Multi-index notation

A **multi-index**, or **multi-order**, is an element α of $\mathbf{N}^n$. That is, α is an n-tuple of non-negative integers: $\alpha = (\alpha_1, \cdots, \alpha_n)$.

The **order**, or **length**, of α is $|\alpha| := \alpha_1 + \cdots + \alpha_n$.

If α and β are multi-indices, we say that $\boldsymbol{\alpha \le \beta}$ if $\alpha_k \le \beta_k$ for $k = 1, \ldots, n$.

If α is a multi-index, $\boldsymbol{\alpha!} := \alpha_1! \cdots \alpha_n!$. This is read “alpha factorial.”

If α and γ are multi-indices, $\binom{\boldsymbol{\alpha}}{\boldsymbol{\gamma}} := \dfrac{\alpha!}{(\alpha-\gamma)!\gamma!}$; this is read “alpha choose gamma.”

If x is in $\mathbf{R}^n$, and α is a multi-index, $\mathbf{x}^{\boldsymbol{\alpha}} := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. We read this as “x to the alpha (power).”

We use multi-orders to express the operator of (mixed) partial differentiation. Two fairly common notations we will use are

$$\left(\frac{\boldsymbol{\partial}}{\boldsymbol{\partial x}}\right)^{\boldsymbol{\alpha}} \text{ and } \mathbf{D}^{\boldsymbol{\alpha}} \text{ to denote } \frac{\partial^{\alpha_1 + \cdots + \alpha_n}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} .$$

We also express $D^\alpha f$ as $\mathbf{f}^{(\boldsymbol{\alpha})}$. A wide variety of notations is used for differentiation.

A **polynomial of degree d** in $x \in \mathbf{R}^n$ is a function of the form $P(x) := \sum_{|\alpha| \le d} c_\alpha x^\alpha$, where the **coefficients** c_α are complex numbers and d is in $\mathbf{N}$, and at least one c_α with $|\alpha| = d$ is non-zero.

We will use the notations P_α and $\overline{P}_\alpha$, respectively, to denote the **multiplication operators** defined by $\mathbf{P}_{\boldsymbol{\alpha}}\mathbf{f(x)} := (ix)^\alpha f(x)$ and $\overline{\mathbf{P}}_{\boldsymbol{\alpha}}\mathbf{f(x)} := (-ix)^\alpha f(x)$.

Exercise: Prove that, for d in $\mathbf{N}$, $(x_1 + \cdots + x_n)^d = \sum_{|\alpha| = d} \frac{d!}{\alpha!} x^\alpha$. This is also known as a “multinomial theorem.”

Exercise: Prove **Leibniz’ Rule**: $(fg)^{(\alpha)} = \sum_{\gamma \le \alpha} \binom{\alpha}{\gamma} f^{(\gamma)} g^{(\alpha-\gamma)}$. Assume all the indicated partial derivatives are continuous functions.

Exercise: Determine how many multi-indices α there are, for a given n, with $|\alpha| = d$.

Exercise: Find the formula for $(x + y)^\alpha$ in terms of powers x^β and y^γ. Here, x and y are in $\mathbf{R}^n$.

Let $\mathcal{S}$ denote the (Laurent) **Schwartz class** of “rapidly decreasing functions,” defined to be the set of all C^∞ complex-valued functions $f(x)$ defined on $\mathbf{R}^n$ such that, for all multi-indices α and multi-orders β, $x^\alpha f^{(\beta)}(x)$ is a bounded function of x. The bound can depend on α and β.

Theorem: $S \subseteq L^1(\mathbf{R}^n) \cap C_o(\mathbf{R}^n)$.

Proof: Since $(1 + |x|^2)^n$ is a polynomial, and $(1 + |x|^2)^n f(x)$ is bounded if f is in S, the result follows at once because $(1 + |x|^2)^{-n}$ is integrable and in $C_o(\mathbf{R}^n)$.

Theorem: S is a complex linear space that is closed under partial differentiations of all orders, multiplication by polynomials, and pointwise products (so that S is an algebra).

Proof is straightforward, and uses Leibniz' Rule.

Let M denote the set of all C^∞ complex-valued functions $m(x)$ defined on $\mathbf{R}^n$ such that, for each multi-order α, $m^{(\alpha)}(x)$ is bounded by a polynomial in x. The bound can depend on α.

Theorem: $MS \subseteq S$.

Proof is straightforward, and uses Leibniz' Rule.

Lemma (on integration by parts): If f and g are in $M \cup S$, and at least one of them is in S, then for $k = 1, \ldots, n$, $\int \frac{\partial f}{\partial x_k} g \, dx = -\int f \frac{\partial g}{\partial x_k} dx$.

Proof: Without loss of generality, we may assume $k = n$. Using Fubini's theorem we write the integral as an iterated integral with the dx_n integral performed first. If we prove the theorem for $n = 1$, the rest now follows. Since the integral is the limit, as $R \to +\infty$, of the integrals over $(-R, R)$, we can use ordinary integration by parts on each such integral, and note that the boundary terms tend to 0, by the previous theorem.

The following lemma *could* be phrased in terms making little use of Lebesgue theory. Since we have it, though, we will use it, and probably achieve supefluous generality.

Lemma (on differentiation under the integral sign): Let Y be a measure space, with measure μ, and suppose that

(i) $F(x, y)$ is measurable on $I \times Y$, where I is an open interval contained in $\mathbf{R}$;
(ii) for each x in I, $F(x, y)$ is an integrable function of y;
(iii) for a.e. y in Y, $F(x, y)$ is differentiable in I;
(iv) there exists a function g in $L^1(d\mu)$ such that, for all x in I, and a.e. y in Y, $\left|\frac{\partial F}{\partial x}(x,y)\right| \leq g(y)$.

Then, for all x in I, $\frac{d}{dx}\int_Y F(x, y)\, d\mu(y) = \int_Y \frac{\partial F}{\partial x}(x,y)\, d\mu(y)$.

Scholium: The equation implicitly states that the derivative exists!

Proof of the lemma amounts to setting up the difference quotient, applying the Mean Value Theorem for a.e. y, and using Lebesgue's Dominated Convergence Theorem.

The Fourier transformation and the Schwartz class

There is a natural affinity between Fourier transforms and $\mathcal{S}$. Our first objective is to show that the Fourier transformation is an invertible linear mapping on $\mathcal{S}$. The "affinity" referred to will become clear as we derive several formulas relating the composition of the Fourier transformation with operators involved in the definition of $\mathcal{S}$.

We define the Fourier transform, $\hat{f}$ of f in $\mathcal{S}$ to be the function given by

$\hat{f}(\xi) := \int e^{-i\xi\bullet x} f(x)\, dx$. We will sometimes write $\mathcal{F} f$ for $\hat{f}$.

We have to show, for each f in $\mathcal{S}$, that $\xi^{\alpha}\hat{f}^{(\beta)}(\xi)$ is a bounded function of ξ.
The first step is to show that $\hat{\mathcal{S}} \subseteq C^{\infty}$.
Theorem: If $f \in \mathcal{S}$, then $\hat{f}$ is a C^{∞} function.

Proof: We can apply the differentiation lemma with $Y = \mathbf{R}^n$, $d\mu(y) := |f(y)|\, dy$, and $F(\xi, y) := e^{-i\xi\bullet y} \operatorname{sgn} f(y)$. Recall that $\operatorname{sgn} z := z/|z|$ for a complex number $z \neq 0$; $\operatorname{sgn} 0 = 0$.
We find, inductively, that $D^{\alpha}\hat{f}(\xi) = \int e^{-i\xi\bullet y} (-iy)^{\alpha} f(y)\, dy$, or

$$\boxed{D^{\alpha} \mathcal{F} = \mathcal{F}\, \overline{P}_{\alpha}\,.}$$

Since f in $\mathcal{S}$ implies $\overline{P}_{\alpha} f$ is in $\mathcal{S}$, and $\mathcal{S}$ is contained in L^1, the desired result follows since the Fourier transform of an integrable function is continuous.

Theorem: If $f \in \mathcal{S}$, then for each multi-index α, $P_{\alpha}\hat{f}$ is a bounded function.

Proof: Since S is closed under partial differentiations, we can prove the theorem by verifying the formula

$$\boxed{\mathcal{F} D^{\alpha} = P_{\alpha} \mathcal{F}.}$$

We do this by applying the integration by parts lemma, inductively:
$\int e^{-i\xi\bullet x} D^{\alpha} f(x)\, dx = \int (i\xi)^{\alpha} e^{-i\xi\bullet x} f(x)\, dx$. The minus signs in the integrations by parts are cancelled.

Theorem: $\mathcal{F}\mathcal{S} \subseteq \mathcal{S}$.

Proof: We have $\xi^{\alpha}\hat{f}^{(\beta)}(\xi) = \xi^{\alpha}\left((-ix)^{\beta} f(x)\right)^{\wedge}(\xi) = (-i)^{\alpha}\left(D^{\alpha}\left((-ix)^{\beta} f(x)\right)\right)^{\wedge}(\xi)$, by the last two theorems (we have to juggle powers of $-i$, and abuse notation; $(-i)^{\alpha}$ "should" be
$(-i)^{|\alpha|}$).

Next we prove the inversion theorem for S, by using an approximate identity based on an important example of an S function.

Example: $\int e^{-i\xi\bullet x} e^{-|x|^2/2}\, dx = (2\pi)^{n/2} e^{-|\xi|^2/2}$.

It is routine to verify that $e^{-|x|^2/2}$ is in S, using the fact that a partial derivative of this function is a polynomial times $e^{-|x|^2/2}$. The formula is derived by Jones: FT16g, pp 301-302. We need a variant of this, that uses the formula for an affine change of variables:

$$\int e^{-i\xi\bullet x} e^{-\varepsilon|x|^2/2}\, dx = \left(\frac{2\pi}{\varepsilon}\right)^{n/2} e^{-|\xi|^2/2\varepsilon}.$$

The change of variables is $Tx = x/\sqrt{\varepsilon}$. The general formula in this case is

$$f(Tx)\hat{}(\xi) = \varepsilon^{-n/2} \hat{f}(\varepsilon^{1/2}\xi) .$$

The Inversion Formula we want is

$$\boxed{f(x) = (2\pi)^{-n}\int e^{i\xi\bullet x} \hat{f}(\xi)\, d\xi.}$$

We cannot replace $\hat{f}(\xi)$ by the integral and change the order of integration. But we can take the limit, as $\varepsilon \to 0$, of $f_\varepsilon(x) := (2\pi)^{-n}\int e^{-\varepsilon|\xi|^2/2} e^{i\xi\bullet x} \hat{f}(\xi)\, d\xi$, and show that it is $f(x)$. In the last formula, we *can* change the order of integration:

$$f_\varepsilon(x) := (2\pi)^{-n}\int e^{-\varepsilon|\xi|^2/2} e^{i\xi\bullet x} \left(\int e^{-i\xi\bullet y} f(y)\, dy\right) d\xi$$

$$= (2\pi)^{-n}\int \left(\int e^{-\varepsilon|\xi|^2/2} e^{-i\xi\bullet(y-x)}\, d\xi\right) f(y)\, dy$$

$$= \int (2\pi\varepsilon)^{-n/2} e^{-|y-x|^2/2\varepsilon} f(y)\, dy.$$

This is one of the approximate identity formulas. Since f is continuous and bounded, we can apply the substitute result for L^∞ (see Jones, Ch 12 # 17). In our case, the approximate identity has integral 1, so $f_\varepsilon(x)$ converges to $f(x)$, as desired. And the integral that defines $f_\varepsilon(x)$ converges to the integral on the right in the Inversion Formula.

We have proved:

Theorem: F is invertible on S, and $F^{-1}f(x) = (2\pi)^{-n} (F f)(-x)$.

Some applications of the Fourier Inversion Theorem for S.

1). **The Fourier transformation is one-to-one on S.** This follows from the invertibility.

It is tempting to try to prove this without using the inversion theorem.

2). Let us define some semi-norms on $\mathcal{S}$.

Definition: For f in $\mathcal{S}$, and α and β in $\mathbf{N}^n$, let $\| f \|_{\alpha\beta} := \sup_{x \in \mathbf{R}^n} |x^\alpha f^{(\beta)}(x)|$.

Theorem: For each α and β in $\mathbf{N}^n$, $\| f \|_{\alpha\beta}$ is a norm on $\mathcal{S}$.

To see this, we note that $\| f \|_{\alpha\beta} = 0$ implies that $x^\alpha f^{(\beta)}(x)$ is identically 0. Since x^α is 0 in a closed set of measure 0, and $f^{(\beta)}(x)$ is continuous, we conclude that $f^{(\beta)}(x)$ is identically 0. Then, since $\mathcal{F} D^\beta = P_\beta \mathcal{F}$, we conclude, as before, that the Fourier transform of f is identically 0. It follows that f is identically 0, by application 1).

3). **Theorem:** $\mathcal{S}$ is closed under convolution.

Proof: This can be shown directly from the definition of $\mathcal{S}$. To show that $\| f*g \|_{\alpha\beta} < \infty$ we can use the differentiation lemma to differentiate, say, f. Then we write
$x^\alpha (f*g)^{(\beta)}(x) = x^\alpha (f^{(\beta)}*g)(x) = x^\alpha \int f^{(\beta)}(x-y)\, g(y)\, dy =$

$$\int (x-y+y)^{\alpha} f^{(\beta)}(x-y)\, g(y)\, dy = \int (x-y+y)^{\alpha} f^{(\beta)}(x-y)\, g(y)\, dy.$$

This is equal to a sum of terms of the form const. $\int (x-y)^{\alpha-\gamma} f^{(\beta)}(x-y)\, y^\gamma g(y)\, dy$. We recognize that each of these is the convolution of two integrable functions, hence is bounded.

We can also prove this using the inversion theorem and the equation $\int \hat{f}(\xi)\; g(\xi)\, d\xi =$ $\int f(x)\, \hat{g}(x)\, dx$, valid for pairs f, g of integrable functions. Suppose f and g are in S. Then $f*g(x) = \int f(x-y)\, g(y)\, dy = (2\pi)^{-n} \int f(x-y)\, \hat{\hat{g}}(-y)\, dy =$ $(2\pi)^{-n} \int f(x+y)\, \hat{\hat{g}}(y)\, dy = (2\pi)^{-n} \int f(x+\bullet)^\wedge(\eta)\, \hat{g}(\eta)\, d\eta = (2\pi)^{-n} \int e^{i\eta\bullet x}\, \hat{f}(\eta)\, \hat{g}(\eta)$ $d\eta$. We recognize from this equation that $f*g$ is the Fourier transform of a function in $\mathcal{S}$, namely the product of the two functions $\hat{f}$ and $\hat{g}$, and hence $f*g$ is in $\mathcal{S}$.

4). **A formula for the Fourier transform of a product.**

If f and g are in $\mathcal{S}$, then $\boxed{\mathcal{F}\, fg = (2\pi)^{-n}\, \mathcal{F} f * \mathcal{F} g.}$

To show this, we use the second proof that $\mathcal{S}$ is closed under convolution, with f and g replaced by their Fourier transforms. In the last member of the equations we recognize the Fourier inversion formula.

5). **Parseval's formula for $\mathcal{S}$.**

$$\int |f(x)|^2\, dx = (2\pi)^{-n} \int |\hat{f}(\xi)|^2\, d\xi.$$

This is an application of 4), and recognition that the integral of a function is the value of its Fourier transform at 0. To carry it out we also need a formula for the Fourier transform of the complex conjugate of a function in terms of its Fourier transform. Let $Cf(x) := \overline{f(x)}$. This is not a linear map, but is *conjugate* linear. Such mappings have the same continuity properties as linear ones. Let $Rf(x) := f(-x)$. Then $RC = CR$, and $\mathcal{F}R = R\mathcal{F}$, but

$$\int e^{-i\xi\bullet x}\, \overline{f(x)}\, dx = \overline{\int e^{i\xi\bullet x} f(x)\, dx} = \overline{\hat{f}(-\xi)} .$$

That is, $\mathcal{F}C = RC\mathcal{F}$. Then, in 4), $\mathcal{F} fCf = (2\pi)^{-n} \mathcal{F} f * \mathcal{F} Cf = (2\pi)^{-n} \mathcal{F} f * RC\mathcal{F} f$. That is, $\int e^{-i\xi\bullet x} |f(x)|^2\, dx = (2\pi)^{-n} \int \hat{f}(\xi-\eta)\, \overline{\hat{f}(-\eta)}\, d\eta = (2\pi)^{-n} \int \hat{f}(\xi+\eta)\, \overline{\hat{f}(\eta)}\, d\eta.$ Finally, we set $\xi = 0$ to get the desired result. This result says that, when properly normalized, the Fourier transformation is an isometry with respect to the L^2 norm on $\mathcal{S}$.

Some applications of these properties of $\mathcal{S}$ outside $\mathcal{S}$.

1. **The Fourier Inversion Theorem on L^1.** If f and $\hat{f}$ are both in $L^1(\mathbf{R}^n)$, then f is equal almost everywhere to a continuous function, and for almost all x,

$$f(x) = (2\pi)^{-n} \int e^{i\xi\bullet x} \hat{f}(\xi)\, d\xi.$$

This says that if the Fourier transform of f is integrable, then f is the Fourier transform of an integrable function. The set of all such functions is an algebra with respect to both convolution and pointwise multiplication.

Proof 1(outline): Examine the proof for $\mathcal{S}$, and note that all we used there was an approximate identity in $\mathcal{S}$ and the fact that, for f in $\mathcal{S}$, $\hat{f}$ is in L^1. Here, we assume that $\hat{f}$ is in L^1. In fact, we can use approximate identities that are not in $\mathcal{S}$.

Proof 2: Since $\mathcal{S}$ contains the infinitely differentiable functions with compact support, $\mathcal{S}$ is dense in L^1. Thus there exists a sequence $f_k \to f$, f_k in $\mathcal{S}$. As in the proof of the completeness of L^p, a subsequence exists that converges pointwise, a.e. Without loss (by renumbering) we may assume the sequence converges both pointwise a.e. and in L^1. Then $\hat{f}_k - \hat{f}$ converges uniformly to zero.
By the inversion theorem,

$$(2\pi)^{-n}\hat{\hat{f}}_k(-x) - (2\pi)^{-n}\hat{\hat{f}}(-x) \text{ converges a.e. to } f(x) - (2\pi)^{-n}\hat{\hat{f}}(-x) .$$

Thus, for any C^∞ function g with compact support, $\int \hat{g}(\hat{f}_k - \hat{f}) \to 0$.

Hence $\int g(\hat{f}_k - \hat{f})^\wedge \to 0$. That is, $\int g((2\pi)^n f - \hat{\hat{f}}) = 0$ for all such g. In any ball there is a sequence of such g's that converges pointwise a.e. to $\operatorname{sgn}((2\pi)^n f - \hat{\hat{f}})$.

Thus $\int_{|x|<R} \left|(2\pi)^n f - \hat{\hat{f}}\right| dx = 0$ for all R, so $f = (2\pi)^{-n}\hat{\hat{f}}$, as desired.

Remark: Four sentences in the proof, beginning with "Thus, for any C^∞ function g," comprise a very useful tool that deserves to be emphasized. One consequence is:

Theorem: If f is locally integrable, then for $1 \le p \le \infty$,

$$\|f\|_p = \sup\left\{ \int f g : g \in C_c^\infty, \text{ and } \|g\|_{p'} \le 1 \right\}.$$

The proof is essentially covered in those three sentences.

Example: The Fourier transform of $e^{-|x|}$ is $\dfrac{2}{1+\xi^2}$. This is an integrable function. By the inversion theorem, its Fourier transform is $2\pi\, e^{-|-x|} = 2\pi\, e^{-|x|}$. That is,

$$\int \frac{e^{-i\xi x}}{1+x^2}\, dx = \pi\, e^{-|\xi|}.$$

2. **Definition of the Fourier transformation on L^2.** Parseval's formula for $\mathcal{S}$ can be rewritten $\int |\hat{f}(\xi)|^2\, d\xi = (2\pi)^n \int |f(x)|^2\, dx$. Parseval's formula shows that $\mathcal{F}$ is continuous on $\mathcal{S}$ with respect to the L^2 norm. Let us define an operator $\mathcal{F}_2$ on L^2 by a procedure known as "extension by continuity from a dense subset." We will normalize the operator so that $\mathcal{F}_2$ becomes an isometry. Let $f \in L^2$. Since $\mathcal{S}$ is dense in L^2, there exists a sequence $f_k \to f$, f_k in $\mathcal{S}$. The convergence is in L^2. Then $(2\pi)^{-n/2}\mathcal{F} f_k$ is a Cauchy sequence in L^2. If some other sequence $g_k \to f$, g_k in $\mathcal{S}$, the sequence $(2\pi)^{-n}\mathcal{F}(f_k - g_k)$ converges to 0 in L^2. Thus

$$\mathcal{F}_2 f := L^2\text{-}\lim_{k\to\infty} (2\pi)^{-n/2}\mathcal{F} f_k$$

is well defined, and $\left\|\mathcal{F}_2 f\right\|_2 = \lim_{k\to\infty} \left\|(2\pi)^{-n/2}\mathcal{F} f_k\right\|_2 = \lim_{k\to\infty} \left\|f_k\right\|_2 = \|f\|_2$.

This argument shows also that $\mathcal{F}_2$ is invertible, with $\mathcal{F}_2^{-1} = R\mathcal{F}_2$.

Example: The Fourier transform of the characteristic function of $(-1, 1)$ is $\dfrac{2\sin\xi}{\xi}$. We have, since this last function is in L^2, that $\dfrac{2}{\pi}\int \dfrac{\sin^2\xi}{\xi^2}\, d\xi = 2$, by the isometric property of $\mathcal{F}_2$, or $\int \dfrac{\sin^2\xi}{\xi^2}\, d\xi = \pi$.

3. **The Fourier transform of a product.**

If f and g are in L^2, or if f and g are in L^1, and the Fourier transform of one of them is in L^1, then $\mathcal{F}\, fg = (2\pi)^{-n} \mathcal{F} f * \mathcal{F} g$.

In either case, the product fg is integrable. If $f_k \to f$, f_k in $\mathcal{S}$, and $g_k \to g$, g_k in $\mathcal{S}$, the convergence taking place in L^2, it follows from $\mathcal{F}\, f_k g_k = (2\pi)^{-n} \mathcal{F} f_k * \mathcal{F} g_k$ that the desired result holds. In the second case we take $f_k \to f$, f_k in $\mathcal{S}$, $h_k \to (2\pi)^{-n} R\hat{g}$, h_k in $\mathcal{S}$, the convergence of both sequences taking place in L^1. Then $\hat{h}_k \to g$, uniformly. In the equation $\mathcal{F}\, f_k \hat{h}_k = (2\pi)^{-n} \mathcal{F} f_k * \mathcal{F}\, \hat{h}_k = \mathcal{F} f_k * Rh_k$, $f_k \hat{h}_k$ converges, in L^1, to fg, $\mathcal{F} f_k$ converges uniformly to $\mathcal{F} f$, and Rh_k converges in L^1 to $(2\pi)^{-n}\hat{g}$. Therefore both sides converge uniformly to the desired equation.

Exercise: Find the convolution of $\dfrac{1}{a^2 + x^2}$ and $\dfrac{1}{b^2 + x^2}$, where $a > 0, b > 0$.