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 VIII

PARTIAL INTEGRALS AND MULTIPLE INTEGRALS. CONVOLUTION.

8.1. Partial limits for distributions of two variables.

Let I and J be two intervals in $/R$, and suppose that J is unbounded on the right. Given two functions $f(x, y)$ and $g(x)$ respectively on $I\times J$ and I, $f(x, y)$ is said to **converge uniformly** on I to $g(x)$ as $y \to +\infty$, if and only if for every $\varepsilon > 0$, there exists a $\eta \in J$ (independent of x), such that: $|f(x, y) - g(x)| < \varepsilon$ for all $y > \eta$ and $x \in I$.

On the other hand, if α is any real, we write $f(x, y) \in o(y^\alpha)$ **uniformly** on I as $y \to +\infty$, if and only if (iff) $\dfrac{f(x, y)}{y^\alpha} \to 0$ uniformly on I as $y \to +\infty$. Put for every $f \in C(I \times J)$:

$$\mathfrak{I}_x f(x, y) \equiv \int_{x_0}^{x} f(\xi, y)\,d\xi, \quad \mathfrak{I}_y f(x, y) \equiv \int_{y_0}^{y} f(x, \eta)\,d\eta,$$

where x_0, (respectively y_0) is a fixed point, arbitrarily chosen in I (respectively J). The following lemma is easily proved (cf. 6.1.3.):

8.1.1. LEMMA. *If $\alpha > -1$ and $f(x, y) \in o(y^\alpha)$ uniformly on I as* $y \to +\infty$, then

$$\mathfrak{I}_x f \in o(y^\alpha) \quad \text{and} \quad \mathfrak{I}_y f \in o(y^{\alpha+1})$$

uniformly on I as $y \to +\infty$.

This lemma leads to the following:

8.1.2. DEFINITION. If $f \in \mathscr{D}(I \times J)$ and $\alpha > -1$, we write $f(x, y) \in o(y^\alpha)$ on I as $y \to +\infty$, iff there exist $m, n \in I\!N_0$ and $F \in C(I \times J)$ such that:

(1) $f(x, y) = D_x^m D_y^n F(x, y)$;

(2) $F(x, y) \in o(y^{\alpha+n})$ uniformly on each compact interval $I^* \subseteq I$ as $y \to +\infty$.(7)

Applying this definition and the lemma, the following properties are easily shown:

8.1.3. *If $f(x, y) \in o(y^\alpha)$ and $g(x, y) \in o(y^\alpha)$ on I as $y \to +\infty$, then, for all $\lambda, \mu \in \mathbb{C}$:*

$$\lambda f + \mu g \in o(y^\alpha) \text{ on } I \text{ as } y \to +\infty.$$

8.1.4. *If $f(x, y) \in o(y^\alpha)$ on I as $y \to +\infty$, then $D_x f(x, y) \in o(y^\alpha)$ on I as $y \to +\infty$.*

8.1.5. *If $f(x, y) \in o(y^\alpha)$ on I as $y \to +\infty$ and $\varphi(x)$ is multipliable by $f(x, y)$,* then *$\varphi(x) f(x, y) \in o(y^\alpha)$ on I as $y \to +\infty$.*

Consider now $g \in \mathscr{D}(I)$ and $f \in \mathscr{D}(I \times J)$; then:

8.1.6. DEFINITION. We say that $f(x, y)$ **converges** on I to $g(x)$ as $y \to +\infty$, if and only if $f(x, y) - g(x) \in o(1)$ on I as $y \to +\infty$. The distribution $f(x, y)$ is said to be **convergent** on I as $y \to +\infty$, iff there exists a distribution g on I satisfying the preceding condition.

---

(7) A more general condition could be required instead of (2), but this definition is quite sufficient for applications.

The uniqueness property as well as the linearity property of convergence are in this case immediate consequences of 8.1.3. Then we can write:

$$g(x) = \lim_{y \to +\infty} f(x, y) \quad \text{or} \quad g(x) = f(x, +\infty) \quad \text{on} \quad I,$$

to express that $f(x, y) \to g(x)$ on I as $y \to +\infty$.

On the other hand, the following important property, *which does not hold in classical analysis,* is an immediate consequence of 8.1.4.:

8.1.7. DIFFERENTIATION PROPERTY. *If* $f(x, y) \to g(x)$ *on* I *as* $y \to +\infty$, *then* $D_x f(x, y) \to D_x g(x)$ *on* I *as* $y \to +\infty$, that is:

$$D_x \lim_{y \to +\infty} f(x, y) = \lim_{y \to +\infty} D_x f(x, y) \textit{ on } I.$$

It turn, from 8.1.5., follows

8.1.7'. MULTIPLICATION PROPERTY. *If* $f(x, y) \to g(x)$ *on* I *as* $y \to +\infty$ *and* $\varphi(x)$ *is multipliable by* $f(x, y)$, *then*:

$$\lim_{y \to +\infty} [\varphi(x) f(x, y)] = \varphi(x) \lim_{y \to +\infty} f(x, y) \textit{ on } I.$$

Moreover, applying 8.1.7. and the linearity property, it is easily shown:

8.1.8. SUBSTITUTION PROPERTY. *If* $f(x, y) \to g(x)$ *on* I *as* $y \to +\infty$, *and if* h *is a mapping of an interval* I^* *into* I, *such that* $f(h(t), y)$ *exists* (cf. 6.8.), *then* $g(h(t))$ *exists too and* $f(h(t), y) \to$ $\to g(h(t))$ *on* I^* *as* $y \to +\infty$.

This substitution rule concerns the *parameter* x. Substitution rules concerning the *converging variable* y can be easily found as generalization of the criteria given in 6.6.

The "O" symbol is extended to distributions $f(x, y)$ on $I\times J$, with respect to y, in the following way:

8.1.9. DEFINITION. If $\alpha>-1$, we write $f(x, y)\in O(y^{\alpha})$ on I as $y\to+\infty$, if there exist $m, n\in IN_0$ and $F\in C(I\times J)$ such that:

(i) $f(x, y)=D_x^m D_y^n F(x, y)$;

(ii) for every compact interval $I^*\subseteq I$, there exists a number M such that

$$\frac{F(x, y)}{(1+|y|)^{\alpha+n}}\leq M \quad \text{on} \quad I^*\times J.^{(8)}$$

More generally, if $\varphi\in C^{\infty}(J)$, we write $f(x, y)\in O(\varphi(y))$ on I as $y\to+\infty$, if and only if there exists a real y_0 and a distribution $f_0(x, y)\in O(1)$ on I as $y\to+\infty$ such that $f(x, y)=\varphi(y)f_0(x, y)$, for $y>y_0$ and $x\in I$.

Besides the linearity property, it is easily shown:

8.1.10. DIFFERENTIATION PROPERTY. *If α is any real and $f(x, y)\in O(y^{\alpha})$ on I as $y\to+\infty$, then $D_x f(x, y)\in O(y^{\alpha})$ and $D_y f(x, y)\in O(y^{\alpha-1})$ on I as $y\to+\infty$.*

Obviously all preceding considerations extend to the case when J is an interval unbounded on the left, and $y\to-\infty$.

8.2. Partial integrals for distributions of two variables.

Let I and J be any two intervals in IR, $f(x, y)$ a distribution on $I\times J$. A distribution $\varphi(x, y)$ such that $D_y\varphi(x, y)=f(x, y)$ will be called a (**partial**) **primitive** of $f(x, y)$ with respect to y. On the other hand, a distribution $u(x, y)$ on $I\times J$ is said to be **independent** of y, if

---

(8) The choice of $(1+|y|)^{\alpha+n}$ instead of $y^{\alpha+n}$ is only to make the quotient continuous on $I^*\times J$.

and only if it reduces to a distribution g of the variable x only, i.e., iff it is of the form $u(x, y) = D_x^m G(x)$ with $m \in \mathbb{N}_0$, $G \in C(I)$.

8.2.1. LEMMA. *A distribution u on $I \times J$ is independent of y, iff $D_y u = 0$.*

PROOF. It is readily seen that, if u is independent of y, then $D_y u = 0$. Suppose now conversely that $D_y u = 0$ and assume $u = D_x^m D_y^n U$ with $m, n \in \mathbb{N}_0$ and $U \in C(I)$. Then $D_y u = D_x^m D_y^{n+1} U = 0$ and therefore (cf. 7.2. axiom 4) U must be of the form

$$U(x, y) = \sum_0^{m-1} x^i a_i(y) + \sum_0^{n} y^j b_j(x), \text{ with } a_i \in C(J) \text{ and } b_j \in C(I).$$

Hence $u(x, y) = D_x^m D_y^n U(x, y) = n!\, D_x^m b_n(x)$. ♦

That being so, it is easily proved, as in the case of one variable:

8.2.2. THEOREM. *Every distribution f on $I \times J$ has infinitely many primitives with respect to y and two such primitives differ by a distribution independent of y.*

We are now able to define, in a natural way, the concept of **partial** (or **parametric**) **integral** of a distribution $f(x, y)$. It will be sufficient to consider integrals on $\mathbb{R}$. Let I be any interval in $\mathbb{R}$ and $f \in \mathscr{D}(I \times \mathbb{R})$; then:

8.2.3. DEFINITION. The integral $\int_{\mathbb{R}} f(x, y)\,dy$ is said to be **convergent** on I, if and only if there exists a primitive φ of f with respect to y wich is convergent on I as $y \to +\infty$ and as $y \to -\infty$. Then, we write $\int_{\mathbb{R}} f(x, y)\,dy = \varphi(x, +\infty) - \varphi(x, -\infty)$ on I.

From 8.2.2., follows at once the uniqueness of the partial integral. From the properties of partial limits we can deduce the *linearity property for partial integrals*, as well as the following properties:

8.2.4. DIFFERENTIATION PROPERTY. *If* $\int_{/R} f(x, y)\,dy$ *is convergent on I, so is* $\int_{/R} f_x'(x, y)\,dy$ *and*

$$D_x \int_{/R} f(x, y)\,dy = \int_{/R} D_x f(x, y)\,dy \text{ on } I.$$

8.2.5. SUBSTITUTION PROPERTY. *If* $\int_{/R} f(x, y)\,dy = g(x)$ *on I and if h is any continuous mapping of an interval I* into I such that* $f(h(t), y)$ *exists, then*

$$\int_{/R} f(h(t), y)\,dy = g(h(t)) \text{ on } I^*.$$

As for substitutions concerning the integration variable y, the criteria established in 6.6. can be easily extended to partial integrals. In particular, we have, for all $h \in /R$:

8.2.6. $$\int_{/R} f(x, y+h)\,dy = \int_{/R} f(x, y)\,dy.$$

8.2.7. $$\int_{/R} f(x, hy)\,dy = \frac{1}{|h|} \int_{/R} f(x, y)\,dy.$$

Criterium 6.3.6. can also be extended to partial integrals:

8.2.8. THEOREM. *If for any compact interval* $I^* \subseteq I$, *there exists a compact interval K such that* $f(x, y) = 0$ *on* $I^* \times (/R - K)$, *then the integral* $\int_{/R} f(x, y)\,dy$ *is convergent on I.*

PROOF. Suppose $f = D_x^m D_y^n F$, with $F \in C(I \times /R)$. The hypothesis implies that, in a set $\{x \in I^*, y < -y_0\}$, $F(x, y)$ reduces to a pseudo polynomial P of degree $< (m, n)$, which we can assume to be zero, otherwise we could subtract P from F (remember that P is uniquely defined by the value of $F(x, y)$ for m values of x in I^* and n values of y in $/R$). Then there exists a primitive of f with respect to y, say φ, which is zero for $x \in I^*$, $y < -y_0$ and reduces to a function ψ independent of y for $x \in I^*$, $y > y_0$. Now, it is easily seen that $\varphi \to \psi$ on I as $y \to +\infty$ and $\varphi \to 0$ on I as $y \to -\infty$, so that

$$\int_{/R} f(x, y)\, dy = \psi(x). \ \blacklozenge$$

Finally, the following extensions of 6.5.1. and 6.5.2. are easily proved:

8.2.9. THEOREM. *If* $\int_{/R} f(x, y)\, dy$ *is convergent on I, then* $f \in O(y^{-1})$ *on I as* $y \to \infty$. *On the other hand, if there exists* $\alpha < -1$ *such that* $f \in O(y^{\alpha})$ *on I as* $y \to \infty$, *then* $\int_{/R} f(x, y)\, dy$ *is convergent on I.*

Remarks. 1 – If $f(x, y)$ is a function, then for the convergence of $\int_{/R} f(x, y)\, dy$ on *I in distributional sense*, it is not sufficient (nor necessary) that the integral be convergent *for each* $x \in I$. Obviously, a sufficient condition is that the integral be uniformly convergent on each compact subinterval contained in I. More generally, it can be proved that, if f *is summable on each set* $I^* \times /R$, *where* I^* *is a compact interval contained in I, then* $\int_{/R} f(x, y)\, dy$ *is convergent on I in distributional sense.*

2 – The differentiation property can be associated with the linearity property in a more general property. Let $p(D)$ be a *derivation polynomial,* that is an operator of the form $p(D) = \sum_{1}^{n} a_k D^k$, with $a_1, \ldots, a_n \in \mathbb{C}$. Then we have

$$p(D_x)\int_{/R} f(x, y)\,dy = \int_{/R} p(D_x) f(x, y)\,dy \text{ on } I,$$

whenever the first integral is convergent on I.

Example. The preceding remarks offer a simple justification of formula in 6.3.5. - 2. Observe that:

8.2.10. $$\int_{/R} \frac{e^{ixy}}{1+y^2}\,dy = \pi e^{-|x|} \textit{ for each } x \in /R.$$

This can be easily found by the *method of residues*. Besides, as x, y are real variables, we have $|e^{ixy}| = 1$ and

$$\left|\frac{e^{ixy}}{1+y^2}\right| = \frac{1}{1+y^2} \text{ for all } x, y \in /R.$$

Thus the integral 8.2.10. is dominated, for all $x \in /R$, by the integral $\int_{/R}(1+y^2)^{-1}dy$, which is obviously convergent. Hence, according to the Weierstrass test, the first integral is *uniformly convergent on /R and therefore convergent on /R in distributional sense*. Consequently,

$$(1-D_x^2)\int_{/R} \frac{e^{ixy}}{1+y^2}\,dy = \int_{/R} \frac{(1-D_x^2)e^{ixy}}{1+y^2}\,dy = \int_{/R} e^{ixy}dy \quad \text{on} \quad /R.$$

On the other hand,

$$D_x^2 e^{-|x|} = -D_x(e^{-|x|} sgn\ x) = e^{-|x|} - 2e^{-|x|}\delta(x),$$

so that

$$(1-D_x^2)e^{-|x|} = 2\delta(x).$$

Hence, from 8.2.10. follows:

8.2.11. $$\int_{/R} e^{ixy}dy = 2\pi\delta(x) \text{ on } /R.$$

8.3. Multiple integrals (on $/R^n$).

Let f be a distribution on $/R^n$ and λ any complex number. We say that $f(\boldsymbol{x})$ **converges** to λ as $\boldsymbol{x} \to +\infty_n$ if and only if there exist $\boldsymbol{r} \in /N_0^n$ and $F \in C(/R^n)$ such that $f = D^{\boldsymbol{r}}F$ and

$$\frac{F(\boldsymbol{x})}{x_1^{r_1} \cdots x_n^{r_n}} \to \frac{\lambda}{r_1! \cdots r_n!} \quad \text{as } x_1 \to +\infty, \ldots, x_n \to +\infty.$$

Then, we write $\lambda = \lim\limits_{x \to +\infty_n} f(\boldsymbol{x})$ or $\lambda = f(+\infty_n)$.

The uniqueness of the limit, as well as the linearity property can be proved by an argument similar to the one used in the case $n=1$. The concept of convergence as $\boldsymbol{x} \to -\infty$ is analogously defined.

On the other hand, every distribution φ such that $\bar{D}\varphi = f$ (where $\bar{D} = D_1 \cdots D_n$) will be called a **pure mixed primitive** of f. It is easily seen that:

8.3.1. THEOREM. *Every $f \in \mathscr{D}'(/R^n)$ has infinitely many pure mixed primitives and two such primitives differ necessarily by a distribution of the form $\sum_1^n u_j$ where u_j is a distribution independent of x_j (that is, of the form $D^{\boldsymbol{r}}\boldsymbol{u}$ where $\boldsymbol{u}$ is a continuous function on $/R^n$ independent of x_j).*

That being so, we shall write by definition.

8.3.2. $$\int_{\boldsymbol{x}}^{\boldsymbol{x}'} f(\boldsymbol{\xi})\,d\boldsymbol{\xi} = \bar{\boldsymbol{\Delta}}_{\boldsymbol{x}'-\boldsymbol{x}}\,\varphi(\boldsymbol{x}),$$

where φ is any pure mixed primitive of f and $\bar{\Delta}_h$ is the mixed difference operator $\Delta_{1h_1} \cdots \Delta_{nh_n}$.

From 8.3.1. follows that formula 8.3.2. defines actually a distribution $\phi(x, x')$ on $/R^{2n}$ independent of the choice of the pure mixed primitive φ. To see this it is sufficient to observe that $\Delta_{jh_j} u_j = 0$ for every distribution u_j independent of x_j.

8.3.3. DEFINITION. A distribution f is said to be **integrable** on $/R^n$, iff $\int_x^{x'} f(\xi)\,d\xi$ is convergent as $(x, x') \to (-\infty_n, +\infty_n)$. Then we write:

8.3.4.
$$\int_{/R^n} f(x)\,dx = \lim_{\substack{x \to -\infty_n \\ x' \to +\infty_n}} \int_x^{x'} f(\xi)\,d\xi.$$

For example, if $n = 2$

$$\int_{/R^2} f(x_1, x_2)\,dx_1\,dx_2 = \varphi(+\infty, +\infty) - \varphi(+\infty, -\infty) - \varphi(-\infty, +\infty) + \varphi(-\infty, -\infty)$$

where φ is a primitive of f with respect to x.

The integral of f on $/R^n$ can also be denoted by $\int_{/R^n} f$ or simply by $\int f$. Uniqueness and linearity properties are immediate consequences of the corresponding properties for limits. In order to obtain further criteria it is convenient to introduce a suitable definition of bounded distributions.

8.3.5. DEFINITION. A distribution f is said to be **bounded on** $/R^n$, if and only if there exists $r \in /N_0^n$ and $F \in C(/R^n)$ such that:

(i) $f = D^r F$;

(ii) for every regular matrix A of order n, the function $x_1^{-r_1} \cdots x_n^{-r_n} F(Ax)$ is bounded on $/R^n$.

The linearity property of boundedness is easily proved.

8.3.6. DEFINITION. Given $f \in \mathscr{D}(/R^n)$ and $\varphi \in C^\infty(/R^n)$, we write $f \in O(\varphi)$ as $|\boldsymbol{x}| \to \infty$ or simply $f \in O(\varphi)$, if and only if there exists a distribution f_0 bounded on $/R^n$ and a real $\varepsilon > 0$, such that $f = \varphi f_0$, for $|\boldsymbol{x}| > \varepsilon$.

That being so, the following generalization of 6.5.1. is easily obtained:

8.3.7. THEOREM. *If there exists $\alpha < -n$ such that $f \in O(|\boldsymbol{x}|^\alpha)$, then f is integrable on $/R^n$.*

On the other hand:

8.3.8. THEOREM. *Suppose $f \in O(|\boldsymbol{x}|^\alpha)$ with $\alpha < -n$ and let $\boldsymbol{h}$ be a C^∞ one-to-one mapping of $/R^n$ onto itself such that*

(i) the Jacobian matrix $[D_i h_j]$ of $\boldsymbol{h}$ is regular on $/R^n$ and converges to a regular matrix as $|\boldsymbol{t}| \to \infty$,

(ii) $D^{\boldsymbol{r}} D_i h_j \in o(\boldsymbol{t}^{\boldsymbol{r}})$, for all $\boldsymbol{r} \in /N_0^n$, $i, j = 1, \ldots, n$ [(9)].

Then the classical substitution rule applies:

$$\int_{/R^n} f(\boldsymbol{x})\, d\boldsymbol{x} = \int_{/R^n} f(\boldsymbol{h}(\boldsymbol{t})) \left| J\begin{pmatrix} \boldsymbol{h} \\ \boldsymbol{t} \end{pmatrix} \right| d\boldsymbol{t}.$$

We shall outline the proof only in the case when $\boldsymbol{h}$ is a non-degenerate *affine mapping,* that is a mapping of the form $\boldsymbol{h}(\boldsymbol{t}) = \boldsymbol{c} + \boldsymbol{M}(\boldsymbol{t})$, where $\boldsymbol{c}$ is any vector in $/R^n$ and M is a regular matrix of order n. This case may be taken as a model for the general case since $\boldsymbol{h}$ behaves *asymptotically* just as an affine mapping according to *(i).*

(9) As far as functions are concerned is understood that the stated conditions are to be taken in ordinary sense.

Put $\varphi(\boldsymbol{x})=(1+x_1^2+\cdots+x_2^n)^{1/2}$ and suppose $f\in O(|\boldsymbol{x}|^{\alpha})$, with $\alpha<-n$. Then, it is readily seen that $f\in O(\varphi^{\alpha})$, i.e. there exists $\boldsymbol{r}\in/N_0^n$ and $F\in C$, such that $f=\varphi^{\alpha}D^{\boldsymbol{r}}F$, with $x_1^{-r_1}\cdots x_n^{-r_n}\,F(A\boldsymbol{x})$ bounded on $/R^n$ for every regular matrix A of order n. In such conditions it is easily found:

$$\int f(\boldsymbol{x})\,d\boldsymbol{x}=(-1)^{\|r\|}\int \varphi^{(r)}(\boldsymbol{x})\,F(\boldsymbol{x})\,d\boldsymbol{x}\,,\ \text{where } \|\boldsymbol{r}\|=r_1+\cdots+r_n.$$

Now:

$$\int f(\boldsymbol{x})\,d\boldsymbol{x}=(-1)^{\|r\|}\int \varphi^{(r)}(\boldsymbol{x})F(\boldsymbol{x})d\boldsymbol{x}=\int (-1)^{\|r\|}\varphi^{(r)}(\boldsymbol{h}(\boldsymbol{t}))F(\boldsymbol{h}(\boldsymbol{t}))\,|det\ M|\,dt$$

and it can be seen, without difficulty, that the last integral is just equal to

$$(-1)^{\|r\|}\int f(\boldsymbol{h}(\boldsymbol{t}))\,|det\ M|\,d\boldsymbol{t}.$$

8.4. Partial and multiple integrals.

Let us consider a distribution $f(\boldsymbol{x},\boldsymbol{y})$ on $/R^{m+n}$, with $\boldsymbol{x}\in/R^m$ and $\boldsymbol{y}\in/R^n$ (m, $n=1, 2,\ldots$). The concept of partial integral $\int_{/R^n} f(\boldsymbol{x},\boldsymbol{y})\,d\boldsymbol{y}$ can be easily defined as a generalization of preceding concepts of partial and multiple integral, with similar properties. But there is a new property:

8.4.1. THEOREM. *If $f(\boldsymbol{x},\boldsymbol{y})$ is integrable on $/R^{m+n}$ and in addition the integral $\int_{/R^n} f(\boldsymbol{x},\boldsymbol{y})\,d\boldsymbol{y}$ is convergent on $/R^m$, then*

$$\int_{/R^{m+n}} f(\boldsymbol{x},\boldsymbol{y})d\boldsymbol{x}\,d\boldsymbol{y}=\int_{/R^m}\left(\int_{/R^n} f(\boldsymbol{x},\boldsymbol{y})\,d\boldsymbol{y}\right)d\boldsymbol{x}\,.$$

This is a consequence of a property for limits that we can state as follows:

8.4.2. THEOREM. *If $f(\boldsymbol{x}, \boldsymbol{y})$ is convergent as $(\boldsymbol{x}, \boldsymbol{y}) \to (+\infty_m, +\infty_n)$ and if in addition $f(\boldsymbol{x}, \boldsymbol{y})$ is convergent on $\mathbb{R}^m$ as $\boldsymbol{y} \to +\infty_n$, then*

$$\lim_{\substack{x \to +\infty_m \\ y \to +\infty_n}} f(\boldsymbol{x}, \boldsymbol{y}) = \lim_{x \to +\infty_m} \left(\lim_{y \to +\infty_n} f(\boldsymbol{x}, \boldsymbol{y}) \right).$$

PROOF. It is sufficient to prove this rule in the case $m=n=1$. Suppose that the hypothesis holds. Then there exist four integers r, s, t, u, two functions $F_1, F_2 \in C(\mathbb{R}^2)$, a function $G \in C(\mathbb{R})$ and a number λ, such that $f = D_x^r D_y^s F_1 = D_x^t D_y^u F_2$ and

(i) $\dfrac{F_1(x, y)}{x^r y^s} \to \dfrac{\lambda}{r!\, s!}$ as $(x, y) \to (+\infty, +\infty)$;

(ii) $\dfrac{F_2(x, y)}{y^u} \to \dfrac{G(x)}{u!}$, uniformly on each compact set in $\mathbb{R}^m$ as $y \to +\infty$;

We can assume that $t=r$, $u=s$. Take $\varepsilon > 0$, then according to (i) there exist a, $b > 0$ such that:

8.4.3. $$\left| \frac{F_1(x, y)}{x^r y^s} - \frac{\lambda}{r!\, s!} \right| < \varepsilon \text{ for } x > a, y > b.$$

Take now r additional points $x_j > a$, s additional points $y_k > b$ and consider two pseudo-polynomials $\mathcal{P}_1(x, y)$, $\mathcal{P}_2(x, y)$ of degree (m, n) such that $F_1 - \mathcal{P}_1$ and $F_2 - \mathcal{P}_2$ vanish on the lines $x = x_j$, $y = y_k$. Then if we put $F_0 = F_1 - \mathcal{P}_1$, we have $f = D_x^r D_y^s F_0$, $F_0 = F_2 - \mathcal{P}_2$ and it is easily seen that (i), (ii) are again satisfied with F_0 in the place of F_1 and F_2 ($t = r$, $u = s$), since the coefficients of the pseudo-polynomials are obtained as linear combinations of the values of $F_1(x, y)$ and $F_2(x, y)$ on the lines

$x=x_j$, $y=y_k$. Hence from 8.4.3. follows, with F_0 in the place of F_1, and taking the limit as $y \to +\infty$:

$$\left| \frac{G(x)}{x^r} - \frac{\lambda}{r!} \right| \leq s!\varepsilon \text{ for } x>a.$$

The number ε being arbitrary, this implies that $\dfrac{G(x)}{x^r} \to \dfrac{\lambda}{r!}$ as $x \to +\infty$, which means that $\lambda = \lim_{x\to+\infty} \lim_{y\to+\infty} f(x, y)$. ♦

More generally:

8.4.4. *If* $f(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$, *with* $\boldsymbol{x} \in /R^m$, $\boldsymbol{y} \in /R^n$, $\boldsymbol{z} \in /R^p$ *is convergent on* $/R^{m+n}$ *as* $\boldsymbol{z} \to +\infty_p$ *and if* $f(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ *is convergent on* $/R^m$ *as* $(\boldsymbol{y}, \boldsymbol{z}) \to (+\infty_n, +\infty_p)$, *then*

$$\lim_{\substack{\boldsymbol{y}\to+\infty_n \\ \boldsymbol{z}\to+\infty_p}} f(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = \lim_{\boldsymbol{y}\to+\infty_n} \left(\lim_{\boldsymbol{z}\to+\infty_p} f(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \right)$$

8.5. Convolution of two distributions on /R.

Consider two distributions $f = D^m F$ and $g = D^n G$, where $F, G \in C(/R)$. Then we have:

$$f(x-t) = D_x^m F(x-t) = (-1)^m D_t^m F(x-t)$$

so that, for every $k = 0, 1, \ldots$

$$D_t^k f(x-t) = (-1)^k D_x^k f(x-t).$$

This suggests to write by definition

$$f(x-t)\, g(t) = f(x-t)\, D_t^n G(t) = \sum_{k=0}^{n} \binom{n}{k} D_t^{n-k}(G(t) D_x^k f(x-t))$$

with

$$G(t)D_x^k f(x-t) = D_x^{m+k}(F(x-t)G(t)),$$

that is

8.5.1. $$f(x-t)g(t) = \sum_{k=0}^{n} \binom{n}{k} D_x^{m+k} D_t^{n-k}(F(x-t)G(t)).$$

It is easily seen that the "product" $f(x-t)g(t)$ does not depend on the representation of the distributions f and g. We can prove it as we have done for the product of a C^n function with a C_n distribution in 4.1. The analogy between these two situations comes from the following proposition, which can be proved without difficulty, but which is not essential for the following subject: *The mapping* $t \to f(x-t)$ *of* $I\!R$ *into the space* $\mathscr{D}(I\!R)$ *is infinitely differentiable.*

Consider now the expression $f(x-t)g(t-y)$. We have two possible interpretations:

$$f(x-t)g(t-y) = \sum_{k=0}^{n} \binom{n}{k} D_x^{m+k} D_t^{n-k}(F(x-t)G(t-y))$$

8.5.2.

$$f(x-t)g(t-y) = \sum_{k=0}^{m} \binom{m}{k} (-1)^k D_t^{m-k} D_y^{n+k}(F(x-t)G(t-y)).$$

Remembering that the functions F and G can be approached by two sequences $\{F_n\}$ and $\{G_n\}$ of C^∞ functions converging uniformly on each compact interval, it is readily seen that:

8.5.3. *The right members of the formulas* 8.5.2. *represent the same distribution.*

A direct proof of this proposition does not seem to be easy.

8.5.4. DEFINITION. If the integral $\int_{-\infty}^{+\infty} f(x-t)g(t)dt$ is convergent on $/R$, the distribution

$$h(x)=\int_{/R} f(x-t)g(t)dt$$

is called the **convolution** of f and g and denoted by $f*g$.

From this definition, taking into account the linearity property of the partial integral, as well as 8.5.1., follows immediately that the *convolution is bilinear,* that is, we have:

8.5.5. $\quad (\alpha f_1+\beta f_2)*g=\alpha(f_1*g)+\beta(f_2*g),\ \forall\alpha,\beta\in\mathbb{C},$

whenever f_1*g and f_2*g exist, and analogously for the right side. Moreover

8.5.6. COMMUTATIVE LAW: *If $f*g$ exists, $g*f$ exists too, and $f*g=g*f$.*

PROOF. Suppose that $f*g$ exists and put $h=f*g$, that is

$h(x)=\int_{/R} f(x-t)g(t)dt$. Then for each $y\in /R$, we have:

$$h(x-y)=\int_{/R} f(x-y-t)g(t)dt$$

and it is obvious that the last integral is still convergent with respect to (x, y) on $/R^2$. On the other hand, *for each* $y\in /R$, we may perform on this integral the substitution $t=u-y$, which gives:

$$h(x-y)=\int_{/R} f(x-u)g(u-y)du\,.$$

Now, taking 8.5.3. into account, it can be seen that the last integral is also convergent with respect to y *for each* $x\in /R$. In particular, for $x=0$, we have:

$$h(-y)=\int_{IR} f(-u)\,g(u-y)\,du\,.$$

Hence by the substitutions $y=-x,\ u=-t$:

$$h(x)=\int_{IR} g(x-t)\,f(t)\,dt,$$

that is, $h=g*f$. ♦

In the general case the convolution is not associative. But the following criterion can be used in several cases:

8.5.7. *If* $\int_{IR^2} f(x-y)\,g(y-t)\,h(t)\,dy\,dt$, *where* $f,\ g,\ h\in\mathscr{D}$, *is convergent on* IR, *then*

$$(f*g)*h=f*(g*h)=\int_{IR^2} f(x-y)\,g(y-t)\,h(t)\,dy\,dt.$$

This is an immediate consequence of 8.4.4.

In turn, from the differentiation and substitution properties for partial integrals and from 8.5.6., follows immediately, taking definition 8.5.4. into account:

8.5.8. DIFFERENTIATION PROPERTY. *If* $f*g$ *exists, then* $D(f*g)$ exists too, and

$$D(f*g)=(Df)*g=f*(Dg).$$

8.5.9. TRANSLATION PROPERTY. *If* $f*g$ *exists, then for every* $h\in IR$

$$\tau_h(f*g)=(\tau_h f)*g=f*(\tau_h g).$$

On the other hand:

8.5.10. *If $f*g$ and $f*(\hat{x}g)$ exists, then $(\hat{x}f)*g$ exists too and*

$$\hat{x}(f*g)=(\hat{x}f)*g+f*(\hat{x}g).$$

PROOF. It is sufficient to observe that $(\hat{x}f)*g$ is given by

$$\int_{/R}(x-t)f(x-t)g(t)dt=x\int_{/R}f(x-t)g(t)dt-\int_{/R}f(x-t)tg(t)dt. \ \blacklozenge$$

This important property shows that multiplication by x, with respect to convolution, behaves like a derivation operator.

Finally, we can analogously prove that

8.5.11. *If $f*g$ exists, then*

$$e^{ax}(f*g)=(e^{ax}f)*(e^{ax}g) \quad \forall a\in\mathbb{C}.$$

8.6. Convolution of distributions whose carrier is bounded on the left and (or) on the right.

We shall denote by $\mathscr{D}_*(/R)$ or simply $\mathscr{D}_*$ the vector space of all distributions on $/R$ with bounded carrier.

8.6.1. THEOREM. *The convolution $f*g$ exists whenever $f\in\mathscr{D}_*$ and $g\in\mathscr{D}$. Besides,*

(i) $f*(g*h)=(f*g)*h$, whenever f, $g\in\mathscr{D}_*$, $h\in\mathscr{D}$;

(ii) $\delta*f=f$, for every $f\in\mathscr{D}$.

PROOF. a) Suppose $f\in\mathscr{D}_*$, $g\in\mathscr{D}$. Then there exists a bounded interval J such that $g(x-t)f(t)$ vanishes for $(x, t)\notin /R\times J$. Hence

$\int_{/R}g(x-t)f(t)\,dt$ is convergent on $/R$ and gives $f*g$.

b) Suppose f, $g\in\mathscr{D}_*$, $h\in\mathscr{D}$. Then by an argument similar to the preceding it is shown that the integral

$$\int_{/R^2} f(x-y)\,g(y-t)\,h(t)\,dy\,dt$$

is convergent on $/R$, and this according to 8.5.7. implies *(i)*.

c) Consider $f=D^nF$, where $F\in C(/R)$, and put $F_1=FH$, $F_2=F-F_1$. Now:

$$H*F_1=\int_0^\infty H(x-t)F_1(t)\,dt=\int_0^x F_1(t)\,dt.$$

Hence $\delta*D^nF_1=D^{n+1}(H*F_1)=D^nF_1$. It is seen analogously that $\delta*D^nF_2=D^nF_2$, so that $\delta*f=f$. ♦

This theorem along with 8.5.5. can be expressed by saying:

8.6.2. *The space $\mathscr{D}_*$ is an algebra under convolution and $\mathscr{D}$ is a module over that algebra, having δ as unit element.*

Property *(ii)* in 8.6.1. can be expressed explicity by the important formula

$$f(x)=\int_{/R}\delta(x-t)\,f(t)\,dt \quad \text{(DIRAC'S FORMULA)}.$$

We shall denote by $\mathscr{D}_+$ (respectively $\mathscr{D}_-$) the vector space of all distributions vanishing on the left (resp. on the right) of 0 and by $\widetilde{\mathscr{D}}_+$ (resp. $\widetilde{\mathscr{D}}_-$) the space of all distributions whose carrier is bounded on the left (resp. on the right) of 0.

8.6.3. THEOREM. *The space $\widetilde{\mathscr{D}}_+$ (resp. $\widetilde{\mathscr{D}}_-$) is an algebra under convolution and $\mathscr{D}_+$ (resp. $\mathscr{D}_-$) is a subalgebra of $\widetilde{\mathscr{D}}_+$ (resp. $\widetilde{\mathscr{D}}_-$).*

In fact, if $f, g\in\widetilde{\mathscr{D}}_+$, there exists a real c such that f and g vanish for $x<c$. Then $f(x-t)g(t)$ vanish for $t<c$ and $t>x-c$. Hence $\int_{/R} f(x-t)g(t)dt$ is convergent on $/R$ and vanishes for $x<2c$. The remaining parts of the theorem are easily proved. ♦

8.7. Convolution and order of growth, tempered distributions and rapidly decreasing distributions (on /R).

Several criteria can be found, connecting convolution with order of growth of distributions. One of these criteria is the following:

8.7.1. THEOREM. *Let α and β be two real numbers satisfying one of the following conditions*

(i) $\alpha+\beta<-1$ *and* $\alpha\geq 0$;

(ii) $\alpha+\beta<-3$ *and* $\beta\leq\alpha<0$.

On the other hand, let f and g be two continuous functions on /R such that $f\in O(x^{\alpha})$ and $g\in O(x^{\beta})$[(10)]. *Then* $f*g$ *exists and* $f*g\in O(x^{\alpha})$.

PROOF. a) Suppose $\alpha+\beta<-1$ with $\alpha\geq 0$. Then as $f\in O(x^{\alpha})$, there exists a number M such that $|f(x)|\leq M(1+|x|)^{\alpha}$ for all $x\in/R$. Hence

$$|f(x-t)|\leq M(1+|x-t|)^{\alpha}\leq M(1+|x|)^{\alpha}(1+|t|)^{\alpha}\ \forall x,\, t\in/R,$$

since $\alpha\geq 0$.

So the integral

$\int_{/R} f(x-t)g(t)dt$ is dominated by $M(1+|x|)^{\alpha}\int_{/R}(1+|t|)^{\alpha}|g(t)|dt.$

Since $g\in O(x^{\beta})$ and $\alpha+\beta<-1$, the last integral exists. Hence the first integral is uniformly convergent on each compact interval in /R and

its absolute value is $\leq MK(1+|x|)^{\alpha}$ where $K=\int_{/R}(1+|t|)^{\alpha}|g(t)|dt.$ Consequentely, $f*g\in O(x^{\alpha})$.

b) Suppose now $\alpha+\beta<-3$, with $\beta\leq\alpha<0$, and consider the integer n such that $0\leq\alpha+n<1$. Then it is easily seen that $\alpha+\beta+n<-1$ so that $x^k f*x^{n-k}g$ exists and is $O(x^{\alpha+k})$ for $k=0, 1,\ldots, n$ according to the previous conclusion. Hence (cf. 8.5.10):

---

(10) It is understood: "in ordinary sense as $x\to\infty$".

$$x^n(f*g)=\sum_{k=0}^{n}\binom{n}{k}(x^kf*x^{n-k}g)\in O(x^{\alpha+n})$$

so that $f*g\in O(x^\alpha)$. ♦

8.7.2. COROLLARY. *Let α be a real <-2, A_α the set of all continuous functions f on IR such that $f\in O(x^\alpha)$ as $x\to\infty$ and B_α the set of all continuous functions g on R such that there exists a real number $\beta>0$ (depending on g) satisfying the conditions $\alpha+\beta<-1$ and $g\in O(x^\beta)$. Then A_α is an algebra under convolution and B_α is a module over that algebra.*

PROOF. Applying to the theorem (changing the roles of α and β), it is readily seen that $f*g$ exists and belongs to B_α whenever $f\in A_\alpha$ and $g\in B_\alpha$; and that $f*g\in A_\alpha$ whenever f, $g\in A_\alpha$. So we have only to prove the associative law: $f*(g*h)=(f*g)*h$, f, $g\in A_\alpha$, $h\in B_\alpha$. But this can be easily seen applying 8.5.7. as we did for 8.6.1. ♦

8.7.3. Remark: The preceding theorem and corollary can be extended to locally summable functions according to the following criterium (FUBINI-TONELLI THEOREM): *If f, $g\in L(IR)$*, then

$\int_{IR} f(x-t)g(t)dt$ *is convergent almost everywhere in IR and defines a function $h\in L(IR)$*. It can still be stated that the preceding integral is convergent in the mean on IR, so that $f*g$ exists in the distributional sense. Applying 8.5.11. and taking the Fubini-Tonelli theorem into account, it is a simple matter to obtain the following generalization of 8.7.1.:

8.7.4. THEOREM. *Let α, β be two real numbers satisfying the conditions (i) or (ii) of* 8.7.1., *α', β' two real numbers such that $\alpha'+\beta'\leq 0$ and f, g two locally summable functions such that $f\in O(x^\alpha e^{\alpha'|x|})$* and *$g\in O(x^\beta e^{\beta'|x|})$. Then $f*g$ exists and $f*g\in O(x^\alpha e^{\gamma|x|})$*, where $\gamma=max(\alpha',\beta')$.

For the proof it is convenient to consider f and g in the form $f=f_1+f_2$, $g=g_1+g_2$, with f_1, $g_1 \in C_+$, f_2, $g_2 \in C_-$, remembering that $f_1 * g_1 \in C_+$, $f_2 * g_2 \in C_-$.

From 8.7.4. is easily deduced a corresponding generalization of 8.7.2.

Now, applying the differentiation property, we can derive from the preceding criteria corresponding rules for distributions. For example, let us denote by $\widetilde{A}_\alpha$ for every $\alpha<-2$, the set of all distributions of the form $f=\sum_{k=0}^{p} D^{n_k} F_k$, where p, $n_1, \ldots, n_p$ are arbitrary integers and F_k locally summable functions such that $F_k \in O(x^\alpha)$, and by $\widetilde{B}_\alpha$ the set of all distributions of the form $g=\sum_{k=0}^{q} D^{r_k} G_k$ where q, $r_1, \ldots, r_q$ are arbitrary integers and G_k locally summable functions such that $G_k \in O(x^\beta)$ with $\alpha+\beta<-3$ and $\alpha \leq \beta$ (β depending on g). Then it is easily seen that $\widetilde{A}_\alpha$ is an algebra under convolution and $\widetilde{B}_\alpha$ a module over $\widetilde{A}_\alpha$.

8.7.5. DEFINITION. A distribution f on $/R$ is said to be **tempered** (**slowly increasing** or of **polynomial type**) if there exists a real α such that $f \in O(x^\alpha)$ (in distributional sense).

An equivalent definition to this is the following: *f is tempered if and only if there exist two integers n, k and a function $F \in C(/R)$ such that $f=D^n F$ and $F \in O(x^k)$ in ordinary sense.*

We shall denote by $\breve{\mathscr{D}}(/R)$ or simply by $\breve{\mathscr{D}}$ the set of all tempered distributions. It is readily seen that *$\breve{\mathscr{D}}$ is a vector space closed under D.*

8.7.6. DEFINITION. A distribution f on $/R$ is said to be **rapidly decreasing** if and only if for every $\alpha<0$, f can be represented in the form $f=\sum_{k=1}^{p} D^{n_k}F_k$, where p, $n_1,\ldots, n_p$ are arbitrary integers ($n_k \geq 0$, $p \geq 1$) and F_k continuous functions such that $F_k \in O(x^\alpha)$ in ordinary sense.

We shall denote by $\widehat{\mathscr{D}}$ the set of all rapidly decreasing distributions on $I\!R$. From preceding results it is easily deduced:

8.7.7. COROLLARY. $\widehat{\mathscr{D}}$ *is an algebra under convolution and* $\breve{\mathscr{D}}$ *a module over* $\widehat{\mathscr{D}}$.

A similar result can be obtained concerning the space $\overset{\vee}{\mathscr{D}}$ of all distributions of **exponential** type (that is, of the form $f=D^nF$, where F is a continuous function on $I\!R$ such that $F \in O(e^{\alpha|x|})$ for *some* real α) and the space $\overset{\wedge}{\mathscr{D}}$ of all **exponentially decreasing** distributions (that is, of the form $f=D^nF$, where F is a continuous function such that $F \in O(e^{\alpha|x|})$ for *all* real number α).

Observe that $\mathscr{D}_* \subset \overset{\wedge}{\mathscr{D}} \subset \widehat{\mathscr{D}} \subset \breve{\mathscr{D}} \subset \overset{\vee}{\mathscr{D}} \subset \mathscr{D}$.

8.7.8. Convolution in $I\!R^n$. The concept of convolution of distributions on $I\!R$ is readily extended to the case of distributions on $I\!R^n$, and all preceding properties of convolutions can be generalized to this case: only we are now concerned with derivation operators, translation operators, etc., corresponding to the different variables.

Theorem 8.6.1. is readily extended to distributions of several variables. As for theorem 8.6.3. it gives place to new possibilities in the case of n variables.

Let $\boldsymbol{\Gamma}$ be any convex cone in $I\!R^n$ whose vertex is at the origin and not reducing to a half space. We shall denote by $\widetilde{\mathscr{D}}_{\boldsymbol{\Gamma}}$ the set of all distributions on $I\!R^n$ vanishing outside some cone $\boldsymbol{a}+\boldsymbol{\Gamma}$ with $\boldsymbol{a} \in I\!R^n$. Then it is easily seen that $\widetilde{\mathscr{D}}_{\boldsymbol{\Gamma}}$ *is an algebra under convolution and* $\mathscr{D}_{\boldsymbol{\Gamma}}$ *a subalbegra of* $\widetilde{\mathscr{D}}_{\boldsymbol{\Gamma}}$; besides, there exists a maximal subspace of $\mathscr{D}$ distinct from $\mathscr{D}_{\boldsymbol{\Gamma}}$ which is a module over $\widetilde{\mathscr{D}}_{\boldsymbol{\Gamma}}$.

Finally, the criteria given in 8.7. can also be extended to the case of n variables and combined between them and the preceding ones, according to the different variables.