

1.3. Fourier transforms in crystallography: theory, algorithms and applications

BY G. BRICOGNE

1.3.1. General introduction

Since the publication of Volume II of *International Tables*, most aspects of the theory, computation and applications of Fourier transforms have undergone considerable development, often to the point of being hardly recognizable.

The mathematical analysis of the Fourier transformation has been extensively reformulated within the framework of distribution theory, following Schwartz's work in the early 1950s.

The computation of Fourier transforms has been revolutionized by the advent of digital computers and of the Cooley–Tukey algorithm, and progress has been made at an ever-accelerating pace in the design of new types of algorithms and in optimizing their interplay with machine architecture.

These advances have transformed both theory and practice in several fields which rely heavily on Fourier methods; much of electrical engineering, for instance, has become digital signal processing.

By contrast, crystallography has remained relatively unaffected by these developments. From the conceptual point of view, old-fashioned Fourier series are still adequate for the quantitative description of X-ray diffraction, as this rarely entails consideration of molecular transforms between reciprocal-lattice points. From the practical point of view, three-dimensional Fourier transforms have mostly been used as a tool for visualizing electron-density maps, so that only moderate urgency was given to trying to achieve ultimate efficiency in these relatively infrequent calculations.

Recent advances in phasing and refinement methods, however, have placed renewed emphasis on concepts and techniques long used in digital signal processing, *e.g.* flexible sampling, Shannon interpolation, linear filtering, and interchange between convolution and multiplication. These methods are iterative in nature, and thus generate a strong incentive to design new crystallographic Fourier transform algorithms making the fullest possible use of all available symmetry to save both storage and computation.

As a result, need has arisen for a modern and coherent account of Fourier transform methods in crystallography which would provide:

(i) a simple and foolproof means of switching between the three different guises in which the Fourier transformation is encountered (Fourier transforms, Fourier series and discrete Fourier transforms), both formally and computationally;

(ii) an up-to-date presentation of the most important algorithms for the efficient numerical calculation of discrete Fourier transforms;

(iii) a systematic study of the incorporation of symmetry into the calculation of crystallographic discrete Fourier transforms;

(iv) a survey of the main types of crystallographic computations based on the Fourier transformation.

The rapid pace of progress in these fields implies that such an account would be struck by quasi-immediate obsolescence if it were written solely for the purpose of compiling a catalogue of results and formulae 'customized' for crystallographic use. Instead, the emphasis has been placed on a mode of presentation in which most results and formulae are *derived* rather than listed. This does entail a substantial mathematical overhead, but has the advantage of preserving in its 'native' form the context within which these results are obtained. It is this context, rather than any particular set of results, which constitutes the most fertile source

of new ideas and new applications, and as such can have any hope at all of remaining useful in the long run.

These conditions have led to the following choices:

(i) the mathematical theory of the Fourier transformation has been cast in the language of Schwartz's theory of distributions which has long been adopted in several applied fields, in particular electrical engineering, with considerable success; the extra work involved handsomely pays for itself by allowing the three different types of Fourier transformations to be treated together, and by making all properties of the Fourier transform consequences of a single property (the convolution theorem). This is particularly useful in all questions related to the sampling theorem;

(ii) the various numerical algorithms have been presented as the consequences of basic algebraic phenomena involving Abelian groups, rings and finite fields; this degree of formalization greatly helps the subsequent incorporation of symmetry;

(iii) the algebraic nature of space groups has been re-emphasized so as to build up a framework which can accommodate both the phenomena used to factor the discrete Fourier transform and those which underlie the existence (and lead to the classification) of space groups; this common ground is found in the notion of *module over a group ring* (*i.e.* integral representation theory), which is then applied to the formulation of a large number of algorithms, many of which are new;

(iv) the survey of the main types of crystallographic computations has tried to highlight the roles played by various properties of the Fourier transformation, and the ways in which a better exploitation of these properties has been the driving force behind the discovery of more powerful methods.

In keeping with this philosophy, the theory is presented first, followed by the crystallographic applications. There are 'forward references' from mathematical results to the applications which later invoke them (thus giving 'real-life' examples rather than artificial ones), and 'backward references' as usual. In this way, the internal logic of the mathematical developments – the surest guide to future innovations – can be preserved, whereas the alternative solution of relegating these to appendices tends on the contrary to obscure that logic by subordinating it to that of the applications.

It is hoped that this attempt at an overall presentation of the main features of Fourier transforms and of their ubiquitous role in crystallography will be found useful by scientists both within and outside the field.

1.3.2. The mathematical theory of the Fourier transformation

1.3.2.1. Introduction

The Fourier transformation and the practical applications to which it gives rise occur in three different forms which, although they display a similar range of phenomena, normally require distinct formulations and different proof techniques:

(i) *Fourier transforms*, in which both function and transform depend on continuous variables;

(ii) *Fourier series*, which relate a periodic function to a discrete set of coefficients indexed by n -tuples of integers;

(iii) *discrete Fourier transforms*, which relate finite-dimensional vectors by linear operations representable by matrices.

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

At the same time, the most useful property of the Fourier transformation – the exchange between multiplication and convolution – is mathematically the most elusive and the one which requires the greatest caution in order to avoid writing down meaningless expressions.

It is the unique merit of Schwartz's theory of distributions (Schwartz, 1966) that it affords complete control over all the troublesome phenomena which had previously forced mathematicians to settle for a piecemeal, fragmented theory of the Fourier transformation. By its ability to handle rigorously highly 'singular' objects (especially δ -functions, their derivatives, their tensor products, their products with smooth functions, their translates and lattices of these translates), distribution theory can deal with all the major properties of the Fourier transformation as particular instances of a single basic result (the exchange between multiplication and convolution), and can at the same time accommodate the three previously distinct types of Fourier theories within a unique framework. This brings great simplification to matters of central importance in crystallography, such as the relations between

- (a) periodization, and sampling or decimation;
- (b) Shannon interpolation, and masking by an indicator function;
- (c) section, and projection;
- (d) differentiation, and multiplication by a monomial;
- (e) translation, and phase shift.

All these properties become subsumed under the same theorem.

This striking synthesis comes at a slight price, which is the relative complexity of the notion of distribution. It is first necessary to establish the notion of topological vector space and to gain sufficient control (or, at least, understanding) over convergence behaviour in certain of these spaces. The key notion of *metrizability* cannot be circumvented, as it underlies most of the constructs and many of the proof techniques used in distribution theory. Most of Section 1.3.2.2 builds up to the fundamental result at the end of Section 1.3.2.2.6.2, which is basic to the definition of a distribution in Section 1.3.2.3.4 and to all subsequent developments.

The reader mostly interested in applications will probably want to reach this section by starting with his or her favourite topic in Section 1.3.4, and following the backward references to the relevant properties of the Fourier transformation, then to the proof of these properties, and finally to the definitions of the objects involved. Hopefully, he or she will then feel inclined to follow the forward references and thus explore the subject from the abstract to the practical. The books by Dieudonné (1969) and Lang (1965) are particularly recommended as general references for all aspects of analysis and algebra.

1.3.2.2. Preliminary notions and notation

Throughout this text, \mathbb{R} will denote the set of real numbers, \mathbb{Z} the set of rational (signed) integers and \mathbb{N} the set of natural (unsigned) integers. The symbol \mathbb{R}^n will denote the Cartesian product of n copies of \mathbb{R} :

$$\mathbb{R}^n = \mathbb{R} \times \dots \times \mathbb{R} \quad (n \text{ times}, n \geq 1),$$

so that an element \mathbf{x} of \mathbb{R}^n is an n -tuple of real numbers:

$$\mathbf{x} = (x_1, \dots, x_n).$$

Similar meanings will be attached to \mathbb{Z}^n and \mathbb{N}^n .

The symbol \mathbb{C} will denote the set of complex numbers. If $z \in \mathbb{C}$, its modulus will be denoted by $|z|$, its conjugate by \bar{z} (not z^*), and its real and imaginary parts by $\operatorname{Re}(z)$ and $\operatorname{Im}(z)$:

$$\operatorname{Re}(z) = \frac{1}{2}(z + \bar{z}), \quad \operatorname{Im}(z) = \frac{1}{2i}(z - \bar{z}).$$

If X is a finite set, then $|X|$ will denote the number of its elements. If mapping f sends an element x of set X to the element $f(x)$ of set Y , the notation

$$f : x \mapsto f(x)$$

will be used; the plain arrow \rightarrow will be reserved for denoting limits, as in

$$\lim_{p \rightarrow \infty} \left(1 + \frac{x}{p}\right)^p = e^x.$$

If X is any set and S is a subset of X , the *indicator function* χ_S of S is the real-valued function on X defined by

$$\chi_S(x) = 1 \quad \text{if } x \in S \\ = 0 \quad \text{if } x \notin S.$$

1.3.2.2.1. Metric and topological notions in \mathbb{R}^n

The set \mathbb{R}^n can be endowed with the structure of a vector space of dimension n over \mathbb{R} , and can be made into a Euclidean space by treating its standard basis as an orthonormal basis and defining the Euclidean norm:

$$\|\mathbf{x}\| = \left(\sum_{i=1}^n x_i^2\right)^{1/2}.$$

By misuse of notation, \mathbf{x} will sometimes also designate the column vector of coordinates of $\mathbf{x} \in \mathbb{R}^n$; if these coordinates are referred to an orthonormal basis of \mathbb{R}^n , then

$$\|\mathbf{x}\| = (\mathbf{x}^T \mathbf{x})^{1/2},$$

where \mathbf{x}^T denotes the transpose of \mathbf{x} .

The distance between two points \mathbf{x} and \mathbf{y} defined by $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ allows the topological structure of \mathbb{R} to be transferred to \mathbb{R}^n , making it a *metric space*. The basic notions in a metric space are those of neighbourhoods, of open and closed sets, of limit, of continuity, and of convergence (see Section 1.3.2.2.6.1).

A subset S of \mathbb{R}^n is *bounded* if $\sup \|\mathbf{x} - \mathbf{y}\| < \infty$ as \mathbf{x} and \mathbf{y} run through S ; it is *closed* if it contains the limits of all convergent sequences with elements in S . A subset K of \mathbb{R}^n which is both bounded and closed has the property of being *compact*, *i.e.* that whenever K has been covered by a family of open sets, a finite subfamily can be found which suffices to cover K . Compactness is a very useful topological property for the purpose of proof, since it allows one to reduce the task of examining infinitely many local situations to that of examining only finitely many of them.

1.3.2.2.2. Functions over \mathbb{R}^n

Let φ be a complex-valued function over \mathbb{R}^n . The *support* of φ , denoted $\operatorname{Supp} \varphi$, is the smallest closed subset of \mathbb{R}^n outside which φ vanishes identically. If $\operatorname{Supp} \varphi$ is compact, φ is said to have *compact support*.

1. GENERAL RELATIONSHIPS AND TECHNIQUES

If $\mathbf{t} \in \mathbb{R}^n$, the *translate* of φ by \mathbf{t} , denoted $\tau_{\mathbf{t}}\varphi$, is defined by

$$(\tau_{\mathbf{t}}\varphi)(\mathbf{x}) = \varphi(\mathbf{x} - \mathbf{t}).$$

Its support is the geometric translate of that of φ :

$$\text{Supp } \tau_{\mathbf{t}}\varphi = \{\mathbf{x} + \mathbf{t} | \mathbf{x} \in \text{Supp } \varphi\}.$$

If A is a nonsingular linear transformation in \mathbb{R}^n , the *image* of φ by A , denoted $A^{\#}\varphi$, is defined by

$$(A^{\#}\varphi)(\mathbf{x}) = \varphi[A^{-1}(\mathbf{x})].$$

Its support is the geometric image of $\text{Supp } \varphi$ under A :

$$\text{Supp } A^{\#}\varphi = \{A(\mathbf{x}) | \mathbf{x} \in \text{Supp } \varphi\}.$$

If S is a nonsingular affine transformation in \mathbb{R}^n of the form

$$S(\mathbf{x}) = A(\mathbf{x}) + \mathbf{b}$$

with A linear, the image of φ by S is $S^{\#}\varphi = \tau_{\mathbf{b}}(A^{\#}\varphi)$, *i.e.*

$$(S^{\#}\varphi)(\mathbf{x}) = \varphi[A^{-1}(\mathbf{x} - \mathbf{b})].$$

Its support is the geometric image of $\text{Supp } \varphi$ under S :

$$\text{Supp } S^{\#}\varphi = \{S(\mathbf{x}) | \mathbf{x} \in \text{Supp } \varphi\}.$$

It may be helpful to visualize the process of forming the image of a function by a geometric operation as consisting of applying that operation to the *graph* of that function, which is equivalent to applying the *inverse* transformation to the coordinates \mathbf{x} . This use of the inverse later affords the ‘left-representation property’ [see Section 1.3.4.2.2.2(e)] when the geometric operations form a group, which is of fundamental importance in the treatment of crystallographic symmetry (Sections 1.3.4.2.2.4, 1.3.4.2.2.5).

1.3.2.2.3. Multi-index notation

When dealing with functions in n variables and their derivatives, considerable abbreviation of notation can be obtained through the use of multi-indices.

A *multi-index* $\mathbf{p} \in \mathbb{N}^n$ is an n -tuple of natural integers: $\mathbf{p} = (p_1, \dots, p_n)$. The *length* of \mathbf{p} is defined as

$$|\mathbf{p}| = \sum_{i=1}^n p_i,$$

and the following abbreviations will be used:

- (i) $\mathbf{x}^{\mathbf{p}} = x_1^{p_1} \dots x_n^{p_n}$
- (ii) $D_i f = \frac{\partial f}{\partial x_i} = \partial_i f$
- (iii) $D^{\mathbf{p}} f = D_1^{p_1} \dots D_n^{p_n} f = \frac{\partial^{|\mathbf{p}|} f}{\partial x_1^{p_1} \dots \partial x_n^{p_n}}$
- (iv) $\mathbf{q} \leq \mathbf{p}$ if and only if $q_i \leq p_i$ for all $i = 1, \dots, n$
- (v) $\mathbf{p} - \mathbf{q} = (p_1 - q_1, \dots, p_n - q_n)$
- (vi) $\mathbf{p}! = p_1! \times \dots \times p_n!$
- (vii) $\binom{\mathbf{p}}{\mathbf{q}} = \binom{p_1}{q_1} \times \dots \times \binom{p_n}{q_n}$.

Leibniz’s formula for the repeated differentiation of products then assumes the concise form

$$D^{\mathbf{p}}(fg) = \sum_{\mathbf{q} \leq \mathbf{p}} \binom{\mathbf{p}}{\mathbf{q}} D^{\mathbf{p}-\mathbf{q}} f D^{\mathbf{q}} g,$$

while the Taylor expansion of f to order m about $\mathbf{x} = \mathbf{a}$ reads

$$f(\mathbf{x}) = \sum_{|\mathbf{p}| \leq m} \frac{1}{\mathbf{p}!} [D^{\mathbf{p}} f(\mathbf{a})](\mathbf{x} - \mathbf{a})^{\mathbf{p}} + o(\|\mathbf{x} - \mathbf{a}\|^m).$$

In certain sections the notation ∇f will be used for the gradient vector of f , and the notation $(\nabla \nabla^T) f$ for the Hessian matrix of its mixed second-order partial derivatives:

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}, \quad \nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix},$$

$$(\nabla \nabla^T) f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}.$$

1.3.2.2.4. Integration, L^p spaces

The Riemann integral used in elementary calculus suffers from the drawback that vector spaces of Riemann-integrable functions over \mathbb{R}^n are not *complete* for the topology of convergence in the mean: a Cauchy sequence of integrable functions may converge to a nonintegrable function.

To obtain the property of completeness, which is fundamental in functional analysis, it was necessary to extend the notion of integral. This was accomplished by Lebesgue [see Berberian (1962), Dieudonné (1970), or Chapter 1 of Dym & McKean (1972) and the references therein, or Chapter 9 of Sprecher (1970)], and entailed identifying functions which differed only on a subset of zero measure in \mathbb{R}^n (such functions are said to be equal ‘almost everywhere’). The vector spaces $L^p(\mathbb{R}^n)$ consisting of function classes f modulo this identification for which

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

$$\|f\|_p = \left(\int_{\mathbb{R}^n} |f(\mathbf{x})|^p d^n \mathbf{x} \right)^{1/p} < \infty$$

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

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

are then complete for the topology induced by the norm $\|\cdot\|_p$; the limit of every Cauchy sequence of functions in L^p is itself a function in L^p (Riesz–Fischer theorem).

The space $L^1(\mathbb{R}^n)$ consists of those function classes f such that

$$\|f\|_1 = \int_{\mathbb{R}^n} |f(\mathbf{x})| d^n \mathbf{x} < \infty$$

which are called *summable* or *absolutely integrable*. The convolution product:

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

is well defined; combined with the vector space structure of L^1 , it makes L^1 into a (commutative) *convolution algebra*. However, this algebra has no unit element: there is no $f \in L^1$ such that $f * g = g$ for all $g \in L^1$; it has only approximate units, *i.e.* sequences (f_ν) such that $f_\nu * g$ tends to g in the L^1 topology as $\nu \rightarrow \infty$. This is one of the starting points of distribution theory.

The space $L^2(\mathbb{R}^n)$ of *square-integrable* functions can be endowed with a scalar product

$$(f, g) = \int_{\mathbb{R}^n} \overline{f(\mathbf{x})}g(\mathbf{x}) d^n \mathbf{x}$$

which makes it into a *Hilbert space*. The Cauchy–Schwarz inequality

$$|(f, g)| \leq [(f, f)(g, g)]^{1/2}$$

generalizes the fact that the absolute value of the cosine of an angle is less than or equal to 1.

The space $L^\infty(\mathbb{R}^n)$ is defined as the space of functions f such that

$$\|f\|_\infty = \lim_{p \rightarrow \infty} \|f\|_p = \lim_{p \rightarrow \infty} \left(\int_{\mathbb{R}^n} |f(\mathbf{x})|^p d^n \mathbf{x} \right)^{1/p} < \infty.$$

The quantity $\|f\|_\infty$ is called the ‘essential sup norm’ of f , as it is the smallest positive number which $|f(\mathbf{x})|$ exceeds only on a subset of zero measure in \mathbb{R}^n . A function $f \in L^\infty$ is called *essentially bounded*.

1.3.2.2.5. Tensor products. Fubini’s theorem

Let $f \in L^1(\mathbb{R}^m)$, $g \in L^1(\mathbb{R}^n)$. Then the function

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

is called the *tensor product* of f and g , and belongs to $L^1(\mathbb{R}^m \times \mathbb{R}^n)$. The finite linear combinations of functions of the form $f \otimes g$ span a subspace of $L^1(\mathbb{R}^m \times \mathbb{R}^n)$ called the tensor product of $L^1(\mathbb{R}^m)$ and $L^1(\mathbb{R}^n)$ and denoted $L^1(\mathbb{R}^m) \otimes L^1(\mathbb{R}^n)$.

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

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

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

1. GENERAL RELATIONSHIPS AND TECHNIQUES

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 nonmetrizable 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_α 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_α 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* v on a vector space E is a non-negative real-valued function on $E \times E$ such that

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

Subsets of E defined by conditions of the form $v(\mathbf{x}) \leq r$ 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 $d(\mathbf{x}, \mathbf{y}) = v(\mathbf{x} - \mathbf{y})$. 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* σ 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 Σ 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 Σ 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 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 δ -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 δ -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), δ 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 φ 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_ν) 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 φ 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 δ is the limit of the sequence (f_ν) as $\nu \rightarrow \infty$. However,

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

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

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

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

because the sequence (f_ν) 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_ν) , 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 δ . 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_ν) [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 φ 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_ν 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_ν at infinity, the φ 's will be chosen to have *compact support*;

(ii) to minimize restrictions on the *local* behaviour of the f_ν , the φ 's will be chosen *infinitely differentiable*.

To ensure further the *continuity* of functionals such as T_ν with respect to the test function φ as the f_ν go increasingly wild, very strong control will have to be exercised in the way in which a sequence (φ_j) of test functions will be said to converge towards a limiting φ : conditions will have to be imposed not only on the values of the functions φ_j , but also on those of all their *derivatives*. Hence, defining a strong enough topology on the space of test functions φ 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 Ω 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 Ω 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 Ω .

When Ω 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 Ω . A 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 ε , and compact subset K of Ω . Since a *countable* family of compact subsets K suffices to cover Ω , and since restricted values of ε 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 (φ_ν) in \mathcal{E} will be said to converge to 0 if for any given $V(m, \varepsilon, K)$ there exists ν_0 such that $\varphi_\nu \in V(m, \varepsilon, K)$ whenever $\nu > \nu_0$; in other words, if the φ_ν and all their derivatives $D^{\mathbf{p}} \varphi_\nu$ converge to 0 uniformly on any given compact K in Ω .

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

It is defined by the family of semi-norms

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

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

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

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

Convergence in \mathcal{D}_K may thus be defined by means of sequences. A sequence (φ_ν) in \mathcal{D}_K will be said to converge to 0 if for any given $V(m, \varepsilon)$ there exists ν_0 such that $\varphi_\nu \in V(m, \varepsilon)$ whenever $\nu > \nu_0$; in other words, if the φ_ν and all their derivatives $D^{\mathbf{p}} \varphi_\nu$ converge to 0 uniformly in K .

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

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

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

1. GENERAL RELATIONSHIPS AND TECHNIQUES

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

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

A sequence (φ_ν) in \mathcal{D} will thus be said to converge to 0 in \mathcal{D} if all the φ_ν belong to some \mathcal{D}_K (with K a compact subset of Ω independent of ν) and if (φ_ν) converges to 0 in \mathcal{D}_K .

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

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

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

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 Ω is a *linear form* over $\mathcal{D}(\Omega)$, *i.e.* a map

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

which associates linearly a complex number $\langle T, \varphi \rangle$ to any $\varphi \in \mathcal{D}(\Omega)$, 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 $\mathcal{D}'(\Omega)$, the *topological dual* of $\mathcal{D}(\Omega)$.

Continuity over \mathcal{D} is equivalent to continuity over \mathcal{D}_K for all compact K contained in Ω , and hence to the condition that for any sequence (φ_ν) in \mathcal{D} such that

- (i) $\text{Supp } \varphi_\nu$ is contained in some compact K independent of ν ,
 - (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 $\mathcal{D}^{(m)}$ and thus $T \in \mathcal{D}'^{(m)}$; T is said to be a *distribution of finite order* m . In particular, for $m = 0$, $\mathcal{D}^{(0)}$ is the space of continuous functions with compact support, and a distribution $T \in \mathcal{D}'^{(0)}$ 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:

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

This clearly results from the observation that if the φ '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 φ 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 ' δ -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 φ (because φ has compact support) but is not as φ varies.

(v) If (\mathbf{p}_ν) 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 Ω such that $\int_K |f(\mathbf{x})| d^n \mathbf{x}$ exists for any given compact K in Ω ; f is then called *locally integrable*.

The linear mapping from $\mathcal{D}(\Omega)$ 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 $\mathcal{D}(\Omega)$. It thus defines a *distribution* $T_f \in \mathcal{D}'(\Omega)$:

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

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

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 ω of Ω 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 ω on which T vanishes; or equivalently as the smallest closed subset of Ω outside which T vanishes.

When $T = T_f$ for $f \in L^1_{\text{loc}}(\Omega)$, 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 Ω , 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 φ , 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 .

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

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 φ ? 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 (φ_ν) 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 δ . If (f_ν) is a sequence of locally summable functions on \mathbb{R}^n such that

- (i) $\int_{\|\mathbf{x}\| < b} f_\nu(\mathbf{x}) d^n \mathbf{x} \rightarrow 1$ as $\nu \rightarrow \infty$ for all $b > 0$;
 - (ii) $\int_{a \leq \|\mathbf{x}\| \leq 1/a} |f_\nu(\mathbf{x})| d^n \mathbf{x} \rightarrow 0$ as $\nu \rightarrow \infty$ 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 ν ;
- then the sequence (T_{f_ν}) of distributions converges to δ 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

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

the integrated term vanishes, since φ 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,

$$\begin{aligned} \langle \partial_{ij}^2 T, \varphi \rangle &= -\langle \partial_j T, \partial_i \varphi \rangle = \langle T, \partial_{ij}^2 \varphi \rangle, \\ \langle D^{\mathbf{p}} T, \varphi \rangle &= (-1)^{|\mathbf{p}|} \langle T, D^{\mathbf{p}} \varphi \rangle, \\ \langle \Delta T, \varphi \rangle &= \langle T, \Delta \varphi \rangle, \end{aligned}$$

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 \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

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

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 $\langle \cdot, \cdot \rangle$ 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 $D^{\mathbf{p}} T_f$ 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 $Y [Y(x) = 0 \text{ for } x < 0, Y(x) = 1 \text{ for } x \geq 0]$:

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

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

1. GENERAL RELATIONSHIPS AND TECHNIQUES

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:

$$\begin{aligned} g_0 &= f - \sigma_0 Y \\ g_1 &= g'_0 - \sigma_1 Y \end{aligned}$$

$$g_k = g'_{k-1} - \sigma_k Y.$$

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:

$$(T_f)' = T_{f'} + \sigma_0 \delta$$

$$(T_f)'' = T_{f''} + \sigma_0 \delta' + \sigma_1 \delta$$

$$(T_f)^{(m)} = T_{f^{(m)}} + \sigma_0 \delta^{(m-1)} + \dots + \sigma_{m-1} \delta.$$

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

$$\langle \delta_{(S)}, \varphi \rangle = \int_S \varphi d^{n-1} S$$

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

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 $\chi \in \mathcal{D}$ 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 ψ 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 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:

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

In order that αT be a distribution, the mapping $\varphi \mapsto \alpha \varphi$ must send $\mathcal{D}(\mathbb{R}^n)$ 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 $\alpha \in \mathcal{G}^{(m)}$ 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 $\alpha T = S$.

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

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 nonsingular 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 $\varphi \in \mathcal{D}(\Omega')$ we may write:

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

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 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 $\sigma : \mathbf{x} \mapsto \mathbf{x} + \mathbf{a}$ 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 nonsingular 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 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 $L_{\text{loc}}^1(\mathbb{R}^m) \otimes L_{\text{loc}}^1(\mathbb{R}^n)$ 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

1. GENERAL RELATIONSHIPS AND TECHNIQUES

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

$S \otimes T$ 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:

$$\langle S_x \otimes T_y, \varphi_{x,y} \rangle = \langle S_x, \langle T_y, \varphi_{x,y} \rangle \rangle = \langle T_y, \langle S_x, \varphi_{x,y} \rangle \rangle.$$

The tensor product of distributions is associative:

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

Derivatives may be calculated by

$$D_x^p D_y^q (S_x \otimes T_y) = (D_x^p S_x) \otimes (D_y^q T_y).$$

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

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

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 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:

$$\langle S_x \otimes T_y, \varphi \circ \sigma \rangle$$

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

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

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 $\text{Supp}(S \otimes T) = (\text{Supp } S) \times (\text{Supp } T)$ 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

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

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_i) to convergent sequences $(S * T_i)$. 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) $\delta * T = T$: 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) $(D^p \delta) * T = D^p T$: 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: $\mathcal{D}(\mathbb{R}^n)$ 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 θ is in \mathcal{D} and is normalized), and put

$$\theta_\varepsilon(x) = \frac{1}{\varepsilon} \theta\left(\frac{x}{\varepsilon}\right) \quad \text{in dimension 1,}$$

$$\theta_\varepsilon(\mathbf{x}) = \prod_{j=1}^n \theta_\varepsilon(x_j) \quad \text{in dimension } n.$$

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

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 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 $\mathcal{F}[f]$ 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.

1.3.2.4.2. Fourier transforms in L^1

1.3.2.4.2.1. Linearity

Both transformations \mathcal{F} and $\bar{\mathcal{F}}$ are obviously linear maps from L^1 to L^∞ when these spaces are viewed as vector spaces over the field \mathbb{C} of complex numbers.

1.3.2.4.2.2. Effect of affine coordinate transformations

\mathcal{F} and $\bar{\mathcal{F}}$ turn translations into phase shifts:

$$\begin{aligned}\mathcal{F}[\tau_{\mathbf{a}} f](\xi) &= \exp(-2\pi i \xi \cdot \mathbf{a}) \mathcal{F}[f](\xi) \\ \bar{\mathcal{F}}[\tau_{\mathbf{a}} f](\xi) &= \exp(+2\pi i \xi \cdot \mathbf{a}) \bar{\mathcal{F}}[f](\xi).\end{aligned}$$

Under a general linear change of variable $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ with nonsingular matrix \mathbf{A} , the transform of $A^\#f$ is

$$\begin{aligned}\mathcal{F}[A^\#f](\xi) &= \int_{\mathbb{R}^n} f(\mathbf{A}^{-1}\mathbf{x}) \exp(-2\pi i \xi \cdot \mathbf{x}) d^n \mathbf{x} \\ &= \int_{\mathbb{R}^n} f(\mathbf{y}) \exp(-2\pi i (\mathbf{A}^T \xi) \cdot \mathbf{y}) |\det \mathbf{A}| d^n \mathbf{y} \\ &\quad \text{by } \mathbf{x} = \mathbf{A}\mathbf{y} \\ &= |\det \mathbf{A}| \mathcal{F}[f](\mathbf{A}^T \xi)\end{aligned}$$

i.e.

$$\mathcal{F}[A^\#f] = |\det \mathbf{A}| [(\mathbf{A}^{-1})^T]^\# \mathcal{F}[f]$$

and similarly for $\bar{\mathcal{F}}$. The matrix $(\mathbf{A}^{-1})^T$ is called the *contragredient* of matrix \mathbf{A} .

Under an affine change of coordinates $\mathbf{x} \mapsto S(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ with nonsingular matrix \mathbf{A} , the transform of $S^\#f$ is given by

$$\begin{aligned}\mathcal{F}[S^\#f](\xi) &= \mathcal{F}[\tau_{\mathbf{b}}(A^\#f)](\xi) \\ &= \exp(-2\pi i \xi \cdot \mathbf{b}) \mathcal{F}[A^\#f](\xi) \\ &= \exp(-2\pi i \xi \cdot \mathbf{b}) |\det \mathbf{A}| \mathcal{F}[f](\mathbf{A}^T \xi)\end{aligned}$$

with a similar result for $\bar{\mathcal{F}}$, replacing $-i$ by $+i$.

1.3.2.4.2.3. Conjugate symmetry

The kernels of the Fourier transformations \mathcal{F} and $\bar{\mathcal{F}}$ satisfy the following identities:

$$\exp(\pm 2\pi i \xi \cdot \mathbf{x}) = \exp[\overline{\pm 2\pi i \xi \cdot (-\mathbf{x})}] = \exp[\overline{\pm 2\pi i (-\xi) \cdot \mathbf{x}}].$$

As a result the transformations \mathcal{F} and $\bar{\mathcal{F}}$ themselves have the following 'conjugate symmetry' properties [where the notation $\check{f}(\mathbf{x}) = f(-\mathbf{x})$ of Section 1.3.2.2.2 will be used]:

$$\begin{aligned}\mathcal{F}[f](\xi) &= \overline{\mathcal{F}[\check{f}](-\xi)} = \overline{\mathcal{F}[\check{f}](\xi)} \\ \bar{\mathcal{F}}[f](\xi) &= \mathcal{F}[\check{f}](\xi).\end{aligned}$$

Therefore,

(i) f real $\Leftrightarrow f = \check{f} \Leftrightarrow \mathcal{F}[f] = \overline{\mathcal{F}[\check{f}]} \Leftrightarrow \mathcal{F}[f](\xi) = \overline{\mathcal{F}[f](-\xi)}$: $\mathcal{F}[f]$ is said to possess *Hermitian symmetry*;

(ii) f centrosymmetric $\Leftrightarrow f = \check{f} \Leftrightarrow \mathcal{F}[f] = \overline{\mathcal{F}[f]}$;

(iii) f real centrosymmetric $\Leftrightarrow f = \check{f} = \bar{\check{f}} \Leftrightarrow \mathcal{F}[f] = \overline{\mathcal{F}[f]} = \overline{\overline{\mathcal{F}[f]}} \Leftrightarrow \mathcal{F}[f]$ real centrosymmetric.

Conjugate symmetry is the basis of Friedel's law (Section 1.3.4.2.1.4) in crystallography.

1.3.2.4.2.4. Tensor product property

Another elementary property of \mathcal{F} is its naturality with respect to tensor products. Let $u \in L^1(\mathbb{R}^m)$ and $v \in L^1(\mathbb{R}^n)$, and let $\mathcal{F}_{\mathbf{x}}, \mathcal{F}_{\mathbf{y}}, \mathcal{F}_{\mathbf{x},\mathbf{y}}$ denote the Fourier transformations in $L^1(\mathbb{R}^m), L^1(\mathbb{R}^n)$ and $L^1(\mathbb{R}^m \times \mathbb{R}^n)$, respectively. Then

$$\mathcal{F}_{\mathbf{x},\mathbf{y}}[u \otimes v] = \mathcal{F}_{\mathbf{x}}[u] \otimes \mathcal{F}_{\mathbf{y}}[v].$$

Furthermore, if $f \in L^1(\mathbb{R}^m \times \mathbb{R}^n)$, then $\mathcal{F}_{\mathbf{y}}[f] \in L^1(\mathbb{R}^n)$ as a function of \mathbf{x} and $\mathcal{F}_{\mathbf{x}}[f] \in L^1(\mathbb{R}^m)$ as a function of \mathbf{y} , and

$$\mathcal{F}_{\mathbf{x},\mathbf{y}}[f] = \mathcal{F}_{\mathbf{x}}[\mathcal{F}_{\mathbf{y}}[f]] = \mathcal{F}_{\mathbf{y}}[\mathcal{F}_{\mathbf{x}}[f]].$$

This is easily proved by using Fubini's theorem and the fact that $(\xi, \eta) \cdot (\mathbf{x}, \mathbf{y}) = \xi \cdot \mathbf{x} + \eta \cdot \mathbf{y}$, where $\mathbf{x}, \xi \in \mathbb{R}^m, \mathbf{y}, \eta \in \mathbb{R}^n$. This property may be written:

$$\mathcal{F}_{\mathbf{x},\mathbf{y}} = \mathcal{F}_{\mathbf{x}} \otimes \mathcal{F}_{\mathbf{y}}.$$

1.3.2.4.2.5. Convolution property

If f and g are summable, their convolution $f * g$ exists and is summable, and

1. GENERAL RELATIONSHIPS AND TECHNIQUES

$$\mathcal{F}[f * g](\xi) = \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} f(\mathbf{y})g(\mathbf{x} - \mathbf{y}) d^n \mathbf{y} \right] \exp(-2\pi i \xi \cdot \mathbf{x}) d^n \mathbf{x}.$$

With $\mathbf{x} = \mathbf{y} + \mathbf{z}$, so that

$$\exp(-2\pi i \xi \cdot \mathbf{x}) = \exp(-2\pi i \xi \cdot \mathbf{y}) \exp(-2\pi i \xi \cdot \mathbf{z}),$$

and with Fubini's theorem, rearrangement of the double integral gives:

$$\mathcal{F}[f * g] = \mathcal{F}[f] \times \mathcal{F}[g]$$

and similarly

$$\bar{\mathcal{F}}[f * g] = \bar{\mathcal{F}}[f] \times \bar{\mathcal{F}}[g].$$

Thus the Fourier transform and cotransform turn convolution into multiplication.

1.3.2.4.2.6. Reciprocity property

In general, $\mathcal{F}[f]$ and $\bar{\mathcal{F}}[f]$ are not summable, and hence cannot be further transformed; however, as they are essentially bounded, their products with the Gaussians $G_t(\xi) = \exp(-2\pi^2 \|\xi\|^2 t)$ are summable for all $t > 0$, and it can be shown that

$$f = \lim_{t \rightarrow 0} \bar{\mathcal{F}}[G_t \mathcal{F}[f]] = \lim_{t \rightarrow 0} \mathcal{F}[G_t \bar{\mathcal{F}}[f]],$$

where the limit is taken in the topology of the L^1 norm $\|\cdot\|_1$. Thus \mathcal{F} and $\bar{\mathcal{F}}$ are (in a sense) mutually inverse, which justifies the common practice of calling $\bar{\mathcal{F}}$ the 'inverse Fourier transformation'.

1.3.2.4.2.7. Riemann–Lebesgue lemma

If $f \in L^1(\mathbb{R}^n)$, i.e. is summable, then $\mathcal{F}[f]$ and $\bar{\mathcal{F}}[f]$ exist and are continuous and essentially bounded:

$$\|\mathcal{F}[f]\|_\infty = \|\bar{\mathcal{F}}[f]\|_\infty \leq \|f\|_1.$$

In fact one has the much stronger property, whose statement constitutes the *Riemann–Lebesgue lemma*, that $\mathcal{F}[f](\xi)$ and $\bar{\mathcal{F}}[f](\xi)$ both tend to zero as $\|\xi\| \rightarrow \infty$.

1.3.2.4.2.8. Differentiation

Let us now suppose that $n = 1$ and that $f \in L^1(\mathbb{R})$ is differentiable with $f' \in L^1(\mathbb{R})$. Integration by parts yields

$$\begin{aligned} \mathcal{F}[f'](\xi) &= \int_{-\infty}^{+\infty} f'(x) \exp(-2\pi i \xi \cdot x) dx \\ &= [f(x) \exp(-2\pi i \xi \cdot x)]_{-\infty}^{+\infty} \\ &\quad + 2\pi i \xi \int_{-\infty}^{+\infty} f(x) \exp(-2\pi i \xi \cdot x) dx. \end{aligned}$$

Since f' is summable, f has a limit when $x \rightarrow \pm\infty$, and this limit must be 0 since f is summable. Therefore

$$\mathcal{F}[f'](\xi) = (2\pi i \xi) \mathcal{F}[f](\xi)$$

with the bound

$$\|2\pi \xi \mathcal{F}[f]\|_\infty \leq \|f'\|_1$$

so that $|\mathcal{F}[f](\xi)|$ decreases faster than $1/|\xi| \rightarrow \infty$.

This result can be easily extended to several dimensions and to any multi-index \mathbf{m} : if f is summable and has continuous summable partial derivatives up to order $|\mathbf{m}|$, then

$$\mathcal{F}[D^{\mathbf{m}}f](\xi) = (2\pi i \xi)^{\mathbf{m}} \mathcal{F}[f](\xi)$$

and

$$\|(2\pi \xi)^{\mathbf{m}} \mathcal{F}[f]\|_\infty \leq \|D^{\mathbf{m}}f\|_1.$$

Similar results hold for $\bar{\mathcal{F}}$, with $2\pi i \xi$ replaced by $-2\pi i \xi$. Thus, the more differentiable f is, with summable derivatives, the faster $\mathcal{F}[f]$ and $\bar{\mathcal{F}}[f]$ decrease at infinity.

The property of turning differentiation into multiplication by a monomial has many important applications in crystallography, for instance differential syntheses (Sections 1.3.4.2.1.9, 1.3.4.4.7.2, 1.3.4.4.7.5) and moment-generating functions [Section 1.3.4.5.2.1(c)].

1.3.2.4.2.9. Decrease at infinity

Conversely, assume that f is summable on \mathbb{R}^n and that f decreases fast enough at infinity for $\mathbf{x}^{\mathbf{m}}f$ also to be summable, for some multi-index \mathbf{m} . Then the integral defining $\mathcal{F}[f]$ may be subjected to the differential operator $D^{\mathbf{m}}$, still yielding a convergent integral: therefore $D^{\mathbf{m}}\mathcal{F}[f]$ exists, and

$$D^{\mathbf{m}}(\mathcal{F}[f])(\xi) = \mathcal{F}[(-2\pi i \mathbf{x})^{\mathbf{m}}f](\xi)$$

with the bound

$$\|D^{\mathbf{m}}(\mathcal{F}[f])\|_\infty = \|(2\pi \mathbf{x})^{\mathbf{m}}f\|_1.$$

Similar results hold for $\bar{\mathcal{F}}$, with $-2\pi i \mathbf{x}$ replaced by $2\pi i \mathbf{x}$. Thus, the faster f decreases at infinity, the more $\mathcal{F}[f]$ and $\bar{\mathcal{F}}[f]$ are differentiable, with bounded derivatives. This property is the converse of that described in Section 1.3.2.4.2.8, and their combination is fundamental in the definition of the function space \mathcal{S} in Section 1.3.2.4.4.1, of tempered distributions in Section 1.3.2.5, and in the extension of the Fourier transformation to them.

1.3.2.4.2.10. The Paley–Wiener theorem

An extreme case of the last instance occurs when f has *compact support*: then $\mathcal{F}[f]$ and $\bar{\mathcal{F}}[f]$ are so regular that they may be analytically continued from \mathbb{R}^n to \mathbb{C}^n where they are *entire* functions, i.e. have no singularities at finite distance (Paley & Wiener, 1934). This is easily seen for $\mathcal{F}[f]$: giving vector $\xi \in \mathbb{R}^n$ a vector $\boldsymbol{\eta} \in \mathbb{R}^n$ of imaginary parts leads to

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

where the latter transform always exists since $\exp(2\pi \boldsymbol{\eta} \cdot \mathbf{x})f$ is summable with respect to \mathbf{x} for all values of $\boldsymbol{\eta}$. This analytic continuation forms the basis of the saddlepoint method in probability theory [Section 1.3.4.5.2.1(f)] and leads to the use of maximum-entropy distributions in the statistical theory of direct phase determination [Section 1.3.4.5.2.2(e)].

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

By Liouville's theorem, an entire function in \mathbb{C}^n cannot vanish identically on the complement of a compact subset of \mathbb{R}^n without vanishing everywhere: therefore $\mathcal{F}[f]$ cannot have compact support if f has, and hence $\mathcal{D}(\mathbb{R}^n)$ is not stable by Fourier transformation.

1.3.2.4.3. Fourier transforms in L^2

Let f belong to $L^2(\mathbb{R}^n)$, i.e. be such that

$$\|f\|_2 = \left(\int_{\mathbb{R}^n} |f(\mathbf{x})|^2 d^n \mathbf{x} \right)^{1/2} < \infty.$$

1.3.2.4.3.1. Invariance of L^2

$\mathcal{F}[f]$ and $\bar{\mathcal{F}}[f]$ exist and are functions in L^2 , i.e. $\mathcal{F}L^2 = L^2$, $\bar{\mathcal{F}}L^2 = L^2$.

1.3.2.4.3.2. Reciprocity

$\mathcal{F}[\bar{\mathcal{F}}[f]] = f$ and $\bar{\mathcal{F}}[\mathcal{F}[f]] = f$, equality being taken as 'almost everywhere' equality. This again leads to calling $\bar{\mathcal{F}}$ the 'inverse Fourier transformation' rather than the Fourier cotransformation.

1.3.2.4.3.3. Isometry

\mathcal{F} and $\bar{\mathcal{F}}$ preserve the L^2 norm:

$$\|\mathcal{F}[f]\|_2 = \|\bar{\mathcal{F}}[f]\|_2 = \|f\|_2 \text{ (Parseval's/Plancherel's theorem).}$$

This property, which may be written in terms of the inner product (\cdot, \cdot) in $L^2(\mathbb{R}^n)$ as

$$(\mathcal{F}[f], \mathcal{F}[g]) = (\bar{\mathcal{F}}[f], \bar{\mathcal{F}}[g]) = (f, g),$$

implies that \mathcal{F} and $\bar{\mathcal{F}}$ are unitary transformations of $L^2(\mathbb{R}^n)$ into itself, i.e. infinite-dimensional 'rotations'.

1.3.2.4.3.4. Eigenspace decomposition of L^2

Some light can be shed on the geometric structure of these rotations by the following simple considerations. Note that

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

so that \mathcal{F}^4 (and similarly $\bar{\mathcal{F}}^4$) is the identity map. Any eigenvalue of \mathcal{F} or $\bar{\mathcal{F}}$ is therefore a fourth root of unity, i.e. ± 1 or $\pm i$, and $L^2(\mathbb{R}^n)$ splits into an orthogonal direct sum

$$\mathbf{H}_0 \otimes \mathbf{H}_1 \otimes \mathbf{H}_2 \otimes \mathbf{H}_3,$$

where \mathcal{F} (respectively $\bar{\mathcal{F}}$) acts in each subspace \mathbf{H}_k ($k = 0, 1, 2, 3$) by multiplication by $(-i)^k$. Orthonormal bases for these subspaces can be constructed from Hermite functions (cf. Section 1.3.2.4.4.2) This method was used by Wiener (1933, pp. 51–71).

1.3.2.4.3.5. The convolution theorem and the isometry property

In L^2 , the convolution theorem (when applicable) and the Parseval/Plancherel theorem are not independent. Suppose that $f, g, f \times g$ and $f * g$ are all in L^2 (without questioning whether these properties are independent). Then $f * g$ may be written in terms of the inner product in L^2 as follows:

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

i.e.

$$(f * g)(\mathbf{x}) = (\tau_{\mathbf{x}} \check{f}, g).$$

Invoking the isometry property, we may rewrite the right-hand side as

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

so that the initial identity yields the convolution theorem.

To obtain the converse implication, note that

$$\begin{aligned} (f, g) &= \int_{\mathbb{R}^n} \overline{f(\mathbf{y})}g(\mathbf{y}) d^n \mathbf{y} = (\check{f} * g)(\mathbf{0}) \\ &= \bar{\mathcal{F}}[\mathcal{F}[\check{f}] \times \mathcal{F}[g]](\mathbf{0}) \\ &= \int_{\mathbb{R}^n} \overline{\mathcal{F}[f](\xi)} \mathcal{F}[g](\xi) d^n \xi = (\mathcal{F}[f], \mathcal{F}[g]), \end{aligned}$$

where conjugate symmetry (Section 1.3.2.4.2.2) has been used.

These relations have an important application in the calculation by Fourier transform methods of the derivatives used in the refinement of macromolecular structures (Section 1.3.4.4.7).

1.3.2.4.4. Fourier transforms in \mathcal{S}

1.3.2.4.4.1. Definition and properties of \mathcal{S}

The duality established in Sections 1.3.2.4.2.8 and 1.3.2.4.2.9 between the local differentiability of a function and the rate of decrease at infinity of its Fourier transform prompts one to consider the space $\mathcal{S}(\mathbb{R}^n)$ of functions f on \mathbb{R}^n which are infinitely differentiable and all of whose derivatives are rapidly decreasing, so that for all multi-indices \mathbf{k} and \mathbf{p}

$$(\mathbf{x}^{\mathbf{k}} D^{\mathbf{p}} f)(\mathbf{x}) \rightarrow 0 \text{ as } \|\mathbf{x}\| \rightarrow \infty.$$

The product of $f \in \mathcal{S}$ by any polynomial over \mathbb{R}^n is still in \mathcal{S} (\mathcal{S} is an algebra over the ring of polynomials). Furthermore, \mathcal{S} is invariant under translations and differentiation.

If $f \in \mathcal{S}$, then its transforms $\mathcal{F}[f]$ and $\bar{\mathcal{F}}[f]$ are

- (i) infinitely differentiable because f is rapidly decreasing;
 - (ii) rapidly decreasing because f is infinitely differentiable;
- hence $\mathcal{F}[f]$ and $\bar{\mathcal{F}}[f]$ are in \mathcal{S} : \mathcal{S} is invariant under \mathcal{F} and $\bar{\mathcal{F}}$.

Since $L^1 \supset \mathcal{S}$ and $L^2 \supset \mathcal{S}$, all properties of \mathcal{F} and $\bar{\mathcal{F}}$ already encountered above are enjoyed by functions of \mathcal{S} , with all restrictions on differentiability and/or integrability lifted. For instance, given two functions f and g in \mathcal{S} , then both fg and $f * g$ are in \mathcal{S} (which was not the case with L^1 nor with L^2) so that the reciprocity theorem inherited from L^2

$$\mathcal{F}[\bar{\mathcal{F}}[f]] = f \quad \text{and} \quad \bar{\mathcal{F}}[\mathcal{F}[f]] = f$$

allows one to state the reverse of the convolution theorem first established in L^1 :

1. GENERAL RELATIONSHIPS AND TECHNIQUES

$$\begin{aligned}\mathcal{F}[fg] &= \mathcal{F}[f] * \mathcal{F}[g] \\ \bar{\mathcal{F}}[fg] &= \bar{\mathcal{F}}[f] * \bar{\mathcal{F}}[g].\end{aligned}$$

1.3.2.4.4.2. Gaussian functions and Hermite functions

Gaussian functions are particularly important elements of \mathcal{S} . In dimension 1, a well known contour integration (Schwartz, 1965, p. 184) yields

$$\mathcal{F}[\exp(-\pi x^2)](\xi) = \bar{\mathcal{F}}[\exp(-\pi x^2)](\xi) = \exp(-\pi \xi^2),$$

which shows that the ‘standard Gaussian’ $\exp(-\pi x^2)$ is invariant under \mathcal{F} and $\bar{\mathcal{F}}$. By a tensor product construction, it follows that the same is true of the standard Gaussian

$$G(\mathbf{x}) = \exp(-\pi \|\mathbf{x}\|^2)$$

in dimension n :

$$\mathcal{F}[G](\xi) = \bar{\mathcal{F}}[G](\xi) = G(\xi).$$

In other words, G is an eigenfunction of \mathcal{F} and $\bar{\mathcal{F}}$ for eigenvalue 1 (Section 1.3.2.4.3.4).

A complete system of eigenfunctions may be constructed as follows. In dimension 1, consider the family of functions

$$H_m = \frac{D^m G^2}{G} \quad (m \geq 0),$$

where D denotes the differentiation operator. The first two members of the family

$$H_0 = G, \quad H_1 = 2DG,$$

are such that $\mathcal{F}[H_0] = H_0$, as shown above, and

$$DG(x) = -2\pi x G(x) = i(2\pi i x)G(x) = i\mathcal{F}[DG](x),$$

hence

$$\mathcal{F}[H_1] = (-i)H_1.$$

We may thus take as an induction hypothesis that

$$\mathcal{F}[H_m] = (-i)^m H_m.$$

The identity

$$D\left(\frac{D^m G^2}{G}\right) = \frac{D^{m+1} G^2}{G} - \frac{DG}{G} \frac{D^m G^2}{G}$$

may be written

$$H_{m+1}(x) = (DH_m)(x) - 2\pi x H_m(x),$$

and the two differentiation theorems give:

$$\begin{aligned}\mathcal{F}[DH_m](\xi) &= (2\pi i \xi) \mathcal{F}[H_m](\xi) \\ \mathcal{F}[-2\pi x H_m](\xi) &= -iD(\mathcal{F}[H_m])(\xi).\end{aligned}$$

Combination of this with the induction hypothesis yields

$$\begin{aligned}\mathcal{F}[H_{m+1}](\xi) &= (-i)^{m+1} [(DH_m)(\xi) - 2\pi \xi H_m(\xi)] \\ &= (-i)^{m+1} H_{m+1}(\xi),\end{aligned}$$

thus proving that H_m is an eigenfunction of \mathcal{F} for eigenvalue $(-i)^m$ for all $m \geq 0$. The same proof holds for $\bar{\mathcal{F}}$, with eigenvalue i^m . If these eigenfunctions are normalized as

$$\mathcal{H}_m(x) = \frac{(-1)^m 2^{1/4}}{\sqrt{m! 2^m \pi^{m/2}}} H_m(x),$$

then it can be shown that the collection of *Hermite functions* $\{\mathcal{H}_m(x)\}_{m \geq 0}$ constitutes an orthonormal basis of $L^2(\mathbb{R})$ such that \mathcal{H}_m is an eigenfunction of \mathcal{F} (respectively $\bar{\mathcal{F}}$) for eigenvalue $(-i)^m$ (respectively i^m).

In dimension n , the same construction can be extended by tensor product to yield the multivariate Hermite functions

$$\mathcal{H}_{\mathbf{m}}(\mathbf{x}) = \mathcal{H}_{m_1}(x_1) \times \mathcal{H}_{m_2}(x_2) \times \dots \times \mathcal{H}_{m_n}(x_n)$$

(where $\mathbf{m} \geq \mathbf{0}$ is a multi-index). These constitute an orthonormal basis of $L^2(\mathbb{R}^n)$, with $\mathcal{H}_{\mathbf{m}}$ an eigenfunction of \mathcal{F} (respectively $\bar{\mathcal{F}}$) for eigenvalue $(-i)^{|\mathbf{m}|}$ (respectively $i^{|\mathbf{m}|}$). Thus the subspaces \mathbf{H}_k of Section 1.3.2.4.3.4 are spanned by those $\mathcal{H}_{\mathbf{m}}$ with $|\mathbf{m}| \equiv k \pmod{4}$ ($k = 0, 1, 2, 3$).

General multivariate Gaussians are usually encountered in the nonstandard form

$$G_{\mathbf{A}}(\mathbf{x}) = \exp\left(-\frac{1}{2} \mathbf{x}^T \cdot \mathbf{A} \mathbf{x}\right),$$

where \mathbf{A} is a symmetric positive-definite matrix. Diagonalizing \mathbf{A} as $\mathbf{E} \mathbf{\Lambda} \mathbf{E}^T$ with $\mathbf{E} \mathbf{E}^T$ the identity matrix, and putting $\mathbf{A}^{1/2} = \mathbf{E} \mathbf{\Lambda}^{1/2} \mathbf{E}^T$, we may write

$$G_{\mathbf{A}}(\mathbf{x}) = G\left[\left(\frac{\mathbf{A}}{2\pi}\right)^{1/2} \mathbf{x}\right]$$

i.e.

$$G_{\mathbf{A}} = [(2\pi \mathbf{A}^{-1})^{1/2}]^{\#} G;$$

hence (by Section 1.3.2.4.2.3)

$$\mathcal{F}[G_{\mathbf{A}}] = |\det(2\pi \mathbf{A}^{-1})|^{1/2} \left[\left(\frac{\mathbf{A}}{2\pi}\right)^{1/2}\right]^{\#} G,$$

i.e.

$$\mathcal{F}[G_{\mathbf{A}}](\xi) = |\det(2\pi \mathbf{A}^{-1})|^{1/2} G[(2\pi \mathbf{A}^{-1})^{1/2} \xi],$$

i.e. finally

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

$$\mathcal{F}[G_{\mathbf{A}}] = |\det(2\pi\mathbf{A}^{-1})|^{1/2} G_{4\pi^2\mathbf{A}^{-1}}.$$

This result is widely used in crystallography, *e.g.* to calculate form factors for anisotropic atoms (Section 1.3.4.2.2.6) and to obtain transforms of derivatives of Gaussian atomic densities (Section 1.3.4.4.7.10).

1.3.2.4.4.3. Heisenberg's inequality, Hardy's theorem

The result just obtained, which also holds for $\bar{\mathcal{F}}$, shows that the 'peakier' $G_{\mathbf{A}}$, the 'broader' $\mathcal{F}[G_{\mathbf{A}}]$. This is a general property of the Fourier transformation, expressed in dimension 1 by the *Heisenberg inequality* (Weyl, 1931):

$$\left(\int x^2 |f(x)|^2 dx \right) \left(\int \xi^2 |\mathcal{F}[f](\xi)|^2 d\xi \right) \geq \frac{1}{16\pi^2} \left(\int |f(x)|^2 dx \right)^2,$$

where, by a beautiful theorem of Hardy (1933), equality can only be attained for f Gaussian. Hardy's theorem is even stronger: if both f and $\mathcal{F}[f]$ behave at infinity as constant multiples of G , then each of them is *everywhere* a constant multiple of G ; if both f and $\mathcal{F}[f]$ behave at infinity as constant multiples of $G \times$ monomial, then each of them is a finite linear combination of Hermite functions. Hardy's theorem is invoked in Section 1.3.4.4.5 to derive the optimal procedure for spreading atoms on a sampling grid in order to obtain the most accurate structure factors.

The search for optimal compromises between the confinement of f to a compact domain in x -space and of $\mathcal{F}[f]$ to a compact domain in ξ -space leads to consideration of prolate spheroidal wavefunctions (Pollack & Slepian, 1961; Landau & Pollack, 1961, 1962).

1.3.2.4.4.4. Symmetry property

A final formal property of the Fourier transform, best established in \mathcal{S} , is its *symmetry*: if f and g are in \mathcal{S} , then by Fubini's theorem

$$\begin{aligned} \langle \mathcal{F}[f], g \rangle &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f(\mathbf{x}) \exp(-2\pi i \boldsymbol{\zeta} \cdot \mathbf{x}) d^n \mathbf{x} \right) g(\boldsymbol{\zeta}) d^n \boldsymbol{\zeta} \\ &= \int_{\mathbb{R}^n} f(\mathbf{x}) \left(\int_{\mathbb{R}^n} g(\boldsymbol{\zeta}) \exp(-2\pi i \boldsymbol{\zeta} \cdot \mathbf{x}) d^n \boldsymbol{\zeta} \right) d^n \mathbf{x} \\ &= \langle f, \mathcal{F}[g] \rangle. \end{aligned}$$

This possibility of 'transposing' \mathcal{F} (and $\bar{\mathcal{F}}$) from the left to the right of the duality bracket will be used in Section 1.3.2.5.4 to extend the Fourier transformation to distributions.

1.3.2.4.5. Various writings of Fourier transforms

Other ways of writing Fourier transforms in \mathbb{R}^n exist besides the one used here. All have the form

$$\bar{\mathcal{F}}_{h,\omega}[f](\boldsymbol{\zeta}) = \frac{1}{h^n} \int_{\mathbb{R}^n} f(\mathbf{x}) \exp(-i\omega \boldsymbol{\zeta} \cdot \mathbf{x}) d^n \mathbf{x},$$

where h is real positive and ω real nonzero, with the reciprocity formula written:

$$f(\mathbf{x}) = \frac{1}{k^n} \int_{\mathbb{R}^n} \bar{\mathcal{F}}_{h,\omega}[f](\boldsymbol{\zeta}) \exp(+i\omega \boldsymbol{\zeta} \cdot \mathbf{x}) d^n \boldsymbol{\zeta}$$

with k real positive. The consistency condition between h, k and ω is

$$hk = \frac{2\pi}{|\omega|}.$$

The usual choices are:

- (i) $\omega = \pm 2\pi, h = k = 1$ (as here);
- (ii) $\omega = \pm 1, h = 1, k = 2\pi$ (in probability theory and in solid-state physics);
- (iii) $\omega = \pm 1, h = k = \sqrt{2\pi}$ (in much of classical analysis).

It should be noted that conventions (ii) and (iii) introduce numerical factors of 2π in convolution and Parseval formulae, while (ii) breaks the symmetry between \mathcal{F} and $\bar{\mathcal{F}}$.

1.3.2.4.6. Tables of Fourier transforms

The books by Campbell & Foster (1948), Erdélyi (1954) and Magnus *et al.* (1966) contain extensive tables listing pairs of functions and their Fourier transforms. Bracewell (1986) lists those pairs particularly relevant to electrical engineering applications.

1.3.2.5. Fourier transforms of tempered distributions

1.3.2.5.1. Introduction

It was found in Section 1.3.2.4.2 that the usual space of test functions \mathcal{D} is not invariant under \mathcal{F} and $\bar{\mathcal{F}}$. By contrast, the space \mathcal{S} of infinitely differentiable rapidly decreasing functions *is* invariant under \mathcal{F} and $\bar{\mathcal{F}}$, and furthermore transposition formulae such as

$$\langle \mathcal{F}[f], g \rangle = \langle f, \mathcal{F}[g] \rangle$$

hold for all $f, g \in \mathcal{S}$. It is precisely this type of transposition which was used successfully in Sections 1.3.2.3.9.1 and 1.3.2.3.9.3 to define the derivatives of distributions and their products with smooth functions.

This suggests using \mathcal{S} instead of \mathcal{D} as a space of test functions φ , and defining the Fourier transform $\mathcal{F}[T]$ of a distribution T by

$$\langle \mathcal{F}[T], \varphi \rangle = \langle T, \mathcal{F}[\varphi] \rangle$$

whenever T is capable of being extended from \mathcal{D} to \mathcal{S} while remaining continuous. It is this latter proviso which will be subsumed under the adjective 'tempered'. As was the case with the construction of \mathcal{D}' , it is the definition of a sufficiently strong topology (*i.e.* notion of convergence) in \mathcal{S} which will play a key role in transferring to the elements of its topological dual \mathcal{S}' (called tempered distributions) all the properties of the Fourier transformation.

Besides the general references to distribution theory mentioned in Section 1.3.2.3.1 the reader may consult the books by Zemanian (1965, 1968). Lavoine (1963) contains tables of Fourier transforms of distributions.

1. GENERAL RELATIONSHIPS AND TECHNIQUES

1.3.2.5.2. \mathcal{S} as a test-function space

A notion of convergence has to be introduced in $\mathcal{S}(\mathbb{R}^n)$ in order to be able to define and test the continuity of linear functionals on it.

A sequence (φ_j) of functions in \mathcal{S} will be said to converge to 0 if, for any given multi-indices \mathbf{k} and \mathbf{p} , the sequence $(\mathbf{x}^{\mathbf{k}} D^{\mathbf{p}} \varphi_j)$ tends to 0 uniformly on \mathbb{R}^n .

It can be shown that $\mathcal{D}(\mathbb{R}^n)$ is dense in $\mathcal{S}(\mathbb{R}^n)$. Translation is continuous for this topology. For any linear differential operator $P(D) = \sum_{\mathbf{p}} a_{\mathbf{p}} D^{\mathbf{p}}$ and any polynomial $Q(\mathbf{x})$ over \mathbb{R}^n , $(\varphi_j) \rightarrow 0$ implies $[Q(\mathbf{x}) \times P(D)\varphi_j] \rightarrow 0$ in the topology of \mathcal{S} . Therefore, differentiation and multiplication by polynomials are continuous for the topology on \mathcal{S} .

The Fourier transformations \mathcal{F} and $\bar{\mathcal{F}}$ are also continuous for the topology of \mathcal{S} . Indeed, let (φ_j) converge to 0 for the topology on \mathcal{S} . Then, by Section 1.3.2.4.2,

$$\|(2\pi\xi)^{\mathbf{m}} D^{\mathbf{p}}(\mathcal{F}[\varphi_j])\|_{\infty} \leq \|D^{\mathbf{m}}[(2\pi\mathbf{x})^{\mathbf{p}}\varphi_j]\|_1.$$

The right-hand side tends to 0 as $j \rightarrow \infty$ by definition of convergence in \mathcal{S} , hence $\|\xi\|^{\mathbf{m}} D^{\mathbf{p}}(\mathcal{F}[\varphi_j]) \rightarrow 0$ uniformly, so that $(\mathcal{F}[\varphi_j]) \rightarrow 0$ in \mathcal{S} as $j \rightarrow \infty$. The same proof applies to $\bar{\mathcal{F}}$.

1.3.2.5.3. Definition and examples of tempered distributions

A distribution $T \in \mathcal{D}'(\mathbb{R}^n)$ is said to be *tempered* if it can be extended into a continuous linear functional on \mathcal{S} .

If $\mathcal{S}'(\mathbb{R}^n)$ is the topological dual of $\mathcal{S}(\mathbb{R}^n)$, and if $S \in \mathcal{S}'(\mathbb{R}^n)$, then its restriction to \mathcal{D} is a tempered distribution; conversely, if $T \in \mathcal{D}'$ is tempered, then its extension to \mathcal{S} is unique (because \mathcal{D} is dense in \mathcal{S}), hence it defines an element S of \mathcal{S}' . We may therefore identify \mathcal{S}' and the space of tempered distributions.

A distribution with compact support is tempered, *i.e.* $\mathcal{S}' \supset \mathcal{E}'$. By transposition of the corresponding properties of \mathcal{S} , it is readily established that the derivative, translate or product by a polynomial of a tempered distribution is still a tempered distribution.

These inclusion relations may be summarized as follows: since \mathcal{S} contains \mathcal{D} but is contained in \mathcal{E} , the reverse inclusions hold for the topological duals, and hence \mathcal{S}' contains \mathcal{E}' but is contained in \mathcal{D}' .

A locally summable function f on \mathbb{R}^n will be said to be of *polynomial growth* if $|f(\mathbf{x})|$ can be majorized by a polynomial in $\|\mathbf{x}\|$ as $\|\mathbf{x}\| \rightarrow \infty$. It is easily shown that such a function f defines a tempered distribution T_f via

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

In particular, polynomials over \mathbb{R}^n define tempered distributions, and so do functions in \mathcal{S} . The latter remark, together with the transposition identity (Section 1.3.2.4.4), invites the extension of \mathcal{F} and $\bar{\mathcal{F}}$ from \mathcal{S} to \mathcal{S}' .

1.3.2.5.4. Fourier transforms of tempered distributions

The Fourier transform $\mathcal{F}[T]$ and cotransform $\bar{\mathcal{F}}[T]$ of a tempered distribution T are defined by

$$\begin{aligned} \langle \mathcal{F}[T], \varphi \rangle &= \langle T, \mathcal{F}[\varphi] \rangle \\ \langle \bar{\mathcal{F}}[T], \varphi \rangle &= \langle T, \bar{\mathcal{F}}[\varphi] \rangle \end{aligned}$$

for all test functions $\varphi \in \mathcal{S}$. Both $\mathcal{F}[T]$ and $\bar{\mathcal{F}}[T]$ are themselves tempered distributions, since the maps $\varphi \mapsto \mathcal{F}[\varphi]$ and $\varphi \mapsto \bar{\mathcal{F}}[\varphi]$ are both linear and continuous for the topology of \mathcal{S} . In the same way that \mathbf{x} and ξ have been used consistently as

arguments for φ and $\mathcal{F}[\varphi]$, respectively, the notation $T_{\mathbf{x}}$ and $\mathcal{F}[T]_{\xi}$ will be used to indicate which variables are involved.

When T is a distribution with compact support, its Fourier transform may be written

$$\mathcal{F}[T_{\mathbf{x}}]_{\xi} = \langle T_{\mathbf{x}}, \exp(-2\pi i \xi \cdot \mathbf{x}) \rangle$$

since the function $\mathbf{x} \mapsto \exp(-2\pi i \xi \cdot \mathbf{x})$ is in \mathcal{E} while $T_{\mathbf{x}} \in \mathcal{E}'$. It can be shown, as in Section 1.3.2.4.2, to be analytically continuable into an entire function over \mathbb{C}^n .

1.3.2.5.5. Transposition of basic properties

The duality between differentiation and multiplication by a monomial extends from \mathcal{S} to \mathcal{S}' by transposition:

$$\begin{aligned} \mathcal{F}[D_{\mathbf{x}}^{\mathbf{p}} T_{\mathbf{x}}]_{\xi} &= (2\pi i \xi)^{\mathbf{p}} \mathcal{F}[T_{\mathbf{x}}]_{\xi} \\ D_{\xi}^{\mathbf{p}} (\mathcal{F}[T_{\mathbf{x}}]_{\xi}) &= \mathcal{F}[(-2\pi i \mathbf{x})^{\mathbf{p}} T_{\mathbf{x}}]_{\xi}. \end{aligned}$$

Analogous formulae hold for $\bar{\mathcal{F}}$, with i replaced by $-i$.

The formulae expressing the duality between translation and phase shift, *e.g.*

$$\begin{aligned} \mathcal{F}[\tau_{\mathbf{a}} T_{\mathbf{x}}]_{\xi} &= \exp(-2\pi i \mathbf{a} \cdot \xi) \mathcal{F}[T_{\mathbf{x}}]_{\xi} \\ \tau_{\mathbf{a}} (\mathcal{F}[T_{\mathbf{x}}]_{\xi}) &= \mathcal{F}[\exp(2\pi i \mathbf{a} \cdot \mathbf{x}) T_{\mathbf{x}}]_{\xi}; \end{aligned}$$

between a linear change of variable and its contragredient, *e.g.*

$$\mathcal{F}[A^{\#} T] = |\det \mathbf{A}| [(\mathbf{A}^{-1})^T]^{\#} \mathcal{F}[T];$$

are obtained similarly by transposition from the corresponding identities in \mathcal{S} . They give a transposition formula for an affine change of variables $\mathbf{x} \mapsto S(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ with nonsingular matrix \mathbf{A} :

$$\begin{aligned} \mathcal{F}[S^{\#} T] &= \exp(-2\pi i \xi \cdot \mathbf{b}) \mathcal{F}[A^{\#} T] \\ &= \exp(-2\pi i \xi \cdot \mathbf{b}) |\det \mathbf{A}| [(\mathbf{A}^{-1})^T]^{\#} \mathcal{F}[T], \end{aligned}$$

with a similar result for $\bar{\mathcal{F}}$, replacing $-i$ by $+i$.

Conjugate symmetry is obtained similarly:

$$\mathcal{F}[\bar{T}] = \overline{\mathcal{F}[T]}, \quad \bar{\mathcal{F}}[\bar{T}] = \overline{\mathcal{F}[T]},$$

with the same identities for $\bar{\mathcal{F}}$.

The tensor product property also transposes to tempered distributions: if $U \in \mathcal{S}'(\mathbb{R}^m)$, $V \in \mathcal{S}'(\mathbb{R}^n)$,

$$\begin{aligned} \mathcal{F}[U_{\mathbf{x}} \otimes V_{\mathbf{y}}] &= \mathcal{F}[U]_{\xi} \otimes \mathcal{F}[V]_{\eta} \\ \bar{\mathcal{F}}[U_{\mathbf{x}} \otimes V_{\mathbf{y}}] &= \bar{\mathcal{F}}[U]_{\xi} \otimes \bar{\mathcal{F}}[V]_{\eta}. \end{aligned}$$

1.3.2.5.6. Transforms of δ -functions

Since δ has compact support,

$$\mathcal{F}[\delta_{\mathbf{x}}]_{\xi} = \langle \delta_{\mathbf{x}}, \exp(-2\pi i \xi \cdot \mathbf{x}) \rangle = 1_{\xi}, \quad \text{i.e. } \mathcal{F}[\delta] = 1.$$

It is instructive to show that conversely $\mathcal{F}[1] = \delta$ without invoking the reciprocity theorem. Since $\partial_j 1 = 0$ for all $j = 1, \dots, n$, it follows from Section 1.3.2.3.9.4 that $\mathcal{F}[1] = c\delta$; the

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

constant c can be determined by using the invariance of the standard Gaussian G established in Section 1.3.2.4.3:

$$\langle \mathcal{F}[1]_{\mathbf{x}}, G_{\mathbf{x}} \rangle = \langle 1_{\xi}, G_{\xi} \rangle = 1;$$

hence $c = 1$. Thus, $\mathcal{F}[1] = \delta$.

The basic properties above then read (using multi-indices to denote differentiation):

$$\begin{aligned} \mathcal{F}[\delta_{\mathbf{x}}^{(\mathbf{m})}]_{\xi} &= (2\pi i \xi)^{\mathbf{m}}, & \mathcal{F}[\mathbf{x}^{\mathbf{m}}]_{\xi} &= (-2\pi i)^{-|\mathbf{m}|} \delta_{\xi}^{(\mathbf{m})}; \\ \mathcal{F}[\delta_{\mathbf{a}}]_{\xi} &= \exp(-2\pi i \mathbf{a} \cdot \xi), & \mathcal{F}[\exp(2\pi i \mathbf{a} \cdot \mathbf{x})]_{\xi} &= \delta_{\mathbf{a}}, \end{aligned}$$

with analogous relations for $\bar{\mathcal{F}}$, i becoming $-i$. Thus derivatives of δ are mapped to monomials (and *vice versa*), while translates of δ are mapped to ‘phase factors’ (and *vice versa*).

1.3.2.5.7. Reciprocity theorem

The previous results now allow a self-contained and rigorous proof of the reciprocity theorem between \mathcal{F} and $\bar{\mathcal{F}}$ to be given, whereas in traditional settings (*i.e.* in L^1 and L^2) the implicit handling of δ through a limiting process is always the sticking point.

Reciprocity is first established in \mathcal{S} as follows:

$$\begin{aligned} \bar{\mathcal{F}}[\mathcal{F}[\varphi]](\mathbf{x}) &= \int_{\mathbb{R}^n} \mathcal{F}[\varphi](\xi) \exp(2\pi i \xi \cdot \mathbf{x}) \, d^n \xi \\ &= \int_{\mathbb{R}^n} \mathcal{F}[\tau_{-\mathbf{x}}\varphi](\xi) \, d^n \xi \\ &= \langle 1, \mathcal{F}[\tau_{-\mathbf{x}}\varphi] \rangle \\ &= \langle \mathcal{F}[1], \tau_{-\mathbf{x}}\varphi \rangle \\ &= \langle \tau_{\mathbf{x}}\delta, \varphi \rangle \\ &= \varphi(\mathbf{x}) \end{aligned}$$

and similarly

$$\mathcal{F}[\bar{\mathcal{F}}[\varphi]](\mathbf{x}) = \varphi(\mathbf{x}).$$

The reciprocity theorem is then proved in \mathcal{S}' by transposition:

$$\bar{\mathcal{F}}[\mathcal{F}[T]] = \mathcal{F}[\bar{\mathcal{F}}[T]] = T \quad \text{for all } T \in \mathcal{S}'.$$

Thus the Fourier cotransformation $\bar{\mathcal{F}}$ in \mathcal{S}' may legitimately be called the ‘inverse Fourier transformation’.

The method of Section 1.3.2.4.3 may then be used to show that \mathcal{F} and $\bar{\mathcal{F}}$ both have period 4 in \mathcal{S}' .

1.3.2.5.8. Multiplication and convolution

Multiplier functions $\alpha(\mathbf{x})$ for tempered distributions must be infinitely differentiable, as for ordinary distributions; furthermore, they must grow sufficiently slowly as $\|\mathbf{x}\| \rightarrow \infty$ to ensure that $\alpha\varphi \in \mathcal{S}$ for all $\varphi \in \mathcal{S}$ and that the map $\varphi \mapsto \alpha\varphi$ is continuous for the topology of \mathcal{S} . This leads to choosing for multipliers the subspace \mathcal{O}_M consisting of functions $\alpha \in \mathcal{E}$ of *polynomial growth*. It can be shown that if f is in \mathcal{O}_M , then the associated distribution T_f is in \mathcal{S}' (*i.e.* is a tempered distribution); and that conversely if T is in \mathcal{S}' , $\mu * T$ is in \mathcal{O}_M for all $\mu \in \mathcal{D}$.

Corresponding restrictions must be imposed to define the space \mathcal{O}'_C of those distributions T whose convolution $S * T$ with a tempered distribution S is still a tempered distribution: T must be such that, for all $\varphi \in \mathcal{S}$, $\theta(\mathbf{x}) = \langle T_{\mathbf{y}}, \varphi(\mathbf{x} + \mathbf{y}) \rangle$ is in \mathcal{S} ; and such that the map $\varphi \mapsto \theta$ be continuous for the topology of \mathcal{S} . This implies that S is ‘rapidly decreasing’. It can be shown that if f is in

\mathcal{S} , then the associated distribution T_f is in \mathcal{O}'_C ; and that conversely if T is in \mathcal{O}'_C , $\mu * T$ is in \mathcal{S} for all $\mu \in \mathcal{D}$.

The two spaces \mathcal{O}_M and \mathcal{O}'_C are mapped into each other by the Fourier transformation

$$\begin{aligned} \mathcal{F}(\mathcal{O}_M) &= \bar{\mathcal{F}}(\mathcal{O}_M) = \mathcal{O}'_C \\ \mathcal{F}(\mathcal{O}'_C) &= \bar{\mathcal{F}}(\mathcal{O}'_C) = \mathcal{O}_M \end{aligned}$$

and the convolution theorem takes the form

$$\begin{aligned} \mathcal{F}[\alpha S] &= \mathcal{F}[\alpha] * \mathcal{F}[S] & S \in \mathcal{S}', \alpha \in \mathcal{O}_M, \mathcal{F}[\alpha] \in \mathcal{O}'_C; \\ \mathcal{F}[S * T] &= \mathcal{F}[S] \times \mathcal{F}[T] & S \in \mathcal{S}', T \in \mathcal{O}'_C, \mathcal{F}[T] \in \mathcal{O}_M. \end{aligned}$$

The same identities hold for $\bar{\mathcal{F}}$. Taken together with the reciprocity theorem, these show that \mathcal{F} and $\bar{\mathcal{F}}$ establish mutually inverse isomorphisms between \mathcal{O}_M and \mathcal{O}'_C , and exchange multiplication for convolution in \mathcal{S}' .

It may be noticed that most of the basic properties of \mathcal{F} and $\bar{\mathcal{F}}$ may be deduced from this theorem and from the properties of δ . Differentiation operators $D^{\mathbf{m}}$ and translation operators $\tau_{\mathbf{a}}$ are convolutions with $D^{\mathbf{m}}\delta$ and $\tau_{\mathbf{a}}\delta$; they are turned, respectively, into multiplication by monomials $(\pm 2\pi i \xi)^{\mathbf{m}}$ (the transforms of $D^{\mathbf{m}}\delta$) or by phase factors $\exp(\pm 2\pi i \xi \cdot \mathbf{a})$ (the transforms of $\tau_{\mathbf{a}}\delta$).

Another consequence of the convolution theorem is the duality established by the Fourier transformation between sections and projections of a function and its transform. For instance, in \mathbb{R}^3 , the *projection* of $f(x, y, z)$ on the x, y plane along the z axis may be written

$$(\delta_x \otimes \delta_y \otimes 1_z) * f;$$

its Fourier transform is then

$$(1_{\xi} \otimes 1_{\eta} \otimes \delta_{\zeta}) \times \mathcal{F}[f],$$

which is the *section* of $\mathcal{F}[f]$ by the plane $\zeta = 0$, orthogonal to the z axis used for projection. There are numerous applications of this property in crystallography (Section 1.3.4.2.1.8) and in fibre diffraction (Section 1.3.4.5.1.3).

1.3.2.5.9. L^2 aspects, Sobolev spaces

The special properties of \mathcal{F} in the space of square-integrable functions $L^2(\mathbb{R}^n)$, such as Parseval’s identity, can be accommodated within distribution theory: if $u \in L^2(\mathbb{R}^n)$, then T_u is a tempered distribution in \mathcal{S}' (the map $u \mapsto T_u$ being continuous) and it can be shown that $S = \mathcal{F}[T_u]$ is of the form S_v , where $u = \mathcal{F}[u]$ is the Fourier transform of u in $L^2(\mathbb{R}^n)$. By Plancherel’s theorem, $\|u\|_2 = \|v\|_2$.

This embedding of L^2 into \mathcal{S}' can be used to derive the convolution theorem for L^2 . If u and v are in $L^2(\mathbb{R}^n)$, then $u * v$ can be shown to be a bounded continuous function; thus $u * v$ is not in L^2 , but it is in \mathcal{S}' , so that its Fourier transform is a distribution, and

$$\mathcal{F}[u * v] = \mathcal{F}[u] \times \mathcal{F}[v].$$

Spaces of tempered distributions related to $L^2(\mathbb{R}^n)$ can be defined as follows. For any real s , define the Sobolev space $H_s(\mathbb{R}^n)$ to consist of all tempered distributions $S \in \mathcal{S}'(\mathbb{R}^n)$ such that

$$(1 + |\xi|^2)^{s/2} \mathcal{F}[S]_{\xi} \in L^2(\mathbb{R}^n).$$

1. GENERAL RELATIONSHIPS AND TECHNIQUES

These spaces play a fundamental role in the theory of partial differential equations, and in the mathematical theory of tomographic reconstruction – a subject not unrelated to the crystallographic phase problem (Natterer, 1986).

1.3.2.6. Periodic distributions and Fourier series

1.3.2.6.1. Terminology

Let \mathbb{Z}^n be the subset of \mathbb{R}^n consisting of those points with (signed) integer coordinates; it is an n -dimensional *lattice*, i.e. a free Abelian group on n generators. A particularly simple set of n generators is given by the standard basis of \mathbb{R}^n , and hence \mathbb{Z}^n will be called the *standard lattice in \mathbb{R}^n* . Any other ‘nonstandard’ n -dimensional lattice Λ in \mathbb{R}^n is the image of this standard lattice by a general linear transformation.

If we identify any two points in \mathbb{R}^n whose coordinates are congruent modulo \mathbb{Z}^n , i.e. differ by a vector in \mathbb{Z}^n , we obtain the *standard n -torus $\mathbb{R}^n/\mathbb{Z}^n$* . The latter may be viewed as $(\mathbb{R}/\mathbb{Z})^n$, i.e. as the Cartesian product of n circles. The same identification may be carried out modulo a nonstandard lattice Λ , yielding a *nonstandard n -torus \mathbb{R}^n/Λ* . The correspondence to crystallographic terminology is that ‘standard’ coordinates over the standard 3-torus $\mathbb{R}^3/\mathbb{Z}^3$ are called ‘fractional’ coordinates over the unit cell; while Cartesian coordinates, e.g. in ångströms, constitute a set of nonstandard coordinates.

Finally, we will denote by I the unit cube $[0, 1]^n$ and by C_ε the subset

$$C_\varepsilon = \{\mathbf{x} \in \mathbb{R}^n \mid |x_j| < \varepsilon \text{ for all } j = 1, \dots, n\}.$$

1.3.2.6.2. \mathbb{Z}^n -periodic distributions in \mathbb{R}^n

A distribution $T \in \mathcal{D}'(\mathbb{R}^n)$ is called *periodic with period lattice \mathbb{Z}^n* (or \mathbb{Z}^n -periodic) if $\tau_{\mathbf{m}}T = T$ for all $\mathbf{m} \in \mathbb{Z}^n$ (in crystallography the period lattice is the *direct* lattice).

Given a distribution with compact support $T^0 \in \mathcal{E}'(\mathbb{R}^n)$, then $T = \sum_{\mathbf{m} \in \mathbb{Z}^n} \tau_{\mathbf{m}}T^0$ is a \mathbb{Z}^n -periodic distribution. Note that we may write $T = r * T^0$, where $r = \sum_{\mathbf{m} \in \mathbb{Z}^n} \delta_{(\mathbf{m})}$ consists of Dirac δ 's at all nodes of the period lattice \mathbb{Z}^n .

Conversely, any \mathbb{Z}^n -periodic distribution T may be written as $r * T^0$ for some $T^0 \in \mathcal{E}'$. To retrieve such a ‘motif’ T^0 from T , a function ψ will be constructed in such a way that $\psi \in \mathcal{D}$ (hence has compact support) and $r * \psi = 1$; then $T^0 = \psi T$. Indicator functions (Section 1.3.2.2) such as χ_1 or $\chi_{C_{1/2}}$ cannot be used directly, since they are discontinuous; but regularized versions of them may be constructed by convolution (see Section 1.3.2.3.9.7) as $\psi_0 = \chi_{C_\varepsilon} * \theta_\eta$, with ε and η such that $\psi_0(\mathbf{x}) = 1$ on $C_{1/2}$ and $\psi_0(\mathbf{x}) = 0$ outside $C_{3/4}$. Then the function

$$\psi = \frac{\psi_0}{\sum_{\mathbf{m} \in \mathbb{Z}^n} \tau_{\mathbf{m}}\psi_0}$$

has the desired property. The sum in the denominator contains at most 2^n nonzero terms at any given point \mathbf{x} and acts as a smoothly varying ‘multiplicity correction’.

1.3.2.6.3. Identification with distributions over $\mathbb{R}^n/\mathbb{Z}^n$

Throughout this section, ‘periodic’ will mean ‘ \mathbb{Z}^n -periodic’.

Let $s \in \mathbb{R}$, and let $[s]$ denote the largest integer $\leq s$. For $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, let $\tilde{\mathbf{x}}$ be the unique vector $(\tilde{x}_1, \dots, \tilde{x}_n)$ with $\tilde{x}_j = x_j - [x_j]$. If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then $\tilde{\mathbf{x}} = \tilde{\mathbf{y}}$ if and only if $\mathbf{x} - \mathbf{y} \in \mathbb{Z}^n$. The image of the map $\mathbf{x} \mapsto \tilde{\mathbf{x}}$ is thus \mathbb{R}^n modulo \mathbb{Z}^n , or $\mathbb{R}^n/\mathbb{Z}^n$.

If f is a periodic function over \mathbb{R}^n , then $\tilde{\mathbf{x}} = \tilde{\mathbf{y}}$ implies $f(\mathbf{x}) = f(\mathbf{y})$; we may thus define a function \tilde{f} over $\mathbb{R}^n/\mathbb{Z}^n$ by putting $\tilde{f}(\tilde{\mathbf{x}}) = f(\mathbf{x})$ for any $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{x} - \tilde{\mathbf{x}} \in \mathbb{Z}^n$. Conversely, if \tilde{f} is a function over $\mathbb{R}^n/\mathbb{Z}^n$, then we may define a function f over \mathbb{R}^n by putting $f(\mathbf{x}) = \tilde{f}(\tilde{\mathbf{x}})$, and f will be

periodic. Periodic functions over \mathbb{R}^n may thus be identified with functions over $\mathbb{R}^n/\mathbb{Z}^n$, and this identification preserves the notions of convergence, local summability and differentiability.

Given $\varphi^0 \in \mathcal{D}(\mathbb{R}^n)$, we may define

$$\varphi(\mathbf{x}) = \sum_{\mathbf{m} \in \mathbb{Z}^n} (\tau_{\mathbf{m}}\varphi^0)(\mathbf{x})$$

since the sum only contains finitely many nonzero terms; φ is periodic, and $\tilde{\varphi} \in \mathcal{D}(\mathbb{R}^n/\mathbb{Z}^n)$. Conversely, if $\tilde{\varphi} \in \mathcal{D}(\mathbb{R}^n/\mathbb{Z}^n)$ we may define $\varphi \in \mathcal{E}(\mathbb{R}^n)$ periodic by $\varphi(\mathbf{x}) = \tilde{\varphi}(\tilde{\mathbf{x}})$, and $\varphi^0 \in \mathcal{D}(\mathbb{R}^n)$ by putting $\varphi^0 = \psi\varphi$ with ψ constructed as above.

By transposition, a distribution $\tilde{T} \in \mathcal{D}'(\mathbb{R}^n/\mathbb{Z}^n)$ defines a unique periodic distribution $T \in \mathcal{D}'(\mathbb{R}^n)$ by $\langle T, \varphi^0 \rangle = \langle \tilde{T}, \tilde{\varphi} \rangle$; conversely, $T \in \mathcal{D}'(\mathbb{R}^n)$ periodic defines uniquely $\tilde{T} \in \mathcal{D}'(\mathbb{R}^n/\mathbb{Z}^n)$ by $\langle \tilde{T}, \tilde{\varphi} \rangle = \langle T, \varphi^0 \rangle$.

We may therefore identify \mathbb{Z}^n -periodic distributions over \mathbb{R}^n with distributions over $\mathbb{R}^n/\mathbb{Z}^n$. We will, however, use mostly the former presentation, as it is more closely related to the crystallographer’s perception of periodicity (see Section 1.3.4.1).

1.3.2.6.4. Fourier transforms of periodic distributions

The content of this section is perhaps the central result in the relation between Fourier theory and crystallography (Section 1.3.4.2.1.1).

Let $T = r * T^0$ with r defined as in Section 1.3.2.6.2. Then $r \in \mathcal{D}'$, $T^0 \in \mathcal{E}'$ hence $T^0 \in \mathcal{O}'_C$, so that $T \in \mathcal{D}'$: \mathbb{Z}^n -periodic distributions are tempered, hence have a Fourier transform. The convolution theorem (Section 1.3.2.5.8) is applicable, giving:

$$\mathcal{F}[T] = \mathcal{F}[r] \times \mathcal{F}[T^0]$$

and similarly for $\tilde{\mathcal{F}}$.

Since $\mathcal{F}[\delta_{(\mathbf{m})}](\xi) = \exp(-2\pi i \xi \cdot \mathbf{m})$, formally

$$\mathcal{F}[r]_\xi = \sum_{\mathbf{m} \in \mathbb{Z}^n} \exp(-2\pi i \xi \cdot \mathbf{m}) = Q,$$

say.

It is readily shown that Q is tempered and periodic, so that $Q = \sum_{\mu \in \mathbb{Z}^n} \tau_\mu(\psi Q)$, while the periodicity of r implies that

$$[\exp(-2\pi i \xi_j) - 1]\psi Q = 0, \quad j = 1, \dots, n.$$

Since the first factors have single isolated zeros at $\xi_j = 0$ in $C_{3/4}$, $\psi Q = c\delta$ (see Section 1.3.2.3.9.4) and hence by periodicity $Q = cr$; convoluting with χ_{C_1} shows that $c = 1$. Thus we have the fundamental result:

$$\boxed{\mathcal{F}[r] = r}$$

so that

$$\mathcal{F}[T] = r \times \mathcal{F}[T^0];$$

i.e., according to Section 1.3.2.3.9.3,

$$\mathcal{F}[T]_\xi = \sum_{\mu \in \mathbb{Z}^n} \mathcal{F}[T^0](\mu) \times \delta_{(\mu)}.$$

The right-hand side is a *weighted* lattice distribution, whose nodes $\mu \in \mathbb{Z}^n$ are weighted by the *sample values* $\mathcal{F}[T^0](\mu)$ of the transform of the motif T^0 at those nodes. Since $T^0 \in \mathcal{E}'$, the latter values may be written

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

$$\mathcal{F}[T^0](\boldsymbol{\mu}) = \langle T_{\mathbf{x}}^0, \exp(-2\pi i \boldsymbol{\mu} \cdot \mathbf{x}) \rangle.$$

By the structure theorem for distributions with compact support (Section 1.3.2.3.9.7), T^0 is a derivative of finite order of a continuous function; therefore, from Section 1.3.2.4.2.8 and Section 1.3.2.5.8, $\mathcal{F}[T^0](\boldsymbol{\mu})$ grows at most polynomially as $\|\boldsymbol{\mu}\| \rightarrow \infty$ (see also Section 1.3.2.6.10.3 about this property). Conversely, let $W = \sum_{\boldsymbol{\mu} \in \mathbb{Z}^n} w_{\boldsymbol{\mu}} \delta_{(\boldsymbol{\mu})}$ be a weighted lattice distribution such that the weights $w_{\boldsymbol{\mu}}$ grow at most polynomially as $\|\boldsymbol{\mu}\| \rightarrow \infty$. Then W is a tempered distribution, whose Fourier cotransform $T_{\mathbf{x}} = \sum_{\boldsymbol{\mu} \in \mathbb{Z}^n} w_{\boldsymbol{\mu}} \exp(+2\pi i \boldsymbol{\mu} \cdot \mathbf{x})$ is periodic. If T is now written as $r * T^0$ for some $T^0 \in \mathcal{E}'$, then by the reciprocity theorem

$$w_{\boldsymbol{\mu}} = \mathcal{F}[T^0](\boldsymbol{\mu}) = \langle T_{\mathbf{x}}^0, \exp(-2\pi i \boldsymbol{\mu} \cdot \mathbf{x}) \rangle.$$

Although the choice of T^0 is not unique, and need not yield back the same motif as may have been used to build T initially, different choices of T^0 will lead to the same coefficients $w_{\boldsymbol{\mu}}$ because of the periodicity of $\exp(-2\pi i \boldsymbol{\mu} \cdot \mathbf{x})$.

The Fourier transformation thus establishes a duality between periodic distributions and weighted lattice distributions. The pair of relations

$$\begin{aligned} \text{(i)} \quad & w_{\boldsymbol{\mu}} = \langle T_{\mathbf{x}}^0, \exp(-2\pi i \boldsymbol{\mu} \cdot \mathbf{x}) \rangle \\ \text{(ii)} \quad & T_{\mathbf{x}} = \sum_{\boldsymbol{\mu} \in \mathbb{Z}^n} w_{\boldsymbol{\mu}} \exp(+2\pi i \boldsymbol{\mu} \cdot \mathbf{x}) \end{aligned}$$

are referred to as the *Fourier analysis* and the *Fourier synthesis* of T , respectively (there is a discrepancy between this terminology and the crystallographic one, see Section 1.3.4.2.1.1). In other words, any periodic distribution $T \in \mathcal{S}'$ may be represented by a Fourier series (ii), whose coefficients are calculated by (i). The convergence of (ii) towards T in \mathcal{S}' will be investigated later (Section 1.3.2.6.10).

1.3.2.6.5. The case of nonstandard period lattices

Let Λ denote the nonstandard lattice consisting of all vectors of the form $\sum_{j=1}^n m_j \mathbf{a}_j$, where the m_j are rational integers and $\mathbf{a}_1, \dots, \mathbf{a}_n$ are n linearly independent vectors in \mathbb{R}^n . Let R be the corresponding lattice distribution: $R = \sum_{\mathbf{x} \in \Lambda} \delta_{(\mathbf{x})}$.

Let \mathbf{A} be the nonsingular $n \times n$ matrix whose successive columns are the coordinates of vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$ in the standard basis of \mathbb{R}^n ; \mathbf{A} will be called the *period matrix* of Λ , and the mapping $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ will be denoted by A . According to Section 1.3.2.3.9.5 we have

$$\langle R, \varphi \rangle = \sum_{\mathbf{m} \in \mathbb{Z}^n} \varphi(\mathbf{A}\mathbf{m}) = \langle r, (A^{-1})^{\#} \varphi \rangle = |\det \mathbf{A}|^{-1} \langle A^{\#} r, \varphi \rangle$$

for any $\varphi \in \mathcal{S}$, and hence $R = |\det \mathbf{A}|^{-1} A^{\#} r$. By Fourier transformation, according to Section 1.3.2.5.5,

$$\mathcal{F}[R] = |\det \mathbf{A}|^{-1} \mathcal{F}[A^{\#} r] = [(\mathbf{A}^{-1})^T]^{\#} \mathcal{F}[r] = [(\mathbf{A}^{-1})^T]^{\#} r,$$

which we write:

$$\mathcal{F}[R] = |\det \mathbf{A}|^{-1} R^*$$

with

$$R^* = |\det \mathbf{A}| [(\mathbf{A}^{-1})^T]^{\#} r.$$

R^* is a lattice distribution:

$$R^* = \sum_{\boldsymbol{\mu} \in \mathbb{Z}^n} \delta_{[(\mathbf{A}^{-1})^T \boldsymbol{\mu}]} = \sum_{\boldsymbol{\xi} \in \Lambda^*} \delta_{(\boldsymbol{\xi})}$$

associated with the *reciprocal lattice* Λ^* whose basis vectors $\mathbf{a}_1^*, \dots, \mathbf{a}_n^*$ are the columns of $(\mathbf{A}^{-1})^T$. Since the latter matrix is equal to the adjoint matrix (*i.e.* the matrix of co-factors) of \mathbf{A} divided by $\det \mathbf{A}$, the components of the reciprocal basis vectors can be written down explicitly (see Section 1.3.4.2.1.1 for the crystallographic case $n = 3$).

A distribution T will be called Λ -*periodic* if $\tau_{\boldsymbol{\xi}} T = T$ for all $\boldsymbol{\xi} \in \Lambda$; as previously, T may be written $R * T^0$ for some motif distribution T^0 with compact support. By Fourier transformation,

$$\begin{aligned} \mathcal{F}[T] &= |\det \mathbf{A}|^{-1} R^* \cdot \mathcal{F}[T^0] \\ &= |\det \mathbf{A}|^{-1} \sum_{\boldsymbol{\xi} \in \Lambda^*} \mathcal{F}[T^0](\boldsymbol{\xi}) \delta_{(\boldsymbol{\xi})} \\ &= |\det \mathbf{A}|^{-1} \sum_{\boldsymbol{\mu} \in \mathbb{Z}^n} \mathcal{F}[T^0][(\mathbf{A}^{-1})^T \boldsymbol{\mu}] \delta_{[(\mathbf{A}^{-1})^T \boldsymbol{\mu}]} \end{aligned}$$

so that $\mathcal{F}[T]$ is a weighted reciprocal-lattice distribution, the weight attached to node $\boldsymbol{\xi} \in \Lambda^*$ being $|\det \mathbf{A}|^{-1}$ times the value $\mathcal{F}[T^0](\boldsymbol{\xi})$ of the Fourier transform of the motif T^0 .

This result may be further simplified if T and its motif T^0 are referred to the standard period lattice \mathbb{Z}^n by defining t and t^0 so that $T = A^{\#} t$, $T^0 = A^{\#} t^0$, $t = r * t^0$. Then

$$\mathcal{F}[T^0](\boldsymbol{\xi}) = |\det \mathbf{A}| \mathcal{F}[t^0](\mathbf{A}^T \boldsymbol{\xi}),$$

hence

$$\mathcal{F}[T^0][(\mathbf{A}^{-1})^T \boldsymbol{\mu}] = |\det \mathbf{A}| \mathcal{F}[t^0](\boldsymbol{\mu}),$$

so that

$$\mathcal{F}[T] = \sum_{\boldsymbol{\mu} \in \mathbb{Z}^n} \mathcal{F}[t^0](\boldsymbol{\mu}) \delta_{[(\mathbf{A}^{-1})^T \boldsymbol{\mu}]}$$

in nonstandard coordinates, while

$$\mathcal{F}[t] = \sum_{\boldsymbol{\mu} \in \mathbb{Z}^n} \mathcal{F}[t^0](\boldsymbol{\mu}) \delta_{(\boldsymbol{\mu})}$$

in standard coordinates.

The reciprocity theorem may then be written:

$$\begin{aligned} \text{(iii)} \quad & W_{\boldsymbol{\xi}} = |\det \mathbf{A}|^{-1} \langle T_{\mathbf{x}}^0, \exp(-2\pi i \boldsymbol{\xi} \cdot \mathbf{x}) \rangle, \quad \boldsymbol{\xi} \in \Lambda^* \\ \text{(iv)} \quad & T_{\mathbf{x}} = \sum_{\boldsymbol{\xi} \in \Lambda^*} W_{\boldsymbol{\xi}} \exp(+2\pi i \boldsymbol{\xi} \cdot \mathbf{x}) \end{aligned}$$

in nonstandard coordinates, or equivalently:

$$\begin{aligned} \text{(v)} \quad & w_{\boldsymbol{\mu}} = \langle t_{\mathbf{x}}^0, \exp(-2\pi i \boldsymbol{\mu} \cdot \mathbf{x}) \rangle, \quad \boldsymbol{\mu} \in \mathbb{Z}^n \\ \text{(vi)} \quad & t_{\mathbf{x}} = \sum_{\boldsymbol{\mu} \in \mathbb{Z}^n} w_{\boldsymbol{\mu}} \exp(+2\pi i \boldsymbol{\mu} \cdot \mathbf{x}) \end{aligned}$$

1. GENERAL RELATIONSHIPS AND TECHNIQUES

in standard coordinates. It gives an n -dimensional Fourier series representation for any periodic distribution over \mathbb{R}^n . The convergence of such series in $\mathcal{S}'(\mathbb{R}^n)$ will be examined in Section 1.3.2.6.10.

1.3.2.6.6. Duality between periodization and sampling

Let T^0 be a distribution with compact support (the ‘motif’). Its Fourier transform $\tilde{\mathcal{F}}[T^0]$ is analytic (Section 1.3.2.5.4) and may thus be used as a multiplier.

We may rephrase the preceding results as follows:

- (i) if T^0 is ‘periodized by R ’ to give $R * T^0$, then $\tilde{\mathcal{F}}[T^0]$ is ‘sampled by R^* ’ to give $|\det \mathbf{A}|^{-1} R^* \cdot \tilde{\mathcal{F}}[T^0]$;
- (ii) if $\tilde{\mathcal{F}}[T^0]$ is ‘sampled by R^* ’ to give $R^* \cdot \tilde{\mathcal{F}}[T^0]$, then T^0 is ‘periodized by R ’ to give $|\det \mathbf{A}| R * T^0$.

Thus the Fourier transformation establishes a duality between the periodization of a distribution by a period lattice Λ and the sampling of its transform at the nodes of lattice Λ^* reciprocal to Λ . This is a particular instance of the convolution theorem of Section 1.3.2.5.8.

At this point it is traditional to break the symmetry between \mathcal{F} and $\tilde{\mathcal{F}}$ which distribution theory has enabled us to preserve even in the presence of periodicity, and to perform two distinct identifications:

- (i) a Λ -periodic distribution T will be handled as a distribution \tilde{T} on \mathbb{R}^n / Λ , was done in Section 1.3.2.6.3;
- (ii) a weighted lattice distribution $W = \sum_{\mu \in \mathbb{Z}^n} W_\mu \delta_{[(\mathbf{A}^{-1})^T \mu]}$ will be identified with the collection $\{W_\mu | \mu \in \mathbb{Z}^n\}$ of its n -tuply indexed coefficients.

1.3.2.6.7. The Poisson summation formula

Let $\varphi \in \mathcal{S}$, so that $\mathcal{F}[\varphi] \in \mathcal{S}$. Let R be the lattice distribution associated to lattice Λ , with period matrix \mathbf{A} , and let R^* be associated to the reciprocal lattice Λ^* . Then we may write:

$$\begin{aligned} \langle R, \varphi \rangle &= \langle R, \tilde{\mathcal{F}}[\mathcal{F}[\varphi]] \rangle \\ &= \langle \tilde{\mathcal{F}}[R], \mathcal{F}[\varphi] \rangle \\ &= |\det \mathbf{A}|^{-1} \langle R^*, \mathcal{F}[\varphi] \rangle \end{aligned}$$

i.e.

$$\sum_{\mathbf{x} \in \Lambda} \varphi(\mathbf{x}) = |\det \mathbf{A}|^{-1} \sum_{\xi \in \Lambda^*} \mathcal{F}[\varphi](\xi).$$

This identity, which also holds for $\tilde{\mathcal{F}}$, is called the *Poisson summation formula*. Its usefulness follows from the fact that the speed of decrease at infinity of φ and $\mathcal{F}[\varphi]$ are inversely related (Section 1.3.2.4.4.3), so that if one of the series (say, the left-hand side) is slowly convergent, the other (say, the right-hand side) will be rapidly convergent. This procedure has been used by Ewald (1921) [see also Bertaut (1952), Born & Huang (1954)] to evaluate lattice sums (Madelung constants) involved in the calculation of the internal electrostatic energy of crystals (see Chapter 3.4 in this volume on convergence acceleration techniques for crystallographic lattice sums and Chapter 3.5 on modern extensions of the Ewald summation method).

When φ is a multivariate Gaussian

$$\varphi(\mathbf{x}) = G_{\mathbf{B}}(\mathbf{x}) = \exp(-\frac{1}{2} \mathbf{x}^T \mathbf{B} \mathbf{x}),$$

then

$$\mathcal{F}[\varphi](\xi) = |\det(2\pi \mathbf{B}^{-1})|^{1/2} G_{\mathbf{B}^{-1}}(\xi),$$

and Poisson’s summation formula for a lattice with period matrix \mathbf{A} reads:

$$\begin{aligned} \sum_{\mathbf{m} \in \mathbb{Z}^n} G_{\mathbf{B}}(\mathbf{A} \mathbf{m}) &= |\det \mathbf{A}|^{-1} |\det(2\pi \mathbf{B}^{-1})|^{1/2} \\ &\times \sum_{\mu \in \mathbb{Z}^n} G_{4\pi^2 \mathbf{B}^{-1}}[(\mathbf{A}^{-1})^T \mu] \end{aligned}$$

or equivalently

$$\sum_{\mathbf{m} \in \mathbb{Z}^n} G_{\mathbf{C}}(\mathbf{m}) = |\det(2\pi \mathbf{C}^{-1})|^{1/2} \sum_{\mu \in \mathbb{Z}^n} G_{4\pi^2 \mathbf{C}^{-1}}(\mu)$$

with $\mathbf{C} = \mathbf{A}^T \mathbf{B} \mathbf{A}$.

1.3.2.6.8. Convolution of Fourier series

Let $S = R * S^0$ and $T = R * T^0$ be two Λ -periodic distributions, the motifs S^0 and T^0 having compact support. The convolution $S * T$ does not exist, because S and T do not satisfy the support condition (Section 1.3.2.3.9.7). However, the three distributions R , S^0 and T^0 do satisfy the generalized support condition, so that their convolution is defined; then, by associativity and commutativity:

$$R * S^0 * T^0 = S * T^0 = S^0 * T.$$

By Fourier transformation and by the convolution theorem:

$$\begin{aligned} R^* \times \mathcal{F}[S^0 * T^0] &= (R^* \times \mathcal{F}[S^0]) \times \mathcal{F}[T^0] \\ &= \mathcal{F}[T^0] \times (R^* \times \mathcal{F}[S^0]). \end{aligned}$$

Let $\{U_\xi\}_{\xi \in \Lambda^*}$, $\{V_\xi\}_{\xi \in \Lambda^*}$ and $\{W_\xi\}_{\xi \in \Lambda^*}$ be the sets of Fourier coefficients associated to S , T and $S * T^0 (= S^0 * T)$, respectively. Identifying the coefficients of δ_ξ for $\xi \in \Lambda^*$ yields the forward version of the convolution theorem for Fourier series:

$$W_\xi = |\det \mathbf{A}| U_\xi V_\xi.$$

The backward version of the theorem requires that T be infinitely differentiable. The distribution $S \times T$ is then well defined and its Fourier coefficients $\{Q_\xi\}_{\xi \in \Lambda^*}$ are given by

$$Q_\xi = \sum_{\eta \in \Lambda^*} U_\eta V_{\xi - \eta}.$$

1.3.2.6.9. Toeplitz forms, Szegő’s theorem

Toeplitz forms were first investigated by Toeplitz (1907, 1910, 1911a). They occur in connection with the ‘trigonometric moment problem’ (Shohat & Tamarkin, 1943; Akhiezer, 1965) and probability theory (Grenander, 1952) and play an important role in several direct approaches to the crystallographic phase problem [see Sections 1.3.4.2.1.10, 1.3.4.5.2.2(e)]. Many aspects of their theory and applications are presented in the book by Grenander & Szegő (1958).

1.3.2.6.9.1. Toeplitz forms

Let $f \in L^1(\mathbb{R}/\mathbb{Z})$ be real-valued, so that its Fourier coefficients satisfy the relations $c_{-m}(f) = \overline{c_m(f)}$. The Hermitian form in $n + 1$ complex variables

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

$$T_n[f](\mathbf{u}) = \sum_{\mu=0}^n \sum_{\nu=0}^n \overline{u_\mu} c_{\mu-\nu} u_\nu$$

is called the n th *Toeplitz form* associated to f . It is a straightforward consequence of the convolution theorem and of Parseval's identity that $T_n[f]$ may be written:

$$T_n[f](\mathbf{u}) = \int_0^1 \left| \sum_{\nu=0}^n u_\nu \exp(2\pi i \nu x) \right|^2 f(x) dx.$$

1.3.2.6.9.2. The Toeplitz–Carathéodory–Herglotz theorem

It was shown independently by Toeplitz (1911b), Carathéodory (1911) and Herglotz (1911) that a function $f \in L^1$ is almost everywhere non-negative if and only if the Toeplitz forms $T_n[f]$ associated to f are positive semidefinite for all values of n .

This is equivalent to the infinite system of determinantal inequalities

$$D_n = \det \begin{pmatrix} c_0 & c_{-1} & \cdot & \cdot & c_{-n} \\ c_1 & c_0 & c_{-1} & \cdot & \cdot \\ \cdot & c_1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & c_{-1} \\ c_n & \cdot & \cdot & c_1 & c_0 \end{pmatrix} \geq 0 \quad \text{for all } n.$$

The D_n are called *Toeplitz determinants*. Their application to the crystallographic phase problem is described in Section 1.3.4.2.1.10.

1.3.2.6.9.3. Asymptotic distribution of eigenvalues of Toeplitz forms

The eigenvalues of the Hermitian form $T_n[f]$ are defined as the $n+1$ real roots of the characteristic equation $\det\{T_n[f - \lambda]\} = 0$. They will be denoted by

$$\lambda_1^{(n)}, \lambda_2^{(n)}, \dots, \lambda_{n+1}^{(n)}.$$

It is easily shown that if $m \leq f(x) \leq M$ for all x , then $m \leq \lambda_\nu^{(n)} \leq M$ for all n and all $\nu = 1, \dots, n+1$. As $n \rightarrow \infty$ these bounds, and the distribution of the $\lambda^{(n)}$ within these bounds, can be made more precise by introducing two new notions.

(i) *Essential bounds*: define $\text{ess inf } f$ as the largest m such that $f(x) \geq m$ except for values of x forming a set of measure 0; and define $\text{ess sup } f$ similarly.

(ii) *Equal distribution*. For each n , consider two sets of $n+1$ real numbers:

$$a_1^{(n)}, a_2^{(n)}, \dots, a_{n+1}^{(n)}, \quad \text{and} \quad b_1^{(n)}, b_2^{(n)}, \dots, b_{n+1}^{(n)}.$$

Assume that for each ν and each n , $|a_\nu^{(n)}| < K$ and $|b_\nu^{(n)}| < K$ with K independent of ν and n . The sets $\{a_\nu^{(n)}\}$ and $\{b_\nu^{(n)}\}$ are said to be equally distributed in $[-K, +K]$ if, for any function F over $[-K, +K]$,

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{\nu=1}^{n+1} [F(a_\nu^{(n)}) - F(b_\nu^{(n)})] = 0.$$

We may now state an important theorem of Szegő (1915, 1920). Let $f \in L^1$, and put $m = \text{ess inf } f$, $M = \text{ess sup } f$. If m and M are

finite, then for any continuous function $F(\lambda)$ defined in the interval $[m, M]$ we have

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{\nu=1}^{n+1} F(\lambda_\nu^{(n)}) = \int_0^1 F[f(x)] dx.$$

In other words, the eigenvalues $\lambda_\nu^{(n)}$ of the T_n and the values $f[\nu/(n+2)]$ of f on a regular subdivision of $]0, 1[$ are equally distributed.

Further investigations into the spectra of Toeplitz matrices may be found in papers by Hartman & Wintner (1950, 1954), Kac *et al.* (1953), Widom (1965), and in the notes by Hirschman & Hughes (1977).

1.3.2.6.9.4. Consequences of Szegő's theorem

(i) If the λ 's are ordered in ascending order, then

$$\lim_{n \rightarrow \infty} \lambda_1^{(n)} = m = \text{ess inf } f, \quad \lim_{n \rightarrow \infty} \lambda_{n+1}^{(n)} = M = \text{ess sup } f.$$

Thus, when $f \geq 0$, the condition number $\lambda_{n+1}^{(n)}/\lambda_1^{(n)}$ of $T_n[f]$ tends towards the 'essential dynamic range' M/m of f .

(ii) Let $F(\lambda) = \lambda^s$ where s is a positive integer. Then

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{\nu=1}^{n+1} [\lambda_\nu^{(n)}]^s = \int_0^1 [f(x)]^s dx.$$

(iii) Let $m > 0$, so that $\lambda_\nu^{(n)} > 0$, and let $D_n(f) = \det T_n(f)$. Then

$$D_n(f) = \prod_{\nu=1}^{n+1} \lambda_\nu^{(n)},$$

hence

$$\log D_n(f) = \sum_{\nu=1}^{n+1} \log \lambda_\nu^{(n)}.$$

Putting $F(\lambda) = \log \lambda$, it follows that

$$\lim_{n \rightarrow \infty} [D_n(f)]^{1/(n+1)} = \exp \left\{ \int_0^1 \log f(x) dx \right\}.$$

Further terms in this limit were obtained by Szegő (1952) and interpreted in probabilistic terms by Kac (1954).

1.3.2.6.10. Convergence of Fourier series

The investigation of the convergence of Fourier series and of more general trigonometric series has been the subject of intense study for over 150 years [see *e.g.* Zygmund (1976)]. It has been a constant source of new mathematical ideas and theories, being directly responsible for the birth of such fields as set theory, topology and functional analysis.

This section will briefly survey those aspects of the classical results in dimension 1 which are relevant to the practical use of Fourier series in crystallography. The books by Zygmund (1959), Tolstov (1962) and Katznelson (1968) are standard references in the field, and Dym & McKean (1972) is recommended as a stimulant.

1. GENERAL RELATIONSHIPS AND TECHNIQUES

1.3.2.6.10.1. Classical L^1 theory

The space $L^1(\mathbb{R}/\mathbb{Z})$ consists of (equivalence classes of) complex-valued functions f on the circle which are summable, *i.e.* for which

$$\|f\|_1 \equiv \int_0^1 |f(x)| dx < +\infty.$$

It is a convolution algebra: If f and g are in L^1 , then $f * g$ is in L^1 . The m th Fourier coefficient $c_m(f)$ of f ,

$$c_m(f) = \int_0^1 f(x) \exp(-2\pi imx) dx$$

is bounded: $|c_m(f)| \leq \|f\|_1$, and by the Riemann–Lebesgue lemma $c_m(f) \rightarrow 0$ as $m \rightarrow \infty$. By the convolution theorem, $c_m(f * g) = c_m(f)c_m(g)$.

The p th partial sum $S_p(f)$ of the Fourier series of f ,

$$S_p(f)(x) = \sum_{|m| \leq p} c_m(f) \exp(2\pi imx),$$

may be written, by virtue of the convolution theorem, as $S_p(f) = D_p * f$, where

$$D_p(x) = \sum_{|m| \leq p} \exp(2\pi imx) = \frac{\sin[(2p+1)\pi x]}{\sin \pi x}$$

is the *Dirichlet kernel*. Because D_p comprises numerous slowly decaying oscillations, both positive and negative, $S_p(f)$ may not converge towards f in a strong sense as $p \rightarrow \infty$. Indeed, spectacular pathologies are known to exist where the partial sums, examined pointwise, diverge everywhere (Zygmund, 1959, Chapter VIII). When f is piecewise continuous, but presents isolated jumps, convergence near these jumps is marred by the *Gibbs phenomenon*: $S_p(f)$ always ‘overshoots the mark’ by about 9%, the area under the spurious peak tending to 0 as $p \rightarrow \infty$ but not its height [see Larmor (1934) for the history of this phenomenon].

By contrast, the *arithmetic mean* of the partial sums, also called the p th Cesàro sum,

$$C_p(f) = \frac{1}{p+1} [S_0(f) + \dots + S_p(f)],$$

converges to f in the sense of the L^1 norm: $\|C_p(f) - f\|_1 \rightarrow 0$ as $p \rightarrow \infty$. If furthermore f is *continuous*, then the convergence is *uniform*, *i.e.* the error is bounded everywhere by a quantity which goes to 0 as $p \rightarrow \infty$. It may be shown that

$$C_p(f) = F_p * f,$$

where

$$\begin{aligned} F_p(x) &= \sum_{|m| \leq p} \left(1 - \frac{|m|}{p+1}\right) \exp(2\pi imx) \\ &= \frac{1}{p+1} \left[\frac{\sin(p+1)\pi x}{\sin \pi x} \right]^2 \end{aligned}$$

is the *Fejér kernel*. F_p has over D_p the advantage of being everywhere positive, so that the Cesàro sums $C_p(f)$ of a positive function f are always positive.

The de la Vallée Poussin kernel

$$V_p(x) = 2F_{2p+1}(x) - F_p(x)$$

has a trapezoidal distribution of coefficients and is such that $c_m(V_p) = 1$ if $|m| \leq p+1$; therefore $V_p * f$ is a trigonometric polynomial with the same Fourier coefficients as f over that range of values of m .

The Poisson kernel

$$\begin{aligned} P_r(x) &= 1 + 2 \sum_{m=1}^{\infty} r^m \cos 2\pi mx \\ &= \frac{1 - r^2}{1 - 2r \cos 2\pi mx + r^2} \end{aligned}$$

with $0 \leq r < 1$ gives rise to an Abel summation procedure [Tolstov (1962, p. 162); Whittaker & Watson (1927, p. 57)] since

$$(P_r * f)(x) = \sum_{m \in \mathbb{Z}} c_m(f) r^{|m|} \exp(2\pi imx).$$

Compared with the other kernels, P_r has the disadvantage of not being a trigonometric polynomial; however, P_r is the real part of the Cauchy kernel (Cartan, 1961; Ahlfors, 1966):

$$P_r(x) = \Re \left[\frac{1 + r \exp(2\pi ix)}{1 - r \exp(2\pi ix)} \right]$$

and hence provides a link between trigonometric series and analytic functions of a complex variable.

Other methods of summation involve forming a moving average of f by convolution with other sequences of functions $\alpha_p(x)$ besides D_p or F_p which ‘tend towards δ ’ as $p \rightarrow \infty$. The convolution is performed by multiplying the Fourier coefficients of f by those of α_p , so that one forms the quantities

$$S'_p(f)(x) = \sum_{|m| \leq p} c_m(\alpha_p) c_m(f) \exp(2\pi imx).$$

For instance the ‘sigma factors’ of Lanczos (Lanczos, 1966, p. 65), defined by

$$\sigma_m = \frac{\sin[m\pi/p]}{m\pi/p},$$

lead to a summation procedure whose behaviour is intermediate between those using the Dirichlet and the Fejér kernels; it corresponds to forming a moving average of f by convolution with

$$\alpha_p = p \chi_{[-1/(2p), 1/(2p)]} * D_p,$$

which is itself the convolution of a ‘rectangular pulse’ of width $1/p$ and of the Dirichlet kernel of order p .

A review of the summation problem in crystallography is given in Section 1.3.4.2.1.3.

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

1.3.2.6.10.2. Classical L^2 theory

The space $L^2(\mathbb{R}/\mathbb{Z})$ (of equivalence classes of) square-integrable complex-valued functions f on the circle is contained in $L^1(\mathbb{R}/\mathbb{Z})$, since by the Cauchy–Schwarz inequality

$$\begin{aligned} \|f\|_1^2 &= \left(\int_0^1 |f(x)| \times 1 \, dx \right)^2 \\ &\leq \left(\int_0^1 |f(x)|^2 \, dx \right) \left(\int_0^1 1^2 \, dx \right) = \|f\|_2^2 \leq \infty. \end{aligned}$$

Thus all the results derived for L^1 hold for L^2 , a great simplification over the situation in \mathbb{R} or \mathbb{R}^n where neither L^1 nor L^2 was contained in the other.

However, more can be proved in L^2 , because L^2 is a Hilbert space (Section 1.3.2.2.4) for the inner product

$$(f, g) = \int_0^1 \overline{f(x)}g(x) \, dx,$$

and because the family of functions $\{\exp(2\pi imx)\}_{m \in \mathbb{Z}}$ constitutes an orthonormal Hilbert basis for L^2 .

The sequence of Fourier coefficients $c_m(f)$ of $f \in L^2$ belongs to the space $\ell^2(\mathbb{Z})$ of square-summable sequences:

$$\sum_{m \in \mathbb{Z}} |c_m(f)|^2 < \infty.$$

Conversely, every element $c = (c_m)$ of ℓ^2 is the sequence of Fourier coefficients of a unique function in L^2 . The inner product

$$(c, d) = \sum_{m \in \mathbb{Z}} \overline{c_m}d_m$$

makes ℓ^2 into a Hilbert space, and the map from L^2 to ℓ^2 established by the Fourier transformation is an isometry (Parseval/Plancherel):

$$\|f\|_{L^2} = \|c(f)\|_{\ell^2}$$

or equivalently:

$$(f, g) = (c(f), c(g)).$$

This is a useful property in applications, since (f, g) may be calculated either from f and g themselves, or from their Fourier coefficients $c(f)$ and $c(g)$ (see Section 1.3.4.4.6) for crystallographic applications).

By virtue of the orthogonality of the basis $\{\exp(2\pi imx)\}_{m \in \mathbb{Z}}$, the partial sum $S_p(f)$ is the best mean-square fit to f in the linear subspace of L^2 spanned by $\{\exp(2\pi imx)\}_{|m| \leq p}$, and hence (Bessel's inequality)

$$\sum_{|m| \leq p} |c_m(f)|^2 = \|f\|_2^2 - \sum_{|M| \geq p} |c_M(f)|^2 \leq \|f\|_2^2.$$

1.3.2.6.10.3. The viewpoint of distribution theory

The use of distributions enlarges considerably the range of behaviour which can be accommodated in a Fourier series, even in the case of general dimension n where classical theories meet with even more difficulties than in dimension 1.

Let $\{w_m\}_{m \in \mathbb{Z}}$ be a sequence of complex numbers with $|w_m|$ growing at most polynomially as $|m| \rightarrow \infty$, say $|w_m| \leq C|m|^K$. Then the sequence $\{w_m/(2\pi im)^{K+2}\}_{m \in \mathbb{Z}}$ is in ℓ^2 and even defines a *continuous* function $f \in L^2(\mathbb{R}/\mathbb{Z})$ and an associated tempered distribution $T_f \in \mathcal{D}'(\mathbb{R}/\mathbb{Z})$. Differentiation of T_f ($K+2$) times then yields a tempered distribution whose Fourier transform leads to the original sequence of coefficients. Conversely, by the structure theorem for distributions with compact support (Section 1.3.2.3.9.7), the motif T^0 of a \mathbb{Z} -periodic distribution is a derivative of finite order of a continuous function; hence its Fourier coefficients will grow at most polynomially with $|m|$ as $|m| \rightarrow \infty$.

Thus distribution theory allows the manipulation of Fourier series whose coefficients exhibit polynomial growth as their order goes to infinity, while those derived from functions had to tend to 0 by virtue of the Riemann–Lebesgue lemma. The distribution-theoretic approach to Fourier series holds even in the case of general dimension n , where classical theories meet with even more difficulties (see Ash, 1976) than in dimension 1.

1.3.2.7. The discrete Fourier transformation

1.3.2.7.1. Shannon's sampling theorem and interpolation formula

Let $\varphi \in \mathcal{E}(\mathbb{R}^n)$ be such that $\Phi = \mathcal{F}[\varphi]$ has compact support K . Let φ be sampled at the nodes of a lattice Λ^* , yielding the lattice distribution $R^* \times \varphi$. The Fourier transform of this sampled version of φ is

$$\mathcal{F}[R^* \times \varphi] = |\det \mathbf{A}|(R * \Phi),$$

which is essentially Φ periodized by period lattice $\Lambda = (\Lambda^*)^*$, with period matrix \mathbf{A} .

Let us assume that Λ is such that the translates of K by different period vectors of Λ are disjoint. Then we may recover Φ from $R * \Phi$ by masking the contents of a 'unit cell' \mathcal{V} of Λ (*i.e.* a fundamental domain for the action of Λ in \mathbb{R}^n) whose boundary does not meet K . If $\chi_{\mathcal{V}}$ is the indicator function of \mathcal{V} , then

$$\Phi = \chi_{\mathcal{V}} \times (R * \Phi).$$

Transforming both sides by $\bar{\mathcal{F}}$ yields

$$\varphi = \bar{\mathcal{F}} \left[\chi_{\mathcal{V}} \times \frac{1}{|\det \mathbf{A}|} \mathcal{F}[R^* \times \varphi] \right],$$

i.e.

$$\varphi = \left(\frac{1}{V} \bar{\mathcal{F}}[\chi_{\mathcal{V}}] \right) * (R^* \times \varphi)$$

since $|\det \mathbf{A}|$ is the volume V of \mathcal{V} .

This interpolation formula is traditionally credited to Shannon (1949), although it was discovered much earlier by Whittaker (1915). It shows that φ may be recovered from its sample values on Λ^* (*i.e.* from $R^* \times \varphi$) provided Λ^* is sufficiently fine that no overlap (or 'aliasing') occurs in the periodization of Φ by the dual lattice Λ . The interpolation kernel is the transform of the normalized indicator function of a unit cell of Λ containing the support K of Φ .

If K is contained in a sphere of radius $1/\Delta$ and if Λ and Λ^* are rectangular, the length of each basis vector of Λ must be greater than $2/\Delta$, and thus the sampling interval must be smaller than $\Delta/2$. This requirement constitutes the Shannon sampling criterion.

1. GENERAL RELATIONSHIPS AND TECHNIQUES

1.3.2.7.2. *Duality between subdivision and decimation of period lattices*

1.3.2.7.2.1. *Geometric description of sublattices*

Let $\Lambda_{\mathbf{A}}$ be a period lattice in \mathbb{R}^n with matrix \mathbf{A} , and let $\Lambda_{\mathbf{A}}^*$ be the lattice reciprocal to $\Lambda_{\mathbf{A}}$, with period matrix $(\mathbf{A}^{-1})^T$. Let $\Lambda_{\mathbf{B}}, \mathbf{B}, \Lambda_{\mathbf{B}}^*$ be defined similarly, and let us suppose that $\Lambda_{\mathbf{A}}$ is a sublattice of $\Lambda_{\mathbf{B}}$, i.e. that $\Lambda_{\mathbf{B}} \supset \Lambda_{\mathbf{A}}$ as a set.

The relation between $\Lambda_{\mathbf{A}}$ and $\Lambda_{\mathbf{B}}$ may be described in two different fashions: (i) multiplicatively, and (ii) additively.

(i) We may write $\mathbf{A} = \mathbf{B}\mathbf{N}$ for some nonsingular matrix \mathbf{N} with integer entries. \mathbf{N} may be viewed as the period matrix of the coarser lattice $\Lambda_{\mathbf{A}}$ with respect to the period basis of the finer lattice $\Lambda_{\mathbf{B}}$. It will be more convenient to write $\mathbf{A} = \mathbf{D}\mathbf{B}$, where $\mathbf{D} = \mathbf{B}\mathbf{N}\mathbf{B}^{-1}$ is a rational matrix (with integer determinant since $\det \mathbf{D} = \det \mathbf{N}$) in terms of which the two lattices are related by

$$\Lambda_{\mathbf{A}} = \mathbf{D}\Lambda_{\mathbf{B}}.$$

(ii) Call two vectors in $\Lambda_{\mathbf{B}}$ *congruent modulo* $\Lambda_{\mathbf{A}}$ if their difference lies in $\Lambda_{\mathbf{A}}$. Denote the set of congruence classes (or ‘cosets’) by $\Lambda_{\mathbf{B}}/\Lambda_{\mathbf{A}}$, and the number of these classes by $[\Lambda_{\mathbf{B}} : \Lambda_{\mathbf{A}}]$. The ‘coset decomposition’

$$\Lambda_{\mathbf{B}} = \bigcup_{\ell \in \Lambda_{\mathbf{B}}/\Lambda_{\mathbf{A}}} (\ell + \Lambda_{\mathbf{A}})$$

represents $\Lambda_{\mathbf{B}}$ as the disjoint union of $[\Lambda_{\mathbf{B}} : \Lambda_{\mathbf{A}}]$ translates of $\Lambda_{\mathbf{A}}$. $\Lambda_{\mathbf{B}}/\Lambda_{\mathbf{A}}$ is a finite lattice with $[\Lambda_{\mathbf{B}} : \Lambda_{\mathbf{A}}]$ elements, called the *residual lattice* of $\Lambda_{\mathbf{B}}$ modulo $\Lambda_{\mathbf{A}}$.

The two descriptions are connected by the relation $[\Lambda_{\mathbf{B}} : \Lambda_{\mathbf{A}}] = \det \mathbf{D} = \det \mathbf{N}$, which follows from a volume calculation. We may also combine (i) and (ii) into

$$(iii) \quad \Lambda_{\mathbf{B}} = \bigcup_{\ell \in \Lambda_{\mathbf{B}}/\Lambda_{\mathbf{A}}} (\ell + \mathbf{D}\Lambda_{\mathbf{B}})$$

which may be viewed as the n -dimensional equivalent of the Euclidean algorithm for integer division: ℓ is the ‘remainder’ of the division by $\Lambda_{\mathbf{A}}$ of a vector in $\Lambda_{\mathbf{B}}$, the quotient being the matrix \mathbf{D} .

1.3.2.7.2.2. *Sublattice relations for reciprocal lattices*

Let us now consider the two *reciprocal lattices* $\Lambda_{\mathbf{A}}^*$ and $\Lambda_{\mathbf{B}}^*$. Their period matrices $(\mathbf{A}^{-1})^T$ and $(\mathbf{B}^{-1})^T$ are related by: $(\mathbf{B}^{-1})^T = (\mathbf{A}^{-1})^T \mathbf{N}^T$, where \mathbf{N}^T is an integer matrix; or equivalently by $(\mathbf{B}^{-1})^T = \mathbf{D}^T (\mathbf{A}^{-1})^T$. This shows that the roles are reversed in that $\Lambda_{\mathbf{B}}^*$ is a sublattice of $\Lambda_{\mathbf{A}}^*$, which we may write:

$$(i)^* \quad \Lambda_{\mathbf{B}}^* = \mathbf{D}^T \Lambda_{\mathbf{A}}^*$$

$$(ii)^* \quad \Lambda_{\mathbf{A}}^* = \bigcup_{\ell^* \in \Lambda_{\mathbf{A}}^*/\Lambda_{\mathbf{B}}^*} (\ell^* + \Lambda_{\mathbf{B}}^*).$$

The residual lattice $\Lambda_{\mathbf{A}}^*/\Lambda_{\mathbf{B}}^*$ is finite, with $[\Lambda_{\mathbf{A}}^* : \Lambda_{\mathbf{B}}^*] = \det \mathbf{D} = \det \mathbf{N} = [\Lambda_{\mathbf{B}} : \Lambda_{\mathbf{A}}]$, and we may again combine (i)^{*} and (ii)^{*} into

$$(iii)^* \quad \Lambda_{\mathbf{A}}^* = \bigcup_{\ell^* \in \Lambda_{\mathbf{A}}^*/\Lambda_{\mathbf{B}}^*} (\ell^* + \mathbf{D}^T \Lambda_{\mathbf{A}}^*).$$

1.3.2.7.2.3. *Relation between lattice distributions*

The above relations between lattices may be rewritten in terms of the corresponding *lattice distributions* as follows:

$$(i) \quad R_{\mathbf{A}} = \frac{1}{|\det \mathbf{D}|} \mathbf{D}^{\#} R_{\mathbf{B}}^*$$

$$(ii) \quad R_{\mathbf{B}} = T_{\mathbf{B}/\mathbf{A}} * R_{\mathbf{A}}$$

$$(i)^* \quad R_{\mathbf{B}}^* = \frac{1}{|\det \mathbf{D}|} (\mathbf{D}^T)^{\#} R_{\mathbf{A}}^*$$

$$(ii)^* \quad R_{\mathbf{A}}^* = T_{\mathbf{A}/\mathbf{B}}^* * R_{\mathbf{B}}^*$$

where

$$T_{\mathbf{B}/\mathbf{A}} = \sum_{\ell \in \Lambda_{\mathbf{B}}/\Lambda_{\mathbf{A}}} \delta_{(\ell)}$$

and

$$T_{\mathbf{A}/\mathbf{B}}^* = \sum_{\ell^* \in \Lambda_{\mathbf{A}}^*/\Lambda_{\mathbf{B}}^*} \delta_{(\ell^*)}$$

are (finite) residual-lattice distributions. We may incorporate the factor $1/|\det \mathbf{D}|$ in (i) and (i)^{*} into these distributions and define

$$S_{\mathbf{B}/\mathbf{A}} = \frac{1}{|\det \mathbf{D}|} T_{\mathbf{B}/\mathbf{A}}, \quad S_{\mathbf{A}/\mathbf{B}}^* = \frac{1}{|\det \mathbf{D}|} T_{\mathbf{A}/\mathbf{B}}^*.$$

Since $|\det \mathbf{D}| = [\Lambda_{\mathbf{B}} : \Lambda_{\mathbf{A}}] = [\Lambda_{\mathbf{A}}^* : \Lambda_{\mathbf{B}}^*]$, convolution with $S_{\mathbf{B}/\mathbf{A}}$ and $S_{\mathbf{A}/\mathbf{B}}^*$ has the effect of *averaging* the translates of a distribution under the elements (or ‘cosets’) of the residual lattices $\Lambda_{\mathbf{B}}/\Lambda_{\mathbf{A}}$ and $\Lambda_{\mathbf{A}}^*/\Lambda_{\mathbf{B}}^*$, respectively. This process will be called ‘coset averaging’. Eliminating $R_{\mathbf{A}}$ and $R_{\mathbf{B}}$ between (i) and (ii), and $R_{\mathbf{A}}^*$ and $R_{\mathbf{B}}^*$ between (i)^{*} and (ii)^{*}, we may write:

$$(i') \quad R_{\mathbf{A}} = \mathbf{D}^{\#} (S_{\mathbf{B}/\mathbf{A}} * R_{\mathbf{A}})$$

$$(ii') \quad R_{\mathbf{B}} = S_{\mathbf{B}/\mathbf{A}} * (\mathbf{D}^{\#} R_{\mathbf{B}})$$

$$(i')^* \quad R_{\mathbf{B}}^* = (\mathbf{D}^T)^{\#} (S_{\mathbf{A}/\mathbf{B}}^* * R_{\mathbf{B}}^*)$$

$$(ii')^* \quad R_{\mathbf{A}}^* = S_{\mathbf{A}/\mathbf{B}}^* * [(\mathbf{D}^T)^{\#} R_{\mathbf{A}}^*].$$

These identities show that *period subdivision* by convolution with $S_{\mathbf{B}/\mathbf{A}}$ (respectively $S_{\mathbf{A}/\mathbf{B}}^*$) on the one hand, and *period decimation* by ‘dilation’ by $\mathbf{D}^{\#}$ on the other hand, are mutually inverse operations on $R_{\mathbf{A}}$ and $R_{\mathbf{B}}$ (respectively $R_{\mathbf{A}}^*$ and $R_{\mathbf{B}}^*$).

1.3.2.7.2.4. *Relation between Fourier transforms*

Finally, let us consider the relations between the *Fourier transforms* of these lattice distributions. Recalling the basic relation of Section 1.3.2.6.5,

$$\begin{aligned} \mathcal{F}[R_{\mathbf{A}}] &= \frac{1}{|\det \mathbf{A}|} R_{\mathbf{A}}^* \\ &= \frac{1}{|\det \mathbf{D}\mathbf{B}|} T_{\mathbf{A}/\mathbf{B}}^* * R_{\mathbf{B}}^* && \text{by (ii)^*} \\ &= \left(\frac{1}{|\det \mathbf{D}|} T_{\mathbf{A}/\mathbf{B}}^* \right) * \left(\frac{1}{|\det \mathbf{B}|} R_{\mathbf{B}}^* \right) \end{aligned}$$

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

i.e.

$$(iv) \quad \mathcal{F}[R_{\mathbf{A}}] = S_{\mathbf{A}/\mathbf{B}}^* * \mathcal{F}[R_{\mathbf{B}}]$$

and similarly:

$$(v) \quad \mathcal{F}[R_{\mathbf{B}}^*] = S_{\mathbf{B}/\mathbf{A}} * \mathcal{F}[R_{\mathbf{A}}^*].$$

Thus $R_{\mathbf{A}}$ (respectively $R_{\mathbf{B}}^*$), a *decimated* version of $R_{\mathbf{B}}$ (respectively $R_{\mathbf{A}}^*$), is transformed by \mathcal{F} into a *subdivided* version of $\mathcal{F}[R_{\mathbf{B}}]$ (respectively $\mathcal{F}[R_{\mathbf{A}}^*]$).

The converse is also true:

$$\begin{aligned} \mathcal{F}[R_{\mathbf{B}}] &= \frac{1}{|\det \mathbf{B}|} R_{\mathbf{B}}^* \\ &= \frac{1}{|\det \mathbf{B}|} \frac{1}{|\det \mathbf{D}|} (\mathbf{D}^T)^{\#} R_{\mathbf{A}}^* && \text{by (i)*} \\ &= (\mathbf{D}^T)^{\#} \left(\frac{1}{|\det \mathbf{A}|} R_{\mathbf{A}}^* \right) \end{aligned}$$

i.e.

$$(iv') \quad \mathcal{F}[R_{\mathbf{B}}] = (\mathbf{D}^T)^{\#} \mathcal{F}[R_{\mathbf{A}}]$$

and similarly

$$(v') \quad \mathcal{F}[R_{\mathbf{A}}^*] = \mathbf{D}^{\#} \mathcal{F}[R_{\mathbf{B}}^*].$$

Thus $R_{\mathbf{B}}$ (respectively $R_{\mathbf{A}}^*$), a *subdivided* version of $R_{\mathbf{A}}$ (respectively $R_{\mathbf{B}}^*$) is transformed by \mathcal{F} into a *decimated* version of $\mathcal{F}[R_{\mathbf{A}}]$ (respectively $\mathcal{F}[R_{\mathbf{B}}^*]$). Therefore, *the Fourier transform exchanges subdivision and decimation of period lattices for lattice distributions.*

Further insight into this phenomenon is provided by applying $\bar{\mathcal{F}}$ to both sides of (iv) and (v) and invoking the convolution theorem:

$$(iv'') \quad R_{\mathbf{A}} = \bar{\mathcal{F}}[S_{\mathbf{A}/\mathbf{B}}^*] \times R_{\mathbf{B}}$$

$$(v'') \quad R_{\mathbf{B}}^* = \bar{\mathcal{F}}[S_{\mathbf{B}/\mathbf{A}}] \times R_{\mathbf{A}}^*.$$

These identities show that multiplication by the transform of the period-subdividing distribution $S_{\mathbf{A}/\mathbf{B}}^*$ (respectively $S_{\mathbf{B}/\mathbf{A}}$) has the effect of decimating $R_{\mathbf{B}}$ to $R_{\mathbf{A}}$ (respectively $R_{\mathbf{A}}^*$ to $R_{\mathbf{B}}^*$). They clearly imply that, if $\ell \in \Lambda_{\mathbf{B}/\Lambda_{\mathbf{A}}}$ and $\ell^* \in \Lambda_{\mathbf{A}}^*/\Lambda_{\mathbf{B}}^*$, then

$$\begin{aligned} \bar{\mathcal{F}}[S_{\mathbf{A}/\mathbf{B}}^*](\ell) &= 1 \text{ if } \ell = \mathbf{0} && \text{(i.e. if } \ell \text{ belongs} \\ & && \text{to the class of } \Lambda_{\mathbf{A}}), \\ &= 0 \text{ if } \ell \neq \mathbf{0}; \\ \bar{\mathcal{F}}[S_{\mathbf{B}/\mathbf{A}}](\ell^*) &= 1 \text{ if } \ell^* = \mathbf{0} && \text{(i.e. if } \ell^* \text{ belongs} \\ & && \text{to the class of } \Lambda_{\mathbf{B}}^*), \\ &= 0 \text{ if } \ell^* \neq \mathbf{0}. \end{aligned}$$

Therefore, the duality between subdivision and decimation may be viewed as another aspect of that between convolution and multiplication.

There is clearly a strong analogy between the sampling/periodization duality of Section 1.3.2.6.6 and the decimation/subdivision duality, which is viewed most naturally in terms of subgroup relationships: both sampling and decimation involve

restricting a function to a *discrete additive subgroup* of the domain over which it is initially given.

1.3.2.7.2.5. Sublattice relations in terms of periodic distributions

The usual presentation of this duality is not in terms of lattice distributions, but of periodic distributions obtained by convolving them with a motif.

Given $T^0 \in \mathcal{E}'(\mathbb{R}^n)$, let us form $R_{\mathbf{A}} * T^0$, then *decimate* its transform $(1/|\det \mathbf{A}|)R_{\mathbf{A}}^* \times \mathcal{F}[T^0]$ by keeping only its values at the points of the coarser lattice $\Lambda_{\mathbf{B}}^* = \mathbf{D}^T \Lambda_{\mathbf{A}}^*$; as a result, $R_{\mathbf{A}}^*$ is replaced by $(1/|\det \mathbf{D}|)R_{\mathbf{B}}^*$, and the reverse transform then yields

$$\frac{1}{|\det \mathbf{D}|} R_{\mathbf{B}} * T^0 = S_{\mathbf{B}/\mathbf{A}} * (R_{\mathbf{A}} * T^0) \quad \text{by (ii),}$$

which is the *coset-averaged* version of the original $R_{\mathbf{A}} * T^0$. The converse situation is analogous to that of Shannon's sampling theorem. Let a function $\varphi \in \mathcal{E}(\mathbb{R}^n)$ whose transform $\Phi = \mathcal{F}[\varphi]$ has compact support be sampled as $R_{\mathbf{B}} \times \varphi$ at the nodes of $\Lambda_{\mathbf{B}}$. Then

$$\mathcal{F}[R_{\mathbf{B}} \times \varphi] = \frac{1}{|\det \mathbf{B}|} (R_{\mathbf{B}}^* * \Phi)$$

is periodic with period lattice $\Lambda_{\mathbf{B}}^*$. If the sampling lattice $\Lambda_{\mathbf{B}}$ is decimated to $\Lambda_{\mathbf{A}} = \mathbf{D}\Lambda_{\mathbf{B}}$, the inverse transform becomes

$$\begin{aligned} \mathcal{F}[R_{\mathbf{A}} \times \varphi] &= \frac{1}{|\det \mathbf{D}|} (R_{\mathbf{A}}^* * \Phi) \\ &= S_{\mathbf{A}/\mathbf{B}}^* * (R_{\mathbf{B}}^* * \Phi) && \text{by (ii)*,} \end{aligned}$$

hence becomes periodized more finely by averaging over the cosets of $\Lambda_{\mathbf{A}}^*/\Lambda_{\mathbf{B}}^*$. With this finer periodization, the various copies of $\text{Supp } \Phi$ may start to overlap (a phenomenon called 'aliasing'), indicating that decimation has produced too coarse a sampling of φ .

1.3.2.7.3. Discretization of the Fourier transformation

Let $\varphi^0 \in \mathcal{E}(\mathbb{R}^n)$ be such that $\Phi^0 = \mathcal{F}[\varphi^0]$ has compact support (φ^0 is said to be *band-limited*). Then $\varphi = R_{\mathbf{A}} * \varphi^0$ is $\Lambda_{\mathbf{A}}$ -periodic, and $\Phi = \mathcal{F}[\varphi] = (1/|\det \mathbf{A}|)R_{\mathbf{A}}^* \times \Phi^0$ is such that only a finite number of points $\lambda_{\mathbf{A}}^*$ of $\Lambda_{\mathbf{A}}^*$ have a nonzero Fourier coefficient $\Phi^0(\lambda_{\mathbf{A}}^*)$ attached to them. We may therefore find a *decimation* $\Lambda_{\mathbf{B}}^* = \mathbf{D}^T \Lambda_{\mathbf{A}}^*$ of $\Lambda_{\mathbf{A}}^*$ such that the distinct translates of $\text{Supp } \Phi^0$ by vectors of $\Lambda_{\mathbf{B}}^*$ do not intersect.

The distribution Φ can be uniquely recovered from $R_{\mathbf{B}}^* * \Phi$ by the procedure of Section 1.3.2.7.1, and we may write:

$$\begin{aligned} R_{\mathbf{B}}^* * \Phi &= \frac{1}{|\det \mathbf{A}|} R_{\mathbf{B}}^* * (R_{\mathbf{A}}^* \times \Phi^0) \\ &= \frac{1}{|\det \mathbf{A}|} R_{\mathbf{A}}^* \times (R_{\mathbf{B}}^* * \Phi^0) \\ &= \frac{1}{|\det \mathbf{A}|} R_{\mathbf{B}}^* * [T_{\mathbf{A}/\mathbf{B}}^* \times (R_{\mathbf{B}}^* * \Phi^0)]; \end{aligned}$$

these rearrangements being legitimate because Φ^0 and $T_{\mathbf{A}/\mathbf{B}}^*$ have compact supports which are intersection-free under the action of $\Lambda_{\mathbf{B}}^*$. By virtue of its $\Lambda_{\mathbf{B}}^*$ -periodicity, this distribution is entirely characterized by its 'motif' $\tilde{\Phi}$ with respect to $\Lambda_{\mathbf{B}}^*$.

$$\tilde{\Phi} = \frac{1}{|\det \mathbf{A}|} T_{\mathbf{A}/\mathbf{B}}^* \times (R_{\mathbf{B}}^* * \Phi^0).$$

1. GENERAL RELATIONSHIPS AND TECHNIQUES

Similarly, φ may be uniquely recovered by Shannon interpolation from the distribution sampling its values at the nodes of $\Lambda_{\mathbf{B}} = \mathbf{D}^{-1}\Lambda_{\mathbf{A}}$ ($\Lambda_{\mathbf{B}}$ is a *subdivision* of $\Lambda_{\mathbf{B}}$). By virtue of its $\Lambda_{\mathbf{A}}$ -periodicity, this distribution is completely characterized by its motif:

$$\tilde{\varphi} = T_{\mathbf{B}/\mathbf{A}} \times \varphi = T_{\mathbf{B}/\mathbf{A}} \times (R_{\mathbf{A}}^* * \varphi^0).$$

Let $\ell \in \Lambda_{\mathbf{B}}/\Lambda_{\mathbf{A}}$ and $\ell^* \in \Lambda_{\mathbf{A}}^*/\Lambda_{\mathbf{B}}^*$, and define the two sets of coefficients

- (1) $\tilde{\varphi}(\ell) = \varphi(\ell + \lambda_{\mathbf{A}})$ for any $\lambda_{\mathbf{A}} \in \Lambda_{\mathbf{A}}$
(all choices of $\lambda_{\mathbf{A}}$ give the same $\tilde{\varphi}$),
- (2) $\tilde{\Phi}(\ell^*) = \Phi^0(\ell^* + \lambda_{\mathbf{B}}^*)$ for the unique $\lambda_{\mathbf{B}}^*$ (if it exists)
such that $\ell^* + \lambda_{\mathbf{B}}^* \in \text{Supp } \Phi^0$,
 $= 0$ if no such $\lambda_{\mathbf{B}}^*$ exists.

Define the two distributions

$$\omega = \sum_{\ell \in \Lambda_{\mathbf{B}}/\Lambda_{\mathbf{A}}} \tilde{\varphi}(\ell) \delta_{(\ell)}$$

and

$$\Omega = \sum_{\ell^* \in \Lambda_{\mathbf{A}}^*/\Lambda_{\mathbf{B}}^*} \tilde{\Phi}(\ell^*) \delta_{(\ell^*)}.$$

The relation between ω and Ω has two equivalent forms:

- (i) $R_{\mathbf{A}} * \omega = \mathcal{F}[R_{\mathbf{B}}^* * \Omega]$
- (ii) $\tilde{\mathcal{F}}[R_{\mathbf{A}} * \omega] = R_{\mathbf{B}}^* * \Omega.$

By (i), $R_{\mathbf{A}} * \omega = |\det \mathbf{B}| R_{\mathbf{B}} \times \mathcal{F}[\Omega]$. Both sides are weighted lattice distributions concentrated at the nodes of $\Lambda_{\mathbf{B}}$, and equating the weights at $\lambda_{\mathbf{B}} = \ell + \lambda_{\mathbf{A}}$ gives

$$\tilde{\varphi}(\ell) = \frac{1}{|\det \mathbf{D}|} \sum_{\ell^* \in \Lambda_{\mathbf{A}}^*/\Lambda_{\mathbf{B}}^*} \tilde{\Phi}(\ell^*) \exp[-2\pi i \ell^* \cdot (\ell + \lambda_{\mathbf{A}})].$$

Since $\ell^* \in \Lambda_{\mathbf{A}}^*$, $\ell^* \cdot \lambda_{\mathbf{A}}$ is an integer, hence

$$\tilde{\varphi}(\ell) = \frac{1}{|\det \mathbf{D}|} \sum_{\ell^* \in \Lambda_{\mathbf{A}}^*/\Lambda_{\mathbf{B}}^*} \tilde{\Phi}(\ell^*) \exp(-2\pi i \ell^* \cdot \ell).$$

By (ii), we have

$$\frac{1}{|\det \mathbf{A}|} R_{\mathbf{B}}^* * [T_{\mathbf{A}/\mathbf{B}}^* \times (R_{\mathbf{B}}^* * \Phi^0)] = \frac{1}{|\det \mathbf{A}|} \tilde{\mathcal{F}}[R_{\mathbf{A}} * \omega].$$

Both sides are weighted lattice distributions concentrated at the nodes of $\Lambda_{\mathbf{B}}^*$, and equating the weights at $\lambda_{\mathbf{A}}^* = \ell^* + \lambda_{\mathbf{B}}^*$ gives

$$\tilde{\Phi}(\ell^*) = \sum_{\ell \in \Lambda_{\mathbf{B}}/\Lambda_{\mathbf{A}}} \tilde{\varphi}(\ell) \exp[+2\pi i \ell \cdot (\ell^* + \lambda_{\mathbf{B}}^*)].$$

Since $\ell \in \Lambda_{\mathbf{B}}$, $\ell \cdot \lambda_{\mathbf{B}}^*$ is an integer, hence

$$\tilde{\Phi}(\ell^*) = \sum_{\ell \in \Lambda_{\mathbf{B}}/\Lambda_{\mathbf{A}}} \tilde{\varphi}(\ell) \exp(+2\pi i \ell \cdot \ell^*).$$

Now the decimation/subdivision relations between $\Lambda_{\mathbf{A}}$ and $\Lambda_{\mathbf{B}}$ may be written:

$$\mathbf{A} = \mathbf{D}\mathbf{B} = \mathbf{B}\mathbf{N},$$

so that

$$\begin{aligned} \ell &= \mathbf{B}k & \text{for } k \in \mathbb{Z}^n \\ \ell^* &= (\mathbf{A}^{-1})^T k^* & \text{for } k^* \in \mathbb{Z}^n \end{aligned}$$

with $(\mathbf{A}^{-1})^T = (\mathbf{B}^{-1})^T (\mathbf{N}^{-1})^T$, hence finally

$$\ell^* \cdot \ell = \ell \cdot \ell^* = k^* \cdot (\mathbf{N}^{-1}k).$$

Denoting $\tilde{\varphi}(\mathbf{B}k)$ by $\psi(k)$ and $\tilde{\Phi}[(\mathbf{A}^{-1})^T k^*]$ by $\Psi(k^*)$, the relation between ω and Ω may be written in the equivalent form

- (i) $\psi(k) = \frac{1}{|\det \mathbf{N}|} \sum_{k^* \in \mathbb{Z}^n / \mathbf{N}^T \mathbb{Z}^n} \Psi(k^*) \exp[-2\pi i k^* \cdot (\mathbf{N}^{-1}k)]$
- (ii) $\Psi(k^*) = \sum_{k \in \mathbb{Z}^n / \mathbf{N}\mathbb{Z}^n} \psi(k) \exp[+2\pi i k^* \cdot (\mathbf{N}^{-1}k)],$

where the summations are now over *finite* residual lattices in standard form.

Equations (i) and (ii) describe two mutually inverse linear transformations $\mathcal{F}(\mathbf{N})$ and $\tilde{\mathcal{F}}(\mathbf{N})$ between two vector spaces $W_{\mathbf{N}}$ and $W_{\mathbf{N}}^*$ of dimension $|\det \mathbf{N}|$. $\mathcal{F}(\mathbf{N})$ [respectively $\tilde{\mathcal{F}}(\mathbf{N})$] is the *discrete* Fourier (respectively inverse Fourier) transform associated to matrix \mathbf{N} .

The vector spaces $W_{\mathbf{N}}$ and $W_{\mathbf{N}}^*$ may be viewed from two different standpoints:

(1) as vector spaces of *weighted residual-lattice distributions*, of the form $\alpha(\mathbf{x})T_{\mathbf{B}/\mathbf{A}}$ and $\beta(\mathbf{x})T_{\mathbf{A}/\mathbf{B}}^*$; the canonical basis of $W_{\mathbf{N}}$ (respectively $W_{\mathbf{N}}^*$) then consists of the $\delta_{(\ell)}$ for $\ell \in \mathbb{Z}^n / \mathbf{N}\mathbb{Z}^n$ [respectively $\delta_{(\ell^*)}$ for $\ell^* \in \mathbb{Z}^n / \mathbf{N}^T \mathbb{Z}^n$];

(2) as vector spaces of *weight vectors* for the $|\det \mathbf{N}|$ δ -functions involved in the expression for $T_{\mathbf{B}/\mathbf{A}}$ (respectively $T_{\mathbf{A}/\mathbf{B}}^*$); the canonical basis of $W_{\mathbf{N}}$ (respectively $W_{\mathbf{N}}^*$) consists of weight vectors \mathbf{u}_k (respectively \mathbf{v}_{k^*}) giving weight 1 to element k (respectively k^*) and 0 to the others.

These two spaces are said to be ‘isomorphic’ (a relation denoted \cong), the isomorphism being given by the one-to-one correspondence:

$$\begin{aligned} \omega = \sum_k \psi(k) \delta_{(\ell)} &\leftrightarrow \psi = \sum_k \psi(k) \mathbf{u}_k \\ \Omega = \sum_{k^*} \Psi(k^*) \delta_{(\ell^*)} &\leftrightarrow \Psi = \sum_{k^*} \Psi(k^*) \mathbf{v}_{k^*}. \end{aligned}$$

The second viewpoint will be adopted, as it involves only linear algebra. However, it is most helpful to keep the first one in mind and to think of the data or results of a discrete Fourier transform as representing (through their sets of unique weights) two periodic lattice distributions related by the full, distribution-theoretic Fourier transform.

We therefore view $W_{\mathbf{N}}$ (respectively $W_{\mathbf{N}}^*$) as the vector space of complex-valued functions over the finite residual lattice $\Lambda_{\mathbf{B}}/\Lambda_{\mathbf{A}}$ (respectively $\Lambda_{\mathbf{A}}^*/\Lambda_{\mathbf{B}}^*$) and write:

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

$$W_{\mathbf{N}} \cong L(\Lambda_{\mathbf{B}}/\Lambda_{\mathbf{A}}) \cong L(\mathbb{Z}^n/\mathbf{N}\mathbb{Z}^n)$$

$$W_{\mathbf{N}}^* \cong L(\Lambda_{\mathbf{A}}^*/\Lambda_{\mathbf{B}}^*) \cong L(\mathbb{Z}^n/\mathbf{N}^T\mathbb{Z}^n)$$

since a vector such as ψ is in fact the function $k \mapsto \psi(k)$.

The two spaces $W_{\mathbf{N}}$ and $W_{\mathbf{N}}^*$ may be equipped with the following Hermitian inner products:

$$(\varphi, \psi)_W = \sum_k \overline{\varphi(k)} \psi(k)$$

$$(\Phi, \Psi)_{W^*} = \sum_k \overline{\Phi(k^*)} \Psi(k^*),$$

which makes each of them into a *Hilbert space*. The canonical bases $\{\mathbf{u}_k | k \in \mathbb{Z}^n/\mathbf{N}\mathbb{Z}^n\}$ and $\{\mathbf{v}_{k^*} | k^* \in \mathbb{Z}^n/\mathbf{N}^T\mathbb{Z}^n\}$ and $W_{\mathbf{N}}$ and $W_{\mathbf{N}}^*$ are *orthonormal* for their respective product.

1.3.2.7.4. Matrix representation of the discrete Fourier transform (DFT)

By virtue of definitions (i) and (ii),

$$\tilde{\mathcal{F}}(\mathbf{N})\mathbf{v}_{k^*} = \frac{1}{|\det \mathbf{N}|} \sum_k \exp[-2\pi i k^* \cdot (\mathbf{N}^{-1}k)] \mathbf{u}_k$$

$$\tilde{\mathcal{F}}(\mathbf{N})\mathbf{u}_k = \sum_{k^*} \exp[+2\pi i k^* \cdot (\mathbf{N}^{-1}k)] \mathbf{v}_{k^*}$$

so that $\mathcal{F}(\mathbf{N})$ and $\tilde{\mathcal{F}}(\mathbf{N})$ may be represented, in the canonical bases of $W_{\mathbf{N}}$ and $W_{\mathbf{N}}^*$, by the following matrices:

$$[\mathcal{F}(\mathbf{N})]_{kk^*} = \frac{1}{|\det \mathbf{N}|} \exp[-2\pi i k^* \cdot (\mathbf{N}^{-1}k)]$$

$$[\tilde{\mathcal{F}}(\mathbf{N})]_{k^*k} = \exp[+2\pi i k^* \cdot (\mathbf{N}^{-1}k)].$$

When \mathbf{N} is symmetric, $\mathbb{Z}^n/\mathbf{N}\mathbb{Z}^n$ and $\mathbb{Z}^n/\mathbf{N}^T\mathbb{Z}^n$ may be identified in a natural manner, and the above matrices are symmetric.

When \mathbf{N} is diagonal, say $\mathbf{N} = \text{diag}(v_1, v_2, \dots, v_n)$, then the tensor product structure of the full multidimensional Fourier transform (Section 1.3.2.4.2.4)

$$\tilde{\mathcal{F}}_{\mathbf{x}} = \tilde{\mathcal{F}}_{x_1} \otimes \tilde{\mathcal{F}}_{x_2} \otimes \dots \otimes \tilde{\mathcal{F}}_{x_n}$$

gives rise to a tensor product structure for the DFT matrices. The tensor product of matrices is defined as follows:

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & \dots & a_{1n}\mathbf{B} \\ \vdots & & \vdots \\ a_{n1}\mathbf{B} & \dots & a_{nn}\mathbf{B} \end{pmatrix}.$$

Let the index vectors k and k^* be ordered in the same way as the elements in a Fortran array, e.g. for k with k_1 increasing fastest, k_2 next fastest, \dots , k_n slowest; then

$$\tilde{\mathcal{F}}(\mathbf{N}) = \tilde{\mathcal{F}}(v_1) \otimes \tilde{\mathcal{F}}(v_2) \otimes \dots \otimes \tilde{\mathcal{F}}(v_n),$$

where

$$[\tilde{\mathcal{F}}(v_j)]_{k_j, k_j^*} = \frac{1}{v_j} \exp\left(-2\pi i \frac{k_j^* k_j}{v_j}\right),$$

and

$$\tilde{\mathcal{F}}(\mathbf{N}) = \tilde{\mathcal{F}}(v_1) \otimes \tilde{\mathcal{F}}(v_2) \otimes \dots \otimes \tilde{\mathcal{F}}(v_n),$$

where

$$[\tilde{\mathcal{F}}(v_j)]_{k_j^*, k_j} = \exp\left(+2\pi i \frac{k_j^* k_j}{v_j}\right).$$

1.3.2.7.5. Properties of the discrete Fourier transform

The DFT inherits most of the properties of the Fourier transforms, but with certain numerical factors ('Jacobians') due to the transition from continuous to discrete measure.

(1) *Linearity* is obvious.

(2) *Shift property*. If $(\tau_a \psi)(k) = \psi(k - a)$ and $(\tau_{a^*} \Psi)(k^*) = \Psi(k^* - a^*)$, where subtraction takes place by modular vector arithmetic in $\mathbb{Z}^n/\mathbf{N}\mathbb{Z}^n$ and $\mathbb{Z}^n/\mathbf{N}^T\mathbb{Z}^n$, respectively, then the following identities hold:

$$\tilde{\mathcal{F}}(\mathbf{N})[\tau_a \psi](k^*) = \exp[+2\pi i k^* \cdot (\mathbf{N}^{-1}a)] \tilde{\mathcal{F}}(\mathbf{N})[\psi](k^*)$$

$$\mathcal{F}(\mathbf{N})[\tau_{a^*} \Psi](k) = \exp[-2\pi i k^* \cdot (\mathbf{N}^{-1}a^*)] \mathcal{F}(\mathbf{N})[\Psi](k).$$

(3) *Differentiation identities*. Let vectors $\boldsymbol{\psi}$ and $\boldsymbol{\Psi}$ be constructed from $\varphi^0 \in \mathcal{E}(\mathbb{R}^n)$ as in Section 1.3.2.7.3, hence be related by the DFT. If $D^{\mathbf{p}}\boldsymbol{\psi}$ designates the vector of sample values of $D_{\xi}^{\mathbf{p}}\varphi^0$ at the points of $\Lambda_{\mathbf{B}}/\Lambda_{\mathbf{A}}$, and $D^{\mathbf{p}}\boldsymbol{\Psi}$ the vector of values of $D_{\xi}^{\mathbf{p}}\Phi^0$ at points of $\Lambda_{\mathbf{A}}^*/\Lambda_{\mathbf{B}}^*$, then for all multi-indices $\mathbf{p} = (p_1, p_2, \dots, p_n)$

$$(D^{\mathbf{p}}\boldsymbol{\psi})(k) = \tilde{\mathcal{F}}(\mathbf{N})[(+2\pi i k^*)^{\mathbf{p}}\boldsymbol{\Psi}](k)$$

$$(D^{\mathbf{p}}\boldsymbol{\Psi})(k^*) = \mathcal{F}(\mathbf{N})[(-2\pi i k)^{\mathbf{p}}\boldsymbol{\psi}](k^*)$$

or equivalently

$$\mathcal{F}(\mathbf{N})[D^{\mathbf{p}}\boldsymbol{\psi}](k^*) = (+2\pi i k^*)^{\mathbf{p}}\boldsymbol{\Psi}(k^*)$$

$$\tilde{\mathcal{F}}(\mathbf{N})[D^{\mathbf{p}}\boldsymbol{\Psi}](k) = (-2\pi i k)^{\mathbf{p}}\boldsymbol{\psi}(k).$$

(4) *Convolution property*. Let $\boldsymbol{\varphi} \in W_{\mathbf{N}}$ and $\boldsymbol{\Phi} \in W_{\mathbf{N}}^*$ (respectively $\boldsymbol{\psi}$ and $\boldsymbol{\Psi}$) be related by the DFT, and define

$$(\boldsymbol{\varphi} * \boldsymbol{\psi})(k) = \sum_{k' \in \mathbb{Z}^n/\mathbf{N}\mathbb{Z}^n} \boldsymbol{\varphi}(k') \boldsymbol{\psi}(k - k')$$

$$(\boldsymbol{\Phi} * \boldsymbol{\Psi})(k^*) = \sum_{k' \in \mathbb{Z}^n/\mathbf{N}^T\mathbb{Z}^n} \boldsymbol{\Phi}(k') \boldsymbol{\Psi}(k^* - k').$$

Then

$$\tilde{\mathcal{F}}(\mathbf{N})[\boldsymbol{\Phi} * \boldsymbol{\Psi}](k) = |\det \mathbf{N}| \boldsymbol{\varphi}(k) \boldsymbol{\psi}(k)$$

$$\mathcal{F}(\mathbf{N})[\boldsymbol{\varphi} * \boldsymbol{\psi}](k^*) = \boldsymbol{\Phi}(k^*) \boldsymbol{\Psi}(k^*)$$

and

$$\tilde{\mathcal{F}}(\mathbf{N})[\boldsymbol{\varphi} \times \boldsymbol{\psi}](k^*) = \frac{1}{|\det \mathbf{N}|} (\boldsymbol{\Phi} * \boldsymbol{\Psi})(k^*)$$

$$\mathcal{F}(\mathbf{N})[\boldsymbol{\Phi} \times \boldsymbol{\Psi}](k) = (\boldsymbol{\varphi} * \boldsymbol{\psi})(k).$$

Since addition on $\mathbb{Z}^n/\mathbf{N}\mathbb{Z}^n$ and $\mathbb{Z}^n/\mathbf{N}^T\mathbb{Z}^n$ is modular, this type of convolution is called *cyclic convolution*.

1. GENERAL RELATIONSHIPS AND TECHNIQUES

(5) *Parseval/Plancherel property.* If ϕ, ψ, Φ, Ψ are as above, then

$$\begin{aligned} (\mathcal{F}(\mathbf{N})[\Phi], \mathcal{F}(\mathbf{N})[\Psi])_W &= \frac{1}{|\det \mathbf{N}|} (\Phi, \Psi)_{W^*} \\ (\tilde{\mathcal{F}}(\mathbf{N})[\phi], \tilde{\mathcal{F}}(\mathbf{N})[\psi])_W &= \frac{1}{|\det \mathbf{N}|} (\phi, \psi)_W. \end{aligned}$$

(6) *Period 4.* When \mathbf{N} is symmetric, so that the ranges of indices k and k^* can be identified, it makes sense to speak of powers of $\mathcal{F}(\mathbf{N})$ and $\tilde{\mathcal{F}}(\mathbf{N})$. Then the ‘standardized’ matrices $(1/|\det \mathbf{N}|^{1/2})\mathcal{F}(\mathbf{N})$ and $(1/|\det \mathbf{N}|^{1/2})\tilde{\mathcal{F}}(\mathbf{N})$ are *unitary* matrices whose fourth power is the identity matrix (Section 1.3.2.4.3.4); their eigenvalues are therefore ± 1 and $\pm i$.

1.3.3. Numerical computation of the discrete Fourier transform

1.3.3.1. Introduction

The Fourier transformation’s most remarkable property is undoubtedly that of turning convolution into multiplication. As distribution theory has shown, other valuable properties – such as the shift property, the conversion of differentiation into multiplication by monomials, and the duality between periodicity and sampling – are special instances of the convolution theorem.

This property is exploited in many areas of applied mathematics and engineering (Campbell & Foster, 1948; Sneddon, 1951; Champeney, 1973; Bracewell, 1986). For example, the passing of a signal through a linear filter, which results in its being convolved with the response of the filter to a δ -function ‘impulse’, may be modelled as a multiplication of the signal’s transform by the transform of the impulse response (also called transfer function). Similarly, the solution of systems of partial differential equations may be turned by Fourier transformation into a division problem for distributions. In both cases, the formulations obtained after Fourier transformation are considerably simpler than the initial ones, and lend themselves to constructive solution techniques.

Whenever the functions to which the Fourier transform is applied are band-limited, or can be well approximated by band-limited functions, the discrete Fourier transform (DFT) provides a means of constructing explicit numerical solutions to the problems at hand. A great variety of investigations in physics, engineering and applied mathematics thus lead to DFT calculations, to such a degree that, at the time of writing, about 50% of all supercomputer CPU time is alleged to be spent calculating DFTs.

The straightforward use of the defining formulae for the DFT leads to calculations of size N^2 for N sample points, which become unfeasible for any but the smallest problems. Much ingenuity has therefore been exerted on the design and implementation of faster algorithms for calculating the DFT (McClellan & Rader, 1979; Nussbaumer, 1981; Blahut, 1985; Brigham, 1988). The most famous is that of Cooley & Tukey (1965) which heralded the age of digital signal processing. However, it had been preceded by the prime factor algorithm of Good (1958, 1960), which has lately been the basis of many new developments. Recent historical research (Goldstine, 1977, pp. 249–253; Heideman *et al.*, 1984) has shown that Gauss essentially knew the Cooley–Tukey algorithm as early as 1805 (before Fourier’s 1807 work on harmonic analysis!); while it has long been clear that Dirichlet knew of the basis of the prime factor algorithm and used it extensively in his theory of multiplicative characters [see *e.g.* Chapter I of Ayoub (1963), and Chapters 6 and 8 of Apostol (1976)]. Thus the computation of the DFT, far from being a purely technical and rather narrow piece of specialized numerical analysis, turns out to have very rich connections with such central areas of pure mathematics as number theory (algebraic and analytic), the representation

theory of certain Lie groups and coding theory – to list only a few. The interested reader may consult Auslander & Tolimieri (1979); Auslander, Feig & Winograd (1982, 1984); Auslander & Tolimieri (1985); Tolimieri (1985).

One-dimensional algorithms are examined first. The Sande mixed-radix version of the Cooley–Tukey algorithm only calls upon the additive structure of congruence classes of integers. The prime factor algorithm of Good begins to exploit some of their multiplicative structure, and the use of relatively prime factors leads to a stronger factorization than that of Sande. Fuller use of the multiplicative structure, *via* the group of units, leads to the Rader algorithm; and the factorization of short convolutions then yields the Winograd algorithms.

Multidimensional algorithms are at first built as tensor products of one-dimensional elements. The problem of factoring the DFT in several dimensions simultaneously is then examined. The section ends with a survey of attempts at formalizing the interplay between algorithm structure and computer architecture for the purpose of automating the design of optimal DFT code.

It was originally intended to incorporate into this section a survey of all the basic notions and results of abstract algebra which are called upon in the course of these developments, but time limitations have made this impossible. This material, however, is adequately covered by the first chapter of Tolimieri *et al.* (1989) in a form tailored for the same purposes. Similarly, the inclusion of numerous detailed examples of the algorithms described here has had to be postponed to a later edition, but an abundant supply of such examples may be found in the signal processing literature, for instance in the books by McClellan & Rader (1979), Blahut (1985), and Tolimieri *et al.* (1989).

1.3.3.2. One-dimensional algorithms

Throughout this section we will denote by $e(t)$ the expression $\exp(2\pi it)$, $t \in \mathbb{R}$. The mapping $t \mapsto e(t)$ has the following properties:

$$\begin{aligned} e(t_1 + t_2) &= e(t_1)e(t_2) \\ e(-t) &= \overline{e(t)} = [e(t)]^{-1} \\ e(t) &= 1 \Leftrightarrow t \in \mathbb{Z}. \end{aligned}$$

Thus e defines an isomorphism between the additive group \mathbb{R}/\mathbb{Z} (the reals modulo the integers) and the multiplicative group of complex numbers of modulus 1. It follows that the mapping $\ell \mapsto e(\ell/N)$, where $\ell \in \mathbb{Z}$ and N is a positive integer, defines an isomorphism between the one-dimensional residual lattice $\mathbb{Z}/N\mathbb{Z}$ and the multiplicative group of N th roots of unity.

The DFT on N points then relates vectors \mathbf{X} and \mathbf{X}^* in W and W^* through the linear transformations:

$$\begin{aligned} F(N) : \quad X(k) &= \frac{1}{N} \sum_{k^* \in \mathbb{Z}/N\mathbb{Z}} X^*(k^*)e(-k^*k/N) \\ \bar{F}(N) : \quad X^*(k^*) &= \sum_{k \in \mathbb{Z}/N\mathbb{Z}} X(k)e(k^*k/N). \end{aligned}$$

1.3.3.2.1. The Cooley–Tukey algorithm

The presentation of Gentleman & Sande (1966) will be followed first [see also Cochran *et al.* (1967)]. It will then be reinterpreted in geometric terms which will prepare the way for the treatment of multidimensional transforms in Section 1.3.3.3.

Suppose that the number of sample points N is composite, say $N = N_1 N_2$. We may write k to the base N_1 and k^* to the base N_2 as follows: