International Tables for Crystallography (2006). Vol. B, Section 1.3.2.3.9, pp. 31–34.

It can be shown that two locally integrable functions f and g define the same distribution, *i.e.*

$$\langle T_f, \varphi\rangle = \langle T_K, \varphi\rangle \quad \text{for all } \varphi \in \mathscr{D},$$

if and only if they are equal almost everywhere. The classes of locally integrable functions modulo this equivalence form a vector space denoted $L^1_{\rm loc}(\Omega)$; each element of $L^1_{\rm loc}(\Omega)$ may therefore be identified with the distribution T_f defined by any one of its representatives f.

1.3.2.3.7. *Support of a distribution*

A distribution $T \in \mathscr{D}'(\Omega)$ is said to *vanish* on an open subset ω of Ω if it vanishes on all functions in $\mathscr{D}(\omega)$, *i.e.* if $\langle T, \varphi\rangle = 0$ whenever $\varphi \in \mathscr{D}(\omega)$.

The *support* of a distribution T, denoted Supp T, is then defined as the complement of the set-theoretic union of those open subsets ω on which T vanishes; or equivalently as the smallest closed subset of Ω outside which T vanishes.

When $T = T_f$ for $f \in L^1_{\rm loc}(\Omega)$, then Supp T = Supp f, so that the two notions coincide. Clearly, if Supp T and Supp φ are disjoint subsets of Ω, then $\langle T, \varphi\rangle = 0$.

It can be shown that any distribution $T \in \mathscr{D}'$ with compact support may be extended from $\mathscr{D}$ to $\mathscr{E}$ while remaining continuous, so that $T \in \mathscr{E}'$; and that conversely, if $S \in \mathscr{E}'$, then its restriction T to $\mathscr{D}$ is a distribution with compact support. Thus, *the topological dual $\mathscr{E}'$ of $\mathscr{E}$ consists of those distributions in $\mathscr{D}'$ which have compact support*. This is intuitively clear since, if the condition of having compact support is fulfilled by T, it needs no longer be required of φ, which may then roam through $\mathscr{E}$ rather than $\mathscr{D}$.

1.3.2.3.8. *Convergence of distributions*

A *sequence* (T_j) of distributions will be said to converge in $\mathscr{D}'$ to a distribution T as $j \to \infty$ if, for any given $\varphi \in \mathscr{D}$, the sequence of complex numbers $(\langle T_j, \varphi\rangle)$ converges in $\mathbb{C}$ to the complex number $\langle T, \varphi\rangle$.

A *series* $\sum_{j=0}^{\infty} T_j$ of distributions will be said to converge in $\mathscr{D}'$ and to have distribution S as its sum if the sequence of partial sums $S_k = \sum_{j=0}^{k}$ converges to S.

These definitions of convergence in $\mathscr{D}'$ assume that the limits T and S are known in advance, and are distributions. This raises the question of the *completeness* of $\mathscr{D}'$: if a sequence (T_j) in $\mathscr{D}'$ is such that the sequence $(\langle T_j, \varphi\rangle)$ has a limit in $\mathbb{C}$ for all $\varphi \in \mathscr{D}$, does the map

$$\varphi \longmapsto \lim_{j\to\infty} \langle T_j, \varphi\rangle$$

define a distribution $T \in \mathscr{D}'$? In other words, does the limiting process preserve continuity with respect to φ? It is a remarkable theorem that, because of the strong topology on $\mathscr{D}$, this is actually the case. An analogous statement holds for series. This notion of convergence does not coincide with any of the classical notions used for ordinary functions: for example, the sequence (φ_ν) with $\varphi_\nu(x) = \cos \nu x$ converges to 0 in $\mathscr{D}'(\mathbb{R})$, but fails to do so by any of the standard criteria.

An example of convergent sequences of distributions is provided by sequences which converge to δ. If (f_ν) is a sequence of locally summable functions on $\mathbb{R}^n$ such that

(i) $\int_{\|\mathbf{x}\|< b} f_\nu(\mathbf{x})\, {\rm d}^n\mathbf{x} \to 1$ as $\nu \to \infty$ for all $b > 0$;

(ii) $\int_{a\le\|\mathbf{x}\|\le 1/a} |f_\nu(\mathbf{x})|\, {\rm d}^n\mathbf{x} \to 0$ as $\nu \to \infty$ for all $0 < a < 1$;

(iii) there exists $d > 0$ and $M > 0$ such that $\int_{\|\mathbf{x}\|< d} |f_\nu(\mathbf{x})|\, {\rm d}^n\mathbf{x} < M$ for all ν;

then the sequence (T_{f_ν}) of distributions converges to δ in $\mathscr{D}'(\mathbb{R}^n)$.

1.3.2.3.9. *Operations on distributions*

As a general rule, the definitions are chosen so that the operations coincide with those on functions whenever a distribution is associated to a function.

Most definitions consist in transferring to a distribution T an operation which is well defined on $\varphi \in \mathscr{D}$ by 'transposing' it in the duality product $\langle T, \varphi\rangle$; this procedure will map T to a new distribution provided the original operation maps $\mathscr{D}$ *continuously* into itself.

1.3.2.3.9.1. *Differentiation*

(*a*) *Definition and elementary properties*

If T is a distribution on $\mathbb{R}^n$, its partial derivative $\partial_i T$ with respect to x_i is defined by

$$\langle \partial_i T, \varphi\rangle = -\langle T, \partial_i \varphi\rangle$$

for all $\varphi \in \mathscr{D}$. This does define a distribution, because the partial differentiations $\varphi \longmapsto \partial_i\varphi$ are continuous for the topology of $\mathscr{D}$.

Suppose that $T = T_f$ with f a locally integrable function such that $\partial_i f$ exists and is almost everywhere continuous. Then integration by parts along the x_i axis gives

$$\begin{aligned}\int_{\mathbb{R}^n} &\partial_i f(x_1, \ldots, x_i, \ldots, x_n)\varphi(x_1, \ldots, x_i, \ldots, x_n)\, {\rm d}x_i \\ &= (f\varphi)(x_1, \ldots, +\infty, \ldots, x_n) - (f\varphi)(x_1, \ldots, -\infty, \ldots, x_n) \\ &\quad - \int_{\mathbb{R}^n} f(x_1, \ldots, x_i, \ldots, x_n)\partial_i\varphi(x_1, \ldots, x_i, \ldots, x_n)\, {\rm d}x_i;\end{aligned}$$

the integrated term vanishes, since φ has compact support, showing that $\partial_i T_f = T_{\partial_i f}$.

The test functions $\varphi \in \mathscr{D}$ are infinitely differentiable. Therefore, transpositions like that used to define $\partial_i T$ may be repeated, so that *any distribution is infinitely differentiable*. For instance,

$$\langle \partial^2_{ij} T, \varphi\rangle = -\langle \partial_j T, \partial_i\varphi\rangle = \langle T, \partial^2_{ij}\varphi\rangle,$$

$$\langle D^{\mathbf{p}} T, \varphi\rangle = (-1)^{|\mathbf{p}|}\langle T, D^{\mathbf{p}}\varphi\rangle,$$

$$\langle \Delta T, \varphi\rangle = \langle T, \Delta\varphi\rangle,$$

where Δ is the Laplacian operator. The derivatives of Dirac's δ distribution are

$$\langle D^{\mathbf{p}}\delta, \varphi\rangle = (-1)^{|\mathbf{p}|}\langle \delta, D^{\mathbf{p}}\varphi\rangle = (-1)^{|\mathbf{p}|} D^{\mathbf{p}}\varphi(\mathbf{0}).$$

It is remarkable that *differentiation is a continuous operation* for the topology on $\mathscr{D}'$: if a sequence (T_j) of distributions converges to distribution T, then the sequence $(D^{\mathbf{p}} T_j)$ of derivatives converges to $D^{\mathbf{p}} T$ for any multi-index $\mathbf{p}$, since as $j \to \infty$

$$\langle D^{\mathbf{p}} T_j, \varphi\rangle = (-1)^{|\mathbf{p}|}\langle T_j, D^{\mathbf{p}}\varphi\rangle \to (-1)^{|\mathbf{p}|}\langle T, D^{\mathbf{p}}\varphi\rangle = \langle D^{\mathbf{p}} T, \varphi\rangle.$$

An analogous statement holds for series: any convergent series of distributions may be differentiated termwise to all orders. This illustrates how 'robust' the constructs of distribution theory are in comparison with those of ordinary function theory, where similar statements are notoriously untrue.

(*b*) *Differentiation under the duality bracket*

Limiting processes and differentiation may also be carried out under the duality bracket $\langle , \rangle$ as under the integral sign with ordinary functions. Let the function $\varphi = \varphi(\mathbf{x}, \lambda)$ depend on a parameter $\lambda \in \Lambda$ and a vector $\mathbf{x} \in \mathbb{R}^n$ in such a way that all functions

$$\varphi_\lambda : \mathbf{x} \longmapsto \varphi(\mathbf{x}, \lambda)$$

be in $\mathscr{D}(\mathbb{R}^n)$ for all $\lambda \in \Lambda$. Let $T \in \mathscr{D}'(\mathbb{R}^n)$ be a distribution, let

$$I(\lambda) = \langle T, \varphi_\lambda\rangle$$

and let $\lambda_0 \in \Lambda$ be given parameter value. Suppose that, as λ runs through a small enough neighbourhood of λ_0,

(i) all the φ_λ have their supports in a fixed compact subset K of $\mathbb{R}^n$;

(ii) all the derivatives $D^{\mathbf{p}}\varphi_\lambda$ have a partial derivative with respect to λ which is continuous with respect to $\mathbf{x}$ and λ.

Under these hypotheses, $I(\lambda)$ is differentiable (in the usual sense) with respect to λ near λ_0, and its derivative may be obtained by 'differentiation under the $\langle , \rangle$ sign':

$$\frac{\mathrm{d}I}{\mathrm{d}\lambda} = \langle T, \partial_\lambda \varphi_\lambda \rangle.$$

(*c*) *Effect of discontinuities*

When a function *f* or its derivatives are no longer continuous, the derivatives $D^{\mathbf{p}}T_f$ of the associated distribution T_f may no longer coincide with the distributions associated to the functions $D^{\mathbf{p}}f$.

In dimension 1, the simplest example is Heaviside's unit step function Y [$Y(x) = 0$ for $x < 0, Y(x) = 1$ for $x \geq 0$]:

$$\langle (T_Y)', \varphi \rangle = -\langle (T_Y), \varphi' \rangle = -\int\limits_0^{+\infty} \varphi'(x)\,\mathrm{d}x = \varphi(0) = \langle \delta, \varphi \rangle.$$

Hence $(T_Y)' = \delta$, a result long used 'heuristically' by electrical engineers [see also Dirac (1958)].

Let *f* be infinitely differentiable for $x < 0$ and $x > 0$ but have discontinuous derivatives $f^{(m)}$ at $x = 0$ [$f^{(0)}$ being *f* itself] with jumps $\sigma_m = f^{(m)}(0+) - f^{(m)}(0-)$. Consider the functions:

$$\begin{aligned} g_0 &= f - \sigma_0 Y \\ g_1 &= g_0' - \sigma_1 Y \\ &- - - - - - - - - - \\ g_k &= g_{k-1}' - \sigma_k Y. \end{aligned}$$

The g_k are continuous, their derivatives g_k' are continuous almost everywhere [which implies that $(T_{g_k})' = T_{g_k'}$ and $g_k' = f^{(k+1)}$ almost everywhere]. This yields immediately:

$$\begin{aligned} (T_f)' &= T_{f'} + \sigma_0 \delta \\ (T_f)'' &= T_{f''} + \sigma_0 \delta' + \sigma_1 \delta \\ &- - - - - - - - - - - - - - - - - - \\ (T_f)^{(m)} &= T_{f^{(m)}} + \sigma_0 \delta^{(m-1)} + \ldots + \sigma_{m-1}\delta. \\ &- - - - - - - - - - - - - - - - - - \end{aligned}$$

Thus the 'distributional derivatives' $(T_f)^{(m)}$ differ from the usual functional derivatives $T_{f^{(m)}}$ by singular terms associated with discontinuities.

In dimension n, let f be infinitely differentiable everywhere except on a smooth hypersurface S, across which its partial derivatives show discontinuities. Let σ_0 and σ_ν denote the discontinuities of f and its normal derivative $\partial_\nu \varphi$ across S (both σ_0 and σ_ν are functions of position on S), and let $\delta_{(S)}$ and $\partial_\nu \delta_{(S)}$ be defined by

$$\begin{aligned} \langle \delta_{(S)}, \varphi \rangle &= \int\limits_S \varphi \,\mathrm{d}^{n-1}S \\ \langle \partial_\nu \delta_{(S)}, \varphi \rangle &= -\int\limits_S \partial_\nu \varphi \,\mathrm{d}^{n-1}S. \end{aligned}$$

Integration by parts shows that

$$\partial_i T_f = T_{\partial_i f} + \sigma_0 \cos\theta_i \delta_{(S)},$$

where θ_i is the angle between the x_i axis and the normal to S along which the jump σ_0 occurs, and that the Laplacian of T_f is given by

$$\Delta(T_f) = T_{\Delta f} + \sigma_\nu \delta_{(S)} + \partial_\nu[\sigma_0 \delta_{(S)}].$$

The latter result is a statement of Green's theorem in terms of distributions. It will be used in Section 1.3.4.4.3.5 to calculate the Fourier transform of the indicator function of a molecular envelope.

1.3.2.3.9.2. *Integration of distributions in dimension 1*

The reverse operation from differentiation, namely calculating the 'indefinite integral' of a distribution *S*, consists in finding a distribution *T* such that $T' = S$.

For all $\chi \in \mathscr{D}$ such that $\chi = \psi'$ with $\psi \in \mathscr{D}$, we must have

$$\langle T, \chi \rangle = -\langle S, \psi \rangle.$$

This condition defines *T* in a 'hyperplane' $\mathscr{H}$ of $\mathscr{D}$, whose equation

$$\langle 1, \chi \rangle \equiv \langle 1, \psi' \rangle = 0$$

reflects the fact that ψ has compact support.

To specify *T* in the whole of $\mathscr{D}$, it suffices to specify the value of $\langle T, \varphi_0 \rangle$ where $\varphi_0 \in \mathscr{D}$ is such that $\langle 1, \varphi_0 \rangle = 1$: then any $\varphi \in \mathscr{D}$ may be written uniquely as

$$\varphi = \lambda \varphi_0 + \psi'$$

with

$$\lambda = \langle 1, \varphi \rangle, \qquad \chi = \varphi - \lambda\varphi_0, \qquad \psi(x) = \int\limits_0^x \chi(t)\,\mathrm{d}t,$$

and *T* is defined by

$$\langle T, \varphi \rangle = \lambda \langle T, \varphi_0 \rangle - \langle S, \psi \rangle.$$

The freedom in the choice of φ_0 means that *T* is defined up to an additive constant.

1.3.2.3.9.3. *Multiplication of distributions by functions*

The product αT of a distribution *T* on $\mathbb{R}^n$ by a function α over $\mathbb{R}^n$ will be defined by transposition:

$$\langle \alpha T, \varphi \rangle = \langle T, \alpha\varphi \rangle \quad \text{for all } \varphi \in \mathscr{D}.$$

In order that αT be a distribution, the mapping $\varphi \longmapsto \alpha\varphi$ must send $\mathscr{D}(\mathbb{R}^n)$ continuously into itself; hence *the multipliers* α *must be infinitely differentiable*. The product of two general distributions cannot be defined. The need for a careful treatment of multipliers of distributions will become clear when it is later shown (Section 1.3.2.5.8) that the Fourier transformation turns convolutions into multiplications and *vice versa*.

If *T* is a distribution of order *m*, then α needs only have continuous derivatives up to order *m*. For instance, δ is a distribution of order zero, and $\alpha\delta = \alpha(\mathbf{0})\delta$ is a distribution provided α is continuous; this relation is of fundamental importance in the theory of sampling and of the properties of the Fourier transformation related to sampling (Sections 1.3.2.6.4, 1.3.2.6.6). More generally, $D^{\mathbf{p}}\delta$ is a distribution of order $|\mathbf{p}|$, and the following formula holds for all $\alpha \in \mathscr{D}^{(m)}$ with $m = |\mathbf{p}|$:

$$\alpha(D^{\mathbf{p}}\delta) = \sum_{\mathbf{q} \leq \mathbf{p}} (-1)^{|\mathbf{p}-\mathbf{q}|} \binom{\mathbf{p}}{\mathbf{q}} (D^{\mathbf{p}-\mathbf{q}}\alpha)(\mathbf{0}) D^{\mathbf{q}}\delta.$$

The derivative of a product is easily shown to be

$$\partial_i(\alpha T) = (\partial_i \alpha)T + \alpha(\partial_i T)$$

and generally for any multi-index $\mathbf{p}$

$$D^{\mathbf{p}}(\alpha T) = \sum_{\mathbf{q} \leq \mathbf{p}} \binom{\mathbf{p}}{\mathbf{q}} (D^{\mathbf{p}-\mathbf{q}}\alpha)(\mathbf{0}) D^{\mathbf{q}} T.$$

1.3.2.3.9.4. *Division of distributions by functions*

Given a distribution S on $\mathbb{R}^n$ and an infinitely differentiable multiplier function α, the division problem consists in finding a distribution T such that $\alpha T = S$.

If α never vanishes, $T = S/\alpha$ is the unique answer. If $n = 1$, and if α has only isolated zeros of finite order, it can be reduced to a collection of cases where the multiplier is x^m, for which the general solution can be shown to be of the form

$$T = U + \sum_{i=0}^{m-1} c_i \delta^{(i)},$$

where U is a particular solution of the division problem $x^m U = S$ and the c_i are arbitrary constants.

In dimension $n > 1$, the problem is much more difficult, but is of fundamental importance in the theory of linear partial differential equations, since the Fourier transformation turns the problem of solving these into a division problem for distributions [see Hörmander (1963)].

1.3.2.3.9.5. *Transformation of coordinates*

Let σ be a smooth non-singular change of variables in $\mathbb{R}^n$, *i.e.* an infinitely differentiable mapping from an open subset Ω of $\mathbb{R}^n$ to Ω' in $\mathbb{R}^n$, whose Jacobian

$$J(\sigma) = \det \left[\frac{\partial \sigma(\mathbf{x})}{\partial \mathbf{x}}\right]$$

vanishes nowhere in Ω. By the implicit function theorem, the inverse mapping σ^{-1} from Ω' to Ω is well defined.

If f is a locally summable function on Ω, then the function $\sigma^{\#} f$ defined by

$$(\sigma^{\#} f)(\mathbf{x}) = f[\sigma^{-1}(\mathbf{x})]$$

is a locally summable function on Ω', and for any $\varphi \in \mathscr{D}(\Omega')$ we may write:

$$\begin{aligned}\int_{\Omega'} (\sigma^{\#} f)(\mathbf{x})\varphi(\mathbf{x}) \,\mathrm{d}^n\mathbf{x} &= \int_{\Omega'} f[\sigma^{-1}(\mathbf{x})]\varphi(\mathbf{x}) \,\mathrm{d}^n\mathbf{x}\\ &= \int_{\Omega'} f(\mathbf{y})\varphi[\sigma(\mathbf{y})]|J(\sigma)| \,\mathrm{d}^n\mathbf{y} \quad \text{by } \mathbf{x} = \sigma(\mathbf{y}).\end{aligned}$$

In terms of the associated distributions

$$\langle T_{\sigma^{\#} f}, \varphi\rangle = \langle T_f, |J(\sigma)|(\sigma^{-1})^{\#}\varphi\rangle.$$

This operation can be extended to an arbitrary distribution T by defining its *image* $\sigma^{\#} T$ under coordinate transformation σ through

$$\langle \sigma^{\#} T, \varphi\rangle = \langle T, |J(\sigma)|(\sigma^{-1})^{\#}\varphi\rangle,$$

which is well defined provided that σ is *proper*, *i.e.* that $\sigma^{-1}(K)$ is compact whenever K is compact.

For instance, if $\sigma : \mathbf{x} \longmapsto \mathbf{x} + \mathbf{a}$ is a *translation* by a vector $\mathbf{a}$ in $\mathbb{R}^n$, then $|J(\sigma)| = 1$; $\sigma^{\#}$ is denoted by $\tau_{\mathbf{a}}$, and the translate $\tau_{\mathbf{a}} T$ of a distribution T is defined by

$$\langle \tau_{\mathbf{a}} T, \varphi\rangle = \langle T, \tau_{-\mathbf{a}}\varphi\rangle.$$

Let $A : \mathbf{x} \longmapsto \mathbf{A}\mathbf{x}$ be a linear transformation defined by a non-singular matrix $\mathbf{A}$. Then $J(A) = \det \mathbf{A}$, and

$$\langle A^{\#} T, \varphi\rangle = |\det \mathbf{A}|\langle T, (A^{-1})^{\#}\varphi\rangle.$$

This formula will be shown later (Sections 1.3.2.6.5, 1.3.4.2.1.1) to be the basis for the definition of the reciprocal lattice.

In particular, if $\mathbf{A} = -\mathbf{I}$, where $\mathbf{I}$ is the identity matrix, A is an inversion through a centre of symmetry at the origin, and denoting $A^{\#}\varphi$ by $\breve{\varphi}$ we have:

$$\langle \breve{T}, \varphi\rangle = \langle T, \breve{\varphi}\rangle.$$

T is called an even distribution if $\breve{T} = T$, an odd distribution if $\breve{T} = -T$.

If $\mathbf{A} = \lambda\mathbf{I}$ with $\lambda > 0$, A is called a *dilation* and

$$\langle A^{\#} T, \varphi\rangle = \lambda^n \langle T, (A^{-1})^{\#}\varphi\rangle.$$

Writing symbolically δ as $\delta(\mathbf{x})$ and $A^{\#}\delta$ as $\delta(\mathbf{x}/\lambda)$, we have:

$$\delta(\mathbf{x}/\lambda) = \lambda^n \delta(\mathbf{x}).$$

If $n = 1$ and f is a function with isolated simple zeros x_j, then in the same symbolic notation

$$\delta[f(x)] = \sum_j \frac{1}{|f'(x_j)|}\delta(x_j),$$

where each $\lambda_j = 1/|f'(x_j)|$ is analogous to a 'Lorentz factor' at zero x_j.

1.3.2.3.9.6. *Tensor product of distributions*

The purpose of this construction is to extend Fubini's theorem to distributions. Following Section 1.3.2.2.5, we may define the tensor product $L^1_{\text{loc}}(\mathbb{R}^m) \otimes L^1_{\text{loc}}(\mathbb{R}^n)$ as the vector space of finite linear combinations of functions of the form

$$f \otimes g : (\mathbf{x}, \mathbf{y}) \longmapsto f(\mathbf{x})g(\mathbf{y}),$$

where $\mathbf{x} \in \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^n$, $f \in L^1_{\text{loc}}(\mathbb{R}^m)$ and $g \in L^1_{\text{loc}}(\mathbb{R}^n)$.

Let $S_{\mathbf{x}}$ and $T_{\mathbf{y}}$ denote the distributions associated to f and g, respectively, the subscripts $\mathbf{x}$ and $\mathbf{y}$ acting as mnemonics for $\mathbb{R}^m$ and $\mathbb{R}^n$. It follows from Fubini's theorem (Section 1.3.2.2.5) that $f \otimes g \in L^1_{\text{loc}}(\mathbb{R}^m \times \mathbb{R}^n)$, and hence defines a distribution over $\mathbb{R}^m \times \mathbb{R}^n$; the rearrangement of integral signs gives

$$\langle S_{\mathbf{x}} \otimes T_{\mathbf{y}}, \varphi_{\mathbf{x},\mathbf{y}}\rangle = \langle S_{\mathbf{x}}, \langle T_{\mathbf{y}}, \varphi_{\mathbf{x},\mathbf{y}}\rangle\rangle = \langle T_{\mathbf{y}}, \langle S_{\mathbf{x}}, \varphi_{\mathbf{x},\mathbf{y}}\rangle\rangle$$

for all $\varphi_{\mathbf{x},\mathbf{y}} \in \mathscr{D}(\mathbb{R}^m \times \mathbb{R}^n)$. In particular, if $\varphi(\mathbf{x}, \mathbf{y}) = u(\mathbf{x})v(\mathbf{y})$ with $u \in \mathscr{D}(\mathbb{R}^m)$, $v \in \mathscr{D}(\mathbb{R}^n)$, then

$$\langle S \otimes T, u \otimes v\rangle = \langle S, u\rangle\langle T, v\rangle.$$

This construction can be extended to general distributions $S \in \mathscr{D}'(\mathbb{R}^m)$ and $T \in \mathscr{D}'(\mathbb{R}^n)$. Given any test function $\varphi \in \mathscr{D}(\mathbb{R}^m \times \mathbb{R}^n)$, let $\varphi_{\mathbf{x}}$ denote the map $\mathbf{y} \longmapsto \varphi(\mathbf{x}, \mathbf{y})$; let $\varphi_{\mathbf{y}}$ denote the map $\mathbf{x} \longmapsto \varphi(\mathbf{x}, \mathbf{y})$; and define the two functions $\theta(\mathbf{x}) = \langle T, \varphi_{\mathbf{x}}\rangle$ and $\omega(\mathbf{y}) = \langle S, \varphi_{\mathbf{y}}\rangle$. Then, by the lemma on differentiation under the $\langle , \rangle$ sign of Section 1.3.2.3.9.1, $\theta \in \mathscr{D}(\mathbb{R}^m)$, $\omega \in \mathscr{D}(\mathbb{R}^n)$, and there exists a unique distribution $S \otimes T$ such that

$$\langle S \otimes T, \varphi\rangle = \langle S, \theta\rangle = \langle T, \omega\rangle.$$

$S \otimes T$ is called the *tensor product* of S and T.

With the mnemonic introduced above, this definition reads identically to that given above for distributions associated to locally integrable functions:

$$\langle S_{\mathbf{x}} \otimes T_{\mathbf{y}}, \varphi_{\mathbf{x},\mathbf{y}}\rangle = \langle S_{\mathbf{x}}, \langle T_{\mathbf{y}}, \varphi_{\mathbf{x},\mathbf{y}}\rangle\rangle = \langle T_{\mathbf{y}}, \langle S_{\mathbf{x}}, \varphi_{\mathbf{x},\mathbf{y}}\rangle\rangle.$$

The tensor product of distributions is associative:

$$(R \otimes S) \otimes T = R \otimes (S \otimes T).$$

Derivatives may be calculated by

$$D^{\mathbf{p}}_{\mathbf{x}} D^{\mathbf{q}}_{\mathbf{y}} (S_{\mathbf{x}} \otimes T_{\mathbf{y}}) = (D^{\mathbf{p}}_{\mathbf{x}} S_{\mathbf{x}}) \otimes (D^{\mathbf{q}}_{\mathbf{y}} T_{\mathbf{y}}).$$

The support of a tensor product is the Cartesian product of the supports of the two factors.

1.3.2.3.9.7. *Convolution of distributions*

The convolution $f * g$ of two functions f and g on $\mathbb{R}^n$ is defined by

$$(f * g)(\mathbf{x}) = \int_{\mathbb{R}^n} f(\mathbf{y})g(\mathbf{x} - \mathbf{y}) \,\mathrm{d}^n\mathbf{y} = \int_{\mathbb{R}^n} f(\mathbf{x} - \mathbf{y})g(\mathbf{y}) \,\mathrm{d}^n\mathbf{y}$$

whenever the integral exists. This is the case when f and g are both in $L^1(\mathbb{R}^n)$; then $f * g$ is also in $L^1(\mathbb{R}^n)$. Let S, T and W denote the distributions associated to f, g and $f * g$, respectively: a change of variable immediately shows that for any $\varphi \in \mathscr{D}(\mathbb{R}^n)$,

$$\langle W, \varphi \rangle = \int\limits_{\mathbb{R}^n \times \mathbb{R}^n} f(\mathbf{x}) g(\mathbf{y}) \varphi(\mathbf{x} + \mathbf{y}) \,\mathrm{d}^n\mathbf{x} \,\mathrm{d}^n\mathbf{y}.$$

Introducing the map σ from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}^n$ defined by $\sigma(\mathbf{x}, \mathbf{y}) = \mathbf{x} + \mathbf{y}$, the latter expression may be written:

$$\langle S_{\mathbf{x}} \otimes T_{\mathbf{y}}, \varphi \circ \sigma \rangle$$

(where $\circ$ denotes the composition of mappings) or by a slight abuse of notation:

$$\langle W, \varphi \rangle = \langle S_{\mathbf{x}} \otimes T_{\mathbf{y}}, \varphi(\mathbf{x} + \mathbf{y}) \rangle.$$

A difficulty arises in extending this definition to general distributions S and T because the mapping σ is not proper: if K is compact in $\mathbb{R}^n$, then $\sigma^{-1}(K)$ is a cylinder with base K and generator the 'second bisector' $\mathbf{x} + \mathbf{y} = \mathbf{0}$ in $\mathbb{R}^n \times \mathbb{R}^n$. However, $\langle S \otimes T, \varphi \circ \sigma \rangle$ is defined whenever the intersection between Supp $(S \otimes T) =$ (Supp S) $\times$ (Supp T) and σ^{-1}(Supp φ) is compact.

We may therefore define the *convolution* $S * T$ of two distributions S and T on $\mathbb{R}^n$ by

$$\langle S * T, \varphi \rangle = \langle S \otimes T, \varphi \circ \sigma \rangle = \langle S_{\mathbf{x}} \otimes T_{\mathbf{y}}, \varphi(\mathbf{x} + \mathbf{y}) \rangle$$

whenever the following *support condition* is fulfilled:

'the set $\{(\mathbf{x}, \mathbf{y}) | \mathbf{x} \in A, \mathbf{y} \in B, \mathbf{x} + \mathbf{y} \in K\}$ is compact in $\mathbb{R}^n \times \mathbb{R}^n$ for all K compact in $\mathbb{R}^n$'.

The latter condition is met, in particular, if S or T has compact support. The support of $S * T$ is easily seen to be contained in the closure of the vector sum

$$A + B = \{\mathbf{x} + \mathbf{y} | \mathbf{x} \in A, \mathbf{y} \in B\}.$$

Convolution by a fixed distribution S is a *continuous* operation for the topology on $\mathscr{D}'$: it maps convergent sequences (T_j) to convergent sequences $(S * T_j)$. Convolution is commutative: $S * T = T * S$.

The convolution of p distributions $T_1, \ldots, T_p$ with supports $A_1, \ldots, A_p$ can be defined by

$$\langle T_1 * \ldots * T_p, \varphi \rangle = \langle (T_1)_{\mathbf{x}_1} \otimes \ldots \otimes (T_p)_{\mathbf{x}_p}, \varphi(\mathbf{x}_1 + \ldots + \mathbf{x}_p) \rangle$$

whenever the following generalized support condition:

'the set $\{(\mathbf{x}_1, \ldots, \mathbf{x}_p) | \mathbf{x}_1 \in A_1, \ldots, \mathbf{x}_p \in A_p, \mathbf{x}_1 + \ldots + \mathbf{x}_p \in K\}$ is compact in $(\mathbb{R}^n)^p$ for all K compact in $\mathbb{R}^n$'

is satisfied. It is then associative. Interesting examples of associativity failure, which can be traced back to violations of the support condition, may be found in Bracewell (1986, pp. 436–437).

It follows from previous definitions that, for all distributions $T \in \mathscr{D}'$, the following identities hold:

(i) $\delta * T = T$: δ is the unit convolution;

(ii) $\delta_{(\mathbf{a})} * T = \tau_{\mathbf{a}} T$: translation is a convolution with the corresponding translate of δ;

(iii) $(D^{\mathbf{p}} \delta) * T = D^{\mathbf{p}} T$: differentiation is a convolution with the corresponding derivative of δ;

(iv) translates or derivatives of a convolution may be obtained by translating or differentiating any one of the factors: convolution 'commutes' with translation and differentiation, a property used in Section 1.3.4.4.7.7 to speed up least-squares model refinement for macromolecules.

The latter property is frequently used for the purpose of *regularization*: if T is a distribution, α an infinitely differentiable function, and at least one of the two has compact support, then $T * \alpha$ is an infinitely differentiable ordinary function. Since sequences (α_ν) of such functions α can be constructed which have compact support and converge to δ, it follows that any distribution T can be obtained as the limit of infinitely differentiable functions $T * \alpha_\nu$. In topological jargon: $\mathscr{D}(\mathbb{R}^n)$ is 'everywhere dense' in $\mathscr{D}'(\mathbb{R}^n)$. A standard function in $\mathscr{D}$ which is often used for such proofs is defined as follows: put

$$\begin{aligned} \theta(x) &= \frac{1}{A} \exp\left(-\frac{1}{1 - x^2}\right) \quad &&\text{for } |x| \leq 1, \\ &= 0 &&\text{for } |x| \geq 1, \end{aligned}$$

with

$$A = \int\limits_{-1}^{+1} \exp\left(-\frac{1}{1 - x^2}\right) \mathrm{d}x$$

(so that θ is in $\mathscr{D}$ and is normalized), and put

$$\begin{aligned} \theta_\varepsilon(x) &= \frac{1}{\varepsilon} \theta\left(\frac{x}{\varepsilon}\right) \quad &&\text{in dimension 1,} \\ \theta_\varepsilon(\mathbf{x}) &= \prod_{j=1}^{n} \theta_\varepsilon(x_j) &&\text{in dimension } n. \end{aligned}$$

Another related result, also proved by convolution, is the *structure theorem*: the restriction of a distribution $T \in \mathscr{D}'(\mathbb{R}^n)$ to a bounded open set Ω in $\mathbb{R}^n$ is a derivative of finite order of a continuous function.

Properties (i) to (iv) are the basis of the symbolic or operational calculus (see Carslaw & Jaeger, 1948; Van der Pol & Bremmer, 1955; Churchill, 1958; Erdélyi, 1962; Moore, 1971) for solving integro-differential equations with constant coefficients by turning them into convolution equations, then using factorization methods for convolution algebras (Schwartz, 1965).

1.3.2.4. *Fourier transforms of functions*

1.3.2.4.1. *Introduction*

Given a complex-valued function f on $\mathbb{R}^n$ subject to suitable regularity conditions, its Fourier transform $\mathscr{F}[f]$ and Fourier cotransform $\bar{\mathscr{F}}[f]$ are defined as follows:

$$\begin{aligned} \mathscr{F}[f](\xi) &= \int\limits_{\mathbb{R}^n} f(\mathbf{x}) \exp(-2\pi i \boldsymbol{\xi} \cdot \mathbf{x}) \,\mathrm{d}^n\mathbf{x} \\ \bar{\mathscr{F}}[f](\xi) &= \int\limits_{\mathbb{R}^n} f(\mathbf{x}) \exp(+2\pi i \boldsymbol{\xi} \cdot \mathbf{x}) \,\mathrm{d}^n\mathbf{x}, \end{aligned}$$

where $\boldsymbol{\xi} \cdot \mathbf{x} = \sum_{i=1}^n \xi_i x_i$ is the ordinary scalar product. The terminology and sign conventions given above are the standard ones in mathematics; those used in crystallography are slightly different (see Section 1.3.4.2.1.1). These transforms enjoy a number of remarkable properties, whose natural settings entail different regularity assumptions on f: for instance, properties relating to convolution are best treated in $L^1(\mathbb{R}^n)$, while Parseval's theorem requires the Hilbert space structure of $L^2(\mathbb{R}^n)$. After a brief review of these classical properties, the Fourier transformation will be examined in a space $\mathscr{S}(\mathbb{R}^n)$ particularly well suited to accommodating the full range of its properties, which will later serve as a space of test functions to extend the Fourier transformation to distributions.

There exists an abundant literature on the 'Fourier integral'. The books by Carslaw (1930), Wiener (1933), Titchmarsh (1948), Katznelson (1968), Sneddon (1951, 1972), and Dym & McKean (1972) are particularly recommended.