

1. GENERAL RELATIONSHIPS AND TECHNIQUES

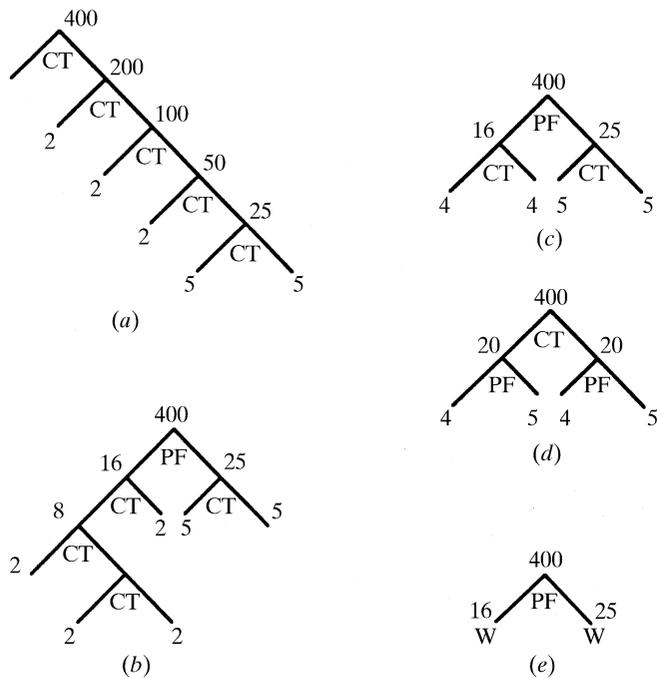


Fig. 1.3.3.1. A few global algorithms for computing a 400-point DFT. CT: Cooley–Tukey factorization. PF: prime factor (or Good) factorization. W: Winograd algorithm.

The simplest DFT may then be carried out into a global algorithm in many different ways. The diagrams in Fig. 1.3.3.1 illustrate a few of the options available to compute a 400-point DFT. They may differ greatly in their arithmetic operation counts.

1.3.3.3.2. Computer architecture considerations

To obtain a truly useful measure of the computational complexity of a DFT algorithm, its arithmetic operation count must be tempered by computer architecture considerations. Three main types of trade-offs must be borne in mind:

- (i) reductions in floating-point (f.p.) arithmetic count are obtained by reindexing, hence at the cost of an increase in integer arithmetic on addresses, although some shortcuts may be found (Uhrich, 1969; Burrus & Eschenbacher, 1981);
- (ii) reduction in the f.p. multiplication count usually leads to a large increase in the f.p. addition count (Morris, 1978);
- (iii) nesting can increase execution speed, but causes a loss of modularity and hence complicates program development (Silverman, 1977; Kolba & Parks, 1977).

Many of the mathematical developments above took place in the context of single-processor serial computers, where f.p. addition is substantially cheaper than f.p. multiplication but where integer address arithmetic has to compete with f.p. arithmetic for processor cycles. As a result, the alternatives to the Cooley–Tukey algorithm hardly ever led to particularly favourable trade-offs, thus creating the impression that there was little to gain by switching to more exotic algorithms.

The advent of new machine architectures with vector and/or parallel processing features has greatly altered this picture (Pease, 1968; Korn & Lambiotte, 1979; Fornberg, 1981; Swartzrauber, 1984):

- (i) *pipelining* equalizes the cost of f.p. addition and f.p. multiplication, and the ideal ‘blend’ of the two types of operations depends solely on the number of adder and multiplier units available in each machine;
- (ii) integer address arithmetic is delegated to specialized arithmetic and logical units (ALUs) operating concurrently with

the f.p. units, so that complex reindexing schemes may be used without loss of overall efficiency.

Another major consideration is that of data flow [see *e.g.* Nawab & McClellan (1979)]. Serial machines only have few registers and few paths connecting them, and allow little or no overlap between computation and data movement. New architectures, on the other hand, comprise banks of vector registers (or ‘cache memory’) besides the usual internal registers, and dedicated ALUs can service data transfers between several of them simultaneously and concurrently with computation.

In this new context, the devices described in Sections 1.3.3.2 and 1.3.3.3 for altering the balance between the various types of arithmetic operations, and reshaping the data flow during the computation, are invaluable. The field of machine-dependent DFT algorithm design is thriving on them [see *e.g.* Temperton (1983*a,b,c*, 1985); Agarwal & Cooley (1986, 1987)].

1.3.3.3.3. The Johnson–Burrus family of algorithms

In order to explore systematically all possible algorithms for carrying out a given DFT computation, and to pick the one best suited to a given machine, attempts have been made to develop:

- (i) a high-level notation of describing all the ingredients of a DFT computation, including data permutation and data flow;
- (ii) a formal calculus capable of operating on these descriptions so as to represent all possible reorganizations of the computation;
- (iii) an automatic procedure for evaluating the performance of a given algorithm on a specific architecture.

Task (i) can be accomplished by systematic use of a tensor product notation to represent the various stages into which the DFT can be factored (reindexing, small transforms on subsets of indices, twiddle factors, digit-reversal permutations).

Task (ii) may for instance use the Winograd **CBA** normal form for each small transform, then apply the rules governing the rearrangement of tensor product \otimes and ordinary product \times operations on matrices. The matching of these rearrangements to the architecture of a vector and/or parallel computer can be formalized algebraically [see *e.g.* Chapter 2 of Tolimieri *et al.* (1989)].

Task (iii) is a complex search which requires techniques such as dynamic programming (Bellman, 1958).

Johnson & Burrus (1983) have proposed and tested such a scheme to identify the optimal trade-offs between prime factor nesting and Winograd nesting of small Winograd transforms. In step (ii), they further decomposed the pre-addition matrix **A** and post-addition matrix **C** into several factors, so that the number of design options available becomes very large: the *N*-point DFT when *N* has four factors can be calculated in over 10^{12} distinct ways.

This large family of nested algorithms contains the prime factor algorithm and the Winograd algorithms as particular cases, but usually achieves greater efficiency than either by reducing the f.p. multiplication count while keeping the number of f.p. additions small.

There is little doubt that this systematic approach will be extended so as to incorporate all available methods of restructuring the DFT.

1.3.4. Crystallographic applications of Fourier transforms

1.3.4.1. Introduction

The central role of the Fourier transformation in X-ray crystallography is a consequence of the kinematic approximation used in the description of the scattering of X-rays by a distribution of electrons (Bragg, 1915; Duane, 1925; Havighurst, 1925*a,b*; Zachariasen, 1945; James, 1948*a*, Chapters 1 and 2; Lipson & Cochran, 1953, Chapter 1; Bragg, 1975).

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

Let $\rho(\mathbf{X})$ be the density of electrons in a sample of matter contained in a finite region V which is being illuminated by a parallel monochromatic X-ray beam with wavevector \mathbf{K}_0 . Then the far-field amplitude scattered in a direction corresponding to wavevector $\mathbf{K} = \mathbf{K}_0 + \mathbf{H}$ is proportional to

$$\begin{aligned} F(\mathbf{H}) &= \int_V \rho(\mathbf{X}) \exp(2\pi i \mathbf{H} \cdot \mathbf{X}) \, d^3 \mathbf{X} \\ &= \tilde{\mathcal{F}}[\rho](\mathbf{H}) \\ &= \langle \rho_{\mathbf{x}}, \exp(2\pi i \mathbf{H} \cdot \mathbf{X}) \rangle. \end{aligned}$$

In certain model calculations, the ‘sample’ may contain not only volume charges, but also point, line and surface charges. These singularities may be accommodated by letting ρ be a distribution, and writing

$$F(\mathbf{H}) = \tilde{\mathcal{F}}[\rho](\mathbf{H}) = \langle \rho_{\mathbf{x}}, \exp(2\pi i \mathbf{H} \cdot \mathbf{X}) \rangle.$$

F is still a well behaved function (analytic, by Section 1.3.2.4.2.10) because ρ has been assumed to have compact support.

If the sample is assumed to be an infinite crystal, so that ρ is now a *periodic* distribution, the customary limiting process by which it is shown that F becomes a discrete series of peaks at reciprocal-lattice points (see e.g. von Laue, 1936; Ewald, 1940; James, 1948a p. 9; Lipson & Taylor, 1958, pp. 14–27; Ewald, 1962, pp. 82–101; Warren, 1969, pp. 27–30) is already subsumed under the treatment of Section 1.3.2.6.

1.3.4.2. Crystallographic Fourier transform theory

1.3.4.2.1. Crystal periodicity

1.3.4.2.1.1. Period lattice, reciprocal lattice and structure factors

Let ρ be the distribution of electrons in a crystal. Then, by definition of a crystal, ρ is Λ -periodic for some period lattice Λ (Section 1.3.2.6.5) so that there exists a motif distribution ρ^0 with compact support such that

$$\rho = R * \rho^0,$$

where $R = \sum_{\mathbf{x} \in \Lambda} \delta(\mathbf{x})$. The lattice Λ is usually taken to be the finest for which the above representation holds.

Let Λ have a basis $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ over the integers, these basis vectors being expressed in terms of a standard orthonormal basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ as

$$\mathbf{a}_k = \sum_{j=1}^3 a_{jk} \mathbf{e}_j.$$

Then the matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

is the period matrix of Λ (Section 1.3.2.6.5) with respect to the unit lattice with basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, and the volume V of the unit cell is given by $V = |\det \mathbf{A}|$.

By Fourier transformation

$$\tilde{\mathcal{F}}[\rho] = R^* \times \tilde{\mathcal{F}}[\rho^0],$$

where $R^* = \sum_{\mathbf{H} \in \Lambda^*} \delta(\mathbf{H})$ is the lattice distribution associated to the reciprocal lattice Λ^* . The basis vectors $(\mathbf{a}_1^*, \mathbf{a}_2^*, \mathbf{a}_3^*)$ have coordinates in $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ given by the columns of $(\mathbf{A}^{-1})^T$, whose expression in terms of the cofactors of \mathbf{A} (see Section 1.3.2.6.5) gives the familiar formulae involving the cross product of vectors for $n = 3$. The \mathbf{H} -distribution F of scattered amplitudes may be written

$$F = \tilde{\mathcal{F}}[\rho]_{\mathbf{H}} = \sum_{\mathbf{H} \in \Lambda^*} \tilde{\mathcal{F}}[\rho^0](\mathbf{H}) \delta(\mathbf{H}) = \sum_{\mathbf{H} \in \Lambda^*} F_{\mathbf{H}} \delta(\mathbf{H})$$

and is thus a weighted reciprocal-lattice distribution, the weight $F_{\mathbf{H}}$ attached to each node $\mathbf{H} \in \Lambda^*$ being the value at \mathbf{H} of the transform $\tilde{\mathcal{F}}[\rho^0]$ of the motif ρ^0 . Taken in conjunction with the assumption that the scattering is elastic, i.e. that \mathbf{H} only changes the direction but not the magnitude of the incident wavevector \mathbf{K}_0 , this result yields the usual forms (Laue or Bragg) of the diffraction conditions: $\mathbf{H} \in \Lambda^*$, and simultaneously \mathbf{H} lies on the Ewald sphere.

By the reciprocity theorem, ρ^0 can be recovered if F is known for all $\mathbf{H} \in \Lambda^*$ as follows [Section 1.3.2.6.5, e.g. (iv)]:

$$\rho_{\mathbf{x}} = \frac{1}{V} \sum_{\mathbf{H} \in \Lambda^*} F_{\mathbf{H}} \exp(-2\pi i \mathbf{H} \cdot \mathbf{X}).$$

These relations may be rewritten in terms of standard, or ‘fractional crystallographic’, coordinates by putting

$$\mathbf{X} = \mathbf{A} \mathbf{x}, \quad \mathbf{H} = (\mathbf{A}^{-1})^T \mathbf{h},$$

so that a unit cell of the crystal corresponds to $\mathbf{x} \in \mathbb{R}^3 / \mathbb{Z}^3$, and that $\mathbf{h} \in \mathbb{Z}^3$. Defining ρ and ρ^0 by

$$\rho = \frac{1}{V} A^{\#} \rho, \quad \rho^0 = \frac{1}{V} A^{\#} \rho^0$$

so that

$$\rho(\mathbf{X}) \, d^3 \mathbf{X} = \rho(\mathbf{x}) \, d^3 \mathbf{x}, \quad \rho^0(\mathbf{X}) \, d^3 \mathbf{X} = \rho^0(\mathbf{x}) \, d^3 \mathbf{x},$$

we have

$$\begin{aligned} \tilde{\mathcal{F}}[\rho]_{\mathbf{h}} &= \sum_{\mathbf{h} \in \mathbb{Z}^3} F(\mathbf{h}) \delta(\mathbf{h}), \\ F(\mathbf{h}) &= \langle \rho_{\mathbf{x}}^0, \exp(2\pi i \mathbf{h} \cdot \mathbf{x}) \rangle \\ &= \int_{\mathbb{R}^3 / \mathbb{Z}^3} \rho^0(\mathbf{x}) \exp(2\pi i \mathbf{h} \cdot \mathbf{x}) \, d^3 \mathbf{x} \quad \text{if } \rho^0 \in L_{\text{loc}}^1(\mathbb{R}^3 / \mathbb{Z}^3), \\ \rho_{\mathbf{x}} &= \sum_{\mathbf{h} \in \mathbb{Z}^3} F(\mathbf{h}) \exp(-2\pi i \mathbf{h} \cdot \mathbf{x}). \end{aligned}$$

These formulae are valid for an arbitrary motif distribution ρ^0 , provided the convergence of the Fourier series for ρ is considered from the viewpoint of distribution theory (Section 1.3.2.6.10.3).

The experienced crystallographer may notice the absence of the familiar factor $1/V$ from the expression for ρ just given. This is because we use the (mathematically) natural unit for ρ , the *electron per unit cell*, which matches the dimensionless nature of the crystallographic coordinates \mathbf{x} and of the associated volume element $d^3 \mathbf{x}$. The traditional factor $1/V$ was the result of the somewhat inconsistent use of \mathbf{x} as an argument but of $d^3 \mathbf{X}$ as a volume element to obtain ρ in electrons per unit volume (e.g. \AA^{-3}). A fortunate consequence of the present convention is that nuisance factors of V or $1/V$, which used to abound in convolution or scalar product formulae, are now absent.

It should be noted at this point that the crystallographic terminology regarding $\tilde{\mathcal{F}}$ and $\tilde{\mathcal{F}}$ differs from the standard mathematical terminology introduced in Section 1.3.2.4.1 and applied to periodic distributions in Section 1.3.2.6.4: F is the *inverse* Fourier transform of ρ rather than its Fourier transform, and the calculation of ρ is called a Fourier *synthesis* in crystallography even though it is mathematically a Fourier analysis. The origin of this discrepancy may be traced to the fact that the mathematical theory of the Fourier transformation originated with the study of *temporal* periodicity, while crystallography deals with *spatial* periodicity; since the expression for the phase factor of a plane wave is $\exp[2\pi i(\nu t - \mathbf{K} \cdot \mathbf{X})]$, the difference in sign between the