J. Korean Math. Soc. **37** (2000), No. 3, pp. 371–389

LOCALIZATION AND MULTIPLICATION OF DISTRIBUTIONS

IAN RICHARDS AND HEEKYUNG K. YOUN

ABSTRACT. Working within classical distribution theory, we develop notions of multiplication and convolution for tempered distributions which are general enough to encompass the classical cases – such as pointwise multiplication of continuous functions or the convolution of L^1 functions – which most textbook treatments of distribution theory leave out. Pains are taken to develop a theory which satisfies the commutative and associative laws.

0. Introduction

In this note we have attempted to present, in a compact and usable form, the theory of localization and multiplication for tempered distributions. There is a substantial research literature on this problem ([cf. the bibliography]), and it can be fairly said that the essential problems have been solved. However, many of these solutions – and even the definitions which give rise to them – are quite intricate.

What are the problems? Well, consider multiplication. At the text book level, the product gT is only defined when g is a C^∞ function. This gives an extremely lopsided theory of multiplication – allowing complete freedom for T, while placing severe restrictions on g. For many purposes in analysis one needs symmetrical definitions. Thus the product of two continuous functions $g(x)h(x)$ is well defined, whether g is C^∞ or not. The "good" qualities of h (being continuous) balance the "bad" qualities of g (not being C^∞).

It would be nice if any two distributions were multiplicable, but a little thought suggests that certain multiplications – e.g. the square of the Dirac delta function – make no sense within distribution theory.

Received March 25, 1999.
2000 Mathematics Subject Classification: 46F30.
Key words and phrases: tempered distributions, localization, multiplication, Fourier transform, partial integration, convolution.

Our objective, then, is to give a definition of multiplication which is general enough to cover the important cases. We would also like to prove commutative and associative laws for this multiplication. The commutative law turns out to be trivial. However, the associative law is not. The following example illustrates the difficulties. Consider

$$(1/x) \cdot x \cdot \delta(x).$$

If we group the term as $[(1/x) \cdot x] \cdot \delta(x)$ we get $1 \cdot \delta(x) = \delta(x)$. But if we group the terms as $(1/x) \cdot [x \cdot \delta(x)]$ we get $(1/x) \cdot 0 = 0$. Clearly $\delta(x) \neq 0$. So the associative law appears to fail.

The way our of this difficulty is to insist that the product of three distributions must exist as a "whole" or "global" 3-fold operation, and not merely as a concatenation of 2-fold operations. Then it turns out that the associative law does hold, and the existence of the product $R(x) \cdot S(x) \cdot T(x)$ guarantees the existence of $S(x) \cdot T(x)$, subject to the obvious side condition that $R(x)$ is not identically zero. In fact, and perhaps surprisingly, $R(x) \not\equiv 0$ is the only extra hypothesis that is necessary.

[A similar situation occurs with Fubini's theorem. the "global" definition of the product is analogous to the global definition of the double integral $\iint_{R^2} f(x, y)\mathrm{d}(\text{area})$, whereas the iterated product corresponds to the iterated integral $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)dydx$. It is the global integral whose existence guarantees Fubini's theorem.]

We remark in passing that there is a theory of "new generalized functions" (or super distributions), due to Colombeau ([1984, 1985]), Rosinger ([1978, 1987]) and others, in which multiplication is universally defined. However, this theory is not symmetrical under the Fourier transform, and it leaves out convolution. In fact, no universal theory which covers multiplication, the Fourier transform, *and* convolution can exist. To see this, the previous counterexample – $(1/x) \cdot x \cdot \delta(x)$ – suffices. If we consider only multiplication, we can allow $x \cdot \delta(x) \neq 0$; indeed that must happen to preserve the associative law. But the Fourier transform of $x \cdot \delta(x)$ is just (d/dx) Const; and surely the calculus law that the derivative of a constant is zero is one that we should wish to preserve.

Thus we choose here to work within the traditional theory of distributions, with a view towards completing the application of that theory to functional analysis. Our aim is that all of the traditional constructs of functional analysis – e.g. the multiplication of continuous functions or the convolution of L^1 functions – should be included in distribution theory. Nonlinear theories such as Colombeau's lie outside of our scope.

Outline of the sections: In Section 1 we set the notations. Section 2 deals with *localization*, i.e., the restriction of a distribution $T(u, v)$ to a linear subspace $U = \{v = 0\}$ of $\mathbb{R}^q$. This forms the basis for the definition of multiplication, which is given in Section 3. That section includes a variety of nontrivial examples and also contains the statement and proof of the associative law. Section 4 gives the proofs of some technical results about localization, which were stated without proof in Section 2. Finally, Section 5 contains some brief remarks about the parallel problem of convolution.

It seems worthwhile to list the distinctive features of our approach, even though this involves notions to be defined later. These features are: the fact that in dealing with convolution, we are able to get by with the classical textbook case of it (cf. Section 1); the fact that localization is defined globally, not locally, which is what makes the associative law possible (cf. Section 2 and 3); and finally the use of Lemmas A-C in Section 4, which produce a considerable simplification in the proofs.

All approaches to the problem of distribution multiplication seem to have certain features in common: 1. They use tensor products, and 2. they are complicated. Ours is no exception. However, we believe that the particular mix of operations presented here reduces the complexity to a substantial degree. Distribution operations have to be tolerably simple, or else no one will use them. It is in that spirit that we offer this communication.

1. Notations and Standing Conventions

We consider only tempered distributions on $\mathbf{R}^q$. As usual, $\mathcal{S}(\mathbf{R}^q)$ denotes the space of "open support" test functions, i.e., C^∞ functions which are rapidly decreasing together with all of their derivatives.

[Recall that a function is called rapidly decreasing if it decreases faster that any negative power of $|x|$ as $x \to \infty$. It is called slowly increasing if it is dominated by some positive power of $|x|$ as $|x| \to \infty$.]

The standard textbook definition of convolution for distributions (most commonly given for $\mathcal{D}'(\mathbf{R}^q)$ extends immediately to tempered distributions (i.e., to $\mathcal{S}'(\mathbf{R}^q)$ if we make the following obvious modifications. We say that a tempered distribution S is rapidly decreasing if, for every test function $\varphi \in \mathcal{S}(\mathbf{R}^q)$,

$$(S * \varphi)(x) = \langle S(y), \varphi(x-y)\rangle_y$$

is also a test function, and the mapping is continuous from $\mathcal{S}$ to $\mathcal{S}$. Then, as is usual, we define, for any tempered distribution T:

$$\langle T * S, \varphi\rangle = \langle T, S(-x) * \varphi\rangle.$$

We write $\mathbf{R}^q = U + V$ to mean that U and V are complementary linear subspaces of $\mathbf{R}^q$, i.e., that U and V together span $\mathbf{R}^q$, and $U \cap V = \{0\}$. Together with U and V we have vector variables u, v such that the (u, v) span $\mathbf{R}^q, U = \{v = 0\}$, and $V = \{u = 0\}$. We do not require U and V to be orthogonal. However, the measures du and dv on U and V should be normalized so that $dudv = dx$ =the Lebesgue measure on $\mathbf{R}^q$.

The *tensor product* $S(u)T(v)$ (often written $S(u) \otimes T(v)$) of two distributions $S \in \mathcal{S}'(U)$ and $T \in \mathcal{S}'(V)$ is defined by

$$\langle S(u)T(v), \Psi(u,v)\rangle = \langle S(u), \langle T(v), \Psi(u,v)\rangle_v\rangle_u.$$

We recall that the tensor product always exists, that it is also a tempered distribution, and that this operation is commutative.

REMARK. The tensor product always exists, essentially because it involves *independent* variables u and v. By contrast, product of the form $S(x)T(x)$, involving a repetition of the same variable, may or may not exist. For example, the square of the Dirac delta function, $\delta(x)^2$, has no distribution theoretic meaning. Determining with reasonable generality the cases where $S(x)T(x)$ does exist is the main point of this paper.

2. Localization

Let $\mathbf{R}^q = U + V$ as above. We ask, under what conditions can a tempered distribution $T(u,v)$ be localized to the subspace U? That is, when does it make sense to set $v = 0$ and speak of $T(u,0)$? Since distributions are not defined pointwise, this is not always possible. As a first step, we take an arbitrary test function $\varphi(u) \in \mathcal{S}(U)$, and then form the tensor product $\varphi(-u)\delta(v)$ $(= \varphi(-u) \otimes \delta(v))$, which is a rapidly decreasing tempered distribution. Then we take the convolution $T_\varphi(u,v) = T(u,v) * [\varphi(-u)\delta(v)]$, which always exists as a distribution. It is the structure of $T_\varphi(u,v)$ which determines whether or not $T(u,v)$ is localizable to U.

Definition. Let $\mathbf{R}^q = U + V$. A tempered distribution $T(u,v)$ is *localizable* to U if, for all $\varphi(u) \in \mathcal{S}(U)$, the distribution

$$T_\varphi(u,v) = T(u,v) * [\varphi(-u)\delta(v)]$$

is a slowly increasing continuous function. In that case we define the distribution $T(u,0)$ on U by

$$\langle T(u,0), \varphi(u)\rangle = T_\varphi(0,0).$$

Remarks. (1) To see the motivation for this definition, we consider some extreme cases.

If we replace $\varphi(-u)\delta(v)$ by a test function $\Psi(-u,-v)$ (equivalent to setting $U = \mathbf{R}^q$), then the convolution $T(u,v) * \Psi(-u,-v)$ is always continuous and slowly increasing; its value at $(0,0)$ is $\langle T, \Psi\rangle$. At the opposite extreme: if we replace $\varphi(-u)\delta(v)$ by $\delta(u)\delta(v)$ (setting $U = \{0\}$), then $T(u,v) * \delta(u)\delta(v)$ is just $T(u,v)$. Obviously the convolution $T(u,v) * \varphi(-u)\delta(v)$ smooths out $T(u,v)$ somewhat, but to an intermediate degree. The smoothing is in the u-direction, whereas some sort of continuity in the v-direction must already be present.

(2) Suppose that $T(u,v)$ itself is a continuous slowly increasing function. Then the distribution theoretic convolution can be replaced by

an ordinary integral:

$$\begin{aligned}T_\varphi(u,v) &= T(u,v) * [\varphi(-u)\delta(v)]\\ &= \int_U \int_V T(a,b)\varphi(a-u)\delta(v-b)\,db\,da\\ &= \int_U T(u+a,v)\varphi(a)\,da.\end{aligned}$$

Then setting $(u,v)=(0,0)$, we obtain $\langle T(u,0),\varphi(u)\rangle = \int_U T(a,0)\varphi(a)da$, just as we would expect.

(3) The restriction of $T(u,v)$ to U depends only on U, not on the complementary subspace V. Thus let V and V' be subspaces complementary to U. Let the corresponding coordinates be (u,v) and (u',v') respectively. Then $v=v'$ everywhere, and $u=u'$ on the subspace U. Hence with $T_\varphi(u,v)$ as above, the corresponding T_φ for the decomposition $\mathbf{R}^q = U+V'$ is simply $T_\varphi(u',v)$.

(4) In our definition, we require $T_\varphi(u,v)$ to be continuous at all points (u,v) and not merely at $(0,0)$. This is a global criterion: it means that if $T(u,v)$ is localizable to U, then so is any translate $T(u-u_0, v-v_0)$. It is precisely this feature which allows the associative law to hold. (See also the proof of Theorem 2 in Section 4.)

EXAMPLES. 1. Let $\mathbf{R}^q=\mathbf{R}^1$. Then the delta function $\delta(x)$ cannot be localized to the 0-dimensional subspace $\{0\}$ (since $\delta(x)$ is not a continuous function).

2. Let $\mathbf{R}^q=\mathbf{R}^2$ with the usual xy-coordinates. Then the tensor product $\delta(x)1(y)$ can be localized to the x-axis (trivially, since $1(y)$ is constant); but $\delta(x)1(y)$ cannot be localized to the y-axis.

Less trivial examples will be given below.

We now state three theorems concerning localization.

THEOREM 1 (Continuity). *The functional $T(u,0)$, when it exists, is a distribution. That is, if $\varphi_n \to 0$ in $\mathcal{S}(U)$, then $\langle T(u,0),\varphi_n(u)\rangle \to 0$ in $\mathcal{C}$.*

THEOREM 2 ("Fubini Theorem"). *Let $\mathbf{R}^q = X_1+X_2+X_3$ be a decomposition of $\mathbf{R}^q$ into three complementary subspaces. Let*

$T(x_1, x_2, x_3)$ be localizable to X_1. Then $T(x_1, x_2, x_3)$ is localizable to $X_1 + X_2$, and the localization $S(x_1, x_2) = T(x_1, x_2, 0)$ is further localizable to X_1, and finally $S(x_1, 0) = T(x_1, 0, 0)$.

THEOREM 3 (Variable Constants Theorem). *Let $\mathbf{R}^q$ be decomposed into three complementary subspaces X_1, X_2 and X_3 as above. Let $T(X_1) \not\equiv 0$ be in $\mathcal{S}'(X_1)$ and let $S(x_2, x_3)$ be in $\mathcal{S}'(X_2 + X_3)$. Then $S(x_2, x_3)$is localizable to X_2 if and only if $T(x_1)S(x_2, x_3)$ is localizable to $X_1 + X_2$. When that happens*

$$T(x_1)[S(x_2, x_3)|_{x_3=0}] = (T(x_1)S(x_2, x_3))|_{x_3=0}.$$

[As noted in the Introduction, the proofs of Theorems 1-3 are given in Section 4.]

3. Multiplication

To define the product $S(x)T(x)$ of two tempered distributions in $\mathcal{S}'(\mathbf{R}^q)$, we first form the cartesian product $\mathbf{R}^{2q} = \mathbf{R}^q * \mathbf{R}^q$ of $\mathbf{R}^q$ with itself. For convenience, let x denote the first q variables and y the second set of q variables, so that (x, y) represents an arbitrary point in $\mathbf{R}^{2q}$. Then the tensor product $S(x)T(y)$ (or $S(x) \otimes T(y)$) always exists, no matter what the distributions S and T might be.

DEFINITION. Two tempered distributions $S(x)$ and $T(x)$ in $\mathcal{S}'(\mathbf{R}^q)$ are *multiplicable* if the tensor product $S(x)T(y)$ is localizable to the subspace $U = \{y = x\}$. When that happens, we define the *product* $S(x)T(x)$ to be that localization.

REMARK. (1) Of course the product may not exist, because the localization does not always exist.

(2) On the subspace $U = \{y = x\}$, it makes no difference whether we use the variable x or y or $\frac{(x+y)}{2}$.

(3) For the complementary space V, it is useful to take $V = \{x + y = 0\}$. Then we use the variables $u = \frac{(x+y)}{2}, v = y - x$, so that $dudv = dxdy$. The tensor product $S(x)T(y)$ becomes

$$S\left[u - \frac{v}{2}\right] T\left[u + \frac{v}{2}\right],$$

and the product $S(u)T(u)$ (if it exists) is just the localization of this tensor product to $\{v = 0\}$. This formulation is useful in applications.

EXAMPLES. We now give what are probably the most significant examples for applications. Of these, only the first is considered in most standard textbook treatments of distributions. The proofs are routine, although a bit tedious, and we omit them. For details (albeit based on a somewhat different definition) see Richards-Youn ([1990]).

1. Let $S(x) = g(x)$ be a C^∞ function which is slowly increasing together with all of its derivatives, and let $T(x)$ be an arbitrary tempered distribution. Then the product $g(x)T(x)$, as defined above, exists. Furthermore, this definition coincides with the traditional one: $\langle gT, \varphi\rangle = \langle T, g\varphi\rangle$ for any test function $\varphi \in \mathcal{S}(\mathbf{R}^q)$.

2. Let $f(x)$ and $g(x)$ be slowly increasing continuous functions. Then the product $f(x)g(x)$, as defined above, exists and equals the usual product. [Even this simple case is not included in the standard textbook treatments of multiplication for distributions.]

3. In example 2 above, the function $g(x)$ can be replaced by a slowly increasing complex measure $\mu(x)$, i.e., a measure for which $\int_{|x|\leq r} |d\mu(x)|$ is a slowly increasing function of r.

4. Let $f \in L^p(\mathbf{R}^q)$ and $g \in L^r(\mathbf{R}^q)$ with $p^{-1}+r^{-1} = t^{-1} \leqq 1$. Then, by a well known extension of Hölder's inequality, $f(x)g(x) \in L^t(\mathbf{R}^q)$ and $||fg||_t \leqq ||f||_p||g||_r$. Does the product $f(x)g(x)$ exist within distribution theory? By the traditional definition, no – it makes no sense. But this case is also covered by the more general definition above.

We turn now to n-fold products. As stated in the Introduction, these must be defined by a single n-ary operation, and not as a concatenation of binary operations. This time we pass from $\mathbf{R}^q$ to $\mathbf{R}^{nq}$, and let $x_1, \cdots, x_n$ denote the q-dimensional vectors corresponding to each of the components $\mathbf{R}^q$ in the cartesian product $\mathbf{R}^{nq} = \mathbf{R}^q \times \cdots \times \mathbf{R}^q$ (n times).

DEFINITION. The tempered distributions $T_1(x), \cdots, T_n(x)$ in $\mathcal{S}'$ $(\mathbf{R}^q)$ are *multiplicable* if the tensor product (on $\mathbf{R}^{nq}$)

$$T_1(x_1)T_2(x_2)\cdots T_n(x_n)$$

is localizable to the subspace

$$U = \{x_1 = x_2 = \cdots = x_n\}.$$

When that happens, we define the product $T_1(x)\cdots T_n(x)$, with $x = x_1 = \cdots = x_n$, to be that localization.

PROPOSITION (Commutative Law). *Let σ be any permutation of the integers $1, \cdots, n$. Then the product $T_1(x)\cdots T_n(x)$ exists if and only if the product $T_{\sigma(1)}\cdots T_{\sigma(n)}$ exists, and when they exist they are the same.*

Proof. Simply observe that the tensor product is commutative, and that our definition is symmetric in the variables x_i. □

Now we come to our main theorem.

THEOREM 4 (Associative Law). *Let $T_1(x), \cdots, T_n(x)$ be multiplicable, and suppose that none of the $T_i(x)$ is identically zero. Then, for any $r < n$, $T_1(x), \cdots, T_r(x)$ are multiplicable; the $(n-r+1)$-fold product below also exists, and we have:*

$$T_1 \cdots T_n = (T_1 \cdots T_r) \cdot T_{r+1} \cdots T_n.$$

REMARK. By combining Theorem 4 with the commutative law above, it is easy to obtain any desired associativity relation, *provided* that the n-fold product exists. Thus for example,

$$T_1 \cdots T_n = (T_1 \cdots T_r) \cdot (T_{r+1} \cdots T_n).$$

Proof. The proof is based on Theorems 2 and 3 from Section 2. In applying these theorems, it is important to remember that the spaces X_i are not required to be orthogonal.

That $T_1, \cdots, T_n$ are multiplicable means $T_1(x_1)\cdots T_n(x_n)$ is localizable to $\{x_1 = \cdots = x_n\}$. By the "Fubini Theorem", it follows that $T_1(x_1)\cdots T_n(x_n)$ is localizable to the larger subspace $\{x_1 = \cdots = x_r,\ x_i \text{ arbitrary for } i > r\}$. Now we observe that (unlike the ordinary product) the tensor product of nonzero distributions is always nonzero. Hence, by the "Variable Constants Theorem", $T_1(x_1)\cdots T_r(x_r)$ is localizable to $\{x_1 = \cdots = x_r,\ x_i = 0 \text{ for } i > r\}$ in $\mathbf{R}^{rq}$. Thus we have the existence of $T_1 \cdots T_r$.

We now show that $(T_1 \cdots T_r), T_{r+1}, \cdots, T_n$ are multiplicable, and that their product is what it ought to be. Again using the Variable Constants Theorem, we see that

$$(*) \qquad (T_1 \cdots T_r)(x) T_{r+1}(x_{r+1}) \cdots T_n(x_n),$$

where $x = x_1 = x_2 = \cdots = x_r$, is the localization of $T_1(x_1) \cdots T_n(x_n)$ to the subspace

$$\{x_1 = \cdots = x_r,\ x_i \text{ arbitrary for } i > r\} \text{ in } \mathbf{R}^{nq}.$$

Now by applying the Fubini Theorem one more time, the localization of (*) to $\{x = x_{r+1} = \cdots = x_n\}$ exists and coincides with the localization of $T_1(x_1) \cdots T_n(x_n)$, to $\{x_1 = \cdots = x_n\}$.

By definition, the localization of (*) is $(T_1 \cdots T_r)(x)\, T_{r+1}(x) \cdots T_n(x)$, whereas the localization of $T_1(x_1) \cdots T_n(x_n)$ is $T_1(x) \cdots T_n(x)$. □

4. Proof of Theorems 1–3

We now give the proofs which were postponed in Section 2.

Recall that a subset K of $\mathcal{S}$ is called *bounded* if for each of the seminorms $\mathcal{N}_{\alpha,N}$ for $\mathcal{S}$ (cf. the proof below), the set of values $\{\mathcal{N}_{\alpha,N}(\varphi) : \varphi \in K\}$ is bounded.

LEMMA A. *Let $\{\varphi_n\}$ be a bounded family of functions in $\mathcal{S}$. (The set of indices n need not be countable.) Then there exists a function $\theta > 0$ in $\mathcal{S}$ such that, if we write $\tau_n = \frac{\varphi_n}{\theta}$, then $\tau_n \in \mathcal{S}$ for all n, and the family $\{\tau_n\}$ is bounded.*

Proof. For the sake of completeness, we sketch the construction. The topology on $\mathcal{S}$ is determined by the seminorms

$$\mathcal{N}_{\alpha,N}(\varphi) = \|(1+|x|)^N \varphi^{(\alpha)}(x)\|_\infty,$$

or equivalently

$$\mathcal{N}_N(\varphi) = \sum_{|\alpha| \leqq N} \|(1+|x|)^N \varphi^{(\alpha)}(x)\|_\infty.$$

Let

$$\mathcal{M}_N = \sup_n \mathcal{N}_N(\varphi_n).$$

Since $\{\varphi_n\}$ is bounded in $\mathcal{S}$, $\mathcal{M}_N < \infty$ for $N = 0, 1, 2, \cdots$.

Now to construct θ. The function $\theta(x)$ will have the form $r^{-\sigma(r)}$, $r = |x|$, where $\sigma(r)$ is to be defined below.

Let $r_0 = 1$, and let the sequence of positive reals $\{r_i\}$ be defined by the conditions

$$\begin{aligned} r_{i+1} &\geqq 2r_i, \\ r_i &\geqq \mathcal{M}_{2i}. \end{aligned}$$

Let $\sigma(r)$, $r \geqq 0$, be a monotone nondecreasing C^∞ function such that $\sigma(0) = 0$ and $\sigma(r_i) = i$ for all i. The spacing of the r_i guarantees that, by standard C^∞ patching procedures, we can construct $\sigma(r)$ so that

$$|\sigma^{(\alpha)}(r)| \leqq \mathrm{Const}_\alpha \cdot r^{-|\alpha|} \text{ for all } \alpha \neq 0.$$

Now, as noted above, we let

$$\theta(x) = |x|^{-\sigma(|x|)}.$$

Clearly $\theta \in \mathcal{S}$. We next show that $\tau_n = \frac{\varphi_n}{\theta} \in \mathcal{S}$ for all n. Define

$$\mathcal{N}^\theta_{\alpha,N,n} = \left\| (1+|x|)^N \left(\frac{\varphi_n}{\theta}\right)^{(\alpha)} \right\|_\infty.$$

We need to show that $\mathcal{N}^\theta_{\alpha,N,n} < \infty$ for all α, N, n. We will do the case $\alpha = 0$ and leave the computation of the derivatives to the reader. [Note however that division by θ is multiplication by $\frac{1}{\theta} = r^{\sigma(r)}$, so of course one takes the derivative of the product, not a quotient. It turns out that the $\alpha = 0$ term, which we compute below, is the dominant one. That is $(\frac{\partial}{\partial r})^{|\alpha|}(\frac{1}{\theta(r)})$ is dominated by $(\frac{1}{\theta(r)})$ itself.]

To show that $\mathcal{N}^\theta_{0,N,n} < \infty$. Take $i = N + 2$ in the sequence $\{r_i\}$ defined above. On the compact disk $\{|x| \leqq r_{N+2}\}$ the functions $(1 + |x|)^N(\frac{\varphi_n}{\theta})$ are uniformly bounded. For $|x| \geqq r_{N+2}$, $|x|$ must lie in one of the intervals $r_i \leqq |x| \leqq r_{i+1}$, $i \geqq N+2$. On that interval, $\sigma(r) \leqq i+1$, so $\frac{1}{\theta(r)} = r^{\sigma(r)} \leqq r^{i+1}$. On the other hand, by definition,

$$r_i \geqq \mathcal{M}_{2i},$$
$$\mathcal{M}_{2i} \geqq (1+|x|)^{2i}|\varphi_n(x)|,$$

so that since $|x| \geqq r_i$ and $N+i+1 \leqq 2i-1$,

$$(1+|x|)^N \theta^{-1}(x)|\varphi_n(x)| \leqq (1+|x|)^{N+i+1}|\varphi_n(x)| \leqq \frac{\mathcal{M}_{2i}}{(1+|x|)} \leqq 1,$$

on the interval $r_i \leqq |x| \leqq r_{i+1}$. This, together with the corresponding calculations for $\alpha \neq 0$, shows that $\frac{\varphi_n}{\theta} \in \mathcal{S}$.

Since the above estimates are uniform in n, it also follows that the family $\{\frac{\varphi_n}{\theta}\}$ is bounded in $\mathcal{S}$ □

By almost identical arguments one can prove:

LEMMA B. *Let $\varphi_n \in \mathcal{S}$ and $\varphi_n \to 0$ in $\mathcal{S}$. Then there exists a function $\theta > 0$ in $\mathcal{S}$ such that*

$$\tau_n = \frac{\varphi_n}{\theta} \in \mathcal{S} \text{ and } \tau_n \to 0 \text{ in } \mathcal{S}.$$

LEMMA C. *Let $\mathbf{R}^q = U + V$ as above. Then the function $\theta(x)$ above can be replaced by a function of the form $\theta_1(u)\theta_2(v)$, $\theta_1 \in \mathcal{S}(U)$, $\theta_2 \in \mathcal{S}(V)$. Thus instead of $\varphi_n(x) = \theta(x)\tau_n(x)$, we have $\varphi_n(u,v) = \theta_1(u)\theta_2(v)\tau_n(u,v)$.*

Finally we observe that the space $\mathcal{S}$ is symmetrical under the Fourier transform. [In fact, that is our main reason for using $\mathcal{S}$.] Hence the products in the above lemmas can be replaced by convolutions. That is what will happen in the proofs which follow.

We turn now to the proofs of Theorems 1-3. We state Theorem 1 in a slightly stronger form which will be useful.

THEOREM 1*. *Let* $\mathbf{R}^q = X_1 + X_2 + X_3$ *be a decomposition of* $\mathbf{R}^q$ *into three complementary subspaces, and let* $U = X_1 + X_2$. *Let* $T(x_1, x_2, x_3)$ *be a tempered distribution on* $\mathbf{R}^q$. *Suppose that for all test functions in* $\mathcal{S}(U)$ *of the form* $\varphi_1(x_1)\varphi_2(x_2), T_{\varphi_1\varphi_2}(x_1, x_2, x_3)$ *is a slowly increasing continuous function. Then the same thing holds for all test functions* $\Psi(x_1, x_2) \in \mathcal{S}(U)$, *i.e.,* T *is localizable to* U. *The localization is a distribution. Finally, this distribution is uniquely determined by its action on the special test functions* $\varphi_1(x_1)\varphi_2(x_2)$.

Proof. Let $\Psi_n(x_1, x_2)$ be any sequence of test functions approaching zero in $\mathcal{S}(U)$. We use Lemmas A-C above. Furthermore, since the space $\mathcal{S}(U)$ is symmetric under the Fourier transform, the products in Lemmas A-C can be replaced by convolutions. Thus there exist test functions $\theta_1(x_1)$ and $\theta_2(x_2)$ and a sequence $\tau_n(x_1, x_2) \in \mathcal{S}(U)$ such that

$$\Psi_n(x_1, x_2) = [\theta_1(x_1)\theta_2(x_2)] * \tau_n(x_1, x_2)$$

and

$$\tau_n(x_1, x_2) \to 0 \text{ in } \mathcal{S}(U).$$

Then

$$\begin{aligned}
&T_{\Psi_n}(x_1, x_2, x_3)\\
&= T(x_1, x_2, x_3) * \Psi_n(-x_1, -x_2)\delta(x_3)\\
&= T(x_1, x_2, x_3) * [[\theta_1(-x_1)\theta_2(-x_2) * \tau_n(-x_1, -x_2)]\delta(x_3)]\\
&= T(x_1, x_2, x_3) * [(\theta_1(-x_1)\theta_2(-x_2)\delta(x_3)] * [\tau_n(-x_1, -x_2)\delta(x_3)],
\end{aligned}$$

since the delta function is an identity under convolution. Furthermore, except for T, all of the terms in the triple convolution above are rapidly decreasing, so the associative law for convolution holds. Thus we arrive at

$$T_{\Psi_n}(x_1, x_2, x_3) = T_{\theta_1\theta_2}(x_1, x_2, x_3) * \tau_n(-x_1, -x_2)\delta(x_3).$$

By hypothesis, $T_{\theta_1\theta_2}$ is already a slowly increasing continuous function. Since $\tau_n(-x_1, -x_2)\delta(x_3)$ is rapidly decreasing, T_{Ψ_n} is also continuous and slowly increasing. Thus T is localizable to U. To show that $\langle T(x_1, x_2, 0), \Psi_n(x_1, x_2)\rangle \to 0$:

$$\begin{aligned}
\langle T(x_1, x_2, 0), \Psi_n\rangle &= T_{\Psi_n}(0, 0, 0)\\
&= [T_{\theta_1\theta_2}(x_1, x_2, x_3) * \tau_n(-x_1, -x_2)\delta(x_3)](0, 0, 0).
\end{aligned}$$

Since we already know that $T_{\theta_1\theta_2}$ is continuous and slowly increasing, and since $\tau_n \to 0$ in $\mathcal{S}(U)$, it follows that $[(T_{\theta_1\theta_2}) * \tau_n\delta](0,0,0) \to 0$.

Finally, since the linear span of all products $\varphi_1(x_1)\varphi_2(x_2)$ is dense in $\mathcal{S}(U)$, the action of $T(x_1,x_2,0)$ on these products determines the distribution uniquely. □

Proof of Theorem 2 ("Fubini"). Recall that we have $\mathbf{R}^q = X_1+X_2+X_3$, with $T(x_1,x_2,x_3)$ localizable to X_1. We want to show, among other things, that this implies a localization to the larger space X_1+X_2. By Theorem 1*, in dealing with $\mathcal{S}(X_1+X_2)$, it suffices to consider only test functions of the form $\varphi_1(x_1)\varphi_2(x_2)$. Recall that by definition

$$T_{\varphi_1}(x_1,x_2,x_3) = T(x_1,x_2,x_3) * \varphi_1(-x_1)\delta(x_2)\delta(x_3), \tag{1}$$

$$T_{\varphi_1\varphi_2}(x_1,x_2,x_3) = T(x_1,x_2,x_3) * \varphi_1(-x_1)\varphi_2(-x_2)\delta(x_3). \tag{2}$$

Now since the delta function is an identity under convolution, (2) is equivalent to

$$T(x_1,x_2,x_3) * \varphi_1(-x_1)\delta(x_2)\delta(x_3) * \delta(x_1)\varphi_2(-x_2)\delta(x_3). \tag{3}$$

Except for T, all of the terms in this convolution are rapidly decreasing, so the associative law for convolution holds. The first two terms in (3) coincide with (1), and thus we have

$$T_{\varphi_1\varphi_2} = [T_{\varphi_1}]_{\varphi_2}. \tag{4}$$

Since T is localizable to X_1, T_{φ_1} is continuous and slowly increasing, and hence so is the convolution $[T_{\varphi_1}]_{\varphi_2}$ given by (3). Since $T_{\varphi_1\varphi_2} = [T_{\varphi_1}]_{\varphi_2}$, this implies that T is localizable to X_1+X_2.

REMARK. Here it is essential that, in the definition of localization, $T_\varphi(u,v)$ is continuous at all points (u,v) – and not just at $(0,0)$. For convolutions do not leave the origin invariant.

Now let $S(x_1,x_2) = T(x_1,x_2,0)$. We want to show that

$$S_{\varphi_1}(x_1,x_2) = T_{\varphi_1}(x_1,x_2,0). \tag{*}$$

[This is not self evident, since it involves a Fubini – like interchange of the restriction to X_1+X_2 and the $\varphi_1(x_1)$ – operation.] Apply both of the distributions in (*) to an arbitrary test function $\xi(x_1, x_2) \in \mathcal{S}(X_1 + X_2)$:

$$\begin{aligned}&\langle S_{\varphi_1}(x_1, x_2), \xi(x_1, x_2)\rangle \\ &= \langle S(x_1, x_2) * \varphi_1(-x_1)\delta(x_2), \xi(x_1, x_2)\rangle \\ &= \langle S(x_1, x_2), \varphi_1(x_1)\delta(x_2) * \xi(x_1, x_2)\rangle\end{aligned}$$

which by definition of the localization S is

(5)
$$\begin{aligned}&[T(x_1, x_2, x_3) * [\varphi_1(-x_1)\delta(x_2) * \xi(-x_1, -x_2)]\delta(x_3)](0,0,0) \\ &= [T(x_1, x_2, x_3) * [\varphi_1(-x_1)\delta(x_2)\delta(x_3) * \xi(-x_1, -x_2)\delta(x_3)]](0,0,0).\end{aligned}$$

On the other hand, $T_{\varphi_1}(x_1, x_2, x_3)$ can be viewed either as a slowly increasing continuous function or as a distribution. We have

(6)
$$\begin{aligned}&\langle T_{\varphi_1}(x_1, x_2, 0), \xi(x_1, x_2)\rangle \\ &= \langle [T(x_1, x_2, x_3) * \varphi_1(-x_1)\delta(x_2)\delta(x_3)](x_1, x_2, 0), \xi(x_1, x_2)\rangle \\ &= [[T(x_1, x_2, x_3) * \varphi_1(-x_1)\delta(x_2)\delta(x_3)] * \xi(-x_1, -x_2)\delta(x_3)](0,0,0).\end{aligned}$$

Now by the associativity of convolution (5) equals (6). This proves (*).

Once we have (*), everything is easy. Since T_{φ_1} is continuous and slowly increasing, so is S_{φ_1}. Hence S is localizable to X_1. Finally by (*), $S_{\varphi_1}(0,0) = T_{\varphi_1}(0,0,0)$, which means by definition that

$$\langle S(x_1, 0), \varphi_1(x_1)\rangle = \langle T(x_1, 0, 0), \varphi_1(x_1)\rangle. \qquad \square$$

Proof of Theorem 3 (Variable Constants). Again we have $\mathbf{R}^q = X_1 + X_2 + X_3$, and we are considering the tensor product $T(x_1)S(x_2, x_3)$ with $T \not\equiv 0$. Again it suffices to consider only test functions of the form $\varphi_1(x_1)\varphi_2(x_2)$. Since $T \not\equiv 0$, there exists some φ_1 with $\langle T, \varphi_1\rangle \neq 0$, which implies that $T_{\varphi_1} \neq 0$. Recall that by definition

(1) $$S_{\varphi_2}(x_2, x_3) = S(x_2, x_3) * \varphi_2(-x_2)\delta(x_3),$$

(2) $$(TS)_{\varphi_1\varphi_2}(x_1, x_2, x_3) = T(x_1)S(x_2, x_3) * \varphi_1(-x_1)\varphi_2(-x_2)\delta(x_3),$$

and since convolution involving independent variables commutes with the tensor product, (2) is equivalent to

$$(T(x_1) * \varphi_1(-x_1))(S(x_2, x_3) * \varphi_2(-x_2)\delta(x_3)). \tag{3}$$

In other words,

$$(TS)_{\varphi_1\varphi_2} = (T_{\varphi_1})(S_{\varphi_2}). \tag{4}$$

Now $T_{\varphi_1}(x_1)$ is always a continuous slowly increasing function (since $\varphi_1(x_1)$ involves the whole space X_1), and for certain $\varphi_1, T_{\varphi_1} \neq 0$. The formula (4) written out in full becomes

$$(TS)_{\varphi_1\varphi_2}(x_1, x_2, x_3) = T_{\varphi_1}(x_1)S_{\varphi_2}(x_2, x_3), \tag{5}$$

and now it is obvious that the continuity of $(TS)_{\varphi_1\varphi_2}$ is equivalent to the continuity of S_{φ_2}. Also $(TS)_{\varphi_1\varphi_2}(0,0,0) = T_{\varphi_1}(0)S_{\varphi_2}(0,0)$ which by definition means that

$$(T(x_1)S(x_2, x_3))|_{x_3=0} = T(x_1)(S(x_2, x_3)|_{x_3=0}). \qquad \square$$

5. Convolution

We now indicate briefly how these procedures lead to a general definition of convolution. [Of course we have used convolutions extensively in this paper – but only in the classical case where all but one of the distributions are rapidly decreasing.] The idea, essentially, is to take the Fourier transform of our definition for multiplication.

Let $\mathbf{R}^q = U + V$ as above. Let $\hat{U} = V^\perp$ and $\hat{V} = U^\perp$. Adjust the dual variables $\hat{u}, \hat{v}$ so that the inner product

$$(u, v) \cdot (\hat{u}, \hat{v}) = u \cdot \hat{u} + v \cdot \hat{v}.$$

Then the Fourier transform of the tensor product

$$[S(u)T(v)]^\wedge = \hat{S}(\hat{u})\hat{T}(\hat{v}).$$

DEFINITION. A tempered distribution $T(x)$ in $\mathcal{S}'(\mathbf{R}^q)$ is *integrable* if its Fourier transform $\hat{T}$ is a slowly increasing continuous function. Then we define the integral of T by

$$\int\int_{\mathbf{R}^q} T(x)dx = \hat{T}(0).$$

Corresponding to localization we have:

DEFINITION. A tempered distribution $T(u,v)$ is *partially integrable* over V if, for all $\varphi(u) \in \mathcal{S}(U)$, the product $\varphi(u)T(u,v)$ is integrable over $\mathbf{R}^q$. In that case we define the partial integral over V by

$$\left\langle \int_V T(u,v)dv, \varphi(u) \right\rangle = \int\int_{\mathbf{R}^q} \varphi(u)T(u,v)dvdu.$$

To define convolution on $\mathbf{R}^q$, we again form the cartesian product $\mathbf{R}^{2q} = \mathbf{R}^q * \mathbf{R}^q$, and then specify that:

DEFINITION. Two tempered distributions $S(x)$ and $T(x)$ in $\mathcal{S}'(\mathbf{R}^q)$ are *convolvable* if, for all $\varphi(x) \in \mathcal{S}(\mathbf{R}^q)$, the product

$$\varphi(x+y)S(x)T(y)$$

is integrable over $\mathbf{R}^{2q}$. In that case we define the convolution $S * T$ by

$$\langle (S*T)(x), \varphi(x)\rangle = \int\int_{\mathbf{R}^{2q}} \varphi(x+y)S(x)T(y)dydx.$$

Now it is easy to check that these definitions of partial integration/convolution are simply the duals under the Fourier transform of the definitions of localization/multiplication given above. For instance for convolution: The product $\hat{S}\hat{T}$ naturally led to the subspaces $\hat{U} = \{x = y\}$ and $\hat{V} = \{x + y = 0\}$, and to the variables $\hat{u} = \frac{(x+y)}{2}$, $\hat{v} = y - x$. Then if we define u, v so as to preserve inner products, we arrive at $u = x + y$, $v = \frac{(y-x)}{2}$. This explains the factor $\varphi(u) = \varphi(x+y)$ in the formulas above.

Since we have complete duality under the Fourier transform, we have the trivial but important

THEOREM. *Two distributions $S(x)$ and $T(x)$ in $S'(\mathbf{R}^q)$ are convolvable if and only if their Fourier transforms are multiplicable. In that case*

$$[S * T]\hat{} = \hat{S}\hat{T}.$$

REMARKS CONCERNING OTHER DEFINITIONS. Shiraishi ([1959]), Horváth ([1974]), and others have proposed a definition of convolution which is very similar in spirit to the above, but which is less general. Essentially, these authors define a distribution $T(x)$ to be *integrable* only if T has a certain explicit structure: T is required to be a finite sum of derivatives of bounded complex measures.

[Clearly if T has the Shiraishi/Horváth form, then $\hat{T}$ is continuous and slowly increasing – the converse is false.]

In fact, the Shiraishi/Horváth definition is extremely natural, and it suffices for all applications. More precisely, it suffices for all applications which involve only convolution. Duality under the Fourier transform is lost.

References

[1] Ambrose, W., *Products of distributions with values in distributions*, J. Reine Agnew. Math. **315** (1980), 73–91.

[2] Bredimas, A., *La differentiation d'ordre complexe et les produits canoniques et la convolution generalise*, C. R. Acad. Sc. Paris 283, **A** (1976), 1095–1098.

[3] Bremermann, H. J., *On analytic continuation, multiplication and Fourier transform of Schwartz distributions*, J. Math. Phys. 2 **2** (1961), 240–258.

[4] Colombeau, J. F., *New Generalized Functions and Multiplication of Distributions*, North-Holland Math. Studies 84, Amsterdam (1984).

[5] ———, *Elementary Introduction to New Generalized Functions*, North-Holland Math. Studies 113, Amsterdam (1985).

[6] Fisher, B., *The product of distributions*, Quart. J. Math. Oxford 22 **2** (1971), 291–298.

[7] ———, *Some results on the products of distributions*, Proc. Camb. Phil. Soc **73** (1973a), 317–325.

[8] ———, *Singular products of distributions*, Math. Ann. **203** (1973b), 103–116.

[9] Horváth, J., *Sur la convolution des distributions*, Bull Sc. Math. 2e Série 98 (1974), 183–192.

[10] Kaminski, A., *Convolution, product and Fourier transform of distributions*, Studia Math. **74** (1982), 83–96.

[11] König, H., *Multiplikation von Distributionen*, Math. Ann. **128** (1955), 420–452.

[12] Mikusinski, J., *Irregular operations on distributions*, Stud. Math. **20** (1961), 163–169.

[13] ———, *Criteria of the existence and of the associativity of the product of distributions*, Stud. Math. **21** (1962), 253–259.

[14] Oberguggenberger, M., *Products of distributions*, J. Reine Agnew. Math. **365** (1986), 1–11.

[15] Richards, I. and Youn, H. K., *On the general definition of convolution for distributions*, J. Korean Math. Soc. **17** (1980), no. 1, 13–37.

[16] ———, *On the general definition of convolution for several distributions*, J. Korean Math. Soc. **17** (1981), no. 2, 161–168.

[17] ———, *Theory of Distributions–a nontechnical introduction*, Cambridge Univ. Press, 1990.

[18] Rosinger, E. E., *A distribution multiplication theory*, Techinion's Preprint Series No. AMT-31, Haifa, Israel, October, 1974.

[19] ———, *Distributions and Nonlinear Partial Differential Equations*, Springer Lecture Notes 684, 1978.

[20] ———, *Generalized Solutions of Nonlinear Partial Differential Equations*, North-Holland Math. Studies 146, Amsterdam, 1987.

[21] Schwartz, L., *Sur l'impossibilite de la multiplication des distributions*, C. R. Acad. Sc. Paris 239, 1954, pp. 847–848.

[22] Shiraishi, R., *On the definition of convolution for distributions*, J. Sc. Hirosh. Univ. **A-23** (1959), 19–32.

[23] ———, *On the multiplicative products of distributions*, J. Sc. Hirosh. Univ. **28** (1964), 223–235.

[24] ———, *On the value of distributions at a point and the multiplicative products*, J. Sc. Hirosh. Univ. **31** (1967), 89–104.

[25] Walter, H. F., *Ueber die Multiplikation von distributionen in einem Folgenmodell*, Math. Ann. **189** (1970), 211–221.

Heekyung K. Youn
Department of Mathematics
University of St. Thomas
Mail #OSS 201
St. Paul, MN 55101
USA
E-mail: hkyoun@stthomas.edu

Ian Richards
University of Minnesota
Minneapolis, MN
USA