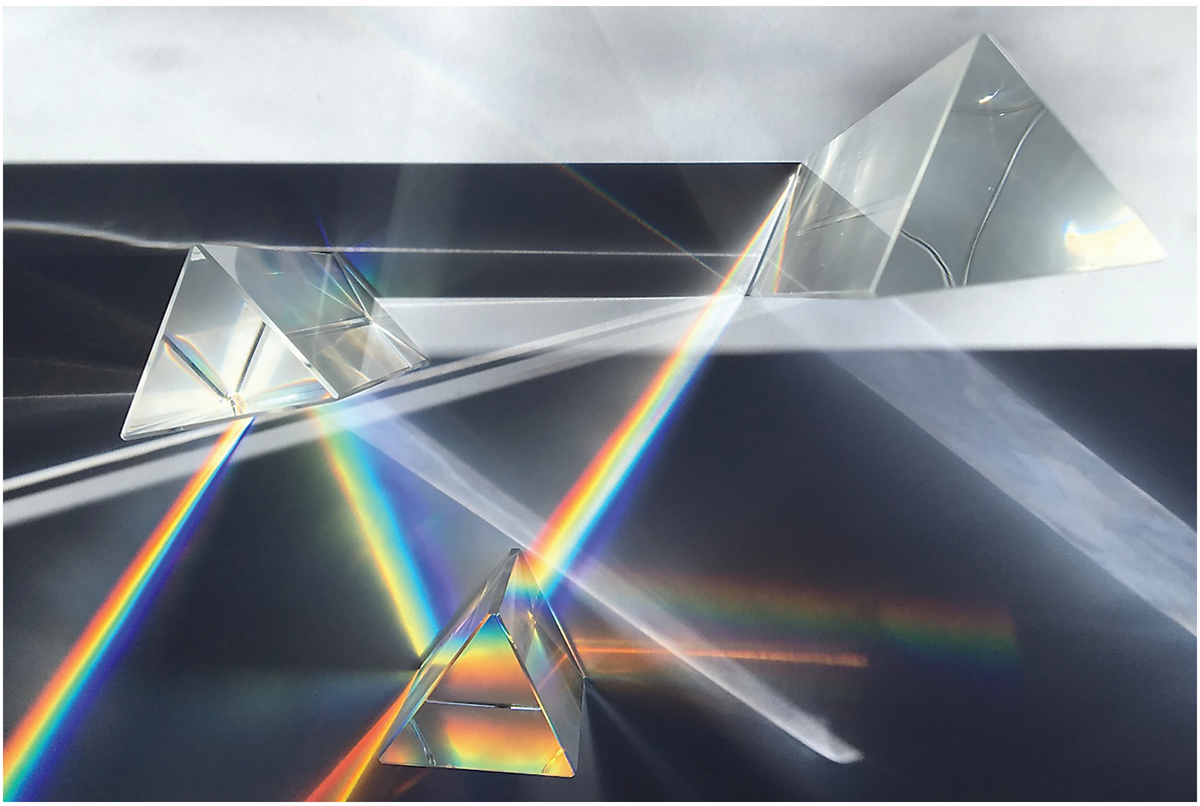
Almost-Orthogonality in Fourier Analysis

From Discrete Characterizations of Function Spaces, to Singular Integrals, to Leibniz Rules for Fractional Derivatives



Rodolfo H. Torres

Introduction

From early on in their learning of elementary calculus students get troubled, puzzled, and also occasionally amazed by the Leibniz rule for the derivative of the product of two functions. Namely, say in dimension one,

$$(f \cdot g)' = f' \cdot g + f \cdot g'.$$

Later on in our mathematics and physics education we also learn of the ubiquitous roles of both differentiation and what is perhaps the simplest of all nonlinearities, the powers of functions. To understand the combined effect of smoothness and nonlinearities involving square or cubic terms in the statements and solutions of differential equations that model many physical situations, we often

Rodolfo H. Torres is Vice Chancellor for Research and Economic Development and a Distinguished Professor of Mathematics at the University of California, Riverside. His email address is rodolfo.h.torres@ucr.edu.

This article is dedicated to the memory of Richard L. Wheeden.

Communicated by Notices *Associate Editor Daniela De Silva.*

For permission to reprint this article, please contact: reprint-permission@ams.org.

DOI: https://doi.org/10.1090/noti2133

find ourselves on the harmonic analysis side with the need to study the regularity properties of the product of functions. In many contexts, it is sometimes necessary to look at the regularity of functions in terms of Hölder continuity or other related properties which are a *fractional* or intermediate measure of smoothness between standard integer numbers of derivatives. The product of functions is also the simplest example of a *bilinear operator*, that is, an operator $T(f,g)$ which is linear in each of its entries. Another such example is of course

$$T(f,g) = f' \cdot g + f \cdot g'.$$

Borrowing from linear algebra (think of matrices to represent linear operators) it is very convenient to try to *diagonalize* bilinear operators. Since we will be doing analysis, we will need to quantify continuity, smoothness, and other properties of products and derivatives, and this is best done in terms of *norm estimates* in function spaces. The *discretization* of such spaces using appropriate expansions permits one to diagonalize and estimate the operators in question.

We will attempt to better detail all of the above in the subsequent narrative. In fact, the purpose of this expository article is to illustrate how well-established tools together with more recent developments in Fourier analysis led to the proof of fractional-power versions of the classical Leibniz rule for derivatives of the product of functions.

The ideas used circle around the natural but powerful concept that, in many situations, waves with different frequencies (and a lot of cancellation) are *almost-orthogonal*; that is, they do not interact much. This can be clearly observed, for example, in the scattering of light in the atmosphere. Due to the presence of small particles, waves with shorter wavelengths scatter much more than the ones with larger wavelengths, hence the blue color of the sky we observe. A cartoonish physical representation, as in Figure 1, is usually employed to explain this. A big basketball (representing large red wavelengths) will roll unaffected by a tiny crack on the floor (representing a variation in index of refraction) because of the difference in size. A small blue ball (representing shorter wavelengths), on the other hand, will be scattered by the crack because of comparable sizes. In this sense, the basketball and the tiny crack are said to be almost-orthogonal to each other, as are the large red wavelengths and the small particles in the atmosphere. We will come back to this analogy in a mathematical context later in the presentation.

We will describe how Fourier analysis techniques allow us to study bilinear operators and obtain appropriate size estimates on their fractional derivatives. We will use a rather informal nontechnical but hopefully intuitive language at first, to later make more mathematically precise statements. We will borrow some parts of our narrative

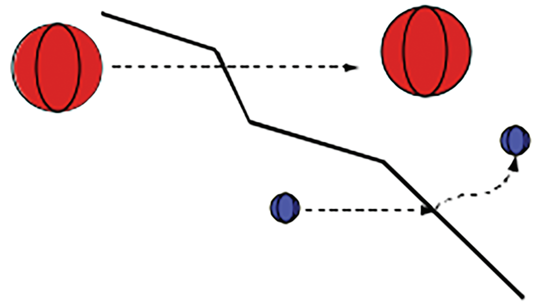


Figure 1. Waves with different frequencies do not interact too much.

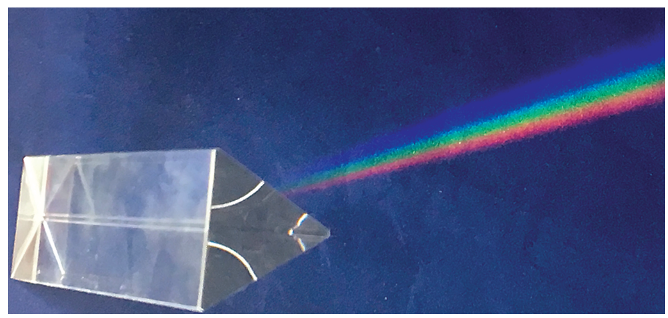


Figure 2. The FT behaves like a prism decomposing a signal into waves of different frequencies.

from what has become part of the folklore of the subject and can also be found in many of the articles cited, but we will also strive to present these perspectives in a simple way accessible to nonexperts. It is hard to do justice to all the contributions by many colleagues in the subject in this short note. We encourage the interested reader to visit the extensive and far more comprehensive bibliographies found in the small sample of articles we shall cite. Our goal is to start from very simple principles and arrive at recent results for fractional Leibniz rules in mixed Lebesgue spaces.

Fourier analysis permits the decomposition of a function into a combination of oscillating waves of different frequencies and amplitudes, very much in the same way that a prism separates a beam of light into a rainbow of colors of different wavelengths; see Figure 2. Mathematically, this was first codified in the classical Fourier series

$$\begin{aligned} g(x) &= a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) \\ &= \sum_{n=-\infty}^{\infty} c_n e^{inx}. \end{aligned}$$

We will work, however, with functions defined in $\mathbb{R}^n$, so we will use instead the Fourier Transform (FT),

$$F(f)(\xi) = \widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-ix\cdot\xi}\, dx,$$

with inverse given by

$$f(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} \widehat{f}(\xi) e^{ix\cdot\xi}\, d\xi.$$

Fourier series are useful in solving heat and wave equations via separation of variables in the periodic case because the sines and cosines are eigenfunctions of $L = -\frac{\partial^2}{\partial x^2}$. Likewise, the FT diagonalizes differential operators,

$$\widehat{\partial^\alpha f}(\xi) = i^{|\alpha|}\xi^\alpha \widehat{f}(\xi).$$

Here we have employed the usual notation for $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$, $|\alpha| = \alpha_1 + \cdots + \alpha_n$, $\xi^\alpha = \prod_{j=1}^n \xi_j^{\alpha_j}$, and $\partial^\alpha = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n}$. For example, for $\triangle = \sum_{j=1}^n \partial_j^2$,

$$\widehat{\triangle f}(\xi) = -|\xi|^2 \widehat{f}(\xi).$$

We will be interested in *fractional derivatives*, so we introduce then the fractional derivative operators $D^s f = (-\Delta)^{s/2} f$ and $J^s f = (I - \Delta)^{s/2} f$ defined for $s \in \mathbb{R}$ via the FT by

$$\widehat{D^s f}(\xi) = |\xi|^s \widehat{f}(\xi) \quad \text{and} \quad \widehat{J^s f} = (1 + |\xi|^2)^{s/2} \widehat{f}(\xi).$$

The FT also diagonalizes translation invariant operations such as *convolutions*. That is,

$$f * g(x) = \int_{\mathbb{R}^n} f(x-y) g(y)\, dy \Rightarrow \widehat{f * g}(\xi) = \widehat{f}(\xi)\widehat{g}(\xi).$$

The practical effect of a *convolution* operation as above is that of *filtering* different frequencies in a function or signal. This is illustrated in Figure 3, where we cover the spectrum of a given function with two functions, $\widehat{\phi}$ supported at low frequencies around 0 and $\widehat{\psi}$ supported at higher frequencies. The convolution with ϕ recovers then the general shape of the function, while the one with ψ encodes the higher oscillations in it.

The FT interacts nicely with the group of translations and dilations in $\mathbb{R}^n$ as follows:

$$\text{if } \tau_h f(x) = f(x-h), \text{ then } \widehat{\tau_h f}(\xi) = e^{-ih\cdot\xi} \widehat{f}(\xi);$$

$$\text{if } f_\epsilon(x) = \epsilon^{-n} f(\epsilon^{-1}x), \text{ then } \widehat{f_\epsilon}(\xi) = \widehat{f}(\epsilon\xi).$$

Many of the estimates we will consider have some translation and dilation invariance built into them, and so these properties of the FT are very well suited and simplify a lot of the analysis.

Discretization of Linear Operators

Suppose that a countable family $\{\varphi_\lambda\}_\lambda$ *generates* a space of functions in a similar fashion to the way a complete orthonormal basis expands elements in a Hilbert space (though that does not have to be the case here). That is, every element in the space can be written in the form

$$f = \sum_\lambda \langle f, \varphi_\lambda \rangle \varphi_\lambda,$$

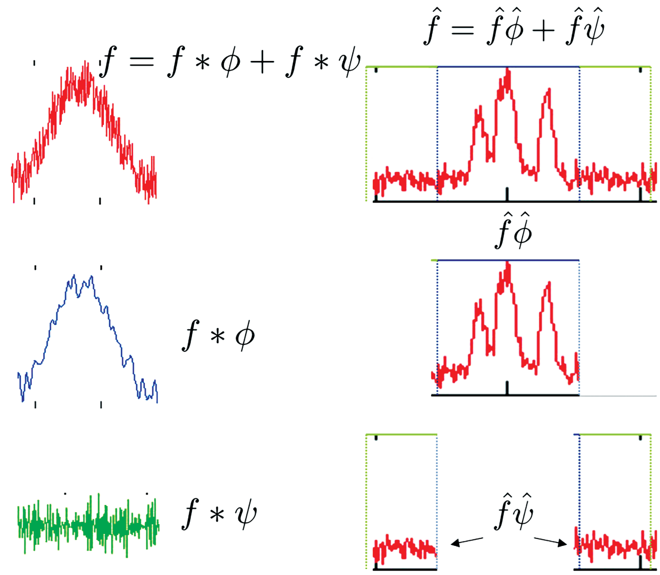


Figure 3. Pictorial representation of convolutions as filters.

where $\langle \cdot, \cdot \rangle$ will typically denote the integral of the product of the functions involved or the action of a distribution against a test function.

Let T be a linear operator acting on such space. We can write $T(f)$ in terms of the generating family and then expanding f, and without considering issues about convergence, we arrive at a familiar formula:

$$\begin{aligned} T(f) &= \sum_\nu \langle T(f), \varphi_\nu \rangle \varphi_\nu \\ &= \sum_\nu \sum_\lambda \langle T(\varphi_\lambda), \varphi_\nu \rangle \langle f, \varphi_\lambda \rangle \varphi_\nu. \end{aligned}$$

In this way, we can associate to T, as we do in finite-dimensional linear algebra, a *matrix*

$$A_T = (\langle T(\varphi_\lambda), \varphi_\nu \rangle)_{\lambda,\nu}.$$

If T *preserves the building blocks* φ_λ (as if they were eigenfunctions), then $T(\varphi_\lambda)$ will be a function very similar to φ_λ and, if we also have some orthogonality between the φ_ν's, then (wishfully thinking) A_T will be *almost diagonal*. That is to say, the off-diagonal entries will satisfy a size estimate

$$|\langle T(\varphi_\lambda), \varphi_\nu \rangle| \sim small_{\lambda,\nu}$$

when λ is *far away from* ν.[1]

Similarly, along these lines of thought, if T is a bilinear operator we can write for a pair of functions

$$\begin{aligned} T(f,g) &= \sum_\nu \langle T(f,g), \varphi_\nu \rangle \varphi_\nu \\ &= \sum_\nu \sum_\lambda \sum_\mu \langle T(\varphi_\lambda, \varphi_\mu), \varphi_\nu \rangle \langle f, \varphi_\lambda \rangle \langle g, \varphi_\mu \rangle \varphi_\nu. \end{aligned}$$

[1] *Although we are not assigning a precise meaning to the notation $A \sim B$, in the next sections we will use the notation $A \lesssim B$ to mean $A \le cB$ for some $c > 0$ independent of A and B. Likewise, we will write $A \approx B$ when $A \lesssim B$ and $B \lesssim A$ for appropriate constants.*

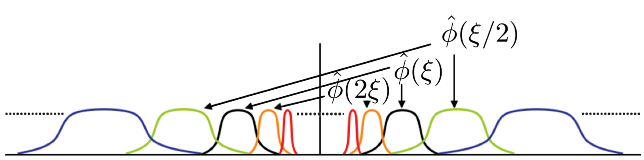


Figure 4. LPS function.

We can now associate to T a tensor

$$A_T = \left(\langle T(\varphi_\lambda, \varphi_\mu), \varphi_\nu\rangle\right)_{\lambda,\mu,\nu}.$$

Again, we would like to have some almost-orthogonality for A_T of the form

$$|\langle T(\varphi_\lambda, \varphi_\mu), \varphi_\nu\rangle| \sim small_{\lambda,\mu,\nu}$$

whenever λ, μ, and ν are far apart from each other. Of course, all this needs to be precisely defined and quantified, which we will do later. The idea is to give up on the notion of eigenfunctions for a particular operator to use instead *approximate eigenfunctions* for a whole family of operators. To do so, instead of the *pure tones* $x \to e^{i\lambda\cdot x}$ of classical Fourier analysis we will use time-frequency localized smooth bumps φ_λ to expand the spaces. These bump functions are elements of the Schwartz spaces $\mathcal{S}$ of rapidly decreasing C^∞ functions. We can even take them with their FT compactly supported away from the origin, and hence φ_λ will have a lot of cancellation too. We will look at particular classes of such generating families of elements that are called *wavelets*.

Littlewood–Paley–Stein and Wavelet Decompositions

With our definition of the FT we have, up to a multiplicative constant, an L^2-isometry, $\|f\|_{L^2} = (2\pi)^{-n/2}\|\widehat{f}\|_{L^2}$. For $1 < p < \infty$ and $p \neq 2$, however, we cannot estimate the L^p-norm of a function directly in terms of the L^p-norm of its FT. There is nonetheless a practical substitute that allows us, to some extent, to estimate the L^p-norm of a function in terms of pieces whose FTs do not interact much with each other. This is called the *Littlewood–Paley–Stein (LPS) decomposition or discrete square function*. Let $\phi \in \mathcal{S}$ be such that $\widehat{\phi}$ has compact support away from the origin and satisfies

$$\sum_{\nu\in\mathbb{Z}} \widehat{\phi}(2^{-\nu}\xi)^2 = 1, \quad \xi \neq 0, \tag{1}$$

as in Figure 4.

Define $\phi_{2^{-\nu}}(x) = 2^{\nu n}\phi(2^\nu x)$, so $\widehat{\phi}_{2^{-\nu}}(\xi) = \widehat{\phi}(2^{-\nu}\xi)$. We have that

$$\begin{aligned}\|f\|_{L^p} &= \left(\int_{\mathbb{R}^n} |f(x)|^p\, dx\right)^{1/p} \\ &\approx \left\|\left(\sum_\nu |f * \phi_{2^{-\nu}}|^2\right)^{1/2}\right\|_{L^p}. \end{aligned}\tag{2}$$

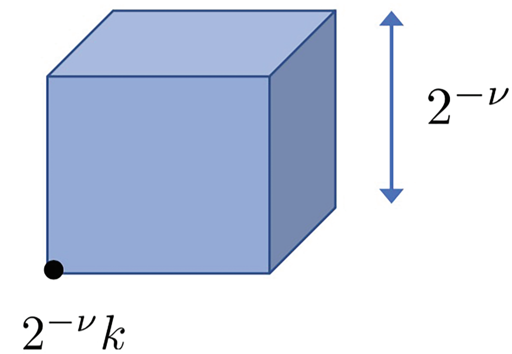


Figure 5. Dyadic cube at position $2^{-\nu}k$ and scale $2^{-\nu}$.

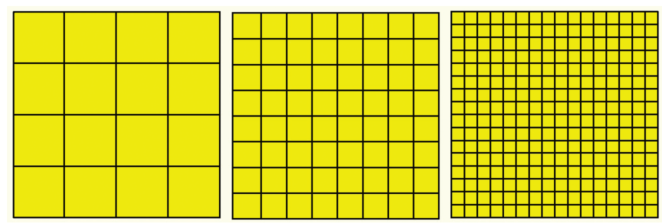


Figure 6. Dyadic squares at different scales.

The reason why such an apparently complicated formula (which is trivially verified in L^2) may also be true in L^p is given by the orthogonality of (most of) the LPS pieces $f * \phi_{2^{-\nu}}$. Moreover, the LPS pieces have excellent localization in frequency since they are supported where $|\xi| \approx 2^\nu$, and so each $\phi_{2^{-\nu}}$ filters out all frequencies but those in such a specific band. Using this, one can verify that $\partial^\alpha(f * \phi_{2^{-\nu}})$ behaves approximately like $2^{\nu|\alpha|} f * \phi_{2^{-\nu}}$. This is very useful when studying estimates for derivatives. In fact one can also characterize the Sobolev spaces $L^p_s(\mathbb{R}^n)$ of *functions with s derivative in* $L^p(\mathbb{R}^n)$ in the form[2]

$$\|f\|_{L^p_s} \approx \left\|\left(\sum_\nu |2^{\nu s} f * \phi_{2^{-\nu}}|^2\right)^{1/2}\right\|_{L^p} + \|f\|_{L^p}.$$

The LPS decomposition can be further discretized. Our choice of function ϕ allows us to write the discrete Calderón reproducing formula

$$f = \sum_\nu (f * \phi_{2^{-\nu}}) * \phi_{2^{-\nu}}. \tag{3}$$

Let $k = (k_1, \dots, k_n) \in \mathbb{Z}^n$, $\nu \in \mathbb{Z}$, and consider the usual collection of all dyadic cubes $Q_{\nu k}$ given by *position* $2^{-\nu}k$ and *scale* $2^{-\nu}$. More precisely, $Q_{\nu k} = \{\xi \in \mathbb{R}^n : 2^{-\nu}k_j \leq \xi_j < 2^{-\nu}(k_j + 1), j = 1, \dots, n\}$. For each ν, the corresponding family $\{Q_{\nu k}\}_k$ provides a covering of $\mathbb{R}^n$. Varying the scales, one can check that the cubes that overlap have very useful hierarchy properties in terms of how they intersect; see Figures 5 and 6.

[2] *More precisely,* $f \in L^p_s$ *if and only if* $J^s f \in L^p$*. The Sobolev space norm is defined then by* $\|f\|_{L^p_s} = \|J^s f\|_{L^p}$*.*

A discretization of (3) at different scales leads, formally, to wavelet-type decomposition of functions. In fact, since each of the functions $f * \phi_{2^{-\nu}}$ does not have frequencies larger than 2^{ν}, it does not oscillate much in an interval of length $2^{-\nu}$. It can then be efficiently *sampled* at that rate. So, formally discretizing as a Riemann sum the second convolution in Calderón's formula,

$$
\begin{aligned}
&(f * \phi_{2^{-\nu}}) * \phi_{2^{-\nu}}(x) \\
&= \int_{\bigcup Q_{\nu k}} (f * \phi_{2^{-\nu}})(y)\, \phi_{2^{-\nu}}(x-y)\, dy \\
&\sim \sum_k (f * \phi_{2^{-\nu}})(2^{-\nu}k)\, \phi_{2^{-\nu}}(x - 2^{-\nu}k)\, 2^{-\nu n},
\end{aligned}
$$

yields the representation

$$f \sim \sum_{\nu,k} \langle f, \phi_{\nu k}\rangle\, \phi_{\nu k},$$

where now $\phi_{\nu k}(x) = 2^{\nu n/2}\phi(2^{\nu}x - k)$.

Of course, there is a lot more work to do to justify the above empirical approach. Y. Meyer [23] (constructing orthonormal wavelet expansions) and M. Frazier jointly with B. Jawerth [13] (following a sampling based on the nonorthogonal ϕ-transform approach) obtained in the 1980s rigorous characterizations of L^p spaces of the form

$$f = \sum_{\nu\in\mathbb{Z}}\sum_{k\in\mathbb{Z}^n} \langle f, \phi_{\nu k}\rangle \phi_{\nu k}$$

with

$$\|f\|_{L^p} \approx \left\| \Big(\sum_{\nu} \Big(\sum_{k} |\langle f, \phi_{\nu k}\rangle| 2^{\nu n/2} \chi_{Q_{\nu k}} \Big)^2 \Big)^{1/2} \right\|_{L^p}$$

(here $\chi_{Q_{\nu k}}$ is the indicator function of $Q_{\nu k}$). We will call both representations in the characterizations by Meyer and Frazier–Jawerth wavelet representations.[3] Remarkably, other spaces of functions (Sobolev, Hardy, Besov, Triebel–Lizorkin, and essentially any space which admits a discrete LPS decomposition) can be characterized in a similar way using only size estimates of the wavelet coefficients $\langle f, \phi_{\nu k}\rangle$!

Without stating a technical definition, we will say that a function is *essentially supported* or *concentrated on a cube* if it decays fast to zero outside the cube. Note that if the function ϕ is essentially supported at the origin on a cube of side length 1, then by dilation and translation $\phi_{\nu k}$ is essentially supported on the cube $Q_{\nu k}$. Hence, as is nowadays well understood, the wavelet coefficient $\langle f, \phi_{\nu k}\rangle$ provides a snapshot of f at scale or resolution $2^{-\nu}$ (frequency $\approx 2^{\nu}$) and position $\approx 2^{-\nu}k$. This multiresolution analysis helps explain the success of wavelets in the characterization of function spaces, as well as in many of their science, engineering, and everyday-life applications we now enjoy.

[3] *There is a lot more to be said about wavelets and Calderón reproducing formula which we cannot include here; see* [5].

Calderón–Zygmund Operators

One of the original motivations for the seminal work of A. P. Calderón and A. Zygmund on singular integrals in higher dimensions arises from potential theory and the Poisson equation $\triangle u = f$. In $\mathbb{R}^n$, with $n \geq 3$, the solution of such an equation is given for an appropriate constant c_n by

$$u(x) = c_n \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|^{n-2}}\, dy.$$

If we take two derivatives of u (and differentiate inside the integral) we need to study integral operators of the type

$$T(f) = \int_{\mathbb{R}^n} K(x,y) f(y)\, dy$$

where

$$K(x,y) \approx |x-y|^{-n}.$$

Because of the singularity of the kernels, such operators are not given by absolutely convergent Lebesgue integrals. They have to be interpreted in an appropriate way, typically as principal-value integrals or in a distributional sense. Many operators of interest in analysis are of this form. The following are relevant examples.[4]

- The Hilbert transform:

$$
\begin{aligned}
Hf(x) &= c_1\, p.v.\!\int_{\mathbb{R}} \frac{f(y)}{x-y}\, dy \\
&= -i \int_{\mathbb{R}} \operatorname{sgn}(\xi) \widehat{f}(\xi) e^{ix\xi}\, d\xi
\end{aligned}
$$

- The Riesz transforms:

$$
\begin{aligned}
R_j f(x) &= c_n\, p.v.\!\int_{\mathbb{R}^n} \frac{x_j - y_j}{|x-y|^{n+1}} f(y)\, dy \\
&= -i \int_{\mathbb{R}^n} \frac{\xi_j}{|\xi|} \widehat{f}(\xi) e^{ix\cdot\xi}\, d\xi
\end{aligned}
$$

- The Calderón commutators with a Lipschitz function A:

$$T_m f(x) = p.v.\!\int_{\mathbb{R}} \frac{(A(x)-A(y))^m}{(x-y)^{m+1}} f(y)\, dy$$

- The Cauchy integral on a Lipschitz graph:

$$Cf(x) = p.v.\!\int_{\mathbb{R}} \frac{f(y)}{(x-y) + i(A(x)-A(y))}\, dy$$

- The pseudodifferential operators with symbols in the class S^0:

$$T_\sigma f(x) = \int_{\mathbb{R}^n} \sigma(x,\xi) e^{ix\cdot\xi} \widehat{f}(\xi)\, d\xi,$$

  where $|\partial_x^\beta \partial_\xi^\alpha \sigma(x,\xi)| \leq C_{\alpha,\beta}(1+|\xi|)^{-|\alpha|}$ for all multiindices α and β.

[4] *The constants c_n depend on the definition of the FT but are of no relevance here.*

These operators are examples of what are generically called Calderón–Zygmund operators (CZOs). It would take one expository article for each of the above examples to talk about their rich history and applications. In any case, the characteristic feature of all CZOs is that if $T : L^2 \to L^2$ and

$$|\partial^\alpha K(x,y)| \le C|x-y|^{-(n+|\alpha|)} \quad \text{for } |\alpha| \le 1,$$

then $T : L^p \to L^p, 1 < p < \infty$.[5] The main question in studying such operators becomes then their boundedness on L^2. Although for convolution operators like the first two examples the answer is immediate,[6] for nonconvolution ones the difficult question of characterizing L^2 boundedness was finally answered by G. David and J. L. Journé in the mid-1980s with the T(1)-Theorem. The name of the result is due to a characterization of boundedness in terms of the action of a CZO and its transpose on the constant function 1, plus some weak boundedness condition. However, the authors also obtained a different formulation which is more convenient for our purposes.

T(1)-Theorem (David–Journé [11]). *Assume that the kernel of the operator T satisfies*

$$|\partial^\alpha K(x,y)| \le C|x-y|^{-(n+|\alpha|)}, \quad |\alpha| \le 1.$$

Then,

$$T : L^2 \to L^2, \quad \textit{and hence} \quad T : L^p \to L^p, 1 < p < \infty,$$

if and only if[7]

$$\sup_\xi (\|T(e^{ix\cdot\xi})\|_{BMO} + \|T^*(e^{ix\cdot\xi})\|_{BMO}) \le C.$$

Although a cumbersome way to state the theorem, as it needs to make sense a priori of the action of the operator on exponentials (the operators are usually only initially defined on a dense set of L^2 which of course does not contain the exponential functions), the above has a beautiful interpretation: we only need to understand the operator on the basic elements of Fourier analysis. In a way, we like to say that to understand the music played by T we need to understand how it plays each note or pure tone $e^{ix\cdot\xi}$.

We mentioned early on that sometimes we want to study operators by discretizing them and studying their matrices. So instead of using the exponential pure tones to check T, we can use wavelets. Although this was not the original proof of the theorem, one can prove it by showing that the matrix associated to a CZO with respect to a family of wavelets is almost diagonal.

[5] *CZOs also satisfy some appropriate end-point estimates when $p = 1$ or $p = \infty$, which we will not recount here.*

[6] *If $K(x,y) = k(x-y)$, then $Tf(x) = k * f(x)$ or $\widehat{Tf}(\xi) = \hat{k}(\xi)\hat{f}(\xi)$, so $T : L^2 \to L^2 \iff \hat{k} \in L^\infty$.*

[7] *The John–Nirenberg space of functions of bounded mean oscillation, BMO, is the collection of f (modulo constants) for which $\frac{1}{|Q|}\int_Q |f(x) - f_Q|\,dx < C < \infty$ for all cubes Q, and where f_Q is the average of the function over Q.*

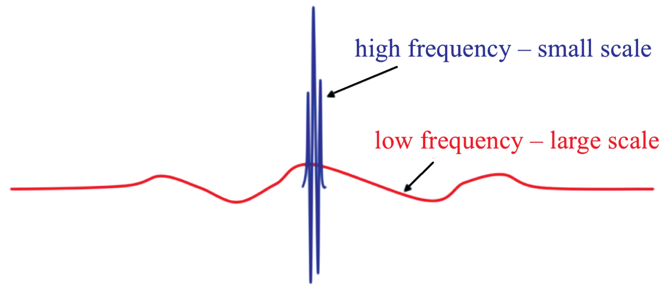


Figure 7. Waves with different frequencies do not interact much.

In fact, under suitable cancellations on the operator (to which the general theorem can be reduced), one can show that for some $\varepsilon > 0$, and, say, $\nu \le \mu$,

$$|\langle T(\phi_{\nu k}), \phi_{\mu l}\rangle| \lesssim \frac{2^{-(\mu-\nu)(\varepsilon+n/2)}}{(1+2^\nu|2^{-\nu}k-2^{-\mu}l|)^{n+\epsilon}}. \tag{4}$$

Or, in a different format, for $Q = Q_{\nu k}$ and $Q' = Q_{\mu l}$,

$$|\langle T(\phi_{\nu k}), \phi_{\mu l}\rangle| \lesssim \left(\frac{|Q'|}{|Q|}\right)^{\varepsilon+1/n} \frac{1}{(1+|Q|^{-1}\operatorname{diam}(Q\cup Q'))^{1+\varepsilon}},$$

since $1 + \frac{\operatorname{diam}(Q\cup Q')}{|Q|} \approx 1 + \frac{|2^{-\nu}k-2^{-\mu}l|}{2^{-\nu n}}$. We can see that as the cubes Q and Q' get far apart (the associated wavelets are concentrated away from each other) or as Q gets much larger than Q' (the wavelets live at very different scales, i.e., have very different frequencies), the entries on the matrix for T get very small. These estimates can then be easily used to show via the Schur test (a result about positive matrices) that the operator is bounded on L^2.

The reason behind the estimate (4) is the fact that the wavelets $\phi_{\nu k}$ are almost eigenfunctions for all CZOs. That is, they get transformed into functions which have very similar properties (they are called *smooth molecules* in the language of Frazier and Jawerth [13]). Roughly speaking, the functions $\{T(\phi_{\nu k})\}$ are concentrated on appropriate cubes and oscillate at certain sets of frequencies as wavelets do. Clearly $T(\phi_{\nu k})$ and $\phi_{\mu l}$ do not interact much if the cubes $Q_{\nu k}$ and $Q_{\mu l}$ are far apart, since they die out fast outside the cubes. When the cubes overlap the analysis is more delicate. We borrow Figure 7 from [5] and some of the narrative therein to illustrate the almost-orthogonality of the functions $T(\phi_{\nu k})$ and $\phi_{\mu l}$ when they are concentrated around overlapping cubes at very different scales. Suppose that, say, $\nu \ll \mu$. Then $T(\phi_{\nu k})$ (red in the figure) oscillates at a much larger scale than $\phi_{\mu l}$ (blue), so as in the cartoon about balls at the beginning of this article, they do not interact much. In fact, $\phi_{\mu l}$ highly oscillates where $T(\phi_{\nu k})$ does not vary much, and hence

$$\langle T(\phi_{\nu k}), \phi_{\mu l}\rangle = \int_{\mathbb{R}^n} T(\phi_{\nu k})(x)\phi_{\mu l}(x)\,dx \sim 0,$$

since we are integrating a function with mean zero against a function which is essentially constant. Here both the regularity of the wavelets and their cancellation play a crucial role. The vanishing moments of $\phi_{\mu l}$ can be combined with the smoothness of $T(\phi_{\nu k})$ to quantify the off-diagonal decay in the estimate (4).

Bilinear Pseudodifferential Operators

The best example of a bilinear operator we want to look at is again provided by the product of two functions. Let us write the product $f \cdot g$ in multiplier form, that is,

$$\begin{aligned}&(f \cdot g)(x)\\ &= (2\pi)^{-n} \int_{\mathbb{R}^n} \widehat{f}(\xi) e^{ix\cdot\xi}\, d\xi \cdot (2\pi)^{-n} \int_{\mathbb{R}^n} \hat{g}(\eta) e^{ix\cdot\eta}\, d\eta \\ &= (2\pi)^{-2n} \int_{\mathbb{R}^{2n}} \widehat{f}(\xi)\hat{g}(\eta) e^{ix\cdot(\xi+\eta)}\, d\xi d\eta.\end{aligned}$$

Similarly,

$$\begin{aligned}&\partial^\alpha (f.g)(x)\\ &= (2\pi)^{-2n} i^{|\alpha|} \int_{\mathbb{R}^{2n}} (\xi+\eta)^\alpha \widehat{f}(\xi)\hat{g}(\eta) e^{ix\cdot(\xi+\eta)} d\xi d\eta.\end{aligned}$$

More generally, and in analogy with the classical Hörmander classes of linear pseudodifferential operators, we can consider bilinear pseudodifferential operators of the form

$$T_\sigma(f,g)(x) = \int_{\mathbb{R}^{2n}} \sigma(x,\xi,\eta) \widehat{f}(\xi)\hat{g}(\eta) e^{ix\cdot(\xi+\eta)}\, d\xi d\eta.$$

For the purpose of this note, we say that the symbol σ is in the class BS^m, $m \in \mathbb{R}$, if

$$|\partial_x^\beta \partial_{\xi,\eta}^\alpha \sigma(x,\xi,\eta)| \le C_{\alpha,\beta}(1+|\xi|+|\eta|)^{m-|\alpha|}.$$

For example, when $m = 0$ the above estimates tell us that the more derivatives we take on the symbol in the frequency variables ξ and η, the faster the resulting function decays at infinity. On the other hand, derivatives in the space variable x do not have much of an impact on the decay of the symbols. The order m indicates the original number of pseudoderivatives the operator represents. These operators enjoy nice boundedness properties which could be obtained in several ways. At least for $m = 0$, they were pioneered by R. R. Coifman and Y. Meyer in the 1970s and 1980s; see for example [10]. We shall proceed to explain boundedness results for operators with symbols in BS^m along with some more recent developments.

It turns out that when the operators are of order zero ($m = 0$), they also admit a kernel representation of the form

$$T_\sigma(f,g)(x) = \int_{\mathbb{R}^{2n}} K(x,y,z) f(y) g(z)\, dydz,$$

at least for $x \notin \operatorname{supp} f \cap \operatorname{supp} g$. Here, K satisfies the estimates

$$|\partial^\alpha K(x,y,z)| \le C_\alpha (|x-y|+|x-z|)^{-(2n+|\alpha|)}. \qquad (5)$$

Note that if we write $\tilde{K}((x,x),(y,z)) = K(x,y,z)$, then $\tilde{K}$ satisfies the same estimates of a Calderón–Zygmund kernel in $\mathbb{R}^{2n}$ when its first $2n$-variables are of the form (x,x).

For a bilinear operator there are two naturally associated transposes, T^{*1} and T^{*2}, which are defined via the formulas

$$\langle T^{*1}(f,g), h\rangle = \langle T(h,g), f\rangle,$$

$$\langle T^{*2}(f,g), h\rangle = \langle T(f,h), g\rangle.$$

Bounded bilinear operators whose kernels and those of its transposes satisfy the estimates (5) are called bilinear Calderón–Zygmund operators (BCZOs). They are well understood by now through the works of R. R. Coifman and Y. Meyer [10], M. Christ and J. L. Journé [8], C. Kenig and E. Stein [21], and L. Grafakos and the author of this note [17].

Recall that by Hölder's inequality,

$$\|f\cdot g\|_{L^r} \le \|f\|_{L^p}\|g\|_{L^q}$$

for $1 \le p, q \le \infty$, and $1/p + 1/q = 1/r$. Similarly, for operators of order zero, and more general BCZOs, one expects the same estimates (except for the end-point ones when p and/or q are 1 or ∞). Indeed, a T(1,1)-Theorem for BCZOs exists. One version of it is the following.

T(1,1)-Theorem ([17]). *Let T be a bilinear operator so that its kernel and those of its transposes satisfy* (5), $1 < p, q < \infty$, *and* $1/r = 1/p + 1/q$. *Then*

$$T : L^p \times L^q \to L^r$$

if and only if

$$\sup_{\xi_1,\xi_2} (\|T^{*j}(e^{ix\cdot\xi_1}, e^{ix\cdot\xi_2})\|_{BMO}) \le C, \qquad j = 0,1,2,$$

where $T^{*0} = T$.

The classes of pseudodifferential operators with symbols in BS^m are closed by transpositions, so it is easy to apply the T(1,1)-Theorem to BS^0. One can check that $x \to T(e^{ix\cdot\xi_1}, e^{ix\cdot\xi_2}) = \sigma(x,\xi_1,\xi_2)$ is in $L^\infty \subset BMO$. This is clear at a formal level, since $\widehat{e^{ix\cdot\xi_j}}(\cdot) = \delta_{\xi_j}(\cdot)$, but it holds also in a rigorous way. Since T^{*1} and T^{*2} are also given through symbols in BS^0, similar computations apply to them. It follows then that $T_\sigma : L^p \times L^q \to L^r$ in the full range of exponents of the theorem.

For $m > 0$, results for BS^m could be obtained from the case of operators of order zero using the following technique, which can be traced back to Coifman–Meyer's work.

Let $\sigma \in BS^m$, $\phi \in C^\infty(\mathbb{R})$ such that $0 \le \phi \le 1$, $\operatorname{supp}\phi \subset [-2,2]$, and $\phi(r) + \phi(\frac{1}{r}) = 1$ on $[0,\infty)$, and let us write

$\langle \xi \rangle = (1+|\xi|^2)^{1/2}$. Then,

$$\begin{aligned} &T_\sigma(f,g)(x) \\ &= \int \sigma(x,\xi,\eta)\left(\phi\left(\frac{\langle\eta\rangle^2}{\langle\xi\rangle^2}\right)+\phi\left(\frac{\langle\xi\rangle^2}{\langle\eta\rangle^2}\right)\right) \\ &\qquad\qquad \times \widehat{f}(\xi)\hat{g}(\eta)e^{ix\cdot(\xi+\eta)}d\xi d\eta \\ &= \int \sigma(x,\xi,\eta)\phi\left(\frac{\langle\eta\rangle^2}{\langle\xi\rangle^2}\right)\frac{\widehat{J^m f}(\xi)}{\langle\xi\rangle^m}\hat{g}(\eta)e^{ix\cdot(\xi+\eta)}d\xi d\eta \\ &+ \int \sigma(x,\xi,\eta)\phi\left(\frac{\langle\xi\rangle^2}{\langle\eta\rangle^2}\right)\widehat{f}(\xi)\frac{\widehat{J^m g}(\eta)}{\langle\eta\rangle^m}e^{ix\cdot(\xi+\eta)}d\xi d\eta. \end{aligned}$$

That is, we have

$$T_\sigma(f,g) = T_{\sigma_1}(J^m f, g) + T_{\sigma_2}(f, J^m g), \tag{6}$$

where the symbols

$$\sigma_1(x,\xi,\eta) = \sigma(x,\xi,\eta)\phi\left(\frac{\langle\eta\rangle^2}{\langle\xi\rangle^2}\right)\langle\xi\rangle^{-m}$$

and

$$\sigma_2(x,\xi,\eta) = \sigma(x,\xi,\eta)\phi\left(\frac{\langle\xi\rangle^2}{\langle\eta\rangle^2}\right)\langle\eta\rangle^{-m}$$

are easily seen to be in BS^0. From the boundedness properties of T_{σ_j}, $j = 1,2$, it then follows that

$$\|T_\sigma(f,g)\|_{L^r} \lesssim \|f\|_{L^p_m}\|g\|_{L^q} + \|f\|_{L^p}\|g\|_{L^q_m}. \tag{7}$$

This was obtained in joint work with Á. Bényi and A. R. Nahmod [4]. Many other similar results for bilinear pseudodifferential operators exist for a variety of classes of symbols (see for example the work with Á. Bényi, F. Bernicot, D. Maldonado, and V. Naibo [2] for additional references).

There are also results for operators resembling the Bilinear Hilbert Transform (BHT) having multipliers (i.e., x-independent symbols) with singularities when $\xi = \eta$. The BHT is a much more singular bilinear operator whose study requires a sophisticated refinement of almost-orthogonality techniques beyond what is described here. The analysis of the BHT was developed in ground-breaking achievements of M. Lacey and C. Thiele [22] and several subsequent works. For reasons of space we shall not present more singular pseudodifferential operators with singularities similar to the BHT. Additional work on them has been done in numerous articles by many authors. See the works of F. Bernicot [6] and A. Bényi, C. Demeter, A. Nahmod, C. Thiele, P. Villarroya, and the author [3] for further references.

It is of interest to note, to draw a parallel to the linear situation, that under suitable cancellation conditions, boundedness properties for multipliers $\sigma(\xi,\eta)$ of order zero can also be obtained through the discretization techniques leading to almost-diagonal tensors. These methods also apply to other function spaces admitting LPS characterizations. This was first studied in [16] and other works followed. In particular, for Lebesgue spaces the estimates in the matrix of T_σ take the much more complicated form[8]

$$\begin{aligned} |\langle T_\sigma(\phi_{\nu k},\phi_{\mu l}),\phi_{\lambda m}\rangle| &\lesssim 2^{-(\max(\mu,\nu,\lambda)-\min(\mu,\nu,\lambda))\varepsilon} \\ &\times \frac{2^{(-\max(\mu,\nu,\lambda)+\mathrm{med}(\mu,\nu,\lambda)+\min(\mu,\nu,\lambda))n/2}}{(G(\nu,k,\mu,l)\,G(\mu,l,\lambda,m)\,G(\lambda,m,\nu,k))^N} \end{aligned}$$

for some $\varepsilon > 0$ and $N > n$, and where

$$G(x,a,y,b) = (1+2^{\min(x,y)}|2^{-x}a - 2^{-y}b|).$$

Despite its complicated appearance, the first factor on the right shows that the entries of the tensor get smaller as the sizes of the cubes where the wavelets are essentially supported differ from each other. On the other hand, the presence of the functions G on the denominator shows that the entries also get smaller as those cubes move farther away from each other.

Leibniz Rule for the Fractional Derivatives

The estimate (7) already has a Leibniz rule flavor as it estimates m pseudodifferential derivatives of the product in terms of two terms, each of which has only one function pseudodifferentiated. In general, similar estimates also hold for a large collection of spaces (such as Triebel–Lizorkin, Besov, Hardy, etc.) which have a *p-size* estimate for its elements and an *s-smoothness scale* (like L^p_s) . Given a triplet of such spaces X^p_s, Y^q_s, Z^r_s and T_σ of order m the estimates expected and often obtained are of the form

$$\|T_\sigma(f,g)\|_{Z^r_s} \lesssim \|f\|_{X^p_{m+s}}\|g\|_{Y^q_0} + \|f\|_{X^p_0}\|g\|_{Y^q_{m+s}},$$

provided $1/r = 1/p + 1/q$ and suitable restrictions on the other parameters hold (depending, in particular, on the class of pseudodifferential operators used). We refer to the work of V. Naibo and A. Thomson [26] for a more complete account of known results.

Let us go back to the actual fractional derivatives of a product, $D^s(f\cdot g)$. The estimate

$$\|D^s(f\cdot g)\|_{L^r} \lesssim \|D^s f\|_{L^p}\|g\|_{L^q} + \|f\|_{L^p}\|D^s g\|_{L^q} \tag{8}$$

has a long history preceding most of what we have described so far. It has also been successfully revisited over the years to extend the sets of exponents for which it holds. The original versions go back to the work of T. Kato and G. Ponce [19] and M. Christ and M. I. Weinstein [9]. The optimal range of exponents $1/2 < r < \infty$, $1 < p,q \le \infty$, $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$, and $s > \max(0, \frac{n}{r} - n)$ or s a positive even integer was obtained by C. Muscalu and W. Schlag [25] and L. Grafakos and S. Oh [15], together with appropriate weak-type estimates when $p = 1$ or $q = 1$. A detailed analysis of counterexamples for the limitations on s can be found also in [15]. The other end-point case $p = q = r = \infty$ was resolved by J. Bourgain and D. Li [7].

[8] *For* $a \le b \le c$, $\mathrm{med}(a,b,c) = b$.

The simple approach we showed leading to (6) does not quite work for $D^s(f\cdot g)$ for the optimal range of parameters involved in the estimate (7), so a different decomposition of the bilinear operator is used. Let ψ be a function such that $\widehat{\psi} = \widehat{\phi}^2$ as in (1). Writing $\widehat{f}(\xi) = \sum_j \widehat{f}(\xi)\widehat{\psi}(2^{-j}\xi)$, $\hat{g}(\eta) = \sum_k \hat{g}(\eta)\widehat{\psi}(2^{-k}\eta)$, grouping appropriate terms together, and introducing powers of the frequency variables and corresponding fractional derivatives, we can compute

$$
\begin{aligned}
&D^s(f\cdot g)(x)\\
&= \sum_j \sum_{k<j-3} \int e^{i(\xi+\eta)x} \frac{|\xi+\eta|^s}{|\xi|^s}\\
&\qquad \times \widehat{\psi}(2^{-j}\xi)\widehat{D^s f}(\xi)\widehat{\psi}(2^{-k}\eta)\hat{g}(\eta)d\xi\, d\eta\\
&+ \sum_k \sum_{j<k-3} \int e^{i(\xi+\eta)x} \frac{|\xi+\eta|^s}{|\eta|^s}\\
&\qquad \times \widehat{\psi}(2^{-j}\xi)\widehat{f}(\xi)\widehat{\psi}(2^{-k}\eta)\widehat{D^s g}(\eta)d\xi\, d\eta\\
&+ \sum_{\substack{j,k\\ |j-k|\le 2}} \int e^{i(\xi+\eta)x} \frac{|\xi+\eta|^s}{|\eta|^s}\\
&\qquad \times \widehat{\psi}(2^{-j}\xi)\widehat{f}(\xi)\widehat{\psi}(2^{-k}\xi)\widehat{D^s g}(\eta)d\xi\, d\eta\\
&= T_1(D^s f, g)(x) + T_2(f, D^s g)(x) + T_3(f, D^s g)(x),
\end{aligned}
$$

where the T_i for $i = 1, 2, 3$ are defined by the symbols

$$
\begin{aligned}
\sigma_1(\xi,\eta) &= \sum_j \sum_{k<j-3} \frac{|\xi+\eta|^s}{|\xi|^s}\widehat{\psi}(2^{-j}\xi)\widehat{\psi}(2^{-k}\eta),\\
\sigma_2(\xi,\eta) &= \sum_k \sum_{j<k-3} \frac{|\xi+\eta|^s}{|\eta|^s}\widehat{\psi}(2^{-j}\xi)\widehat{\psi}(2^{-k}\eta),\\
\sigma_3(\xi,\eta) &= \sum_{\substack{j,k\\ |j-k|\le 2}} \frac{|\xi+\eta|^s}{|\eta|^s}\widehat{\psi}(2^{-j}\xi)\widehat{\psi}(2^{-k}\eta).
\end{aligned}
$$

It will be enough then to show that for $i = 1, 2, 3$,

$$\|T_i(f,g)\|_{L^p} \lesssim \|f\|_{L^{p_1}}\|g\|_{L^{p_2}}. \tag{9}$$

Note that in the regions where the first two symbols are supported, we have, respectively, $|\xi| \gg |\eta|$ and $|\xi| \ll |\eta|$. Using this it is not hard to see they are actually Coifman–Meyer multipliers, that is,

$$|\partial^\alpha_{\xi,\eta}\sigma_i(\xi,\eta)| \le C_\alpha(|\xi|+|\eta|)^{-|\alpha|}, \tag{10}$$

$i = 1, 2$, which in turn are also bilinear Calderón–Zygmund operators. Hence T_1 and T_2 satisfy the right L^r estimates for the full range of exponents. The troublesome operator is the last one. If s is not an even integer, then its symbol develops a singularity along $\xi + \eta = 0$ for large derivatives, since now $|\xi| \approx |\eta|$. As proved in [15], T_3 belongs to a class of pseudodifferential operators introduced in [4] for which there is no general theory, and hence it needs to be studied on its own. Nonetheless, in both [25] and [15] it is shown that T_3 is bounded on the optimal range of exponents mentioned before. A crucial technical ingredient in the proof is an estimate on the dependence of the L^r norm of discrete square functions obtained by translates of a given LPS-function. We refer to the mentioned works for details.

It is of interest to note that the same approach works if instead of $D^\alpha(f,g)$ we consider $D^\alpha T_\sigma(f,g)$ with σ satisfying the m order version of (10). Moreover, using techniques introduced by M. Fujita and N. Tomita [14], one can even treat symbols where the pointwise regularity estimates (10) are replaced by Sobolev-type ones as well as symbols of different order; see [18]. J. Brummer and V. Naibo obtained similar results for smooth symbols using smooth molecular decomposition techniques and also weighted versions for multipliers (see again [26] for full references).

Fractional Derivatives in the Mixed Lebesgue Space Setting

Consider now the *mixed Lebesgue spaces* defined by the norm (quasi-norm if either $p < 1$ or $q < 1$)

$$\|f\|_{L^p_t L^q_x(\mathbb{R}^{n+1})} = \left(\int_{\mathbb{R}}\left(\int_{\mathbb{R}^n}|f(t,x)|^q\, dx\right)^{p/q} dt\right)^{1/p}.$$

There is a long history within harmonic analysis of results on mixed Lebesgue spaces going back to works of A. Benedek, A. P. Calderón, and R. Panzone; J. Rubio de Francia, A. Ruiz, and J. L. Torrea; and D. L. Fernández, to name a few. Recent interest arose from bilinear estimates needed in partial differential equations involving functions in n spatial variables x and one time variable t.

In this context, C. Kenig, G. Ponce, and L. Vega [20] established in their study of well-posedness for the generalized KdV equation the estimate

$$
\begin{aligned}
&\|D^s_x(fg) - fD^s_x(g) - gD^s_x(f)\|_{L^p_t L^q_x}\\
&\lesssim \|D^{s_1}_x(f)\|_{L^{p_1}_t L^{q_1}_x}\|D^{s_2}_x(g)\|_{L^{p_2}_t L^{q_2}_x},
\end{aligned}
$$

where now

$$\widehat{D^s_x f}(\xi_0,\xi_n) = |\xi_n|^s\hat{f}(\xi_0,\xi_n)$$

for $\xi_0 \in \mathbb{R}$, $\xi_n \in \mathbb{R}^n$. They proved this estimate for $0 < s_1, s_2 < 1$, $s = s_1 + s_2$, $1 < p, p_1, p_2, q, q_1, q_2 < \infty$, $1/p = 1/p_1 + 1/p_2$, and $1/q = 1/q_1 + 1/q_2$. Note that the above estimates only involve derivatives in x, but it is of interest to have Leibniz rules in mixed Lebesgue spaces involving derivatives in x and t.

For any s, $D^s_{t,x}$ is defined in $\mathbb{R}^{n+1}$ as before,

$$\widehat{D^s_{t,x}f}(\xi_0,\xi_n) = |(\xi_0,\xi_n)|^s\hat{f}(\xi_0,\xi_n) = |\xi|^s\hat{f}(\xi),$$

where now $\xi \in \mathbb{R}^{n+1}$ and the FT is taken also in $\mathbb{R}^{n+1}$. Clearly, this corresponds to fractional differentiation in $\mathbb{R}^{n+1}$, but the mixed norms used to estimate them create many difficulties.

Consider then a bilinear multiplier in $\mathbb{R}^{2(n+1)}$,

$$\mathbf{T}_\sigma(f,g)(x) = \int_{\mathbb{R}^{2(n+1)}} \sigma(\xi,\eta) e^{ix\cdot(\xi+\eta)} \hat{f}(\xi)\hat{g}(\eta)\, d\xi d\eta,$$

where the symbol satisfies

$$|\partial^\beta_{\xi,\eta}\sigma(\xi,\eta)| \lesssim_\beta (|\xi|+|\eta|)^{-|\beta|}. \tag{11}$$

In the author's joint work with J. Hart and X. Wu [18], it was shown that

$$\begin{aligned}\|D^s_{t,x}T_\sigma(f,g)\|_{L^pL^q} &\lesssim \|D^s_{t,x}f\|_{L^{p_1}L^{q_1}}\|g\|_{L^{p_2}L^{q_2}} \\ &\quad + \|f\|_{L^{p_1}L^{q_1}}\|D^s_{t,x}g\|_{L^{p_2}L^{q_2}}\end{aligned} \tag{12}$$

for $1 < p_i, q_i < \infty$, $i = 1,2$, $1/p = 1/p_1 + 1/p_2$, $1/q = 1/q_1 + 1/q_2$, and some restrictions in s which are rather technical to explain in this article.[9] Of course a particular case is that of $\sigma = 1$, which gives again the Leibniz rule for the product but now in mixed Lebesgue spaces,

$$\begin{aligned}\|D^s_{t,x}(f\cdot g)\|_{L^pL^q} &\lesssim \|D^s_{t,x}f\|_{L^{p_1}L^{q_1}}\|g\|_{L^{p_2}L^{q_2}} \\ &\quad + \|f\|_{L^{p_1}L^{q_1}}\|D^s_{t,x}g\|_{L^{p_2}L^{q_2}}.\end{aligned} \tag{13}$$

The estimate in (12) holds for less regular symbols in the sense of [14] too.

The approach to the estimate (12) follows the one described in the Lebesgue case but making use of the vector-value Calderón–Zygmund theory. In particular, one also needs to use the square function estimate

$$\|f\|_{L^p_tL^q_x(\mathbb{R}\times\mathbb{R}^n)} \approx \left\| \left(\sum_{j\in\mathbb{Z}} |f*\phi_{2^{-j}}|^2 \right)^{1/2} \right\|_{L^p_tL^q_x(\mathbb{R}\times\mathbb{R}^n)} \tag{14}$$

for $1 < p,q < \infty$ and an appropriate LPS-function ϕ in $\mathbb{R}^{n+1}$.

A limited range of exponents when $p,q > 1$ in (13) was previously obtained in joint work with E. Ward [27]. A main obstacle to reach the full range of exponent was the need to establish one side of the estimate (14) when either $0 < p \le 1$ or $0 < q \le 1$.

It is a classical result in the Lebesgue context that

$$\|f\|_{L^p} \lesssim \left\| \left(\sum_{\nu} |f*\phi_{2^{-\nu}}|^2 \right)^{1/2} \right\|_{L^p} \tag{15}$$

also holds for $0 < p \le 1$. In fact, the right-hand side of (15) gives a characterization of the Hardy space H^p (and hence the reverse estimate cannot hold in this range). The analogous estimates for mixed Lebesgue spaces,

$$\|f\|_{L^p_tL^q_x(\mathbb{R}\times\mathbb{R}^n)} \lesssim \left\| \left(\sum_{j\in\mathbb{Z}} |f*\phi_{2^{-j}}|^2 \right)^{1/2} \right\|_{L^p_tL^q_x(\mathbb{R}\times\mathbb{R}^n)}$$

when $p \le 1$ or $q \le 1$, was established in [18]. The details are rather technical to be sketched in this note, but it is of interest to point out that the proof presented therein relies again on almost-orthogonality and wavelet-type decompositions in the mixed norm setting, circling back to some of the ideas we already described in this presentation.

Before we conclude we would like to point to other important results in this context. Several authors have considered iterated fractional derivative variations of the Leibniz rule instead. In the case of Lebesgue spaces, the first iterated derivative versions were studied by C. Muscalu, J. Pipher, T. Tao, and C. Thiele [24]. In the mixed norm setting, iterated derivative versions of (13) exist too. For example, let D^s_x and D^s_t be the fractional derivatives in the respective one-dimensional variables x and t. In a series of works culminating with C. Benea and C. Muscalu [1] and F. Di Plinio and Y. Ou [12], it was proved that

$$\begin{aligned}\|D^\beta_tD^\alpha_x(fg)\|_{L^pL^q} &\lesssim \|f\|_{L^{p_1}L^{q_1}}\|D^\beta_tD^\alpha_xg\|_{L^{p_2}L^{q_2}} \\ &\quad + \|D^\beta_tf\|_{L^{p_1}L^{q_1}}\|D^\alpha_xg\|_{L^{p_2}L^{q_2}} \\ &\quad + \|D^\alpha_xf\|_{L^{p_1}L^{q_1}}\|D^\beta_tg\|_{L^{p_2}L^{q_2}} \\ &\quad + \|D^\beta_tD^\alpha_xf\|_{L^{p_1}L^{q_1}}\|g\|_{L^{p_2}L^{q_2}}.\end{aligned} \tag{16}$$

The largest range can be found in [1], where for $n = 1$, $1/2 < p < \infty$, $1/2 < q < \infty$, $\frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2}$, $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$, one has $\alpha > \max(0, \frac{1}{p} - 1, \frac{1}{q} - 1)$ and $\beta > \max(0, \frac{1}{q} - 1)$. The methods used in [1] or [12] are substantially different from the direct ones we have described here and allow also for some end-point estimates, which we do not know if the work in [18] can produce. On the other hand, (16) does not imply (13) or the other way around.

It would be of interest to consider versions of (12) or the other ones presented in spaces of smooth functions defined also in terms of mixed norms. We believe that the successful almost-orthogonality techniques we have tried to illustrate here have still a lot of potential in the characterization of function spaces and the study of operators with less regular kernels, commutators, weighted estimates, and other topics of harmonic analysis in the mixed norm context.

ACKNOWLEDGMENTS. The author would like to thank Árpád Bényi, Estela A. Gavosto, Virginia Naibo, and the anonymous reviewers for comments and suggestions which greatly improved this presentation.

[9] *The actual restrictions are* $s \in 2\mathbb{N}$ *or*

$$s > \max\left(0, \frac{n+1}{p} - (n+1), \frac{n+1}{q} - (n+1)\right).$$

References

[1] C. Benea and C. Muscalu, *Quasi-Banach valued inequalities via the helicoidal method*, J. Funct. Anal. **273** (2017), no. 4, 1295–1353, DOI 10.1016/j.jfa.2017.04.014. MR3661402

[2] Á. Bényi, F. Bernicot, D. Maldonado, V. Naibo, and R. H. Torres, *On the Hörmander classes of bilinear pseudodifferential operators II*, Indiana Univ. Math. J. **62** (2013), no. 6, 1733–1764, DOI 10.1512/iumj.2013.62.5168. MR3205530

[3] Á. Bényi, C. Demeter, A. R. Nahmod, C. M. Thiele, R. H. Torres, and P. Villarroya, *Modulation invariant bilinear T(1) theorem*, J. Anal. Math. **109** (2009), 279–352, DOI 10.1007/s11854-009-0034-z. MR2585397

[4] Á. Bényi, A. R. Nahmod, and R. H. Torres, *Sobolev space estimates and symbolic calculus for bilinear pseudodifferential operators*, J. Geom. Anal. **16** (2006), no. 3, 431–453, DOI 10.1007/BF02922061. MR2250054

[5] Á. Bényi and R. H. Torres, *The discrete Calderón reproducing formula of Frazier and Jawerth*, Functional analysis, harmonic analysis, and image processing: a collection of papers in honor of Björn Jawerth, Contemp. Math., vol. 693, Amer. Math. Soc., Providence, RI, 2017, pp. 79–107, DOI 10.1090/conm/693. MR3682606

[6] F. Bernicot, *A bilinear pseudodifferential calculus*, J. Geom. Anal. **20** (2010), no. 1, 39–62, DOI 10.1007/s12220-009-9105-8. MR2574719

[7] J. Bourgain and D. Li, *On an endpoint Kato-Ponce inequality*, Differential Integral Equations **27** (2014), no. 11-12, 1037–1072. MR3263081

[8] M. Christ and J.-L. Journé, *Polynomial growth estimates for multilinear singular integral operators*, Acta Math. **159** (1987), no. 1-2, 51–80, DOI 10.1007/BF02392554. MR906525

[9] F. M. Christ and M. I. Weinstein, *Dispersion of small amplitude solutions of the generalized Korteweg-de Vries equation*, J. Funct. Anal. **100** (1991), no. 1, 87–109, DOI 10.1016/0022-1236(91)90103-C. MR1124294

[10] R. R. Coifman and Y. Meyer, *Au delà des opérateurs pseudo-différentiels* (French), Astérisque, vol. 57, Société Mathématique de France, Paris, 1978. With an English summary. MR518170

[11] G. David and J.-L. Journé, *A boundedness criterion for generalized Calderón-Zygmund operators*, Ann. of Math. (2) **120** (1984), no. 2, 371–397, DOI 10.2307/2006946. MR763911

[12] F. Di Plinio and Y. Ou, *Banach-valued multilinear singular integrals*, Indiana Univ. Math. J. **67** (2018), no. 5, 1711–1763, DOI 10.1512/iumj.2018.67.7466. MR3875242

[13] M. Frazier and B. Jawerth, *A discrete transform and decompositions of distribution spaces*, J. Funct. Anal. **93** (1990), no. 1, 34–170, DOI 10.1016/0022-1236(90)90137-A. MR1070037

[14] M. Fujita and N. Tomita, *Weighted norm inequalities for multilinear Fourier multipliers*, Trans. Amer. Math. Soc. **364** (2012), no. 12, 6335–6353, DOI 10.1090/S0002-9947-2012-05700-X. MR2958938

[15] L. Grafakos and S. Oh, *The Kato-Ponce inequality*, Comm. Partial Differential Equations **39** (2014), no. 6, 1128–1157, DOI 10.1080/03605302.2013.822885. MR3200091

[16] L. Grafakos and R. H. Torres, *Discrete decompositions for bilinear operators and almost diagonal conditions*, Trans. Amer. Math. Soc. **354** (2002), no. 3, 1153–1176, DOI 10.1090/S0002-9947-01-02912-9. MR1867376

[17] L. Grafakos and R. H. Torres, *Multilinear Calderón-Zygmund theory*, Adv. Math. **165** (2002), no. 1, 124–164, DOI 10.1006/aima.2001.2028. MR1880324

[18] J. Hart, R. H. Torres, and X. Wu, *Smoothing properties of bilinear operators and Leibniz-type rules in Lebesgue and mixed Lebesgue spaces*, Trans. Amer. Math. Soc. **370** (2018), no. 12, 8581–8612, DOI 10.1090/tran/7312. MR3864388

[19] T. Kato and G. Ponce, *Commutator estimates and the Euler and Navier-Stokes equations*, Comm. Pure Appl. Math. **41** (1988), no. 7, 891–907, DOI 10.1002/cpa.3160410704. MR951744

[20] C. E. Kenig, G. Ponce, and L. Vega, *Well-posedness and scattering results for the generalized Korteweg-de Vries equation via the contraction principle*, Comm. Pure Appl. Math. **46** (1993), no. 4, 527–620, DOI 10.1002/cpa.3160460405. MR1211741

[21] C. E. Kenig and E. M. Stein, *Multilinear estimates and fractional integration*, Math. Res. Lett. **6** (1999), no. 1, 1–15, DOI 10.4310/MRL.1999.v6.n1.a1. MR1682725

[22] M. Lacey and C. Thiele, *L^p estimates for the bilinear Hilbert transform*, Proc. Nat. Acad. Sci. U.S.A. **94** (1997), no. 1, 33–35, DOI 10.1073/pnas.94.1.33. MR1425870

[23] Y. Meyer, *Ondelettes et opérateurs. I: Ondelettes. [Wavelets]* (French), Actualités Mathématiques. [Current Mathematical Topics], Hermann, Paris, 1990. MR1085487

[24] C. Muscalu, J. Pipher, T. Tao, and C. Thiele, *Bi-parameter paraproducts*, Acta Math. **193** (2004), no. 2, 269–296, DOI 10.1007/BF02392566. MR2134868

[25] C. Muscalu and W. Schlag, *Classical and multilinear harmonic analysis. Vol. I*, Cambridge Studies in Advanced Mathematics, vol. 137, Cambridge University Press, Cambridge, 2013. MR3052498

[26] V. Naibo and A. Thomson, *Coifman-Meyer multipliers: Leibniz-type rules and applications to scattering of solutions to PDEs*, Trans. Amer. Math. Soc. **372** (2019), no. 8, 5453–5481, DOI 10.1090/tran/7866. MR4014283

[27] R. H. Torres and E. L. Ward, *Leibniz's rule, sampling and wavelets on mixed Lebesgue spaces*, J. Fourier Anal. Appl. **21** (2015), no. 5, 1053–1076, DOI 10.1007/s00041-015-9397-y. MR3393695

Rodolfo H. Torres

Credits

Opening photo and Figures 2–6 are courtesy of Rodolfo H. Torres.

Figure 1 is courtesy of Rodolfo H. Torres. Previously published in Rodolfo H. Torres and Richard O. Prum, *Análisis espectral de nanoestructuras en tejidos biológicos*, Matematicalia.net, vol. 1 (2), 2005.

Figure 7 is courtesy of Rodolfo H. Torres. Previously published in M. Cwikel and M. Milman, ed., *Functional Analysis, Harmonic Analysis, and Image Processing: A Collection of Papers in Honor of Björn Jawerth*, Amer. Math. Soc., 2017.

Author photo is courtesy of Stan Lim / University of California, Riverside.