J. Korean Math. Soc.
Vol. 17, No. 1, 1980

ON THE GENERAL DEFINITION OF CONVOLUTION FOR DISTRIBUTIONS

BY HEEKYUNG KANG YOUN AND IAN RICHARDS

Introduction

The theory of distributions is, among other things, a language. With its aid, many of the operations of classical analysis can be performed universally. When applied to convolutions, however, the theory has a serious flaw: the convolution of two distributions is not defined unless one of the distributions has compact support. (We refer here to the definition as it is most commonly given.) Hence the theory leaves out important topics, such as L^1 algebras, where the compact support hypothesis is not satisfied.

This article attempts to remedy that defect, while at the same time staying close to standard treatments of the subject. It is our aim to show that, by making a fairly simple modification in the definition of convolution, we obtain a comprehensive notion given once and for all which covers most of the classical cases.

Previous work in this direction has been done by Chevalley ([2]; see also [4], p. 498), Shiraishi [17], and others [8, 11]. It seems that Shiraishi has given the best definition, and our development is based on his approach. Here and in [21] we have attempted a systematic development, tying together integrability and the operational calculus connected with it, convolution, associative laws, and some applications. There is one detail in which our development differs from others, and this may be worth mentioning. While all of the above mentioned approaches (including ours) use the same space of test functions, we use a different topology, the so called "*strict topology*". (A description of the test function space, $\mathscr{B}(R^k)$, and its topology is given in Section II.)

Now to return to our problem. Ideally, one would like to define the convolution for two *arbitrary* distributions. Unfortunately there does not seem to be any natural way to do this. All de finitions proposed so far, including ours, apply only in certain cases. Moreover there are counterexamples, involving *e. g.* the question of associativity for triple convolutions, which suggest

Received Jan. 30, 1980

that such limitations are inevitable. The standard definition of convolution for distributions requires all but one of the distributions to have compact support. But there are other cases, such as the convolution of measures, where quite different principles are invoked. Thus we have several different instances of convolution, all obviously related, yet each one requiring slightly different treatment. What is wanted is a single notion of *convolvability* that combines all of these cases, and that specifies at the outset which distributions are convolvable and which are not.

Our approach is to start with the well known fact that, formally speaking, the convolution $f*g$ acting on a test function $\varphi\in\mathcal{D}(R^k)$ is equal to the $2k$-fold integral

$$\int_{R^k}\varphi(x)\int_{R^k}f(x-y)g(y)\,dy\,dx$$
$$=\int_{R^k}\int_{R^k}\varphi(x+y)f(x)g(y)\,dx\,dy,$$

where the variables $x, y\in R^k$. The question is: what should this integral mean when $f\in\mathcal{D}'(R^k)$ and $g\in\mathcal{D}'(R^k)$ are distributions? In particular, for which pairs f, g is it defined? Looked at this way, the question reduces to defining $\int_{R^k}h(x)\,dx$, h being a distribution on R^k and k an integer (here k replaces the $2k$ above). Now, still proceding formally, $\int_{R^k}h(x)\,dx$ is equal to the "inner product" of the distribution h tested against the "test function" 1. Thus we are led to introduce a test function space which contains the constant function 1, namely:

$\mathcal{B}(R^k)$ = the space of C^∞ functions on R^k which are bounded together with all of their mixed partial derivatives. [In Section II we will give an appropriate topology on $\mathcal{B}(R^k)$ so that $\mathcal{D}(R^k)$ is dense in $\mathcal{B}(R^k)$ and hence its dual space $\mathcal{B}'(R^k)\subset\mathcal{D}'(R^k)$.]

Now we define a distribution $h\in\mathcal{D}'(R^k)$ to be "integrable" if it has an extension to a continuous linear functional on $\mathcal{B}(R^k)$ (i. e. if $h\in\mathcal{B}'(R^k)$). Of course we set $\int_{R^k}h(x)\,dx=\langle h, 1\rangle$. Finally we will define convolvability of two arbitrary distributions $f, g\in\mathcal{D}'(R^k)$ in terms of this notion of integrability: Briefly, f and g in $\mathcal{D}'(R^k)$ are called *convolvable* if the distribution $\varphi(x+y)f(x)g(y)\in\mathcal{D}'(R^{2k})$ is integrable for every test function $\varphi\in\mathcal{D}(R^k)$. Then $f*g$ is given by the formula

$$\langle f*g, \varphi\rangle=\int_{R^{2k}}\varphi(x+y)f(x)g(y)\,dx\,dy.$$

It is important to note that the expression $\varphi(x+y)f(x)g(y)$ always exists as a distribution on R^{2k}: for $f(x)g(y)$ is a tensor product with respect to the independent variables x, y, and $\varphi(x+y)$ is a C^∞ function, so that all of the multiplications indicated here can be carried out. Only the integrability of $\varphi(x+y)f(x)g(y)$ is in question, and this is decisive for the convolvability of f and g.

The space $\mathcal{B}'(R^k)$ of integrable distributions turns out to consist of all finite sums of derivatives of bounded complex measures: *i. e.* every $f\in \mathcal{B}'(R^k)$ has a representation as a finite sum $f=\sum_\alpha \mu_\alpha^{(\alpha)}$, where the μ_α are bounded complex measures. This is proved in an Appendix, which deals with structural questions. (We prefer to keep the core of our development, in Section II, independent of such questions.) Recall that we have defined the convolvability of two distributions f and g to mean integrability of $\varphi(x+y)$ $f(x)g(y)$ for every test function φ. Now $\varphi(x)\in\mathcal{D}(R^k)$ has compact support, but of course $\varphi(x+y)$ does not; its support is an "infinite strip" in the product space $R^k\times R^k$. Thus the convolvability of f and g depends on the behavior of $f(x)g(y)$ along such strips: briefly, the restriction of $f(x)g(y)$ to these "strips" must be *integrable.*

Our definition of convolvability can be looked at from a different viewpoint. In Section II we will introduce the notion of the "partial integral" of a distribution $f\in\mathcal{D}'(R^k)$ over a linear subspace $V\subset R^k$. Then it will turn out that $f\in\mathcal{D}'(R^k)$ and $g\in\mathcal{D}'(R^k)$ are *convolvable* if and only if the partial integral

$$(f*g)(x)=\int_{R^k} f(x-y)g(y)dy$$

exists (where the tensor product $f(x-y)g(y)$ is a distribution on R^{2k}, and the partial integral is taken over the subspace R^k corresponding to the y variable).

We obtain a Fubini-type theorem, which states that the partial integral can be iterated, and a "variable constants theorem". These will be used in a subsequent paper ([21]) to extend the convolution operation to three or more distributions.

Having outlined our approach, this may be a good place to mention one of the main difficulties in generalizing the definition of $f*g$. Recall that the standard definition requires one of the distributions, say g, to have compact support. This suggests that we use a sequence of g_n with compact support to approximate an arbitrary g. The extension to more general g would then be made via continuity. Unfortunately, however, under the distribution topology the convolution $f*g$ is not even a *closed* operator as a function of both variables, and hence is certainly not continuous (For a counterexample, see the end of Section III.). Of course in certain classical cases, *e. g.* in convolving functions in L^p with those in L^q, $p^{-1}+q^{-1}\geq 1$, there is another topology-namely the $L^p\times L^q$ topology-in terms of which the convolution is continuous. But this would require introducing a different topology for each special case; and that is precisely what we want to avoid.

A key property of the definition $\langle f*g, \varphi\rangle = \int_{R^{2k}} \varphi(x+y) f(x) g(y)\, dxdy$ given above is that it is symmetric in f and g; thus it automatically allows various growth and decay conditions on f and g to balance each other out. To complete this approach, we need to put a topology on the test function space $\mathcal{B}(R^k)$. The appropriate choice is the so called "strict topology". Loosely speaking, this is the topology of *uniform convergence on compact subsets, together with global uniform boundedness*: $\varphi_n \longrightarrow 0$ in $\mathcal{B}(R^k)$ means that, for all derivatives $\varphi_n^{(\alpha)}$ (with α fixed), $\varphi_n^{(\alpha)}(x) \longrightarrow 0$ uniformly on compact subsets as $n \longrightarrow \infty$, and there exists some constant M_α such that $|\varphi_n^{(\alpha)}(x)| \leq M_\alpha$ for all x, n. Actually, the strict topology is not first-countable, but for most purposes one can treat it as if it were. It has been studied by Buck, Herz, Rubel, Collins [1, 6, 13, 3] and others.

The introduction of a "new" (i. e. not completely standard) space $\mathcal{B}(R^k)$ may seem a heavy price to pay for defining convolution more generally. However, the resulting development is completely straightforward. Only a preliminary knowledge of the theory of distributions is needed to follow our demonstrations (we do not even use the closed graph theorem). Instead of topological vector space arguments, we use "mesa functions" (C^∞ functions with compact support which are identically 1 over a given interval) to localize the problems.

As for the applications, which will be discussed in Section III, our goal is to show that various classical cases of convolution fit into our framework. In addition to the cases mentioned above, we will find that the Hilbert transform and the generalized Hilbert transforms can be defined as convolutions between their kernels and L^p functions. We do not prove here hard analytic theorems like the boundedness of the Hilbert transform in L^p, because our purpose is merely to provide an abstract framework for studying such questions.

The reasons for seeking such generality are, of course, mainly linguistic. Thus consider the one dimensional Hilbert transform: formally it may be defined as a convolution, $T(f)=(1/x)*f$. But, under the classical definition, since the kernel $1/x$ is not L^1 and does not have compact support, $(1/x)*f$

is only meaningful when the function f is smooth. The classical method for avoiding this difficulty is to state the key theorem very carefully; thus it might read:

(1) *The operator* $T(f)=(1/x)*f$ *is well defined for functions* f *which are* C^∞ *with compact support*;

(2) *if* $1<p<\infty$, *the resulting operator is bounded in terms of the* L^p*-norm*;

(3) *therefore* T *has a unique extension to a bounded operator on* L^p.

However, as we shall see, it is easy to show that $(1/x)*f$ exists as a distribution. Then the above theorem can be stated:

For $1<p<\infty$, *if* $f\in L^p$, *then* $(1/x)*f\in L^p$, *and* $\|(1/x)*f\|_p\leqq C_p\|f\|_p$.

I. Standard preliminary results

We will establish some notations and definitions and state some basic facts from classical distribution theory. For more details, we refer the reader to Edwards [4] and Yosida [20].

Notations. Let $\varphi(x)$ be a complex valued C^∞ function on R^k and $\alpha=(\alpha_1, \ldots, \alpha_k)$ be a multi-index with $\alpha_i\geq 0$. Then $|\alpha|$ denotes $\alpha_1+\cdots+\alpha_k$ and

$$\varphi^{(\alpha)}(x)=(\partial^{|\alpha|}/\partial x_1^{\alpha_1}\ldots\partial x_k^{\alpha_k})\varphi(x),\quad x=(x_1,\ldots,x_k)\in R^k.$$

Let $\alpha=(\alpha_1,\ldots,\alpha_k)$ and $\beta=(\beta_1,\ldots,\beta_k)$. We say $\alpha\geq\beta$, if $\alpha_i\geq\beta_i$.

Let $x=(x_1,\ldots,x_k)\in R^k$. Then $|x|=\max(|x_1|,\ldots,|x_k|)$.

Let f be a continuous linear form on a topological vector space X. We will denote $f(x)$ by $\langle f, x\rangle$, $x\in X$.

The following test function spaces are standard and very well known except for $\mathscr{B}(R^k)$.

DEFINITIONS. We will work on real Euclidean space R^k and consider the following families of test functions:

1. 1. $\mathcal{D}(R^k)$ is the space of C^∞ functions with compact support on R^k, endowed with the following topology: A family of functions U in $\mathcal{D}(R^k)$ is open if and only if $U\cap\mathcal{D}_K$ is open in $\mathcal{D}_K$ for any compact subset K in R^k; here $\mathcal{D}_K$ consists of all C^∞ functions with support on K with the topology given by the family of semi-norms $\|\varphi\|_{\alpha,K}=\max\limits_{0\leqq\beta\leqq\alpha}\sup\limits_{x\in K}|\varphi^{(\beta)}(x)|$, $\varphi\in\mathcal{D}_K$ and α a multi-index.

1. 2. $\mathcal{E}(R^k)$ is the space of *all* C^∞ functions on R^k, with the topology given by the family of semi-norms $\|\varphi\|_{\alpha,K}$ defined in (1. 1) above (however here these semi-norms determine the topology directly; the much stronger

topology in (1. 1) is the direct limit of restrictions to the subspaces $\mathcal{D}_K$). [The dual spaces $\mathcal{D}'(R^k)$ and $\mathcal{E}'(R^k)$, defined below, are just the space of all Schwartz distributions and the subspace of distributions with compact support.]

We will consider $\mathcal{B}(R^k)$, which is not so well known, but will be essential for our development of convolutions between distributions.

1. 3. $\mathcal{B}(R^k)$ is the space of C^∞ functions on R^k which are bounded together with all of their mixed partial derivatives. We will define a topology on $\mathcal{B}(R^k)$ later, in Section II.

1. 4. We denote by $\mathcal{D}'(R^k)$, $\mathcal{E}'(R^k)$, and $\mathcal{B}(R^k)$, the spaces of distributions formed by taking the duals of $\mathcal{D}(R^k)$, $\mathcal{E}'(R^k)$ and $\mathcal{B}'(R^k)$, respectively.

It is evident that the spaces $\mathcal{D}(R^k)$, $\mathcal{E}(R^k)$, $\mathcal{B}(R^k)$, $\mathcal{D}'(R^k)$, $\mathcal{E}'(R^k)$, and $\mathcal{B}'(R^k)$ defined as above are tcpological vector spaces. Furthermore it will turn out that $\mathcal{D}(R^k)\subset\mathcal{B}(R^k)\subset\mathcal{E}(R^k)$ and that $\mathcal{E}'(R^k)\subset\mathcal{B}'(R^k)\subset\mathcal{D}'(R^k)$.

A notational remark. Very often we will consider spaces $\mathcal{D}(U)$, $\mathcal{E}(U)$, etc., where U is a linear subspace of R^k. We will not give any detailed description of these spaces, since they are self explanatory. There is only one point worth mentioning. The identification of ordinary fuctions $f(x)$ with distributions is via the inner product $\langle f(x), \varphi(x)\rangle=\int_{R^k} f(x)\ \ \varphi(x)\,dx$. When we break up R^k into the direct sum of two complementary subspaces, $R^k=U\oplus V$, then we want the Haar measures du, dv and dx to be normalized so that $dx=dudv$. We will always assume that this holds.

Let us now define the support of a distribution.

1. 5 Definition. For a distribution f, we define the *support* of f to be the smallest closed set F in R^k such that, for any $\varphi\in C^\infty$ which belongs to a proper test function space and is supported outside of F, $\langle f, \varphi\rangle=0$.

1. 6 Definition. Let $\varphi_n\in\mathcal{D}(R^k)$. Then we say that $\varphi_n\longrightarrow 0$ in $\mathcal{D}(R^k)$, if there is a compact set K in R^k such that the supports of the φ_n's are all in K, and $\varphi_n{}^{(\alpha)}\longrightarrow 0$ uniformly for any multi-index α.

The following theorems are well known. So we will skip the proofs, with a reference to Edwards [4] and Yosida [20].

1. 7 Theorem. *$f\in\mathcal{D}'(R^k)$ if and only if $\langle f, \varphi_n\rangle\longrightarrow 0$ whenever $\varphi_n\longrightarrow 0$ in $\mathcal{D}(R^k)$.*

Remark. The topology on $\mathcal{D}(R^k)$ is not first countable. That is why Theorem 1. 7 is interesting.

1.8 THEOREM. $\mathcal{E}'(R^k)$ *consists of distributions which have compact support.*

Finally we will define the tensor product of two distributions.

1.9 DEFINITION. Let U and V be complementary subspaces of R^k, (i.e., $U\oplus V=R^k$ and $U\cap V=\{0\}$) and let $f(u)$ and $g(v)$ be arbitrary distributions in $\mathcal{D}'(U)$ and $\mathcal{D}'(V)$, respectively. Then the *tensor product* $f(u)g(v)$ is that distribution on R^k given by $\langle f(u)g(v), \varphi(u,v)\rangle=\langle f(u), \langle g(v), \varphi(u,v)\rangle\rangle$ for any $\varphi(u,v)\in\mathcal{D}(U\oplus V)=\mathcal{D}(R^k)$.

REMARK. Since g is continuous and φ has compact support, $\theta(u)=\langle g(v), \varphi(u,v)\rangle$ is a C^∞ function with compact support. Hence the above is well defined.

1.10 THEOREM. *The tensor products is commutative*: $f(u)g(v)=g(v)f(u)$.

1.11 DEFINITION. Let U and V be complementary subspaces of R^k, and let $f(u)$ be a distribution on U (i.e., $f\in\mathcal{D}'(U)$). Then we define the *extension along* V, $f(u,v)$ of $f(u)$, to be the tensor product of $f(u)$ with the distribution $g(v)=1$ on V. (This corresponds, in the case of ordinary functions, to setting $f(u,v)=f(u)$.)

II. Convolution of two distributions in $\mathcal{D}'(R^k)$.

In this section we will define convolution between two distributions in a natural way. We will use the fact that $1\in\mathcal{B}(R^k)$, and if f and g are in $L^1(R^k)$ and $\varphi\in\mathcal{D}(R^k)$ then

$$\langle f*g, \varphi\rangle=\int_{R^k}\int_{R^k}f(x-y)g(y)\varphi(x)\ dydx$$
$$=\int_{R^k}\int_{R^k}f(x)g(y)\varphi(x+y)dxdy,$$

which we can formally describe as $\langle f(x)g(y)\varphi(x+y), 1\rangle$. That is, $\langle f*g, \varphi\rangle$ is $f(x)g(y)\varphi(x+y)$, as a functional on $\mathcal{B}(R^{2k})$, evaluated at 1. (Recall that $\mathcal{B}(R^k)$ denotes the family of C^∞ functions on R^k which are bounded together with all of their derivatives.)

In the following, we will give an appropriate topology on $\mathcal{B}$ so that $\mathcal{D}$ is dense in $\mathcal{B}$. Then we will define a distribution $f\in\mathcal{D}'$ to be "integrable" if it has an extension to a continuous linear functional on $\mathcal{B}$ (i.e. if $f\in\mathcal{B}'$). Finally we will define "convolvability" of two arbitrary distributions $f, g\in\mathcal{D}'$ in terms of this notion of integrability.

2.1 DEFINITION. The topology on $\mathcal{B}(R^k)$ is given by the family of semi-norms: for $\varphi\in\mathcal{B}(R^k)$,

$$\|\varphi\|_{\rho,\alpha}=\sup_{x\in R^k}|\rho(x)\varphi^{(\alpha)}(x)|,$$

where $\rho(x)$ is a continuous function on R^k which approaches zero as $|x|\to\infty$, and α is a multi-index.

This topology, sometimes known as "the strict topology", has been studied by various authors [1, 6, 13, 3].

2.2 DEFINITION. We denote by $\mathcal{B}'(R^k)$ the space of distributions formed by taking the dual of $\mathcal{B}(R^k)$.

REMARKS. $\varphi_n\longrightarrow 0$ in $\mathcal{B}(R^k)$ (i.e. $\|\varphi_n\|_{\rho,\alpha}\to 0$ as $n\to\infty$ for all semi-norms $\|\ \|_{\rho,\alpha}$) if and only if, for each multi-index α, the sequence $\varphi_n^{(\alpha)}$ is uniformly bounded on R^k and converges to zero uniformly on compact sets. Even though the space $\mathcal{B}(R^k)$ is not first countable, this description in terms of sequences is adequate for most purposes (*cf.* Proposition 2.3 below).

2.3 PROPOSITION. *The space $\mathcal{B}(R^k)$ coincides with the set of all finite sums of derivatives of bounded complex measures. Thus every $f\in\mathcal{B}'(R^k)$ has a representation as a finite sum $f=\sum_\alpha\mu_\alpha^{(\alpha)}$, where the μ_α are bounded complex measures.*

Furthermore, a distribution $f\in\mathcal{D}'(R^k)$ is continuous on $B(R^k)$ (i.e. $f\in\mathcal{B}'(R^k)$) if and only if, for every sequence $\varphi_n\in\mathcal{D}(R^k)$ (not $\mathcal{B}(R^k)$) : $\varphi_n\to 0$ in the topology of $\mathcal{B}(R^k)$ implies $\langle f,\varphi_n\rangle\longrightarrow 0$.

We will give the proof of this proposition in an Appendix. Because it is somewhat less elementary, Proposition 2.3 will not be used in the rest of the paper.

Now let us see that $\mathcal{D}(R^k)$ is dense in $\mathcal{B}(R^k)$ in the topology of $\mathcal{B}(R^k)$.

2.4 PROPOSITION. *The space $\mathcal{D}(R^k)$ of test functions is dense in $\mathcal{B}(R^k)$ in the topology of $\mathcal{B}(R^k)$.*

We note that Proposition 2.4 would fail if, instead of the strict topology, we had used the topology of uniform convergence on R^k.

Proof. Choose "mesa functions" $\theta_n\in\mathcal{D}(R^k)$ so that:

(1) $\theta_n(x)=\begin{cases}1 \text{ if } |x|\le n,\\ 0 \text{ if } |x|\ge n+1,\end{cases}\qquad (|x|=\max|x_i|)$;

(2) for each α, $|\theta_n^{(\alpha)}(x)|$ is bounded independent of n, x.

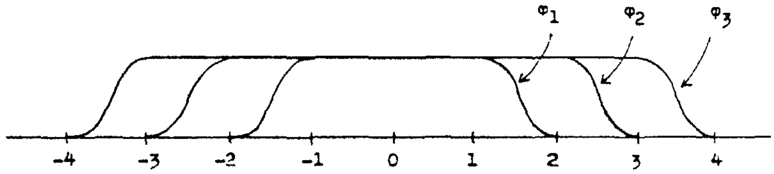

Figure 1. A sequence of test functions converging to 1 in the manner of Proposition 2.4.

(This is easy to achieve: for a 1-dimensional picture, see Figure 1; the k-dimensional case is handled by taking products of functions like those in the figure.)

Let $\varphi \in \mathcal{B}(R^k)$. Then since $\theta_n(x)\varphi(x) \in \mathcal{D}(R^k)$, it suffices to show that $\theta_n\varphi \to \varphi$ in $\mathcal{B}(R^k)$. But (1) implies, for any multi-index α, $(\theta_n\varphi)^{(\alpha)}(x) = \varphi^{(\alpha)}(x)$ for $|x| < n$, which in turn gives uniform convergence on compact sets. Also (2) together with the Leibniz formula for $(\theta_n\varphi)^{(\alpha)}$ gives us uniform boundedness of $(\theta_n\varphi)^{(\alpha)}$. Hence we have $\theta_n\varphi \to \varphi$ in $\mathcal{B}(R^k)$ with $\theta_n\varphi \in \mathcal{D}(R^k)$, which proves the proposition. Q. E. D.

It is evident that the topology of $\mathcal{D}(R^k)$ is stronger than the topology on $\mathcal{B}(R^k)$. Hence we have that every distribution in $\mathcal{B}'(R^k)$ corresponds to a unique distribution in $\mathcal{D}'(R^k)$. The same relation holds between $\mathcal{E}(R^k)$ and $\mathcal{B}(R^k)$. That is, $\mathcal{B}(R^k)$ is dense in $\mathcal{E}(R^k)$ in the topology of $\mathcal{E}(R^k)$, and the topology on $B(R^k)$ is stronger than the topology on $\mathcal{E}(R^k)$. Hence we have the following proposition.

2.5 Proposition. $\mathcal{E}'(R^k) \subseteq \mathcal{B}'(R^k) \subseteq \mathcal{D}'(R^k)$.

The next proposition is about the tensor product in $\mathcal{B}(R^k)$.

2.6 Proposition (Closure). (a) *Let U and V be complementary subspaces of R^k, so that $R^k = U \oplus V$. Let $f(u)$ and $g(v)$ be distributions in $\mathcal{B}'(U)$ and $\mathcal{B}'(V)$ respectively. Then the tensor product $f(u)g(v) \in \mathcal{B}'(R^k)$.*

(b) *$f \in B'(R^k)$, $\varphi \in B(R^k)$ implies $\varphi f \in \mathcal{B}'(R^k)$; furthermore, if $\psi \in \mathcal{B}(R^k)$, then $\langle \varphi f, \psi \rangle = \langle f, \varphi\psi \rangle$.*

(c) *$f \in \mathcal{B}'(R^k)$ implies $f^{(\alpha)} \in \mathcal{B}'(R^k)$ for each multi-index α; if $\psi \in \mathcal{B}(R^k)$, then $\langle \partial f/\partial x_1, \psi \rangle = \langle f, -\partial\psi/\partial x_1 \rangle$, and similarly for other partial derivatives.*

Proof. The proofs of (b) and (c) are almost trivial, but there is one logical point worth mentioning. Take (c): if the test function ψ were in $\mathcal{D}(R^k)$, then the identity $\langle \partial f/\partial x_1, \psi \rangle = \langle f, -\partial\psi/\partial x_1 \rangle$ would be true by definition. But since $\psi \in \mathcal{B}(R^k)$, we must approximate ψ by a sequence of functions $\psi_n \in \mathcal{D}(R^k)$ such that $\psi_n \to \psi$ in the topology of $\mathcal{B}(R^k)$. Then also $\partial\psi_n/\partial x_1 \to \partial\psi/\partial x_1$ in $\mathcal{B}(R^k)$, and we are done.

For (a), we could use the structure theorem (2.3) (proved in an Appendix). However we prefer to give a more conceptual proof.

LEMMA. *Let $\varphi\in\mathcal{B}(R^k)$. Then for each partial derivative, say $\partial/\partial x_1$, we have*

$$\partial\varphi/\partial x_1=\lim_{h\to 0}[\varphi(x_1+h, x_2, \dots, x_k)-\varphi(x_1, \dots x_k)]/h$$

in the topology of $\mathcal{B}(R^k)$.

Proof. Routine.

Now we go back to the definition of the tensor product (1.9).

$$\langle f(u)g(v),\ \varphi(u, v)\rangle=f'(u), \langle g(v), \varphi(u, v)\rangle\rangle.$$

This definition applies just as well to "test-functions" $\varphi\in\mathcal{B}(R^k)$ as to those in $\mathcal{D}(R^k)$ (here is where the Lemma is needed), and thus the tensor product $f(u)g(v)$ gives a continuous linear functional on $\mathcal{B}(R^k)$.

Now we define integrability:

2.7 DEFINITION. A distribution $f\in\mathcal{D}'$ is ***integrable*** if it can be extended to a continuous linear functional on $\mathcal{B}$ (i. e. $f\in\mathcal{B}'$). The integral $\int_{R^k}f(x)\,dx$ is defined to be $\langle f, 1\rangle$.

With the above definitions and propositions in hand, we are ready to define our main object.

2.8 DEFINITION. Two distributions f and g in $\mathcal{D}'(R^k)$ are *convolvable* if, for every test function $\varphi\in\mathcal{D}(R^k)$, the distribution $\varphi(x+y)f(x)g(y)$ $(\in\mathcal{D}'(R^{2k}))$ is integrable over R^{2k}, i. e. $\varphi(x+y)f(x)g(y)\in\mathcal{B}'(R^{2k})$. In that case we define the convolution $f*g$ by;

$$\langle f*g, \varphi\rangle=\langle\varphi(x+y)f(x)g(y), 1\rangle.$$

The following proposition will show that the convolution defined by Definition 2.8 is in fact a distribution in $\mathcal{D}'(R^k)$.

2.9 PROPOSITION. *If $\varphi(x+y)f(x)g(y)\in\mathcal{B}'(R^{2k})$ for every $\varphi\in\mathcal{D}(R^k)$, then $f*g$ defined by Definition 2.8 is continuous on $\mathcal{D}(R^k)$.*

Proof. Applying Theorem 1.7, let $\varphi_n\in\mathcal{D}(R^k)$ be such that $\varphi_n\to 0$ in $\mathcal{D}(R^k)$ and let K be a compact set containing the support of φ_n for every n. Choose a mesa function $\theta_k\in\mathcal{D}(R^k)$ with $\theta_K(x)=1$ for $x\in K$. Then $\varphi_n(x+y)=\varphi_n(x+y)\theta_K(x+y)$. Hence

$$\langle f*g, \varphi_n\rangle=\langle\varphi_n(x+y)f(x)g(y), 1\rangle$$

$$=\langle\theta_K(x+y)\varphi_n(x+y)f(x)g(y),1\rangle$$
$$=\langle\theta_K(x+y)f(x)g(y),\ \varphi_n(x+y)\rangle$$
$$\to\langle\theta_K(x+y)f(x)g(y),0\rangle=0 \text{ as } n\to\infty,$$

because $\theta_K(x+y)f(x)g(y)\in\mathcal{B}'(R^{2k})$ and $\varphi_n\to 0$ in $\mathcal{D}(R^k)$ implies $\varphi_n(x+y)\to 0$ in $\mathcal{B}(R^{2k})$.
Hence $f*g$ is continuous. Q. E. D.

Our last definition introduces the notion of partial integrability; this leads to an "operational calculus" on distribution space.

2. 10 DEFINITION. Let U and V be complementary subspaces of R^k (i. e. $R^k=U\oplus V$ and $U\cap V=\{0\}$). We say that a distribution $f\in\mathcal{D}'(R^k)$ is *partially integrable* over V if for each $\varphi(u)\in\mathcal{D}(U)$, the product $\varphi(u)f(u,v)$ is an integrable distribution on R^k. We define $\int_V f(u,v)\,dv$ to be the distribution on U given by

$$\varphi(u)\to\langle\varphi(u)f(u,v),1\rangle.$$

REMARK 1. By choosing a $\theta_K\in\mathcal{D}(U)$ as in the proof of Proposition 2. 9, we can easily verify that $\varphi(u)\to\langle\varphi(u)f(u,v),1\rangle$ is continuous on $\mathcal{D}(U)$. Another question is whether the existence of $\int_V f(u,v)\,dv$ is independent of U. But if U' is another subspace complementary to V, then the corresponding variables u and u' are related by a non-singular linear transformation. The resulting distributions on U and U' are transformed in the same manner, of course. A complete structural description of when $\int_V f(u,v)\,dv$ exists is given in an Appendix. We observe that f is "integrable over R^k" if and only if f is "integrable" in the sense of Definition 2. 7.

REMARK 2. It is natural to choose the Haar measure on U, V, and R^k so that the measure on R^k is the product measure corresponding to those on U and V. Actually, this assumption plays no part in the purely *distribution-theoretic* results below, because the notion of an abstract distribution (as a linear functional on a test-function space) does not involve Haar measure. It is in the transition *from functions to distributions*, via the formula $\langle f,\varphi\rangle=\int f(x)\varphi(x)\,dx$, that the Haar measure plays a role.

REMARK 3. Of course the partial integral of f over V may exist even when f is not integrable over R^k. A simple example is $f(u,v)=g(u)h(v)$, where g is any distribution in $\mathcal{D}'(U)$ and $h\in\mathcal{B}'(V)$.

2.11 LEMMA (Consistency). *Let the subspaces U and V be as above, and let μ be a finite complex measure defined on R^k. Then the "partial integral" difined in* (2.10) *corresponds to the ordinary partial integral of measure theory.*

Proof. For convenience, let us denote by ν_m and ν_d repectively the measure-theoretic and distribution-theoretic (Definition 2.10) entities defined formally by the partial integral $\int_V \mu(u, v)\,dv$. We have to show that the measure ν_m corresponds to the distribution ν_d in the usual way, i.e. that, for all test-functions $\varphi(u) \in \mathcal{D}(U)$, $\int_U \varphi(u)\,d\nu_m = \langle \nu_d, \varphi \rangle$.
By measure theory we have

$$\int_U \varphi(u)\,d\nu_m = \int_{U\oplus V} \varphi(u)\,d\mu(u, v),$$

where the $\varphi(u)$ in the right hand integral is the natural extension of $\varphi(u)$ from U to $R^k = U \oplus V$. From Definition 2.10 we have $\langle \nu_d, \varphi \rangle = \langle \varphi(u)\mu(u, v), 1 \rangle$, and since μ is a finite measure and φ is bounded, this coincides with $\int_{U\oplus V} \varphi(u)\,d\mu(u, v)$, as desired. Q. E. D.

The above definition of the partial integral enables us to describe $f*g$ in a very function-like way.

2.12 THEOREM. *Two distributions f and g on R^k are convolvable if and only if the following partial integral exists (and then that integral gives the convolution)*

$$(f*g)(x) = \int_{R^k} f(x-y)g(y)\,dy,$$

where the tensor product $f(x-y)g(y)$ is difined on R^{2k}.

Since non-singular linear changes of variables are permissible in distribution theory, it follows that tensor products such as $f(x-y)g(y)$ are well defined.

Proof. The partial integral $\int_{R^k} f(x-y)g(y)\,dy$ exists if and only if $\varphi(x) f(x-y)g(y) \in \mathcal{B}'(R^{2k})$ for all $\varphi(x) \in \mathcal{D}(R^k)$. But, since non-singular linear changes of variables map $\mathcal{B}'(R^k)$ into itself, $\varphi(x)f(x-y)g(y) \in \mathcal{B}'(R^{2k})$ is equivalent to $\varphi(x+y)f(x)g(y) \in \mathcal{B}'(R^{2k})$, which in turn implies that f and g are convolvable. Q. E. D.

The following extension lemma is based on the fact that if $\varphi(v) \in \mathcal{B}(V)$,

$X=U\oplus V$, then $\varphi(u, v)\in\mathcal{B}(X)$, where $\varphi(u, v)=\varphi(v)$ (i. e. $\varphi(u, v)$ is constant along all lines parallel to the subspace U). We note that the corresponding statement for $\mathcal{D}(V)$ is false.

2.13 EXTENSION LEMMA. *Let $X=U\oplus V$. Then the map defined by mapping each function $\varphi(v)$ on V into $\varphi(u, v)=\varphi(v)$ is continuous in the topology of $\mathcal{B}(X)$.*

Proof. Clear.

As a corollary of (2.13), we see that the following is well defined (it depends on U as well as V):

2.14 DEFINITION. Let $X=U\oplus V$, $f\in\mathcal{B}'(X)$. Then the *restriction* $f|\mathcal{B}(V)$ of f to $\mathcal{B}(V)$ is defined by applying f to the extensions $\varphi(u, v)=\varphi(v)$ described in the above lemma.

REMARKS. In other words, the restriction of f to $\mathcal{B}(V)$ is the dual map $\mathcal{B}'(X)\to\mathcal{B}'(V)$ to the map given in the lemma. This is equal to the partial integral $\int_U f(u, v)\,du$. We observe that if we considered the standard Schwartz spaces $\mathcal{D}$ and $\mathcal{D}'$, instead of $\mathcal{B}$ and $\mathcal{B}'$, then nothing like the extension or restriction lemmas would be true.

We will now construct "mesa-functions" which will play the role of identity and will be of great help in the rest of the section.

2.15 EXISTENCE LEMMA. *For any $\varphi(u, v)\in\mathcal{D}(U\oplus V)$ there are mesa functions $\theta(u)\in\mathcal{D}(U)$ and $\eta(v)\in\mathcal{D}(V)$ such that*

$$\varphi(u, v)=\theta(u)\varphi(u, v)=\theta(u)\eta(v)\varphi(u, v).$$

Proof. Let $a>0$ be such that $\varphi(u, v)=0$ for $|u|\geq a$ or $|v|\geq a$. Choose $\theta(u)\in\mathcal{D}(U)$, $\eta(v)\in\mathcal{D}(V)$ so that $\theta(u)=\eta(v)=1$ for $|u|\leq a$ and $|v|\leq a$. Then $\varphi(u, v)=\theta(u)\varphi(u, v)=\theta(u)\eta(v)\varphi(u, v)$. Q. E. D.

The next theorem, together with the one which follows, forms the main tool in extending the notion of convolution to three or more distributions. This is carried out in a sequel to the present paper [21].

2.16 THEOREM (Fubini's theorem). *Let U and V be complementary subspaces of R^k, and let V be the sum of two complementary subspaces S, T. Suppose that $\int_V f(u, v)\,dv$ exists. Then the iterated integral $\int_S\int_T f(u, s, t)\,dtds$ exists and equals $\int_V f(u, v)\,dv$.*

Proof. Recall the definition of the partial integral (2.10): Given a decomposition $X=Y\oplus Z$ of a finite dimensional real vector space X, and a distribution $f\in\mathcal{D}'(X)$, we say $\int_Z f(y,z)dz$ *exists if and only if* $\varphi(y)f(y,z)\in\mathcal{B}'(X)$ *for all* $\varphi(y)\in D(Y)$; *then* $\langle\int_Z f(y,z)dz,\ \varphi(y)\rangle=\langle\varphi(y)f(y,z),1\rangle$.

[In order to understand the minutae of this argument, it is important to recall that $\langle\varphi(y)f(y,z),1\rangle$ is defined only indirectly. Originally, $\varphi(y)f(y,z)$ is a distribution $\in\mathcal{D}'(X)$ (so that it acts on the subspace $\mathcal{D}(X)\subset\mathcal{B}(X)$), but it is continuous in terms of the $\mathcal{B}$-topology, and so it can be extended to $\mathcal{B}(X)$. Recall that the constant function $1\in\mathcal{B}(X)$, but $1\notin\mathcal{D}(X)$.]

First we will show that $\int_T f(u,s,t)dt$ exists; its existence implies that it defines a distribution $g(u,s)\in\mathcal{D}'(U\oplus S)$ (*cf.* Remark 1 after 2.10).

Let $\varphi(u,s)\in\mathcal{D}(U\oplus S)$ and choose a mesa function $\theta(u)\in\mathcal{D}(U)$ so that $\varphi(u,s)=\varphi(u,s)\theta(u)$ (*cf.* Existence lemma 2.15). Then

$$\varphi(u,s)f(u,s,t)=\varphi(u,s)\theta(u)f(u,s,t)\in\mathcal{B}'(R^k),$$

because $\varphi(u,s)\in\mathcal{B}(R^k)$ (clearly $\varphi(u,s)$ and its derivatives are bounded), and $\theta(u)f(u,s,t)\in\mathcal{B}'(R^k)$ since f is partially integrable over V. (Recall that $\varphi\in\mathcal{B}$, $F\in\mathcal{B}'$ implies $\varphi F\in\mathcal{B}'$, after (2.6).) Thus $g(u,s)=\int_T f(u,s,t)dt$ exists.

Now let us show that $g(u,s)$ is integrable over S.

LEMMA. *With the above assumptions on* f, *let* $g(u,s)=\int_T f(u,s,t)\ dt$, *and let* $\varphi(u)\in\mathcal{D}(U)$. *Then* $\varphi(u)g(u,s)\in\mathcal{B}'(U\oplus S)$ (*which means* $\int_S g(u,s)ds$ *exists*). *Furthermore, if* $\psi(u,s)\in\mathcal{B}(U\oplus S)$, *and* $\psi(u,s,t)$ *denotes the natural extension of* ψ *from* $U\oplus S$ *to* R^k, *then*:

$$\langle\varphi(u)g(u,s),\psi(u,s)\rangle=\langle\varphi(u)f(u,s,t),\psi(u,s,t)\rangle.$$

Proof of lemma. Begin by taking $\psi(u,s)\in\mathcal{D}(U\oplus S)$ (and not $\mathcal{B}(U\oplus S)$). Then:

$$\langle\varphi(u)g(u,s),\psi(u,s)\rangle=\langle g(u,s),\varphi(u)\psi(u,s)\rangle$$

$$=\langle\int_T f(u,s,t)dt,\ \varphi(u)\psi(u,s)\rangle=\langle\varphi(u)\psi(u,s)f(u,s,t),1_{R^k}\rangle.$$

So far we have only used the fact that $g(u,s)$ exists (i.e. that f is partially integrable over T). Now we recall that, by hypothesis, f is partially integrable over V, and hence $\varphi(u)f(u,s,t)\in\mathcal{B}'(R^k)$. Thus we may write:

$$\langle \varphi(u)g(u,s),\psi(u,s)\rangle=\langle \varphi(u)f(u,s,t),\psi(u,s,t)\rangle,$$

still assuming that $\psi\in\mathcal{D}(U\oplus S)$. By the Extension lemma 2. 13, the natural injection of $\mathcal{B}(U\oplus S)$ into $\mathcal{B}(R^k)$ given by setting $\psi(u,s,t)=\psi(u,s)$ is continuous in the $\mathcal{B}$-topology. Then, since $\varphi(u)f(u,s,t)\in\mathcal{B}'(R^k)$, it follows that the mapping

$$\psi \rightarrow \langle \varphi(u)g(u,s),\psi(u,s)\rangle=\langle \varphi(u)f(u,s,t),\psi(u,s,t)\rangle$$

is continuous in the topology of $\mathcal{B}(U\oplus S)$. Hence $\varphi(u)g(u,s)\in\mathcal{B}'(U\oplus S)$. This in turn implies that the preceding identity is valid for all $\psi\in\mathcal{B}(U\oplus S)$, proving the lemma.

To prove the theorem, we apply the lemma to the functions $\psi(u,s)=1_{U\oplus S}$ and $\psi(u,s,t)=1_{R^k}$ (and in addition we use the definition of partial integration several times). Thus:

$$\langle\int_S\int_T f(u,s,t)\,dt,\varphi(u)\rangle=\langle\int_S g(u,s)\,ds,\ \varphi(u)\rangle$$

$$=\langle\varphi(u)g(u,s),1_{U\oplus S}\rangle=\langle\varphi(u)f(u,s,t),1_{R^k}\rangle=\langle\int_V f(u,v)\,dv,\varphi(u)\rangle. \text{ Q. E. D.}$$

2. 17 THEOREM (Variable constants theorem). *Let* $R^k=U\oplus V$, $V=S\oplus T$. *Let* $f(u)\in\mathcal{D}'(U)$, $f\neq 0$, *and* $g(s,t)\in\mathcal{D}'(V)$. *Then*

$$H(u,s)=\int_T f(u)g(s,t)\,dt$$

exists as a distribution on $U\oplus S$ *if and only if*

$$G(s)=\int_T g(s,t)\,dt$$

exists as adistribution on S, *in which case* $H(u,s)=f(u)G(s)$.

Note. The products involved here are tensor products. This explains why the multiplier must have the form $f(u)$ and not $f(u,s)$. In defining $G(s)$, we have regarded T as a subspace of V, not R^k. It is easy to see what happens if we extend the distribution $g(s,t)$ from V to R^k, by forming its tensor product with the distribution $f(u)=1$ on U. There is a similar extension of $G(s)$ from S to $S\oplus U$. Then the above theorem, applied to $f(u)=1$, shows that these extensions correspond in the obvious way.

REMARK. The above theorem fails if $f=0$.

Proof of theorem. Let us assume that $H(u,s)$ is defined and show that $G(s)$ is defined. Thus we assume that $\psi(u,s)f(u)g(s,t)\in\mathcal{B}'(R^k)$ for all

$\psi(u,s)\in\mathcal{D}(U\oplus S)$, and we want to show that $\varphi(s)g(s,t)\in\mathcal{B}'(V)$ for all $\varphi(s)\in\mathcal{D}(S)$. Choose functions as follows:

(1) $\varepsilon(u)\in\mathcal{D}(U)$ such that $\langle f(u),\varepsilon(u)\rangle\neq 0$;

(2) a mesa function $\theta(u)\in\mathcal{D}(U)$ such that $\theta(u)\varepsilon(u)=\varepsilon(u)$.

Take any $\alpha(s,t)\in\mathcal{D}(V)$ and $\varphi(s)\in\mathcal{D}(S)$. Then since we have $\theta(u)\alpha(s,t)\in\mathcal{D}(R^k)$, we can compute:

$$\begin{aligned}&\langle\varepsilon(u)\varphi(s)f(u)g(s,t),\theta(u)\alpha(s,t)\rangle\\&=\langle\varepsilon(u)f(u),\ \theta(u)\rangle\ \ \langle\varphi(s)g(s,t),\ \alpha(s,t)\rangle\\&=f(u),\varepsilon(u)\rangle\ \ \langle\varphi(s)g(s,t),\alpha(s,t)\rangle\end{aligned}$$

by (1) and (2). Hence

$$\langle\varphi(s)g(s,t),\alpha(s,t)\rangle=\text{Const}\ \langle\varepsilon(u)\varphi(s)f(u)g(s,t),\ \theta(u)\alpha(s,t)\rangle,$$

where the constant $=\langle f(u),\varepsilon(u)\rangle^{-1}$.

Now $\varepsilon(u)\varphi(s)f(u)g(s,t)\in\mathcal{B}'(R^k)$, since $\varepsilon(u)\varphi(s)\in\mathcal{D}(U\oplus S)$ and $f(u)g(s,t)$ is partially integrable over T. This shows that the mapping $\alpha\to\langle\varphi(s)g(s,t),\alpha(s,t)\rangle$ is continuous in terms of the $\mathcal{B}$-topology, i. e. that $\varphi(s)g(s,t)\in\mathcal{B}'(V)$, as desired. Hence $G(s)$ exists.

Now for the converse: assume that $G(s)$ exists. Let $\psi(u,s)\in\mathcal{D}(U\oplus S)$. Choose $\theta(u)\in\mathcal{D}(U)$ and $\varphi(s)\in\mathcal{D}(S)$ so that $\psi(u,s)=\psi(u,s)\theta(u)\varphi(s)$, by Existence lemma 2. 15. Then using Proposition2. 6, the tensor product

$$\psi(u,s)f(u)g(s,t)=\psi(u,s)\theta(u)f(u)\varphi(s)g(s,t)\in\mathcal{B}'(R^k),$$

because $\psi(u,s)\in\mathcal{B}(R^k)$ (by the Extension lemma 2. 13), $\theta(u)f(u)\in\mathcal{B}'(U)$ (since it has compact support), and $\varphi(s)g(s,t)\in\mathcal{B}'(V)$ (since $G(s)$ exists). Hence $\int_T f(u)g(s,t)\,dt$ exists.

Proof of equality: Take $\psi(u)\in\mathcal{D}(U)$ and $\varphi(s)\in\mathcal{D}(S)$. We use the fact that finite linear combinations of products like $\psi(u)\varphi(s)$ are dense in $\mathcal{D}(U\oplus S)$. Then $\psi(u)f(u)\in\mathcal{B}'(U)$, $\varphi(s)g(s,t)\in\mathcal{B}'(V)$, and after Proposition 2. 6:

$$\begin{aligned}&\langle H(u,s),\psi(u)\varphi(s)\rangle=\langle\psi(u)\varphi(s)f(u)g(s,t),1_{R^k}\rangle\\&=\langle\psi(u)f(u),1_U\rangle\ \ \langle\varphi(s)g(s,t),1_V\rangle\\&=\langle f(u),\psi(u)\rangle\ \ \langle G(s),\varphi(s)\rangle\\&=\langle f(u)G(s),\ \psi(u)\varphi(s)\rangle.\end{aligned}$$

Hence $H(u,s)=f(u)G(s)$. Q. E. D.

III. Applications

It was our aim to define a notion of convolution between distributions which encompasses as many as possible of the classical cases.

3.1 THEOREM. *Two distributions $f, g \in \mathcal{D}'(R^k)$ are convolvable if any of the following holds:*

(a) *g has compact support;*

(b) *$R^k = R^1$ and f and g have support on $[0, \infty]$;*

(c) *$f = \mu$ and $g = \nu$ are complex measures, and the convolution of the positive measures $\int_{R^k} |\mu|(x-y)\, d|\nu|(y)$ is locally finite. In particular, if $f \in L^p$ and $g \in L^q$ and $1/p + 1/q \geq 1$, then $f*g$ exists.*

Furthermore in each of (a), (b) *and* (c), *the convolution coincides with the classical ones.*

REMARK. Theorem 3.1 could be extended to include many other examples, among them: tempered distributions convolved with rapidly decreasing distributions, exponentially bounded distributions convolved with exponentially decaying ones, distributions of L^p type, etc. In fact, all of these cases can be gotten from Theorem 3.1 combined with Theorem 3.4 below.

Proof of (3.1).

(a). Let $\varphi \in \mathcal{D}(R^k)$. Then $\varphi(x+y)f(x)g(y)$ has compact support on R^{2k} because g and φ have compact support on R^k. Hence by Proposition 2.5, $\varphi(x+y)f(x)g(y) \in \mathcal{B}'(R^{2k})$. So f and g are convolvable.

Now let us compute $\langle f*g, \varphi\rangle$ for $\varphi \in \mathcal{D}(R^k)$. Since g and φ have compact support on R^k, there is a mesa function $\psi(x, y) \in \mathcal{D}(R^{2k})$ such that $\varphi(x+y) g(y) = \varphi(x+y) g(y) \psi(x, y)$. Hence

$$\begin{aligned}\langle f*g, \varphi\rangle &= \langle \varphi(x+y)f(x)g(y), 1\rangle \\ &= \langle \psi(x, y)\varphi(x+y)f(x)g(y), 1\rangle = \langle f(x)g(y),\ \varphi(x+y)\psi(x, y)\rangle \\ &= \langle f(x), \langle g(y), \varphi(x+y)\psi(x, y)\rangle\rangle = \langle f(x), \langle g(y), \varphi(x+y)\rangle\rangle.\end{aligned}$$

Note that these equalities follow because $\varphi(x+y)\psi(x, y) \in \mathcal{D}(R^{2k})$, and $\varphi(x+y)\psi(x, y) = \varphi(x+y)$ on the support of g. So $\langle f*g, \varphi\rangle$ coincides with classical definition.

(b). For any $\varphi \in \mathcal{D}(R^1)$, $\varphi(x+y)f(x)g(y)$ has compact support on R^2 because $\varphi(x+y)f(x)g(y) \neq 0$ only if $0 \leq x, 0 \leq y$ and $x+y \leq M$ for some $M > 0$. Hence $\varphi(x+y)f(x)g(y) \in \mathcal{B}'(R^2)$ which implies that f and g are convolvable.

To compare $\langle f*g, \varphi\rangle$ with classical definition, let us again choose a mesa function $\psi(x, y) \in \mathcal{D}(R^2)$ so that $\varphi(x+y)f(x)g(y)=\varphi(x+y)f(x)g(y)\psi(x, y)$. Then

$$\langle f*g, \varphi\rangle = \langle \varphi(x+y)f(x)g(y), 1\rangle$$
$$= \langle f(x)g(y),\ \varphi(x+y)\psi(x, y)\rangle = \langle f(x), \langle g(y), \varphi(x+y)\psi(x, y)\rangle\rangle.$$

Now let us see that $\langle f(x), \langle g(y), \varphi(x+y)\psi(x, y)\rangle\rangle = \langle f(x), \langle g(y), \varphi(x+y)\rangle\rangle$. Let $\theta(x) = \langle g(y), \varphi(x+y)\psi(x, y)\rangle$ and $\eta(x) = \langle g(y), \varphi(x+y)\rangle$. Then over the support of f, $\theta(x)=\eta(x)$ because $\varphi(x+y)f(x)g(y)=\varphi(x+y)f(x)g(y)\psi(x, y)$. Hence $\langle f(x), \theta(x)\rangle = \langle f(x), \eta(x)\rangle$. So we have $\langle f*g, \varphi\rangle = \langle f(x), \langle g(y), \varphi(x+y)\rangle\rangle$.

(c). Let $\varphi \in \mathcal{D}(R^k)$. Then $\varphi(x+y)\mu(x)\nu(y)$ is a finite complex measure on R^{2k} because

$$\int_{R^{2k}} |\varphi(x+y)|\, d|\mu|(x)\, d|\nu|(y) = \int_{R^{2k}} |\varphi(x)|\, d|\mu|(x-y)\, d|\nu|(y)$$
$$= \int_{R^k}\int_{R^k} |\varphi(x)|\, d|\mu|(x-y)\, d|\nu|(y),$$

and $\theta(x) = \int_{R^k} |\mu|(x-y)\, d|\nu|(y)$ is a locally finite measure. Hence $\varphi(x+y)\mu(x)\nu(y) \in \mathcal{B}'(R^{2k})$ (*cf.* Lemma A. 1 in the Appendix) and

$$\langle \varphi(x+y)\, \mu(x)\nu(y), 1\rangle = \int_{R^{2k}} \varphi(x+y)\, d\mu(x)\, d\nu(y),$$

where the integral is an ordinary measure theoretic one. So $\mu*\nu$ is defined and coincides with classical definition. Q. E. D.

A nice application is that we can now formally express the Hilbert Transform as a convolution.

3. 2 THEOREM. *Let* $f(x)=1/x$, $x \in R^1$. *Then* $f(x)$ *is convolvable with* $L^p(R^1)$ *functions for* $1 \leq p < \infty$.

Proof. Let

$$f_1(x) = \begin{cases} 1/x & \text{if } 0<|x|<1 \\ 0 & \text{if } 1 \leq |x|, \end{cases} \quad \text{and} \quad f_2(x) = \begin{cases} 0 & \text{if } |x| \leq 1 \\ 1/x & \text{if } |x|>1. \end{cases}$$

Then $f_2(x) \in L^q$ for $1<q\leq\infty$ and hence by the previous Theorem (c), f_2 is convolvable with L^p functions for $1\leq p<\infty$.

Now $f_1(x)$ is a distribution because $f_1(x) = d/dx \log|x|$, $|x|\leq 1$; and $\log x$ being locally integrable, is a distribution. Since $f_1(x)$ has compact support, it is convolvable with any other distribution, particularly with L^p functions. Q. E. D.

We extend the application to singular integrals.

3.3 THEOREM. *Let $K(x)$ be a function on R^k such that $K(x)=\Omega(x)/|x|^k$, where $\Omega(x)$ is a bounded measurable function, $\Omega(\lambda x)=\Omega(x)$ for $\lambda>0$, and $\int_{|x|=1}\Omega(x)dx=0$. Then $K(x)$ is convolvable with L^p functions for $1\leq p<\infty$.*

Proof. The idea of the proof is the same as that of the proof for the Hilbert Transform. Let

$$K_1(x)=\begin{cases}K(x) & \text{if } 0<|x|<1,\\ 0 & \text{if } |x|\geq 1,\end{cases} \qquad \text{and} \qquad K_2(x)=\begin{cases}0 & \text{if } |x|<1,\\ K(x) & \text{if } |x|\geq 1.\end{cases}$$

Now we must show that $K_1(x)$ defines a distribution. Firstly, since $\int_{|x|=1}\Omega(x)dx=0$, the Cauchy principal value $\lim_{\delta\to 0}\int_{\delta<|x|<1}K(x)dx=0$. Hence for any $\varphi\in\mathcal{D}(R^k)$,

$$\lim_{\delta\to 0}\int_{\delta<|x|<1}K(x)\varphi(x)dx=\int_{0<|x|<1}K(x)(\varphi(x)-\varphi(0))dx.$$

Now $|\varphi(x)-\varphi(0)|\leq|x|\sup_z|\varphi'(z)|$, where φ' is the gradient of φ. So

$$\begin{aligned}|\langle K_1(x),\varphi(x)\rangle|&=\left|\int_{0<|x|<1}K(x)\ (\varphi(x)-\varphi(0))dx\right|\\ &\leq\left(\int_{0<|x|<1}|K(x)|\,|x|dx\right)\sup_z|\varphi'(z)|.\end{aligned}$$

Since $\int_{0<|x|<1}|K(x)|\,|x|dx=\int_{0<|x|<1}(|\Omega(x)|/|x|^k)\,|x|dx$
(which in polar coordinates $r=|x|$, $dx\cong r^{k-1}dr$)

$$=\int_0^1(r/r^k)(r^{k-1})dr\ \int_{|x|=1}|\Omega(x)|dx=\int_{|x|=1}|\Omega(x)|dx<\infty,$$

we have that $K_1(x)$ defines a distribution on $\mathcal{D}(R^k)$. But $K_1(x)$ has compact support, hence is convolvable with any other distribution.

$|K_2(x)|\leq$Constant $|x|^{-k}$ for $|x|>1$. Hence $K_2(x)\in L^q$ for $1<q\leq\infty$, and $K_2(x)$ is convolvable with L^p functions for $1\leq p<\infty$. Hence $K(x)$ is convolvable with L^p functions. Q. E. D.

3.4 THEOREM. *If two distributions f and g are convolvable, then for any multi-index α, the derivative $f^{(\alpha)}$ is convolvable with g; furthermore*

$$f^{(\alpha)}*g=(f*g)^{(\alpha)}=f*g^{(\alpha)}.$$

Proof. This result belongs more properly to the theory of n-fold convolutions (*cf.* [21]), but because of its importance we give a separate proof here.

There is no loss of generality in assuming that the differentiation operation is $\partial/\partial x_1$, and for convenience we shall write, for example, φ' in place of $\partial\varphi/\partial x_1$. We need to show that, for every test function $\varphi\in\mathcal{D}(R^k)$, $\varphi(x+y)f'(x)g(y)\in\mathcal{B}'(R^{2k})$, and that

$$\langle\varphi(x+y)f'(x)g(y),1\rangle=\langle-\varphi'(x+y)f(x)g(y),1\rangle.$$

Now we know, since f and g are convolvable, that $\varphi(x+y)f(x)g(y)\in\mathcal{B}'(R^{2k})$, and then Proposition 2.6 implies that

$$\begin{aligned}&[\varphi(x+y)f(x)g(y)]'\\&\quad=\varphi'(x+y)f(x)g(y)+\varphi(x+y)f'(x)g(y)\in\mathcal{B}'(R^{2k}).\end{aligned}$$

(Note that the "$'$" operation involves only the x variables.) But φ' is also a test function, and so $\varphi'(x+y)f(x)g(y)$ also lies in $\mathcal{B}'(R^{2k})$; hence so does $\varphi(x+y)f'(x)g(y)$.

To prove the required identity, we simply observe that by Proposition 2.6,

$$\langle[\varphi(x+y)f(x)g(y)]',\ 1\rangle=\langle\varphi(x+y)f(x)g(y),\ -1'\rangle=0$$

since $1'=0$. Q.E.D.

The following situation comes up in the study of time-series. It is related to the Poisson summation formula. [A further development would require bringing in Fourier transforms and products. We plan to study these questions in a subsequent paper.]

3.5 THEOREM. *Let $f\in\mathcal{B}'(R^1)$ be any integrable distribution on R^1, and let g be the "periodic δ-distribution" $g(x)=\sum_{n=-\infty}^{\infty}\delta(x-n)$. Then the convolution $f*g$ exists.*

Proof. By the structure theorem (A.4), proved in the Appendix, f is a finite sum of derivatives of bounded compex measures. Such measures have a measure-theoretic convolution with g, and thus Theorem 3.1c combined with Theorem 3.4 gives us what we want.

3.6 COUNTEREXAMPLE. Convolution is not a closed operator in terms of either the weak-$*$ or the strong topology on $\mathcal{D}'(R^k)\times\mathcal{D}'(R^k)$. Furthermore, even if one of the variables, say g, is held fixed, the convolution $f*g$ is not closed as a function of the other variable.

Thus let g be the distribution on R^1 correspondintg to the constant 1, and let $f_n(x)=\delta(x-n)$, $n=0,1,2,\ldots$ Then $f_n\to 0$ in either the weak or the strong topology, but $f_n*g=g$ for all n. [To see that $f_n\to 0$ in the strong topology, we use the fact (*cf.* Yosida [20]) that every bounded subset of $\mathcal{D}(R^k)$ consists of functions with uniformly bounded supports]

IV. Multiplication

Here we will give only the barest outline of a theory. Other approaches to this question have been given by König [9, 10] and Rosinger [12].

Our definition of multiplication is formally the dual, under the Fourier transform, of the definition of convolution given in Section II. The duality is only formal, however, since there is no good characterization of the space $\mathcal{B}'(R^k)^\wedge$ of Fourier transforms of distributions in $\mathcal{B}'(R^k)$. Probably our definition of convolvability will have to be modified in order to bring in Fourier transforms. This modification would at the same time be more general, since the space $\mathcal{B}'(R^k)^\wedge$ is smaller than the space we would use to replace it (see below). On the other hand, in treating convolutions by themselves, as in this paper, $\mathcal{B}'(R^k)$ seems to be the natural space to use, and additional generality is not always desirable when it entails at the same time additional complications.

In Section II we said that two distributions $f, g\in\mathcal{D}'(R^k)$ are convolvable if the tensor product $\varphi(x+y)f(x)g(y)\in\mathcal{B}'(R^{2k})$ for all $\varphi\in\mathcal{D}(R^k)$. Under the Fourier transform, tensor products go over into the corresponding tensor products, while ordinary multiplication corresponds to convolution. Furthermore, the inverse Fourier transform of the function $\varphi(x+y)$ on R^{2k} is the distribution $\hat{\varphi}[(x+y)/2]\delta(x-y)$. [Note: the $\varphi[-(x+y)/2]$ could be replaced by either $\hat{\varphi}(-x)$ or $\hat{\varphi}(-y)$, since $\delta(x-y)$ concentrates the mass along the subspace $\{x=y\}$.]

Thus we are led to:

4.1 Definition. Two distributions $f, g\in\mathcal{D}'(R^k)$ are *multiplicable* if, for every test function $\varphi\in\mathcal{D}(R^k)$, the distribution

$$[f(x)g(y)]*[\varphi(-x)\delta(x-y)]$$

can be represented by a continuous function. If so, we define $\langle fg, \varphi\rangle$ to be the value of that function at $x=0$.

Notes. Since $\varphi(-x)\delta(x-y)$ is a distribution with compact support, the convolution in the preceding definition exists. (Of course, it may or may not be a continuous function.) For the reasons mentioned above, $\varphi(-x)$ could be replaced by $\varphi(-y)$ or $\varphi[-(x+y)/2]$. In treating Fourier transforms, it would probably be better to alter the definition, replacing "continuous" by "continuous and slowly increasing" (and of course replacing $\mathcal{D}(R^k)$ by $\mathcal{S}(R^k)$). We plan to work out these ideas in a subsequent paper.

Appendix.

The objective of this section is to describe $\mathscr{B}'(R^k)$. (For a slightly different approach to this question, see Horváth [7].)

DEFINITION. $\mathscr{B}_0(R^k)$ is the space of functions in $\mathscr{B}(R^k)$ which vanish at infinity together with all of their mixed partial derivatives, endowed with topology given by the following family of semi-norms;

$$\|\varphi\|_m = \max_{0\le|\alpha|\le m} \sup_{x\in R^k} |\varphi^{(\alpha)}(x)|,$$

where m is a positive integer and α is a multi-index.

Immediately, we notice that $\mathscr{B}_0(R^k)$ is a Fréchet space and is dense in $\mathscr{B}(R^k)$ in the topology of $\mathscr{B}(R^k)$. Also if $\varphi\in\mathscr{B}_0(R^k)$, then $\varphi^{(\alpha)}\in\mathscr{B}_0(R^k)$ for any multi-index α.

A. 1 LEMMA. *Let $f=\mu^{(\alpha)}$, where μ is a bounded complex measure and α is a multi-index. Then $f\in\mathscr{B}'(R^k)$ and hence any finite sum $\sum\mu^{(\alpha)}$ belongs to $\mathscr{B}'(R^k)$.*

Proof. Let K_m be compact sets such that
(1) $|\mu|(K_m^c)<1/4^m$, $m=1, 2, \dots$
(2) $\bigcup_{m=0}^{\infty} K_m=R^k$ with $\{\phi\}=K_0\subset K_1\subset K_2\subset\cdots$
Construct a continuous function $\rho(x)$ as follows;

$$\min_{x\in K_{m+1}/K^m} |\rho(x)|>1/2^m, \quad m=0, 1, 2, \dots$$

and $o(x)\to 0$ as $|x|\to\infty$. Then for all $\varphi\in\mathscr{B}(R^k)$,

$$|\langle f, \varphi\rangle| = \left|\int_{R^k}\varphi(x)\,d\mu^{(\alpha)}(x)\right| \le \int_{R^k}|\varphi^{(\alpha)}(x)|\,d|\mu|(x)$$

$$\le \sum_{m=0}^{\infty}\left(\max_{x\in K_{m+1}/K_m}|\varphi^{(\alpha)}(x)|\int_{K_{m+1}/K_m} d|\mu|(x)\right)$$

$$\le \sum_{m=1}^{\infty}\max_{x\in K_{m+1}K_m}|\varphi^{(\alpha)}(x)|\cdot 1/4^m+\max_{x\in K_1}|\varphi^{(\alpha)}(x)|\int_{K_1} d|\mu|(x)$$

$$\le \sum_{m=1}^{\infty}\max_{x\in K_{m+1}/K_m}|\varphi^{(\alpha)}(x)\rho(x)|\cdot 1/2^m+\max_{x\in K_1}|\varphi^{(\alpha)}(x)|\,|\mu|(R^k)$$

$$\le \left(\max_{x\in R^k}|\varphi^{(\alpha)}(x)\rho(x)|\right)\left(\sum_{m=1}^{\infty}1/2^m+|\mu|(R^k)\right)=\text{Constant }\|\varphi\|_{\rho,\alpha}.$$

Hence $f\in\mathscr{B}'(R^k)$. Q. E. D.

A. 2 LEMMA. *Let X be a Banach space and $X^k=\{(x_1, \dots, x_k) : x_i \in X\}$ with $\|(x_1, \dots, x_k)\| = \sum_{i=1}^{k} \|x_i\|$.*

Then if f is a bounded linear functional on X^k, there are bounded linear functionals $f_1, \dots, f_k$ on X such that $f(x_1, \dots, x_k) = \sum_{i=1}^{k} f_i(x_i)$.

Proof. Define $f_i : X \to \mathbf{C}$ by $f_i(x) = f(\bar{x}_i)$, where $x \in X$, and $\bar{x}_i \in X^k$ has its i^{th} component equal to x and all other components zero. Then it is clear that f_i's are bounded linear functionals on X and $f(x_1, \dots, x_k) = \sum_{i=1}^{k} f_i(x_i)$. Q. E. D.

With the above lemma and some well known theorems from analysis, we can prove the following.

A. 3 LEMMA. *Let $f \in \mathscr{B}_0'(R^k)$. Then there are a finite number of bounded complex measures μ_α such that $f = \sum \mu_\alpha^{(\alpha)}$, where α's are multi-indices.*

Proof. Without loss of generality, we assume that $k=1$, because the proof for $k>1$ goes exactly the same way, with a slight modification in the notation.

Let $f \in \mathscr{B}_0'(R)$. Then there is a constant C and an integer $N>0$ such that

$$|\langle f, \varphi \rangle| \leq C \sum_{m=0}^{N} \|\varphi\|_m, \quad \text{for all } \varphi \in \mathscr{B}_0(R).$$

We build some Banach spaces:

Let $C_0=$ {the space of continuous functions on R vanishing at infinity, endowed with the sup norm}. Let $C_0^N = \{(\theta_0, \dots, \theta_N) : \theta_i \in C_0\}$ with $\|(\theta_0, \dots, \theta_N)\| = \sum_{i=0}^{N} \|\theta_i\|$, and take the subspace $A = \{(\varphi, \varphi', \dots, \varphi^{(N)}) : \varphi \in \mathscr{B}_0\}$. Define $\bar{f}$ on A by $\bar{f}(\varphi, \dots, \varphi^{(N)}) = \langle f, \varphi \rangle$. Then since $f \in \mathscr{B}_0'(R)$, $\bar{f}$ is a continuous linear functional on A. Hence by the Hahn-Banach Theorem, $\bar{f}$ can be extended to C_0^N. We will still call the extended function $\bar{f}$. Now by Lemma A. 2, there are bounded linear functionals $f_0, f_1, \dots_1, \dots, f_N$ on C_0 such that $\bar{f}(\varphi_0, \dots, \varphi_N) = \sum_{i=0}^{N} f_i(\varphi_i)$.

The Riesz-Representation Theorem gives $f_i(\varphi_i) = \int \varphi_i d\nu_i$ for some bounded complex measure ν_i. Hence

$$\langle f, \varphi \rangle = \bar{f}(\varphi, \varphi', \dots \varphi^{(N)}) = \sum_{i=0}^{N} f_i(\varphi^{(i)})$$
$$= \sum_{0i=}^{N} \int \varphi^{(i)} d\nu_i = \sum_{i=0}^{N} (-1)^i \int \varphi d\nu_i^{(i)}.$$

Hence $f = \sum_{i=0}^{N} \mu_i^{(i)}$, where $\mu_i = (-1)^i \nu_i$ with $|\mu_i|(R) = |\nu_i|(R) < \infty$. Q. E. D.

By putting Lemma A. 1 and Lemma A. 3 together with a remark that $\mathscr{B}_0'(R^k) \supseteq \mathscr{B}R^k)$, we have the following theorem.

A. 4 THEOREM. *The space $\mathcal{B}'(R^k)$ coincides with the set of all finite sums of derivatives of bounded complex measures. Furthermore $\mathcal{B}'(R^k)=\mathcal{B}_0'(R^k)$.*

Another thing that we want to know is that $f\in\mathcal{B}'(R^k)$ if and only if $\langle f, \varphi_n\rangle \to 0$ whenever $\varphi_n\in\mathcal{D}(R^k)$ and $\varphi_n\to 0$ in the topology of $\mathcal{B}(R^k)$. The proof is not difficult to see, because $\mathcal{B}_0'(R^k)=\mathcal{B}'(R^k)$, and $\mathcal{B}_0(R^k)$ is a Fréchet space in which $\mathcal{D}(R^k)$ is a dense subset.

Our last theorem characterizes convolvable pairs of distributions in terms of their structure. First recall that a distribution $f\in\mathcal{D}'(R^k)$ can be restricted to an open subset $\Omega\subseteq R^k$ by applying f only to those test functions whose supports are in Ω. A set is called "precompact" if its closure is compact. Finally we define a *diagonal strip* in $R^{2k}=R^k\oplus R^k$ to be a subset of the form

$$\Delta_K=\{(x,y)\in R^{2k} : (x+y)\in K\},$$

where K is a preassigned subset of R^k.

A. 5 THEOREM. *Two distributions $f, g\in\mathcal{D}'(R^k)$ are convolvable if and only if the following condition holds: For every precompact open subset $K\subseteq R^k$, the restriction of the tensor product $f(x)g(y)$ to the diagonal strip Δ_K is representable as a finite sum of derivatives of bounded complex measures.*

Proof. Recall that after Definition 2. 8, f and g are convolvable when, for every test function $\varphi\in\mathcal{D}(R^k)$, the product $\varphi(x+y)f(x)g(y)\in\mathcal{B}'(R^{2k})$. Now if $K\subseteq R^k$ contains the support of φ, then the support of $\varphi(x+y)$ lies in the strip Δ_K in R^{2k}. Hence the sufficiency follows from Lemma A. 1 above. For the necessity, given any precompact set $K\subseteq R^k$, choose a "mesa function" $\varphi\in\mathcal{D}(R^k)$ such that $\varphi(x)=1$ for $x\in K$. Then applying Theorem A. 4 to the product $\varphi(x+y)f(x)g(y)$ gives the desired result. Q. E. D.

EXAMPLES. Most of the theorems of Section III ("Applications") could also be proved using Theorem A. 5. Thus for Theorem 3. 1a : if f and g are distributions, g has compact support, and K is a precompact set in R^k, then the restriction of $f(x)g(y)$ to the strip Δ_K has compact support. By a well known representation theorem, all distributions with compact support are given by derivatives of measures, as required. Or take Theorem 3. 1 c : if μ and ν are complex measures such that $|\mu|$ and $|\nu|$ have a locally finite convolution, then the tensor product $\mu(x)\nu(y)$ gives a finite measure on each strip Δ_K for which K is precompact.

References

1. R. C. Buck, *Algebraic properties of classes of analytic functions* Seminars on Analytic Functions, vol. II, Princeton Univ. Press, Princeton, 1957, 175–188.
2. C. Chevalley, *Theory of Distributions*, Lectures at Columbia Univ., Columbia Univ. Press, New York, 1950.
3. H. S. Collins, *Strict topologies and applications*, to appear in Advances in Mathematics.
4. R. E. Edwards, *Functional Analysis*, Holt-Rinehart-Winston, New York, 1965.
5. I. M. Gelfand and G. Šilov, *Generalized Functions* (several vols.), Academic Press, New York, 1964.
6. C. S. Herz, *The spectral theory of bounded functions*, Trans. Amer. Math. Soc. **94** (1960), 181–232.
7. J. Horváth, *Topological Vector Spaces and Distributions*, vol. I, Addison-Wesley, Reading, Mass., 1966.
8. _______, *Sur la convolution des distributions*, Bull. Sc. Math., 2^e Série, **98** (1974), 183–192.
9. H. König, *Multiplikation von Distributionen*, Math. Ann. **128** (1955), 420–452.
10. _______, *Multiplikationstheorie der verallgemeinerten Functionen*, Bayrische Akad. Wiss. Math. Nat. Kl. Abh. (N. F.), **82** (1957), München.
11. B. Roider, *Sur la convolution des distributions* (to appear).
12. E. Rosinger, *A Distribution Multiplication Theory*, Technion's Preprint Series, No. AMT-31, Haifa, Israel, October, 1974.
13. L. A. Rubel, *Bounded convergence of analytic functions*, Bull. Amer. Math. Soc. **77** (1971), 13–24.
14. L. Schwartz, *Théorie des Distributions* (2 vols.), Hermann et Cie, Paris, 1950.
15. _______, *Produits tensoriels topologiques et espaces nucléaires*, *Séminaire*, Institut Henri Poincarè, Paris, 1953–54.
16. _______, *Théorie des distributions à valeurs vectcrielles* (I), Ann. Inst. Fourier Grenoble **7** (1957), 1–138.
17. R. Shiraishi, *On the definition of convolutions for distributions*, J. Sci. Hiroshima Univ., Ser. A-**23** (1959), 19–32.
18. F. Trèves, *Topological Vector Spaces, Distributions and Kernels*, Academic Press, New York, 1967.
19. K. Yoshinaga and H. Ogata, *On convolutions*, J. Hiroshima Univ. **22** (1958), 15–24.
20. K. Yosida, *Functional Analysis* (several ed.), Springer-Verlag, Berlin, 1965.
21. H. K. Youn and I. Richards, *On the general definition of convolution for several distributions* (to appear).
22. A. H. Zemanian, *Distribution Theory and Transform Analysis*, McGraw-Hill, New York, 1965.

University of Toronto,
University of Minnesota