ACKNOWLEDGMENT

The author is grateful to Prof. J. Moore who suggested to look at the Baum–Weich and EM algorithms, and to V. Krishnamurthy for interesting conversations. The hospitality of the Systems Engineering Department, Research School of Physical Sciences, Australian National University is also acknowledged.

REFERENCES


Computing $m$ DFT’s Over GF($q$) with One DFT Over GF($q^m$)

Jonathan Hong and Martin Vetterli

Abstract—Over the field of complex numbers, it is well-known that if the input is real then it is possible to compute two real DFT’s with one complex DFT. We extend the result to finite fields and show how to compute $m$ DFT’s over GF($q$) with one DFT over GF($q^m$).

Index Terms—Finite fields, DFT, complexity.

I. INTRODUCTION

The discrete Fourier transform and its inverse are defined as

$$X_k = \sum_{n=0}^{N-1} x_n W_N^{nk}, \quad (1)$$

$$x_n = \sum_{k=0}^{N-1} X_k W_N^{-nk}, \quad (2)$$

Manuscript received February 5, 1991; revised March 1, 1992. This work was supported in part by the National Science Foundation under Grants CDR-84-21402 and MIP-90-14189.

The authors are with the Department of Electrical Engineering and Center for Telecommunications Research, Columbia University, Room 1342 S.W. Mudd Building, 500 W. 120th Street, New York, NY 10027–6699.

IEEE Log Number 9203039.

The definition is valid over any field $F$ provided that $W_N$, an element of order $N$, exists in $F$ and $N^{-1}$ exists in $F$.

For $F = C$, the field of complex numbers, it is well known that if the input $x_n$ is real then it is possible to compute two real DFT’s with one complex DFT [6], thereby showing that the complexity of a real DFT is approximately half that of a complex DFT. This reduction in complexity is due essentially to the fact that the $C$ is an extension of degree 2 over $R$, i.e., $[C:R] = 2$. In this correspondence, we will show that this result extends to the case where $F$ is a finite field. Specifically, if $F$ is an extension field of $K$ with $[F:K] = m$, then the complexity of the DFT of a $F$-sequence is approximately $m$ times that of a $K$-sequence. We will demonstrate this by showing how to compute the DFT of $m$ $K$-sequences with the DFT of 1 $F$-sequence.

The procedure is quite general and will be shown to apply to the real/complex case as well. As is customary, we will assume that the constants in the algorithm are precomputed and thus the cost of their computation excluded from the overall complexity of the algorithm.

II. CONJUGACY RELATIONS

We derive, in this section, the well-known conjugacy relationship that exists among the elements of the DFT of a subfield sequence. The derivation differs (conceptually) from the usual derivation in that it makes use of the automorphism group associated with the fields $F$ and $K$. While it is possible to present our algorithm without reference to the automorphism group; we choose to introduce it in order to keep the derivation as general as possible. The generality eliminates the need to rederive the algorithm for the real and complex case in Section IV.

Let $F$ and $K$ be finite fields such that $K \subset F$, then there are two ways we can view $F$ relative to $K$. First, with respect to the addition operation in $F$, $F$ is a $[F:K]$-dimensional vector space over $K$ (notation: $F_K$). Thus if $K = GF(q)$ and $[F:K] = m$, then $F = GF(q^m)$; from which it follows that if $\{\beta_0, \beta_1, \ldots, \beta_{m-1}\}$ is any basis of $F_K$ then any $x \in F$ can be written uniquely in the form

$$x = x_0 \beta_0 + x_1 \beta_1 + \cdots + x_{m-1} \beta_{m-1}$$

with the coefficients $x_i$ in $K$. Alternatively we can view $F$ as an extension field of $K$. Associated with this field extension is the group $\text{Aut}_K F$ which is the group of automorphisms of $F$ that leave $K$ fixed. It is well known that this group is cyclic of order $[F:K]$ and has as its generator $\psi: a \mapsto a^\alpha$, [1]–[3]. In other words, $\text{Aut}_K F = \{\psi^l\}_{l=0}^{m-1}$ where

$$\psi^l = \psi^l: a \mapsto a^{l}, \quad (3)$$

Consider

$$X_k = \sum_{n=0}^{N-1} x_n W_N^{nk},$$

with $x_n \in K$ for every $n$. Since $\psi^l$ is a homomorphism which fixes elements of $K$, we see that

$$\psi^l(X_k) = \psi^l \left( \sum_{n=0}^{N-1} x_n W_N^{nk} \right) = \sum_{n=0}^{N-1} \psi^l(x_n W_N^{nk})$$

$$= \sum_{n=0}^{N-1} \psi^l(x_n) \psi^l(W_N^{nk}) = \sum_{n=0}^{N-1} x_n \psi^l(W_N^{nk}),$$

$$\psi^l(x_n W_N^{nk}) = \psi^l(x_n) \psi^l(W_N^{nk}),$$

$$\psi^l(W_N^{nk}) = W_N^{lk},$$

$$\psi^l(x_n) = x_{n^l},$$

$$x_{n^l} W_N^{nk} = x_{n^l} W_N^{nk}.$$
therefore,
\[ X_k^q = \sum_{n=0}^{N-1} x_n W_n^q = X_k, \]  
Equation (4) is the familiar conjugacy relation [4], [5]. We call \( X_k^q \) a conjugate of \( X_k \) and \( \{ X_k^q \}_{q=0}^{m-1} \) the conjugacy class of \( X_k \).

Define
\[ X_i \sim X_j \quad \text{if } j = i^q, \quad \text{for some } l \in Z; \]
then \( \sim \) is an equivalence relation on \( \{ X_k \} \) with \( \{ X_k^q \}_{q=0}^{m-1} \) the equivalence class of \( X_k \). Equation (4) says that when \( x_n \in K \), then the conjugacy class of \( X_k \) is the same as the equivalence class of \( X_k \), i.e.,
\[ \{ \phi_0(X_k), \phi_1(X_k), \ldots, \phi_{m-1}(X_k) \} = \{ X_k^0, X_k^1, \ldots, X_k^{m-1} \} = \{ X_k, X_k^q, \ldots, X_k^{q(m-1)} \}. \]

III. COMPUTING \( m \) DFT'S OVER \( K \) WITH 1 DFT OVER \( F \)

Given \( m \) sequences in \( K \)
\[ \{ x_n^{(0)} \}, \{ x_n^{(1)} \}, \ldots, \{ x_n^{(m-1)} \}, \quad n = 0, 1, \ldots, N - 1, \]
we can compute the DFT of all \( m \) sequences simultaneously with one DFT as follows. Choose a basis \( \{ \beta_0, \beta_1, \ldots, \beta_{m-1} \} \) of \( F_K \). Form a new sequence \( \{ x_n \} \) from the \( m \) sequences by defining
\[ x_n = x_n^{(0)} \beta_0 + x_n^{(1)} \beta_1 + \cdots + x_n^{(m-1)} \beta_{m-1}, \quad n = 0, 1, \ldots, N - 1 \]
Note that the sequences \( \{ x_n^{(i)} \} \) are in \( K \) while the new sequence \( \{ x_n \} \) is in \( F \). From (7), we have that
\[ x_n W_n^k = \sum_{n=0}^{N-1} x_n W_n^k + \sum_{n=0}^{N-1} x_n^{(1)} W_n^k + \cdots + \sum_{n=0}^{N-1} x_n^{(m-1)} W_n^k. \]
Thus, if we denote the DFT of \( \{ x_n \} \) by \( \{ X_k \} \) and the DFT of \( \{ x_n^{(i)} \} \) by \( \{ X_k^{(i)} \} \), then the previous equation can be restated as
\[ X_k = \beta_0 X_k^{(0)} + \beta_1 X_k^{(1)} + \cdots + \beta_{m-1} X_k^{(m-1)}. \]
The problem is to generate the \( X_k^{(i)} \)'s from \( X_k \)'s.

Consider the equivalence class (5) of \( X_k \)
\[ X_k = \beta_0 X_k^{(0)} + \beta_1 X_k^{(1)} + \cdots + \beta_{m-1} X_k^{(m-1)}, \]
\[ X_k^q = \beta_0 X_k^{q(0)} + \beta_1 X_k^{q(1)} + \cdots + \beta_{m-1} X_k^{q(m-1)}, \]
\[ X_k^{q(m-1)} = \beta_0 X_k^{q(m-1)} + \beta_1 X_k^{q(1)} + \cdots + \beta_{m-1} X_k^{q(m-1)}. \]
Since \( \{ x_n \} \) is an \( F \)-sequence, the conjugacy relation (4) does not hold for the \( X_k \)'s. However, the conjugacy relation does hold for the \( X_k^q \)'s since \( \{ x_n^q \} \) are \( K \)-sequences. Therefore, the above can be rewritten as
\[ X_k = \beta_0 X_k^{(0)} + \beta_1 X_k^{(1)} + \cdots + \beta_{m-1} X_k^{(m-1)}, \]
\[ X_k^q = \beta_0 X_k^{(0)} + \beta_1 X_k^{(1)} + \cdots + \beta_{m-1} X_k^{(m-1)}, \]
\[ X_k^{q(m-1)} = \beta_0 X_k^{q(m-1)} + \beta_1 X_k^{q(1)} + \cdots + \beta_{m-1} X_k^{q(m-1)}. \]

Applying \( \varphi_m \) to \( X_k^q \) and remembering that in a finite field an element raised to the order of field is the element itself, we have
\[ \varphi_0(X_k) = \varphi_0(\beta_0) X_k^{(0)} + \varphi_0(\beta_1) X_k^{(1)} + \cdots + \varphi_0(\beta_{m-1}) X_k^{(m-1)} \]
\[ = \beta_0 X_k^{(0)} + \beta_1 X_k^{(1)} + \cdots + \beta_{m-1} X_k^{(m-1)} = X_k, \]
\[ \varphi_m(X_k^q) = \varphi_m(\beta_0) X_k^{q(0)} + \varphi_m(\beta_1) X_k^{q(1)} + \cdots + \varphi_m(\beta_{m-1}) X_k^{q(m-1)} \]
\[ = \beta_0 X_k^{q(0)} + \beta_1 X_k^{q(1)} + \cdots + \beta_{m-1} X_k^{q(m-1)} = X_k^{q(m-1)}, \]
Note that the sequences \( \{ x_n^{(i)} \} \) are in \( K \) while the new sequence \( \{ x_n \} \) is in \( F \). From (7), we have that
\[ \left( \begin{array}{c} X_k^{q(m-1)} \\ X_k^{q(m-2)} \\ \vdots \\ X_k^{q(m-1)} \\ \end{array} \right) = \left( \begin{array}{ccccc} \beta_0 & \beta_1 & \cdots & \beta_{m-1} \\ \beta_0^m & \beta_1^m & \cdots & \beta_{m-1}^m \\ \vdots & \vdots & \ddots & \vdots \\ \beta_0^{m-1} & \beta_1^{m-1} & \cdots & \beta_{m-1}^{m-1} \end{array} \right) \left( \begin{array}{c} X_k^{(0)} \\ X_k^{(1)} \\ \vdots \\ X_k^{(m-1)} \end{array} \right), \]
From this it is clear that the \( X_k^{(i)} \)'s are recoverable from the \( X_k \)'s iff the matrix
\[ M \triangleq \left( \begin{array}{cccc} \varphi_0(\beta_0) & \varphi_0(\beta_1) & \cdots & \varphi_0(\beta_{m-1}) \\ \varphi_m(\beta_0) & \varphi_m(\beta_1) & \cdots & \varphi_m(\beta_{m-1}) \\ \vdots & \vdots & \ddots & \vdots \\ \beta_0 & \beta_1 & \cdots & \beta_{m-1} \end{array} \right) \]
is invertible. Fortunately, this is the case.

Theorem 1 [1]: \( M \) is invertible iff \( \{ \beta_0, \beta_1, \ldots, \beta_{m-1} \} \) is a basis of \( F_K \).

Therefore,
Proof:
\[
(MM^{-1})_{ij} = \sum_{k=0}^{m-1} \alpha^{m-i+k} \eta^{m-j+k} = \delta_{ij},
\]
\[\Box\]

Tables of finite fields and normal bases can be found in [1], [7], [8].

IV. THE REAL/COMPLEX CASE REVISITED

Though the previous procedure is developed in the context of finite fields, the derivation is valid, mutatis mutandis, for the case \(K = \mathbb{R}\) and \(F = \mathbb{C}\). By making use of the automorphism group \(\text{Aut}_{KF}\), we have ensured the validity of the derivation in all respects except for the statement of the conjugacy relationship. By replacing all expressions of conjugacy relations in finite fields by their counterparts in the complex field, we give an algorithm for computing two real DFT's with one complex DFT. It will be seen that the algorithm includes, as a special case, the usual method for performing this computation.

For \(K = \mathbb{R}\) and \(F = \mathbb{C}\), we have \([C:K] = 2\) and \(\text{Aut}_{RC} = \{\varphi_0, \varphi_1\}\), where
\[
\varphi_0 : a + ib \mapsto a + ib,
\]
\[
\varphi_1 : a + ib \mapsto a - ib.
\]

Let \(\{\beta_0, \beta_1\}\) be a basis for \(CR\). Given 2 real sequences \(\{x_n^{(0)}\}\) and \(\{x_n^{(1)}\}\), form the complex sequence \(\{x_n\}\) by defining
\[
x_n = x_n^{(0)} \beta_0 + x_n^{(1)} \beta_1, \quad n = 0, 1, \ldots, N - 1.
\]

Denote the DFT of \(\{x_n^{(0)}\}\), \(\{x_n^{(1)}\}\), and \(\{x_n\}\), respectively. Then the three DFT's are related by (11) as
\[
\begin{pmatrix}
X_k^{(0)} \\
X_k^{(1)}
\end{pmatrix} = M^{-1}
\begin{pmatrix}
\varphi_0(X_k) \\
\varphi_1(X_k)
\end{pmatrix},
\]
where (see (10))
\[
M = \begin{pmatrix}
\varphi_0(\beta_0) & \varphi_0(\beta_1) \\
\varphi_1(\beta_0) & \varphi_1(\beta_1)
\end{pmatrix}.
\]

In other words,
\[
\begin{pmatrix}
X_k^{(0)} \\
X_k^{(1)}
\end{pmatrix} = \frac{1}{\beta_0 \beta_1 - \beta_0 \beta_1} \begin{pmatrix}
\beta_1 & -\beta_0 \\
-\beta_0 & \beta_1
\end{pmatrix}
\begin{pmatrix}
X_k \\
X_{-k}
\end{pmatrix}.
\]

In particular, for \(\beta_0 = 1\) and \(\beta_1 = i\), we have
\[
\begin{pmatrix}
X_k^{(0)} \\
X_k^{(1)}
\end{pmatrix} = \frac{1}{-2i}
\begin{pmatrix}
-i & -i \\
-i & i
\end{pmatrix}
\begin{pmatrix}
X_k \\
X_{-k}
\end{pmatrix} = \frac{1}{2}
\begin{pmatrix}
1 & 1 \\
-1 & i
\end{pmatrix}
\begin{pmatrix}
X_k \\
X_{-k}
\end{pmatrix},
\]
which is the usual method of computing two real DFT's with one complex DFT [6].

V. CONCLUSION

For finite fields \(F\) and \(K\) with \(K \subset F\) and \([F:K] = m\) we have shown that a single DFT over \(F\) has complexity approximately \(m\) times that of DFT over \(K\) by showing how to compute \(m\) \(K\)-DFT's simultaneously with one \(F\)-DFT. The result can be generalized to arbitrary fields \(F\) and \(K\) by assuming additionally the normality and separability of \(F\) over \(K\). Thus, the precise condition required for this argument to work for arbitrary fields \(F\) and \(K\) is that \(F\) is a finite Galois extension of \(K\). In the case considered here, since all finite extensions of finite fields are Galois, the only assumption needed is that \([F:K]\) is finite. For details, see [9].
Random Interactions in Higher Order Neural Networks

Pierre Baldi and Santosh S. Venkatesh

Abstract—Recurrent networks of polynomial threshold elements with random symmetric interactions are studied. Precise asymptotic estimates are used to specify the expected number of fixed points as a function of the margin of stability. In particular, it is shown that for a critical range of margins of stability (depending on the degree of polynomial interaction) such that the expected number of fixed points with margins below the critical range grows exponentially with the number of nodes in the network, while the expected number of fixed points with margins above the critical range decreases exponentially with the number of nodes in the network. The random energy model is also briefly examined and links with higher order neural networks and higher order spin glass models are made explicit.

Index Terms—Neural networks, spin glasses, polynomial threshold elements, fixed points, Laplace's method.

I. INTRODUCTION

Recurrent networks of formal neurons have been popular in a variety of computational applications. The model neurons in such structures are typically linear threshold elements which compute the sign of a linear form of the inputs. A recurrent network results when such elements are fully interconnected, and as in any recurrent system, the fixed points are important in the characterization of the computations done by the structure. A particular case of interest results when the interconnections between neurons are symmetric:

in such cases network dynamics are regulated by a Hamiltonian or energy function (cf. Hopfield [1], for instance). In such an instance, we can imagine the state space of the network to be embedded in an energy landscape with fixed points residing at energy minima.

A classical application of such networks is in associative memory physics (spin glasses), or of neural networks, before any learning has occurred, or in the limiting case when too much learning has occurred (the onset of senility!). These will be our focus of analysis in this paper: in particular, we consider recurrent, higher order neural networks with symmetric, random polynomial interactions. We characterize the fixed points of these structures according to their margin of stability1 that is a measure of how stable a fixed point is with respect to perturbations. Our main result may be informally stated as follows:

There exists a critical range of margins of stability (depending on the degree of polynomial interaction) such that the expected number of fixed points with margins of stability below the critical range increases exponentially in the size of the network while the expected number of fixed points with margins of stability above the critical range decays exponentially as the size of the network is increased.

There is thus a threshold phenomenon in evidence for the expected number of fixed points around the critical range of the margin of stability. The fact that for a certain range of margins the expected number of fixed points grows exponentially with the number of nodes in the network is not unexpected; more counter-intuitive, perhaps, is the existence of a critical margin of stability above which the expected number of fixed points actually decays as more nodes are added. We also provide exact asymptotic expressions for the coefficients and exponents in the regime of exponential behavior, and evaluate the critical margins of stability. While considerable attention has been focused on spin glass models in the statistical physics literature, at

REFERENCES


1In this context, this notion is due to Komlós and Paturi [9].