A New Prime Edge Length Crystallographic FFT

Jaime Seguel[1], Dorothy Bollman[2], and Edusmildo Orozco[3]

[1] Electrical and Computer Engineering Department, UPRM
Mayagüez, Puerto Rico
jaime.seguel@ece.uprm.edu
[2] Department of Mathematics, UPRM
Mayagüez, Puerto Rico
bollman@cs.uprm.edu
[3] Doctoral Program in CISE, UPRM
Mayagüez, Puerto Rico
eorozco@cs.uprm.edu

Abstract. A new method for computing the discrete Fourier transform (DFT) of data endowed with linear symmetries is presented. The method minimizes operations and memory space requirements by eliminating redundant data and computations induced by the symmetry on the DFT equations. The arithmetic complexity of the new method is always lower, and in some cases significantly lower than that of its predecesor. A parallel version of the new method is also discussed. Symmetry-aware DFTs are crucial in the computer determination of the atomic structure of crystals from x-ray diffraction intensity patterns.

1 Preliminares

Let N be a positive integer, Z/N the set of integers modulo N, $Z^d/N = Z/N \times \cdots \times Z/N$, the cartesian product of d copies of Z/N, and f a real- or complex-valued mapping defined in Z^d/N. The *d-dimensional discrete Fourier transform* (DFT) of *edge length* N is defined by

$$\hat{f}(\boldsymbol{k}) = \frac{1}{\sqrt{N}} \sum_{\boldsymbol{l} \in Z^d/N} f(\boldsymbol{l}) w_N^{\boldsymbol{k}\cdot\boldsymbol{l}}, \quad \boldsymbol{k} \in Z^d/N; \tag{1}$$

where $w_N = \exp(-2\pi i/N)$, $\cdot$ denotes the dot product, and $i = \sqrt{-1}$. A *fast Fourier transform* [3] (FFT) for $d = 1$ computes (1) in $O(N \log N)$ operations. For $d \geq 2$, the usual FFT method consists of applying N^{d-1} one-dimensional FFTs along each of the d dimensions. This yields $O(N^d \log N)$ operations. Although this complexity bound cannot be improved for general DFT computations, some attempts to reduce the actual operation count and memory space requirements have been made for problems whose data is endowed with redundancies, such as x-ray crystal diffraction intensity data. In this article we review

P.M.A. Sloot et al. (Eds.): ICCS 2002, LNCS 2330, pp. 548–557, 2002.

a method for computing prime edge length DFTs of crystallographic data introduced by Auslander and Shenefelt [1],[2] and propose a new method with lower arithmetic complexity.

We first establish some notation. An expression of the form $[V_k]$ denotes a column vector. A bracketed expression with an ordered pair as a subscript denotes a matrix, and its row and and the column indices, respectively. If A is a set and g a mapping defined in A, its image set is denoted $g(A) = \{g(a) : a \in A\}$. By a *partition* of a set A we understand a collection $\{A_1, ..., A_m\}$ of subsets of A such that for each $j \neq k$, $A_j \cap A_k = \emptyset$, and $A = A_1 \cup \cdots A_m$. An *equivalence relation* on A is a relation $a \approx b$ defined for pairs in $A \times A$ which satisfies reflexivity, symmetry, and transitivity. An equivalence relation *induces* the partition of A consisting of the sets $O(a) = \{b \in A : a \approx b\}$. A subset of A formed by selecting one and only one element from each set in a partition of A is called *fundamental set*. The number of elements in a set A is denoted $|A|$.

A matrix $[W(k,l)]_{(k,l)}$, $0 \leq k < N_1$, $0 \leq l < N_2$ is *Hankel* if $W(k,l) = W(k',l')$ whenever $k + l = k' + l'$. An $N \times N$ matrix $[H(k,l)]_{(k,l)}$ is *Hankel-circulant* if $H(k,l) = H(k',l')$ whenever $k + l = k' + l'$ modulo N. Since a Hankel-circulant matrix is completely determined by its first row, we write $H = hc$(first row of H). A Hankel matrix W is said to be *embedded* in a Hankel-circulant H if W is the upper leftmost corner of H. Given an $N_1 \times N_2$ Hankel matrix W there exists at least one $N \times N$ Hankel-circulant into which W can be embedded, for each $N \geq N_1 + N_2 - 1$. For $N = N_1 + N_2 - 1$, this Hankel-circulant is

$$hc(W(0,0), \cdots, W(0, N_2 - 1), W(1, N_2 - 1), \cdots, W(N_1 - 1, N_2 - 1)). \quad (2)$$

For $N > N_1 + N_2 - 1$, the Hankel-circulant is obtained just by padding the argument of hc with $N - (N_1 + N_2 - 1)$ zeros to the right. An N-point *cyclic convolution* is a product of the form $\boldsymbol{U} = H\boldsymbol{V}$, where H is an $N \times N$ Hankel-circulant. We represent the N-point one-dimensional DFT by its complex matrix

$$F_N = \left[w_N^{kl}\right]_{(k,l)}, \quad 0 \leq k, l \leq N - 1. \quad (3)$$

It has been shown that for a Hankel-circulant H

$$\Delta(H) = F_N H F_N \quad (4)$$

is a diagonal matrix whose main diagonal is the DFT of the first row of H. Thus, cyclic convolutions can be computed in $O(N \log N)$ operations through

$$\boldsymbol{U} = F_N^{-1} \Delta(H) F_N^{-1} \boldsymbol{V}. \quad (5)$$

Equation (5) is often referred as *fast cyclic convolution* algorithm.

2 Symmetries and Crystallographic FFTs

Crystalline structures are determined at atomic level by computing several three-dimensional DFTs of their energy spectrum sampled every few angstroms. Since

crystal structures consist of repeating symmetric unit cells, their spectral data is highly redundant. A fast Fourier transform (FFT) method that uses these redundancies to lower the arithmetic count and memory space requirements of a three-dimensional FFT is called a *crystallographic FFT*. In this section we review some basic properties of matrix representations of crystal symmetries, use them to eliminate redundancies from the DFT equations, and outline the Auslander-Shenefelt crystallographic FFT.

We assume throughout that P is prime. A square matrix is a *matrix over* Z/P if all its entries are elements of the set Z/P of integers modulo P. The product of two such matrices modulo P is a matrix over Z/P. A matrix over Z/P is nonsingular if and only if its determinant is not zero modulo P. For a matrix M over Z/P and a nonegative integer j, M^j denotes the product of M by itself modulo P, j times. In particular, $M^0 = I$, where I is the identity matrix of appropriate size.

A real- or complex-valued mapping f defined on the set Z^d/P of d-dimensional vectors over Z/P is said to be S-symmetric if there exists a $d \times d$ nonsingular matrix S over Z/P such that

$$f(\boldsymbol{l}) = f(S\boldsymbol{l}) \quad \text{for all } \boldsymbol{l} \in Z^d/P. \tag{6}$$

For example, the mapping f defined in $Z^2/5$ by the two-dimensional data array $[f(k,l)]_{(k,l)}$

$$f(Z^2/5) = \begin{bmatrix} 2.9 & 2.3 & 1.5 & 1.5 & 2.3 \\ 1.2 & 6.0 & 4.3 & 4.6 & 2.8 \\ 1.4 & 3.3 & 5.1 & 4.2 & 1.7 \\ 1.4 & 1.7 & 4.2 & 5.1 & 3.3 \\ 1.2 & 2.8 & 4.6 & 4.3 & 6.0 \end{bmatrix} \tag{7}$$

is S-symmetric where

$$S = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \quad \text{modulo } 5. \tag{8}$$

It is clear from (6) that $f(S^j\boldsymbol{a}) = f(\boldsymbol{a})$ for all j. Thus, the data redundancies of a S-symmetric mapping constitute subsets of Z^d/P of the form

$$O_S(\boldsymbol{a}) = \{S^j\boldsymbol{a} \text{ modulo } P : \ j \text{ integer } \}. \tag{9}$$

Such a subset is called an S *orbit* of $\boldsymbol{a}$. It is readily verified that the relation defined by $\boldsymbol{a} \approx_S \boldsymbol{b}$ if and only if $\boldsymbol{b} = S^j\boldsymbol{a}$ for some integer j is an equivalence relation over Z^d/P. Thus, the set of all S-orbits is a partition of Z^d/P. The number of elements $|O_S(\boldsymbol{a})|$ is called the *orbit length*. A subset of the form $O_S^{(l_{\boldsymbol{a}})}(\boldsymbol{a}) = \{S^j\boldsymbol{a} : 0 \le j \le l_{\boldsymbol{a}} - 1\}$ where $l_{\boldsymbol{a}} \le |O_S(\boldsymbol{a})|$ is said to be an S *segment*. An S-symmetric function is completely determined by its values on a fundamental set $\mathcal{F}_S$ induced by $\approx_S$. For example, the S-orbits and their images under f for (7) are

$$O_S(0,0) = \{(0,0)\}\ ,\ f(O_S(0,0)) = \{2.9\} \tag{10}$$

$$O_S(0,1) = \{(0,1),\ (0,4)\}\ ,\ f(O_S(0,1)) = \{2.3\} \tag{11}$$
$$O_S(0,2) = \{(0,2),\ (0,3)\}\ ,\ f(O_S(0,2)) = \{1.5\} \tag{12}$$
$$O_S(1,0) = \{(1,0),\ (4,0)\}\ ,\ f(O_S(1,0)) = \{1.2\} \tag{13}$$
$$O_S(1,1) = \{(1,1),\ (4,4)\}\ ,\ f(O_S(1,1)) = \{6.0\} \tag{14}$$
$$O_S(1,2) = \{(1,2),\ (4,3)\}\ ,\ f(O_S(1,2)) = \{4.3\} \tag{15}$$
$$O_S(1,3) = \{(1,3),\ (4,2)\}\ ,\ f(O_S(1,3)) = \{4.6\} \tag{16}$$
$$O_S(1,4) = \{(1,4),\ (4,1)\}\ ,\ f(O_S(1,4)) = \{2.8\} \tag{17}$$
$$O_S(2,0) = \{(2,0),\ (3,0)\}\ ,\ f(O_S(2,0)) = \{1.4\} \tag{18}$$
$$O_S(2,1) = \{(2,1),\ (3,4)\}\ ,\ f(O_S(2,1)) = \{3.3\} \tag{19}$$
$$O_S(2,2) = \{(2,2),\ (3,3)\}\ ,\ f(O_S(2,2)) = \{5.1\} \tag{20}$$
$$O_S(2,3) = \{(2,3),\ (3,2)\}\ ,\ f(O_S(2,3)) = \{4.2\} \tag{21}$$
$$O_S(2,4) = \{(2,4),\ (3,1)\}\ ,\ f(O_S(2,4)) = \{1.7\}. \tag{22}$$

Thus, we may choose $\mathcal{F}_S = \{(0,0),\ (0,1),\ (0,2),\ (1,0),\ (1,1),\ (1,2),\ (1,3),\ (1,4),\ (2,0),\ (2,1),\ (2,2),\ (2,3),\ (2,4)\}$.

It is easy to show that if f is S-symmetric, then $\hat{f}$ is S_*-symmetric, where S_* is the transpose of the inverse of S. Therefore, $\hat{f}$ is also completely determined by its values in a fundamental set $\mathcal{F}_{S_*}$ induced by $\approx_{S_*}$. We call $f\,(\mathcal{F}_S)$ and $\hat{f}\,(\mathcal{F}_{S_*})$ *fundamental input data* and *fundamental output data*, respectively. In example (7) the fundamental input data set is

$$f(\mathcal{F}_S) = \{2.9,\ 2.3,\ 1.5,\ 1.2,\ 6.0,\ 4.3,\ 4.6,\ 2.8,\ 1.4,\ 3.3,\ 5.1,\ 4.2,\ 1.7\} \tag{23}$$

Reducing the original input data to a fundamental input data set requires a modification in the DFT equations. Since $f(\mathcal{F}_S(\boldsymbol{a}))$ contains just one $f(\boldsymbol{a})$ from each $f(O_S(\boldsymbol{a}))$ orbit and each $\boldsymbol{k} \in Z^d/P$, the input datum $f(\boldsymbol{a})$ is factored out of the terms of $\hat{f}(\boldsymbol{k})$ indexed by $O_S(\boldsymbol{a})$ in equation (1). That is,

$$\sum_{\boldsymbol{l}\in O_S(\boldsymbol{a})} f(\boldsymbol{l}) w_P^{\boldsymbol{k}\cdot\boldsymbol{l}} = f(\boldsymbol{a}) \left(\sum_{\boldsymbol{l}\in O_S(\boldsymbol{a})} w_P^{\boldsymbol{k}\cdot\boldsymbol{l}} \right) = f(\boldsymbol{a}) K_P(\boldsymbol{k}, \boldsymbol{a}). \tag{24}$$

$K_P(\boldsymbol{k}, \boldsymbol{a}) = \sum_{\boldsymbol{l}\in O_S(\boldsymbol{a})} w_P^{\boldsymbol{k}\cdot\boldsymbol{l}}$ is called a *symmetrized DFT kernel*. The linear transformation

$$\hat{f}(\boldsymbol{k}) = \sum_{\boldsymbol{a}\in\mathcal{F}_S} K_P(\boldsymbol{k}, \boldsymbol{a}) f(\boldsymbol{a}), \quad \boldsymbol{k} \in \mathcal{F}_{S_*} \tag{25}$$

where the output has also been restricted to a fundamental output set, is called a *symmetrized DFT*. Equation (25) involves

$$\sum_{\boldsymbol{k}\in\mathcal{F}_{S_*}} \sum_{l\in\mathcal{F}_S} |O_S(\boldsymbol{l})| = \sum_{\boldsymbol{k}\in\mathcal{F}_{S_*}} P^d \le P^{2d} \tag{26}$$

arithmetic operations. Thus, in general, the mere reduction of redundant data does not yield a superior FFT method.

Auslander and Shenefelt [1] proposed computing (25) through *fast cyclic convolutions*. Crucial to this aim is the fact the set of all non-null elements in Z/P, denoted Z/P^*, is a *cyclic group* under multiplication. In particular, there is an element $g \in Z/P^*$ called *generator*, such that for each $a \in Z/P^*$, an integer j can always be found for which $g^j = a$ modulo P. The action of g on Z^d/P produces *g-orbits* of the form $O_g(\boldsymbol{a}) = \{g^j \boldsymbol{a} : 0 \leq j \leq P-2\}$. Furthermore, the action of g produces partitions of $\mathcal{F}_S$ and $\mathcal{F}_{S_*}$ formed by g-segments. Let $\mathcal{F}_{gS}$ and $\mathcal{F}_{gS_*}$ be fundamental sets for these partitions. For a pair $(\boldsymbol{a}, \boldsymbol{b}) \in \mathcal{F}_{gS_*} \times \mathcal{F}_{gS}$ let $l_{\boldsymbol{a}}$ and $l_{\boldsymbol{b}}$ be the lengths of the g-segments in $\mathcal{F}_{gS_*}$ and $\mathcal{F}_{gS}$, respectively. Then,

$$W_{(\boldsymbol{a},\boldsymbol{b})}(k,l) = K_P(g^k \boldsymbol{a}, g^l \boldsymbol{b}), \quad 0 \leq k \leq l_{\boldsymbol{a}} - 1.\ 0 \leq l_{\boldsymbol{b}} - 1, \tag{27}$$

is a block in (25). Since clearly for $k + l = k' + l'$, $W_{(\boldsymbol{a},\boldsymbol{b})}(k,l) = W_{(\boldsymbol{a},\boldsymbol{b})}(k',l')$, each block $W_{(\boldsymbol{a},\boldsymbol{b})}$ is Hankel. Each one of these blocks can be computed with the fast cyclic convolution algorithm by embeddings in Hankel-circulants.

Since the length of a g-segment is at most $P-1$, the Auslander-Shenefelt method involves $O(P^{d-1})$ cyclic convolutions unless the average length of the S-orbits is also a power of P. The latter condition is not satisfied by most of the symmetries of practical importance.

Our strategy towards a more efficient crystallographic FFT is to minimize the amount of cyclic convolutions, even at the cost of increasing their sizes. This is achieved by using M-segments instead of g-segments, where M is a nonsingular matrix over Z/P. This modification and its impact in the computation of (25) are described in the next section.

3 A Framework for Designing Crystallographic FFTs

Let S and M be nonsingular matrices over Z/P. Let k, l, k', and l' be nonegative integers such that $k + l = k' + l'$. Then for any pair $(\boldsymbol{a}, \boldsymbol{b}) \in \mathcal{F}_{S_*} \times \mathcal{F}_{\mathcal{S}}$,

$$K_P(M^k \boldsymbol{a}, (M^t)^l \boldsymbol{b}) = K_P(M^{k+l} \boldsymbol{a}, \boldsymbol{b}) \tag{28}$$

$$= K_P(M^{k'+l'} \boldsymbol{a}, \boldsymbol{b}) \tag{29}$$

$$= K_P(M^{k'} \boldsymbol{a}, (M^t)^{l'} \boldsymbol{b}). \tag{30}$$

It follows that $W_{(\boldsymbol{a},\boldsymbol{b})}(k,l) = K_P(M^k \boldsymbol{a}, (M^t)^l \boldsymbol{b})$ is Hankel. These matrices will represent all the computations in (25) if and only if $\mathcal{F}_{\mathcal{S}_*}$ and $\mathcal{F}_S$ can be partitioned into M- and M^t-segments, respectively. It can easily be shown that this is the case if $MS_* = S_*M$. Let $\mathcal{F}_{MS_*}$ and $\mathcal{F}_{M^tS}$ be fundamental sets for these partitions. Then (25) can be rewritten as

$$\left[\hat{f}\left(\mathcal{F}_{S_*}\right)\right] = \left[\left[W_{(\boldsymbol{a},\boldsymbol{b})}(k,l)\right]_{(k,l)}\right]_{(\boldsymbol{a},\boldsymbol{b})} \left[f\left(\mathcal{F}_S\right)\right], \tag{31}$$

where the nested brackets denote a block matrix, $(\boldsymbol{a}, \boldsymbol{b}) \in \mathcal{F}_{MS_*} \times \mathcal{F}_{M^tS}$, and $0 \leq k < l_{\boldsymbol{a}}$, $0 \leq l < l_{\boldsymbol{b}}$.

Before embedding each $W_{(\boldsymbol{a},\boldsymbol{b})}$ in a Hankel-circulant we make two considerations. First, no blocks corresponding to pairs $(\mathbf{0},\boldsymbol{b})$ and $(\boldsymbol{a},\mathbf{0})$ will be embedded. These are $1 \times l_{\boldsymbol{b}}$ and $l_{\boldsymbol{a}} \times 1$ matrices and therefore, embedding them in Hankel-circulants is not practical. We call these matrices *border blocks* and compute with them separately. We call the remaining set of blocks the *core symmetric DFT* computations. Second, although other choices are available, we propose a common size $N \times N$ for the Hankel-circulants into which the $W_{(\boldsymbol{a},\boldsymbol{b})}$ blocks will be embedded. This N is a positive integer greater than or equal to the sum of the length of the largest input segment plus the length of the largest output segment minus 1. It turns out that the maximum length of the input segments is always equal to the maximum length of the output segments. We denote this common maximum length by L. On the other hand, since the choice of N is guided by the efficiency of the N-point FFT, it is reasonable to assume that $2L-1 \leq N \leq 2^{\lceil \log_2(2L-1) \rceil}$, where $\lceil x \rceil$ is the smallest integer that is greater than or equal to x.

Let $\mathcal{F}^*_{M^tS}$ and $\mathcal{F}^*_{MS_*}$ be the original fundamental sets without $\mathbf{0}$. Let $H_{(\boldsymbol{a},\boldsymbol{b})}$ be the $N \times N$ Hankel-circulant in which $W_{(\boldsymbol{a},\boldsymbol{b})}$ is embedded. Then, the core symmetric DFT computations are represented as $\boldsymbol{U} = H\boldsymbol{V}$ where

$$H = \Big[\big[H_{(\boldsymbol{a},\boldsymbol{b})}(k,l) \big]_{(k,l)} \Big]_{(\boldsymbol{a},\boldsymbol{b})}, \tag{32}$$

$(\boldsymbol{a},\boldsymbol{b}) \in {\mathcal{F}^*}_{MS_*} \times \mathcal{F}^*_{M^tS}$, and $0 \leq k,l < N$. The input vector $\boldsymbol{V}$ is composed of N-point vector segments $\boldsymbol{V}_{\boldsymbol{b}}$, each consisting of the $l_{\boldsymbol{b}}$ values of $f(O^{(l_{\boldsymbol{b}})}_{M^tS}(\boldsymbol{b}))$ followed by $N - l_{\boldsymbol{b}}$ zeros. Vector $\boldsymbol{U}$, in turn, is composed by the N-point segments $\boldsymbol{U}_{\boldsymbol{a}}$, $\boldsymbol{a} \in \mathcal{F}^*_{MS_*}$, that result from

$$\boldsymbol{U}_{\boldsymbol{a}} = \sum_{\boldsymbol{b} \in \mathcal{F}^*_{M^tS}} H_{(\boldsymbol{a},\boldsymbol{b})} \boldsymbol{V}_{\boldsymbol{b}} \tag{33}$$

$$= \sum_{\boldsymbol{b} \in \mathcal{F}^*_{M^tS}} F_N^{-1} \Delta \left(H_{(\boldsymbol{a},\boldsymbol{b})} \right) F_M^{-1} \boldsymbol{V}_{\boldsymbol{b}} \tag{34}$$

$$= F_N^{-1} \left(\sum_{\boldsymbol{b} \in \mathcal{F}^*_{M^tS}} \Delta \left(H_{(\boldsymbol{a},\boldsymbol{b})} \right) F_M^{-1} \boldsymbol{V}_{\boldsymbol{b}} \right). \tag{35}$$

Following is a prime edge length symmetric FFT framework based on (35):

Core computations:
Step 1. For each $\boldsymbol{b} \in \mathcal{F}^*_{M^tS}$ compute $\boldsymbol{Y}_{\boldsymbol{b}} = F_N^{-1} \boldsymbol{V}_{\boldsymbol{b}}$.
Step 2. For each pair $(\boldsymbol{a},\boldsymbol{b}) \in \mathcal{F}^*_{MS_*} \times {\mathcal{F}^*}_{M^tS}$ compute the Hadamard or component-wise product $\boldsymbol{Z}_{(\boldsymbol{a},\boldsymbol{b})} = \Delta \left(H_{(\boldsymbol{a},\boldsymbol{b})} \right) \boldsymbol{Y}_{\boldsymbol{b}}$.
Step 3. For each $\boldsymbol{a} \in \mathcal{F}^*_{MS_*}$ compute $\boldsymbol{X}_{\boldsymbol{a}} = \sum_{\boldsymbol{b}} \boldsymbol{Z}_{(\boldsymbol{a},\boldsymbol{b})}$.
Step 4. For each $\boldsymbol{a} \in {\mathcal{F}^*}_{MS_*}$ compute $\boldsymbol{U}_{\boldsymbol{a}} = F_N^{-1} \boldsymbol{X}_{\boldsymbol{a}}$.

Border computations:
Step 5. $\hat{f}(\mathbf{0}) = \frac{1}{\sqrt{P}} \sum_{\boldsymbol{b} \in \mathcal{F}_S} f(\boldsymbol{b}) |O_S(\boldsymbol{b})|$

Step 6. For each $\boldsymbol{a} \in \mathcal{F}^*{}_{MS_*}$ compute $\hat{f}(O_M^{l_{\boldsymbol{a}}}(\boldsymbol{a})) = \frac{1}{\sqrt{P}}\,[f(\mathbf{0})]_{l_{\boldsymbol{a}}} + \boldsymbol{U}^*_{\boldsymbol{a}}$. Here $\boldsymbol{U}^*_{\boldsymbol{a}}$ is the column vector formed by the $l_{\boldsymbol{a}}$ first entries of the vector $\boldsymbol{U}_{\boldsymbol{a}}$ computed in step 4 .

All parameters required for the actual implementation of the method are determined in a separate precomputation phase. This phase includes the computation of the M^t- and M-orbit segments, and the diagonals $\Delta(H_{(\boldsymbol{a},\boldsymbol{b})})$.

Since the border computations involve $O(|\mathcal{F}_S|)$ sums and products, their contribution to the arithmetic complexity of the method is marginal. As for the core computations, let $\Lambda = |\mathcal{F}^*_{M^tS}|$. Then, using N-point FFTs, steps 1 and 4 involve $O(2\Lambda N \log\, N)$ operations. Step 2 involves $\Lambda^2 N$ complex multiplications and step 3, $\Lambda N(\Lambda - 1)$ complex additions. Therefore, the complexity of the core computation phase is modeled by the two-parameter mapping

$$c(\Lambda, N) = 2\Lambda N\,(\kappa \log\, N + \Lambda - 1)\,, \tag{36}$$

where κ is a positive constant. It follows that the order of the arithmetic complexity of a crystallographic FFT derived from the general framework is $O(\Lambda N\ \log\ N)$ if

$$\rho_1 = \frac{(\Lambda - 1)}{\log_2\,(2L-1)} \leq 1, \tag{37}$$

and $O(\Lambda^2 N)$ if

$$\rho_2 = \frac{(\Lambda - 1)}{\lceil \log_2\,(2L-1)\rceil} > 1. \tag{38}$$

The best case, $\Lambda = 1$, ensures an $O(N\ \log\, N)$ symmetric FFT. If ρ_2 is greater than 1, but close to 1 and Λ is relatively small, it is still possible to get a competitive crystallographic FFT. However, the arithmetic count of the crystallographic FFT increases rapidly with Λ, as shown in table 1. The parallel version of the method that is outlined in section 5 is intended to reduce the execution time for cases in which $\Lambda > 1$.

The Auslander-Shenefelt method can be derived from the general framework by setting $M = gI_d$, where I_d is the $d \times d$ identity matrix. From a previous remark we know that the Auslander-Shenefelt algorithm is likely to produce large values for Λ for most symmetries of practical importance. One such example is the symmetry

$$S = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}. \tag{39}$$

In this case, the Auslander-Shenefelt crystallographic FFT gives $\Lambda = P^2 + P$ and $L = (P-1)/2$. Thus, $c(\Lambda, N)$ is always $O(\Lambda^2 N)$. Using the minimum $N = 2L-1$ we see that $\Lambda^2(2L-1) = (P^2+P)^2(P-2)$. Therefore, step 2 of the Auslander-Shenefelt FFT is $O(P^5)$ which is greater than $O(P^3\ \log\ P)$, the complexity order of the usual three-dimensional FFT.

4 Existence of $O(N \log N)$ Crystallographic FFTs

Some theoretical results concerning the existence of $O(N \log N)$ crystallographic FFTs are presented in this section. The lemmas, proofs of which can be found, for instance, in [4], are well-known results that have been used in coding theory, digital signal processing, and the theory of linear autonomous sequential machines, among others.

Lemma 1. For any polynomial $\phi(x)$ over Z/P, there is a positive integer k for which $\phi(x)$ divides $x^k - 1$.

The smallest such k is called the *period* of $\phi(x)$.

Lemma 2. For each irreducible polynomial $\phi(x)$ of degree n over Z/P, $\phi(x)$ divides $x^{P^n-1} - 1$.

A polynomial is *maximal* if its period is $P^n - 1$.

Lemma 3. The number of maximal polynomials of degree n over Z/P is $\varphi(P^n - 1)/n$, where φ denotes Euler's $\varphi-$function.

Similarly, for any nonsingular square matrix M over Z/P, there is a least positive integer k such that $M^k = I$, where I denotes an identity matrix of appropriate size. This k is called the *period* of M. A $d \times d$ M is *maximal* if its period is $P^d - 1$. Thus, the action of a maximal M on Z^d/P produces only two orbits: one consisting of $\{\mathbf{0}\}$, and one consisting of all non-null vectors in Z^d/P. The *characteristic polynomial* ϕ_M of a square matrix M is defined by the determinant $M - xI$ modulo P. The Cayley Hamilton Theorem states that every square matrix satisfies its own characteristic equation, i.e., $\phi_M(M) = 0$.

Lemma 4. A nonsingular matrix M over Z/P is maximal if and only if its characteristic polynomial $\phi_M(x)$ is irreducible and maximal.

Corollary 1. All M-orbits are of equal length if the characteristic polynomial of M is irreducible.

These results provide a first example of the existence of an $O(N \log N)$ crystallographic FFT for a symmetry of practical importance. In fact, for any P there will be a maximal matrix M over Z/P. Since the symmetry $S = S_*$ defined in (39) commutes with any 3×3 matrix, it commutes, in particular, with any maximal matrix M over Z^3/P. But since M is maximal, the partitions induced by M^t and M on $\mathcal{F}_S^*$ and $\mathcal{F}_{S_*}^*$, respectively, consist of just one segment. Therefore, since $\Lambda = 1$, the crystallographic FFT derived from the general framework is $O(N \log N)$.

Another interesting case is given by the next theorem.

Theorem 1. Let S be a matrix over Z/P with irreducible characteristic polynomial and let M be a maximal matrix that commutes with S. Then there is exactly one M-segment in $\mathcal{F}_{S_*}$ and its size is $m = (P^d - 1)/k$ where k is the size of the S_*-orbits.

Proof. From the previous corollary, all orbits induced by S_* are of equal length. Let x be any nonzero vector in Z^d/P. Then the sequence $x, Mx, M^2, \cdots, M^{P^d-2}x$ contains at most k S_*-equivalent elements. Suppose it contains fewer than m S_*-equivalent elements. Let a and b, $a < b$ be the least positive integers for which $M^a x$ and $M^b x$ are S_*-equivalent and so $M^a x = S^c M^b x$ for some c.

Hence ${S_*}^{-c}x = M^{b-a}x$ and so $M^{(b-a)k}M^{b-a}\cdots M^{b-a}x = ({S_*}^{-c})^k x = x$ where $(b-a)k < P^d - 1$, which contradicts that M is maximal.

5 Further Examples and Conclusions

As remarked before, some crystallographic symmetries will not yield $O(N \log N)$ crystallographic FFTs. For such symmetries, the parameter Λ is minimized by finding a matrix M satisfying $MS_* = S_*M$, and whose action produces the largest possible M-segments in $\mathcal{F}_{\mathcal{S}_*}$. Such M is called *optimal*. In this section we show a symmetry whose parameter $\Lambda > 1$ and compare the complexity of the crystallographic FFT built with the optimal M with that of the Auslander-Shenefelt algorithm. We also describe a natural parallel version of the method for symmetries whose parameter Λ is greater than 1.

Let us consider the symmetry represented by

$$S = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}. \tag{40}$$

Table 1 compares the parameters and the arithmetic complexity of the second step of the crystallographic FFT derived from the general framework using an optimal M with the Auslander-Shenefelt method. We have chosen $N = 2L-1$ for these comparisons. To our knowledge there is no existing method for computing

Table 1. Crystallographic FFT derived from an optimal segmentation of the fundamental sets versus Auslander-Shenefelt crystallographic FFT (ASCFFT) for symmetry (40)

Problem size	Optimal M	Λ	N	$\Lambda^2 N$	ASCFFT Λ	ASCFFT N	ASCFFT $\Lambda^2 E$
$23^3 = 12167$	$\begin{bmatrix} 2 & 0 & 1 \\ 0 & 5 & 0 \\ 22 & 0 & 2 \end{bmatrix}$	24	263	151,488	145	43	904,075
$31^3 = 29791$	$\begin{bmatrix} 4 & 0 & 1 \\ 0 & 3 & 0 \\ 30 & 0 & 4 \end{bmatrix}$	32	479	490,496	257	59	3,896,891
$43^3 = 79507$	$\begin{bmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 42 & 0 & 2 \end{bmatrix}$	44	923	1,786,928	485	83	19,523,675
$47^3 = 103823$	$\begin{bmatrix} 2 & 0 & 1 \\ 0 & 5 & 0 \\ 46 & 0 & 2 \end{bmatrix}$	48	1103	2,541,312	577	91	30,296,539

the optimal M directly. Our results have been produced computationally.

For $\Lambda > 1$, in a distributed memory system with Λ processors, the symmetric FFT can be performed in parallel, as follows:

Parallel core symmetric FFT

Parallel step 1. For each $\boldsymbol{b} \in \mathcal{F}^*_{M^t S}$ compute in parallel $\boldsymbol{Y_b} = F_N^{-1} V_{\boldsymbol{b}}$.
Parallel step 2. For each $\boldsymbol{b}) \in \mathcal{F}^*_{MS_*}$ compute in parallel $\boldsymbol{Z}_{(\boldsymbol{a},\boldsymbol{b})} = \Delta\left(H_{(\boldsymbol{a},\boldsymbol{b})}\right) \boldsymbol{Y_b}$, for all $\boldsymbol{a} \in \mathcal{F}^*{}_{M^t S}$.
Communication step. For each $\boldsymbol{a} \in \mathcal{F}^*_{MS_*}$, gather all vector segments $\boldsymbol{Z}_{(\boldsymbol{a},\boldsymbol{b})}$ in a single processor.
Step 3. For each $\boldsymbol{a} \in \mathcal{F}^*_{MS_*}$ compute in parallel $\boldsymbol{X_a} = \sum_{\boldsymbol{b}} \boldsymbol{Z}_{(\boldsymbol{a},\boldsymbol{b})}$.
Step 4. For each $\boldsymbol{a} \in \mathcal{F}^*{}_{MS_*}$ compute in parallel $\boldsymbol{U_a} = F_N^{-1} \boldsymbol{X_a}$.

Parallel border computations:

Step 5. $\hat{f}(\mathbf{0}) = \frac{1}{\sqrt{P}} \sum_{\boldsymbol{b} \in \mathcal{F}_S} f(\boldsymbol{b}) |O_S(\boldsymbol{b})|$
Step 6. For each $\boldsymbol{a} \in \mathcal{F}^*{}_{MS_*}$ compute in parallel $\hat{f}(O_M^{l_{\boldsymbol{a}}}(\boldsymbol{a}) = \frac{1}{\sqrt{P}} \left[f(\mathbf{0})\right]_{l_{\boldsymbol{a}}} + \boldsymbol{U}^*_{\boldsymbol{a}}$. Here $\boldsymbol{U}^*_{\boldsymbol{a}}$ is the column vector formed by the $l_{\boldsymbol{a}}$ first entries of the vector $\boldsymbol{U_a}$ computed in step 4 .

The parallel method reduces the time complexity of steps 1 and 4 to $O(N \log N)$ and that of steps 2 and 3 to $O(\Lambda N)$, to the cost of sending $\Lambda N(\Lambda - 1)$ complex number between processors. For large values of Λ, an adquate balance between communications and computations can be achieved by aggregating parallel computations. The symmetric FFT framework can be implemented as a meta-algorithm able to generate crystallographic FFTs that are tailored to a target computer system. The inputs of the meta-algorithm will be the symmetry S and the edge length P of the DFT. The output will be a crystallographic FFT computer program that optimally uses the hardware and the software of the target system, very much in the spirit of the fast Fourier transform method known as the FFTW [5]. Experimental work exploring the potential of this idea is currently underway.

Acknowledgements. This work was partially supported by NIH/NIGMS grant No. S06GM08103 and the NSF PRECISE project of the University of Puerto Rico at Mayagüez

References

1. L. Auslander, and M. Shenefelt, *Fourier transforms that respect crystallogaphic symmetries*, IBM J. Res. and Dev., 31, (1987), pp. 213-223.
2. M. An, J. W. Cooley, and R. Tolimeri, *Factorization methods for crystallographic Fourier transforms*, Advances in Appl. Math., 11 (1990), pp. 358-371.
3. J. Cooley, and J. Tukey, *An algorithm for the machine calculation of complex Fourier series*, Math. Comp., 19 (1965), pp. 297-301.
4. B. Elspas, *The Theory of Autonomous Linear Sequential Networks*, Linear Sequential Switching Circuits, ed. W. Kautz, Holden-Day inc., 1965, 21-61.
5. M. Frigo, S. G. Johnson *An adaptive software architecture for the FFT* ICASSP Conference Proceedings, 3 (1998), pp 1381-1384.