach[2, Chapter 8]. One of the problems multiresolution analyses and wavelets can solve is a non-homogeneous heat equation on the torus $\mathbb{T}$:

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(c(x)\frac{\partial}{\partial x}u\right).$$

where $0 < c < 1$ is a non-uniform function, for instance the function $c(x) = \sin^2(x)$. By sampling the torus in various equidistant points, and considering the space spanned by the Kronecker delta functions in each point, one can construct a multiresolution analysis by applying a discretized version of the right hand operator "T" of above equation to this space of functions, and then for level j apply T^{2^j-1} to the set of kronecker delta functions, thus obtaining the various resolutions. By looking at the functions in the level $j = 4$, it turns out that the dimension of the space is reduced to two. The "scaling" functions in this set are smooth on the place where there is high conductivity (c is close to 1), while they oscillate at the point where there is low-conductivity. By looking at the points above a specified precision, one can see that the points with very high conductivity are very tightly clustered together, the points with low conductivity are much further apart.

A different example of the construction of wavelets, is in the case of a noisy image of a white disk on a black background. It will follow that the scaling functions and the wavelets are actually images. The approach that is taken in

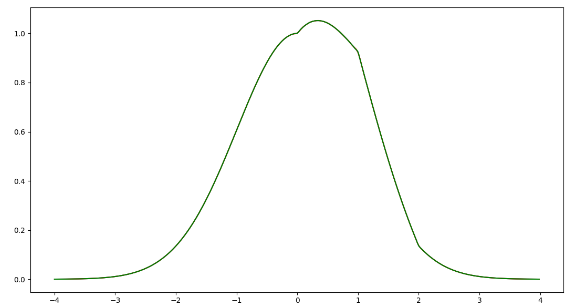

Figure 2: $f(x) = e^{-\frac{x^2}{2}} + \frac{1}{2\pi}(||x-1|-1| - |x-1| + 1)$

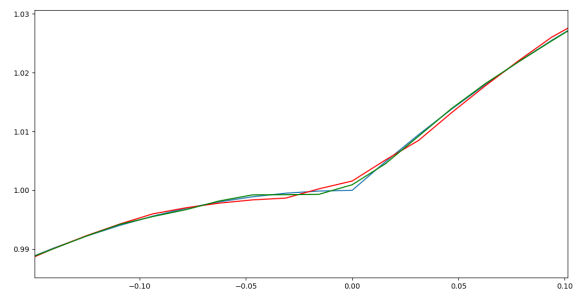

Figure 3: Above picture zoomed in at $x = 0$. Blue is the original signal, red is the signal after cutting off the highest frequencies in the Fourier transferm, green is the signal after removing the contributions of the finest layer of wavelets

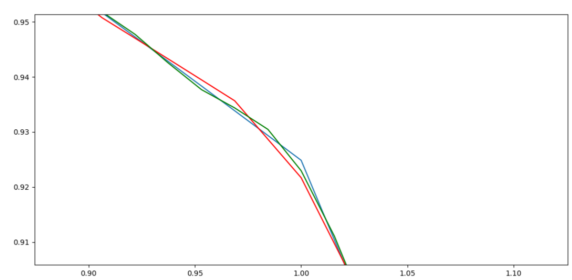

Figure 4: Above picture zoomed in at $x = 1$.

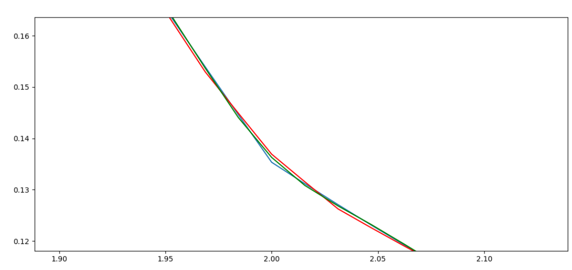

Figure 5: Above picture zoomed in at $x = 2$.

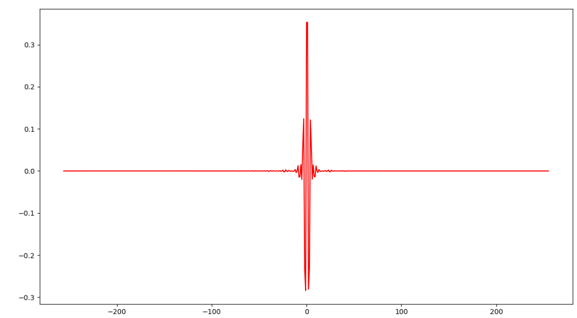

Figure 6: Fourier transform of f.

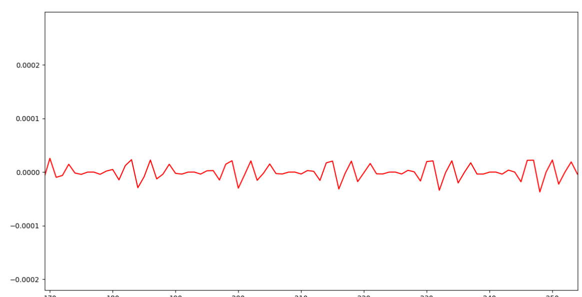

Figure 7: The Fourier transform, zoomed in at the high frequencies. The wiggles are necessary for the behavior of the above signal at $x \in \{0, 1, 2\}$.

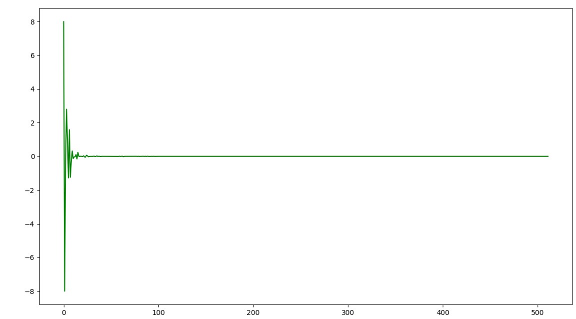

Figure 8: Wavelet transform of f. The wavelets get scaled every 2^k points, and then translated by integer amounts.

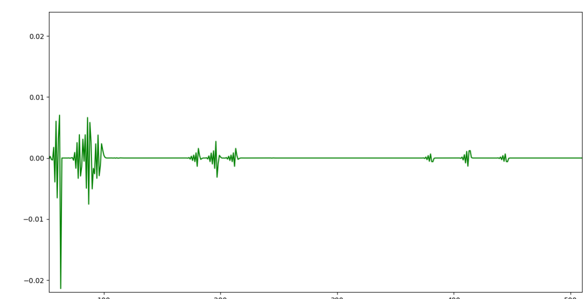

Figure 9: Wavelet transform, zoomed in at the higher "frequency" contributions. The repeating three peaks are the contributions for the behaviour at $x \in \{0, 1, 2\}$.

this example is to get 5×5 squares, centered around a given (x, y) coordinate, and seen as a point in $\mathbb{R}^{25}$. The idea of these patches is that it is clear that it is centered on a point in the disk, on the boundary, or outside the disk, while still having a small enough patch such that the computer can still handle it. Continuing with the example, the patches around points which lie close together are considered, and among them the patches with the lowest distance in $\mathbb{R}^{25}$ (the patches look similar). By making a new metric which takes these "edges" into account, we can make sure that the wavelet denoising algorithm will preserve the boundary of the disk. The pictures of the scaling functions made this way, show the ring around the disk clearly [2, Fig. 17]. This is a nonlinear way of denoising the image.

1.2 Organization of the thesis

In the next chapter, some of the used theory of functional analysis and integral theory is mentioned, together with some basic results in complex and harmonic analysis. In chapter 3, the theory in chapter 2 is applied to get some interpolation and extrapolation results on boundedness and compactness of operators. In chapter 4, we start with the first definitions and properties of the type of space we work with. The definition of wavelets, and a short introduction on the classical methods can be found in chapter 5. Then in chapter 6, the algorithm of Coifman and Maggioni is covered, which is used to make a multiresolution analysis and therefore wavelets out of compact semigroups. For the noncompact semigroups, chapter 7 gives the approach of Coulhon, Kerkyacharian and Petrushev to make a wavelet frame using a partition of unity, and a sampling theorem. Finally, we relax the assumptions of chapter 7 in chapter 8. We modify Coulhon's approach for semigroups which satisfy the Davies-Gaffney estimation, which is a weaker condition than the one their original work assumes.

2 Preliminaries

2.1 Functional calculus and the Bochner integral.

In this section, we describe some fundamental theory in functional analysis we will use throughout the thesis.

We will use the Fourier transform regularly, where we will use the following definition:

$$\hat{f}(\xi) := \mathcal{F}f(\xi) := \int_{\mathbb{R}} f(x)e^{-i\xi x}\,\mathrm{d}x, \qquad \xi \in \mathbb{R}$$

If we use a different domain of integration in a particular case, we will comment on that. In some cases, the Fourier inversion formula also holds, which is

$$f(x) = \mathcal{F}^{-1}\hat{f}(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}(\xi)e^{i\xi x}\,\mathrm{d}\xi, \qquad x \in \mathbb{R}.$$

Moreover, we will need a functional calculus for selfadjoint operators. We will do so by spectral theory [8, Chapter 3.3 and Theorem 4.3] and [7]: For a selfadjoint operator L on a Hilbert space H there exists a spectral resolution $(E_\lambda)_\lambda \in \mathbb{R}$, i.e. each E_λ is a projection, $E_{-\infty} = 0$, $E_\infty = I$, and $E_\lambda E_\mu = E_\mu$ for $\mu \leq \lambda$. From this spectral resolution, we can define a projection-valued measure such that $E_{[a,b]} := E_{b^+} - E_{a^-}, E_{(a,b]} := E_{b^+} - E_{a^+}$, etc. For a $x \in H$, we can define the Lebesgue-Stieltjes measure $\mu_{\|E_\lambda x\|^2}(A) := \|E_A x\|^2$. Note that this measure is finite, as $\mu_{\|E_\lambda x\|^2}(\mathbb{R}) = \|E_\infty x - E_{-\infty}x\|^2 = \|x - 0\|^2 < \infty$. Now, for $f \in L^2(\mathbb{R}, \mu_{\|E_\lambda x\|^2})$, we have that

$$\|f(L)x\|^2 = \int_{\mathbb{R}} |f(\lambda)|^2\,\mathrm{d}\mu_{\|E_\lambda x\|^2} < \infty.$$

We denote $\mathcal{D}(f(L)) := \{x \in H : f \in L^2(\mathbb{R}, \mu_{\|E_\lambda x\|^2})\}$. Then for $x \in \mathcal{D}(f(L))$, we have

$$f(L)x = \int_{\mathbb{R}} f(\lambda)\,\mathrm{d}E_\lambda x$$

Here, E_λ refers to the projection-valued measure, not to the spectral resolution. We will also write

$$f(L) = \int_{\mathbb{R}} f(\lambda)\,\mathrm{d}E_\lambda.$$

As an example, for bounded measurable functions f, we have that $\mathcal{D}(f(L)) = H$, as $\|f(L)x\|^2 \leq \|f\|_\infty^2\|x\|^2$ for each $x \in H$. If f is continuous, this definition is compatible with the definition of $f(L)$ in the continuous calculus.

Note that by above definition of the functional calculus, the spectral resolution is constructed is such a way that

$$L = \int_{\mathbb{R}} \lambda\,\mathrm{d}E_\lambda.$$

Furthermore, we need the Bochner integral, i.e. a Banach space valued integral [11]: Let X be a Banach space, and (Y, ν) be a measure space. Then for step functions $\varphi := \sum_{n=1}^{N} x_n \mathbb{1}_{A_n}$, with A_n ν-measurable sets, and $x_n \in X$ for $n = 1, 2, \ldots, N$, we define $\int_Y \varphi \, \mathrm{d}\nu = \sum_{n=1}^{N} x_n \nu(A_n)$. Then for strongly measurable functions $f : Y \to X$ (i.e. there exist step functions f_n converging almost everywhere to f, such that $\|f_n\| \leq \|f\|$ a.e.), which satisfy $\int_Y \|f\| \, \mathrm{d}\nu < \infty$, we have that

$$\int_Y \|f - f_n\| \, \mathrm{d}\nu \to 0$$

by the Dominated convergence theorem. From this, it also follows that $\int_Y f_n \, \mathrm{d}\nu$ is a Cauchy sequence, and so by the completeness of X we have that

$$\int_Y f \, \mathrm{d}\nu := \lim_{n \to \infty} \int_Y f_n \, \mathrm{d}\nu \in X.$$

This integral has some nice properties: For a functional $x^* \in X^*$ we have the Pettis integral

$$\left\langle \int_Y f \, \mathrm{d}\nu, x^* \right\rangle = \int_Y \langle f, x^* \rangle \, \mathrm{d}\nu.$$

Moreover, for a Bochner integrable function f, and a closed linear operator A on X such that $f(Y) \subseteq \mathcal{D}(A)$ and such that Af is Bochner integrable, we have that

$$A \int_Y f(y) \, \mathrm{d}\nu(y) = \int_Y A(f(y)) \, \mathrm{d}\nu(y).$$

2.2 Interpolation estimates for holomorphic functions

In this subsection, we state some results about estimates of holomorphic functions in terms of bounds of two enclosing (half-)lines. The first one is the Phragmén-Lindelöf Theorem on sectors, and after that we prove a couple of corallaries, one of which is the Hadamard Three Lines Lemma, which is needed for the Riesz-Thorin Interpolation Theorem down below.

For the proof of the Phragmén-Lindelöf Theorem, we refer to [13, p. 108].

Theorem 2.1. *Let $S = \{z \in \mathbb{C} : 0 < (\theta_0 + \arg(z - z_0)) \mod 2\pi < \pi/\alpha\}$ be the open sector bounded by two rays meeting at an angle π/α in a point $z_0 \in \mathbb{C}$ for some $\alpha > \frac{1}{2}$ and some $\theta_0 \in (-\pi, \pi]$. Suppose that F is analytic on S, continuous on $\overline{S}$ and satisfies $|F(z)| \leq Ce^{c|z|^\beta}$ for some $\beta \in [0, \alpha)$ and for all $z \in S$. Then the condition $|F(z)| \leq B$ on the two bounding rays implies that $|F(z)| \leq B$ for all $z \in S$.*

Using this theorem, we now prove the Hadamard Three Lines Lemma.

Corollary 2.2 (Hadamard Three Lines Lemma)**.** *Suppose $\Phi : \mathbb{C} \to \mathbb{C}$ is holomorphic on the inside of the strip $0 \leq \operatorname{Re} z \leq 1$ and continuous and bounded on the closure of the strip. Furthermore, suppose $|\Phi(z)| \leq M_0$ on the boundary $\operatorname{Re} z = 0$ and $|\Phi(z)| \leq M_1$ on the boundary $\operatorname{Re} z = 1$. Then, for all $y \in \mathbb{R}$, $x \in (0, 1)$,*

$$|\Phi(x + iy)| \leq M_0^{1-x} M_1^x.$$

Proof. We define a function $\phi(z) := \Phi(z)M_0^{z-1}M_1^{-z}$. This function is holomorphic on the open strip $0 < \operatorname{Re} z < 1$, and continuous on the closure. Moreover, as $|M_0^{iy}| = |M_1^{iy}| = 1$ for all $y \in \mathbb{R}$, it follows that $\phi(z)$ is bounded on the closure on the strip, with $|\phi(iy)| = |\phi(1+iy)| = 1$ for all $y \in \mathbb{R}$.

Now consider the bijective mapping $\xi : z \mapsto -i \log z$ on $0 \leq \arg z \leq 1$. This maps the halfline $(0, \infty)$ to the line $\operatorname{Re} z = 0$, the halfline $\{re^i : r \in (0, \infty)\}$ to the line $\operatorname{Re} z = 1$, and the open sector bounded by those halflines to the open strip $0 < \operatorname{Re} z < 1$. Now it follows that $\phi \circ \xi$ satisfies the assumptions of Theorem 2.1, with $\alpha = \pi$, $\beta = 0$, and $B = 1$. Hence $|(\phi \circ \xi)(z)| \leq 1$ for all $z \in \mathbb{C}\backslash\{0\}$ with $0 < \arg z < 1$. But ξ maps $\{z : 0 < \arg z < 1\}$ into the open strip surjectively. Hence $|\phi(z)| < 1$ for $0 < \operatorname{Re} z < 1$. And so $|\Phi(x+iy)| \leq |M_0^{1-x-iy}M_1^{x+iy}| = M_0^{1-x}M_1^x$ for $y \in \mathbb{R}$, $x \in (0,1)$. $\square$

The other corollary which we are going to state will be used to prove the equivalence between the Davies-Gaffney estimate and the Finite Speed Propagation property. The corollary gives an exponential bound for right half of the complex plane given that the function is bounded on this half, and the function satisfies a similar bound on the nonnegative real numbers:

Corollary 2.3 ([4, Prop 2.2])**.** *Suppose F is an analytic function on $\mathbb{C}_+ := \{x + iy : x > 0, y \in \mathbb{R}\}$. Assume that, for given number $A, B, \gamma > 0$ and $\alpha \geq 0$ we have $|F(\cdot)| \leq B$ on $\mathbb{C}_+$ and*

$$|F(t)| \leq Ae^{at}e^{-\gamma/t} \qquad \forall\, t \in \mathbb{R}_+.$$

Then we have

$$|F(z)| \leq Be^{-\operatorname{Re}\frac{\gamma}{z}} \qquad \forall\, z \in \mathbb{C}_+.$$

Proof. We consider the function u on $\mathbb{C}_+$:

$$u(\zeta) := F(\gamma/\zeta).$$

Fix $\varepsilon > 0$. By the first condition, we have $|u(\zeta)| \leq B$, hence $|u(\zeta)e^\zeta| \leq Be^{\operatorname{Re}\zeta} \leq Be^{|\zeta|}$. Moreover, by evaluating $\zeta \mapsto u(\zeta)e^\zeta$ on the line $\operatorname{Re} z = \varepsilon$, it follows from the hypothesis that

$$\sup_{\operatorname{Re}\zeta = \varepsilon} |u(\zeta)e^\zeta| \leq Be^\varepsilon$$

Moreover, from the second condition, we have for $\zeta \in [\varepsilon, \infty)$

$$\begin{aligned} |u(\zeta)| &= |F(\gamma/\zeta)| \\ &\leq Ae^{a\gamma/\zeta}e^{-\gamma/(\gamma/\zeta)} \\ &= Ae^{a\gamma/\zeta}e^{-\zeta}. \end{aligned}$$

So

$$|u(\zeta)e^\zeta| \leq Ae^{a\gamma/\zeta}$$

and

$$\begin{aligned}\sup_{\zeta\geq\varepsilon}|u(\zeta)e^{\zeta}| &\leq \sup_{\zeta\geq\varepsilon} Ae^{a\gamma/\zeta}\\ &= Ae^{a\gamma/\varepsilon}\end{aligned}$$

Now we can apply the Phragmén-Lindelöf Theorem 2.1 twice with the regions bounded by $[\varepsilon,\infty)$ and $\operatorname{Re} z=\varepsilon$, where we choose $\alpha=2$, $z_0=\varepsilon$, $\beta=1$ and $\theta_0\in\{0,\pi/2\}$ and we apply it to the function $\zeta\mapsto u(\zeta)e^{\zeta}$. This function is bounded on the rays by $\max\{Ae^{a\gamma/\varepsilon},Be^{\varepsilon}\}$. By the theorem, we have

$$\sup_{\operatorname{Re}\zeta\geq\varepsilon}|u(\zeta)e^{\zeta}|\leq\max\{Ae^{\alpha\gamma/\varepsilon},Be^{\varepsilon}\}$$

Now we can use Theorem 2.1 again, now with $\alpha=1$, $\beta=0$ and bound Be^{ε} on the halfplane $\{\operatorname{Re} z\geq\varepsilon\}$. We get

$$\sup_{\operatorname{Re}\zeta\geq\varepsilon}|u(\zeta)e^{\zeta}|\leq Be^{\varepsilon}.$$

The reason why we had to do this in two steps was that we otherwise had $\alpha=\beta=1$, while according to the hypothesis of the Phragmén-Lindelöf Theorem we must have $\beta<\alpha$. Now we let $\varepsilon\to 0$, and we obtain

$$\sup_{\operatorname{Re}\zeta>0}|u(\zeta)e^{\zeta}|\leq B.$$

Lastly we substitute $\zeta=\gamma/z$, and we have

$$|F(z)|=|F(\gamma/\zeta)|=|u(\zeta)|\leq B|e^{-\gamma/z}|=Be^{-\operatorname{Re}\frac{\gamma}{z}}$$ □

2.3 The Riesz-Thorin Interpolation Theorem

Theorem 2.4 (The Riesz-Thorin Interpolation Theorem)**.** *Consider a linear functional T, which maps the step functions in the measure space (X,μ) to the step functions in the measure space (Y,ν). Suppose $p_0,q_0,p_1,q_1\in[1,\infty]$ and*

$$\frac{1}{p}=\frac{1-t}{p_0}+\frac{t}{p_1}\qquad\frac{1}{q}=\frac{1-t}{q_0}+\frac{t}{q_1}$$

for $t\in(0,1)$. If $q_0=q_1=\infty$, we suppose further that ν is semifinite. Suppose T can be extended to a bounded operator $T:L^{p_0}(X,\mu)\to L^{q_0}(X,\mu)$ with norm $M_0>0$ and that it can be extended to a bounded operator $T:L^{p_1}(\mu)\to L^{q_1}(\nu)$ wit hrom $M_1>0$. Then T can be extended to a bounded operator $T:L^p(X,\mu)\to L^q(X,\mu)$. Furthermore, $\|Tf\|_q\leq M_0^{1-t}M_1^t\|f\|_p$ for all $f\in L^p$.

Proof. First assume that $p_0=p_1$. Then we can use Hölder's inequality with

$\frac{1}{q} = \frac{1}{q_0/(1-t)} + \frac{1}{q/t}$ to obtain

$$\begin{aligned}
\|Tf\|_q &= \left(\int_X |Tf|^q \,\mathrm{d}\mu\right)^{1/q} \\
&= \left(\int_X |(Tf)^{1-t}(Tf)^t|^q \,\mathrm{d}\mu\right)^{1/q} \\
&\leq \left(\int_X |(Tf)^{1-t}|^{q_0/(1-t)} \,\mathrm{d}\mu\right)^{(1-t)/q_0} \left(\int_X |(Tf)^t|^{q_1/t} \,\mathrm{d}\mu\right)^{t/q_1} \\
&\leq \left(\int_X |(Tf)|^{q_0} \,\mathrm{d}\mu\right)^{(1-t)/q_0} \left(\int_X |(Tf)|^{q_1} \,\mathrm{d}\mu\right)^{t/q_1} \\
&= \|Tf\|_{q_0}^{(1-t)} \|Tf\|_{q_1}^t \\
&\leq M_0^{1-t} \|f\|_p^{1-t} M_1^t \|f\|_p^t \\
&= M_0^{1-t} M_1^{1-t} \|f\|_p
\end{aligned}$$

We now continue with the case $p_0 \neq p_1$, where we assume without loss of generality that $p_0 < p < p_1$. We fix simple functions f on X and g on Y. To prove that T is of strong type (p, q), we will define an holomorphic function, which is bounded by $M_0 C(f, g)$ on $\operatorname{Re} z = 0$, bounded by $M_1 C'(f, g)$ on $\operatorname{Re} z = 1$ and such that $|\Phi(t)| = |\langle Tf, g\rangle|$, the duality operation of L^q. Then we can use Hahn-Banach to find a bound of $\|Tf\|_q$ in terms of M_0, M_1 and f. Finally we extend this result to general $f \in L^p$.

Write $f := \sum_{j=1}^n a_j e^{i\theta_j} \mathbb{1}_{E_j}$, where $\theta_j \in \mathbb{R}, a_j \geq 0$ for $j = 1, \ldots, n$, and the E_j are pairwise disjoint. Similarily, we write $g := \sum_{k=1}^m b_k e^{i\phi_k} \mathbb{1}_{F_k}$, where $\phi_k \in \mathbb{R}, b_k \geq 0$ for $k = 1, \ldots, m$. We now define two complex functions α, β on the strip $\{0 \leq \operatorname{Re} z \leq 1\}$ by

$$\alpha(z) := \frac{1-z}{p_0} + \frac{z}{p_1} \qquad \beta(z) := \frac{1-z}{q_0} + \frac{z}{q_1}.$$

Note that $\alpha(t) = \frac{1}{p}$ and $\beta(t) = \frac{1}{q}$. Using these definitions we can define f_z and g_z by

$$\begin{aligned}
f_z &= \sum_{j=1}^n a_j^{\alpha(z)/\alpha(t)} e^{i\theta_j} \mathbb{1}_{E_j} \\
g_z &= \begin{cases} \sum_{k=1}^m b_k^{(1-\beta(z))/(1-\beta(t))} e^{i\theta_k} \mathbb{1}_{F_k} & \text{if } \beta(t) \neq 1 \\ g & \text{if } \beta(t) = 1 \end{cases}
\end{aligned}$$

As $p < \infty$, we have $\alpha(t) > 0$, hence $z \mapsto f_z : \{0 \leq \operatorname{Re} z \leq 1\} \to \operatorname{Simple}(X)$ is well-defined. Finally, we are able to define Φ:

$$\Phi(z) := \int_Y (Tf_z) g_z \,\mathrm{d}\nu$$

By writing out the definitions of f_z and g_z, and using linearity of T, we can rewrite Φ as

$$\Phi(z) = \begin{cases} \sum_{j,k} a_j^{\alpha(z)/\alpha(t)} b_k^{(1-\beta(z))/(1-\beta(t))} e^{i(\theta_j+\phi_k)} \int_Y (T\mathbb{1}_{E_j})\mathbb{1}_{F_k}\,\mathrm{d}\nu & \text{if } \beta(t) \neq 1 \\ \sum_{j,k} \alpha^{\alpha(z)/\alpha(t)} b_k e^{i(\theta_j+\phi_k)} \int_Y (T\mathbb{1}_{E_j})\mathbb{1}_{F_k}\,\mathrm{d}\nu & \text{if } \beta(t) = 1 \end{cases}$$

It is not hard to see that $\Phi(z)$ is holomorphic on the open strip $\{0 < \operatorname{Re} z < 1\}$: that follows easily from the fact that $z \mapsto e^z$ is holomorphic on $\mathbb{C}$. Now $\Phi(z)$ is bounded on the strip, as $|\Phi(z)| = |\Phi(\operatorname{Re} z)| \leq \sup_{x\in[0,1]} |\Phi(x)| < \infty$, where the equality is due to the fact that $|c^{iy}| = 1$ for real c and y. The last thing we have to check before we are able to use the Hadamard Three Lines Lemma 2.2 is the bound on the boundaries $\operatorname{Re} z = 0$ and $\operatorname{Re} z = 1$. We claim that these bounds are equal to respectively $M_0\|f\|_p^{p/p_0}\|g\|_{q'}^{q'/q'_0}$ and $M_1\|f\|_p^{p/p_1}\|g\|_{q'}^{q'/q'_1}$.

Indeed, let $z := iy$, $y \in \mathbb{R}$, fix $x \in X$, and let $j \in \{1, \dots, n\}$ such that $x \in E_j$. This j is unique, as all those E_js are mutually disjoint. We have

$$\begin{aligned} |f_{iy}(x)| &= |a_j^{\alpha(iy)/\alpha(t)} e^{i\theta_j} \mathbb{1}_{E_j}(x)| \\ &= |a_j^{\alpha(iy)/\alpha(t)}| \\ &= |e^{p((1-iy)p_0^{-1}+iyp_1^{-1})\log(a_j)}| \\ &= e^{p/p_0 \log(a_j)} = a_j^{p/p_0} \\ &= (a_j\mathbb{1}_{E_j}(x))^{p/p_0} \\ &= |f(x)|^{p/p_0}. \end{aligned}$$

Similarily, by fixing an $x \in Y$, and then finding k such that $x \in F_k$, we can do exactly the same with $g_{iy}(x)$ and find that $|g_{iy}(x)| \leq |g(x)|^{(1-q_0^{-1})/(1-q^{-1})} = |g(x)|^{q'/q'_0}$, where q' and q'_0 denote the Hölder conjugates of q and q_0 respectively. Finally, in the definition of $\Phi(iy)$, we can use Hölder, and obtain

$$\begin{aligned} |\Phi(iy)| &= \int_Y (Tf_{iy})g_{iy}\,\mathrm{d}\nu \\ &\leq \|Tf_{iy}\|_{q_0}\|g_{iy}\|_{q'_0} \\ &\leq \|T\|_{p_0\to q_0}\|f_{iy}\|_{p_0}\|g_{iy}\|_{q'_0} \\ &\leq M_0 \left(\int_X (|f(x)|^{p/p_0})^{p_0}\,\mathrm{d}\mu\right)^{1/p_0} \left(\int_Y (|g(x)|^{q'/q'_0})^{q'_0}\,\mathrm{d}\nu\right)^{1/q'_0} \\ &= M_0\|f\|_p^{p/p_0}\|g\|_{q'}^{q'/q'_0} \end{aligned}$$

For $\Phi(1+iy)$ we can do the same calculations, where we now have that

$$\begin{aligned} |a^{\alpha(1+iy)/\alpha(t)}| &= |a^{p(iyp_0^{-1}+(1+iy)p_1^{-1})}| \\ &= |a^{p/p_1}| \end{aligned}$$

and similarly $|b^{(1-\beta(1+iy))/(1-\beta(t))}| = |b^{q'/q'_1}|$. For the estimation of $|\Phi(1+iy)|$ we now use Hölder with the functions in L^{q_1} and $L^{q'_1}$, and using the same estimations we finally arrive at $|\Phi(1+iy)| \leq M_1\|f\|_p^{p/p_1}\|g\|_{q'}^{q'/q'_1}$.

We have now shown that Φ satisfies the assumptions of the Hadamard Three Lines Lemma 2.2. Using this lemma, we have that

$$\begin{aligned}|\Phi(x+iy)| &\leq (M_0\|f\|_p^{p/p_0}\|g\|_{q'}^{q'/q'_0})^{1-x}(M_1\|f\|_p^{p/p_1}\|g\|_{q'}^{q'/q'_0})^x \\ &= M_0^{1-x}M_1^x\|f\|_p^{p(\frac{1-x}{p_0}+\frac{x}{p_1})}\|g\|_{q'}^{q'((1-x)(1-\frac{1}{q_0})+x(1-\frac{1}{q_1}))} \\ &= M_0^{1-x}M_1^x\|f\|_p^{p(\frac{1-x}{p_0}+\frac{x}{p_1})}\|g\|_{q'}^{q'(1-\frac{1-x}{q_0}-\frac{x}{q_1})}\end{aligned}$$

In particular, we have

$$\Phi(t) \leq M_0^{1-t}M_1^t\|f\|_p^{p(\frac{1}{p})}\|g\|_{q'}^{q'(1-\frac{1}{q})} = M_0^{1-t}M_1^t\|f\|_p\|g\|_{q'}$$

Noting that $\Phi(t) = \int_Y (Tf_t)g_t \,\mathrm{d}\nu = \int_Y (Tf)g \,\mathrm{d}\nu$, and using Hahn-Banach, we have

$$\begin{aligned}\|Tf\|_q &= \sup_{\|g\|_{q'}\leq 1} |\langle Tf, g\rangle| \\ &= \sup_{\|g\|_{q'}\leq 1} \Phi(t) \\ &\leq \sup_{\|g\|_{q'}\leq 1} M_0^{1-t}M_1^t\|f\|_p\|g\|_{q'} \\ &= M_0^{1-t}M_1^t\|f\|_p\end{aligned}$$

So we have proved the theorem for simple functions f. To prove the theorem for the whole of $L^p(X)$, fix an $f \in L^p(X)$. Then there exists a sequence of measurable simple functions $(f_n)_{n=1}^\infty$ such that $f_n \to f$ pointwise almost everywhere, and $|f_n| \leq |f|$ for all $n \in \mathbb{N}$. We let $E := \{x : |f(x)| > 1\}$, $g := f\mathbb{1}_E, g_n := f_n\mathbb{1}_E, h := f\mathbb{1}_{E^c}, h_n := f_n\mathbb{1}_{E^c}$ such that $f = g + h$ and $f_n = g_n + h_n$ for all $n \in \mathbb{N}$. Furthermore, we have that $g_n \to g$ almost everywhere and $h_n \to h$ almost everywhere.

Note that $\mu(E) < \infty$, because

$$\begin{aligned}\infty &> \|f\|_p^p \\ &= \int_X |f|^p \,\mathrm{d}\mu \\ &\geq \int_E |f|^p \,\mathrm{d}\mu \\ &\geq \int_E 1^p \,\mathrm{d}\mu \\ &= \mu(E)\end{aligned}$$

using the definition of E in the fourth line. Therefore, by using Hölder, we have $L^p(E) \subseteq L^q(E)$ when $\infty \geq p \geq q \geq 1$. So we automatically get that $g \in L^{p_0}(X)$, using that its support is contained in E.

On the other hand, we have $\|h\|_\infty \leq 1$, so that $h \in L^{p_1}$ if $p_1 = \infty$. If $p_1 < \infty$, we have

$$\begin{aligned} \infty &> \|f\|_p \\ &\geq \|f\|_p \operatorname*{ess\,sup}_{x \in E^c} |f|^{p_1 - p} \\ &\geq \int_X |f|^p |f|^{p_1-p} \mathbb{1}_{E^c} \, \mathrm{d}\mu \\ &= \|h\|_{p_1}^{p_1} \end{aligned}$$

So $h \in L^{p_1}(X)$. Moreover, using that $|g_n| \leq |g|$ and $|h_n| \leq |h|$, which follow immediately from this estimate of f, we can use the Dominated Convergence Theorem three times to obtain

$$\begin{aligned} \|f_n - f\|_p &\to 0 \\ \|g_n - g\|_{p_0} &\to 0 \\ \|h_n - h\|_{p_1} &\to 0. \end{aligned}$$

Now we can use that T is bounded on $L^{p_0}(X)$ and $L^{p_1}(X)$, so we have that

$$\begin{aligned} \|Tg_n - Tg\|_{q_0} &\to 0 \\ \|Th_n - Th\|_{q_1} &\to 0. \end{aligned}$$

By measure theory, there exists a subsequence such that $Tg_{n_k} \to Tg$ almost everywhere as $k \to \infty$, and a subsequence of $(n_k)_{k=1}^\infty$ such that $Th_{n_{k_l}} \to Th$ almost everywhere as $l \to \infty$. On this subsequence, we have $Tf_{n_{k_l}} = Tg_{n_{k_l}} + Th_{n_{k_l}} \to Tg + Th = Tf$ almost everywhere. Using Fatou's lemma, we finally get

$$\|Tf\|_q \leq \liminf_{l \to \infty} \|Tf_{n_{k_l}}\|_q \leq \liminf_{l \to \infty} M_0^{1-t} M_1^t \|f_{n_{k_l}}\|_p \leq M_0^{1-t} M_1^t \|f\|_p$$

And so we have that T is of strong type (p, q), with $\|T\|_{p \to q} \leq M_0^{1-t} M_1^t$. □

2.4 Finite speed of propagation

In this subsection, we show that the Davies-Gaffney estimate and the finite speed of propagation property are equivalent for a nonnegative selfadjoint operator L are equivalent. Then we show that we can modify the Finite Speed of Propagation property to replace the cosine in its definition with a partition of unity. This will be useful when we look at noncompact Markovian semigroups to see if they generate wavelets. We start with the definitions of both aforementioned properties.

Definition 2.5 ([4, p. 513 and 517]). *Let (X, d, μ) be a metric measure space.*

1. *Suppose that L is a nonnegative self-adjoint operator on $L^2(X, \mu)$, such that $\{e^{-tL}\}_{t \geq 0}$ is a bounded semigroup of linear operators. We let $d(U_1, U_2) :=$*

$\inf_{x\in U_1,y\in U_2} d(x,y)$. *We say that* $\{e^{-zL}\}_{z\in\mathbb{C}_+}$ *satisfies the Davies-Gaffney estimate if*

$$|\left\langle e^{-tL}f_1, f_2\right\rangle| \leq Ce^{-\frac{r^2}{4t}}\|f_1\|_2\|f_2\|_2$$

for some $C>0$*, and all* $f_1, f_2 \in L^2(X,\mu)$*,* $U_i \subseteq X, (i=1,2)$ *disjoint sets such that* $r := d(U_1,U_2) > 0$*,* $\operatorname{supp} f_i \subseteq U_i$ *and for all* $t>0$*.*

2. *A nonnegative selfadjoint operator* L *satisfies the finite speed of propagation property with respect to a smooth function* ϕ *and speed* v *if*

$$\left\langle \phi(t\sqrt{L})f_1, f_2\right\rangle = 0$$

for $0 < t < r/v$*, where* $r := d(U_1,U_2)$*, with* $U_i \subseteq X$ *open sets, and* $f_i \in L^2(U_i,\mu)$*. Here* $\phi(t\sqrt{L})$ *is defined using the spectral theory discussed in section 2.1*

The reason for the name "finite speed of propagation", is that when you apply the cosine function $\cos(t\sqrt{L})$, you obtain the solution of the wave equation corresponding to the (second order) operator L. Then this property says that in a point of distance r from the initial wave, before the time r/v, you cannot have effects of the wave of speed v. This v is dependent on the operator, and in particular finite. Hence the name

Now we can state the theorem.

Theorem 2.6 ([4, Thm 3.4])**.** *Let* (X,d,μ) *be a metric measure space, and let* L *be a self-adjoint non-negative operator acting on* L^2*. Then the finite speed propagation property with respect to the cosine and the Davies-Gaffney estimate are equivalent.*

Proof. Firstly assume the Davies-Gaffney estimate. Fix two open sets $U_1, U_2 \subseteq X$. Let $f_i \in L^2(X,\mu)$ such that $\operatorname{supp} f_i \subseteq U_i$ $(i=1,2)$. We define a function $F:\mathbb{C}_+\to\mathbb{C}$ by

$$F(z) := \left\langle e^{-zL}f_1, f_2\right\rangle.$$

Since by assumption e^{-zL} is bounded on $L^2(X,\mu)$, it follows that F is a bounded analytic function on $\mathbb{C}_+$:

$$\begin{aligned}\sup_{z\in\mathbb{C}_+}|F(z)| &\leq \sup_{z\in\mathbb{C}_+}\int |e^{-z\lambda}|\,\mathrm{d}|\langle E_\lambda f_1, f_2\rangle| \\ &= \sup_{t>0}\sup_{y\in\mathbb{R}}\int |e^{-t\lambda}||e^{-iy\lambda}|\,\mathrm{d}|\langle E_\lambda f_1, f_2\rangle| \\ &= \sup_{t>0}\int |e^{-t\lambda}|\,\mathrm{d}|\langle E_\lambda f_1, f_2\rangle| \\ &\leq C\|f_1\|\|f_2\|.\end{aligned}$$

The fact that F is holomorphic follows from the holomorphicity of $e^{-z\lambda}$, and the dominated convergence theorem (with dominating function $\langle f_1, f_2\rangle$).

It satisfies the assumptions of Corollary 2.3, with

$$B := C\|f_1\|_2\|f_2\|_2$$

by Cauchy-Schwarz and

$$A := C\|f_1\|_2\|f_1\|_2 \qquad a := 0 \qquad \gamma := r^2/4 := d(U_1, U_2)^2/4$$

by the definition of the Davies-Gaffney estimate. So by Corollary 2.3 we have

$$|F(z)| \leq C\|f_1\|_2\|f_2\|_2 e^{-r^2 \operatorname{Re}\frac{1}{4z}}.$$

Next, write the Hadamard Transmutation Formula for $s > 0$:

$$\left\langle e^{-sL} f_1, f_2 \right\rangle = \int_0^\infty \left\langle \cos(t\sqrt{L}) f_1, f_2 \right\rangle \frac{e^{-t^2/(4s)}}{\sqrt{\pi s}} \,\mathrm{d}t.$$

This follows from the Fourier transform of the heat kernel:

$$\mathcal{F}(\omega \mapsto e^{-\omega^2})(t) = \sqrt{\pi} e^{-t^2/4}$$

So

$$\begin{aligned} \mathcal{F}(\omega \mapsto e^{-s\omega^2})(t) &= \mathcal{F}(e^{-(\cdot)^2}(s^{1/2}\omega))(t) \\ &= \frac{1}{\sqrt{s}} \mathcal{F}(\omega \mapsto e^{-\omega^2})(\frac{t}{s^{1/2}}) \\ &= \sqrt{\frac{\pi}{s}} e^{-\frac{t^2}{4s}}. \end{aligned}$$

Now taking the inverse, we have $e^{-s\omega^2} = \mathcal{F}^{-1}(\sqrt{\frac{\pi}{s}} e^{-\frac{t^2}{4s}})(\omega)$, or, written out as an integral, we have,

$$e^{-s\omega^2} = \frac{1}{2\pi}\sqrt{\frac{\pi}{s}} \int_\infty^\infty e^{-t^2/(4s)} e^{i\omega t} \,\mathrm{d}t.$$

Which implies, with $\lambda := \omega^2$,

$$\begin{aligned} e^{-s\lambda} &= \frac{1}{2\sqrt{\pi s}} \int_\infty^\infty e^{-t^2/(4s)} e^{it\sqrt{\lambda}} \,\mathrm{d}t \\ &= \frac{1}{\sqrt{\pi s}} \int_0^\infty \cos(t\sqrt{\lambda}) e^{-t^2/(4s)} \,\mathrm{d}t. \end{aligned} \qquad (*)$$

We view both sides of this equation as a function of λ. Now we note that by the spectral theorem which was explained above we have

$$e^{-sL} = \int e^{-s\lambda} \,\mathrm{d}E_\lambda$$

We write out e^{-sL} and use Fubini:

$$\begin{aligned}
\left\langle e^{-sL} f_1, f_2 \right\rangle &= \int e^{-s\lambda}\, \mathrm{d} \left\langle E_\lambda f_1, f_2 \right\rangle \\
&= \int \frac{1}{\sqrt{\pi s}} \int_0^\infty \cos(t\sqrt{\lambda}) e^{-t^2/(4s)}\, \mathrm{d}t\, \mathrm{d} \left\langle E_\lambda f_1, f_2 \right\rangle \\
&= \frac{1}{\sqrt{\pi s}} \int_0^\infty \int \cos(t\sqrt{\lambda}) e^{-t^2/(4s)}\, \mathrm{d} \left\langle E_\lambda f_1, f_2 \right\rangle \mathrm{d}t \\
&= \int_0^\infty \left\langle \cos(t\sqrt{L}) f_1, f_2 \right\rangle \frac{e^{-t^2/(4s)}}{\sqrt{\pi s}}\, \mathrm{d}t
\end{aligned}$$

Where Fubini is justified as the integrated function is continuous in both t and λ, hence measurable. Moreover, the Lebesgue-Stieltjes measure $\mu_{\langle E_\lambda f_1, f_2\rangle}$ is finite, and the function is uniformly bounded in λ: $|\cos(t\sqrt{\lambda})e^{-t^2/(4s)}| \leq e^{-t^2/(4s)}$. Morover, this function is integrable with respect to the Lebesgue measure in the variable t. So we have

$$(t,\lambda) \mapsto \cos(t\sqrt{\lambda})e^{-t^2/(4s)} \in L^1([0,\infty) \times [0,\infty), \mathcal{L} \otimes \mu_{\left\langle E_{(\cdot)} f_1, f_2 \right\rangle}),$$

where $\mathcal{L}$ is the Lebesgue measure. This proves the Hadamard Transmutation formula.

Moving on, we change in the formula $t \to \sqrt{t}$ and $s \to \frac{1}{4s}$, and we get, after applying the substitution rule,

$$s^{-1/2} \left\langle e^{-L/(4s)} f_1, f_2 \right\rangle = \int_0^\infty \frac{1}{\sqrt{\pi t}} \left\langle \cos(\sqrt{t}\sqrt{L} f_1, f_2 \right\rangle e^{-st}\, \mathrm{d}t$$

for $s > 0$. We now extend the expression above holomorphically to $\mathbb{C}_+$ and we get for $\zeta \in \mathbb{C}_+$

$$\zeta^{-1/2} \left\langle e^{-L/(4\zeta)} f_1, f_2 \right\rangle = \int_0^\infty \frac{1}{\sqrt{\pi t}} \left\langle \cos(\sqrt{t}\sqrt{L} f_1, f_2 \right\rangle e^{-\zeta t}\, \mathrm{d}t$$

Note that we have just proved that $u(\zeta) := \zeta^{-1/2} F(\frac{1}{4\zeta}) = \zeta^{-1/2} \left\langle e^{-L/(4\zeta)} f_1, f_2 \right\rangle$ is the Laplace transform of $v(t) := \frac{1}{\sqrt{\pi t}} \left\langle \cos(\sqrt{t}\sqrt{L}) f_1, f_2 \right\rangle$! Moreover, by the estimate of F we gave above, we have

$$|u(\zeta)| \leq C\|f_1\|\|f_2\||\zeta|^{-1/2} e^{-r^2 \operatorname{Re}\zeta}$$

To prove that L admits the finite speed of propagation property, we want to show that $\operatorname{supp} v \subseteq [r^2, \infty)$. To do this, we use a version of the Paley-Wiener Theorem [10, Theorem 7.4.3], which states that

$$\operatorname{supp} f \subseteq K \iff |\hat{f}(\xi)| \leq C e^{\sup_{x\in K} \langle x, \operatorname{Im}\xi \rangle} \qquad \forall \xi \in \mathbb{C}^n$$

for convex closed sets $K \subseteq \mathbb{R}^n$, and $\hat{f}(\xi)$ is the Fourier transform of f, extended to the whole of $\mathbb{C}^n$. We have the Laplace transform instead of the Fourier transform, so we need to substitute $\zeta = i\xi$. Then $\xi = -i\zeta$ and $\operatorname{Im}\xi = -\operatorname{Re}\zeta$. The right hand side becomes

$$|\mathcal{L}(f)(\zeta)| \leq Ce^{\sup_{x\in K}\langle x, -\operatorname{Re}\zeta\rangle}.$$

We want to apply this theorem for $K := [r^2, \infty)$ and $\zeta \in \mathbb{C}_+$. Then the inner product simplifies to a product, and we can estimate

$$e^{\sup_{x\geq r^2}(-x\operatorname{Re}\zeta)} = e^{-\inf_{x\geq r^2}(x\operatorname{Re}\zeta)} = e^{-r^2\operatorname{Re}\zeta}$$

where we used that $\operatorname{Re}\zeta > 0$. We have from before that

$$|u(\zeta)| \leq C\|f_1\|\|f_2\||\zeta|^{-1/2}e^{-r^2\operatorname{Re}\zeta}$$

For a fixed $\varepsilon > 0$ we can estimate for all $\zeta \in \mathbb{C}_+$ such that $|\zeta| > \varepsilon^2$

$$|u(\zeta)| \leq C\|f_1\|\|f_2\|\varepsilon^{-1}e^{-r^2\operatorname{Re}\zeta}$$

So we get that $\operatorname{supp}\mathcal{L}^{-1}(u\phi_\varepsilon) \subseteq [r^2, \infty)$ for ϕ_ε a holomophic function which is 0 for $|\zeta| < \varepsilon^2$, 1 for $|\zeta| > 2\varepsilon^2$ and $0 \leq \phi_\varepsilon \leq 1$. Now letting $\varepsilon \to 0$, we have $\operatorname{supp} v \subseteq [r^2, \infty)$, that is $\left\langle \cos(\sqrt{t}\sqrt{L})f_1, f_2 \right\rangle = 0$ for $t < r^2$. Putting that differently, we have $\left\langle \cos(t\sqrt{L})f_1, f_2 \right\rangle = 0$ for $t < r$. So we have proved the finite speed of propagation.

Conversely, assume the finite speed of propagation property. Then by the Hadamard Transmutation Formula, and the fact that for $(x, y) \in [r, \infty) \times [r, \infty)$

it holds that $|(x,y)| > r\sqrt{2}$ we have

$$
\begin{aligned}
|\left\langle e^{-sL} f_1, f_2 \right\rangle| &\leq \int_0^\infty |\left\langle \cos(t\sqrt{L}) f_1, f_2 \right\rangle| \frac{e^{-t^2/(4s)}}{\sqrt{\pi s}} \, \mathrm{d}t \\
&= \int_r^\infty |\left\langle \cos(t\sqrt{L}) f_1, f_2 \right\rangle| \frac{e^{-t^2/(4s)}}{\sqrt{\pi s}} \, \mathrm{d}t \\
&\leq \| \cos(t\sqrt{L}) f_1 \|_2 \|f_2\|_2 \int_r^\infty \frac{e^{-t^2/(4s)}}{\sqrt{\pi s}} \, \mathrm{d}t \\
&\leq C\|f_1\|_2\|f_2\|_2 \left(\int_r^\infty \frac{e^{-x^2/(4s)}}{\sqrt{\pi s}} \, \mathrm{d}x \int_r^\infty \frac{e^{-y^2/(4s)}}{\sqrt{\pi s}} \, \mathrm{d}y \right)^{1/2} \\
&\leq C\|f_1\|\|f_2\| \left(\int_r^\infty \int_r^\infty \frac{e^{-(x^2+y^2)/(4s)}}{\pi s} \, \mathrm{d}x \, \mathrm{d}y \right)^{1/2} \\
&\leq C\|f_1\|\|f_2\| \left(\int_0^{\pi/2} \frac{1}{\pi s} \int_{r\sqrt{2}}^\infty e^{-\rho^2/(4s)} \rho \, \mathrm{d}\rho \, \mathrm{d}\theta \right)^{1/2} \\
&= C\|f_1\|\|f_2\| \left(\frac{1}{2s} \int_{2r^2}^\infty e^{-u/(4s)} \frac{1}{2} \, \mathrm{d}u \right)^{1/2} \\
&= C\|f_1\|\|f_2\| \left(\frac{1}{4s} \left[-4se^{-u/(4s)} \right]_{2r^2}^\infty \right)^{1/2} \\
&= C\|f_1\|\|f_2\| \left(e^{-2r^2/(4s)} \right)^{1/2} \\
&= C\|f_1\|\|f_2\| e^{-r^2/(4s)}.
\end{aligned}
$$

This is the Davies-Gaffney estimate. □

Now we want to define a property which looks like the finite speed of propagation property, but we have replaced the cosine with a partition of unity. First I am going to repeat the definition of a partition of unity:

Definition 2.7. *Let $\psi \in C^\infty(\mathbb{R})$ have exponential decay. It is said to generate a partition of unity if*

$$\sum_{j=-\infty}^{\infty} \psi_j = 1 \qquad \textit{pointwise,}$$

where $\psi_j(x) := \psi(2^{-j}x)$.

Theorem 2.8. *Let L be a nonnegative self-adjoint linear operator on a metric measure space (X, d, μ). Let $\psi \in C^\infty(\mathbb{R})$ generate a partition of unity, such that $\operatorname{supp} \mathcal{F}\psi$ is compact. Then if L has the finite speed of propagation property with speed v with respect to the cosine, it has the Finite Speed of Propagation property with speed vR with respect to ψ, where R is such that $\hat{\psi}(\xi) = 0$ for $|\xi| > R$.*

Proof. As it is possible to rewrite the cosine by complex exponential powers, we can rerun the argument of last proof with $e^{it\sqrt{L}}$ instead of $\cos(t\sqrt{L})$:

$$\int_{-\infty}^{\infty} e^{it\sqrt{\lambda}} \dots \, \mathrm{d}t = \frac{1}{2}\int_{0}^{\infty} \cos(t\sqrt{\lambda}) \dots \, \mathrm{d}t$$

where the omitted part is an even function. We would end up with a property like

$$\left\langle e^{it\sqrt{L}} f_1, f_2 \right\rangle = 0 \qquad 0 < t < r$$

So, if L admits the finite speed of propagation property with respect to the cosine, it satisfies the Finite Speed of Propagation property with respect to the complex exponential. By a simple scaling argument, we can see that this will hold as well if L has the Finite Speed of Propagation property with speed v. (i.e., scale L accordingly).

We start the actual proof by writing ψ in terms of its Fourier transform, by using the Fourier Inversion Formula at time $t\sqrt{\lambda}$, $\lambda \geq 0$.:

$$\psi(t\sqrt{\lambda}) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{\psi}(\xi) e^{i\xi t\sqrt{\lambda}} \, \mathrm{d}\xi.$$

If we let $f_1, f_2 \in L^2(X, \mu)$, with $\operatorname{supp} f_i \subseteq U_i \subseteq X$, U_i open and $r := d(U_1, U_2)$ (as before), we have

$$\begin{aligned}
\left\langle \psi(t\sqrt{L}) f_1, f_2 \right\rangle &= \int \psi(t\sqrt{\lambda}) \, \mathrm{d} \left\langle E_\lambda f_1, f_2 \right\rangle \\
&= \int \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{\psi}(\xi) e^{i\xi t\sqrt{\lambda}} \, \mathrm{d}\xi \, \mathrm{d} \left\langle E_\lambda f_1, f_2 \right\rangle \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \int \hat{\psi}(\xi) e^{i\xi t\sqrt{\lambda}} \, \mathrm{d} \left\langle E_\lambda f_1, f_2 \right\rangle \, \mathrm{d}\xi \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{\psi}(\xi) \left\langle e^{i\xi t\sqrt{L}} f_1, f_2 \right\rangle \mathrm{d}\xi
\end{aligned}$$

We were able to use Fubini in the third line because $|e^{i\xi t\sqrt{\lambda}}| = 1$ for all $\xi \in \mathbb{R}$ and all $\lambda \geq 0$, and $\hat{\psi} \in L^2(\mathbb{R}, \mathcal{L})$ which follows from the exponential decay of ψ.

Now notice that we have the definition of $e^{i\xi t\sqrt{L}}$ in the last line. We know that when we apply this to f_1, and take the dot product with f_2, that the result is zero for $0 < \xi t < r/v$, that is, when $t < \frac{r}{v\xi}$. To get a bound which is uniform in ξ, we now use the compact support of $\hat{\psi}$:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{\psi}(\xi) \left\langle e^{i\xi t\sqrt{L}} f_1, f_2 \right\rangle \mathrm{d}\xi = \frac{1}{2\pi}\int_{-R}^{R} \hat{\psi}(\xi) \left\langle e^{i\xi t\sqrt{L}} f_1, f_2 \right\rangle \mathrm{d}\xi$$

As $\xi < R$, we have $0 < \frac{r}{vR} \leq \frac{r}{v\xi}$. So for $0 < t < \frac{r}{vR}$, we have

$$0 = \frac{1}{2\pi}\int_{-R}^{R} \hat{\psi}(\xi) \left\langle e^{i\xi t\sqrt{L}} f_1, f_2 \right\rangle \mathrm{d}\xi = \left\langle \psi(t\sqrt{L}) f_1, f_2 \right\rangle$$

But that means that L has the finite speed of propagation property with respect to ψ, with speed vR. $\square$

3 Generalising operator properties from one L^p space to the other

In this chapter we show that properties for an operator T, like being a contraction on L^∞ or being compact on L^2 will extend to all L^p spaces. This can be used to prove that a certain operator semigroup is a symmetric diffusion semigroup. We start with extending the property of being a contraction for self-adjoint (in the L^2-sense) operators T:

Theorem 3.1. *Let (X,μ) be a measure space. Let T be a self-adjoint operator on $L^2(X,\mu)$. Then if T can be extended to a bounded operator $L^\infty(X,\mu)$ which is contractive on $L^\infty(X,\mu)$, then it is a contraction on $L^p(X,\mu)$ for all $p \in [1,\infty]$.*

Proof. The first thing to prove is that T is a contraction on $L^1(X,\mu)$. To that end, note that $L^\infty(X,\mu)\cap L^2(X,\mu)$ is dense in $L^\infty(X,\mu)$. We will use that to estimate the L^1-norm in terms of $L^\infty(X,\mu)\cap L^2(X,\mu)$ functions.

Fix $\varepsilon > 0$. Then for $f \in L^1(X,\mu)$ there exists a $g \in L^\infty(X,\mu)$ with $\|g\|_\infty \leq 1$ such that $\|f\|_1 \leq |\langle f,g\rangle| + \varepsilon/2$. Moreover, there exists an $h \in L^\infty(X,\mu)\cap L^2(X,\mu)$ such that $\|h-g\|_\infty < \varepsilon/2\|f\|_1$ and $\|h\|_\infty \leq 1$. Then

$$\begin{aligned}
\|f\|_1 &\leq |\langle f,g\rangle| + \varepsilon/2\\
&\leq |\langle f,h\rangle| + |\langle f,g-h\rangle| + \varepsilon/2\\
&\leq |\langle f,h\rangle| + \|f\|_1\|g-h\|_\infty + \varepsilon/2\\
&\leq |\langle f,h\rangle| + \varepsilon/2 + \varepsilon/2\\
&\leq |\langle f,h\rangle| + \varepsilon.
\end{aligned}$$

As ε was chosen arbitrarily, we have just proven that

$$\|f\|_1 = \sup\{|\langle f,h\rangle| : h \in L^\infty(X,\mu)\cap L^2(X,\mu), \|h\|_\infty \leq 1\}$$

Now back to the proof in case $p=1$. By the above remark and the fact that T is self-adjoint, we get for $f \in L^1(X,\mu)\cap L^2(X,\mu)$

$$\begin{aligned}
\|Tf\|_1 &= \sup\{|\langle Tf,g\rangle| : g \in L^\infty(X,\mu)\cap L^2(X,\mu), \|g\|_\infty \leq 1\}\\
&= \sup\{|\langle f,T^*g\rangle| : g \in L^\infty(X,\mu)\cap L^2(X,\mu), \|g\|_\infty \leq 1\}\\
&= \sup\{|\langle f,Tg\rangle| : g \in L^\infty(X,\mu)\cap L^2(X,\mu), \|g\|_\infty \leq 1\}\\
&\leq \sup\{\|f\|_1\|Tg\|_\infty : g \in L^\infty(X,\mu)\cap L^2(X,\mu), \|g\|_\infty \leq 1\}\\
&\leq \|f\|_1 \sup\{\|g\|_\infty : g \in L^\infty(X,\mu)\cap L^2(X,\mu), \|g\|_\infty \leq 1\}\\
&= \|f\|_1
\end{aligned}$$

Hence T can be extended to a contraction on $L^1(X,\mu)$.

Let $p \in (1,\infty)$. Then we can apply the Riesz-Thorin Interpolation Theorem 2.4 with $p_0 = q_0 = 1$, $p_1 = q_1 = \infty$, $q = p$, $t = 1 - \frac{1}{p}$ to obtain that

$$\|Tf\|_p \leq \|T\|_1^{1/p}\|T\|_\infty^{1-1/p}\|f\|_p = 1^{1/p}1^{1-1/p}\|f\|_p = \|f\|_p$$

This concludes the proof. □

The following results are heavily based on Davies.[6, Thm 1.6.1 - 1.6.3]

Theorem 3.2. *Fix $1 \leq p_0 < p < p_1 \leq \infty$, and suppose that (X, μ) is a σ-finite measure space. Furthermore, suppose that the linear operator $T : L^{p_0}(X, \mu) \cap L^{p_1}(X, \mu) \to L^{p_0}(X, \mu) \cap L^{p_1}(X, \mu)$ can be extended to a compact operator on $L^{p_0}(X, \mu)$, and to a bounded operator on $L^{p_1}(X, \mu)$. Then T is compact on $L^p(X, \mu)$.*

Proof. If $\{E_r\}_{r=1}^n$ is a sequence of pairwise disjoint subsets of finite positive measure, then we can define a projection P on $L^q(X, \mu)$ for all $1 \leq q \leq \infty$ by

$$Pf := \sum_{r=1}^{n} \mu(E_r)^{-1} \mathbb{1}_{E_r} \int_{E_r} f \, \mathrm{d}\mu$$

This projection is of finite rank (its dimension is n), and it is a contraction on $L^q(X, \mu)$ for all $1 \leq q \leq \infty$. Denote q' for the Hölder conjugate of q. Then

$$\begin{aligned}
\|Pf\|_q^q &= \left\| \sum_{r=1}^{n} \mu(E_r)^{-1} \mathbb{1}_{E_r} \int_{E_r} f \, \mathrm{d}\mu \right\|_q^q \\
&= \int_X \left| \sum_{r=1}^{n} \mu(E_r)^{-1} \mathbb{1}_{E_r} \int_{E_r} f \, \mathrm{d}\mu \right|^q \mathrm{d}\mu \\
&= \sum_{r=1}^{n} \int_{E_r} \left| \sum_{k=1}^{n} \mu(E_k)^{-1} \mathbb{1}_{E_k} \int_{E_k} f \, \mathrm{d}\mu \right|^q \mathrm{d}\mu \\
&= \sum_{r=1}^{n} \int_{E_r} \left| \mu(E_r)^{-1} \mathbb{1}_{E_r} \int_{E_r} f \, \mathrm{d}\mu \right|^q \mathrm{d}\mu \\
&= \sum_{r=1}^{n} \int_X \left| \mu(E_r)^{-1} \mathbb{1}_{E_r}(y) \int_{E_r} f(x) \, \mathrm{d}\mu(x) \right|^q \mathrm{d}\mu(y) \\
&= \sum_{r=1}^{n} \mu(E_r)^{-q} \left| \int_{E_r} f \, \mathrm{d}\mu \right|^q \int_X \mathbb{1}_{E_r} \, \mathrm{d}\mu \\
&= \sum_{r=1}^{n} \mu(E_r)^{1-q} \left| \int_{E_r} \mathbb{1} \cdot f \, \mathrm{d}\mu \right|^q \\
&\leq \sum_{r=1}^{n} \mu(E_r)^{1-q} \left(\int_{E_r} |\mathbb{1}|^{q'} \, \mathrm{d}\mu \right)^{q/q'} \int_{E_r} |f|^q \, \mathrm{d}\mu \\
&= \sum_{r=1}^{n} \mu(E_r)^{1-q} \mu(E_r)^{q/q'} \int_{E_r} |f|^q \, \mathrm{d}\mu \\
&\leq \|f\|_q^q
\end{aligned}$$

In the third and fourth line, we used that the E_r are disjoint, and that $\operatorname{supp} \mathbb{1}_{E_r} = E_r$ in order to change the domain of integration in the fifth line.

The fact that X is σ-finite ensures that there is a sequence P_n of such projections such that $P_n \to I$ strongly on $L^q(X,\mu)$, for all $1 \leq q < \infty$: By the σ-finiteness we have a sequence of pairwise disjoint sets of finite positive measure $\{E_r\}_{r=1}^\infty$ such that $\mu(E_r) < \infty$ and $\bigcup_r E_r = X$. Then we set $P_1 = \mu(E_1)^{-1}\mathbb{1}_{E_1}\int_{E_1} f\,\mathrm{d}\mu$ and if we constructed P_n, we can construct P_{n+1} as follows: We split up all sets $\{F_k\}_{k=1}^N$ constructing P_n in two, obtaining new sets of positive measure $\{G_k\}_{k=1}^{2N}$. Then we set

$$P_{n+1}f = \sum_{k=1}^{2N} \mu(G_k)^{-1}\mathbb{1}_{G_k}\int_{G_k} f\,\mathrm{d}\mu + \mu(E_{n+1})^{-1}\mathbb{1}_{E_{n+1}}\int_{E_{n+1}} f\,\mathrm{d}\mu.$$

By the Lebesgue differentiation theorem, we have that each of those integrals converge to f almost everywhere if $n \to \infty$. Moreover, we have that $\bigcup G_k \cup E_{n+1} \uparrow X$ when $n \to \infty$. Hence $P_{n+1}f \to f$ strongly in $L^q(X,\mu)$.

As T is compact on $L^{p_0}(X,\mu)$, we have $\|T - P_nT\|_{p_0\to p_0} \to 0$ as $n \to \infty$:

As $\|T - P_nT\|_{p_0\to p_0} = \sup_{\|f\|_{p_0}\leq 1} \|(T - P_nT)f\|_{p_0}$, we can find a sequence $(f_k)_{k=1}^\infty \subset L^{p_0}(X,\mu)$ such that $\|f_k\|_{p_0} \leq 1$ for all $k \in \mathbb{N}$, and such that $\|(I - P_n)Tf_k\|_{p_0} = \|(T - P_nT)f_k\|_{p_0} \to \|T - P_nT\|_{p_0\to p_0}$. Now using the compactness of T, we can find a subsequence $(f_{k_l})_{l=1}^\infty$ such that $Tf_{k_l} \to g \in L^{p_0}(X,\mu)$ when $l \to \infty$. We get

$$\begin{aligned}
&\lim_{n\to\infty} \|T - P_nT\|_{p_0\to p_0}\\
&\qquad= \lim_{n\to\infty}\lim_{l\to\infty} \|(T - P_nT)f_{k_l}\|_{p_0}\\
&\qquad\leq \lim_{n\to\infty}\lim_{l\to\infty} (\|Tf_{k_l} - g\|_{p_0} + \|g - P_ng\|_{p_0} + \|P_ng - P_nTf_{k_l}\|_{p_0})\\
&\qquad\leq \lim_{n\to\infty}\lim_{l\to\infty} (\|Tf_{k_l} - g\|_{p_0} + \|g - P_ng\|_{p_0} + \|P_n\|\|g - Tf_{k_l}\|_{p_0})\\
&\qquad\leq \lim_{n\to\infty}\lim_{l\to\infty} (2\|Tf_{k_l} - g\|_{p_0} + \|g - P_ng\|_{p_0})\\
&\qquad= 2\cdot 0 + 0 = 0
\end{aligned}$$

So $P_nT \to T$ in operator norm.

Furthermore, as T is bounded on $L^{p_1}(X,\mu)$, it holds that $\|T - P_nT\|_{p_1\to p_1} \leq \|T\| + \|P_n\|\|T\| \leq 2\|T\|_{p_1\to p_1}$, and so these terms are uniformly bounded. Now the Riesz-Thorin Interpolation Theorem 2.4 implies that

$$\lim_{n\to\infty} \|T - P_nT\|_{p\to p} \leq \lim_{n\to\infty} \|T - P_nT\|_{p_0\to p_0}^{\theta} \|T - P_nT\|_{p_1\to p_1}^{1-\theta} = 0$$

But P_nT is of finite rank, so T is a limit of finite rank operators in $L^p(X,\mu)$, hence it is compact on $L^p(X,\mu)$. □

Note 3.3. We didn't use the assumption that $p_0 < p_1$. In fact, the theorem is even true when $1 \leq p_1 < p < p_0 < \infty$, with exactly the same proof! We only used that $L^p(X,\mu)$ could be interpolated by $L^{p_0}(X,\mu)$ and $L^{p_1}(X,\mu)$.

With this compactness, it turns out the spectrum of T is independent of p! The following proof is followed almost literally from [6, Theorem 1.6.2]

Proposition 3.4. *Suppose that $1 \leq p_0 < p_1 \leq \infty$, that (X, μ) is a σ-finite measure space, and that $T : L^{p_0}(X, \mu) \cap L^{p_1}(X, \mu) \to L^{p_0}(X, \mu) \cap L^{p_1}(X, \mu)$ can be extended to a compact operator $L^{p_0}(X, \mu) \to L^{p_0}(X, \mu)$ and to a bounded operator $L^{p_1}(X, \mu) \to L^{p_1}(X, \mu)$. In this case, the spectrum of T is the same for all $p \in [p_0, p_1)$, and the spectral projections corresponding to non-zero eigenvalues are independent of p on $L^{p_0}(X, \mu) \cap L^{p_1}(X, \mu)$.*

Proof. We will show that for arbitrary $p \in (p_0, p_1)$ the spectrum and spectral projections coincide with the spectrum and spectral projections of T on $L^{p_0}(X, \mu)$. Denote the extension of T to $L^p(X, \mu)$ by T_p, and define T_{p_0} similarly. By above theorem, the set

$$S := \sigma(A_{p_0}) \cup \sigma(A_p) \subseteq \mathbb{C}$$

is countable and closed with 0 as only limit point. Now, it holds for $|z| > \max(\|T_{p_0}\|, \|T_p\|)$ and $f \in L^{p_0}(X, \mu) \cap L^{p_1}(X, \mu)$ that

$$\begin{aligned}
(T_p - zI)^{-1} f &= z^{-1}(z^{-1}T_p - I)^{-1} f \\
&= z^{-1} \sum_{n=1}^{\infty} z^{-n} T_p^n f \\
&= z^{-1} \sum_{n=1}^{\infty} z^{-n} T_{p_0}^n f \\
&= (T_{p_0} - zI)^{-1} f
\end{aligned}$$

and it holds for all $z \notin S$ by holomorphic extension (recalling that the resolvent of an operator A as function of z ($z \mapsto R(z, A)$) is holomorphic on the resolvent set $\rho(A)$). If $0 \neq s \in S$, and γ is a small enough contour looping once around s, we can use the Dunford-Riesz calculus [8, Chapter 1] to get the spectral projection for s for such f:

$$\begin{aligned}
P_p f &= \frac{1}{2\pi i} \int_\gamma (z - T_p)^{-1} f \, \mathrm{d}z \\
&= \frac{1}{2\pi i} \int_\gamma (z - T_{p_0})^{-1} f \, \mathrm{d}z \\
&= P_{p_0} f
\end{aligned}$$

As $L^{p_0}(X, \mu) \cap L^{p_1}(X, \mu)$ is dense in both $L^p(X, \mu)$ and $L^{p_0}(X, \mu)$, it follows that P_p and P_{p_0} have the same finite dimensional range. By some basic functional calculus, and noticing that $P_p = \mathbb{1}_{\{s\}}(T_p)$, we have $T_p P_p = (z \mapsto z\mathbb{1}_{\{s\}}(z))(T_p)$. Now the spectral mapping theorem tells that $\sigma(T_p P_p) = \{\lambda \mathbb{1}_{\{s\}}(\lambda) : \lambda \in \sigma(T_p)\}$. As P_p and P_{p_0} have the same range, and T_p and T_{p_0} coincide on $L^{p_0}(X, \mu) \cap L^p(X, \mu)$, we have that $\sigma(T_p P_p) = \sigma(T_{p_0} P_{p_0}) = \{\lambda \mathbb{1}_{\{s\}}(\lambda) : \lambda \in \sigma(T_p)\}$. Moreover, $\sigma(T_p(I - P_p)) = \sigma(T_p) \backslash \{s\}$, and the same statement is true for T_{p_0}. Hence if $s \notin \sigma(T_p)$, we have $\sigma(T_{p_0} P_{p_0}) = \{0\}$, which means that $\sigma(T_{p_0}) = \sigma(T_{p_0} P_{p_0}) \cup \sigma(T_{p_0}(I - P_{p_0})) = \sigma(T_{p_0}) \backslash \{s\} \cup \{0\}$, in other words,

$s \notin \sigma(T_{p_0})$. However, if $s \in \sigma(T_p)$ we have $\sigma(T_{p_0} P_{p_0}) = \{s, 0\}$, from which it follows that $s \in \sigma(T_{p_0})$. So $s \in \sigma(T_p) \iff s \in \sigma(T_{p_0})$.

We have not considered the case $s = 0$ yet. But 0 always lies in both $\sigma(T_p)$ and $\sigma(T_{p_0})$, unless X is finite, in which case the theorem is trivial, as it then holds that $L^p(X, \mu) \subseteq L^{p_0}(X, \mu)$. □

Note 3.5. Once again, this theorem also holds when $p_1 < p_0$. We again have that T_p is compact, because of the last remark. The only thing that fails when $p < p_0$, is the last statement about finite-dimensional X. But in this case $L^{p_0}(X, \mu) \subseteq L^p(X, \mu)$ for each $p_1 < p < p_0$, which will give a similar statement about the case $s = 0$.

At last, we tie everything together.

Corollary 3.6. *Let (X, μ) be a σ-finite measure, and let $(T_t)_{t \geq 0}$ be a semigroup on (X, μ), which is bounded on every $L^p(X, \mu)$, $p \in [1, \infty]$, such that T_t is compact on $L^2(X, \mu)$ for all $t \geq 0$. Then T_t is compact on every $L^p(X, \mu)$, $1 < p < \infty$. Moreover, the spectrum of T_t is independent of $p \in (1, \infty)$, for all $t \geq 0$, and every L^2 eigenfunction of T_t is in $L^p(X, \mu)$ for all $1 < p < \infty$.*

Proof. This follows from above theorems and remarks by choosing $p_0 = 2$ and $p_1 = \infty$ or $p_1 = 1$. The statement about the eigenfunctions follows from the identity of the spectral projections: For each $s \in \sigma(T)$, the spectral projections P_2 and P_p, such as defined above, have the property that $P_p f = P_2 f = f$ for each eigenfunction f corresponding to s in $L^2(X, \mu)$. But this also says that f is an eigenfunction for s in $L^p(X, \mu)$. □

4 Symmetric Diffusion Semigroups

4.1 Basic definitions

To introduce the symmetric diffusion semigroup, we need to describe the space we work with.

Definition 4.1. *Let (X, d, μ) be a (quasi-)metric measure space, and let $x \in X$, $r > 0$. We denote the open ball around x with radius r with*

$$B(x, r) := \{y \in X : d(y, x) < r\}.$$

We say that (X, d, μ) is of homogeneous type *if μ is a non-negative doubling Borel measure, which means that there exists a constant $C_X > 0$ such that for every $x \in X$, $\delta > 0$ the following inequality holds:*

$$\mu(B(x, 2\delta)) \leq C_X \mu(B(x, \delta)).$$

We will assume that $\mu(B(x, \delta)) < \infty$ for all $x \in X$ and all $\delta > 0$. Note that this automatically implies that μ is σ-finite. If we consider the sequence $(B(x, 2^n\delta))_{n=1}^{\infty}$, we see that, by the above estimate, all of its elements have finite measure. Moreover, it holds that $B(x, 2^n\delta) \uparrow X$.

In the following, we assume that (X, d, μ) is a space of homogeneous type, where d is assumed to be a metric. In case where a theorem or lemma only specifies (X, μ) in its hypothesis, we assume that this measure space is σ-finite.

Notation 4.2. *We denote the distance between a point $x \in X$ and a subset $S \subseteq X$ by $d(x, S) := \operatorname{dist}(x, S)$. Moreover, the diameter of a set $S \subseteq X$ is denoted by $d(S) := \operatorname{diam}(S)$. The context will make clear what "d" means.*

The benefit of this notation is that it works if you have multiple metrics on the same space. So if you have a different metric ρ, then the diameter of S with respect to ρ is denoted by $\rho(S)$.

We are finally ready to introduce the symmetric diffusion semigroup:

Definition 4.3 ([2, Def 13])**.** *Let $(T_t)_{t\geq 0}$ be a semigroup of linear operators on (X, μ), with each T_t mapping $L^2(X, \mu)$ into itself. Suppose furthermore that $(T_t)_{t\geq 0}$ is strongly continuous in $L^2(X, \mu)$, which means that for all $f \in L^2(X, \mu)$ it holds that $T_t f \to f$ if $t \to 0$. Then $(T_t)_{t\geq 0}$ is called a symmetric diffusion semigroup if it satisfies the following properties:*

(i) For each $t \geq 0$, T_t can be extended to a contraction on $L^p(X, \mu)$ for every $1 \leq p \leq \infty$, i.e. $\|T_t\|_{p\to p} \leq 1$.

(ii) Each T_t is self-adjoint.

(iii) T_t is positivity preserving: $Tf \geq 0$ for every $0 \leq f \in L^2(X, \mu)$.

(iv) The semigroup has a positive self-adjoint generator.

We call $(T_t)_{t\geq 0}$ *a compact symmetric diffusion semigroup if it also satisfies the following property:*

(v) T_t *is a compact operator on* $L^2(X,\mu)$ *for each* $t>0$.

Definition 4.4 ([2, Def 8, 16 and 18])**.** *Let* $(T_t)_{t\geq 0}$ *be a symmetric diffusion semigroup.*

1. *If* $(T_t)_t$ *is compact, with spectrum* $\lambda_0 \geq \lambda_1 \geq \ldots \geq 0$, *it is said to have* γ*-strong decay, for a* $\gamma > 0$, *if there exists a constant* $C > 0$ *such that for each* $\lambda \in (0,1)$:
$$\#\{k : \lambda_k \geq \lambda\} \leq C\left({}^2\log\frac{1}{\lambda}\right)^{\gamma}.$$

2. *A function* f *is called* η*-local around* $x \in X$ *for some* $\eta > 0$, *if* $\operatorname{supp} f \subseteq B(x,\eta)$. *Likewise, a family of functions* $\{f_k\}_{k\in K}$ *is called* η*-local, if there exists a set* $\{x_k\}_{k\in K}$ *such that* $\operatorname{supp} f_k \subseteq B(x_k,\eta)$ *for each* $k \in K$. *This set* $\{x_k\}_{k\in K}$ *is called the center set for this family.*

3. $(T_t)_{t\geq 0}$ *acts* η*-locally for some* $\eta > 0$, *if for every* $x \in X$, *every* $t \geq 0$ *and every* f δ*-local around* x *the function* $T_t f$ *is* $\eta+\delta$*-local around* x.

Note 4.5. Some of the properties are redundant. By the theorems in the last chapter, we know that if T_t is a contraction on $L^\infty(X,\mu)$, then it is a contraction on every $L^p(X,\mu)$ for $p \in [1,\infty]$. Moreover, we have proven that if T_t is compact on $L^2(X,\mu)$, then it is compact on all $L^p(X,\mu)$ for $1 < p < \infty$. And the spectrum is independent of $p \in (1,\infty)$. Lastly, every L^2 eigenfunction is in $L^p(X,\mu)$, $1 < p < \infty$.

Notation 4.6. *We will denote the spectrum of* $T := T_1$ *by* $\sigma(T)$. *Then by the spectral mapping theorem, it follows that* $\sigma(T_t) = \{\lambda^t\}_{\lambda\in\sigma(T)}$. *We will denote* $\{\xi_\lambda\}_{\lambda\in\sigma(T)}$ *for the corresponding (orthogonal!) basis of eigenvectors, where we assume without loss of generality that they are normalized in* $L^2(X,\mu)$.

In the upcoming few sections, we assume that (T_t) is a compact symmetric diffusion semigroup with γ-strong decay, that acts η-locally on X, unless stated otherwise.

4.2 Multiresolution analysis

The following construction is courtesy of Coifman and Maggioni [2, Sec. 4.3]. To set up the wavelets, we need to set up a multiresolution analysis. To this end, we discretize the symmetric diffusion semigroup $(T_t)_t$ at the times $t_j = 2^j - 1 = 1 + 2 + \ldots + 2^{j-1}$. Here the "2" is arbitrary, we could have used any other factor > 1.

With this discretization, and a "precision" $\varepsilon \in (0,1)$, we define portions of the spectrum by $\sigma_j(T) = \{\lambda \in \sigma(T) : \lambda^{t_j} \geq \varepsilon\}$. Moreover, we define subspaces V_j by $V_j := \operatorname{span}\{\xi_\lambda : \lambda \in \sigma_j(T)\}$. If we let $V_{-1} := L^2(X,\mu)$, then the set of subspaces $\{V_j\}_{j\geq -1}$ is a multiresolution analysis, that is, it satisfies the following properties:

(i) $V_{-1} = L^2(X, \mu)$, $\bigcap_j V_j = \text{span}\{\xi_i : \lambda_i = 1\}$.

(ii) $V_{j+1} \subseteq V_j \, \forall \, j \in \mathbb{Z}, j \geq -1$.

(iii) $\{\xi_\lambda : \lambda^{t_j} \geq \varepsilon\}$ is an orthonormal basis for V_j.

Indeed, the first part of the first statement is true by definition. For the second part, one has to note that for any $0 < \alpha < 1$, the sequence $(\alpha^n)_{n=1}^\infty$ is strictly decreasing. Thus it holds that $\sigma_{j+1}(T) \subseteq \sigma_j(T)$, and so for the intersection it is true that $\bigcup_{j=0}^\infty \sigma_j(T) = \{\lambda \in \sigma(T) : \lambda^{t_j} > \varepsilon \, \forall \, j \in \mathbb{N}_0\} = \{1\} \cup \sigma(T)$. So when one looks at the corresponding spans, it holds that $V_{j+1} \subseteq V_j$, which proves statement (ii), and $\lim_{j\to\infty} V_j = \bigcup_{j=-1}^\infty V_j = \text{span}\{\xi_i : \lambda_i = 1\}$. We use this notation as the eigenvalue '1' may have multiple eigenvectors, and the above introduced (abuse of) notation ξ_λ does not take that into account.

For the third statement, note that $\{\xi_\lambda : \lambda^{t_j} \geq \varepsilon\}$ is a basis for V_j, as we defined the latter to be the span of these functions. Moreover, we chose the basis of eigenfunctions $\{\xi_\lambda\}_{\lambda\in\sigma(T)}$ to be orthonormal, so that is also true for the basis of V_j, since it is a subset of this orthonormal basis.

The next two chapters of the paper will cover how to get a better orthonormal basis than the one given above. This is necessary, as the wavelets should be localized, and the basis $\{\xi_\lambda\}_{\lambda\in\sigma_j(T)}$ is in general not so. With this better base of V_j, it will be easy to construct the wavelets.

5 Wavelets

In this chapter, we define what a wavelet is.

Definition 5.1. *A family of functions Ψ is said to consist of p-wavelets for some $p > 0$ if for each $\psi \in \Psi$ the following conditions hold*

i) *ψ has exponential decay.*

ii) *For all $k \in \mathbb{N}_0$, $k < p$,*

$$\int_X x^k \psi(x)\,\mathrm{d}\mu(x) = 0$$

iii) $L^2(X,\mu) = \overline{\operatorname{span}\Psi}$

The classical way to construct these wavelets, as described in [5, 12], is using a multiresolutional analyses, as described in last section. In these publications, there is an additional requirement, that there exists a function ϕ, such that the translations $(\phi(x-n))_{n\in\mathbb{N}}$ forms a Riesz basis for V_0, and the other V_j are there defined as $f \in V_j \iff f(2\cdot) \in V_{j-1}$. It follows that $V_{j+1} \subseteq V_j$. Then it follows that the dilated and translated system $2^{-m/2}\phi(2^{-m}x - n)$ forms a Riesz basis for V_m.

We now take a look at the orthogonal complement of V_{j+1} inside V_j: $V_{j+1} = V_j \oplus W_j$. Then it follows that W_j is a space spanned by wavelets. These wavelets can be constructed as follows: Note that $V_0 \subseteq V_{-1}$. So in particular, ϕ can be written in terms of the basis V_{-1}: $\phi = \sum_n c_n \phi(2\cdot - n)$. Then one can proof that $\psi := \sum_n (-1)^n c_{n+1}\phi(2\cdot + n)$ is a wavelet, and $(\psi(2^{-j}\cdot - n))_n$ forms a Riesz basis for W_j. By the property that $L^2 = \bigoplus_j V_j$, and the property that $V_0 = W_1 \oplus V_1 = W_1 \oplus W_2 \oplus V_2$, etc. it follows that $L^2 = \bigoplus_j W_j$. And so $(\psi(2^{-j}\cdot - n))_{n,j}$ spans L^2. The proofs of the previous statements, and the more rigorous construction can be found in [12].

We now go back to the multiresolution analysis we defined in previous section. As above, we define the spaces W_j to be the orthogonal complement of V_{j+1} within V_j, such that the relation $V_j = V_{j+1} \oplus W_j$ holds. Then the direct orthogonal sum

$$L^2(X,\mu) = \bigoplus_{j=-1}^{\infty} W_j$$

is a wavelet decomposition of the space. Note that in this setting, we do not have a "master" wavelet ψ, as we constructed above, because there is no equivalence to the function ϕ above. But W_j is spanned by wavelets.

6 Construction of an orthogonal basis

In this chapter, we will follow the construction by Coifman and Maggioni [2] to build an orthogonal basis of localized functions, which will span the V_j, which we defined in Section 4.2, up to a precision ε. To be able to do this, we will introduce a way to orthonormalize a set of functions at different scales and locations. But first, we need a few definitions.

Definition 6.1 ([2, Thm 6])**.** *Let (X, d, μ) be a space of homogenous type. A collection of open subsets $\mathcal{Q} := \left\{\{Q_{j,k}\}_{k\in\mathcal{K}_j}\right\}_{j\in\mathbb{Z}} \subseteq \mathcal{P}(X)$ is called a family of dyadic cubes for X if there exist constants $\delta_X > 1, \eta > 0, c_1, c_2 \in (0,\infty)$, depending on C_X, such that the following properties hold:*

(i) For every $j \in \mathbb{Z}$, $\mu(X\backslash \bigcup_{k\in\mathcal{K}_j} Q_{j,k}) = 0$.

(ii) For $j \geq j_0$, either $Q_{j_0,k} \subseteq Q_{j,k'}$ or $\mu(Q_{j,k} \cap Q_{j_0,k'}) = 0$.

(iii) For each $j \in \mathbb{Z}, k \in \mathcal{K}_j$ and $j' < j$, there exists a unique k' such that $Q_{j',k} \subseteq Q_{j,k'}$.

(iv) Each $Q_{j,k}$ contains a point $x_{j,k}$, called the center of $Q_{j,k}$, such that

$$B(x_{j,k}, \min\{c_1\delta_X^j, d(X)\}) \subseteq Q_{j,k} \subseteq B(x_{j,k}, c_2\delta_X^j)$$

(v) For each $j \in \mathbb{Z}$ and $k \in \mathcal{K}_j$, if we define

$$\partial_t Q_{j,k} := \{x \in Q_{j,k} : d(x, X\backslash Q_{j,k}) \leq t\delta_X^j\},$$

then $\mu(\partial_t Q_{j,k}) \leq c_2 t^\eta \mu(Q_{j,k})$.

For each $j \in \mathbb{Z}$, the set $\{Q_{j,k}\}_{k\in\mathcal{K}_j}$ is called the set of dyadic cubes at scale j, and the set of ponts $\Gamma_j := \{x_{j,k}\}_{k\in\mathcal{K}_j}$ is called the set of dyadic centers at scale j. For each $j \in \mathbb{Z}$, the unique dyadic cube at scale j containing $x \in X$ will be denoted by $Q_j(x)$.

One can prove that such a family of dyadic cubes always exists for a space of homogeneous type. The uniqueness of $Q_j(x)$ follows from property (ii) and (iii): property (ii) gives that either Q_{j,k_1} and Q_{j,k_2} coincide, or that they are disjoint up to a null set. Then property (iii) says that they must be disjoint, otherwise the choice of k' is not unique.

For the Euclidean space $\mathbb{R}^n$, we get the classical dyadic cubes back with choices $\delta_X = 2, \eta = 1, c_1 = 1, c_2 = \sqrt{n}$ and $\mathcal{K}_j = 2^j\mathbb{Z}^n$.

From now on, we will redefine the t_j from the last chapter as $\delta_X^j - 1$ instead of $2^j - 1$.

Definition 6.2 ([2, Def 12])**.** *Let H be a Hilbert space and $\{v_k\}_{k\in\mathcal{K}} \subseteq H$. Fix $\varepsilon > 0$. A set of vectors $\{\xi_i\}_{i\in\mathcal{I}}$ ε-spans $\{v_k\}_{k\in\mathcal{K}}$ if for every $k \in \mathcal{K}$:*

$$\|P_{\text{span}\{\xi_i : i\in\mathcal{I}\}} v_k - v_k\|_H \leq \varepsilon$$

where P_V denotes the orthogonal projection onto the subspace V. We will use the notation $\langle \{v_k\}_{k\in\mathcal{K}}\rangle \subseteq \langle \{\xi_i\}_{i\in\mathcal{I}}\rangle_\varepsilon$. We define

$$\dim_\varepsilon(\{v_k\}_{k\in\mathcal{K}}) := \inf\{\dim(V') : V' \varepsilon\text{-spans } \{v_k\}_{k\in\mathcal{K}}\}$$

Remarks 6.3. Note that this definition is more general than just a different basis for the same span. This also means that there are situations where $\dim_\varepsilon(\{v_k\}_{k\in\mathcal{K}}) \leq \dim(\{v_k\}_{k\in\mathcal{K}})$. For instance, in $\mathbb{R}^2$, $\{\hat{x}\}$ ε-spans $\{\hat{x}, \varepsilon\hat{y}\}$, and so $\dim_\varepsilon(\{\hat{x}, \varepsilon\hat{y}\}) \leq 1 < 2 = \dim(\{\hat{x}, \varepsilon\hat{y}\})$. Also note that in the definition of an ε-span, $\{\xi_i\}_i$ is a basis for each of the vectors v_k, not for their span. While $\{\hat{x}\}$ ε-spans $\{\hat{x}, \varepsilon\hat{y}\}$, it does not ε-span $\{\hat{x}, \hat{y}\}$, while these two sets of vectors have the same span.

The precision ε comes from the definition of V_j in the last chapter. It turns out that $\langle V_j\rangle \subseteq \langle T_{t_j} V_0\rangle_\varepsilon$, hence we have to deal with this ε in the rest of the proofs. But in some situations we can just take a different basis for the same span, and this basis is a ε-span for each $\varepsilon > 0$.

Notation 6.4. *If Ψ is a family of functions and $S \subseteq X$, we let*

$$\Psi|_S := \{\psi \in \Psi : \operatorname{supp}\psi \subseteq S\}$$

We are finally ready to present the orthogonalization procedure. As the proposition and its proof are both notation-heavy, we first explain what happens in a more comprehendable way:

We have a set of functions we want to orthogonalize: Ψ. In the proposition, we assume that all of those functions have their support in sets with the diameter smaller than a given constant. We split the space up into dyadic cubes, starting from cubes with a diameter larger than this constant. Then we only look at the functions which support are in precisely one cube (that is, their support does not cross the boundary to a neighbor of this cube). Looking at each individual cube in this layer, we orthogonalize the functions with support in this cube. This makes sure that they will still have support in the cube. By considering the union of these orthogonalized functions, we have a basis for this layer.

Then we proceed with the parents of the cubes. As we have considered all functions whose support lies fully in one of the children of the new cubes we consider, we look at the functions which lie fully in one of these cubes, but their support "cross" the boundary of the children of this cube. Then we orthogonalize these functions with respect to the new basis for the last layer and to themselves. Once we are done with a layer, we go to the parents, and consider the functions whose support lies fully in one of these cubes, but whose support intersect with multiple children of the considered cube. At the end, we have considered all the functions, and we have made a set which is mutually orthogonal.

The last thing that the proof does is checking the dimensions of these orthogonal sets, and making sure that this dimension is still comparable against the size of the dyadic cubes in the corresponding layer.

The proof is followed almost literally from [2, Prop. 22]

Proposition 6.5. *Let (X, d, μ) be a space of homogeneous type, $d(X) < \infty$ and $\mathcal{Q} = \{Q_{j,k_{k\in\mathcal{K}_j}}\}_{j\in\mathbb{Z}}$ a family of dyadic cubes, and $\delta_X > 1, \eta > 0, c_1, c_2 > 0$ as in Definition 6.1. Fix $J > 0$ and assume $X \in \mathcal{Q}$, more precisely $X = Q_{J+j_X,k}$ for a $j_X \leq {}^{\delta_X}\log(c_1^{-1}\delta_X^{-J}d(X))$. Fix $\varepsilon > 0$. Let $\Psi = \{\psi_x\}_{x\in\Gamma}$ be an $\alpha\delta_X^J$-local family, $\alpha \leq c_1$, with center set Γ. Suppose Ψ is "uniformly locally finite-dimensional", in the sense that there exist $c'_\varepsilon, c''_\varepsilon > 0$ such that for all $k \in K_J$,*

$$\dim_{c_1\delta_X^J d(X)^{-1}\varepsilon}(\Psi|Q_{J,k}) \leq c'_\varepsilon \mu(Q_{J,k}),$$

and for all $l \geq 0, k \in \mathcal{K}_{J+l}$,

$$\dim_{c_1\delta_X^{J+l}(2d(X))^{-1}\varepsilon}(\{\psi_x\}_{x\in\Gamma\cup\partial_{\alpha\delta_X^{-l}}(Q_{J+l,k})}) \leq c''_\varepsilon(\alpha\delta_X^{-l})^\eta \mu(Q_{J+l,k}).$$

Then there exists an orthonormal basis

$$\Phi := \{\Phi_l\}_{l=0,\ldots,L} = \Big\{\{\{\phi_{l,k,i}\}_{i\in\mathcal{I}_{l,k}}\}_{k\in\mathcal{K}_{J+l}}\Big\}_{l=0,\ldots,L}$$

where $L \leq j_X$, such that

(i) $\langle\Psi\rangle \subseteq \langle\Phi\rangle_{(j_X+1)\varepsilon}$.

(ii) For $l = 0, \ldots, L$, all $k \in \mathcal{K}_{J+l}$ and all $i \in \mathcal{I}_{l,k}$, $\operatorname{supp}\phi_{l,k,i} \subseteq Q_{J+l,k}$. In particular, $\phi_{l,k,i}$ is $c_2\delta_X^{J+l}$-local.

(iii) For $l > 0$,

$$\left\langle \bigcup_{k\in\mathcal{K}_{J+l}} \Psi|_{Q_{J+l,k}} \right\rangle \subseteq \left\langle \bigcup_{l'=0}^{l} \Phi_{l'} \right\rangle_{(l+1)\varepsilon}$$

and $\#I_{l,k} \leq c''_\varepsilon(\alpha\min\{\delta_X^{-(l-1)}, 1\})^\eta\mu(Q_{J+l,k})$.

Proof. We will need to orthonormalize various sets of functions V, to get an orthonormal basis for their ε-span, containing $\dim_\varepsilon(V)$ elements. For this, one can use a modified version of the Gram-Schmidt algorithm. We will call this process ε-orthogonalization.

The first "layer" $l = 0$ is constructed as follows: For each dyadic cube $Q_{J,k}$ at scale J, consider $\tilde{\Psi}_{0,k} := \Psi|_{Q_{J,k}}$, and $c_1\delta_X^J d(X)^{-1}\varepsilon$-orthonormalize this set of functions to obtain $\Phi_0 : \{\phi_{0,k,i}\}_{i\in\mathcal{I}_{0,k}}$. Then property (ii) is satisfied by construction (when you orthonormalize only functions with support in $Q_{J,k}$, the basis must have support in $Q_{J,k}$ again). Also observe that

$$\left\langle \bigcup_{k\in\mathcal{K}_J} \tilde{\Psi}_{0,k} \right\rangle \subseteq \langle\Phi_0\rangle_\varepsilon$$

as the restriction of every function on the left side to every dyadic cube $Q_{J,k}$ is $c_1\delta_X^J d(X)^{-1}\varepsilon$ approximated in Φ_0 by construction, and as all the $Q_{J,k}$ are disjoint, and have size at least $c_1\delta_X^J$ it follows that there are at most $(c_1\delta_X^J)^{-1}d(X)$ such cubes, hence every function on the left hand side is ε-approximated.

We move on with the second "layer", $l = 1$. There we consider

$$\begin{aligned}\tilde{\Psi}_{1,k} &:= \Psi|_{Q_{J+1,k}} \setminus \bigcup_{Q_{J,k'} \subseteq Q_{J+1,k}} \Psi|_{Q_{J,k'}} \\ &= \{\psi \in \Psi : \operatorname{supp}\psi \subseteq Q_{J+1,k} \text{ but } \operatorname{supp}\psi \not\subseteq Q_{J,k'}\ \forall Q_{J,k'} \subseteq Q_{J+1,k}\} \\ &\subseteq \bigcup_{Q_{J,k'} \subseteq Q_{J+1,k}} \{\psi_x \in \Psi : x \in \partial_\alpha Q_{J,k'}\}.\end{aligned}$$

This last inclusion holds, as $\operatorname{supp}\psi_x \subseteq Q_{J+l,k}$ implies $x \in Q_{J,k'}$ for some $Q_{J,k'} \subseteq Q_{J+1,k}$. Now, $\operatorname{supp}\psi \not\subseteq Q_{J,k'}$ together with the $\alpha\delta_X^J$-locality forces that $d(x, Q_{J+1,k} \backslash Q_{J,k'}) < \alpha\delta_X^J$, or put differently, $x \in \partial_\alpha Q_{J,k'}$.

From the assumptions (with $l = 0$) we deduce that

$$\begin{aligned}\dim_{c_1\delta_X^J(2d(X))^{-1}\varepsilon}(\tilde{\Psi}_{1,k}) &\leq \sum_{Q_{J,k'} \subseteq Q_{J+1,k}} \dim_{c_1\delta_X^J(2d(X))^{-1}\varepsilon}(\{\psi_x \in \Psi : x \in \partial_{\alpha\delta_X^0} Q_{J,k'}\}) \\ &\leq \sum_{Q_{J,k'} \subseteq Q_{J+1,k}} c''_\varepsilon(\alpha\delta_X^0)^\eta \mu(Q_{J,k}) \\ &= c''_\varepsilon(\alpha\delta_X^0)^\eta \mu(\bigcup_{Q_{J,k'} \subseteq Q_{J+1,k}} Q_{J,k'}) \\ &\leq c''_\varepsilon(\alpha\delta_X^0)^\eta \mu(Q_{J+1,k}).\end{aligned}$$

We $c_1\delta_X^{J+1}(2d(X))^{-1}\varepsilon$-orthonormalize $\tilde{\Psi}_{1,k}$ to the functions in Φ_0, obtaining $\tilde{\Phi}_{1,k}$, and then we $c_1\delta_X^{J+1}(2d(X))^{-1}\varepsilon$-orthonormalize this set again to itself to obtain $\Phi_{1,k} := \{\phi_{1,k,i}\}_{i\in\mathcal{I}_{1,k}}$ for each $k \in \mathcal{K}_{J+1}$. As every function in both $\tilde{\Psi}_{1,k}$ and $\tilde{\Phi}_{1,k}$ have support in $Q_{J+1,k}$, this is also true for every function in $\Phi_{1,k}$. This proves property (ii). To see that (iii) also holds, observe that $\left\langle \bigcup_{k\in\mathcal{K}_{J+1}} \tilde{\Psi}_{1,k} \right\rangle \subseteq \langle\Phi_1\rangle_\varepsilon$, as the functions in $\tilde{\Psi}_{1,k}$ are $2c_1\delta_X^{J+1}(2d(X))^{-1}\varepsilon$ approximated in Φ_1, and there are $d(X)(c_1\delta_X^{J+1})^{-1}$ such k, hence their union is ε-approximated. Moreover,

$$\left\langle \bigcup_{k\in\mathcal{K}_J} \tilde{\Psi}_{0,k} \cup \bigcup_{k\in\mathcal{K}_{J+1}} \tilde{\Psi}_{1,k} \right\rangle \subseteq \langle\Phi_0\rangle_\varepsilon + \langle\Phi_1\rangle_\varepsilon \subseteq \langle\Phi_0 \cup \Phi_1\rangle_{2\varepsilon}$$

At last,

$$\begin{aligned}\#\mathcal{I}_{1,k} &\leq \dim_{c_1\delta_X^{J+1}(2d(X))^{-1}\varepsilon}(\tilde{\Psi}_{1,k}) \\ &\leq \dim_{c_1\delta_X^J(2d(X))^{-1}\varepsilon}(\tilde{\Psi}_{1,k}) \\ &\leq c''_\varepsilon(\alpha\delta_X^0)^\eta \mu(Q_{J+1,k}).\end{aligned}$$

The first inequality follows from the assumption on the orthogonalization procedure. The second inequality follows from the observation that in general more subspaces will $(\varepsilon+\delta)$-span a given subspace than ε-span the same subspace. As every ε-span is also a $(\varepsilon+\delta)$-span, the "inf" in the definiton of $\dim_{\varepsilon+\delta}$ wil have more elements, hence the dimension could be lower.

We now proceed for the layers $l \geq 1$. This will mostly be a repetition from the case $l = 1$. We consider for $k \in \mathcal{K}_{J+l}$,

$$\begin{aligned}\tilde{\Psi}_{l,k} &:= \Psi|_{Q_{J+l,k}} \setminus \bigcup_{Q_{J+l-1,k'} \subseteq Q_{J+l,k}} \Psi|_{Q_{J+l-1,k'}} \\ &= \{\psi \in \Psi : \operatorname{supp}\psi \subseteq Q_{J+l,k} \text{ but } \operatorname{supp}\psi \not\subseteq Q_{J+l-1,k'} \,\forall\, Q_{J+l-1,k'} \subseteq Q_{J+l,k}\} \\ &\subseteq \bigcup_{Q_{J+l-1,k'} \subseteq Q_{J+l,k}} \{\psi_x \in \Psi : x \in \partial_{\alpha\delta_X^{1-l}} Q_{J+l-1,k}\}.\end{aligned}$$

Once again, the inclusion follows because $\operatorname{supp}\psi_x \in Q_{J+l,k}$ implies $x \in Q_{J+l-1,k'}$ for some $Q_{J+l-1,k'} \subseteq Q_{J+l,k}$. Moreover, $\operatorname{supp}\psi_x \not\subseteq Q_{J+l-1,k'}$, and the $\alpha\delta_X^J$-locality enforce that $d(x, Q_{J+l,k} \backslash Q_{J+l-1,k'}) < \alpha\delta_X^J = \alpha\delta_X^{1-l}\delta_X^{J+l-1}$. From this it follows that $x \in \partial_{\alpha\delta_X^{1-l}}(Q_{J+l-1,k'})$.

By assumption it follows that

$$\dim_{c_1\delta_X^{J+l-1}(2d(X))^{-1}\varepsilon}(\tilde{\Psi}_{l,k}) \subseteq c''_\varepsilon(\alpha\delta_X^{-l+1})^\eta \mu(Q_{J+l,k})$$

We $c_1\delta_X^{J+l}(2d(X))^{-1}\varepsilon$-orthonormalize $\tilde{\Psi}_{l,k}$ to the functions in $\Phi_0, \dots, \Phi_{l-1}$, obtaining $\tilde{\Phi}_{l,k}$, and then we $c_1\delta_X^{J+l}(2d(X))^{-1}\varepsilon$-orthonormalize this set to obtain $\Phi_{l,k} := \{\phi_{l,k,i}\}_{i \in \mathcal{I}_{l,k}}$. Once again by construction, these function have support in $Q_{J+l,k}$. This proves property (ii). To see that (iii) also holds, first observe that $\left\langle \bigcup_{k \in \mathcal{K}_{J+l}} \tilde{\Psi}_{l,k} \right\rangle \subseteq \langle \Phi_l \rangle \, \varepsilon$, since these functions are $c_1\delta_X^{J+l} d(X)^{-1}\varepsilon$-approximated in each dyadic cube of scale $J+l$, and there are maximal $d(X)(c_1\delta_X^{J+l})^{-1}$ such cubes. In fact,

$$\left\langle \bigcup_{l'=0}^{l} \bigcup_{k \in \mathcal{K}_{J+l'}} \tilde{\Psi}_{l',k} \right\rangle \subseteq \bigoplus_{l'=0}^{l} \langle \Phi_{l'} \rangle_\varepsilon \subseteq \left\langle \bigcup_{l'=0}^{l} \Phi_{l'} \right\rangle$$

and secondly,

$$\begin{aligned}\#\mathcal{I}_{l,k} &\leq \dim_{c_1\delta_X^{J+l}(2d(X))^{-1}\varepsilon}(\tilde{\Psi}_{l,k}) \\ &\leq \dim_{c_1\delta_X^{J}(2d(X))^{-1}\varepsilon}(\tilde{\Psi}_{l,k}) \\ &\leq c''_\varepsilon(\alpha\delta_X^{1-l})^\eta \mu(Q_{J+l,k})\end{aligned}$$

We stop if $\#I_{l,k} = 0 \,\forall\, k \in \mathcal{K}_{J+l}$. That means we have actually considered all functions and orthonormalized the whole space. Otherwise, as $X \in \mathcal{Q}$, we eventually have $X = Q_{J+L,k}$ is the only dyadic cube left at scale $J+L$. We simply orthonormalize the functions in Ψ to $\Phi_0, \dots, \Phi_{L-1}$ which have not been considered yet, and that finishes the construction. This happens at most at scale $L \leq j_X$ by assumption.

The last thing left to check is property (i). This follows from property (iii), because

$$\langle \Psi \rangle = \langle \Psi|_X \rangle = \left\langle \bigcup_{k\in\mathcal{K}_{J+L}} \Psi|_{Q_{J+L,k}} \right\rangle \subseteq \left\langle \bigcup_{l'=0}^{L} \Phi_{l'} \right\rangle_{(L+1)\varepsilon}$$

And so $\langle \Psi \rangle \subseteq \langle \Phi \rangle_{(j_X+1)\varepsilon}$. □

We will now apply above orthogonaliation method to functions of the form $T^{\delta_X^j-1}\phi$:

Theorem 6.6 ([2, Thm. 27])**.** *Let (X,d,μ) be a space of homogeneous type with $d(X)<\infty$, and let $\mathcal{Q}$ be a family of dyadic cubes, such that $X = Q_{J_0+j_X,k}$ for some $j_X \geq 0$ and $k \in \mathcal{K}_{J_0+j_X}$. Moreover, let δ_X, η, c_1 and c_2 be as in the definition of dyadic cubes. Furthermore, let $(T_t)_{t\geq 0}$ be a symmetric diffusion semigroup, acting $\delta_X^{J_0}$-locally on X. Fix $\varepsilon > 0$, and let $\Phi_0 := \{\phi_x\}_{x\in\Gamma}$ be a $\alpha_0\delta_X^{J_0}$-local family with center set Γ (see Definition 4.4.2) ($\alpha_0 < c_1$), such that $V_0 \subseteq \langle \Phi_0 \rangle_\varepsilon$. Assume that there exists constants $c'_\varepsilon, c''_\varepsilon$ such that, when we set $\varepsilon' := c_1\delta_X^{J_0}(2\mu(X))^{-1}\varepsilon$, then for all $l \geq 0$ and $k \in \mathcal{K}_{J_0+l}$,*

$$\dim_{\varepsilon'}(\Phi_0|Q_{J_0+l,k}) \leq c'_\varepsilon \mu(Q_{J_0+l,k}) \text{ and}$$
$$\dim_{\varepsilon'}(\{\phi_x : x \in \Gamma \cap \partial_{\alpha_0\delta_X^{-l}} Q_{J_0+l,k}\}) \leq c''_\varepsilon(\alpha_0\delta_X^{-l})^\eta \mu(Q_{J_0+l,k}).$$

Then there exists a sequence of orthonormal bases $\{\Phi_j\}_{j=1,\ldots,j_X}$:

$$\Phi_j := \{\{\{\phi_{j,l,k,i}\}_{i\in\mathcal{I}(j,l,k)}\}_{k\in\mathcal{K}_{J_0+j+l}}\}_{l=0,\ldots,j_X-j}$$

with the following properties:

(i) $\langle V_j \rangle \subseteq \langle \Phi_j \rangle_{(j_X-j+2)\varepsilon}$

(ii) $\operatorname{supp}\phi_{j,l,k,i} \subseteq Q_{J_0+j+l,k}$ *for all* $l = 0,\ldots,j_X - j$, $k \in \mathcal{K}_{J_0+j+l}$, $i \in \mathcal{I}(j,l,k)$.

(iii) $\#\mathcal{I}(j,l,k) \leq c''_\varepsilon(1+\delta_X^{-j}(\alpha_0-1)\min\{\delta_X^{-(l-1)},1\})^\eta \mu(Q_{J_0+j+l,k})$.

Proof. We define for every $j \geq 0$ $\tilde{\Phi}_j := T_{\delta_X^j-1}\Phi_0 = T_{t_j}\Phi_0$. Then using that $\langle V_0 \rangle \subseteq \langle \Phi_0 \rangle_\varepsilon$ we obtain for the basis $\{\lambda^{t_j}\xi_\lambda\}_{\lambda^{t_j}\geq\varepsilon}$ of V_j

$$\begin{aligned}\|P_{\tilde{\Phi}_j}\lambda^{t_j}\xi_\lambda - \lambda^{t_j}\xi_\lambda\| &= \|P_{T_{t_j}\Phi_0}T_{t_j}\xi_\lambda - T_{t_j}\xi_\lambda\| \\ &= \|T_{t_j}(P_{\Phi_0}\xi_\lambda - \xi_\lambda)\| \\ &\leq \|T_{t_j}\|\|P_{\Phi_0}\xi_\lambda - \xi_\lambda\| \\ &\leq \varepsilon\end{aligned}$$

so that $\tilde{\Phi}_j$ ε-spans V_j. We want to apply Proposition 6.5 to obtain an ε-orthonormal basis for V_j. So we need to check the hypotheses for this proposition, for $\alpha_j := 1 + \delta_X^{-j}(\alpha_0 - 1)$ and $J_j := J_0 + j$. To do this, we have the following properties of T:

(P1) We have that $\dim_\varepsilon T(S) \leq \dim_\varepsilon S$ for any finite dimensional subspace $S \subseteq X$. This follows from the fact that T is a contraction: For a fixed $\eta > 0$, we can find an ε-span V of S, denoting $\{s_k\}_k$ as the corresponding basis of S, such that $\dim_V \leq \dim_\varepsilon(S) + \eta$. We have

$$\begin{aligned}\|P_{T(S)}T(V) - T(V)\| &= \|TP_SV - T(V)\| \\ &\leq \|T\|\|P_SV - V\| \leq \varepsilon\end{aligned}$$

And so for V ε-spanning S, we have $T(V)$ as a ε-span of $T(S)$. We get that $\dim_\varepsilon T(S) \leq \dim T(V) \leq \dim V \leq \dim_\varepsilon(S) + \eta$. Now letting $\eta \to 0$ we have established the claim.

(P2) As T acts $\delta_X^{J_0}$-locally, it follows that $T_{\delta_X^j - 1}$ acts $(\delta_X^j - 1)\delta_X^{J_0}$-locally. Now, as Φ_0 is $\alpha_0\delta_X^{J_0}$-local, we obtain that $\tilde{\Phi}_J$ is $\delta_X^{J_j} - \delta_X^{J_0} + \alpha_0\delta_X^{J_0} = \delta_X^{J_j}(1 - \delta_X^{-j}(\alpha_0 - 1)) = \alpha_j\delta_X^{J_j}$-local.

(P3) By definition of acting locally, T preserves the center of each function in Φ_0.

Now we check the hypotheses. We are given a space of homogeneous type, and dyadic cubes, and we know that $X \in \mathcal{Q}$. We have checked that $\tilde{\Phi}_j$ is $\alpha_j\delta_X^{J_j}$-local, what we have not checked yet, is that $\alpha_j \leq c_1$. The following inequalities are equivalent.

$$\begin{aligned}\alpha_0 &\leq c_1 \\ \alpha_0 - 1 &\leq c_1 - 1 \\ \delta_X^{-j}(\alpha_0 - 1) &\leq \delta_X^{-j}(c_1 - 1) \\ 1 + \delta_X^{-j}(\alpha_0 - 1) &\leq 1 + \delta_X^{-j}(c_1 - 1).\end{aligned}$$

Now,

$$1 + \delta_X^{-j}(c_1 - 1) \leq 1 + 1 \cdot (c_1 - 1) = c_1.$$

The last thing to check is that $\tilde{\Phi}_j$ is uniformly finite dimensional:

$$\begin{aligned}\dim_{2\varepsilon'}(\tilde{\Phi}_j|Q_{J_j,k}) &\leq \dim_{\varepsilon'}(T_{\delta_X^j - 1}\Phi_0|Q_{J_j,k}) \\ &\leq \dim_{\varepsilon'}(T_{\delta_X^j - 1}(\Phi_0|Q_{J_j,k}) \\ &\leq \dim_{\varepsilon'}(\Phi_0|Q_{J_j,k}) \\ &\leq c'_\varepsilon\mu(Q_{J_j,k})\end{aligned}$$

The first lines follows because the set of $2\varepsilon'$-spans is bigger than the set of ε'-spans, hence the infimum over the dimension can be lower. The third line comes from property P1, the last one from the assumptions. For the second line, note that for $\phi_x \in \Phi_0$ we have $\operatorname{supp}\phi_x \subseteq \operatorname{supp}T_{t_j}\phi_x \subseteq B(x, \alpha_j\delta_X^{J_j})$, hence $\operatorname{supp}T_{t_j}\phi_x \subseteq Q_{J_j,k} \Rightarrow \operatorname{supp}\phi_x \subseteq Q_{J_j,k}$, and so $T_{\delta_X^j - 1}\Phi_0|_{Q_{J_j,k}} \subseteq T_{\delta_X^j - 1}(\Phi_0|_{Q_{J_j,k}})$. So the dimension of ($\varepsilon$-spans of) this first subspace will be lower than the dimension of the second one.

Now for the second inequality

$$\begin{aligned}
&\dim_{\varepsilon'}(\{T_{\delta_X^j-1}\phi_x : x \in \Gamma \cap \partial_{\alpha_j \delta_X^{-l}} Q_{J_j+l,k}\}) \\
&\qquad = \dim_{\varepsilon'}(T_{\delta_X^j-1}\{\phi_x : x \in \Gamma \cap \partial_{\alpha_j \delta_X^{-l}} Q_{J_j+l,k}\}) \\
&\qquad \leq \dim_{\varepsilon'}(\{\phi_x : x \in \Gamma \cap \partial_{\alpha_j \delta_X^{-l}} Q_{J_j+l,k}\}) \\
&\qquad \leq c''_\varepsilon(\alpha_j \delta_X^{-l})^\eta \mu(Q_{J_j+l,k})
\end{aligned}$$

Now we can use Proposition 6.5, and obtain for each j an orthonormal set Φ_j. Properties (ii) and (iii) from the theorem are satisfied automatically, and for property (i) one need a triangle inequality to get $\langle V_j \rangle \subseteq \left\langle \tilde{\Phi}_j \right\rangle_\varepsilon \subseteq \langle \Phi_j \rangle_{(j_x - j+1)\varepsilon+\varepsilon}$. This concludes the proof. □

Using these Φ_j, we can construct the wavelet basis. We apply the multiscale orthogonalization technique in Proposition 6.5 to the set $\{(P_j - P_{j+1})\phi_{j,l,k,i}\}$, where P_j is the projection onto V_j we constructed in Section 4.2. We obtain an orthonormal basis of wavelets which spans the orthogonal complement W_j of V_{j+1} in V_j. By construction, these wavelets have compact support.

7 Construction using partitions of unity

In this section, we will use the approach of Coulhon, Kerkyacharian and Petrushev [3] to build a wavelet frame using a sampling theorem.

7.1 Assumptions

In this section, we assume that (X, d, μ) is a space of homogeneous type, where in this case, (X, d) is assumed to be a metric space. Moreover, we assume that the reverse doubling condition also holds, such that $\mu(B(x, \lambda r)) \geq C\lambda^d \mu(B(x, r))$ for some $C > 0$. Thirdly, we assume that balls cannot vanish at some place under the measure μ, that is $\inf_{x\in X} \mu(B(x, 1)) \geq c > 0$.

We will assume that the operator L is a self-adjoint non-negative operator on $L^2(X, \mu)$. By [8, Theorem 9.11], it generates an analytic semigroup $(P_z)_{z\in\mathbb{C}_+}$. We assume that each P_z is an integral operator with kernel $p_z(\cdot, \cdot)$, obeying the following conditions:

i) If $x, y \in X$ and $0 < t \leq 1$, then

$$p_t(x, y) \leq C \frac{e^{-\frac{cd^2(x,y)}{t}}}{\sqrt{\mu(B(x, \sqrt{t}))\mu(B(y, \sqrt{t}))}}. \tag{7.1}$$

Additionality, if $z \in \mathbb{C}$ such that $t := \operatorname{Re} z \in (0, 1]$, then

$$|p_z(x, y)| \leq C \frac{e^{-c\mathrm{Re}\,\frac{d^2(x,y)}{z}}}{\sqrt{\mu(B(x, \sqrt{t}))\mu(B(y, \sqrt{t}))}} \tag{7.2}$$

(Gaussian estimate)

ii) There exists $\alpha > 0$ such that if $x, y, y' \in X$, $t > 0$ and $d(y, y') \leq \sqrt{t}$, then

$$|p_t(x, y) - p_t(x, y')| \leq C \left(\frac{d(y, y')}{\sqrt{t}}\right)^\alpha \frac{e^{-\frac{cd^2(x,y)}{t}}}{\sqrt{\mu(B(x, \sqrt{t}))\mu(B(y, \sqrt{t}))}}. \tag{7.3}$$

(Hölder continuity)

iii) For all $z \in \mathbb{C}$ such that $\operatorname{Re} z > 0$ and all $x \in X$,

$$\int_X p_z(x, y)\, \mathrm{d}\mu(y) = 1. \tag{7.4}$$

These assumptions on L will have a few consequences. Before we can get to those, we state some technical details.

7.2 Integral operators and kernels

We start with a proposition that tells us which conditions an operator $T : L^1 \to L^\infty$ must have in order for it to be an integral operator.

Proposition 7.1. *A linear operator $T : L^1(X, \mu) \to L^\infty(X, \mu)$ is bounded if and only if T is an integral operator with kernel $K \in L^\infty(X \times X)$, i.e.*

$$(Tf)(x) = \int_X K(x, y) f(y) \, \mathrm{d}\mu(y)$$

for almost every $x \in X$. If this is the case, $\|T\|_{1\to\infty} = \|K\|_\infty$. Moreover, the boundedness of T can be expressed in the bilinear form

$$| \langle Tf, g\rangle | \leq c\|f\|_1\|g\|_1$$

for all $f, g \in L^1(X, \mu)$.

Proof. Note that the last statement is a direct consequence of Hölder's inequality and Hahn-Banach's extension theorem. First assume that T is an integral operator with kernel $K \in L^\infty(X \times X)$. Then

$$\begin{aligned} \|Tf\|_\infty &= \sup_{x\in X} |(Tf)(x)| \\ &= \sup_{x\in X} \left| \int_X K(x, y) f(y) \, \mathrm{d}\mu(y) \right| \\ &\leq \sup_{x\in X} \int_X |K(x, y)||f(y)| \, \mathrm{d}\mu(y) \\ &\leq \sup_{x\in X} \sup_{y\in Y} |K(x, y)| \cdot \int_X |f(y)| \, \mathrm{d}\mu(y) \\ &= \|K\|_\infty\|f\|_1. \end{aligned}$$

So $\|T\|_{1\to\infty} \leq \|K\|_\infty$.

Conversely, assume that $T : L^1(X, \mu) \to L^\infty(X, \mu)$ is bounded. Then consider the space

$$L^1(X, \mu) \otimes L^1(X, \mu) := \left\{ \sum_{n=1}^{N} f_n g_n : N \in \mathbb{N}, (f_n)_{n=1}^N \subseteq L^1(X, \mu), (g_n)_{n=1}^N \subseteq L^1(X, \mu) \right\}$$

and the operator $S : L^1(X, \mu) \otimes L^1(X, \mu) \to \mathbb{C}$ defined by

$$S\left(\sum_{n=1}^{N} f_n g_n\right) = \sum_{n=1}^{N} \langle Tf_n, g_n\rangle$$

Now note that $L^1(X, \mu) \otimes L^1(X, \mu)$ is dense in $L^1(X \times X, \mu \otimes \mu)$. Moreover, note that S can be extended to a bounded operator $L^1(X \times X, \mu \otimes \mu) \to \mathbb{C}$:

$$|S(f \cdot g)| = | \langle Tf, g\rangle | \leq \|T\|\|f\|_1\|g\|_1$$

where we used the last statement of the proposition.

Now, by duality, there exists a function $K \in L^\infty(X \times X)$ such that

$$S(f \cdot g) = \int_X \int_X K(x,y) g(x) f(y) \, \mathrm{d}\mu(y) \, \mathrm{d}\mu(x)$$

and $\|K\|_\infty \leq \|S\|_{L^1(X\times X, \mu\otimes\mu)} \leq \|T\|_{1\to\infty}$ But $S(f \cdot g)$ was given by $\langle Tf, g\rangle$, which means that for all $g \in L^1(X,\mu)$ we have the identity

$$\int_X (Tf)(x) g(x) \, \mathrm{d}\mu(x) = \int_X \int_X K(x,y) f(y) \, \mathrm{d}\mu(y) g(x) \, \mathrm{d}\mu(x)$$

which tells us that $(Tf)(x) = \int_X K(x,y) f(y) \, \mathrm{d}\mu(y)$ as claimed. Lastly, we combine both directions of the proof to find that $\|T\|_{1\to\infty} = \|K\|_\infty$. □

We need a definition of a function that will dominate most of the kernels in the remainder of this section:

Definition 7.2. *For $\delta, \sigma > 0$, define the function $D_{\delta,\sigma} : X \times X \to \mathbb{R}$ by*

$$D_{\delta,\sigma}(x,y) := \frac{\left(1 + \frac{d(x,y)}{\delta}\right)^{-\sigma}}{\sqrt{\mu(B(x,\delta))\mu(B(y,\delta))}}$$

Later on, we will need some properties of these functions. We will state these properties next.

Lemma 7.3. *a) $D_{\delta,\sigma}(x,y) \leq 2^{d/2} \mu(B(x,\delta))^{-1} \left(1 + \frac{d(x,y)}{\delta}\right)^{d/2-\sigma}$ for each $x, y \in X$, $\delta, \sigma > 0$, by the doubling property of μ.*

b) $D_{\lambda\delta,\sigma}(x,y) \leq (2/\lambda)^d D_{\delta,\sigma}(x,y)$ for $0 < \lambda < 1$.

c) $D_{\lambda\delta,\sigma}(x,y) \leq \lambda^\sigma D_{\delta,\sigma}(x,y)$ for $\lambda > 1$.

d) If $0 < p < \infty$ and $\sigma > d(1/2 + 1/p)$, we have

$$\|D_{\delta,\sigma}(x,\cdot)\|_p \leq c(p) \mu(B(x,\delta))^{1/p-1}$$

where $c(p) := \left(\frac{2^{dp/2}}{2^{-d} - 2^{-(\sigma-\delta/2)p}}\right)^{1/p}$.

e) We have

$$\int_X D_{\delta,\sigma}(x,u) D_{\delta,\sigma}(u,y) \, \mathrm{d}\mu(u) \leq c D_{\delta,\sigma}(x,y)$$

if $\sigma > 2d$, where $c := \frac{2^{\sigma+d+1}}{2^{-d} - 2^{d-\sigma}}$.

Proof. a) Note that $B(x,\delta) \subseteq B(y, d(x,y)+\delta)$ as a very simple consequence of the triangle inequality. Together with a consequence of the doubling condition,

$$\mu(B(x,\lambda r)) \leq (2\lambda)^d \mu(B(x,r))$$

for $\lambda > 1, r > 0$, it follows that

$$\begin{aligned}\mu(B(x,\delta)) &\leq \mu(B(y, d(x,y)+\delta)) \\ &= \mu(B(y, \delta(\frac{d(x,y)}{\delta}+1))) \\ &\leq \left(2\left(\frac{d(x,y)}{\delta}+1\right)\right)^d \mu(B(y,\delta)).\end{aligned}$$

And so

$$\left(\sqrt{\mu(B(y,\delta))}\right)^{-1} \leq 2^{d/2}\left(1+\frac{d(x,y)}{\delta}\right)^{d/2}\left(\sqrt{\mu(B(x,\delta))}\right)^{-1}$$

Putting this together with the definition of $D_{\delta,\sigma}$, we end up with

$$D_{\delta,\sigma}(x,y) \leq \frac{2^{d/2}}{B(x,\delta)}\left(1+\frac{d(x,y)}{\delta}\right)^{d/2-\sigma}$$

b) Note that above mentioned doubling conditions implies for $0 < \lambda < 1$ and $r > 0$ that

$$\mu(B(x,r)) \leq (2/\lambda)^d \mu(B(x,\lambda r))$$

by applying this condition with $1/\lambda$. So for these λ and $\delta, \sigma > 0$ we have

$$\begin{aligned}D_{\lambda\delta,\sigma}(x,y) &= \frac{\left(1+\frac{d(x,y)}{\lambda\delta}\right)^{-\sigma}}{\sqrt{\mu(B(x,\lambda\delta))\mu(B(y,\lambda\delta))}} \\ &\leq \sqrt{(2/\lambda)^d(2/\lambda)^d}\frac{\left(1+\frac{d(x,y)}{\lambda\delta}\right)^{-\sigma}}{\sqrt{\mu(B(x,\delta))\mu(B(y,\delta))}} \\ &\leq \left(\frac{2}{\lambda}\right)^d D_{\delta,\sigma}(x,y).\end{aligned}$$

In the last line, we used that $\lambda < 1$ implies that $1/\lambda > 1$, hence $(1 + d(x,y)/(\lambda\delta))^{-\sigma} \leq (1+d(x,y)/\delta)^{-\sigma}$ because of the negative exponent.

c) For $\lambda > 1$, we get that

$$\begin{aligned}D_{\lambda\delta,\sigma}(x,y) &= \frac{\left(1+\frac{d(x,y)}{\lambda\delta}\right)^{-\sigma}}{\sqrt{\mu(B(x,\lambda\delta))\mu(B(y,\lambda\delta))}} \\ &\leq \frac{\lambda^\sigma\left(\lambda+\frac{d(x,y)}{\delta}\right)^{-\sigma}}{\sqrt{\mu(B(x,\delta))\mu(B(y,\delta))}} \\ &\leq \lambda^\sigma D_{\delta,\sigma}(x,y),\end{aligned}$$

where we used that $\lambda > 1$ implies that $(\lambda + d(x,y)/\delta)^{-\sigma} \leq (1 + d(x,y)/\delta)^{-\sigma}$. Moreover, we used that $B(x,\delta) \subseteq B(x,\lambda\delta)$ in order to estimate the denominator.

d) In order to prove the bound of the L^p-norm of $D_{\delta,\sigma}$, we will need the following result. For $\sigma > d$, we have

$$\int_X (1 + \delta^{-1} d(x,y))^{-\sigma} \, \mathrm{d}\mu(y) \leq (2^{-d} - 2^{-\sigma})^{-1} \mu(B(x,\delta)). \tag{7.5}$$

To proof this equality, split the space up in annuli $\{E_j\}_{j=0}^{\infty}$, where $E_0 = B(x,\delta)$ and $E_j = B(x, 2^j\delta) \backslash B(x, 2^{j-1}\delta)$ for $j \in \mathbb{N}$. Then using the doubling property we get

$$\begin{aligned}
\int_X (1 + \delta^{-1} d(x,y))^{-\sigma} \, \mathrm{d}\mu(y) &= \sum_{j=0}^{\infty} \int_{E_j} (1 + \delta^{-1} d(x,y))^{-\sigma} \, \mathrm{d}\mu(y) \\
&\leq \mu(B(x,\delta)) + \sum_{j=1}^{\infty} \frac{\mu(B(x,2^j\delta) \backslash B(x,2^{j-1}\delta))}{(1 + 2^{j-1})^{\sigma}} \\
&= \mu(B(x,\delta)) + \sum_{j=1}^{\infty} \frac{2^d \mu(B(x,2^{j-1}\delta)) - \mu(B(x,2^{j-1}\delta))}{(1 + 2^{j-1})^{\sigma}} \\
&\leq \mu(B(x,\delta)) + \sum_{j=0}^{\infty} (2^d - 1) \frac{\mu(B(x,2^j\delta))}{(1 + 2^j)^{\sigma}} \\
&\leq \mu(B(x,\delta)) \left(1 + (2^d - 1) \sum_{j=0}^{\infty} \frac{2^{jd}}{(1 + 2^j)^{\sigma}} \right) \\
&\leq \mu(B(x,\delta)) \left(1 + (2^d - 1) \sum_{j=0}^{\infty} 2^{jd - j\sigma} \right) \\
&= \mu(B(x,\delta)) \left(1 + (2^d - 1) \frac{1}{1 - 2^{d-\sigma}} \right) \\
&= \mu(B(x,\delta)) \left(1 + \frac{1 - 2^{-d}}{2^{-d} - 2^{-\sigma}} \right) \\
&= \mu(B(x,\delta)) \frac{1 - 2^{-\sigma}}{2^{-d} - 2^{-\sigma}} \\
&\leq \frac{\mu(B(x,\delta))}{2^{-d} - 2^{-\sigma}}.
\end{aligned}$$

This proves the claim. Now we have

$$\begin{aligned}
\|D_{\delta,\sigma}(x,\cdot)\|_p &= \left(\int_X (D_{\delta,\sigma}(x,y))^p \,\mathrm{d}\mu(y)\right)^{1/p} \\
&\leq \left(\int_X \left(\frac{2^{d/2}}{\mu(B(x,\delta))}\left(1+\frac{d(x,y)}{\delta}\right)^{d/2-\sigma}\right)^p \mathrm{d}\mu(y)\right)^{1/p} \\
&\leq \frac{2^{d/2}}{\mu(B(x,\delta))}\left(\int_X \left(1+\frac{d(x,y)}{\delta}\right)^{dp/2-\sigma p} \mathrm{d}\mu(y)\right)^{1/p} \\
&\leq \frac{2^{d/2}}{\mu(B(x,\delta))}\left(\frac{\mu(B(x,\delta))}{2^{-d}-2^{dp/2-\sigma p}}\right)^{1/p} \\
&= \left(\frac{2^{dp/2}}{2^{-d}-2^{dp/2-\sigma p}}\right)^{1/p} \mu(B(x,\delta))^{1/p-1}.
\end{aligned}$$

using $\sigma > d(\frac{1}{2}+\frac{1}{p})$, which implies that $\sigma p - dp/2 > d$. This validates the use of the claim in the fourth line.

e) Note that the triangle inequality implies that

$$\begin{aligned}
\frac{1+\delta^{-1}d(x,y)}{(1+\delta^{-1}d(x,u))(1+\delta^{-1}d(y,u))} &\leq \frac{1+\delta^{-1}d(y,u)+1+\delta^{-1}d(x,u)}{(1+\delta^{-1}d(x,u))(1+\delta^{-1}d(y,u))} \\
&= \frac{1}{1+\delta^{-1}d(x,u)} + \frac{1}{1+\delta^{-1}d(y,u)}
\end{aligned}$$

Now we use that $(a+b)^\sigma \leq 2^\sigma(a^\sigma + b^\sigma)$ for nonnegative a and b, which is true because if $a \leq b$, then $(a+b)^\sigma \leq (2b)^\sigma = 2^\sigma b^\sigma$, and if $b \leq a$ then $(a+b)^\sigma \leq 2^\sigma a^\sigma$, so either way $(a+b)^\sigma \leq 2^\sigma(a^\sigma + b^\sigma)$. We end up with

$$\frac{(1+\delta^{-1}d(x,y))^\sigma}{(1+\delta^{-1}d(x,u))^\sigma(1+\delta^{-1}d(y,u))^\sigma} \leq \frac{2^\sigma}{(1+\delta^{-1}d(x,u))^\sigma} + \frac{2^\sigma}{(1+\delta^{-1}d(y,u))^\sigma}$$

Moving on, we use the same inequality $\mu(B(x,\delta)) \leq 2^d(1+\delta^{-1}d(x,u))^d\mu(B(u,\delta))$ we used before, and obtain

$$\begin{aligned}
\frac{(1+\delta^{-1}d(x,y))^\sigma}{\mu(B(u,\delta))(1+\delta^{-1}d(x,u))^\sigma(1+\delta^{-1}d(y,u))^\sigma} &\leq \frac{2^{\sigma+d}}{\mu(B(x,\delta))(1+\delta^{-1}d(x,u))^{\sigma-d}} \\
&\quad + \frac{2^{\sigma+d}}{\mu(B(y,\delta))(1+\delta^{-1}d(y,u))^{\sigma-d}}
\end{aligned}$$

We integrate left and right, with respect to u and obtain

$$\begin{aligned}\int_X D_{\delta,\sigma}(x,u)D_{\delta,\sigma}(u,y)\,\mathrm{d}\mu(u) &\leq \frac{2^{\sigma+d}(1+\delta^{-1}d(x,y))^{-\sigma}}{\sqrt{\mu(B(x,\delta))\mu(B(y,\delta))}} \\ &\cdot\left(\int_X \frac{1}{\mu(B(x,\delta))(1+\delta^{-1}d(x,u)^{\sigma-d})}\,\mathrm{d}\mu(u)\right. \\ &\left.+\int_X \frac{1}{\mu(B(y,\delta))(1+\delta^{-1}d(y,u)^{\sigma-d})}\,\mathrm{d}\mu(u)\right) \\ &\leq 2^{\sigma+d}D_{\delta,\sigma}(x,y)\left(\frac{1}{2^{-d}-2^{d-\sigma}}+\frac{1}{2^{-d}-2^{d-\sigma}}\right) \\ &= D_{\delta,\sigma}(x,y)\frac{2^{\sigma+d+1}}{2^{-d}-2^{d-\sigma}}\end{aligned}$$

Where we have used estimate (7.5) for the both integrals on the right hand side.

□

Now that we have $D_{\delta,\sigma}$, we can use it to estimate kernels of operators. The next thing we prove is that certain functions generate integral operators whose kernels are dominated by this $D_{\delta,\sigma}$. The proof is followed almost literally from [3, Thm 3.1]

Theorem 7.4. *Let $g : \mathbb{R} \to \mathbb{C}$ be a measurable function such that its Fourier transform $\hat{g}$ exists and that it satisfies for some $\sigma > 2d$*

$$\|g\|_* := \int_{\mathbb{R}} |\hat{g}(\xi)|(1+|\xi|)^\sigma\,\mathrm{d}\xi < \infty.$$

Then $g(\delta^2 L)e^{-\delta^2 L}$, $0 < \delta \leq 1$ is an integral operator with kernel $g(\delta^2 L)e^{-\delta^2 L}(x,y)$ satisfying

$$|g(\delta^2 L)e^{-\delta^2 L}(x,y)| \leq c_\sigma \|g\|_* D_{\delta,\sigma}(x,y) \qquad \forall\, x,y \in X$$

and

$$|g(\delta^2 L)e^{-\delta^2 L}(x,y) - g(\delta^2 L)e^{-\delta^2 L}(x,y')| \leq c_\sigma \|g\|_* \left(\frac{\rho(y,y')}{\delta}\right)^\alpha D_{\delta,\sigma}(x,y)$$

for all $x,y,y' \in X$ such that $\rho(y,y') \leq \delta$. $\alpha > 0$ is the constant from Assumption (7.3) on p_t. c_σ is a constant depending on the structural constants in Assumptions (7.1, 7.3) on p_t and depending on $\sigma > 0$. Moreover, it holds that

$$\int_X g(\delta^2 L)e^{-\delta^2 L}(x,y)\,\mathrm{d}\mu(y) = g(0) \qquad \forall\, x \in X.$$

Proof. Let $(E_\lambda)_{\lambda\geq 0}$ be the spectral resolution associated to L, as we have defined in Section 2.1. Then the operator $g(\delta^2 L)e^{-\delta^2 L}$ is defined as

$$g(\delta^2 L)e^{-\delta^2 L} = \int_0^\infty g(\delta^2\lambda)e^{-\delta^2\lambda}\,\mathrm{d}E_\lambda.$$

Note that the Fourier inversion formula holds for g, as $\|\hat{g}\|_1 \leq \|g\|_* < \infty$. In particular, this means that $\|g\|_\infty \leq \frac{1}{2\pi}\|\hat{g}\|_1$. Hence

$$\|g(\delta^2 L)e^{-\delta^2 L}\|_{2\to 2} \leq \|\lambda \mapsto g(\delta^2\lambda)e^{-\delta^2\lambda}\|_\infty \leq \frac{1}{2\pi}\|\hat{g}\|_1.$$

This shows that $g(\delta^2 L)e^{-\delta^2 L}$ is a bounded operator on $L^2(X,\mu)$. Hence its domain is all of L^2, and we can write for $\phi, \psi \in L^1(X,\mu) \cap L^2(X,\mu)$

$$\begin{aligned}
\left\langle g(\delta^2 L)e^{-\delta^2 L}\phi, \psi\right\rangle &= \int_0^\infty g(\delta^2\lambda)e^{-\delta^2\lambda}\,\mathrm{d}\left\langle E_\lambda \phi, \psi\right\rangle \\
&= \int_0^\infty \mathcal{F}^{-1}(\hat{g})(\delta^2\lambda)e^{-\delta^2\lambda}\,\mathrm{d}\left\langle E_\lambda \phi, \psi\right\rangle \\
&= \frac{1}{2\pi}\int_0^\infty \int_{\mathbb{R}} \hat{g}(\xi)e^{i\xi\delta^2\lambda}e^{-\delta^2\lambda}\,\mathrm{d}\xi\,\mathrm{d}\left\langle E_\lambda \phi, \psi\right\rangle \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \int_0^\infty \hat{g}(\xi)e^{-\delta^2\lambda(1-i\xi)}\,\mathrm{d}\left\langle E_\lambda \phi, \psi\right\rangle\,\mathrm{d}\xi \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \hat{g}(\xi)\left\langle e^{-\delta^2(1-i\xi)L}\phi, \psi\right\rangle\,\mathrm{d}\xi \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \hat{g}(\xi)\int_X\int_X p_{\delta^2(1-i\xi)}(x,y)\phi(x)\overline{\psi}(y)\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)\,\mathrm{d}\xi \\
&= \int_X\int_X \frac{1}{2\pi}\int_{\mathbb{R}} \hat{g}(\xi)p_{\delta^2(1-i\xi)}(x,y)\,\mathrm{d}\xi\phi(x)\overline{\psi}(y)\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y).
\end{aligned}$$

In the sixth line, we used that e^{-zL} is a kernel operator with kernel p_z. Moreover, we have used Fubini twice in the above calculations: in the fourth line and in the seventh line. We will now justify the use of Fubini. Let $h \in \mathcal{L}^2(X,\mu)$. Then

$$\begin{aligned}
\int_{\mathbb{R}}\int_0^\infty |\hat{g}(\xi)||e^{-\delta^2\lambda(1-i\xi)}|\,\mathrm{d}\|E_\lambda h\|_2^2\,\mathrm{d}\xi &= \int_{\mathbb{R}} |\hat{g}(\xi)|\,\mathrm{d}\xi \int_0^\infty e^{-\delta^2\lambda}\,\mathrm{d}\|E_\lambda h\|_2^2 \\
&\leq \|\hat{g}\|_1\|h\|_2^2 < \infty
\end{aligned}$$

and, using that $\|p_z\|_\infty \leq c$, uniformly in $\operatorname{Im} z$, for some $c > 0$,

$$\begin{aligned}
&\int_{\mathbb{R}}\int_X\int_X |\hat{g}(\xi)||p_{\delta^2(1-i\xi)}(x,y)||\phi(x)||\overline{\psi}(y)|\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)\,\mathrm{d}\xi \\
\leq c\int_{\mathbb{R}} |\hat{g}(\xi)|\,\mathrm{d}\xi &\int_X |\phi(x)|\,\mathrm{d}\mu(x)\int_X |\overline{\psi}(y)|\,\mathrm{d}\mu(y) = c\|\hat{g}\|_1\|\phi\|_1\|\psi\|_1 < \infty
\end{aligned}$$

Combining the above calculations with this estimation, we have that

$$|\left\langle g(\delta^2 L)e^{-\delta^2 L}\phi, \psi\right\rangle| \lesssim \|\hat{g}\|\|\phi\|_1\|\psi\|_1.$$

This enables us to use Proposition 7.1 to obtain that $g(\delta^2 L)e^{-\delta^2 L}$ is a kernel operator. Moreover, from the calculations, we have that the kernel is given by

$$g(\delta^2 L)e^{-\delta^2 L}(x,y) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{g}(\xi)p_{\delta^2(1-i\xi)}(x,y)\,\mathrm{d}\xi$$

The next step is to show the estimations for this kernel. By assumption (7.2), we have that

$$|g(\delta^2 L)e^{-\delta^2 L}(x,y)| \leq \frac{c'}{\sqrt{\mu(B(x,\delta))\mu(B(y,\delta))}} \int_{\mathbb{R}} |\hat{g}(u)| e^{-\frac{cd^2(x,y)}{\delta^2(1+u^2)}} \, \mathrm{d}u \tag{7.6}$$

In order to estimate the kernel, we split up in two cases: $d(x,y) < \delta$ and $d(x,y) \geq \delta$. We start with the case $d(x,y) < \delta$:

For $\sigma > 0$, it follows that $2^{-\sigma} < (1 + \frac{d(x,y)}{\delta})^{-\sigma}$. Hence

$$\begin{aligned}\int_0^\infty |\hat{g}(u)| e^{-\frac{cd^2(x,y)}{\delta^2(1+u^2)}} \, \mathrm{d}u &\leq \int_0^\infty |\hat{g}(u)| \, \mathrm{d}u \\ &\leq c \int_0^\infty |\hat{g}(u)|(1+|u|)^\sigma \, \mathrm{d}u (1 + \frac{d(x,y)}{\delta})^{-\sigma}\end{aligned}$$

Putting everything together, we have that

$$|g(\delta^2 L)e^{-\delta^2 L}(x,y)| \leq c D_{\delta,\sigma}(x,y) \|g\|_*$$

In the case $d(x,y) \geq \delta$, we use an elementary estimation: by differentiation, it follows easily that $\sup_{x \geq 0} x^\beta e^{-x} = (\beta/e)^\beta$ for $\beta > 0$. Hence $e^{-x} \leq (\beta/e)^\beta x^{-\beta}$ for $x \geq 0$. We will use this estimation for $\beta = \sigma/2$, together with the estimation $2\frac{d^2(x,y)}{\delta^2} \geq (1 + \frac{d(x,y)^2}{\delta^2})$ to prove the claim:

$$\begin{aligned}\exp\left\{-\frac{cd^2(x,y)}{\delta^2(1+u^2)}\right\} &\leq \exp\left\{-\left(1+\frac{d^2(x,y)}{\delta^2}\right)\frac{c}{2(1+u^2)}\right\} \\ &\lesssim \left(1+\frac{d^2(x,y)}{\delta^2}\right)^{-\sigma/2}(1+u^2)^{\sigma/2} \\ &\lesssim \left(1+\frac{d(x,y)}{\delta}\right)^{-\sigma}(1+|u|)^\sigma\end{aligned}$$

In the third line, we used that $a^2 + b^2 \leq (a+b)^2$ for $ab > 0$. Plugging this in in above estimation (7.6), we obtain that

$$|g(\delta^2 L)e^{-\delta^2 L}(x,y)| \lesssim D_{\delta,\sigma}(x,y)\|g\|_*$$

This proves the first estimation.

For the Hölder continuity of the kernel, we will use assumption (7.3), of the heat kernel. To be able to do this, we split $g(\delta^2 L)e^{-\delta^2 L} = g(\delta^2 L)e^{-\delta^2 L/2}e^{-\delta^2 L/2}$. Then by a simple use of Fubini, the kernels are related by

$$g(\delta^2 L)e^{-\delta^2 L}(x,y) = \int_X g(\delta^2 L)e^{-\frac{1}{2}\delta^2 L}(x,u)e^{-\frac{1}{2}\delta^2 L}(u,y) \, \mathrm{d}\mu(u)$$

and this enables us to use above estimations to estimate the first part. We write $g_2 = g(2\cdot)$, and use the first part on this dilated function. Moreover, we use that $e^{-c\frac{d^2(x,y)}{\delta^2}} \le c_\sigma(1+\frac{d(x,y)}{\delta})^{-\sigma}$, which follows from the estimations in the two cases above. For $d(y,y') < \delta$, we have that

$$\begin{aligned}
&|g(\delta^2 L)e^{-\delta^2 L}(x,y) - g(\delta^2 L)e^{-\delta^2 L}(x,y')| \\
&\le \int_X |g(\delta^2 L)e^{-\frac{1}{2}\delta^2 L}(x,u)||p_{\delta^2/2}(u,y) - p_{\delta^2/2}(u,y')| \,\mathrm{d}\mu(u) \\
&= \int_X |g_2(\left(\frac{\delta}{\sqrt{2}}\right)^2 L)e^{-\left(\frac{\delta}{\sqrt{2}}\right)^2 L}(x,u)||p_{\delta^2/2}(u,y) - p_{\delta^2/2}(u,y')| \,\mathrm{d}\mu(u) \\
&\le c\|g_2\|_* \left(\frac{d(y,y')}{\delta}\right)^\alpha \int_X D_{\delta,\sigma}(x,u)D_{\delta,\sigma}(u,y)\,\mathrm{d}\mu(u) \\
&\lesssim c\|g\|_* \left(\frac{d(y,y')}{\delta}\right)^\alpha D_{\delta,\sigma}(x,y)
\end{aligned}$$

In the last line, we used Lemma 7.3 e), and the fact that $\sigma > 2d$.

The last thing to prove is that $g(\delta^2 L)e^{-\delta^2 L}\mathbb{1} = g(0)\mathbb{1}$. This follows directly from assumption (7.4): For $x \in X$, we have

$$\begin{aligned}
\int_X g(\delta^2 L)e^{-\delta^2 L}(x,y)\mathbb{1}(y)\,\mathrm{d}\mu(y) &= \frac{1}{2\pi}\int_{\mathbb{R}} \hat{g}(u)\int_X p_{\delta^2(1-iu)}(x,y)\,\mathrm{d}\mu(y)\,\mathrm{d}u \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \hat{g}(u)\,\mathrm{d}u = g(0).
\end{aligned}$$

The swap of integral is just a straightforward application of Fubini, we omit the justification of it. □

Note that by Cauchy-Schwarz, and rewriting $(1+|\xi|^2)^{\sigma/2} = (1+|\xi|^2)^{(\sigma+1)/2}(1+|\xi|^2)^{-1/2}$, we can show that $\|g\|_* \le \|g\|_{H^{\sigma+1}}$. Furthermore, we can rewrite the norm $\|g\|_*$ in terms of the derivatives of g. By integration by parts, one can prove that $\xi^k\hat{g}(\xi) = (-i)^k\mathcal{F}(g^{(k)})(\xi)$. Moreover, we have that $|\hat{g}(\xi)| \le \int_{\mathbb{R}}|g(x)||e^{ix\xi}|\,\mathrm{d}x = \|g\|_1$. Hence we get that $|\xi|^k|\hat{g}(\xi)| \le \|g^{(k)}\|_1$. Pick $k \ge \sigma > 2d$, $k \in \mathbb{N}$. Then we have that

$$\begin{aligned}
(1+|\xi|)^{k+2}|\hat{g}(\xi)| &\le 2^{k+1}(|\hat{g}(\xi)| + |\xi|^{k+2}|\hat{g}(\xi)|) \\
&\le 2^{k+1}(\|g\|_1 + \|g^{(k+2)}\|_1)
\end{aligned}$$

which implies that

$$\begin{aligned}
\|g\|_* &= \int_{\mathbb{R}} |\hat{g}(\xi)|(1+|\xi|)^k\,\mathrm{d}\xi \\
&= \int_{\mathbb{R}} |\hat{g}(\xi)|(1+|\xi|)^{k+2}(1+|\xi|)^{-2}\,\mathrm{d}\xi \\
&\le c(\|g\|_1 + \|g^{(k+2)}\|_1).
\end{aligned}$$

When we want to use the theorem, we will check this condition instead of the one given there.

We have to set up yet another theorem, which gives different conditions for functions to generate integral operators with kernels satisfying the same kind of estimates as the ones in last theorem. This one will be used by functions with compact support, which are constant around 0. In particular, we will use it with a partition of unity.

The proof is followed almost literally from [3, Thm 3.4]

Theorem 7.5. *Let $f \in C^{2k+4}(\mathbb{R}^d)$, $k > 2d$ with $\operatorname{supp} f \subseteq [0, R]$ for some $R \geq 1$, and $f^{(2n+1)}(0) = 0$ for $n \in \{0, 1, \dots, k+1\}$. Then $f(\delta\sqrt{L})$, $0 < \delta \leq 1$ is an integral operator with kernel $f(\delta\sqrt{L})(x, y)$ satisfiying*

$$|f(\delta\sqrt{L})(x, y)| \leq c_k D_{\delta,k}(x, y)$$

with $c_k := c_k(f) := \tilde{c_k} R^{2k+d+4}(\|f\|_\infty + \|f^{(2k+4)}\|_\infty + \max_{n \leq 2k+4} |f^{(n)}(0)|)$ where $\tilde{c_k}$ is a constant depending on k, d and the constants in Assumptions (7.1, 7.3) on p_t. Moreover, it satisfies

$$|f(\delta\sqrt{L})(x, y) - f(\delta\sqrt{L})(x, y')| \leq c'_k \left(\frac{d(y, y')}{\delta}\right)^\alpha D_{\delta,k}(x, y)$$

if $d(y, y') < \delta$, where $c'_k := c_k R^\alpha$ and α is the constant from Assumption (7.3) on p_t. Lastly,

$$\int_X f(\delta\sqrt{L})(x, y) \, \mathrm{d}\mu(y) = f(0) \qquad \forall\, x \in M.$$

Note that it follows from [7, Theorem 3.1] that $f(L)$ is a bounded operator on $L^p(X, \mu)$ for $1 < p < \infty$.

Proof. We start with proving the theorem in the case when $R = 1$. The idea of the proof is to apply Theorem 7.4 to the function $\lambda \mapsto f(\sqrt{\lambda})e^\lambda$.

Choose a $\theta \in C^\infty(\mathbb{R})$ such that θ is an even function, with $\operatorname{supp}\theta \subseteq [-1, 1]$, $\theta(\lambda) = 1$ for all $\lambda \in [-1/2, 1/2]$ and $0 \leq \theta \leq 1$. Denote $P_k(\lambda) := \sum_{j=0}^{k+2} \frac{f^{(2j)}(0)}{(2j)!}\lambda^{2j}$ and let

$$\begin{aligned} g_0(\lambda) &:= \theta(\sqrt{\lambda})P_k(\sqrt{\lambda})e^\lambda \\ f_1(\lambda) &:= f(\lambda) - \theta(\lambda)P_k(\lambda) \\ g_1(\lambda) &:= f_1(\sqrt{\lambda})e^\lambda \end{aligned}$$

for $\lambda \geq 0$. We extend g_0 to $\lambda < 0$ by setting $g_0(\lambda) := \theta(\sqrt{|\lambda|})P_k(\sqrt{|\lambda|})e^\lambda$. We have $g_0 \in C^\infty(\mathbb{R})$, $\operatorname{supp} g_0 \subseteq [-1, 1]$ and by Leibniz' differentiation rule we have for its derivative for almost every $\lambda \in [-1,1]$

$$\begin{aligned}
g_0^{(k+2)}(\lambda) &= \sum_{i=0}^{k+2} \frac{(k+2)!}{(k+2-i)!i!} \frac{d^i}{d\lambda^i}\theta(\sqrt{|\lambda|}) \frac{d^{k+2-i}}{d\lambda^{k+2-i}}(P_k(\sqrt{|\lambda|})e^{\lambda}) \\
&= \sum_{i=0}^{k+2}\sum_{j=0}^{k+2-i} \frac{(k+2)!}{(k+2-i)!(i-j)!j!} \left(\frac{d^i}{d\lambda^i}\theta(\sqrt{|\lambda|})\right)\left(\frac{d^j}{d\lambda^j}P_k(\sqrt{|\lambda|})\right)e^{\lambda} \\
&= \sum_{i=0}^{k+2}\sum_{j=0}^{k+2-i} \frac{(k+2)!}{(k+2-i)!(i-j)!j!} \sum_{m=1}^{i}\left(c_m \operatorname{sign}(\lambda)^m |\lambda|^{-i+m/2}\theta^{(m)}(\sqrt{|\lambda|})\right) \\
&\quad \cdot \left(\frac{d^j}{d\lambda^j}\sum_{l=0}^{k+2}\frac{f^{(2l)}(0)}{(2l)!}|\lambda|^l\right)e^{\lambda} \\
&= \sum_{i=0}^{k+2}\sum_{j=0}^{k+2-i} \frac{(k+2)!}{(k+2-i)!(i-j)!j!} \sum_{m=1}^{i}\left(c_m \operatorname{sign}(\lambda)^m |\lambda|^{-i+m/2}\theta^{(m)}(\sqrt{|\lambda|})\right) \\
&\quad \cdot \left(\sum_{l=j}^{k+2}\frac{f^{(2l)}(0)}{(2l)!}\frac{l!}{(l-j)!}|\lambda|^{l-j}\operatorname{sign}(\lambda)^j\right)e^{\lambda}
\end{aligned}$$

We get, after estimating all the constants and using that $\lambda \in [-1,1]$, that

$$\begin{aligned}
|g_0^{(k+2)}(\lambda)| &\le c(k)\sum_{i=0}^{k+2}\sum_{j=0}^{k+2-i}\sum_{m=1}^{i}\sum_{l=j}^{k+2} f^{(2l)}(0)|\lambda|^{l-j-i+m/2}\operatorname{sign}(\lambda)^{m+j}e^{\lambda}\theta^{(m)}(\sqrt{|\lambda|}) \\
&\le c(k)\sum_{i=0}^{k+2}\sum_{j=0}^{k+2-i}\sum_{m=1}^{i}\sum_{l=j}^{k+2} f^{(2l)}(0)e\theta^{(m)}(\sqrt{|\lambda|}) \\
&\le c'(k)\sup_{0\le m\le k+2}\|\theta^{(m)}\|_\infty e \sup_{l\le k+2}|f^{(2l)}(0)|
\end{aligned}$$

Note that $|g_0(\lambda)|$ is also bounded by this constant, perhaps with a different $c'(k)$. As g_0's support is compact, we have $\|g_0\|_1 \le \mathcal{L}([-1,1])\|g_0\|_\infty$, and similar for the $(k+2)$nd derivative of g_0. In total, we finally get

$$\|g_0\|_1 + \|g^{(k+2)}\|_1 \le c(k)\sup_{n\le 2k+4}|f^{(n)}(0)|$$

Hence, by the remark of Theorem 7.4, we have that $g_0(\delta^2 L)e^{-\delta^2 L} = \theta(\delta\sqrt{L})P_k(\delta\sqrt{L})$ is an integral operator satisfying the required inequalities with $R = 1$.

We proceed with $g_1(\lambda) = f_1(\sqrt{|\lambda|})$ for $\lambda \in \mathbb{R}$. Then $f_1(\delta\sqrt{L}) = g_1(\delta^2 L)e^{-\delta^2 L}$. Moreover, because both f and θ have support inside $[-1,1]$, the same holds for both f_1 and g_1. Furthermore, $f_1 \in C^{2k+4}(\mathbb{R}_+)$ as θ and P_k are infinitely many times differentiable, and $f \in C^{2k+4}(\mathbb{R}_+)$. Due to the fact that θ is even, we have that $f_1^{(n)}(0) = 0$ for all $n = 0,\dots,2k+4$ and that

$$\|f_1^{(j)}\|_\infty \le \|f^{(j)}\|_\infty + c\max_{n\le 2k+4}|f^{(n)}(0)| \qquad 0 \le j \le 2k+4.$$

This follows directly from the estimates of g_0 above. We next show that $g_1 \in C^{k+2}(\mathbb{R})$ and estimate the derivatives of g_1, just like what we did above. By Leibniz' differentiation rule, we have for $1 \leq m \leq k+2$ and $\lambda > 0$

$$\begin{aligned} g_1^{(m)}(\lambda) &= \sum_{n=0}^{m} \frac{m!}{n!(m-n)!} e^{\lambda} \frac{d^n}{d\lambda^n}(f_1(\sqrt{\lambda})) \\ &= \sum_{n=0}^{m} \frac{m!}{n!(m-n)!} e^{\lambda} \sum_{j=1}^{n} c_j \lambda^{-n+j/2} f_1^{(j)}(\sqrt{\lambda}). \end{aligned}$$

Here it holds that $|c_j| \leq n!$. Now by Taylor's theorem we get that $|f_1^{(j)}| \leq |\lambda|^{(2m-j)/2} \|f_1^{(2m)}\|_\infty$ and so

$$\left| \left(\frac{d^n}{d\lambda^n} \right) (f_1(\sqrt{|\lambda|})) \right| \leq c|\lambda|^{m-n} \|f_1^{(2m)}\|_\infty \qquad \text{for } 1 \leq n \leq m.$$

And we get the same estimate for $\lambda < 0$. Now note that above calculations and estimations imply that $\lambda \mapsto f_1(\sqrt{|\lambda|}) \in C^{k+2}(\mathbb{R})$. From this it follows that $g_1 \in C^{k+2}(\mathbb{R})$. We also get that

$$|g_1^{(m)}(\lambda)| \leq c \sum_{n=0}^{m} e^{\lambda} |\lambda|^{m-n} \|f_1^{(2m)}\|_\infty \leq c(m+1) \|f_1^{(2m)}\|_\infty$$

using the compact support of g_1. This applied with $m = k+2$ implies that $\|g_1^{(k+2)}\|_1 \leq c(k+3)\|f_1^{(2k+4)}\|_\infty$, and it follows immediately that $\|g_1\|_1 \leq e\|f_1\|_\infty$. We apply Theorem 7.4 to conclude that $f_1(\delta\sqrt{L})$ is an integral operator with a kernel satisfying the required estimations, and the constants c_k and c'_k are of the correct form based on all the estimations we did.

Putting everything together, we have $f(\delta\sqrt{L}) = f_1(\delta\sqrt{L}) + \theta(\delta\sqrt{L})P_k(\delta\sqrt{L})$ is an integral operator with the kernel satisfying the estimations with $R = 1$. Lastly, the identity $\int_X f(\delta\sqrt{L})(x,y)\,\mathrm{d}\mu(y) = f(0)$ follows directly from the same statement in the last theorem.

We will now extend the result for the case $R \neq 1$, by using a dilation argument. Assume that f satisfies the hypotheses of the theorem, and set $h(\lambda) := f(R\lambda)$. Then h satisfies the hypotheses with $R = 1$. We have that

$$\begin{aligned} |f(\delta\sqrt{L})(x,y)| &= |h(\delta R^{-1}\sqrt{L})(x,y)| \\ &\leq c_k(h) D_{\delta/R,k}(x,y) \\ &\leq (2R)^d c_k(h) D_{\delta,k}(x,y) \end{aligned}$$

using the properties of the function $D_{\delta,k}$. Similarly we have

$$\begin{aligned}|f(\delta\sqrt{L})(x,y) - f(\delta\sqrt{L})(x,y')| &= |h(\delta R^{-1}\sqrt{L})(x,y) - h(\delta R^{-1}\sqrt{L})(x,y')| \\ &\leq c_k'(h)\left(\frac{d(y,y')}{\delta/R}\right)^\alpha D_{\delta/R,k}(x,y) \\ &= c_k'(h)R^\alpha\left(\frac{d(y,y')}{\delta}\right)^\alpha (2R)^d D_{\delta,k}(x,y) \\ &\leq (2R)^{d+\alpha}c_k'(h)\left(\frac{d(y,y')}{\delta}\right)^\alpha D_{\delta,k}(x,y)\end{aligned}$$

if $d(y,y') \leq \delta/R$. For the case $\delta/R < d(y,y') \leq \delta$, we have

$$\begin{aligned}|f(\delta\sqrt{L})(x,y) - f(\delta\sqrt{L})(x,y')| &\leq (2R)^d c_k(h)|D_{\delta,k}(x,y) - D_{\delta,k}(x,y')| \\ &\leq (2R)^d c_k'(h) D_{\delta,k}(x,y) \\ &\leq (2R)^d c_k'(h) 2^\alpha D_{\delta,k}(x,y)\left(R\frac{d(y,y')}{\delta}\right)^\alpha\end{aligned}$$

In the second line, we used if $D_{\delta,k}(x,y) \geq D_{\delta,k}(x,y')$ that $D_{\delta,k} \geq 0$, and in case $D_{\delta,k}(x,y) \leq D_{\delta,k}(x,y')$, it follows from the fact that $a^{-k} - b^{-k} \leq (a+b)^{-k} \leq b^{-k}$ if $a \geq b$, together with the doubling property of the measure. The last thing to observe is that

$$\begin{aligned}c_k(h) &= \tilde{c_k}\left(\|f\|_\infty + R^{2k+4}\|f^{(2k+4)}\|_\infty + \max_{n\leq 2k+4} R^n|f^{(n)}(0)|\right) \\ &\leq \tilde{c_k}R^{2k+4}(\|f\|_\infty + \|f^{(2k+4)}\|_\infty + \max_{n\leq 2k+4}|f^{(n)}(0)|)\end{aligned}$$

using the chain rule. So the theorem holds in general. □

7.3 The sampling theorem

In this section, we set up a sampling theorem, in which we estimate L^p-norms of certain functions by their behaviour in countably many points inside the space X. This estimation can be made as precise as necessary, but you would need points which are closer to each other. We start with introducing the function space.

Definition 7.6. *For $1 \leq p \leq \infty$ we define the space*

$$\Sigma_\lambda^p := \{f \in L^p : \theta(\sqrt{L})f = f \ \ \forall \theta \in C_c^\infty(\mathbb{R}_+), \theta \equiv 1 \text{ on } [0,\lambda]\}$$

Similarly we define Σ_K^p, for compact sets $K \subseteq [0,\infty)$, where we require that $\theta \equiv 1$ on K instead. These spaces are called the spectral spaces *of $\sqrt{L}$.*

The idea behind this definition is the spectral projections E_{λ^2} of L. These projections need not be continuous however for $p > 2$, and so we correct it with functions which are smooth outside the interval $[0, \lambda]$. This will make sure that we get a continuous projection again for all $1 \leq p \leq \infty$. By requiring the statement holds for all such functions, we correct for the part $[\lambda, \infty)$.

We proceed with the definition of a *maximal δ-net*:

Definition 7.7. *We say that $\mathcal{X} \subseteq X$ is a δ-*net *on X ($\delta > 0$) if $d(\xi, \eta) \geq \delta$ for all $\eta, \xi \in \mathcal{X}$, and we call $\mathcal{X} \subseteq X$ a* maximal δ-net *on X if $\mathcal{X}$ is a δ-net on X that cannot be enlarged, that is, there is no $x \in X \backslash \mathcal{X}$ such that $d(\xi, x) \geq \delta$ for all $\xi \in \mathcal{X}$.*

These maximal δ-nets have some simple properties, which we will state next.

Proposition 7.8. *Suppose that (X, d, μ) is a space of homogeneous type and let $\delta > 0$.*

(a) A maximal δ-net on X always exists.

(b) If $\mathcal{X}$ is a maximal δ-net on X, then

$$X = \bigcup_{\xi \in \mathcal{X}} B(\xi, \delta)$$

and

$$B(\xi, \delta/2) \cup B(\eta, \delta/2) = \emptyset \qquad \text{if } \eta \neq \xi, \eta, \xi \in \mathcal{X}$$

(c) Let $\mathcal{X}$ ba a maximal δ-net on X. Then $\mathcal{X}$ is at most countable and there exists a disjoint partition $\{A_\xi\}_{\xi \in \mathcal{X}}$ of X consisting of measurable sets such that

$$B(\xi, \delta/2) \subseteq A_\xi \subseteq B(\xi, \delta), \qquad \xi \in \mathcal{X}.$$

The partition $\{A_\xi\}_{\xi \in \mathcal{X}}$ is called a companion disjoint partition *of X.*

Proof. For (a) note that we can partially order all the δ-nets on X by inclusion. We like to use Zorn's lemma to prove that there exists a maximal δ-net, so we have to prove that every totally ordered set of δ-nets $\mathcal{C}$ has an upper bound that is again a δ-net. We denote $I := \bigcup \mathcal{C}$, the union of those δ-nets. It is immediately clear that I is an upper bound for each δ-net in $\mathcal{C}$, so we only need to prove that I is again a δ-net. Let $\xi, \eta \in I$. Then there exists $J, K \in \mathcal{C}$ such that $\xi \in J$ and $\eta \in K$. Now using that $\mathcal{C}$ is totally ordered, we have that either $J \subseteq K$ or $K \subseteq J$. In the first case, we have that $\xi, \eta \in K$, and in the second case that $\xi, \eta \in J$. In both cases, it holds that $d(\xi, \eta) \geq \delta$. This proves that I is a δ-net, and as we noted earlier an upper bound of $\mathcal{C}$. We have now proven the assumptions on Zorn's lemma, and obtain that there exists a maximal δ-net on X.

Part (b) follows immediately from the definition, and the triangle inequality.

We start the proof of part (c) by showing that $\mathcal{X}$ is at most countable. We cut off $\mathcal{X}$ by balls of larger radii, and show that each of those sets is finite. Let $y \in X$. Then consider $\mathcal{X} \cap B(y, n)$ for $n \in \mathbb{N}$, and let ξ be an element of this set. Now for $\delta < n$, it follows from the doubling property and property (b) of this proposition, denoting "#" for the counting measure, that

$$\begin{aligned}
\#(\mathcal{X} \cap B(y,n))\mu(B(y,n)) &= \sum_{\xi \in \mathcal{X} \cap B(y,n)} \mu(B(y,n)) \\
&\leq \sum_{\xi \in \mathcal{X} \cap B(y,n)} 2^d \left(1 + \frac{d(\xi,y)}{n}\right)^d \mu(B(\xi,n)) \\
&\leq \sum_{\xi \in \mathcal{X} \cap B(y,n)} 2^d (1+1)^d (2\frac{n}{\delta/2})^d \mu(B(\xi,\delta/2)) \\
&=: \sum_{\xi \in \mathcal{X} \cap B(y,n)} c\mu(B(\xi,\delta/2)) \\
&\leq c\mu(B(y,2n)) \\
&\leq c2^d \mu(B(y,n))
\end{aligned}$$

By dividing left and right by $\mu(B(y,n))$, we have that $\mathcal{X} \cap B(y,n)$ is finite for each $n \in \mathbb{N}$. Since $\mathcal{X} = \bigcup_{n=1}^{\infty} \mathcal{X} \cap B(y,n)$, it follows that $\mathcal{X}$ is at most countable.

Next we will construct the partition $\{A_\xi\}_{\xi \in \mathcal{X}}$. As we have proved that $\mathcal{X}$ is finite or countable, we can order it as $\mathcal{X} = \{\xi_1, \xi_2, \ldots\}$. Now we can define the sets A_{ξ_i} for $i \in \mathbb{N}$. We define $A_{\xi_1} := B(\xi_1, \delta) \backslash \bigcup_{\eta \in \mathcal{X}, \eta \neq \xi_1} B(\eta, \delta/2)$ and for $i > 1$

$$A_{\xi_i} := B(\xi_i, \delta) \backslash \left(\bigcup_{j \leq i-1} A_{\xi_j} \cup \bigcup_{\eta \in \mathcal{X}, \eta \neq \xi_j} B(\eta, \delta/2) \right).$$

Then by definition $A_{\xi_i} \subseteq B(\xi_i, \delta)$, and by the disjointness of $B(\eta, \delta/2)$, we have that $B(\xi_i, \delta/2) \subseteq X \backslash \bigcup_{\eta \neq \xi_i} B(\eta, \delta/2)$, so in particular $B(\xi_i, \delta/2) \subseteq A_{\xi_i}$. Moreover, by construction, A_{ξ_i} and A_{ξ_j} are disjoint for $j \neq i$, and as $\{B(\xi_i, \delta)\}_{i \in \mathbb{N}}$ is a cover for X, it follows that $\{A_\xi\}_{\xi \in \mathcal{X}}$ is a cover for X as well. □

For the rest of the section, we assume that ϕ is a smooth cutoff function, i.e. $\phi \in C^\infty(\mathbb{R}_+)$, $\operatorname{supp} \phi \subseteq [0,b]$ with $b > 1$, $0 \leq \phi \leq 1$ and $\phi = 1$ on $[0,1]$. By Theorem 7.5, we know that there exists a constant $\alpha > 0$ such that for any $\delta > 0$ and $x, y, x' \in X$ we have

$$|\phi(\delta\sqrt{L})(x,y)| \leq K(\sigma) D_{\delta,\sigma}(x,y) \qquad \text{and}$$
$$|\phi(\delta\sqrt{L})(x,y) - \phi(\delta\sqrt{L})(x',y)| \leq K(\sigma) \left(\frac{d(x,x')}{\delta}\right)^\alpha D_{\delta,\sigma}(x,y)$$

where $d(x,x') \leq \delta$, the Hölder continuity of $\phi(\delta\sqrt{L})(x,y)$. Here, the constant $K(\sigma) > 1$ is independent of x, y, x', δ, but it depends on ϕ, σ and the other parameters. $\sigma > 2d$ describes the smoothness of $\phi(\delta\sqrt{L})(x,y)$

We next present a proposition for Σ_λ^p functions, in which the L^p-distance between those functions on the sets A_ξ and in the point ξ are estimated. This will be needed for the sampling theorem.

The proof is followed almost literally from [3, Prop. 4.1]

Proposition 7.9 (Marcinkiewicz-Zygmund inequality)**.** *Fix $\lambda \geq 1$. Let $\mathcal{X}_\delta$ be a maximal δ-net on X where $\delta := \gamma/\lambda$ with $0 < \gamma \leq 1$. Suppose $\{A_\xi\}_{\xi \in \mathcal{X}_\delta}$ is a disjoint partition of X consisting of measurable sets such that $B(\xi, \delta/2) \subseteq A_\xi \subseteq B(\xi, \delta)$ for each $\xi \in \mathcal{X}_\delta$. Then for every $f \in \Sigma^p_\lambda$ with $1 \leq p < \infty$ it holds that*

$$\sum_{\xi \in \mathcal{X}_\delta} \int_{A_\xi} |f(x) - f(\xi)|^p \, \mathrm{d}\mu(x) \leq (K(\sigma_*)\gamma^\alpha c^\diamond)^p \|f\|_p^p,$$

and for $f \in \Sigma^\infty_\lambda$

$$\sup_{\xi \in \mathcal{X}_\delta} \sup_{x \in A_\xi} |f(x) - f(\xi)| \leq K(\sigma_*)\gamma^\alpha c^\diamond \|f\|_\infty$$

Where $K(\sigma_*)$ is the constant from above, $\sigma_* := 2d + 1$ and $c^\diamond = 2^{2d+1}$.

Proof. Let ϕ be the cutoff function as above. Then we have

$$f = \phi(\lambda^{-1}\sqrt{L})f = \int_X \phi(\lambda^{-1}\sqrt{L})(\cdot, y) f(y) \, \mathrm{d}y$$

for $f \in \Sigma^p_\lambda$, $1 \leq p \leq \infty$. Now we use the Hölder continuity of ϕ as stated above (with $\delta = \lambda^{-1}$). If $f \in \Sigma^\infty_\lambda$, we have,

$$\begin{aligned}
&\sup_{\xi \in \mathcal{X}_\delta} \sup_{x \in A_\xi} |f(x) - f(\xi)| \\
&\qquad = \sup_{\xi \in \mathcal{X}_\delta} \sup_{x \in A_\xi} \left| \int_X (\phi(\lambda^{-1}\sqrt{L})(x, y) - \phi(\lambda^{-1}\sqrt{L})(\xi, y)) f(y) \, \mathrm{d}\mu(y) \right| \\
&\qquad \leq \sup_{\xi \in \mathcal{X}_\delta} \sup_{x \in A_\xi} \int_X K(\sigma_*)(\lambda d(x, \xi))^\alpha D_{\delta,\sigma_*}(x, y) |f(y)| \, \mathrm{d}\mu(y) \\
&\qquad \leq K(\sigma_*)\gamma^\alpha \int_X D_{\delta,\sigma_*}(x, y) |f(y)| \, \mathrm{d}\mu(y) \\
&\qquad \leq K(\sigma_*)\gamma^\alpha c^\diamond \|f\|_\infty,
\end{aligned}$$

where we used Proposition 7.11 b) with $p = q = \infty$, which is stated down below. For $1 \leq p < \infty$, we get using similar calculations

$$\begin{aligned}
&\sum_{\xi \in \mathcal{X}_\delta} \int_{A_\xi} |f(x) - f(\xi)|^p \, \mathrm{d}\mu(x) \\
&\qquad = \sum_{\xi \in \mathcal{X}_\delta} \int_{A_\xi} \left| \int_X (\phi(\lambda^{-1}\sqrt{L})(x, y) - \phi(\lambda^{-1}\sqrt{L})(\xi, y)) f(y) \, \mathrm{d}\mu(y) \right|^p \mathrm{d}\mu(x) \\
&\qquad \leq K(\sigma_*)^p \sum_{\xi \in \mathcal{X}_\delta} \int_{A_\xi} \left(\int_X (\lambda d(x, \xi))^\alpha D_{\delta,\sigma_*}(x, y) |f(y)| \, \mathrm{d}\mu(y) \right)^p \mathrm{d}x \\
&\qquad \leq K(\sigma_*)^p \gamma^{\alpha p} \int_X \left(\int_X D_{\delta,\sigma_*}(x, y) |f(y)| \, \mathrm{d}\mu(y) \right)^p \mathrm{d}\mu(x) \\
&\qquad \leq (K(\sigma_*)\gamma^\alpha c^\diamond)^p \|f\|_p^p
\end{aligned}$$

For the last inequalty we used Proposition 7.11 b) again, this time with $q := p$. □

We are finally ready to state and prove the sampling theorem. The proof is followod almost literally from [3, Thm 4.2].

Theorem 7.10. *Let $0 < \gamma < 1$ such that*

$$K(\sigma_*)\gamma^\alpha c^\diamond \leq \frac{1}{2},$$

where the constants on the left hand side are the same as in last proposition. For a given $\lambda \geq 1$, let $\mathcal{X}_\delta$ be a maximal δ-net on X with $\delta := \frac{\gamma}{\lambda}$. Suppose $\{A_\xi\}_{\xi\in\mathcal{X}_\delta}$ is a companion disjoint partition of X consisting of measurable sets such that $B(\xi, \delta/2) \subseteq A_\xi \subseteq B(\xi, \delta)$, $\xi \in \mathcal{X}_\delta$. Then for any $f \in \Sigma_\lambda^p$ for $1 \leq p < \infty$,

$$\frac{1}{2}\|f\|_p \leq \left(\sum_{\xi\in\mathcal{X}_\delta} \mu(A_\xi)|f(\xi)|^p\right)^{1/p} \leq 2\|f\|_p$$

and for $f \in \Sigma_\lambda^\infty$,

$$\frac{1}{2}\|f\|_\infty \leq \sup_{\xi\in\mathcal{X}_\delta} |f(\xi)| \leq \|f\|_\infty.$$

Furthermore, if we fix $0 < \varepsilon < 1$ and have $0 < \gamma < 1$ such that

$$K(\sigma_*)\gamma^\alpha c^\diamond \leq \frac{\varepsilon}{3}$$

instead, then for $f \in \Sigma_\lambda^p$, $1 \leq p \leq 2$, we have

$$(1-\varepsilon)\|f\|_p^p \leq \sum_{\xi\in\mathcal{X}_\delta} \mu(A_\xi)|f(\xi)|^p \leq (1+\varepsilon)\|f\|_p^p.$$

Proof. We start with some elementary calculus: Note that $(\cdot)^p$ is a convex function for $p \geq 1$, so in particular we have for $0 < t < 1$

$$(ta + (1-t)b)^p \leq ta^p + (1-t)b^p \qquad a, b \geq 0$$

We substitute $c := ta$ and $d := (1-t)b$, and have

$$(c+d)^p \leq t^{1-p}c^p + (1-t)^{1-p}d^p$$

Therefore,

$$(1-t)^{p-1}(c+d)^p \leq \left(\frac{t}{1-t}\right)^{1-p} c^p + d^p$$

By substituting $\delta := \frac{t}{1-t}$ (which means $1 + \delta = \frac{1}{1-t}$, hence $1 - t = \frac{1}{1+\delta}$) for $0 < \delta \leq 1$, and substituting $a := c + d$ and $b := d$ we have

$$\frac{1}{(1+\delta)^{p-1}} a^p \leq \frac{1}{\delta^{p-1}}|a-b|^p + b^p$$

We extend this result for $a, b \in \mathbb{C}$ by

$$\frac{1}{(1+\delta)^{p-1}}|a|^p \leq \frac{1}{\delta^{p-1}}|a-b|^p + |b|^p \tag{7.7}$$

We assume that $K(\sigma_*)\gamma^\alpha c^\diamond < 1/2$. Note that

$$\frac{1}{2^{p-1}}|a|^p \leq |a-b|^p + |b|^p$$

for $1 \leq p < \infty$ by filling in $\delta = 1$ in estimate (7.7). We use this with $a = f(x)$, $b = f(\xi)$ and we integrate the inequality over A_ξ. This yields

$$\frac{1}{2^{p-1}} \int_{A_\xi} |f(x)|^p \, \mathrm{d}\mu(x) \leq \int_{A_\xi} |f(x) - f(\xi)|^p \, \mathrm{d}\mu(x) + \mu(A_\xi)|f(\xi)|^p$$

and

$$\frac{1}{2^{p-1}} \mu(A_\xi)|f(\xi)|^p \leq \int_{A_\xi} |f(x) - f(\xi)|^p \, \mathrm{d}\mu(x) + \int_{A_\xi} |f(x)|^p \, \mathrm{d}\mu(x).$$

We sum up over $\xi \in \mathcal{X}_\delta$ and use the Marcinkiewicz-Zygmund inequality Proposition 7.9 and end up with

$$\begin{aligned}
\frac{1}{2^{p-1}}\|f\|_p^p &\leq \sum_{\xi \in \mathcal{X}_\delta} \int_{A_\xi} |f(x) - f(\xi)|^p \, \mathrm{d}\mu(x) + \sum_{\xi \in \mathcal{X}_\delta} \mu(A_\xi)|f(\xi)|^p \\
&\leq (K(\sigma_*)\gamma^\alpha c^\diamond)^p \|f\|_p^p + \sum_{\xi \in \mathcal{X}_\delta} \mu(A_\xi)|f(\xi)|^p \\
&\leq \frac{1}{2^p}\|f\|_p^p + \sum_{\xi \in \mathcal{X}_\delta} \mu(A_\xi)|f(\xi)|^p,
\end{aligned}$$

and by moving the first term on the right hand side to the left, we have

$$\frac{1}{2^p}\|f\|_p^p \leq \sum_{\xi \in \mathcal{X}_\delta} \mu(A_\xi)|f(\xi)|^p$$

After taking left and right to the $1/p$'th power, we have proven the left hand inequality for $1 \leq p < \infty$. For the right hand inequality, we start with the other estimate:

$$\begin{aligned}
\frac{1}{2^{p-1}} \sum_{\xi \in \mathcal{X}_\delta} \mu(A_\xi)|f(\xi)|^p &\leq \sum_{\xi \in \mathcal{X}_\delta} \int_{A_\xi} |f(x) - f(\xi)|^p \, \mathrm{d}\mu(x) + \|f\|_p^p \\
&\leq (K(\sigma_*)\gamma^\alpha c^\diamond)^p \|f\|_p^p + \|f\|_p^p \\
&\leq (\frac{1}{2^p} + 1)\|f\|_p^p
\end{aligned}$$

This implies that

$$\sum_{\xi\in\mathcal{X}_\delta}\mu(A_\xi)|f(\xi)|^p \le (\frac{1}{2}+2^{p-1})\|f\|_p^p$$
$$\le 2^p\|f\|_p^p$$

Again, taking left and right to the $1/p$'th power proves the right hand inequality for $1 \le p < \infty$.

The next thing to prove is the case $p = \infty$. By the triangle inequality, we have

$$|a| \le |a-b| + |b|$$

which is just the same inequality we used before in case $p = 1$ and $\delta = 1$. Once again, we fill in $a = f(x)$ and $b = f(\xi)$ and take the supremum over $x \in A_\xi$, and obtain

$$\sup_{x\in A_\xi}|f(x)| \le \sup_{x\in A_\xi}|f(x)-f(\xi)| + |f(\xi)|.$$

After taking the supremum over $\xi \in \mathcal{X}_\delta$, and apply the Marcinkiewicz-Zygmund inequality yet again, we end up with

$$\|f\|_\infty \le \sup_{\xi\in\mathcal{X}_\delta}\sup_{x\in A_\xi}|f(x)-f(\xi)| + \sup_{\xi\in\mathcal{X}_\delta}|f(\xi)|$$
$$\le K(\sigma_*)\gamma^\alpha c^\diamond\|f\|_\infty + \sup_{\xi\in\mathcal{X}_\delta}|f(\xi)|$$
$$\le \frac{1}{2}\|f\|_\infty + \sup_{\xi\in\mathcal{X}_\delta}|f(\xi)|$$

and move the first term on the right hand side to the left:

$$\frac{1}{2}\|f\|_\infty \le \sup_{\xi\in\mathcal{X}_\delta}|f(\xi)|.$$

The other estimate is proved directly: It is easily seen that

$$\sup_{\xi\in\mathcal{X}_\delta}|f(\xi)| = \sup\{|f(\xi)| : \xi\in\mathcal{X}_\delta\subseteq X\} \le \sup\{|f(x)| : x\in X\} = \|f\|_\infty$$

For the last estimate, we need a little more elementary calculus. If we restrict $p \le 2$ in (7.7), we claim that

$$(1-\delta)|a|^p \le \frac{1}{\delta^{p-1}}|a-b|^p + |b|^p$$

as we have for $\delta \to 0$ that $(1-\delta) = \frac{1}{(1+\delta)^{p-1}} = 1$, and for $\delta > 0$ we have

$$\frac{d}{d\delta}\frac{1}{(1+\delta)^{p-1}} = \frac{d}{d\delta}(1+\delta)^{1-p}$$
$$= (1-p)(1+\delta)^{-p}$$
$$= \frac{1-p}{(1+\delta)^p}$$

The derivative is bounded by $\frac{1-p}{2^p} < \frac{1-p}{(1+\delta)^p} < 1-p$ for $0 < \delta < 1$ and for $1 \leq p \leq 2$. We can make these bounds uniform in p:

$$-\frac{1}{4} < \frac{1-p}{(1+\delta)^p} < 0.$$

This shows that the function $\frac{1}{(1+\delta)^{p-1}}$ decays slower than the function $1-\delta$ for all $1 \leq p \leq 2$. As they are equal for $\delta = 0$, we have that $1 - \delta < \frac{1}{(1+\delta)^{p-1}}$ for $0 < \delta < 1$. This proves the claim.

We now substitute $\delta := \varepsilon/3$ and fill in $a = f(x)$, $b = f(\xi)$ ($\xi \in \mathcal{X}_\delta$) and integrate x over A_ξ. Then we have

$$(1-\varepsilon/3)\int_{A_\xi} |f(x)|^p \,\mathrm{d}\mu(x) \leq \frac{1}{(\varepsilon/3)^{p-1}}\int_{A_\xi} |f(x)-f(\xi)|^p \,\mathrm{d}\mu(x) + \mu(A_\xi)|f(\xi)|^p,$$

and by exchanging a and b

$$(1-\varepsilon/3)\mu(A_\xi)|f(\xi)|^p \leq \frac{1}{(\varepsilon/3)^{p-1}}\int_{A_\xi} |f(x)-f(\xi)|^p \,\mathrm{d}\mu(x) + \int_{A_\xi} |f(x)| \,\mathrm{d}\mu(x).$$

We sum these estimates up over $\xi \in \mathcal{X}_\delta$, and use the Marcinkiewicz-Zygmund inequality Proposition 7.9:

$$\begin{aligned}
(1-\varepsilon/3)\|f\|_p &\leq \frac{1}{(\varepsilon/3)^{p-1}}\sum_{\xi\in\mathcal{X}_\delta}\int_{A_\xi} |f(x)-f(\xi)|^p \,\mathrm{d}\mu(x) + \sum_{\xi\in\mathcal{X}_\delta}\mu(A_\xi)|f(\xi)|^p \\
&\leq \frac{1}{(\varepsilon/3)^{p-1}}(K(\sigma_*)\gamma^\alpha c^\diamond)^p\|f\|_p^p + \sum_{\xi\in\mathcal{X}_\delta}\mu(A_\xi)|f(\xi)|^p \\
&\leq \frac{\varepsilon}{3}\|f\|_p^p + \sum_{\xi\in\mathcal{X}_\delta}\mu(A_\xi)|f(\xi)|^p
\end{aligned}$$

And so

$$(1-\varepsilon)\|f\|_p^p \leq (1-\frac{2}{3}\varepsilon)\|f\|_p^p \leq \sum_{\xi\in\mathcal{X}_\delta}\mu(A_\xi)|f(\xi)|^p.$$

Moreover,

$$\begin{aligned}
(1-\varepsilon/3)\sum_{\xi\in\mathcal{X}_\delta}\mu(A_\xi)|f(\xi)|^p &\leq \frac{1}{(\varepsilon/3)^{p-1}}\sum_{\xi\in\mathcal{X}_\delta}\int_{A_\xi} |f(x)-f(\xi)|^p \,\mathrm{d}\mu(x) + \|f\|_p \\
&\leq \frac{1}{(\varepsilon/3)^{p-1}}(K(\sigma_*)\gamma^\alpha c^\diamond)^p\|f\|_p^p + \|f\|_p^p \\
&\leq (1+\varepsilon/3)\|f\|_p^p.
\end{aligned}$$

And so

$$\begin{aligned}\sum_{\xi\in\mathcal{X}_\delta}\mu(A_\xi)|f(\xi)|^p &\le \frac{1+\varepsilon/3}{1-\varepsilon/3}\|f\|_p\\ &= (1+\frac{2\varepsilon/3}{1-\varepsilon/3})\|f\|_p\\ &\le (1+\frac{2\varepsilon/3}{1-1/3})\|f\|_p\\ &= (1+\frac{3}{2}\frac{2}{3}\varepsilon)\|f\|_p\\ &= (1+\varepsilon)\|f\|_p.\end{aligned}$$

This proves the theorem. □

7.4 Construction of the wavelets

Now that we have all the technical results, we can state the properties of the kernels p_t. Moreover, we will use a partition of unity to create a wavelet basis.

We start with with some L^p estimates for integral kernels, in particular we will give a bound for integral operators $\|H\|_{p\to q}$ with $1\le p\le q\le\infty$. In order to do so, we need the following version of Schur's inequality:

Proposition 7.11. *a) Suppose $\frac{1}{p}-\frac{1}{q}=1-\frac{1}{r}$ with $1\le p,q,r\le\infty$, and let $H(x,y)$ be a measurable kernel verifying*

$$\|H(\cdot,y)\|_r\le K \qquad \text{and} \qquad \|H(x,\cdot)\|_r\le K \qquad \forall\, x,y\in X$$

Then for the operator $(Hf)(x)=\int_X H(x,y)f(y)\,\mathrm{d}\mu(y)$ we have that $\|Hf\|_q\le K\|f\|_p$ for $f\in L^p$.

b) Let H be an integral operator with kernel $H(x,y)$ such that $|H(x,y)|\le c'D_{\delta,\sigma}(x,y)$ for some $0<\delta\le 1$ and $\sigma\ge 2d+1$. If $1\le p\le q\le\infty$ then

$$\|Hf\|_q\le c\delta^{d(\frac{1}{p}-\frac{1}{q})}\|f\|_p \qquad f\in L^p(X,\mu)$$

with $c=c'(k2^{-d})^{d(1/q-1/p)}2^{2d+1}$, where $k>0$ is chosen such that $\inf_{x\in X}\mu(B(x,1))\ge k$.

Proof. a) We will use Hahn-Banach to express the $L^q(X,\mu)$ norm in terms of functions in the dual space $L^{q'}(X,\mu)$:

$$\|Hf\|_q=\sup_{\|g\|_{q'}\le 1}|\langle Hf,g\rangle|$$

If we denote the Hölder conjugates of p and r by p' and r' respectively, we have

$$\frac{1}{p'}+\frac{1}{q}+\frac{1}{r'}=1-\frac{1}{p}+\frac{1}{q}+1-\frac{1}{r}=1-(1-\frac{1}{r})+1-\frac{1}{r}=1$$

This implies that $\frac{1}{p} = \frac{1}{q} + \frac{1}{r'}$, $\frac{1}{q'} = \frac{1}{p'} + \frac{1}{r'}$ and $\frac{1}{r} = \frac{1}{p'} + \frac{1}{q}$.
We will split up f, g, H, then use Hölder's inequality for three functions, and in the end we have the L^q-norm of H:

$$\begin{aligned}
|\langle Hf, g\rangle| &= \left|\int_X \int_X H(x,y)f(y)\,\mathrm{d}\mu(y)g(x)\,\mathrm{d}\mu(x)\right| \\
&\leq \int_X \int_X |f(y)||g(x)||H(x,y)|\,\mathrm{d}\mu(y)\,\mathrm{d}\mu(x) \\
&= \int_X \int_X (|g(x)|^{q'}|H(x,y)|^r)^{1/p'}(|f(y)|^p|H(x,y)|^r)^{1/q} \\
&\quad \cdot (|f(y)|^p|g(x)|^{q'})^{1/r'}\,\mathrm{d}\mu(y)\,\mathrm{d}\mu(x) \\
&\leq \int_X \left(\int_X |g(x)|^{q'}|H(x,y)|^r\,\mathrm{d}\mu(y)\right)^{1/p'} \left(\int_X |f(y)|^p|H(x,y)|^r\,\mathrm{d}\mu(y)\right)^{1/q} \\
&\quad \cdot \left(\int_X |f(y)|^p|g(x)|^{q'}\,\mathrm{d}\mu(y)\right)^{1/r'}\,\mathrm{d}\mu(x) \\
&\leq \left(\int_X \int_X |g(x)|^{q'}H(x,y)|^r\,\mathrm{d}\mu(y)\,\mathrm{d}\mu(x)\right)^{1/p'} \\
&\quad \cdot \left(\int_X \int_X |f(y)|^p|H(x,y)|^r\,\mathrm{d}\mu(y)\,\mathrm{d}\mu(x)\right)^{1/q} \\
&\quad \cdot \left(\int_X \int_X |f(y)|^p|g(x)|^{q'}\,\mathrm{d}\mu(y)\,\mathrm{d}\mu(x)\right)^{1/r'} \\
&\leq \left(K^r\|g\|_{q'}^{q'}\right)^{1/p'} \left(\|f\|_p^p K^r\right)^{1/q} \left(\|f\|_p^p\|g\|_{q'}^{q'}\right)^{1/r'} \\
&= K\|f\|_p\|g\|_{q'} < \infty
\end{aligned}$$

We used Fubini in the 6th line, which is always allowed for nonnegative functions except when the result is infinite, which is not the case. Now we have

$$\|Hf\|_q = \sup_{\|g\|_{q'}\leq 1} |\langle Hf, g\rangle| \leq \sup_{\|g\|_{q'}\leq 1} K\|f\|_p\|g\|_{q'} = K\|f\|_p$$

b) Pick $1 \leq r \leq \infty$ such that $1/p - 1/q = 1 - 1/r$. Then by Lemma 7.3 (a) and (d) we have

$$\|H(\cdot, y)\|_r \leq c'c(r)\mu(B(y,\delta))^{1/r-1} \leq c'c(1)(2^{-d}k\delta)^{d(1/r-1)}.$$

Here we used the doubling property in the last inequality, and the fact that $c(p)$ is decreasing. We have a similar estimate for $\|H(x,\cdot)\|_r$, and so we can use part a). This gives the required estimate and proves the theorem. □

Note in particular that $p_t(x,y) \leq c'D_{\delta,\sigma}$, so this proposition proves that e^{-tL} is a bounded operator on $L^p(X,\mu)$ for every $p \in [1,\infty]$.

The next thing to prove is a lower bound on the heat kernel p_t. The proof is followed almost literally from [3, Lemma 3.19]

Proposition 7.12. *Let $x \in X$. Then $p_t(x,x) \geq c'\mu(B(x,\sqrt{t}))$ for some $c' > 0$ and $0 < t \leq 1$. Moreover, when $d(x,y) < \sqrt{t}$, then*

$$p_t(x,y) \geq c'\mu(B(x,\sqrt{t})) - C\left(\frac{d(x,y)}{\sqrt{t}}\right)^{\alpha} \frac{e^{-cd^2(x,y)}}{\sqrt{\mu(B(x,\sqrt{t}))\mu(B(y,\sqrt{t}))}}$$

Proof. The last part is immediate from assumption (7.3) For the first part, we rewrite $e^{-tL} = e^{-t/2L}e^{-t/2L}$ in terms of the kernels. By using Fubini, we get

$$p_t(x,u) = \int_X p_{t/2}(x,y)p_{t/2}(y,u)\,\mathrm{d}\mu(y).$$

By replacing $u = x$, and using that $p_{t/2}(x,y) = \overline{p_{t/2}(y,x)} = p_{t/2}(y,x)$, we have that

$$p_t(x,x) = \int_X p_{t/2}(x,y)^2\,\mathrm{d}\mu(y)$$

We will now estimate the right hand side, using Cauchy-Schwarz. Let $l > 1$, then

$$\begin{aligned}
\int_X p_{t/2}(x,y)^2\,\mathrm{d}\mu(y) &\geq \int_{B(x,2^l\sqrt{t})} p_{t/2}(x,y)^2\,\mathrm{d}\mu(y) \\
&\geq \frac{1}{\mu(B(x,2^l\sqrt{t}))}\left(\int_{B(x,2^l\sqrt{t})} p_{t/2}(x,y)\,\mathrm{d}\mu(y)\right)^2 \\
&\geq \frac{2^{-ld}}{\mu(B(x,\sqrt{t}))}\left(1 - \int_{X\backslash B(x,2^l\sqrt{t})} p_{t/2}(x,y)\,\mathrm{d}\mu(y)\right)^2.
\end{aligned}$$

We estimate the integral on the right hand side. We split up the space into annuli:

$$X\backslash B(x,2^l\sqrt{t}) = \bigcup_{n=1}^{\infty} B(x,2^{l+n}\sqrt{t})\backslash B(x,2^{l+n-1}\sqrt{t}) =: \bigcup_{n=1}^{\infty} E_n.$$

We can now estimate each part separately, using that $p_{t/2}(x,y) \leq D_{\sqrt{t},\sigma}(x,y)$ for any $\sigma > 3d/2 + 1$, and Lemma 7.3 a):

$$\begin{aligned}
\int_{E_n} p_{t/2}(x,y)\,\mathrm{d}\mu(y) &\leq c\frac{1}{\mu(B(x,\sqrt{t}))}\int_{E_n}(1 + \frac{d(x,y)}{\sqrt{t}})^{-\sigma+d/2}\,\mathrm{d}\mu(y) \\
&\leq \frac{c}{\mu(B(x,\sqrt{t}))}\int_{B(x,2^{l+n}\sqrt{t})}(1 + 2^{l+n}\sqrt{t}t^{-1/2})^{-\sigma+d/2} \\
&\leq c2^{(l+n)d}\frac{1}{(1+2^{l+n})^{\sigma-d/2}} \\
&\leq \frac{c}{(2^{l+n})^{\sigma-3d/2}}
\end{aligned}$$

By summing over n, we have that

$$\int_{X\setminus B(x,2^l\sqrt{t})} p_{t/2}(x,y)\,\mathrm{d}\mu(y) \lesssim \sum_{n=1}^{\infty} 2^{(l+n)(3d/2-\sigma)} \leq \sum_{n=1}^{\infty} 2^{-(l+n)} = 2^{-l}$$

As the constant in the inequality is independent of l, we can now choose l so large that the integral is bounded by $1/2$. Then $p_t(x,x) \geq \frac{2^{-ld-2}}{\mu(B(x,\sqrt{t}))}$ for this choice of l. This proves the proposition □

In the remainder of the section, we present another way of making a wavelet frame for $L^2(X,\mu)$, by using the sampling theorem above. We start with the definition of a frame.

Definition 7.13. *Let H be a Hilbert space. A system $(h_n)_{n=1}^{\infty} \subseteq H$ is a* frame *for H if and only if there exist constants $c, C > 0$ such that*

$$c\|x\|^2 \leq \sum_{n=1}^{\infty} |\langle x, h_n\rangle|^2 \leq C\|x\|^2 \qquad \forall\, x \in H$$

A frame is a generalization of an orthogonal basis: note that any orthonormal basis spanning a Hilbert space H is a frame: by Pythagoras' theorem we can choose $c = C = 1$.

Just like in the last section, we start the construction with a function $\Phi \in C^{\infty}(\mathbb{R}_+)$, which is 1 on $[0,1]$, $0 \leq \Phi \leq 1$ and $\operatorname{supp}\Phi \subseteq [0,b]$, $b > 1$ as in Lemma 7.16 down below. We set $\Psi(u) := \Phi(u) - \Phi(bu)$. It then holds that $0 \leq \Psi \leq 1$ and $\operatorname{supp}\Psi \subseteq [b^{-1}, b]$. Moreover, we assume that Φ is selected such that $\Psi(u) \geq c > 0$ for $u \in [b^{-3/4}, b^{3/4}]$. We define

$$\Psi_0(u) := \Phi(u), \qquad \Psi_j := \Psi(b^{-j}u)$$

Then $(\Psi_j)_{j=0}^{\infty}$ is a partition of unity, and $\Psi_j \in C^{\infty}(\mathbb{R}_+)$, $\operatorname{supp}\Psi_j \subseteq [b^{j-1}, b^{j+1}]$ for $j \geq 1$, $\operatorname{supp}\Psi_0 \subseteq [0,b]$ and $0 \leq \Psi_j \leq 1$ for all $j \geq 0$.

The following result is a discrete version of the Calderón reproducing formula [1], from this lemma it follows that we can write $f = \sum_{j\geq 0} \Psi_j(\sqrt{L})f$ for $f \in L^p(X,\mu)$, $1 \leq p \leq \infty$:

Lemma 7.14. *Let $\varphi_0, \varphi \in C^{\infty}(\mathbb{R}_+)$, $\operatorname{supp}\varphi_0 \subseteq [0,b]$ and $\operatorname{supp}\varphi \subseteq [b^{-1}, b]$ for some $b > 1$, $\varphi_0(0) = 1$, $\varphi_0^{(2n+1)}(0) = 0$ for $n \geq 0$ and $\varphi_0(\lambda) + \sum_{j\geq 1}\varphi(b^{-j}\lambda) = 1$ for all $\lambda \in \mathbb{R}_+$. Then for any $f \in L^p(X,\mu)$, $1 \leq p \leq \infty$,*

$$f = \varphi_0(\sqrt{L})f + \sum_{j=1}^{\infty} \varphi(b^{-j}\sqrt{L})f$$

in $L^p(X,\mu)$.

Proof. For $p = 2$, we have for $N \in \mathbb{N}$ that

$$\varphi_0(\sqrt{L})f + \sum_{j=1}^{N} \varphi(b^{-j}\sqrt{L})f = \int_0^\infty \left(\varphi_0(\sqrt{\lambda}) + \sum_{j=1}^{N} \varphi(b^{-j}\sqrt{\lambda}) \right) \mathrm{d}E_\lambda f$$
$$\to \int_0^\infty 1 \,\mathrm{d}E_\lambda f = f. \qquad (N \to \infty)$$

Here we used that only two terms of the infinite sum are nonzero for any given $\lambda > 0$. The convergence is in the L^2 sense. For $p \neq 2$, we look at $f \in L^2(X,\mu) \cap L^p(X,\mu)$. For $N > 0$, we have that

$$\left\| \varphi_0(\sqrt{L})f + \sum_{j=1}^{N} \varphi(b^{-j}\sqrt{L})f - f \right\|_p$$
$$= \left\| \int_0^\infty \left(\phi_0(\sqrt{\lambda}) + \sum_{j=1}^{N} \phi(b^{-j}\sqrt{\lambda}) - 1 \right) \mathrm{d}E_\lambda f \right\|_p$$

From the L^2-convergence above, we know that there exists a subsequence $(N_k)_{k=1}^\infty \subseteq \mathbb{N}$ such that $\varphi_0(\sqrt{\lambda}) + \sum_{j=1}^{N_k} \varphi(b^{-j}\sqrt{\lambda})$ converges almost everywhere to 1. Moreover, from the choice of φ_0 and φ, we have for each λ that only two terms of the sum are nonzero, so we have that the sum is uniform bounded by $\|\varphi_0\|_\infty + 2\|\varphi\|_\infty$ for each $\lambda \in \mathbb{R}_+$. Using that $\| \int_0^\infty g \,\mathrm{d}E_\lambda \| = \|g\|_\infty$ for bounded functions g, we can use the dominated convergence theorem along $(N_k)_k$, and we end up with

$$\left\| \varphi_0(\sqrt{L})f + \sum_{j=1}^{\infty} \varphi(b^{-j}\sqrt{L})f - f \right\|_p$$
$$= \left\| \int_0^\infty \left(\phi_0(\sqrt{\lambda}) + \sum_{j=1}^{\infty} \phi(b^{-j}\sqrt{\lambda}) - 1 \right) \mathrm{d}E_\lambda f \right\|_p$$
$$= \left\| \int_0^\infty (1-1) \,\mathrm{d}E_\lambda f \right\|_p = 0$$

For general $f \in L^p(X,\mu)$, we can use a density argument: There exists a sequence $(f_n)_{n=1}^\infty \subseteq L^p(X,\mu) \cap L^2(X,\mu)$ such that $f_n \to f$ in $L^p(X,\mu)$ if $n \to \infty$. Then for a fixed $\varepsilon > 0$, we can find an $M \in \mathbb{N}$ such that $\|f_n - f\| \leq \varepsilon$ for all

$n \geq M$. We have that

$$\begin{aligned}
&\left\| \varphi_0(\sqrt{L})f + \sum_{j=1}^{N} \varphi(b^{-j}\sqrt{L})f - f \right\|_p \\
&\quad \leq \left\| \left(\varphi_0(\sqrt{L}) + \sum_{j=1}^{N} \varphi(b^{-j}\sqrt{L}) \right) (f - f_n) \right\|_p \\
&\qquad + \left\| \varphi_0(\sqrt{L})f_n + \sum_{j=1}^{N} \varphi(b^{-j}\sqrt{L})f_n - f_n \right\|_p + \|f - f_n\|_p
\end{aligned}$$

$$\begin{aligned}
&\leq \left(\left\| \varphi_0(\sqrt{L}) + \sum_{j=1}^{N} \varphi(b^{-j}\sqrt{L}) \right\|_{p \to p} + 1 \right) \|f - f_n\|_p \\
&\quad + \left\| \varphi_0(\sqrt{L})f_n + \sum_{j=1}^{N} \varphi(b^{-j}\sqrt{L})f_n - f_n \right\|_p \\
&< K\varepsilon
\end{aligned}$$

if we choose N large enough. Here K depends on the L^p norm of the finite sum of the partition of unity, which is finite because of Proposition 7.11. This proves the lemma. □

So now we know that for $f \in L^p(X, \mu)$, $f = \sum_{j=1}^{\infty} \Psi_j(\sqrt{L})f$ in the L^p sense.

The next claim we make is that $\frac{1}{2} \leq \sum_{j=1}^{\infty} \Psi_j^2(u) \leq 1$. Note first that for $u \in (b^m, b^{m+1})$ only two terms in the sum are nonzero: We have $1 = \Psi_m(u) + \Psi_{m+1}(u)$. So we have

$$\sum_{j=1}^{\infty} \Psi_j^2(u) = \Psi_m^2(u) + \Psi_{m+1}^2(u) = \Psi_m^2(u) + (1 - \Psi_m(u))^2$$

We use this to prove the bounds:

$$\begin{aligned}
\sum_{j=1}^{\infty} \Psi_j^2(u) &= \Psi_m^2(u) + (1 - \Psi_m(u))^2 \\
&= \Psi_m^2(u) + 1 - 2\Psi_m(u) + \Psi_m^2(u) \\
&= 2\Psi_m^2(u) - 2\Psi_m(u) + 1 \\
&= 2(\Psi_m^2(u) - \frac{1}{2})^2 + \frac{1}{2}
\end{aligned}$$

This expression has a minimum at $\Psi_m(u) = 1/2$, with value $\sum_{j=1}^{\infty} \Psi_j^2(u) = \frac{1}{2}$. Moreover, on the interval $[0,1]$, it has its maxima on $\Psi_m(u) \in \{0,1\}$, where in both cases it holds that $\sum_{j=1}^{\infty} \Psi_j^2(u) = 1$. This establishes the bounds.

Now we introduce the spectral resolution $(F_\lambda)_{\lambda \geq 0}$ of $\sqrt{L}$, such that $F_\lambda = E_{\lambda^2}$:

$$\int_0^\infty \lambda \,\mathrm{d}F_\lambda = \sqrt{L} = \int_0^\infty \sqrt{\lambda}\,\mathrm{d}E_\lambda = \int_0^\infty \lambda \,\mathrm{d}E_{\lambda^2}$$

and we note that $\|\Psi_j(\sqrt{L})f\|_2^2 = \left\langle \Psi_j(\sqrt{L})f, \Psi_j(\sqrt{L})f \right\rangle = \left\langle \Psi_j^2(\sqrt{L})f, f \right\rangle$ for $f \in L^2(X,\mu)$, because Ψ_j is self-adjoint. This implies that

$$\sum_{j=0}^{\infty} \|\Psi_j(\sqrt{L})f\|_2^2 = \sum_{j=0}^{\infty} \left\langle \Psi_j^2(\sqrt{L})f, f \right\rangle = \int_0^\infty \sum_{j=0}^{\infty} \Psi_j^2(u)\,\mathrm{d}\left\langle F_u f, f\right\rangle$$

where we could switch sum and integral because of the upper bound we just proved, the fact that $\mathrm{d}\left\langle F_u f, f\right\rangle$ is a finite measure, and the dominated convergence theorem. Using the upper and lower bounds, we get

$$\frac{1}{2}\|f\|_2^2 = \int_0^\infty \frac{1}{2}\,\mathrm{d}\left\langle F_u f, f\right\rangle \leq \sum_{j=0}^{\infty} \|\Psi_j(\sqrt{L})f\|_2^2 \leq \int_0^\infty \mathrm{d}\left\langle F_u f, f\right\rangle = \|f\|_2^2$$

for $f \in L^2(X,\mu)$. We fix $0 < \varepsilon < 1$, and depending on that, we fix $0 < \gamma < 1$ such that

$$K(\sigma_*)\gamma^\alpha c^\diamond = \varepsilon/3,$$

where these constants are the same ones as in the sampling theorem and the Marcinkiewicz-Zygmund inequality. We let $\delta_j := \gamma b^{-j-2}$ for each $j \geq 0$, and let $\mathcal{X}_j \subseteq X$ be maximal δ_j-nets on X, and we suppose $\{A_\xi^j\}_{\xi \in \mathcal{X}_j}$ are companion disjoint partitions of X, with respect to the maximal δ_j-nets $\mathcal{X}_j$. By Theorem 7.10, we have for $f \in \Sigma^2_{b^{j+2}}$

$$(1-\varepsilon)\|f\|_2^2 \leq \sum_{\xi \in \mathcal{X}_j} \mu(A_\xi^j)|f(\xi)|^2 \leq (1+\varepsilon)\|f\|_2^2.$$

By the definition of Ψ_j it follows that $\Psi_j(\sqrt{L})f \in \Sigma^2_{b^{j+1}}$ for $f \in L^2(X,\mu)$: Let $\theta \in C_c^\infty(\mathbb{R}_+)$ such that $\theta(u) = 1$ for all $u \in [0, b^{j+1}]$. Then $\theta(\sqrt{L})\Psi_j(\sqrt{L})f = (\theta\Psi_j)(\sqrt{L})f = \Psi_j(\sqrt{L})$, as $\theta \equiv 1$ on the support of Ψ_j. On further inspection, it follows that $\Psi_j(\sqrt{L})f \in \Sigma^2_{b^{j+2}}$ by the same argument. So we can apply the previous estimation to $\Psi_j(\sqrt{L})$, and we get

$$\begin{aligned}\frac{1}{4}\|f\|_2^2 \leq \frac{1}{2}(1-\varepsilon)\|f\|_2 \leq (1-\varepsilon)\sum_{j=0}^{\infty}\|\Psi_j(\sqrt{L})f\|_2^2 &\leq \sum_{j=0}^{\infty}\sum_{\xi \in \mathcal{X}_j} \mu(A_\xi^j)|\Psi_j(\sqrt{L})f(\xi)|^2 \\ &\leq (1+\varepsilon)\sum_{j=0}^{\infty}\|\Psi_j(\sqrt{L})f\|_2^2 \leq (1+\varepsilon)\|f\|_2^2 \leq 2\|f\|_2^2\end{aligned}$$

Where we have silently assumed that $\varepsilon \le 1/2$. We can now apply Theorem 7.5 to Ψ_j, from which it follows that $\Psi_j(\sqrt{L})$ is a kernel operator, and so we obtain for $f \in L^2(X,\mu)$

$$\begin{aligned}\Psi_j(\sqrt{L})f(\xi) &= \int_X f(u)\Psi_j(\sqrt{L})(\xi,u)\,\mathrm{d}\mu(u)\\ &= \int_X f(u)\overline{\Psi_j(\sqrt{L})(u,\xi)}\,\mathrm{d}\mu(u)\\ &= \left\langle f, \Psi_j(\sqrt{L})(\cdot,\xi)\right\rangle,\end{aligned}$$

where the second line can be obtained by taking the inner product with any $L^2(X,\mu)$ function with support on $\mathcal{X}_j$, and then using the self-adjointness of $\Psi_j(\sqrt{L})$. Now we define the functions which will turn out to be the wavelets: We define the system $\{\{\psi_{j\xi}\}_{\xi\in\mathcal{X}_j}\}_{j\ge 0}$ by

$$\psi_{j\xi}(x) := \sqrt{\mu(A^j_\xi)}\Psi_j(\sqrt{L})(x,\xi)$$

We end this section with the main properties of this system. We start with the boundedness of the operators $\Psi_j(\sqrt{L})$. Using Theorem 7.5 and [7, Theorem 1.2] (where $F := \Psi_j(\sqrt{\cdot})$), we have that $\Psi_j(\sqrt{L})$ is a bounded operator on L^p.

For the estimations in the remainder of the properties, we need a couple of results, to show that some functions which are not (weakly) differentiable can still generate integral operators. The proof of part a) is heavily based on the proof of [3, Prop 2.9]. For part b), the first part of the the proof of [3, Th 3.7] is followed.

Proposition 7.15. *a) Let $U, V : L^2(X,\mu) \to L^2(X,\mu)$ be integral operators such that for some $0 < \delta \le 1$ and $\sigma \ge d+1$ we have*

$$|U(x,y)| \le c_1 D_{\delta,\sigma}(x,y) \qquad |V(x,y)| \le c_2 D_{\delta,\sigma}(x,y).$$

Let $R : L^2(X,\mu) \to L^2(X,\mu)$ be a bounded operator. Then URV is an integral operator whose kernel is bounded by

$$|URV(x,y)| \le \|U(x,\cdot)\|_2\|R\|_{2\to2}\|V(\cdot,y)\|_2 \le c_1c_2c^{\diamond}\frac{\|R\|_{2\to2}}{\sqrt{\mu(B(x,\delta))\mu(B(y,\delta))}}$$

b) Let f be a bounded measurable function with $\operatorname{supp} f \subseteq [0,\tau]$ *for $\tau \ge 1$. Then $f(\sqrt{L})$ is an integral operator with kernel $f(\sqrt{L})(x,y)$ satisfying*

$$|f(\sqrt{L})(x,y)| \le c\frac{\|f\|_\infty}{\sqrt{\mu(B(x,\tau^{-1}))\mu(B(y,\tau^{-1}))}}$$

Proof. For a), note that by Proposition 7.11 b), we have

$$\|URV\|_{1\to\infty} \le \|U\|_{2\to\infty}\|R\|_{2\to2}\|V\|_{1\to2} \le c\delta^{-d}\|R\|_{2\to2}$$

So URV is an integral operator. In order to estimate the kernel, we will find an expression for it. For this, take $f \in L^2(X,\mu)$ such that $\operatorname{supp} f \subseteq B(a,R)$, an arbitrary ball on X.

For the next part, we need Bochner integrals, which are briefly explained in Section 2.1. By the bound on $|V(x,y)|$, the doubling property applied to $B(a,\delta+d(y,a)) \supseteq B(y,\delta)$, and Lemma 7.3 d) we have that

$$\begin{aligned}\int_X \|V(\cdot,y)f(y)\|_2 \,\mathrm{d}\mu(y) &\leq c\int_{B(a,R)} |f(y)|\mu(B(y,\delta))^{-1/2}\,\mathrm{d}\mu(y)\\ &\leq c\|f\|_2\left(\int_{B(a,R)}\mu(B(y,\delta))^{-1}\,\mathrm{d}\mu(y)\right)^{1/2}\\ &\leq \frac{c\|f\|_2}{\sqrt{\mu(B(a,\delta))}}\left(\int_{B(a,R)}(1+\delta^{-1}d(y,a))^d\,\mathrm{d}\mu(y)\right)^{1/2}\\ &< \infty\end{aligned}$$

So the Bochner integral $\int_X V(\cdot,y)f(y)\,\mathrm{d}\mu(y)$ exists. We will now show that it is in fact equal to Vf: For any $g \in L^2(X,\mu)$ we have

$$\begin{aligned}\left\langle \int_X V(\cdot,y)f(y)\,\mathrm{d}\mu(y), g\right\rangle &= \int_X \langle V(\cdot,y)f(y),g\rangle\,\mathrm{d}\mu(y)\\ &= \int_X\left(\int_X \overline{g}(x)V(x,y)\,\mathrm{d}\mu(x)\right)f(y)\,\mathrm{d}\mu(y)\\ &= \int_X \overline{g}(x)\left(\int_X V(x,y)f(y)\,\mathrm{d}\mu(y)\right)\mathrm{d}\mu(x)\\ &= \langle Vf,g\rangle\end{aligned}$$

We were able to use Fubini, because

$$\begin{aligned}\iint |V(x,y)||f(y)||g(x)|\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) &\leq \|g\|_2\left\|\int_X |V(\cdot,y)||f(y)|\,\mathrm{d}\mu(y)\right\|_2\\ &\leq \|g\|_2\int_X \|V(\cdot,y)f(y)\|_2\,\mathrm{d}\mu(y) < \infty\end{aligned}$$

by what we already showed. By varying $g \in L^2(X,\mu)$, we have that

$$Vf = \int_X V(\cdot,y)f(y)\,\mathrm{d}\mu(y).$$

By using the property of Bochner integrals that one may exchange operator and integral, we get

$$RVf = R\int_X V(\cdot,y)f(y)\,\mathrm{d}\mu(y) = \int_X R(V(\cdot,y))f(y)\,\mathrm{d}\mu(y).$$

And so, by applying U to above expression and using some properties from Pettis integrals, we obtain

$$\begin{aligned}(URV)f(x) &= \int_X U(x,u)(RV)f(u)\,\mathrm{d}\mu(u)\\ &= \left\langle (RV)f, \overline{U}(x,\cdot)\right\rangle\\ &= \left\langle \int_X R(V(\cdot,y))f(y)\,\mathrm{d}\mu(y), \overline{U}(x,\cdot)\right\rangle\\ &= \int_X \left\langle R(V(\cdot,y))f(y), \overline{U}(x,\cdot)\right\rangle \mathrm{d}\mu(y)\\ &= \int_X\int_X U(x,u)R(V(\cdot,y))(u)\,\mathrm{d}\mu(u)f(y)\,\mathrm{d}\mu(y).\end{aligned}$$

For arbitrary $f \in L^2(X,\mu)$, we can use a partition of unity to split up $f = \sum \phi_j f$. As those ϕ_j each have compact support, we can show this way that indeed

$$(URV)f(x) = \int_X\int_X U(x,u)R(V(\cdot,y))(u)\,\mathrm{d}\mu(u)f(y)\,\mathrm{d}\mu(y)$$

by using the boundedness of the operator. Now it follows that the kernel $URV(x,y)$ is given by

$$URV(x,y) = \int_X U(x,u)R(V(\cdot,y))(u)\,\mathrm{d}\mu(u)$$

By using the bounds on $|U(x,y)|$ and $|V(x,y)|$, and some properties of the function $D_{\delta,\sigma}$, we have

$$|URV(x,y)| \le \|U(x,\cdot)\|_2\|R(V(\cdot,y))\|_2 \le \frac{c_1c_2(c(2))^2\|R\|_{2\to 2}}{\mu(B(x,\delta))^{1/2}\mu(B(y,\delta))^{1/2}}$$

By noticing that $(c(2))^2 \le c^\diamond$, this part of the lemma is proved.

For part b), let $\theta \in C^\infty$ be a function such that $\theta(x) = 1$ for $0 \le x \le 1$ and $0 \le \theta \le 1$ and such that $\operatorname{supp}\theta \subseteq [0,2]$. Then $\theta(\tau^{-1}x)f(x)\theta(\tau^{-1}x) = f(x)$ for $x \ge 0$. Furthermore, by Theorem 7.5 we have that $\theta(\tau^{-1}\sqrt{L})$ is an integral operator, whose kernel satisfies

$$|\theta(\tau^{-1}\sqrt{L})(x,y)| \le cD_{\tau^{-1},\sigma}(x,y)$$

Using some basic properties of functional calculi, we have that

$$\begin{aligned}\theta(\tau^{-1}\sqrt{L})f(\sqrt{L})\theta(\tau^{-1}\sqrt{L}) &= (\lambda \mapsto \theta(\tau^{-1}\lambda)f(\lambda)\theta(\tau^{-1}\lambda))(\sqrt{L})\\ &= f(\cdot)(\sqrt{L}) = f(\sqrt{L})\end{aligned}$$

So we can apply part a) with $U, V = \theta(\tau^{-1}\sqrt{L})$ to see that $f(\sqrt{L})$ is an integral operator whose kernel satisfies

$$|f(\sqrt{L})(x,y)| \le c\frac{\|f(\sqrt{L})\|_{2\to 2}}{\sqrt{\mu(B(x,\tau^{-1}))\mu(B(y,\tau^{-1}))}} \le c\frac{\|f\|_\infty}{\sqrt{\mu(B(x,\tau^{-1}))\mu(B(y,\tau^{-1}))}}$$

This proves the proposition. □

Now we want to give bounds for kernels of projections which are generated by indicator functions of specific intervals. The existence of these kernels was proved above. The proof is followod almoest literally from [3, Lem 3.19]

Lemma 7.16. *a) There exist constants $c_3, c_4 > 0$ such that for any $\tau \geq 1$*

$$c_3 \mu(B(x, \tau^{-1}))^{-1} \leq \mathbb{1}_{[0,\tau]}(\sqrt{L})(x,x) \leq c_4 \mu(B(x,\tau^{-1}))^{-1}.$$

b) There exists a $b > 1$ such that if $\tau \geq 1$ and $\tau^{-1} \leq \frac{d(X)}{3}$ (see Notation 4.2) if $d(X) < \infty$, then

$$c_5 \mu(B(x, \tau^{-1}))^{-1} \leq \mathbb{1}_{[\tau,b\tau]}(\sqrt{L})(x,x) \leq c_6 \mu(B(x,\tau^{-1}))^{-1}.$$

Here $c_5, c_6 > 0$ only depend on the parameters of the space.

Proof. For part a), we first note that $\mathbb{1}_{[0,\tau]}(\sqrt{L})e^{-tL}$ is a kernel operator by part b) of Proposition 7.15. Then it follows that $(\mathbb{1}_{[0,\tau]}(\sqrt{L})e^{-tL})(x,y) \to p_t(x,y)$ when $\tau \to \infty$. This is because for $f, g \in L^2(X,\mu)$, we have

$$\begin{aligned}
\lim_{\tau\to\infty} \int_X \int_X & (\mathbb{1}_{[0,\tau]}(\sqrt{L})e^{-tL})(x,y) f(y)\overline{g(x)} \,\mathrm{d}\mu(y)\,\mathrm{d}\mu(x) \\
&= \lim_{\tau\to\infty} \left\langle \mathbb{1}_{[0,\tau]}(\sqrt{L})e^{-tL} f, g \right\rangle \\
&= \lim_{\tau\to\infty} \int_0^\infty \mathbb{1}_{[0,\tau]}(\sqrt{\lambda})e^{-t\lambda} \,\mathrm{d}\left\langle E_\lambda f, g\right\rangle \\
&= \int_0^\infty e^{-t\lambda}\,\mathrm{d}\left\langle E_\lambda f, g\right\rangle \\
&= \left\langle e^{-tL} f, g\right\rangle \\
&= \int_X \int_X p_t(x,y) f(y)\overline{g(x)}\,\mathrm{d}\mu(y)\,\mathrm{d}\mu(x)
\end{aligned}$$

In the fourth line, we used the Dominated Convergence Theorem.

In order to prove the bounds on $\mathbb{1}_{[0,\tau]}(\sqrt{L})(x,x)$, we need the same type of bounds for $p_t(x,x)$:

$$c' \mu(B(x,\sqrt{t}))^{-1} \leq p_t(x,x) \leq c\mu(B(x,\sqrt{t}))^{-1}$$

We already proved the lower bound in Proposition 7.12. The upper bound follows immediately from the assumptions on p_t.

We now prove the upper bound of $\mathbb{1}_{[0,\tau]}(\sqrt{L})(x,x)$. As $\mathbb{1}_{[0,\tau]}(u) \leq e^{-\tau^{-2}u^2+1}$, we get that

$$\mathbb{1}_{[0,\tau]}(\sqrt{L})(x,x) \leq e e^{-\tau^{-2}L}(x,x) \leq c\mu(B(x,\tau^{-1}))^{-1}.$$

This proves the upper bound.

For the lower bound, we rewrite

$$\begin{aligned} e^{-tu^2} &= \mathbb{1}_{[0,\tau]}(u)e^{-tu^2} + \sum_{k=0}^{\infty} \mathbb{1}_{(2^k\tau,2^{k+1}\tau]}(u)e^{-tu^2} \\ &\leq \mathbb{1}_{[0,\tau]}(u) + \sum_{k=0}^{\infty} \mathbb{1}_{(2^k\tau,2^{k+1}\tau]}(u)e^{-t2^{2k}\tau^2} \\ &\leq \mathbb{1}_{[0,\tau]}(u) + \sum_{k=0}^{\infty} \mathbb{1}_{[0,2^{k+1}\tau]}(u)e^{-t2^{2k}\tau^2} \end{aligned}$$

From this, the lower bound on $p_t(x,x)$ and the upper bound on $\mathbb{1}_{[0,\tau]}(\sqrt{L})(x,x)$ and the doubling property we have

$$\begin{aligned} c'\mu(B(x,\sqrt{t}))^{-1} &\leq p_t(x,x) \\ &\leq \mathbb{1}_{[0,\tau]}(\sqrt{L})(x,x) + \sum_{k=0}^{\infty} \mathbb{1}_{[0,2^{k+1}\tau]}(\sqrt{L})(x,x)e^{-t2^{2k}\tau^2} \\ &\leq \mathbb{1}_{[0,\tau]}(\sqrt{L})(x,x) + c_4 \sum_{k=0}^{\infty} e^{-t2^{2k}\tau^2}\mu(B(x,2^{-k-1}\tau^{-1}))^{-1} \\ &\leq \mathbb{1}_{[0,\tau]}(\sqrt{L})(x,x) + c_4\mu(B(x,\tau^{-1}))^{-1} \sum_{k=0}^{\infty} e^{-t2^{2k}\tau^2}2^{(k+1)d}. \end{aligned}$$

Now we fix $r \in \mathbb{N}$ and choose t such that $\tau\sqrt{t} = 2^r$. Then we have that

$$\begin{aligned} \frac{c'2^{-rd}}{\mu(B(x,\tau^{-1}))} &\leq \frac{c'}{\mu(B(x,\sqrt{t}))} \\ &\leq \mathbb{1}_{[0,\tau]}(\sqrt{L})(x,x) + \frac{c_4 2^d 2^{-rd}}{\mu(B(x,\tau^{-1}))} \sum_{k=0}^{\infty} e^{-2^{2k}2^{2r}}2^{(k+r)d} \\ &\leq \mathbb{1}_{[0,\tau]}(\sqrt{L})(x,x) + \frac{c_4 2^d 2^{-rd}}{\mu(B(x,\tau^{-1}))} \sum_{k=r}^{\infty} e^{-2^{2k}}2^{kd}. \end{aligned}$$

Hence we have that

$$\frac{2^{-rd}}{\mu(B(x,\tau^{-1}))}\left(c' - c_4 2^d \sum_{k=r}^{\infty} e^{-2^{2k}}2^{kd}\right) \leq \mathbb{1}_{[0,\tau]}(\sqrt{L})(x,x).$$

We can take $r \in \mathbb{N}$ sufficiently large, such that the term on the left hand side is actually positive. Then this implies the lower bound of $\mathbb{1}_{[0,\tau]}(\sqrt{L})(x,x)$ and proves part a).

The upper bound in part b) follows directly from part a). Using part a) and the reverse doubling condition, with $\tau^{-1} \leq \frac{d(X)}{3}$, we have for $l \in \mathbb{N}$

$$\begin{aligned}\mathbb{1}_{[\tau,2^l\tau]}(\sqrt{L})(x,x) &= \mathbb{1}_{[0,2^l\tau]}(\sqrt{L})(x,x) - \mathbb{1}_{[0,\tau]}(\sqrt{L})(x,x) \\ &\geq \frac{c_3}{\mu(B(x,2^{-l}\tau^{-1}))} - \frac{c_4}{\mu(B(x,\tau^{-1}))} \\ &\geq \frac{c_3 c 2^{ld} - c_4}{\mu(B(x,\tau^{-1}))}\end{aligned}$$

This proves the lemma with $b = 2^l$ for sufficiently large l such that the bound is positive. □

Now we can list and prove the properties of the system $\{\psi_{j\xi}\}$

Theorem 7.17. *The system $\{\psi_{j\xi}\}_{j\in\mathbb{N}_0,\xi\in\mathcal{X}_j}$ obeys the following properties.*

(a) Localization: For any $\sigma > 0$ there exists a constant $c_\sigma > 0$ such that for any $\xi \in \mathcal{X}_j$, $j \geq 0$, and $x \in X$ we have

$$|\psi_{j\xi}(x)| \leq c_\sigma \mu(B(\xi, b^{-j}))^{-1/2}(1 + b^j d(x,\xi))^{-\sigma}$$

and if $y \in X$ is such that $d(x,y) < b^{-j}$, then

$$|\psi_{j\xi}(x) - \psi_{j\xi}(y)| \leq c_\sigma \mu(B(\xi, b^{-j}))^{-1/2}(b^j d(x,y))^{\alpha}(1 + b^j d(x,\xi))^{-\sigma}$$

with $\alpha > 0$.

(b) Norm:

$$\|\psi_{j\xi}\|_p \simeq \mu(B(\xi, b^{-j}))^{\frac{1}{p}-\frac{1}{2}}$$

where $1 < p \leq \infty$. The constants in the equivalence only depend of p.

(c) Spectral localization ("frequency localization"): $\psi_{0\xi} \in \Sigma_b^p$ for $\xi \in \mathcal{X}_0$ and $\psi_{j\xi} \in \Sigma^p_{[b^{j-1},b^{j+1}]}$ if $\xi \in \mathcal{X}_j$, $j \geq 1$ and $1 < p \leq \infty$.

(d) The system $(\psi_{j\xi})$ is a frame for $L^2(X,\mu)$ as in Definition 7.13, with constants $c = \frac{1}{4}$ and $C = 2$.

Proof. d) follows immediately from the construction we did above. a) follows from Theorem 7.5, so only parts b) and c) remains to be proven. We start with the easier part c).

For $j = 0$ let $\theta \in C_c^\infty(\mathbb{R}_+)$ such that $\theta(u) = 1$ for $0 \leq u \leq b$, and for $j \geq 1$ let $\theta \in C_c^\infty(\mathbb{R}_+)$ be such that $\theta(u) = 1$ for all $b^{j-1} \leq u \leq b^{j+1}$. Then we have for all $j \geq 0$ and all $f \in L^{p'}(X,\mu)$, $p^{-1} + p'^{-1} = 1$,

$$\begin{aligned}
\left\langle f, \theta(\sqrt{L})\psi_{j\xi} \right\rangle &= \mu(A_\xi^j)^{1/2} \left\langle f, \theta(\sqrt{L})\Psi_j(\sqrt{L})(\cdot,\xi) \right\rangle \\
&= \mu(A_\xi^j)^{1/2} \int_X f(x)\overline{(\theta(\sqrt{L})\Psi_j(\sqrt{L})(\cdot,\xi))}(x)\,\mathrm{d}\mu(x) \\
&= \mu(A_\xi^j)^{1/2} \int_X \overline{\int_X \theta(\sqrt{L})(x,y)\Psi_j(\sqrt{L})(y,\xi)\,\mathrm{d}\mu(y)} f(x)\,\mathrm{d}\mu(x) \\
&= \mu(A_\xi^j)^{1/2} \int_X \overline{(\theta(\sqrt{L})\Psi_j(\sqrt{L}))(x,\xi)} f(x)\,\mathrm{d}\mu(x) \\
&= \mu(A_\xi^j)^{1/2} \int_X (\theta(\sqrt{L})\Psi_j(\sqrt{L}))(\xi,x) f(x)\,\mathrm{d}\mu(x) \\
&= \mu(A_\xi^j)^{1/2}\theta(\sqrt{L})\Psi_j(\sqrt{L})f(\xi) \\
&= \mu(A_\xi^j)^{1/2}((\theta\Psi_j)(\sqrt{L})f)(\xi) \\
&= \mu(A_\xi^j)^{1/2}(\Psi_j(\sqrt{L})f)(\xi) \\
&= \mu(A_\xi^j)^{1/2} \left\langle f, \Psi_j(\sqrt{L})(\cdot,\xi) \right\rangle .
\end{aligned}$$

As this is true for all $f \in L^{p'}(X,\mu)$, it follows that $\psi_{0\xi} \in \Sigma_b^p$ and $\psi_{j\xi} \in \Sigma^p_{[b^{j-1},b^{j+1}]}$.

It now remains to prove b). For this, notice that by Theorem 7.5 we have $|\Psi_j(\sqrt{L})(x,y)| \leq c'\mu(A_\xi^j)^{1/2} D_{b^j,\sigma}(x,y)$, and so we can use part d) of Lemma 7.3 to obtain that

$$\begin{aligned}
\|\psi_{j\xi}\|_p = \|\mu(A_\xi^j)^{1/2}\overline{\Psi_j(\sqrt{L})}(\xi,\cdot)\|_p &\leq c(p)c'\mu(A_\xi^j)^{1/2}\mu(B(\xi,b^{-j}))^{1/p-1} \\
&\leq c\mu(B\xi,b^{-j})^{1/p-1/2}.
\end{aligned}$$

For the other direction, we note that $\Psi_0 = 1$ on $[0,1]$ and that we assumed that Φ was selected such that $\Psi \geq c > 0$ on $[b^{-3/4}, b^{3/4}]$. We want to give lower bound estimates for $\|\Psi_0(\sqrt{L})(x,y)\|_2^2$ and $\|\Psi_j(\sqrt{L})(x,y)\|_2^2$. We can estimate these norms by the values in (ξ,ξ): From the identity

$$\begin{aligned}
\Psi_0(\sqrt{L})^2 f(\xi) &= \Psi_0(\sqrt{L})\Psi_0(\sqrt{L})f(\xi) \\
&= \iint \Psi_0(\sqrt{L})(\xi,u)\Psi_0(\sqrt{L})(u,y)f(y)\,\mathrm{d}\mu(y)\,\mathrm{d}\mu(u)
\end{aligned}$$

and Fubini we obtain that

$$\Psi_0^2(\sqrt{L})(\xi,y) = \int_X \Psi_0(\sqrt{L})(\xi,u)\Psi_0(\sqrt{L})(u,y)\,\mathrm{d}\mu(u)$$

and by filling in $y = \xi$ we have

$$\Psi_0^2(\sqrt{L})(\xi,\xi) = \int_X |\Psi_0(\sqrt{L})(\xi,u)|^2\,\mathrm{d}\mu(u) = \|\Psi_0(\sqrt{L})(\xi,\cdot)\|_2^2$$

This is also true for Ψ_j. Now we can use that $\Psi_0 \geq \mathbb{1}_{[0,1]}$ and $\Psi_j \geq \mathbb{1}_{[b^{j-3/4},b^{j+3/4}]}$ together with the lower bound estimates in part a) and b) of Lemma 7.16. We have that

$$\begin{aligned}
\psi_{0\xi}^2(\xi) &\geq \mu(A_\xi^j)\mathbb{1}_{[0,1]}(\sqrt{L})(\xi,\xi) \geq c_3 \\
\Psi_{j\xi}^2(\xi) &\geq \mu(A_\xi^j)c^2\mathbb{1}_{[b^{j-3/4},b^{j+3/4}]}(\sqrt{L})(\xi,\xi) \\
&\geq c^2\mu(A_\xi^j)\mathbb{1}_{[b^{j-3/4},b^{j+1/4}]}(\sqrt{L})(\xi,\xi) = c^2\mu(A_\xi^j)\mathbb{1}_{[b^j,b^{j+1}]}(b^{3/4}\sqrt{L}) \geq c^2 c_5
\end{aligned}$$

By the remark before we have established the lower bounds for $\|\psi_{0\xi}\|_2$ and $\|\psi_{j\xi}\|_2$. We will extrapolate this to all $0 < p < \infty$ using the upper bound estimates. To make the upcoming estimations easier to read, we notate $f_j := \psi_{j\xi}$ for $j \geq 0$. Now for $0 < p < 2$ we have

$$c_3 \leq \|f_0\|_2^2 \leq \|f_0\|_p^p \|f_0\|_\infty^{2-p} \leq c\|f_0\|_p^p \mu(B(\xi,1))^{1/2(p-2)}$$

This implies that $\|f_0\|_p \geq c'\mu(B(\xi,1))^{1/p-1/2}$. Similar estimates hold for f_j, $j \neq 0$, with a different constant and $B(\xi,b^j)$ instead of $B(\xi,1)$. For the case $p > 2$, we have by Hölder's inequality, with $\frac{1}{p'} := 1 - \frac{1}{p}$

$$c_3 \leq \|f_0\|_2^2 \leq \|f\|_p \|f\|_{p'} \leq c\|f\|_p^p \mu(B(\xi,1))^{1/p'-1/2}$$

And so $\|f_0\|_p \geq c'\mu(B(\xi,1))^{1/2-1/p'} = c'\mu(B(\xi,1))^{1/2-1+1/p} = c'\mu(B(\xi,1))^{1/p-1/2}$. Again, similar estimates follow for f_j with $j \geq 0$. This proves the equivalence and the proposition. □

8 Wavelets and the finite speed of propagation

8.1 Partition of unity without compact support

In the last chapter, we showed a way to construct wavelets out of positive selfadjoint operators using a partition of unity. We like to redo the construction, but this time, we want to use the finite speed of propagation property we discussed in Section 2.4. We can do this, as the Davies-Gaffney estimate is a weaker conditions than the Gaussian estimate (7.2). The partition of unity will not have compact support in this new setting, but the finite speed of propagation property will give the wavelets compact support.

We start again with a partition of unity, like in last chapter, but this time, we assume the functions Ψ_j to satisfy the conditions of Theorem 2.8, that is, $\operatorname{supp} \mathcal{F}\Psi \subseteq B(0,R)$ for some $R>0$, and Ψ_j has exponential decay of the form $\Psi(x) = g(x)e^{-\alpha x}$ with $\alpha > 1$ and $g(x)$ bounded. We would like to prove that $\Psi_j(\sqrt{L})$ is an integral operator. However, in the last chapter we used Theorem 7.5 to prove this, but we cannot do that anymore, since the estimations in the proof heavily relied on the compact support of the function. Instead, we will state a theorem which is similar to Theorem 7.4 which is modified to our new setting. For this, we first need to prove that $e^{-t\sqrt{L}}$ is a kernel operator for each $t>0$. We will do so by the H^∞-calculus, briefly covered in Section 2.1.

In this chapter, we rely on the assumptions of the last chapter. We will not explicitly state them in each theorem.

Proposition 8.1. *For each $t > 0$, the operator $e^{-t\sqrt{L}}$ is an integral operator.*

Proof. We start by proving that $R(z, L)$ is an integral operator for $\operatorname{Re} z < 0$. For this, we note that

$$(z-\lambda)^{-1} = -\int_0^\infty e^{uz}e^{-u\lambda}\,\mathrm{d}u.$$

Integrating left and right over λ, with respect to the spectral resolution $(E_\lambda)_{\lambda\geq 0}$ of L, we have that

$$(z-L)^{-1} = -\int_0^\infty e^{uz}e^{-uL}\,\mathrm{d}u.$$

Now we use that fact that e^{-uL} is an integral operator with kernel p_u, and get that

$$\begin{aligned} R(z,L)f_1(x) &= -\int_0^\infty e^{uz}e^{-uL}f_1(x)\,\mathrm{d}u \\ &= -\int_0^\infty e^{uz}\int_X p_u(x,y)f_1(y)\,\mathrm{d}\mu(y)\,\mathrm{d}u \\ &= -\int_X\int_0^\infty e^{uz}p_u(x,y)\,\mathrm{d}u f_1(y)\,\mathrm{d}\mu(y) \end{aligned}$$

This would imply that $R(z,L)(x,y) = -\int_0^\infty e^{uz} p_u(x,y)\,\mathrm{d}u$, if we can justify Fubini in the last line:

$$\begin{aligned}\iint |e^{uz}||p_u(x,y)||f_1(y)|\,\mathrm{d}\mu(y)\,\mathrm{d}u &= \int |e^{uz}| \int |p_u(x,y)f_1(y)|\,\mathrm{d}\mu(y)\,\mathrm{d}u \\ &\le \int |e^{uz}|\,\mathrm{d}u C\|f_1\| \\ &< \infty\end{aligned}$$

Now note that we have $e^{-t\sqrt{\lambda}} \le 1/(1+\lambda)$, and so $\|e^{-t\sqrt{L}}\|_{2\to 2} \le \|(1+L)^{-1}\|_{2\to 2}$. This implies that $\|e^{-t\sqrt{L}}\|_{1\to\infty} \le \|(1+L)^{-1}\|_{1\to\infty}$, and so $e^{-t\sqrt{L}}$ is an integral operator.

Now we use the Cauchy representation of $e^{-t\sqrt{L}}$ via the sectorial functional calculus [8, Chapter 9], and write for $\theta \in (\pi/2, \pi)$

$$\begin{aligned}e^{-t\sqrt{z}} f_1(x) &= \int_{\partial S_\theta} e^{-t\sqrt{z}} R(z,L) f_1(x)\,\mathrm{d}z \\ &= -\int_{\partial S_\theta}\int_X \int_0^\infty e^{-t\sqrt{z}+uz} p_u(x,y)\,\mathrm{d}u f_1(y)\,\mathrm{d}\mu(y)\,\mathrm{d}z\end{aligned}$$

We want to apply Fubini a second time. That is allowed, because for $f_1 \in L^1(X,\mu)$, we have

$$\begin{aligned}\iiint |e^{-t\sqrt{z}}||e^{uz}||p_u(x,y)||f_1(y)|\,\mathrm{d}z\,\mathrm{d}u\,\mathrm{d}\mu(y) &\lesssim \iint |\frac{1}{u} p_u(x,y)||f_1(y)|\,\mathrm{d}u\,\mathrm{d}\mu(y) \\ &\lesssim \int |f_1(y)|\,\mathrm{d}\mu(y) < \infty\end{aligned}$$

After applying Fubini, to the integral $-\int_{\partial S_\theta}\int_X \int_0^\infty e^{-t\sqrt{z}+uz} p_u(x,y)\,\mathrm{d}u f_1(y)\,\mathrm{d}\mu(y)\,\mathrm{d}z$ to exchange the outer two integrals, we end up with

$$e^{-t\sqrt{L}}(x,y) = -\int_{\partial S_\theta}\int_0^\infty e^{uz-t\sqrt{z}} p_u(x,y)\,\mathrm{d}u\,\mathrm{d}z$$

Hence $e^{-t\sqrt{L}}$ is an integral operator. □

Theorem 8.2. *Let $g : \mathbb{R} \to \mathbb{C}$ be a measurable function such that for some $\sigma > 0$*

$$\|g\|_* := \int_{\mathbb{R}} |\hat{g}(\xi)|(1+|\xi|)^\sigma\,\mathrm{d}\xi < \infty$$

Then $g(\delta\sqrt{L})e^{-\delta\sqrt{L}}$ is an integral operator for $0 < \delta \le 1$.

Proof. Note that just like in the proof of Theorem 7.4, we have that $\|\hat{g}\|_1 < \infty$, hence $\|g\|_\infty < \infty$ and the Fourier inversion formula holds. Let $\phi, \psi \in C_c^\infty(X,d)$.

Then we have for $g(\delta\sqrt{L})e^{-\delta\sqrt{L}}$

$$\begin{aligned}
\left\langle g(\delta\sqrt{L})e^{-\delta\sqrt{L}}\phi,\psi\right\rangle &= \int_0^\infty g(\delta\lambda)e^{-\delta\lambda}\,\mathrm{d}\left\langle F_\lambda\phi,\psi\right\rangle \\
&= \int_0^\infty \frac{1}{2\pi}\int_{\mathbb{R}} \hat{g}(\xi)e^{i\delta\lambda\xi}\,\mathrm{d}\xi e^{-\delta\lambda}\,\mathrm{d}\left\langle F_\lambda\phi,\psi\right\rangle \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \hat{g}(\xi)\int_0^\infty e^{-\delta\lambda(1-i\xi)}\,\mathrm{d}\left\langle F_\lambda\phi,\psi\right\rangle\,\mathrm{d}\xi \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \hat{g}(\xi)\left\langle e^{-\delta(1-i\xi)\sqrt{L}}\phi,\psi\right\rangle\mathrm{d}\xi \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \hat{g}(\xi)\int_X\int_X e^{-\delta(1-i\xi)\sqrt{L}}(x,y)\phi(x)\psi(y)\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)\,\mathrm{d}\xi \\
&= \int_X\int_X \frac{1}{2\pi}\int_{\mathbb{R}} \hat{g}(\xi)e^{-\delta(1-i\xi)\sqrt{L}}(x,y)\,\mathrm{d}\xi\phi(x)\psi(y)\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)
\end{aligned}$$

We used Fubini twice: in the third line and in the sixth line. In the third line, it is allowed because

$$\iint |\hat{g}(\xi)||e^{-\delta\lambda(1-i\xi)}|\,\mathrm{d}\|F_\lambda\phi\|_2^2\,\mathrm{d}\xi \le \iint |\hat{g}(\xi)|\,\mathrm{d}\|F_\lambda\phi\|_2^2\,\mathrm{d}\xi = \|\phi\|_2^2\|\hat{g}\|_1 < \infty$$

In the sixth line, we have that

$$\begin{aligned}
&\iiint |\hat{g}(\xi)|e^{-\delta(1-i\xi)\sqrt{L}}(x,y)||\phi(x)||\psi(y)|\,\mathrm{d}\xi\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \\
&\qquad \le \int |\hat{g}(\xi)|\,\mathrm{d}\xi\int|\phi(x)|\,\mathrm{d}\mu(x)\int|\psi(y)|\,\mathrm{d}\mu(y) < \infty
\end{aligned}$$

This actually shows that $\left\langle g(\delta\sqrt{L})e^{-\delta\sqrt{L}}\phi,\psi\right\rangle \lesssim \|\phi\|_1\|\psi\|_1$, so that by Proposition 7.1 we have that $g(\delta\sqrt{L})e^{-\delta\sqrt{L}}$ is an integral operator. By the calculations above, we have that the integral kernel is given by

$$g(\delta\sqrt{L})e^{-\delta\sqrt{L}}(x,y) = \frac{1}{2\pi}\int_{\mathbb{R}}\hat{g}(\xi)e^{-\delta(1-i\xi)\sqrt{L}}(x,y)\,\mathrm{d}\xi. \qquad \square$$

As by assumption we have that Ψ has exponential decay, we can write it in the form $\Psi(x) = g(x)e^{-\alpha x} = g(x)e^{-(\alpha-1)x}e^{-x}$ with $\alpha > 1$, where $g(x)$ is a bounded function. It is then easy to verify that $\|x \mapsto g(x)e^{-(\alpha-1)x}\|_* < \infty$. Hence we can apply the above theorem and obtain that $\Psi(\delta\sqrt{L})$ is an integral operator.

Furthermore, as [7, Theorem 1.2] does not require F to have compact support, we can apply it again. From this, it follows that $\Psi(\delta\sqrt{L})$ is a bounded operator on $L^p(X,\mu)$.

Note that by the assumptions on L we have that the semigroup $(e^{-zL})_{z\in\mathbb{C}_+}$ satisfies the Davies-Gaffney estimate. So by Theorems 2.6 and 2.8 we have that L has the finite speed of propagation property with respect to Ψ. We can use this to prove that $y \mapsto \Psi_j(\sqrt{L})(x,y)$ has compact support. Indeed, take $U_1, U_2 \subseteq X$ such that $d(U_1,U_2) > \frac{R}{b^j}$. Then let $f_1, f_2 \in C_c^\infty(\mathbb{R})$ such that $\operatorname{supp} f_i \subseteq U_i$, $i = 1, 2$. Then we have

$$\begin{aligned}\left\langle \Psi_j(\sqrt{L})f_1, f_2 \right\rangle &= \int_X \Psi_j(\sqrt{L})f_1(x)\overline{f_2}(x)\,\mathrm{d}\mu(x) \\ &= \int_X \int_X \Psi_j(\sqrt{L})(x,y)f_1(y)\,\mathrm{d}\mu(y)\overline{f_2}(x)\,\mathrm{d}\mu(x) \\ &= \int_{U_1}\int_{U_2} \Psi_j(\sqrt{L})(x,y)f_1(y)\overline{f_2}(x)\,\mathrm{d}\mu(y)\,\mathrm{d}\mu(x) \\ &= 0\end{aligned}$$

By varying f_1 and f_2, we can show that $\operatorname{supp}(y \mapsto \Psi_j(\sqrt{L})(x,y)) \subseteq U_2^c$ Now by varying U_2 such that $d(U_1,U_2) > \frac{R}{b^j}$, we can show that $\operatorname{supp}(y \mapsto \Psi_j(\sqrt{L})(x,y)) \subseteq B(x, \frac{R}{b^j})$, hence the kernel has compact support. By definition of $\psi_{j\xi}$, it follows that $\psi_{j\xi}$ also has compact support. From this it follows that tail estimations like the first estimation in Theorem 7.17 a) are trivial. However, by construction of the kernel, the Hölder continuity in Theorem 7.17 a) may not be true. For this reason, we discuss a different kind of smoothness assumption on L in the next section.

For the sake of argument in the remainder of this section, we assume that the metric space (X,d) is either $\mathbb{R}^d$ with the Euclidean metric, or a Riemannian manifold. In these cases, we let ∇ be the (Riemannian) gradient. We will assume that this X is selected such that the $p-$*Poincaré*-inequality

$$\int_B |f - \langle f \rangle_B|\,\mathrm{d}\mu \leq Cr\mu(B)^{1-1/p}\left(\int_B |\nabla f|^p\,\mathrm{d}\mu\right)^{1/p}, \qquad \langle f \rangle_B := \frac{1}{\mu(B)}\int_B f\,\mathrm{d}\mu \tag{8.1}$$

holds for any ball $B \subseteq X$, with $r := \operatorname{rad}(B)$ and $f \in L^1_{loc}(X,\mu)$ (cf. [9, (5)]).

8.2 Gradient estimates

In this subsection, we relax the assumptions on L. Instead of the Hölder continuity assumption on the kernels of the semigroup, we require a uniform bound on the gradient of the semigroup:

$$\sup_{t>0} \|t^{1/2}\nabla e^{-tL}\|_{p\to p} < \infty \tag{8.2}$$

for some $p \in (1,\infty)$. In the case $\alpha = 1$ in (7.3), this assumption follows immediately from the Hölder continuity: multiply left and right by $\sqrt{t}/d(y,y')$, and take $y' \to y$. Then note that the remainder term on the right hand side is in $L^p(X,\mu)$

As we needed the Hölder continuity assumption for Proposition 7.9, we have to give an alternative to this proposition. For this, we need to show that the same estimate holds for $\phi(\sqrt{L})$, used in this inequality.

We will show that the gradient estimate given above also holds for all other kernel operators derived from this one, especially for $e^{-t\sqrt{L}}$ and $\Psi(\sqrt{L})$:

Proposition 8.3. *Assume that (8.2) holds for some $p \in (1, \infty)$. Then*

$$\sup_{t>0} \|t\nabla e^{-t\sqrt{L}}\|_{p\to p} < \infty.$$

Proof. From the proof of Proposition 8.1 we have that the operator $e^{-t\sqrt{L}}$ is given by

$$e^{-t\sqrt{L}} = -\int_{\partial S_\theta}\int_0^\infty e^{uz-t\sqrt{z}} e^{-uL}\,\mathrm{d}u\,\mathrm{d}z$$

as a Bochner integral in the space of bounded operators. By the theory on Bochner integrals, briefly covered in Section 2.1, we have that

$$\nabla e^{-t\sqrt{L}} = -\int_{\partial S_\theta}\int_0^\infty e^{uz-t\sqrt{z}}\nabla e^{-uL}\,\mathrm{d}u\,\mathrm{d}z.$$

In order to show that $t\nabla e^{-t\sqrt{L}}$ is bounded on $L^p(X,\mu)$, we want to estimate $|\left\langle t\nabla e^{-t\sqrt{L}} f, g\right\rangle|$ for $f \in L^p(X,\mu)$, $g \in L^{p'}(X,\mu)$, where $\frac{1}{p'} = 1 - \frac{1}{p}$.

$$\begin{aligned}
\left|\left\langle t\nabla e^{-t\sqrt{L}} f, g\right\rangle\right| &= \left|\left\langle -\iint te^{uz-t\sqrt{z}}\nabla e^{-uL}\,\mathrm{d}u\,\mathrm{d}z f, g\right\rangle\right| \\
&= \left|\iiint tu^{-1/2}e^{uz-t\sqrt{z}}(u^{1/2}\nabla e^{-uL}f)(x)\,\mathrm{d}u\,\mathrm{d}z g(x)\,\mathrm{d}\mu(x)\right| \\
&= \left|\iiint tu^{-1/2}e^{uz-t\sqrt{z}}(u^{1/2}\nabla e^{-uL}f)(x) g(x)\,\mathrm{d}\mu(x)\,\mathrm{d}u\,\mathrm{d}z\right| \\
&\le \left|\iint tu^{-1/2}e^{uz-t\sqrt{z}}\left|\left\langle u^{1/2}\nabla e^{-uL}f, g\right\rangle\right|\,\mathrm{d}u\,\mathrm{d}z\right| \\
&\le \left|\iint tu^{-1/2}e^{uz-t\sqrt{z}}\,\mathrm{d}u\,\mathrm{d}z\right| \sup_{t>0}\|t^{1/2}\nabla e^{-tL}\|_{p\to p}\|f\|_p\|g\|_{p'}
\end{aligned}$$

So if the first integral is uniformly bounded in $t > 0$, we have that $t\nabla e^{-t\sqrt{L}}$ is a bounded operator on $L^p(X,\mu)$, moreover, we have that $\sup_{t>0}\|t\nabla e^{-t\sqrt{L}}\|_{p\to p} < \infty$. Lastly, that also justifies Fubini in the third line.

We choose $\theta \in (\pi/2, \pi)$. Then we have that

$$\begin{aligned}
\left|\int_{\partial S_\theta}\int_0^\infty tu^{-1/2}e^{uz-t\sqrt{z}}\,\mathrm{d}u\,\mathrm{d}z\right| &= \left|\int_{\partial S_\theta}\int_0^\infty 2te^{u^2z-t\sqrt{z}}\,\mathrm{d}u\,\mathrm{d}z\right| \\
&= \left|\int_{\partial S_\theta} 2t\frac{\sqrt{\pi}}{2\sqrt{z}}e^{-t\sqrt{z}}\,\mathrm{d}z\right| \\
&= \left|2\sqrt{\pi}\int_{\partial S_{\theta/2}} te^{-tz}\,\mathrm{d}z\right| \\
&= \left|2\sqrt{\pi}\left(\int_0^\infty te^{-tre^{-i\theta/2}}\,\mathrm{d}r - \int_0^\infty te^{-tre^{i\theta/2}}\,\mathrm{d}r\right)\right| \\
&= \left|2\sqrt{\pi}\left(e^{i\theta/2} - e^{-i\theta/2}\right)\right| \\
&= \sqrt{\pi}\sin(\theta/2)
\end{aligned}$$

We have $\theta/2 \in (\pi/4, \pi/2)$. This means that $\sin(\theta/2) < \sin(\pi/2) = 1$, and so the integral is uniformly bounded by $\sqrt{\pi}$. This, together with the earlier estimations, means that $\sup_{t>0}\|t\nabla e^{-t\sqrt{L}}\|_{p\to p} < \infty$. □

Using this result, we can show that $b^{-j}\nabla\Psi(\sqrt{L})$ is also a bounded operator.

Proposition 8.4. *Assume that (8.2) holds for some $p \in (1,\infty)$. Then $b^{-j}\nabla\Psi_j(\sqrt{L})$ is a bounded operator on $L^p(X,\mu)$.*

Proof. We use the same principle as in the proof of Theorem 8.2. We once again write $\Psi(b^{-j}\sqrt{L}) = g(b^{-j}\sqrt{L})e^{-b^{-j}\sqrt{L}}$. Then we write for $f \in L^p(X,\mu)$

$$g(b^{-j}\sqrt{L})e^{-b^{-j}\sqrt{L}}f = \frac{1}{2\pi}\int_{\mathbb{R}}\hat{g}(\xi)e^{-b^{-j}(1-i\xi)\sqrt{L}}f\,\mathrm{d}\xi$$

as a Bochner integral. By appling the gradient to $\Psi(b^{-j}\sqrt{L})$, and testing against arbitrary $h \in L^{p'}(X,\mu)$ with $\|h\|_{p'} \le 1$, p' being the Hölder conjugate of p, we have

$$\begin{aligned}
&\left\|b^{-j}\nabla g(b^{-j}\sqrt{L})e^{b^{-j}\sqrt{L}}f\right\|_p \\
&\qquad= \left|\left\langle\frac{1}{2\pi}\int_{\mathbb{R}}\hat{g}(\xi)b^{-j}\nabla e^{-b^{-j}(1-i\xi)\sqrt{L}}f\,\mathrm{d}\xi, h\right\rangle\right| \\
&\qquad= \left|\int_X\frac{1}{2\pi}\int_{\mathbb{R}}\hat{g}(\xi)(b^{-j}\nabla e^{-b^{-j}(1-i\xi)\sqrt{L}}f)(x)\,\mathrm{d}\xi h(x)\,\mathrm{d}\mu(x)\right| \\
&\qquad= \left|\frac{1}{2\pi}\int_{\mathbb{R}}\hat{g}(\xi)\int_X(b^{-j}\nabla e^{-b^{-j}(1-i\xi)\sqrt{L}}f)(x)h(x)\,\mathrm{d}\mu(x)\,\mathrm{d}\xi\right| \\
&\qquad\le \frac{1}{2\pi}\int_{\mathbb{R}}|\hat{g}(\xi)|\|b^{-j}\nabla e^{-b^{-j}(1-i\xi)\sqrt{L}}\|_{p\to p}\|f\|_p\|h\|_{p'}\,\mathrm{d}\xi \\
&\qquad\le \frac{\|\hat{g}\|_1}{2\pi}\sup_t\|t\nabla e^{-t\sqrt{L}}\|_{p\to p}\|f\|_p
\end{aligned}$$

□

We can adapt previous theorem for $\phi \in C_c^\infty$, $\phi \equiv 1$ on $[0,1]$, $0 \le \phi \le 1$, by writing $\phi(\delta\sqrt{L}) = \phi(\delta\sqrt{L})e^{\sqrt{L}}e^{-\sqrt{L}} =: g_\delta(\sqrt{L})e^{-\sqrt{L}}$, and then we use that g_δ has compact support, so it is bounded and in $L^1(\mathbb{R})$. From this, we obtain that $\|\nabla\phi(\delta\sqrt{t})\|_{p\to p} < \infty$. With this information, we can give the replacement for Proposition 7.9.

For the estimates needed in that proposition, we will need Theorem 3.2 by Hajłasz-Koskela [9]. We repeat the proof here.

Theorem 8.5. *Assume that X is a Riemannian manifold, or $\mathbb{R}^n$, and that μ is a doubling measure, such that the p-Poincaré inequality (8.1) holds, with $u \in L^1_{loc}(X,\mu)$ and $p \in [1,\infty)$. Then*

$$|u(x) - u(y)| \le Cd(x,y)\left((M_{2d(x,y)}(\nabla u)^p(x))^{1/p} + (M_{2d(x,y)}(\nabla u)^p(y))^{1/p}\right)$$

for almost every $x, y \in X$, where $M_R f(x) := \sup_{0<r<R} \frac{1}{\mu(B(x,r))}\int_{B(x,r)} |f| \,\mathrm{d}\mu$.

Proof. We first note that the Lebesgue differentiation theorem holds for any space of homogeneous type: At first, the Vitali covering lemma works for any separable measure space, and the doubling property actually implies that X is separable. Then, by restricting the maximal function to a maximal radius R, we can redo the argument in case of the Lebesgue measure to show that the maximal function is weak type (1,1), (we need the R to get a maximum radius for the Vitali covering lemma). Using this, and the fact that continuous functions are dense in L^1, the proof of the Lebesgue differentiation theorem holds without change. More details can be found in [9, Appendix 14.6].

So by the Lebesgue differentiation theorem, almost all points in X are Lebesgue points of u. Let $x, y \in X$ be such points. We write $B_i(x) := B(x, 2^{-i}r) := B(x, 2^{-i}d(x,y))$ for $i \in \mathbb{N}_0$. By the definition of Lebesgue points, we have that $\langle u\rangle_{B_i(x)} \to u(x)$ as $i \to \infty$, where $\langle u\rangle_{B_i(x)} := \mu(B_i(x))^{-1}\int_{B_i(x)} u \,\mathrm{d}\mu$. We will now use the p-Poincaré inequality and the doubling of μ to obtain that

$$\begin{aligned}
|u(x) - \langle u\rangle_{B_0(x)}| &\le \sum_{i=0}^\infty |\langle u\rangle_{B_i(x)} - \langle u\rangle_{B_{i+1}(x)}| \\
&= \sum_{i=0}^\infty \left|\left\langle u - \langle u\rangle_{B_i(x)}\right\rangle_{B_{i+1}(x)}\right| \\
&\le \sum_{i=0}^\infty \frac{1}{\mu(B_{i+1}(x))}\int_{B_{i+1}(x)} |u - \langle u\rangle_{B_i(x)}| \,\mathrm{d}\mu \\
&\le \sum_{i=0}^\infty \frac{C}{\mu(B_i(x))}\int_{B_i(x)} |u - \langle u\rangle_{B_i(x)}| \,\mathrm{d}\mu \\
&\le \sum_{i=0}^\infty C2^{-i}r\left(\frac{1}{\mu(B_i(x))}\int_{B_i(x)} (\nabla u)^p \,\mathrm{d}\mu\right)^{1/p}
\end{aligned}$$

$$\begin{aligned}&\le \sum_{i=0}^{\infty} C2^{-i} r(M_{d(x,y)}(\nabla u)^p(x))^{1/p}\\&= Cr(M_{d(x,y)}(\nabla u)^p(x))^{1/p}\end{aligned}$$

Similar estimates also hold for y. We want to write $|u(x)-u(y)| \le |u(x) - \langle u\rangle_{B_0(x)}| + |u(y)-\langle u\rangle_{B_0(y)}| + |\langle u\rangle_{B_0(x)} - \langle u\rangle_{B_0(y)}|$. We have estimated the first two terms, now for the last one:

$$\begin{aligned}|\langle u\rangle_{B_0(x)} - \langle u\rangle_{B_0(y)}| &\le |\langle u\rangle_{B_0(x)} - \langle u\rangle_{2B_0(x)}| + |\langle u\rangle_{B_0(y)} - \langle u\rangle_{2B_0(x)}|\\&\le \frac{K}{\mu(2B_0(x))}\int_{2B_0(x)} |u-\langle u\rangle_{2B_0(x)}|\,\mathrm{d}\mu\\&\le K\left(\frac{1}{\mu(2B_0(x))}\int_{2B_0(x)}(\nabla u)^p\,\mathrm{d}\mu\right)^{1/p}\\&\le K(M_{2d(x,y)}(\nabla u)^p(x))^{1/p}\end{aligned}$$

By noting that $M_{d(x,y)}(\nabla u)^p \le M_{2d(x,y)}(\nabla u)^p$, and combining the three estimates, we have proven the theorem. □

Using above theorem, we can prove the equivalence to Proposition 7.9

Proposition 8.6. *Assume that the q-Poincaré inequality holds for some $q \ge 1$, and that (8.2) holds for some $p \in (q,\infty)$. Fix $\lambda \ge 1$, and let $\mathcal{X}$ be a maximal δ-net on X with $\delta := \gamma/\lambda$, with $0<\gamma<1$. Suppose $\{A_\xi\}_{\xi\in\mathcal{X}}$ is a companion disjoint partition of X. Then for any $f \in \Sigma^p_\lambda$,*

$$\sum_{\xi\in\mathcal{X}}\int_{A_\xi}|f(x)-f(\xi)|^p\,\mathrm{d}\mu(x) \le K\gamma^p\|\lambda^{-1}\nabla\phi(\lambda^{-1}\sqrt{L})\|^p_{p\to p}\|f\|^p_p.$$

Proof. We use that $f\in\Sigma^p_\lambda$, so that we can apply $\phi(\lambda^{-1}\sqrt{L})f = f$. We use the notation $\langle f\rangle_Q := \mu(Q)^{-1}\int_Q f\,\mathrm{d}\mu$, and write $u := \phi(\lambda^{-1}\sqrt{L})f$. Now using Theorem 8.5, we have

$$\begin{aligned}|u(x)-u(\xi)| &\le C\delta((M_{2\delta}(\nabla u)^q(x))^{1/q} + (M_{2\delta}(\nabla u)^q(\xi))^{1/q})\\&\le C\delta((M_{2\delta}(\nabla u)^q(x))^{1/q} + (M'_{2\delta}(\nabla u)^q(\xi))^{1/q})\\&\le C\delta((M_{2\delta}(\nabla u)^q(x))^{1/q} + K(M_{2\delta}(\nabla u)^q(x))^{1/q})\end{aligned}$$

where

$$M_R f(x) := \sup_{0<r<R}\frac{1}{\mu(B(x,r))}\int_{B(x,r)} f\,\mathrm{d}\mu,$$

and

$$M'_R f(x) := \sup\{\frac{1}{\mu(B)}\int_B f\,\mathrm{d}\mu : x\in B = B(y,r), 0<r<R, y\in X\}$$

Moreover, we used that $\xi \in B(x, 2\delta)$ for each $x \in A_\xi$, and the equivalence

$$M_R f \le M'_R f \le K M_R f,$$

which follows from the doubling property of the measure. So we have that $M_R u(\xi) \le M'_R u(\xi) = M'_R u(x) \le K M_R u(x)$. Continuing with the argument, we have

$$\begin{aligned}
\sum_{\xi \in \mathcal{X}} \int_{A_\xi} |u(x) - u(\xi)|^p &\lesssim \delta^p \sum_{\xi \in \mathcal{X}} \int_{A_\xi} (M_{2\delta}(\nabla u)^q(x))^{p/q} \,\mathrm{d}\mu(x) \\
&= \delta^p \int_X (M_{2\delta}(\nabla u)^q(x))^{p/q} \,\mathrm{d}\mu(x) \\
&\le \delta^p \int_X ((\nabla u)^q)^{p/q} \,\mathrm{d}\mu(x) \\
&= \delta^p \|\nabla u\|_p^p \\
&= \gamma^p \|\lambda^{-1} \nabla \phi(\lambda^{-1}\sqrt{L}) f\|_p^p
\end{aligned}$$

□

In this case, we have to be careful with our δ-net $\mathcal{X}$. Theorem 8.5 is true for all Lebesgue points x and y of f, and so we need to have that all $\xi \in \mathcal{X}$ are Lebesgue points. But as these are dense, it does not give a problem during the construction of $\mathcal{X}$. This does mean that $\mathcal{X}$ depends on f. That is not a problem, however, as we will apply this proposition to the fixed functions Ψ_j and not to any arbitrary function.

With this proposition, we can apply Theorem 7.10 again without big changes (the constants inside the calculations change a little bit). So the construction of the wavelets does not change.

In the new situation, we have a theorem similar to 7.17

Theorem 8.7. *Assume that the q-Poincaré inequality holds for some $q \ge 1$, and that (8.2) holds for some $p \in (q, \infty)$. Then for the wavelets $\psi_{j\xi}$ and the operators $\Psi_j(\sqrt{L})$, the following statements hold:*

(a) Localization: Every $\psi_{j\xi}$ has compact support. Moreover, we have that $\|b^{-j}\nabla\Psi_j(\sqrt{L})\|_{p\to p} < \infty$, uniformly in $j \ge 0$.

(b) Norms:
$$\|\psi_{j\xi}\|_p \simeq \mu(B(\xi, b^{-j}))^{1/p-1/2}$$

(c) Spectral localization: Ψ_j have exponential decay, therefore we can find $\varepsilon > 0$ such that
$$\|\phi(b^{-1}\sqrt{L})\psi_{0\xi} - \psi_{0\xi}\|_p < \varepsilon \qquad \|\varphi(b^{-j}\sqrt{L})\psi_{j\xi} - \psi_{j\xi}\|_p < \varepsilon$$
for $\phi, \varphi \in C_c^\infty(\mathbb{R})$, $\phi \equiv 1$ on $[0,1]$, $\varphi \equiv 1$ on $[b^{-1}, b]$.

(d) The system $\{\psi_{j\xi}\}$ is a frame for $L^2(X, \mu)$.

Proof. (a) follows from the finite speed of propagation in the first subsection, and the gradient estimate of Proposition 8.4.

(b) follows using the same proof of Theorem 7.17 (b): By the compact support of $\Psi_j(\sqrt{L})(x,\xi)$, we have that there exists c' such that

$$|\psi_{j\xi}(x)| \leq c'\mu(A_\xi^j)^{1/2} D_{b^j,\sigma}(x,\xi).$$

Hence we can use Lemma 7.3 d) again to obtain that $|\psi_{j\xi}(x)| \leq c\mu(B(\xi, b^{-j}))^{1/p-1/2}$. For the other inequality, we once again use the trick to write $\|\psi_{j\xi}\|_2^2 = \psi_{j\xi}^2(\xi)$ and use Lemma 7.16 again to prove the lower bound.

For (c), using the same calculations as in Theorem 7.17 (c), we have that $\left\langle f, \theta(\sqrt{L})\psi_{j\xi}\right\rangle = \mu(A_\xi^j))^{1/2}((\theta\Psi_j)(\sqrt{L})f)(\xi)$ for all $f \in L^{p'}(X,\mu)$. Similarily, we have that $\langle f, \psi_{j\xi}\rangle = \mu(A_\xi^j))^{1/2}(\Psi_j)(\sqrt{L})f)(\xi)$. Therefore we have

$$\begin{aligned}\|\theta(b^{-j}\sqrt{L})\psi_{j\xi} - \psi_{j\xi}\|_p = \sup\{&\mu(A_\xi^j))^{1/2}((\theta\Psi_j - \Psi_j)(\sqrt{L})f)(\xi) : \\ &f \in L^{p'}(X,\mu), \|f\|_{p'} \leq 1\}\end{aligned}$$

If we choose $\theta = \phi(b^{-1}\cdot)$ for $j = 0$, or $\theta = \varphi$ for $j \geq 1$, we have that only the tail is left over in above expression. But by the exponential decay, we have that the tail is small, and so it must follow that $\|\theta(b^{-j}\sqrt{L})\psi_{j\xi} - \psi_{j\xi}\|_p < \varepsilon$ for some $\varepsilon > 0$.

(d) follows from Proposition 8.6 and the sampling theorem 7.10, which we modify by requiring that $K\gamma < 1/2$ respectively $K\gamma < \varepsilon/3$, where K is the constant in Proposition 8.6. □

9 Conclusion

In this paper, we presented two algorithms to construct wavelets out of semigroups on spaces of homogeneous type. The first approach requires the semigroup to consist of compact operators, and build a multiresolution analysis, out of which the wavelet spaces and bases can be constructed via the classical ways. The second way requires conditions on the heat kernel, and build a wavelet frame using a partition of unity. The last thing that is done in this paper, is modifying the algorithm, by using that the Davies-Gaffney estimate holds, and so the finite speed of propagation property holds. Then the partition of unity cannot have compact support anymore, but instead has exponential decay. By this finite speed of propagation however, it turned out that the wavelets have compact support instead. As the Hölder continuity may not hold anymore, we replaced the Hölder continuity assumption for the heat semigroup by a Poincaré inequality and an L^p-gradient estimation.

9.1 Future work

In the future, one may examine if the requirements on the semigroups can be reduced even further. Or one could research if all the space requirements are strictly necessary, so that wavelets could be build on spaces where for instance the doubling property does not hold.

References

[1] A.-P. Calderón. Intermediate spaces and interpolation, the complex method. *Studia Math.*, 24:113–190, 1964.

[2] Ronald R. Coifman and Mauro Maggioni. Diffusion wavelets. *Applied and Computational Harmonic Analysis*, 21(1):53 – 94, 2006. Special Issue: Diffusion Maps and Wavelets.

[3] T. Coulhon, G. Kerkyacharian, and P. Petrushev. Heat kernel generated frames in the setting of Dirichlet spaces. *Fourier Analysis and Applications*, 18:995 – 1066, 2012.

[4] Thierry Coulhon and Adam Sikora. Gaussian heat kernel upper bounds via the Phragmén-Lindelöf theorem. *Proc. London Math. Soc.*, 96(3):707 – 744, 2008.

[5] Ingrid Daubechies. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 41(7):909–996, 1988.

[6] E. B. Davies. *Heat Kernels and Spectral Theory.* Cambridge Tracts in Mathematics. Cambridge University Press, 1989.

[7] Xuan Thinh Duong, El Maati Ouhabaz, and Adam Sikora. Plancherel-type estimates and sharp spectral multipliers. *Journal of Functional Analysis*, 196(2):443–485, 2002.

[8] Markus Haase. Lectures on functional calculus. Technical report, Kiel University, 2018. Retrieved from http://www.math.uni-kiel.de/isem21/en/course/phase1/isem21-lectures-on-functional-calculus.

[9] P. Hajłasz and P. Koskela. *Sobolev Met Poincare*. Number nr. 688 in American Mathematical Society: Memoirs of the American Mathematical Society. American Mathematical Society, 2000.

[10] L. Hörmander. *The Analysis of Linear Partial Differential Operators I: Distribution Theory and Fourier Analysis*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 1990.

[11] Tuomas Hytönen, Jan van Neerven, Mark Veraar, and Lutz Weis. *Analysis in Banach spaces*. Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer, Cham, 2016.

[12] Yves Meyer. *Wavelets and operators*, volume 137 of *London Math. Soc. Lecture Note Ser.* Cambridge Univ. Press, Cambridge, 1989.

[13] E.M. Stein and G.L. Weiss. *Introduction to Fourier Analysis on Euclidean Spaces*. Mathematical Series. Princeton University Press, 1971.