

### 1. GENERAL RELATIONSHIPS AND TECHNIQUES

The integration of a general function over  $\mathbb{R}^m \times \mathbb{R}^n$  may be accomplished in two steps according to *Fubini's theorem*. Given  $F \in L^1(\mathbb{R}^m \times \mathbb{R}^n)$ , the functions

$$F_1 : \mathbf{x} \mapsto \int_{\mathbb{R}^n} F(\mathbf{x}, \mathbf{y}) \, d^n \mathbf{y}$$

$$F_2 : \mathbf{y} \mapsto \int_{\mathbb{R}^m} F(\mathbf{x}, \mathbf{y}) \, d^m \mathbf{x}$$

exist for almost all  $\mathbf{x} \in \mathbb{R}^m$  and almost all  $\mathbf{y} \in \mathbb{R}^n$ , respectively, are integrable, and

$$\int_{\mathbb{R}^m \times \mathbb{R}^n} F(\mathbf{x}, \mathbf{y}) \, d^m \mathbf{x} \, d^n \mathbf{y} = \int_{\mathbb{R}^m} F_1(\mathbf{x}) \, d^m \mathbf{x} = \int_{\mathbb{R}^n} F_2(\mathbf{y}) \, d^n \mathbf{y}.$$

Conversely, if any one of the integrals

$$(i) \quad \int_{\mathbb{R}^m \times \mathbb{R}^n} |F(\mathbf{x}, \mathbf{y})| \, d^m \mathbf{x} \, d^n \mathbf{y}$$

$$(ii) \quad \int_{\mathbb{R}^m} \left( \int_{\mathbb{R}^n} |F(\mathbf{x}, \mathbf{y})| \, d^n \mathbf{y} \right) \, d^m \mathbf{x}$$

$$(iii) \quad \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^m} |F(\mathbf{x}, \mathbf{y})| \, d^m \mathbf{x} \right) \, d^n \mathbf{y}$$

is finite, then so are the other two, and the identity above holds. It is then (and only then) permissible to change the order of integrations.

Fubini's theorem is of fundamental importance in the study of tensor products and convolutions of distributions.

#### 1.3.2.2.6. Topology in function spaces

Geometric intuition, which often makes 'obvious' the topological properties of the real line and of ordinary space, cannot be relied upon in the study of function spaces: the latter are infinite-dimensional, and several inequivalent notions of convergence may exist. A careful analysis of topological concepts and of their interrelationship is thus a necessary prerequisite to the study of these spaces. The reader may consult Dieudonné (1969, 1970), Friedman (1970), Trèves (1967) and Yosida (1965) for detailed expositions.

##### 1.3.2.2.6.1. General topology

Most topological notions are first encountered in the setting of *metric spaces*. A metric space  $E$  is a set equipped with a *distance function*  $d$  from  $E \times E$  to the non-negative reals which satisfies:

- (i)  $d(x, y) = d(y, x) \quad \forall x, y \in E$  (symmetry);
- (ii)  $d(x, y) = 0$  iff  $x = y$  (separation);
- (iii)  $d(x, z) \leq d(x, y) + d(y, z) \quad \forall x, y, z \in E$  (triangular inequality).

By means of  $d$ , the following notions can be defined: open balls, neighbourhoods; open and closed sets, interior and closure; convergence of sequences, continuity of mappings; Cauchy sequences and completeness; compactness; connectedness. They suffice for the investigation of a great number of questions in analysis and geometry (see e.g. Dieudonné, 1969).

Many of these notions turn out to depend only on the properties of the collection  $\mathcal{O}(E)$  of open subsets of  $E$ : two distance functions leading to the same  $\mathcal{O}(E)$  lead to identical topological properties. An axiomatic reformulation of topological notions is thus possible: a *topology* in  $E$  is a collection  $\mathcal{O}(E)$  of subsets of  $E$  which satisfy suitable axioms and are deemed open irrespective of the way they are obtained. From the practical standpoint, however, a topology which can be obtained from a distance function (called a *metrizable topology*) has the very useful property that *the notions of closure,*

*limit and continuity may be defined by means of sequences.* For non-metrizable topologies, these notions are much more difficult to handle, requiring the use of 'filters' instead of sequences.

In some spaces  $E$ , a topology may be most naturally defined by a family of *pseudo-distances*  $(d_\alpha)_{\alpha \in A}$ , where each  $d_\alpha$  satisfies (i) and (iii) but not (ii). Such spaces are called *uniformizable*. If for every pair  $(x, y) \in E \times E$  there exists  $\alpha \in A$  such that  $d_\alpha(x, y) \neq 0$ , then the separation property can be recovered. If furthermore a *countable* subfamily of the  $d_\alpha$  suffices to define the topology of  $E$ , the latter can be shown to be *metrizable*, so that limiting processes in  $E$  may be studied by means of sequences.

##### 1.3.2.2.6.2. Topological vector spaces

The function spaces  $E$  of interest in Fourier analysis have an underlying vector space structure over the field  $\mathbb{C}$  of complex numbers. A topology on  $E$  is said to be *compatible* with a vector space structure on  $E$  if vector addition [*i.e.* the map  $(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} + \mathbf{y}$ ] and scalar multiplication [*i.e.* the map  $(\lambda, \mathbf{x}) \mapsto \lambda \mathbf{x}$ ] are both *continuous*;  $E$  is then called a *topological vector space*. Such a topology may be defined by specifying a 'fundamental system  $S$  of neighbourhoods of  $\mathbf{0}$ ', which can then be translated by vector addition to construct neighbourhoods of other points  $\mathbf{x} \neq \mathbf{0}$ .

A *norm*  $\nu$  on a vector space  $E$  is a non-negative real-valued function on  $E \times E$  such that

- (i')  $\nu(\lambda \mathbf{x}) = |\lambda| \nu(\mathbf{x})$  for all  $\lambda \in \mathbb{C}$  and  $\mathbf{x} \in E$ ;
- (ii')  $\nu(\mathbf{x}) = 0$  if and only if  $\mathbf{x} = \mathbf{0}$ ;
- (iii')  $\nu(\mathbf{x} + \mathbf{y}) \leq \nu(\mathbf{x}) + \nu(\mathbf{y})$  for all  $\mathbf{x}, \mathbf{y} \in E$ .

Subsets of  $E$  defined by conditions of the form  $\nu(\mathbf{x}) \leq r$  with  $r > 0$  form a fundamental system of neighbourhoods of  $\mathbf{0}$ . The corresponding topology makes  $E$  a *normed space*. This topology is *metrizable*, since it is equivalent to that derived from the translation-invariant distance  $d(\mathbf{x}, \mathbf{y}) = \nu(\mathbf{x} - \mathbf{y})$ . Normed spaces which are *complete*, *i.e.* in which all Cauchy sequences converge, are called *Banach spaces*; they constitute the natural setting for the study of differential calculus.

A *semi-norm*  $\sigma$  on a vector space  $E$  is a positive real-valued function on  $E \times E$  which satisfies (i') and (iii') but not (ii'). Given a set  $\Sigma$  of semi-norms on  $E$  such that any pair  $(\mathbf{x}, \mathbf{y})$  in  $E \times E$  is separated by at least one  $\sigma \in \Sigma$ , let  $B$  be the set of those subsets  $\Gamma_{\sigma, r}$  of  $E$  defined by a condition of the form  $\sigma(\mathbf{x}) \leq r$  with  $\sigma \in \Sigma$  and  $r > 0$ ; and let  $S$  be the set of finite intersections of elements of  $B$ . Then there exists a unique topology on  $E$  for which  $S$  is a fundamental system of neighbourhoods of  $\mathbf{0}$ . This topology is *uniformizable* since it is equivalent to that derived from the family of translation-invariant pseudo-distances  $(\mathbf{x}, \mathbf{y}) \mapsto \sigma(\mathbf{x} - \mathbf{y})$ . It is *metrizable* if and only if it can be constructed by the above procedure with  $\Sigma$  a *countable* set of semi-norms. If furthermore  $E$  is complete,  $E$  is called a *Fréchet space*.

If  $E$  is a topological vector space over  $\mathbb{C}$ , its *dual*  $E^*$  is the set of all linear mappings from  $E$  to  $\mathbb{C}$  (which are also called *linear forms*, or *linear functionals*, over  $E$ ). The subspace of  $E^*$  consisting of all linear forms which are *continuous* for the topology of  $E$  is called the *topological dual* of  $E$  and is denoted  $E'$ . If the topology on  $E$  is metrizable, then the continuity of a linear form  $T \in E'$  at  $f \in E$  can be ascertained by means of sequences, *i.e.* by checking that the sequence  $[T(f_j)]$  of complex numbers converges to  $T(f)$  in  $\mathbb{C}$  whenever the sequence  $(f_j)$  converges to  $f$  in  $E$ .

#### 1.3.2.3. Elements of the theory of distributions

##### 1.3.2.3.1. Origins

At the end of the 19th century, Heaviside proposed under the name of 'operational calculus' a set of rules for solving a class of

### 1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

differential, partial differential and integral equations encountered in electrical engineering (today's 'signal processing'). These rules worked remarkably well but were devoid of mathematical justification (see Whittaker, 1928). In 1926, Dirac introduced his famous  $\delta$ -function [see Dirac (1958), pp. 58–61], which was found to be related to Heaviside's constructs. Other singular objects, together with procedures to handle them, had already appeared in several branches of analysis [Cauchy's 'principal values'; Hadamard's 'finite parts' (Hadamard, 1932, 1952); Riesz's regularization methods for certain divergent integrals (Riesz, 1938, 1949)] as well as in the theories of Fourier series and integrals (see *e.g.* Bochner, 1932, 1959). Their very definition often verged on violating the rigorous rules governing limiting processes in analysis, so that subsequent recourse to limiting processes could lead to erroneous results; *ad hoc* precautions thus had to be observed to avoid mistakes in handling these objects.

In 1945–1950, Laurent Schwartz proposed his theory of distributions (see Schwartz, 1966), which provided a unified and definitive treatment of all these questions, with a striking combination of rigour and simplicity. Schwartz's treatment of Dirac's  $\delta$ -function illustrates his approach in a most direct fashion. Dirac's original definition reads:

- (i)  $\delta(\mathbf{x}) = 0$  for  $\mathbf{x} \neq \mathbf{0}$ ,
- (ii)  $\int_{\mathbb{R}^n} \delta(\mathbf{x}) d^n \mathbf{x} = 1$ .

These two conditions are irreconcilable with Lebesgue's theory of integration: by (i),  $\delta$  vanishes almost everywhere, so that its integral in (ii) must be 0, not 1.

A better definition consists in specifying that

- (iii)  $\int_{\mathbb{R}^n} \delta(\mathbf{x}) \varphi(\mathbf{x}) d^n \mathbf{x} = \varphi(\mathbf{0})$

for any function  $\varphi$  sufficiently well behaved near  $\mathbf{x} = \mathbf{0}$ . This is related to the problem of finding a unit for convolution (Section 1.3.2.2.4). As will now be seen, this definition is still unsatisfactory. Let the sequence  $(f_\nu)$  in  $L^1(\mathbb{R}^n)$  be an approximate convolution unit, *e.g.*

$$f_\nu(\mathbf{x}) = \left(\frac{\nu}{2\pi}\right)^{1/2} \exp(-\frac{1}{2}\nu^2 \|\mathbf{x}\|^2).$$

Then for any well behaved function  $\varphi$  the integrals

$$\int_{\mathbb{R}^n} f_\nu(\mathbf{x}) \varphi(\mathbf{x}) d^n \mathbf{x}$$

exist, and the sequence of their numerical values tends to  $\varphi(\mathbf{0})$ . It is tempting to combine this with (iii) to conclude that  $\delta$  is the limit of the sequence  $(f_\nu)$  as  $\nu \rightarrow \infty$ . However,

$$\lim_{\nu \rightarrow \infty} f_\nu(\mathbf{x}) = 0 \quad \text{as } \nu \rightarrow \infty$$

almost everywhere in  $\mathbb{R}^n$  and the crux of the problem is that

$$\begin{aligned} \varphi(\mathbf{0}) &= \lim_{\nu \rightarrow \infty} \int_{\mathbb{R}^n} f_\nu(\mathbf{x}) \varphi(\mathbf{x}) d^n \mathbf{x} \\ &\neq \int_{\mathbb{R}^n} \left[ \lim_{\nu \rightarrow \infty} f_\nu(\mathbf{x}) \right] \varphi(\mathbf{x}) d^n \mathbf{x} = 0 \end{aligned}$$

because the sequence  $(f_\nu)$  does not satisfy the hypotheses of Lebesgue's dominated convergence theorem.

Schwartz's solution to this problem is deceptively simple: the regular behaviour one is trying to capture is an attribute not of the sequence of functions  $(f_\nu)$ , but of the sequence of continuous linear functionals

$$T_\nu : \varphi \mapsto \int_{\mathbb{R}^n} f_\nu(\mathbf{x}) \varphi(\mathbf{x}) d^n \mathbf{x}$$

which has as a limit the continuous functional

$$T : \varphi \mapsto \varphi(\mathbf{0}).$$

It is the latter functional which constitutes the proper definition of  $\delta$ . The previous paradoxes arose because one insisted on writing down the simple linear operation  $T$  in terms of an integral.

The essence of Schwartz's theory of distributions is thus that, rather than try to define and handle 'generalized functions' *via* sequences such as  $(f_\nu)$  [an approach adopted *e.g.* by Lighthill (1958) and Erdélyi (1962)], one should instead look at them as continuous linear functionals over spaces of well behaved functions.

There are many books on distribution theory and its applications. The reader may consult in particular Schwartz (1965, 1966), Gel'fand & Shilov (1964), Bremermann (1965), Trèves (1967), Challifour (1972), Friedlander (1982), and the relevant chapters of Hörmander (1963) and Yosida (1965). Schwartz (1965) is especially recommended as an introduction.

#### 1.3.2.3.2. Rationale

The guiding principle which leads to requiring that the functions  $\varphi$  above (traditionally called 'test functions') should be well behaved is that correspondingly 'wilder' behaviour can then be accommodated in the limiting behaviour of the  $f_\nu$  while still keeping the integrals  $\int_{\mathbb{R}^n} f_\nu \varphi d^n \mathbf{x}$  under control. Thus

- (i) to minimize restrictions on the limiting behaviour of the  $f_\nu$  at infinity, the  $\varphi$ 's will be chosen to have *compact support*;
- (ii) to minimize restrictions on the *local* behaviour of the  $f_\nu$ , the  $\varphi$ 's will be chosen *infinitely differentiable*.

To ensure further the *continuity* of functionals such as  $T_\nu$  with respect to the test function  $\varphi$  as the  $f_\nu$  go increasingly wild, very strong control will have to be exercised in the way in which a sequence  $(\varphi_j)$  of test functions will be said to converge towards a limiting  $\varphi$ : conditions will have to be imposed not only on the values of the functions  $\varphi_j$ , but also on those of all their *derivatives*. Hence, defining a strong enough topology on the space of test functions  $\varphi$  is an *essential* prerequisite to the development of a satisfactory theory of distributions.

#### 1.3.2.3.3. Test-function spaces

With this rationale in mind, the following function spaces will be defined for any open subset  $\Omega$  of  $\mathbb{R}^n$  (which may be the whole of  $\mathbb{R}^n$ ):

- (a)  $\mathcal{E}(\Omega)$  is the space of complex-valued functions over  $\Omega$  which are indefinitely differentiable;
- (b)  $\mathcal{D}(\Omega)$  is the subspace of  $\mathcal{E}(\Omega)$  consisting of functions with (unspecified) compact support contained in  $\mathbb{R}^n$ ;
- (c)  $\mathcal{D}_K(\Omega)$  is the subspace of  $\mathcal{D}(\Omega)$  consisting of functions whose (compact) support is contained within a fixed compact subset  $K$  of  $\Omega$ .

When  $\Omega$  is unambiguously defined by the context, we will simply write  $\mathcal{E}, \mathcal{D}, \mathcal{D}_K$ .

It sometimes suffices to require the existence of continuous derivatives only up to finite order  $m$  inclusive. The corresponding spaces are then denoted  $\mathcal{E}^{(m)}, \mathcal{D}^{(m)}, \mathcal{D}_K^{(m)}$  with the convention that if  $m = 0$ , only continuity is required.

The *topologies* on these spaces constitute the most important ingredients of distribution theory, and will be outlined in some detail.

##### 1.3.2.3.3.1. Topology on $\mathcal{E}(\Omega)$

It is defined by the family of semi-norms

$$\varphi \in \mathcal{E}(\Omega) \mapsto \sigma_{\mathbf{p}, K}(\varphi) = \sup_{\mathbf{x} \in K} |D^{\mathbf{p}} \varphi(\mathbf{x})|,$$

where  $\mathbf{p}$  is a multi-index and  $K$  a compact subset of  $\Omega$ . A

## 1. GENERAL RELATIONSHIPS AND TECHNIQUES

fundamental system  $S$  of neighbourhoods of the origin in  $\mathcal{E}(\Omega)$  is given by subsets of  $\mathcal{E}(\Omega)$  of the form

$$V(m, \varepsilon, K) = \{\varphi \in \mathcal{E}(\Omega) \mid |\mathbf{p}| \leq m \Rightarrow \sigma_{\mathbf{p}, K}(\varphi) < \varepsilon\}$$

for all natural integers  $m$ , positive real  $\varepsilon$ , and compact subset  $K$  of  $\Omega$ . Since a *countable* family of compact subsets  $K$  suffices to cover  $\Omega$ , and since restricted values of  $\varepsilon$  of the form  $\varepsilon = 1/N$  lead to the same topology,  $S$  is equivalent to a *countable* system of neighbourhoods and hence  $\mathcal{E}(\Omega)$  is *metrizable*.

Convergence in  $\mathcal{E}$  may thus be defined by means of sequences. A sequence  $(\varphi_\nu)$  in  $\mathcal{E}$  will be said to converge to 0 if for any given  $V(m, \varepsilon, K)$  there exists  $\nu_0$  such that  $\varphi_\nu \in V(m, \varepsilon, K)$  whenever  $\nu > \nu_0$ ; in other words, if the  $\varphi_\nu$  and all their derivatives  $D^{\mathbf{p}}\varphi_\nu$  converge to 0 uniformly on any given compact  $K$  in  $\Omega$ .

### 1.3.2.3.3.2. Topology on $\mathcal{Q}_K(\Omega)$

It is defined by the family of semi-norms

$$\varphi \in \mathcal{Q}_K(\Omega) \mapsto \sigma_{\mathbf{p}}(\varphi) = \sup_{\mathbf{x} \in K} |D^{\mathbf{p}}\varphi(\mathbf{x})|,$$

where  $K$  is now fixed. The fundamental system  $S$  of neighbourhoods of the origin in  $\mathcal{Q}_K$  is given by sets of the form

$$V(m, \varepsilon) = \{\varphi \in \mathcal{Q}_K(\Omega) \mid |\mathbf{p}| \leq m \Rightarrow \sigma_{\mathbf{p}}(\varphi) < \varepsilon\}.$$

It is equivalent to the countable subsystem of the  $V(m, 1/N)$ , hence  $\mathcal{Q}_K(\Omega)$  is *metrizable*.

Convergence in  $\mathcal{Q}_K$  may thus be defined by means of sequences. A sequence  $(\varphi_\nu)$  in  $\mathcal{Q}_K$  will be said to converge to 0 if for any given  $V(m, \varepsilon)$  there exists  $\nu_0$  such that  $\varphi_\nu \in V(m, \varepsilon)$  whenever  $\nu > \nu_0$ ; in other words, if the  $\varphi_\nu$  and all their derivatives  $D^{\mathbf{p}}\varphi_\nu$  converge to 0 uniformly in  $K$ .

### 1.3.2.3.3.3. Topology on $\mathcal{Q}(\Omega)$

It is defined by the fundamental system of neighbourhoods of the origin consisting of sets of the form

$$V((m), (\varepsilon)) = \left\{ \varphi \in \mathcal{Q}(\Omega) \mid |\mathbf{p}| \leq m_\nu \Rightarrow \sup_{\|\mathbf{x}\| \leq \nu} |D^{\mathbf{p}}\varphi(\mathbf{x})| < \varepsilon_\nu \text{ for all } \nu \right\},$$

where  $(m)$  is an increasing sequence  $(m_\nu)$  of integers tending to  $+\infty$  and  $(\varepsilon)$  is a decreasing sequence  $(\varepsilon_\nu)$  of positive reals tending to 0, as  $\nu \rightarrow \infty$ .

This topology is *not metrizable*, because the sets of sequences  $(m)$  and  $(\varepsilon)$  are essentially uncountable. It can, however, be shown to be the *inductive limit* of the topology of the subspaces  $\mathcal{Q}_K$ , in the following sense:  $V$  is a neighbourhood of the origin in  $\mathcal{Q}$  if and only if its intersection with  $\mathcal{Q}_K$  is a neighbourhood of the origin in  $\mathcal{Q}_K$  for any given compact  $K$  in  $\Omega$ .

A sequence  $(\varphi_\nu)$  in  $\mathcal{Q}$  will thus be said to converge to 0 in  $\mathcal{Q}$  if all the  $\varphi_\nu$  belong to some  $\mathcal{Q}_K$  (with  $K$  a compact subset of  $\Omega$  independent of  $\nu$ ) and if  $(\varphi_\nu)$  converges to 0 in  $\mathcal{Q}_K$ .

As a result, a complex-valued functional  $T$  on  $\mathcal{Q}$  will be said to be continuous for the topology of  $\mathcal{Q}$  if and only if, for any given compact  $K$  in  $\Omega$ , its restriction to  $\mathcal{Q}_K$  is continuous for the topology of  $\mathcal{Q}_K$ , *i.e.* maps convergent sequences in  $\mathcal{Q}_K$  to convergent sequences in  $\mathbb{C}$ .

This property of  $\mathcal{Q}$ , *i.e.* having a non-metrizable topology which is the inductive limit of metrizable topologies in its subspaces  $\mathcal{Q}_K$ , conditions the whole structure of distribution theory and dictates that of many of its proofs.

### 1.3.2.3.3.4. Topologies on $\mathcal{E}^{(m)}$ , $\mathcal{Q}_K^{(m)}$ , $\mathcal{Q}^{(m)}$

These are defined similarly, but only involve conditions on derivatives up to order  $m$ .

### 1.3.2.3.4. Definition of distributions

A *distribution*  $T$  on  $\Omega$  is a *linear form* over  $\mathcal{Q}(\Omega)$ , *i.e.* a map

$$T : \varphi \mapsto \langle T, \varphi \rangle$$

which associates linearly a complex number  $\langle T, \varphi \rangle$  to any  $\varphi \in \mathcal{Q}(\Omega)$ , and which is *continuous* for the topology of that space. In the terminology of Section 1.3.2.2.6.2,  $T$  is an element of  $\mathcal{Q}'(\Omega)$ , the *topological dual* of  $\mathcal{Q}(\Omega)$ .

Continuity over  $\mathcal{Q}$  is equivalent to continuity over  $\mathcal{Q}_K$  for all compact  $K$  contained in  $\Omega$ , and hence to the condition that for any sequence  $(\varphi_\nu)$  in  $\mathcal{Q}$  such that

- (i)  $\text{Supp } \varphi_\nu$  is contained in some compact  $K$  independent of  $\nu$ ,
- (ii) the sequences  $(|D^{\mathbf{p}}\varphi_\nu|)$  converge uniformly to 0 on  $K$  for all multi-indices  $\mathbf{p}$ ;

then the sequence of complex numbers  $\langle T, \varphi_\nu \rangle$  converges to 0 in  $\mathbb{C}$ .

If the continuity of a distribution  $T$  requires (ii) for  $|\mathbf{p}| \leq m$  only,  $T$  may be defined over  $\mathcal{Q}^{(m)}$  and thus  $T \in \mathcal{Q}'^{(m)}$ ;  $T$  is said to be a *distribution of finite order*  $m$ . In particular, for  $m = 0$ ,  $\mathcal{Q}^{(0)}$  is the space of continuous functions with compact support, and a distribution  $T \in \mathcal{Q}'^{(0)}$  is a (Radon) *measure* as used in the theory of integration. Thus measures are particular cases of distributions.

Generally speaking, the *larger* a space of test functions, the *smaller* its topological dual:

$$m < n \Rightarrow \mathcal{Q}^{(m)} \supset \mathcal{Q}^{(n)} \Rightarrow \mathcal{Q}'^{(n)} \supset \mathcal{Q}'^{(m)}.$$

This clearly results from the observation that if the  $\varphi$ 's are allowed to be less regular, then less wildness can be accommodated in  $T$  if the continuity of the map  $\varphi \mapsto \langle T, \varphi \rangle$  with respect to  $\varphi$  is to be preserved.

### 1.3.2.3.5. First examples of distributions

(i) The linear map  $\varphi \mapsto \langle \delta, \varphi \rangle = \varphi(\mathbf{0})$  is a measure (*i.e.* a zeroth-order distribution) called Dirac's measure or (improperly) Dirac's ' $\delta$ -function'.

(ii) The linear map  $\varphi \mapsto \langle \delta_{(\mathbf{a})}, \varphi \rangle = \varphi(\mathbf{a})$  is called Dirac's measure at point  $\mathbf{a} \in \mathbb{R}^n$ .

(iii) The linear map  $\varphi \mapsto (-1)^{|\mathbf{p}|} D^{\mathbf{p}}\varphi(\mathbf{a})$  is a distribution of order  $m = |\mathbf{p}| > 0$ , and hence is not a measure.

(iv) The linear map  $\varphi \mapsto \sum_{\nu > 0} \varphi^{(\nu)}(\nu)$  is a distribution of infinite order on  $\mathbb{R}$ : the order of differentiation is bounded for each  $\varphi$  (because  $\varphi$  has compact support) but is not as  $\varphi$  varies.

(v) If  $(\mathbf{p}_\nu)$  is a sequence of multi-indices  $\mathbf{p}_\nu = (p_{1\nu}, \dots, p_{n\nu})$  such that  $|\mathbf{p}_\nu| \rightarrow \infty$  as  $\nu \rightarrow \infty$ , then the linear map  $\varphi \mapsto \sum_{\nu > 0} (D^{\mathbf{p}_\nu}\varphi)(\mathbf{p}_\nu)$  is a distribution of infinite order on  $\mathbb{R}^n$ .

### 1.3.2.3.6. Distributions associated to locally integrable functions

Let  $f$  be a complex-valued function over  $\Omega$  such that  $\int_K |f(\mathbf{x})| d^n \mathbf{x}$  exists for any given compact  $K$  in  $\Omega$ ;  $f$  is then called *locally integrable*.

The linear mapping from  $\mathcal{Q}(\Omega)$  to  $\mathbb{C}$  defined by

$$\varphi \mapsto \int_{\Omega} f(\mathbf{x})\varphi(\mathbf{x}) d^n \mathbf{x}$$

may then be shown to be continuous over  $\mathcal{Q}(\Omega)$ . It thus defines a *distribution*  $T_f \in \mathcal{Q}'(\Omega)$ :

$$\langle T_f, \varphi \rangle = \int_{\Omega} f(\mathbf{x})\varphi(\mathbf{x}) d^n \mathbf{x}.$$

As the continuity of  $T_f$  only requires that  $\varphi \in \mathcal{Q}^{(0)}(\Omega)$ ,  $T_f$  is actually a Radon measure.

### 1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

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

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

if and only if they are equal almost everywhere. The classes of locally integrable functions modulo this equivalence form a vector space denoted  $L^1_{\text{loc}}(\Omega)$ ; each element of  $L^1_{\text{loc}}(\Omega)$  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  $T \in \mathcal{D}'(\Omega)$  is said to *vanish* on an open subset  $\omega$  of  $\Omega$  if it vanishes on all functions in  $\mathcal{D}(\omega)$ , *i.e.* if  $\langle T, \varphi \rangle = 0$  whenever  $\varphi \in \mathcal{D}(\omega)$ .

The *support* of a distribution  $T$ , denoted  $\text{Supp } T$ , is then defined as the complement of the set-theoretic union of those open subsets  $\omega$  on which  $T$  vanishes; or equivalently as the smallest closed subset of  $\Omega$  outside which  $T$  vanishes.

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

It can be shown that any distribution  $T \in \mathcal{D}'$  with compact support may be extended from  $\mathcal{D}$  to  $\mathcal{E}$  while remaining continuous, so that  $T \in \mathcal{E}'$ ; and that conversely, if  $S \in \mathcal{E}'$ , then its restriction  $T$  to  $\mathcal{D}$  is a distribution with compact support. Thus, *the topological dual  $\mathcal{E}'$  of  $\mathcal{E}$  consists of those distributions in  $\mathcal{D}'$  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  $\varphi$ , which may then roam through  $\mathcal{E}$  rather than  $\mathcal{D}$ .

#### 1.3.2.3.8. Convergence of distributions

A sequence  $(T_j)$  of distributions will be said to converge in  $\mathcal{D}'$  to a distribution  $T$  as  $j \rightarrow \infty$  if, for any given  $\varphi \in \mathcal{D}$ , 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  $\mathcal{D}'$  and to have distribution  $S$  as its sum if the sequence of partial sums  $S_k = \sum_{j=0}^k T_j$  converges to  $S$ .

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

$$\varphi \mapsto \lim_{j \rightarrow \infty} \langle T_j, \varphi \rangle$$

define a distribution  $T \in \mathcal{D}'$ ? In other words, does the limiting process preserve continuity with respect to  $\varphi$ ? It is a remarkable theorem that, because of the strong topology on  $\mathcal{D}$ , 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  $(\varphi_\nu)$  with  $\varphi_\nu(x) = \cos \nu x$  converges to 0 in  $\mathcal{D}'(\mathbb{R})$ , 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  $\delta$ . If  $(f_\nu)$  is a sequence of locally summable functions on  $\mathbb{R}^n$  such that

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

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

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

then the sequence  $(T_{f_\nu})$  of distributions converges to  $\delta$  in  $\mathcal{D}'(\mathbb{R}^n)$ .

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

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

$$\begin{aligned} \int_{\mathbb{R}^n} \partial_i f(x_1, \dots, x_i, \dots, x_n) \varphi(x_1, \dots, x_i, \dots, x_n) \, dx_i \\ = (f\varphi)(x_1, \dots, +\infty, \dots, x_n) - (f\varphi)(x_1, \dots, -\infty, \dots, x_n) \\ - \int_{\mathbb{R}^n} f(x_1, \dots, x_i, \dots, x_n) \partial_i \varphi(x_1, \dots, x_i, \dots, x_n) \, dx_i; \end{aligned}$$

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

The test functions  $\varphi \in \mathcal{D}$  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,

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

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

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

where  $\Delta$  is the Laplacian operator. The derivatives of Dirac’s  $\delta$  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  $\mathcal{D}'$ : 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  $j \rightarrow \infty$

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

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  $\langle \cdot, \cdot \rangle$  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

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

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

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

# 1. GENERAL RELATIONSHIPS AND TECHNIQUES

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

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

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

Under these hypotheses,  $I(\lambda)$  is differentiable (in the usual sense) with respect to  $\lambda$  near  $\lambda_0$ , and its derivative may be obtained by ‘differentiation under the  $\langle \cdot, \cdot \rangle$  sign’:

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

### (c) Effect of discontinuities

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

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

$$\langle (T_Y)', \varphi \rangle = -\langle (T_Y), \varphi' \rangle = -\int_0^{+\infty} \varphi'(x) dx = \varphi(0) = \langle \delta, \varphi \rangle.$$

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:

$$\begin{aligned} g_0 &= f - \sigma_0 Y \\ g_1 &= g'_0 - \sigma_1 Y \\ &\text{-----} \\ g_k &= g'_{k-1} - \sigma_k Y. \end{aligned}$$

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:

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

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  $\sigma_0$  and  $\sigma_\nu$  denote the discontinuities of  $f$  and its normal derivative  $\partial_\nu \varphi$  across  $S$  (both  $\sigma_0$  and  $\sigma_\nu$  are functions of position on  $S$ ), and let  $\delta_{(S)}$  and  $\partial_\nu \delta_{(S)}$  be defined by

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

Integration by parts shows that

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

where  $\theta_i$  is the angle between the  $x_i$  axis and the normal to  $S$  along which the jump  $\sigma_0$  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  $\chi \in \mathcal{D}$  such that  $\chi = \psi'$  with  $\psi \in \mathcal{D}$ , we must have

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

This condition defines  $T$  in a ‘hyperplane’  $\mathcal{H}$  of  $\mathcal{D}$ , whose equation

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

reflects the fact that  $\psi$  has compact support.

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

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

with

$$\lambda = \langle 1, \varphi \rangle, \quad \chi = \varphi - \lambda \varphi_0, \quad \psi(x) = \int_0^x \chi(t) dt,$$

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  $\varphi_0$  means that  $T$  is defined up to an additive constant.

### 1.3.2.3.9.3. Multiplication of distributions by functions

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

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

In order that  $\alpha T$  be a distribution, the mapping  $\varphi \mapsto \alpha \varphi$  must send  $\mathcal{D}(\mathbb{R}^n)$  continuously into itself; hence *the multipliers  $\alpha$  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  $\alpha$  needs only have continuous derivatives up to order  $m$ . For instance,  $\delta$  is a distribution of order zero, and  $\alpha \delta = \alpha(\mathbf{0}) \delta$  is a distribution provided  $\alpha$  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,  $D^p \delta$  is a distribution of order  $|\mathbf{p}|$ , and the following formula holds for all  $\alpha \in \mathcal{D}^{(m)}$  with  $m = |\mathbf{p}|$ :

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

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  $\mathbf{p}$

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

### 1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

#### 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  $\alpha$ , the division problem consists in finding a distribution  $T$  such that  $\alpha T = S$ .

If  $\alpha$  never vanishes,  $T = S/\alpha$  is the unique answer. If  $n = 1$ , and if  $\alpha$  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  $\sigma$  be a smooth non-singular change of variables in  $\mathbb{R}^n$ , i.e. an infinitely differentiable mapping from an open subset  $\Omega$  of  $\mathbb{R}^n$  to  $\Omega'$  in  $\mathbb{R}^n$ , whose Jacobian

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

vanishes nowhere in  $\Omega$ . By the implicit function theorem, the inverse mapping  $\sigma^{-1}$  from  $\Omega'$  to  $\Omega$  is well defined.

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

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

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

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

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  $\sigma$  through

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

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

For instance, if  $\sigma: \mathbf{x} \mapsto \mathbf{x} + \mathbf{a}$  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  $A: \mathbf{x} \mapsto \mathbf{A}\mathbf{x}$  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  $\check{\varphi}$  we have:

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

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

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  $\delta$  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  $L_{\text{loc}}^1(\mathbb{R}^m) \otimes L_{\text{loc}}^1(\mathbb{R}^n)$  as the vector space of finite linear combinations of functions of the form

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

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

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  $f \otimes g \in L_{\text{loc}}^1(\mathbb{R}^m \times \mathbb{R}^n)$ , 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  $\varphi_{\mathbf{x}, \mathbf{y}} \in \mathcal{D}(\mathbb{R}^m \times \mathbb{R}^n)$ . In particular, if  $\varphi(\mathbf{x}, \mathbf{y}) = u(\mathbf{x})v(\mathbf{y})$  with  $u \in \mathcal{D}(\mathbb{R}^m), v \in \mathcal{D}(\mathbb{R}^n)$ , 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  $S \in \mathcal{D}'(\mathbb{R}^m)$  and  $T \in \mathcal{D}'(\mathbb{R}^n)$ . Given any test function  $\varphi \in \mathcal{D}(\mathbb{R}^m \times \mathbb{R}^n)$ , let  $\varphi_{\mathbf{x}}$  denote the map  $\mathbf{y} \mapsto \varphi(\mathbf{x}, \mathbf{y})$ ; let  $\varphi_{\mathbf{y}}$  denote the map  $\mathbf{x} \mapsto \varphi(\mathbf{x}, \mathbf{y})$ ; 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,  $\theta \in \mathcal{D}'(\mathbb{R}^m), \omega \in \mathcal{D}'(\mathbb{R}^n)$ , 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

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

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

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

## 1. GENERAL RELATIONSHIPS AND TECHNIQUES

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  $\varphi \in \mathcal{D}(\mathbb{R}^n)$ ,

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

Introducing the map  $\sigma$  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  $\sigma$  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  $\text{Supp}(S \otimes T) = (\text{Supp } S) \times (\text{Supp } T)$  and  $\sigma^{-1}(\text{Supp } \varphi)$  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  $\mathcal{D}'$ : it maps convergent sequences  $(T_j)$  to convergent sequences  $(S * T_j)$ . Convolution is commutative:  $S * T = T * S$ .

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

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

whenever the following generalized support condition:

‘the set  $\{(\mathbf{x}_1, \dots, \mathbf{x}_p) | \mathbf{x}_1 \in A_1, \dots, \mathbf{x}_p \in A_p, \mathbf{x}_1 + \dots + \mathbf{x}_p \in K\}$  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  $T \in \mathcal{D}'$ , the following identities hold:

- (i)  $\delta * T = T$ :  $\delta$  is the unit convolution;
- (ii)  $\delta_{(\mathbf{a})} * T = \tau_{\mathbf{a}}T$ : translation is a convolution with the corresponding translate of  $\delta$ ;
- (iii)  $(D^{\mathbf{p}}\delta) * T = D^{\mathbf{p}}T$ : differentiation is a convolution with the corresponding derivative of  $\delta$ ;
- (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,  $\alpha$  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

$(\alpha_\nu)$  of such functions  $\alpha$  can be constructed which have compact support and converge to  $\delta$ , it follows that any distribution  $T$  can be obtained as the limit of infinitely differentiable functions  $T * \alpha_\nu$ . In topological jargon:  $\mathcal{D}(\mathbb{R}^n)$  is ‘everywhere dense’ in  $\mathcal{D}'(\mathbb{R}^n)$ . A standard function in  $\mathcal{D}$  which is often used for such proofs is defined as follows: put

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

with

$$A = \int_{-1}^{+1} \exp\left(-\frac{1}{1-x^2}\right) dx$$

(so that  $\theta$  is in  $\mathcal{D}$  and is normalized), and put

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

Another related result, also proved by convolution, is the *structure theorem*: the restriction of a distribution  $T \in \mathcal{D}'(\mathbb{R}^n)$  to a bounded open set  $\Omega$  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  $\mathcal{F}[f]$  and Fourier cotransform  $\bar{\mathcal{F}}[f]$  are defined as follows:

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

where  $\xi \cdot \mathbf{x} = \sum_{i=1}^n \xi_i x_i$  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  $\mathcal{S}'(\mathbb{R}^n)$  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.