Università di Pisa

Faculty of Mathematical, Physical, and Natural Sciences
Master in Mathematics

Master Thesis

Multi-parameter Singular Integrals: Product and Flag Kernels

Candidate:
Gennady N. Uraltsev

Advisor:
Prof. Fulvio Ricci

Academic Year 2012/2013

Contents

Summary

This thesis is concerned with the study of multi-parameter singular integrals on the Euclidean space. The Schwartz Kernel Theorem states that translation invariant continuous linear operators with minimal smoothness conditions are convolution operators. Singular integral operator theory is concerned with the study of the singular kernels associated with such operators. A well developed theory exists for the class of Calderón-Zygmund operators and the associated kernels. This kind of kernels can be seen as a natural generalization of the Hilbert kernel on $\mathbb{R}$, of the Riesz kernels in $\mathbb{R}^n$, and, more generally, of kernels of homogeneous degree $-n$ in $\mathbb{R}^n$. Calderón-Zygmund theory is a one-parameter homogeneous theory since the kernels of interest are well-behaved with respect to a family of homogeneous dilation with one parameter. Calderón-Zygmund kernels arise from many problems in linear PDEs and complex analysis. L^p boundedness for $p \in (1, +\infty)$ and stability under composition are well known results for such kernels.

Product-type kernels arise naturally in analysis in several complex variables and PDEs. As a matter of fact joint spectral functional calculus for more than one differential operator naturally produce to product structures. Product spaces occur naturally in the heat equation or in the Shrödinger equation.

Abstractly, product kernels are the result of the extension of Calderón-Zygmund theory to product spaces. The tensor product of two or more Calderón-Zygmund kernels gives a singular kernel defined on the product space. The new kernel has a singularity not only in the origin but also along all coordinate sub-spaces. From the point of view of the associated operators, the tensor product corresponds to the composition of the original operators acting independently on the coordinates of the product space. Product kernel theory aims to extend the space of tensor products of Calderón-Zygmund kernels to a suitably defined completion. This is done mainly by using multi-parameter dilation techniques, with one parameter for each factor of the product space. An other idea that is pursued is that product theory can be inspired by vector valued functional analysis and integration. While avoiding a too abstract approach to such functional analysis in this thesis, some ideas are shown to be very useful.

This thesis illustrates the adaptation of some important results inspired by Calderón-Zygmund theory to product kernels. These include decomposing a kernel into a multi-parameter dyadic series of homogeneous dilates of smooth functions concentrated on essentially disjoint scales and, conversely, finding conditions when such dyadic sums converge to product kernels. Furthermore, since tensor products of bounded operators on L^p remain bounded on L^p one can suppose that this remains true for general product kernels. However, the proof usually used for Calderón-Zygmund operators does not seem to be generalized to product kernels since weak $L^1 - L^1_w$ boundedness fails. A finer technique based on product square function estimates and product Littlewood-Payley theory is developed to solve this problem. This idea is based on the quasi-orthogonality of the dyadic decomposition for the kernels.

The second part of this thesis deals with a certain sub-class of product kernels given by flag kernels. While product kernels are the most intuitive generalization of Calderón-Zygmund

theory to a multi-parameter setting, the singularities are generally too many to work with directly. Flag kernels have singularities concentrated on a flag or filtration of the space, and not along all coordinate subspaces. Kernels with such an ordered structure of singularities appear more often from concrete problems than general product kernels. A multi-parameter theory for flag kernels similar to the one for product kernels is developed. We also show that even though flag kernels form a sub-class of product kernels any product kernel can be written as a sum of flag kernels adapted to different flags.

These results were already present in literature. Flag kernels were introduced by Nagel, Ricci, and Stein in "Singular integrals with flag kernels and analysis on quadratic CR manifolds", *Journal of Functional Analysis*, 2001. A large portion of the above paper is dedicated to applications of product-type singular integral operators. Here we develop the results and provide detailed proofs based on the ideas contained in the part of that paper dedicated to the general theory of flag kernels.

In this thesis we also establish several new results. While the question of whether changes of variables conserve product and flag kernels will be addressed in a forthcoming paper by Alexander Nagel, Fulvio Ricci, Elias Stein, and Richard Wainger, they deal only with polynomial changes of variable. We show that the classes of Calderón-Zygmund, product and flag kernels with compact support are stable with respect to generic smooth changes of variable that have the geometric property of fixing the singular subspaces. These results and the techniques we use can be the first step to studying product-type singular integral operators on manifolds.

Furthermore an attempt is made to develop a basic functional calculus for product singular integrals with respect to derivation, multiplication and convolution. This is done by introducing kernels of generic pseudo-differential order. We establish some useful facts but show that some properties may fail except for a restricted range of pseudo-differential orders.

Finally we show how this functional calculus can be used to establish an approximation result for kernels composed with changes of variable.

Chapter 1

Introduction

1.1 Singular integrals: generalities

The study of singular integral operators, in most generality comes from the Schwartz Kernel Theorem.

Theorem 1.1.1 (Schwartz kernel Theorem)**.**
Let $T : S(\mathbb{R}^n) \to S'(\mathbb{R}^m)$ be a linear continuous map. Then T has an associated distribution $K \in S'(\mathbb{R}^n \times \mathbb{R}^m)$ such that for any $\varphi \in S(\mathbb{R}^n)$ the distribution $T\varphi$ acts in the following way:

$$\langle T\varphi; \psi\rangle = \langle K; \varphi \otimes \psi\rangle$$

for all $\psi \in S(\mathbb{R}^m)$. Such a distribution K is called the kernel of T. Vice-versa any kernel $K \in S'(\mathbb{R}^n \times \mathbb{R}^m)$ defines a continuous linear map from $S(\mathbb{R}^n)$ to $S'(\mathbb{R}^m)$ by the above relation.

It often occurs that one is presented with a linear map T whose functional properties are not known. If a very weak type of continuity on the map can be assumed, in particular the component-wise continuity of the bilinear map

$$(\varphi, \psi) \mapsto \langle T\varphi; \psi\rangle$$

on $S(\mathbb{R}^n) \times S(\mathbb{R}^m)$, then by using Theorem 1.1.1 one can get the existence of the associated kernel K and then try to get a priori estimates by studying the kernel directly. This kind of formulation occurs as the natural generalization of many problems in PDEs, complex analysis, and spectral theory.

In spectral theory, we can consider the following easily stated problem. Consider the self adjoint operator $i\frac{\mathrm{d}}{\mathrm{d}x}$ on $H^1(\mathbb{R})$ and the associated spectral measure μ. Any Borel, a.e. finite function m on $\mathbb{R}$ defines, by functional calculus, a possibly unbounded operator on

$$m\left(i\frac{\mathrm{d}}{\mathrm{d}x}\right) \stackrel{\text{def}}{=} \int_{\mathbb{R}} m(\lambda)\mathrm{d}\mu(\lambda)$$

on $L^2(\mathbb{R})$. Using the Fourier transform one can prove that if m has some minimal regularity then $m(i\frac{\mathrm{d}}{\mathrm{d}x})$ acts by multiplication on the Fourier transform side. We have

$$m\left(i\frac{\mathrm{d}}{\mathrm{d}x}\right)\varphi = \mathcal{F}^{-1}\left(m(\xi)\widehat{\varphi}(\xi)\right)$$

and this holds for those m for which the right hand side is well defined. The hypothesis of Theorem 1.1.1 are satisfied so $m\left(i\frac{\mathrm{d}}{\mathrm{d}x}\right)$ has an associated kernel. More in general, since $i\frac{\mathrm{d}}{\mathrm{d}x}$

commutes with translations, $m\left(i\partial_x\right)$ also commutes, and by an extension of Theorem 1.1.1 the kernel is determined by the diagonal and so $m\left(i\frac{\mathrm{d}}{\mathrm{d}x}\right)$ is a convolution operator

$$m\left(i\frac{\mathrm{d}}{\mathrm{d}x}\right)\varphi = \varphi * \widetilde{K}$$

for some $\widetilde{K} \in S'(\mathbb{R}^n)$. It is fairly easy to see when $m\left(i\frac{\mathrm{d}}{\mathrm{d}x}\right)$ is bounded on $L^2(\mathbb{R})$. Both spectral theory and Fourier theory tell us that we have boundedness if and only if $m \in L^\infty(\mathbb{R})$. A question that naturally arises is when do we have boundedness for other functional spaces like $L^p(\mathbb{R})$ with $p \neq 2$.

A fairly basic question that arises in PDEs and also has a natural formulation in terms of singular integrals comes from the definition of Sobolev spaces. By definition, $W^{m,p}(\mathbb{R}^n)$ is the space of functions in $L^p(\mathbb{R}^n)$ such that all the weak derivatives of order up to n are also in $L^p(\mathbb{R}^n)$. A norm on this space is given by

$$\|f\|_{W^{m,p}(\mathbb{R}^n)} \stackrel{\text{def}}{=} \sum_{|\alpha|\leq m} \|\partial_x^\alpha f\|_{L^p(\mathbb{R}^n)}.$$

But is it sufficient to control L^p norm of f and the L^p norm of its the highest order derivatives to control the $W^{m,p}(\mathbb{R}^n)$ norm? In particular if $m = 2k$ is the norm

$$\|f\|_{L^p(\mathbb{R}^n)} + \left\|\Delta^k f\right\|_{L^p(\mathbb{R}^n)}$$

equivalent to the $W^{m,p}(\mathbb{R}^n)$ norm? It is evident that the latter is coarser. So we need to see if the identity mapping is bounded from the latter norm to $W^{m,p}(\mathbb{R}^n)$. Suppose that $f \in S(\mathbb{R}^n)$. Taking the Fourier transform we ask if $\mathcal{F}^{-1}\left(|\xi|^{2k}\widehat{f}(\xi)\right)$ and f are in L^p then can we deduce that $\mathcal{F}^{-1}(\xi^\alpha \widehat{f}(\xi))$ is in L^p for $|\alpha| \leq m$? This would follow from the boundedness of the convolution operator $f \mapsto f * K$ with K such that

$$\widehat{K} = \frac{(i\xi)^\alpha}{1 + |\xi|^m}.$$

Finally, singular integrals arise frequently in complex analysis. The most well known example is the Hilbert transform on $\mathbb{R}$ given by the convolution with the kernel $\mathrm{PV}\,{}^1\!/_x$. Suppose that we have a real function $f(x) \in C_c^\infty(\mathbb{R})$. There is a unique extension to a harmonic function on the upper half plane with f as its boundary condition. Such a function u is given by convolution with the Poisson kernel:

$$u(x,t) = f * P_t(x) = \frac{1}{\pi}\int_{-\infty}^{+\infty} f(s)\frac{t}{(x-s)^2 + t^2}\mathrm{d}s$$

where

$$P_t(x) = \frac{1}{\pi}\frac{t}{x^2 + t^2}.$$

To study when the boundary value of $u(x,t)$ as $t \to 0$ is f in some adequate sense it is sufficient to study the kernel P_t. As $t \to 0$ it is an approximate identity and so for $u \in L^p(\mathbb{R})$ for all $p \in [1, +\infty)$ we have norm and poinwise a.e. convergence. A natural problem that arises if f is complex-valued and one interprets the upper half plane as the complex domain $\{z \in \mathbb{C} \mid \Im z > 0\}$ is when there is a holomorphic function $F(z)$ on the upper half plane that has f as its boundary value in some sense. We know that if u is harmonic there is a unique harmonic conjugate function $v(z)$ such that $u + iv$ is holomorphic on the upper half plane. Since the holomorphic extension to the upper half plane is unique and is given by the Cauchy integral formula

$$F(z_0) = \frac{i}{\pi} \int_{-\infty}^{+\infty} \frac{f(z)}{z_0 - z} \mathrm{d}z$$

we have that the harmonic conjugate of $u(z)$ is $v(z)$ that is given by the convolution with the conjugate Poisson kernel Q_t such that

$$P_t + iQ_t(x) = \frac{i}{\pi(x + it)}$$

so

$$Q_t(x) = \frac{1}{\pi} \frac{x}{x^2 + t^2}.$$

However Q_t is not integrable. For $t \to 0$ it is not an approximate identity but it tends to the Hilbert kernel

$$Q_t \to \mathrm{PV}\, \frac{1}{\pi x} \qquad \text{as } t \to 0.$$

For f to have a holomorphic extension to the upper half plane it is thus necessary that the $\mathcal{H}\ \Re f = -\Im f$ where $\mathcal{H}$ is the Hilbert transform, that is given by the convolution with the Hilbert kernel $1/\pi x$. The Hilbert kernel is the test case and historically the most important singular integral operator. Most of the theory illustrated in this thesis deals with kernels that have properties that are generalizations of the ones that are typical of the Hilbert kernel.

We have made some examples of singular integral operators. Usually one cannot study this kind of operators in a completely abstract way without any hypothesis on K. We usually require some kind of additional regularity conditions on K. We usually suppose that K at least coincides with a L^1_{loc} function on a "large" set, for example on $\mathbb{R}^n \setminus \{0\}$. In this thesis we will do a brief overview of the classical translation invariant Calderón-Zygmund theory showing how the Calderón-Zygmund kernels are well behaved with respect to dilation structure of $\mathbb{R}^n$. The thesis will concentrate on developing a multi-parameter generalization of this theory by introducing product spaces and generalizing tensor products of Calderón-Zygmund kernels. We will prove a boundedness result similar to the one available for Calderón -Zygmund kernels and illustrate some of the difficulties that arise when developing a functional calculus for such kernels. Finally we will present an original result about the stability of the above classes with respect to changes of variable with certain geometric properties.

1.2 Calderón-Zygmund theory: motivation

The Hilbert kernel has a set of very important properties. It is a homogeneous kernel of degree -1, that is minus the dimension of the space. For a distribution to be homogeneous it is in general not sufficient for it to coincide with a homogeneous function outside the origin. As a matter of fact, the Hilbert kernel is homogeneous of degree -1 because for any test function φ we have that

$$\left\langle \mathrm{PV}\, \frac{1}{\pi x}; \varphi(Rx) \right\rangle = \left\langle \mathrm{PV}\, \frac{1}{\pi x}; \varphi(x) \right\rangle.$$

The Hilbert transform is also smooth away from the origin. Finally it is an odd kernel and this provides for a "cancellation condition". As a matter of fact $\frac{1}{x}$ is not integrable neither in 0 nor at ∞ but since the kernel is odd

$$\lim_{\varepsilon \to 0} \int_{\varepsilon < |x| < \varepsilon^{-1}} \frac{1}{\pi x} \varphi(x)$$

exists for any bounded function with at least Hölder regularity. It is easy to see that $\mathrm{PV}\, {}^{1}\!/_{\pi x}$ and δ_0 are the only kernel on $\mathbb{R}$ with the above properties (homogeneity, cancellation). However this does provide us with an idea of how to generalize this type of kernel.

On $\mathbb{R}^n$ a kernel of homogeneous dimension $-n$ is given by $K = \mathrm{PV}\, {}^{1}\!/_{|x|^n}\Omega(\omega) + c\delta_0$ with $c \in \mathbb{C}$, where Ω is a distribution on the unit sphere S^{n-1} and

$$\left\langle \mathrm{PV}\, \frac{1}{|x|^n}\Omega(\omega);\, \varphi(x) \right\rangle = \lim_{\varepsilon \to 0^+} \int_{r>\varepsilon} \frac{1}{r} \left\langle \Omega(\omega);\, \varphi(r\omega) \right\rangle_{S^{n-1}} \mathrm{d}r.$$

The above expression is homogeneous of degree $-n$ however it is not always well-defined. In particular the limit on the right hand side exists only if

$$\langle \Omega;\, 1 \rangle_{S^{n-1}} = 0.$$

Some examples of often encountered distributions of homogeneous degree $-n$ are the Riesz kernels

$$\mathcal{R}_k = \mathrm{PV}\, \frac{x_k}{|x|^{n+1}}$$

that are the equivalent to the Hilbert kernel in higher dimension.

However we would like to generalize this class of kernels further. In particular the assumption on the homogeneity of the kernel can be too restrictive. Calderón-Zygmund kernels are a class of kernels that are not, per se, homogeneous but are homogeneous as a class. In particular a homogeneous rescaling of order $-n$ transforms a Calderón-Zygmund kernel into an other Calderón-Zygmund kernel with similar properties.

1.2.1 Homogeneity and dilations on $\mathbb{R}^n$

First of all we will be introducing some basic terminology on the Euclidean spaces on which Calderón-Zygmund kernels are defined and on different homogeneous structures they can posses.

First we define the necessary properties and structures a set must possess for it to be possible to formulate a theory that relies on homogeneity. We will omit most of the proofs that are classical.

Definition 1.2.1 (Quasi-distance)**.**
Let X be a set. A quasi-distance is a mapping $\mathrm{d} : X \times X \to \mathbb{R}^+ \cup \{0\}$ that satisfies the following properties:

Positivity $\mathrm{d}(x, y) \geq 0$ *for all $x, y \in X$;*

Coincidence $\mathrm{d}(x, y) = 0$ *if and only is $x = y$;*

Symmetry $\mathrm{d}(x, y) = \mathrm{d}(y, x)$ *for all $x, y \in X$;*

Relaxed triangle inequality *there exists a $c \geq 1$ such that*

$$\mathrm{d}(x, z) \leq c \big(\mathrm{d}(x, y) + \mathrm{d}(y, z)\big) \tag{1.2.1}$$

for all $x, y, z \in X$.

The lower bound of the admissible constants in (1.2.1) *is called the triangle constant.*

A space (X, d) with a quasi-distance d is said to be a quasi-metric space. The usual notions of generated topology, completeness, etc. apply also to quasi-metric spaces in the expected way.

Definition 1.2.2 (Doubling measures and spaces of homogeneous type).
Let (X, d) *be a quasi-metric space. A measure* μ *on* (X, d) *that satisfies*

$$\mu\left(B(x, 2r)\right) \leq c\mu\left(B(x, r)\right) \tag{1.2.2}$$

for some $c \geq 1$ *and all* $x \in X$ *and* $r > 0$ *is called a doubling measure.* (X, d, μ) *where* (X, d) *is a quasi-metric space endowed with a doubling measure* μ *is called a space of homogeneous type.*

We now illustrate the natural way to endow the Euclidean space with such a structure of a homogeneous space.

Definition 1.2.3 (Non-isotropic family of dilations).
Consider the d-dimensional Euclidean space $\mathbb{R}^d$*. Given some exponents* $\lambda_1, \dots \lambda_d \in \mathbb{N}$*, a family of linear mappings*

$$r \cdot (x_1, \dots, x_d) \mapsto (r^{\lambda_1} x_1, \dots, r^{\lambda_d} x_d)$$

with $r \in \mathbb{R}^+$ *is called a non-isotropic family of dilations on* X*. We have that such that for any* $r_1, r_2 \in \mathbb{R}^+$

$$r_1 \cdot \left(r_2 \cdot (x_1, \dots, x_d)\right) = r_1 r_2 \cdot (x_1, \dots, x_d)$$

A vector $v \in \mathbb{R}^n$ *such that* $(r \cdot v)_j = r^\lambda v_j$ *for some* $\lambda \in \mathbb{N}$ *for all* $r > 0$ *and for all coordinates* $j \in \{1, \dots, d\}$ *is called an eigen-vector of the family of dilations*

Definition 1.2.4 (Homogeneous norms).
A homogeneous norm on R^d *is a function* $\|\cdot\| : X \to \mathbb{R}^+ \cup \{0\}$ *that satisfies the following properties:*

Continuity $\|\cdot\|$ *is continuous on* R^d*;*

Coincidence $\|x\| = 0$ *if and only if* $x = 0$*;*

Symmety $\|x\| = \|-x\|$ *for all* $x \in \mathbb{R}^d$*;*

Homogeneity $\|r \cdot x\| = r\|x\|$ *for all* $x \in \mathbb{R}^d$ *and* $r \in \mathbb{R}^+ \cup \{0\}$*.*

Proposition 1.2.5 (Properties of homogeneous norms).

1. *Given a non-isotropic family of dilations there exists a compatible homogeneous norm. Furthermore there exists a choice of a homogeneous norm that is smooth away from the origin.*

2. *A homogeneous norm satisfies*

$$\|x - z\| \leq c\left(\|x - y\| + \|y - z\|\right)$$

   for some $c > 1$*. The function*

$$\mathrm{d}(x, y) = \|x - y\|$$

   is a semi-distance.

3. *The homogeneous norm and the induced semi-distance determines a topology on* $\mathbb{R}^n$ *that coincides with the standard Euclidean topology.*

4. *The Lebesgue measure on* $\mathbb{R}^n$ *is doubling with respect to the semi-distance. In particular*

$$\mathcal{L}^n(B(x, 2r)) = 2^q \mathcal{L}^n(B(x, r))$$

   for any $x \in \mathbb{R}^n$ *and* $r > 0$*.* $q = \sum_{j=1}^n \lambda_i$ *is called the homogeneous dimension of* $\mathbb{R}^n$*.*

The above properties show that $\mathbb{R}^n$, endowed with a non-isotropic family of dilations and the corresponding homogeneous norm, is a space of homogeneous type.

From now on $\|x\|$ will indicate a certain (smooth) homogeneous norm, $r \cdot x$ will indicate the action of the dilation with parameter $r > 0$ on $x \in \mathbb{R}^n$. $|x|$ indicates the standard Euclidean norm of x. Dealing with a homogeneous theory we will recur heavily to dilations of functions. Given a function f on $\mathbb{R}^n$, for $R \in \mathbb{R}^+$ we define the scaled function in the following way:

$$f^{(R)}(x) \stackrel{\text{def}}{=} R^{-q} f(R^{-1} \cdot x).$$

This rescaling dilates the function f by a factor R but it maintains its L^1 norm and its integral. Consequently we say that a distribution $K \in D'(\mathbb{R}^n)$ is homogeneous of order μ if

$$\left\langle K;\, \varphi^{(r)} \right\rangle = r^{\mu} \left\langle K;\, \varphi \right\rangle$$

for any test function φ. If we choose a multi-index of derivation $\alpha \in \mathbb{N}^n$ then we indicate by $|\alpha| = \sum_{j=1}^n \alpha_j$ its standard length and by $\|\alpha\| = \sum_{j=1}^n \lambda_j \alpha_j$ its homogeneous length. The rationale for this notation is given by the fact that we have

$$\partial_x^\alpha \varphi^{(r)}(x) = r^{-\|\alpha\|} (\partial_x^\alpha \varphi)^{(r)}(x)$$

We also introduce functional spaces that take in account the non-isotropic family of dilations present on $\mathbb{R}^n$. While classical Sobolev spaces treat derivation in each direction in the same way we can generalize Sobolev spaces to account for different dilation exponents along different directions. It is useful to notice that a non-isotropic family of dilations on $\mathbb{R}^n$ naturally induces a non-isotropic family of dilations on the dual space $(\mathbb{R}^n)^* \simeq (\mathbb{R}^n)$. Let $\xi \in (\mathbb{R}^n)^*$ then $(R \cdot \xi)(x) = \xi(R \cdot x)$.

On $\mathbb{R}^n$ endowed with a non-isotropic family of dilations we define the differential operator $\mathcal{L}$ by setting

$$-\mathcal{L}\varphi(x) \stackrel{\text{def}}{=} \mathcal{F}^{-1}\left(\|\xi\|^2 \, \widehat{\varphi}(\xi) \right)$$

where we have chosen $\|x\|$ and thus $\|\xi\|$ to be a smooth homogeneous norm. $\mathcal{L}$ is a densely defined self-adjoint negative operator on L^2.

Definition 1.2.6 (Non-isotropic Sobolev spaces)**.**
Let $f \in L^2(\mathbb{R}^n)$. For $s \geq 0$ we say that $f \in \mathcal{H}^s(\mathbb{R}^n)$ if $\widehat{f}$ is locally integrable and

$$\int_{\mathbb{R}^n} \left(1 + \|\xi\|^2\right)^s |\widehat{f}(\xi)|^2 \mathrm{d}\xi \leq \infty.$$

Furthermore $\mathcal{H}^s$ is a Hilbert space with the norm given by

$$\|f\|^2_{\mathcal{H}^s} = \int_{\mathbb{R}^n} \left(1 + \|\xi\|^2\right)^s |\widehat{f}(\xi)|^2 \mathrm{d}\xi.$$

For $s < 0$ the spaces $\mathcal{H}^s$ are defined as the duals to $\mathcal{H}^{-s}$. They are given by the distributions $f \in S'(\mathbb{R}^n)$ such that $\widehat{f}$ is locally integrable and

$$\int_{\mathbb{R}^n} \left(1 + \|\xi\|^2\right)^{-|s|} |\widehat{f}(\xi)|^2 \mathrm{d}\xi \leq \infty.$$

The space $\mathcal{H}^s$ is the domain of the self adjoint operator $(-\mathcal{L})^{s/2}$.

We finally define cutoff functions, that turn out to be very useful in the study of singular integrals.

Definition 1.2.7 (Cutoff function).
Let $\Xi \in C_c^\infty\big([0,1)\big)$ *such that* $0 \le \Xi(t) \le 1$ *and* $\Xi(t) = 1$ *for* $t \in [0, 1/2]$. *A cutoff function* η *on* $\mathbb{R}^n$ *is*

$$\eta(x) \stackrel{def}{=} \Xi(\|x\|). \tag{1.2.3}$$

From now on, with an abuse of notation, we will indicate the distribution - test function duality pairing in the following way:

$$\int_{\mathbb{R}^n} K(x)\varphi(x)\mathrm{d}x \stackrel{\text{def}}{=} \langle K;\, \varphi\rangle$$

1.2.2 Definition of Calderón-Zygmund kernels

We will now see the definition of Calderón-Zygmund kernels. Low regularity theory is beyond the scope and interest of this thesis so we will define the class of smooth Calderón-Zygmund kernels. In literature the definition is often given in a different manner and involves much fewer smoothness assumptions. Furthermore, sometimes it is asked that the operator associated with a Calderón-Zygmund kernel K be L^2 bounded. For convolution operators this is equivalent to asking the boundedness of the multiplier $\widehat{K}$. However we avoid this approach and characterize smooth Calderón-Zygmund kernel directly in terms of the cancellation property that is expressed as the boundedness of K on dilates of a certain class of test functions.

Definition 1.2.8 (Normalized bump function).
A smooth function $\varphi \in C_c^\infty(\mathbb{R}^n)$ *is a b-normalized bump function if it is supported on the unit ball* $B_{\mathbb{R}^n}(0,1)$ *of* $\mathbb{R}^n$ *and is normalized with respect to the* C^b *norm i.e.* $\|\varphi\|_{C^b} < 1$.

Definition 1.2.9 (Calderón-Zygmund kernel).
Consider the Euclidean space $\mathbb{R}^n$ *endowed with a family of dilations and with the associated homogeneous norm. Let* $K \in S'(\mathbb{R}^n)$. K *is in the class* $\mathcal{CZ}$ *of Calderón-Zygmund kernels if, away from the origin,* K *coincides with a smooth function i.e.*

$$K\big|_{\mathbb{R}^n\setminus\{0\}} \in C^\infty\big(\mathbb{R}^n \setminus \{0\}\big)$$

and it satisfies the following two kinds of conditions:

Size conditions

$$\left|\partial_x^\alpha K(x)\right| \le C_\alpha \|x\|^{-q-\|\alpha\|} \tag{1.2.4}$$

for all $x \neq 0$ *and for any multi-index* α.

Cancellation condition *For all* $R > 0$ *and all b-normalized bump functions* φ *there is a constant* $C_c \ge$ *so that the inequality*

$$\left|\int_{\mathbb{R}^n} K(x)\varphi(R^{-1}\cdot x)\mathrm{d}x\right| < C_c \tag{1.2.5}$$

holds. It can be shown that the definition does not depend on the order of normalization b *as long as* $b \ge 1$. *For a proof in the more general case of product kernels refer to Proposition 2.4.11 in this thesis.*

We say that a family of kernels in $\mathcal{CZ}$ *is uniformly bounded if there is a choice of constants so that the inequalities* (1.2.4) *and* (1.2.5) *hold uniformly for the whole family.*

We can easily see from the definition that the class is homogeneous of degree $-q$ in the sense that if $K \in \mathcal{CZ}$, then for $r > 0$ the kernels

$$K^{(r)}(x) = r^{-q}K(r^{-1}\cdot x)$$

are a bounded family in $\mathcal{CZ}$. It suffices to check (1.2.5) and (1.2.4) explicitly.

1.2.3 Fourier transforms of $\mathcal{CZ}$ kernels and the dyadic decompositions

A useful theorem characterizes Fourier transforms of $\mathcal{CZ}$ kernels. As a matter of fact the L^2 boundedness usually required in other definitions of Calderón-Zygmund kernels follows from the boundedness of the multipliers of the operators associated with $\mathcal{CZ}$ kernels as defined here. We have the following theorem.

Theorem 1.2.10 (Fourier transforms of $\mathcal{CZ}$ kernels).
Let $K \in \mathcal{CZ}$ be a kernel. Then its multiplier $\widehat{K}$ is a smooth Mihlin-Hörmander multiplier i.e.

$$\widehat{K} \in L^\infty \cap C^\infty(\mathbb{R}^n \setminus \{0\})$$

and

$$\left|\partial_\xi^\alpha \widehat{K}(\xi)\right| < C_\alpha \|\xi\|^{-\|\alpha\|}$$

for $\xi \neq 0$ and any multi-index α.

Vice-versa the inverse Fourier transform of a smooth Mihlin-Hörmander multiplier is a $\mathcal{CZ}$ kernel.

The homogeneity of the class of Calderón-Zygmund kernels is also reflected in the property of the dyadic decomposition and convergence of dyadic sums.

Theorem 1.2.11 (Dyadic sums).
Let $\{\varphi_i\}_{i\in\mathbb{Z}} \subset S(\mathbb{R}^n)$ be a uniformly bounded sequence of Schwartz functions with zero mean i.e.

$$\int_{\mathbb{R}^n} \varphi_i(x)\mathrm{d}x = 0.$$

Then the dyadic sum

$$\sum_{i\in Z} \varphi_i^{(2^i)}(x) = \sum_{i\in Z} 2^{-iq}\varphi_i(2^{-i}\cdot x)$$

converges in the sense of distributions to a $\mathcal{CZ}$ kernel. Furthermore the partial sums of the series are all uniformly bounded in $\mathcal{CZ}$.

This theorem can be proven by taking the Fourier transform of the series and verifying, via size estimates, that the sum converges to a Mihlin-Hörmander multiplier.

1.2.4 Calderón-Zygmund operators

Calderón-Zygmund operators, the convolution operators associated with $\mathcal{CZ}$ kernels, are bounded on L^p for all $p > 1$. Theorem 1.2.10 provides us with the boundedness of the multiplier associated with a kernel in $\mathcal{CZ}$ so we have L^2 boundedness of the associated operator.

For L^1 there is no strong boundedness result. However we can introduce the space L^1_w.

Definition 1.2.12 (L^1_w space).
The space $L^1_w\left((X,\mu)\right)$ is the space of those measurable functions f for which the following quasi-norm is finite

$$\|f\|_{L^1_w} = \sup_{t>0} t^{-1}\mu\left(\{|f| > t\}\right).$$

We say that an operator $\mathcal{T}$ is $L^1 - L^1_w$ bounded if there exists a constant C such that for all $f \in L^1(X)$

$$\|\mathcal{T}f\|_{L^1_w} \leq C\,\|f\|_{L^1}$$

We can prove $L^1 - L^1_w$ boundedness for $\mathcal{CZ}$ operators using the Calderón-Zygmund decomposition (see [Duo01, Chapter 5]). Strong L^p boundedness for $p \in (1, \infty)$ is obtained by the use of interpolation and duality theorems.

Theorem 1.2.13 (L^p boundedness of Calderón-Zygmund operators)**.**
Let $K \in \mathcal{CZ}$ and let $\mathcal{T}$ be the convolution operator associated to K:

$$\mathcal{T}\varphi = \varphi * K$$

for all $\varphi \in S(\mathbb{R}^n)$. Then the following inequality holds for all $p \in (1; +\infty)$:

$$\|\mathcal{T}\varphi\|_{L^p} < C_p \|\varphi\|_{L^p}$$

and T admits a unique extension to a bounded operator on $L^p(\mathbb{R}^n)$.

1.2.5 Kernels of arbitrary order

It is easy to see that the operators associated to kernels in $\mathcal{CZ}$ admit an extension to bounded operators on $\mathcal{H}^s$ for any $s \in \mathbb{R}$. For this reason it makes sense to define the class of distributional derivatives of such kernels. As a matter of fact for $s \in \{z \in \mathbb{C} \mid \Re z > -q\}$ the expression $(-\mathcal{L})^{s/2}$ defines an unbounded operator on L^2 that is well defined for $\varphi \in S(\mathbb{R}^n)$. By duality we can define the operators $(-\mathcal{L})^{s/2}K$ so that for all $\varphi \in S(\mathbb{R}^n)$ we have

$$\int_{\mathbb{R}^n} \left((-\mathcal{L})^{s/2} K(x)\right) \varphi(x) \mathrm{d}x \stackrel{\text{def}}{=} \int_{\mathbb{R}^n} K(x) \left((-\mathcal{L})^{s/2} \varphi(x)\right) \mathrm{d}x.$$

Using this property we can define the class $\mathcal{CZ}(\nu)$ of Calderón-Zygmund kernels of orders $\nu > -q$, $\nu \in \mathbb{R}$, by imposing that $K \in \mathcal{CZ}(\nu)$ if and only if there exists a kernel $\widetilde{K} \in \mathcal{CZ}$ such that $K = (-\mathcal{L})^{\nu/2}\widetilde{K}$.

It can be verified that the following is an equivalent definition for $\mathcal{CZ}(\nu)$ if one supposes that $\nu > -q$.

Definition 1.2.14 (Calderón-Zygmund kernels of non-zero order)**.**
Consider the Euclidean space $\mathbb{R}^n$ endowed with a family of dilations and with the associated homogeneous norm. Let $K \in S'(\mathbb{R}^n)$. We say that K is in the class $\mathcal{CZ}(\nu)$ of Calderón-Zygmund kernels if, away from the origin, K coincides with a smooth function i.e.

$$K\big|_{\mathbb{R}^n \setminus \{0\}} \in C^\infty\left(\mathbb{R}^N \setminus \{0\}\right)$$

and it satisfies the following two kinds of conditions:

Size conditions

$$\left|\partial_x^\alpha K(x)\right| \le C_\alpha \|x\|^{-q-\nu-\|\alpha\|} \tag{1.2.6}$$

for all $x \neq 0$ and for any multi-index α.

Cancellation condition *For all $R > 0$ and all b-normalized bump functions φ the inequality*

$$R^\nu \left|\int_{\mathbb{R}^n} K(x)\varphi(R^{-1} \cdot x)\mathrm{d}x\right| < C_c \tag{1.2.7}$$

It can be shown that the definition does not depend on the order of normalization b as long as $b > \nu$. The proof is given in a more general case in Proposition 2.4.11.

We say that a family of kernels in $\mathcal{CZ}(\nu)$ is uniformly bounded if there is a choice of constants so that the inequalities (1.2.6) and (1.2.7) hold uniformly for the whole family.

Furthermore it can be checked that, for $-q < \nu < 0$, if $K \in \mathcal{CZ}(\nu)$ then $\widehat{K} \in \mathcal{CZ}(-q-\nu)$. Furthermore, dropping the condition $\nu > -q$ in Definition 1.2.14 allows us to define the classes $\mathcal{CZ}(\nu)$ for all $\nu \in \mathbb{R}$ and the property of the Fourier transform still holds.

Theorem 1.2.15 (Fourier transforms of $\mathcal{CZ}$ kernels of arbitrary order)**.**
For any $\nu \in \mathbb{R}$ the Fourier transform maps $\mathcal{CZ}(\nu) \mapsto \mathcal{CZ}(-q-\nu)$.

Notice however that by defining $\mathcal{CZ}(\nu)$ like this it is in general not true that $K \in \mathcal{CZ}(\nu) \Rightarrow (-\mathcal{L})^{s/2} K \in \mathcal{CZ}(\nu+s)$. In particular the implication can fails if either $\nu \leq -q$ and $\nu + \Re s \leq -q$. In any case the above expression is well defined only for $\Re s > -q$.

For a certain range of orders $\nu \in \mathbb{R}$ we can establish some useful boundedness properties of the associated convolution operators.

Theorem 1.2.16 (Boundedness on Sobolev space spaces)**.**
Let $K \in \mathcal{CZ}(\nu)$ with $\nu \geq 0$. Then the convolution operator $\mathcal{T}$ associated to K extends to a bounded operator

$$\mathcal{T} : \mathcal{H}^s(\mathbb{R}^n) \to \mathcal{H}^{s-\nu}(\mathbb{R}^n).$$

for any $s \in \mathbb{R}$.

Theorem 1.2.17 (Convolution algebra)**.**
Let $K_1 \in \mathcal{CZ}(\nu)$ and $K_2 \in \mathcal{CZ}(\mu)$ with $\nu, \mu > -q$ and $\nu + \mu > -q$. Then the convolution of the two kernels is well defined. If $\mathcal{T}_1$ and $\mathcal{T}_2$ are the convolution operators associated to K_1 and to K_2 the the composition operator $\mathcal{T}_1 \circ \mathcal{T}_2$ is well defined on $S(\mathbb{R}^N)$ and the kernel of the associated operator is $K_1 * K_2$.

We will prove these last two properties in the more general case of product kernels in Sections 2.9 and 2.10.

Chapter 2

Product kernels

2.1 Product spaces

Product type operators and distributions rely on the properties of $\mathbb{R}^N$ considered as a product space. This section is dedicated to some basic properties and notation necessary to start developing a theory of product and flag singular integrals.

From the most general point of view we can consider two sigma-finite measure spaces $(X, \mathscr{X}, \mu_X)$ and $(Y, \mathscr{Y}, \mu_Y)$. Classical measure theory provides us with the possibility of defining a sigma algebra and a measure on the space $X \times Y$ compatible with the product structure. The following is a consequence of Carathéodory's extension and uniqueness theorems.

Corollary 2.1.1.
There exists a unique measure $\mu_X \otimes \mu_Y$ on $(X \times Y, \mathscr{X} \otimes \mathscr{Y})$ defined on the sigma algebra $\mathscr{X} \otimes \mathscr{Y}$ such that for any rectangular sets $A \times B \subset X \times Y$ with $A \in \mathscr{X}$ and $B \in \mathscr{Y}$ we have that $\mu_X \otimes \mu_Y(A \times B) = \mu_X(A)\mu_Y(B)$.

The sigma algebra $\mathscr{X} \otimes \mathscr{Y}$ is the sigma algebra generated by the rectangular sets $\{A \times B \mid A \in \mathscr{X},\ B \in \mathscr{Y}\}$.

We also recall the famous Fubini-Tonelli theorem that concerns product measure spaces.

Theorem 2.1.2 (Fubini-Tonelli)**.**
Let $f : X \times Y \to \mathbb{R}$ be a measurable function. For any $x \in X$ the function $y \mapsto f(x,y)$ is measurable. The same holds for any $y \in Y$ and the function $x \mapsto f(x,y)$. Furthermore if f is non-negative or integrable then the functions

$$x \mapsto \int_Y f(x,y)\mathrm{d}\mu_Y(y) \qquad\qquad y \mapsto \int_X f(x,y)\mathrm{d}\mu_X(x)$$

are measurable and the following equality holds.

$$\int_{X\times Y} f(x,y)\mathrm{d}x\mathrm{d}y = \int_Y \left(\int_X f(x,y)\mathrm{d}x\right)\mathrm{d}y = \int_X \left(\int_Y f(x,y)\mathrm{d}y\right)\mathrm{d}x$$

The proof of the above and an introduction to abstract measure theory can be found in [EG92, Chapter 1].

The study of spaces of functions on product spaces can be carried out using vector-valued Bochner integral theory. The definition and basic properties of the Bochner integrals can be found in [**??**]. In particular we have the following result.

Theorem 2.1.3 (Product L^p spaces).
For any $p \in [1, +\infty]$, the functional space $L^p(X \times Y, \mathscr{X} \otimes \mathscr{Y}, \mu_X \otimes \mu_Y; \mathbb{R})$ is naturally isometric to the vector-valued L^p spaces $L^p(X, \mathscr{X}, \mu_X; L^p(Y, \mathscr{Y}, \mu_Y))$ and $L^p(Y, \mathscr{Y}, \mu_Y; L^p(X, \mathscr{X}, \mu_X))$. Explicitly, the isometries associate the mappings $x \mapsto f(x, \cdot)$ and $y \mapsto f(\cdot, y)$ to a function $f \in L^p(X \times Y)$.

Proof. It suffices to prove that the mappings $x \mapsto f(x, \cdot)$ and $y \mapsto f(\cdot, y)$ are strongly Bochner-measurable. Fubini-Tonelli theorem guarantees that the correspondence is then an isometry. More precisely the linear mapping $f \mapsto \big(x \mapsto f(x, \cdot)\big)$ defined on the functions $f(x, y) \in L^p(X \times Y)$ is such that $\big(x \mapsto f(x, \cdot)\big) \in L^p(X; L^p(Y))$ is a linear surjective isometry since

$$\|f\|^p_{L^p(X \times Y)} = \int_{X \times Y} |f(x, y)|^p \mathrm{d}x\mathrm{d}y = \int_X \|f(x, \cdot)\|^p_{L^p(Y)} \,\mathrm{d}x = \|f\|^p_{L^p(X; L^p(Y))} \,.$$

We must prove that the above mapping is densely defined. Let $f \in L^p(X \times Y)$, then there is a sequence of simple functions $f_n \in L^p(X \times Y)$ that converge poinwise a.e. and in L^p to f. Let us write

$$f_n = \sum_{i=1}^{N_n} a_{n,i} \mathbb{1}_{C_{n,i}}$$

where $a_{n_i} \in \mathbb{R} \setminus \{0\}$ and $C_{n,i}$ are $\mathscr{X} \otimes \mathscr{Y}$ measurable sets that are pairwise disjoint for every n in the sense that for all $n \in \mathbb{N}$ if $i \neq i'$ then $C_{n,i} \cap C_{n,i'} = \emptyset$. However, by Carathéodory's theorem, on $\mathscr{X} \otimes \mathscr{Y}$ the outer measure generated by rectangular sets coincides with $\mu_X \otimes \mu_Y$. This allows us to approximate f with simple functions built up using only rectangular sets.

As a matter of fact for any ε and for any $C_{n,i}$ we can choose a sequence of pairwise disjoint rectangular sets $D_{n,i,j} = A_{n,i,j} \times B_{n,i,j}$ such that

$$\left| \mu_X \otimes \mu_Y(C_{n,i}) - \sum_{j=1}^{M_{n,i}} \mu_X \otimes \mu_Y(D_{n,i,j}) \right| < 2^{-i} \varepsilon |a_{n,i}|^{-1}.$$

Setting

$$\widetilde{f}_n = \sum_{i=1}^{N_n} \sum_{j=1}^{M_{n,i}} a_{n,i} \mathbb{1}_{D_{n,i,j}}$$

we have that

$$\left\| f_n - \widetilde{f}_n \right\|_{L^p(X \times Y)} < \varepsilon.$$

Using a diagonal procedure setting $\varepsilon = \varepsilon_n = 2^{-n}$ we get a sequence of functions $g_n \in L^p(X \times Y)$ converging to f in L^p. The functions g_n are simple and are built using characteristic functions of rectangular sets. In particular, using the Fubini-Tonelli theorem, we get that the mappings $x \to g_n(x, \cdot)$ are simple vector-valued functions in $L^p(X; L^p(Y))$. This concludes the proof. □

All the above properties and proofs are valid by induction for any finite product of measure spaces. If the factors are not merely measure spaces but have additional structures the product space can inherit these structures and present more complex properties. We avoid intermediate steps and we concentrate directly on spaces that are involved in classical Calderón-Zygmund theory i.e. spaces of homogeneous type and especially Euclidean spaces.

2.1.1 Multi-parameter structures on product spaces

In this section we illustrate what we mean by multi-parameter structures on product spaces. The contents herein mainly concern notation.

Suppose $\{H_k\}_{k\in\{1,\dots,d\}}$ is a family of spaces of homogeneous type. The product space $\bigoplus_{k=i}^{d} H_k$ is also a space of homogeneous type. The product measure is doubling and the doubling constant in (1.2.2) for $\bigoplus_{k=1}^{d} H_k$ is bounded from above by $\prod_{k=1}^{d} c_k$, where c_k is the doubling constants for H_k for each k respectively.

Now suppose that each H_k is an Euclidean space. We indicate vectors in H_k by writing $x_k \in H_k$. The topological dimensions of the spaces $H_1,\dots,H_d$ are $n_1,\dots,n_d$ respectively and so $H_k \simeq \mathbb{R}^{n_k}$ and $\bigoplus_{k=1}^{d} H_k \simeq \mathbb{R}^N$ with $N = \sum_{k=1}^{d} n_k$. We indicate vectors in the product space as $\boldsymbol{x} \in \bigoplus_{k=1}^{d} H_k$. We write

$$\boldsymbol{x} = (x_1,\dots,x_d) \quad \text{or} \quad \boldsymbol{x} = x_1 \oplus \dots \oplus x_k. \tag{2.1.1}$$

Let each space H_k be endowed with a (non-isotropic) family of dilations as in Definition 1.2.3. By choosing homogeneous coordinates on H_k for every $k \in \{1,\dots,d\}$ we can write $x_k = (x_{k,1},\dots,x_{k,n_k})$ and $r \cdot x_k = (r^{\lambda_{k,1}} x_{k,1},\dots,r^{\lambda_{k,n_k}} x_{k,n_k})$ where $\lambda_{k,l}$ are the exponents of the dilations relative to the l^{th} coordinate in H_k. The product space $\bigoplus_{k=1}^{d} H_k$ inherits the dilation structure. As a matter of fact a natural family of dilations on $\bigoplus_{k=1}^{d} H_k$ is given by

$$r \cdot \boldsymbol{x} = (r \cdot x_1,\dots,r \cdot x_d).$$

Let q_k be the homogeneous dimensions of H_k given by $q_k = \sum_{l=1}^{n_k} \lambda_{k,l}$. Then the homogeneous dimension of $\bigoplus_{k=1}^{d} H_k$ is $Q = \sum_{k=1}^{d} q_k$ and $x_{k,l}$ with $k \in \{1,\dots,d\}$ and $l \le n_k$ are a set of homogeneous coordinates.

However $\bigoplus_{k=1}^{d} H_k$ has a more complex homogeneous structure. The dilations on each factor are independent so we can introduce a multi-parameter family of dilations by setting

$$\boldsymbol{r} \cdot \boldsymbol{x} \overset{\text{def}}{=} (r_1 \cdot x_1,\dots,r_d \cdot x_d) \tag{2.1.2}$$

for every $\boldsymbol{r} = (r_1,\dots,r_d) \in (\mathbb{R}^+)^d$. In this multi-parameter notation we will indicate the set of topological and homogeneous dimensions of the factors as $\boldsymbol{n} = (n_1,\dots,n_d)$ and $\boldsymbol{q} = (q_1,\dots,q_d)$ respectively. By abuse of notation we will indicate the 96 homogeneous norms on the factors of $\bigoplus_{k=1}^{d} H_k$ by $\|\cdot\|$. It will be clear to which norm we are referring to because of the different vectors they are applied to. Possible equivalent homogeneous norms on $\bigoplus_{k=1}^{d} H_k$ are

$$\|\boldsymbol{x}\| = \|x_1\| + \dots + \|x_d\|$$
$$\|\boldsymbol{x}\| = \max_{k\in\{1,\dots d\}} \|x_k\|$$
$$\|\boldsymbol{x}\| = \left(\|x_1\|^2 + \dots + \|x_d\|^2\right)^{1/2}.$$

The last of these three is smooth if the homogeneous norms on each space are smooth.

It is useful to notice that one can reason conversely. Suppose that the Euclidean space $\mathbb{R}^N$ is endowed with a one-parameter (non-isotropic) family of dilations as in 1.2.3. A multi-parameter, or product space, structure can be introduced by choosing a product decomposition $\mathbb{R}^N = \oplus_{k=1}^{d} H_k$ where $\{H_k\}_{k\in\{1,\dots,d\}}$ are a set of subspaces invariant with respect to the given family of dilations. The dilations on $\mathbb{R}^N$ restrict to each H_k independently and provide each subspace H_k with a family of dilations. These structures in turn provide $\mathbb{R}^N$ with a multi-parameter structure.

Finally, we also use the following notations when dealing with product spaces. It is useful to introduce multi-indexes when working with derivatives on product spaces. In particular if $\alpha_k \in \mathbb{N}^{n_k}$ are multi-indexes on each factor H_k then we write the multi-index

$$\boldsymbol{\alpha} \stackrel{\text{def}}{=} (\alpha_1, \dots, \alpha_d) \tag{2.1.3}$$

where we admit the expressions

$$\boldsymbol{x}^{\boldsymbol{\alpha}} \stackrel{\text{def}}{=} \prod_{k=1}^{d} x_k^{\alpha_k} = \prod_{k=1}^{d} \prod_{j=1}^{n_k} x_{k,j}^{\alpha_{k,j}} \qquad \text{and} \qquad \boldsymbol{\partial}_{\boldsymbol{x}}^{\boldsymbol{\alpha}} \varphi \stackrel{\text{def}}{=} \partial_{x_1}^{\alpha_1} \dots \partial_{x_d}^{\alpha_d} \varphi \tag{2.1.4}$$

for any smooth function φ on $\bigoplus_{k=1}^{d} H_k$.

Let $J \subset \{1, \dots, d\}$ be a subset of indexes relative to a given product structure on $\mathbb{R}^N$. We indicate by H_J the product space $\bigoplus_{j \in J} H_j$ intended both as a subspace of $\bigoplus_{k=1}^{d} H_k \simeq \mathbb{R}^N$ and separately as an abstract vector space. If $\boldsymbol{x} \in \bigoplus_{k=1}^{d} H_k$ then $\boldsymbol{x}_J \stackrel{\text{def}}{=} \pi_{H_J}(\boldsymbol{x})$ where π_{H_J} is the canonical projection from $\bigoplus_{k=1}^{d} H_k$ onto H_J. In this case also $\boldsymbol{x}_J$ can be seen as a vector in $\mathbb{R}^N$ or it can be naturally identified as an abstract vector in H_J. We indicate by $H_{(J)} \stackrel{\text{def}}{=} H_{J^c}$ with $J^c = \{1, \dots, d\} \setminus J$. Once again if we identify H_J and $H_{(J)}$ as subspaces of $\mathbb{R}^N$ then we can write for any $\boldsymbol{x} \in R^N$ the equality $\boldsymbol{x} = \boldsymbol{x}_J \oplus \boldsymbol{x}_{(J)}$. If J consists of only one element j then we have $\boldsymbol{x}_J = x_j$. Analogously we will write $\boldsymbol{x}_{(j)} \stackrel{\text{def}}{=} \boldsymbol{x}_{(J)}$ with $J = \{j\}$. A similar notation applies to multi-indexes. The multi-index $\boldsymbol{\alpha}_J \in \mathbb{N}^{|J|}$ is such that

$$\boldsymbol{\alpha}_{\boldsymbol{J}} = (\alpha_{j_1}, \dots, \alpha_{j_{|J|}}) \tag{2.1.5}$$

with $J = \{j_1, \dots, j_{|J|}\}$. However, by a slight abuse of notation, when we write the multi-index $\boldsymbol{\alpha}_J$ applied to vectors in the product space $\boldsymbol{x}$ we mean

$$\boldsymbol{\alpha}_{Jk} = \begin{cases} \alpha_k & \text{if } k \in J \\ 0 & \text{if } k \notin J \end{cases}$$

so that for example the equality

$$\boldsymbol{x}^{\boldsymbol{\alpha}} = \boldsymbol{x}^{\boldsymbol{\alpha}_J} \boldsymbol{x}^{\boldsymbol{\alpha}_{(J)}}$$

holds. We can also easily define the size of multi-indexes. We have that $\|\boldsymbol{\alpha}\| = \|\alpha_1\| + \ldots + \|\alpha_d\|$ is the homogeneous size of the multi-index $\boldsymbol{\alpha}$ and $|\boldsymbol{\alpha}| = |\alpha_1| + \ldots + |\alpha_d|$ is the standard size.

Dealing with multi-parameter homogeneous theory we recur heavily to dilations of functions. Given a function f on $\bigoplus_{k=1}^{d} \simeq \mathbb{R}^N$, for $\boldsymbol{R} \in (\mathbb{R}^+)^d$ we define the rescaled function in the following way:

$$f^{(\boldsymbol{R})}(\boldsymbol{x}) \stackrel{\text{def}}{=} \boldsymbol{R}^{-\boldsymbol{q}} f(\boldsymbol{R}^{-1} \cdot \boldsymbol{x}) = R_1^{-q_1} \dots R_d^{-q_d} f(R_1^{-1} \cdot x_1, \dots, R_d^{-1} \cdot x_d). \tag{2.1.6}$$

This rescaling dilates the function f by a factor R but it maintains its L^1 norm and its integral.

Product Sobolev spaces

On Euclidean product spaces we also introduce functional spaces that take in account the differentiability properties of functions. Since we are dealing with Euclidean spaces it is straightforward to generalize $\mathcal{H}^s$ Sobolev spaces.

Definition 2.1.4 (Product Sobolev spaces).
Let $\mathbb{R}^N = \bigoplus_{k=1}^d H_k$ be a product space and let $\boldsymbol{s} = (s_1, \ldots, s_d) \in \mathbb{R}^d$ be a multi-order. The product Sobolev space $\mathcal{H}^{\boldsymbol{s}}\left(\bigoplus_{k=1}^d H_k\right)$ is the space of tempered distributions f such that $\widehat{f}$ is locally integrable and has finite $H^{\boldsymbol{s}}$ norm

$$\|f\|^2_{\mathcal{H}^{\boldsymbol{s}}} \stackrel{def}{=} \int_{\mathbb{R}^N} \prod_{k=1}^d (1 + \|\xi_k\|^2)^{s_k} |\widehat{f}(\boldsymbol{\xi})|^2 \mathrm{d}\boldsymbol{\xi}. \tag{2.1.7}$$

The space $\mathcal{H}^{\boldsymbol{s}}\left(\bigoplus_{k=1}^d H_k\right)$ endowed the norm (2.1.7) *is a Hilbert space.*

Notice that the Sobolev norm makes use of the homogeneous norm for the weight on the Fourier transform side. It is also useful to introduce standard isotropic product Sobolev spaces even when working with factors with non-isotropic delations.

For product spaces we have the following lemma that is a direct adaptation of results for classical Sobolev spaces.

Lemma 2.1.5.
Let $\widehat{f} \in \mathcal{H}^{\boldsymbol{s}}(\mathbb{R}^N)$ with $\boldsymbol{s}$ such that $s_k > {}^{q_k}\!/_2$ for all $k \in \{1, \ldots, d\}$. Then $f \in L^1(\mathbb{R}^N)$ and if $\boldsymbol{\varepsilon} \in (\mathbb{R}^+)^d$ is such that $s_k > {}^{q_k}\!/_2 + \varepsilon_k$ then

$$\int_{\mathbb{R}^N} |f(x)|(1 + \|x_1\|)^{\varepsilon_1} \ldots (1 + \|x_d\|)^{\varepsilon_d} \mathrm{d}\boldsymbol{x} < C_{\boldsymbol{\varepsilon}} \left\|\widehat{f}\right\|_{\mathcal{H}^{\boldsymbol{s}}}$$

In particular, for such $\boldsymbol{s}$ and $\boldsymbol{\varepsilon}$, $\mathcal{H}^{\boldsymbol{s}}(\mathbb{R}^N) \subset C_0(\mathbb{R}^N)$ and the inclusion is continuous.

2.1.2 Distributions, product distributions and (partial) duality

We now proceed with some preliminary questions of notation and basic properties of distributions on product spaces. Suppose, as before, that we are working on a product space $\bigoplus_{k=1}^d H_k \simeq \mathbb{R}^N$. Given a distribution $K \in S'(\mathbb{R}^N)$, by an abuse of notation, we indicate with

$$\int_{\mathbb{R}^N} K(\boldsymbol{x})\varphi(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \stackrel{\text{def}}{=} \langle K;\, \varphi\rangle \tag{2.1.8}$$

the duality pairing with any $\varphi \in S(\mathbb{R}^N)$. On the other hand given a test functions $\varphi_k \in S(H_k)$ for all $k \in \{1, \ldots, d\}$ we can naturally construct the test function

$$\varphi_1 \otimes \ldots \otimes \varphi_d(\boldsymbol{x}) \stackrel{\text{def}}{=} \prod_{k=1}^d \varphi_k(x_k). \tag{2.1.9}$$

The mapping $\bigotimes : \bigtimes_{k=1}^d S(H_k) \to S(\mathbb{R}^N)$ thus defined is continuous and multi-linear.

Since $\mathbb{R}^N$ is endowed with a product structure it is often useful to consider partial duality pairings. For any $\psi \in S(H_k)$ we define

$$\int_{H_k} K(\boldsymbol{x})\psi(x_k)\mathrm{d}x_k$$

as the distribution in $S'(H_{(k)})$ given by the relation

$$\left\langle \int_{H_k} K(\boldsymbol{x})\psi(x_k)\mathrm{d}x_k;\, \varphi \right\rangle \stackrel{\text{def}}{=} \langle K;\, \varphi \otimes \psi\rangle = \int_{\mathbb{R}^N} K(\boldsymbol{x})\psi(x_k)\varphi(\boldsymbol{x}_{(k)})\, \mathrm{d}\boldsymbol{x} \tag{2.1.10}$$

for any $\varphi \in S(H_{(k)})$. More in general for any subset of indexes $J \subset \{1,\dots d\}$ and for any $\psi(\boldsymbol{x}_J) \in S(H_J)$ and $\varphi(\boldsymbol{x}_{(J)}) \in S(H_{(J)})$ we have

$$\left\langle \int_{H_{(J)}} K(\boldsymbol{x})\psi(\boldsymbol{x}_J)\mathrm{d}\boldsymbol{x}_J;\, \varphi \right\rangle = \left\langle \int_{H_J} K(\boldsymbol{x})\varphi(\boldsymbol{x}_{(J)})\mathrm{d}\boldsymbol{x}_{(J)};\, \psi \right\rangle \stackrel{\text{def}}{=} \int K(\boldsymbol{x})\psi(\boldsymbol{x}_J)\varphi(\boldsymbol{x}_{(J)})\mathrm{d}\boldsymbol{x} \tag{2.1.11}$$

Many results from standard distribution theory can be adapted to this product setting. We define partial convolution with a Schwartz function $\varphi \in S(H_J)$ as

$$\varphi *_J K(\boldsymbol{x}) = \int_{H_J} \varphi(\boldsymbol{x}_J - \boldsymbol{y}_J)K\left(\boldsymbol{y}_J \oplus \boldsymbol{x}_{(J)}\right)\mathrm{d}\boldsymbol{y}_J. \tag{2.1.12}$$

We have that the duality relation for convolution assumes the following form:

$$\int_{\mathbb{R}^N} \left(\varphi *_J K\right)(\boldsymbol{x})\psi(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^N} K(\boldsymbol{x})\left(\check{\varphi} *_J \psi\right)(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \tag{2.1.13}$$

where $\check{\varphi} \stackrel{\text{def}}{=} \varphi \circ (-\operatorname{Id})$ and the relation holds for $\varphi \in S(H_J)$ and $\psi \in S(\mathbb{R}^N)$. It is easy to see that if $\varphi_n \in S(H_J)$ is an approximate identity then $\varphi_n *_J K \rightharpoonup K$.

Special product test functions

In the study of singular integrals we frequently use some test functions with special properties. Most of the times the special properties we ask are not strictly necessary, but are included to simplify otherwise more cumbersome proofs or properties. Here we recall some of these cases and ask some additional properties to make these classes of functions well suited for a multi-parameter theory.

We begin with cutoff functions.

Definition 2.1.6 (Product cutoff functions)**.**
Let $\eta_k(x_k) \in D(H_k)$ be cutoff functions on H_k for every $k \in \{1,\dots,d\}$ as defined in 1.2.7. We say that

$$\boldsymbol{\eta}(\boldsymbol{x}) \stackrel{def}{=} \left(\bigotimes_{k=1}^{d} \eta_k\right)(\boldsymbol{x}) = \prod_{k=1}^{d} \eta_k(x_k) \tag{2.1.14}$$

is a product cutoff function.

Now we turn to approximate identities.

Definition 2.1.7 (Product approximate identities)**.**
Let $(\varphi_{k,n})_{n\in\mathbb{N}} \in D(H_k)$ be an approximate identity on H_k for every $k \in \{1,\dots,d\}$. We call $(\boldsymbol{\varphi}_n)_{n\in\mathbb{N}} \in D(\mathbb{R}^N)$ a product approximate identity when

$$\boldsymbol{\varphi}_n(\boldsymbol{x}) \stackrel{def}{=} \left(\bigotimes_{k=1}^{d} \varphi_{k,n}\right)(\boldsymbol{x}) = \prod_{k=1}^{d} \varphi_{k,n}(x_k). \tag{2.1.15}$$

As approximate identities we will take rescaled L^1-normalized cutoff functions:

$$\varphi_{k,n}(x_k) = \frac{\eta_k^{(2^{-n})}(x_k)}{\int_{H_k} \eta_k(x_k)\mathrm{d}x_k} \qquad\qquad \boldsymbol{\varphi}_n(\boldsymbol{x}) = \boldsymbol{\varphi}^{(2^{-n})}(\boldsymbol{x}). \tag{2.1.16}$$

2.2 Product kernels

We will now proceed to introduce new class of singular integrals that generalize Calderón-Zygmund kernels to product space. In Section 1.2 of the introduction we have seen that the class of Calderón-Zygmund kernels is naturally adapted to a one-parameter family of dilations. On product spaces it is natural to introduce a multi-parameter dilation structure and to study multi-parameter homogeneity. The most direct approach to multi-parameter theory is to start with tensor products of Calderón-Zygmund kernels on product spaces.

From the point of view of the associated operators, product kernels also tend to appear naturally. Consider the Euclidean space $\mathbb{R}^2 = \mathbb{R} \oplus \mathbb{R}$ considered as a product space and the two operators given by the Hilbert transforms along the coordinates

$$\mathcal{H}_1\varphi \stackrel{\text{def}}{=} \varphi *_{x_1} \mathrm{PV}\,{}^1\!/_{x_1}$$
$$\mathcal{H}_2\varphi \stackrel{\text{def}}{=} \varphi *_{x_2} \mathrm{PV}\,{}^1\!/_{x_2}$$

where

$$\left(\varphi *_{x_1} \mathrm{PV}\,{}^1\!/_{x_1}\right)(x_1, x_2) = \lim_{\varepsilon\to 0} \int_{|s|>\varepsilon} \varphi(x_1 - s, x_2)s^{-1}\mathrm{d}s \tag{2.2.1}$$

and $\mathcal{H}_2$ is defined similarly. These operators are initially defined on test functions in $S(\mathbb{R}^2)$ and the relation (2.2.1) is meaningless for functions φ in other spaces like L^p since any line has zero Lebesgue measure. However these operators do admit a bounded extension to $L^p(\mathbb{R}^2)$. Even though the one-coordinate Hilbert transforms are not, per se, Calderón-Zygmund operators, we can adapt the result 1.2.13 that gives the L^p boundedness of the Hilbert transform to our case. We can interpret x_2 as a parameter and seeing $\mathcal{H}_1$ as the one-dimensional Hilbert transform $\mathcal{H}$. For $\varphi \in S(\mathbb{R}^2)$ we write

$$\|\mathcal{H}_1\varphi\|_{L^p(\mathbb{R}^2)} = \left(\int_{\mathbb{R}^2} \left|\mathrm{PV}\int \varphi(x_1 - s, x_2)s^{-1}\mathrm{d}s\right|^p \mathrm{d}\boldsymbol{x}\right)^{1/p} = \tag{2.2.2}$$

$$\left(\int_{\mathbb{R}} \|\mathcal{H}\varphi(\cdot, x_2)\|^p_{L^p(\mathbb{R})} \mathrm{d}x_2\right)^{1/p} < C\left(\int_{\mathbb{R}} \|\varphi(\cdot, x_2)\|^p_{L^p(\mathbb{R})} \mathrm{d}x_2\right)^{1/p} = \tag{2.2.3}$$

$$C\left(\int_{\mathbb{R}^2} |\varphi(x_1, x_2)|^p \mathrm{d}\boldsymbol{x}\right)^{1/p} = C\,\|\varphi\|_{L^p(\mathbb{R}^2)}\,. \tag{2.2.4}$$

Otherwise the existence of a bounded extension follows using the formalism expressed in Section 2.1, dedicated to product spaces. Theorem 2.1.3 allows us to see $L^p(\mathbb{R}^2)$ as $L^p(\mathbb{R}; L^p(\mathbb{R}))$ through the mapping $y \mapsto \varphi(\cdot; y)$. $\mathcal{H}_1$ acts as a Hilbert transform on $L^p(\mathbb{R})$, the target space of the above mapping. So $\mathcal{H}_1\varphi$ is associated with the mapping $y \mapsto \mathcal{H}\varphi(\cdot; y)$. Since $\mathcal{H}$ is a bounded linear operator the mapping $y \mapsto \mathcal{H}\varphi(\cdot; y)$ is in $L^p(\mathbb{R}; L^p(\mathbb{R}))$ and

$$\|\mathcal{H}_1\varphi\|_{L^p(\mathbb{R}^2)} = \left\|y \mapsto \mathcal{H}\varphi(\cdot; y)\right\|_{L^p(\mathbb{R};L^p(\mathbb{R}))} \le C\left\|y \mapsto \varphi(\cdot; y)\right\|_{L^p(\mathbb{R};L^p(\mathbb{R}))} = \|\varphi\|_{L^p(\mathbb{R}^2)} \tag{2.2.5}$$

where C is the $L^p - L^p$ norm of the Hilbert transform. The same boundedness result can be stated for $\mathcal{H}_2$ thus the composition operators $H_1 \circ H_2$ and $H_2 \circ H_1$ are well defined. Furthermore both $\mathcal{H}_1$ and $\mathcal{H}_2$ are translation invariant and so are the composition operators. By Schwartz's kernel theorem $H_1 \circ H_2$ and $H_2 \circ H_1$ are given by convolutions with some distribution. It is natural to try to determine the properties of these kernels. By taking the Fourier transform of φ we see that $\mathcal{H}_1$ and $\mathcal{H}_2$ act by multiplication by $-i\,\mathrm{sign}(\xi_1)$ and $-i\,\mathrm{sign}(\xi_2)$ respectively. Thus the two operators commute

$$\mathcal{H}_1 \circ \mathcal{H}_2 = \mathcal{H}_2 \circ \mathcal{H}_1$$

and the associated multiplier is

$$-\operatorname{sign}(\xi_1\xi_2)$$

Since $\mathcal{H}_1$ and $\mathcal{H}_2$ commute and they act independently on the variables x_1 and x_2 i.e. they respect the product space nature of $\mathbb{R}^2$ we will write the composition operator as

$$\mathcal{H}_1 \circ \mathcal{H}_2 = \mathcal{H}_2 \circ \mathcal{H}_1 = \mathcal{H}(x_1) \otimes \mathcal{H}(x_2).$$

It is tempting to write that $\mathcal{H}(x_1) \otimes \mathcal{H}(x_2)$ is given by integration with the kernel ${}^1\!/\!_{x_1x_2}$, however this kernel has a non-integrable singularity at the origin. Since the associated kernel K is odd in each variable one can actually write something similar to principal value integral

$$\langle K; \varphi\rangle = \lim_{\substack{\varepsilon_1\to 0^+\\ \varepsilon_2\to 0^+}} \int_{\substack{|x_1|>\varepsilon_1\\ |x_1|>\varepsilon_2}} \frac{\varphi(x_1,x_2)}{x_1x_2} \mathrm{d}x_1\mathrm{d}x_2 =$$

$$\frac{1}{4}\lim_{\substack{\varepsilon_1\to 0^+\\ \varepsilon_2\to 0^+}} \int_{\substack{|x_1|>\varepsilon_1\\ |x_1|>\varepsilon_2}} \frac{\varphi(x_1,x_2)-\varphi(-x_1,x_2)-\varphi(x_1,-x_2)+\varphi(-x_1,-x_2)}{x_1x_2} \mathrm{d}x_1\mathrm{d}x_2.$$

The last expression is clearly a convergent integral so the limit is well defined.

The Hilbert transform was the starting point and the test case for Calderón-Zygmund theory. In much the same way as above we could define tensor products of Calderón-Zygmund kernels and operators. Product kernels arise when we try to define a class that includes tensor products of Calderón-Zygmund kernels but are not necessarily representable as independently acting on the variables of a product space. One can think of product kernels as a completion of tensor products of Calderón-Zygmund kernels in a suitable topology.

2.3 Definition

The definition of product kernels on $\mathbb{R}^N$ is closely related to the product space structure of the space. On the other hand product kernels also possess a natural homogeneity so they depend on the family of dilations defined on $\mathbb{R}^N$. For this reason we begin by fixing a product decomposition of $\mathbb{R}^N$ compatible with its homogeneous structure as described in Section 2.1.1. Let us from now on suppose $\mathbb{R}^N = \bigoplus_{k=1}^d H_d$ where $\{H_k\}_{k\in\{1,\dots,d\}}$ are Euclidean spaces that are invariant with respect to the family of dilations on $\mathbb{R}^N$.

The definition of product kernels has an inductive nature and it is based on the definition of the class $\mathcal{CZ}(\nu)$. The properties and theorems relative of these kernels will also be inspired by the classical results and ideas relative to Calderón-Zygmund kernels. In the same way as we did for Calderón-Zygmund kernels in 1.2.9 we do not use the Fourier transform in the definition but we make use of the cancellation property that is expressed using the duality pairing with normalized bump function defined in 1.2.8.

In much the same way as we did for multi-indexes in (2.1.3) we define multi-orders for kernels. A multi-order relative to the product decomposition $\mathbb{R}^N = \bigoplus_{k=1}^d H_k$ is

$$\boldsymbol{\nu} = (\nu_1, \dots, \nu_d) \in \mathbb{R}^d. \tag{2.3.1}$$

As usual for any subset $J \subset \{1,\dots,d\}$ we define the multi-order $\boldsymbol{\nu_J} \in \mathbb{R}^{|J|}$ relative to the product space H_J so that

$$\boldsymbol{\nu_J} = (\nu_{j_1}, \dots, \nu_{j_{|J|}}). \tag{2.3.2}$$

For any $j \in \{1,\dots,d\}$ we have that the multi-order $\boldsymbol{\nu}_{(j)}$ relative to $H_{(j)}$ is $(\nu_1,\dots,\nu_{j-1},\nu_{j+1},\dots,\nu_d)$. The other definitions and remarks made for multi-indexes in Section 2.1.1 hold accordingly.

Definition 2.3.1 (Product Kernel)**.**
Consider a product decomposition $\{H_k\}_{k\in\{1,\dots,d\}}$ *of* R^N *of length d and an associated multi-order* $\boldsymbol{\nu} = (\nu_1, \dots \nu_d) \in \mathbb{R}^d$.

For $d = 1$*, the class of product kernels* $\mathcal{PK}_{\{H_k\}}(\boldsymbol{\nu})$ *coincides with* $\mathcal{CZ}(\nu_1)$ *on* $\mathbb{R}^N$ *as defined in* 1.2.9*. We say that a family of kernels in* $\mathcal{CZ}(\nu_1)$ *is uniformly bounded if the inequalities* (1.2.4) *and* (1.2.5) *hold with uniformly bounded constants.*

For $d > 1$*, we say that a distribution* $K \in S'(\mathbb{R}^N)$ *is of class* $\mathcal{PK}_{\{H_k\}}(\boldsymbol{\nu})$ *if, away from all the coordinate subspaces* $H_k^\perp = \{\boldsymbol{x} \in \mathbb{R}^N \mid x_k = 0\}$ *with* $k \in \{1, \dots, d\}$*,* K *coincides with a smooth function i.e.*

$$K\big|_{\mathbb{R}^N\setminus\left(\bigcup_{k=1}^d H_k^\perp\right)} \in C^\infty\left(\mathbb{R}^N \setminus \bigcup_{k=1}^d H_k^\perp\right)$$

and it satisfies the following two kinds of conditions:

Size conditions

$$\left|\partial_{\boldsymbol{x}}^{\boldsymbol{\alpha}} K(\boldsymbol{x})\right| = \left|\partial_{x_1}^{\alpha_1} \dots \partial_{x_d}^{\alpha_d} K(\boldsymbol{x})\right| \leq C_{\boldsymbol{\alpha}} \|x_1\|^{-q_1-\nu_1-\|\alpha_1\|} \dots \|x_d\|^{-q_d-\nu_d-\|\alpha_d\|} \tag{2.3.3}$$

for all $\boldsymbol{x} \notin \bigcup_{k=1}^d H_k^\perp$ *and for any multi-index* $\boldsymbol{\alpha}$.

Cancellation conditions *For every* $k \in \{1, \dots, d\}$ *the distributions*

$$R^{\nu_k} \int K(\boldsymbol{x})\varphi(R^{-1} \cdot x_k)\mathrm{d}x_k \tag{2.3.4}$$

obtained by contracting K *with all possible rescaled* b_k*-normalized bump functions* φ *on the subspace* H_k*, where*

$$b_k = b(\nu_k) = \min\{b \in \mathbb{N} \mid \ b > \nu_k\},$$

are a family of product kernels of order $\boldsymbol{\nu}_{(k)}$ *on* $H_{(k)}$ *relative to the decomposition* $\left\{H_j\right\}_{j\neq k}$*, uniformly bounded with respect to* $R > 0$ *and to* φ.

We say that a family of kernels in $\mathcal{PK}_{\{H_k\}}(\boldsymbol{\nu})$ *is uniformly bounded if the bounds arising inductively from* (2.3.3) *and* (2.3.4) *on all the kernels of the family are uniformly bounded.*

For any given product decomposition $\{H_k\}$ and for any multi-order $\boldsymbol{\nu}$ the class $\mathcal{PK}_{\{H_k\}}(\boldsymbol{\nu})$ is a vector space. We will usually omit the dependence of the class on the given product decomposition by simply writing $\mathcal{PK}(\boldsymbol{\nu})$. This class, as a vector space, has a natural Fréchet space topology. As a matter of fact one can get a countable family of semi-norms by taking the lower bound of the constants that appear inductively from the conditions (2.3.3) and (2.3.4) in the definition. Explicitly these semi-norms are given in a non-inductive manner by the following expressions.

Definition 2.3.2 (Semi-norms on product kernels)**.**
Let $K \in \mathcal{PK}_{\{H_k\}}(\boldsymbol{\nu})$ *on* $\mathbb{R}^N$ *associated with a product decomposition* $\{H_k\}_{k\in\{1,\dots,d\}}$ *with d terms.*

For any $J \subset \{1, \dots, d\}$ *let* $\{j_1, \dots, j_{|J|}\}$ *be the elements of* J *and* $\{j_{|J|+1}, \dots, j_d\}$ *be the elements of* $\{1, \dots, d\} \setminus J$*. Then for any* $N \in \mathbb{N}$ *we define the semi-norm*

$$P_{J,N}^{\boldsymbol{\nu}}(K) = \sup \sum_{\|\boldsymbol{\alpha}\|\leq N} \Bigg| \left\|x_{j_{|J|+1}}\right\|^{q_{j_{|J|+1}}+\nu_{j_{|J|+1}}+\left\|\alpha_{j_{|J|+1}}\right\|} \dots \left\|x_{j_d}\right\|^{q_{j_d}+\nu_{j_d}+\left\|\alpha_{j_d}\right\|} R_1^{\nu_{j_1}} \dots R_{|J|}^{\nu_{j_{|J|}}}$$

$$\left.\int_{H_J} \partial_{x_{j_{|J|+1}}}^{\alpha_{j_{|J|+1}}} \dots \partial_{x_{j_d}}^{\alpha_{j_d}} K(\boldsymbol{x}) \varphi_{j_1}(R_1^{-1} \cdot x_{j_1}) \dots \varphi_{j_{|J|}}(R_{|J|}^{-1} \cdot x_{j_{|J|}}) \mathrm{d}x_{j_1} \dots \mathrm{d}x_{j_{|J|}} \right|$$

where the upper bound is taken over all independent scales $\boldsymbol{R} \in (\mathbb{R}^+)^{|J|}$ *and over all* b_j*-normalized bump functions* $\varphi_j \in C_c^\infty\left(B_{H_j}(0,1)\right)$ *with* $j \in J$ *and* $b_j = b(\nu_j) = \min\{b \in \mathbb{N} \mid b > \nu_j\}$*, and over all points* $(x_{j_{|J|+1}}, \dots, x_{j_d})$ *such that* $x_j \in H_j \setminus \{0\}$ *for* $j \notin J$*.*

As J varies over the subsets of $\{1, \dots, d\}$ and $N \in \mathbb{N}$ we obtain a countable family of semi-norms on $\mathcal{PK}(\boldsymbol{\nu})$. For details on Fréchet spaces we refer to [Rud91].

It is actually necessary to check that this family of semi-norms separate points and also that the resulting metric space is complete. We will get to proving these two facts in the next section dedicated specifically to questions of topology. These considerations allows us to refer to a family with uniformly bounded constants appearing inductively from conditions (2.3.3) and (2.3.4) as a bounded family of $\mathcal{PK}(\boldsymbol{\nu})$ kernels.

Remark 2.3.3.
Due to the inductive nature of the definition of product kernels the proofs of many properties will also be inductive. For ease of notation it is useful to notice that the inductive definition can be stated starting from $d = 0$*. As a matter of fact for* $d = 0$ *let us define* $\mathcal{PK}(\boldsymbol{\nu})$ *as simply the space of constants. Then Definition 1.2.9 for* $\mathcal{CZ}(\nu)$ *coincides with the first inductive step in Definition 2.3.1. In most of the successive proofs, we will use an induction argument starting from* $d = 0$*.*

2.4 Topology and basic properties

This section is dedicated to the study of the topology of the spaces $\mathcal{PK}(\boldsymbol{\nu})$. $\mathcal{PK}(\boldsymbol{\nu})$ as vector space possesses two important topologies. The first one, that originates from the already mentioned the semi-norms defined in 2.3.2, is the so called strong topology. However $\mathcal{PK}(\boldsymbol{\nu})$ also inherits the weak-* topology as a subspace of $S'(\mathbb{R}^N)$. We deal here with some properties of these topologies. We state some useful approximation properties and well-behaved approximation methods. After that we deal with the relationship between the weak-* and strong topologies on $\mathcal{PK}(\boldsymbol{\nu})$. Finally we prove that the definition of product kernels is tolerant towards the normalization orders in the cancellation condition (2.3.4).

A very important result useful to study product kernels is given by a generalization of Hadamard's Lemma for product test functions. As a matter of fact this version of the Hadamard's lemma enables us see how close a test function on a product space is to a tensor product of test functions on the factors.

Lemma 2.4.1 (Generalized Hadamard's lemma)**.**
Let $\varphi \in S(R^N)$ *with* $\mathbb{R}^N = \bigoplus_{k=1}^d H_k$*. Let* η_k *be some cutoff functions on* H_k *and let* $\boldsymbol{\eta}$ *be the associated product cutoff function. For arbitrary orders* $m_1, \dots, m_d \in \mathbb{N}$ *the following decomposition holds:*

$$\varphi(\boldsymbol{x}) = \sum_{\substack{|\alpha_1|<m_1 \\ \dots \\ |\alpha_d|<m_d}} \frac{\partial_{\boldsymbol{x}}^{\boldsymbol{\alpha}} \varphi(\boldsymbol{0})}{\boldsymbol{\alpha}!} \boldsymbol{x}^{\boldsymbol{\alpha}} \boldsymbol{\eta}(\boldsymbol{x}) + \sum_{\substack{J \subset \{1,\dots,d\} \\ J \neq \emptyset}} \sum_{\substack{|\alpha_j|=m_j \; j \in J \\ |\alpha_k|<m_k \; k \notin J}} \boldsymbol{x}_{(J)}^{\boldsymbol{\alpha}_{(J)}} \boldsymbol{\eta}_{(J)}(\boldsymbol{x}_{(J)}) \boldsymbol{x}_J^{\boldsymbol{\alpha}_J} \psi_{\boldsymbol{\alpha}}(\boldsymbol{x}) \qquad (2.4.1)$$

where $\psi_{\boldsymbol{\alpha}}$ *are smooth remainder terms that depend only on the coordinates* x_j *with* j *such that* $|\alpha_j| = m_j$*. The remainder terms* $\psi_{\boldsymbol{\alpha}}$ *depend linearly on* φ *and they are normalized with respect to* φ *in the sense that we have the following estimate for the Schwartz norms:*

$$\left\|(1+|\boldsymbol{x}|)^a \psi_{\boldsymbol{\alpha}}\right\|_{C^b} \leq C_{b,\boldsymbol{\alpha}} \left\|(1+|\boldsymbol{x}|)^a \varphi\right\|_{C^{b+|\boldsymbol{\alpha}|}} \qquad a, b \in \mathbb{N}.$$

Furthermore if φ is compactly supported then $\psi_{\boldsymbol{\alpha}}$ are also supported on compact sets dependent on the support of φ but not on φ itself.

It is important to notice that here we are dealing with the standard size of the multi-indexes. This decomposition is independent of the homogeneous structure on $\mathbb{R}^N$.

Proof. We do a triple induction the number of product factors d in the decomposition of $\mathbb{R}^N$, on the orders m_k and on the dimension of the single spaces H_k.

We proceed in the inverse order. First let us suppose $d = 1$ and $m_1 = 1$. Let the dimension of the single factor be $n_1 = 1$. Take

$$\psi(x) = \begin{cases} \frac{\varphi(x)-\varphi(0)\eta(x)}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

and write the decomposition $\varphi(x) = \varphi(0)\eta(x) + x\psi(x)$. If φ is compactly supported on $[-R, R]$ for some $R > 0$ then the support of ψ is contained in the set $(\operatorname{spt}\varphi \cup \operatorname{spt}\eta) \subset [-(R\vee 1); +(R\vee 1)]$. By setting $f(x) = x\psi(x) = \varphi(x) - \varphi(0)\eta(x)$ we get

$$\psi(x) = \frac{f(x)}{x} = \frac{\int_0^x f'(s)\mathrm{d}s}{x} = \int_0^1 f'(tx)\mathrm{d}t$$

and

$$D^b\psi(x) = D^b\left(\frac{f(x)}{x}\right) = \int_0^1 t^b D^{b+1} f(tx)\mathrm{d}t.$$

It now suffices to notice that

$$\left\|(1+|x|)^a f(x)\right\|_{C^b} = \left\|(1+|x|)^a \left(\varphi(x) - \varphi(0)\eta(x)\right)\right\|_{C^b} \approx \left\|(1+|x|)^a\varphi\right\|_{C^b}.$$

The linearity of the mapping $\varphi \mapsto \psi$ that associate to a given φ the remainder term ψ is given by construction.

Now suppose that we are working on $\mathbb{R}^N$ with $N = n_1 > 1$. Write down $x = (x_1, x')$ where x_1 is the first coordinate and x' are the last $n_1 - 1$ coordinates that we consider parameters. By the induction hypothesis we have the decomposition

$$\varphi(x_1, x') = \varphi(0, x')\eta(x_1) + x_1\psi_1(x_1, x') \tag{2.4.2}$$

We can now apply the induction hypothesis again to $\varphi(0, x')$ to get

$$\varphi(x_1, x') = \varphi(0)\eta(x_1)\eta(x') + \sum_{j=2}^{n_1} \eta(x_1)x_j\psi_j(0, x') + x_1\psi_1(x_1, x').$$

The normalization of ψ_1 is evident by setting $f(x) = \varphi(x_1, x') - \varphi(0, x')\eta(x_1)$ and writing

$$\partial_x^\alpha\left(\frac{f(x_1, x')}{x_1}\right) = \int_0^1 t^{\alpha_1} D^\alpha D_1 f(tx_1, x')\mathrm{d}t.$$

The normalization for the other terms follows by the induction hypothesis. Once again linearity is obvious by construction. We have proved the lemma for order $m = 1$ and for trivial product decompositions of $\mathbb{R}^N$ with arbitrary N.

If $m > 1$ then by the induction hypothesis we have the decomposition

$$\varphi(x) = \sum_{|\alpha|<m-1} \frac{\partial^\alpha\varphi(0)}{\alpha!} x^\alpha\eta(x) + \sum_{|\alpha|=m-1} x^\alpha\psi_\alpha(x) \tag{2.4.3}$$

with ψ_α normalized with respect to φ in the sense that

$$\left\|(1+|x|)^a\psi_\alpha\right\|_{C^b} < C_{b,\alpha}\left\|(1+|x|)^a\varphi\right\|_{C^{b+|\alpha|}}.$$

Applying the induction hypothesis again with $m=1$ to the remainder terms ψ_α we get

$$\psi_\alpha(x) = \psi_\alpha(0)\eta(x) + \sum_{j=1}^{n_1} x_j\widetilde{\psi}_{\alpha,j}(x)$$

and substituting into (2.4.3) we obtain

$$\varphi(x) = \sum_{|\alpha|<m-1}\frac{\partial^\alpha\varphi(0)}{\alpha!}x^\alpha\eta(x) + \sum_{|\alpha|=m-1}\psi_\alpha(0)x^\alpha\eta(x) + \sum_{|\alpha|=m-1}x_jx^\alpha\widetilde{\psi}_{\alpha,j}(x). \tag{2.4.4}$$

The remainder terms are normalized since

$$\left\|(1+|x|)^a\widetilde{\psi}_{\alpha,j}\right\|_{C^b} < C_{b,1}\left\|(1+|x|)^a\psi_\alpha\right\|_{C^{b+1}} < C_{b,m}\left\|(1+|x|)^a\varphi\right\|_{C^{b+m}}$$

by the induction hypothesis. For any given α such that $|\alpha|=m-1$ taking the α derivative in 0 of (2.4.4) yields

$$\partial^\alpha\varphi(0) = \alpha!\psi_\alpha(0)$$

and this gives the needed equality. The procedure adapted in the proof was constructive and linear. So the mapping from φ to the remainder terms is linear.

Now we prove the induction on d. Suppose $d>1$ and write $\boldsymbol{x}=(x_1,\boldsymbol{x}_{(1)})$ where x_1 are the first subspace coordinates and $\boldsymbol{x}_{(1)}$ are the last $d-1$ coordinates that we initially see as parameters. By applying the induction hypothesis along x_1 we get the decomposition

$$\varphi(\boldsymbol{x}) = \sum_{|\alpha_1|<m_1}\frac{\partial_{x_1}^{\alpha_1}\varphi(0,\boldsymbol{x}_{(1)})}{\alpha_1!}x_1^{\alpha_1}\eta_1(x_1) + \sum_{|\alpha_1|=m_1}x_1^{\alpha_1}\psi_{\alpha_1}(x_1,\boldsymbol{x}_{(1)}). \tag{2.4.5}$$

For any given α_1 such that $|\alpha_1|=m_1$ the mapping $\varphi(\cdot,\boldsymbol{x}_{(1)})\mapsto\psi_{\alpha_1}(\cdot,\boldsymbol{x}_{(1)})$ is linear. As a consequence $\left\|\psi_{\alpha_1}\right\|_{C^b} < C_{b,\alpha_1}\left\|\varphi\right\|_{C^{b+m}}$. As a matter of fact by boundedness and linearity $\partial^{\boldsymbol{\beta}}_{\boldsymbol{x}_{(1)}}\psi_{\alpha_1}(\cdot,\boldsymbol{x}_{(1)})$ is the remainder obtained by applying the decomposition to $\partial^{\boldsymbol{\beta}}_{\boldsymbol{x}_{(1)}}\varphi(\cdot,\boldsymbol{x}_{(1)})$ so

$$\left\|\prod_{k=1}^d(1+|x_k|)^a\partial^{\boldsymbol{\beta}}_{\boldsymbol{x}_{(1)}}\psi_{\alpha_1}(\cdot,\boldsymbol{x}_{(1)})\right\|_{C^b} < C_{b,\alpha_1}\left\|\prod_{k=1}^d(1+|x_k|)^a\partial^{\boldsymbol{\beta}}_{\boldsymbol{x}_{(1)}}\varphi(\cdot,\boldsymbol{x}_{(1)})\right\|_{C^{b+m_1}} < C_{b,\alpha_1}\left\|\prod_{k=1}^d(1+|x_k|)^a\varphi\right\|_{C^{b+m_1+|\boldsymbol{\beta}|}}$$

and this gives the normalization

$$\left\|\prod_{k=1}^d(1+|x_k|)^a\psi_{\alpha_1}(\boldsymbol{x})\right\|_{C^b} < C_{b,\alpha_1}\left\|\prod_{k=1}^d(1+|x_k|)^a\varphi(\boldsymbol{x})\right\|_{C^{b+m_1}}.$$

Now we apply the induction hypothesis and a reasoning similar to the above to ψ_{α_1} for all $|\alpha_1|=m_1$ keeping x_1 as a parameter. We also apply the induction hypothesis to the terms of the sum

$$\sum_{|\alpha_1|<m_1}\frac{\partial_{x_1}^{\alpha_1}\varphi(0,\boldsymbol{x}_{(1)})}{\alpha_1!}x_1^{\alpha_1}\eta_1(x_1).$$

We get

$$\varphi(\boldsymbol{x}) = \sum_{\substack{|\alpha_1|<m_1\\ \cdots \\ |\alpha_d|<m_d}} \frac{\partial_{\boldsymbol{x}}^{\boldsymbol{\alpha}}\varphi(\mathbf{0})}{\boldsymbol{\alpha}!}\boldsymbol{x}^{\alpha}\boldsymbol{\eta}(\boldsymbol{x}) +$$

$$\sum_{|\alpha_1|<m_1} \sum_{\substack{J\subset\{2,\ldots,d\}\\ J\neq\emptyset}} \prod_{k\in\{1,\ldots,d\}\setminus J} \eta_k(x_k) \sum_{\substack{|\alpha_k|<m_k\ k\in\{2,\ldots,d\}\setminus J\\ |\alpha_j|=m_j\ j\in J}} \boldsymbol{x}^{\alpha}\widetilde{\psi}_{\boldsymbol{\alpha}}(0,\boldsymbol{x}_{(1)}) +$$

$$\sum_{|\alpha_1|=m_1} \sum_{\substack{J\subset\{2,\ldots,d\}\\ J\neq\emptyset}} \prod_{k\in\{2,\ldots,d\}\setminus J} \eta_k(x_k) \sum_{\substack{|\alpha_k|<m_k\ k\in\{2,\ldots,d\}\setminus J\\ |\alpha_j|=m_j\ j\in J}} \boldsymbol{x}^{\alpha}\widetilde{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{x}) \tag{2.4.6}$$

The inductive hypothesis gives the normalization property for the remainders, the compactness of the support as functions of $\boldsymbol{x}_{(1)}$, and finally the independence along those coordinates for which $\boldsymbol{\alpha}_k < m_k$. The linearity of the mapping to the remainders is given by construction. Furthermore the normalization property gives us that the remainder terms have compact support. In fact $\psi_{\alpha_1}(\cdot, \boldsymbol{x}_{(1)})$ are non-zero functions only for a compact set of parameters $\boldsymbol{x}_{(1)}$. □

Together with Hadamard's Lemma we need a technical preliminary result before we proceed to prove the properties in the next sections.

Lemma 2.4.2 (Complete local integrability for all negative orders)**.**
Let $K \in \mathcal{PK}(\boldsymbol{\nu})$. Then given a multi-index $\boldsymbol{\alpha}$ such that $\nu_k - \|\alpha_k\| < 0$ for all $k \in \{1,\ldots,d\}$ the kernel $\boldsymbol{x}^{\boldsymbol{\alpha}}K(\boldsymbol{x})$ coincides with a locally integrable function on the whole $\mathbb{R}^N$ that is given by $\boldsymbol{x}^{\boldsymbol{\alpha}}K(\boldsymbol{x})$ on $\mathbb{R}^N \setminus \bigcup_{k=1}^d H_k^{\perp}$.

Proof. Consider the function $F(\boldsymbol{x}) = \boldsymbol{x}^{\boldsymbol{\alpha}}K(\boldsymbol{x})$ defined everywhere on $\mathbb{R}^N$ except on the set $\bigcup_{k=1}^d H_k^{\perp}$ of zero Lebesgue measure. Due to the size condition (2.3.3) on $K(\boldsymbol{x})$ we have that F is locally integrable on $\mathbb{R}^N$. We now prove that $\boldsymbol{x}^{\boldsymbol{\alpha}}K(\boldsymbol{x})$ coincides with integrating against F on $\mathbb{R}^N$ by showing this to be true for all test functions $\boldsymbol{\varphi}(\boldsymbol{x}) \in D(\mathbb{R}^N)$ of the type $\bigotimes_{k=1}^d \varphi_k(x_k)$ with $\varphi_k \in D(H_k)$. Since linear combinations of these types of functions are dense in $S(\mathbb{R}^N)$ this proves our lemma. Choose cutoff functions $\eta_k(x_k)$ on H_k for all $k \in \{1,\ldots,d\}$ and a parameter $0 < r \leq 1$. We write

$$\int_{\mathbb{R}^N} \boldsymbol{x}^{\boldsymbol{\alpha}}K(\boldsymbol{x})\boldsymbol{\varphi}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^N} K(\boldsymbol{x})\prod_{k=1}^d (x_k)^{\alpha_k}(1-\eta_k)(r^{-1}\cdot x_k)\varphi_k(x_k)\mathrm{d}\boldsymbol{x} +$$

$$\sum_{\substack{J\subset\{1,\ldots,d\}\\ J\neq\emptyset}} \int_{H_{(J)}}\int_{H_J} r^{\|\boldsymbol{\alpha}_J\|}K(\boldsymbol{x})\prod_{j\in J}(r^{-1}\cdot x_j)^{\alpha_j}\eta_k(r^{-1}\cdot x_j)\varphi_j(x_j)\mathrm{d}\boldsymbol{x}_J$$

$$\prod_{k\notin J}(x_k)^{\alpha_k}(1-\eta_k)(r^{-1}\cdot x_k)\varphi_k(x_k)\mathrm{d}\boldsymbol{x}_{(J)}.$$

For the first term we have

$$\int_{\mathbb{R}^N} K(\boldsymbol{x})\prod_{k=1}^d (x_k)^{\alpha_k}(1-\eta_k)(r^{-1}\cdot x_k)\varphi_k(x_k)\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^N} F(\boldsymbol{x})\prod_{k=1}^d (1-\eta_k)(r^{-1}\cdot x_k)\varphi_k(x_k)\mathrm{d}\boldsymbol{x}$$

and if $r \to 0$ then

$$\int_{\mathbb{R}^N} K(\boldsymbol{x})\prod_{k=1}^d (x_k)^{\alpha_k}(1-\eta_k)(r^{-1}\cdot x_k)\varphi_k(x_k)\mathrm{d}\boldsymbol{x} \to \int_{\mathbb{R}^N} F(\boldsymbol{x})\boldsymbol{\varphi}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}.$$

For all the other terms notice that if $j \in J$ then $(r^{-1} \cdot x_j)^{\alpha_j}\eta_k(r^{-1} \cdot x_j)\varphi_j(x_j)$ are r-rescaled versions of $x_j^{\alpha_j}\eta_k(x_j)\varphi_j(r \cdot x_j)$ that, up to a constant depending on φ_j, are b_j-normalized bump functions.

For each $J \subset \{1, \ldots, d\}$ with $J \neq \emptyset$

$$\int_{H_J} K(\boldsymbol{x}) \prod_{j \in J} r^{\nu_j}(r^{-1} \cdot x_j)^{\alpha_j}\eta_k(r^{-1} \cdot x_j)\varphi_j(x_j)\mathrm{d}\boldsymbol{x}_J$$

are uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu}_{(J)})$ for $r \in (0;1]$ and since $\nu_j - \|\alpha_j\| < 0$ the kernels

$$\int_{H_J} r^{\|\boldsymbol{\alpha}_J\|}K(\boldsymbol{x}) \prod_{j \in J}(r^{-1} \cdot x_j)^{\alpha_j}\eta_k(r^{-1} \cdot x_j)\varphi_j(x_j)\mathrm{d}\boldsymbol{x}_J \to 0$$

in the strong topology on $\mathcal{PK}(\boldsymbol{\nu}_{(J)})$ as $r \to 0$. This means that on $H_{(J)} \bigcup_{k \notin J} H_k^{\perp}$ the kernels

$$\boldsymbol{x}_{(J)}^{\boldsymbol{\alpha}_{(J)}} \int_{H_J} r^{\|\boldsymbol{\alpha}_J\|}K(\boldsymbol{x}) \prod_{j \in J}(r^{-1} \cdot x_j)^{\alpha_j}\eta_k(r^{-1} \cdot x_j)\varphi_j(x_j)\mathrm{d}\boldsymbol{x}_J$$

are, up to a constant, dominated by the locally integrable function $\prod_{k \notin J} x_k^{\alpha_k} \|x_k\|^{-q_k-\nu_k}$ and converge a.e. to 0. Since $\prod_{k \notin J}(1 - \eta_k)(r^{-1} \cdot x_k)\varphi_k(x_k)$ are supported away from $\bigcup_{k \notin J} H_k^{\perp}$ and are L^{∞} bounded we have by Lebesgue's dominated convergence theorem that

$$\int_{H_{(J)}} \int_{H_J} r^{\|\boldsymbol{\alpha}_J\|}K(\boldsymbol{x}) \prod_{j \in J}(r^{-1} \cdot x_j)^{\alpha_j}\eta_k(r^{-1} \cdot x_j)\varphi_j(x_j)\mathrm{d}\boldsymbol{x}_J \prod_{k \notin J}(x_k)^{\alpha_k}(1 - \eta_k)(r^{-1} \cdot x_k)\varphi_k(x_k)\mathrm{d}\boldsymbol{x}_{(J)} \to 0$$

as $r \to 0$ for any $J \subset \{1, \ldots, d\}$. So passing to the limit we have

$$\int_{\mathbb{R}^N} \boldsymbol{x}^{\boldsymbol{\alpha}}K(\boldsymbol{x})\boldsymbol{\varphi}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^N} F(\boldsymbol{x})\boldsymbol{\varphi}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}.$$

□

2.4.1 Weak and strong topologies

In this section we turn to the relationship between different vector space topologies on $\mathcal{PK}(\boldsymbol{\nu})$. As shown above, the vector space $\mathcal{PK}(\boldsymbol{\nu})$ has both a strong topology and the weak-* topology inherited as a subspace of $S'(\mathbb{R}^N)$. We write $K_n \rightharpoonup K$ if K_n converges in the weak-* (distributional) topology to K and distinguish from the strong convergence indicated as $K_n \to K$.

The first useful, albeit trivial, fact is that the family of semi-norms on $\mathcal{PK}(\boldsymbol{\nu})$ separates points. If all constant in the inequalities that arise inductively from (2.3.4) and (2.3.3) are identically zero then the kernel cannot be non-zero. It is sufficient to approximate such a kernel K by convolving with an product approximate identity as defined in 2.1.7.

Proposition 2.4.3 (The semi-norms separate points).
Let K be a product kernel with all norms in 2.3.2 equal to 0. Then the kernel is trivial.

Proof. Let $\boldsymbol{\varphi}_n$ be a product approximate identity. $\boldsymbol{\varphi}_n * K \rightharpoonup K$ and $\boldsymbol{\varphi}_n * K \rightharpoonup K$ are smooth functions given by

$$\boldsymbol{\varphi}_n * K(\boldsymbol{x}) = \int_{\mathbb{R}^N} \boldsymbol{\varphi}(\boldsymbol{x} - \boldsymbol{y})K(\boldsymbol{y})\mathrm{d}\boldsymbol{y}$$

But $\boldsymbol{y} \mapsto \boldsymbol{\varphi}(\boldsymbol{x}-\boldsymbol{y}) = \prod_{k=1}^{d} \varphi_k(x_k - y_k)$ is a tensor product of some-how rescaled bump functions. Using the fact that the norm $P^{\boldsymbol{\nu}}_{J,0}$ from Definition 2.3.2 with $J = \{1, \dots, d\}$ is 0 we get $\boldsymbol{\varphi}_n * K = 0$. □

Proposition 2.4.4 (Weak-* and strong topologies)**.**
For any $\boldsymbol{\nu}$ the following relations between the strong topology and the weak- topology hold on $\mathcal{PK}(\boldsymbol{\nu})$:*

1. *The strong topology is finer than the weak-* topology on $\mathcal{PK}(\boldsymbol{\nu})$.*
2. *For any bounded set $V \in S(\mathbb{R}^N)$ and any $\varepsilon > 0$ there is a neighborhood $U_{V,\varepsilon}$ of 0 in $\mathcal{PK}(\boldsymbol{\nu})$ such that if $\varphi \in V$ and $K \in U_{V,\varepsilon}$ then $\left|\int_{R^N} K(\boldsymbol{x})\varphi(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\right| < \varepsilon$. Vice-versa for any bounded set $U \in \mathcal{PK}(\boldsymbol{\nu})$ and any $\varepsilon > 0$ there is a neighborhood $V_{U,\varepsilon}$ of 0 in $S(\mathbb{R}^N)$ such that if $\varphi \in V_{U,\varepsilon}$ and $K \in U$ then $\left|\int_{R^N} K(\boldsymbol{x})\varphi(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\right| < \varepsilon$.*
3. *On bounded sets of $\mathcal{PK}(\boldsymbol{\nu})$ weak-* convergence can be verified only on a dense subset of test functions.*
4. *On bounded sets of $\mathcal{PK}(\boldsymbol{\nu})$ the weak-* convergence implies strong convergence.*
5. *The weak-* closure in $S'(\mathbb{R}^N)$ of a bounded set in $\mathcal{PK}(\boldsymbol{\nu})$ is a closed subset of $\mathcal{PK}(\boldsymbol{\nu})$.*
6. *The semi-norms on $\mathcal{PK}(\boldsymbol{\nu})$ are lower semi-continuous with respect to the weak-* convergence.*

Proof. 1. The strong topology on $\mathcal{PK}(\boldsymbol{\nu})$ is a metric vector space topology so all we need to check is that $K_n \to 0$ implies $K_n \rightharpoonup K$ for any sequence $(K_n)_{n\in N}$. For any fixed test function $\varphi(\boldsymbol{x}) \in S(\mathbb{R}^N)$ we apply Lemma 2.4.1 and write

$$\int_{\mathbb{R}^N} K_n(\boldsymbol{x})\varphi(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \sum_{\substack{|\alpha_1|<m_1\\ \dots \\ |\alpha_d|<m_d}} \frac{\partial_{\boldsymbol{x}}^{\boldsymbol{\alpha}}\varphi(\boldsymbol{0})}{\boldsymbol{\alpha}!} \int_{\mathbb{R}^N} K_n(\boldsymbol{x})\boldsymbol{x}^{\boldsymbol{\alpha}}\boldsymbol{\eta}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} +$$
$$\sum_{\substack{J\subset\{1,\dots,d\}\\ J\neq\emptyset}} \sum_{\substack{|\alpha_j|=m_j\ j\in J\\ |\alpha_k|<m_k\ k\notin J}} \int_{H_J}\int_{H_{(J)}} K_n(\boldsymbol{x})\boldsymbol{x}_{(J)}^{\boldsymbol{\alpha}_{(J)}}\boldsymbol{\eta}_{(J)}(\boldsymbol{x}_{(J)})\mathrm{d}\boldsymbol{x}_{(J)}\boldsymbol{x}_J^{\boldsymbol{\alpha}_J}\psi_{\boldsymbol{\alpha}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}_J.$$

Since $P^{\boldsymbol{\nu}}_{\{1,\dots d\},0}(K_n) \to 0$ and $\boldsymbol{x}^{\boldsymbol{\alpha}}\boldsymbol{\eta}(\boldsymbol{x})$, are up to a constant, normalized bump functions we have that the terms

$$\frac{\partial_{\boldsymbol{x}}^{\boldsymbol{\alpha}}\varphi(\boldsymbol{0})}{\boldsymbol{\alpha}!} \int_{\mathbb{R}^N} K_n(\boldsymbol{x})\boldsymbol{x}^{\boldsymbol{\alpha}}\boldsymbol{\eta}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$$

tend to 0. For the remainder terms we have that

$$\boldsymbol{x}_J^{\boldsymbol{\alpha}_J} \int_{H_{(J)}} K_n(\boldsymbol{x})\boldsymbol{x}_{(J)}^{\boldsymbol{\alpha}_{(J)}}\boldsymbol{\eta}_{(J)}(\boldsymbol{x}_{(J)})\mathrm{d}\boldsymbol{x}_{(J)}$$

tend strongly to 0 in $\mathcal{PK}(\boldsymbol{\nu}_J)$ and by Lemma 2.4.2 these kernels coincide with locally integrable functions on H_J. Since

$$P^{\boldsymbol{\nu}_J}_{\emptyset,0}\left(\int_{H_{(J)}} K_n(\boldsymbol{x})\boldsymbol{x}_{(J)}^{\boldsymbol{\alpha}_{(J)}}\boldsymbol{\eta}_{(J)}(\boldsymbol{x}_{(J)})\mathrm{d}\boldsymbol{x}_{(J)}\right) \to 0$$

by Lebesgue's dominated convergence theorem we have that

$$\int_{H_J}\int_{H_{(J)}} K_n(\boldsymbol{x})\boldsymbol{x}_{(J)}^{\boldsymbol{\alpha}_{(J)}}\boldsymbol{\eta}_{(J)}(\boldsymbol{x}_{(J)})\mathrm{d}\boldsymbol{x}_{(J)}\boldsymbol{x}_J^{\boldsymbol{\alpha}_J}\psi_{\boldsymbol{\alpha}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}_J \to 0$$

as required.

2. The statement follows from the above proof and the normalization property of the Hadamard decomposition 2.4.1.

3. This also is a consequence of the previous point.

4. Let us prove the statement by induction on the number of terms d in the product decomposition of $\mathbb{R}^N$. For $d = 0$ there is nothing to prove. Suppose that $d \geq 1$ and the statement is true for all product spaces of up to $d-1$ factors.

   Suppose that $(K_n)_{n\in\mathbb{N}}$ is a bounded sequence in $\mathcal{PK}(\boldsymbol{\nu})$ and $K_n \rightharpoonup K$. By Ascoli-Arzelá the sequence $K_n\big|_{\mathbb{R}^N\setminus\bigcup_{k=1}^d H_k^\perp}$ is pre-compact in $C^\infty_{loc}(\mathbb{R}^N \setminus \bigcup_{k=1}^d H_k^\perp)$. This means that a subsequence converges on that domain to a C^∞ function that respects the size conditions (2.3.3). Furthermore any subsequence has a converging sub-subsequence and the limit is unique by weak-* convergence of K_n and coincides with K on that domain. For any $k \in \{1,\dots,d\}$, any b_k-normalized bump function φ on H_k and any dilation parameter $R > 0$ the kernels

$$R^{\nu_k}\int_{H_k} K_n(\boldsymbol{x})\varphi(R^{-1}\cdot x_k)\mathrm{d}x_k$$

   are uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu}_{(k)})$ and weakly converge to

$$R^{\nu_k}\int_{H_k} K(\boldsymbol{x})\varphi(R^{-1}\cdot x_k)\mathrm{d}x_k.$$

   By induction we have that

$$R^{\nu_k}\int_{H_k} K_n(\boldsymbol{x})\varphi(R^{-1}\cdot x_k)\mathrm{d}x_k \to R^{\nu_k}\int_{H_k} K(\boldsymbol{x})\varphi(R^{-1}\cdot x_k)\mathrm{d}x_k.$$

   in the strong topology and the above kernel is in $\mathcal{PK}(\boldsymbol{\nu}_{(k)})$ with the same bounds as the sequence. This proves the statement.

5. The weak-* closure of a bounded subset in $\mathcal{PK}(\boldsymbol{\nu})$ is still a subset of $\mathcal{PK}(\boldsymbol{\nu})$ because of the previous statement. The weak-* topology is coarser than the strong topology so a weak-* closed set is also closed in the strong topology.

6. This statement follows directly from the above points.

□

As a consequence we have the metric completeness of $\mathcal{PK}(\boldsymbol{\nu})$

Corollary 2.4.5 (Fréchet space topology).
$\mathcal{PK}(\boldsymbol{\nu})$ with the topology given by the family of semi-norms arising as the lower bounds of the constants in (2.3.3) *and* (2.3.4) *is a Fréchet space.*

Proof. The countable family of semi-norms separates points. All we need to prove is the metric completeness i.e. that all Cauchy sequences in $\mathcal{PK}(\boldsymbol{\nu})$ converge to a kernel in $\mathcal{PK}(\boldsymbol{\nu})$ in the strong topology.

Consider a Cauchy sequence $K_n \in \mathcal{PK}(\boldsymbol{\nu})$. By proposition 2.4.4 the strong topology on $\mathcal{PK}(\boldsymbol{\nu})$ is finer than the weak-* topology so, for any test function φ on $\mathbb{R}^N$, the sequence $\int K_n(\boldsymbol{x})\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ is also a Cauchy sequence. The relation

$$\int K(\boldsymbol{x})\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \lim_{n\to\infty} \int K_n(\boldsymbol{x})\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$$

defines an element $K \in S'(\mathbb{R}^N)$ that is the weak-* limit of K_n. A Cauchy sequence in $\mathcal{PK}(\boldsymbol{\nu})$ is necessarily bounded and so its weak-* closure is a subset of $\mathcal{PK}(\boldsymbol{\nu})$ and so $K_n \rightharpoonup K \in \mathcal{PK}(\boldsymbol{\nu})$. By lower semi-continuity of the semi-norms we get that $K_n \to K$ strongly in $\mathcal{PK}(\boldsymbol{\nu})$ and this concludes the proof. □

2.4.2 Approximation

A very useful method to study singular integrals is approximating kernels in the weak-* topology by smooth functions (possibly with compact support). This is usually done via convolution with an approximate identity and possibly via multiplication by a cutoff function. Since product kernels are in particular tempered distributions, the convolution with any approximate identity yields a C^∞ approximating sequence in the distributional sense. However, taking in account that $\mathcal{PK}(\boldsymbol{\nu})$ has also a strong topology on $\mathcal{PK}(\boldsymbol{\nu})$, we are encouraged to develop an approximation procedure that is well-behaved with respect to the strong topology. In particular we will aim to approximate product kernels with C^∞ functions that are uniformly bounded distributions in $\mathcal{PK}(\boldsymbol{\nu})$.

It is important to notice that the approximation properties for kernels in $\mathcal{PK}(\boldsymbol{\nu})$ depend on the order $\boldsymbol{\nu}$. In particular important properties depend on whether $\nu_k > 0$, $\nu_k < 0$, or $\nu_k > -q_k$ for any given $k \in \{1, \dots, d\}$

Lemma 2.4.6 (Approximation for $\nu_k > -q_k$)**.**
Let $G = \{g \in \{1, \dots, d\} \mid \nu_g \geq -q_g\}$ and let $(\boldsymbol{\varphi}_n)_{n\in N}$, be a compactly supported product approximate identity on H_G obtained by rescaling an L_1-normalized product cutoff function as described in (2.1.16)*. Then $\boldsymbol{\varphi}_n *_G K$ is a sequence of product kernels uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu})$ that coincide with smooth functions away from $\bigcup_{k\notin G} H_k^\perp$ and $\boldsymbol{\varphi}_n *_G K \rightharpoonup K$.*

Proof. By the general theory of distributions $\boldsymbol{\varphi}_n *_G K$ weak-* converge to K since for any $\psi \in S(\mathbb{R}^N)$, $\check{\varphi}_n *_G \psi \to \psi$ in $S(\mathbb{R}^N)$. The distributions $\varphi_n *_G K$ coincide with C^∞ functions away from $\bigcup_{k\notin G} H_k^\perp$. As a matter of fact the mappings

$$\boldsymbol{x}_G \mapsto \int_{H_G} \boldsymbol{\varphi}_n(\boldsymbol{x}_G - \boldsymbol{y}_G)K(\boldsymbol{x}_{(G)} \oplus \boldsymbol{y}_G)\mathrm{d}\boldsymbol{y}_G \quad \text{and} \quad \boldsymbol{x}_{(G)} \mapsto \int_{H_G} \boldsymbol{\varphi}_n(\boldsymbol{x}_G - \boldsymbol{y}_G)K(\boldsymbol{x}_{(G)} \oplus \boldsymbol{y}_G)\mathrm{d}\boldsymbol{y}_G$$

are smooth for all $\boldsymbol{x}_G$ and all $\boldsymbol{x}_{(G)} \notin \bigcup_{k\notin G} H_k^\perp$. The smoothness in the variable $\boldsymbol{x}_G$ is the standard result about smooth dependence of a test function - distribution duality pairing on the translation of the test function. The smoothness in $\boldsymbol{x}_{(G)}$ comes from the fact that by definition

$$\int_{H_G} \boldsymbol{\varphi}_n(\boldsymbol{x}_G - \boldsymbol{y}_G)K(\boldsymbol{x}_{(G)} \oplus \boldsymbol{y}_G)\,\mathrm{d}\boldsymbol{y}_G$$

is a product kernel in $\mathcal{PK}_{\{H_{(G)}\}}(\boldsymbol{\nu}_{(G)})$ and thus is smooth away from $\bigcup_{k\notin G} H_k^\perp$. Thus the mapping

$$\boldsymbol{x}_G \oplus \boldsymbol{x}_{(G)} \mapsto \int_{H_G} \boldsymbol{\varphi}_n(\boldsymbol{x}_G - \boldsymbol{y}_G)K(\boldsymbol{x}_{(G)} \oplus \boldsymbol{y}_G)\,\mathrm{d}\boldsymbol{y}_G$$

is a mapping $\mathbb{R}^N \setminus \bigcup_{k\notin G} H_k^{\perp} \to \mathbb{C}$ that is separately smooth along the multi-coordinates $\boldsymbol{x}_G$ and $\boldsymbol{x}_{(G)}$. However by explicitly calculating the limit we have that

$$\partial_{\boldsymbol{x}_G}^{\boldsymbol{\alpha}_G} \left(\int_{H_G} \boldsymbol{\varphi}_n(\boldsymbol{x}_G - \boldsymbol{y}_G) K(\boldsymbol{x}_{(G)} \oplus \boldsymbol{y}_G) \, \mathrm{d}\boldsymbol{y}_G \right) = \left(\int_{H_G} \partial_{\boldsymbol{x}_G}^{\boldsymbol{\alpha}_G} \boldsymbol{\varphi}_n(\boldsymbol{x}_G - \boldsymbol{y}_G) K(\boldsymbol{x}_{(G)} \oplus \boldsymbol{y}_G) \, \mathrm{d}\boldsymbol{y}_G \right).$$

Since $\partial_{\boldsymbol{x}_G}^{\boldsymbol{\alpha}_G} \boldsymbol{\varphi}_n(\boldsymbol{x}_G - \boldsymbol{\cdot})$ is a tensor product of bump functions on H_G

$$\boldsymbol{x}_{(G)} \mapsto \int_{H_G} \partial_{\boldsymbol{x}_G}^{\boldsymbol{\alpha}_G} \boldsymbol{\varphi}_n(\boldsymbol{x}_G - \boldsymbol{y}_G) K(\boldsymbol{x}_{(G)} \oplus \boldsymbol{y}_G) \, \mathrm{d}\boldsymbol{y}_G$$

is continuous and thereby the whole mapping $\varphi_n *_G K$ is continuous on the need domain.

We now verify that $\varphi_n * K$ are uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu})$.

We reason by induction on the length d of the product decomposition of $\mathbb{R}^N$. For $d = 0$ there is nothing to prove. Suppose $\mathbb{R}^N = \oplus_{k\in\{1,\ldots,d\}} H_k$ and the above proposition is true for all decompositions of up to $d-1$ terms. For ease of notation we indicate $r = 2^{-n}$ and recall that, as explained in (2.1.16), $\boldsymbol{\varphi}_n$ is the tensor product of renormalized cutoff functions.

Size conditions (2.3.3) To check the size conditions we distinguish the cases when $\|x_k\| \lesssim r$ and $\|x_k\| >\!\!\gg r$. For any $\boldsymbol{x} \in \mathbb{R}^N$ let $J_{\boldsymbol{x}} = \left\{ j \in G \,\middle|\, \left\|x_j\right\| \leq cr \right\}$ for some sufficiently large $c > 0$ to be chosen subsequently depending on the triangular constant in (1.2.1) and let the complimentary set restricted to the indexes in G be $J_{\boldsymbol{x}}^c = G \setminus J_{\boldsymbol{x}}$. Depending on the value of $\boldsymbol{x}$ we write

$$\begin{aligned} \varphi_n *_G K(\boldsymbol{x}) = &\int_{H_G} \varphi_n(\boldsymbol{x}_G - \boldsymbol{y}_G) K\left(\boldsymbol{x}_{(G)} \oplus \boldsymbol{y}_G\right) \, \mathrm{d}\boldsymbol{y}_G = \\ &\int_{H_{J_{\boldsymbol{x}}^c}} \left(\int_{H_{J_{\boldsymbol{x}}}} K\left(\boldsymbol{x}_{(G)} \oplus \boldsymbol{y}_G\right) \prod_{j\in J_{\boldsymbol{x}}} r^{-q_j} \eta_j \left(r^{-1} \cdot (x_j - y_j)\right) \, \mathrm{d}\boldsymbol{y}_{J_{\boldsymbol{x}}} \right) \\ &\qquad \prod_{k\in J_{\boldsymbol{x}}^c} r^{-q_k} \eta_k \left(r^{-1} \cdot (x_k - y_k)\right) \, \mathrm{d}\boldsymbol{y}_{J_{\boldsymbol{x}}^c}. \end{aligned}$$

We have that $c > 0$ can be chosen sufficiently large so that $\varphi_n(\boldsymbol{x}_G - \boldsymbol{y}_G)$ as a function of $\boldsymbol{y}_G$ is supported away from $\bigcup_{k\in J_{\boldsymbol{x}}^c} H_k^{\perp}$ and in particular so that

$$\operatorname{spt} \eta_k \left(r^{-1} \cdot (x_k - y_k)\right) \subset \left\{ y_k \,\middle|\, r < \|y_k\| < C \|x_k\| \right\}$$

for some large $C > 0$. Vice-versa for $j \in J_{\boldsymbol{x}}$, we have that φ_n is supported close to the respective coordinate sub-planes:

$$\operatorname{spt} \eta_j \left(r^{-1} \cdot (x_j - y_j)\right) \subset \left\{ y_j \,\middle|\, \left\|y_j\right\| < C' r \right\}$$

for some sufficiently large $C' > 0$. However this means that $\eta_j(r^{-1} \cdot x_j - C' \cdot y_j)$ is, up to a constant, a b_j-normalized bump function in the y_j variable.

Notice that

$$\partial^{\boldsymbol{\alpha}} \left(\varphi_n *_G K\right) = \left(\partial^{\boldsymbol{\alpha}_{J_{\boldsymbol{x}}}} \varphi_n\right) *_G \left(\partial^{\boldsymbol{\alpha}_{(J_{\boldsymbol{x}})}} K\right)$$

and

$$\partial^{\boldsymbol{\alpha}_{J_{\boldsymbol{x}}}} \boldsymbol{\varphi}_n = r^{-\boldsymbol{q} - \left\|\boldsymbol{\alpha}_{J_{\boldsymbol{x}}}\right\|} (\partial^{\boldsymbol{\alpha}_{J_{\boldsymbol{x}}}} \boldsymbol{\eta})(r^{-1} \cdot \boldsymbol{x}) =$$

$$= \otimes_{j\in J_{\boldsymbol{x}}} r^{-q_j-\|\alpha_j\|}(\partial^{\alpha_j}\eta_j)(r^{-1}\cdot x_j) \otimes_{k\in J^c_{\boldsymbol{x}}} r^{-q_k}\eta_k(r^{-1}\cdot x_k).$$

By much the same reasoning as above $(\partial^{\alpha_j}\eta)(r^{-1}\cdot x_j - C'\cdot y_j)$ are, up to a constant, b_j-normalized bump functions in the y_j variable for $j \in J_{\boldsymbol{x}}$ while $(\partial^{\alpha_k}\eta)\left(r^{-1}\cdot(x_j - \cdot y_k)\right)$ are supported away from $H_k^{\perp}$ for $k \in J^c_{\boldsymbol{x}}$

Using the cancellation conditions (2.3.4) along H_j with $j \in J_{\boldsymbol{x}}$ for any $\boldsymbol{\alpha}$ we get that

$$\prod_{j\in J_{\boldsymbol{x}}} r^{q_j+\nu_j+\|\alpha_j\|}\int_{H_{J_{\boldsymbol{x}}}} K\left(\boldsymbol{x}_{(G)}\oplus \boldsymbol{y}_G\right)\prod_{j\in J_{\boldsymbol{x}}} r^{-q_j-\|\alpha_j\|}\left(\partial^{\alpha_j}\eta_j\right)\left(r^{-1}\cdot(x_j-y_j)\right)\,\mathrm{d}\boldsymbol{y}_{J_{\boldsymbol{x}}}$$

are a uniformly bounded family of product kernels in $\mathcal{PK}(\boldsymbol{\nu}_{(J_{\boldsymbol{x}})})$. By the previous remarks $r^{-q_k-\|\alpha_k\|}\left(\partial^{\alpha_k}\eta_k\right)\left(r^{-1}\cdot(x_k-y_k)\right)$ are supported away from the coordinate sub-planes so we get by integrating

$$\begin{aligned}
\left|\partial^{\boldsymbol{\alpha}}\left(\varphi_n *_G K\right)(\boldsymbol{x})\right| &= \left|\left(\partial^{\boldsymbol{\alpha}_{J_{\boldsymbol{x}}}}\varphi_n\right) *_G \left(\partial^{\boldsymbol{\alpha}_{(J_{\boldsymbol{x}})}}K\right)(\boldsymbol{x})\right| \leq \\
&C''\prod_{i\notin G}\|x_i\|^{-q_i-\nu_i-\|\alpha_i\|}\prod_{j\in J_{\boldsymbol{x}}} r^{-q_j-\nu_j-\|\alpha_j\|} \\
&\quad \prod_{k\in J^c_{\boldsymbol{x}}}\int_{r<\|y_k\|<C\|x_k\|}\|y_k\|^{-q_k-\nu_k-\|\alpha_k\|}\, r^{-q_k}\eta_k\left(r^{-1}\cdot(x_k-y_k)\right)\,\mathrm{d}y_k \leq \\
&C''\prod_{i\notin G}\|x_i\|^{-q_i-\nu_i-\|\alpha_i\|}\prod_{j\in J_{\boldsymbol{x}}}\|x_j\|^{-q_j-\nu_j-\|\alpha_j\|}\prod_{k\in J^c_{\boldsymbol{x}}}\|x_k\|^{-q_k-\nu_k-\|\alpha_k\|} \leq \\
&C''\prod_{k\in\{1,\dots,d\}}\|x_k\|^{-q_k-\nu_k-\|\alpha_k\|}.
\end{aligned}$$

These are the necessary size conditions.

Cancellation conditions (2.3.3) We reason by induction. We need to prove that for any $k \in \{1,\dots,d\}$ the family of kernels

$$R^{\nu_k}\int_{H_k}\boldsymbol{\varphi}_n *_G K(\boldsymbol{x})\psi(R^{-1}\cdot x_k)\,\mathrm{d}x_k$$

is uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu}_{(k)})$ for all b_k-normalized bump functions ψ on H_k and for all $R > 0$.

If the cancellation occurs along H_k with $k \notin G$ it suffices to notice that

$$R^{\nu_k}\int_{H_k}\boldsymbol{\varphi}_n *_G K(\boldsymbol{x})\psi(R^{-1}\cdot x_k)\,\mathrm{d}x_k = \boldsymbol{\varphi}_n *_G \left(R^{\nu_k}\int_{H_k}K(\boldsymbol{x})\psi(R^{-1}\cdot x_k)\,\mathrm{d}x_k\right).$$

The term inside the parenthesis is uniformly bounded so it suffices to apply the induction hypotheses.

If, on the other hand, the cancellation occurs along H_k with $k \in G$, indicating $\eta_g^{(r)}(x_g) = r^{-q_g}\eta_g(r^{-1}\cdot x_g)$ one can write

$$\begin{aligned}
&R^{\nu_k}\int_{H_k}\boldsymbol{\varphi}_n *_G K(\boldsymbol{x})\psi(R^{-1}\cdot x_k)\mathrm{d}x_k = \\
&\qquad\left(\bigotimes_{g\in(G\setminus\{k\})}\eta_g^{(r)}\right) *_{G\setminus\{k\}} R^{\nu_k}\int_{H_k}\eta_k^{(r)} *_{\{k\}} K(\boldsymbol{x})\psi(R^{-1}\cdot x_k)\,\mathrm{d}x_k =
\end{aligned}$$

$$\left(\bigotimes_{g\in(G\setminus\{k\})}\eta_g^{(r)}\right) *_{G\setminus\{k\}} R^{\nu_k}\int_{H_k} K(\boldsymbol{x})\left(\eta_k^{(r/R)} *_{\{k\}}\psi\right)(R^{-1}\cdot x_k)\,\mathrm{d}x_k =$$
$$\left(\bigotimes_{g\in(G\setminus\{k\})}\eta_g^{(r)}\right) *_{G\setminus\{k\}} R^{\nu_k+q_k}r^{-q_k}\int_{H_k} K(\boldsymbol{x})\left(\eta_k *_{\{k\}}\psi^{(R/r)}\right)(r^{-1}\cdot x_k)\,\mathrm{d}x_k.$$

Now it suffices to notice that if $r \leq R$ then $\left(\eta_k^{(r/R)} *_{\{k\}}\psi\right)(2\cdot x_k)$ is a b_k-normalized bump function and so

$$R^{\nu_k}\int_{H_k} K(\boldsymbol{x})\left(\eta_k^{(r/R)} *_{\{k\}}\psi\right)(R^{-1}\cdot x_k)\,\mathrm{d}x_k$$

is a uniformly bounded family of product kernel. Otherwise, if $r > R$, we have that $\left(\eta_k *_{\{k\}}\psi^{(R/r)}\right)(2\cdot x_k)$ is a b_k-normalized bump function and so

$$R^{\nu_k+q_k}r^{-q_k}\int_{H_k} K(\boldsymbol{x})\left(\eta_k *_{\{k\}}\psi^{(R/r)}\right)(r^{-1}\cdot x_k)\,\mathrm{d}x_k$$

is uniformly bounded since $\nu_k > -q_k$. The proof follows by the inductions hypothesis.

□

This second lemma addresses the question of approximation along those subspaces H_k for which $\nu_k < 0$.

Lemma 2.4.7 (Approximation for $\nu_k < 0$).
Let $K \in \mathcal{PK}(\boldsymbol{\nu})$ be a product kernel and let $J = \{j \in \{1,\dots,d\} \mid \nu_j < 0\}$. Then K can be approximated in the weak- topology by a uniformly bounded sequence of product kernels $K_n(\boldsymbol{x})$ in $\mathcal{PK}(\boldsymbol{\nu})$ with $n \in \mathbb{N}$ that are supported away from $\bigcup_{j\in J} H_j^\perp$. In particular* $\operatorname{spt} K_n \subset \{\boldsymbol{x} \mid \|x_j\| > 2^{-n-1}\}$ *for all $j \in J$ and K_n is smooth away from $\bigcup_{k\notin J} H_k^\perp$.*

Proof. Consider a set of cutoff functions η_j on H_j for each $j \in J$. We prove that

$$K_n(\boldsymbol{x}) \stackrel{\text{def}}{=} \prod_{j\in J}(1-\eta_j)(2^{-n}\cdot x_j)K(\boldsymbol{x})$$

are uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu})$ and $K_n \rightharpoonup K$. More generally this holds for any subset of indexes J as long as $\nu_j < 0$ for all $j \in J$.

We begin by proving the uniform boundedness. As usual we proceed by induction on the number of term d in the product decomposition on $\mathbb{R}^N$. For $d = 0$ there is nothing to prove. Now suppose that $d > 1$ and the boundedness is true for all product decompositions of order up to $d-1$. For ease of notation we indicate $r = 2^{-n}$

Size condition (2.3.3) Write

$$\partial^{\boldsymbol{\alpha}} K_n(\boldsymbol{x}) = \sum_{\boldsymbol{\beta}+\boldsymbol{\gamma}=\boldsymbol{\alpha}_J}\partial^{\boldsymbol{\beta}}\left(\prod_{j\in J}(1-\eta_j)(r^{-1}\cdot x_j)\right)\partial^{\boldsymbol{\gamma}}\partial^{\boldsymbol{\alpha}_{(J)}}K_n(\boldsymbol{x}) =$$
$$\sum_{\boldsymbol{\beta}+\boldsymbol{\gamma}=\boldsymbol{\alpha}_J}\prod_{j\in J}\left(r^{-|\beta_j|}\left(\partial^{\beta_j}(1-\eta_j)\right)(r^{-1}\cdot x_j)\partial^{\gamma_j}\partial^{\boldsymbol{\alpha}_{(J)}}K_n(\boldsymbol{x})\right).$$

If $\beta_j \neq 0$ then $\left(\partial^{\beta_j}(1-\eta_j)\right)(r^{-1}\cdot x_j)$ is supported on $r/2 \leq \|x_j\| \leq r$ while if $\beta_j = 0$ the bound is trivial so we have

$$\left|\partial^{\boldsymbol{\alpha}} K_n(\boldsymbol{x})\right| < C \prod_{k\in\{1,\dots,d\}} \|x_k\|^{-q_k-\nu_k-|\alpha_k|}$$

as required.

Cancellation conditions (2.3.4) Suppose that the cancellation occurs along H_k, so we must prove that for any b_k-normalized bump function φ on H_k and for any $R > 0$

$$R^{\nu_k}\int_{H_k}\prod_{j\in J}(1-\eta_j)(r^{-1}\cdot x_j)K(\boldsymbol{x})\varphi_k(R^{-1}\cdot x_k)\,\mathrm{d}x_k$$

is uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu}_{(k)})$. If $k \notin J$ this is trivial since

$$R^{\nu_k}\int_{H_k}\prod_{j\in J}(1-\eta_j)(r^{-1}\cdot x_j)K(\boldsymbol{x})\varphi_k(R^{-1}\cdot x_k)\,\mathrm{d}x_k =$$
$$\prod_{j\in J}(1-\eta_j)(r^{-1}\cdot x_j)R^{\nu_k}\int_{H_k}K(\boldsymbol{x})\varphi_k(R^{-1}\cdot x_k)\,\mathrm{d}x_k$$

but $R^{\nu_k}\int_{H_k}K(\boldsymbol{x})\varphi_k(R^{-1}\cdot x_k)\,\mathrm{d}x_k$ is uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu}_{(k)})$ and is a product kernel on $H_{(k)}$, a space with a product decomposition of $d-1$ factors. Using the inductive hypothesis we get the required boundedness.

If $k \in J$ then we write

$$R^{\nu_k}\int_{H_k}\prod_{j\in J}(1-\eta_j)(r^{-1}\cdot x_j)K(\boldsymbol{x})\varphi_k(R^{-1}\cdot x_k)\,\mathrm{d}x_k =$$
$$\prod_{j\in J\setminus\{k\}}(1-\eta_j)(r^{-1}\cdot x_j)R^{\nu_k}\int_{H_k}K(\boldsymbol{x})\varphi_k(R^{-1}\cdot x_k)(1-\eta_k)(r^{-1}\cdot x_k)\,\mathrm{d}x_k$$

Notice that $\varphi_k(x_k)(1-\eta)(Rr^{-1})$ is a 0-normalized bump function (b_k=0). So

$$R^{\nu_k}\int_{H_k}K(\boldsymbol{x})\varphi_k(R^{-1}\cdot x_k)(1-\eta_k)(r^{-1}\cdot x_k)\,\mathrm{d}x_k$$

is uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu}_{(k)})$ and the boundedness follows again from the induction hypothesis.

Now we prove that $K_n \rightharpoonup K$. We have that

$$K(\boldsymbol{x}) - K_n(\boldsymbol{x}) = \sum_{\widetilde{J}\subsetneq J}\prod_{k\in J\setminus\widetilde{J}}\eta_k(r^{-1}\cdot x_k)\prod_{j\in\widetilde{J}}(1-\eta_j)(r^{-1}\cdot x_j)K(\boldsymbol{x}).$$

The terms

$$\prod_{k\in J\setminus\widetilde{J}}\eta_k(r^{-1}\cdot x_k)\prod_{j\in\widetilde{J}}(1-\eta_j)(r^{-1}\cdot x_j)K(\boldsymbol{x})$$

are all uniformly bounded for $r > 0$ in $\mathcal{PK}(\boldsymbol{\nu})$ by induction on $|J\setminus\widetilde{J}|$. We will now prove that each term of the sum weak-* converges to 0. For any given term it is sufficient to prove that convergence to 0 holds for tensor products of compactly supported test functions. This is true due to Proposition 2.4.4 since linear combinations of tensor products of test functions are dense in $S(\mathbb{R}^N)$ and the the terms are uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu})$. For

$$\boldsymbol{\varphi}(\boldsymbol{x}) = \bigotimes_{k=1}^{d} \varphi_k(x_k)$$

we have that for any $\widetilde{J} \subsetneq J$

$$\int_{\mathbb{R}^N} \prod_{k \in J\setminus\widetilde{J}} \eta_k(r^{-1} \cdot x_k) \prod_{j \in \widetilde{J}} (1-\eta_j)(r^{-1} \cdot x_j) K(\boldsymbol{x}) \boldsymbol{\varphi}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \to 0$$

because $\varphi_k(r \cdot x_k)\eta_k(x_k)$ are 0-normalized bump functions and so, as $r \to 0$,

$$\left(\prod_{k \in J\setminus\widetilde{J}} r^{\nu_k} \right) \int_{H_{J\setminus\widetilde{J}}} \prod_{k \in J\setminus\widetilde{J}} \eta_k(r^{-1} \cdot x_k)\varphi_k(x_k) \prod_{j \in \widetilde{J}} (1-\eta_j)(r^{-1} \cdot x_j) K(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}_{J\setminus\widetilde{J}}$$

are uniformly bounded.

□

From these two lemmas we obtain several useful results about approximating kernels in $\mathcal{PK}(\boldsymbol{\nu})$

Corollary 2.4.8 (Approximation by smooth functions)**.**
Any kernel $K \in \mathcal{PK}(\boldsymbol{\nu})$ can be weak- approximated by a sequence of $C^\infty(\mathbb{R}^N)$ functions that are uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu})$.*

Proof. Let $K \in \mathcal{PK}(\boldsymbol{\nu})$ and let $J = \{j \in \{1, \ldots, d\} \mid \nu_j < -q_j\}$. Using lemma 2.4.7 construct an approximating sequence $K_n \rightharpoonup K$ of kernels supported away from $\bigcup_{j \in J} H_j^\perp$. For every K_n construct an approximating sequence $K_{n,m} \rightharpoonup K_n$ by convolving with a product approximate identity on H_{J^c}. The kernels $K_{n,m}$ are uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu})$ and coincide with smooth functions away from $\bigcup_{j \in J} H_j^\perp$. However they are also supported away from $\bigcup_{j \in J} H_j^\perp$ since convolving along H_{J^c} doesn't alter this property. Thus $K_{n,m}$ are smooth on the whole $\mathbb{R}^N$. By a diagonal argument one extracts a sequence weakly-* converging to K. □

Finally we note a basic approximation property for product kernels with kernels of bounded support. Since we are dealing with a product structure it makes sense to distinguish the properties of the support along different subspaces.

Definition 2.4.9 (Kernels with bounded support along subspaces)**.**
We say a kernel K has bounded or compact support along H_k if $\pi_k(\operatorname{spt} K)$ is compact in H_k, where π_k is the projection operator on the subspace H_k of $\mathbb{R}^N$. Equivalently K has compact support along H_k if $\operatorname{spt} K \subset \{\boldsymbol{x} \mid \|x_k\| < R\}$ for some $R > 0$.

A rudimentary approximation property with kernels of bounded support is given by the following proposition. A more fine result holds allowing kernels of arbitrary order to be approximated with compactly supported smooth function uniformly, however the proof involves more delicate results on the structure and cancellation properties of kernels of $\mathcal{PK}(\boldsymbol{\nu})$.

Lemma 2.4.10 (Approximation with kernels with bounded support)**.**
Let K be a product kernel of order $\boldsymbol{\nu}$. If $\nu_k \leq 0$ for some k then K can be uniformly approximated with product kernels in $\mathcal{PK}(\boldsymbol{\nu})$ with compact support along H_k. In other words, there exists a sequence K_n uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu})$ such that $K_n \rightharpoonup K$ in $S'(\mathbb{R}^N)$ as $n \to \infty$ and such that the support of every K_n is bounded along H_k for every k such that $\nu_k \leq 0$.

Proof. Let J be the set of indexes k such that $\nu_k \leq 0$. For every $j \in J$ consider a cutoff function $\eta_j(x_j)$ on H_j and let $\boldsymbol{\eta}(\boldsymbol{x}) = \bigotimes_{j\in J} \eta_j(x_j)$ be a product cutoff function on H_J. Setting $\boldsymbol{r} = (r, \dots, r)$ for $r > 0$, we have that $K_r(\boldsymbol{x}) \stackrel{\text{def}}{=} K(\boldsymbol{x})\boldsymbol{\eta}(\boldsymbol{r}^{-1} \cdot \boldsymbol{x}) \rightharpoonup K(\boldsymbol{x})$ in $S'(\mathbb{R}^N)$ as $r \to +\infty$ and every K_r has bounded support along all H_j for $j \in J$.

We will pass to proving the uniform boundedness of the sequence K_r in $\mathcal{PK}(\boldsymbol{\nu})$. We will prove this inductively on the number d of terms in the product decomposition of $\mathbb{R}^N$. For $d = 0$ there is nothing to prove. Now suppose that $\mathbb{R}^N = \bigoplus_{k\in\{1,\dots,d\}} H_k$ and the above proposition holds for any product kernel on a product space with less than d terms.

Size conditions (2.3.3) We check the size conditions directly

$$\partial^{\boldsymbol{\alpha}} K_r(\boldsymbol{x}) = \sum_{\boldsymbol{\beta}+\boldsymbol{\gamma}=\boldsymbol{\alpha}_J} \partial^{\boldsymbol{\beta}}\partial^{\boldsymbol{\alpha}_{(J)}} K(\boldsymbol{x}) r^{-\|\boldsymbol{\gamma}\|} \left(\partial^{\boldsymbol{\gamma}}\boldsymbol{\eta}\right)(\boldsymbol{r}^{-1} \cdot \boldsymbol{x}).$$

On the support of $\boldsymbol{\eta}(\boldsymbol{r}^{-1} \cdot \boldsymbol{x})$, we have that $\|x_k\| \leq r$ so for every addend we can write

$$\left|\partial^{\boldsymbol{\beta}} K(\boldsymbol{x}) r^{-\|\boldsymbol{\gamma}\|} \left(\partial^{\boldsymbol{\gamma}}\boldsymbol{\eta}\right)(\boldsymbol{r}^{-1} \cdot \boldsymbol{x})\right| \leq C \|x_1\|^{-q_1-\nu_1-\|\beta_1\|-\|\gamma_1\|} \dots \|x_d\|^{-q_d-\nu_d-\|\beta_d\|-\|\gamma_d\|}$$

as required.

Cancellation conditions (2.3.4) Multiplication by a cutoff function and the duality pairing along different subspaces commute. In particular if the cancellation occurs along H_k with $k \notin J$ then for any b_k-normalized bump function on H_k and for any $R > 0$ we write

$$\int_{H_k} K_r(\boldsymbol{x})\varphi(R^{-1} \cdot x_k)\,\mathrm{d}x_k = \boldsymbol{\eta}(\boldsymbol{r}^{-1} \cdot \boldsymbol{x}) \int_{H_k} K(\boldsymbol{x})\varphi(R^{-1} \cdot x_k)\,\mathrm{d}x_k.$$

But $\int_{H_k} K(\boldsymbol{x})\varphi(R^{-1}\cdot x_k)\,\mathrm{d}x_k$ is uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu}_{(k)}$ by the cancellation property of K and we can directly use the induction hypothesis to have the required boundedness.

If the cancellation occurs along H_k with $k \in J$ then let $\widetilde{\boldsymbol{\eta}}(\boldsymbol{x}) = \bigotimes_{\substack{j\in J\\ j\neq k}} \eta_j(x_j)$. We then write

$$R^{\nu_k} \int_{H_k} K_r(\boldsymbol{x})\varphi(R^{-1} \cdot x_k)\,\mathrm{d}x_k = \widetilde{\boldsymbol{\eta}}(\boldsymbol{x}) R^{\nu_k} \int_{H_k} K(\boldsymbol{x})\eta_k(r^{-1} \cdot x_k)\varphi(R^{-1} \cdot x_k)\,\mathrm{d}x_k.$$

Notice that for $2R < r$ we have that $\varphi(R^{-1}x_k)$ is supported where $\eta_k(r^{-1}x_k)$ is constantly equal to 1. In particular, for any b_k-normalized bump function φ on H_k and any $0 < R < {}^{r}\!/_{2}$, the family of kernels

$$R^{\nu_k} \int_{H_k} K(\boldsymbol{x})\eta_k(r^{-1} \cdot x_k)\varphi(R^{-1} \cdot x_k)\,\mathrm{d}x_k$$

is uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu}_{(k)})$. Then by the inductive hypothesis the above kernels are uniformly bounded.

Finally if $2R \geq r$ then $\eta_k(x_k)\varphi(rR^{-1} \cdot x_k)$ are b_k-normalized bump functions so

$$r^{\nu_k} \int_{H_k} K(\boldsymbol{x})\eta_k(r^{-1} \cdot x_k)\varphi(R^{-1} \cdot x_k)\,\mathrm{d}x_k$$

are uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu}_{(k)})$. Using the assumption on R and r and that $\nu_k \leq 0$ we have that

$$R^{\nu_k} \int_{H_k} K(\boldsymbol{x})\eta_k(r^{-1} \cdot x_k)\varphi(R^{-1} \cdot x_k)\,\mathrm{d}x_k$$

is also uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu}_{(k)})$. The proof is concluded using the induction hypothesis.

□

2.4.3 Bump function normalization orders

An other very useful fact is that the cancellation conditions 2.3.4 are tolerant on how the bump functions are normalized. In particular the bump functions φ in definition 2.3.1 can be normalized with respect to any C^b norm with b large enough. More precisely if in the cancellation condition 2.3.4 in definition 2.3.1 we ask that the bump functions be $\widetilde{b}_k$-normalized with some other order $\widetilde{b}_k \in \mathbb{N}$ greater than $b_k = \min\{b \in \mathbb{N} \mid b > \nu_k\}$ then the resulting class $\widetilde{\mathcal{PK}}(\boldsymbol{\nu})$ coincides with $\mathcal{PK}(\boldsymbol{\nu})$ and the respective semi-norms and topology are equivalent to the original ones.

One inclusion and inequality between norms is obvious: the new class is per-se larger and the topology coarser. The other inclusion and inequality follow from the next lemma.

Proposition 2.4.11 (Bump function normalization orders)**.**
Let $\widetilde{\mathcal{PK}}(\boldsymbol{\nu})$ be the class of product kernels of order $\boldsymbol{\nu}$ with the exception that in the definition 2.3.1 we require that bump function normalization orders in condition (2.3.4) *be $\widetilde{b}_k \geq b(\nu_k)$. Then the classes $\widetilde{\mathcal{PK}}(\boldsymbol{\nu})$ and $K \in \mathcal{PK}(\boldsymbol{\nu})$ coincide as vector spaces and have the same topologies.*

Proof. The inclusion $\mathcal{PK}(\boldsymbol{\nu}) \subset \widetilde{\mathcal{PK}}(\boldsymbol{\nu})$ is obvious. As a matter of fact with $\widetilde{b}_k \geq b_k$ the conditions on $\mathcal{PK}(\boldsymbol{\nu})$ are more restrictive than on $\widetilde{\mathcal{PK}}(\boldsymbol{\nu})$. The constants in the conditions (2.3.3) and (2.3.4) are larger for $\mathcal{PK}(\boldsymbol{\nu})$ and so are the respective semi-norms. This means that the inclusion is continuous.

We prove the converse by induction on the number of factors d in the product decomposition of $\mathbb{R}^N$. For $d = 0$ there is nothing to prove. We indicate the class of product kernels with modified bump function normalization orders as $\widetilde{\mathcal{PK}}(\boldsymbol{\nu})$ and the original class simply as $\mathcal{PK}(\boldsymbol{\nu})$.

Suppose that the statement is true for d i.e. suppose that $\widetilde{\mathcal{PK}}(\boldsymbol{\nu}) = \mathcal{PK}(\boldsymbol{\nu})$ if the number of product terms in the decomposition of $\mathbb{R}^N$ is less or equal to d. Let $K \in \widetilde{\mathcal{PK}}(\boldsymbol{\nu})$ on $\mathbb{R}^N = \oplus_{k=1}^{d+1} H_k$. We need to check the cancellation condition 2.3.4 on K. Suppose that the cancellation occurs along H_k. We must check that

$$R^{\nu_k} \int K(\boldsymbol{x})\varphi(R^{-1} \cdot x_k)\mathrm{d}x_k$$

is an uniformly bounded family on $\mathcal{PK}(\boldsymbol{\nu}_{(k)})$ for all b_k-normalized bump functions φ on H_k with $b_k = \min(b \in \mathbb{N} \mid b > \nu_k)$. Since $K \in \widetilde{\mathcal{PK}}(\boldsymbol{\nu})$ the above quantity is uniformly bounded in $\widetilde{\mathcal{PK}}(\boldsymbol{\nu}_{(k)})$ if φ is $\widetilde{b}_k$-normalized. However the class $\widetilde{\mathcal{PK}}(\boldsymbol{\nu}_{(k)})$ coincides with the class $\mathcal{PK}(\boldsymbol{\nu})$ by the induction hypothesis, since kernels in $\widetilde{\mathcal{PK}}(\boldsymbol{\nu}_{(k)})$ are relative to decompositions with d factors. So it is sufficient to prove that for any $k \in \{1, \ldots, d\}$ and any b_k-normalized bump function the quantity (2.4.3) is uniformly bounded in $\widetilde{\mathcal{PK}}(\boldsymbol{\nu}_{(k}).$

By Lemma 2.4.1 we can write for any given b_k-normalized bump function φ on H_k the decomposition

$$\varphi(x_k) = \sum_{|\alpha_k| \leq b_k - 1} \frac{1}{\alpha_k!} \partial_k^{\alpha} \varphi(0) x_k^{\alpha_k} \eta(x_k) + \sum_{|\beta_k| = b_k} x_k^{\beta_k} \varphi_{\beta_k}(x_k).$$

Set $\widetilde{\varphi}(x_k) = \sum_{|\alpha_k| \leq b_k - 1} \frac{1}{\alpha_k!} \partial_{x_k}^{\alpha_k} \varphi(0) x^{\alpha} \eta(x_k)$. We have that $\widetilde{\varphi}$ is, up to a constant independent of φ, a $\widetilde{b}_k$-normalized bump function since it depends only on the first b_k derivatives of φ. Thus

$$R^{\nu_k} \int_{H_k} K(\boldsymbol{x})\widetilde{\varphi}(R^{-1} \cdot x_k)\mathrm{d}x_k$$

is a uniformly bounded family of $\widetilde{\mathcal{PK}}(\boldsymbol{\nu}_{(k)})$ kernels as required. Now we look at the remainder term

$$\sum_{|\beta_k|=b_k} R^{\nu_k} \int_{H_k} K(\boldsymbol{x}) R^{-\|\beta_k\|} x_k^{\beta_k} \varphi_\beta(R^{-1}\cdot x_k)\mathrm{d}x_k =$$

$$R^{\nu_k} \int_{H_k} K(\boldsymbol{x})\varphi(R^{-1}\cdot x_k)\mathrm{d}x_k - R^{\nu_k}\int_{H_k} K(\boldsymbol{x})\widetilde{\varphi}(R^{-1}\cdot x_k)\mathrm{d}x_k.$$

By Lemma 2.4.1 the functions $\varphi_\beta(R_0^{-1}\cdot x_k)$ are 0-normalized bump functions for some $R_0 > 0$ independent of φ. To check that this quantity is uniformly bounded in $\widetilde{\mathcal{PK}}(\boldsymbol{\nu}_{(k)})$ the key consideration is that since $\|\beta_k\| \geq b_k$, $x_k^\beta K(\boldsymbol{x})$ coincides with an L^1_{loc} function on $\mathbb{R}^N \setminus \bigcup_{j\neq k} H_j^\perp$. More specifically we look separately on the size (2.3.3) and cancellation (2.3.4) on the kernel.

Size We look separately at every term of the sum

$$R^{\nu_k}\int_{H_k} K(\boldsymbol{x}) R^{-\|\beta\|} x_k^\beta \; \varphi_\beta(R^{-1}\cdot x_k)\mathrm{d}x_k$$

away from the coordinate subspaces $H_j^\perp$. We see that $x_k^\beta K(\boldsymbol{x})$ is locally integrable across $H_k^\perp$ by rewriting

$$R^{\nu_k-\|\beta\|}\int_{H_k} K(\boldsymbol{x}) x_k^\beta \varphi_\beta(R^{-1}\cdot x_k)\mathrm{d}x_k =$$

$$R^{\nu_k-\|\beta\|}\int_{H_k} K(\boldsymbol{x}) x_k^\beta \eta(\rho^{-1}\cdot x_k)\varphi_\beta(R^{-1}\cdot x_k)\,\mathrm{d}x_k +$$

$$R^{\nu_k-\|\beta\|}\int_{H_k} K(\boldsymbol{x}) x_k^\beta \left(1-\eta(\rho^{-1}\cdot x_k)\right)\varphi_\beta(R^{-1}\cdot x_k)\,\mathrm{d}x_k$$

for a cutoff function η on H_k. Passing to the limit $\rho \to 0$ for the first term we get that

$$\int_{H_k} K(\boldsymbol{x}) x_k^\beta \eta(\rho^{-1}\cdot x_k)\varphi_\beta(R^{-1}\cdot x_k)\,\mathrm{d}x_k =$$

$$\rho^{-\nu_k+\|\beta\|}\rho^{\nu_k}\int_{H_k} K(\boldsymbol{x})(\rho^{-1}\cdot x_k)^\beta \eta(\rho^{-1}\cdot x_k)\varphi_\beta(\rho^{-1}\cdot {}^{\rho}\!/_{R}\cdot x_k)\,\mathrm{d}x_k \to 0$$

since $\|\beta\| - \nu_k > 0$ by construction, $x_k^\beta \eta(x_k)\varphi_\beta({}^{\rho}\!/_{R}\cdot x_k)$ is $\widetilde{b}_k$-normalized up to a constant that can depend on φ_β but not on ρ, and so

$$\rho^{\nu_k}\int_{H_k} K(\boldsymbol{x})(\rho^{-1}\cdot x_k)^\beta \eta(\rho^{-1}\cdot x_k)\varphi_\beta(\rho^{-1}\cdot {}^{\rho}\!/_{R}\cdot x_k)\,\mathrm{d}x_k$$

is uniformly bounded in $\widetilde{\mathcal{PK}}(\boldsymbol{\nu}_{(k)})$. This means that

$$R^{\nu_k-|\beta|}\int_{H_k} K(\boldsymbol{x}) x_k^\beta \varphi_\beta(R^{-1}\cdot x_k)\mathrm{d}x_k =$$

$$= \lim_{\rho\to 0} R^{\nu_k-|\beta|}\int_{H_k} K(\boldsymbol{x}) x_k^\beta \left(1-\eta(\rho^{-1}\cdot x_k)\right)\varphi_\beta(R^{-1}\cdot x_k)\,\mathrm{d}x_k$$

so by dominated convergence we have that for any multi-index $\boldsymbol{\alpha}$ on $\{H_j\}_{j\neq k}$

$$\left| R^{\nu_k-\|\beta\|}\int_{H_k} \partial^{\boldsymbol{\alpha}}_{\boldsymbol{x}_{(k)}} K(\boldsymbol{x}) x_k^\beta \varphi_\beta(R^{-1}\cdot x_k)\mathrm{d}x_k \right| <$$

$$< CR^{\nu_k - \|\beta\|} \prod_{j \neq k} \|\boldsymbol{x}_j\|^{-q_j-\nu_j-|\alpha_j|} \int_{B_{H_k}(0,R)} \left| \|x_k\|^{-q_k-\nu_k+|b_k|} \right| |\varphi_\beta|_{C^0} \, \mathrm{d}x_k <$$

$$< C \prod_{j \neq k} \|\boldsymbol{x}_j\|^{-q_j-\nu_j-|\alpha_j|}$$

for $\left(\boldsymbol{x}_{(k)}\right)_j \neq 0$. This proves the size estimates for every term of the sum.

Cancellation We prove the cancellation directly for

$$R_j^{\nu_k} \int_{H_k} K(\boldsymbol{x}) \varphi(R^{-1} \cdot x_k) \mathrm{d}x_k.$$

This follows by induction on d.

Let us fix $j \neq k$ and check the cancellation condition (2.3.4) along H_j. Let ψ be a $\widetilde{b}_j$-normalized bump function and $R_j > 0$ and we need to prove that

$$R_j^{\nu_j+\nu_k} \int_{H_j} \int_{H_k} K(\boldsymbol{x}) \varphi(R_k^{-1} \cdot x_k) \psi(R_j^{-1} \cdot x_j) \mathrm{d}x_k \mathrm{d}x_j$$

is uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu}_{(\{j,k\})})$ Contractions with functions along different subspaces commute so we first first apply the duality pairing along x_j and since ψ is $\widetilde{b}_j$-normalized we get that

$$R_j^{\nu_j} \int_{H_j} K(\boldsymbol{x}) \psi(R_j^{-1} \cdot x_j) \, \mathrm{d}x_j$$

is uniformly bounded in $\widetilde{\mathcal{PK}}(\boldsymbol{\nu}_{(j)})$. Since we are now dealing with a product kernel on $H_{(j)}$, a space with a $d-1$ factors in the product decomposition, the inductive hypothesis gives us the result.

□

2.5 Fourier transform

A fundamental result for dealing with product kernels is given by the characterization of the associated multipliers. As we mentioned in Section 1.2 on Calderón-Zygmund kernels, the properties of the Fourier transform also justify the definition we give for kernels of large negative multi-order.

In [NRS01, Theorem 2.1.11 p.38] it is shown that the Fourier transform of a product kernel is a product multiplier and vice-versa. It is easy to check that product multipliers defined in [NRS01, Definition 2.1.10 p.37] are none other than product kernels of order $\boldsymbol{\nu} = -\boldsymbol{q} = (-q_1, \ldots, -q_d)$. We extend this result to kernels of arbitrary order and refine it according to the product structure.

Definition 2.5.1.
Let $\mathbb{R}^N$ be a product space. Define the partial Fourier transform along H_k to be $\mathcal{F}_k : S(\mathbb{R}^N) \to S(\mathbb{R}^N)$ such that

$$\mathcal{F}_k \varphi(x_1, \ldots, x_{k-1}, \xi_k, x_{k+1}, \ldots, x_d) = \int_{H_k} \varphi(\boldsymbol{x}) e^{-i\xi_k x_k} \mathrm{d}x_k.$$

Accordingly define the partial Fourier transform for distributions so that

$$\int_{\mathbb{R}^N} (\mathcal{F}_k K)(\boldsymbol{x})\varphi(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^N} K(\boldsymbol{x})(\mathcal{F}_k\varphi)(\boldsymbol{x})\mathrm{d}x$$

for any $\varphi \in S(R^N)$.

Theorem 2.5.2.
Let $K \in \mathcal{PK}(\boldsymbol{\nu})$ *then* $\mathcal{F}_k K \in \mathcal{PK}\left((\nu_1, \ldots, \nu_{k-1}, q_k - \nu_k, \nu_{k+1}, \ldots, \nu_d)\right)$. *Furthermore* $\mathcal{F}_k : \mathcal{PK}(\boldsymbol{\nu}) \to \mathcal{PK}\left((\nu_1, \ldots, \nu_{k-1}, q_k - \nu_k, \nu_{k+1}, \ldots, \nu_d)\right)$ *is continuous between the strong topologies.*

Before proving this theorem we state a product-space generalization of a useful lemma about when taking a Fourier transform coincides with integrating against $e^{-i\xi x}$.

Lemma 2.5.3.
Suppose that $K \in S'(\mathbb{R}^N)$ *is a tempered distribution with bounded support along* H_k. *Then the partial Fourier transform* $\mathcal{F}_k K$ *is given by the distribution* $\int_{H_k} K(\boldsymbol{x})\eta(\rho^{-1}\cdot x_k)e^{-i\xi_k x_k}\,\mathrm{d}x_k$ *for a cutoff function* η *and* ρ *large enough so that* $\eta(\rho^{-1}\cdot x_k)$ *is identically* 1 *on the support of* K. *The derivatives of* $\partial_{\xi_k}^{\alpha_k}\mathcal{F}_k K$ *of the partial Fourier transform are given by* $\int_{H_k} K(\boldsymbol{x})\eta(\rho^{-1}\cdot x_k)(-ix_k)^{\alpha_k}e^{-i\xi_k x_k}\,\mathrm{d}x_k$.

The expression $\int_{H_k} K(\boldsymbol{x})\eta(\rho^{-1}\cdot x_k)(-ix_k)^{\alpha_k}e^{-i\xi_k x_k}\,\mathrm{d}x_k$ *defines a distribution on* $\mathbb{R}^N$ *by*

$$\left\langle \int_{H_k} K(\boldsymbol{x})\eta(\rho^{-1}\cdot x_k)(-ix_k)^{\alpha_k}e^{-i\xi_k x_k}\,\mathrm{d}x_k;\, \varphi(x_1, \ldots, x_{k-1}, \xi_k, x_{k+1}, \ldots, x_d)\right\rangle_{\mathbb{R}^N} =$$
$$\int_{H_k}\left\langle \int_{H_k} K(\boldsymbol{x})\eta(\rho^{-1}\cdot x_k)(-ix_k)^{\alpha_k}e^{-i\xi_k x_k}\,\mathrm{d}x_k;\, \varphi(x_1, \ldots, x_{k-1}, \xi_k, x_{k+1}, \ldots, x_d)\right\rangle_{\oplus_{j\neq k}H_j} \mathrm{d}\xi_k. \tag{2.5.1}$$

Proof. Notice that

$$\left\langle \int_{H_k} K(\boldsymbol{x})\eta(\rho^{-1}\cdot x_k)(-ix_k)^{\alpha_k}e^{-i\xi_k x_k}\,\mathrm{d}x_k;\, \varphi(x_1, \ldots, x_{k-1}, \xi_k, x_{k+1}, \ldots, x_d)\right\rangle_{\oplus_{j\neq k}H_j}$$

coincides by definition with

$$\left\langle K(\boldsymbol{x});\, \eta(\rho^{-1}\cdot x_k)(-ix_k)^{\alpha_k}e^{-i\xi_k x_k}\varphi(x_1, \ldots, x_{k-1}, \xi_k, x_{k+1}, \ldots, x_d)\right\rangle_{\mathbb{R}^N}. \tag{2.5.2}$$

The mapping $\xi_k \mapsto \eta(\rho^{-1}\cdot x_k)(-ix_k)^{\alpha_k}e^{-i\xi_k x_k}\varphi(x_1, \ldots, x_{k-1}, \xi_k, x_{k+1}, \ldots, x_d)$ is continuous from H_k to $S(\mathbb{R}^N)$ with all Schwartz norms decaying rapidly as $\xi_k \to \infty$. Since K, as tempered distributions in general, has finite order, the quantity (2.5.2) is continuous, rapidly decaying, and thus integrable. Continuity with respect to $\varphi \in S(\mathbb{R}^N)$ of the integral can also be easily checked.

The integral and the duality pairing on the right hand side of (2.5.1) commute. As a matter of fact, since the quantity in (2.5.2) is continuous in ξ_k and decays rapidly one can approximate the integral with Riemann sums. By linearity the Riemann sums commute with the duality pairing and passing to the limit again yields

$$\int_{H_k}\left\langle K(\boldsymbol{x});\, \eta(\rho^{-1}\cdot x_k)(-ix_k)^{\alpha_k}e^{-i\xi_k x_k}\varphi(x_1, \ldots, x_{k-1}, \xi_k, x_{k+1}, \ldots, x_d)\right\rangle_{\mathbb{R}^N}\mathrm{d}\xi_k =$$
$$\left\langle K(\boldsymbol{x});\, \int_{H_k}\eta(\rho^{-1}\cdot x_k)(-ix_k)^{\alpha_k}e^{-i\xi_k x_k}\varphi(x_1, \ldots, x_{k-1}, \xi_k, x_{k+1}, \ldots, x_d)\,\mathrm{d}\xi_k\right\rangle_{\mathbb{R}^N}$$

Given the condition on the support of K the term $\eta(\rho^{-1}\cdot x_k)$ is identically equal to 1, so above quantity is equal to

$$\left\langle K(\boldsymbol{x});\int_{H_k}(-ix_k)^{\alpha_k}e^{-i\xi_k x_k}\varphi(x_1,\dots,x_{k-1},\xi_k,x_{k+1},\dots,x_d)\,\mathrm{d}\xi_k\right\rangle_{\mathbb{R}^N}=$$
$$\left\langle K(\boldsymbol{x});\,(-1)^{\alpha_k}\mathcal{F}_k\partial^{\alpha_k}_{\xi_k}\varphi(x_1,\dots,x_{k-1},\xi_k,x_{k+1},\dots,x_d)\right\rangle_{\mathbb{R}^N}=\left\langle\partial^{\alpha_k}_{\xi_k}\mathcal{F}_kK;\,\varphi\right\rangle_{\mathbb{R}^N}.$$

□

Proof of Theorem 2.5.2. To prove the theorem we must check the cancellation conditions (2.3.4) and the size estimates (2.3.3) on the kernel $\mathcal{F}_kK$. Notice that partial duality pairing with a test function $\varphi\in S(H_j)$ commutes with the partial Fourier transform along H_k for $k\neq j$. This facilitates the proof allowing us to reason by induction on the number d of factors in the product decomposition of $\mathbb{R}^N$. For $d=0$ there is nothing to prove. Now suppose that $\mathbb{R}^N=\oplus_{k\in\{1,\dots,d\}}H_k$ and that the theorem is valid for all product kernels on products of up to $d-1$ spaces.

Size condition (2.3.3) We begin with the size estimates on $\mathcal{F}_kK$. Consider a cutoff function $\eta(x_k)$ on H_k as explained in 1.2.7. Decompose the kernel as $K(\boldsymbol{x})=K^1(\boldsymbol{x})+K^2(\boldsymbol{x})$ with $K^1(\boldsymbol{x})=K(\boldsymbol{x})\eta(\rho^{-1}\cdot x_k)$ and $K^2(\boldsymbol{x})=K(\boldsymbol{x})\left(1-\eta(\rho^{-1}\cdot x_k)\right)$ for some $\rho>0$ to be chosen subsequently.

The support of K^1 is bounded along H_k, so by lemma 2.5.3 any derivative of its partial Fourier transform is given by

$$\partial^{\alpha_k}_{\xi_k}\mathcal{F}_kK^1(x_1,\dots,x_{k-1},\xi_k,x_{k+1},x_d)=\int_{H_k}K^1(\boldsymbol{x})\eta\left((2\rho)^{-1}\cdot x_k\right)(-ix_k)^{\alpha_k}e^{-i\xi_kx_k}\,\mathrm{d}x_k=$$
$$\rho^{\|\alpha_k\|}\int_{H_k}K(\boldsymbol{x})\eta\left(\rho^{-1}\cdot x_k\right)(-i\rho^{-1}\cdot x_k)^{\alpha_k}e^{-i(\rho\cdot\xi_k)(\rho^{-1}\cdot x_k)}\,\mathrm{d}x_k.$$

The function $\eta(x_k)(-ix_k)^{\alpha_k}e^{-i(\rho\cdot\xi_k)x_k}$ is supported in the unit ball of H_k. By choosing $\rho=\|\xi_k\|^{-1}$ for any given ξ_k the above function is also b_k-normalized so the cancellation condition on K gives that

$$\rho^{\nu_k-\|\alpha_k\|}\partial^{\alpha_k}_{\xi_k}\mathcal{F}_kK^1$$

is an bounded family of product kernels in $\mathcal{PK}(\boldsymbol{\nu}_{(k)})$. In particular from $H_j^\perp$ for any $j\neq k$ we have that

$$\left|\partial^{\alpha_k}_{\xi_k}\partial^{\alpha(k)}_{\boldsymbol{x}_{(k)}}\mathcal{F}_kK^1(x_1,\dots,x_{k-1},\xi_k,x_{k+1},\dots,x_d)\right|\le C\|\xi_k\|^{\nu_k-\|\alpha_k\|}\prod_{j\neq k}\|x_j\|^{-q_j-\nu_j-\|\alpha_j\|}.$$

Having chosen $\rho=\|\xi_k\|^{-1}$ we must now prove the estimate for $\mathcal{F}_kK^2$. K^2 is supported away from the plane $H_k^\perp=\{\boldsymbol{x}\in\mathbb{R}^N\mid x_k=0\}$. This means that K^2 coincides with a smooth function on the set $\mathbb{R}^N\setminus\left(\bigcup_{j\neq k}H_j^\perp\right)$. We have as usual that $\partial^{\alpha_k}_{\xi_k}\mathcal{F}_kK^2=\mathcal{F}_k\left((-ix_k)^{\alpha_k}K^2\right)$ and $(-ix_k)^{\alpha_k}K^2$ is also a smooth function on the above stated domain.

We are concerned with the size estimates when $\xi_k\neq0$ and $x_j\neq0$. For a large constant $C>1$ and for every coordinate $x_{k,j}$ in H_k consider the open set such that $C^{-1}\|\xi_k\|<|\xi_{k,j}|^{1/\lambda_{k,j}}<C\|\xi_k\|$. Since $\|\xi_k\|$ and $\max_{j\in\{1,\dots,n_k\}}\{|\xi_{k,j}|^{1/\lambda_{k,j}}\}$ are two homogeneous norms on H_k, they are equivalent, so C can be chosen large enough so that the above open sets

cover the set $\mathbb{R}^N \setminus H_k^{\perp}$. We prove that $\mathcal{F}_k K^2$ coincides with a function and satisfies the size estimates on these open sets.

Fix a $j \in \{1, \ldots, n_k\}$ and the corresponding coordinate $x_{k,j}$ with its dual coordinate $\xi_{k,j}$. Consider a test function φ compactly supported on $\mathbb{R}^N \setminus \left(\bigcup_{j\neq k} H_j^{\perp}\right)$ where $C^{-1}\|x_k\| < |\xi_{k,j}|^{1/\lambda_{k,j}} < C\|x_k\|$. Its partial Fourier transform can be written as

$$\mathcal{F}_k\varphi = \int_{H_k} \varphi(x_1, \ldots, \xi_k, \ldots, x_d) e^{-i\xi_k x_k}\, \mathrm{d}\xi_k =$$

$$\int_{H_k} \varphi(x_1, \ldots, \xi_k, \ldots, x_d) \frac{\partial^{\gamma}_{x_{k,j}} e^{-i\xi_k x_k}}{(-i\xi_{k,j})^{\gamma}}\, \mathrm{d}\xi_k = \partial^{\gamma}_{x_{k,j}} \mathcal{F}_k \left(\frac{\varphi}{(-i\xi_{k,j})^{\gamma}}\right)(\boldsymbol{x})$$

for any $\gamma \in \mathbb{N}$. Writing down the duality relation one gets

$$\left\langle \partial^{\alpha_k}_{\xi_k} \mathcal{F}_k K^2;\, \varphi \right\rangle = \left\langle (-ix_k)^{\alpha_k} K^2;\, \partial^{\gamma}_{x_{k,j}} \mathcal{F}_k \left(\frac{\varphi}{(-i\xi_{k,j})^{\gamma}}\right)\right\rangle =$$

$$(-1)^{\gamma} \left\langle \partial^{\gamma}_{x_{k,j}} \left((-ix_k)^{\alpha_k} K(\boldsymbol{x}) \left(1 - \eta(\rho^{-1}\cdot x_k)\right)\right);\, \mathcal{F}_k \left(\frac{\varphi}{(-i\xi_{k,j})^{\gamma}}\right)\right\rangle =$$

$$(-1)^{\gamma} \left\langle \frac{1}{(-i\xi_{k,j})^{\gamma}} \mathcal{F}_k \left(\partial^{\gamma}_{x_{k,j}} \left((-ix_k)^{\alpha_k} K(\boldsymbol{x}) \left(1 - \eta(\rho^{-1}\cdot x_k)\right)\right)\right);\, \varphi \right\rangle$$

So on the open set above we have

$$\partial^{\alpha_k}_{\xi_k} \mathcal{F}_k K^2 = (-1)^{\gamma} \frac{1}{(-i\xi_{k,j})^{\gamma}} \mathcal{F}_k \left(\partial^{\gamma}_{x_{k,j}} \left((-ix_k)^{\alpha_k} K(\boldsymbol{x}) \left(1 - \eta(\rho^{-1}\cdot x_k)\right)\right)\right)$$

We now analyze the derivative on the kernel and distinguish the cases when at least one derivative falls on the term $(1-\eta)$.

$$\partial^{\gamma}_{x_{k,j}} \left((-ix_k)^{\alpha_k} K(\boldsymbol{x}) \left(1 - \eta(\rho^{-1}\cdot x_k)\right)\right) =$$

$$\partial^{\gamma}_{x_{k,j}} \left((-ix_k)^{\alpha_k} K(\boldsymbol{x})\right) \left(1 - \eta(\rho^{-1}\cdot x_k)\right) +$$

$$\sum_{\substack{\gamma_1+\gamma_2=\gamma \\ \gamma_2\geq 1}} \partial^{\gamma_1}_{x_{k,j}} \left((-ix_k)^{\alpha_k} K(\boldsymbol{x})\right) \partial^{\gamma_2}_{x_{k,j}} \left(1 - \eta(\rho^{-1}\cdot x_k)\right)$$

Choosing $\gamma > \alpha_{k,j} - \nu_k$first term on the right hand side is integrable in x_k so its Fourier transform is given by the integral and

$$\left|\mathcal{F}_k \left(\partial^{\gamma}_{x_{k,j}} \left((-ix_k)^{\alpha_k} K(\boldsymbol{x})\right) \left(1 - \eta(\rho^{-1}\cdot x_k)\right)\right)\right| =$$

$$\left|\int_{H_k} \partial^{\gamma}_{x_{k,j}} \left((-ix_k)^{\alpha_k} K(\boldsymbol{x})\right) \left(1 - \eta(\rho^{-1}\cdot x_k)\right) e^{-i\xi_k x_k}\, \mathrm{d}x_k\right| =$$

$$\left|\int_{\|x_k\|>\rho/2} \partial^{\gamma}_{x_{k,j}} \left((-ix_k)^{\alpha_k} K(\boldsymbol{x})\right) \left(1 - \eta(\rho^{-1}\cdot x_k)\right) e^{-i\xi_k x_k}\, \mathrm{d}x_k\right| <$$

$$C \int_{\|x_k\|>\rho/2} \|x_k\|^{-q_k-\nu_k+\|\alpha_k\|-\lambda_{k,j}\gamma}\, \mathrm{d}x_k \prod_{j\neq k} \|x_j\|^{-q_j-\nu_j} <$$

$$C' \rho^{-\nu_k+\|\alpha_k\|-\lambda_{k,j}\gamma} \prod_{j\neq k} \|x_j\|^{-q_j-\nu_j}.$$

Since we chose $\rho = \|\xi_k\|^{-1}$ and $\left|\xi_{k,j}\right| \approx \|\xi_k\|^{\lambda_{k,j}}$ we have the necessary estimate.

For the terms with derivatives on $1-\eta$ it suffices to notice that they are compactly supported along H_k so one can use the cancellation property in a manner similar to the estimates for K^1

By a partition of unity argument $\mathcal{F}_k K^2$ coincides with a smooth function on the whole $\mathbb{R}^N \setminus H_k^\perp$. Since the size estimates we obtained are point-wise they hold on that domain.

Cancellation conditions (2.3.4) For the cancellation conditions (2.3.4) it suffices to check the size estimates for the family

$$R^{q_k-\nu_k}\int_{H_k}\mathcal{F}_k K(x_1,\dots,x_{k-1},\xi_k,x_{k+1},\dots,x_d)\varphi(R^{-1}\cdot\xi_k)\,\mathrm{d}\xi_k$$

away from $x_j = 0$ for any $j \neq k$, when φ is a b_k-normalized bump function and for any $R > 0$. When checking the cancellation condition long H_j with $j \neq k$ we can use the inductive hypothesis. Since

$$R^{\nu_j}\int_{H_j}\mathcal{F}_k K(x_1,\dots,x_{k-1},\xi_k,x_{k+1},\dots,x_d)\varphi(R^{-1}\cdot x_j)\,\mathrm{d}x_j = \mathcal{F}_k\left(R^{\nu_j}\int_{H_j}K(\boldsymbol{x})\varphi(R^{-1}\cdot x_j)\,\mathrm{d}x_j\right)$$

and for any b_j-normalized bump function φ on H_j and any $R > 0$

$$R^{\nu_j}\int_{H_j}K(\boldsymbol{x})\varphi(R^{-1}\cdot x_j)\,\mathrm{d}x_j$$

is uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu}_{(j)})$ then the required boundedness follows by the inductive hypothesis.

When checking the uniform boundedness of the family

$$R_k^{q_k-\nu_k}\int_{H_k}\mathcal{F}_k K(x_1,\dots,x_{k-1},\xi_k,x_{k+1},\dots,x_d)\varphi(R^{-1}\cdot\xi_k)\,\mathrm{d}\xi_k$$

for any b_k-normalized bump function φ_k on H_k and any $R_k > 0$ we approach the cancellation conditions in a similar manner. As a matter of fact for any $j \neq k$

$$\begin{aligned}R_j^{\nu_j}\int_{H_j}R_k^{q_k-\nu_k}\int_{H_k}&\mathcal{F}_k K(x_1,\dots,x_{k-1},\xi_k,x_{k+1},\dots,x_d)\\ &\varphi_k(R_k^{-1}\cdot\xi_k)\,\mathrm{d}\xi_k\;\varphi_j(R_j^{-1}\cdot x_j)\,\mathrm{d}x_j = \\ &R_k^{q_k-\nu_k}\int_{H_k}\mathcal{F}_k\left(R_j^{\nu_j}\int_{H_j}K(\boldsymbol{x})\varphi_j(R_j^{-1}\cdot x_j)\,\mathrm{d}x_j\right)\varphi_k(R_k^{-1}\cdot\xi_k)\,\mathrm{d}\xi_k.\end{aligned}$$

For b_j-normalized bump functions φ_j on H_j and any $R_j > 0$ the inner member

$$R_j^{\nu_j}\int_{H_j}K(\boldsymbol{x})\varphi_j(R_j^{-1}\cdot x_j)\mathrm{d}x_j$$

is uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu}_{(j)})$ and by the inductive hypothesis so is its partial Fourier transform. The required boundedness is the consequence of the cancellation property.

Now we proceed to check the size estimates 2.3.3 for

$$R^{-q_k-\nu_k}\int_{H_i}\mathcal{F}_kK(x_1,\dots,x_{k-1},\xi_k,x_{k+1},\dots,x_d)\varphi(R^{-1}\cdot\xi_k)\,\mathrm{d}\xi_k$$

uniformly for all $R>0$ and all normalized bump functions. Thanks to Proposition 2.4.11 we can take φ normalized with respect to an arbitrarily large C^b norm.

By the duality and by the scaling property of the Fourier transform we have

$$\begin{aligned}R^{-q_k-\nu_k}\int_{H_k}\mathcal{F}_kK(x_1,\dots,x_{k-1},\xi_k,x_{k+1},\dots,x_d)\varphi(R^{-1}\cdot\xi_k)\,\mathrm{d}\xi_k =\\ R^{-q_k-\nu_k}\int_{H_k}K(\boldsymbol{x})R^{q_k}\left(\mathcal{F}_k\varphi\right)(R\cdot x_k)\,\mathrm{d}x_k =\\ R^{-q_k-\nu_k}\int_{H_k}K(\boldsymbol{x})\eta(R\cdot x_k)R^{q_k}\left(\mathcal{F}_k\varphi\right)(R\cdot x_k)\,\mathrm{d}x_k+\\ R^{-q_k-\nu_k}\int_{H_k}K(\boldsymbol{x})\left(1-\eta(R\cdot x_k)\right)R^{q_k}\left(\mathcal{F}_k\varphi\right)(R\cdot x_k)\,\mathrm{d}x_k\end{aligned}$$

where η is a cutoff function. For the first term we can apply the cancellation property with respect to the bump function $\eta\mathcal{F}_k\varphi$. This bump function is normalized since $\|\eta\mathcal{F}_k\varphi\|_{C^b}<C\,\|\mathcal{F}_k\varphi\|_{C^b}$ and by L^1-L^∞ boundedness of the Fourier transform $\|\mathcal{F}_k\varphi\|_{C^b}<C\left\|\left|1+|x|^2\right|^{b/2}\varphi\right\|_{L^1}<C\,\|\varphi\|_{C^0}$ where the last equality holds since φ is continuous and supported on the unit ball of H_k

Thus we have that

$$\begin{aligned}R^{-q_k-\nu_k}\int_{H_k}K(\boldsymbol{x})\eta(R\cdot x_k)R^{q_k}\left(\mathcal{F}_k\varphi\right)(R\cdot x_k)\,\mathrm{d}x_k =\\ R^{-\nu_k}\int_{H_k}K(\boldsymbol{x})\eta(R\cdot x_k)\left(\mathcal{F}_k\varphi\right)(R\cdot x_k)\,\mathrm{d}x_k\end{aligned}$$

is a uniformly bounded family of kernels and thus they satisfy the needed size estimates uniformly. For the second term notice that $K(\boldsymbol{x})\left(1-\eta(R\cdot x_k)\right)$ coincides with an L^1_{loc} function on $\mathbb{R}^N\setminus\bigcup_{j\neq k}H_j^\perp$. So away from $\bigcup_{j\neq k}H_j^\perp$ we have

$$\begin{aligned}\left|R^{-q_k-\nu_k}\int_{H_k}K(\boldsymbol{x})\left(1-\eta(R\cdot x_k)\right)R^{q_k}\left(\mathcal{F}_k\varphi\right)(R\cdot x_k)\,\mathrm{d}x_k\right|<\\ R^{-q_k-\nu_k}\prod_{j\neq k}\|x_j\|^{-q_j-\nu_j}\int_{\|R\cdot x_k\|>1/2}\frac{\|x_k\|^{-q_k-\nu_k}}{\left(1+|R\cdot x_k|^2\right)^{-s/2}}\\ \left(1+|R\cdot x_k|^2\right)^{s/2}\left|\left(\mathcal{F}_k\varphi\right)(R\cdot x_k)\right|R^{q_k}\,\mathrm{d}x_k<\\ \prod_{j\neq k}\|x_j\|^{-q_j-\nu_j}\int_{\|x_k\|>1/2}\frac{\|x_k\|^{-q_k-\nu_k}}{\left(1+|x_k|^2\right)^{-s/2}}\left(1+|x_k|^2\right)^{s/2}\left|\left(\mathcal{F}_k\varphi\right)(x_k)\right|\,\mathrm{d}x_k.\end{aligned}$$

By choosing $s>0$ large enough one can apply Cauchy-Schwartz and get

$$\begin{aligned}\left|R^{-q_k-\nu_k}\int_{H_k}K(\boldsymbol{x})\left(1-\eta(R\cdot x_k)\right)R^{q_k}\left(\mathcal{F}_k\varphi\right)(R\cdot x_k)\,\mathrm{d}x_k\right|<\\ C\prod_{j\neq k}\|x_j\|^{-q_j-\nu_j}\|\varphi\|_{H^s(\xi_k)}<C\prod_{j\neq k}\|x_j\|^{-q_j-\nu_j}\|\varphi\|_{C^b}\end{aligned}$$

for some $b \in \mathbb{N}$ and $b > s$, since φ is supported on the unit ball of H_k. The size estimates for the derivatives are identical.

□

As a corollary we have the following result for the full Fourier transform

Corollary 2.5.4.
The full Fourier transform acts in the following way $\mathcal{F} : \mathcal{PK}(\boldsymbol{\nu}) \to \mathcal{PK}(\boldsymbol{q} - \boldsymbol{\nu})$ for any $\boldsymbol{\nu}$ and it is strongly continuous.

2.6 Basic functional calculus

The motivation for introducing kernels of non-zero order is given by the need of a basic functional calculus. A reasonable requirement is that the classes we are working with be well behaved with respect to derivation and multiplication by homogeneous polynomials. As we have already illustrated in the Section 1.2 for Calderón-Zygmund kernels, the cancellation conditions and the size conditions we require on product kernels are those that naturally arise when considering distributional derivatives of **0** order product kernels. For any multi-indexes $\boldsymbol{\alpha}$ and for any multi-order $\boldsymbol{\nu}$ let

$$\boldsymbol{\nu} + \boldsymbol{\alpha} \stackrel{\text{def}}{=} (\nu_1 + \|\alpha_1\|, \ldots, \nu_d + \|\alpha_d\|).$$

We have the following proposition that establishes a basic functional calculus.

Proposition 2.6.1 (Derivation and multiplication of product kernels)**.**
Let $K \in \mathcal{PK}(\boldsymbol{\nu})$, then the following propositions hold. For any multi-indexes $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ the distribution $\boldsymbol{x}^{\boldsymbol{\alpha}} \partial^{\boldsymbol{\beta}} K \in \mathcal{PK}(\boldsymbol{\nu} - \boldsymbol{\alpha} + \boldsymbol{\beta})$. Furthermore the mapping $K \mapsto \boldsymbol{x}^{\boldsymbol{\alpha}} \partial^{\boldsymbol{\beta}} K$ is continuous from $\mathcal{PK}(\boldsymbol{\nu})$ to $\mathcal{PK}(\boldsymbol{\nu} - \boldsymbol{\alpha} + \boldsymbol{\beta})$.

Proof. As usual the proof goes by induction on the number of factors d of the product decomposition of $\mathbb{R}^N$. For $d = 0$ there is nothing to prove.

Suppose the statement is true for any decomposition in up to d factors and $K \in \widetilde{\mathcal{PK}}(\boldsymbol{\nu})$ on $\mathbb{R}^N = \oplus_{k=1}^{d+1} H_k$. We prove this statement when the derivations and the multiplication concerns only one subspace. Since

$$\boldsymbol{x}^{\boldsymbol{\alpha}} \partial_{\boldsymbol{x}}^{\boldsymbol{\beta}} K(\boldsymbol{x}) = x_1^{\alpha_1} \partial_{x_1}^{\beta_1} \ldots x_{d+1}^{\alpha_{d+1}} \partial_{x_{d+1}}^{\beta_{d+1}} K(\boldsymbol{x})$$

this, by an iteration argument, is equivalent to proving the statement. We suppose that $\boldsymbol{x}^{\boldsymbol{\alpha}} = x_k^{\alpha_k}$ and $\partial_{\boldsymbol{x}}^{\boldsymbol{\beta}} = \partial_{x_k}^{\beta_k}$ for some $k \in \{1, \ldots, d\}$.

Given a multi-index $\boldsymbol{\gamma}$ we indicate by $\boldsymbol{\gamma}_{(k)}$ the multi-index such that $\gamma_{(k)\,k} = 0$ and $\gamma_{(k)\,j} = \gamma_j$ for $j \neq k$.

Size conditions (2.3.3) First we check the size conditions. If K coincides with a C^∞ function on an open set then its derivatives and products by polynomials coincide with the derivatives and products of the smooth function on that open set.

$$\left|\partial^{\gamma} x_k^{\alpha_k} \partial_{x_k}^{\beta_k} K(\boldsymbol{x})\right| < C \sum_{\substack{\sigma_k + \tau_k = \gamma_k \\ \alpha_k \geq \sigma_k}} \left|\boldsymbol{x}_k^{\alpha_k - \sigma_k} \partial_{x_k}^{\beta_k + \tau_k} \partial^{\gamma_{(k)}} K\right| \leq$$

$$\leq C \prod_{j \neq k} \|x_j\|^{-q_j - \nu_j - \gamma_j} \left(\sum_{\substack{\sigma_k + \tau_k = \gamma_k \\ \alpha_k \geq \sigma_k}} \|x_k\|^{\|\alpha_k\| - \|\sigma_k\|} \|x_k\|^{-q_k - \nu_k - \|\tau_k + \beta_k\|} \right) \leq$$

$$\leq C \|x_k\|^{-q_k - \nu_k - \|\beta_k\| + \|\alpha_k\| - \|\gamma_k\|} \prod_{j \neq k} \|x_j\|^{-q_j - \nu_j - \gamma_j}.$$

Cancellation along H_j with $j \neq k$ Now we pass to the cancellation conditions (2.3.4). If the cancellation occurs along a subspace H_j with $j \neq k$ then we can use the inductive hypothesis. As a matter of fact when φ is a normalized bump function

$$R^{\nu_j} \int_{H_j} x_k^{\alpha_k} \partial_{x_k}^{b_k} K(\boldsymbol{x}) \varphi(R^{-1} \cdot x_j) \, \mathrm{d}x_j = x_k^{\alpha_k} \partial_{x_k}^{b_k} R^{\nu_j} \int_{H_j} K(\boldsymbol{x}) \varphi(R^{-1} \cdot x_j) \, \mathrm{d}x_j$$

for any $R > 0$. But $R^{\nu_j} \int_{H_j} K(\boldsymbol{x}) \varphi(R^{-1} \cdot x_j) \, \mathrm{d}x_j$ is a uniformly bounded family of product kernels and by the inductive hypothesis

$$x_k^{\alpha_k} \partial_{x_k}^{b_k} R^{\nu_j} \int_{H_j} K(\boldsymbol{x}) \varphi(R^{-1} \cdot x_j) \, \mathrm{d}x_j$$

is uniformly bounded in $(\boldsymbol{\nu} - \boldsymbol{\alpha} + \boldsymbol{\beta})_{(j)}$.

Cancellation along H_k Suppose $\varphi \in C_c^\infty\big(B_{H_k}(0,1)\big)$ is a $\widetilde{b}_k$-normalized bump function with $\widetilde{b}_k$ large enough to be chosen afterwards. Proposition 2.4.11 guarantees that we can choose an arbitrary high order of normalization.

$$\begin{aligned} R^{\nu_k - \|\alpha_k\| + \|\beta_k\|} \int_{H_k} x_k^{\boldsymbol{\alpha}_k} \partial_{x_k}^{\boldsymbol{\beta}_k} K(\boldsymbol{x}) \varphi(R^{-1} \cdot x_k) \, \mathrm{d}x_k = \\ R^{\nu_k + \|\beta_k\|} \int_{H_k} K(\boldsymbol{x}) \partial_{x_k}^{\beta_k} \Big((R^{-1} \cdot x_k)^{\alpha_k} \varphi(R^{-1} \cdot x_k) \Big) \, \mathrm{d}x_k \end{aligned}$$

Notice that $R^{\|\beta_k\|} \partial_{x_k}^{\beta_k} \big((R^{-1} \cdot x_k)^{\alpha_k} \varphi(R^{-1} \cdot x_k) \big)$ is the rescaled version of the bump function $\partial_{x_k}^{\beta_k} \big(x_k^{\alpha_k} \varphi(x_k) \big)$. Since φ is $\widetilde{b}_k$-normalized the above bump function is $(\widetilde{b}_k - |\beta_k|)$-normalized as long as $\widetilde{b}_k \geq |\beta_k|$. So as long as we choose $\widetilde{b}_k > |\beta_k| + b_k$ the expression above is uniformly bounded. This concludes the proof.

□

2.7 Dyadic decomposition

As in the case of classic Calderón-Zygmund kernels, the class of product-kernels is homogeneous with respect to multi-parameter dilation in the sense that for any bounded family of kernels $K \in \mathcal{PK}(\boldsymbol{\nu})$ the kernels

$$\boldsymbol{R}^{-\boldsymbol{q} - \boldsymbol{\nu}} K(\boldsymbol{R}^{-1} \cdot \boldsymbol{x})$$

for all $\boldsymbol{R} \in (\mathbb{R}^+)^d$ are uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu})$. To better capture this idea we introduce the multi-parameter dyadic decompositions for product kernels. We start by defining the "building blocks".

Definition 2.7.1.
For every $\boldsymbol{l} \in \mathbb{R}^d$ we define

$$\begin{aligned} S_0^{\boldsymbol{l}}(\mathbb{R}^N) &\overset{def}{=} \left\{ \varphi \in S(\mathbb{R}^N) \,\middle|\, \int_{H_k} x_k^{\alpha_k} \varphi(\boldsymbol{x}) \mathrm{d}x_k = 0 \ \ \forall \|\alpha\|_k \leq l_k \ \ \forall k \in \{1, \dots, d\} \right\} \\ \widehat{S}_0^{\boldsymbol{l}}(\mathbb{R}^N) &\overset{def}{=} \left\{ \varphi \in S(\mathbb{R}^N) \,\middle|\, \int_{H_k} \xi_k^{\alpha_k} \widehat{\varphi}(\boldsymbol{\xi}) \mathrm{d}\xi_k = 0 \ \ \forall \|\alpha\|_k \leq l_k \ \ \forall k \in \{1, \dots, d\} \right\}. \end{aligned}$$

If, for some $k \in \{1, \dots, d\}$, $l_k < 0$ then no condition along H_k needs to be satisfied. Using the Fourier transform, the same sets can be described as

$$S_0^{\boldsymbol{l}}(\mathbb{R}^N) = \left\{ \varphi \in S(\mathbb{R}^N) \,\middle|\, \partial_{\xi_k}^{\alpha_k} \widehat{\varphi}(\boldsymbol{\xi}_{(k)} \oplus 0_k) = 0 \;\; \forall \|\alpha_k\| \le l_k \;\; \forall k \in \{1, \dots, d\} \right\}$$
$$\widehat{S}_0^{\boldsymbol{l}}(\mathbb{R}^N) = \left\{ \varphi \in S(\mathbb{R}^N) \,\middle|\, \partial_{x_k}^{\alpha_k} \varphi(\boldsymbol{x}_{(k)} \oplus 0_k) = 0 \;\; \forall \|\alpha_k\| \le l_k \;\; \forall k \in \{1, \dots, d\} \right\}$$

We indicate $S_0(\mathbb{R}^N) \overset{def}{=} S_0^{\boldsymbol{\infty}}(\mathbb{R}^N)$ and $\widehat{S}_0(\mathbb{R}^N) \overset{def}{=} \widehat{S}_0^{\boldsymbol{\infty}}(\mathbb{R}^N)$.

We will now look at a series of theorems that concentrate on the possibility of writing dyadic decompositions for product kernels in terms of rescaled versions of the above-defined functions. However we will start by proving a somewhat easier theorem which answers the converse question of whether a certain dyadic sum converges to a product kernel of some order $\boldsymbol{\nu}$. Finally we will show how similar results can be obtained concentrating on the space localization of the dyadic "building blocks".

We begin by stating a condition when a dyadic sum converges to a product kernel. The following is the product equivalent of Theorem 1.2.11.

Theorem 2.7.2 (Sufficient conditions for convergence of dyadic sums)**.**
Consider the product space $\mathbb{R}^N = \bigoplus_{k=1}^d H_k$ and a multi-order $\boldsymbol{\nu} \in R^d$. Let $\{\varphi_{\boldsymbol{i}}\}_{\boldsymbol{i} \in \mathbb{Z}^d}$ be a bounded family of Schwartz functions that satisfy the following conditions:

- *If $k \in \{1, \dots, d\}$ is such that $\nu_k \ge 0$ then*

$$\int_{H_k} x_k^{\alpha_k} \varphi_{\boldsymbol{i}}(\boldsymbol{x}) \mathrm{d}x_k = 0 \qquad \forall \|\alpha_k\| \le \nu_k.$$

- *If $k \in \{1, \dots, d\}$ is such that $\nu_k \le -q_k$ then*

$$\partial_{x_k}^{\alpha_k} \varphi_i(0_k \oplus \boldsymbol{x}_{(k)}) = \int_{H_k} \xi_k^{\alpha_k} \widehat{\varphi}_{\boldsymbol{i}}(\boldsymbol{\xi}) \mathrm{d}\xi_k = 0 \qquad \forall \|\alpha_k\| \le -q_k - \nu_k.$$

- *if $k \in \{1, \dots, d\}$ is such that $-q_k < \nu_k < 0$, no condition along H_k needs to be satisfied.*

Then the dyadic sum

$$\sum_{\boldsymbol{i} \in \mathbb{Z}^d} 2^{-\boldsymbol{i}\boldsymbol{\nu}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}) = \sum_{\boldsymbol{i} \in \mathbb{Z}^d} 2^{-i_1(q_1+\nu_1)} \dots 2^{-i_d(q_d+\nu_d)} \varphi_{\boldsymbol{i}}(2^{-i_1} \cdot x_1, \dots, 2^{-i_d} \cdot x_d) \tag{2.7.1}$$

has all (finite and infinite) partial sums uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu})$ and it weak- converges to a kernel $K \in \mathcal{PK}(\boldsymbol{\nu})$*

This dyadic sum has a continuity property in the sense that for any neighborhood U of 0 in $\mathcal{PK}(\boldsymbol{\nu})$ there is a neighborhood V_U of 0 in $S(\mathbb{R}^N)$ such that if $\{\varphi_{\boldsymbol{i}}\} \subset V_U$ then the series and all the partial sums of (2.7.1) are in U.

We can also state the conditions on $\{\varphi_{\boldsymbol{i}}\}$ in the following manner.

$$\{\varphi_{\boldsymbol{i}}\} \subset S_0^{\boldsymbol{\nu}}(\mathbb{R}^N) \cap \widehat{S}_0^{-\boldsymbol{q}-\boldsymbol{\nu}}(\mathbb{R}^N)$$

Proof. The continuity property of the statement will follow from the fact that the estimates we make on the partial sums and on the series depend only on the Schwartz semi-norms bounds of the family $\{\varphi_{\boldsymbol{i}}\}$.

Let $L = \{ l \in \{1, \dots, d\} \,|\, \nu_l \geq 0 \}$. Consider any partial finite sum of the dyadic series (2.7.1). By Theorem 2.5.2 the family of partial sums is in $\mathcal{PK}(\boldsymbol{\nu})$ and is uniformly bounded if and only if the partial Fourier transform along H_L are uniformly bounded in $\mathcal{PK}(\widetilde{\boldsymbol{\nu}})$ where $\widetilde{\boldsymbol{\nu}}$ is such that $\widetilde{\nu}_l = -q_l - \nu_l$ if $l \in L$ and $\widetilde{\nu}_l = \nu_l$ otherwise. Set $\widetilde{\varphi}_{\widetilde{\boldsymbol{i}}} = \mathcal{F}_L \varphi_{\boldsymbol{i}}$ where $\widetilde{i}_l = -i_l$ if $l \in L$ and $\widetilde{l}_l = i_l$ otherwise. For any given finite sum we have

$$\mathcal{F}_L \left(\sum_{\boldsymbol{i} \in \mathbb{Z}^d} 2^{-\boldsymbol{i\nu}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}) \right) = \sum_{\widetilde{\boldsymbol{i}} \in \mathbb{Z}^d} 2^{-\widetilde{\boldsymbol{i}}\widetilde{\boldsymbol{\nu}}} \widetilde{\varphi}_{\widetilde{\boldsymbol{i}}}^{(2^{\widetilde{\boldsymbol{i}}})}(\boldsymbol{x}).$$

(2.3.3) So it is sufficient to prove the statement only for those $\boldsymbol{\nu}$ so that $\nu_k < 0$ for all $k \in \{1, \dots, d\}$. Furthermore, in this case, it is sufficient to check the uniformity of the size conditions (2.3.3) and the point-wise convergence a.e. As a matter of fact, if this is proved, we can use the Lebesgue dominated convergence Theorem to affirm that the series converges to an L^1_{loc} function that satisfies (2.3.3) that by Lemma 2.4.2 is a $\mathcal{PK}(\boldsymbol{\nu})$ kernel.

Suppose now that $\nu_k < 0$ for all $k \in \{1, \dots, d\}$. The condition that $\{\varphi_{\boldsymbol{i}}\} \subset \widehat{S}_0^{-\boldsymbol{q}-\boldsymbol{\nu}}(\mathbb{R}^N)$ on guarantees that for all $k \in \{1, \dots, d\}$ and for any α_k such that $\|\alpha_k\| \leq -q_k - \nu_k$ we have

$$\partial_{x_k}^{\alpha_k} \varphi_{\boldsymbol{i}}(\boldsymbol{x}_k \oplus 0_k) = 0$$

But this means that for all $\varphi_{\boldsymbol{i}}$ the following estimates hold uniformly:

$$\left| \partial_{\boldsymbol{x}}^{\boldsymbol{\beta}} \varphi_{\boldsymbol{i}}(\boldsymbol{x}) \right| \leq C_{\boldsymbol{\beta}} \frac{\|x_1\|^{1+\lfloor -q_1-\nu_1 \rfloor - \|\beta_1\|}}{(1+\|x_1\|)^{Q_1}} \cdots \frac{\|x_d\|^{1-\lfloor -q_d-\nu_d \rfloor - \|\beta_d\|}}{(1+\|x_d\|)^{Q_d}}$$

for any fixed, arbitrarily large, $Q_1, \dots, Q_d \in \mathbb{N}$. Using these estimates we write

$$\left| \partial_{\boldsymbol{x}}^{\boldsymbol{\beta}} \sum_{\boldsymbol{i} \in \mathbb{Z}^d} 2^{-\boldsymbol{i\nu}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}) \right| \leq$$

$$C_{\boldsymbol{\beta}} \sum_{\boldsymbol{i} \in \mathbb{Z}^d} 2^{-\boldsymbol{i}\left(\boldsymbol{q}+\boldsymbol{\nu}+(\|\beta_1\|,\dots,\|\beta_d\|)\right)} \frac{\left\|2^{-i_1} \cdot x_1\right\|^{1+\lfloor -q_1-\nu_1 \rfloor - \|\beta_1\|}}{(1+\left\|2^{-i_1} \cdot x_1\right\|)^{Q_1}} \cdots \frac{\left\|2^{-i_d} \cdot x_d\right\|^{1+\lfloor -q_d-\nu_d \rfloor - \|\beta_d\|}}{(1+\left\|2^{-i_d} \cdot x_d\right\|)^{Q_d}} <$$

$$C_{\boldsymbol{\beta}} \left(\prod_{k=1}^d \|x_k\|^{-q_k-\nu_k-\|\beta_k\|} \right) \sum_{\boldsymbol{i} \in \mathbb{Z}^d} \frac{\left\|2^{-i_1} \cdot x_1\right\|^{1+\lfloor -q_1-\nu_1 \rfloor - (-q_1-\nu_1)}}{(1+\left\|2^{-i_1} \cdot x_1\right\|)^{Q_1}} \cdots \frac{\left\|2^{-i_d} \cdot x_d\right\|^{1+\lfloor -q_d-\nu_d \rfloor - (-q_d-\nu_d)}}{(1+\left\|2^{-i_d} \cdot x_d\right\|)^{Q_d}}$$

and the sum on the right hand side converges to a quantity uniformly bounded in $\boldsymbol{x}$. This is true by homogeneity. As a matter of fact the quantity

$$\sum_{\boldsymbol{i} \in \mathbb{Z}^d} \frac{\left\|2^{-i_1} \cdot x_1\right\|^{1+\lfloor -q_1-\nu_1 \rfloor - (-q_1-\nu_1)}}{(1+\left\|2^{-i_1} \cdot x_1\right\|)^{Q_1}} \cdots \frac{\left\|2^{-i_d} \cdot x_d\right\|^{1+\lfloor -q_d-\nu_d \rfloor - (-q_d-\nu_d)}}{(1+\left\|2^{-i_d} \cdot x_d\right\|)^{Q_d}}$$

if invariant under the transformation $\boldsymbol{x} \to 2^{\boldsymbol{j}} \cdot x$ for any $\boldsymbol{j} \in \mathbb{Z}^d$. So we must check the boundedness only on the product of compact dyadic coronas. The series converges absolutely on such sets because the exponents $1 + \lfloor -q_k - \nu_k \rfloor - (-q_k - \nu_k)$ are positive and so the limit is continuous. This concludes the proof. □

We now proceed to prove that any kernel of $\mathcal{PK}(\boldsymbol{\nu})$ admits a dyadic decomposition i.e. there is a family of Schwartz functions $\{\varphi_{\boldsymbol{i}}\}$ such that (2.7.1) holds. This theorem is the inverse of Theorem 2.7.2

Theorem 2.7.3 (Dyadic decomposition for product kernels)**.**
For any kernel $K \in \mathcal{PK}(\boldsymbol{\nu})$ there is a uniformly bounded set of Schwartz functions $\{\varphi_{\boldsymbol{i}}\}_{\boldsymbol{i}\in\mathbb{Z}^d}$ such that the decomposition (2.7.1) *holds:*

$$K(\boldsymbol{x}) = \sum_{\boldsymbol{i}\in\mathbb{Z}^d} 2^{-\boldsymbol{i}\boldsymbol{\nu}}\varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}) = \sum_{\boldsymbol{i}\in\mathbb{Z}^d} 2^{-i_1(q_1+\nu_1)}\dots 2^{-i_d(q_d+\nu_d)}\varphi_{\boldsymbol{i}}(2^{-i_1}\cdot x_1,\dots,2^{-i_d}\cdot x_d).$$

The sum is intended in the weak (distributional) sense and the partial sums are a uniformly bounded family of $\mathcal{PK}(\boldsymbol{\nu})$ kernels.

The family of functions $\{\varphi_{\boldsymbol{i}}\}$ satisfy the following conditions.

- *If $k \in \{1,\dots,d\}$ is such that $\nu_k > -q_k$ then $\{\widehat{\varphi}_{\boldsymbol{i}}\}$ are supported on the set $\left\{\boldsymbol{\xi}\,\middle|\,{}^1\!/\!_2 < \|\xi_k\| < 2\right\}$. In particular for all multi-indexes α_k*

$$\int_{H_k} x_k^{\alpha_k}\varphi_{\boldsymbol{i}}(\boldsymbol{x})\mathrm{d}x_k = 0.$$

- *If $k \in \{1,\dots,d\}$ is such that $\nu_k \le -q_k$ then $\{\varphi_{\boldsymbol{i}}\}$ are supported on the set $\left\{\boldsymbol{x}\,\middle|\,{}^1\!/\!_2 < \|x_k\| < 2\right\}$ and in particular for all multi-indexes α_k*

$$\int_{H_k} \xi_k^{\alpha_k}\widehat{\varphi}_{\boldsymbol{i}}(\boldsymbol{\xi})\mathrm{d}\xi_k = 0.$$

This dyadic decomposition has a continuity property in the sense that for any neighborhood V of 0 in $S(\mathbb{R}^N)$ there is a neighborhood U_V of 0 in $\mathcal{PK}(\boldsymbol{\nu})$ such that if $K \in U_V$ then it admits a dyadic decomposition $\{\varphi_{\boldsymbol{i}}\} \subset V$.

Proof. Let $G = \left\{k \in \{1,\dots,d\}\,|\,\nu_k > -q_k\right\}$. Begin by applying the partial Fourier transform along all subspaces H_k with $k \in G$. In particular consider

$$\mathcal{F}_G K(\boldsymbol{x}) \overset{\text{def}}{=} \mathcal{F}_{k_1}\dots\mathcal{F}_{k_{|G|}}K(\boldsymbol{x})$$

where $k_1,\dots,k_{|G|}$ are indexes of the factors in G. Let $\widetilde{\boldsymbol{\nu}} \in \mathbb{R}^d$ be the multi-order such that $\widetilde{\nu}_k = -q_k - \nu_k$ if $k \in G$ and $\widetilde{\nu}_k = \nu_k$ otherwise. By Theorem 2.5.2 $\mathcal{F}_G K \in \mathcal{PK}(\widetilde{\boldsymbol{\nu}})$ and in particular, by Lemma 2.4.2 it coincides with an $L^1_{loc}(\mathbb{R}^N)$ function on the whole $\mathbb{R}^N$. Consider a cutoff function $\boldsymbol{\eta}(\boldsymbol{x}) = \prod_{k=1}^d \eta(x_k)$ and a product dyadic corona $\boldsymbol{\psi}(\boldsymbol{x}) = \prod_{k=1}^d \left(\eta_k(2^{-1}\cdot x_k) - \eta_k(x_k)\right)$ that is supported on the corona product set $\left\{\boldsymbol{x}\mid {}^1\!/\!_2 < \left\|x_j\right\| < 2\ \forall j \in \{1,\dots,d\}\right\}$. We have that

$$\sum_{\boldsymbol{i}\in\mathbb{Z}^d} \psi(2^{-\boldsymbol{i}}\cdot\boldsymbol{x}) \to 1$$

point-wise a.e. . Since $\mathcal{F}_G K$ is locally integrable we have that

$$\sum_{\boldsymbol{i}\in\mathbb{Z}^d} \psi(2^{-\boldsymbol{i}}\cdot\boldsymbol{x})\mathcal{F}_G K(\boldsymbol{x}) \rightharpoonup \mathcal{F}_G K.$$

Now set

$$\widetilde{\varphi}_{\boldsymbol{i}}(\boldsymbol{x}) = 2^{\boldsymbol{i}(\boldsymbol{q}+\widetilde{\boldsymbol{\nu}})}\psi(\boldsymbol{x})\mathcal{F}_G K(2^{\boldsymbol{i}}\cdot\boldsymbol{x})$$

for any $\boldsymbol{i} \in \mathbb{Z}^d$ so that

$$\mathcal{F}_G K = \sum_{\boldsymbol{i}\in\mathbb{Z}^d} 2^{-\boldsymbol{i}(\boldsymbol{q}+\widetilde{\boldsymbol{\nu}})}\widetilde{\varphi}_{\boldsymbol{i}}(2^{-\boldsymbol{i}}\cdot\boldsymbol{x}) = \sum_{\boldsymbol{i}\in\mathbb{Z}^d} 2^{-\boldsymbol{i}\widetilde{\boldsymbol{\nu}}}\widetilde{\varphi}_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}).$$

The functions $\widetilde{\varphi}_{\boldsymbol{i}}$ are smooth and supported on the corona product set. Using the size estimates on $\mathcal{F}_G K$ we get

$$\|\widetilde{\varphi}_{\boldsymbol{i}}\|_{C^b} = 2^{\boldsymbol{i}(\boldsymbol{q}+\widetilde{\boldsymbol{\nu}})} \left\| \psi(\boldsymbol{x}) \mathcal{F}_G K(2^{\boldsymbol{i}} \cdot \boldsymbol{x}) \right\|_{C^b} < C 2^{\boldsymbol{i}(\boldsymbol{q}+\widetilde{\boldsymbol{\nu}})} \left\| \mathcal{F}_G K(2^{\boldsymbol{i}} \cdot \boldsymbol{x})\big|_{\{1/2 < \|x_j\| < 1\}} \right\|_{C^b} < C_b$$

so the functions $\widetilde{\varphi}_{\boldsymbol{i}}$ are equi-bounded in $S(\mathbb{R}^N)$. Finally set $\varphi_{\boldsymbol{i}} = \mathcal{F}_G^{-1} \widetilde{\varphi}_{\widetilde{\boldsymbol{i}}}$ where $\widetilde{\boldsymbol{i}} \in \mathbb{Z}^d$ such that $\widetilde{i}_k = -i_k$ if $k \in G$ and $\widetilde{i}_k = i_k$ otherwise. Notice that such defined building blocks $\varphi_{\boldsymbol{i}}$ satisfy the conditions of theorem 2.7.2 and so the partial sums are uniformly bounded in the space of product kernels. Furthermore since $\mathcal{F}_G K$ is locally integrable on the whole $\mathbb{R}^N$ the limit of the partial Fourier transforms of the partial sums converge to the $\mathcal{F}_G K$ itself and so do the actual partial sums converge to K. □

Now we turn to a result that allows us to have a dyadic decomposition with "building blocks" localized in space. In particular, the procedure used in Theorem 2.7.3 guarantees the localization of the bump functions on circular coronas along those subspaces for which the order ν_k is less than $-q_k$. The same thing can be done for all subspaces and all orders. However if one asks for the functions to be localized one cannot ask for cancellation of all orders on that subspace. As a matter of fact any localized function with all moments vanishing is identically zero because Paley-Wiener's Theorem states that the Fourier transform of a compactly supported smooth function is analytic.

Theorem 2.7.4 (Dyadic decomposition with localized building blocks)**.**
For any kernel $K \in \mathcal{PK}(\boldsymbol{\nu})$ there is a uniformly bounded family of Schwartz functions $\{\varphi_{\boldsymbol{i}}\}_{\boldsymbol{i} \in \mathbb{Z}^d}$ such that the decomposition (2.7.1) *holds:*

$$K(\boldsymbol{x}) = \sum_{\boldsymbol{i} \in \mathbb{Z}^d} 2^{-\boldsymbol{i}\boldsymbol{\nu}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}).$$

The sum is intended in the weak (distributional) sense and the partial sums are a uniformly bounded family of $\mathcal{PK}(\boldsymbol{\nu})$ kernels.

The family of functions $\{\varphi_{\boldsymbol{i}}\}$ are supported on $\{\boldsymbol{x} \mid 1/4 < \|x_k\| < 4 \; \forall k \in \{1, \ldots, d\}\}$ and for any $\boldsymbol{g} \in \mathbb{N}^d$ they can be chosen to satisfy the following cancellation condition:

$$\int_{H_k} x_k^{\alpha_k} \varphi_{\boldsymbol{i}}(\boldsymbol{x}) \mathrm{d}x_k \qquad \forall \, \|\alpha_k\| < g_k$$

for all indexes k such that $\nu_k > -q_k$.

The proof is the consequence of the following lemma about the possibility of localizing in space a "building block" on a given scale, without losing cancellation conditions up to small corrections.

Lemma 2.7.5 ("Building block" localization)**.**
Let $\varphi \in S_0^{\boldsymbol{g}}(\mathbb{R}^N)$ for some some $\boldsymbol{g}$ with $g_k \geq 0$. Then there is a family $\{\psi_{\boldsymbol{i}}\}_{\boldsymbol{i} \in \mathbb{Z}^d} \subset S_0^{\boldsymbol{g}}(\mathbb{R}^N)$ supported on the sets where $\frac{1}{4} < \|x_k\| < 4$ for any $k \in \{1, \ldots, d\}$ such that

$$\varphi(\boldsymbol{x}) = \sum_{\boldsymbol{i} \in \mathbb{Z}^d} \psi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x})$$

holds in the weak sense and the Schwartz norms of $\psi_{\boldsymbol{i}}$ decay at least like the homogeneous dimension

$$\|\boldsymbol{x}^{\boldsymbol{\gamma}} \partial^{\boldsymbol{\tau}} \psi_{\boldsymbol{i}}\|_\infty < C_{\boldsymbol{\gamma},\boldsymbol{\tau}} \frac{2^{i_1 q_1}}{\left(1 + 2^{i_1}\right)^{Q_1}} \cdots \frac{2^{i_d q_d}}{\left(1 + 2^{i_d}\right)^{Q_d}}$$

where $\boldsymbol{Q} \in \mathbb{N}^d$ are some arbitrarily large integers and $C_{\boldsymbol{\gamma},\boldsymbol{\tau}}$ depends only on $\boldsymbol{\gamma}$, $\boldsymbol{\tau}$ and the original bounds on φ.

Proof. We prove this by induction on the length d of the product decomposition of $\mathbb{R}^N$. We reason along H_1 and regard $\boldsymbol{x}' = \boldsymbol{x}_{(1)} = (x_2, \ldots, x_d)$ as parameters. Let η be a cutoff function along H_1 and set $\theta(x_1) = \eta(2^{-1} \cdot x_1) - \eta(x_1)$. Thus θ is a smooth function supported on the corona $1/2 < \|x_1\| < 4$. Now for any index α on H_1 such that $|\alpha| \leq g_1$ choose a smooth function $f_\alpha \in \mathbb{C}_c^\infty\left(B(0,4) \setminus B(0,1/2)\right)$ on H_1 such that

$$\int_{H_1} x_1^\beta f_\alpha(x_1)\,\mathrm{d}x_1 = \delta_\alpha^\beta = \begin{cases} 1 & \text{if } \alpha = \beta \\ 0 & \text{if } \alpha \neq \beta \end{cases}.$$

This can always be done inductively. Define

$$\begin{aligned} a_i^\alpha(\boldsymbol{x}') &= \int_{H_1} x_1^\alpha \varphi(x_1, \boldsymbol{x}')\theta(2^{-1} \cdot x_1)\,\mathrm{d}x_1 \\ A_i^\alpha(\boldsymbol{x}') &= \sum_{j \geq i} a_j^\alpha(\boldsymbol{x}') = \sum_{j<i} a_j^\alpha(\boldsymbol{x}') \end{aligned}$$

because $\sum_{j\in\mathbb{Z}} a_j^\alpha(\boldsymbol{x}') = 0$ by hypothesis. Furthermore we have that

$$\left|a_i^\alpha(\boldsymbol{x}')\right| < C_\alpha \int_{B(0,4\cdot 2^i)} \|x\|^{|\alpha|} \left|\theta(2^{-1} \cdot x_1)\right| \mathrm{d}x_1 < C_\alpha 2^{i(q_1+|\alpha|)}$$

and for an arbitrarily large Q_1

$$\left|a_i^\alpha(\boldsymbol{x}')\right| < C_\alpha \int_{H_1 \setminus B(0,2^{i-1})} \|x\|^{|\alpha|} \left(1 + \|x\|\right)^{-Q_1-q_1-|\alpha|} \left|\theta(2^{-1} \cdot x_1)\right| \mathrm{d}x_1 < C_\alpha(1+2^i)^{Q_1}$$

so

$$\left|a_i^\alpha(\boldsymbol{x}')\right| < C_\alpha \frac{2^{i(q_1+|\alpha|)}}{(1+2^i)^{Q_1}}$$

and also

$$\left|A_i^\alpha(\boldsymbol{x}')\right| < C_\alpha \frac{2^{i(q_1+|\alpha|)}}{(1+2^i)^{Q_1}}.$$

Since $\sum_{i\in\mathbb{Z}} \theta(2^{-i} \cdot x_1) \to 1$ a.e. and in the distributional sense we have

$$\begin{aligned} \varphi(\boldsymbol{x}) =& \sum_{i\in\mathbb{Z}} \theta(2^{-i} \cdot x_1)\varphi(\boldsymbol{x}) = \\ =& \sum_{i\in\mathbb{Z}} \Bigg(\varphi(\boldsymbol{x})\theta(2^{-i} \cdot x_1) - \sum_{|\alpha|\leq l_1} a_i^\alpha(\boldsymbol{x}') 2^{-i(q_1+|\alpha|)} f_\alpha(2^{-i} \cdot x_1) + \\ & \sum_{|\alpha|\leq l_1} \left(A_{i+1}^\alpha(\boldsymbol{x}') - A_i^\alpha(\boldsymbol{x}')\right) 2^{-i(q_1+|\alpha|)} f_\alpha(2^{-i} \cdot x_1) \Bigg) = \\ & \sum_{i\in\mathbb{Z}} \Bigg(\varphi(\boldsymbol{x})\theta(2^{-i} \cdot x_1) - \sum_{|\alpha|\leq l_1} a_i^\alpha(\boldsymbol{x}') 2^{-i(q_1+|\alpha|)} f_\alpha(2^{-i} \cdot x_1) + \\ & \sum_{|\alpha|\leq l_1} A_i^\alpha(\boldsymbol{x}') \left(2^{-(i-1)(q_1+|\alpha|)} f_\alpha(2^{-(i-1)} \cdot x_1) - 2^{-i(q_1+|\alpha|)} f_\alpha(2^{-i} \cdot x_1)\right) \Bigg). \end{aligned}$$

Now set

$$\psi_i^{(2^i)}(x_1,\boldsymbol{x}') \stackrel{\text{def}}{=} \Bigg(\varphi(\boldsymbol{x})\theta(2^{-i}\cdot x_1) - \sum_{|\alpha|\le l_1} a_i^\alpha(\boldsymbol{x}')2^{-i(q_1+|\alpha|)}f_\alpha(2^{-i}\cdot x_1) +$$

$$+\sum_{|\alpha|\le l_1} A_i^\alpha(\boldsymbol{x}')\left(2^{-(i-1)(q_1+|\alpha|)}f_\alpha(2^{-(i-1)}\cdot x_1) - 2^{-i(q_1+|\alpha|)}f_\alpha(2^{-i}\cdot x_1)\right)\Bigg)$$

so that

$$\psi_i(x_1,\boldsymbol{x}') \stackrel{\text{def}}{=} \Bigg(2^{iq_1}\varphi(2^i\cdot x_1,\boldsymbol{x}')\theta(x_1) - \sum_{|\alpha|\le l_1} a_i^\alpha(\boldsymbol{x}')2^{-i|\alpha|}f_\alpha(x_1) +$$

$$+\sum_{|\alpha|\le l_1} 2^{-i|\alpha|}A_i^\alpha(\boldsymbol{x}')\left(2^{q_1+|\alpha|}f_\alpha(2\cdot x_1) - f_\alpha(x_1)\right)\Bigg).$$

By the inequalities on $a_i^\alpha(\boldsymbol{x}')$ and $A_i^\alpha(\boldsymbol{x}')$ we get the needed decay on all except the first term. For the term $2^{iq_1}\varphi(2^i\cdot\ x_1,\boldsymbol{x}')\theta(x_1)$ notice that

$$\left\|x_1^\gamma\partial_{x_1}^\tau 2^{iq_1}\varphi(2^i\cdot\ x_1,\boldsymbol{x}')\theta(x_1)\right\|_\infty < C_{\gamma,\tau}2^{i(q_1+|\tau|)}\left\|\partial_{x_1}^\tau\varphi(x_1,\boldsymbol{x}')\right\|_\infty$$

and

$$\left\|x_1^\gamma\partial_{x_1}^\tau 2^{iq_1}\varphi(2^i\cdot\ x_1,\boldsymbol{x}')\theta(x_1)\right\|_\infty < C_{\gamma,\tau}\left\|\partial_{x_1}^\tau\varphi(x_1,\boldsymbol{x}')|_{H_1\setminus B(0,2^{i-2})}\right\|_\infty < C_{\gamma,\tau}(1+2^i)^{-Q_1}.$$

All of the addends are supported on the needed dyadic corona and it can be easily checked that

$$\int_{H_1} x_1^\alpha\psi_i(x_1,\boldsymbol{x}')\,\mathrm{d}x_1 = 0$$

for any $|\alpha|\le g_1$.

The proof is then concluded by induction. As a matter of fact we can now write

$$\varphi(\boldsymbol{x}) = \sum_{\boldsymbol{i}_1\in\mathbb{Z}} 2^{-i_1q_i}\psi_{i_1}(2^{-i_1}\cdot x_1,\boldsymbol{x}')$$

and apply the induction hypothesis along $\boldsymbol{x}'$ to each ψ_{i_1} with x_1 as a parameter. □

We now proceed to the proof of Theorem 2.7.4.

Proof of Theorem 2.7.4. Take $K\in\mathcal{PK}(\boldsymbol{\nu})$ and write the associated dyadic decomposition (2.7.1) given by Theorem 2.7.3:

$$K(\boldsymbol{x}) = \sum_{\boldsymbol{i}\in\mathbb{Z}^d} 2^{-\boldsymbol{i}\boldsymbol{\nu}}\varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}).$$

Let $G = \{k\in\{1,\dots,d\}\mid \nu_k > -q_k\}$. The functions $\varphi_{\boldsymbol{i}}$ are uniformly bounded on $S(\mathbb{R}^N)$, are supported on $\{\boldsymbol{x}\mid 1/2 < \|x_k\| < 2\ \forall j\notin G\}$. Without loss of generality, suppose that $G=\{1,\dots,|G|\}$ are the first $|G|$ subspace indexes and indicate $\boldsymbol{x}_G = (x_1,\dots,x_{|G|})$ and $\boldsymbol{x}' = \boldsymbol{x}_{(G)} = (x_{|G|+1},\dots,x_d)$. Recall that for $k\in G$ we have that

$$\int_{H_k} x_k^{\alpha_k}\varphi_{\boldsymbol{i}}(\boldsymbol{x})\mathrm{d}x_k = 0 \qquad \forall\alpha_k$$

To prove the theorem we will have to localize the dyadic decomposition along the subspaces H_k for $k \in G$ so from now on we will think of the coordinates $\boldsymbol{x}'$ as parameters.

Choose $\boldsymbol{l}$ such that $l_k > \nu_k$ and $\boldsymbol{Q}$ such that $Q_k > \nu_k - q_k$ for all $k \in G$. For every $\varphi_{\boldsymbol{i}}$ apply lemma 2.7.5 and write

$$\varphi_{\boldsymbol{i}}(\boldsymbol{x}) = \sum_{\widetilde{\boldsymbol{i}} \in \mathbb{Z}^{|G|}} \psi_{\boldsymbol{i},\widetilde{\boldsymbol{i}}}^{(2^{\widetilde{\boldsymbol{i}}})}(\boldsymbol{x}_{|G|}, \boldsymbol{x}')$$

where we intend that the rescaling occurs only along $\boldsymbol{x}_{|G|}$. The sum holds in the distributional sense and $\psi_{\boldsymbol{i},\widetilde{\boldsymbol{i}}}$ are uniformly bounded functions in $S(\mathbb{R}^N)$ and are supported on $\{\boldsymbol{x} \mid 1/2 < \|x_k\| < 4 \; \forall k \in \{1, \ldots, d\}\}$.

As a consequence, any finite partial sum we have

$$\sum_{\boldsymbol{i} \in \mathbb{Z}^d} 2^{-\boldsymbol{i}\boldsymbol{\nu}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{i}=(\boldsymbol{i}_G,\boldsymbol{i}') \in (\mathbb{Z}^{|G|} \times \mathbb{Z}^{d-|G|}) \\ \widetilde{\boldsymbol{i}} \in (\mathbb{Z}^{|G|} \times \{0\}^{d-|G|})}} 2^{-\boldsymbol{i}\boldsymbol{\nu}} \psi_{\boldsymbol{i},\widetilde{\boldsymbol{i}}}^{(2^{\boldsymbol{i}+\widetilde{\boldsymbol{i}}})}(\boldsymbol{x})$$

Now, changing the set of summation so as to group "building blocks" on the same scale, we get with $\boldsymbol{h} = \boldsymbol{i} + \widetilde{\boldsymbol{i}}$

$$\sum_{\boldsymbol{i} \in \mathbb{Z}^d} 2^{-\boldsymbol{i}\boldsymbol{\nu}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{h}=(\boldsymbol{h}_G,\boldsymbol{h}') \in (\mathbb{Z}^{|G|} \times \mathbb{Z}^{d-|G|}) \\ \widetilde{\boldsymbol{i}} \in (\mathbb{Z}^{|G|} \times \{0\}^{d-|G|})}} 2^{-\boldsymbol{h}\boldsymbol{\nu}} 2^{\widetilde{\boldsymbol{i}}\boldsymbol{\nu}} \psi_{\boldsymbol{h}+\widetilde{\boldsymbol{i}},\widetilde{\boldsymbol{i}}}^{(2^{\boldsymbol{h}})}(\boldsymbol{x}).$$

Notice that $2^{\widetilde{\boldsymbol{i}}\boldsymbol{\nu}} = 2^{\widetilde{\boldsymbol{i}}_G \boldsymbol{\nu}_G}$ where $\widetilde{\boldsymbol{i}}_G = (\widetilde{i}_1, \ldots, \widetilde{i}_{|G|})$ and $\boldsymbol{\nu}_G = (\nu_1, \ldots, \nu_{|G|})$ so just set

$$\widetilde{\varphi}_{\boldsymbol{i}}(\boldsymbol{x}) = \sum_{\widetilde{\boldsymbol{i}}_G \in \mathbb{Z}^{|G|}} 2^{\widetilde{\boldsymbol{i}}_G \boldsymbol{\nu}_G} \psi_{(\boldsymbol{i}_G+\widetilde{\boldsymbol{i}}_G,\boldsymbol{i}'),\widetilde{\boldsymbol{i}}_G}(\boldsymbol{x}).$$

However since for any $k \in G$ we have that $\nu_k + q_k > 0$ and $\nu_k + q_k - Q_k < 0$, by the decay given by 2.7.5 the sum above converges uniformly with all derivatives to a compactly supported smooth function. So

$$\int_{H_k} x_k^{\alpha_k} \widetilde{\varphi}_{\boldsymbol{i}}(\boldsymbol{x}) \mathrm{d}x_k = 0 \qquad \forall \|\alpha_k\| < l_k$$

for all $k \in G$, and all these "building blocks" are supported on $\{\boldsymbol{x} \mid 1/2 < \|x_k\| < 4 \; \forall k \in \{1, \ldots, d\}\}$. Furthermore the dyadic sum holds

$$K(\boldsymbol{x}) = \sum_{\boldsymbol{i} \in \mathbb{Z}^d} 2^{-\boldsymbol{i}\boldsymbol{\nu}} \widetilde{\varphi}_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x})$$

and the partial sums are equi-bounded since the hypotheses of theorem 2.7.2 hold. □

As a consequence of these lemmas we have a uniform approximation property for product kernels of arbitrary order.

Corollary 2.7.6.
Let $K \in \mathcal{PK}(\boldsymbol{\nu})$. There exists a uniformly bounded sequence K_n of kernels in $\mathcal{PK}(\boldsymbol{\nu})$ that coincide with smooth functions with compact support and that weakly converge to K.

Proof. Use theorem 2.7.4 to write a dyadic decomposition for K with localized "building blocks". The partial sums are kernels with compact support uniformly in $\mathcal{PK}(\boldsymbol{\nu})$. □

2.8 Further properties

2.8.1 Different product decompositions

We now deal with the relationship between product kernels adapted to different product decompositions of $\mathbb{R}^N$.

Definition 2.8.1 (Coarser product decomposition).
Consider two product decompositions $\{\widetilde{H}_j\}_{j\in\{1,\dots,\widetilde{d}\}}$ *and* $\{H_k\}_{k\in\{1,\dots,d\}}$. *The decomposition* $\{H_k\}_{k\in\{1,\dots,d\}}$ *is coarser than* $\{\widetilde{H}_j\}_{j\in\{1,\dots,\widetilde{d}\}}$ *if there is a partition* $J_1 \cup \ldots \cup J_d = \{1,\dots,\widetilde{d}\}$ *consisting of non-empty sets* J_k *such that* $H_k = \bigoplus_{j\in J_k} \widetilde{H}_j$. *Vice-versa* $\{\widetilde{H}_j\}$ *is said to be finer than* $\{H_k\}$.

When working with product kernels we need to take in account what naturally happens to the multi-orders if we are considering with a coarser product decomposition.

Definition 2.8.2 (Admissibly coarser orders).
Let $\{\widetilde{H}_j\}_{j\in\{1,\dots,\widetilde{d}\}}$ *and* $\{H_k\}_{k\in\{1,\dots,d\}}$ *be two product decompositions of* $\mathbb{R}^N$ *such that* $\{H_k\}$ *is coarser than* $\{\widetilde{H}_j\}$ *and* $H_k = \oplus_{j\in J_k}\widetilde{H}_j$. *Given two multi-orders* $\boldsymbol{\nu} \in R^d$ *and* $\widetilde{\boldsymbol{\nu}} \in \mathbb{R}^{\widetilde{d}}$ *relative to the decompositions* $\{H_k\}$ *and* $\{\widetilde{H}_j\}$ *respectively, we say that* $\boldsymbol{\nu}$ *is admissibly coarser than the order* $\widetilde{\boldsymbol{\nu}}$ *(*$\widetilde{\boldsymbol{\nu}}$ *is admissibly finer than* $\boldsymbol{\nu}$*) if* $\nu_k = \sum_{j\in J_k}\widetilde{\nu}_j$ *and for any* $k \in \{1,\dots,d\}$ *such that* $|J_k| \geq 2$ *we have that* $-\widetilde{q}_j \leq \widetilde{\nu}_j < 0$ *or* $-\widetilde{q}_j < \widetilde{\nu}_j \leq 0$ *for all* $j \in J_k$.

We have the following result about product kernels adapted to different product decompositions

Proposition 2.8.3 (Product kernels adapted to different product decompositions).
Given two product decompositions $\{\widetilde{H}_j\}_{j\in\{1,\dots,\widetilde{d}\}}$ *and* $\{H_k\}_{k\in\{1,\dots,d\}}$ *of* $\mathbb{R}^N$ *with respective multi-orders* $\widetilde{\boldsymbol{\nu}} \in \mathbb{R}^{\widetilde{d}}$ *and* $\boldsymbol{\nu} \in R^d$ *such that* $\{H_k\}$ *is coarser than* $\{\widetilde{H}_j\}$ *and* $\boldsymbol{\nu}$ *is admissibly coarser than* $\widetilde{\boldsymbol{\nu}}$ *then*

$$\mathcal{PK}_{\{H_k\}}(\boldsymbol{\nu}) \subset \mathcal{PK}_{\{\widetilde{H}_j\}}(\widetilde{\boldsymbol{\nu}})$$

and the inclusion is continuous.

Proof. Let $H_k = \bigoplus_{j\in J_k} \widetilde{H}_j$. First let us suppose that $\nu_k < 0$ for all $k \in \{1,\dots,d\}$ and that for any k such that $|J_k| \geq 2$ and for any $j \in J_k$ the inequality $-\widetilde{q}_j \leq \widetilde{\nu}_j < 0$ holds. Lemma 2.4.2 is applicable to kernels in $\mathcal{PK}_{\{H_k\}}(\boldsymbol{\nu})$ and $\mathcal{PK}_{\{\widetilde{H}_j\}}(\widetilde{\boldsymbol{\nu}})$. It states that such kernels coincides with an L^1_{loc} function on the whole $\mathbb{R}^N$. To check that a $K \in \mathcal{PK}_{\{H_k\}}(\boldsymbol{\nu})$, being locally integrable, is in $\widetilde{\mathcal{PK}}(\widetilde{\boldsymbol{\nu}})$ it suffices to check the size conditions (2.3.3), so all we need to prove is that

$$\prod_{k=1}^{d} \|x_k\|^{-q_k-\nu_k-\|\alpha_k\|} \leq C \prod_{k=1}^{d}\prod_{j\in J_k} \left\|\widetilde{x}_j\right\|^{-\widetilde{q}_j-\widetilde{\nu}_j-\left\|\widetilde{\alpha}_j\right\|}$$

where, supposing $J_k = \{j_{k,1},\dots,j_{k,|J_k|}\}$, $\alpha_k = (\widetilde{\alpha}_{j_{k,1}},\dots,\widetilde{\alpha}_{j_{k,|J_k|}})$, $q_k = \sum_{j\in J_k}\widetilde{q}_j$, and $\nu_k = \sum_{j\in J_k}\widetilde{\nu}_j$ as usual. Bearing in mind that $\|x_k\| \approx \sum_{j\in J_k}\left\|\widetilde{x}_j\right\|$, and $\|\alpha_k\| = \sum_{j\in J_k}\left\|\widetilde{\alpha}_j\right\|$, the generalized Young's inequality gives that

$$\prod_{j\in J_k} \left\|\widetilde{x}_j\right\|^{\widetilde{q}_j+\widetilde{\nu}_j+\left\|\widetilde{\alpha}_j\right\|} \leq C\left(\sum_{j\in J_k}\left\|\widetilde{x}_j\right\|\right)^{q_k+\nu_k+\|\alpha_k\|} \approx \|x_k\|^{q_k+\nu_k+\|\alpha_k\|}$$

as long as all the exponents are positive (which is the case by hypotheses). This gives the above inequality and thus

$$\left|\partial_x^\alpha K(\boldsymbol{x})\right| \le C \prod_{k=1}^d \|x_k\|^{-q_k-\nu_k-\|\alpha_k\|} \le C \prod_{k=1}^d \prod_{j\in J_k} \|\widetilde{x}_j\|^{-\widetilde{q}_j-\widetilde{\nu}_j-\|\widetilde{\alpha}_j\|}$$

proving the statement.

In the more general case we use the Fourier transform. Let

$$L = \left\{k \,\middle|\, |J_k| \ge 2 \text{ and } -q_j < \widetilde{\nu}_j, \le 0 \quad \forall j \in J_k\right\} \cup \left\{k \,\middle|\, |J_k| = 1 \text{ and } \nu_k \ge 0\right\}$$

and let $\mathcal{F}_L$ be the Fourier transform along the subspaces indexed by L. Let $\boldsymbol{\mu}$ be such that $\mu_k = -q_k - \nu_k$ if $k \in L$ and $\mu_k = \nu_k$ otherwise. Correspondingly let $\widetilde{\mu}_j = -\widetilde{q}_j - \widetilde{\nu}_j$ if $j \in J_k$ with $k \in L$ and $\widetilde{\mu}_j = \widetilde{\nu}_j$ otherwise. By Theorem 2.5.2 $\mathcal{F}_L : \mathcal{PK}_{\{H_k\}}(\boldsymbol{\nu}) \to \mathcal{PK}_{\{H_k\}}(\boldsymbol{\mu})$ and $\mathcal{F}_L^{-1} : \mathcal{PK}_{\{\widetilde{H}_j\}}(\widetilde{\boldsymbol{\mu}}) \to \mathcal{PK}_{\{H_k\}}(\widetilde{\boldsymbol{\nu}})$ and both mappings are continuous. Notice that $\boldsymbol{\mu}$ is admissibly finer than $\widetilde{\boldsymbol{\mu}}$ and in particular $-q_j \le \widetilde{\mu}_j < 0$ for any $j \in J_k$ with k such that $|J_k| \ge 2$ while for any $j \in J_k$ the inequality $\widetilde{\mu}_j < 0$ holds. Applying the first part of the proof one gets that

$$\mathcal{F}_L^{-1} \circ \mathcal{F}_L : \mathcal{PK}_{\{H_k\}}(\boldsymbol{\nu}) \to \mathcal{PK}_{\{\widetilde{H}_j\}}(\widetilde{\boldsymbol{\nu}})$$

and is continuous. This concludes the proof. □

2.8.2 Kernels with bounded support

In Section 2.4 we talked about kernels with bounded support along certain subspaces in the product decomposition of $\mathbb{R}^N$. These kinds of kernels have some interesting properties. In general there are no non-trivial inclusions between classes of product kernels of different orders. However this is not true if we are dealing with kernels with bounded support

Proposition 2.8.4 (Kernels of bounded support of different orders)**.**
Let $\boldsymbol{\nu}$ and $\boldsymbol{\mu}$ be two multi-orders such that $\nu_k \le \mu_k$ for all $k \in \{1,\dots d\}$ and if $\nu_j \lneqq \mu_j$ for some j then $\mu_j \le 0$. Let $J \subset \{1,\dots,d\}$ be the subset of those indexes for which $\nu_j \lneqq \mu_j$. For any fixed constants $M_j > 0$ with $j \in J$ we have that

$$\left\{K \in \mathcal{PK}(\boldsymbol{\nu}) \,\middle|\, \operatorname{spt} K \subset \left\{\boldsymbol{x} \,\middle|\, \|x_j\| < M_j \ \forall j \in J\right\}\right\} \subset \mathcal{PK}(\boldsymbol{\mu})$$

and the inclusion is continuous.

This proposition can be proved by verifying the necessary cancellation and size conditions by hand. However it also follows immediately from the following important property.

Theorem 2.7.4 does provide us with a dyadic decomposition for kernels in $\mathcal{PK}(\boldsymbol{\nu})$ but, even though it uses localized "building blocks", the decomposition does not respect the boundedness of the support of a given kernel. As a matter of fact even if the support of K is bounded along a certain subspace H_k the dyadic sum could span all the indexes $i_k \in \mathbb{Z}$. However, to address this shortcomings in dealing with kernels with bounded support, we have the following two results

Theorem 2.8.5 (Convergence of dyadic series with bounded indexes)**.**
Given a subset of indexes $J \subset \{1,\dots,d\}$ and bounds on indexes $m_j \in Z$ for $j \in J$ let

$$\{\varphi_{\boldsymbol{i}}\}_{\substack{\boldsymbol{i}\in\mathbb{Z}^d \\ i_j \le m_j \ \forall j\in J}}$$

be a bounded family of Schwartz functions such that $\varphi_{\boldsymbol{i}}$ satisfy the hypothesis of Theorem 2.7.2 along subspaces H_k with $k \notin J$. Let us also suppose that the functions $\{\varphi_{\boldsymbol{i}}\}$ satisfy the following properties along the subspaces H_j with $j \in J$ for which $\nu_j \ge 0$.

- *If $i_j < m_j$ then*
$$\int_{H_j} x_j^{\alpha_j} \varphi_{\boldsymbol{i}}(\boldsymbol{x}) \mathrm{d}x_j = 0 \qquad \forall \left\|\alpha_j\right\| \leq \nu_j.$$
- *If $i_j = m_j$ then*
$$\int_{H_j} x_j^{\alpha_j} \varphi_{\boldsymbol{i}}(\boldsymbol{x}) \mathrm{d}x_j = 0 \qquad \forall \left\|\alpha_j\right\| < \nu_j.$$

Then the dyadic sum
$$\sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ i_j \leq m_j \ \forall j \in J}} 2^{-\boldsymbol{i}\boldsymbol{\nu}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x})$$
converges to a product kernel in the distributional sense and all the partial sums are a uniformly bounded family of product kernels.

The dyadic series possesses a continuity property in the sense that for any neighborhood U of 0 in $\mathcal{PK}(\boldsymbol{\nu})$ there is a neighborhood V_U of 0 in $S(\mathbb{R}^N)$ such that if $\{\varphi_{\boldsymbol{i}}\} \subset V_U$ then all the partial sums and the whole series are in U.

Theorem 2.8.6 (Localized dyadic decomposition for kernels of bounded support)**.**
For any $K \in \mathcal{PK}(\boldsymbol{\nu})$ such that for a certain subset of indexes $J \subset \{1, \dots, d\}$ and for some $m_j \in \mathbb{Z}$ we have that
$$\operatorname{spt} K \subset \left\{ \boldsymbol{x} \,\middle|\, \left\|x_j\right\| < 2^{m_j} \ \ \forall j \in J \right\}.$$
There is a uniformly bounded set of Schwartz functions
$$\{\varphi_{\boldsymbol{i}}\}_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ i_j \leq m_j - 2 \ \forall j \in J}}$$
such that the dyadic decomposition
$$K(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ i_j \leq m_j - 2 \ \forall j \in J}} 2^{-\boldsymbol{i}\boldsymbol{\nu}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}) \tag{2.8.1}$$
holds. The series converges in the weak (distributional) sense and the partial sums are a uniformly bounded family of $\mathcal{PK}(\boldsymbol{\nu})$ kernels.

The functions $\{\varphi_{\boldsymbol{i}}\}$ are supported on the set $\left\{\boldsymbol{x} \,\middle|\, 1/4 < \|x_k\| < 4 \ \ \forall k \in \{1, \dots, d\}\right\}$. For any $\boldsymbol{l} \in N^d$ the family $\{\varphi_{\boldsymbol{i}}\}$ can be chosen so that along those spaces H_k with $k \notin J$, $\{\varphi_{\boldsymbol{i}}\}$ possess the properties given by Theorem 2.7.4 with cancellation order $\boldsymbol{l}$. For every $j \in J$ the functions $\{\varphi_{\boldsymbol{i}}\}$ possess the following cancellation property along H_j.

- *If $i_j < m_j - 2$ then*
$$\int_{H_j} x_j^{\alpha_j} \varphi_{\boldsymbol{i}}(\boldsymbol{x}) \mathrm{d}x_j = 0 \qquad \forall \left\|\alpha_j\right\| \leq l_j.$$
- *If $i_j = m_j - 2$ then*
$$\int_{H_j} x_j^{\alpha_j} \varphi_{\boldsymbol{i}}(\boldsymbol{x}) \mathrm{d}x_j = 0 \qquad \forall \left\|\alpha_j\right\| < \nu_j.$$

The decomposition possesses a continuity property in the sense that for any neighborhood V of 0 in $S(\mathbb{R}^N)$ there is a neighborhood U_V of 0 in $\mathcal{PK}(\boldsymbol{\nu})$ such that if $K \in U_V$ and K its support bounded as in the hypothesis then K admits a dyadic decomposition with $\{\varphi_{\boldsymbol{i}}\} \subset U$.

Proof. Let us reason by induction on the number of factors H_k along which the support of K is bounded. If the factors along which the support is bounded are 0 there is nothing to prove. Suppose the proposition is true for all kernels with supports bounded along up to $D-1$ factors H_k with $D \geq 1$.

Let the support of K be bounded along D subspaces one of which is H_k for a certain k. We have that

$$\eta_k(2^{-m_k} \cdot x_k)K(\boldsymbol{x}) = K(\boldsymbol{x})$$

where η_k is a cutoff function on H_k. Let us write the localized dyadic decomposition for K as given by Theorem 2.7.4 and apply the above observation:

$$K(\boldsymbol{x}) = \eta_k(2^{-m_k} \cdot x_k)K(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{i}\in\mathbb{Z}^d \\ i_j \leq m_j - 2\ \forall j \in J\setminus\{k\}}} 2^{-\boldsymbol{i}\boldsymbol{\nu}}\eta_k(2^{-m_k} \cdot x_k)\varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}).$$

For $i_k \geq m_k + 2$ the terms are identically 0, for $i_k < m_k - 2$ the cutoff function $\eta_k(2^{-m_k} \cdot x_k)$ is identically 1 on the support of $\varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x})$. Let us put

$$\widetilde{\varphi}_{\widetilde{\boldsymbol{i}}}(\boldsymbol{x}) \stackrel{\text{def}}{=} \begin{cases} \varphi_{\widetilde{\boldsymbol{i}}}(\boldsymbol{x}) & i_k < m_k - 2 \\ \displaystyle\sum_{\substack{i_j = \widetilde{i}_j\ j \neq k \\ i_k \in \{\widetilde{i}_k - 2, \dots, \widetilde{i}_k + 2\}}} 2^{-(i_k - \widetilde{i}_k)(q_k + \nu_k)}\varphi_{\boldsymbol{i}}\left(\boldsymbol{x}_{(k)} \oplus 2^{-(i_k - \widetilde{i}_k)} \cdot x_k\right) & i_k = m_k - 2 \\ 0 & i_k > m_k - 2 \end{cases}$$

The cancellation conditions for $i_k < m_k - 2$ are immediate while the ones for $i_k = m_k - 2$ can be checked using the fact that the series must converge to a product kernels of the correct order. □

2.9 Convolution algebra and functional calculus

We will now see how product kernels of different orders behave with respect to convolution.

Theorem 2.9.1.
Let $K_1 \in \mathcal{PK}(\boldsymbol{\nu})$ and $K_2 \in \mathcal{PK}(\boldsymbol{\mu})$ such that $\nu_k, \mu_k > -q_k$ and $\nu_k + \mu_k > -q_k$ for all $k \in \{1, \dots, d\}$ and let $\mathcal{T}_1$ and $\mathcal{T}_2$ be the convolution operators associated to K_1 and K_2 respectively. Then the operator $\mathcal{T}_2 \circ \mathcal{T}_1$ is well defined on $S(\mathbb{R}^N)$. The associated kernel is in $\mathcal{PK}(\boldsymbol{\nu} + \boldsymbol{\mu})$ and we indicate it as $K_2 * K_1$. Furthermore, let

$$K_1(\boldsymbol{x}) = \sum_{\boldsymbol{i}\in\mathbb{Z}^d} 2^{-\boldsymbol{i}\boldsymbol{q}}\varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}) \qquad K_2(\boldsymbol{x}) = \sum_{\boldsymbol{i}'\in\mathbb{Z}^d} 2^{-\boldsymbol{i}'\boldsymbol{q}}\psi_{\boldsymbol{i}'}^{(2^{\boldsymbol{i}'})}(\boldsymbol{x})$$

be two dyadic decompositions of K_1 and K_2 satisfying the hypotheses of Theorem 2.7.2. Then the kernel $K_1 * K_2$ is given as the weak- limit of the double dyadic series*

$$\sum_{\boldsymbol{i}\in\mathbb{Z}^d}\sum_{\boldsymbol{i}'\in\mathbb{Z}^d}\left(2^{-\boldsymbol{i}\boldsymbol{q}}\varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})} * 2^{-\boldsymbol{i}'\boldsymbol{q}}\psi_{\boldsymbol{i}'}^{(2^{\boldsymbol{i}'})}\right)(\boldsymbol{x}).$$

Thus the operation $* : \mathcal{PK}(\boldsymbol{\nu}) \times \mathcal{PK}(\boldsymbol{\mu}) \to \mathcal{PK}(\boldsymbol{\nu} + \boldsymbol{\mu})$ is a continuous bilinear map.

In general $\mathcal{T}_1$ and $\mathcal{T}_2$ are continuous operators from $S(\mathbb{R}^N)$ to $S'(\mathbb{R}^N)$ so there is no immediate way to define their composition. However suppose there exists a X such that $S(\mathbb{R}^N) \subset X \subset S'(\mathbb{R}^N)$ endowed with its own topology (X, τ) such that $S(\mathbb{R}^N)$ is dense in (X, τ). If $\mathcal{T}_1$ is a continuous operator from $S(\mathbb{R}^N)$ to (X, τ) and T_2 extends to a bounded operator from (X, τ) to $S'(\mathbb{R}^N)$ then $\mathcal{T}_2 \circ \mathcal{T}_1$ can be defined as the composition of the extension of $\mathcal{T}_2$ with $\mathcal{T}_1$. We will use this idea in the proof of Theorem 2.9.1.

Proof of Theorem 2.9.1. On $S(\mathbb{R}^N)$ let us introduce the norm

$$\|\varphi\|_X \stackrel{\text{def}}{=} \left\| \prod_{k=1}^{d} (\|\xi_k\|^{\mu_k} + 1)\, \widehat{\varphi}(\xi) \right\|_{L^1(\mathbb{R}^N)}$$

and let X be the completion of $S(\mathbb{R}^N)$ with respect to this norm. It is easy to see that $\mathcal{T}_1$ is a bounded operator from $S(\mathbb{R}^N)$ to X. As a matter of fact $\widehat{K}_1 \in L^1_{loc}$ and so if $\varphi \in S(\mathbb{R}^N)$ is such that $\widehat{\varphi} \in D(\mathbb{R}^N)$ we have that $\mathcal{F}(\mathcal{T}_1\varphi)$ is given by the function $\widehat{K}_1(\boldsymbol{\xi})\widehat{\varphi}(\boldsymbol{\xi})$ and

$$\left\|\mathcal{F}(\mathcal{T}_1\varphi)\right\|_X < \infty$$

because $\nu_k + \mu_k > -q_k$. On the other hand $\mathcal{T}_2$ extends to a bounded operator from X to $S'(\mathbb{R}^N)$. $\widehat{K}_2$ is locally integrable and the size inequalities (2.3.3) on $\widehat{K}_2$ guarantee that the inequality

$$\left\|\mathcal{F}(\mathcal{T}_2\varphi)\right\|_{L^1} < C\, \|\varphi\|_X$$

for all $\varphi \in S(\mathbb{R}^N)$. For these reasons we can define $\mathcal{T}_2 \circ \mathcal{T}_1$ as the composition of the extension of $\mathcal{T}_2$ with $\mathcal{T}_1$. Let (X', τ') be the completion of $S(\mathbb{R}^N)$ with respect to the norm

$$\|\varphi\|_{X'} \stackrel{\text{def}}{=} \|\widehat{\varphi}\|_{L^1}\,.$$

Convergence in X' implies convergence in $S'(\mathbb{R}^N)$.

The dyadic decompositions of K_1 and of K_2 converge poinwise a.e. and the Fourier transforms of the dyadic series converge poinwise a.e. Since the partial sums of the dyadic series of K_1 are uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu})$, by Lebesgue dominated convergence theorem, the operators $\mathcal{T}_{1,J}$ associated to the partial sums of the series

$$K_{1,J}(\boldsymbol{x}) \stackrel{\text{def}}{=} \sum_{\boldsymbol{i} \in J \subset \mathbb{Z}^d} 2^{-\boldsymbol{i}\boldsymbol{q}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x})$$

converge in the strong operator topology to $\mathcal{T}_1$ and for all finite $J \subset \mathbb{Z}^d$ the image of the operators $\mathcal{T}_{1,J}$ are in $S(\mathbb{R}^N)$ so the composition $\mathcal{T}_2 \circ \mathcal{T}_{1,J}$ is well defined. Furthermore the operators $\mathcal{T}_{2,J'}$ associated to the partial sums

$$K_{2,J'}(\boldsymbol{x}) \stackrel{\text{def}}{=} \sum_{\boldsymbol{i}' \subset J' \in \mathbb{Z}^d} 2^{-\boldsymbol{i}'\boldsymbol{q}} \psi_{\boldsymbol{i}'}^{(2^{\boldsymbol{i}'})}(\boldsymbol{x})$$

extend to uniformly bounded operators from (X, τ) to (X', τ') and the extensions of $\mathcal{T}_{2,J'}$ converge in the strong operator norm to $\mathcal{T}_2$. This means that $\mathcal{T}_{2,J'} \circ \mathcal{T}_{1,J}$ are associated with the kernels $K_{2,J'} * K_{1,J}$ and for $J, J' \uparrow \mathbb{Z}^d$ the operators $\mathcal{T}_{2,J'} \circ \mathcal{T}_{1,J}$ converge in the strong operator topology to $\mathcal{T}_2 \circ \mathcal{T}_1$. Notice that

$$K_{2,J'} * K_{1,J}(\boldsymbol{x}) = \sum_{\boldsymbol{i} \in J} \sum_{\boldsymbol{i}' \in J'} \left(2^{-\boldsymbol{i}\boldsymbol{q}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})} * 2^{-\boldsymbol{i}'\boldsymbol{q}} \psi_{\boldsymbol{i}'}^{(2^{\boldsymbol{i}'})} \right)(\boldsymbol{x}).$$

Furthermore the partial sums can be rewritten by setting

$$\widetilde{\varphi}_{\boldsymbol{i}''}(\boldsymbol{x}) \stackrel{\text{def}}{=} \sum_{\substack{(\boldsymbol{i},\boldsymbol{i}')\in J\times J' \\ \max(\boldsymbol{i},\boldsymbol{i}')=\boldsymbol{i}''}} \left(2^{-\boldsymbol{i}\boldsymbol{q}}\varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})} * 2^{-\boldsymbol{i}'\boldsymbol{q}}\psi_{\boldsymbol{i}'}^{(2^{\boldsymbol{i}'})}\right)(\boldsymbol{x})$$

where $\{\widetilde{\varphi}_{\boldsymbol{i}''}\}$ are a uniformly bounded family in $S(\mathbb{R}^N)$ with the required strong cancellation conditions.

The boundedness of the bilinear mapping $* : \mathcal{PK}(\boldsymbol{\nu}) \times \mathcal{PK}(\boldsymbol{\mu}) \to \mathcal{PK}(\boldsymbol{\nu}+\boldsymbol{\mu})$ is due to the fact that Theorem 2.7.3 allows writing dyadic decompositions with a small dyadic family of $S(\mathbb{R}^N)$ functions if the kernel itself is small. □

It is important to notice that the above Theorem actually guarantees that if K_1 and K_2 coincide with L^1_{loc} functions for which the convolution is defined and is an L^1_{loc} kernel then it is associated to the composition of the two operators. Using the Fourier transform we can also define the multiplication of product kernels for a certain range of orders.

Corollary 2.9.2 (Multiplication of product kernels)**.**
Let $\boldsymbol{\nu}$ and $\boldsymbol{\mu}$ be two multi-orders such that $\nu_k, \mu_k < 0$ and $\nu_k + \mu_k < -q$. Then the multiplication operator $(K_1, K_2) \mapsto K_1K_2$ is well defined and continuous from $\mathcal{PK}(\boldsymbol{\nu})\times\mathcal{PK}(\boldsymbol{\mu})$ to $\mathcal{PK}(\boldsymbol{\nu}+\boldsymbol{\mu}+\boldsymbol{q})$

2.9.1 Further functional calculus

Let us define the differential operators $\mathcal{L}_k$ on H_k given by

$$\mathcal{L}_k\varphi(x_k) \stackrel{\text{def}}{=} \mathcal{F}^{-1}\left(-\|\xi_k\|^2\,\widehat{\varphi}(\xi_k)\right)$$

for all $\varphi \in S(H_k)$. Using spectral theory we have for all $s \in \{z \in \mathbb{C} \mid \Re z > -q_k\}$ the operators

$$(-\mathcal{L})^s = \mathcal{F}^{-1}\left(\|\xi_k\|^{2s}\,\widehat{\varphi}(\xi_k)\right).$$

For $\boldsymbol{s} = (s_1, \dots, s_d) \in \mathbb{C}^d$ such that $\Re s_k > -q_k$ the product operators $(-\boldsymbol{\mathcal{L}})^{\boldsymbol{s}}$ are given by

$$(-\boldsymbol{\mathcal{L}})^{\boldsymbol{s}} = \mathcal{F}^{-1}\left(\prod_{k=1}^{d}\|\xi_k\|^{2s_k}\,\widehat{\varphi}(\boldsymbol{\xi})\right)$$

for any $\varphi \in S'(\mathbb{R}^N)$.

We have the following theorem about the relationship between product kernels of different orders and differential operators $(-\boldsymbol{\mathcal{L}})^{\boldsymbol{s}}$.

Corollary 2.9.3 (The action of $(-\boldsymbol{\mathcal{L}})^{\boldsymbol{s}}$)**.**
Let $\boldsymbol{\nu}$ be a multi-order such that $\nu_k > -q_k$ and $\boldsymbol{s} \in \mathbb{C}^d$ such that $2\Re s_k > -q_k$ and $\nu_k + 2\Re s_k > -q_k$. Then the map $K \mapsto (-\boldsymbol{\mathcal{L}})^{\boldsymbol{s}}K$ is continuous from $\mathcal{PK}(\boldsymbol{\nu})$ to $\mathcal{PK}(\boldsymbol{\nu} + 2\Re\boldsymbol{s})$.

Proof. If $\mathcal{T}$ is the convolution operator associated with K then $(-\boldsymbol{\mathcal{L}})^{\boldsymbol{s}}K$ is the kernel associated with the operator $\mathcal{T}\circ(-\boldsymbol{\mathcal{L}})^{\boldsymbol{s}}$. $(-\boldsymbol{\mathcal{L}})^{\boldsymbol{s}}$ acts by multiplication on the Fourier side by $\prod_{k=1}^{d}(\|\xi_k\|)^{2s_k}$ and since $2\Re s_k > -q_k$ we have that the function $\prod_{k=1}^{d}(\|\xi_k\|)^{2s_k}$ is locally integrable so it is in $\mathcal{PK}(-\boldsymbol{q} - 2\Re\boldsymbol{s})$ with $-q_k - 2\Re s_k < 0$. This means that $(-\boldsymbol{\mathcal{L}})^{\boldsymbol{s}}$ is a convolution operator relative to a kernel $K_{(-\boldsymbol{\mathcal{L}})^{\boldsymbol{s}}} \in \mathcal{PK}(2\Re\boldsymbol{s})$. By Theorem 2.9.1 we have that

$$(-\boldsymbol{\mathcal{L}})^{\boldsymbol{s}}K = K * K_{(-\boldsymbol{\mathcal{L}})^{\boldsymbol{s}}} \in \mathcal{PK}(\boldsymbol{\nu} + 2\Re\boldsymbol{s})$$

□

As a result we have that for multi-orders $\boldsymbol{\nu}$ such that $\nu_k > -q_k$ for all $k \in \{1, \dots, d\}$. $K \in \mathcal{PK}(\boldsymbol{\nu})$ if and only if there exists a kernel $\widetilde{K} \in \mathcal{PK}(\boldsymbol{0})$ such that $K = (-\boldsymbol{\mathcal{L}})^{\boldsymbol{\nu}/2}\widetilde{K}$.

2.10 Operators associated with product kernels

We now pass to considering the convolution operators associated with product kernels. The theory developed earlier is particularly useful to prove boundedness results on appropriate functional spaces.

2.10.1 Littlewood-Paley theory

A technique to prove L^p boundedness that we will use is the Littlewood-Paley estimates and square function estimates. This approach is based on the idea of studying series of "quasi-orthogonal" operators in L^p. We develop this technique directly in the product setting. The theory of one parameter square function estimates can be found in [Gra08].

Definition 2.10.1 (Rademacher square functions).
Suppose we are working on the tensor product of unit intervals $I^d = [0;1]^d$. Let $\boldsymbol{i} \in \mathbb{N}^d$ then the Rademacher function is

$$r_{\boldsymbol{i}}(\boldsymbol{t}) \stackrel{def}{=} \prod_{k=1}^{d} (-1)^{\lfloor 2^{i_k} t_k \rfloor}$$

where $\lfloor t \rfloor$ is the greatest integer less or equal to t.

Rademacher functions have some very important properties.

Proposition 2.10.2.
The Rademacher functions $r_{\boldsymbol{i}}(\boldsymbol{t})$ as $\boldsymbol{i} \in \mathbb{N}^d$ are an orthonormal non-complete family in $L^2(I^d)$.

Proof. Orthonormality can be checked by explicit calculation. Notice that choosing $\boldsymbol{i} \neq \boldsymbol{i}' \in \mathbb{N}^d$ we have that $r_{\boldsymbol{i}}(\boldsymbol{t}) r_{\boldsymbol{i}'}(\boldsymbol{t})$ is orthogonal to all Rademacher functions. □

Furthermore we have the following crucial equivalence of norms known as Khintchin's Theorem.

Theorem 2.10.3 (Khintchin's Theorem).
On the L^2 subspace generated by the Rademacher functions all L^p norms for $p \in [1, +\infty)$ are equivalent.

Proof. First suppose that $2 < p < +\infty$. Since we are working on a space of total measure 1, Hölder's inequality gives us that

$$\|f\|_{L^2} < \|f\|_{L^p}.$$

Let us prove the inequality in the other sense. Suppose that $p = 2k$ for some $m \in \mathbb{N}$, $k > 1$. For any

$$f(\boldsymbol{t}) = \sum_{\substack{\boldsymbol{i} \in \mathbb{N}^d \\ i_k \le l}} a_{\boldsymbol{i}} r_{\boldsymbol{i}}(\boldsymbol{t})$$

we have that

$$\|f\|_{L^{2m}}^{2m} = \sum_{\substack{\boldsymbol{i}^1 \in \mathbb{N}^d \\ \vdots \\ \boldsymbol{i}^{2m} \in \mathbb{N}^d}} \int_{I^d} a_{\boldsymbol{i}^1} \dots a_{\boldsymbol{i}^{2m}} r_{\boldsymbol{i}^1}(\boldsymbol{t}) \dots r_{\boldsymbol{i}^{2m}}(\boldsymbol{t}) \mathrm{d}\boldsymbol{t} \le C_m \sum_{\substack{\boldsymbol{i}^1 \in \mathbb{N}^d \\ \vdots \\ \boldsymbol{i}^{m} \in \mathbb{N}^d}} a_{\boldsymbol{i}^1}^2 \dots a_{\boldsymbol{i}^m}^2 = \|f\|_{L^2}^{2m}.$$

The theorem for $1 < p < 2$ follows by convex interpolation of norms. As a matter of fact

$$\|f\|_{L^p} < \|f\|_{L^2}$$

and

$$\|f\|_{L^2}^2 \leq \|f\|_{L^p} \|f\|_{L^{p'}}$$

But since the $p' = {}^p\!/_{p-1}$ norm is equivalent to the L^2 norm this gives us what we need. Finally for $p = 1$ one can do use the interpolation of norms with respect to some norm L^q with $1 < q < 2$. □

As a consequence of these two properties we have the following lemmas that characterize "quasi-orthogonal" operators on L^p.

Lemma 2.10.4 (First boundedness criterion for series of quasi-orthogonal operators)**.** *Let $\{\mathcal{T}_{\boldsymbol{i}}\}_{\boldsymbol{i}}$ with $\boldsymbol{i} \in \mathbb{N}^d$ be a family of bounded operators on L^p such that for any choice of a multi-sequence*

$$\boldsymbol{\varepsilon_i} = \prod_{k=1}^{d} \varepsilon_{k,i_k},$$

where ε_k are some sequences with values in $\{+1, -1\}$, any finite sum of operators satisfies

$$\left\| \sum_{\boldsymbol{i} \in J \subset \mathbb{N}^d} \boldsymbol{\varepsilon_i} \mathcal{T}_{\boldsymbol{i}} \right\|_{L^p - L^p} < A$$

for some constant A independent of J and of $\boldsymbol{\varepsilon}$. Then the operator

$$\boldsymbol{\mathcal{T}} f \stackrel{def}{=} \{T_{\boldsymbol{i}} f\}_{\boldsymbol{i} \in \mathbb{N}^d} \tag{2.10.1}$$

is bounded from L^p to $L^p(l^2(\mathbb{N}^d))$. In other words we have the following inequality

$$\left\| \left(\sum_{\boldsymbol{i} \in \mathbb{N}^d} |\mathcal{T}_{\boldsymbol{i}} f|^2 \right)^{1/2} \right\|_{L^p} < C_p A \|f\|_{L^p} \tag{2.10.2}$$

Proof. We prove inequality (2.10.2) for finite partial sums. Since all the terms are positive, the boundedness for the series follows from monotone convergence. Let $J \subset \mathbb{N}^d$ be a finite subset. Let us write the expression

$$\sum_{\boldsymbol{i} \in J} r_{\boldsymbol{i}}(\boldsymbol{t}) \mathcal{T}_{\boldsymbol{i}} f(x).$$

Taking the L^p norm in the x variable and then integrating in t we have

$$\int_{I^d} \left\| \sum_{\boldsymbol{i} \in J} r_{\boldsymbol{i}}(\boldsymbol{t}) \mathcal{T}_{\boldsymbol{i}} f(x) \right\|_{L^p}^p \mathrm{d}\boldsymbol{t} \leq A^p \|f\|_{L^p}^p .$$

Exchanging the order of integration and using the equivalence between the L^2 and L^p norms in $\boldsymbol{t}$ given by Theorem 2.10.3 we get

$$\int_{I^d} \left\| \sum_{\boldsymbol{i} \in J} r_{\boldsymbol{i}}(\boldsymbol{t}) \mathcal{T}_{\boldsymbol{i}} f(x) \right\|_{L^p}^p \mathrm{d}\boldsymbol{t} = \int \int_{I^d} \left| \sum_{\boldsymbol{i} \in J} r_{\boldsymbol{i}}(\boldsymbol{t}) \mathcal{T}_{\boldsymbol{i}} f(x) \right|^p \mathrm{d}\boldsymbol{t} \mathrm{d}x <$$

$$C_p \int \left(\sum_{\boldsymbol{i} \in J} |\mathcal{T}_{\boldsymbol{i}} f(x)|^2 \right)^{p/2} \mathrm{d}x = \left\| \left(\sum_{\boldsymbol{i} \in \mathbb{N}^d} |\mathcal{T}_{\boldsymbol{i}} f|^2 \right)^{1/2} \right\|_{L^p}$$

and this gives us inequality (2.10.2) for all finite partial sums and for the whole series by a limiting procedure. Furthermore using inequality (2.10.2) all truncated operators

$$\boldsymbol{\mathcal{T}}_J f = \{\mathcal{T}_{\boldsymbol{i}} f\}_{\boldsymbol{i}\in J\subset\mathbb{N}^d}$$

are uniformly bounded from L^p to $L^p(l^2(\mathbb{N}^d))$ and converge in the strong operator topology to $\boldsymbol{\mathcal{T}}$. □

A consequence of the above lemma is a boundedness results for the adjoint operators. It is easy to see that for $1 < p < \infty$ the dual space of $L^p(l^2(\mathbb{N}^d))$ is the space $L^{p'}(l^2(\mathbb{N}^d))$.

Lemma 2.10.5 (Second boundedness criterion for series of quasi-orthonormal operators)**.**
Let $1 < p < \infty$ and let $\{\mathcal{T}_{\boldsymbol{i}}\}$ be a family of operators with the same hypothesis as in Lemma 2.10.4. Then the operator

$$\widetilde{\boldsymbol{\mathcal{T}}}\{f_{\boldsymbol{i}}\} \overset{def}{=} \sum_{\boldsymbol{i}\in\mathbb{N}^d} \mathcal{T}_{\boldsymbol{i}} f_{\boldsymbol{i}}$$

is bounded from $L^p(l^2(\mathbb{N}^d))$ to L^p.

Proof. Since $\Big(L^p(l^2(\mathbb{N}^d))\Big)^* = L^{p'}(l^2(\mathbb{N}^d))$ this statement follows by duality to Lemma 2.10.4. As a matter of fact the family $\{\mathcal{T}_{\boldsymbol{i}}^*\}$ satisfies the hypothesis of Lemma 2.10.4 for p'. Let

$$\boldsymbol{\mathcal{T}} f \overset{\text{def}}{=} \{T_{\boldsymbol{i}}^* f\}_{\boldsymbol{i}\in\mathbb{N}^d}.$$

It is sufficient to notice that $\widetilde{\boldsymbol{\mathcal{T}}}\{f_{\boldsymbol{i}}\} = \boldsymbol{\mathcal{T}}^*$ and since by Lemma 2.10.4 $\boldsymbol{\mathcal{T}}$ is bounded from $L^{p'}$ to $L^{p'}(l^2(\mathbb{N}^d))$ we have the result. □

Notice that the exact same statements hold true if the indexes $\boldsymbol{i}$ are in $\mathbb{Z}^d$ rather than $\mathbb{N}^d$.

Definition 2.10.6 (Littlewood-Paley functions)**.**
Consider a product space $\mathbb{R}^N = \bigoplus_{k=1}^d H_k$ and for each k let $\varphi_k(x_k) \overset{def}{=} \mathcal{F}^{-1}\eta_k(\xi_k)$ where η_k are cutoff functions on H_k. Let $\psi_k(x_k) \overset{def}{=} 2^{q_k}\varphi_k(2\cdot x_k) - \varphi_k(x_k)$ and $\boldsymbol{\psi}(\boldsymbol{x}) = \bigotimes_{k=1}^d \psi_k(x_k)$. We say that the family $\{\psi_{\boldsymbol{i}}\}_{\boldsymbol{i}\in\mathbb{Z}^d}$ defined by the relation

$$\psi_{\boldsymbol{i}} \overset{def}{=} \boldsymbol{\psi}^{(2^{\boldsymbol{i}})} = 2^{-\boldsymbol{i}\boldsymbol{q}}\boldsymbol{\psi}(2^{-\boldsymbol{i}}\cdot\boldsymbol{x})$$

is a family of Littlewood-Paley functions.

We have the following lemma for L^p functions.

Lemma 2.10.7 (Strong convergence in L^p)**.**
Let $f \in L^p(\mathbb{R}^N; X)$ then

$$\lim_{M\to\infty} \sum_{\substack{\boldsymbol{i}\in\mathbb{Z}^d \\ |i_k|<M}} \boldsymbol{\psi}_{\boldsymbol{i}} * f \to f$$

in L^p norm.

Proof. Notice that

$$\sum_{|i_k|<M} 2^{-q_k}\psi_k(2^{-i}\cdot x_k) = 2^{Mq_k}\varphi_k(2^M\cdot x_k) - 2^{-(M-1)q_k}\varphi_k(2^{-(M-1)}\cdot x_k)$$

so

$$\sum_{\substack{\boldsymbol{i}\in\mathbb{Z}^d\\ |i_k|<M}} \boldsymbol{\psi_i} = -1^{(d-|J|)} \sum_{J\subset\{1,\ldots,d\}} \prod_{j\in J} \varphi_j^{(2^{-M})}(x_j) \prod_{k\notin J} \varphi_k^{(2^{M-1})}(x_j)$$

We have that

$$\int_{H_k} \varphi_k(x_k)\mathrm{d}x_k = 1$$

so for $J = \{1, \ldots, d\}$ the term

$$\prod_{j\in J} \varphi_j^{(2^{-M})}(x_j) \prod_{k\notin J} \varphi_k^{(2^{M-1})}(x_j) = \prod_{j\in\{1,\ldots,d\}} \varphi_j^{(2^{-M})}(x_j)$$

is an approximate identity as $M \to \infty$. On the other hand, if $J \neq \{1, \ldots, d\}$ then the term

$$\prod_{j\in J} \varphi_j^{(2^{-M})}(x_j) \prod_{k\notin J} \varphi_k^{(2^{M-1})}(x_j)$$

tends to 0 in L^p as $M \to \infty$. Since all the partial sums are bounded on L^p by 2^d we have that the statement hold for all $f \in L^p \cap L^1$. The statement follows by density for all $f \in L^p$. □

We can now proceed to the fundamental result of norm equivalence of the Littlewood-Paley theory.

Theorem 2.10.8 (Littlewood-Paley decomposition)**.**
Let $\{\boldsymbol{\psi_i}\}$ be a family of Littlewood-Paley functions on the product space $\mathbb{R}^N = \bigoplus_{k=1}^d H_k$. Then we have the equivalence of norms

$$\left\| \{f * \boldsymbol{\psi_i}\}_{\boldsymbol{i}\in\mathbb{Z}^d} \right\|_{L^p(l^2(\mathbb{Z}^d))} \approx \|f\|_{L^p}.$$

Proof. The family of convolution operators

$$\mathcal{T}_{\boldsymbol{i}} f = f * \boldsymbol{\psi_i}$$

associated to the Littlewood-Paley functions $\{\boldsymbol{\psi_i}\}$ satisfy the conditions of Lemma 2.10.4 and thus of Lemma 2.10.5. As a matter of fact the sum

$$\sum_{\boldsymbol{i}\in J\subset\mathbb{N}^d} \boldsymbol{\varepsilon_i} \mathcal{T}_{\boldsymbol{i}}$$

for any sequence

$$\boldsymbol{\varepsilon_i} = \prod_{k=1}^d \varepsilon_{k,i_k}$$

is a tensor product of $\mathcal{CZ}$ kernels. Tensor products of convolution operators coincide with the composition of the operators acting along the single subspaces. Thanks to Theorem 1.2.13 each Calderón-Zygmund operator is bounded on L^p and using elementary Bochner integration the tensor product is also bounded on L^p. The uniformity is given by the fact that the bounds on the dyadic elements do not depend on the choice of the signs $\boldsymbol{\varepsilon}$.

Using Lemma 2.10.4 we have

$$\left\| \{f * \boldsymbol{\psi_i}\}_{\boldsymbol{i}\in\mathbb{Z}^d} \right\|_{L^p(l^2(\mathbb{Z}^d))} \leq C \|f\|_{L^p}.$$

To see the converse notice that the family $\{\psi_{\boldsymbol{i}} * \psi_{\boldsymbol{i}}\}$ also satisfies the conditions of Lemma 2.10.4 and of Lemma 2.10.5. As a matter of fact $\psi_{\boldsymbol{i}} * \psi_{\boldsymbol{i}}$ are the functions $\psi * \psi$ rescaled. Using Lemma 2.10.5 and Lemma 2.10.7 we have that

$$\left\| \sum_{\boldsymbol{i}\in\mathbb{Z}^d} \boldsymbol{\psi_i} * \boldsymbol{\psi_i} * f \right\|_{L^p} \leq C \left\| \{f * \boldsymbol{\psi_i}\}_{\boldsymbol{i}\in\mathbb{Z}^d} \right\|_{L^p(l^2(\mathbb{Z}^d))}$$

but

$$\left\| \sum_{\boldsymbol{i}\in\mathbb{Z}^d} \boldsymbol{\psi_i} * \boldsymbol{\psi_i} * f \right\|_{L^p} \approx \|f\|_{L^p}$$

because taking the Fourier transform both $\left(\sum_{\boldsymbol{i}\in\mathbb{Z}^d} \widehat{\boldsymbol{\psi}}_{\boldsymbol{i}}^2\right)^{-1}$ and $\sum_{\boldsymbol{i}\in\mathbb{Z}^d} \widehat{\boldsymbol{\psi}}_{\boldsymbol{i}}^2$ converge to Mihlin-Hörmander multipliers. □

2.10.2 Boundedness of product kernel operators

We begin this section with results on boundedness on Sobolev spaces.

Theorem 2.10.9 (Boundedness on Sobolev spaces)**.**
Let $K \in \mathcal{PK}(\boldsymbol{\nu})$ and suppose that for those j for which $\nu_j < 0$ the support of K is bounded along H_j. Then the convolution operator $\mathcal{T}$ associated to K extends to a bounded operator between the (anisotropic) product Sobolev spaces

$$\mathcal{T} : \mathcal{H}^{\boldsymbol{s}} \to \mathcal{H}^{\boldsymbol{s}-\boldsymbol{\nu}}.$$

The operator norm of $\mathcal{T}$ depends on the bounds on K and in particular for any fixed choice $M_j \in \mathbb{R}^+$ for those indexes $j \in \{1, \dots, d\}$ for which $\nu_j < 0$ and for any constant C there is a neighborhood V_C of 0 in $\mathcal{PK}(\boldsymbol{\nu})$ such that if $K \in V_C$ and $\operatorname{spt} K \subset \left\{\boldsymbol{x} \,\middle|\, \|x_j\| < M_j\right\}$ then the associated convolution operator $\mathcal{T}$ has operator norm bounded by C.

Proof. Since K has bounded support along the spaces H_j with $\nu_j < 0$ we can set $\widetilde{\boldsymbol{\nu}} = \boldsymbol{\nu} \vee \boldsymbol{0}$ and Proposition 2.8.4 guarantees that $K \in \mathcal{PK}(\widetilde{\boldsymbol{\nu}})$. Using theorem 2.5.2 we can see that the multiplier m of $\mathcal{T}$ is a locally bounded function that satisfies

$$|m(\boldsymbol{\xi})| < C_K \prod_{k=1}^{d} \|\xi_k\|^{\widetilde{\nu}_k}.$$

However, since $K \in \mathcal{PK}(\boldsymbol{\nu})$, m away from the origin satisfies the size conditions on $\mathcal{PK}(-\boldsymbol{q}-\boldsymbol{\nu})$ so we have that

$$|m(\boldsymbol{\xi})| < C'_K \prod_{k=1}^{d} (1 + \|\xi_k\|)^{\nu_k}.$$

Recalling the definition of $\mathcal{H}^{\boldsymbol{s}}$ in terms of multipliers gives us the desired result. □

We now pass to proving boundedness on L^p for $1 < p < \infty$. To do this we will use the multi-parameter Littlewood-Paley theory illustrated in Section 2.10.1. Before proving the positive result we show via a counterexample why the approach used for proving boundedness for $\mathcal{CZ}$ operators does not work. In particular product kernels are not $L^1 - L^1_w$ bounded.

Example 2.10.10.
Consider $\mathbb{R}^2$ *seen as the product space* $\mathbb{R}\times\mathbb{R}$ *and let* $\varphi_\varepsilon(\boldsymbol{x})=\varepsilon^{-2}\mathbb{1}_{[-\varepsilon,\varepsilon]^2}(\boldsymbol{x})$*. The double Hilbert transform* $\mathcal{H}_1\otimes\mathcal{H}_2$ *is not bounded from* $L^1(\mathbb{R}^2)$ *to* $L^1_w(\mathbb{R}^2)$*. In particular*

$$\lim_{\varepsilon\to 0}\mathscr{L}^2\left(\left\{\boldsymbol{x}\,\middle|\,|\mathcal{H}_1\otimes\mathcal{H}_2\varphi_\varepsilon(\boldsymbol{x})|>1\right\}\right)=\infty.$$

As a matter of fact

$$\mathcal{H}_1\otimes\mathcal{H}_2\varphi_\varepsilon=\varepsilon^{-2}\mathcal{H}\mathbb{1}_{[-\varepsilon,\varepsilon]}\otimes\mathcal{H}\mathbb{1}_{[-\varepsilon,\varepsilon]}.$$

A simple calculation yields

$$\mathcal{H}\mathbb{1}_{[-\varepsilon,\varepsilon]}(x)=\frac{1}{\pi}\log\left|\frac{x+\varepsilon}{x-\varepsilon}\right|$$

so for $x>2\varepsilon$

$$\frac{1}{\pi}\log\left|\frac{x+\varepsilon}{x-\varepsilon}\right|>\frac{\varepsilon}{\pi x}$$

and thus

$$\left\{\boldsymbol{x}\,\middle|\,|\mathcal{H}_1\otimes\mathcal{H}_2\varphi_\varepsilon(\boldsymbol{x})|>1\right\}\supset\left\{(x_1,x_2)\,\middle|\,x_1x_2<\pi^{-2}\quad x_1,x_2>2\varepsilon\right\}.$$

For $\varepsilon\to 0$ *the right hand side has unbounded measure.*

Theorem 2.10.11 (L^p boundedness of product kernel convolution operators)**.**
Let $K\in\mathcal{PK}(\mathbf{0})$ *then the associated convolution operator* $\mathcal{T}$ *extends to a bounded operator on* $L^p(\mathbb{R}^N)$*.*

Proof. We need to prove that for any $f\in D(\mathbb{R}^N)$ we have

$$\|f*K\|_{L^p}\le C_p\,\|f\|_{L^p}\,.$$

Let $\{\boldsymbol{\psi_i}\}$ be Littlewood-Paley functions. We have seen in the proof of Theorem 2.10.8 that the convolution operators associated with the functions $\{\psi_{\boldsymbol{i}}*\psi_{\boldsymbol{i}}\}$ satisfy the hypotheses of Lemma 2.10.4 and of Lemma 2.10.5. Using the second lemma we have

$$\|f*K\|_{L^p}\le\left\|\{f*K*\psi_{\boldsymbol{i}}*\psi_{\boldsymbol{i}}\}\right\|_{L^p(l^2(\mathbb{Z}^d))}$$

Setting

$$\varphi_{\boldsymbol{i}}\stackrel{\text{def}}{=}(K*\psi_{\boldsymbol{i}})^{(2^{\boldsymbol{i}})}=\mathcal{F}^{-1}\left(\widehat{K}(2^{\boldsymbol{i}}\cdot\boldsymbol{\xi})\prod_{k=1}^{d}\left(\eta_k(2^{\cdot}\xi_k)-\eta_k(\xi_k)\right)\right)$$

we have that $\{\varphi_{\boldsymbol{i}}\}$ is a uniformly bounded family of functions in $S(\mathbb{R}^N)$. We also indicate

$$f_{\boldsymbol{i}}\stackrel{\text{def}}{=}f*\psi_{\boldsymbol{i}}.$$

Suppose for now that $p>2$. We expand

$$\left\|\{f*K*\psi_{\boldsymbol{i}}*\psi_{\boldsymbol{i}}\}\right\|_{L^p(l^2(\mathbb{Z}^d))}=\left\|\sum_{\boldsymbol{i}\in Z^d}\left|f_{\boldsymbol{i}}*\varphi_{\boldsymbol{i}}^{(2^{-\boldsymbol{i}})}\right|^2\right\|_{L^{p/2}}^{1/2}$$

but

$$\left|f_{\boldsymbol{i}}*\varphi_{\boldsymbol{i}}^{(2^{-\boldsymbol{i}})}\right|^2\le\int|\varphi_{\boldsymbol{i}}^{(2^{-\boldsymbol{i}})}(x-y)|\mathrm{d}y\int|\varphi_{\boldsymbol{i}}^{(2^{-\boldsymbol{i}})}(x-y)||f_{\boldsymbol{i}}(y)|^2\mathrm{d}y\le C_K\int\left|\varphi_{\boldsymbol{i}}^{(2^{-\boldsymbol{i}})}(x-y)\right|\left|f_{\boldsymbol{i}}(y)\right|^2\mathrm{d}y.$$

We can write

$$\begin{aligned}\left\|\sum_{\boldsymbol{i}\in Z^d}\left|f_{\boldsymbol{i}}*\varphi_{\boldsymbol{i}}^{(2^{-\boldsymbol{i}})}\right|^2\right\|_{L^{p/2}} \leq & C_K \sup_{\substack{g\in L^{(p/2)'}\\ \|g\|_{L^{(p/2)'}}<1}} \int_{\mathbb{R}^N}\sum_{\boldsymbol{i}\in Z^d}\left(|f_{\boldsymbol{i}}|^2*\left|\varphi_{\boldsymbol{i}}^{(2^{-\boldsymbol{i}})}\right|(\boldsymbol{x})\right)g(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \\ & C_K \sup_{\substack{g\in L^{(p/2)'}\\ \|g\|_{L^{(p/2)'}}<1}} \sum_{\boldsymbol{i}\in Z^d}\int_{\mathbb{R}^N}|f_{\boldsymbol{i}}|^2(\boldsymbol{x})\left(\left|\check{\varphi}_{\boldsymbol{i}}^{(2^{-\boldsymbol{i}})}\right|*g(\boldsymbol{x})\right)\mathrm{d}\boldsymbol{x} \leq \\ & C_K \sup_{\substack{g\in L^{(p/2)'}\\ \|g\|_{L^{(p/2)'}}<1}} \int_{\mathbb{R}^N}\sum_{\boldsymbol{i}\in \mathbb{Z}^d}|f_{\boldsymbol{i}}|^2(\boldsymbol{x})\sup_{\boldsymbol{i}\in\mathbb{Z}^d}\left(\left|\check{\varphi}_{\boldsymbol{i}}^{(2^{-\boldsymbol{i}})}\right|*g(\boldsymbol{x})\right)\mathrm{d}\boldsymbol{x} \leq \\ & \left\|\sum_{\boldsymbol{i}\in\mathbb{Z}^d}|f_{\boldsymbol{i}}|^2\right\|_{L^{p/2}}\left\|\sup_{\boldsymbol{i}\in\mathbb{Z}^d}\left(\left|\check{\varphi}_{\boldsymbol{i}}^{(2^{-\boldsymbol{i}})}\right|*|g|\right)\right\|_{L^{(p/2)'}}\end{aligned}$$

The maximal type operator

$$g\mapsto \sup_{\boldsymbol{i}\in\mathbb{Z}^d}\left(\left|\check{\varphi}_{\boldsymbol{i}}^{(2^{-\boldsymbol{i}})}\right|*|g|(\boldsymbol{x})\right)$$

is bounded on $L^{(p/2)'}$. This is true because we can bound $\varphi_{\boldsymbol{i}}$ from above by

$$\left|\varphi_{\boldsymbol{i}}(\boldsymbol{x})\right|\leq\omega(\boldsymbol{x})\stackrel{\text{def}}{=}C'_k\prod_{k=1}^{d}\prod_{j=1}^{n_k}(1+|x_{k,j}|)^{-2}$$

where $x_{k,j}$ are the eigenvector coordinates with respect to the dilations on H_k. Let us introduce a finer multi-parameter structure on $\mathbb{R}^N$. Let each eigenvector $x_{k,j}$ be considered a product factor so we have

$$\sup_{\boldsymbol{i}\in\mathbb{Z}^d}\left(\left|\check{\varphi}_{\boldsymbol{i}}^{(2^{-\boldsymbol{i}})}\right|*|g|(\boldsymbol{x})\right) < \sup_{\boldsymbol{i}'\in\mathbb{Z}^N}\left(\left|\omega^{(2^{-\boldsymbol{i}})}\right|*|g|(\boldsymbol{x})\right).$$

The operator

$$g\mapsto \sup_{\boldsymbol{i}'\in\mathbb{Z}^N}\left(\left|\omega^{(2^{-\boldsymbol{i}})}\right|*|g|(\boldsymbol{x})\right)$$

is the composition of maximal type operators along all coordinates. It is sufficient to prove boundedness on $L^{p/2'}$ for the one-dimensional operator

$$g(t)\mapsto\sup_{i'\in Z}|g|*\left(2^{\lambda_{k,j}}(1+|2^{\lambda_{k,j}}x_{k,j}|)^{-2}\right)$$

But this holds because it can be controlled from above by the Hardy-Littlewood maximal operator.

As a consequence we have that

$$\|f*K\|_{L^p}\approx\left\|\{f*K*\psi_{\boldsymbol{i}}*\psi_{\boldsymbol{i}}\}\right\|_{L^p(l^2(\mathbb{Z}^d))}\leq C''_K\left\|\{f*\psi_{\boldsymbol{i}}\}\right\|_{L^p(l^2(\mathbb{Z}^d))}\approx\|f\|_{L^p}$$

for $p>2$ as required. The result for $1<p\leq 2$ follows by duality and interpolation. □

2.11 Changes of variables and diffeomorphisms

We will now study the effect of changes of variable on product kernels. Suppose that $\Phi : \Omega \to \mathbb{R}^N$ is a diffeomorphism of a certain open domain Ω onto its image in $\mathbb{R}^N$. Given some distribution $T \in D'(\mathbb{R}^N)$ we can define the distribution $T \circ \Phi$ in $D'(\Omega)$ by setting

$$\int_\Omega T \circ \Phi(x)\varphi(x)\mathrm{d}x \overset{\text{def}}{=} \int_{R^N} T(x') \left(\det D\Phi^{-1}\right)(x')\, \varphi\left(\Phi^{-1}(x')\right) \mathrm{d}x'$$

for all test functions $\varphi \in D(\Omega)$. If Φ is a sufficiently well behaved diffeomorphism then if $T \in S'(\mathbb{R}^N)$ then $T \circ \Phi \in S'(\mathbb{R}^N)$ with $\operatorname{spt} T \subset \Omega$. This is true for all diffeomorphisms if Ω is compact or if T has a fixed compact support contained in $\Phi(\Omega)$. In other cases one needs to impose additional requirements on the diffeomorphism. We will now deal with the stability of product kernels with respect to changes of variables. A geometrically relevant fact that we must ask of the diffeomorphisms we are dealing with is that they conserve the singular subspaces, and that the same hold for the inverse diffeomorphism Φ^{-1}.

Definition 2.11.1 (Product diffeomorphism)**.**
Let $\Phi : \Omega \to \Phi(\Omega) \subset \mathbb{R}^N$ be a diffeomorphism of the open domain $\Omega \subset \mathbb{R}^N$ with its image. We say that Φ is a product diffeomorphism if

$$\Phi(\Omega \cap H_k^\perp) \subset H_k^\perp \qquad \forall k \in \{1, \ldots, d\}$$

and

$$\Phi^{-1}(\Phi(\Omega) \cap H_k^\perp) \subset H_k^\perp \qquad \forall k \in \{1, \ldots, d\}.$$

We say that Φ is a product diffeomorphism of a compact domain if it is defined on a compact set Ω and it extends to a product diffeomorphism of an open neighborhood of Ω.

The product diffeomorphisms of a fixed domain Ω onto itself form a group under composition.

We now concentrate on some properties of product diffeomorphisms that are useful in relation to product kernels. However, this study can be carried out only on Euclidean spaces with a standard system of dilations.

Lemma 2.11.2 (Differential inequalities for diffeomorphisms)**.**
Let Φ be a product diffeomorphism of a compact domain and let Φ^k be the coordinate along H_k of Φ for any given $k \in \{1, \ldots, d\}$. We have the following differential inequalities.

$$\begin{aligned} C_k^{-1}\,|x_k| < \left|\Phi^k(\boldsymbol{x})\right| < C_k\,|x_k| \\ \left|\partial_{\boldsymbol{x}}^{\boldsymbol{\alpha}} \Phi^k\right|(\boldsymbol{x}) < C_{\boldsymbol{\alpha},k}(1 + |x_k|)^{1-|\alpha_k|} \end{aligned} \tag{2.11.1}$$

Based on the previous inequalities we introduce the following terminology.

Definition 2.11.3 (Uniformly bounded product diffeomorphisms)**.**
Given a domain $\Omega \supset \bigotimes \overline{B_k(0, 2^m)}$ let us consider product diffeomorphisms of Ω into $\mathbb{R}^N$. We say that a family of such diffeomorphisms is uniformly bounded if there are constants C_k and $C_{\boldsymbol{\alpha},k}$ such that the inequalities (2.11.1) *hold uniformly for the whole family.*

Definition 2.11.4 (Stretching of product diffeomorphisms)**.**
Given a domain $\Omega \supset \bigotimes \overline{B_k(0, 2^m)}$ let us consider a product diffeomorphisms Φ of Ω into $\mathbb{R}^N$.

For $G \subset \{1,\dots,d\}$ we call the stretched diffeomorphism along the spaces H_G with parameter s the diffeomorphism $\Phi_{s,G}$ defined for $s>0$ by the relation

$$\Phi^k_{s,G} \overset{def}{=} \begin{cases} s^{-1}\cdot \Phi^k\left(\boldsymbol{x}_{(G)} \oplus s\cdot \boldsymbol{x}_G\right) & \text{if } k\in G \\ \Phi^k\left(\boldsymbol{x}_{(G)} \oplus s\cdot \boldsymbol{x}_G\right) & \text{if } k\notin G \end{cases}$$

and for $s=0$ by the relation

$$\Phi^k_{0,G} \overset{def}{=} \begin{cases} \partial_{\boldsymbol{x}_G}\Phi^k\left(\boldsymbol{x}_{(G)} \oplus \mathbf{0}_G\right)\boldsymbol{x}_G & \text{if } k\in G \\ \Phi^k\left(\boldsymbol{x}_{(G)} \oplus \mathbf{0}_G\right) & \text{if } k\notin G \end{cases}$$

on the domain where the right-hand side is defined.

It is easy to check by direct calculation that for any given $G \subset \{1,\dots,d\}$ the mapping $(s,x) \mapsto \Phi_{s,G}(\boldsymbol{x})$ with $s \geq 0$ is smooth on the domain of definition. It is also useful to notice that for $k\in G$ we have the equality

$$\partial_{\boldsymbol{x}_G}\Phi^k\left(\boldsymbol{x}_{(G)} \oplus \mathbf{0}_G\right)\boldsymbol{x}_G = \partial_{x_k}\Phi^k\left(\boldsymbol{x}_{(G)} \oplus \mathbf{0}_G\right)x_k$$

because Φ_0, as also Φ, conserves the singular subspaces.

Lemma 2.11.5 (Boundedness of stretched product diffeomorphisms).
Let us fix a domain $\Omega \supset \bigotimes \overline{B_k(0,2^m)}$ and let us consider a bounded family of product diffeomorphisms of Ω into $\mathbb{R}^N$. For any given $G \subset \{1,\dots,d\}$ the product diffeomorphisms obtained by stretching all the kernels of the family along H_G are defined at least on $\bigotimes \overline{B_k(0,2^m)}$ and are uniformly bounded for the parameter s in the range $0 \leq s \leq 1$.

Proof. The proof of this property can also be obtained by verifying (2.11.1) directly. □

We also need a lemma on the convergence of dyadic series with building that do not have complete cancellation along all subspaces but only a weak equivalent.

Definition 2.11.6 (Weak cancellation).
Let

$$\{\varphi_{\boldsymbol{i}}\}_{\boldsymbol{i}\in\mathbb{Z}^d} \subset S(\mathbb{R}^N)$$

be a family of functions uniformly bounded in $S(\mathbb{R}^N)$. We say that this family has weak cancellation (with parameter ε) along subspaces H_j with $j \in J \subset \{1,\dots,d\}$ if there exists an $\varepsilon>0$ such that for any subset of indexes $\widetilde{J} \subset J$

$$2^{-\varepsilon \min\{i_j | j\in\widetilde{J}\}} \int_{H_{\widetilde{J}}} \varphi_{\boldsymbol{i}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}_{\widetilde{J}}$$

are a uniformly bounded family of functions in $S(H_{(J)})$.

Proposition 2.11.7 (Dyadic sums with weak cancellation).
Consider a multi-order $\boldsymbol{\nu}$ and let $J \subset \{1,\dots,d\}$ such that $\nu_j = 0$ for $j\in J$. Let

$$\{\varphi_{\boldsymbol{i}}\}_{\boldsymbol{i}\in\mathbb{Z}^d}$$

be a uniformly bounded family of functions in $S(\mathbb{R}^N)$ that possesses the following properties.

- *The family $\{\varphi_{\boldsymbol{i}}\}$ possesses weak cancellation with some parameter ε along subspace H_j with $j \in J$.*
- *If $k \in \{1, \dots, d\} \setminus J$ is such that $\nu_k \geq 0$ then*
$$\int_{H_k} x_k^{\alpha_k} \varphi_{\boldsymbol{i}}(\boldsymbol{x}) \mathrm{d}x_k = 0 \qquad \forall \|\alpha_k\| \leq l_k$$
for some $l_k \in \mathbb{N}$, $l_k \geq \nu_k$.
- *If $k \in \{1, \dots, d\}$ is such that $\nu_k \leq -q_k$ then*
$$\partial_{x_k}^{\alpha_k} \varphi_i(0_k \oplus \boldsymbol{x}_{(k)}) = \int_{H_k} \xi_k^{\alpha_k} \widehat{\varphi}_{\boldsymbol{i}}(\boldsymbol{\xi}) \mathrm{d}\xi_k = 0 \qquad \forall \|\alpha_k\| \leq l'_k$$
for some $l'_k \in \mathbb{N}$, $l'_k \geq -q_k - \nu_k \geq 0$.
- *if $k \in \{1, \dots, d\}$ is such that $-q_k < \nu_k < 0$, no condition along H_k needs to be satisfied.*

Then for any fixed $M \in Z$ the dyadic sum
$$\sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ i_j \leq M \ \forall j \in J}} 2^{-\boldsymbol{i}\boldsymbol{q}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x})$$
converges to a kernel in $\mathcal{PK}(\boldsymbol{\nu})$. All the partial sums are uniformly bounded in $\mathcal{PK}(\boldsymbol{\nu})$.

Proof. This proof is closely related to the proof of Theorem 2.7.2. We reason by induction on $|J|$. For $J = \emptyset$ the statement reduces to Theorem 2.7.2. Suppose that the statement is true when $|J| \leq D - 1$ for some $D \geq 1$. Let us suppose that $k \in J$ and let
$$\widetilde{\eta}_k(x_k) \stackrel{\text{def}}{=} \frac{\eta_k(x_k)}{\int_{H_k} \eta_k(x_k) \mathrm{d}x_k}.$$
Now set
$$\widetilde{\varphi}_{\boldsymbol{i}}(\boldsymbol{x}) \stackrel{\text{def}}{=} 2^{-\varepsilon/2 i_k} \eta_k(x_k) \int_{H_k} \varphi_{\boldsymbol{i}}(\boldsymbol{x}) \mathrm{d}x_k.$$
We have that
$$\{\varphi_{\boldsymbol{i}} - 2^{\varepsilon/2 i_k} \widetilde{\varphi}_{\boldsymbol{i}}\}$$
is a family of Schwartz functions that satisfy the hypothesis of the proposition and have weak cancellation along $J \setminus \{k\}$. Using the induction hypothesis the dyadic sum relative to this family converges to a kernel in $\mathcal{PK}(\boldsymbol{\nu})$. On the other hand let $\widetilde{\boldsymbol{\nu}}$ be the multi-order such that $\widetilde{\nu}_l = \nu_l$ for $l \neq k$ and $\widetilde{\nu}_k = -\varepsilon/2$. Then $\{\widetilde{\varphi}_{\boldsymbol{i}}\}$ is a family that satisfies the hypothesis of this preposition for the multi-order $\widetilde{\boldsymbol{\nu}}$ with weak cancellation of parameter $\varepsilon/2$ along $J \setminus \{k\}$. Furthermore the resulting kernel is bounded along H_k so Proposition 2.8.4 states that there is a continuous inclusion $\mathcal{PK}(\widetilde{\boldsymbol{\nu}}) \hookrightarrow \mathcal{PK}(\boldsymbol{\nu})$. This concludes the proof.

□

We have the following stability theorem for product kernels.

Theorem 2.11.8 (Product kernel stability w.r.t diffeomorphisms)**.**
Let $K \in \mathcal{PK}(\boldsymbol{\nu})$ be a product kernel on $\mathbb{R}^N = \bigoplus_{k=1}^d H_k$ with multi-order $\boldsymbol{\nu}$ such that $\nu_k \leq 0$ for all $k \in \{1, \dots, d\}$ and with bounded support along all subspaces H_k. Let Φ be a product diffeomorphism of a compact domain $\Omega \supset \bigotimes_{k=1}^d \overline{B_k(0, 2^{m_k})}$ such that $\operatorname{spt} K \subset \bigotimes_{k=1}^d \overline{B_k(0, 2^{m'_k})} \subset \Phi(\Omega)$ for some $m'_k \in \mathbb{Z}$. Then $K \circ \Phi$ is a product kernel in $\mathcal{PK}(\boldsymbol{\nu})$.

For any fixed bounds $m_k \in \mathbb{Z}$ the mapping $K \mapsto K \circ \Phi$ is continuous from the kernels in $\mathcal{PK}(\boldsymbol{\nu})$ such that $\operatorname{spt} K \subset \bigotimes_{k=1}^{d} \overline{B_{H_k}(0, 2^{m'_k})} \subset \Phi(\Omega)$ *to $\mathcal{PK}(\boldsymbol{\nu})$. Furthermore if Φ varies among a uniformly bounded family of product diffeomorphisms then the kernels $K \circ \Phi$ are uniformly bounded.*

Proof. We indicate $\Phi^k \stackrel{\text{def}}{=} \pi_k \circ \Phi$. Theorem 2.8.6 gives us the decomposition (2.8.1). Since the support of all the dyadic building blocks is inside the image of Φ we can write

$$K \circ \Phi(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ i_k \leq m'_k - 2}} 2^{-\boldsymbol{i}\boldsymbol{\nu}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\Phi(\boldsymbol{x})) = \sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ i_k \leq m'_k - 2}} 2^{-\boldsymbol{i}\boldsymbol{\nu}} 2^{-\boldsymbol{i}\boldsymbol{q}} \varphi_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \Phi(\boldsymbol{x})).$$

As a matter of fact the above equality holds in the distributional sense because the mapping $\psi(\boldsymbol{x}') \mapsto \left(\det D\Phi^{-1}\right)(\boldsymbol{x}')$ is continuous from $D\left(\Phi(\Omega)\right)$ to $D\left(\Omega\right)$. Ω is compact so $\Phi(\Omega)$ is also compact and Φ extends to a diffeomorphism of an open neighborhood Ω' of Ω. Let $\rho \in C_c^\infty(\mathbb{R}^N)$ be a function that is identically 1 on a neighborhood $\Phi(\Omega)$ and 0 outside $\Phi(\Omega')$. Notice that $\Phi(\Omega)$ is compact. For any $\psi \in S(\mathbb{R}^N)$ we have

$$\begin{aligned}\int_{\mathbb{R}^N} K \circ \Phi(\boldsymbol{x}) \psi(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \stackrel{\text{def}}{=} &\int_{\mathbb{R}^N} K \circ \Phi(\boldsymbol{x}) \rho(\boldsymbol{x}) \psi(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \\ &\int_{\mathbb{R}^N} K(\boldsymbol{x}') \left(\det D\Phi^{-1}\right)(\boldsymbol{x}') \rho \circ \Phi^{-1}(\boldsymbol{x}') \psi \circ \Phi^{-1}(\boldsymbol{x}') \mathrm{d}\boldsymbol{x}'.\end{aligned}$$

The above expression is well-defined and independent of ρ because K is supported in the compact set Ω and Φ is smooth. By a slight abuse of notation, omitting ρ we can write

$$\begin{aligned}\int_{\mathbb{R}^N} K \circ \Phi(\boldsymbol{x}) \psi(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = &\int_{\mathbb{R}^N} K(\boldsymbol{x}') \left(\det D\Phi^{-1}\right)(\boldsymbol{x}') \psi \circ \Phi^{-1}(\boldsymbol{x}') \mathrm{d}\boldsymbol{x}' = \\ &\int_{\mathbb{R}^N} \sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ i_j \leq m'_j - 2}} 2^{-\boldsymbol{i}\boldsymbol{\nu}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}') \left(\det D\Phi^{-1}\right)(\boldsymbol{x}') \psi \circ \Phi^{-1}(\boldsymbol{x}') \mathrm{d}\boldsymbol{x}' = \\ &\sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ i_j \leq m'_j - 2}} \int_{\mathbb{R}^N} 2^{-\boldsymbol{i}\boldsymbol{\nu}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}\left(\Phi(\boldsymbol{x})\right) \psi(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}.\end{aligned}$$

By setting

$$\widetilde{\varphi}_{\boldsymbol{i}}(\boldsymbol{x}) = \varphi_{\boldsymbol{i}}\left(2^{-\boldsymbol{i}} \cdot \Phi(2^{\boldsymbol{i}} \cdot \boldsymbol{x})\right)$$

we have that

$$K \circ \Phi(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ i_k \leq m'_k - 2}} 2^{-\boldsymbol{i}\boldsymbol{\nu}} \widetilde{\varphi}_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}). \tag{2.11.2}$$

Using (2.11.1) we have that the support of $\widetilde{\varphi}_{\boldsymbol{i}}$ is given by

$$\mathcal{D} \stackrel{\text{def}}{=} \left\{\boldsymbol{x} \mid 2^{i_k - 2} < \left|\Phi^k(2^{\boldsymbol{i}} \cdot \boldsymbol{x})\right| < 2^{i_k + 2}\right\} = \left\{\boldsymbol{x} \mid (4C_k)^{-1} < |x_k| < 4C_k\right\}.$$

The differential inequalities (2.11.1) on the diffeomorphism also guarantee that $\widetilde{\varphi}_{\boldsymbol{i}}$ are bounded in $S(\mathbb{R}^N)$.

We will now show that $\widetilde{\varphi}_{\boldsymbol{i}}$ has weak cancellation along $J = \left\{ j \in \{1, \ldots, d\} \,\middle|\, \nu_j = 0 \right\}$. Let $j \in J$ and we write an expansion for Φ^k with an integral remainder term. Taking in account the geometric properties we have

$$2^{-i_k} \Phi^k(2^{\boldsymbol{i}} \cdot \boldsymbol{x}) = \int_0^1 \partial_{x_k} \Phi^k \left(t \cdot 2^{i_k} \cdot x_k \oplus 2^{\boldsymbol{i}_{(k)}} \cdot \boldsymbol{x}_{(k)} \right) x_k \mathrm{d}t.$$

Now expanding in x_j we have for $k = j$

$$2^{-i_j} \Phi^j(2^{\boldsymbol{i}} \cdot \boldsymbol{x}) = \partial_{x_j} \Phi^j \left(0_j \oplus 2^{\boldsymbol{i}_{(j)}} \cdot \boldsymbol{x}_{(j)} \right) x_j +$$

$$2^{i_j} \int_0^1 \int_0^1 \partial^2_{x_j} \Phi^j \left(t_1 t_2 \cdot 2^{i_j} \cdot x_j \oplus 2^{\boldsymbol{i}_{(j)}} \cdot \boldsymbol{x}_{(j)} \right) x_j^2 t_1 \mathrm{d}t_1 \mathrm{d}t_2$$

and for $j \neq k$ we have

$$2^{-i_k} \Phi^k(2^{\boldsymbol{i}} \cdot \boldsymbol{x}) = \int_0^1 \partial_{x_k} \Phi^k \left(t_1 \cdot 2^{i_k} \cdot x_k \oplus 0_j \oplus 2^{\boldsymbol{i}_{(\{j,k\})}} \cdot \boldsymbol{x}_{(\{j,k\})} \right) x_k \mathrm{d}t_1 +$$

$$2^{i_j} \int_0^1 \int_0^1 \partial_{x_k} \partial_{x_j} \Phi^k \left(t_1 \cdot 2^{i_k} \cdot x_k \oplus t_2 \cdot 2^{i_j} \cdot x_j \oplus 2^{\boldsymbol{i}_{(\{j,k\})}} \cdot \boldsymbol{x}_{(\{j,k\})} \right) x_k x_j \mathrm{d}t_1 \mathrm{d}t_2.$$

Since we are working on the domain $\mathcal{D}$ that is compact and since $i_j \leq M$ we have that $2^{i_j} x_j$ is bounded. Using the differential inequalities (2.11.1) we have that the remainder terms

$$Q^j_{\boldsymbol{i}}(\boldsymbol{x}) \stackrel{\text{def}}{=} \int_0^1 \int_0^1 \partial^2_{x_j} \Phi^j \left(t_1 t_2 \cdot 2^{i_j} \cdot x_j \oplus 2^{\boldsymbol{i}_{(j)}} \cdot \boldsymbol{x}_{(j)} \right) x_j^2 t_1 \mathrm{d}t_1 \mathrm{d}t_2$$

and

$$Q^k_{\boldsymbol{i}}(\boldsymbol{x}) \stackrel{\text{def}}{=} \int_0^1 \int_0^1 \partial_{x_k} \partial_{x_j} \Phi^k \left(t_1 \cdot 2^{i_k} \cdot x_k \oplus t_2 \cdot 2^{i_j} \cdot x_j \oplus 2^{\boldsymbol{i}_{(\{j,k\})}} \cdot \boldsymbol{x}_{(\{j,k\})} \right) x_k x_j \mathrm{d}t_1 \mathrm{d}t_2.$$

are uniformly bounded and have uniformly bounded derivatives for all admissible $\boldsymbol{i}$. The inequalities (2.11.1) also guarantee that the map

$$\int_0^1 \partial_{x_k} \Phi^k \left(t_1 \cdot 2^{i_k} \cdot x_k \oplus 0_j \oplus 2^{\boldsymbol{i}_{(\{j,k\})}} \cdot \boldsymbol{x}_{(\{j,k\})} \right) x_k \mathrm{d}t_1$$

is smooth, uniformly bounded, and has all derivatives uniformly bounded. So, by setting

$$\Phi^k_{t,\boldsymbol{i}}(\boldsymbol{x}) \stackrel{\text{def}}{=} 2^{-i_k} \cdot \Phi^k(2^{\boldsymbol{i}}) + (t-1) 2^{i_j} Q^k_{\boldsymbol{i}}(\boldsymbol{x})$$

we can write

$$\widetilde{\varphi}_{\boldsymbol{i}}(\boldsymbol{x}) = \varphi\left(\Phi_{0,\boldsymbol{i}}(\boldsymbol{x})\right) + 2^{i_j} \sum_{k=1}^{d} \int_0^1 (\partial_{x_k} \varphi) \left(\Phi_{t,\boldsymbol{i}}(\boldsymbol{x})\right) Q^k_{\boldsymbol{i}}(\boldsymbol{x}) \mathrm{d}t.$$

$\Phi_{0,\boldsymbol{i}}(\boldsymbol{x})$ has the property that $\pi_k\left(\Phi_{0,\boldsymbol{i}}(\boldsymbol{x})\right)$ is independent of x_j for $j \neq k$ and for a fixed $\boldsymbol{x}_{(j)}$ the $\pi_j\left(\Phi_{0,\boldsymbol{i}}(\boldsymbol{x})\right)$ is a linear non-degenerate map from H_j onto itself. In particular this means that

$$\int_{H_j} \varphi\left(\Phi_{0,\boldsymbol{i}}(\boldsymbol{x})\right) \mathrm{d}x_j = 0$$

and thus

$$2^{-i_j} \int_{H_j} \widetilde{\varphi}_{\boldsymbol{i}}(\boldsymbol{x}) \mathrm{d}x_j$$

is a uniformly bounded family of functions in $S(H_{(j)})$. To see that weak cancellation holds for a generic $\widetilde{J} \subset J$ it is sufficient to choose $j \in \widetilde{J}$ such that i_j is minimal among $i_{\widetilde{j}}$ with $\widetilde{j} \in \widetilde{J}$. It suffices to apply the reasoning above and then to integrate along the remaining variables $\boldsymbol{x}_{\widetilde{J} \setminus \{j\}}$. By Proposition 2.11.7 the resulting dyadic sum (2.11.2) converges to a kernel in $\mathcal{PK}(\boldsymbol{\nu})$. □

2.11.1 An application of functional calculus

We now show how the functional calculus we developed can be applied to the problem of product diffeomorphisms. We start from some lemmas.

Lemma 2.11.9 (Continuity of composition with respect to diffeomorphisms).
Suppose Φ_s with $s \in \mathbb{R}^+ \cup \{0\}$ is an indexed family of product diffeomorphisms of a compact set Ω such that the mapping $(s, \boldsymbol{x}) \mapsto \Phi_s(\boldsymbol{x})$ is smooth and all Φ_s satisfy the conditions of Theorem 2.11.8. Given $K \in \mathcal{PK}(\boldsymbol{\nu})$ that also satisfies the requirements of Theorem 2.11.8, the mapping

$$s \mapsto K \circ \Phi_s$$

is continuous from $\mathbb{R}^+ \cup \{0\}$ to $\mathcal{PK}(\boldsymbol{\nu})$ with the strong topology.

Proof. When s varies over compact subsets of $\mathbb{R}^+ \cup \{0\}$ the smoothness of $(s, \boldsymbol{x}) \mapsto \Phi_s(\boldsymbol{x})$ gives us the uniform boundedness of the family of diffeomorphisms Φ_s. Furthermore for any $\varphi \in D(\Omega)$ we have that $s \mapsto \det\left(\Phi_s^{-1}\right) \varphi \circ \Phi_s^{-1}$ is continuous as a mapping into $D(\mathbb{R}^N)$. This means that

$$s \mapsto K \circ \Phi_s$$

is continuous in the weak-* sense. By uniform boundedness of $K \circ \Phi_s$ and Proposition 2.4.4 the mapping is continuous with respect to the strong topology on $\mathcal{PK}(\boldsymbol{\nu})$. □

Now suppose that we have a smooth, uniformly bounded vector field $W_s(\boldsymbol{x})$ on $\mathbb{R}^N$ such that for any $k \in \{1, \dots, d\}$ we have that $W_s(\boldsymbol{x}) \in H_k^\perp$ if $\boldsymbol{x} \in H_k^\perp$. It follows that the flow Φ_s of such a vector field is a family of diffeomorphisms that satisfy the hypothesis of Lemma 2.11.9. Now suppose that K is a kernel with compact support in $\mathcal{PK}(\boldsymbol{\nu})$ with $\nu_k \le 0$ for all $k \in \{1, \dots, d\}$ and let us also suppose that K coincides with a smooth function. We can write that

$$\frac{\mathrm{d}}{\mathrm{d}s}(K \circ \Phi_s) = \sum_{k=1}^{d} \left((\partial_{x_k} K) \circ \Phi_s\right) \left(\tfrac{\mathrm{d}}{\mathrm{d}s}\Phi^k\right) = (\nabla K) \circ \Phi_s \cdot \frac{\mathrm{d}}{\mathrm{d}s}\Phi_s$$

Notice however that

$$(\nabla K) \circ \Phi = \nabla(K \circ \Phi_s) \circ (D\Phi_s)^{-1} = \nabla(K \circ \Phi_s) \cdot (D\Phi_{-s})(\Phi_s).$$

Taking in account that Φ_s is the flow of W_s, so

$$\frac{\mathrm{d}}{\mathrm{d}s}\Phi_s = W_s \circ \Phi_s,$$

we get the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}s}(K \circ \Phi_s) = \nabla(K \circ \Phi_s) \cdot \left((D\Phi_{-s})W_s\right)(\Phi_s).$$

Let $\widetilde{W}_s \overset{\text{def}}{=} \left((D\Phi_{-s})W_s\right)(\Phi_s)$. It is easy to verify that $\widetilde{W}_s$ is also a smooth, bounded vector field for $s \in [0, T]$ for any finite T and it satisfies $\widetilde{W}_s(\boldsymbol{x}) \in H_k^\perp$ if $\boldsymbol{x} \in H_k^\perp$. This means that $|x_k|^{-1} \pi_k \widetilde{W}_s(\boldsymbol{x})$ is also smooth. So we have

$$K \circ \Phi_1 = K \circ \Phi_0 + \sum_{k=1}^{d} \int_{s=0}^{1} \partial_{x_k}(K \circ \Phi_s)\, |x_k|\, \pi_k \left(|x_k|^{-1}\, \widetilde{W}_s(\boldsymbol{x})\right) \mathrm{d}x. \tag{2.11.3}$$

Since

$$s \mapsto K \circ \Phi_s$$

using Proposition 2.6.1 the mapping

$$s \mapsto \partial_{x_k}(K \circ \Phi_s)\,|x_k|$$

is also continuous from $\mathbb{R}^+ \cup \{0\}$ to $\mathcal{PK}(\boldsymbol{\nu})$. Since

$$\pi_k\left(|x_k|^{-1}\,\widetilde{W}_s(\boldsymbol{x})\right)$$

is a smooth function in $(s, \boldsymbol{x})$ and $K \circ \Phi_s$ has compact support, the whole integrand is continuous function depending on s with values in $\mathcal{PK}(\boldsymbol{\nu})$. For this reason the equation 2.11.3 can be also written for any $K \in \mathcal{PK}(\boldsymbol{\nu})$ and not necessarily for a smooth function. This holds by approximation. Using the continuity of all the maps above the integral can be approximated in the strong topology on $\mathcal{PK}(\boldsymbol{\nu})$ by piecewise sums.

Chapter 3

Flag kernels

While the theory we developed in the previous chapter is interesting in its own right it also happens to be a test case for any type of multi-parameter theory. However the class of product kernels is too large for many applications. In some sense a generic product kernel has "too many" singularities. For this reason we want to concentrate on a similar "multi-parameter" or product class of distributions, flag kernels, whose singularities are in some sense more controlled. In particular product kernels have singularities along all coordinate sub-space corresponding to the product factors. Flag kernels, on the other hand, have singularities concentrated along only one coordinate subspace. Many of these results have been recently published in [Nag+12].

Definition 3.0.10 (Flags and gradations)**.**
A flag on $\mathbb{R}^N$, is a finite sequence of subspaces that we will indicate $(V_k)_{k\in\{0,\dots,d\}}$ such that

$$\{0\} = V_0 \subset V_1 \subset \ldots \subset V_{d-1} \subset V_d = \mathbb{R}^N.$$

We say that an ordered sequence of subspaces $(H_k)_{k\in\{1,\dots,d\}}$ is a gradation adapted to the flag $(V_k)_{k\in\{0,\dots,d\}}$ if

$$V_k = \bigoplus_{j=1}^{k} H_j \qquad \forall k \in \{1,\dots,d\}.$$

Observe that while a gradation naturally determines a flag, the same flag can possess more than one adapted gradation. The gradation associated with a given flag is not unique. A gradation gives us a natural product space structure on $\mathbb{R}^N$. If $\mathbb{R}^N$ is endowed with a family of (non-isotropic) dilations we will only consider flags compatible with that dilation structure. In particular we ask that all subspaces V_k be eigen-spaces of the family of dilations. This is equivalent to asking that there be an associated gradation of eigen-spaces. We extend the usual notation described in Section 2.1. If $J \subset \{1,\dots,d\}$ is a subset of indexes then the flag V_J is the flag on H_J associated to that ordered gradation.

Once a gradation has been chosen for a given flag we can also introduce the product space coordinates so that, as usual, $\boldsymbol{x} = (x_1,\dots,x_d)$ with $x_k \in H_k$. If V_k is a subspace of the flag then $\mathrm{d}(\boldsymbol{x};V_k)$, the distance from the subspace, is equivalent up to a constant to the expression

$$\mathrm{d}(\boldsymbol{x};V_k) \approx \|x_{k+1}\| + \ldots + \|x_d\|.$$

Flag kernels are similar to product kernels and many properties and proofs will be formally similar to the ones in Section 2. However, we note that we introduce flag kernels of order $\mathbf{0}$. As a matter of fact, extending the definition to other pseudo-differential orders is less trivial than what we have done for product kernels, even for "good" orders $\boldsymbol{\nu}$ such that $\nu_k > -q_k$.

3.1 Definition

For a given flag $(V_k)_{k\in\{0,\dots,d\}}$ let us fix a gradation of $\mathbb{R}^N$, $(H_k)_{k\in\{1,\dots,d\}}$, compatible with the flag. We recall that

$$V_k = \oplus_{j=1}^k H_j.$$

We will later show that the definition of flag kernels is actually independent of the gradation and depends only on the flag.

Definition 3.1.1 (Flag kernels).
Consider a flag $(V_k)_{k\in\{0,\dots,d\}}$ with an associated gradation $(H_k)_{k\in\{1,\dots,d\}}$.

For $d = 1$, the class of flag kernels $\mathcal{FK}_{(H_k)}$ coincides with $\mathcal{CZ}(0)$ on $\mathbb{R}^N$ as defined in 1.2.9. *We say that a family of kernels in $\mathcal{CZ}(0)$ is uniformly bounded if the inequalities* (1.2.4) *and* (1.2.5) *hold with uniformly bounded constants.*

For $d > 1$, we say that a distribution $K \in S'(\mathbb{R}^N)$ is of class $\mathcal{FK}_{(H_k)}$ if, away from the subspace $V_{d-1} = H_d^\perp$, K coincides with a smooth function i.e.

$$K\big|_{\mathbb{R}^N\setminus V_{d-1}} \in C^\infty\left(\mathbb{R}^N \setminus V_{d-1}\right)$$

and it satisfies the following two kinds of conditions:

Size conditions

$$\left|\partial_{x_1}^{\alpha_1}\dots\partial_{x_d}^{\alpha_d}K(\boldsymbol{x})\right| \le C_{\boldsymbol{\alpha}}\,\mathrm{d}(\boldsymbol{x},V_0)^{-q_1-\|\alpha_1\|}\dots\mathrm{d}(\boldsymbol{x},V_{d-1})^{-q_d-\|\alpha_d\|} \tag{3.1.1}$$

for any $\boldsymbol{x} \notin V_{d-1}$ and for any multi-index $\boldsymbol{\alpha}$.

Cancellation conditions *For every $k \in \{1,\dots,d\}$ the distributions*

$$\int K(\boldsymbol{x})\varphi(R^{-1}\cdot x_k)\mathrm{d}x_k \tag{3.1.2}$$

obtained by contracting K with all possible rescaled C^1-normalized bump functions φ on the subspace H_k, are a family of flag kernels on $H_{(k)}$ relative to the flag $V_{(k)}$ and to the corresponding gradation $\left(H_j\right)_{j\neq k}$, uniformly bounded with respect to $R > 0$ and to φ.

We say that a family of kernels in $\mathcal{FK}_{(H_k)}$ is uniformly bounded if the bounds arising inductively from (3.1.1) *and* (3.1.2) *on all the kernels of the family are uniformly bounded.*

The formal difference between product and flag kernels consists in the size conditions. The cancellation conditions are expressed in a similar way. On the other hand it is immediately evident that the size conditions (3.1.1) do not depend on the gradation but only on the flag. We will usually omit explicitly writing the dependence on the flag and on the associated gradation by indicating the class of flag kernels simply as $\mathcal{FK}$.

As we did for product kernels, the lower bounds on the constants that appear inductively from (3.1.1) and (3.1.2) can be combined to form a family of semi-norms on $\mathcal{FK}$ and to define a Fréchet space topology. A non-inductive definition of the semi-norms on $\mathcal{FK}$ similar to Definition 2.3.2 is notationally very complicated so prefer to avoid it. From now on we will indicate as $\mathcal{FK}$ the space of flag kernels endowed with its strong topology. We now turn to some basic topology results on $\mathcal{FK}$ that correspond to those obtained in section 2.4.

3.2 Topology

Flag kernels are very closely related to product kernels. As a matter of fact flag kernels form a special subset of product kernels of order $\mathbf{0}$ and the strong topology on $\mathcal{FK}$ is finer than the one on $\mathcal{PK}$. This can be easily checked since by induction on the length of the flag d.

As a consequence, some properties on $\mathcal{PK}(\mathbf{0})$ apply immediately to flag kernels.

Proposition 3.2.1 (The semi-norms separate points).
Let K be a flag kernel with all norms arising inductively from conditions (3.1.1) *and* (3.1.1) *equal to* 0. *Then the kernel is trivial.*

Proposition 3.2.2 (Weak-* and strong topologies).
The following relations between the strong topology and the weak- topology hold on $\mathcal{FK}$:*

1. *The strong topology is finer than the weak-* topology on $\mathcal{FK}$.*
2. *For any bounded set $V \in S(\mathbb{R}^N)$ and any $\varepsilon > 0$ there is a neighborhood $U_{V,\varepsilon}$ of 0 in $\mathcal{FK}$ such that if $\varphi \in V$ and $K \in U_{V,\varepsilon}$ then $\left|\int_{R^N} K(\boldsymbol{x})\varphi(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\right| < \varepsilon$. Vice-versa for any bounded set $U \in \mathcal{FK}$ and any $\varepsilon > 0$ there is a neighborhood $V_{U,\varepsilon}$ of 0 in $S(\mathbb{R}^N)$ such that if $\varphi \in V_{U,\varepsilon}$ and $K \in U$ then $\left|\int_{R^N} K(\boldsymbol{x})\varphi(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\right| < \varepsilon$.*
3. *On bounded sets of $\mathcal{FK}$ weak-* convergence can be verified only on a dense subset of test functions.*
4. *On bounded sets of $\mathcal{FK}$ the weak-* convergence implies strong convergence.*
5. *The weak-* closure in $S'(\mathbb{R}^N)$ of a bounded set in $\mathcal{PK}(\boldsymbol{\nu})$ is a closed subset of $\mathcal{FK}$.*
6. *The semi-norms on $\mathcal{FK}$ are lower semi-continuous with respect to the weak-* convergence.*

As a consequence, the space $\mathcal{FK}$ is, too, complete with respect to the metric given by the family of semi-norms.

Corollary 3.2.3 (Fréchet space topology).
$\mathcal{FK}$ with the topology given by the family of semi-norms arising as the lower bounds of the constants in (3.1.1) *and* (3.1.2) *is a Fréchet space.*

As for product kernels there is a certain flexibility in the choice of the normalization order for bump functions in Definition 2.3.1.

Proposition 3.2.4 (Bump function normalization orders).
Let $\widetilde{\mathcal{FK}}$ be the class of flag kernels with the exception that in the definition 3.1.1 we require that the bump function in condition (3.1.2) *be b_k-normalized for some $b_k \geq 1$. Then the classes $\widetilde{\mathcal{FK}}$ and $\mathcal{FK}$ coincide as vector spaces and have the same topologies.*

Proof. The proof goes along the same lines as that of proposition 2.4.11 for product kernels. We recall the main idea of the proof.

The continuous inclusion $\mathcal{FK}(\boldsymbol{\nu}) \subset \widetilde{\mathcal{FK}}(\boldsymbol{\nu})$ is obvious since the conditions on $\mathcal{FK}(\boldsymbol{\nu})$ are more strict.

To prove the converse we reason by induction on the length of the flag d. For $d = 0$ there is nothing to prove. Suppose that $d \geq 1$ and that the proposition holds for flag of length up to $d-1$. To prove the statement for a flag kernel $K \in \widetilde{\mathcal{FK}}$ adapted to a flag of length d we must only check the cancellation conditions on K. As a matter of fact the size conditions (3.1.1) are the same for both classes $\widetilde{\mathcal{FK}}(\boldsymbol{\nu})$ and $\mathcal{FK}(\boldsymbol{\nu})$. Furthermore it suffices to check that given such a kernel K, for any $k \in \{1, \dots, d\}$ the family of kernels

$$\int_{H_k} K(\boldsymbol{x})\varphi(R^{-1} \cdot x_k)\,\mathrm{d}x_k$$

is uniformly bounded in $\widetilde{\mathcal{FK}}$ when φ is C^1-normalized and $R > 0$. This last class coincides by the induction hypothesis with $\mathcal{FK}$.

To prove the above uniform boundedness one decomposes φ using 2.4.1 into

$$\varphi(x_k) = \widetilde{\varphi}(x_k) + \sum_{|\beta|=b_k+1} x_k^{\beta}\varphi_{\beta}(x_k)$$

where $\widetilde{\varphi}$ is b_k-normalized and φ_β are 0-normalized. It is thus evident that

$$\int_{H_k} K(\boldsymbol{x})\widetilde{\varphi}(R^{-1} \cdot x_k)\,\mathrm{d}x_k$$

is in the correct class $\widetilde{\mathcal{FK}}_{(V_k)}$. For the remaining term

$$\sum_{|\beta|=b_k+1} \int_{H_k} K(\boldsymbol{x})(R^{-1} \cdot x_k)^{\beta}\varphi_{\beta}(R^{-1} \cdot x_k)\,\mathrm{d}x_k$$

one uses an argument based on the integrability of $K(\boldsymbol{x})x_b^{\beta}$ across $H_k^{\perp}$. □

3.3 Fourier transform

A somewhat similar result to 2.5.4 holds for flag kernels. We must illustrate the relationship between flag structures on an Euclidean space $\mathbb{R}^N$ and the induced flag structures on its dual $(\mathbb{R}^N)^*$.

Definition 3.3.1 (Dual flag)**.**
Let $(V_k)_{k\in\{0,\dots,d\}}$ be a flag on R^N. The dual flag $(\widehat{V}_k)_{k\in\{0,\dots,d\}}$ on the dual space to $\mathbb{R}^N$ is given by the relation

$$\widehat{V}_k = \operatorname{Ann} V_k = \{\boldsymbol{\xi} \in (\mathbb{R}^N)^* \mid \boldsymbol{\xi x} = 0 \quad \forall \boldsymbol{x} \in V_k\}.$$

Given a gradation $(H_k)_{k\in\{1,\dots,d\}}$ on $\mathbb{R}^N$ the dual gradation $(\widehat{H}_k)_{k\in\{1,\dots,d\}}$ on $(\mathbb{R}^N)^*$ is given by

$$\widehat{H}_k = \bigcap_{j\neq k} \operatorname{Ann} H_j.$$

Notice that

$$\{0\} \subset \widehat{V}_d \subset \widehat{V}_{d-1} \subset \ldots \subset \widehat{V}_1 \subset \widehat{V}_0 = (\mathbb{R}^N)^*$$

and

$$\widehat{V}_k = \bigoplus_{j=k+1}^{d} \widehat{H}_j$$

so the order of the flag is naturally inverted. Choosing a basis of eigen -vectors for the dilations on $\mathbb{R}^N$ compatible with the gradation $(H_k)_{k\in\{1,\dots,d\}}$ gives a basis and thus a coordinate system on $(\mathbb{R}^N)^*$ such that in particular

$$\langle \boldsymbol{\xi};\, \boldsymbol{x}\rangle = \sum_{k=1}^{d}\sum_{j=1}^{n_k} \xi_{k,j}x_{k,j}.$$

We now define flag multipliers on the dual space $(\mathbb{R}^N)^*$.

Definition 3.3.2.
Let $m \in S'\left((\mathbb{R}^N)^*\right)$ and fix a dual flag $(\widehat{V}_k)_{k\in\{0,\dots,d\}}$. A locally integrable function m is a flag multiplier if it is smooth on $(\mathbb{R}^N)^* \setminus \widehat{V}_1 = (\mathbb{R}^N)^* \setminus \widehat{H}_1^\perp = \left\{\boldsymbol{\xi} \,|\, \xi_1 = 0\right\}$ and it satisfies the following size conditions:

$$\left|\partial_{\boldsymbol{\xi}}^{\boldsymbol{\alpha}} m(\boldsymbol{\xi})\right| < C_{\boldsymbol{\alpha}}\, \mathrm{d}(\boldsymbol{\xi}, \widehat{V}_d)^{-\|\alpha_d\|} \dots \mathrm{d}(\boldsymbol{\xi}, \widehat{V}_1)^{-\|\alpha_1\|} \tag{3.3.1}$$

for all $\boldsymbol{\xi} \notin \widehat{V}_1$ and all multi-indexes $\boldsymbol{\alpha}$

Remark 3.3.3.
The lower bounds on $C_{\boldsymbol{\alpha}}$ in (3.3.1) *are a countable family of semi-norms that define a Fréchet space structure on the space of flag multipliers $\widehat{\mathcal{FK}}$.*

We have the following theorem about the relationship between flag kernels and flag multipliers.

Theorem 3.3.4 (Fourier transform of flag kernels)**.**
Let $K \in \mathcal{FK}$ be a flag kernel on $\mathbb{R}^N$ with respect to a flag $(V_k)_{k\in\{0,\dots,d\}}$. Then the Fourier transform of K is a flag multiplier with respect to the dual flag $(\widehat{V}_k)_{k\in\{0,\dots,d\}}$. Vice-versa the inverse Fourier transform of a flag multiplier with respect to the flag $(\widehat{V}_k)_{k\in\{0,\dots,d\}}$ is a flag kernel in $\mathcal{FK}_{(V_k)}$. The Fourier transform is continuous form $\mathcal{FK}$ to $\widehat{\mathcal{FK}}$ with continuous inverse.

Proof. First we prove that the Fourier transform of a flag kernel is a flag multiplier. We can suppose that K has compact support. We need to prove the differential inequalities on the multiplier.

We reason by induction on the length of the flag d. For $d = 0$ there is nothing to prove. Now suppose that $d \geq 1$ and the proposition is true for all flags with up to $d-1$ terms. Let $E_k \subset (\mathbb{R}^N)^*$ such that $\|\xi_1\| \geq \left\|\xi_j\right\|$ for all $j \leq k$ and if $k \leq d-1$ then $\|\xi_1\| < \|\xi_{k+1}\|$ and let $L_k = \{1,\dots,k\}$. We prove that for any multi-index $\boldsymbol{\alpha}$ we have that

$$\|\xi_1\|^{\|\boldsymbol{\alpha}_{L_k}\|}\, \partial_{\boldsymbol{\xi}_{L_k}}^{\boldsymbol{\alpha}_{L_k}} (\mathcal{F}_{H_{L_k}} K)(\boldsymbol{\xi}_{L_k}, \boldsymbol{x}_{(L_k)})$$

is a family of flag kernel in $\boldsymbol{x}_{(L_k)}$ on $H_{(L_k)}$ with respect to the flag $V_{(L_k)}$ generated by the gradation $H_{(L_k)}$ uniformly bounded in $\boldsymbol{\xi}_{L_k}$.

By the induction hypothesis this means that for any multi-index $\boldsymbol{\alpha}$

$$\left|\partial_{\boldsymbol{\xi}}^{\boldsymbol{\alpha}} \widehat{K}(\boldsymbol{\xi})\right| = \left|\partial_{\boldsymbol{\xi}}^{\boldsymbol{\alpha}} (\mathcal{F}_{H_{(L_k)}} \mathcal{F}_{H_{L_k}} K)(\boldsymbol{\xi})\right| \leq$$
$$C_{\boldsymbol{\alpha}} \|\xi_1\|^{-\|\boldsymbol{\alpha}_{L_k}\|} \|\xi_{k+1}\|^{-\|\boldsymbol{\alpha}_{k+1}\|} \left(\|\xi_{k+1}\| + \|\xi_{k+2}\|\right)^{-\|\boldsymbol{\alpha}_{k+2}\|} \dots \left(\|\xi_{k+1}\| + \dots + \|\xi_d\|\right)^{-\|\boldsymbol{\alpha}_d\|}.$$

Since we are reasoning on the set E_k this gives us the necessary differential inequalities.

To prove the previous statement we write

$$\|\xi_1\|^{\|\boldsymbol{\alpha}_{L_k}\|}\, \partial_{\boldsymbol{\xi}_{L_k}}^{\boldsymbol{\alpha}_{L_k}} (\mathcal{F}_{H_{L_k}} K)(\boldsymbol{\xi}_{L_k}, \boldsymbol{x}_{(L_k)}) = \int_{H_{L_k}} (-i\,\|\xi_1\|\,\boldsymbol{x}_{(L_k)})^{\|\boldsymbol{\alpha}_{L_k}\|} K(\boldsymbol{x}) e^{-i\boldsymbol{x}_{L_k}\boldsymbol{\xi}_{L_k}} \mathrm{d}\boldsymbol{x}_{L_k} =$$
$$\sum_{J\subset L_k} \int_{H_{L_k}} \prod_{j\in J} (-i\,\|\xi_1\|\,\boldsymbol{x}_j)^{\|\boldsymbol{\alpha}_j\|} (1-\eta_j)(\|\xi_1\|\,x_j) e^{-ix_j\xi_j}$$
$$\prod_{l\in L_k\setminus J} (-i\,\|\xi_1\|\,\boldsymbol{x}_l)^{\|\boldsymbol{\alpha}_l\|} \eta_l(\|\xi_1\|\,x_l) e^{-ix_l\xi_l}\, K(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}_{L_k}$$

where $\boldsymbol{\eta}$ is a product cutoff function on H_{L_k}. Notice that $(-ix_l)^{\|\boldsymbol{\alpha}_l\|}\eta_l(\boldsymbol{x}_l)e^{-i\boldsymbol{x}_l(\|\xi_1\|^{-1}\cdot\xi_l)}$ is a normalized bump function on H_l. For each $j \in J$ we choose a direction ξ_{j,a_j} so that $\|\xi_j\| \approx |\xi_{j,a_j}|^{\lambda^{-1}_{j,a_j}}$ and we integrate by parts along x_{j,a_j} a sufficiently large amount of times and we reason as in Theorem 2.5.2.

The proof of the converse follows a similar technique together with the cancellation estimates like the ones carried out in the proof of Theorem 2.5.2. □

It is easy to see that the flag multipliers do not depend on the choice of the gradation for the flag $(\widehat{V}_k)$. Because of this flag kernels also do not depend on the gradation but only on the flag.

Corollary 3.3.5 (Flag kernels do not depend on the specific gradation)**.**
Let $(V_k)_{k\in\{0,\dots,d\}}$ be a flag and let $(H_k)_{k\in\{1,\dots,d\}}$ and $(H'_k)_{k\in\{1,\dots,d\}}$ be two gradations adapted to that flag. Then the two classes $\mathcal{FK}_{(H_k)}$ and $\mathcal{FK}_{(H'_k)}$ given by Definition 3.1.1 coincide and have the same topology.

3.4 Dyadic decompositions

3.4.1 Flag kernels and product kernels

As mentioned before, product kernels are a test case for multi-parameter singular integrals. We have also seen that flag kernels with adapted to a flag $(V_k)_{k\in\{0,\dots,d\}}$ are a subclass of product kernels adapted to any given gradation corresponding to (V_k).

The class of flag kernels adapted to a given flag is strictly smaller than the class of corresponding product kernels. As a matter of fact, flag kernels can be thought of as product kernels with flag-localized supports. This is best illustrated by the following remark.

Proposition 3.4.1 (Product kernels supported near flags)**.**
Suppose that $K \in \mathcal{PK}(\mathbf{0})$ is a product kernel with respect to a product decomposition $\{H_k\}_{k\in\{1,\dots,d\}}$ with arbitrary order $\boldsymbol{\nu}$. If the support of K lies in a flag segment

$$\{\boldsymbol{x} \mid | \, \|x_1\| > C_2 \|x_2\| > \ldots > C_d \|x_d\| \}$$

for some positive constants $C_2, \dots, C_d > 0$ then $K \in \mathcal{FK}_{(H_k)}$ with respect to the flag (V_k) generated by the gradation (H_k). Furthermore the inclusion of the class of product kernels $\mathcal{PK}(\boldsymbol{\nu})$ supported such a flag segment with fixed constants C_k into $\mathcal{FK}(\boldsymbol{\nu})$ is continuous.

Proof. On the flag segment above the quantities $\mathrm{d}(\boldsymbol{x}, V_k) \approx \|x_{k+1}\| + \ldots + \|x_d\|$ are bounded from above and from below up to a constant by $\|x_{k+1}\|$. As a consequence size conditions (2.3.3) on K imply the flag size condition (3.1.1). The cancellation conditions (2.3.4) and (3.1.2) coincide. □

However, even though the class of flag kernels is, per se, smaller than the corresponding class of product kernels, we have an inverse property stating that product kernel can be decomposed into flag kernels, although adapted to different flags. This corresponds to decomposing $\mathbb{R}^N$ into different regions close to different flags.

Theorem 3.4.2 (Decomposition of product kernels into flag kernels)**.**
Let $K(\boldsymbol{x}) \in \mathcal{PK}(\mathbf{0})$. Then K can be represented as a sum of flag kernels with respect to the flag $\{V^\sigma_k\}_{k\in\{0,\dots,d\}}$ such that $V^\sigma_j = \oplus^1_{k=0} H_{\sigma(k)}$ where $\sigma \in S_d$ is a permutation of the indexes $\{1, \dots, d\}$.

Proof. Using Theorem 2.7.4 we write a space localized dyadic decomposition for K. Now write the sum (2.7.1) in the following way:

$$K(\boldsymbol{x}) = \sum_{\sigma \in S_d} \sum_{\boldsymbol{i} \in I_\sigma \subset Z^d} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}).$$

We choose the subsets $I_\sigma \subset \mathbb{Z}^d$ to be a partition of $\mathbb{Z}^d$ so that if $\boldsymbol{i} \in I_\sigma$ then $i_{\sigma(1)} \geq \ldots \geq i_{\sigma(d)}$. Setting

$$K_\sigma(\boldsymbol{x}) \stackrel{\text{def}}{=} \sum_{\boldsymbol{i} \in J_\sigma} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}).$$

is a flag kernel with respect to the flag $(V_k^\sigma)_{k \in \{0,\ldots,d\}}$. As a matter of fact the inner sums converge to product kernels by Theorem 2.7.2. Since the $\varphi_{\boldsymbol{i}}$ are localized on product coronas (not a crucial assumptions, they might have been product balls), by Proposition 3.4.1 K_σ is a flag kernel of the needed type. □

In Section 2.8.1 we have shown that product kernels contain all product kernels adapted to coarser product decomposition of $\mathbb{R}^N$. We will show that a similar remark holds for flag kernels. As a matter of fact the proofs of some of the properties ahead depend on being able to write down flag kernels in a very specific manner up to correction terms that are flag kernels adapted to a coarser flags.

Definition 3.4.3 (Coarser and finer flags).
Consider two flags $(\widetilde{V}_j)_{j \in \{0,\ldots,\widetilde{d}\}}$ and $(V_k)_{k \in \{0,\ldots,d\}}$ on $\mathbb{R}^N$. The flag $(V_k)_{k \in \{1,\ldots,d\}}$ is coarser than the flag $(\widetilde{V}_j)_{j \in \{1,\ldots,\widetilde{d}\}}$ if $\{V_k\} \subset \{\widetilde{V}_j\}$. In particular for every $k \in \{1,\ldots,d\}$ there exists an index $s_k \in \{1,\ldots,\widetilde{d}\}$ such that $V_k = \widetilde{V}_{s_k}$. If $(H_k)_{k \in \{1,\ldots,d\}}$ and $(\widetilde{H}_j)_{j \in \{1,\ldots,\widetilde{d}\}}$ are two gradations associated respectively with the flags (V_k) and $(\widetilde{V}_k)$, then (H_k) is coarser than $(\widetilde{H}_j)$.

Vice-versa the flag $(\widetilde{V}_j)_{j \in \{1,\ldots,\widetilde{d}\}}$ is said to be finer than $(V_k)_{k \in \{1,\ldots,d\}}$.

Remark 3.4.4 (Notation for coarser flags).
Let $(\widetilde{V}_k)_{\widetilde{k} \in \{1,\ldots,\widetilde{d}\}}$ be a flag on $\mathbb{R}^N$ and let $(V_k)_{k \in \{1,\ldots,d\}}$ be a coarser flag. There is a choice of a partition $\{1,\ldots,\widetilde{d}\} = J_1 \cup \ldots \cup J_d$ if $k < l$ and if $a \in J_k$ and $b \in J_l$ then $a < b$. For $k \in \{1,\ldots,d\}$ the gradation (H_k) defined by

$$H_k = \widetilde{H}_{J_k}$$

is a gradation for the flag (V_k). Vice-versa any partition satisfying the above properties defines a coarser gradation and thus a coarser flag.

Proposition 3.4.5 (Kernels adapted to coarser flags).
Let $(\widetilde{V}_j)_{j \in \{1,\ldots,\widetilde{d}\}}$ be a flag and let $(V_k)_{k \in \{1,\ldots,d\}}$ be a flag coarser than $(\widetilde{V}_j)$. Then a flag kernel $K \in \mathcal{FK}_{(V_k)}$ is also a flag kernel in $\mathcal{FK}_{(\widetilde{V}_j)}$. The inclusion

$$\iota : \mathcal{FK}_{(V_k)} \to \mathcal{FK}_{(\widetilde{V}_j)}$$

is continuous.

The same holds for flag multipliers adapted to the respective dual flags. The inclusion mapping

$$\iota : \widehat{\mathcal{FK}}_{(\widehat{V}_k)} \to \widehat{\mathcal{FK}}_{\left(\widehat{\widetilde{V}_j}\right)}$$

is continuous.

Proof. The proposition is trivial for flag multipliers. As a matter of fact it is sufficient to see that the size conditions on flag multipliers (3.3.1) for a coarser dual flag $(\widehat{V}_k)$ imply the ones relative to the finer flag $\left(\widehat{\widetilde{V}}_k\right)$. Using the Fourier transform we get that the inclusion $\iota_{\mathcal{FK}}$ of flag kernels is given by the composition on the continuous maps

$$\iota_{\mathcal{FK}} = \mathcal{F} \circ \iota_{\widehat{\mathcal{FK}}} \circ \mathcal{F}^{-1}$$

where $\iota_{\widehat{\mathcal{FK}}}$ is the inclusion between multipliers. □

3.4.2 Dyadic decomposition for flag kernels

Using the previous characterization of the Fourier transforms of flag kernels we may now proceed to establish properties of dyadic decompositions for flag kernels. In Theorem 3.4.2 we showed how to decompose a product kernels into flag kernels adapted to different flags by separating the dyadic decomposition of a product kernel into separate parts with ordered scale indexes $\boldsymbol{i}$. The next results shows that all flag kernels are essentially of such nature. We begin with a result about when a dyadic sum converges to a flag kernel.

Proposition 3.4.6 (Sufficient conditions for convergence of dyadic flag series)**.**
Consider the flag $(V_k)_{k\in\{1,\dots,d\}}$ on $\mathbb{R}^N$ and an associated gradation $(H_k)_{k\in\{1,\dots,d\}}$. Let

$$\{\varphi_{\boldsymbol{i}}\}_{\substack{\boldsymbol{i}\in\mathbb{Z}^d \\ i_1\geq\ldots\geq i_d}} \subset S(\mathbb{R}^N)$$

be a multi-parameter family of uniformly bounded Schwartz functions such that the cancellation conditions hold:

$$\int_{H_k} \varphi_{\boldsymbol{i}}(\boldsymbol{x})\mathrm{d}x_k = 0 \qquad \forall k \in \{1,\dots,d\}.$$

The dyadic sum

$$K(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{i}\in\mathbb{Z}^d \\ i_1\geq\ldots\geq i_d}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}) \tag{3.4.1}$$

has all (finite and infinite) partial sums uniformly bounded in $\mathcal{FK}$ and it weak- converges to a kernel $K \in \mathcal{FK}$.*

Proof. The proof is similar to Theorem 2.7.2 for order $\mathbf{0}$ kernels. The constraints on the indexes of summation give the necessary size estimates.

The Fourier transform of the dyadic sum is

$$\sum_{\substack{\boldsymbol{i}\in\mathbb{Z}^d \\ i_1\leq\ldots\leq i_d}} \widehat{\varphi}_{-i}(2^{-\boldsymbol{i}}\cdot\boldsymbol{\xi}).$$

For any finite partial sum we can write

$$\partial_{\boldsymbol{\xi}}^{\boldsymbol{\alpha}} \sum_{\substack{\boldsymbol{i}\in\mathbb{Z}^d \\ i_1\leq\ldots\leq i_d}} \widehat{\varphi}_{-i}(2^{-\boldsymbol{i}}\cdot\boldsymbol{\xi}).$$

As in Theorem 2.7.2 we have, due to the cancellation conditions on $\varphi_{\boldsymbol{i}}$, that

$$\widehat{\varphi}_{\boldsymbol{i}}(\boldsymbol{\xi}) = 0 \text{ when } \xi_k = 0.$$

But this means that for all $\varphi_{\boldsymbol{i}}$ the following estimates hold uniformly:

$$\left|\partial^{\boldsymbol{\beta}}_{\boldsymbol{\xi}}\varphi_{\boldsymbol{i}}(\boldsymbol{\xi})\right| < C_{\boldsymbol{\beta}}\frac{1\wedge\|\xi_1\|^{1-\|\beta_1\|}}{(1+\|\xi_1\|)^{Q_1}}\cdots\frac{1\wedge\|\xi_d\|^{1-\|\beta_d\|}}{(1+\|\xi_d\|)^{Q_d}}$$

for any fixed, arbitrarily large, $Q_1,\dots,Q_d\in\mathbb{N}$. We will now prove by induction on the number of product factors d that for any $i_0\in\mathbb{Z}$ we have that

$$\left|\partial^{\boldsymbol{\beta}}_{\boldsymbol{\xi}}\sum_{\substack{\boldsymbol{i}\in\mathbb{Z}^d\\ i_0<i_1\leq\ldots\leq i_d}}\widehat{\varphi}_{-\boldsymbol{i}}(2^{\boldsymbol{i}}\cdot\boldsymbol{\xi})\right| < C'_{\boldsymbol{\beta}}\left(2^{i_0}+\|x_1\|\right)^{-\|\beta_1\|}\cdots\left(2^{i_0}+\|x_1\|+\ldots\|x_d\|\right)^{-\|\beta_d\|}.$$

For $d=0$ there is nothing to prove. Suppose that $d\geq 1$ and the above statement holds for any number of product factors up to $d-1$ Using the above estimates together with an argument similar to the one used for size estimates in the proof of theorem 2.7.2 we can write for any finite partial sum

$$\left|\partial^{\boldsymbol{\beta}}_{\boldsymbol{\xi}}\sum_{\substack{\boldsymbol{i}\in\mathbb{Z}^d\\ i_0<i_1\leq\ldots\leq i_d}}\widehat{\varphi}_{-\boldsymbol{i}}(2^{\boldsymbol{i}}\cdot\boldsymbol{\xi})\right| < C_{\boldsymbol{\beta}}\sum_{\substack{\boldsymbol{i}_1\in\mathbb{Z}^d\\ i_0<i_1\leq i_2\ldots\leq i_d}}2^{-i_1|\beta_1|}\frac{1\wedge\left\|2^{-i_1}\cdot\xi_1\right\|^{1-\|\beta_1\|}}{(1+\left\|2^{-i_1}\cdot\xi_1\right\|)^{Q_1}}$$

$$2^{-\boldsymbol{i}_{(1)}\|\boldsymbol{\beta}_{(1)}\|}\frac{1\wedge\left\|2^{-i_2}\cdot\xi_2\right\|^{1-\|\beta_1\|}}{(1+\left\|2^{-i_2}\cdot\xi_2\right\|)^{Q_1}}\cdots\frac{1\wedge\left\|2^{-i_d}\cdot\xi_d\right\|^{1-\|\beta_d\|}}{(1+\left\|2^{-i_d}\cdot\xi_d\right\|)^{Q_d}} <$$

$$C'_{\boldsymbol{\beta}}\sum_{i_1>i_0}2^{-i_1\|\beta_1\|}\frac{1\wedge\left\|2^{-i_1}\cdot\xi_1\right\|^{1-\|\beta_1\|}}{(1+\left\|2^{-i_1}\cdot\xi_1\right\|)^{Q_1}}$$

$$\left(2^{i_1}+\|x_2\|\right)^{-\|\beta_1\|}\cdots\left(2^{i_1}+\|x_2\|+\ldots\|x_d\|\right)^{-\|\beta_d\|}.$$

Supposing $j_1\in\mathbb{Z}$ is such that $2^{j_1}\leq\|\xi_1\|<2^{j_1+1}$ we write $\widetilde{\xi_1}=2^{-j_1}\cdot\xi_1$. Rewrite the last sum as

$$\sum_{i_1>i_0}2^{-i_1\|\beta_1\|}\frac{1\wedge\left\|2^{-i_1}\cdot\xi_1\right\|^{1-\|\beta_1\|}}{(1+\left\|2^{-i_1}\cdot\xi_1\right\|)^{Q_1}}\left(2^{i_1}+\|\xi_2\|\right)^{-\|\beta_1\|}\cdots\left(2^{i_1}+\|\xi_2\|+\ldots\|x_d\|\right)^{-\|\beta_d\|} =$$

$$2^{-j_1\|\beta_1\|}\sum_{i_1>i_0-j_1}2^{-i_1\|\beta_1\|}\frac{1\wedge\left\|2^{-i_1}\cdot\widetilde{\xi_1}\right\|^{1-\|\beta_1\|}}{(1+\left\|2^{-i_1}\cdot\widetilde{\xi_1}\right\|)^{Q_1}}\left(2^{i_1+j_1}+\|\xi_2\|\right)^{-\|\beta_1\|}\cdots$$

$$\cdots\left(2^{i_1+j_1}+\|\xi_2\|+\ldots\|\xi_d\|\right)^{-\|\beta_d\|}.$$

If $i_0>j_1$ then the term

$$\frac{1\wedge\left\|2^{-i_1}\cdot\widetilde{\xi_1}\right\|^{1-\|\beta_1\|}}{(1+\left\|2^{-i_1}\cdot\widetilde{\xi_1}\right\|)^{Q_1}}\left(2^{i_1+j_1}+\|\xi_2\|\right)^{-\|\beta_1\|}\cdots\left(2^{i_1+j_1}+\|\xi_2\|+\ldots\|\xi_d\|\right)^{-\|\beta_d\|} <$$

$$\left(2^{i_0}+\|\xi_2\|\right)^{-\|\beta_1\|}\cdots\left(2^{i_0}+\|\xi_2\|+\ldots\|\xi_d\|\right)^{-\|\beta_d\|}$$

and the series $2^{-j_1\|\beta_1\|}\sum_{i_1>i_0-j_1} 2^{-i_1\|\beta_1\|}$ is bounded by $2^{-i_0\|\beta_1\|}$. Since $i_0 > j_1$ and $2^{j_1} \approx \|\xi_1\|$ this gives the need inequality. If $j_1 \geq i_0$ then we separate the sum into two pieces

$$2^{-j_1\|\beta_1\|} \sum_{i_1>i_0-j_1} 2^{-i_1\|\beta_1\|} \frac{1 \wedge \left\|2^{-i_1}\cdot\widetilde{\xi_1}\right\|^{1-\|\beta_1\|}}{(1+\left\|2^{-i_1}\cdot\widetilde{\xi_1}\right\|)^{Q_1}} \left(2^{i_1+j_1}+\|\xi_2\|\right)^{-\|\beta_1\|}\dots$$
$$\dots\left(2^{i_1+j_1}+\|\xi_2\|+\dots\|\xi_d\|\right)^{-\|\beta_d\|} =$$
$$2^{-j_1\|\beta_1\|} \sum_{i_1\geq 0} 2^{-i_1\|\beta_1\|} \frac{1 \wedge \left\|2^{-i_1}\cdot\widetilde{\xi_1}\right\|^{1-\|\beta_1\|}}{(1+\left\|2^{-i_1}\cdot\widetilde{\xi_1}\right\|)^{Q_1}} \left(2^{i_1+j_1}+\|\xi_2\|\right)^{-\|\beta_1\|}\dots$$
$$\dots\left(2^{i_1+j_1}+\|\xi_2\|+\dots\|\xi_d\|\right)^{-\|\beta_d\|} +$$
$$2^{-j_1\|\beta_1\|} \sum_{0>i_1>i_0-j_1} 2^{-i_1\|\beta_1\|} \frac{1 \wedge \left\|2^{-i_1}\cdot\widetilde{\xi_1}\right\|^{1-\|\beta_1\|}}{(1+\left\|2^{-i_1}\cdot\widetilde{\xi_1}\right\|)^{Q_1}} \left(2^{i_1+j_1}+\|\xi_2\|\right)^{-\|\beta_1\|}\dots$$
$$\dots\left(2^{i_1+j_1}+\|\xi_2\|+\dots\|\xi_d\|\right)^{-\|\beta_d\|}.$$

The first part follows in a similar manner to the case when $i_0 > j_1$. Otherwise we can write

$$2^{-j_1\|\beta_1\|} \sum_{0>i_1>i_0-j_1} 2^{-i_1\|\beta_1\|} \frac{1 \wedge \left\|2^{-i_1}\cdot\widetilde{\xi_1}\right\|^{1-\|\beta_1\|}}{(1+\left\|2^{-i_1}\cdot\widetilde{\xi_1}\right\|)^{Q_1}} \left(2^{i_1+j_1}+\|\xi_2\|\right)^{-\|\beta_1\|}\dots$$
$$\dots\left(2^{i_1+j_1}+\|\xi_2\|+\dots\|\xi_d\|\right)^{-\|\beta_d\|} <$$
$$C\left(2^{i_0}+2^{j_1}\right)^{-\|\beta_1\|}\left(2^{i_0}+2^{j_1}+\|\xi_2\|\right)^{-\|\beta_1\|}\dots\left(2^{i_0}+2^{j_1}+\|\xi_2\|+\dots+\|\xi_d\|\right)^{-\|\beta_d\|}$$
$$\sum_{0>i_1>i_0-j_1} 2^{-i_1\|\beta_1\|} \frac{1 \wedge \left\|2^{-i_1}\cdot\widetilde{\xi_1}\right\|^{1-\|\beta_1\|}}{(1+\left\|2^{-i_1}\cdot\widetilde{\xi_1}\right\|)^{Q_1}} \left(\frac{2^{i_1+j_1}+\|\xi_2\|}{2^{j_1}+\|\xi_2\|}\right)^{-\|\beta_1\|}\dots$$
$$\dots\left(\frac{2^{i_1+j_1}+\|\xi_2\|+\dots\|\xi_d\|}{2^{j_1}+\|\xi_2\|+\dots\|\xi_d\|}\right)^{-\|\beta_d\|} <$$
$$C\left(2^{i_0}+2^{j_1}\right)^{-\|\beta_1\|}\left(2^{i_0}+2^{j_1}+\|\xi_2\|\right)^{-\|\beta_1\|}\dots\left(2^{i_0}+2^{j_1}+\|\xi_2\|+\dots\|\xi_d\|\right)^{-\|\beta_d\|}$$
$$\sum_{0>i_1>i_0-j_1} 2^{-i_1\|\beta\|}(1+\left\|2^{-i_1}\cdot\widetilde{\xi_1}\right\|)^{-Q_1}.$$

Since $\|\xi_1\| > 1$ the last sum is uniformly bounded and this gives proof. □

We also have a converse theorem expressing a flag kernel as a dyadic series. The idea behind this theorem is similar to the one we used in Theorem 3.4.2. We will divide the dyadic decomposition of a flag kernel K seen as a product kernel into pieces with ordered summation indexes. Indexes not ordered naturally with respect to the flag will add up to terms that are flag kernels adapted to coarser flags.

Theorem 3.4.7 (Dyadic decomposition for flag kernels).
Let $K \in \mathcal{FK}$ be a flag kernel relative to the flag $(V_k)_{k\in\{0,\dots,d\}}$. There is a bounded set of Schwartz

functions

$$\{\varphi_{\boldsymbol{i}}\}_{\substack{\boldsymbol{i}\in\mathbb{Z}^d\\ i_1\geq\ldots\geq i_d}} \subset S(\mathbb{R}^N)$$

such that the following decomposition holds:

$$K(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{i}\in\mathbb{Z}^d\\ i_1\geq\ldots\geq i_d}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}) \tag{3.4.2}$$

The sum is intended in the distributional sense and the partial sums are a uniformly bounded family of $\mathcal{FK}$ kernels.

The dyadic functions $\varphi_{\boldsymbol{i}}$ can be chosen to satisfy either of the following two types of conditions. For any multi-index $\boldsymbol{i}$ let the indexes $\{1,\ldots,d\} = J_1(\boldsymbol{i}) \cup \ldots \cup J_{d'(\boldsymbol{i})}(\boldsymbol{i})$ be subdivided in a partition of non-empty sets such that if $k < k'$ and $a \in J_k$ and $b \in J_{k'}$ then $i_a > i_b$ while if $a, b \in J_k$ then $i_a = i_b$.

Space localization *Every function $\varphi_{\boldsymbol{i}}$ is supported on the set where $1/4 \leq \left\|x_{J_k}\right\| \leq 4$ for all $k \in \{1,\ldots,d'(\boldsymbol{i})\}$. Furthermore for any fixed $l \in \mathbb{N}$ the family $\{\varphi_{\boldsymbol{i}}\}$ can be chosen*

$$\int_{H_{J_k}} \boldsymbol{x}^{\alpha}\varphi_{\boldsymbol{i}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}_{J_k} = 0$$

for all $k \in \{1,\ldots,d'(\boldsymbol{i})\}$ and any multi-index $\boldsymbol{\alpha}$ such that $\|\boldsymbol{\alpha}\| \leq l$.

Frequency localization *Every function $\varphi_{\boldsymbol{i}}$ has its Fourier tranform $\widehat{\varphi}_{\boldsymbol{i}}$ supported on the set where $1/2 \leq \left\|x_{J_k}\right\| \leq 2$ for all $k \in \{1,\ldots,d'(\boldsymbol{i})\}$. In particular*

$$\int_{H_{J_k}} \boldsymbol{x}^{\alpha}\varphi_{\boldsymbol{i}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}_{J_k} = 0$$

for all $k \in \{1,\ldots,d\}$ and any multi-index $\boldsymbol{\alpha}$.

In both cases, the decomposition of K is such that

$$K(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{i}\in\mathbb{Z}^d\\ i_1>\ldots>i_d}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}) + T(\boldsymbol{x}). \tag{3.4.3}$$

The family of functions $\{\varphi_{\boldsymbol{i}}\}$ satisfy strong cancellation conditions and in particular they satisfy the conditions allowed for by the Theorem for $J_k = \{k\}$ for all $k \in \{1,\ldots,d\}$. T is a sum of flag kernels adapted to coarser flags.

Proof. Let $K \in \mathcal{FK}$ be a flag kernel. Using Theorem 2.7.4, let us write a dyadic decomposition of the kernel intended as a product kernel of order $\mathbf{0}$. If we aim for space localization then we should use Theorem 2.7.4 directly, otherwise we should apply Theorem 2.7.4 to the Fourier transform. As we did in the proof of Theorem 3.4.2 we can rewrite the dyadic series in the following way:

$$K(\boldsymbol{x}) = \sum_{\sigma\in S_d}\sum_{\boldsymbol{i}\in I_\sigma\subset\mathbb{Z}^d} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}) = \sum_{\sigma\in S_d} K_\sigma(\boldsymbol{x})$$

where the subsets I_σ form a partition of $\mathbb{Z}^d$ such that if $\boldsymbol{i} \in I_\sigma$ then $i_{\sigma(1)} \geq \ldots \geq i_{\sigma(d)}$. We also require that if σ is the trivial permutation and $\boldsymbol{i} \in I_\sigma$ then $i_1 > \ldots > i_d$.

By Proposition 3.4.6, K_σ is a flag kernel adapted to the flag V^σ. However since the supports of all the inner sums are essentially disjoint each term is also a flag kernel adapted to the original flag V. This means that if σ is not the trivial permutation K_σ adapted to a coarser flag. This is true because of the following consideration.

If we have chosen to do a space-localized dyadic decomposition then it follows that K_σ is supported on the set where

$$\left\|x_{\sigma(1)}\right\| \geq C_2 \left\|x_{\sigma(2)}\right\| \geq \ldots \geq C_d \left\|x_{\sigma(d)}\right\|.$$

However if σ is not trivial then there are two indexes $j, j' \in \{1, \ldots, d\}$ such that $j < j'$ but $\sigma^{-1}(j) > \sigma^{-1}(j')$. This means that $\left\|x_{j'}\right\| \approx \left\|x_{j'}\right\| + \left\|x_j\right\|$ and so

$$\left\|x_{j+1}\right\| + \ldots + \left\|x_{j'}\right\| + \ldots + \left\|x_d\right\| \approx \left\|x_j\right\| + \ldots + \left\|x_{j'}\right\| + \ldots + \left\|x_d\right\|.$$

The size conditions (3.1.1) on K_σ become

$$\left|\partial_{\boldsymbol{x}}^{\boldsymbol{\alpha}} K_\sigma(\boldsymbol{x})\right| < C_{\boldsymbol{\alpha}} \left(\left\|x_1\right\| + \ldots + \left\|x_d\right\|\right)^{-q_1 - \left\|\alpha_1\right\|} \ldots$$
$$\left(\left\|x_j\right\| + \ldots + \left\|x_d\right\|\right)^{-(q_j + q_{j+1}) - (\left\|\alpha_j\right\| + \left\|\alpha_{j+1}\right\|)} \left(\left\|x_{j+2}\right\| + \ldots \left\|x_d\right\|\right)^{-q_{j+2} - \left\|\alpha_{j+2}\right\|} \ldots \left\|x_d\right\|^{-q_d - \alpha_d}$$

that are relative to a coarser flag. The cancellation conditions relative to a coarser flag can also easily be checked. If frequency localization was chosen the same argument can be done for the Fourier transform of the dyadic series.

As a consequence we can write down the following decomposition:

$$K(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ i_1 > \ldots > i_d}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}) + T(\boldsymbol{x})$$

where T are flag kernels adapted to coarser flags and $\{\varphi_{\boldsymbol{i}}\}$ is a uniformly bounded family of Schwartz functions that satisfy the conditions of cancellation and localization (space or frequency respectively). We can further decompose the coarser terms dyadically and add up the results.

Any given coarser flag $(V'_{k'})_{k' \in \{1, \ldots, d'\}}$ is associated to a partition $J_1 \cup \ldots \cup J'_d = \{1, \ldots, d\}$ as specified in Remark 3.4.4. Different partitions give different coarser flags. For any index $\boldsymbol{i} \in Z^d$ such that $i_1 \geq \ldots \geq i_d$ there is a unique partition of such type $J_1 \cup \ldots \cup J_{d'(\boldsymbol{i})}$ such that if $a \in J_k$ and $b \in J_l$ with $k < l$ then $i_a > i_b$ while if $k = l$ then $i_a = i_b$. The dyadic decomposition for a kernel adapted to a coarser flag associated to a given partition is given by

$$K(\boldsymbol{x}) = \sum \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}) + T(\boldsymbol{x})$$

where the summation goes over only those indexes $\boldsymbol{i}$ that correspond to the partition. This effectively concludes the proof. □

Because of the above theorem it is easy to see that Proposition 3.4.6 is not optimal. In particular a dyadic sum is allowed not to have cancellation on the terms with "diagonal" summation indexes $\boldsymbol{i}$. These terms account for flags adapted to coarser degree. To incorporate this theory we introduce weak cancellation.

Definition 3.4.8 (Weak cancellation)**.**
We say that a uniformly bounded family of Schwartz functions

$$\{\varphi_{\boldsymbol{i}}\}_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ i_1 \geq \ldots \geq i_d}} \subset S(\mathbb{R}^N)$$

possesses weak cancellation with parameter $\varepsilon > 0$ if

$$\int_{\mathbb{R}^N} \varphi_{\boldsymbol{i}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}_1 = 0 \tag{3.4.4}$$

and for any subset $J \subset \{2, \dots, d\}$ we have that

$$\int_{H_J} \varphi_{\boldsymbol{i}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}_J = \prod_{j \in J} 2^{\varepsilon(i_j - i_{j-1})} \varphi_{\boldsymbol{i},(J)}(\boldsymbol{x}_{(J)}) \tag{3.4.5}$$

with $\varphi_{\boldsymbol{i},(J)}$ a uniformly bounded family of functions in $S(H_{(J)})$.

Remark 3.4.9.
To check that a family $\{\varphi_{\boldsymbol{i}}\}$ possesses weak cancellation it is actually sufficient to check that condition (3.4.5) holds only for subsets J of cardinality 1. As a matter of fact for an arbitrary J we have

$$\prod_{j \in J} 2^{\varepsilon(i_j - i_{j-1})} \geq 2^{|J| \varepsilon \min_{j \in J}(i_j - i_j - 1)}$$

so if the inequality (3.4.5) is satisfied for all $|J| = 1$ with parameter ε we can evaluate the same inequality for $|J| > 1$ by first integrating along $x_{\widetilde{j}}$ with $\widetilde{j}$ such that $i_{\widetilde{j}} - i_{\widetilde{j}-1} = \min_{j \in J}(i_j - i_j - 1)$ and then in the other variables and thus obtaining weak cancellation with parameter $^{\varepsilon}/_{d}$.

We have the following result for dyadic series with weak cancellation.

Theorem 3.4.10 (Dyadic sums with weak cancellation)**.**
Let

$$\{\varphi_{\boldsymbol{i}}\}_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ i_1 \geq \dots \geq i_d}} \subset S(\mathbb{R}^N)$$

be a uniformly bounded family of Schwartz functions with weak cancellation with some parameter $\varepsilon > 0$. Then the dyadic sum

$$K(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ i_1 \geq \dots \geq i_d}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x})$$

converges in the distributional sense to a flag kernel and all the partial sums are bounded in $\mathcal{FK}$.

Proof. Let us take the Fourier transform of the family $\{\varphi_{\boldsymbol{i}}\}$ and by setting $\psi_{\boldsymbol{i}} \stackrel{\text{def}}{=} \widehat{\varphi}_{-\boldsymbol{i}}$ we need to prove that the dyadic sum

$$\sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ i_1 \leq \dots \leq i_d}} \psi_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \boldsymbol{\xi})$$

converges to a flag multiplier.

We will argue by rewriting the sum in such a way that if $i_{k-1} < i_k$ for a certain k then necessarily $\psi_{\boldsymbol{i}}$ is supported away from $\{\boldsymbol{\xi} \mid \xi_k = 0\}$. It will be easy to see that such a sum possesses strong cancellation if seen as a superposition of dyadic sums of flag kernels adapted to possibly different flags.

Weak cancellation of parameter ε, in terms of the Fourier transform, assumes the following form:

$$\psi_{\boldsymbol{i}}(0_1 \oplus \boldsymbol{\xi}_{(1)}) = 0 \tag{3.4.6}$$

and for any $J \subset \{2, \dots, d\}$

$$\psi_{\boldsymbol{i}}(\mathbf{0}_J \oplus \boldsymbol{\xi}_{(J)}) = \prod_{j \in J} 2^{\varepsilon(i_{j-1} - i_j)} \psi_{\boldsymbol{i},(J)}(\boldsymbol{\xi}_{(J)}) \tag{3.4.7}$$

with $\{\psi_{\boldsymbol{i},(J)}\}$ a uniformly bounded family of functions in $S(H_{(J)})$.

We prove that if $\psi_{\boldsymbol{i}}$ is a uniformly bounded family of Schwartz functions that satisfy conditions (3.4.6) and (3.4.7) there exists a way to rewrite the dyadic sum

$$\sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ i_1 \leq \ldots \leq i_d}} \psi_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \boldsymbol{\xi}) = \sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ i_1 \leq \ldots \leq i_d}} \widetilde{\psi}_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \boldsymbol{\xi})$$

so that $\{\widetilde{\psi}_{\boldsymbol{i}}\}$ is a uniformly bounded family of functions such that if for some $j \in \{2, \dots, d\}$ we have $i_{j-1} < i_j$ then $\widetilde{\psi}_{\boldsymbol{i}}$ is supported on the set $\|\xi_j\| > 1/2$ and $\{\widetilde{\psi}_{\boldsymbol{i}}\}$ satisfies (3.4.6). Formally this is proven for all finite partial sums and then one passes to the limit.

We proceed by iteration. Suppose that the family $\{\psi_{\boldsymbol{i}}\}$ is uniformly bounded in $S(\mathbb{R}^N)$, that it satisfies conditions (3.4.6) and (3.4.7), and there exists a $c \in \{2, \dots, d\}$ that if $i_j > i_{j-1}$ and $c < j$ then $\psi_{\boldsymbol{i}}$ is supported on $\|\xi_j\| > 1/2$. Then there is a way to rewrite the sum in terms of a uniformly bounded family $\{\widetilde{\psi}_{\boldsymbol{i}}\}$ such that it continues to satisfy conditions (3.4.6) and (3.4.7), and also it satisfies the above statement with $c' = c - 1$. Since the condition for $c = d$ is trivial this procedure will allow us to prove the statement for $c = 1$.

Let η_c be a cutoff function on H_c. We write

$$\begin{aligned}\psi_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \boldsymbol{\xi}) = \psi_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \boldsymbol{\xi}) \eta_c(2^{-i_{c-1}} \cdot \xi_c) + \\ \sum_{\substack{i'_c \in \mathbb{Z} \\ i_{c-1} < i'_c < i_c}} \psi_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \boldsymbol{\xi}) \left(\eta_c(2^{-i'_c} \cdot \xi_c) - \eta_c(2^{-i'_c + 1} \cdot \xi_c) \right) + \\ \psi_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \boldsymbol{\xi})(1 - \eta_c)(2^{-i_c} \cdot \xi_c).\end{aligned}$$

Let $\boldsymbol{i}' \in \mathbb{Z}^d$ with $i'_1 \leq \ldots \leq i'_d$. We define the new family $\{\widetilde{\psi}_{\boldsymbol{i}'}\}$ in the following way. For $i'_{c-1} = i'_c$ we set

$$\widetilde{\psi}_{\boldsymbol{i}'}(\boldsymbol{\xi}) = \psi_{\boldsymbol{i}'}(\boldsymbol{\xi}) + \sum_{\substack{\boldsymbol{i} \in Z^d \\ i_k = i'_k \; k \neq c \\ i'_c < i_c \leq i'_{c+1}}} \psi_{\boldsymbol{i}}(\boldsymbol{\xi}_{(c)} \oplus 2^{i'_c - i_c} \cdot \xi_c) \eta_c(\xi_c)$$

while if $i'_{c-1} < i'_c$ we set

$$\widetilde{\psi}_{\boldsymbol{i}'}(\boldsymbol{\xi}) = \left(\eta_c(\xi_c) - \eta_c(2 \cdot \xi_c) \right) \sum_{\substack{\boldsymbol{i} \in Z^d \\ i_k = i'_k \; k \neq c \\ i'_c < i_c < i'_{c+1}}} \psi_{\boldsymbol{i}}(\boldsymbol{\xi}_{(c)} \oplus 2^{i'_c - i_c} \cdot \xi_c) + \psi_{\boldsymbol{i}'}(\boldsymbol{\xi})(1 - \eta_c)(\xi_c).$$

The series

$$\sum_{\substack{\boldsymbol{i} \in Z^d \\ i_k = i'_k \; k \neq c \\ i'_c < i_c < i'_{c+1}}} \psi_{\boldsymbol{i}}(\boldsymbol{\xi}_{(c)} \oplus 2^{i'_c - i_c} \cdot \xi_c)$$

converges to a function in $S(\mathbb{R}^N)$ since

$$\psi_{\boldsymbol{i}}(\boldsymbol{\xi}_{(c)} \oplus 2^{i_c - i'_c} \cdot \xi_c) = \psi_{\boldsymbol{i}}(\boldsymbol{\xi}_{(c)} \oplus 0_c) + 2^{i'_c - i_c} \cdot \xi_c \int_{s=0}^{1} \partial_{\xi_c} \psi_{\boldsymbol{i}}(\boldsymbol{\xi}_{(c)} \oplus s \cdot 2^{i'_c - i_c} \cdot \xi_c) \mathrm{d}s$$

because of the weak cancellation conditions we have that

$$2^{i_c - i'_c} \psi_{\boldsymbol{i}}(\boldsymbol{\xi}_{(c)} \oplus 0_c)$$

is uniformly bounded as well as

$$\xi_c \int_{s=0}^{1} \partial_{\xi_c} \psi_{\boldsymbol{i}}(\boldsymbol{\xi}_{(c)} \oplus s \cdot 2^{i'_c - i_c} \cdot \xi_c) \mathrm{d}s.$$

It is simple to verify that the functiond thus possess the required property i.e. the family $\{\widetilde{\psi}_{\boldsymbol{i}}\}$ possesses weak cancellation and if $i_{k-1} < i_k$ with $c' < k$ then $\psi_{\boldsymbol{i}}$ is supported on $\|\xi_k\| > 1/2$.

It is also true that

$$\sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ i_1 \leq \ldots \leq i_d}} \psi_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \boldsymbol{\xi}) = \sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ i_1 \leq \ldots \leq i_d}} \widetilde{\psi}_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \boldsymbol{\xi}).$$

For any multi-index $\boldsymbol{i}$ let the indexes $\{1, \ldots, d\} = J_1(\boldsymbol{i}) \cup \ldots \cup J_{d'(\boldsymbol{i})}(\boldsymbol{i})$ be subdivided in an partition of indexes such that if $k < l$ and $a \in J_k$ and $b \in J_l$ then $i'_a > i'_b$ while if $a, b \in J_k$ then $i_a = i_b$. The subsets J_k must be non-empty. There is only one the partition for any given $\boldsymbol{i}$ that satisfies the above conditions. For this reason we can split the sum

$$\sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ i_1 \leq \ldots \leq i_d}} \widetilde{\psi}_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \boldsymbol{\xi}).$$

into a sum over all partitions described above of the dyadic sums over those indexes that give rise to a given partition. It is easy to see that for every partition $\{1, \ldots, d\} = J_1(\boldsymbol{i}) \cup \ldots \cup J_{d'(\boldsymbol{i})}(\boldsymbol{i})$ the dyadic sum over those indexes $\boldsymbol{i}$ that give rise to such a partition is a dyadic sum with strong cancellation as required by Proposition 3.4.6 with respect to the flag associated with the gradation $\{H_{J_k}\}_{k \in \{1, \ldots, d'(\boldsymbol{i})\}}$. □

Remark 3.4.11 (Index restrictions for flag kernels)**.**
In this section we have obtained results about the possibility of writing dyadic decompositions and about when dyadic sums converge to flag kernels. In all the above theorems the summation goes over indexes $\boldsymbol{i} \in \mathbb{Z}^d$ *such that* $i_1 \leq \ldots \leq i_d$*. However there is some flexibility on the restriction of summation indexes. For example the above results can be reformulated in the same manner if the summation takes place over indexes* $\boldsymbol{i} \in \mathbb{Z}^d$ *such that* $i_1 + \delta_1 \leq i_2 + \delta_2 \leq \ldots \leq i_d + \delta_d$ *where* $\boldsymbol{\delta} \in \mathbb{Z}^d$ *are some fixed displacement indexes. The same qualitatively different behavior would be reserved for dyadic terms* $\varphi_{\boldsymbol{i}}$ *with* $\boldsymbol{i}$ *"close to the diagonal". While this remark may seem superfluous it will be sometimes useful to suppose these kinds of restrictions on the indexes of summation for ease of notation in some of the subsequent proofs.*

3.4.3 Flag kernels with compact support

Flag kernels with compact support have some additional useful properties. In particular they admit a dyadic decomposition that reflects the geometry of their support. First we state a generalization of the property of convergence of dyadic sums with weak cancellation to the case where the indexes of summation are bounded from above.

Theorem 3.4.12 (Convergence of dyadic series with bounded indexes)**.**
Given a bound on the indexes $m \in \mathbb{Z}$ let

$$\{\varphi_{\boldsymbol{i}}\}_{\substack{\boldsymbol{i}\in\mathbb{Z}^d\\ m\geq i_1\geq\ldots\geq i_d}}$$

be a bounded family of Schwartz functions. Let the functions $\{\varphi_{\boldsymbol{i}}\}$ satisfy property (3.4.5) *with some parameter $\varepsilon > 0$ and a relaxed version of* (3.4.4) *i.e.*

$$\int_{H_1} \varphi_{\boldsymbol{i}}(\boldsymbol{x})\mathrm{d}x_1 = 2^{-\varepsilon(m-i_1)}\varphi_{\boldsymbol{i},(1)}(\boldsymbol{x}_{(1)})$$

where $\{\varphi_{\boldsymbol{i},(1)}\}$ is a bounded family of $S(H_{(1)})$.

Then the dyadic sum

$$\sum_{\substack{\boldsymbol{i}\in\mathbb{Z}^d\\ m\leq i_1\leq\ldots\leq i_d}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x})$$

converges to a flag kernel in the distributional sense and all the partial sums are a uniformly bounded family of product kernels.

The dyadic series possesses a continuity property in the sense that for any neighborhood U of 0 in $\mathcal{PK}(\boldsymbol{\nu})$ there is a neighborhood V_U of 0 in $S(\mathbb{R}^N)$ such that if $\{\varphi_{\boldsymbol{i}}\} \subset V_U$ then all the partial sums and the whole series are in U.

Proof. The proof is very similar to the one of Theorem 3.4.10. Let us take the Fourier transform of the family $\{\varphi_{\boldsymbol{i}}\}$ and by setting $\psi_{\boldsymbol{i}} \overset{\mathrm{def}}{=} \widehat{\varphi}_{-\boldsymbol{i}}$ we need to prove that the dyadic sum

$$\sum_{\substack{\boldsymbol{i}\in\mathbb{Z}^d\\ -m\leq i_1\leq\ldots\leq i_d}} \psi_{\boldsymbol{i}}(2^{-\boldsymbol{i}}\cdot\boldsymbol{\xi})$$

converges to a flag multiplier. The conditions on the family $\{\varphi_{\boldsymbol{i}}\}$ imply the following conditions on the Fourier transforms

$$\psi_{\boldsymbol{i}}(0_1 \oplus \boldsymbol{\xi}_{(1)}) = 2^{-\varepsilon i_1 - m}\psi_{\boldsymbol{i},(1)}(\boldsymbol{\xi}_{(1)}) \tag{3.4.8}$$

and for any $J \subset \{2,\ldots,d\}$

$$\psi_{\boldsymbol{i}}(\boldsymbol{0}_J \oplus \boldsymbol{\xi}_{(J)}) = \prod_{j\in J} 2^{\varepsilon(i_{j-1}-i_j)}\psi_{\boldsymbol{i},(J)}(\boldsymbol{\xi}_{(J)}) \tag{3.4.9}$$

with $\{\psi_{\boldsymbol{i},(J)}\}$ and $\{\psi_{\boldsymbol{i},(1)}\}$ uniformly bounded family of functions in $S(H_{(J)})$.

We argue once again by formally setting $i_0 = -m$ and rewriting the sum in such a way that if $i_{k-1} < i_k$ for a certain $k \in \{1,\ldots,d\}$ then necessarily $\psi_{\boldsymbol{i}}$ is supported away from $\{\boldsymbol{\xi} \mid \xi_k = 0\}$. It will be easy to see that for $i_1 > -m$ such a sum possesses strong cancellation if seen as a superposition of dyadic sums of flag kernels adapted to possibly different flags. The case $i_1 = -m$ will be dealt with separately.

The method of rewriting the sum is identical to the one used in the proof of Theorem 3.4.10. Furthermore that procedure conserves condition (3.4.9) and (3.4.8). Applying the procedure used in that proof we can suppose without loss of generality that $\{\psi_{\boldsymbol{i}}\}$ is a uniformly bounded family of functions such that if for some $j \in \{2,\ldots,d\}$ we have $i_{j-1} < i_j$ then $\psi_{\boldsymbol{i}}$ is supported on the set $\|\xi_j\| > 1/2$ and $\{\psi_{\boldsymbol{i}}\}$ satisfies (3.4.8).

Let η_1 be a cutoff function on H_1. We write

$$\psi_{\boldsymbol{i}}(2^{-\boldsymbol{i}}\cdot\boldsymbol{\xi}) = \psi_{\boldsymbol{i}}(2^{-\boldsymbol{i}}\cdot\boldsymbol{\xi})\eta_1(2^m\cdot\xi_c) + \sum_{\substack{i_1'\in\mathbb{Z}\\ -m<i_1'<i_1}} \psi_{\boldsymbol{i}}(2^{-\boldsymbol{i}}\cdot\boldsymbol{\xi})\left(\eta_1(2^{-i_1'}\cdot\xi_c) - \eta_c(2^{-i_1'+1}\cdot\xi_1)\right) +$$

$$\psi_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \boldsymbol{\xi})(1-\eta_1)(2^{-i_1} \cdot \xi_1).$$

Let $\boldsymbol{i'} \in \mathbb{Z}^d$ with $i'_1 \leq \ldots \leq i'_d$. We define the new family $\{\widetilde{\psi}_{\boldsymbol{i'}}\}$ in the following way. For $-m = i'_1$ we set

$$\widetilde{\psi}_{\boldsymbol{i'}}(\boldsymbol{\xi}) = \psi_{\boldsymbol{i'}}(\boldsymbol{\xi}) + \sum_{\substack{\boldsymbol{i} \in Z^d \\ i_k = i'_k \; k \neq 1 \\ -m < i_1 \leq i'_2}} \psi_{\boldsymbol{i}}(2^{i'_1 - i_1} \cdot \xi_1 \oplus \boldsymbol{\xi}_{(1)}) \eta_1(\xi_1)$$

while if $-m < i'_1$ we set

$$\widetilde{\psi}_{\boldsymbol{i'}}(\boldsymbol{\xi}) = \big(\eta_1(\xi_1) - \eta_c(2 \cdot \xi_1)\big) \sum_{\substack{\boldsymbol{i} \in Z^d \\ i_k = i'_k \; k \neq 1 \\ i'_1 < i_1 < i'_2}} \psi_{\boldsymbol{i}}(2^{i'_1 - i_1} \cdot \xi_1 \oplus \boldsymbol{\xi}_{(1)}) + \psi_{\boldsymbol{i'}}(\boldsymbol{\xi})(1-\eta_1)(\xi_1).$$

The series

$$\sum_{\substack{\boldsymbol{i} \in Z^d \\ i_k = i'_k \; k \neq 1 \\ i'_1 < i_1 < i'_2}} \psi_{\boldsymbol{i}}(2^{i'_1 - i_1} \cdot \xi_1 \oplus \boldsymbol{\xi}_{(1)})$$

converges to a function in $S(\mathbb{R}^N)$ since

$$\psi_{\boldsymbol{i}}(2^{i_1 - i'_1} \cdot \xi_1 \oplus \boldsymbol{\xi}_{(1)}) = \psi_{\boldsymbol{i}}(0_1 \oplus \boldsymbol{\xi}_{(1)}) + 2^{i'_1 - i_1} \cdot \xi_1 \int_{s=0}^{1} \partial_{\xi_1} \psi_{\boldsymbol{i}}(s \cdot 2^{i'_1 - i_1} \cdot \xi_1 \oplus \boldsymbol{\xi}_{(1)}) \mathrm{d}s$$

because of the weak cancellation conditions we have that

$$2^{i_1 - i'_1} \psi_{\boldsymbol{i}}(0_1 \oplus \boldsymbol{\xi}_{(1)})$$

is uniformly bounded as well as

$$\xi_1 \int_{s=0}^{1} \partial_{\xi_1} \psi_{\boldsymbol{i}}(s \cdot 2^{i'_1 - i_1} \cdot \xi_1 \oplus \boldsymbol{\xi}_{(1)}) \mathrm{d}s.$$

It is simple to verify that the functions thus possess the required properties. The family $\{\widetilde{\psi}_{\boldsymbol{i}}\}$ is such that setting formally $i_0 = -m$ we have if $i_{k-1} < i_k$ with $k \in \{1, \ldots, d\}$ then $\widetilde{\psi}_{\boldsymbol{i}}$ is supported on $\|\xi_k\| > 1/2$.

It is also true that

$$\sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ -m \leq i_1 \leq \ldots \leq i_d}} \psi_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \boldsymbol{\xi}) = \sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ -m \leq i_1 \leq \ldots \leq i_d}} \widetilde{\psi}_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \boldsymbol{\xi}).$$

The above sum converges to a flag multiplier because for $i_1 > -m$ the family $\{\widetilde{\psi}_{\boldsymbol{i}}\}$ satisfies the conditions of Proposition 3.4.6. For $i_1 = m$ one can use an estimate like in the proof of Proposition 3.4.6 taking in account that summation does not occur in the index i_1.

□

Theorem 3.4.13 (Localized dyadic decomposition for flag kernels of compact support)**.** *For any $K \in \mathcal{FK}$ such that some $m \in \mathbb{Z}$ we have that*

$$\operatorname{spt} K \subset \left\{ \boldsymbol{x} \,\middle|\, \|x_k\| < 2^m \;\; \forall k \in \{1, \ldots, d\} \right\}.$$

There is a uniformly bounded set of Schwartz functions

$$\{\varphi_{\boldsymbol{i}}\}_{\substack{\boldsymbol{i}\in\mathbb{Z}^d\\ m-2\le i_1\le\ldots\le i_d}}$$

such that the dyadic decomposition

$$K(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{i}\in\mathbb{Z}^d\\ m-2\ge i_1\ge\ldots\ge i_d}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}) \tag{3.4.10}$$

holds. The series converges in the weak (distributional) sense and the partial sums are a uniformly bounded family of $\mathcal{FK}$.

For any multi-index $\boldsymbol{i}$ let the indexes $\{1,\ldots,d\} = J_1(\boldsymbol{i}) \cup \ldots \cup J_{d'(\boldsymbol{i})}(\boldsymbol{i})$ be subdivided in a partition of non-empty sets such that if $k < k'$, $a \in J_k$ and $b \in J_{k'}$ then $i_a > i_b$ while if $a, b \in J_k$ then $i_a = i_b$.

Every function $\varphi_{\boldsymbol{i}}$ is supported on the set where $1/4 \le \|x_{J_k}\| \le 4$ for all $k \in \{1,\ldots,d'(\boldsymbol{i})\}$. Furthermore for any fixed $l \in \mathbb{N}$ the family $\{\varphi_{\boldsymbol{i}}\}$ can be chosen so that

- *If $i_1 < m-2$ then*

  $$\int_{H_{J_k}} \boldsymbol{x}^{\alpha}\varphi_{\boldsymbol{i}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}_{J_k} = 0$$

  for all $k \in \{1,\ldots,d'(\boldsymbol{i})\}$ and any multi-index $\boldsymbol{\alpha}$ such that $\|\boldsymbol{\alpha}\| \le l$.

- *If $i_1 = m-2$ then*

  $$\int_{H_{J_k}} \boldsymbol{x}^{\alpha}\varphi_{\boldsymbol{i}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}_{J_k} = 0$$

  for all $k \in \{1,\ldots,d'(\boldsymbol{i})\}$ and any multi-index $\boldsymbol{\alpha}$ such that $\|\boldsymbol{\alpha}\| \le l$.

The decomposition possesses a continuity property in the sense that for any neighborhood V of 0 in $S(\mathbb{R}^N)$ there is a neighborhood U_V of 0 in $\mathcal{FK}$ such that if $K \in U_V$ and K its support bounded as in the hypothesis then K admits a dyadic decomposition with $\{\varphi_{\boldsymbol{i}}\} \subset U$.

The proof of this Theorem is identical to the one for product kernels (Theorem 2.8.6).

3.5 Operators associated with flag kernels

We now make some remarks about the convolution operators associated with flag kernels. Since flag kernels form a special class of product kernels we have the following results.

Theorem 3.5.1 (L^p boundedness of operators associated to flag kernels)**.**
Let $K \in \mathcal{FK}$ be a flag kernel then the convolution operator

$$\mathcal{T}\varphi = \varphi * K$$

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

Theorem 3.5.2 (Convolution algebras)**.**
$\mathcal{FK}$ is a continuous convolution algebra. Given any two kernels K_1 and K_2 with dyadic decomposition

$$K_1(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{i}\in\mathbb{Z}^d\\ i_1\ge\ldots\ge i_d}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}) \qquad\qquad K_2(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{i}'\in\mathbb{Z}^d\\ i'_1\ge\ldots\ge i'_d}} \psi_{\boldsymbol{i}'}^{(2^{\boldsymbol{i}'})}(\boldsymbol{x})$$

with the families $\{\varphi_{\boldsymbol{i}}\}$ and $\{\varphi_{\boldsymbol{i}'}\}$ satisfying the conditions of Theorem 3.4.10 then kernel $K_1 * K_2$ is given by the weak- limit of the dyadic series*

$$K_1 * K_2(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{j}\in\mathbb{Z}^d \\ j_1\geq\ldots\geq j_d}} \sum_{\substack{\boldsymbol{i}\in\mathbb{Z}^d \\ \boldsymbol{i}'\in\mathbb{Z}^d \\ j_k=i_k\vee i'_k}} \psi_{\boldsymbol{i}'}^{(2^{\boldsymbol{i}'})} * \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}).$$

This last Theorem is a consequence of Theorem 2.9.1 applied to $\mathcal{PK}(\mathbf{0})$.

3.6 Flag diffeomorphisms

In Section 2.11, dedicated to changes of variables for product kernels, we have shown for kernels of compact support, that if a change of variable maintains the singular subspaces then the composition of a product kernel with the diffeomorphism is well defined and is still a product kernel. In this chapter we will show that a similar statement holds for flag kernels.

Definition 3.6.1 (Flag diffeomorphism)**.**
Let $\Phi : \Omega \to \Phi(\Omega) \subset \mathbb{R}^N$ be a diffeomorphism of the open domain $\Omega \subset \mathbb{R}^N$ with its image. We say that Φ is a flag diffeomorphism with respect to the flag $(V_k)_{k\in\{0,\ldots,d\}}$ if

$$\Phi(\Omega\cap V_k) \subset V_k \qquad \forall k \in \{0,\ldots,d\}$$

and

$$\Phi^{-1}(\Phi(\Omega)\cap V_k) \subset V_k \qquad \forall k \in \{0,\ldots,d\}.$$

We say that Φ is a flag diffeomorphism of a compact domain if it is defined on a compact set Ω and it extends to a flag diffeomorphism of an open neighborhood of Ω.

Most of the remarks we made regarding product diffeomorphisms hold for flag diffeomorphisms. Flag diffeomorphisms of a fixed Ω onto itself form a group under composition.

We now concentrate on some properties of flag diffeomorphisms that are useful in relation to flag kernels, however, like in the case of product diffeomorphisms and kernels, this study can be carried out only on Euclidean spaces with a standard system of dilations. We argue mimicking the remarks we made for product kernels.

We begin with some preliminary results on the structure of flag diffeomorphisms. From now on we fix a flag $(V_k)_{k\in\{0,\ldots,d\}}$ on $\mathbb{R}^N$ and we choose a gradation compatible with the flag $(H_k)_{k\in\{1,\ldots,d\}}$. If Φ is a diffeomorphism then for any k we indicate as $\Phi^k \stackrel{\text{def}}{=} \pi_k\Phi$ the coordinate of Φ along the space H_k.

Lemma 3.6.2 (Differential inequalities for flag diffeomorphisms)**.**
Let Φ be a flag diffeomorphism of a compact domain. We have the following differential inequalities:

$$\begin{aligned} C_k^{-1}\,\mathrm{d}(\boldsymbol{x},V_k) < \mathrm{d}(\Phi(\boldsymbol{x}),V_k) < C_k\,\mathrm{d}(\boldsymbol{x},V_k) \\ \left|\partial_{\boldsymbol{x}}^{\boldsymbol{\alpha}}\Phi^k\right|(\boldsymbol{x}) < C_{\boldsymbol{\alpha},k}\left(1+\mathrm{d}(\boldsymbol{x},V_k)\right)^{1-|\alpha_k|} \end{aligned} \tag{3.6.1}$$

for all $k \in \{1,\ldots,d\}$.

Based on the previous inequalities we introduce the following terminology.

Definition 3.6.3 (Uniformly bounded families of diffeomorphisms).
Let us fix an open domain $\Omega \supset \bigotimes \overline{B_k(0,2^m)}$ *and let us consider a flag diffeomorphisms of* Ω *into* $\mathbb{R}^N$. *We say that a family of such diffeomorphisms is uniformly bounded if there are constants* C_k *and* $C_{\boldsymbol{\alpha},k}$ *such that the inequalities* (3.6.1) *hold uniformly for the whole family.*

Definition 3.6.4 (Stretching of flag diffeomorphisms).
Let us fix an open domain $\Omega \supset \bigotimes \overline{B_k(0,2^m)}$ *and let us consider a flag diffeomorphisms* Φ *of* Ω *into* $\mathbb{R}^N$. *We call the stretched diffeomorphism starting from the subspace* V_k *with parameter* s *the diffeomorphism* $\Phi_{s,k}$ *defined for* $s > 0$ *by the relation*

$$\Phi^l_{s,k} \stackrel{def}{=} \begin{cases} t^{-1} \cdot \Phi^l \left(\boldsymbol{x}_{\{1,\dots,k-1\}} \oplus t \cdot \boldsymbol{x}_{\{k,\dots,d\}} \right) & \text{if } l \geq k \\ \Phi^l \left(\boldsymbol{x}_{\{1,\dots,k-1\}} \oplus t \cdot \boldsymbol{x}_{\{k,\dots,d\}} \right) & \text{if } l < k \end{cases}$$

and for $s = 0$ *by the relation*

$$\Phi^l_{0,k} \stackrel{def}{=} \begin{cases} \partial_{\boldsymbol{x}_{\{k,\dots,d\}}} \Phi^l \left(\boldsymbol{x}_{\{1,\dots,k-1\}} \oplus \mathbf{0}_{\{k,\dots,d\}} \right) \boldsymbol{x}_{\{k,\dots,d\}} & \text{if } l \geq k \\ \Phi^l \left(\boldsymbol{x}_{\{1,\dots,k-1\}} \oplus \mathbf{0}_{\{k,\dots,d\}} \right) & \text{if } l < k \end{cases}$$

on the domain where the right-hand side is defined.

One checks by direct calculation that for any given $k \in \{1,\dots,d\}$ the mapping $(s,x) \mapsto \Phi_{(s,k)}(\boldsymbol{x})$ with $s \geq 0$ is smooth on the domain of definition and in particular that for $s \to 0$, $\Phi_{(s,k)}(\boldsymbol{x}) \to \Phi_{(0,k)}(\boldsymbol{x})$. It is also useful to notice that for $l \geq k$ we have the equality

$$\partial_{\boldsymbol{x}_{\{k,\dots,d\}}} \Phi^l \left(\boldsymbol{x}_{\{1,\dots,k-1\}} \oplus \mathbf{0}_{\{k,\dots,d\}} \right) \boldsymbol{x}_{\{k,\dots,d\}} = \partial_{\boldsymbol{x}_{\{l,\dots,d\}}} \Phi^l \left(\boldsymbol{x}_{\{1,\dots,k-1\}} \oplus \mathbf{0}_{\{k,\dots,d\}} \right) \boldsymbol{x}_{\{l,\dots,d\}}$$

because Φ_0, as also Φ, conserves the flag. Finally notice that $\Psi = \Phi^{-1}$ is defined on the domain $\Phi(\Omega)$ that is compact if Ω was compact. For values of the parameter $s > 0$ we have that the inverse of the stretched diffeomorphism $\Phi_{s,k}$ is $\Psi_{s,k}$. This relation holds for $s = 0$ by continuity.

Lemma 3.6.5 (Boundedness of stretched flag diffeomorphisms).
Let us fix an open domain $\Omega \supset \bigotimes \overline{B_k(0,2^m)}$ *and let us consider a bounded family of flag diffeomorphisms of* Ω *into* $\mathbb{R}^N$. *For any given* $k \in \{1,\dots,d\}$ *the flag diffeomorphisms obtained by stretching all the kernels of the family starting from the subspace* V_k *are defined at least on* $\bigotimes \overline{B_k(0,2^m)}$ *and are uniformly bounded for the parameter* s *in the range* $0 \leq s \leq 1$.

Proof. The proof of this property can also be obtained by verifying (3.6.1) directly. □

We have the following stability theorem for product kernels.

Theorem 3.6.6 (Flag kernel stability w.r.t diffeomorphisms).
Given a flag $(V_k)_{k\in\{0,\dots,d\}}$ *and let* $(H_k)_{k\in\{1,\dots,d\}}$ *be some gradation adapted to the flag, let* Φ *be a flag diffeomorphism of a compact domain* $\Omega \supset \bigotimes_{k=1}^d \overline{B_{H_k}(0,2^m)}$ *into* $\mathbb{R}^N$ *such that* $\bigotimes_{k=1}^d \overline{B_k(0,2^{m'})} \subset \Phi(\Omega)$. *For every flag kernel* $K \in \mathcal{FK}$ *with compact support with* $\operatorname{spt} K \subset \bigotimes_{k=1}^d \overline{B_{H_k}(0,2^{m'})}$. *Then* $K \circ \Phi$ *is a flag kernel. Furthermore, for any fixed* $m, m' \in \mathbb{Z}$ *the mapping* $K \mapsto K \circ \Phi$ *is continuous from the kernels in* $\mathcal{FK}$ *such that* $\operatorname{spt} K \subset \bigotimes_{k=1}^d \overline{B_{H_k}(0,2^{m'})}$ *to* $\mathcal{FK}$. *If* Φ *varies over a bounded family of flag diffeomorphisms satisfying the hypothesis then for any given* $K \in \mathcal{FK}$ *with* $\operatorname{spt} K \subset \bigotimes_{k=1}^d \overline{B_{H_k}(0,2^{m'})} \subset \Phi(\Omega)$ *the family* $K \circ \Phi$ *is uniformly bounded in* $\mathcal{FK}$.

Proof. We will prove that for any given diffeomorphism Φ and kernel K the kernel $K \circ \Phi$ is a well defined flag kernel. Since all the quantitative estimates we use are done using inequalities (3.6.1) and the bounds on K, the uniform boundedness and continuity parts of the statement follow.

Theorem 3.4.13 gives us the decomposition (3.4.10). Since the support of all the dyadic building blocks is inside the image of Φ we can write

$$K \circ \Phi(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ m'-2 \geq i_1 \geq \ldots \geq i_d}} \varphi_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\Phi(\boldsymbol{x})) = \sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ m'-2 \geq i_1 \geq \ldots \geq i_d}} 2^{-\boldsymbol{i}\boldsymbol{q}} \varphi_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \Phi(\boldsymbol{x})).$$

By setting

$$\widetilde{\varphi}_{\boldsymbol{i}}(\boldsymbol{x}) = \varphi_{\boldsymbol{i}}\left(2^{-\boldsymbol{i}} \cdot \Phi(2^{\boldsymbol{i}} \cdot \boldsymbol{x})\right)$$

we have that

$$K \circ \Phi(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{i} \in \mathbb{Z}^d \\ m'-2 \geq i_1 \geq \ldots \geq i_d}} 2^{-\boldsymbol{i}\boldsymbol{\nu}} \widetilde{\varphi}_{\boldsymbol{i}}^{(2^{\boldsymbol{i}})}(\boldsymbol{x}). \tag{3.6.2}$$

It is easy to see that $\{\widetilde{\varphi}_{\boldsymbol{i}}\}$ is a bounded family in $S(\mathbb{R}^N)$. Furthermore since the support of $\varphi_{\boldsymbol{i}}$ is contained in the set $\{\boldsymbol{x} |\, |\boldsymbol{x}| < 4\}$ we have that the support of $\widetilde{\varphi}_{\boldsymbol{i}}$ is such that

$$\mathrm{d}(V_k; 2^{-\boldsymbol{i}} \cdot \Phi(2^{\boldsymbol{i}} \cdot \boldsymbol{x})) \approx \sum_{j=k+1}^{d} 2^{-i_j} \left| \Phi^j(2^{\boldsymbol{i}} \cdot \boldsymbol{x}) \right| \lesssim 4.$$

Using inequalities (3.6.1) and an induction argument for some C we have that $|x_k| < C$ for all k.

We will now show that the dyadic family $\{\widetilde{\varphi}_{\boldsymbol{i}}\}$ satisfies the hypothesis of Theorem 3.4.12. We begin by checking that for some $\varepsilon > 0$ any $J \subset \{2, \ldots, d\}$ the family of functions

$$\boldsymbol{x}_{(J)} \mapsto \prod_{j \in J} 2^{-\varepsilon(i_j - i_{j-1})} \int_{H_J} \widetilde{\varphi}_{\boldsymbol{i}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}_J$$

is uniformly bounded. Remark 3.4.9 guarantees that it is sufficient to check the above property with $|J| = 1$.

Let us suppose $J = \{j\}$. Let $\Phi_{s,j}$ be a stretch with parameter $0 \leq s \leq 1$ starting from the subspace V_j of the diffeomorphism Φ. For simplicity we will indicate $\Phi_s \stackrel{\text{def}}{=} \Phi_{s,j}$. We have

$$\int_{H_j} \varphi_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \Phi(2^{\boldsymbol{i}} \cdot \boldsymbol{x})) \mathrm{d}x_j = \int_{H_j} \varphi_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \Phi_1(2^{\boldsymbol{i}} \cdot \boldsymbol{x})) \mathrm{d}x_j =$$

$$\int_{H_j} \varphi_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \Phi_0(2^{\boldsymbol{i}} \cdot \boldsymbol{x})) \mathrm{d}x_j + \int_{s=0}^{1} \partial_s \left(\int_{H_j} \varphi_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \Phi_s(2^{\boldsymbol{i}} \cdot \boldsymbol{x})) \mathrm{d}x_j \right) \mathrm{d}s.$$

For the first term we write

$$\int_{H_J} \varphi_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \Phi_0(2^{\boldsymbol{i}} \cdot \boldsymbol{x})) \mathrm{d}x_j =$$

$$\int_{H_j} \varphi_{\boldsymbol{i}} \left(\pi_{H_{\{1,\ldots,j-1\}}} \left(2^{-\boldsymbol{i}} \cdot \Phi_0(2^{\boldsymbol{i}} \cdot \boldsymbol{x}) \right) \oplus \pi_{H_{\{j,\ldots,d\}}} \left(2^{-\boldsymbol{i}} \cdot \Phi_0(2^{\boldsymbol{i}} \cdot \boldsymbol{x}) \right) \right) \mathrm{d}x_j.$$

$\pi_{H_{\{1,\ldots,j-1\}}}\left(2^{-\boldsymbol{i}}\cdot\Phi_0(2^{\boldsymbol{i}}\cdot\boldsymbol{x})\right)$ does not depend on x_j

$$\pi_{H_{\{j,\ldots,d\}}}\left(2^{-\boldsymbol{i}}\cdot\Phi_0(2^{\boldsymbol{i}}\cdot\boldsymbol{x})\right)=\bigoplus_{l=j}^{d}\sum_{k=l}^{d}2^{i_l-i_k}\partial_l\Phi^k\left(\boldsymbol{x}_{\{1,\ldots,j-1\}}+\boldsymbol{0}_{\{j,\ldots,d\}}\right)x_l.$$

The only term depending on x_j is

$$\partial_j\Phi^j(\boldsymbol{x}_{\{1,\ldots,j-1\}})x_j.$$

The above linear operator on H_j is non-degenerate because Φ_0 is a local diffeomorphism so, since $\varphi_{\boldsymbol{i}}$ had strong cancellation the above integral we were looking at is 0.

We now proceed to the next term.

$$\int_{s=0}^{1}\partial_s\left(\int_{H_j}\varphi_{\boldsymbol{i}}(2^{-\boldsymbol{i}}\cdot\Phi_s(2^{\boldsymbol{i}}\cdot\boldsymbol{x}))\mathrm{d}x_j\right)\mathrm{d}s=$$

$$\sum_{k=1}^{j-1}\int_{s=0}^{1}\int_{H_j}2^{-i_k}\partial_k\varphi_{\boldsymbol{i}}(2^{-\boldsymbol{i}}\cdot\Phi_s(2^{\boldsymbol{i}}\cdot\boldsymbol{x}))\partial_s\Phi_s^k(2^{\boldsymbol{i}}\cdot\boldsymbol{x})\mathrm{d}x_j\mathrm{d}s+$$

$$\sum_{k=j}^{d}\int_{s=0}^{1}\int_{H_j}2^{-i_k}\partial_k\varphi_{\boldsymbol{i}}(2^{-\boldsymbol{i}}\cdot\Phi_s(2^{\boldsymbol{i}}\cdot\boldsymbol{x}))\partial_s\Phi_s^k(2^{\boldsymbol{i}}\cdot\boldsymbol{x})\mathrm{d}x_j\mathrm{d}s.$$

Notice that for $k<j$ we have

$$\partial_s\Phi_s^k(\boldsymbol{x})=\sum_{l=j}^{d}(\partial_{x_l}\Phi^k)(\boldsymbol{x}_{\{1,\ldots,j-1\}}+s\cdot\boldsymbol{x}_{\{j,\ldots,d\}})x_l$$

so for we rewrite the sum as

$$\sum_{k=1}^{j-1}\int_{s=0}^{1}\int_{H_j}2^{-i_k}\partial_k\varphi_{\boldsymbol{i}}(2^{-\boldsymbol{i}}\cdot\Phi_s(2^{\boldsymbol{i}}\cdot\boldsymbol{x}))\partial_s\Phi_s^k(2^{\boldsymbol{i}}\cdot\boldsymbol{x})\mathrm{d}x_j\mathrm{d}s=$$

$$\sum_{k=1}^{j-1}\sum_{l=j}^{d}2^{i_l-i_k}\int_{s=0}^{1}\int_{H_j}\partial_k\varphi_{\boldsymbol{i}}(2^{-\boldsymbol{i}}\cdot\Phi_s(2^{\boldsymbol{i}}\cdot\boldsymbol{x}))(\partial_{x_l}\Phi^k)(2^{\boldsymbol{i}}\cdot(\boldsymbol{x}_{\{1,\ldots,j-1\}}+s\cdot\boldsymbol{x}_{\{j,\ldots,d\}}))x_l\mathrm{d}x_j\mathrm{d}s.$$

Since the integral is uniformly bounded, the coefficient $2^{i_l-i_k}$ gives us the needed estimate.

For $k\geq j$ we have

$$\partial_s\Phi_s^k(\boldsymbol{x})=s^{-2}\left(\sum_{l=j}^{d}(\partial_{x_l}\Phi^k)(\boldsymbol{x}_{\{1,\ldots,j-1\}}+s\cdot\boldsymbol{x}_{\{j,\ldots,d\}})sx_l-\Phi^k(\boldsymbol{x}_{\{1,\ldots,j-1\}}+s\cdot\boldsymbol{x}_{\{j,\ldots,d\}})\right)=$$

$$s^-2\left(\int_0^1\sum_{l=j}^{d}\sum_{l'=j}^{d}(\partial_{x_l}\partial_{x'_l}\Phi^k)(\boldsymbol{x}_{\{1,\ldots,j-1\}}+(1-t)s\cdot\boldsymbol{x}_{\{j,\ldots,d\}})sx_lsx_{l'}\right)$$

The quadratic form for $l,l'<k$

$$\int_0^1(\partial_{x_l}\partial_{x'_l}\Phi^k)(\boldsymbol{x}_{\{1,\ldots,j-1\}}+(1-t)s\cdot\boldsymbol{x}_{\{j,\ldots,d\}})$$

vanishes on V_{k-1} so we finally have that

$$\partial_s \Phi_s^k(\boldsymbol{x}) = O(\mathrm{d}(V_{k-1};\boldsymbol{x})\,\mathrm{d}(V_{j-1};\boldsymbol{x}))$$

and the expression

$$\sum_{k=j}^{d} 2^{i_j} \int_{s=0}^{1} \int_{H_j} \partial_k \varphi_{\boldsymbol{i}}(2^{-\boldsymbol{i}} \cdot \Phi_s(2^{\boldsymbol{i}} \cdot \boldsymbol{x})) 2^{-i_k - i_j} \partial_s \Phi_s^k(2^{\boldsymbol{i}} \cdot \boldsymbol{x}) \mathrm{d}x_j \mathrm{d}s.$$

has a uniformly bounded integrand. The same argument as above, taking $J = \{1\}$, gives that the functions

$$\boldsymbol{x}_{(1)} \to 2^{-i_1} \int_{H_1} \widetilde{\varphi}_{\boldsymbol{i}}(\boldsymbol{x}) \mathrm{d}x_1$$

are uniformly bounded. This provides for the relaxed cancellation condition on x_1. All the estimates necessary for applying Theorem 3.4.12 are satisfied. □

Bibliography

[Duo01] Javier Duoandikoetxea. *Fourier Analysis*. Crm Proceedings & Lecture Notes. American Mathematical Society, 2001. ISBN: 9780821883846.

[EG92] Lawrence C Evans and Ronald F Gariepy. *Measure theory and fine properties of functions*. Vol. 5. CRC press, 1992.

[Gra08] Loukas Grafakos. *Classical fourier analysis*. Vol. 86. Springer, 2008.

[Nag+12] Alexander Nagel et al. "Singular integrals with flag kernels on homogeneous groups, I". In: *Revista matemática iberoamericana* 28.3 (2012), pp. 631–722.

[NRS01] Alexander Nagel, Fulvio Ricci, and Elias M Stein. "Singular integrals with flag kernels and analysis on quadratic CR manifolds". In: *Journal of Functional Analysis* 181.1 (2001), pp. 29–118.

[Rud91] Walter Rudin. *Functional analysis*. McGraw-Hill, Inc., New York, 1991.