

## GENERALIZED FUNCTIONS FOR APPLICATIONS

B. D. CRAVEN<sup>1</sup>

(Received 5 February 1984; revised 20 March 1984)

### Abstract

A simple rigorous approach is given to generalized functions, suitable for applications. Here, a generalized function is defined as a genuine function on a superset of the real line, so that multiplication is unrestricted and associative, and various manipulations retain their classical meanings. The superset is simply constructed, and does not require Robinson's nonstandard real line. The generalized functions go beyond the Schwartz distributions, enabling products and square roots of delta functions to be discussed.

### 1. Introduction

Many questions in physics and engineering, and in partial differential equations, require the Dirac delta function and its relations. The rigorous theory of Schwartz's distributions [9] is widely used, but is too complicated, in concept and detail, for many users. Moreover, multiplication of distributions is restricted, and may be nonassociative, as in Schwartz's famous example [9]:

$$\delta(t) = (t^{-1} \cdot t)\delta(t) \neq t^{-1}(t\delta(t)) = t^{-1} \cdot 0 = 0.$$

This paper presents a simple, though rigorous, approach to generalized functions, in which the delta function is a genuine function, whose domain and range are supersets of the real line. Consequently, expressions like  $\int_{\mathbf{R}} \delta(t)f(t) dt$  and their manipulation have meaning as ordinary integrals, and multiplication of generalized functions is unrestricted, and associative. Engineers have always manipulated generalized functions as if they were ordinary functions; with the present formulation, such manipulations become rigorous. The supersets of the

---

<sup>1</sup>Mathematics Department, University of Melbourne, Parkville, Victoria 3052 and Visiting Scientist at the National Research Institute for Mathematical Sciences of the CSIR, Pretoria, South Africa.  
© Copyright Australian Mathematical Society 1985, Serial-fee code 0334-2700/85

real line used here do *not* require the full panoply of A. Robinson's nonstandard analysis, using symbolic logic [5]; a simpler process suffices, which is constructible (and computable). The class of generalized functions obtained here is wide enough for physics and engineering applications, and is not restricted as in [11]. For the present purpose, there is no need (or advantage) to avoid Cauchy-Weierstrass limits by using infinitesimals. The superset here has long been known; what is novel here is its use for generalized functions, without involving other complications.

The present theory does not preserve the property, that every Schwartz distribution has a (weak) derivative—who needs to differentiate the Weierstrass nondifferentiable function? Most applications call for generalized functions which are smooth ( $C^r$  with  $r \geq 1$ ) except for delta functions or their relatives at some singular points (or curves or surfaces); and these are differentiable, with the present approach. Fractional powers of (suitable) delta functions are possible; these and some other formulas, relevant to quantum theory, which go beyond the Schwartz theory, receive simple treatment. Some generalized functions have more than one reasonable definition, with the values differing by infinitesimals, so that the ambiguity does not affect most applications. The delta distribution is unavoidably multivalued, if extended beyond its natural domain (of functions continuous at 0); the present approach makes the multivaluedness explicit, and makes use of it.

## 2. A superset of the real line

Consider one “infinitesimal” element  $\varepsilon$  adjoined to the real line  $\mathbf{R}$ ; for convenience denote the “infinite value”  $\varepsilon^{-1}$  also by  $\omega$ . In order to include quantities such as

$$u = 1 + 2\varepsilon - 3\varepsilon^2, \quad v = 4\varepsilon + 6\varepsilon^2 + 8\varepsilon^3 + \cdots; \quad y = 3\omega + 2; \quad (1)$$

the superset  $\mathbf{R}^s$  is defined as the field consisting of all formal series of the form

$$z = \sum_{j=-n}^{\infty} z_j \varepsilon^j \quad (\text{all } z_n \in \mathbf{R}, z_n \neq 0), \quad (2)$$

together with zero. Define the *index* of  $z$  of  $q(z) = -n$ , with  $q(0) = 0$ , and the *leading term* of  $z$  as  $z \cdot \varepsilon^{-q(z)}$  with  $z \cdot = z_{-q(z)}$ , and  $0 \cdot = 0$ . Thus, for example,  $q(u) = 0$ ;  $q(v) = -1$  (negative indicating infinitesimal), with leading term  $4\varepsilon$ ; and  $v = 4\varepsilon(1 + \frac{3}{2}\varepsilon + \frac{3}{2}\varepsilon^2 + \cdots)$ , expressing  $v$  as the product of its leading term and a term  $1 + \text{“infinitesimal”}$ ; and  $q(y) = +1$  (positive indicating “infinite value”), with leading term  $3\omega$ , and  $y = 3\omega(1 + \frac{2}{3}\varepsilon)$ . *Level*  $k$  of  $z$  means the term  $z_{-k}\varepsilon^{-k}$ .

The series (2) are formal series, with no convergence requirement. They are added termwise, and multiplied by the Cauchy product rule, namely

$$\left(\sum_{i=m}^{\infty} x_i \varepsilon^i\right) \left(\sum_{j=n}^{\infty} y_j \varepsilon^j\right) = \sum_{k=m+n}^{\infty} \left(\sum_{i+j=k} x_i y_j\right) \varepsilon^k. \tag{3}$$

Since the coefficient of  $\varepsilon^k$  is a finite sum, the right side of (3) is well defined, without any question of convergence arising. From (3), the multiplicative inverse of  $\sum_{i=0}^{\infty} a_i \varepsilon^i$  (where  $a_0 = 1$ ) is  $\sum_{j=0}^{\infty} b_j \varepsilon^j$ , where  $b_0 = 1$ , and the  $b_j$  for  $j \leq 1$  are determined recursively from  $b_j = -\sum_{i=1}^j a_i b_{j-i}$ . The field properties for  $\mathbf{R}^\varepsilon$  then follow directly from the field properties for power series with Cauchy products, since there are no zero divisors, and so need not be proved here. Thus  $\mathbf{R}^\varepsilon$  is an extension field of  $\mathbf{R}$ , which contains  $\varepsilon$ . The zero of  $\mathbf{R}^\varepsilon$  is  $0 = \sum 0 \cdot \varepsilon^j$ . Also  $\mathbf{R}^\varepsilon$  may be linearly ordered, using the leading terms, by

$$x \geq ky \text{ in } \mathbf{R}^\varepsilon \Leftrightarrow x - y \geq 0 \Leftrightarrow (x - y)_{\geq 0}. \tag{4}$$

Thus, setting  $z = x - y$ , exactly one of the three following cases occurs:

- (i)  $0 \neq z \geq 0 \Leftrightarrow z_{> 0}$ ;
  - (ii)  $z = 0 \Leftrightarrow z_{= 0}$ ;
  - (iii)  $0 \neq z \leq 0 \Leftrightarrow z_{< 0}$ .
- (5)

The symbol  $\varepsilon$  for an infinitesimal is chosen following physicists’s usage, and  $\mathbf{R}^\varepsilon$  is chosen to be typable. This field was first used in [1] and [8]; for historical comments, see Section 5. It is not obvious whether  $\mathbf{R}^\varepsilon$  is the smallest extension field of  $\mathbf{R}$  that contains  $\varepsilon$ ; this does not matter for its use. An “infinite value” must not be equated to  $+\infty$  (or to any infinite cardinal); the harmonic series  $\sum 1/n$  diverges to  $+\infty$ , not to any “infinite value” in  $\mathbf{R}^\varepsilon$ . The field  $\mathbf{R}^\varepsilon$  has a graded structure, with some analogy to Chevalley’s “graded algebras”; elements  $r\varepsilon^j$  ( $r \in \mathbf{R}$ ) may be visualized as occupying floor  $j$  of a building, thus the infinitesimals occupy the basements ( $j < 0$ ), with lower basements containing infinitesimals of smaller kind; however, a typical  $z \in \mathbf{R}^\varepsilon$  will have terms on several floors. In this paper, a function with domain  $\mathbf{R}^\varepsilon$  will be assumed to vanish on the “infinite value” domains  $\Omega_- = \{x \in \mathbf{R}^\varepsilon: q(x) > 0, x < 0\}$  and  $\Omega_+ = \{x \in \mathbf{R}^\varepsilon: q(x) > 0, x > 0\}$ . Instead of formal power series (2), coefficient sequences could be used instead; however the familiar manipulations of power series are easier to calculate. Schmiden and Laugwitz [7] used a sequence space, to extend  $\mathbf{R}$  to a space including infinitesimals and infinite values; however, their space is too large for the present purpose, and is not an ordered field.

For  $t \in \mathbf{R}$ , define the *monad*  $[t]^\varepsilon$  of  $t$  as the set of points of  $\mathbf{R}^\varepsilon$  which differ from  $t$  by an infinitesimal; thus  $[t]^\varepsilon = \{t + u \in \mathbf{R}^\varepsilon: q(u) < 0\}$ . Define also

$$[t]^\varepsilon_r = \{t + u \in \mathbf{R}^\varepsilon: -r \leq q(u) < 0\}. \tag{6}$$

Thus  $[t]_r^{\mathfrak{S}}$  is a subset of the monad, restricting the degree of the infinitesimals added to  $t$ . The space  $\mathbf{R}^{\mathfrak{S}}$  may be constructed by replacing each  $t \in \mathbf{R}$  by its monad, then adjoining  $\Omega_-$  and  $\Omega_+$ . It will be convenient later to use subsets of  $\mathbf{R}^{\mathfrak{S}}$ , constructed from  $\mathbf{R}$  by replacing certain  $t \in \mathbf{R}$  by appropriate sets  $[t]_{r(t)}^{\mathfrak{S}}$ .

Sequential convergence in  $\mathbf{R}^{\mathfrak{S}}$  is defined by convergence separately at each level. Thus, let  $x^{(n)} = \sum_j x_j^{(n)} \epsilon^j$  and  $x = \sum_j x_j \epsilon^j$  be in  $\mathbf{R}^{\mathfrak{S}}$ . Then, by definition,

$$\{x^{(n)}\} \rightarrow x \text{ in } \mathbf{R}^{\mathfrak{S}} \Leftrightarrow (\forall j) \{x_j^{(n)}\} \rightarrow x_j \text{ in } \mathbf{R}. \tag{7}$$

A similar definition applies to convergence of nets, replacing  $n (\rightarrow \infty)$  by  $\alpha \in D$ , a directed set, and also when  $n \rightarrow \infty$  is replaced by  $t \rightarrow c$  in  $\mathbf{R}$ . One consequence is that the chain rule for derivatives holds (with the same proof) for differentiable functions with domains and ranges in  $\mathbf{R}^{\mathfrak{S}}$ . Another is that, if  $I$  is a compact interval in  $\mathbf{R}$ , then a function  $g: I \rightarrow \mathbf{R}^{\mathfrak{S}}$  may be expanded as  $g(t) = \sum_j g_j(t) \epsilon^j$  with each  $g_j(t) \in \mathbf{R}$ , and then

$$\int_I g(t) dt = \sum_j \left( \int_I g_j(t) dt \right) \epsilon^j, \tag{8}$$

provided that each integral on the right exists (say as a Riemann or Lebesgue integral.)

The usual (Cauchy-Weierstrass) limits in  $\mathbf{R}$  are used here. If they were to be replaced by use of infinitesimals, a continuum of infinitesimals would presumably be required, such as the Robinson theory provides, rather than the sequence  $\{\epsilon^j: j = 1, 2, \dots\}$ .

### 3. Generalized functions

Consider, as a motivating example, a generalized function  $g$ , described formally by  $g(x) = f(x) + \delta'(x)$  for  $x \in \mathbf{R}$ , where  $f$  is a smooth function, and  $\delta'$  denotes the derivative of the Dirac function. The generalized function  $g$  will be considered as a conventional function from a superset  $E$  of  $\mathbf{R}$  onto  $\mathbf{R}^{\mathfrak{S}}$ . Since  $\delta'(x) = 0$  for  $0 \neq x \in \mathbf{R}$ ,  $\delta'$  may be regarded as “localized” within  $[0]_1^{\mathfrak{S}}$ , or even  $[0]_1^{\mathfrak{S}}$ , and a suitable domain  $E$  for  $g$  is obtained by replacing  $0 \in \mathbf{R}$  by  $[0]_1^{\mathfrak{S}}$ . Then  $g$  is trivially extended to domain  $\mathbf{R}^{\mathfrak{S}}$ .

The delta function can be defined as follows. For  $t \in \mathbf{R}$ , let  $\tilde{\phi}(t) = (2\pi)^{-1} \exp(-\frac{1}{2}t^2)$ . Extend  $\tilde{\phi}$  to  $\phi: \mathbf{R}^{\mathfrak{S}} \rightarrow \mathbf{R}^{\mathfrak{S}}$  by defining  $\phi(x) = \tilde{\phi}(t)$  whenever  $x \in [t]_1^{\mathfrak{S}}$ , and  $\phi(x) = 0$  when  $q(x) > 0$ . Then the delta function may be defined by

$$\delta(x) = \epsilon^{-1} \phi(\epsilon^{-1}x) \quad (x \in \mathbf{R}^{\mathfrak{S}}). \tag{9}$$

Using the chain rule,  $\delta'(x) = \varepsilon^{-2}\phi'(\varepsilon^{-1}x)$  for  $x \in \mathbf{R}^s$ . If  $\psi: \mathbf{R} \rightarrow \mathbf{R}$  is continuous, and has an extension to domain  $\varepsilon\mathbf{R}$  for which  $\psi(\varepsilon v) - \psi(0)$  is infinitesimal for  $v \in \mathbf{R}$ , then

$$\int_{\varepsilon\mathbf{R}} \delta(t)\psi(t) dt = \int_{\mathbf{R}} \varepsilon^{-1}\phi(v)\psi(\varepsilon v)\varepsilon dv = \int_{\mathbf{R}} \phi(v)\psi(0) dv + \theta = \psi(0) + \theta \tag{10}$$

where  $\theta$  is infinitesimal, using (8), with the convention that, for the linear change of variables  $t = \varepsilon v$ , the measure on  $\varepsilon\mathbf{R}$  is  $\varepsilon dv$ . Note that  $\delta(x)$  has been defined so as to be differentiable to all orders on  $\varepsilon\mathbf{R}$ , and continuous on  $\mathbf{R}^s$ . If, moreover,  $\psi$  is differentiable and  $\varepsilon^{-1}[\psi(\varepsilon v) - [\psi(0) + \varepsilon\psi'(0)v]]$  is infinitesimal for  $v \in \mathbf{R}$ , then, for some infinitesimal  $\rho$ ,

$$\begin{aligned} \int_{\varepsilon\mathbf{R}} \delta'(t)\psi(t) dt &= \int_{\mathbf{R}} \varepsilon^{-2}\phi'(v)[\psi(0) + \varepsilon\psi'(0)v]\varepsilon dv + \rho \\ &= \varepsilon^{-1}\phi(0) \int_{\mathbf{R}} \phi'(v) dv + \psi'(0) \int_{\mathbf{R}} \phi'(v)v dv + \rho \\ &= 0 - \psi'(0) + \rho. \end{aligned} \tag{11}$$

Hence

$$\int_{\mathbf{R}^s} \delta(t)\psi(t) dt \in [\psi(0)]^s \quad \text{and} \quad s \int_{\mathbf{R}^s} \delta'(t)\psi(t) dt \in [-\psi'(0)]^s \tag{12}$$

hold, under the stated conditions. Another differentiable function could replace  $\tilde{\phi}$  here, provided that  $\int_{\mathbf{R}} \tilde{\phi}(t) dt = 1$  is retained. Since  $\phi$  is differentiable arbitrarily often,  $\delta$  is also differentiable arbitrarily often on  $\varepsilon\mathbf{R}$ , with

$$\delta^{(k)}(x) = \varepsilon^{-k-1}\phi^{(k)}(\varepsilon^{-1}x) \quad (k = 0, 1, 2, \dots; x \in \mathbf{R}), \tag{13}$$

with  $\phi^{(k)}$  extending  $\tilde{\phi}^{(k)}$  exactly as  $\phi$  extends  $\tilde{\phi}$ . Note that  $\delta(\cdot)$  is continuous at the ‘‘boundary’’ between  $\varepsilon\mathbf{R}$  and  $(0, \infty)$ , since  $\delta(\cdot) = 0$  on  $(0, \infty)$ , and, for  $v \in \mathbf{R}$ ,  $\delta(\varepsilon v) = \varepsilon^{-1}\phi(v) \rightarrow 0$  as  $v \rightarrow \infty$ .

The generalized function  $\delta'$  is an example of a class of local generalized functions, obtained by replacing  $\varepsilon^{-2}\phi'$  above by another suitable function  $h$ .

A local generalized function (l.g.f.) of depth  $r > 0$  at  $x \in \mathbf{R}$  is a function  $p: \mathbf{R}^s \rightarrow \mathbf{R}^s$  defined by

$$p(t) = 0 \quad \text{for } t \notin [x]^s; \quad p(x + y) = h(\varepsilon^{-r}y) \quad \text{for } x + y \in [x]^s; \tag{14}$$

where  $\tilde{h}: \mathbf{R} \rightarrow \mathbf{R}^s$  is a continuous function, with  $h(\pm\infty) = 0$ , extended to  $h: \mathbf{R}^s \rightarrow \mathbf{R}^s$  by

$$h(t) = h(u) \quad \text{if } t \in [u]^s, \quad h(t) = 0 \quad \text{if } q(t) > 0. \tag{15}$$

Then  $p(t + w) = p(t)$  when  $t \in [x]_r^{\mathfrak{S}}$  and  $q(w) < -r$ ; and  $p(x + y) = 0$  when  $q(y) > -r$ . Consequently  $p$  has an *effective domain*  $[x]_r^{\mathfrak{S}}$ ; adding an infinitesimal of lower index to the argument does not affect the function value.

Let  $f: \mathbf{R} \rightarrow \mathbf{R}^{\mathfrak{S}}$  be a conventional function,  $S$  a discrete subset of  $\mathbf{R}$  (*discrete* means that each  $s \in S$  has positive distance from  $S \setminus \{s\}$ ), and  $p_s$  (for  $s \in S$ ) a finite sum of l.g.f. at  $s$ , having maximum depth  $d_s$ ; at  $t \in \mathbf{R}$ , let (each component of)  $f$  have  $k_t (\geq 0)$  derivatives. Define a *generalized function*  $g$  from  $\mathbf{R}$  into  $\mathbf{R}$  as the function from  $\mathbf{R}^{\mathfrak{S}}$  into  $\mathbf{R}^{\mathfrak{S}}$  given, for  $x \in \mathbf{R}$ , and  $x + y \in [x]_r^{\mathfrak{S}}$ , by

$$g(x + y) = \sum_{s \in S} p_s(x + y) + \sum_{j=0}^{k_x} f^{(j)}(x) y^j / j!. \tag{16}$$

The first summation in (16) is finite by construction from l.g.f. The *depth* of  $g$  at  $x$  is  $d_s$  if  $x = s \in S$ , otherwise 0. Denote by  $GF$  the space of generalized functions so defined.

For example, if  $f(t) = t$ , and  $S = \{0\}$ ,  $p_0(t) = \delta(t) + \delta'(t)$ , then  $g$  has depth 1 at 0, and  $g(\varepsilon v) = \varepsilon v + \varepsilon^{-1} \phi(v) + \varepsilon^{-2} \phi'(v)$  for  $v \in \mathbf{R}$ . If, instead,  $g(t) = t + \delta'(\varepsilon^{-1}t)$ , then  $g$  has depth 2 at 0, and, for  $v, w \in \mathbf{R}$ ,

$$g(\varepsilon v + \varepsilon^2 w) = \varepsilon v + \varepsilon^2 w + \varepsilon^{-2} \phi'(w) \theta(v) \tag{17}$$

where  $\theta(v) = 0 (v \neq 0)$ ,  $\theta(v) = 1 (v = 0)$ .

These examples of generalized functions are differentiable arbitrarily often. From (14) and (16), a generalized function  $g$  is differentiable  $\min(k_x, b_x)$  times at  $x$ , where (if  $x \in S$ ) each function  $h$  defining a l.g.f. at  $x$  is differentiable  $b_x$  times, or  $b_x = \infty$  if  $x \notin S$ . If  $f$  is a conventional function, continuously differentiable  $k$  times at 0, then (16) with  $p_s = 0$  defines  $f$  as a generalized function; then a calculation similar to (11) shows that, whenever  $k \geq l$ ,

$$\int_{\mathbf{R}^{\mathfrak{S}}} \delta^{(l)}(t) f(t) dt \in [(-1)^l f^{(l)}(0)]^{\mathfrak{S}}. \tag{18}$$

If  $k$  is replaced in (16) by  $k'$  with  $l \leq k' < k$ , then the generalized function corresponding to the function  $f$  is altered, but (18) is unchanged.

The above definitions extend readily to generalized functions of two or more variables. In defining a l.g.f. on  $\mathbf{R}^2$ , the point  $x \in \mathbf{R}$  is replaced by a curve  $\Gamma \subset \mathbf{R}^2$ . Let  $\Gamma$  be a rectifiable curve in  $\mathbf{R}^2$ , with arc length coordinate  $s$ , and having a normal defined for almost all  $s$ . The local coordinates  $(s, n)$ , where  $n$  denotes distance normal to  $\Gamma$ , are then defined for almost all  $s$ . Then a l.g.f. of depth  $r$  on  $\Gamma$  is a function of form  $p((s, n)) = h(s, \varepsilon^{-r}n)$ , where  $h(s, \cdot)$  has the properties of  $h$  in (15), and  $p(x) = 0$  for  $x \in \mathbf{R}^2 \setminus \Gamma$ . This construction can describe, for example, a distribution of delta functions along a curve, or a double-layer or dipole moment distribution along a curve, and the curve need

only have the restricted amount of smoothness specified here. In three dimensions, the curve is replaced by a surface, for which finite area is defined, and a normal exists for almost all points on the surface. Another kind of l.g.f. may be defined on a curve  $\Gamma \subset \mathbf{R}^3$ , having a normal defined at almost all points; then  $(s, n)$  is replaced by  $(s, n, u)$ , where  $u$  is distance perpendicular to tangent and normal, and a l.g.f. has the form  $p((s, n, u)) = h(s, \varepsilon^{-1}(n, u))$ , where  $h(s, \cdot)$  has the properties of  $h(\cdot)$  in (15), but with argument in  $\mathbf{R}^2$ . There are obvious analogues for dimensions greater than three. Then the definition above of generalized function applies, replacing  $f: \mathbf{R} \rightarrow \mathbf{R}^s$  by  $f: \mathbf{R}^n \rightarrow \mathbf{R}^s$  with  $n = 2, 3, \dots$ . The same definition, applied to  $m$  functions, will define generalized functions from  $\mathbf{R}^n$  into  $\mathbf{R}^m$  (as suitable functions  $g: \mathbf{R}^{sn} \rightarrow \mathbf{R}^{sm}$ ).

The *effective domain* of  $g \in GF$  has as its components the components of the open set  $\mathbf{R} \setminus S$  (or  $\mathbf{R}^n \setminus S$ ), together with the sets  $s + \varepsilon^j \mathbf{R}$  for each  $s \in S$  and  $1 \leq j \leq d_s$ . The integral of  $g$  over  $\mathbf{R}^s$  may be written conventionally as  $\int_{\mathbf{R}} g(t) dt$ ; it reduces to the integral of  $g$  over its effective domain. Using the limit definition (7), integration by parts remains valid, provided that the generalized functions concerned are continuous over the whole domain, and satisfy the usual differentiability requirements within each component of the effective domain of the functions. Integrals over components may be added if there are finitely many components, or (from Lebesgue theory) there are countably many components and the  $g_j(t)$  are nonnegative or satisfy a dominance condition.

A generalized function  $g$  has *compact support* if  $g$  vanishes outside a set  $\bigcup_{t \in K} [t]^s$ , where  $K$  is compact in  $\mathbf{R}$  (or  $\mathbf{R}^n$ ).

#### 4. Some examples and applications of these generalized functions

(a) Using (18) and integration by parts (valid here in consequence of the limit definition (7)), the usual manipulations of  $\delta^{(k)}(\cdot)$  hold, up to an infinitesimal (which makes no difference to calculations with integrals). Also

$$t\delta(t) = 0 \quad (\text{if } 0 \neq t \in \mathbf{R}), \quad t\delta(t) = v\phi(v) \quad (\text{if } t = \varepsilon v, v \in \mathbf{R}). \quad (19)$$

Hence  $t^{-1}(t\delta(t)) = \delta(t)$  in the present theory, verifying associativity of multiplication here. (In view of Schwartz's counterexample, associativity is only possible with  $t\delta(t) \neq 0$ ; however  $\int_{\mathbf{R}} t\delta(t)\psi(t) dt = 0$  for any smooth function  $\psi$ .)

(b) Other delta functions are possible, as well as  $\delta$  from (9). With the same  $\phi$ , consider

$$\Delta_2(x) = \varepsilon^{-2}\phi(\varepsilon^{-2}x) \quad (x \in \mathbf{R}^s). \quad (20)$$

Note that  $\Delta_2$  has depth 2, whereas  $\delta$  from (9) has depth 1. If  $\psi$  is continuous, and is extended to  $\mathbf{R}^s$  so that  $\psi(x + y) - \psi(x)$  is infinitesimal when  $x \in \mathbf{R}$  and  $x + y \in [x]^s$ , then

$$\int_{\mathbf{R}^s} \Delta_2(x)\phi(x) dx = \int_{\varepsilon^2\mathbf{R}} \varepsilon^{-2}\phi(\varepsilon^{-2}x)\psi(x) dx = \int_{\mathbf{R}} \varepsilon^{-2}\phi(w)[\psi(0) + \beta] \varepsilon^2 dw$$

(where  $x = \varepsilon^2w$ ,  $w \in \mathbf{R}$ , and  $\beta$  is infinitesimal)

$$\in \left[ \psi(0) \int_{\mathbf{R}} \phi(w) dw \right]^s = [\psi(0)]^s. \tag{21}$$

(Note that substituting  $x = u + \varepsilon v + \varepsilon^2w + \varepsilon^2z$  with  $u, v, w \in \mathbf{R}$  and  $z$  infinitesimal does not alter (21).) Thus  $\Delta_2$  is also a delta function. Moreover  $\Delta_2$  (unlike  $\delta$  from (9)) possesses a square root in  $GF$ , namely

$$\Delta_2(x)^{1/2} = \varepsilon^{-1}[\phi(\varepsilon^{-2}x)]^{1/2} \quad (x \in \mathbf{R}^s). \tag{22}$$

Similarly, by replacing  $\varepsilon^{-2}$  by  $\varepsilon^{-k}$  in (20), a delta function could be constructed which possesses a  $k$ th root in  $GF$ .

Rational fractional powers of the delta function have been applied in quantum physics [12]. Note that [12] defines the delta function as  $(n/\pi)^{1/2}\exp(-nx^2)$ , with  $n$  an ‘‘infinite value’’ from Robinson’s theory; the choice  $\varepsilon = (2n)^{-1/2}$  recovers  $\delta$  in (9). The fractional powers of the delta function obtained in  $GF$  here are not identical with those in [12], but allow similar calculations.

In  $GF$ ,  $\delta(t)^2$  is well defined, but has values in level 2 of  $\mathbf{R}^s$ , so that the formula  $\delta(t)^2 = \text{const} \cdot \delta(t)$ , sometimes used in quantum physics contexts, is not supported by the present theory. (In view of [12], it need not be.)

(c) If  $f, g \in GF$ , then their convolution

$$(f * g)(t) = \int_0^t f(t - s)g(s) ds \tag{23}$$

is directly defined already, as an integral over  $[0, t]^s$ . The commutative and associative properties of  $*$  then follow if Fubini’s theorem applies; it does if  $f$  and  $g$  are defined, using (16), from local generalized functions whose defining functions  $\tilde{h}$  are locally integrable (at each level) and locally integrable conventional functions from  $\mathbf{R}$  into  $\mathbf{R}^s$ . Most applications fulfil these conditions. (If  $f_0^t$  in (23) is replaced by  $f_{\mathbf{R}}$ , then the generalized functions should all have compact support, or all have support in some half-line.) In particular, since  $[0, t]^s \supset [0]^s$ , substituting  $u = \varepsilon v$ ,  $v \in \mathbf{R}$  shows that

$$(f * \delta)(t) = \int_{\mathbf{R}} f(t - \varepsilon v)\phi(v) dv. \tag{24}$$

Hence  $f * \delta = f$  when  $f$ , in some neighbourhood of 0, is a continuous conventional function. However,  $\delta * \delta \neq \delta$ .

Suppose now that  $f \in GF$  is derived, as in (16), from a conventional function which is continuous at  $s \in \mathbf{R}$  and a l.g.f. having depth  $r > 0$  at  $s$ . (For example,  $f(x) = x + \delta(x)$ ,  $s = 0$ ,  $r = 1$ .) Let

$$\Delta_k(x) = \varepsilon^{-k}\phi(\varepsilon^{-k}x) \tag{25}$$

be a version of the delta function, having depth  $k > r$ . It follows that  $f * \Delta_k = f$ . It suffices to prove the case where  $f(s + z) = h(\varepsilon^{-r}z)$ , as in (14), for  $z \in [0]^{\mathfrak{S}}$ . Then, setting  $u = \varepsilon^k v$ ,  $v \in \mathbf{R}$ ,  $y \in \mathbf{R}$ , (24) and (25) show that

$$\begin{aligned} (f * \Delta_k)(s + \varepsilon^r y) &= \int_{\mathbf{R}} h(\varepsilon^{-r}(\varepsilon^r y + \varepsilon^k v)) \varepsilon^{-k}\phi(v) \varepsilon^k dv \\ &= \int_{\mathbf{R}} h(y + \varepsilon^{k-r}v) \phi(v) dv \\ &= \tilde{h}(y) \int_{\mathbf{R}} \phi(v) dv \quad (\text{from (15)}) \\ &= f(s + \varepsilon^{-r}y). \end{aligned} \tag{26}$$

Thus  $f * \Delta_k = f$  holds with a suitable version  $\Delta_k$  (depending on  $f$ ) of the delta function. A similar argument to (26) shows that, if  $\{f_n\}$  is a sequence of l.g.f. having the same depth for each  $n$ , and the corresponding  $\tilde{h}_n(\cdot) \rightarrow 0$  boundedly, then  $(f_n * g)(\cdot) \rightarrow 0$  pointwise on  $\mathbf{R}^{\mathfrak{S}}$ .

(d) The Fourier and Laplace transforms of  $g \in GF$  are defined directly as integrals:

$$\hat{g}(s) = \int_{\mathbf{R}} e^{ist}g(t) dt \quad \text{and} \quad \bar{g}(\theta) = \int_{\mathbf{R}^+} e^{-\theta t}g(t) dt, \tag{27}$$

when those integrals exist. Consider  $g$  as defined by (16), and  $e^{ist}$ ,  $e^{-\theta t}$  as generalized functions, defined for  $t = x + y \in [x]^{\mathfrak{S}}$  by Taylor series, as in (16) (with  $k_x = \infty$ ). For example, if  $p$  is l.g.f. of depth  $r$  at  $x$ , as in (14), then (27) gives

$$\hat{p}(s) \in [e^{isx}\tilde{h}(0)]^{\mathfrak{S}} \quad \text{and} \quad \bar{g}(\theta) \in [e^{-\theta x}\tilde{h}(0)]^{\mathfrak{S}}. \tag{28}$$

In particular, while  $\tilde{\Delta}_k(s)$  depends on  $k$ ,  $\hat{\Delta}_k(s) \in [1]^{\mathfrak{S}}$  holds for all  $k = 0, 1, 2, \dots$ . However, the transforms of the part of  $g$  derived from a conventional function  $f$  do not necessarily exist in  $GF$ ; for existence, the integral involving  $f$  must converge in the classical sense. For example, the Fourier transform of  $\delta^{(k)}(t)$  is given by  $(-is)^k$ , by the usual integration-by-parts formula; but the integral  $\int_{\mathbf{R}} (-is)^k e^{ist} ds$  is undefined. (In contrast, the Schwartz theory “smooths”  $(-is)^k$  by multiplying it by a “test function” which vanishes suitably at  $\pm \infty$ .) So the usual formulas hold for transforms of derivatives hold here, such as  $(Dg)\hat{\cdot}(s) = -is\hat{g}(s)$ , where  $D = d/dt$ , but not generally Plancherel’s theorem; inverse transforms must replace transforms of transforms.

Consider now an ordinary differential equation with constant coefficients (with independent variable  $t \in \mathbf{R}_+$ ):

$$P(iD)u = g, \tag{29}$$

where  $P$  is a polynomial,  $D = d/dt$ , and  $g \in GF$  is a forcing (generalized) function. In particular, consider

$$P(iD)e = \delta. \tag{30}$$

Well-known methods may be applied. Assuming that  $e$  has a Fourier transform, (30) shows that  $\hat{e}(s) = 1/P(s)$ ; expression of  $1/P(s)$  in partial fractions allows  $e(t)$  to be evaluated, confirming the assumption. Then a solution  $u$  to (29) is obtained as  $u = e * g$ , provided that the convolution exists. Alternatively, “engineers” methods of Laplace transforms may be used to obtain  $u$ , allowing for initial conditons. In the present framework, these methods are rigorous.

It follows that  $e(t) = H(t)\exp(\lambda_j t)$ , where  $H$  is the Heaviside unit function, here defined as  $H(t) = \phi(\varepsilon^{-1}t)$  for  $t \in \mathbf{R}^s$  (so that  $H'(t) = \delta(t)$ ), and the  $\lambda_j$  are the zeros of  $P(i\lambda) = 0$  (with appropriate limiting cases when  $\lambda_j$  coalesce). Similarly, for (29),  $u(t)$  is constructed from integrals of the form

$$\int_0^t e^{i\lambda_j \tau} g(t - \tau) d\tau, \tag{31}$$

together with terms involving  $\exp(\lambda_j t)$  and initial conditions. If each l.g.f. in  $g$  has an absolutely integrable defining function  $\tilde{h}$ , and if the conventional function  $f$  used in defining  $g$  is locally integrable, then the convolution is associative, and the integrals in (31) exist, leading to a solution in  $GF$  for (29).

(e) In some other theories of generalized functions (see *e.g.* Lighthill [3]), the generalized functions corresponding to functions such as  $f(t) = t^{-1}$ , which are not locally integrable, require special treatment. From the present viewpoint, these are ordinary functions like any others. Such formulas as that for the derivative of the logarithm (taking  $\log z$  as the principal value of the logarithm of complex  $z$ ):

$$(d/dz)\log z = z^{-1} - i\pi\delta(z) \quad (z \in \mathbf{R}) \tag{32}$$

are also routine in the present approach, using  $\log z = \log|z| + i\pi H(-z)$ .

### 5. Discussion and further extensions

No detailed bibliography on this topic has been attempted (such a bibliography might appear to be infinite!)

The first use of the series (2) to represent extended real numbers appears to be by Hölder [1] and Schoenflies [8], to represent a non-Archimedean number field. The definition (9) of the delta function may be found in Stroyan and Luxemburg

[10]; but they did not develop the theme, being more concerned to replace limits by infinitesimals. The latter comment also applies to the recent work of Tall [11] and Hoskins [2], who also use extended real numbers of the form (2) to define a delta function. Lighthill [3] has discussed the “delta function and its relations” in detail, from a different viewpoint (defining  $\delta(t)$  by approximating functions). Note, however, that [11] considers only analytic (conventional) functions  $f: \mathbf{R} \rightarrow \mathbf{R}$  for extension to domain  $\mathbf{R}$ , which severely restricts the applications. The present treatment is not so restricted. Generalized functions of several variables still need more investigation.

*Weak convergence* (as used in the Schwartz theory) remains relevant in some contexts. For example, if  $p_\lambda(t) = H(t)t^{\lambda-1}/\Gamma(\lambda)$ , then

$$\int_{\mathbf{R}} p_\lambda(t)f(t) dt \rightarrow \int_{\mathbf{R}} \delta(t)f(t) dt$$

as  $\lambda \downarrow 0$  for each continuous  $f$ , but  $p_\lambda \nrightarrow f$  using (7).

Rosinger [6], following Mikusiński [4], has given several formulas, involving products of generalized functions, said to be relevant in quantum physics. These formulas include

$$\delta(t)^2 - (t^{-1})^2/\pi^2 = -t^{-2}/\pi^2 \tag{33}$$

and

$$t^{-1}\delta(t) = -\frac{1}{2}\delta'(t). \tag{34}$$

These cannot hold exactly in the present theory, both because the generalized functions involved have values on different levels, and also because, with unrestricted multiplication in  $GF$ , (33) would imply that  $\delta(t)^2 = 0$ . However, such formulas can be given meaning by an equivalence concept.

Let  $V$  be a vector space of functions  $f: \mathbf{R} \rightarrow \mathbf{R}$ . Define  $g, G \in GF$  to be *V-weakly equivalent* if, for some infinitesimal  $\beta$ ,

$$(\forall f \in V) \int_{\mathbf{R}} g(t)f(t) dt = \int_{\mathbf{R}} G(t)f(t) dt + \beta. \tag{35}$$

In order to discuss (33), take  $V$  as the space of continuous functions  $f$ , with each  $f \in V$  having support in  $[-c, -\eta] \cup [\eta, c]$  with  $c, \eta > 0$  depending on  $f$ . Now, for  $t \in \mathbf{R}$  (and writing  $D = d/dt$ ),

$$\log t = \log|t| + i\pi H(-t); \tag{36}$$

hence

$$D \log t = t^{-1} - i\pi\delta(t). \tag{37}$$

Similarly, noting that  $\delta - (t) = \delta(t)$ ,

$$D \log(-t) = t^{-1} + i\pi\delta(t). \tag{38}$$

So the left side of (33) may be replaced by the product

$$[\delta(t) - (i\pi t)^{-1}][\delta(t) + (I\pi t)^{-1}] = (D \log t)(D \log(-t))/\pi^2 \equiv g(t). \tag{39}$$

For each  $f \in V$ , for some  $0 < \eta < c$ , depending on  $f$ ,

$$\int_{\mathbf{R}} g(t)f(t) dt = \pi^{-2} \left( \int_{-c}^{-\eta} + \int_{\eta}^c \right) (-t^{-2})f(t) dt.$$

Therefore the left side of (36) is  $V$ -weakly equivalent to the function  $-(\pi t)^{-2}$ . So (33) is confirmed in this restricted sense.

For (34), redefine  $V$  as a space of differentiable functions  $f: \mathbf{R} \rightarrow \mathbf{R}$ , extend to domain  $\varepsilon\mathbf{R}$  by  $f(\varepsilon v) = f(0) + \varepsilon f'(0)v$  for  $v \in \mathbf{R}$ . Then

$$\begin{aligned} \int_{\mathbf{R} \cup \varepsilon\mathbf{R}} t^{-1}\delta(t)f(t) dt &= \int_{\mathbf{R}} [f(0) + \varepsilon f'(0)v](\varepsilon v)^{-1}\varepsilon^{-1}\phi(v) dv \\ &= f(0)\varepsilon^{-1} \int_{\mathbf{R}} v^{-1}\phi(v) dv + f'(0) \int_{\mathbf{R}} \phi(v) dv \\ &= f(0)\varepsilon^{-1} \cdot 0 + f'(0) \cdot 1 \\ &\text{if } \int_{\mathbf{R}} v^{-1}\phi(v) dv \text{ is interpreted as a Cauchy principal value} \\ &= - \int_{\mathbf{R}^{\S}} \delta'(t)f(t) dt. \end{aligned} \tag{41}$$

If, instead,  $f$  is extended to domain  $\mathbf{R}^{\S}$  by  $f(x + z) = f(x) + f'(x)z$  for  $x \in \mathbf{R}$ ,  $x + z \in [s]^{\S}$ , then

$$\int_{\mathbf{R}^{\S}} t^{-1}\delta(t)f(t) dt \in \left[ - \int_{\mathbf{R}^{\S}} \delta'(t)f(t) dt \right]^{\S}. \tag{42}$$

Thus  $t^{-1}\delta(t)$  is  $V$ -weakly equivalent to  $-\delta'(t)$ , but *not* to  $-\frac{1}{2}\delta'(t)$  as in (34).

If a distribution is defined as a linear functional on a subspace  $V$  of a Schwartz space of test functions, then it can be extended by the Hahn-Banach theorem to the whole space, but not uniquely. Only in certain cases is there a canonical extension, which commutes with differentiation. Since  $t^{-1}\delta(t)$  has effectively been defined as a distribution on the subspace  $V$ , other extensions than to  $-\delta'(t)$  would be possible.

One Schwartz distribution, such as the delta function, corresponds to a class of elements of  $GF$ , rather than to a single element. As shown in Section 4(c) for convolutions, the different versions of delta function serve different uses. The apparent arbitrariness in (16) (the definition of  $g$  depends on how many derivatives  $f$  is assumed to have) does not, in fact, disturb the conclusions, as shown by (18). A user of generalized functions may vary somewhat his choice of generalized functions, to suit his application.

### Acknowledgment

The author thanks a referee for various comments on presentation, and for the useful reference [12].

### References

- [1] O. Hölder, "Die Axiome der Quantität und die Lehre von Menge", *Ber. Verh. Sächs. Akad. Wiss.*, Lpz. 53 (1901), 1–69.
- [2] R. F. Hoskins, "Infinitesimals, non-standard analysis, and generalized functions", *Bull. Inst. Math. Appl.* 18 (1982), 49–51.
- [3] M. J. Lighthill, *Introduction to Fourier analysis and generalized functions* (Cambridge Univ. Press, 1958).
- [4] J. Mikusiński, "On the square of the Dirac delta distribution", *Bull. Acad. Polon. Sci.* 14 (9) (1966), 511–513.
- [5] A. Robinson, *Nonstandard analysis* (North-Holland, 1966).
- [6] E. E. Rosinger, *Distributions and nonlinear partial differential equations*, Lecture Notes in Math. 684 (Springer-Verlag, 1978).
- [7] C. Schmieden and D. Laugwitz, "Eine Erweiterung der Infinitesimalrechnung", *Math. Z.* 69 (1958), 1–39.
- [8] A. Schoenflies, "Über die Möglichkeit einer projektiven Geometrie bei transfiniten (nicht archimedischen) Massbestimmungen", *Jber. dt. Mathverein* 15 (1906), 26–31.
- [9] L. Schwartz, *Théorie des distributions* (Hermann, Paris, 1950 and 1966).
- [10] K. D. Stroyan and W. A. J. Luxemburg, *Introduction to the theory of infinitesimals* (Academic Press, New York, 1976).
- [11] D. Tall, "Looking at graphs through infinitesimal microscopes, windows and telescopes", *Math. Gaz.* 64 (1980), 22–49.
- [12] J. K. Thurber and J. Katz, "Applications of fractional powers of delta functions", in *Victoria symposium on nonstandard analysis* (eds. A. Hurd and P. Loeb), Lecture Notes in Mathematics 369 (Springer-Verlag, Berlin, 1974), 272–302.