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 I

HEURISTIC INTRODUCTION

1.1. Intuitive antecedents of the theory of distributions

One of the principal aims of the theory of distributions has been the justification of certain methods of calculation and reasoning used empirically by physicians and engineers, for example in the symbolic calculus of electricians (Heaviside school) and in quantum mechanics (Dirac school).

Although these methods were improvided with a logic foundation, they enabled one to obtain with simplicity and elegance the solution of problems for which the classical resources of analysis were, as a rule, less practical and even, sometimes, not efficient. However, the use of the new methods was a permanent defiance to traditional rigor of mathematics and led eventually to contradiction or to unclear situations, which made more and more desirable a rational foundation of those methods.

The history of mathematics offers many examples of similar situations. The creation of differential and integral calculus, the introduction of complex numbers, etc., were imposed by the necessity of rationalizing some heuristic methods of calculation and reasoning which mathematicians and physicists were naturally induced to apply.

Like those historical examples, the goal of the theory of distributions does not reduce to the justification of heuristic methods to avoid more or less contradictory situations. On the contrary, this theory has opened to mathematics, either pure or applied, new possibilities to an extent that is not yet possible to foresee.

The concept of distribution is a generalization of the concept of function in a similar way in which, for example, the concept of complex number is a generalization of the concept of real number. One of the distributions that first appeared outside the classical domain of mathematics is the entity improperly called the "Dirac's δ-function" (or "unitary impulse function" of electricians). According to physicists and engineers this entity should be, in the case of one variable, a function $\delta(x)$ satisfying the two conditions:

$$(1) \qquad \delta(x) = \begin{cases} +\infty & , \text{ if } x = 0 \\ 0 & , \text{ if } x \neq 0 \end{cases}$$

$$(2) \qquad \int_a^b \delta(x)\, dx = 1 \quad \text{whenever} \quad a < 0 < b.$$

Condition (1) defines, by itself, a function in the real axis – the function that assumes the value $+\infty$ at the point 0 and the value 0 at any point $x \neq 0$. But condition (2) in conjunction with (1), is contrary to any theory of the integral, either classical or modern. In fact the function $\delta(x)$ defined by (1) is integrable, in the Lebesgue sense, on any interval of $/R$, but its integral is always zero, since $\delta(x)$ differs from the zero function on $/R$ only at the point zero.

It might be attempted to define a new concept of integral in order to satisfy condition (1) along with (2). But this is impossible without renouncing the most elementary properties of the integral. For example, we have $2\delta = \delta$ since $2\delta(x) = 0$ for $x \neq 0$, and $2\delta(x) = +2\infty = +\infty$ for $x = 0$. Therefore, if $a < 0 < b$, we should have by (2)

$$\int_a^b 2\delta(x)\, dx = \int_a^b \delta(x)\, dx = 1.$$

But, on the other hand, in order to maintain the elementary property concerning the integral of a function multiplied by a constant, we should have:

$$\int_a^b 2\delta(x)\,dx = 2\int_a^b \delta(x)\,dx = 2.$$

So the integral would not be well defined if condition (2) would be satisfied.

Hence there exists no function $\delta(x)$ *satisfying at the same time* (1) *and* (2) *in any reasonable theory of integration.*

Another solution that might be tried would consist in changing the ordinary meaning of the symbol $+\infty$, according to a consistent theory in which $+c\infty$ would be distinct from $+\infty$, for any number $c \neq 1$. Indeed, some attempts have been made by using such a theory of "hyper numbers", but to our knowledge no significant progress has been obtained in this direction.

So, only one way out is left to us: *trying to define* δ, *not as a function of a real variable, but as an entity of a new kind.* For that purpose, it is convenient to come back to the intuitive considerations, which led to the preceding pseudo definition of δ.

1.2. Intuitive origin of the δ-concept

Let us consider a distribution of matter on the real axis $/R$. Then, to each bounded interval I in $/R$ there corresponds a non-negative number $m(I)$ giving the amount of mass contained in I. In many cases there exist a function $P(x)$ of the real variable x such that:

1.2.1. $$m(I) = \int_I P(x)\,dx$$

for every bounded interval I. Then *for any continuity point* x of this function we have

1.2.2. $$P(x) = \lim_{R\to 0} \frac{m(I_R)}{R} \quad , \quad \text{where } I_R = \left[x - \frac{R}{2}, x + \frac{R}{2}\right].$$

The value $P(x)$ is said to be the *density* (of the mass distribution) at the point x.

Let us now consider, for example, the mass distribution on $\mathbb{R}$ consisting of one single material point of mass 1, placed at the origin, the remaining part of $\mathbb{R}$ being unprovided with matter. In this case $m(I)$ is equal to 1 or 0 according as the point 0 belongs to I or not. Assume that there exist a density function of this distribution and denote by $\delta(x)$ that function. Then we should have according to 1.2.2.

$$\delta(x) = \begin{cases} +\infty\ , & \text{if } x = 0 \\ 0\ , & \text{if } x \neq 0 \end{cases}$$

since

$$m(I_R) = \begin{cases} 1\ , \text{ for every } R > 0 \text{ when } x = 0 \\ \\ 0\ , \text{ if } x \neq 0 \text{ and } 0 < R < 2|x|\,. \end{cases}$$

On the other hand, we should have according to 1.2.1.:

$$m[a,b] = \int_a^b \delta(x)dx = 1\ , \text{ whenever } a < 0 < b.$$

So the density function $\delta(x)$ should satisfy conditions (1) and (2) considered in 1.1. *But we have seen that such a function does not exist.*

It might be argued that the concept of a material point is only an abstraction that is never realized in nature. But so is *every* scientific concept: nothing else but a scheme, i.e. a simplified representation that fits more or less successfully to the described situation. One of the tasks of mathematicians is to organize in a consistent theory everybody of schemes sketched by physicists so that the results obtained by such a theory may be logically correct, free from contradictions.

1.3. The theory of measures

Before the theory of distributions, *measure theory* – or better, *the theory of measures* – afforded already a simple and coherent interpretation of the δ-symbol agreeing with preceding intuitive views.

Observe that a mass distribution on /R is not given, in general, by a function of the *real variable x* (or a function of the *point x*), but by a function $m(I)$ of the *variable interval I.* Besides, this function is additive, i.e. if I is decomposed into two or more intervals $I_1, \ldots, I_n$, mutually disjoint, then $m(I) = m(I_1) + \cdots + m(I_n)$. *It is even natural to assume that this additive property, holding for any decomposition into a finite number of intervals, extends also to any decomposition into a countable system of intervals.*

Another example of an additive function of a variable interval can be given by a *probability distribution* on /R, which assigns to each interval I the probability $p(I)$ corresponding to I. Then we always have $0 \leq p(I) \leq 1$.

A third example may be given in certain cases by a distribution of electric charges on /R. But, there, the number $q(I)$ giving the quantity of electricity contained in each bounded interval is a real number that *may be negative.*

Such concrete models suggested the abstract concept of measure. A measure μ is defined on /R if and only if, to each bounded interval I in /R, there is assigned a real or complex number, denoted by $\mu(I)$ or μI (the μ-measure of I) so that if I is expressed as the union of a countable system of mutually disjoint intervals, $I_1, \ldots, I_n, \ldots$, then the series $\sum \mu(I_n)$ is absolutely convergent and $\mu(I) = \sum \mu(I_n)$. In particular cases, there exists a function $f(x)$ (of the real variable x) such that:

$$\mu(I) = \int_I f\,, \quad \text{for each bounded interval } I\,,$$

where $\int_I f$ denotes the integral of f over I in the ordinary sense. *In such cases giving the measure μ is quite equivalent to give the function f.* Then the function f of the variable x is uniquely extended as a function μ of the variable I, so that the measure μ may be identified with the function f. [If the interval I reduces to a single point a such that $f(a) \neq 0$, we must have at the same time $\mu(I) = f(I) = 0$. But this involves no contradiction; it should be remembered that the interval $[a, a]$ is to be distinguished from the point a itself].

In the general case such a density f does not exist (as we have seen with the δ-measure), but the preceding identification suggests, in such cases, the following definition:

1.3.1. $$\mu(I) = \int_I \mu$$

and to say that $\mu(I)$ is the integral of μ over I. Of course these are only new symbols and a new name for the μ-measure of I, but in doing so, we approach successfully the intuitions of physicists. For example, with respect to the $\delta_{(a)}$-measure we have, for each $a \in /R$, the definition:

$$\int_I \delta_{(a)} = \delta_{(a)}(I) = \begin{cases} 1\,, & \text{if} \quad a \in I \\ 0\,, & \text{if} \quad a \notin I \end{cases}$$

Although $\delta_{(a)}$ is not a function of a real variable, we shall sometimes denote this measure by the functional notation $\delta(x-a)$ used by physicists. This problem of notation will be discussed in chapter III [3.6]. If $a = 0$ we denote $\delta_{(0)}$ only by δ or $\delta(x)$.

1.4. Measures as derivatives of functions of a real variable

Let μ be any measure on $/R$. If we chose any point $c \in /R$ and put

$$F(x) = \begin{cases} \mu[c, x]\,, & \text{if } x \geq c \\ -\mu]x, c[\,, & \text{if } x < c \end{cases}$$

then we obviously define a function $F(x)$ of the real variable x on $/R$. If there exists a density function f of the measure μ, we have, in the ordinary sense:

$$F(x) = \int_c^x f(\xi)\, d\xi\,, \quad \text{for any} \quad x \in /R\,,$$

and hence $f(x)$ is the derivative of $F(x)$ at any continuity point of that function.

In general, the density function f may not exist, but the measure μ can be derived from the function F in a precise way that we shall describe in chapter III. *Thus the measure μ is always determined by the function F; on the other hand, two functions, F and G, determine the same measure μ if and only if $F - G$ reduces to a constant function on /R.* That being so, it will be natural to say (by definition) that μ is the derivative of the functions *F, G,* ... and to write:

$$\mu = DF \ , \quad \mu = DG \ , \ \ldots \ .$$

On the other hand it will be natural to say that *F, G,* ... are the **primitives** of μ. For example, a primitive of δ will be the function:

$$H(x) = \begin{cases} \delta[0, x] \ , & \text{if} \quad x \geq 0 \\ \delta\,]x, 0[\ , & \text{if} \quad x < 0 \end{cases}$$

Hence we have $H(x) = 1$ for $x \geq 0$ and $H(x) = 0$ for $x < 0$ (Heaviside's function), and we may write in the preceding sense:

$$\delta = DH \, .$$

A similar, but a little more elaborate example, is given by the measure:

$$\mu = 3\delta_{(2)} - 2\delta_{(4)} + \delta_{(5)}$$

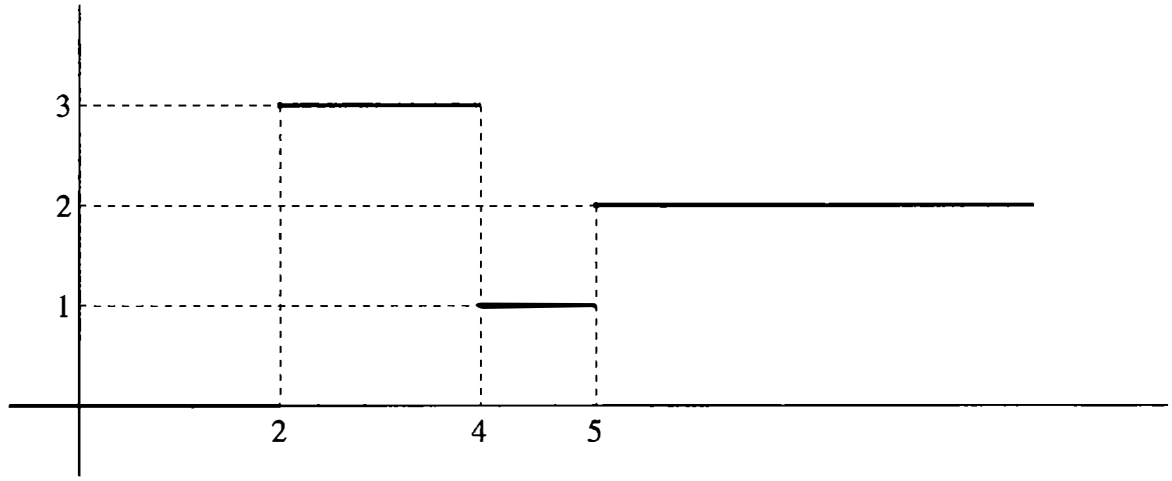
which is defined as follows:

$$\mu(I) = 3\delta_{(2)}(I) - 2\delta_{(4)}(I) + \delta_{(5)}(I) \text{ for any interval } I \, .$$

A concrete model of this measure is given by a system of these electric charges: $q_1 = 3$, $q_2 = -2$ and $q_3 = 1$, placed at the points 2, 4 and 5 respectively. It is readily seen that a primitive of this measure is:

$$F(x) = 3H(x - 2) - 2H(x - 4) + H(x - 5) \, .$$

In this case the charges 3, -2 and 1 are given by the jumps of F at the discontinuity points 2, 4 and 5 respectively.



In general, it is proved that *the primitives of measures are the functions of bounded variation.* So the measures may be characterized as the *derivatives (in the generalized sense) of functions of bounded variation.*

1.5 Insufficiency of the theory of measures

We have just seen that the theory of measures affords a quite satisfactory interpretation of physical schemes such as the δ-concept. But there are physical schemes which may not be interpreted in terms of measures. For example, in the symbolic calculus of electricians, as well as in electromagnetic theory, the use of such entities as the derivatives of δ, which are denoted by δ', δ'',..., are frequent.

If δ' were a measure, then it would be the derivative of some function F of bounded variation; i. e.:

$$\delta' = D\delta = DF.$$

Hence, if we wanted the usual rules for derivatives to hold, we should have $D(\delta - F) = 0$ and δ should be of the form $\delta = F + C$, where C is a constant. But this is impossible since δ is not a function of a real variable. So δ' (as well as δ'', δ''', ...) can not be considered as being

a measure without renouncing the most elementary differentiation rules.

For many cases, symbols like δ', δ'',..., are only calculation devices, comparable to the imaginary numbers, which desappear after helping to find the solution of a problem. However, in other cases, these symbols are directly used for interpreting physical situations. For example, consider the system formed by two electric charges g and $-g$ placed respectively at the points $-h$ and h of $/R$. We then have a charge distribution on $/R$ that is represented by the measures $g\delta(x+h)-g\delta(x-h)$. Suppose now that $h \to 0$ and $g \to +\infty$ in such a way that $2gh$ tends to a finite limite P. Then:

$$\lim_{h\to 0} g[\delta(x+h)-\delta(x-h)] = \lim_{h\to 0}\left[2gh\,\frac{\delta(x+h)-\delta(x-h)}{2h}\right]=P\delta'(x).$$

This being so, $P\delta'(x)$ represents a kind of charge distribution that is not a measure; it is called by physicists a *dipole of momentum P* placed at the origin. Analogously, $\delta''(x)$, which may be described as the limit of the two dipole system $[\delta'(x+R)-\delta'(x)]/R$, as $R \to 0$, is interpreted as a quadripole, and so on. But, of course, since such "distributions" can not be identified with measures, they require an adequate mathematical theory *free from internal contradiction,* extending the theory of measures.

Furthermore, it must be observed that the preceding considerations, that we have related for the sake of simplicity to the real axis, $/R$, are extensible to any $/R^n$ space. For example, if $n=3$, the bounded intervals I are parallelepipeds whose edges are parallel to the three coordinate axes. Then the measure concept may be defined as we did for $/R$; in particular the Dirac measures of $/R^3$ may also be defined as we did for $/R$. For instance, a system of r material points of masses $m_1,\ldots, m_r$, placed at r points $a_1,\ldots, a_r$ of $/R^3$ is represented by the measure:

$$m_1\delta_{(a_1)} + \cdots + m_r\delta_{(a_r)}.$$

On the other hand, if x, y, z are coordinate variables, we have then Dirac measures of one variable $\delta(x)$, $\delta(y)$, $\delta(z)$. For example,

$\delta(x)\,\delta(y)$, represents the measure that assigns to each bounded interval I of $/R^3$ the area of the portion of the xy-plane that is contained in I. It is customary to consider the δ-measure on $/R^3$ as the product of the measures $\delta(x)$, $\delta(y)$, $\delta(z)$:

$$\delta(x, y, z) = \delta(x)\,\delta(y)\,\delta(z)\,.$$

Physics and the theory of probability offer many concrete examples of measures that are not functions of one point. For instance such is the case of a charge distribution on $/R^3$, which reduces to the charge on the surface of one conductor in electrostatic equilibrium; then we may have a *superficial density function,* but not a *spatial density function.*

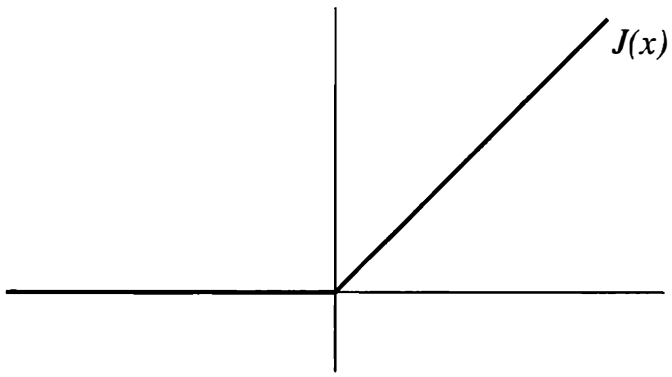
Consider now the symbol $D_z\delta(x, y, z)$. It denotes the derivatives of the δ measure on $/R^3$ with respect to z that may not be identified with a measure and is called by physicists a *dipole* with the *vector momentum* (0, 0, 1) placed at the origin.

Analogously, the symbol $\delta'(z)$ denotes the derivative of $\delta(z)$ with respect to z, which is called a *doublet* on the xy-plan; a plate conductor electrified by induction suggests a scheme of this kind.

These examples and many others that we could draw from several domains of mathematics, whether pure or applied, show the necessity of a theory of distributions.

1.6. Distributions as formal derivatives of continuous functions

Consider again the δ-measure on $/R$. We have seen that $\delta = DH$ where H is the Heaviside function assuming the value 1 for $x \geq 0$ and 0 for $x < 0$. In turn a primitive of H is the function J, such that $J(x) = x$ for $x \geq 0$ and $J(x) = 0$ for $x < 0$. Contrary to the Heaviside function, Y_1 is a continuous function on $/R$.



Observe now that, *in the sense of measure theory, H is* the derivative of J, *but in the ordinary sense,* we have only:

$$H(x) = J'(x), \quad \text{for} \quad x \neq 0$$

since J, has no derivative at 0 (only the right-hand and the left-hand derivatives 1 and 0). Anyway, we have, in the sense of measure theory, $\delta = DH$ and $H = DJ$, hence $\delta = D^2J$. This suggests writing:

$$\delta' = D^3J \quad , \quad \delta'' = D^4J \quad , \ldots$$

so that δ and its derivatives (whatever they may be) come expressed as derivatives of a continuous function on $/R$.

This conclusion holds for any measure μ on $/R$. In fact, a primitive F of μ is always a function of locally bounded variation (hence integrable on any bounded interval), so that a primitive G of F is always a continuous function. Thus we have $\mu = DF = D^2G$ and in general:

$$D^n\mu = D^{n+2}G.$$

In this way, we are always brought back to continuous functions, which are in general more manageable than measures or functions of bounded variation. So, the problem of constructing a theory of distributions can be formulated as follows (for the case of a single variable).

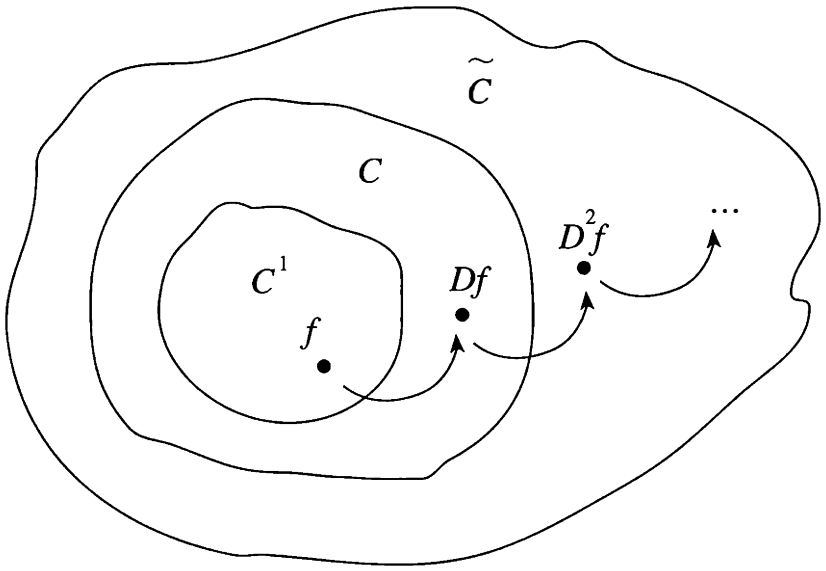
Let us consider the set C of all continuous functions on $/R$ and the set C^1 of all functions f having a continuous derivative, f', on $/R$. We have $C^1 \subset C$, but not $C^1 = C$. So, to each function f in C^1, the derivative operator D assign a function Df (or f') which may not belong to C^1. But if a continuous function does not belong to C^1, then f has no continuous derivative or even *no derivative in the ordinary sense.*

That being so, our purpose is to enlarge the set C, to a set $\tilde{C}$, by adding to it new objects in such a way that the operator D may be extended to the whole set $\tilde{C}$ and that the ordinary formal properties of D may be maintained as much as possible. The elements of $\tilde{C}$ will be called distributions regardless of whether they are functions or

not. If such a set exists, then every continuous function f will have derivatives of all orders

$$Df, D^2f, \ldots, D^nf, \ldots$$

which must be distributions.



In the next chapter this problem will be formulated more precisely and discussed in detail. For the present it is enough to emphasize that the decisive role in defining and solving this problem will be played by the purpose of maintaining the following property of the derivation operator:

If $Df = 0$, on an interval I, then f is a constant function on I.

This implies the more general property:

If $D^nf = 0$, *on an interval I, then f is a polynomial of degree less than n, on I.*

It would be impossible to found an useful theory of distributions without requiring such a condition. In particular all uniqueness theorems for differential equations will depend upon this property.

Obviously the problem of constructing the system $\tilde{C}$ of all distributions on $/R$ must be extended to distributions on an interval I (or even a more general point-set) in any $/R^n$ space. This will be discussed in chapter II (for $n = 1$) and in chapter VII (for $n > 1$).