

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

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

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

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

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:

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

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 $\mathbf{N} = \text{diag}(\nu_1, \nu_2, \dots, \nu_n)$, then the tensor product structure of the full multidimensional Fourier transform (Section 1.3.2.4.2.4)

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

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, ..., k_n slowest; then

$$\mathcal{F}(\mathbf{N}) = \mathcal{F}(\nu_1) \otimes \mathcal{F}(\nu_2) \otimes \dots \otimes \mathcal{F}(\nu_n),$$

where

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

and

$$\tilde{\mathcal{F}}(\mathbf{N}) = \tilde{\mathcal{F}}(\nu_1) \otimes \tilde{\mathcal{F}}(\nu_2) \otimes \dots \otimes \tilde{\mathcal{F}}(\nu_n),$$

where

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

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:

$$\begin{aligned} \tilde{\mathcal{F}}(\mathbf{N})[\tau_k \psi](k^*) &= \exp[+2\pi i k^* \cdot (\mathbf{N}^{-1}k)] \tilde{\mathcal{F}}(\mathbf{N})[\psi](k^*) \\ \mathcal{F}(\mathbf{N})[\tau_{k^*} \Psi](k) &= \exp[-2\pi i k^* \cdot (\mathbf{N}^{-1}k)] \mathcal{F}(\mathbf{N})[\Psi](k). \end{aligned}$$

(3) *Differentiation identities.* Let vectors ψ and Ψ 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 $D^{\mathbf{p}}\psi$ designates the vector of sample values of $D_{\mathbf{x}}^{\mathbf{p}}\varphi^0$ at the points of $\Lambda_{\mathbf{B}}/\Lambda_{\mathbf{A}}$, and $D^{\mathbf{p}}\Psi$ 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 $\mathbf{p} = (p_1, p_2, \dots, p_n)$

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

or equivalently

$$\begin{aligned} \tilde{\mathcal{F}}(\mathbf{N})[D^{\mathbf{p}}\psi](k^*) &= (+2\pi i k^*)^{\mathbf{p}}\Psi(k^*) \\ \mathcal{F}(\mathbf{N})[D^{\mathbf{p}}\Psi](k) &= (-2\pi i k)^{\mathbf{p}}\psi(k). \end{aligned}$$

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

$$\begin{aligned} (\varphi * \psi)(k) &= \sum_{k' \in \mathbb{Z}^n/\mathbf{N}\mathbb{Z}^n} \varphi(k') \psi(k - k') \\ (\Phi * \Psi)(k^*) &= \sum_{k' \in \mathbb{Z}^n/\mathbf{N}^T\mathbb{Z}^n} \Phi(k') \Psi(k^* - k'). \end{aligned}$$

Then

$$\begin{aligned} \tilde{\mathcal{F}}(\mathbf{N})[\Phi * \Psi](k) &= |\det \mathbf{N}| \varphi(k) \psi(k) \\ \mathcal{F}(\mathbf{N})[\varphi * \psi](k^*) &= \Phi(k^*) \Psi(k^*) \end{aligned}$$

and

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

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*.

(5) *Parseval/Plancherel property.* If $\varphi, \psi, \Phi, \Psi$ are as above, then

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

(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.

1. GENERAL RELATIONSHIPS AND 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 $N = N_1N_2$. We may write k to the base N_1 and k^* to the base N_2 as follows:

$$\begin{aligned} k &= k_1 + N_1k_2 & k_1 \in \mathbb{Z}/N_1\mathbb{Z}, & \quad k_2 \in \mathbb{Z}/N_2\mathbb{Z} \\ k^* &= k_2^* + k_1^*N_2 & k_1^* \in \mathbb{Z}/N_1\mathbb{Z}, & \quad k_2^* \in \mathbb{Z}/N_2\mathbb{Z}. \end{aligned}$$

The defining relation for $\bar{F}(N)$ may then be written:

$$\begin{aligned} X^*(k_2^* + k_1^*N_2) &= \sum_{k_1 \in \mathbb{Z}/N_1\mathbb{Z}} \sum_{k_2 \in \mathbb{Z}/N_2\mathbb{Z}} X(k_1 + N_1k_2) \\ &\quad \times e\left[\frac{(k_2^* + k_1^*N_2)(k_1 + N_1k_2)}{N_1N_2}\right]. \end{aligned}$$

The argument of $e[\cdot]$ may be expanded as

$$\frac{k_2^*k_1}{N} + \frac{k_1^*k_1}{N_1} + \frac{k_2^*k_2}{N_2} + k_1^*k_2,$$

and the last summand, being an integer, may be dropped:

$$\begin{aligned} &X^*(k_2^* + k_1^*N_2) \\ &= \sum_{k_1} \left\{ e\left(\frac{k_2^*k_1}{N}\right) \left[\sum_{k_2} X(k_1 + N_1k_2) e\left(\frac{k_2^*k_2}{N_2}\right) \right] \right\} \\ &\quad \times e\left(\frac{k_1^*k_1}{N_1}\right). \end{aligned}$$

This computation may be decomposed into five stages, as follows:

- (i) form the N_1 vectors \mathbf{Y}_{k_1} of length N_2 by the prescription

$$Y_{k_1}(k_2) = X(k_1 + N_1k_2), \quad k_1 \in \mathbb{Z}/N_1\mathbb{Z}, \quad k_2 \in \mathbb{Z}/N_2\mathbb{Z};$$

- (ii) calculate the N_1 transforms $\mathbf{Y}_{k_1}^*$ on N_2 points:

$$\mathbf{Y}_{k_1}^* = \bar{F}(N_2)[\mathbf{Y}_{k_1}], \quad k_1 \in \mathbb{Z}/N_1\mathbb{Z};$$

- (iii) form the N_2 vectors $\mathbf{Z}_{k_2^*}$ of length N_1 by the prescription

$$\mathbf{Z}_{k_2^*}(k_1) = e\left(\frac{k_2^*k_1}{N}\right) Y_{k_1}^*(k_2^*), \quad k_1 \in \mathbb{Z}/N_1\mathbb{Z}, \quad k_2^* \in \mathbb{Z}/N_2\mathbb{Z};$$

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

(iv) calculate the N_2 transforms $\mathbf{Z}_{k_2}^*$ on N_1 points:

$$\mathbf{Z}_{k_2}^* = \bar{F}(N_1)[\mathbf{Z}_{k_2}^*], \quad k_2 \in \mathbb{Z}/N_2\mathbb{Z};$$

(v) collect $X^*(k_2^* + k_1^*N_2)$ as $Z_{k_1}^*(k_1^*)$.

If the intermediate transforms in stages (ii) and (iv) are performed *in place*, i.e. with the results overwriting the data, then at stage (v) the result $X^*(k_2^* + k_1^*N_2)$ will be found at address $k_1^* + N_1k_2^*$. This phenomenon is called *scrambling* by ‘digit reversal’, and stage (v) is accordingly known as *unscrambling*.

The initial N -point transform $\bar{F}(N)$ has thus been performed as N_1 transforms $\bar{F}(N_2)$ on N_2 points, followed by N_2 transforms $\bar{F}(N_1)$ on N_1 points, thereby reducing the arithmetic cost from $(N_1N_2)^2$ to $N_1N_2(N_1 + N_2)$. The phase shifts applied at stage (iii) are traditionally called ‘twiddle factors’, and the transposition between k_1 and k_2^* can be performed by the fast recursive technique of Eklundh (1972). Clearly, this procedure can be applied recursively if N_1 and N_2 are themselves composite, leading to an overall arithmetic cost of order $N \log N$ if N has no large prime factors.

The Cooley–Tukey factorization may also be derived from a geometric rather than arithmetic argument. The decomposition $k = k_1 + N_1k_2$ is associated to a geometric partition of the residual lattice $\mathbb{Z}/N\mathbb{Z}$ into N_1 copies of $\mathbb{Z}/N_2\mathbb{Z}$, each translated by $k_1 \in \mathbb{Z}/N_1\mathbb{Z}$ and ‘blown up’ by a factor N_1 . This partition in turn induces a (direct sum) decomposition of \mathbf{X} as

$$\mathbf{X} = \sum_{k_1} \mathbf{X}_{k_1},$$

where

$$\begin{aligned} X_{k_1}(k) &= X(k) \quad \text{if } k \equiv k_1 \pmod{N_1}, \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

According to (i), \mathbf{X}_{k_1} is related to \mathbf{Y}_{k_1} by *decimation by N_1 and offset by k_1* . By Section 1.3.2.7.2, $\bar{F}(N)[\mathbf{X}_{k_1}]$ is related to $\bar{F}(N_2)[\mathbf{Y}_{k_1}]$ by *periodization by N_2 and phase shift by $e(k^*k_1/N)$* , so that

$$X^*(k^*) = \sum_{k_1} e\left(\frac{k^*k_1}{N}\right) Y_{k_1}^*(k_2^*),$$

the periodization by N_2 being reflected by the fact that $Y_{k_1}^*$ does not depend on k_1^* . Writing $k^* = k_2^* + k_1^*N_2$ and expanding k^*k_1 shows that the phase shift contains both the twiddle factor $e(k_2^*k_1/N)$ and the kernel $e(k_1^*k_1/N_1)$ of $\bar{F}(N_1)$. The Cooley–Tukey algorithm is thus naturally associated to the coset decomposition of a lattice modulo a sublattice (Section 1.3.2.7.2).

It is readily seen that essentially the same factorization can be obtained for $F(N)$, up to the complex conjugation of the twiddle factors. The normalizing constant $1/N$ arises from the normalizing constants $1/N_1$ and $1/N_2$ in $F(N_1)$ and $F(N_2)$, respectively.

Factors of 2 are particularly simple to deal with and give rise to a characteristic computational structure called a ‘butterfly loop’. If $N = 2M$, then two options exist:

(a) using $N_1 = 2$ and $N_2 = M$ leads to collecting the even-numbered coordinates of \mathbf{X} into \mathbf{Y}_0 and the odd-numbered coordinates into \mathbf{Y}_1

$$\begin{aligned} Y_0(k_2) &= X(2k_2), & k_2 &= 0, \dots, M-1, \\ Y_1(k_2) &= X(2k_2+1), & k_2 &= 0, \dots, M-1, \end{aligned}$$

and writing:

$$\begin{aligned} X^*(k_2^*) &= Y_0^*(k_2^*) + e(k_2^*/N)Y_1^*(k_2^*), \\ & \quad k_2^* = 0, \dots, M-1; \\ X^*(k_2^* + M) &= Y_0^*(k_2^*) - e(k_2^*/N)Y_1^*(k_2^*), \\ & \quad k_2^* = 0, \dots, M-1. \end{aligned}$$

This is the original version of Cooley & Tukey, and the process of formation of \mathbf{Y}_0 and \mathbf{Y}_1 is referred to as ‘decimation in time’ (i.e. decimation along the *data* index \mathbf{k}).

(b) using $N_1 = M$ and $N_2 = 2$ leads to forming

$$\begin{aligned} Z_0(k_1) &= X(k_1) + X(k_1 + M), & k_1 &= 0, \dots, M-1, \\ Z_1(k_1) &= [X(k_1) - X(k_1 + M)]e\left(\frac{k_1}{N}\right), & k_1 &= 0, \dots, M-1, \end{aligned}$$

then obtaining separately the even-numbered and odd-numbered components of \mathbf{X}^* by transforming \mathbf{Z}_0 and \mathbf{Z}_1 :

$$\begin{aligned} X^*(2k_1^*) &= Z_0^*(k_1^*), & k_1^* &= 0, \dots, M-1; \\ X^*(2k_1^* + 1) &= Z_1^*(k_1^*), & k_1^* &= 0, \dots, M-1. \end{aligned}$$

This version is due to Sande (Gentleman & Sande, 1966), and the process of separately obtaining even-numbered and odd-numbered results has led to its being referred to as ‘decimation in frequency’ (i.e. decimation along the *result* index k^*).

By repeated factoring of the number N of sample points, the calculation of $F(N)$ and $\bar{F}(N)$ can be reduced to a succession of stages, the smallest of which operate on single prime factors of N . The reader is referred to Gentleman & Sande (1966) for a particularly lucid analysis of the programming considerations which help implement this factorization efficiently; see also Singleton (1969). Powers of two are often grouped together into factors of 4 or 8, which are advantageous in that they require fewer complex multiplications than the repeated use of factors of 2. In this approach, large prime factors P are detrimental, since they require a full P^2 -size computation according to the defining formula.

1.3.3.2.2. The Good (or prime factor) algorithm

1.3.3.2.2.1. Ring structure on $\mathbb{Z}/N\mathbb{Z}$

The set $\mathbb{Z}/N\mathbb{Z}$ of congruence classes of integers modulo an integer N [see e.g. Apostol (1976), Chapter 5] inherits from \mathbb{Z} not only the additive structure used in deriving the Cooley–Tukey factorization, but also a *multiplicative* structure in which the product of two congruence classes mod N is uniquely defined as the class of the ordinary product (in \mathbb{Z}) of representatives of each class. The multiplication can be distributed over addition in the usual way, endowing $\mathbb{Z}/N\mathbb{Z}$ with the structure of a *commutative ring*.

If N is composite, the ring $\mathbb{Z}/N\mathbb{Z}$ has *zero divisors*. For example, let $N = N_1N_2$, let $n_1 \equiv N_1 \pmod{N}$, and let $n_2 \equiv N_2 \pmod{N}$: then $n_1n_2 \equiv 0 \pmod{N}$. In the general case, a product of non-zero elements will be zero whenever these elements collect together all the factors of N . These circumstances give rise to a fundamental theorem in the theory of commutative rings, the *Chinese Remainder Theorem* (CRT), which will now be stated and proved [see Apostol (1976), Chapter 5; Schroeder (1986), Chapter 16].

1.3.3.2.2.2. The Chinese remainder theorem

Let $N = N_1N_2 \dots N_d$ be factored into a product of pairwise coprime integers, so that $\text{g.c.d.}(N_i, N_j) = 1$ for $i \neq j$. Then the system of congruence equations

$$\ell \equiv \ell_j \pmod{N_j}, \quad j = 1, \dots, d,$$

has a unique solution $\ell \pmod{N}$. In other words, each $\ell \in \mathbb{Z}/N\mathbb{Z}$ is

1. GENERAL RELATIONSHIPS AND TECHNIQUES

associated in a one-to-one fashion to the d -tuple $(\ell_1, \ell_2, \dots, \ell_d)$ of its residue classes in $\mathbb{Z}/N_1\mathbb{Z}, \mathbb{Z}/N_2\mathbb{Z}, \dots, \mathbb{Z}/N_d\mathbb{Z}$.

The proof of the CRT goes as follows. Let

$$Q_j = \frac{N}{N_j} = \prod_{i \neq j} N_i.$$

Since g.c.d. $(N_j, Q_j) = 1$ there exist integers n_j and q_j such that

$$n_j N_j + q_j Q_j = 1, \quad j = 1, \dots, d,$$

then the integer

$$\ell = \sum_{i=1}^d \ell_i q_i Q_i \pmod{N}$$

is the solution. Indeed,

$$\ell \equiv \ell_j q_j Q_j \pmod{N_j}$$

because all terms with $i \neq j$ contain N_j as a factor; and

$$q_j Q_j \equiv 1 \pmod{N_j}$$

by the defining relation for q_j .

It may be noted that

$$\begin{aligned} (q_i Q_i)(q_j Q_j) &\equiv 0 \pmod{N} \text{ for } i \neq j, \\ (q_j Q_j)^2 &\equiv q_j Q_j \pmod{N}, \quad j = 1, \dots, d, \end{aligned}$$

so that the $q_j Q_j$ are mutually orthogonal *idempotents* in the ring $\mathbb{Z}/N\mathbb{Z}$, with properties formally similar to those of mutually orthogonal *projectors onto subspaces* in linear algebra. The analogy is exact, since by virtue of the CRT the ring $\mathbb{Z}/N\mathbb{Z}$ may be considered as the direct product

$$\mathbb{Z}/N_1\mathbb{Z} \times \mathbb{Z}/N_2\mathbb{Z} \times \dots \times \mathbb{Z}/N_d\mathbb{Z}$$

via the two mutually inverse mappings:

- (i) $\ell \mapsto (\ell_1, \ell_2, \dots, \ell_d)$ by $\ell \equiv \ell_j \pmod{N_j}$ for each j ;
- (ii) $(\ell_1, \ell_2, \dots, \ell_d) \mapsto \ell$ by $\ell = \sum_{i=1}^d \ell_i q_i Q_i \pmod{N}$.

The mapping defined by (ii) is sometimes called the ‘CRT reconstruction’ of ℓ from the ℓ_j .

These two mappings have the property of sending sums to sums and products to products, i.e.:

- (i) $\ell + \ell' \mapsto (\ell_1 + \ell'_1, \ell_2 + \ell'_2, \dots, \ell_d + \ell'_d)$
 $\ell \ell' \mapsto (\ell_1 \ell'_1, \ell_2 \ell'_2, \dots, \ell_d \ell'_d)$
- (ii) $(\ell_1 + \ell'_1, \ell_2 + \ell'_2, \dots, \ell_d + \ell'_d) \mapsto \ell + \ell'$
 $(\ell_1 \ell'_1, \ell_2 \ell'_2, \dots, \ell_d \ell'_d) \mapsto \ell \ell'$

(the last proof requires using the properties of the idempotents $q_j Q_j$). This may be described formally by stating that the CRT establishes a *ring isomorphism*:

$$\mathbb{Z}/N\mathbb{Z} \cong (\mathbb{Z}/N_1\mathbb{Z}) \times \dots \times (\mathbb{Z}/N_d\mathbb{Z}).$$

1.3.3.2.2.3. The prime factor algorithm

The CRT will now be used to factor the N -point DFT into a tensor product of d transforms, the j th of length N_j .

Let the indices k and k^* be subjected to the following mappings:

- (i) $k \mapsto (k_1, k_2, \dots, k_d), k_j \in \mathbb{Z}/N_j\mathbb{Z}$, by $k_j \equiv k \pmod{N_j}$ for each j , with reconstruction formula

$$k = \sum_{i=1}^d k_i q_i Q_i \pmod{N};$$

- (ii) $k^* \mapsto (k_1^*, k_2^*, \dots, k_d^*), k_j^* \in \mathbb{Z}/N_j\mathbb{Z}$, by $k_j^* \equiv q_j k^* \pmod{N_j}$ for each j , with reconstruction formula

$$k^* = \sum_{i=1}^d k_i^* Q_i \pmod{N}.$$

Then

$$\begin{aligned} k^* k &= \left(\sum_{i=1}^d k_i^* Q_i \right) \left(\sum_{j=1}^d k_j q_j Q_j \right) \pmod{N} \\ &= \sum_{i,j=1}^d k_i^* k_j q_j Q_i Q_j \pmod{N}. \end{aligned}$$

Cross terms with $i \neq j$ vanish since they contain all the factors of N , hence

$$\begin{aligned} k^* k &= \sum_{j=1}^d q_j Q_j^2 k_j^* k_j \pmod{N} \\ &= \sum_{j=1}^d (1 - n_j N_j) Q_j k_j^* k_j \pmod{N}. \end{aligned}$$

Dividing by N , which may be written as $N_j Q_j$ for each j , yields

$$\begin{aligned} \frac{k^* k}{N} &= \sum_{j=1}^d (1 - n_j N_j) \frac{Q_j}{N_j Q_j} k_j^* k_j \pmod{1} \\ &= \sum_{j=1}^d \left(\frac{1}{N_j} - n_j \right) k_j^* k_j \pmod{1}, \end{aligned}$$

and hence

$$\frac{k^* k}{N} \equiv \sum_{j=1}^d \frac{k_j^* k_j}{N_j} \pmod{1}.$$

Therefore, by the multiplicative property of $e(\cdot)$,

$$e\left(\frac{k^* k}{N}\right) \equiv \bigotimes_{j=1}^d e\left(\frac{k_j^* k_j}{N_j}\right).$$

Let $\mathbf{X} \in L(\mathbb{Z}/N\mathbb{Z})$ be described by a one-dimensional array $X(k)$ indexed by k . The index mapping (i) turns \mathbf{X} into an element of $L(\mathbb{Z}/N_1\mathbb{Z} \times \dots \times \mathbb{Z}/N_d\mathbb{Z})$ described by a d -dimensional array $X(k_1, \dots, k_d)$; the latter may be transformed by $\bar{F}(N_1) \otimes \dots \otimes \bar{F}(N_d)$ into a new array $X^*(k_1^*, k_2^*, \dots, k_d^*)$. Finally, the one-dimensional array of results $X^*(k^*)$ will be obtained by reconstructing k^* according to (ii).

The prime factor algorithm, like the Cooley–Tukey algorithm, reindexes a 1D transform to turn it into d separate transforms, but the use of coprime factors and CRT index mapping leads to the further gain that *no twiddle factors* need to be applied between the successive transforms (see Good, 1971). This makes up for the cost of the added complexity of the CRT index mapping.

The natural factorization of N for the prime factor algorithm is thus its factorization into prime powers: $\bar{F}(N)$ is then the tensor product of separate transforms (one for each prime power factor $N_j = p_j^{v_j}$) whose results can be reassembled without twiddle factors. The separate factors p_j within each N_j must then be dealt with by another algorithm (e.g. Cooley–Tukey, which does require twiddle factors). Thus, the DFT on a prime number of points remains undecomposable.

1.3.3.2.3. The Rader algorithm

The previous two algorithms essentially reduce the calculation of the DFT on N points for N composite to the calculation of smaller DFTs on prime numbers of points, the latter remaining irreducible. However, Rader (1968) showed that the p -point DFT for p an odd

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

prime can itself be factored by invoking some extra arithmetic structure present in $\mathbb{Z}/p\mathbb{Z}$.

1.3.3.2.3.1. N an odd prime

The ring $\mathbb{Z}/p\mathbb{Z} = \{0, 1, 2, \dots, p-1\}$ has the property that its $p-1$ non-zero elements, called *units*, form a *multiplicative group* $U(p)$. In particular, all units $r \in U(p)$ have a unique multiplicative inverse in $\mathbb{Z}/p\mathbb{Z}$, i.e. a unit $s \in U(p)$ such that $rs \equiv 1 \pmod{p}$. This endows $\mathbb{Z}/p\mathbb{Z}$ with the structure of a *finite field*.

Furthermore, $U(p)$ is a *cyclic group*, i.e. consists of the successive powers $g^m \pmod{p}$ of a generator g called a *primitive root mod p* (such a g may not be unique, but it always exists). For instance, for $p = 7$, $U(7) = \{1, 2, 3, 4, 5, 6\}$ is generated by $g = 3$, whose successive powers mod 7 are:

$$g^0 = 1, \quad g^1 = 3, \quad g^2 = 2, \quad g^3 = 6, \quad g^4 = 4, \quad g^5 = 5$$

[see Apostol (1976), Chapter 10].

The basis of Rader's algorithm is to bring to light a hidden regularity in the matrix $F(p)$ by permuting the basis vectors \mathbf{u}_k and \mathbf{v}_{k^*} of $L(\mathbb{Z}/p\mathbb{Z})$ as follows:

$$\begin{aligned} \mathbf{u}'_0 &= \mathbf{u}_0 \\ \mathbf{u}'_m &= \mathbf{u}_k \quad \text{with } k = g^m, \quad m = 1, \dots, p-1; \\ \mathbf{v}'_0 &= \mathbf{v}_0 \\ \mathbf{v}'_{m^*} &= \mathbf{v}_{k^*} \quad \text{with } k^* = g^{m^*}, \quad m^* = 1, \dots, p-1; \end{aligned}$$

where g is a primitive root mod p .

With respect to these new bases, the matrix representing $\bar{F}(p)$ will have the following elements:

$$\begin{aligned} \text{element } (0, 0) &= 1 \\ \text{element } (0, m+1) &= 1 \quad \text{for all } m = 0, \dots, p-2, \\ \text{element } (m^*+1, 0) &= 1 \quad \text{for all } m^* = 0, \dots, p-2, \\ \text{element } (m^*+1, m+1) &= e\left(\frac{k^*k}{p}\right) \\ &= e(g^{(m^*+m)/p}) \\ &\quad \text{for all } m^* = 0, \dots, p-2. \end{aligned}$$

Thus the 'core' $\bar{C}(p)$ of matrix $\bar{F}(p)$, of size $(p-1) \times (p-1)$, formed by the elements with two non-zero indices, has a so-called *skew-circulant* structure because element (m^*, m) depends only on $m^* + m$. Simplification may now occur because multiplication by $\bar{C}(p)$ is closely related to a *cyclic convolution*. Introducing the notation $C(m) = e(g^{m/p})$ we may write the relation $\mathbf{Y}^* = \bar{F}(p)\mathbf{Y}$ in the permuted bases as

$$\begin{aligned} Y^*(0) &= \sum_k Y(k) \\ Y^*(m^*+1) &= Y(0) + \sum_{m=0}^{p-2} C(m^*+m)Y(m+1) \\ &= Y(0) + \sum_{m=0}^{p-2} C(m^*-m)Z(m) \\ &= Y(0) + (\mathbf{C} * \mathbf{Z})(m^*), \quad m^* = 0, \dots, p-2, \end{aligned}$$

where \mathbf{Z} is defined by $Z(m) = Y(p-m-2)$, $m = 0, \dots, p-2$.

Thus \mathbf{Y}^* may be obtained by cyclic convolution of \mathbf{C} and \mathbf{Z} , which may for instance be calculated by

$$\mathbf{C} * \mathbf{Z} = F(p-1)[\bar{F}(p-1)[\mathbf{C}] \times \bar{F}(p-1)[\mathbf{Z}]],$$

where \times denotes the component-wise multiplication of vectors. Since p is odd, $p-1$ is always divisible by 2 and may even be

highly composite. In that case, factoring $\bar{F}(p-1)$ by means of the Cooley–Tukey or Good methods leads to an algorithm of complexity $p \log p$ rather than p^2 for $\bar{F}(p)$. An added bonus is that, because $g^{(p-1)/2} = -1$, the elements of $\bar{F}(p-1)[\mathbf{C}]$ can be shown to be either purely real or purely imaginary, which halves the number of real multiplications involved.

1.3.3.2.3.2. N a power of an odd prime

This idea was extended by Winograd (1976, 1978) to the treatment of prime powers $N = p^\nu$, using the cyclic structure of the multiplicative group of units $U(p^\nu)$. The latter consists of all those elements of $\mathbb{Z}/p^\nu\mathbb{Z}$ which are not divisible by p , and thus has $q_\nu = p^{\nu-1}(p-1)$ elements. It is cyclic, and there exist primitive roots g modulo p^ν such that

$$U(p^\nu) = \{1, g, g^2, g^3, \dots, g^{q_\nu-1}\}.$$

The $p^{\nu-1}$ elements divisible by p , which are divisors of zero, have to be treated separately just as 0 had to be treated separately for $N = p$.

When $k^* \notin U(p^\nu)$, then $k^* = pk_1^*$ with $k_1^* \in \mathbb{Z}/p^{\nu-1}\mathbb{Z}$. The results $X^*(pk_1^*)$ are p -decimated, hence can be obtained via the $p^{\nu-1}$ -point DFT of the $p^{\nu-1}$ -periodized data \mathbf{Y} :

$$X^*(pk_1^*) = \bar{F}(p^{\nu-1})[\mathbf{Y}](k_1^*)$$

with

$$Y(k_1) = \sum_{k_2 \in \mathbb{Z}/p\mathbb{Z}} X(k_1 + p^{\nu-1}k_2).$$

When $k^* \in U(p^\nu)$, then we may write

$$X^*(k^*) = X_0^*(k^*) + X_1^*(k^*),$$

where \mathbf{X}_0^* contains the contributions from $k \notin U(p^\nu)$ and \mathbf{X}_1^* those from $k \in U(p^\nu)$. By a converse of the previous calculation, \mathbf{X}_0^* arises from p -decimated data \mathbf{Z} , hence is the $p^{\nu-1}$ -periodization of the $p^{\nu-1}$ -point DFT of these data:

$$X_0^*(p^{\nu-1}k_1^* + k_2^*) = \bar{F}(p^{\nu-1})[\mathbf{Z}](k_2^*)$$

with

$$Z(k_2) = X(pk_2), \quad k_2 \in \mathbb{Z}/p^{\nu-1}\mathbb{Z}$$

(the $p^{\nu-1}$ -periodicity follows implicitly from the fact that the transform on the right-hand side is independent of $k_1^* \in \mathbb{Z}/p\mathbb{Z}$).

Finally, the contribution X_1^* from all $k \in U(p^\nu)$ may be calculated by reindexing by the powers of a primitive root g modulo p^ν , i.e. by writing

$$X_1^*(g^{m^*}) = \sum_{m=0}^{q_\nu-1} X(g^m)e(g^{(m^*+m)/p^\nu})$$

then carrying out the multiplication by the skew-circulant matrix core as a convolution.

Thus the DFT of size p^ν may be reduced to two DFTs of size $p^{\nu-1}$ (dealing, respectively, with p -decimated results and p -decimated data) and a convolution of size $q_\nu = p^{\nu-1}(p-1)$. The latter may be 'diagonalized' into a multiplication by purely real or purely imaginary numbers (because $g^{(q_\nu/2)} = -1$) by two DFTs, whose factoring in turn leads to DFTs of size $p^{\nu-1}$ and $p-1$. This method, applied recursively, allows the complete decomposition of the DFT on p^ν points into arbitrarily small DFTs.

1.3.3.2.3.3. N a power of 2

When $N = 2^\nu$, the same method can be applied, except for a slight modification in the calculation of \mathbf{X}_1^* . There is no primitive root modulo 2^ν for $\nu > 2$: the group $U(2^\nu)$ is the direct product of *two* cyclic groups, the first (of order 2) generated by -1 , the second (of order $N/4$) generated by 3 or 5. One then uses a representation

1. GENERAL RELATIONSHIPS AND TECHNIQUES

$$k = (-1)^{m_1} 5^{m_2}$$

$$k^* = (-1)^{m_1^*} 5^{m_2^*}$$

and the reindexed core matrix gives rise to a two-dimensional convolution. The latter may be carried out by means of two 2D DFTs on $2 \times (N/4)$ points.

1.3.3.2.4. The Winograd algorithms

The cyclic convolutions generated by Rader's multiplicative reindexing may be evaluated more economically than through DFTs if they are re-examined within a new algebraic setting, namely the theory of congruence classes of polynomials [see, for instance, Blahut (1985), Chapter 2; Schroeder (1986), Chapter 24].

The set, denoted $\mathbb{K}[X]$, of polynomials in one variable with coefficients in a given field \mathbb{K} has many of the formal properties of the set \mathbb{Z} of rational integers: it is a *ring* with no zero divisors and has a *Euclidean algorithm* on which a theory of divisibility can be built.

Given a polynomial $P(z)$, then for every $W(z)$ there exist unique polynomials $Q(z)$ and $R(z)$ such that

$$W(z) = P(z)Q(z) + R(z)$$

and

$$\text{degree}(R) < \text{degree}(P).$$

$R(z)$ is called the *residue* of $H(z)$ modulo $P(z)$. Two polynomials $H_1(z)$ and $H_2(z)$ having the same residue modulo $P(z)$ are said to be *congruent* modulo $P(z)$, which is denoted by

$$H_1(z) \equiv H_2(z) \pmod{P(z)}.$$

If $H(z) \equiv 0 \pmod{P(z)}$, $H(z)$ is said to be *divisible* by $P(z)$. If $H(z)$ only has divisors of degree zero in $\mathbb{K}[X]$, it is said to be *irreducible over* \mathbb{K} (this notion depends on \mathbb{K}). Irreducible polynomials play in $\mathbb{K}[X]$ a role analogous to that of prime numbers in \mathbb{Z} , and any polynomial over \mathbb{K} has an essentially unique factorization as a product of irreducible polynomials.

There exists a *Chinese remainder theorem* (CRT) for polynomials. Let $P(z) = P_1(z) \dots P_d(z)$ be factored into a product of pairwise coprime polynomials [*i.e.* $P_i(z)$ and $P_j(z)$ have no common factor for $i \neq j$]. Then the system of congruence equations

$$H(z) \equiv H_j(z) \pmod{P_j(z)}, \quad j = 1, \dots, d,$$

has a unique solution $H(z)$ modulo $P(z)$. This solution may be constructed by a procedure similar to that used for integers. Let

$$Q_j(z) = P(z)/P_j(z) = \prod_{i \neq j} P_i(z).$$

Then P_j and Q_j are coprime, and the Euclidean algorithm may be used to obtain polynomials $p_j(z)$ and $q_j(z)$ such that

$$p_j(z)P_j(z) + q_j(z)Q_j(z) = 1.$$

With $S_i(z) = q_i(z)Q_i(z)$, the polynomial

$$H(z) = \sum_{i=1}^d S_i(z)H_i(z) \pmod{P(z)}$$

is easily shown to be the desired solution.

As with integers, it can be shown that the 1:1 correspondence between $H(z)$ and $H_j(z)$ sends sums to sums and products to products, *i.e.* establishes a *ring isomorphism*:

$$\mathbb{K}[X] \pmod{P} \cong (\mathbb{K}[X] \pmod{P_1}) \times \dots \times (\mathbb{K}[X] \pmod{P_d}).$$

These results will now be applied to the efficient calculation of cyclic convolutions. Let $\mathbf{U} = (u_0, u_1, \dots, u_{N-1})$ and $\mathbf{V} = (v_0, v_1, \dots, v_{N-1})$ be two vectors of length N , and let $\mathbf{W} =$

$(w_0, w_1, \dots, w_{N-1})$ be obtained by cyclic convolution of \mathbf{U} and \mathbf{V} :

$$w_n = \sum_{m=0}^{N-1} u_m v_{n-m}, \quad n = 0, \dots, N-1.$$

The very simple but crucial result is that this cyclic convolution may be carried out by *polynomial multiplication modulo* $(z^N - 1)$: if

$$U(z) = \sum_{l=0}^{N-1} u_l z^l$$

$$V(z) = \sum_{m=0}^{N-1} v_m z^m$$

$$W(z) = \sum_{n=0}^{N-1} w_n z^n$$

then the above relation is equivalent to

$$W(z) \equiv U(z)V(z) \pmod{z^N - 1}.$$

Now the polynomial $z^N - 1$ can be *factored* over the field of rational numbers into irreducible factors called *cyclotomic polynomials*: if d is the number of divisors of N , including 1 and N , then

$$z^N - 1 = \prod_{i=1}^d P_i(z),$$

where the cyclotomics $P_i(z)$ are well known (Nussbaumer, 1981; Schroeder, 1986, Chapter 22). We may now invoke the CRT, and exploit the ring isomorphism it establishes to simplify the calculation of $W(z)$ from $U(z)$ and $V(z)$ as follows:

(i) compute the d residual polynomials

$$\begin{aligned} U_i(z) &\equiv U(z) \pmod{P_i(z)}, & i = 1, \dots, d, \\ V_i(z) &\equiv V(z) \pmod{P_i(z)}, & i = 1, \dots, d; \end{aligned}$$

(ii) compute the d polynomial products

$$W_i(z) \equiv U_i(z)V_i(z) \pmod{P_i(z)}, \quad i = 1, \dots, d;$$

(iii) use the CRT reconstruction formula just proved to recover $W(z)$ from the $W_i(z)$:

$$W(z) \equiv \sum_{i=1}^d S_i(z)W_i(z) \pmod{z^N - 1}.$$

When N is not too large, *i.e.* for 'short cyclic convolutions', the $P_i(z)$ are very simple, with coefficients 0 or ± 1 , so that (i) only involves a small number of additions. Furthermore, special techniques have been developed to multiply general polynomials modulo cyclotomic polynomials, thus helping keep the number of multiplications in (ii) and (iii) to a minimum. As a result, cyclic convolutions can be calculated rapidly when N is sufficiently composite.

It will be recalled that Rader's multiplicative indexing often gives rise to cyclic convolutions of length $p - 1$ for p an odd prime. Since $p - 1$ is highly composite for all $p \leq 50$ other than 23 and 47, these cyclic convolutions can be performed more efficiently by the above procedure than by DFT.

These combined algorithms are due to Winograd (1977, 1978, 1980), and are known collectively as 'Winograd small FFT algorithms'. Winograd also showed that they can be thought of as bringing the DFT matrix \mathbf{F} to the following 'normal form':

$$\mathbf{F} = \mathbf{CBA},$$

where

\mathbf{A} is an integer matrix with entries 0, ± 1 , defining the 'pre-additions',

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

\mathbf{B} is a diagonal matrix of multiplications,

\mathbf{C} is a matrix with entries $0, \pm 1, \pm i$, defining the ‘post-additions’.

The elements on the diagonal of \mathbf{B} can be shown to be either real or pure imaginary, by the same argument as in Section 1.3.2.3.1. Matrices \mathbf{A} and \mathbf{C} may be rectangular rather than square, so that intermediate results may require extra storage space.

1.3.3.3. Multidimensional algorithms

From an algorithmic point of view, the distinction between one-dimensional (1D) and multidimensional DFTs is somewhat blurred by the fact that some factoring techniques turn a 1D transform into a multidimensional one. The distinction made here, however, is a practical one and is based on the dimensionality of the indexing sets for data and results. This section will therefore be concerned with the problem of factoring the DFT when the *indexing sets* for the input data and output results are multidimensional.

1.3.3.3.1. The method of successive one-dimensional transforms

The DFT was defined in Section 1.3.2.7.4 in an n -dimensional setting and it was shown that when the decimation matrix \mathbf{N} is diagonal, say $\mathbf{N} = \text{diag}(N^{(1)}, N^{(2)}, \dots, N^{(n)})$, then $\bar{F}(N)$ has a tensor product structure:

$$\bar{F}(\mathbf{N}) = \bar{F}(N^{(1)}) \otimes \bar{F}(N^{(2)}) \otimes \dots \otimes \bar{F}(N^{(n)}).$$

This may be rewritten as follows:

$$\begin{aligned} \bar{F}(\mathbf{N}) &= [\bar{F}(N^{(1)}) \otimes I_{N^{(2)}} \otimes \dots \otimes I_{N^{(n)}}] \\ &\quad \times [I_{N^{(1)}} \otimes \bar{F}(N^{(2)}) \otimes \dots \otimes I_{N^{(n)}}] \\ &\quad \times \dots \\ &\quad \times [I_{N^{(1)}} \otimes I_{N^{(2)}} \otimes \dots \otimes \bar{F}(N^{(n)})], \end{aligned}$$

where the I 's are identity matrices and \times denotes ordinary matrix multiplication. The matrix within each bracket represents a one-dimensional DFT along one of the n dimensions, the other dimensions being left untransformed. As these matrices commute, the order in which the successive 1D DFTs are performed is immaterial.

This is the most straightforward method for building an n -dimensional algorithm from existing 1D algorithms. It is known in crystallography under the name of ‘Beavers–Lipson factorization’ (Section 1.3.4.3.1), and in signal processing as the ‘row–column method’.

1.3.3.3.2. Multidimensional factorization

Substantial reductions in the arithmetic cost, as well as gains in flexibility, can be obtained if the factoring of the DFT is carried out in several dimensions simultaneously. The presentation given here is a generalization of that of Mersereau & Speake (1981), using the abstract setting established independently by Auslander, Tolimieri & Winograd (1982).

Let us return to the general n -dimensional setting of Section 1.3.2.7.4, where the DFT was defined for an arbitrary decimation matrix \mathbf{N} by the formulae (where $|\mathbf{N}|$ denotes $|\det \mathbf{N}|$):

$$F(\mathbf{N}) : X(\mathbf{k}) = \frac{1}{|\mathbf{N}|} \sum_{\mathbf{k}^*} X^*(\mathbf{k}^*) e[-\mathbf{k}^* \cdot (\mathbf{N}^{-1}\mathbf{k})]$$

$$\bar{F}(\mathbf{N}) : X^*(\mathbf{k}^*) = \sum_{\mathbf{k}} X(\mathbf{k}) e[\mathbf{k}^* \cdot (\mathbf{N}^{-1}\mathbf{k})]$$

with

$$\mathbf{k} \in \mathbb{Z}^n / \mathbf{N}\mathbb{Z}^n, \quad \mathbf{k}^* \in \mathbb{Z}^n / \mathbf{N}^T \mathbb{Z}^n.$$

1.3.3.3.2.1. Multidimensional Cooley–Tukey factorization

Let us now assume that this decimation can be factored into d successive decimations, *i.e.* that

$$\mathbf{N} = \mathbf{N}_1 \mathbf{N}_2 \dots \mathbf{N}_{d-1} \mathbf{N}_d$$

and hence

$$\mathbf{N}^T = \mathbf{N}_d^T \mathbf{N}_{d-1}^T \dots \mathbf{N}_2^T \mathbf{N}_1^T.$$

Then the coset decomposition formulae corresponding to these successive decimations (Section 1.3.2.7.1) can be combined as follows:

$$\begin{aligned} \mathbb{Z}^n &= \bigcup_{\mathbf{k}_1} (\mathbf{k}_1 + \mathbf{N}_1 \mathbb{Z}^n) \\ &= \bigcup_{\mathbf{k}_1} \left\{ \mathbf{k}_1 + \mathbf{N}_1 \left[\bigcup_{\mathbf{k}_2} (\mathbf{k}_2 + \mathbf{N}_2 \mathbb{Z}^n) \right] \right\} \\ &= \dots \\ &= \bigcup_{\mathbf{k}_1} \dots \bigcup_{\mathbf{k}_d} (\mathbf{k}_1 + \mathbf{N}_1 \mathbf{k}_2 + \dots + \mathbf{N}_1 \mathbf{N}_2 \times \dots \times \mathbf{N}_{d-1} \mathbf{k}_d + \mathbf{N} \mathbb{Z}^n) \end{aligned}$$

with $\mathbf{k}_j \in \mathbb{Z}^n / \mathbf{N}_j \mathbb{Z}^n$. Therefore, any $\mathbf{k} \in \mathbb{Z} / \mathbf{N} \mathbb{Z}^n$ may be written uniquely as

$$\mathbf{k} = \mathbf{k}_1 + \mathbf{N}_1 \mathbf{k}_2 + \dots + \mathbf{N}_1 \mathbf{N}_2 \times \dots \times \mathbf{N}_{d-1} \mathbf{k}_d.$$

Similarly:

$$\begin{aligned} \mathbb{Z}^n &= \bigcup_{\mathbf{k}_d^*} (\mathbf{k}_d^* + \mathbf{N}_d^T \mathbb{Z}^n) \\ &= \dots \\ &= \bigcup_{\mathbf{k}_d^*} \dots \bigcup_{\mathbf{k}_1^*} (\mathbf{k}_d^* + \mathbf{N}_d^T \mathbf{k}_{d-1}^* + \dots + \mathbf{N}_d^T \times \dots \times \mathbf{N}_2^T \mathbf{k}_1^* \\ &\quad + \mathbf{N}^T \mathbb{Z}^n) \end{aligned}$$

so that any $\mathbf{k}^* \in \mathbb{Z}^n / \mathbf{N}^T \mathbb{Z}^n$ may be written uniquely as

$$\mathbf{k}^* = \mathbf{k}_d^* + \mathbf{N}_d^T \mathbf{k}_{d-1}^* + \dots + \mathbf{N}_d^T \times \dots \times \mathbf{N}_2^T \mathbf{k}_1^*$$

with $\mathbf{k}_j^* \in \mathbb{Z}^n / \mathbf{N}_j^T \mathbb{Z}^n$. These decompositions are the vector analogues of the multi-radix number representation systems used in the Cooley–Tukey factorization.

We may then write the definition of $\bar{F}(\mathbf{N})$ with $d = 2$ factors as

$$\begin{aligned} X^*(\mathbf{k}_2^* + \mathbf{N}_2^T \mathbf{k}_1^*) &= \sum_{\mathbf{k}_1} \sum_{\mathbf{k}_2} X(\mathbf{k}_1 + \mathbf{N}_1 \mathbf{k}_2) \\ &\quad \times e[(\mathbf{k}_2^{*T} + \mathbf{k}_1^{*T} \mathbf{N}_2) \mathbf{N}_2^{-1} \mathbf{N}_1^{-1} (\mathbf{k}_1 + \mathbf{N}_1 \mathbf{k}_2)]. \end{aligned}$$

The argument of $e(-)$ may be expanded as

$$\mathbf{k}_2^* \cdot (\mathbf{N}_1^{-1} \mathbf{k}_1) + \mathbf{k}_1^* \cdot (\mathbf{N}_1^{-1} \mathbf{k}_1) + \mathbf{k}_2^* \cdot (\mathbf{N}_2^{-1} \mathbf{k}_2) + \mathbf{k}_1^* \cdot \mathbf{k}_2.$$

The first summand may be recognized as a twiddle factor, the second and third as the kernels of $\bar{F}(\mathbf{N}_1)$ and $\bar{F}(\mathbf{N}_2)$, respectively, while the fourth is an integer which may be dropped. We are thus led to a ‘vector-radix’ version of the Cooley–Tukey algorithm, in which the successive decimations may be introduced in all n dimensions simultaneously by general integer matrices. The computation may be decomposed into five stages analogous to those of the one-dimensional algorithm of Section 1.3.3.2.1:

(i) form the $|\mathbf{N}_1|$ vectors $\mathbf{Y}_{\mathbf{k}_1}$ of shape \mathbf{N}_2 by

$$Y_{\mathbf{k}_1}(\mathbf{k}_2) = X(\mathbf{k}_1 + \mathbf{N}_1 \mathbf{k}_2), \quad \mathbf{k}_1 \in \mathbb{Z}^n / \mathbf{N}_1 \mathbb{Z}^n, \quad \mathbf{k}_2 \in \mathbb{Z}^n / \mathbf{N}_2 \mathbb{Z}^n;$$

1. GENERAL RELATIONSHIPS AND TECHNIQUES

(ii) calculate the $|\mathbf{N}_1|$ transforms $\mathbf{Y}_{\mathbf{k}_1}^*$ on $|\mathbf{N}_2|$ points:

$$Y_{\mathbf{k}_1}^*(\mathbf{k}_2^*) = \sum_{\mathbf{k}_2} e[\mathbf{k}_2^* \cdot (\mathbf{N}_2^{-1} \mathbf{k}_2)] Y_{\mathbf{k}_1}(\mathbf{k}_2), \quad \mathbf{k}_1 \in \mathbb{Z}^n / \mathbf{N}_1 \mathbb{Z}^n;$$

(iii) form the $|\mathbf{N}_2|$ vectors $\mathbf{Z}_{\mathbf{k}_2^*}$ of shape \mathbf{N}_1 by

$$\mathbf{Z}_{\mathbf{k}_2^*}(\mathbf{k}_1) = e[\mathbf{k}_2^* \cdot (\mathbf{N}^{-1} \mathbf{k}_1)] Y_{\mathbf{k}_1}^*(\mathbf{k}_2^*), \quad \mathbf{k}_1 \in \mathbb{Z}^n / \mathbf{N}_1 \mathbb{Z}^n, \\ \mathbf{k}_2^* \in \mathbb{Z}^n / \mathbf{N}_2^T \mathbb{Z}^n;$$

(iv) calculate the $|\mathbf{N}_2|$ transforms $\mathbf{Z}_{\mathbf{k}_2^*}$ on $|\mathbf{N}_1|$ points:

$$Z_{\mathbf{k}_2^*}^*(\mathbf{k}_1^*) = \sum_{\mathbf{k}_1} e[\mathbf{k}_1^* \cdot (\mathbf{N}_1^{-1} \mathbf{k}_1)] Z_{\mathbf{k}_2^*}(\mathbf{k}_1), \quad \mathbf{k}_2^* \in \mathbb{Z}^n / \mathbf{N}_2^T \mathbb{Z}^n;$$

(v) collect $X^*(\mathbf{k}_2^* + \mathbf{N}_2^T \mathbf{k}_1^*)$ as $Z_{\mathbf{k}_2^*}^*(\mathbf{k}_1^*)$.

The initial $|\mathbf{N}|$ -point transform $\bar{F}(\mathbf{N})$ can thus be performed as $|\mathbf{N}_1|$ transforms $\bar{F}(\mathbf{N}_2)$ on $|\mathbf{N}_2|$ points, followed by $|\mathbf{N}_2|$ transforms $\bar{F}(\mathbf{N}_1)$ on $|\mathbf{N}_1|$ points. This process can be applied successively to all d factors. The same decomposition applies to $F(\mathbf{N})$, up to the complex conjugation of twiddle factors, the normalization factor $1/|\mathbf{N}|$ being obtained as the product of the factors $1/|\mathbf{N}_j|$ in the successive partial transforms $F(\mathbf{N}_j)$.

The geometric interpretation of this factorization in terms of partial transforms on translates of sublattices applies in full to this n -dimensional setting; in particular, the twiddle factors are seen to be related to the residual translations which place the sublattices in register within the big lattice. If the intermediate transforms are performed *in place*, then the quantity

$$X^*(\mathbf{k}_d^* + \mathbf{N}_d^T \mathbf{k}_{d-1}^* + \dots + \mathbf{N}_d^T \mathbf{N}_{d-1}^T \times \dots \times \mathbf{N}_2^T \mathbf{k}_1^*)$$

will eventually be found at location

$$\mathbf{k}_1^* + \mathbf{N}_1 \mathbf{k}_2^* + \dots + \mathbf{N}_1 \mathbf{N}_2 \times \dots \times \mathbf{N}_{d-1} \mathbf{k}_d^*,$$

so that the final results will have to be *unscrambled* by a process which may be called ‘coset reversal’, the vector equivalent of digit reversal.

Factoring by 2 in all n dimensions simultaneously, *i.e.* taking $\mathbf{N} = 2\mathbf{M}$, leads to ‘ n -dimensional butterflies’. Decimation in time corresponds to the choice $\mathbf{N}_1 = 2\mathbf{I}$, $\mathbf{N}_2 = \mathbf{M}$, so that $\mathbf{k}_1 \in \mathbb{Z}^n / 2\mathbb{Z}^n$ is an n -dimensional parity class; the calculation then proceeds by

$$Y_{\mathbf{k}_1}(\mathbf{k}_2) = X(\mathbf{k}_1 + 2\mathbf{k}_2), \quad \mathbf{k}_1 \in \mathbb{Z}^n / 2\mathbb{Z}^n, \quad \mathbf{k}_2 \in \mathbb{Z}^n / \mathbf{M}\mathbb{Z}^n, \\ Y_{\mathbf{k}_1}^* = \bar{F}(\mathbf{M})[Y_{\mathbf{k}_1}], \quad \mathbf{k}_1 \in \mathbb{Z}^n / 2\mathbb{Z}^n; \\ X^*(\mathbf{k}_2^* + \mathbf{M}^T \mathbf{k}_1^*) = \sum_{\mathbf{k}_1 \in \mathbb{Z}^n / 2\mathbb{Z}^n} (-1)^{\mathbf{k}_1^* \cdot \mathbf{k}_1} \\ \times e[\mathbf{k}_2^* \cdot (\mathbf{N}^{-1} \mathbf{k}_1)] Y_{\mathbf{k}_1}^*(\mathbf{k}_2^*).$$

Decimation in frequency corresponds to the choice $\mathbf{N}_1 = \mathbf{M}$, $\mathbf{N}_2 = 2\mathbf{I}$, so that $\mathbf{k}_2 \in \mathbb{Z}^n / 2\mathbb{Z}^n$ labels ‘octant’ blocks of shape \mathbf{M} ; the calculation then proceeds through the following steps:

$$\mathbf{Z}_{\mathbf{k}_2^*}(\mathbf{k}_1) = \left[\sum_{\mathbf{k}_2 \in \mathbb{Z}^n / 2\mathbb{Z}^n} (-1)^{\mathbf{k}_2^* \cdot \mathbf{k}_2} X(\mathbf{k}_1 + \mathbf{M}\mathbf{k}_2) \right] \\ \times e[\mathbf{k}_2^* \cdot (\mathbf{N}^{-1} \mathbf{k}_1)], \\ \mathbf{Z}_{\mathbf{k}_2^*}^* = \bar{F}(\mathbf{M})[\mathbf{Z}_{\mathbf{k}_2^*}], \\ X^*(\mathbf{k}_2^* + 2\mathbf{k}_1^*) = Z_{\mathbf{k}_2^*}^*(\mathbf{k}_1^*),$$

i.e. the 2^n parity classes of results, corresponding to the different $\mathbf{k}_2^* \in \mathbb{Z}^n / 2\mathbb{Z}^n$, are obtained separately. When the dimension n is 2 and the decimating matrix is diagonal, this analysis reduces to the ‘vector radix FFT’ algorithms proposed by Rivard (1977) and Harris *et al.* (1977). These lead to substantial reductions in the number M of multiplications compared to the row–column method:

M is reduced to $3M/4$ by simultaneous 2×2 factoring, and to $15M/32$ by simultaneous 4×4 factoring.

The use of a non-diagonal decimating matrix may bring savings in computing time if the spectrum of the band-limited function under study is of such a shape as to pack more compactly in a non-rectangular than in a rectangular lattice (Mersereau, 1979). If, for instance, the support K of the spectrum Φ is contained in a sphere, then a decimation matrix producing a close packing of these spheres will yield an aliasing-free DFT algorithm with fewer sample points than the standard algorithm using a rectangular lattice.

1.3.3.3.2.2. Multidimensional prime factor algorithm

Suppose that the decimation matrix \mathbf{N} is diagonal

$$\mathbf{N} = \text{diag} (N^{(1)}, N^{(2)}, \dots, N^{(n)})$$

and let each diagonal element be written in terms of its prime factors:

$$N^{(i)} = \prod_{j=1}^m p_j^{\nu(i,j)},$$

where m is the total number of distinct prime factors present in the $N^{(i)}$.

The CRT may be used to turn each 1D transform along dimension i ($i = 1, \dots, n$) into a multidimensional transform with a separate ‘pseudo-dimension’ for each distinct prime factor of $N^{(i)}$; the number μ_i of these pseudo-dimensions is equal to the cardinality of the set:

$$\{j \in \{1, \dots, m\} | \nu(i,j) > 0 \text{ for some } i\}.$$

The full n -dimensional transform thus becomes μ -dimensional, with $\mu = \sum_{i=1}^n \mu_i$.

We may now permute the μ pseudo-dimensions so as to bring into contiguous position those corresponding to the same prime factor p_j ; the m resulting groups of pseudo-dimensions are said to define ‘ p -primary’ blocks. The initial transform is now written as a tensor product of m p -primary transforms, where transform j is on

$$p_j^{\nu(1,j)} \times p_j^{\nu(2,j)} \times \dots \times p_j^{\nu(n,j)}$$

points [by convention, dimension i is not transformed if $\nu(i,j) = 0$]. These p -primary transforms may be computed, for instance, by multidimensional Cooley–Tukey factorization (Section 1.3.3.3.1), which is faster than the straightforward row–column method. The final results may then be obtained by reversing all the permutations used.

The extra gain with respect to the multidimensional Cooley–Tukey method is that *there are no twiddle factors between p -primary pieces corresponding to different primes p .*

The case where \mathbf{N} is not diagonal has been examined by Guessoum & Mersereau (1986).

1.3.3.3.2.3. Nesting of Winograd small FFTs

Suppose that the CRT has been used as above to map an n -dimensional DFT to a μ -dimensional DFT. For each $\kappa = 1, \dots, \mu$ [κ runs over those pairs (i, j) such that $\nu(i, j) > 0$], the Rader/Winograd procedure may be applied to put the matrix of the κ th 1D DFT in the **CBA** normal form of a Winograd small FFT. The full DFT matrix may then be written, up to permutation of data and results, as

$$\bigotimes_{\kappa=1}^{\mu} (\mathbf{C}_{\kappa} \mathbf{B}_{\kappa} \mathbf{A}_{\kappa}).$$

A well known property of the tensor product of matrices allows this to be rewritten as

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

$$\left(\bigotimes_{\gamma=1}^{\mu} \mathbf{C}_{\gamma} \right) \times \left(\bigotimes_{\beta=1}^{\mu} \mathbf{B}_{\beta} \right) \times \left(\bigotimes_{\alpha=1}^{\mu} \mathbf{A}_{\alpha} \right)$$

and thus to form a matrix in which the *combined* pre-addition, multiplication and post-addition matrices have been *precomputed*. This procedure, called *nesting*, can be shown to afford a reduction of the arithmetic operation count compared to the row–column method (Morris, 1978).

Clearly, the nesting rearrangement need not be applied to all μ dimensions, but can be restricted to any desired subset of them.

1.3.3.3.2.4. The Nussbaumer–Quandalle algorithm

Nussbaumer’s approach views the DFT as the evaluation of certain polynomials constructed from the data (as in Section 1.3.3.2.4). For instance, putting $\omega = e(1/N)$, the 1D N -point DFT

$$X^*(k^*) = \sum_{k=0}^{N-1} X(k) \omega^{k^*k}$$

may be written

$$X^*(k^*) = Q(\omega^{k^*}),$$

where the polynomial Q is defined by

$$Q(z) = \sum_{k=0}^{N-1} X(k) z^k.$$

Let us consider (Nussbaumer & Quandalle, 1979) a 2D transform of size $N \times N$:

$$X^*(k_1^*, k_2^*) = \sum_{k_1=0}^{N-1} \sum_{k_2=0}^{N-1} X(k_1, k_2) \omega^{k_1^*k_1 + k_2^*k_2}.$$

By introduction of the polynomials

$$\begin{aligned} Q_{k_2}(z) &= \sum_{k_1} X(k_1, k_2) z^{k_1} \\ R_{k_2^*}(z) &= \sum_{k_2} \omega^{k_2^*k_2} Q_{k_2}(z), \end{aligned}$$

this may be rewritten:

$$X^*(k_1^*, k_2^*) = R_{k_2^*}(\omega^{k_1^*}) = \sum_{k_2} \omega^{k_2^*k_2} Q_{k_2}(\omega^{k_1^*}).$$

Let us now suppose that k_1^* is coprime to N . Then k_1^* has a unique inverse modulo N (denoted by $1/k_1^*$), so that multiplication by k_1^* simply permutes the elements of $\mathbb{Z}/N\mathbb{Z}$ and hence

$$\sum_{k_2=0}^{N-1} f(k_2) = \sum_{k_2=0}^{N-1} f(k_1^*k_2)$$

for any function f over $\mathbb{Z}/N\mathbb{Z}$. We may thus write:

$$\begin{aligned} X^*(k_1^*, k_2^*) &= \sum_{k_2} \omega^{k_1^*k_2^*k_2} Q_{k_1^*k_2}(\omega^{k_1^*}) \\ &= S_{k_1^*k_2}(\omega^{k_1^*}) \end{aligned}$$

where

$$S_{k^*}(z) = \sum_{k_2} z^{k^*k_2} Q_{k_2}(z).$$

Since only the value of polynomial $S_{k^*}(z)$ at $z = \omega^{k_1^*}$ is involved in the result, the computation of S_{k^*} may be carried out modulo the unique cyclotomic polynomial $P(z)$ such that $P(\omega^{k_1^*}) = 0$. Thus, if we define:

$$T_{k^*}(z) = \sum_{k_2} z^{k^*k_2} Q_{k_2}(z) \bmod P(z)$$

we may write:

$$X^*(k_1^*, k_2^*) = T_{k_1^*k_2^*}(\omega^{k_1^*})$$

or equivalently

$$X^*\left(k_1^*, \frac{k_2^*}{k_1^*}\right) = T_{k_2^*}(\omega^{k_1^*}).$$

For N an odd prime p , all non-zero values of k_1^* are coprime with p so that the $p \times p$ -point DFT may be calculated as follows:

(1) form the polynomials

$$T_{k_2^*}(z) = \sum_{k_1} \sum_{k_2} X(k_1, k_2) z^{k_1 + k_2^*k_2} \bmod P(z)$$

for $k_2^* = 0, \dots, p-1$;

(2) evaluate $T_{k_2^*}(\omega^{k_1^*})$ for $k_1^* = 0, \dots, p-1$;

(3) put $X^*(k_1^*, k_2^*/k_1^*) = T_{k_2^*}(\omega^{k_1^*})$;

(4) calculate the terms for $k_1^* = 0$ separately by

$$X^*(0, k_2^*) = \sum_{k_2} \left[\sum_{k_1} X(k_1, k_2) \right] \omega^{k_2^*k_2}.$$

Step (1) is a set of p ‘polynomial transforms’ involving no multiplications; step (2) consists of p DFTs on p points each since if

$$T_{k_2^*}(z) = \sum_{k_1} Y_{k_2^*}(k_1) z^{k_1}$$

then

$$T_{k_2^*}(\omega^{k_1^*}) = \sum_{k_1} Y_{k_2^*}(k_1) \omega^{k_1^*k_1} = Y_{k_2^*}^*(k_1^*);$$

step (3) is a permutation; and step (4) is a p -point DFT. Thus the 2D DFT on $p \times p$ points, which takes $2p$ p -point DFTs by the row–column method, involves only $(p+1)$ p -point DFTs; the other DFTs have been replaced by polynomial transforms involving only additions.

This procedure can be extended to n dimensions, and reduces the number of 1D p -point DFTs from np^{n-1} for the row–column method to $(p^n - 1)/(p - 1)$, at the cost of introducing extra additions in the polynomial transforms.

A similar algorithm has been formulated by Auslander *et al.* (1983) in terms of Galois theory.

1.3.3.3.3. Global algorithm design

1.3.3.3.3.1. From local pieces to global algorithms

The mathematical analysis of the structure of DFT computations has brought to light a broad variety of possibilities for reducing or reshaping their arithmetic complexity. All of them are ‘analytic’ in that they break down large transforms into a succession of smaller ones.

These results may now be considered from the converse ‘synthetic’ viewpoint as providing a list of procedures for assembling them:

(i) the building blocks are one-dimensional p -point algorithms for p a small prime;

(ii) the low-level connectors are the multiplicative reindexing methods of Rader and Winograd, or the polynomial transform reindexing method of Nussbaumer and Quandalle, which allow the construction of efficient algorithms for larger primes p , for prime powers p^ν , and for p -primary pieces of shape $p^\nu \times \dots \times p^\nu$;

(iii) the high-level connectors are the additive reindexing scheme of Cooley–Tukey, the Chinese remainder theorem reindexing, and the tensor product construction;

(iv) nesting may be viewed as the ‘glue’ which seals all elements.

1. GENERAL RELATIONSHIPS AND TECHNIQUES

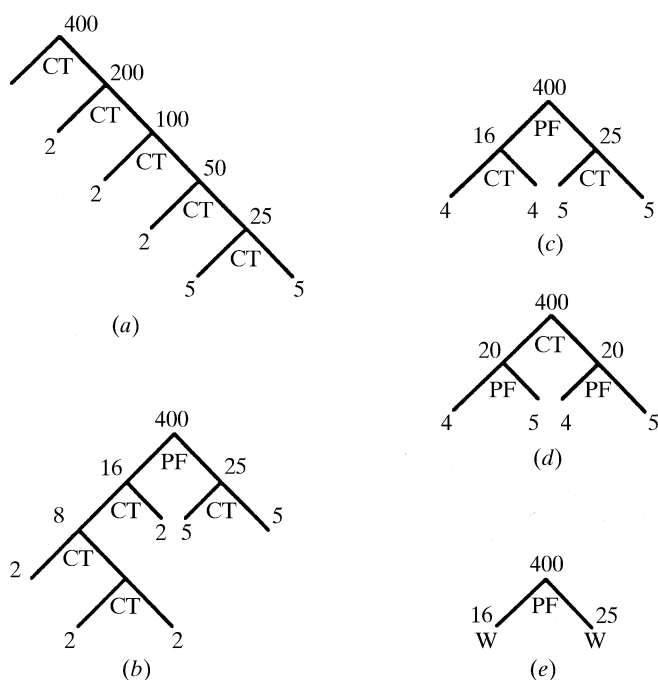


Fig. 1.3.3.1. A few global algorithms for computing a 400-point DFT. CT: Cooley–Tukey factorization. PF: prime factor (or Good) factorization. W: Winograd algorithm.

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).