JOSÉ SEBASTIÃO E SILVA

TEXTOS DIDÁCTICOS

Volume III

SERVIÇO DE EDUCAÇÃO E BOLSAS
FUNDAÇÃO CALOUSTE GULBENKIAN | LISBOA

Edição da
FUNDAÇÃO CALOUSTE GULBENKIAN
Av. de Berna — Lisboa
1999

ISBN 972-31-0971-9
Depósito Legal n.º 148 805/00

III.1

THEORY OF DISTRIBUTIONS*

* Este texto tem por base apontamentos coligidos por diversos alunos de José Sebastião e Silva na sequência de um curso que realizou em 1958 na Universidade de Maryland, e que posteriormente foram utilizados, e por ele revistos, na Faculdade de Ciências de Lisboa.

CHAPTER II

DISTRIBUTIONS OF ONE VARIABLE: FUNDAMENTAL CONCEPTS.

2.1. Terminology and notation

We shall denote by $/R$ the field of real numbers and by $\mathbb{C}$ the field of complex numbers. If $a, b \in /R$, with $a < b$, then

$$[a, b], \ ]a, b], \ [a, b[, \ ]a, b[$$

denote the intervals with extreme points a, b defined respectively by the conditions

$$a \leq x \leq b, \ \ a < x \leq b, \ \ a \leq x < b, \ \ a < x < b$$

They are respectively **closed**, **closed on the right**, **closed on the left**, and **open**. All of them are **bounded**.

An interval (or more generally any set of points of $/R$) is said to be **compact** if it is closed and bounded. By extension of language, any set that reduces to a single point $a \in /R$ is called a **degenerate interval** that is the interval $[a, a]$. However, by an interval we shall always mean a non-degenerate interval, unless the contrary is explicity stated.

On the other hand, the symbols $[a, +\infty[$, $]a, +\infty[$, $]-\infty, a]$, $]-\infty, a[$ denote the **unbounded intervals** defined respectively by $x \geq a, x > a, x \leq a, x < a$.

The first two are bounded on the left and the last two are bounded on the right. The first is closed, the second is open, etc.

Finally the unbounded interval $]-\infty, +\infty[$ is the set $/R$ itself; it is both open and closed (in $/R$).

Let I be an interval in $/R$. We denote by $C(I)$ – or by C if there is no danger of mistake – the set of all complex-valued functions $f(x)$ of the real variable x which are defined and continuous on I. More generally, for any integer $n > 0$, we denote by $C^n(I)$ – or simply C^n – the subset of $C(I)$ formed by those functions f, which have a derivative, $f^{(n)}$, of order n continuous on I; in particular $C^0 = C$. The elements of $C^n(I)$ are said to be **C^n functions on I**.

The term "function" will mean "complex-valued function" wherever the range is not specified.

Instead of $f^{(n)}$ we shall often write $D^n f$. This notation puts in evidence the derivation operator, D, which assigns to each function $f \in C^1$ the function $Df = f' \in C$. So D^n is the n[th] **power** of D.

On the other hand, the symbol $\Im$ denotes an integration operator defined by the formula:

$$\Im f(x) = \int_c^x f(\xi) d\xi, \text{ for all } f \in C$$

where c is an arbitrary fixed point in I. Then $\Im$ is a mapping of the set C into $C^1 \subset C$ such that

$$D\Im f = f, \text{ for all } f \in C.$$

This $\Im$ is a **right-inverse** of D (but not a **left-inverse**). More generally:

2.1.1. $$D^n \Im^n f = f, \text{ for all } f \in C, n = 0, 1, \ldots$$

2.2. Axiomatic introduction of distributions.[1]

Let I be any interval in $/R$; the system of all distributions on I can be described by the following system of fundamental properties:

---

(1) The word "distributions" is used here with a meaning equivalent to that of "distributions of finite order" according to L. Schwartz. We shall further introduce the concept of "global distribution" which is equivalent to that of "distribution" in the sense of Schwartz.

AXIOM 1. *Every function which is defined and continuous on I is a distribution on I.*

AXIOM 2. *To every distribution f on I there corresponds one and only one distribution on I, which is called the "**derivative of f**" and denoted by Df, in such a way that, if f is a C^1 function, then Df is the derivative of f in the ordinary sense.*

DEFINITION: The derivative of order n of a distribution f, which is denoted by $D^n f$, is defined as follows:

$$D^0 f = f\,,\ D^n f = D(D^{n-1} f)\,,\ \text{ for } n = 1, 2, \ldots$$

AXIOM 3. *To every distribution f on I there exists at least one integer $n \geq 0$ and one function F, continuous on I, such that $f = D^n F$.*

AXIOM 4. *If n is an integer, $n \geq 0$, and f, g are two continuous functions on I, then we have $D^n f = D^n g$ if and only if $f - g$ is a polynomial function of degree $< n$.*

We denote by $\mathbb{N}_0$ the set of all integers $n \geq 0$ and by $\mathcal{P}_n$, for each $n \in \mathbb{N}_0$, the set of all polynomial functions of degree $< n$ (restricted to I). Our immediate purpose is to prove that the preceding axioms are:

1° consistent; i.e. there exists at least one structure satisfying the axioms (a model).

2° categorical; i.e. two such models are necessarily isomorphic. This will imply that any statement about distributions on *I* which is not false is a consequence of the axioms, and eventually of some supplementary definitions that have been introduced in order to simplify the language.

We shall begin with the proof of categoricalness because it leads to a natural proof of consistency.

PROOF of categoricalness – Suppose that there is a model *M* satisfying the axioms, i.e. a set of objects $f, g, \ldots$, and an operator *D* such that, if we call these objects the distributions on *I* and Df, Dg, …, the **derivatives** of $f, g, \ldots$, then the axioms are satisfied. The axioms 1 and 2 along with definition 1 imply that, for any couple

(n, F), where $n \in \mathbb{N}_0$ and $F \in C$, there exists one and only one distribution $f = D^n F$ (element of M).[2] Conversely, according to axiom 3, for any $f \in M$ there exists at least one couple (n, F) with $n \in \mathbb{N}_0$, $F \in C$, such that $f = D^n F$.

However there exists more than one couple (n, F) satisfying this condition. Let (m, G) be any couple such that:

2.2.1. $$D^n F = D^m G \quad (m \in \mathbb{N}_0, G \in C),$$

and let k be any integer such that $k \geq m, n$. By axioms 1 and 2, definition 1 and property 2.2.1., we have:

$$D^n F = D^k(\mathfrak{I}^{k-n} F), \quad D^m G = D^k(\mathfrak{I}^{k-m} G)$$

hence

$$D^k(\mathfrak{I}^{k-n} F) = D^k(\mathfrak{I}^{k-m} G)$$

and consequently by axiom 4:

2.2.2. $$\mathfrak{I}^{k-n} F - \mathfrak{I}^{k-m} G \in \mathcal{P}_k.$$

Conversely, axiom 4 shows that 2.2.2. implies 2.2.1. *These two conditions are therefore equivalent.* (For example, if $m \geq n$, we can choose $k = m$, so that condition 2.2.1. is satisfied by all functions G of the form $G = \mathfrak{I}^{m-n} F + P$, where $P \in \mathcal{P}_m$).

Now let us denote by $[n, F]$ the class of all couples (m, G) satisfying 2.2.2., i.e., such that $D^m G = D^n F$, and let us denote by $\tilde{C}$ the set of classes $[n, F]$, with arbitrary $n \in \mathbb{N}_0$, $F \in C$. Then the correspondance:

$$[n, F] \to D^n F$$

is obviously a one-to-one mapping of $\tilde{C}$ onto M such that:

2.2.3. $$[n+1, F] \to D(D^n F).$$

---

(2) Remember that we write C instead of $C(I)$ for the sake of simplicity.

In particular, the correspondence:

$$[0, F] \to F$$

is a one-to-one mapping of a subset C^* of $\tilde{C}$ onto C. Therefore, if we *define*:

2.2.4. $$D[n, F] = [n+1, F]$$

and we identify[3] each element $[0, F]$ of C^* with the function F itself by puting $F = [0, F] = [1, \Im F] = \cdots$, then $\tilde{C}$ becomes a second model, consistent with the axioms, isomorphic to M according to 2.2.3. and 2.2.4.

Thus any model M satisfying the axioms is isomorphic to $\tilde{C}$ and therefore, any two models M and M' are isomorphic (remember that the construction of $\tilde{C}$, based on 2.2.2., is independent of the choice of M). ♦

PROOF of consistency – We have just seen that if there is a model M of the system of axioms then the set $\tilde{C}$, described above, exists too and is also a model. We shall now prove, without assuming the existence of any previous model M, that the set $\tilde{C}$ actually exists and gives us a model of the system of axioms.

Let us consider the set $/N_0 \times C$ of all couples (n, F), where $n \in /N_0$ and $F \in C$. Given two such couples (n, F) and (m, G) we shall write:

$$(n, F) \sim (m, G)$$

if and only if there exists an integer $k \geq m, n$, such that:

2.2.5. $$\Im^{k-n} F - \Im^{k-m} G \in \mathcal{P}_k .$$

---

(3) *By identifying* $[0, F]$ *with* F, we mean in reality that the symbol "$[0, F]$" and its equivalents "$[1, \Im F]$", $[2, \Im^2 F], \ldots$, will denote from now the function F, instead of the class of couples $[0, F]$ equivalent to $(0, F)$. Thus the meaning of the symbol $\tilde{C}$ is also changed.

It is easily seen that the relation "~" just defined is *reflexive* and *symmetrical.* We now prove that it is *transitive*. Observe first that if there exists an integer $k \geq m, n$ satisfying 2.2.5., so does any other integer r, such that $r \geq m, n$. In fact we find that:

$$\mathfrak{I}^{r-n}F - \mathfrak{I}^{r-m}G \in \mathcal{P}_r$$

by applying to both members of 2.2.5. the operator D^{k-r} or $\mathfrak{I}^{r-k}$ according to whether $k \geq r$ or $k < r$. So suppose:

$$(n, F) \sim (m, G) \text{ and } (m, G) \sim (p, H).$$

Then, if we choose $r \geq m, n, p,$ we have:

$$\mathfrak{I}^{r-n}F - \mathfrak{I}^{r-m}G \in \mathcal{P}_r\,, \quad \mathfrak{I}^{r-m}G - \mathfrak{I}^{r-p}H \in \mathcal{P}_r\,,$$

and hence, by addition:

$$\mathfrak{I}^{r-n}F - \mathfrak{I}^{r-p}H \in \mathcal{P}_r\,, \text{ that is } (n, F) \sim (p, H).$$

So the relation ~ is an **equivalence relation** and, as such, it determines a partition of the set $/N_0 \times C$ off all couples (n, F) into equivalence classes. For each couple (n, F), we shall denote by $[n, F]$ the class of all couples which are equivalent to (n, F) and we shall denote by $\tilde{C}$ the set of all of these classes (the "**quotient**" of $/N_0 \times C$ by ~).

The correspondence $[0, F] \rightarrow F$ being a one-to-one mapping of a subset C^* of $\tilde{C}$ onto C, we can identify each element $[0, F]$ of C^* with $F \in C$. Now, let us call the elements of $\tilde{C}$ **distributions on *I*.** *So Axiom 1 is satisfied by* $\tilde{C}$.

Moreover, we shall call $[n+1, F]$ the **derivative** of $[n, F]$ and we shall write:

$$D[n, F] = [n+1, F].$$

According to this definition, there is only one derivative for each $[n, F] \in \tilde{C}$. Indeed, suppose $[n, F] = [m, G]$; this means that

$(n, F) \sim (m, G)$, i.e. $\Im^{k-n}F - \Im^{k-m}G \in \mathcal{P}_k$ for any $k \geq m, n$; hence

$$\Im^{(k+1)-(n+1)}F - \Im^{(k+1)-(m+1)}G \in \mathcal{P}_{k+1},$$

that is $[n+1, F] = [m+1, G]$ which means that $D[n, F] = D[m, G]$.

Moreover if $f \in C^1$, then $D[0, f] = [1, f] = [0, f'] = f'$, since $f - \Im f' \in \mathcal{P}_1$. *So axiom* 2 *is satisfied.*

On the other hand we have:

$$[n, F] = D[n-1, F] = \cdots = D^n[0, F] = D^nF \text{ for every } [n, F] \in \tilde{C}.$$

So axiom 3 *is also satisfied.*

Finally, if $D^nf = D^ng$, with $f, g \in C$, then $[n, f] = [n, g]$, that is $\Im^{k-n}f - \Im^{k-n}g \in \mathcal{P}_k$, for any $k \geq n$. Choosing $k = n$, we see that *axiom* 4 *is also satisfied,* as we have $D^nf = D^ng$ if and only if $f - g \in \mathcal{P}_n$.♦

Conclusion: We have just proved that the set $\tilde{C}$ gives us a model of the preceding system of axioms. We could conceive many other such models, but this would have no essential interest since we have proved that such models are necessarily isomorphic to $\tilde{C}$. The model $\tilde{C}$ itself, after having afforded a simple proof of consistency of axiomatic system, will have no further interest.

From now on, all that matters will be the **rules of calculus of distributions:** that is, *the axioms* 1-4 *and the definitions that will be convenient to add to them, as well as the propositions implied by this system of axioms and definitions.*

In reasoning as well as in calculation, the distributions will be denoted by the notation "D^nf" or by any other that be convenient. But it will no longer be necessary to think of a distribution as a class of couples (n, f). Observe that this situation is quite similar to the one connected with the successive extensions of the number concept.

2.3. Rank of a distribution and further conventions

For each integer $n \geq 0$ we shall denote by $C_n(I)$ – or by C_n, when there will be no danger of mistake – the set of all distributions f on I of the form:

$$f = D^n F, \text{ where } F \in C(I).$$

Observe that $C_0 = C \subset C_1 \subset C_2 \subset \cdots$.

On the other hand, we shall denote by $\mathscr{D}(I)$ – or simply $\mathscr{D}$ – the set of all distributions on I. Thus $\mathscr{D}$ is the union of all the sets $C_n(I)$:

$$\mathscr{D}(I) = \bigcup_{n=0}^{\infty} C_n(I) \quad (C_0 = C),$$

and accordingly we may use the alternative notation C_∞ for $\mathscr{D}$.

2.3.1. We say that a distribution f is of **rank** n if and only if (iff) n is the least integer such that $f \in C_n$.

For example, consider the Dirac δ-distribution which can be defined as follows:

2.3.2. $$\delta = D^2 J, \text{ with } J = \begin{cases} 0, & \text{for } x < 0 \\ x, & \text{for } x \geq 0 \end{cases}.$$

So, $\delta \in C_2$. Suppose there exist a continuous function F, such that $\delta = DF$. Then $DF = D^2 J$, that is $J = \Im F + P$, with $P \in \mathscr{P}_2$. Hence $DJ = F + P'$. But this is impossible as $F + P'$ is continuous and J has no continuous derivative (on $/R$). Consequently the rank of δ is 2.

It follows from this that $\delta^{(n)}$ is of rank $n+2$, for $n = 1, 2, \ldots$.

2.3.3. An interval I is said to be the **domain** of a distribution f iff f is a distribution on I; i.e. $f \in \mathscr{D}(I)$. We also say that f is **defined** on I.

2.4. Addition of distributions

The sum $f+g$, of two distributions f, g, on the same interval I, is to be defined so as to guarantee the following properties:

A1. *If* $f, g \in C(I)$, then $f + g$ *is the sum of the functions in the ordinary sense.*

A2. *If* $f, g \in \mathscr{D}(I)$, *then* $f + g \in \mathscr{D}(I)$ *and* $D(f+g) = Df + Dg$.

Suppose:

2.4.1. $f = D^n F,\ g = D^m G$ with $n, m \in \mathbb{N}_0$, $F, G \in C(I)$.

According to the axioms (2.2), it is possible to represent f and g as derivatives of *the same order* of two continuous functions; indeed, taking $r \geq m, n,$ we have:

2.4.2. $f = D^r \tilde{F},\ g = D^r \tilde{G}$, where $\tilde{F} = \Im^{r-n} F,\ \tilde{G} = \Im^{r-m} G.$

Now, conditions A1 and A2 imply

$$f + g = D^r \tilde{F} + D^r \tilde{G} = D^r(\tilde{F} + \tilde{G}).$$

So,

2.4.3. $$f + g = D^n F + D^m G = D^r(\Im^{r-n} F + \Im^{r-m} G).$$

In this way we assign to each couple (f, g) of distributions on I, *at least* one distribution on I, which is denoted by $f + g$. We shall next prove that there is *only* one distribution $f + g$, for each couple (f, g); i.e., we shall prove the sum $f + g$ *does not depend on the representation* 2.4.1. *of f and g*. Indeed, consider another representation:

$$f = D^\nu \Phi, \quad g = D^\mu \Psi, \quad \text{with} \quad \nu, \mu \in \mathbb{N}_0, \quad \Phi, \Psi \in C.$$

Then, taking $p \geq \nu, \mu$ we get:

$$f + g = D^p(\tilde{\Phi} + \tilde{\Psi}), \quad \text{with} \quad \tilde{\Phi} = \Im^{p-\nu} \Phi, \quad \tilde{\Psi} = \Im^{p-\mu} \Psi.$$

Choose now $k \geq r, p$. Then:

$$D^r(\tilde{F} + \tilde{G}) = D^k(F^* + G^*), \text{ with } F^* = \Im^{k-n} F, \ G^* = \Im^{k-m} G.$$

$$D^p(\tilde{\Phi} + \tilde{\Psi}) = D^k(\Phi^* + \Psi^*), \text{ with } \Phi^* = \Im^{k-\nu} \Phi, \ \Psi^* = \Im^{k-\mu} \Psi.$$

But $D^k F^* = D^k(\Im^{k-n} F) = D^n F = f$ and $D^k \Phi^* = D^k(\Im^{k-\nu} \Phi) = D^\nu \Phi = f.$

Then $D^k F^* = D^k \Phi^*$ and $F^* - \Phi^* \in \mathcal{P}_k$.

Analogously $D^k G^* = D^k \Psi^*$ and $G^* - \Psi^* \in \mathscr{P}_k$; hence $(F^* + G^*) - (\Phi^* + \Psi^*) \in \mathscr{P}_k$, which means:

$$D^k(F^* + G^*) = D^k(\Phi^* + \Psi^*)\ ;\ \text{i.e.,}\ \ D^r(\tilde{F} + \tilde{G}) = D^p(\tilde{\Phi} + \tilde{\Psi}).$$

Thus, we have proved that the sum $f + g$ is uniquely defined for each couple (f, g). Besides, it is obvious that conditions A1 and A2 are actually satisfied by addition defined according to 2.4.3.. Hence addition in $\mathscr{D}(I)$ can be defined either implicity by the properties A1 and A2 or explicity by formula 2.4.3.. Moreover, this operation is:

I. *Associative:* $(f + g) + h = f + (g + h),\ \ \forall f, g, h \in \mathscr{D}$.

II. *Commutative:* $f + g = g + f,\ \ \forall f, g \in \mathscr{D}$.

III. *Reversible:* for any two distributions f, g on I, there exists one and only one distribution ξ on I, such that $f + \xi = g$.

To prove these properties, it is sufficient to represent f, g, h as derivatives of the same order of continuous functions and to apply the corresponding properties of addition in C.

The preceding properties I, II, III along with the existence and uniqueness of $f + g$ in $\mathscr{D}$, for all $f, g \in \mathscr{D}$, can be expressed shortly by saying:

2.4.4. *$\mathscr{D}$ is a commutative group with respect to addition.*

2.5. Multiplication by complex numbers

The product, αf, of a complex number α by a distribution f is to be defined so as to guarantee the two following properties:

P1. – *If $f \in C(I)$, then αf is the product of α by f in the ordinary sense.*

P2. – *If $f \in \mathscr{D}(I)$, then $\alpha f \in \mathscr{D}(I)$ and $D(\alpha f) = \alpha(Df)$.*

Suppose $f = D^n F$, with $n \in \mathbb{N}_0$, $F \in C$. Then P1 and P2 imply the explicit definition:

$$\alpha f = \alpha D^n F = D^n(\alpha F),\ \text{with}\ \alpha F \in C(I).$$

Thus to each couple (α, f), where $\alpha \in \mathbb{C}$ and $f \in C(I)$, there is assigned *at least* one distribution on I, which is denoted by αf. It is easily seen that the product αf is *unique* for each couple (α, f); i.e., *does not depend on the representation of* f. Moreover, it is quit trivial to prove that if $f, g \in \mathscr{D}(I)$ and $\alpha, \beta \in \mathbb{C}$, then:

$$\left.\begin{array}{ll} \text{I.} & \alpha(f+g)=\alpha f+\alpha g \\ \text{II.} & (\alpha+\beta)f=\alpha f+\beta f \end{array}\right\} \textit{(distributive laws)}$$

III. $(\alpha\beta)f=\alpha(\beta f)$ (*associative law*)

IV. $1 \cdot f = f$.

We have seen (2.4.4.) that $\mathscr{D}(I)$ is a module, i.e., a commutative group with respect to addition. As usual, this fact along with properties I-IV, can be expressed by saying:

2.5.1. *$\mathscr{D}(I)$ is a vector space over $\mathbb{C}$ (or a complex vector space).*

On the other hand, the conjunction of the properties $D(f+g)= =Df+Dg$ and $D(\alpha f)=\alpha Df$ is equivalent to the property:

$$D(\alpha f+\beta g)=\alpha Df+\beta Dg, \quad \forall \alpha, \beta \in \mathbb{C}; \; f, g \in \mathscr{D}(I)$$

and it may be expressed by saying:

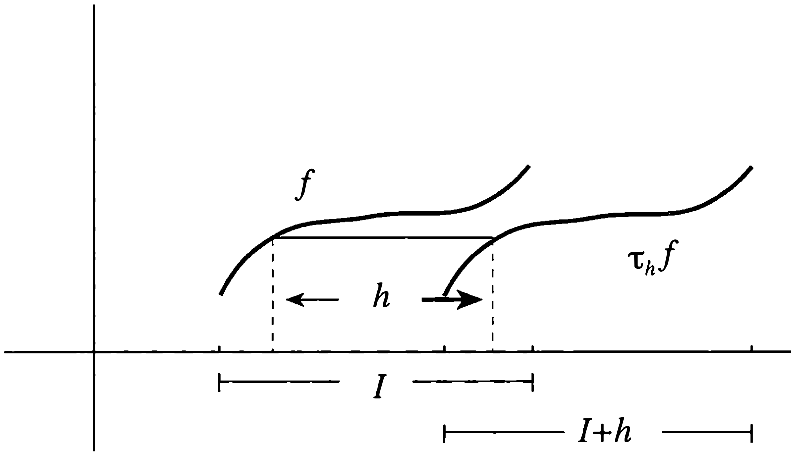
2.5.2. *The operator D is a linear mapping of the space $\mathscr{D}(I)$ into itself.*

We shall further be concerned with the more delicate problem of defining the product of two distributions.

2.6. Translation operators

If $f \in C(I)$, $h \in /R$, then $\tau_h f$ is the function defined as follows:

2.6.1. $$\tau_h f(x)=f(x-h)$$



The graph of $\tau_h f$ is that of f translated by an amplitude h. In particular, the domain of $\tau_h f$ is the interval $J = I + h$. If $h \neq 0$, we have $I=J$ if and only if $I = /R$.

Thus τ_h denotes a one-to-one mapping of $C(I)$ onto $C(J)$, which is natural to call a **translation operator**. Accordingly, we shall call $\tau_h f$ the **h-translate** of f.

Taking 2.6.1. into account it is readily seen that

$$\tau_h (Df) = D(\tau_h f) \text{ if } f \in C^1(I)$$

The extension of the operator τ_h to distributions is defined so as to generalize this property. So we set by *definition*:

$$\tau_h (D^n F) = D^n (\tau_h F), \ \forall n \in /N_0, \ F \in C(I).$$

It is obvious that this formula actually defines a one-to-one mapping τ_h of $\mathscr{D}(I)$ onto $\mathscr{D}(J)$ whose inverse is τ_{-h}. Besides, it is easily seen that for any $h \in /R$, this operation *is linear and interchangeable with D*, that is:

$$\tau_h (Df) = D(\tau_h f) \text{ for any } f \in \mathscr{D}(I).$$

For example, for $\delta = D^2 Y_1$ (cf. 2.3.2.), we have:

$$\tau_h \delta = D^2(\tau_h Y_1), \text{ with } \tau_h Y_1(x) = \begin{cases} x-h, & \text{if } x \geq h \\ 0 & \text{, for } x < h \end{cases}.$$

The distribution $\tau_h \delta$, which we shall also denote by $\delta_{(h)}$ is the **Dirac distribution at the point *h*.**

Remarks about notation. If f is a function and x a point of its domain, the symbol $f\,(x)$ denotes the value that f assumes at this point. When the point x is not specified, we are dealing with a variable and the expression "function $f\,(x)$" is generally used instead of "function f". Now, it must be remembered that this is an *abuse* of language which is certainly convenient in many situations, but which can lead to error in other cases, *especially in functional analysis.* In these cases it is advisable to adopt the convention consisting on writing the accent ˆ over the variable which is then said to be an *apparent or mute variable.* So the symbols f, $f(\hat{x})$, $f(\hat{t})$, … , become equivalent. For example, the expression $3x^2+x$ is only a *variable dependent* on x; meanwhile the expression $3\hat{x}^2+\hat{x}$ denotes properly the function f defined by $f(x)=3x^2+x$, for all $x\in I\!R$.

These conventions can be extended to distributions. If f is a distribution on I and x a point of I, then the symbol $f(x)$ has generally no meaning for there is in general no value of a distribution at a point, as we shall see. But it is often convenient to use the symbol $f(\hat{x})$ for denoting the distribution f. *Accordingly, the distribution* $\tau_h f$ *may be suggestively denoted by* $f(\hat{x}-h)$. *In particular, we may write* $\delta(\hat{x}-a)$ *for* $\tau_a \delta$ *and more generally*

$$\delta^{(n)}(\hat{x}-a) \textit{ instead of } \tau_a \delta^{(n)}.$$

Frequently, we shall write $f(x)$ instead of $f(\hat{x})$ or f. It must be remembered however that this is an *abuse of writing*, which we can admit for the sake of simplicity, *whenever no misunderstanding is possible.*

2.7. Restrictions operators

If $f\in C(I)$, the **restriction** of f to an interval $J\subset I$ is the function f^* whose domain is J and such that:

$$f^*(x)=f\,(x), \text{ for all } x\in J.$$

We denote by $\rho_J f$ the function f^*, which is the restriction of f to J.

It is obvious that the symbol ρ_J denotes a linear mapping of $C(I)$ into $C(J)$, interchangeable with D, that is $\rho_J(Df)=D(\rho_J f)$, for all $f\in C^1(I)$. It is then natural to put by *definition*:

2.7.1. $$\rho_J(D^n F)=D^n(\rho_J F),\ \forall n\in /N_0,\ F\in C(I).$$

Thus the operator ρ_J becomes a linear mapping of $\mathscr{D}(I)$ into $\mathscr{D}(J)$ such that:

2.7.2. $$\rho_J(Df)=D(\rho_J f),\ \forall f\in\mathscr{D}(I).$$

Observe, that *the restriction operator ρ_J may reduce the rank of a distribution.* For example, the distribution $sin\,\hat{x}-3\delta+\delta'(\hat{x}-3)$ which is of rank 3 on $/R$ (cf. 2.3), becomes of rank 2 by restriction to $]-\infty, 3[$ and of rank 0 by restriction to $]-\infty, 0[$.

Another property of the restriction operators, which is easily shown, is the following:

2.7.3. *If I, J, K are three intervals such that $I\supset J\supset K$, then*

$$\rho_K f=\rho_K(\rho_J f),\ \forall f\in\mathscr{D}(I).$$

2.8. Collecting principle. Global distributions (or distributions in the sense of Schwartz)

Let I_1 and I_2 be any two intervals in $/R$ (distinct or coincident). Then:

2.8.1. DEFINITION. Two distributions, $f\in\mathscr{D}(I_1)$ and $g\in\mathscr{D}(I_2)$ are said to be **equal on an interval** $J\subset I_1\cap I_2$ iff $\rho_J f=\rho_J g$. Then we write $f=g$ on J.

2.8.2. LEMMA. *Let I_1, I_2, be two open intersecting intervals in $/R$ and f_1, f_2, two distributions on I_1 and I_2 respectively, such that $f_1=f_2$ on $I_1\cap I_2$. Then there exists one and only one distribution f on the interval $I_1\cup I_2$ such that $f=f_1$ on I_1 and $f=f_2$ on I_2.*

PROOF. Suppose $f_1 = D^n F_1$ and $f_2 = D^n F_2$ with $F_1 \in C(I_1)$ and $F_2 \in C(I_2)$. Then $F_1 - F_2$ equals a polynomial P of degree $< n$ on $I_1 \cap I_2$. Therefore if we put $F = F_1$ on I_1 and $F = F_2 + P$ on I_2, we define a continuous function F on $I_1 \cup I_2$, and the distribution $f = D^n F$ satisfies the condition of the lemma. Conversely, a distribution $g = D^q G$ satisfying this condition coincides necessarily with f, as is readily seen. ♦

2.8.3. Collecting principle (1st form). *Let $I_1, \ldots, I_n$ be n open intervals whose union is again an interval I, and let $f_1, \ldots, f_n$ be n distributions on $I_1, \ldots, I_n$ respectively, satisfying the conditions $f_j = f_k$, on $I_j \cap I_k$, whenever $I_j \cap I_k$ is not empty (j, $k = 1, 2, \ldots, n$). Then there exists one and only one distribution f on I such that $f = f_j$* on *I_j for $j = 1, \ldots, n$.*

PROOF. We can suppose the intervals $I_1, \ldots, I_n$ ordered in such a way that if $j < k$, then the left extremity of I_j precedes the left extremity of I_k or, if these extremities are coincident, the right extremity of I_j precedes that of I_k. Then, the successive unions $I_1 \cup I_2$, $(I_1 \cup I_2) \cup I_3$, ... are again intervals, and we can achieve the proof by repeated application of the lemma along with 2.7.3. ♦

2.8.4. Remark. The conclusion is no longer true, if we consider an infinite system of distributions $f_1, f_2, \ldots$ on intervals $I_1, I_2, \ldots$ instead of a finite system. For example, take: $I_n =]-n, n[$ and $f_n = \rho_{I_n}[\delta(\hat{x}) + \delta'(\hat{x} - 1) + \cdots + \delta^{(n-1)}(\hat{x} - n + 1)]$ for $n = 1, 2, \ldots$. Since the rank of f_n is $n + 1$ for each n, *there is no bound for the ranks of* $f_1, f_2, \ldots$, and, therefore, there is no distribution f on */R,* such that $f = f_n$ on I_n, for every n.

But we can extend the collecting principle in the following way:

2.8.5. Collecting principle (2nd form). *Let A be any set of objects. Suppose that to each $\alpha \in A$ is assigned an open interval I_α and a distribution f_α on I_α, in such a way that:*

(i) the union of all these intervals is again an interval I;

(ii) whenever two of these intervals, I_α and I_β, intersect, then $f_\alpha = f_\beta$ on $I_\alpha \cap I_\beta$;

(iii) there exists an integer γ such that the rank of f_α is $\leq \gamma$ for every $\alpha \in A$.

Then, there exists one and only one distribution f on I such that $f = f_\alpha$ on I_α for every $\alpha \in A$.

PROOF. The open interval I can be expressed as the union of a sequence of compact intervals $K_1 \subset K_2 \subset \cdots$. Now, according to the Heine-Borel principle, there exists for each n a finite system of open intervals $J_n^1, \ldots, J_n^{p_n}$, belonging to the given system (I_α) and *covering* K_n that is such that:

$$K_n \subset J_n, \text{ where } J_n = J_n^1 \cup \cdots \cup J_n^{p_n}.$$

Besides, these intervals may be chosen so that $J_n \subset J_{n+1}$, for any n. Then I is again the union of the increasing sequence of intervals J_n.

Let g_n^k be the distribution of the system (f_α), assigned to J_n^k, for every $n = 1, 2, \ldots$, and $k = 1, \ldots, p_n$. According to 2.8.3., there is one and only one distribution g_n on J_n such that $g_n = g_n^k$ on J_n^k for any n and k. On the other hand, by the hypothesis (iii), there exists necessarily for each n a function $G_n \in C(J_n)$, such that $g_n = D^\gamma G_n$. Since $g_{n+1} = g_n$ on J_n for every n, it is easily seen that we can choose the functions G_n so that $G_{n+1} = G_n$ on J_n for each n. But then it is obvious that there exists one and only one function $G \in C(I)$ such that $G = G_n$ on J_n for $n = 1, \ldots$. Consequently, if there exists a distribution f on I such that $f = f_\alpha$ *on* I_α for every $\alpha \in A$, then necessarily $f = g_n$ on J_n for $n = 1, 2, \ldots$ and therefore $f = D^\gamma G$.

Conversely, the distribution $g = D^\gamma G$ satisfies the condition $g = f_\alpha$ on I_α for every $\alpha \in A$. In fact, if we denote by g_α the restriction of g to I_α and if we put $J_{n_\alpha}^k = J_n^k \cap I_\alpha$, whenever this intersection is not empty, then I_α is the union of all $J_{n_\alpha}^k$ and we have $g_\alpha = f_\alpha$ on each interval $J_{n_\alpha}^k$. Hence, by the uniqueness property just proved, $g_\alpha = f_\alpha$; i.e. $g = f_\alpha$ on I_α (for every α). So the proof is concluded. ♦

Condition (iii) is obviously necessary in theorem 2.8.5. However, the preceding example (2.8.4.) suggests a generalization of the concept of distribution. *First of all we shall consider more generally open sets in /R instead of open intervals.* Remember that every open set Ω in /R, is the union of a finite or countable system of mutually disjoint open intervals (the so-called **components** of Ω). For the same purpose, we could consider still more generally any set which results from any open set Ω by adding one or more *boundary points* to Ω. But this case reduces to the preceding one, as we shall see.

2.8.6. DEFINITION. Let Ω be any open set in $I\!R$ and let us suppose that to each compact interval $I \subset \Omega$ is assigned a distribution f_I on I in such a way that for any two compact intervals I_1 and I_2 contained in Ω, we have $f_{I_1} = f_{I_2}$ on $I_1 \cap I_2$, whenever $I_1 \cap I_2$ is not empty neither degenerate. The system (f_I) of distributions defined in this way is called a **global distribution** on Ω, and Ω is called the **domain** of f. The distributions f_I are the **components** of f.

We shall denote by $\overline{\mathscr{D}}(\Omega)$ the set of all global distributions on Ω.

Given two elements, $f = (f_I)$ and $g = (g_I)$, of $\overline{\mathscr{D}}(\Omega)$, and a complex number α, we shall define $f + g$ (**sum** of f and g), αf (**product** of α by f) and Df (**derivative** of f), by the formulas:

$$f + g = (f_I + g_I), \quad \alpha f = (\alpha f_I), \quad Df = (Df_I).$$

It is immediately seen that $\overline{\mathscr{D}}(\Omega)$ *is a complex vector space with respect to the first two operations* and that *D is a linear mapping of* $\overline{\mathscr{D}}(\Omega)$ *into itself.*

Observe now that to each continuous function f on Ω corresponds the global distribution (f_I), where f_I is the restriction of f to I and that this correspondence is a one-to-one linear mapping of $C(\Omega)$ onto a subspace $\overline{C}(\Omega)$ of $\overline{\mathscr{D}}(\Omega)$ such that if $f \in C^1(\Omega)$, then $D(f_I)$ corresponds to the derivative of f in the usual sense. *Then, we can identify every function* $f \in C(\Omega)$ *with the corresponding element* (f_I) *of* $\overline{\mathscr{D}}(\Omega)$ *so that* $C(\Omega)$ *becomes a subspace of* $\overline{\mathscr{D}}(\Omega)$.

In particular, Ω may be an interval. Then taking 2.8.5. into account, we see that the space $\mathscr{D}(\Omega)$ of all distributions on Ω, can be identified, in the same way, with a subspace of $\overline{\mathscr{D}}(\Omega)$.

In the general case, we shall call any element f of $\overline{\mathscr{D}}(\Omega)$ of the form $f = D^n F$, with $n \in I\!N_0$ and $F \in C(\Omega)$, a **distribution** on Ω. It is easily seen that the set of all distributions on Ω, which we shall denote by $\mathscr{D}(\Omega)$, is then a vector subspace of $\overline{\mathscr{D}}(\Omega)$.

Distributions may be called **global distributions of finite rank**. According, a global distribution which is not distribution is said to be of **infinite rank**.

The concept of restriction, as well as def. 2.8.1., can be extended, in a natural way, to global distributions. Then we can also extend to global distributions the *Collecting Principle* 2.8.5., considering more generally open sets instead of open intervals and suppressing condition (iii).

It should be observed that *there exists actually global distributions of infinite rank*. An example is suggested in 2.8.4.

2.9. Carrier of distribution

Global distributions of infinite rank have rather a theoretical interest mainly connected with the functional theory of L. Schwartz. So we shall hence forth confine our discussion to distributions.

We say that a distribution f on a open set Ω in $I\!R$ is **null on a open set** $O \subset \Omega$ iff f equals the zero function on O.

2.9.1. LEMMA. *The union of all open sets O where a distribution is null, is again an open set Ω_0 where f is null (hence the greatest open set where f is null).*

PROOF. Let I_0 be *any component* of Ω_0. Then I_0 is an open interval which is the union of a system (I_α) of open intervals where f is null. But the zero function is also null on all intervals of the system. Hence, by the collecting principle (2.8.5.), f is equal to the zero function on I_0, and since I_0 is any component of Ω_0, it follows that f is null on Ω_0. ♦

2.9.2. DEFINITION. Let f be a distribution on an open set Ω in $I\!R$ and let Ω_0, be the largest *open* set where f is null. Then the set $\Omega \backslash \Omega_0$ (complement of Ω_0 in Ω) is called the **carrier** of f.

According to this definition, the carrier of f is always *closed relatively to Ω*. In particular, if $\Omega = I\!R$, the carrier of f is a closed set.

Examples: I – If f is a continuous function on $I\!R$, the carrier of f is the *closure* of the set of all points x, such that $f(x) \neq 0$. Thus the function f, such that $f(x) = \sin x$ when $\sin x > 0$ and $f(x) = 0$ when $\sin x \leq 0$, is a continuous function on $I\!R$ whose carrier is the set of all points x such that $\sin x \geq 0$.

II – The carrier of the distribution $3\delta + \delta'(\hat{x}+1)$ reduces to the isolated points 0 and -1.

2.9.3. Proposition. *The carrier of a distribution f on $I\!R$ reduces to a single point a if and only if f is a linear combination of derivatives of $\delta(\hat{x}-a)$, $\sum_{j=0}^{m} c_j \delta^{(j)}(\hat{x}-a)$, where m is an arbitrary integer >0 and $c_0, \ldots, c_m$ are arbitrary complex constants which do not all vanish together.*

PROOF. This condition is obviously sufficient. Let us suppose, conversely, that f is a distribution on $I\!R$ whose carrier reduces to one point a. Then f is of the form $f = D^n F$ with $F \in C(I\!R)$ and, since $f = 0$ on the set of all points $x \neq a$, F is represented by two polynomials, P_1 and P_2, of degree $<n$ for $x<a$ and for $x>a$ respectively. Hence, putting $G = F - P_1$, $P = P_2 - P_1$, we have $f = D^n G$, with $G(x) = P(x)$ for $x > a$ and $G(x) = 0$ for $x < a$. *Since F is continuous on $I\!R$,* so is G and hence $G(a) = P(a) = 0$. Consequently, P must have the form:

$$P(x) = \alpha_1 (x-a) + \alpha_2 (x-a)^2 + \cdots + \alpha_{n-1}(x-a)^{n-1}.$$

Put now, for $k = 0, 1, \ldots$

$$x_+^k = \begin{cases} x^k, & \text{if} \quad x \geq 0 \\ 0\,, & \text{if} \quad x < 0 \end{cases}.$$

Then we have $G(x) = \sum_{k=1}^{n-1} \alpha_k (x-a)_+^k$ and $k!\,\delta(x) = D^{k+1} x_+^k$ for $k = 1, 2, \ldots$. Consequently, putting $n-2 = m$, $(k+1)!\,\alpha_{k+1} = c_{m-k}$, we obtain

$$f = D^n G = \sum_{k=0}^{m} c_k \delta^{(k)}(\hat{x}-a). \ \blacklozenge$$