January 1996

Differentiate and make waves

R. Nillsen
University of Wollongong, nillsen@uow.edu.au

Differentiate and make waves

Abstract

In 1972 Gary Meisters and Wolfgang Schmidt showed that if the Fourier coefficient of a function is zero at the origin, then the function is a sum of three first order differences. They deduced from this that every translation invariant linear form on the space of square integrable functions on the circle group is automatically continuous. Their results were subsequently developed by the author for the case of n-dimensional Euclidean space (Lectures Notes in Mathematics, vol. 1586, Springer-Verlag,1994). The present paper gives an exposition of some of the ideas in this work, with the aim of minimising the technicalities. In particular, it is shown that if we differentiate s times the functions in the Sobolev space of order s on the real line, we obtain as the range of this operator a difference space of order s, where the form of the differences in this space can be explicitly described. This difference space is a Hilbert space, and the operation of differentiation s times is an isometry between these spaces. Whereas finite differences have long been used to approximate derivatives "locally", these results reveal the precise "global" relationship. The main technique involves understanding how the operation of taking a derivative or a finite difference of a function affects the behaviour of the Fourier transform of the function near the origin on the real line.

Keywords

differentiation, derivative, finite difference, Fourier, Fourier transform, Sobolev space, Hilbert space, distribution, wave, convolution, origin, difference space, range, operator

Disciplines

Physical Sciences and Mathematics

Publication Details

This article was originally published as Nillsen, R, Differentiate and make wabes, Expositiones Mathematicae, 14, 1996, 57-84. Original journal avaiable here.

Differentiate And Make Waves

Rodney Nillsen

1. INTRODUCTION. If we differentiate a function, what sort of function do we obtain? Consider the function whose value at x is e^{-x^2}. If we think of its graph as a wave, then it has a crest; that is, a local maximum, at the origin. It has no trough; that is, no local minimum. However, differentiation of e^{-x^2} produces the function $-2xe^{-x^2}$, which has a crest at 1/2 and a trough at $-1/2$. Thus, in this case, differentiation has maintained the crest present in the original wave, but also has produced a trough which was not there before. In this sense, differentiation of the function e^{-x^2} has "made waves", as indicated in Figures 1 and 2.

Another way of "making waves" is by taking finite differences. Let $L^2(\mathbb{R})$ denote the usual space of complex valued, measurable, square integrable functions on the real line $\mathbb{R}$. That is, if f is complex valued and measurable on $\mathbb{R}$, $f \in L^2(\mathbb{R})$ if and only if $\int_{-\infty}^{\infty} |f(x)|^2 dx < \infty$. If $y \in \mathbb{R}$ and $f \in L^2(\mathbb{R})$, $\delta_y * f$ denotes the *translation* of f by y. That is, $(\delta_y * f)(x) = f(x-y)$, for $x \in \mathbb{R}$. Translation of a function by y maintains the shape of the graph of the function, but it changes the position of the graph, which moves through a horizontal amount y. As indicated by the notation, $\delta_y * f$ is the convolution of f with the Dirac measure δ_y at y. A *first order difference* is a function in $L^2(\mathbb{R})$ which is of the form $f - \delta_y * f$, for some $f \in L^2(\mathbb{R})$ and $y \in \mathbb{R}$.

Now consider the function whose value at x is $e^{-x^2} - e^{-(x-1)^2}$. This is a first order difference obtained from the function e^{-x^2}. Like the derivative of e^{-x^2}, this function has a crest and a trough: the crest is at x_0 and the trough is at x_1, where $x_0 < x_1$ and x_0 and x_1 are the solutions of $x/(x-1) = e^{2x-1}$. Thus, in the same sense as differentiation, forming a first order difference from the function e^{-x^2} has "made waves". This is indicated in Figure 3.

In contrast, however, note that if f is a polynomial function of degree n, it may have as many as $n-1$ crests and troughs, while its derivative is a polynomial function of degree $n-1$ which can have at most $n-2$ crests and troughs. In such a case, differentiation has produced a less wave-like function. Also, if g denotes for the moment the usual sine function, $g - \delta_{2\pi} * g$ is the zero function, which obviously has no crests or troughs whatsoever. In this case, taking one finite difference has

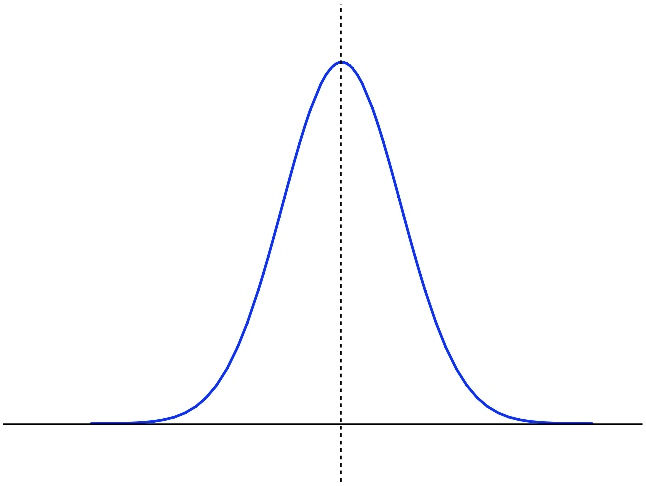

Figure 1: the function e^{-x^2}

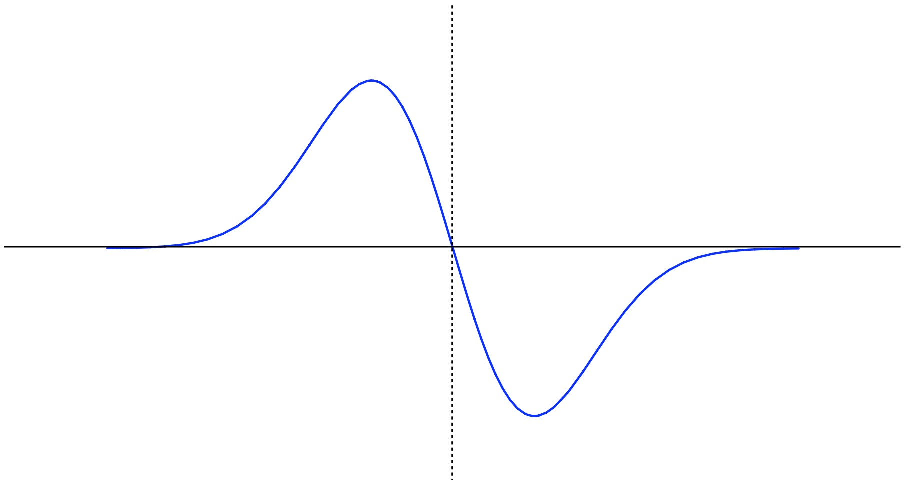

Figure 2: the function $-2xe^{-x^2}$, the derivative of x^2

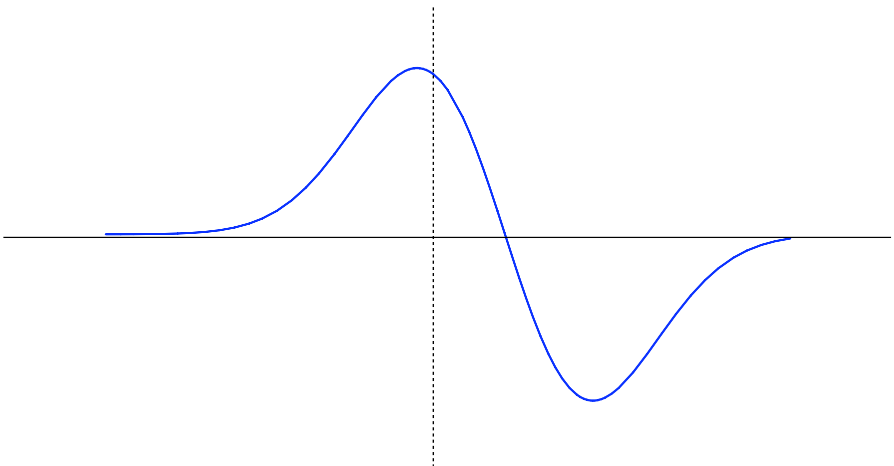

Figure 3: the function $e^{-x^2} - e^{-(x-1)^2}$, a first order difference obtained from e^{-x^2}

completely eliminated the waves in the original function.

The point about such examples as the preceding is that neither the functions involved nor their derivatives are restricted in their behaviour towards ∞ or $-\infty$. To gain some insight into what happens in general when some such restriction is made, we use the fact that a continuous real valued function which vanishes at, or towards, the end points of an interval attains a maximum or a minimum on the interval. Now consider a real valued continuously differentiable function f on $\mathbb{R}$. Let $D(f)$ denote the derivative of f and assume that $\lim_{n\to\infty} D(f)(x) = 0$ and $\lim_{n\to-\infty} D(f)(x) = 0$. These conditions may be thought of as saying "$D(f)(\infty) = 0$" and "$D(f)(-\infty) = 0$". Assume further that there are n distinct points in $\mathbb{R}$, at each of which f has a crest or a trough. At these points, $D(f)$ vanishes but, as well, $D(f)$ also vanishes at ∞ and $-\infty$ in the sense described above. Thus, there is a total of $n+2$ points, including ∞ and $-\infty$, at each of which $D(f)$ vanishes. As $D(f)$ is continuous, $D(f)$ will have a crest or a trough between any two such points, and there are $n+1$ pairs of consecutive such points. This produces $n+1$ distinct points, at each of which $D(f)$ has a crest or a trough. Thus, whereas f had a crest or a trough at n distinct points in $\mathbb{R}$, $D(f)$ has a crest or a trough at $n+1$ distinct points in $\mathbb{R}$. Hence the total number of crests and troughs has increased under differentiation from n to $n+1$.

In a very entertaining article [18], Professor Robert Strichartz described "How to make wavelets". Any readers of his article who may have wished to go even further, by making waves, may be reassured to know that both differentiation and taking finite differences of functions suffice not only to make wavelets, but waves as well! In fact, in a certain precise sense, the procedures of differentiation and taking first order differences produce exactly the same space of functions.

THEOREM 1. *A function in $L^2(\mathbb{R})$ is the derivative of a function in $L^2(\mathbb{R})$ if and only if it is a finite sum of first order differences.*

Here, as explained in Section 2, the derivative of an $L^2(\mathbb{R})$ function is taken in the sense of the theory of distributions, as developed by Laurent Schwartz [17]. If a function in $L^2(\mathbb{R})$ is differentiable in the usual sense and has a derivative in $L^2(\mathbb{R})$, the usual derivative equals the distributional derivative.

When higher order derivatives are considered, it might be expected that these are expressible in terms of higher order differences, and this is indeed the case, as now indicated. A *second order difference* is a function in $L^2(\mathbb{R})$ which is of the form $f - \frac{1}{2}(\delta_y + \delta_{-y}) * f$, for some $f \in L^2(\mathbb{R})$ and $y \in \mathbb{R}$. Figures 4 and 5 indicate that taking the second derivative, or taking a second order difference, again "makes waves" in a similar sense to taking one derivative or a first order difference. More specifically, taking either the second derivative of e^{-x^2} or the given second order difference of e^{-x^2} produces an extra trough as occurred in the case of first derivatives and differences, but produces as well an additional crest or trough.

THEOREM 2. *A function in $L^2(\mathbb{R})$ is the second derivative of some function in $L^2(\mathbb{R})$ if and only if it is a finite sum of second order differences.*

Theorem 1 can be regarded as describing the range of the differentiation

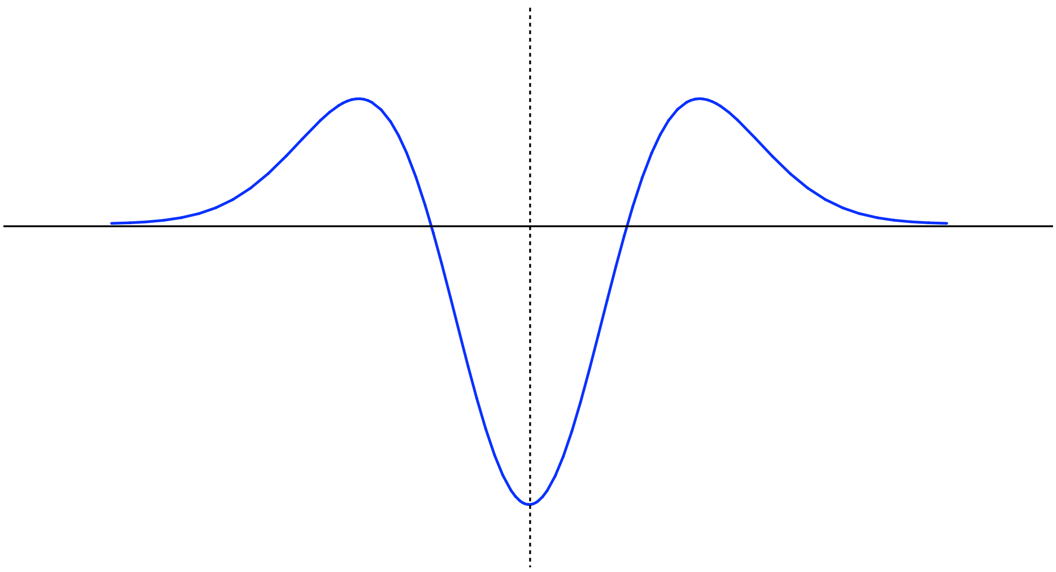

Figure 4: the function $2(2x^2-1)e^{-x^2}$, the second derivative of e^{-x^2}

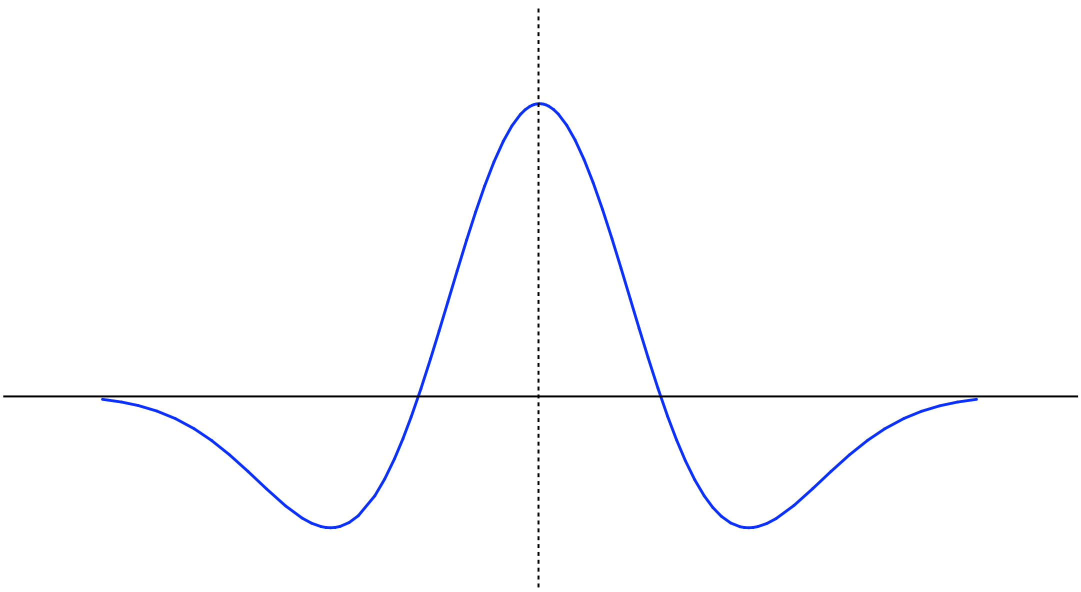

Figure 5: the function $e^{-x^2} - 2^{-1}\left(e^{-(x+1)^2} + e^{-(x-1)^2}\right)$,
a second order difference obtained from e^{-x^2}

operator D on the subspace of $L^2(\mathbb{R})$ comprising those functions whose derivatives are in $L^2(\mathbb{R})$. Theorem 2 gives a corresponding description of the range of D^2. One aim of this paper is to present a more general description of the relationship between derivatives and differences, and this description depends upon Fourier transform techniques. The crux of the matter is indicated by the following observation: if g is in $L^2(\mathbb{R})$, and if either g is the derivative of a function in $L^2(\mathbb{R})$ or if g is a first order difference, then the Fourier transform $\widehat{g}$ of g has a precise behaviour near the origin, as expressed by the fact that $\displaystyle\int_{-\infty}^{\infty} |\widehat{g}(x)|^2 |x|^{-2} dx < \infty$. In fact, as shall be seen, this behaviour characterizes both the derivatives of $L^2(\mathbb{R})$-functions and the functions which are finite sums of first order differences, which leads to the conclusion of Theorem 1.

The classical definition of the derivative suggests that the derivative of a function at a point may be approximated by the value of a first order difference at that point. This idea underlies various methods for calculating approximate solutions of differential equations. A typical illustration of this idea is the follwing: points are selected and the derivatives of the solution of the equation at these points are approximated by finite difference expressions – these approximations are substituted into the original equation, and the new equation so obtained may be solved to obtain approximate values of the solution at the selected points. Such methods are discussed, for example, in Chapter 5 of the book by Uri Ascher, Robert Mattheij and Robert Russell [1]. The underlying idea in these methods is the *local* approximation of derivatives by finite difference expressions. In contradistinction, a main theme here is to describe the relationship between derivatives of functions and finite differences when these are considered *as functions* on $\mathbb{R}$ – that is, the concern is with how derivatives of functions, and finite differences of functions, are related *globally* to each other.

2. DERIVATIVES AND THE FOURIER TRANSFORM. In this section is presented some background on the Fourier transform and Sobolev spaces. Complete descriptions of all necessary results are given, and details of those proofs which are not given may be found in a standard work such as the book by Walter Rudin [16].

Let $\mathbb{Z}$ denote the integers, let $\mathbb{Z}_+$ denote the non-negative integers, and let $\mathbb{N}$ denote the strictly positive integers. A complex valued function φ on $\mathbb{R}$ is said to be *rapidly decreasing* if $p\varphi$ is bounded for each polynomial p. The space $\mathcal{S}(\mathbb{R})$ is defined to consist of all infinitely differentiable functions φ on $\mathbb{R}$ which are rapidly decreasing and whose derivatives of all orders are rapidly decreasing. Then if $k, n \in \mathbb{Z}_+$, the function

$$\varphi \longmapsto \sup \left\{ (1 + |x|^2)^n |D^k(\varphi)(x)| : x \in \mathbb{R} \right\}$$

defines a seminorm on $\mathcal{S}(\mathbb{R})$. The countable family of seminorms so obtained determines a topology on $\mathcal{S}(\mathbb{R})$, the weakest in which these seminorms are continuous, in which $\mathcal{S}(\mathbb{R})$ becomes a complete metrizable, locally convex space; that is, a Fréchet space. A *tempered distribution* is defined to be a continuous linear functional on $\mathcal{S}(\mathbb{R})$. That is, the tempered distributions form the continuous dual $\mathcal{S}(\mathbb{R})'$ of $\mathcal{S}(\mathbb{R})$.

Note that $\mathcal{S}(\mathbb{R}) \subseteq L^2(\mathbb{R})$ and that, if $f \in L^2(\mathbb{R})$, the function on $\mathcal{S}(\mathbb{R})$ given by

$$\varphi \longmapsto \int_{-\infty}^{\infty} f(x)\varphi(x)dx,$$

defines a tempered distribution. In this way, $L^2(\mathbb{R})$ may be considered to be a vector subspace of $\mathcal{S}(\mathbb{R})'$.

If $\varphi_1, \varphi_2 \in \mathcal{S}(\mathbb{R})$, an integration by parts shows that

$$\int_{-\infty}^{\infty} D(\varphi_1)(x)\varphi_2(x)dx = -\int_{-\infty}^{\infty} \varphi_1(x)D(\varphi_2)(x)dx. \tag{2.1}$$

This observation suggests the following definition for the derivative $D(u)$ of a tempered distribution u. Noting that $\varphi \longmapsto D(\varphi)$ is continuous from $\mathcal{S}(\mathbb{R})$ into $\mathcal{S}(\mathbb{R})$, $D(u)$ is defined to be the tempered distribution $-u \circ D$, where $\circ$ denotes composition of functions. That is,

$$D(u)(\varphi) = -u(D(\varphi)), \text{ for all } \varphi \in \mathcal{S}(\mathbb{R}^n).$$

Then (2.1) shows that this definition is consistent with the usual concept of the derivative. Thus, D has been extended from $\mathcal{S}(\mathbb{R})$ to $\mathcal{S}(\mathbb{R})'$ and D maps $\mathcal{S}(\mathbb{R})'$ into $\mathcal{S}(\mathbb{R})'$. This means that $D^k(u)$ may be defined for $k \in \mathbb{Z}_+$ and $u \in \mathcal{S}(\mathbb{R})'$ by taking the k-fold composition of D.

If $f \in L^2(\mathbb{R})$, its derivative is in $L^2(\mathbb{R})$ if there is $g \in L^2(\mathbb{R})$ such that

$$\int_{-\infty}^{\infty} g(x)\varphi(x)\ dx = -\int_{-\infty}^{\infty} f(x)D(\varphi)(x)\ dx,$$

for all $\varphi \in \mathcal{S}(\mathbb{R})$. In this case, $g = D(f)$. If $f \in L^2(\mathbb{R})$ it is not generally true that $D(f) \in L^2(\mathbb{R})$. This leads to consideration of those functions in $L^2(\mathbb{R})$ whose derivatives, up to some given order, *do* belong to $L^2(\mathbb{R})$.

DEFINITION. Let $s \in \mathbb{Z}_+$. Then

$$H^s(\mathbb{R}) = \Big\{ f : f \in L^2(\mathbb{R}),\ D(f) \in L^2(\mathbb{R}), \dots, D^s(f) \in L^2(\mathbb{R}) \Big\}.$$

$H^s(\mathbb{R})$ is called the *Sobolev space of order s*. It is an important fact that the Sobolev spaces are Hilbert spaces, with inner products involving the Fourier transform. Relevant properties of the Fourier transform are now described.

If $\varphi \in \mathcal{S}(\mathbb{R})$, the Fourier transform $\widehat{\varphi}$ of φ is given by

$$\widehat{\varphi}(x) = \int_{-\infty}^{\infty} e^{-ixy}\varphi(y)dy, \quad \text{for all } x \in \mathbb{R}. \tag{2.2}$$

The Fourier transform is a continuous linear bijection on $\mathcal{S}(\mathbb{R})$. Thus, if $u \in \mathcal{S}(\mathbb{R})'$, the function $\varphi \longmapsto u(\widehat{\varphi})$ on $\mathcal{S}(\mathbb{R})$ is in $\mathcal{S}(\mathbb{R})'$ and is taken as the definition of the Fourier transform $\widehat{u}$ of u. If $u \in \mathcal{S}(\mathbb{R})$, this definition of $\widehat{u}$ is consistent with (2.2). The Fourier transform of f may also be denoted by $f^\wedge$. The $L^2(\mathbb{R})$-norm of a

function f in $L^2(\mathbb{R})$ is given by $||f||_2 = (\int_{-\infty}^{\infty} |f(x)|^2 \, dx)^{1/2}$. When the Fourier transform is restricted to $L^2(\mathbb{R})$, the following fundamental result obtains.

PLANCHEREL'S THEOREM. *The Fourier transform is a continuous linear bijection on $L^2(\mathbb{R})$ and*

$$||\widehat{f}||_2 = \sqrt{2\pi}||f||_2, \quad \textit{for all } f \in L^2(\mathbb{R}).$$

By regarding an integral as a form of infinite sum, $||\widehat{f}||_2^2$ can be interpreted as a "sum of squares". Then, if $\widehat{f}$ is regarded as exhibiting the "orthogonal components" of f, the equation $||\widehat{f}||_2 = \sqrt{2\pi}||f||_2$ in Plancherel's Theorem may be regarded as the form of Pythagoras' Theorem which appears in the context of Fourier analysis.

It is now possible to characterize the spaces $H^s(\mathbb{R})$, and to give a Fourier transform description of the action of D^s on $H^s(\mathbb{R})$.

THE SOBOLEV SPACE THEOREM. *Let $s \in \mathbb{Z}_+$. Then if $f \in L^2(\mathbb{R}), f \in H^s(\mathbb{R})$ if and only if $\int_{-\infty}^{\infty} |\widehat{f}(x)|^2 |x|^{2s} dx < \infty$. The space $H^s(\mathbb{R})$ is a Hilbert space in the inner product $\langle \cdot, \cdot \rangle$ given by*

$$\langle f, g \rangle = \int_{-\infty}^{\infty} \widehat{f}(x)\overline{\widehat{g}(x)}(1 + |x|^2)^s dx, \qquad \textit{for } f, g \in H^s(\mathbb{R}).$$

Also, for each $f \in H^s(\mathbb{R})$,

$$[D^s(f)]^\wedge(x) = (ix)^s \widehat{f}(x), \quad \textit{for almost all } x \in \mathbb{R}. \tag{2.3}$$

PROOF. Let $f \in L^2(\mathbb{R})$ and $\varphi \in \mathcal{S}(\mathbb{R})$. Then

$$\begin{aligned}
[D(f)]^\wedge(\varphi) &= D(f)(\widehat{\varphi}), \\
&= -f(D(\widehat{\varphi})), \\
&= -\int_{-\infty}^{\infty} f(x)[y \longmapsto -iy\varphi(y)]^\wedge(x) dx, \qquad \text{by differentiating in (2.2),} \\
&= i \int_{-\infty}^{\infty} \widehat{f}(x) x \varphi(x) dx.
\end{aligned}$$

Now by Plancherel's Theorem, $D(f) \in L^2(\mathbb{R})$ if and only if $D(f)^\wedge \in L^2(\mathbb{R})$, so it follows that $D(f) \in L^2(\mathbb{R})$ if and only if $\int_{-\infty}^{\infty} |\widehat{f}(x)|^2 |x|^2 dx < \infty$, and that in this case $[D(f)]^\wedge(x) = ix\widehat{f}(x)$. This proves the first statement in the Theorem for $s = 1$, and the case of general s follows from the above argument and induction. In particular, this implies that the Fourier transform is a bijection from $H^s(\mathbb{R})$ onto $L^2(\mathbb{R}, (1 + |x|^2)^s dx)$. Since the latter space is an L^2-space it is a Hilbert space, and it follows that $H^s(\mathbb{R})$ is a Hilbert space under the given inner product $\langle \cdot, \cdot \rangle$. The second statement in the Theorem, concerning $[D^s(f)]^\wedge$, has already been proved for $s = 1$, and again the case for general s follows by an induction argument. Q.E.D.

The preceding result has an intuitive interpretation. Namely, if $f \in L^2(\mathbb{R})$ is thought of as a signal, the condition that $\int_{-\infty}^{\infty} |\widehat{f}(x)|^2 |x|^{2s} dx < \infty$ expresses the idea

that f does not contain "too many" high frequencies, in which case its graph will be smoother and more differentiable than would otherwise be the case. The Sobolev Space Theorem has the following important, if immediate, consequence.

PROPOSITION 1. *Let $s \in \mathbb{N}$. Then if D^s is considered as a function with domain $H^s(\mathbb{R})$, D^s becomes an injection from $H^s(\mathbb{R})$ into $L^2(\mathbb{R})$ whose range is*

$$\left\{ g : g \in L^2(\mathbb{R}) \text{ and } \int_{-\infty}^{\infty} \frac{|\widehat{g}(x)|^2}{|x|^{2s}} dx < \infty \right\}. \tag{2.4}$$

PROOF. Let $g \in L^2(\mathbb{R})$. Then, by the Sobolev Space Theorem and Plancherel's Theorem,

$$\begin{aligned} &g = D^s(f) \text{ for some } f \in H^s(\mathbb{R}) \\ &\iff g \in L^2(\mathbb{R}) \text{ and } \widehat{g}(x) = (ix)^s \widehat{f}(x) \quad \text{for some } f \in L^2(\mathbb{R}) \quad \text{and almost all } x \\ &\iff g \in L^2(\mathbb{R}) \text{ and } x \longmapsto \frac{\widehat{g}(x)}{x^s} \text{ is in } L^2(\mathbb{R}) \\ &\iff \int_{-\infty}^{\infty} \frac{|\widehat{g}(x)|^2}{|x|^{2s}}\, dx < \infty, \end{aligned}$$

which proves the statement about the range of D^s. That D^s is an injection is immediate from the identity $(D^s f)^\wedge(x) = (-ix)^s \widehat{f}(x)$, for $f \in H^s(\mathbb{R})$. Q.E.D.

It is shown by (2.4) that the functions in $L^2(\mathbb{R})$ which are in the range of D^s are completely described by the behaviour of their Fourier transforms near the origin. The Fourier transforms of appropriate finite differences exhibit identical behaviour near the origin, as discussed in the next section. It is this identical behaviour which provides the analytical tool for establishing the precise connection between derivatives and differences.

3. DIFFERENCES AND THE FOURIER TRANSFORM. Let $L^1(\mathbb{R})$ denote the complex valued integrable functions on $\mathbb{R}$. Then if $y \in \mathbb{R}$ and $f \in L^1(\mathbb{R})$, for all $x \in \mathbb{R}$

$$(\delta_y * f)^\wedge(x) = \int_{-\infty}^{\infty} e^{-ixt} f(t-y)dt = \int_{-\infty}^{\infty} e^{-ix(u+y)} f(u)du = e^{-ixy}\widehat{f}(x).$$

Since $f \longmapsto \delta_y * f$ is continuous on $L^2(\mathbb{R})$ and since $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ is a dense subspace of $L^2(\mathbb{R})$, it follows from Plancherel's Theorem that for all $f \in L^2(\mathbb{R})$ and all $y \in \mathbb{R}$,

$$(\delta_y * f)^\wedge(x) = e^{-ixy}\widehat{f}(x), \text{ for almost all } x \in \mathbb{R}. \tag{3.1}$$

Now let g be a first order difference in $L^2(\mathbb{R})$. Then $g = f - \delta_y * f$ for some $f \in L^2(\mathbb{R})$ and some $y \in \mathbb{R}$. Then it follows from (3.1) that

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{|\widehat{g}(x)|^2}{|x|^2} dx &= \int_{-\infty}^{\infty} \frac{|(f - \delta_y * f)^\wedge(x)|^2}{|x|^2} dx, \\ &= \int_{-\infty}^{\infty} \frac{|1 - e^{-ixy}|^2 |\widehat{f}(x)|^2}{|x|^2} dx, \end{aligned}$$

$$
\begin{aligned}
&= y^2 \int_{-\infty}^{\infty} \left(\frac{\sin \frac{xy}{2}}{\frac{xy}{2}} \right)^2 |\widehat{f}(x)|^2 dx, \\
&\leq y^2 \int_{-\infty}^{\infty} |\widehat{f}(x)|^2 dx, \quad \text{as } |\sin x| \leq |x|, \\
&< \infty.
\end{aligned}
$$

Thus, if g is a first order difference in $L^2(\mathbb{R})$, its Fourier transform $\widehat{g}$ must be "small" near the origin in the sense that $\int_{-\infty}^{\infty} |\widehat{g}(x)|^2 |x|^{-2} dx < \infty$. A similar argument establishes that if h is a second order difference in $L^2(\mathbb{R})$, then its Fourier transform is again "small" near the origin in the sense that the more stringent condition, $\int_{-\infty}^{\infty} |\widehat{h}(x)|^2 |x|^{-4} dx < \infty$, must be satisfied.

The preceding remarks, together with applications of Proposition 1 for the cases $s = 1$ and $s = 2$, suffice to establish the following result, which is indicative of a close relationship between derivatives and finite differences.

PROPOSITION 2. *If $g \in L^2(\mathbb{R})$ is either the derivative of a function in $H^1(\mathbb{R})$ or is a finite sum of first order differences, then $\int_{-\infty}^{\infty} |\widehat{g}(x)|^2 |x|^{-2} dx < \infty$. Also, if $h \in L^2(\mathbb{R})$ is either the second derivative of a function in $H^2(\mathbb{R})$ or is a finite sum of second order differences, then $\int_{-\infty}^{\infty} |\widehat{h}(x)|^2 |x|^{-4} dx < \infty$.*

An aim now is to prove a form of converse of each statement in Proposition 2 and, in fact, to go further by introducing the concepts of higher order and fractional differences in $L^2(\mathbb{R})$. Then the finite sums of such differences of a given order are characterized in terms of the behaviour of their Fourier transforms near the origin.

Let $C_{2\pi}(\mathbb{R})$ denote the continuous, complex-valued, 2-periodic functions on $\mathbb{R}$. If $\alpha \in C_{2\pi}(\mathbb{R})$, the *Fourier coefficients* $\widehat{\alpha}(k)$ of α are defined by

$$
\widehat{\alpha}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ikx} f(x) dx, \qquad \text{for all } k \in \mathbb{Z}.
$$

The series $\sum_{k=-\infty}^{\infty} \widehat{\alpha}(k) e^{ikx}$ is the Fourier series of α, and it is said to be *absolutely convergent* if $\sum_{k=-\infty}^{\infty} |\widehat{\alpha}(k)| < \infty$. In this case, the Fourier series of α converges uniformly on $\mathbb{R}$ to α.

Now let $\alpha \in C_{2\pi}(\mathbb{R})$ and assume that α has an absolutely convergent Fourier series. Then a function g in $L^2(\mathbb{R})$ is said to be an *α-difference* if there are $y \in \mathbb{R}$ and $f \in L^2(\mathbb{R})$ such that

$$
g = \sum_{k=-\infty}^{\infty} \widehat{\alpha}(k) \delta_{-ky} * f.
$$

Note that the series on the right converges unconditionally in $L^2(\mathbb{R})$, since

$$
\sum_{k=-\infty}^{\infty} ||\widehat{\alpha}(k) \delta_{-ky} * f||_2 \leq \left(\sum_{k=-\infty}^{\infty} |\widehat{\alpha}(k)| \right) ||f||_2 < \infty.
$$

The definition of an α-difference extends the concept of first and second order differences. For, let $\alpha_1(x) = 1 - e^{-ix}$ for $x \in \mathbb{R}$. Then $\alpha_1 \in C_{2\pi}(\mathbb{R})$ and

$$\sum_{k=-\infty}^{\infty} \widehat{\alpha}_1(k)\delta_{-ky} * f = f - \delta_y * f, \tag{3.2}$$

which is the general form of a first order difference. Also, if $\alpha_2(x) = (1 - e^{-ix})^2$ for all $x \in \mathbb{R}$, if $y \in \mathbb{R}$, if $g \in L^2(\mathbb{R})$ and if $f = -2\delta_y * g$, then

$$\sum_{k=-\infty}^{\infty} \widehat{\alpha}_2(k)\delta_{-ky} * g = g - 2\delta_y * g + \delta_{2y} * g = f - 2^{-1}(\delta_y + \delta_{-y}) * f, \tag{3.3}$$

which is the general form of a second order difference.

Given a function α as above, the preceding remarks lead to consideration of the vector subspace $\mathcal{D}_{[\alpha]}(\mathbb{R})$ of $L^2(\mathbb{R})$, which is defined to consist of all finite sums of α-differences. Thus, if $g \in L^2(\mathbb{R}), g \in \mathcal{D}_{[\alpha]}(\mathbb{R})$ if and only if there are $m \in \mathbb{N}$, $y_1, \ldots, y_m \in \mathbb{R}$ and $f_1, \ldots, f_m \in L^2(\mathbb{R})$ such that

$$g = \sum_{j=1}^{m} \left(\sum_{k=-\infty}^{\infty} \widehat{\alpha}(k)\delta_{-ky_j} * f_j \right). \tag{3.4}$$

The following result gives a preliminary characterization of $\mathcal{D}_{[\alpha]}(\mathbb{R})$ in terms of Fourier transforms.

PROPOSITION 3. *Let $\alpha \in C_{2\pi}(\mathbb{R})$ and assume that α has an absolutely convergent Fourier series. Then if $g \in L^2(\mathbb{R}), g \in \mathcal{D}_{[\alpha]}(\mathbb{R})$ if and only if there are $m \in \mathbb{N}$ and $y_1, \ldots, y_m \in \mathbb{R}$ such that*

$$\int_{-\infty}^{\infty} \frac{|\widehat{g}(x)|^2}{\displaystyle\sum_{j=1}^{m} |\alpha(y_j x)|^2} dx < \infty. \tag{3.5}$$

In this case, g is the sum of m α-differences; in fact, there are $f_1, f_2, \ldots, f_m \in L^2(\mathbb{R})$ such that (3.4) holds.

PROOF. It is first necessary to calculate the Fourier transform of an α-difference. As the Fourier series of α is absolutely convergent, the series converges at each point x of $\mathbb{R}$ and has sum $\alpha(x)$. Hence, if $y \in \mathbb{R}$ and $f \in L^2(\mathbb{R})$,

$$\begin{aligned} \left(\sum_{k=-\infty}^{\infty} \widehat{\alpha}(k)\delta_{-ky} * f \right)^{\wedge} (x) &= \left(\sum_{k=-\infty}^{\infty} \widehat{\alpha}(k)e^{ikyx} \right) \widehat{f}(x), \quad \text{by (3.1)}, \\ &= \alpha(yx)\widehat{f}(x), \end{aligned} \tag{3.6}$$

for almost all $x \in \mathbb{R}$.

Now let $g \in \mathcal{D}_{[\alpha]}(\mathbb{R})$. Then there are $m \in \mathbb{N}$, $y_1, \ldots, y_m \in \mathbb{R}$ and $f_1, \ldots, f_m \in L^2(\mathbb{R})$ such that g equals a finite sum of α-differences as in (3.4). Then, for almost all $x \in \mathbb{R}$,

$$
\begin{aligned}
|\widehat{g}(x)|^2 &= \left| \sum_{j=1}^{m} \left(\sum_{k=-} \widehat{\alpha}(k) \delta_{-ky_j} * f_j \right)^{\wedge} (x) \right|^2, \\
&= \left| \sum_{j=1}^{m} \alpha(y_j x) \widehat{f_j}(x) \right|^2, \quad \text{by (3.6)}, \\
&\leq \left(\sum_{j=1}^{m} |\alpha(y_j x)|^2 \right) \left(\sum_{j=1}^{m} |\widehat{f_j}(x)|^2 \right),
\end{aligned}
$$

by the Cauchy-Schwartz inequality. It follows that

$$
\int_{-\infty}^{\infty} \frac{|\widehat{g}(x)|^2}{\sum_{j=1}^{m} |\alpha(y_j x)|^2} dx \leq \sum_{j=1}^{m} \left(\int_{-\infty}^{\infty} |\widehat{f_j}(x)|^2 dx \right) < \infty,
$$

by Plancherel's theorem. Thus if $g \in \mathcal{D}_{[\alpha]}(\mathbb{R})$, (3.5) will hold, which proves part of Proposition 3.

Conversely, let $g \in L^2(\mathbb{R})$, let $m \in \mathbb{N}$ and let $y_1, y_2, \ldots, y_m \in \mathbb{R}$ be such that (3.5) holds. Then define

$$
h(x) = \max \Big\{ |\alpha(y_1 x)|,\ |\alpha(y_2 x)|, \ldots, |\alpha(y_m x)| \Big\},
$$

for each $x \in \mathbb{R}$. Then define subsets $B_0, B_1, \ldots, B_m$ of $\mathbb{R}$ by induction, as follows:

$$
\begin{aligned}
B_0 &= \Big\{ x : x \in \mathbb{R} \text{ and } h(x) = 0 \Big\}, \quad \text{and} \\
B_j &= \Big\{ x : x \in \mathbb{R}, x \notin B_{j-1} \text{ and } h(x) = |\alpha(y_j x)| \Big\},
\end{aligned}
$$

for $j = 1, 2, \ldots, m$. The sets $B_0, B_1, \ldots, B_m$ are disjoint, measurable, and their union is $\mathbb{R}$. Note that if $x \in B_j$ for some $j \in \{1, 2, \ldots, m\}$, then $|\alpha(y_j x)| = h(x) > 0$. Thus, for $j = 1, 2, \ldots, m$, a complex-valued function h_j may be defined by

$$
h_j(x) = \begin{cases} \dfrac{\widehat{g}(x)}{\alpha(y_j x)}, & \text{if } x \in B_j; \\ 0, & \text{if } x \notin B_j. \end{cases} \tag{3.7}
$$

Then

$$
\begin{aligned}
\int_{-\infty}^{\infty} |h_j(x)|^2 dx &= \int_{B_j} \frac{|\widehat{g}(x)|^2}{|\alpha(y_j x)|^2} dx, \\
&= m \int_{B_j} \frac{|\widehat{g}(x)|^2}{m \cdot [\max\{|\alpha(y_1 x)|^2, \ldots, |\alpha(y_m x)|^2\}]} dx, \\
&\leq m \int_{\mathbb{R}} \frac{|\widehat{g}(x)|^2}{\sum_{j=1}^{m} |\alpha(y_j x)|^2} dx, \\
&< \infty,
\end{aligned}
$$

as (3.5) holds. Hence $h_j \in L^2(\mathbb{R})$. By Plancherel's Theorem, there is $f_j \in L^2(\mathbb{R})$ such that $\widehat{f_j} = h_j$, for all $j = 1, 2, \ldots, m$. Then (3.7) gives $\widehat{g}(x) = \alpha(y_j x)\widehat{f_j}(x)$, for all $x \in B_j$. Since $B_1, \ldots, B_m$ are disjoint, and since $\widehat{f_j} = 0$ on B_k if $j \neq k$, it follows that

$$\widehat{g}(x) = \sum_{j=1}^{m} \alpha(y_j x)\widehat{f_j}(x), \qquad \text{for all } x \in \bigcup_{j=1}^{m} B_j.$$

In fact, this equation for $\widehat{g}(x)$ holds for almost all $x \in \mathbb{R}$. For if $x \in B_0, \alpha(y_j x) = 0$ for all $j = 1, 2, \ldots, m$ and so, as (3.5) holds, $\widehat{g}(x) = 0$ for almost all $x \in B_0$. Since $\mathbb{R}$ is the disjoint union of $B_0, B_1, \ldots, B_m$, it follows that for almost all $x \in \mathbb{R}$,

$$\begin{aligned}
\widehat{g}(x) &= \sum_{j=1}^{m} \alpha(y_j x)\widehat{f_j}(x), \\
&= \sum_{j=1}^{m} \left(\sum_{k=-\infty}^{\infty} \widehat{\alpha}(k)\delta_{-ky_j} * f_j \right)^{\wedge} (x), \quad \text{by (3.6)}, \\
&= \left(\sum_{j=1}^{m} \left(\sum_{k=-\infty}^{\infty} \widehat{\alpha}(k)\delta_{-ky_j} * f_j \right) \right)^{\wedge} (x),
\end{aligned}$$

so that

$$g = \sum_{j=1}^{m} \left(\sum_{k=-\infty}^{\infty} \widehat{\alpha}(k)\delta_{-ky_j} * f_j \right).$$

This proves that if g satisfies (3.5), then $g \in \mathcal{D}_{[\alpha]}(\mathbb{R})$.

Thus, $g \in \mathcal{D}_{[\alpha]}(\mathbb{R})$ if and only if (3.5) holds for some $y_1, \ldots, y_m$. The last equation also shows that, in this case, g is the sum of m α-differences as in (3.4).

Q.E.D.

When $g \in \mathcal{D}_{[\alpha]}(\mathbb{R})$, there are $y_1, \ldots, y_m \in \mathbb{R}$ such that, for some $f_1, \ldots, f_m \in L^2(\mathbb{R})$, $g = \sum_{j=1}^{m} (\sum_{k=-\infty}^{\infty} \widehat{\alpha}(k)\delta_{-ky_j} * f_j)$. An examination of the proof of Proposition 3 reveals that these functions $f_1, \ldots, f_m$ in $L^2(\mathbb{R})$ may be chosen so that $\widehat{f_j}\widehat{f_k} = 0$ if $j \neq k$. This means that $(f_j * f_k)^{\wedge} = 0$ for $j \neq k$, where $f_j * f_k$ is a convolution in $L^2(\mathbb{R})$. Thus, the functions $f_1, \ldots, f_m$ which appear in (3.4) may always be chosen so that $f_j * f_k = 0$ if $j \neq k$.

Although Proposition 3 gives a necessary and sufficient condition for a function in $L^2(\mathbb{R})$ to belong to a space $\mathcal{D}_{[\alpha]}(\mathbb{R})$, the criterion it provides is difficult to apply; for if $f \in L^2(\mathbb{R})$ is given, it is not clear under what conditions there will be $y_1, \ldots, y_m \in \mathbb{R}$ such that (3.5) will hold. What is needed is a simpler criterion, of a type suggested by Proposition 2.

If $\alpha_1(x) = 1 - e^{-ix}$ and $\alpha_2(x) = (1 - e^{-ix})^2$, Proposition 2 shows that

$$\begin{aligned}
g \in \mathcal{D}_{[\alpha_1]}(\mathbb{R}) &\Longrightarrow \widehat{g} \in L^2(\mathbb{R}, (1 + |x|^{-2})dx), \quad \text{and} \\
h \in \mathcal{D}_{[\alpha_2]}(\mathbb{R}) &\Longrightarrow \widehat{h} \in L^2(\mathbb{R}, (1 + |x|^{-4})dx).
\end{aligned}$$

This motivates the formulation of the basic problem for characterizing a space $\mathcal{D}_{[\alpha]}(\mathbb{R})$ as follows: describe a Borel measure $\mu_{[\alpha]}$ on $\mathbb{R}$ such that

$$\left\{ \widehat{f} : f \in \mathcal{D}_{[\alpha]}(\mathbb{R}) \right\} = L^2(\mathbb{R}, \mu_{[\alpha]}).$$

We shall see that under appropriate conditions on α, there is such a measure $\mu_{[\alpha]}$. Moreover, $\mu_{[\alpha]}$ is essentially determined not by α itself, but by the behaviour of α near the origin.

LEMMA 1. *Let $s \in [0, \infty)$ and let $m \in \mathbb{N}$. Then there are $\eta_1, \eta_2 \in (0, \infty)$ such that*

$$\eta_1 \left(\sum_{j=1}^{m} |x_j|^2 \right)^s \leq \sum_{j=1}^{m} |x_j|^{2s} \leq \eta_2 \left(\sum_{j=1}^{m} |x_j|^2 \right)^s,$$

for all $x_1, x_2, \ldots, x_m \in \mathbb{R}$.

PROOF. Consider the function φ, defined for all non-zero vectors in $\mathbb{R}^m$, given by

$$\varphi(x_1, x_2, \ldots, x_m) = \frac{\left(\sum_{j=1}^{m} |x_j|^{2s} \right)}{\left(\sum_{j=1}^{m} |x_j|^2 \right)^s}.$$

Then for all non-zero $x \in \mathbb{R}^m$ and all non-zero $\gamma \in \mathbb{R}$, $\varphi(\gamma x) = \varphi(x)$. Thus, it suffices to prove the result only for the case of all unit vectors in $\mathbb{R}^m$, a unit vector being one of the form $(x_1, x_2, \ldots, x_m)$ such that $\sum_{j=1}^{m} |x_j|^2 = 1$. However, the set of all unit vectors in $\mathbb{R}^m$ is a compact space, and the restriction of φ to this space is comtinuous. Thus, φ attains a minimum value η_1 and a maximum value η_2 on this space. Then $\eta_1 \leq \varphi(x) \leq \eta_2$ for all unit vectors x and, as φ does not vanish at any point, $\eta_1 > 0$. The conclusion now follows. Q.E.D.

LEMMA 2. *Let $s \in [0, \infty)$ and let $m \in \mathbb{N}$ be such that $m > 2s$. Let $\eta \in (0, \infty)$ and $\alpha \in C_{2\pi}(\mathbb{R})$ be such that $\eta |x|^s \leq |\alpha(x)|$ for all $x \in [-\pi, \pi]$. Then there is $C \in (0, \infty)$ such that*

$$\int_{(-|x|,|x|)^m} \frac{du_1 du_2 \ldots du_m}{\sum_{j=1}^{m} |\alpha(u_j)|^2} \leq C(|x|^{m-2s} + |x|^m), \quad \textit{for all} \quad x \in \mathbb{R}.$$

PROOF. If $r \in (0, \infty)$, let $B_r(\mathbb{R}^m)$ denote the unit ball of centre 0 and radius r in $\mathbb{R}^m$. Note that $(-|x|, |x|)^m \subseteq B_{\sqrt{m}|x|}(\mathbb{R}^m)$. Now let $x \in [-\pi, \pi]$, and let η_1 be

as in Lemma 1. Then

$$\int_{(-|x|,|x|)^m} \frac{du_1 du_2 \dots du_m}{\sum_{j=1}^{m} |\alpha(u_j)|^2} \leq \frac{1}{\eta^2} \int_{(-|x|,|x|)^m} \frac{du_1 du_2 \dots du_m}{\sum_{j=1}^{m} |u_j|^{2s}},$$

$$\leq \frac{1}{\eta_1 \eta^2} \int_{B_{\sqrt{m}|x|}(\mathbb{R}^m)} \frac{du_1 du_2 \dots du_m}{\left(\sum_{j=1}^{m} |u_j|^2\right)^s},$$

$$= \frac{\mu_{m-1}(S^{m-1})}{\eta_1 \eta^2} \left(\int_0^{\sqrt{m}|x|} r^{m-1-2s} dr \right),$$

$$= \frac{\mu_{m-1}(S^{m-1}) m^{(m-2s)/2} |x|^{m-2s}}{\eta_1 \eta^2}.$$

Here, integration by polar coordinates has been used, by putting $r = \left(\sum_{j=1}^{m} |u_j|^2 \right)^2$ and letting μ_{m-1} be the surface measure on the unit sphere S^{m-1}. It follows that there is $C_1 \in (0,)$ such that for all $x \in [-\pi, \pi]$,

$$\int_{(-|x|,|x|)^m} \frac{du_1 du_2 \dots du_m}{\sum_{j=1}^{m} |\alpha(u_j)|^2} \leq C_1 |x|^{m-2s} \leq C_1 (|x|^{m-2s} + |x|^m). \tag{3.8}$$

In the case where $|x| \in (\pi, \infty)$, there is $\ell \in \mathbb{N}$ such that $\ell < |x|(\ell + 1)$. Then $(-|x|, |x|)^m \subseteq (-2(\ell+1), 2(\ell+1))^m$, and this latter set is the union of $2^m(\ell+1)^m$ disjoint subcubes in $\mathbb{R}^m$, each of whose sides has length 2. As α has period 2, the function $(u_1, \dots, u_m) \longmapsto 1/(\sum_{j=1}^{m} |\alpha(u_j)|^2)$ has the same integral over each of these cubes. It follows that for $|x| \in (\pi, \infty)$,

$$\int_{(-|x|,|x|)^m} \frac{du_1 du_2 \dots du_m}{\sum_{j=1}^{m} |\alpha(u_j)|^2} \leq 2^m (\ell+1)^m \int_{(-,)^m} \frac{du_1 du_2 \dots du_m}{\sum_{j=1}^{m} |\alpha(u_j)|^2},$$

$$\leq C_1 2^m (\ell+1)^m ({}^{m-2s} + {}^m), \quad \text{by (3.8)},$$

$$\leq C_1 2^m (2\ell)^m (1 + {}^{-2s})^m,$$

$$\leq C_1 4^m (1 + {}^{-2s}) |x|^m, \quad \text{as } \ell < |x|,$$

$$\leq C_1 4^m (1 + {}^{-2s}) (|x|^{m-2s} + |x|^m). \tag{3.9}$$

Lemma 2 is immediate from (3.8) and (3.9). Q.E.D.

THEOREM 3. *Let $s \in (0, \infty)$, let $\alpha \in C_{2\pi}(\mathbb{R})$ be such that α has an absolutely convergent Fourier series and for some $\delta_1, \delta_2 \in (0, \infty)$, $\delta_1 |x|^s \leq |\alpha(x)| \leq \delta_2 |x|^s$ for all $x \in [-\pi, \pi]$.*

Then if $g \in L^2(\mathbb{R})$, $g \in \mathcal{D}_{[\alpha]}(\mathbb{R})$ *if and only if* $\int_{-\infty}^{\infty} |\widehat{g}(x)|^2 |x|^{-2s} dx < \infty$. *Equivalently,*

$$\left\{\widehat{g} : g \in \mathcal{D}_{[\alpha]}(\mathbb{R})\right\} = L^2(\mathbb{R}, (1 + |x|^{-2})^s dx).$$

Also, if $m \in \mathbb{N}$ *and* $m > 2s$*, every function in* $\mathcal{D}_{[\alpha]}(\mathbb{R})$ *is the sum of* m α*-differences in* $L^2(\mathbb{R})$*.*

PROOF. Let $g \in \mathcal{D}_{[\alpha]}(\mathbb{R})$. Then by Proposition 3, there are $y_1, y_2, \ldots, y_m \in \mathbb{R}$ such that (3.5) holds. Let $\eta > 0$ be such that $|y_j x| \leq \pi$ for all $x \in [-\eta, \eta]$ and all $j = 1, 2, \ldots, m$. Then

$$\begin{aligned}
\int_{-\eta}^{\eta} \frac{|\widehat{g}(x)|^2}{|x|^{2s}} dx &= \left(\sum_{j=1}^{m} |y_j|^{2s}\right) \int_{-\eta}^{\eta} \frac{|\widehat{g}(x)|^2}{\left(\sum_{j=1}^{m} |y_j x|^{2s}\right)} dx, \\
&\leq \delta_2^2 \left(\sum_{j=1}^{m} |y_j|^{2s}\right) \int_{-\eta}^{\eta} \frac{|\widehat{g}(x)|^2}{\sum_{j=1}^{m} |\alpha(y_j x)|^2} dx, \\
&< \infty, \quad \text{as (3.5) holds.}
\end{aligned}$$

Hence if $g \in \mathcal{D}_{[\alpha]}(\mathbb{R})$, $\int_{-\infty}^{\infty} |\widehat{g}(x)|^2 |x|^{-2s} dx < \infty$, and together with Plancherel's Theorem this implies that $\widehat{g} \in L^2(\mathbb{R}, (1 + |x|^{-2})^s dx)$.

Conversely, let $g \in L^2(\mathbb{R})$. Let $m \in \mathbb{N}$ with $m > 2s$, and let $b \in (0, \infty)$. Recall that in an iterated integral with a non-negative integrand, the value of the integral is independent of the order of the integration. Also, note that the Jacobian of the function $(y_1, \ldots, y_m) \longmapsto (xy_1, \ldots, xy_m)$ on $\mathbb{R}^m$ is $|x|^m$. Using the substitution $u_j = xy_j$ for $j = 1, 2, \ldots, m$, and the formula for integration by substitution in $\mathbb{R}^m$, now gives the following:

$$\begin{aligned}
\int_{(-b,b)^m} \left(\int_{-\infty}^{\infty} \frac{|\widehat{g}(x)|^2}{\sum_{j=1}^{m} |\alpha(y_j x)|^2} dx\right) dy_1 dy_2 \ldots dy_m &= \int_{-\infty}^{\infty} \left[\int_{(-b,b)^m} \frac{dy_1 dy_2 \ldots dy_m}{\sum_{j=1}^{m} |\alpha(y_j x)|^2}\right] |\widehat{g}(x)|^2 dx, \\
&= \int_{-\infty}^{\infty} \left[\int_{(-b|x|, b|x|)^m} \frac{du_1 du_2 \ldots du_m}{\sum_{j=1}^{m} |\alpha(u_j)|^2}\right] |\widehat{g}(x)|^2 |x|^{-m} dx, \\
&\leq D \int_{-\infty}^{\infty} |\widehat{g}(x)|^2 (1 + |x|^{-2s}) dx,
\end{aligned}$$

where $D \in (0, \infty)$ and D comes from applying Lemma 2. It follows that if

$\int_{-\infty}^{\infty} |\widehat{g}(x)|^2 |x|^{-2s} dx < \infty$, then

$$\int_{-\infty}^{\infty} \frac{|\widehat{g}(x)|^2}{\sum_{j=1}^{m} |\alpha(xy_j)|^2} dx < \infty \qquad \text{for almost all } (y_1, \ldots, y_m) \in (-b, b)^m.$$

In particular, this inequality must hold for at least one $(y_1, \ldots, y_m) \in (-b, b)^m$, so by Proposition 3, $g \in \mathcal{D}_{[\alpha]}(\mathbb{R})$ and g is the sum of m α-differences. Q.E.D.

The preceding proof actually demonstrates more than is contained in the statement of Theorem 3. In fact, the proof shows that if $g \in \mathcal{D}_{[\alpha]}(\mathbb{R})$, for almost all $(y_1, y_2, \ldots, y_m) \in \mathbb{R}^m$, there are $f_1, f_2, \ldots, f_m \in L^2(\mathbb{R})$ such that

$$g = \sum_{j=1}^{m} \left(\sum_{k=-\infty}^{\infty} \widehat{\alpha}(k) \delta_{-ky_j} * f_j \right).$$

Further details upon this and related sharpness aspects of Theorem 3 may be found in the work of Wai Lok Lo [9] and also in [14] and [15].

4. DIFFERENCE SPACES AND DERIVATIVES. Another way of thinking of the characterization of $\mathcal{D}_{[\alpha]}(\mathbb{R})$ described in Theorem 3 is as follows: the Fourier transform is a linear bijection from $\mathcal{D}_{[\alpha]}(\mathbb{R})$ onto the weighted L^2-space $L^2(\mathbb{R}, (1 + |x|^{-2})^s dx)$. Under the conditions of Theorem 3, it follows that $\mathcal{D}_{[\alpha]}(\mathbb{R})$ is independent of the function α, and that $\mathcal{D}_{[\alpha]}(\mathbb{R})$ is characterized in terms of s only. This solves what was previously called the basic problem for characterizing the spaces $\mathcal{D}_{[\alpha]}(\mathbb{R})$, and leads to the following.

DEFINITION. Let $s \in (0, \infty)$ be such that there is $\alpha \in C_{2\pi}(\mathbb{R})$ and $\delta_1, \delta_2 \in (0, \infty)$ so that α has an absolutely convergent Fourier series and $\delta_1 |x|^s \le |\alpha(x)| \le \delta_2 |x|^s$ for all $x \in [-\pi, \pi]$. Then $\mathcal{D}_s(\mathbb{R})$ is defined to be the vector subspace $\mathcal{D}_{[\alpha]}(\mathbb{R})$ of $L^2(\mathbb{R})$.

The space $\mathcal{D}_s(\mathbb{R})$ is called the *difference space of order* s of $L^2(\mathbb{R})$.

The question now arises as to the values of $s \in (0, \infty)$ for which $\mathcal{D}_s(\mathbb{R})$ is defined. If $s \in \mathbb{N}$, it suffices to take $\alpha \in C_{2\pi}(\mathbb{R})$ to be given by $\alpha(x) = (1 - e^{-ix})^s$ for $x \in \mathbb{R}$. In fact, in this case the Fourier series of α has only a finite number of non-zero terms and so is absolutely convergent; and since

$$|\alpha(x)| = |1 - e^{-ix}|^s = 2^s \left| \sin \frac{x}{2} \right|^s \quad \text{and} \quad \lim_{x \to 0} \frac{\sin x}{x} = 1,$$

$\alpha(x)|x|^{-s}$ is bounded above and also bounded away from 0 on $[-\pi, \pi]$.

The preceding choice of α when $s \in \mathbb{N}$ motivates the choice of a suitable α in $C_{2\pi}(\mathbb{R})$ in the general case when $s \in (0, \infty)$. Let $s \in (0, \infty)$ be given, and let $\alpha \in C_{2\pi}(\mathbb{R})$ be given by $\alpha(x) = \left| \sin \frac{x}{2} \right|^s$, for $x \in \mathbb{R}$. The preceding argument applies to this case, and it may be deduced that $\alpha(x)|x|^{-s}$ is bounded above and is also bounded away from 0 on $[-\pi, \pi]$. We need to know that α has an absolutely convergent Fourier series. Now the derivative $D(\alpha)$ of α exists at each point of

$(-\pi, 0) \cup (0, \pi)$, and

$$D(\alpha)(x) = \begin{cases} \frac{s}{2} \sin^{s-1} \frac{x}{2} \cos \frac{x}{2}, & \text{if } x \in (0, \pi), \\ -\frac{s}{2} \sin^{s-1} -\frac{x}{2} \cos \frac{x}{2}, & \text{if } x \in (-\pi, 0). \end{cases} \tag{4.1}$$

It follows that $D(\alpha)$ is continuous on $(-\pi, 0) \cup (0, \pi)$ and that $D(\alpha) \in L^1([-\pi, \pi])$. Then $\alpha(x) = \alpha(-\pi) + \int_{-\pi}^{x} D(\alpha)(t)dt$ for all $x \in [-\pi, \pi]$, and so, by a standard result [7, p.286], α is absolutely continuous on $[-\pi, \pi]$.

In order to deduce that α has an absolutely convergent Fourier series it now suffices, by another standard result [19, p.242], to prove that there is $p \in (1, \infty)$ such that $D(\alpha) \in L^p([-\pi, \pi])$. If $s \in [1, \infty)$, (4.1) shows that $D(\alpha)$ is bounded on $[-\pi, \pi]$ and so is in $L^2([-\pi, \pi])$. If $s \in (0, 1)$, let $p \in (1, \infty)$ be such that $p(1-s) < 1$. Then by (4.1),

$$\begin{aligned} \int_{-\pi}^{\pi} |D(\alpha)(x)|^p dx &= \frac{s^p}{2^p} \int_{-\pi}^{\pi} \frac{|\cos \frac{x}{2}|^p}{|\sin \frac{x}{2}|^{p(1-s)}} dx, \\ &\le \frac{s^p}{2^{p-1}} \int_0^{\pi} \frac{\cos \frac{x}{2}}{\sin^{p(1-s)} \frac{x}{2}} dx, \\ &\le \frac{s^p}{2^{p-2}\delta^p} \int_0^1 \frac{du}{u^{p(1-s)}}, \quad \text{substituting } u = \sin \frac{x}{2}, \\ &< \infty, \quad \text{as } p(1-s) < 1. \end{aligned}$$

It follows that $D(\alpha) \in L^p([-\pi, \pi])$. Thus, in either case, α has an absolutely convergent Fourier series.

The preceding remarks show that the difference space $\mathcal{D}_s(\mathbb{R})$ is defined for all $s \in [0, \infty)$. Also, the opening remarks of this section make it clear that the Fourier transform is a linear bijection from $\mathcal{D}_s(\mathbb{R})$ onto $L^2(\mathbb{R}, (1 + |x|^{-2})^s dx)$. Since any L^2-space is a Hilbert space, the space $\mathcal{D}_s(\mathbb{R})$ is also a Hilbert space in the inner product derived from the inner product on $L^2(\mathbb{R}, (1 + |x|^{-2})^s dx)$ under the Fourier transform. These facts are now summarized.

THE DIFFERENCE SPACE THEOREM. *If $s \in (0, \infty)$, the difference space $\mathcal{D}_s(\mathbb{R})$ of order s is defined. If $f \in L^2(\mathbb{R})$, $f \in \mathcal{D}_s(\mathbb{R})$ if and only if $\int_{-\infty}^{\infty} |\widehat{f}(x)|^2 |x|^{-2s} dx < \infty$. The vector space $\mathcal{D}_s(\mathbb{R})$ is a Hilbert space whose inner product is given by*

$$\langle f, g \rangle = \int_{-} \widehat{f}(x)\overline{\widehat{g}(x)}(1 + |x|^{-2})^s dx, \quad \textit{for } f, g \in \mathcal{D}_s(\mathbb{R}).$$

Now it was noted earlier that the definition of an α-difference extended the concept of first and second order differences. In fact, from (3.2) and (3.3) and the

definition of the difference spaces, it follows that $\mathcal{D}_1(\mathbb{R})$ consists of all finite sums of first order differences, and that $\mathcal{D}_2(\mathbb{R})$ consists of a ll finite sums of second order differences. In general, if α satisfies the conditions in the definition of $\mathcal{D}_{[\alpha]}(\mathbb{R})$, it is natural to regard an α-difference as a "fractional difference of order s". Then, $\mathcal{D}_s(\mathbb{R})$ consists of all finite sums of such differences of order s.

The reader may have noted that the Difference Space Theorem has a strong formal similarity to the Sobolev Space Theorem. In fact, whereas the Sobolev Space Theorem characterizes the Hilbert spaces of functions in $L^2(\mathbb{R})$ associated with the behaviour of the Fourier transform *towards infinity*, the Difference Space Theorem characterizes the corresponding Hilbert spaces of functions in $L^2(\mathbb{R})$ associated with the behaviour of the Fourier transform *near the origin*.

A more specific connection between these two classes of spaces is indicated by Proposition 1, where it was shown, in a precise sense, that for $s \in \mathbb{N}$ the functions in $L^2(\mathbb{R})$ which are in the range of D^s are characterized by the behaviour of their Fourier transforms near the origin.

THEOREM 4. *Let $s \in \mathbb{N}$. Then*

$$\mathcal{D}_s(\mathbb{R}) = \Big\{ D^s(f) : f \in H^s(\mathbb{R}) \Big\}.$$

Also, D^s preserves inner products as it maps from the Hilbert apace $H^s(\mathbb{R})$ onto the Hilbert space $\mathcal{D}_s(\mathbb{R})$.

PROOF. It is immediate from (2.4) and the Difference Space Theorem that $\mathcal{D}_s(\mathbb{R}) = \{D^s(f) : f \in H^s(\mathbb{R})\}$. Also, if $f, g \in H^s(\mathbb{R})$,

$$\begin{aligned}
\langle D^s(f), D^s(g) \rangle &= \int_{-\infty}^{\infty} D^s(f)^\wedge(x) \overline{D^s(g)^\wedge(x)} (1 + |x|^{-2})^s dx, \\
&= \int_{-\infty}^{\infty} \widehat{f}(x) \overline{\widehat{g}(x)} |x|^{2s} (1 + |x|^{-2})^s dx, \quad \text{by} \quad (2.3), \\
&= \int_{-\infty}^{\infty} \widehat{f}(x) \overline{\widehat{g}(x)} (1 + |x|^2)^s dx, \\
&= \langle f, g \rangle.
\end{aligned}$$

Here, the inner product on the left is in $\mathcal{D}_s(\mathbb{R})$, while the inner product on the right is in $H^s(\mathbb{R})$. The equation shows that D^s preserves inner products. Q.E.D.

Theorem 4 identifies the range of D^s on $H^s(\mathbb{R})$ as the space $\mathcal{D}_s(\mathbb{R})$. It also implies that D^s is an *isometry* from $H^s(\mathbb{R})$ onto $\mathcal{D}_s(\mathbb{R})$ since, as it preserves inner products, it also preserves the norms in these spaces. The two theorems stated in the Introduction now appear as special cases arising from Theorem 4 by taking $s = 1$ and $s = 2$.

In the case $s = 1$, we may take $\alpha(x) = 1 - e^{-ix}$ and then, as noted before, the α-differences are precisely the first order differences and $\mathcal{D}_1(\mathbb{R})$ consists of the finite sums of first order differences. Since $\mathcal{D}_1(\mathbb{R}) = \{D(f) : f \in H^1(\mathbb{R})\}$, Theorem 1 in the Introduction is proved. Actually, since $3 > 2 = 2.1$, the last statement in Theorem 3 shows that a function in $L^2(\mathbb{R})$ is in $\mathcal{D}_1(\mathbb{R})$ if and only if it is the sum of

three first order differences. Thus, a refined statement of Theorem 1 could read: *a function in $L^2(\mathbb{R})$ is the derivative of a function in $L^2(\mathbb{R})$ if and only if it is a sum of three first order differences.*

A similar argument applies in the case $s = 2$. Let $\alpha(x) = (1 - e^{-ix})^2$ and note, as previously, that the α-differences are precisely the second order differences and that $\mathcal{D}_2(\mathbb{R})$ consists of the finite sums of second order differences. Then Theorem 2 will follow and the corresponding refined statement of it reads: *a function in $L^2(\mathbb{R})$ is the second derivative of a function in $L^2(\mathbb{R})$ if and only if it is a sum of five second order differences.*

A common way of thinking of the derivative of a function is that it measures the *instantaneous* rate of change of the function. Now if $f \in L^2(\mathbb{R})$ and $y > 0$, $(f - \delta_y * f)(x)$ equals $f(x) - f(x - y)$, which is the change in the value of f as the argument increases from $x - y$ to x. Similarly, if $y < 0$, $(f - \delta_y * f)(x)$ is the change in the value of $-f$ as the argument increases from x to $x - y$. Hence, any first order difference may be regarded as a *discrete* rate of change, in the sense that it measures the difference between function values at points which are a given distance apart. From this viewpoint, the above refinement of Theorem 1 can be regarded as saying: *the instantaneous rate of change of a differentiable function in $L^2(\mathbb{R})$ whose derivative is in $L^2(\mathbb{R})$ is a sum of three functions, each of which is a discrete rate of change of some function in $L^2(\mathbb{R})$.*

5. MAKING WAVES. In the introduction, some heuristic remarks were made concerning the making of waves by means of differentiation and taking finite differences. Some further remarks on these lines may now be helpful. The extent to which a function is a wave may be regarded as expressing itself in the number of crests and troughs, or in the oscillations of the function. Now the Fourier transform $\widehat{f}$ of a function f in $L^2(\mathbb{R})$ can be thought of as displaying the frequencies contained in the function g or, as we may prefer to think of it in the immediate context, the signal g. The frequencies correspond to the complex exponential functions $e^{ixy}, y \in \mathbb{R}$ – the larger the value of $|y|$, the more does e^{ixy} oscillate and the more wavy it appears; while the smaller the value of $|y|$, the less does e^{ixy} oscillate and the less wavy it appears. Thus, the function may be expected to oscillate more if it contains many high frequencies, or if it contains few low frequencies, or if it contains more high frequencies in relation to the low frequencies.

In the case of differentiation, we saw earlier that if $f \in H^1(\mathbb{R})$,

$$[D(f)]^\wedge(x) = ix\widehat{f}(x), \qquad \text{for almost all } x,$$

which shows that differentiation simultaneously reduces the magnitude of the low frequencies in f and increases the magnitude of the high frequencies. But if we consider a first order difference of the form $f - \delta_y * f$, for some $f \in L^2(\mathbb{R})$ and $y \in \mathbb{R}$,

$$(f - \delta_y * f)^\wedge(x) = (1 - e^{ixy})\widehat{f}(x) = -2i \sin\frac{xy}{2} e^{ixy/2}\widehat{f}(x).$$

Since $|\sin\dfrac{xy}{2}| \leq 1$, this shows that taking first order differences decreases the magnitude of the low frequencies without increasing the magnitude of the high frequencies.

A similar comment applies in the case of α-differences. However, although differentiation increases the high frequencies when applied to a function f in $H^1(\mathbb{R})$, this effect is not important, in the sense that $x\widehat{f}(x)$ still remains as a function in $L^2(\mathbb{R})$. This may illuminate why, when $f \in H^1(\mathbb{R})$, it is possible to write $D(f)$ as a sum of first order differences even though the first order difference procedure does not increase the high frequencies.

The effect of these remarks is to show that if we wish to restrict our analysis to lie within the space $L^2(\mathbb{R})$ (and there are practical and aesthetic reasons for doing this), making a wave from a given function is done by decreasing the low frequencies in the function. The extent to which the new function so obtained is a wave, or is wave-like, will then be determined by the behaviour of its Fourier transform near the origin.

Now if $f \in L^2(\mathbb{R})$, the behaviour of the Fourier transform $\widehat{f}$ near the origin may be described in different ways. As $\widehat{f}$ is only defined almost everywhere, perhaps the most natural way is the one arising in the Difference Space Theorem: if $s \in (0,\infty)$, $\int_{-\infty}^{\infty} |\widehat{f}(x)|^2|x|^{-2s}dx < \infty$ if and only if $f \in \mathcal{D}_s(\mathbb{R})$. In view of the preceding remarks, it now may seem natural to consider $\mathcal{D}_s(\mathbb{R})$ to be the (Hilbert) space consisting of "waves of order s". However, it may happen that $\widehat{f}$ is several times differentiable near the origin and that some of the derivatives of $\widehat{f}$ at the origin vanish. Then Taylor's Theorem may apply to show that the behaviour of $\widehat{f}$ near the origin may be compared directly with a power of $|x|$.

More specifically, let $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and let $\int_{-\infty}^{\infty} |f(x)| \cdot |x|^j dx < \infty$ for all $j \in \{0, 1, \ldots, s-1\}$, for some given $s \in \mathbb{N}$. The Fourier transform $\widehat{f}$ of f is given by

$$\widehat{f}(x) = \int_{-\infty}^{\infty} e^{-ixy} f(y)dy, \qquad \text{for all } x \in \mathbb{R}.$$

The assumptions on f mean that $\widehat{f}$ is $s-1$ times continuously differentiable and that differentiation under the integral sign may be carried out $s-1$ times to obtain

$$D^j(\widehat{f})(0) = (-i)^j \int_{-\infty}^{\infty} f(x)x^j dx, \quad \text{for all } j \in \{0, 1, \ldots, s-1\}.$$

Thus, $D^j(\widehat{f})(0) = 0$ for all $j \in \{0, 1, \ldots, s-1\}$ if and only if

$$\int_{-\infty}^{\infty} f(x)x^j dx = 0 \quad \text{for all } j \in \{0, 1, \ldots, s-1\}. \tag{5.1}$$

The integrals in (5.1) are called *moments* of f, and on page 67 of his book on wavelets [13], Yves Meyer refers to the vanishing of the moments as the "oscillation" or "cancellation" condition on f. Since this is equivalent to $\widehat{f}$ having a zero of order s at the origin, we see that (5.1) expresses in a precise sense the lack of low frequencies in f.

The following result shows that the two methods of describing the behaviour of the Fourier transform near the origin coincide for functions for which they are applicable simultaneously.

PROPOSITION 4. *Let $s \in \mathbb{N}$ and let $f \in L^2(\mathbb{R})$ be such that $\int_{-\infty}^{\infty} |f(x)| \cdot |x|^j dx < \infty$ for all $j \in \{0, 1, \ldots, s\}$. Then $f \in \mathcal{D}_s(\mathbb{R})$ if and only if $\int_{-\infty}^{\infty} f(x)x^j dx = 0$ for all $j \in \{0, 1, \ldots, s-1\}$.*

PROOF. The assumptions imply that $\widehat{f}$ is s times continuously differentiable. Two cases are considered.

The first is when $D^j(\widehat{f})(0) = 0$ for all $j \in \{0, 1, \ldots, s-1\}$. Then Taylor's Theorem applies to $\widehat{f}$ about the origin, and it follows that $\widehat{f}(x)x^{-s}$ is bounded near the origin, so that $\int_{-\infty}^{\infty} |\widehat{f}(x)|^2 |x|^{-2s} dx < \infty$. Thus, in this case, $f \in \mathcal{D}_s(\mathbb{R})$ by the Difference Space Theorem.

The second case is when, for some $\ell \in \{0, 1, \ldots, s-1\}$, $D^\ell(\widehat{f})(0) \neq 0$. It may be assumed that ℓ is the least integer with these properties. Taylor's Theorem again applies, and we may write

$$\widehat{f}(x) = \frac{D^\ell(\widehat{f})(0)}{\ell!} x^\ell + \rho(x), \quad \text{for all } x \in \mathbb{R},$$

where $\lim_{x \to 0} \rho(x)|x|^{-(\ell+1)} = 0$. It follows that $|\widehat{f}(x)| \cdot |x|^{-\ell}$ is bounded away from 0 on some neighbourhood $(-\eta, \eta)$ of the origin, by δ, say. Then

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{|\widehat{f}(x)|^2}{|x|^{2s}} dx &= \int_{-\eta}^{\eta} \frac{|\widehat{f}(x)|^2}{|x|^{2\ell}} \cdot \frac{1}{|x|^{2(s-\ell)}} dx \\ &\geq \delta^2 \int_{-\eta}^{\eta} \frac{dx}{|x|^{2(s-\ell)}} = \infty,\end{aligned}$$

and so $f \notin \mathcal{D}_s(\mathbb{R})$ by the Difference Space Theorem.

In view of (5.1), the first case establishes that if the first s moments of f vanish, then $f \in \mathcal{D}_s(\mathbb{R})$. The second case establishes the converse statement, by proving it by contradiction. Q.E.D

By now it is hoped that the heuristic remarks made previously, together with Proposition 4, will have made the reader feel comfortable with the idea that a difference space $\mathcal{D}_s(\mathbb{R})$ may be thought of as consisting of "waves of order s". Perhaps $\mathcal{D}_s(\mathbb{R})$ could also be called the "wave space of order s". A space $\mathcal{D}_s(\mathbb{R})$ arises from $L^2(\mathbb{R})$ by a procedure of taking sums of certain α-differences. We can think of this procedure as an "operation $\mathcal{D}_s$". Thus, $\mathcal{D}_s(\mathbb{R})$ is obtained by applying operation $\mathcal{D}_s$ to $L^2(\mathbb{R})$. But what would happen if we applied operation $\mathcal{D}_s$ to a space $\mathcal{D}_t(\mathbb{R})$ in place of $L^2(\mathbb{R})$? Put in another way, what would we obtain by making waves of order s out of other waves of order t? Would we get the waves of order $s+t$? The answer is yes! This fact can be thought of as a semigroup property of the family of operations $\mathcal{D}_s, s \in (0, \infty)$; for it means that

$$\mathcal{D}_s \circ \mathcal{D}_t = \mathcal{D}_{s+t}, \quad \text{for all } s, t \in (0, \infty).$$

The proof of this depends on the fact that under the Fourier transform, products become convolutions. Now the space of absolutely summable sequences on $\mathbb{Z}$ forms an algebra with the usual addition and with convolution as multiplication. It follows

that the functions in $C_{2\pi}(\mathbb{R})$ which have absolutely convergent Fourier series form a subalgebra of $C_{2\pi}(\mathbb{R})$ [5, p.165].

Let $s, t \in (0,)$, let $\alpha, \beta \in C_{2\pi}(\mathbb{R})$, let α, β have absolutely convergent Fourier series, and let there exist $\delta_1, \delta_2 \in (0, \infty)$ such that

$$\delta_1 |x|^s \leq |\alpha(x)| \leq \delta_2 |x|^s \quad \text{and} \quad \delta_1 |x|^t \leq |\beta(t)| \leq \delta_2 |x|^t,$$

for all $x \in [-\pi, \pi]$. Then

$$\delta_1^2 |x|^{s+t} \leq |\alpha(x)\beta(x)| \leq \delta_2^2 |x|^{s+t}, \qquad \text{for all } x \in [-\pi, \pi].$$

Also, $\alpha\beta$ has an absolutely convergent Fourier series. By the definition of the spaces $\mathcal{D}_s(\mathbb{R})$, we have

$$\mathcal{D}_{[\alpha]}(\mathbb{R}) = \mathcal{D}_s(\mathbb{R}),\ \mathcal{D}_{[\beta]}(\mathbb{R}) = \mathcal{D}_t(\mathbb{R}) \text{ and } \mathcal{D}_{[\alpha\beta]}(\mathbb{R}) = \mathcal{D}_{s+t}(\mathbb{R}).$$

Now define the space $\mathcal{D}_{[\alpha]}(\mathcal{D}_{[\beta]}(\mathbb{R}))$ to consist of the functions in $L^2(\mathbb{R})$ which are finite sums of functions of the form $\sum_{k=-} \widehat{\alpha}(k)\delta_{-ky} * h$, for some $y \in \mathbb{R}$ and some β-difference h. That is, $\mathcal{D}_{[\alpha]}(\mathcal{D}_{[\beta]}(\mathbb{R}))$ consists of finite sums of functions which are α-differences obtained from β-differences. Consider a function $g \in L^2(\mathbb{R})$ which is assumed to be an $\alpha\beta$-difference. Then there is $f \in L^2(\mathbb{R})$ and $y \in \mathbb{R}$ such that

$$\begin{aligned}
g &= \sum_{j=-\infty}^{\infty} (\alpha\beta)^\wedge(j)\delta_{-jy} * f, \\
&= \sum_{j=-\infty}^{\infty} \left(\sum_{k+\ell=j} \widehat{\alpha}(k)\widehat{\beta}(\ell)\delta_{-(k+\ell)y} * f \right), \text{ as } (\alpha\beta)^\wedge = \widehat{\alpha} * \widehat{\beta}, \\
&= \sum_{k=-\infty}^{\infty} \widehat{\alpha}(k)\delta_{-ky} * \left[\sum_{\ell=-\infty}^{\infty} \widehat{\beta}(\ell)\delta_{-\ell y} * f \right],
\end{aligned}$$

which belongs to $\mathcal{D}_{[\alpha]}(\mathcal{D}_{[\beta]}(\mathbb{R}))$ as the function in square brackets is a β-difference. Hence,

$$\mathcal{D}_{s+t}(\mathbb{R}) = \mathcal{D}_{[\alpha\beta]}(\mathbb{R}) \subseteq \mathcal{D}_{[\alpha]}(\mathcal{D}_{[\beta]}(\mathbb{R})). \tag{5.2}$$

In fact, the converse inclusion also holds. For consider a function $f \in \mathcal{D}_{[\alpha]}(\mathcal{D}_{[\beta]}(\mathbb{R}))$. Then f is a finite sum of functions, each of the form $\sum_{k=-\infty}^{\infty} \widehat{\alpha}(k)\delta_{-ky} * h$, for some $y \in \mathbb{R}$ and some $h \in \mathcal{D}_{[\beta]}(\mathbb{R})$. Denoting this function by g, and noting that $h \in \mathcal{D}_{[\beta]}(\mathbb{R}) = \mathcal{D}_t(\mathbb{R})$, gives

$$\begin{aligned}
\int_{-\infty}^{\infty} \frac{|\widehat{g}(x)|^2}{|x|^{2(s+t)}} dx &= |y|^{2s} \int_{-\infty}^{\infty} \frac{|\widehat{h}(x)|^2 |\alpha(xy)|^2}{|x|^{2t}|xy|^{2s}} dx, \\
&\leq |y|^{2s}\delta_2^2 \int_{-\infty}^{\infty} \frac{|\widehat{h}(x)|^2}{|x|^{2t}} dx, \quad \text{as } |\alpha(x)| \leq \delta_2 |x|^s, \\
&< \infty, \quad \text{as } h \in \mathcal{D}_t(\mathbb{R}).
\end{aligned}$$

Here the Difference Space Theorem has been used, and it also gives from the above that $g \in \mathcal{D}_{s+t}(\mathbb{R})$. It follows that

$$\mathcal{D}_{[\alpha]}(\mathcal{D}_{[\beta]}(\mathbb{R})) \subseteq \mathcal{D}_{s+t}(\mathbb{R}). \tag{5.3}$$

Thus, (5.2) and (5.3) now give

$$\mathcal{D}_{[\alpha]}(\mathcal{D}_{[\beta]}(\mathbb{R})) = \mathcal{D}_{s+t}(\mathbb{R}), \tag{5.4}$$

which shows that $\mathcal{D}_{[\alpha]}(\mathcal{D}_{[\beta]}(\mathbb{R}))$ is independent of α and β. Since $\mathcal{D}_{[\beta]}(\mathbb{R}) = \mathcal{D}_t(\mathbb{R})$, it is thus consistent to interpret operation $\mathcal{D}_s$, when applied to $\mathcal{D}_t(\mathbb{R})$, as giving the space $\mathcal{D}_{[\alpha]}(\mathcal{D}_{[\beta]}(\mathbb{R}))$. With this definition, (5.4) can be expressed as the following result.

PROPOSITION 5. *If* $s, t \in (0, \infty)$, *then* $\mathcal{D}_s(\mathcal{D}_t(\mathbb{R})) = \mathcal{D}_{s+t}(\mathbb{R})$.

That is, by making waves of order s out of waves of order t, we obtain precisely the waves of order $s + t$. In the case when $s, t \in \mathbb{N}$, a proof of Proposition 5 probably can also be based upon the following observations: D^s is a bijection from $H^s(\mathbb{R})$ onto $\mathcal{D}_s(\mathbb{R})$ by Theorem 4, also D^t is a bijection from $H^t(\mathbb{R})$ onto $\mathcal{D}_t(\mathbb{R})$, and $D^s \circ D^t = D^{s+t}$ on $\mathcal{S}(\mathbb{R})$. The reader is invited to explore this alternative approach.

6. WAVELETS *ARE* WAVES. In wavelet analysis, a major theme is the expansion of functions in $L^2(\mathbb{R})$ in terms of orthonormal functions which arise by integer translations and dyadic dilations of a fixed function in $L^2(\mathbb{R})$. Thus, a function h in $L^2(\mathbb{R})$ is sought so that the family $\{2^{j/2}h(2^{-j}x - k) : j, k \in \mathbb{Z}\}$ is complete and orthonormal in $L^2(\mathbb{R})$. The function h is often required to have a localization property, which means that in some sense it vanishes rapidly at infinity – so that h is a "small" wave or, as it is often called, a *wavelet*. The functions in the family $\{2^{j/2}h(2^{-j}x - k) : j, k \in \mathbb{Z}\}$ also have the scaling property. This refers to the fact that as j varies in $\mathbb{Z}$, the functions $h(2^{-j}x - k)$ have graphs which arise by expanding or contracting the x-axis by powers of 2.

The prototypal example is the Haar wavelet h, given by

$$h(x) = \begin{cases} 1, & \text{if } 0 \le x < \frac{1}{2}, \\ -1 & \text{if } \frac{1}{2} \le x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Then h is localized in the sense that it is actually zero outside a bounded set, and the family $\{2^{j/2}h(2^{-j}x - k) : j, k \in \mathbb{Z}\}$ forms a complete orthonormal set in $L^2(\mathbb{R})$, as described in [4, pp.10-13], for example.

Localization and scaling are characteristic features of wavelet expansions, and this distinguishes them from classical Fourier expansions because the sine and cosine functions are not localized – they are waves rather than wavelets. The connection between wavelets and the difference spaces $\mathcal{D}_s(\mathbb{R})$ introduced earlier is indicated by the following observation: if f is the function $f(x) = 1$ for $x \in [0, \frac{1}{2})$, and $f(x) = 0$ for $x \notin [0, \frac{1}{2})$, then the Haar wavelet h equals $f - \delta_{1/2} * f$ which is in $\mathcal{D}_1(\mathbb{R})$. That is, the Haar wavelet is a "wave of order 1".

Wavelet analysis has recently received very clear and accessible treatments in the books by Ingrid Daubechies [4] and Yves Meyer [13], and elsewhere. In particular there is the expository article [18] by Robert Strichartz. Therefore, no proofs are presented in this section, which is purely descriptive.

Let $\mathbb{R}^*$ denote the set of non-zero real numbers. Let $u \in \mathbb{R}^*$, let $v \in \mathbb{R}$, and let $h_{u,v} \in L^2(\mathbb{R})$ be the function given by

$$h_{u,v}(x) = |u|^{-1/2} h\left(\frac{x-v}{u}\right), \qquad \text{for almost all } x \in \mathbb{R}.$$

Note that $||h_{u,v}||_2 = ||h||_2$. Let $\langle f, g\rangle$ denote the usual inner product of functions $f, g \in L^2(\mathbb{R}) : \langle f, g\rangle = \int_{-\infty}^{\infty} f(x)\overline{g(x)}dx$.

The (continuous) wavelet transform relative to h is the function U_h mapping $L^2(\mathbb{R})$ into the functions defined on $\mathbb{R}^* \times \mathbb{R}$ which is given by

$$U_h(f)(u,v) = \langle f, h_{u,v}\rangle = |u|^{-1/2} \int_{-\infty}^{\infty} f(x)\overline{h\left(\frac{x-v}{u}\right)}dx,$$

for $u \in \mathbb{R}^*,\ v \in \mathbb{R}$. The discrete wavelet transform is the function V_h from $L^2(\mathbb{R})$ into the functions on $\mathbb{Z} \times \mathbb{Z}$ which is given by

$$V_h(f)(j,k) = \langle f, h_{2^j,2^j k}\rangle = 2^{-j/2} \int_{-\infty}^{\infty} f(x)\overline{h(2^{-j}x - k)}dx,$$

for $j, k \in \mathbb{Z}$. The idea in these transforms is that $\langle f, h_{u,v}\rangle$ "samples" f in the vicinity of v at the scale u relative to the given function h. We would like to be able to reconstruct f from the sampled values, as the Fourier inversion theorem reconstructs a function from its Fourier transform, and we would like these transforms to have analogues of the Plancherel theorem.

In the case of the continuous wavelet transform, an important identity is the following, described by A. Grossman and J. Morlet [6], although the formulation here is more like that in [4, p.24]: if $f, h \in L^2(\mathbb{R})$, then

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{|U_h(f)(u,v)|^2}{u^2} dudv = \left(\int_{-\infty}^{\infty} \frac{|\widehat{h}(x)|^2}{|x|}dx\right)\left(\int_{-\infty}^{\infty} |f(x)|^2 dx\right). \tag{6.1}$$

Thus, U_h maps $L^2(\mathbb{R})$ into $L^2(\mathbb{R}^* \times \mathbb{R}, u^{-2}dudv)$ if and only if $\int_{-\infty}^{\infty} |\widehat{h}(x)|^2 |x|^{-1} dx < \infty$, and the latter is often called the *admissibility condition* on h and, when it is satisfied, h is sometimes called an *admissible wavelet*. In view of the Difference Space Theorem, h is an admissible wavelet in this sense if and only if $h \in \mathcal{D}_{1/2}(\mathbb{R})$. That is "$h$ is a wavelet if and only if it is a wave of order $1/2$". In this sense, wavelets *are* waves.

When $h \in L^2(\mathbb{R})$, (6.1) shows that U_h is an isometry, up to a scalar multiplication, from $L^2(\mathbb{R})$ into $L^2(\mathbb{R}^* \times \mathbb{R})$, so (6.1) may be regarded as a Plancherel-type Theorem for U_h. Similar remarks apply in the case of the discrete transform V_h,

for V_h will satisfy a Plancherel-type Theorem if $\{2^{-j/2}h(2^{-j}x - k) : j; k \in \mathbb{Z}\}$ is orthornormal in $L^2(\mathbb{R})$. But if this holds, Daubechies has proved [3, p.975] that again h is admissible, and so is a "wave of order 1/2". The reader is referred to [4, Chapter 3] for a full discussion of the discrete transform case, especially pages 63-66.

Among the spaces $\mathcal{D}_s(\mathbb{R})$, $s \in (0, \infty)$, the space $\mathcal{D}_{1/2}(\mathbb{R})$ plays a special rôle, which hinges upon the fact that $\int_{-1}^{1} |x|^{-2s} dx < \infty$ if and only if $s < \frac{1}{2}$. For if $f \in \mathcal{S}(\mathbb{R})$, its Fourier transform $\widehat{f}$ is also in $\mathcal{S}(\mathbb{R})$ and so is bounded on $\mathbb{R}$. Hence, if $f \in \mathcal{S}(\mathbb{R})$, $\int_{-\infty}^{\infty} |\widehat{f}(x)|^2 |x|^{-2s} dx < \infty$ for all $s \in (0, 1/2)$, so $\mathcal{S}(\mathbb{R}) \subseteq \mathcal{D}_s(\mathbb{R})$ for all $s \in (0, 1/2)$. However, if $f \in \mathcal{S}(\mathbb{R}) \cap \mathcal{D}_{1/2}(\mathbb{R})$, $\int_{\infty}^{\infty} |\widehat{f}(x)|^2 |x|^{-1} dx < \infty$ and so $\int_{-\infty}^{\infty} f(x) dx = \widehat{f}(0) = 0$. Thus, we see that every real valued function $f \in \mathcal{S}(\mathbb{R}) \cap \mathcal{D}_{1/2}(\mathbb{R})$ must have a crest *and* a trough, for otherwise we could not have $\int_{-\infty}^{\infty} f(x) dx = 0$. But although every real valued function in $\mathcal{S}(\mathbb{R})$ must have a crest *or* a trough, many do not have both a crest *and* a trough, although $\mathcal{S}(\mathbb{R}) \subseteq \mathcal{D}_s(\mathbb{R})$ for all $s \in (0, 1/2)$. We might say that the general function in $\mathcal{D}_s(\mathbb{R})$, for $s \in (0, 1/2)$, is a wave in a weaker sense than is the case for each function in $\mathcal{D}_{1/2}(\mathbb{R})$. An extreme point of view might be that the general function in $\mathcal{D}_s(\mathbb{R})$, for $s \in (0, 1/2)$, is not a wave at all. Here, the words of Humpty Dumpty in Lewis Carroll's *Through the Looking-Glass* are apposite: "When I use a word, it means just what I choose it to mean, — neither more nor less".

Finally, although wavelets are waves, not all waves are wavelets; at least, not in the sense that waves are localized. For if $y \in \mathbb{R}$ and $y \neq 0$, it is easy to see that f may be chosen in $L^2(\mathbb{R})$ so that $f - \delta_y * f$ does not vanish rapidly at infinity, even though $f - \delta_y * f$ always does belong to $\mathcal{D}_1(\mathbb{R})$. On the other hand, functions in $L^2(\mathbb{R})$ may be regarded as being "localized", but in an L^2-sense rather than in the sense of rapidly decreasing at infinity.

7. REMARKS, MAINLY CIRCULAR. Let $\mathbb{T}$ denote the circle group, the set of complex numbers of unit modulus. Let μ denote the usual arc-length measure on $\mathbb{T}$, normalised so that $\mu(\mathbb{T}) = 1$. As μ is invariant under translations (rotations) in $\mathbb{T}$, if $f \in L^2(\mathbb{T})$ and $y \in \mathbb{T}$, $\int_{\mathbb{T}} (f - \delta_y * f) d\mu = 0$. Motivated by the fact that $\mathcal{D}_1(\mathbb{R})$ consists of all finite sums of first order differences in $L^2(\mathbb{R})$, let us define $\mathcal{D}_1(\mathbb{T})$ to consist of all finite sums of functions of the form $f - \delta_y * f$ for $f \in L^2(\mathbb{T})$ and $y \in \mathbb{T}$. The previous remark implies that $\int_{\mathbb{T}} h d\mu = 0$ for all $h \in \mathcal{D}_1(\mathbb{T})$. Now Gary Meisters and Wolfgang Schmidt [10] proved the converse of this statement, thus deducing that

$$\mathcal{D}_1(\mathbb{T}) = \left\{ f : f \in L^2(\mathbb{T}) \text{ and } \int_{\mathbb{T}} f d\mu = 0 \right\}. \tag{7.1}$$

They deduced from this that if J is a linear form on $L^2(\mathbb{T})$, then $J(\delta_y * f) = J(f)$ for all $y \in \mathbb{T}$ and $f \in L^2(\mathbb{T})$ if and only if J is a scalar multiple of μ. Thus, they proved that if J is a translation invariant linear form on $L^2(\mathbb{T})$, which is to say that $J(\delta_y * f) = J(f)$ for all $f \in L^2(\mathbb{R})$ and $y \in \mathbb{R}$, then J is a multiple of the Haar measure μ on $\mathbb{T}$ and so J must be continuous. Prior to proving (7.1), Meisters and

Schmidt gave a preliminary characterization of $\mathcal{D}_1(\mathbb{T})$, the proof of which has been adapted here to the non-compact case to give Proposition 3.

In [11], Meisters noted that if $y \in L^2(\mathbb{R})$ is a finite sum of first order differences then $\int_{-\infty}^{\infty} |\widehat{g}(x)|^2 |x|^{-2} dx < \infty$, an observation which is here incorporated into Proposition 2. He deduced from this that there are translation invariant linear forms on $L^2(\mathbb{R})$ which are *not* continuous, in contrast to the circle group case.

The work of the author in [14] and [15] resulted from an attempt to extend the work of Meisters and Schmidt from the compact case of the circle group to the non-compact case of $\mathbb{R}^n$. The present paper considers the case of $\mathbb{R}$ only, and an aim has been to make the underlying ideas of [14] and [15] as accessible and as interesting as possible to anyone familiar with the Fourier transform and the space $L^2(\mathbb{R})$.

The connection between the compact and non-compact case can be understood better if (7.1) is interpreted in terms of the Fourier transform. Recall that if $f \in L^2(\mathbb{T}), \widehat{f}$ is defined on $\mathbb{Z}$ and is given by $\widehat{f}(n) = \int_{\mathbb{T}} f(t) t^{-n} d\mu(t)$, for $n \in \mathbb{Z}$. Then (7.1) may be written as

$$\mathcal{D}_1(\mathbb{T}) = \Big\{ f : f \in L^2(\mathbb{T}) \text{ and } \widehat{f}(0) = 0 \Big\}.$$

Thus, membership of f in $\mathcal{D}_1(\mathbb{T})$ is characterized by the behaviour of $\widehat{f}$ at the origin. The point is that within $\mathbb{Z}$, the origin is an isolated point, so that the behaviour of $\widehat{f}$ *near* the origin reduces to the question of behaviour of $\widehat{f}$ *at* the origin. But this is not the case when $f \in L^2(\mathbb{R})$, because then $\widehat{f}$ is defined (almost everywhere) on $\mathbb{R}$, the origin is not an isolated point within $\mathbb{R}$, and the behaviour of $\widehat{f}$ near the origin does *not* reduce to the question of the behaviour of $\widehat{f}$ at the origin, even if $\widehat{f}$ is continuous. For $f \in L^2(\mathbb{R})$, there is the continuum of possible behaviours of $\widehat{f}$ near the origin, one for each $s \in (0, \infty)$, as expressed by requiring that $\int_{-\infty}^{\infty} |\widehat{f}(x)|^2 |x|^{-2s} dx < \infty$.

The result, contained within Theorem 4, that the differentation operator D is a norm preserving map from $H^1(\mathbb{R})$ onto $\mathcal{D}_1(\mathbb{R})$, suggests the question: what is the corresponding result on the circle group? To describe this, let $C^\infty(\mathbb{T})$ denote the infinitely differentiable complex valued function on $\mathbb{T}$. If $[-\pi, \pi)$ is identified with $\mathbb{T}$ under the map $x \longmapsto e^{ix}, C^\infty(\mathbb{T})$ may be thought of as the C^∞-functions on $[-\pi, \pi]$ all of whose derivatives are equal at $-\pi$ and π.

Now if $\phi_1, \phi_2 \in C^\infty(\mathbb{T})$, and if D denotes differentation,

$$\int_{\mathbb{T}} \phi_1 D(\phi_2) d\mu = \int_{-} \phi_1(t) D(\phi_2)(t) dt = -\int_{-} D(\phi_1)(t) \phi_2(t) dt = -\int_{-} D(\phi_1) \phi_2 d\mu.$$

Then if $f \in L^2(\mathbb{T})$, we say $D(f) \in L^2(\mathbb{T})$ if there is $g \in L^2(\mathbb{T})$ such that

$$\int_{\mathbb{T}} f D(\phi) d\mu = -\int_{\mathbb{T}} g \phi d\mu, \qquad \text{for all } \phi \in C^\infty(\mathbb{T}),$$

and write $g = D(f)$. By analogy with $H^1(\mathbb{R})$, the space $H^1(\mathbb{T})$ is defined to consist of all $f \in L^2(\mathbb{T})$ such that $D(f) \in L^2(\mathbb{T})$. It is straightforward to check that for

$f \in H^1(\mathbb{T})$, $(Df)^\wedge(n) = in\widehat{f}(n)$ for all $n \in \mathbb{Z}$. In fact, if $f \in L^2(\mathbb{T})$, $f \in H^1(\mathbb{T})$ if and only if $\sum_{n=-\infty}^{\infty} |n|^2 |\widehat{f}(n)|^2 < \infty$, and $H^1(\mathbb{T})$ becomes a Hilbert space in the inner product

$$\langle f, g\rangle = \sum_{n=-\infty}^{\infty} (1+n^2)\widehat{f}(n)\overline{\widehat{g}(n)}, \qquad \text{for } f, g \in H^1(\mathbb{T}).$$

Now (7.1) implies that $\mathcal{D}_1(\mathbb{T})$ is closed in $L^2(\mathbb{T})$, so $\mathcal{D}_1(\mathbb{T})$ is a Hilbert space in the $L^2(\mathbb{T})$-inner product. However the inner product on $\mathcal{D}_1(\mathbb{T})$ given by

$$\langle f, g\rangle = \sum_{n\neq 0} (1+n^{-2})\widehat{f}(n)\overline{\widehat{g}(n)}, \qquad \text{for } f, g \in \mathcal{D}_1(\mathbb{T}),$$

gives a norm on $\mathcal{D}_1(\mathbb{T})$ which is equivalent to the $L^2(\mathbb{T})$-norm (recall that $f \in \mathcal{D}_1(\mathbb{T}) \iff \widehat{f}(0) = 0$). With the above inner product on $\mathcal{D}_1(\mathbb{T})$, differentiation becomes a norm preserving linear mapping from $H^1(\mathbb{T})$ onto $\mathcal{D}_1(\mathbb{T})$, so $\mathcal{D}_1(\mathbb{T})$, like $\mathcal{D}_1(\mathbb{R})$, may be regarded as the range of D on its "natural" domain, in this case the space $H^1(\mathbb{T})$.

It should be noted that the above mentioned results of Meisters and Schmidt [10] apply to any compact abelian group with a finite number of components. In [12], Gary Meisters described a result he obtained with Larry Baggett, which gave a necessary condition on a compact abelian group G in order that every translation invariant linear form on $L^2(G)$ be continuous. This necesary condition was shown to be sufficient by Barry Johnson [8]. Jean Bourgain [2] showed that on $\mathbb{T}$, and for $1 < p < \infty$, (7.1) holds for the appropriate formulation for $L^p(\mathbb{T})$ in place of $L^2(\mathbb{T})$, thus deducing that any translation invariant linear form on $L^p(\mathbb{T})$ is continuous.

Finally, it might be wondered if the description of the ranges of D^s as the spaces $\mathcal{D}_s(\mathbb{R})$, as in Theorem 4, has an analogue for partial differential operators. For a reasonably large class of operators, including the Laplace operator and the one-dimensional wave operator, this is the case. The interested reader is referred to [15] for full details. In the case of the one-dimensional wave operator $\dfrac{\partial^2}{\partial x^2} - \dfrac{\partial^2}{\partial y^2}$, its range on a suitable Soblev-type subsuspace of $L^2(\mathbb{R}^2)$ consists of all functions in $L^2(\mathbb{R}^2)$ which can be expressed as the sum of nine functions in $L^2(\mathbb{R}^2)$, each one of which is the form

$$f - f(x+a, y+a) - f(x+b, y-b) + f(x+a+b, x+a-b), \quad \text{for almost all} \quad (x,y) \in \mathbb{R}^2,$$

for some $a, b \in \mathbb{R}$ and $f \in L^2(\mathbb{R}^2)$.

REFERENCES

[1] U . M. ASCHER, R. M. M. MATHEIJ and R. R. RUSSELL, Numerical Solution of Boundary Value Problems for Ordinary Differential Equations, Prentice Hall, Englewood Cliffs, New Jersey, 1988.

[2] J. BOURGAIN, *Translation invariant forms on $L^p(G)(1 < p <)$*, Ann. Inst. Fourier (Grenoble), 36 (1986), 97-104.

[3] I. DAUBECHIES, *The wavelet transform, time-frequency localization and signal analysis*, IEEE Trans. Inform. Theory, 36 (1990), 961-1005.

[4] I. DAUBECHIES, Ten lectures on Wavelets, Society for Industrial and Applied Mathematics, Philadephia, 1992.

[5] R. E. EDWARDS, Fourier Series: a modern introduction, vol. I, Holt, Rinehart and Winston, New York, 1967.

[6] A. GROSSMAN and J. MORLET, *Decomposition of Hardy functions into square integrable wavelets of constant shape*, Siam J. Math. Anal. 15 (1984), 723-736.

[7] E. HEWITT and K. STROMBERG, Real and Abstract Analysis, Springer-Verlag, New York, 1965.

[8] B. E. JOHNSON, *A proof of the translation invariant form conjecture for $L^2(G)$*, Bull. des Sciences Math, 107 (1983), 301-310.

[9] W. L. LO, *Representation of derivatives of functions in Sobolev spaces in terms of finite differences*, Proceedings of the Miniconference on Analysis and Applications held at the University of Queensland, 20-23 September 1993, Centre for Mathematics and its Applications, Australian National University (to appear).

[10] G. H. MEISTERS and W.M. SCHMIDT, *Translation invariant linear forms on $L^2(G)$ for compact abelian groups G*, Jour. Funct. Anal. 11 (1972), 407-424.

[11] G. H. MEISTERS, *Some discontinuous translation invariant linear forms*, Jour. Function Anal. 12 (1973), 199-210.

[12] G. H. MEISTERS, *Some problems and results on translation-invariant linear forms*, Lecture Notes in Mathematics, vol. 975 (J. M. Bachar, W. G. Bade et al, eds.), 423-444, Springer Verlag, New York, 1983.

[13] Y. MEYER, Wavelets and Operators, translated by D. H. Salinger, Cambridge University Press, Cambridge, 1992.

[14] R. NILLSEN, *Banach spaces of functions and distributions characterized by singular integrals involving the Fourier transform*, Journ. Func. Anal. 110 (1992), 73-95.

[15] R. NILLSEN, Difference Spaces and Invariant Linear Forms, Lecture Notes in Mathematics vol.1586, Springer-Verlag, Heidelberg, New York, 1994.

[16] W. RUDIN, Functional Analysis, 2nd ed., McGraw-Hill, New York, 1991.

[17] L. SCHWARTZ, Théorie des Distibutions, Hermann, Paris, 1966.

[18] R. S. STRICHARTZ, *How to make wavelets*, Amer. Math. Monthly, 100 (1993), 539-556.

[19] A. ZYGMUND, Trigonometric Series, Cambridge University Press, Cambridge, 1968.

Department of Mathematics
University of Wollongong
Northfields Avenue
WOLLONGONG NSW 2522
AUSTRALIA