EXACT CONVOLUTIONS BY NUMBER-THEORETIC TRANSFORMS

Philip J. Erdelsky

Data/Ware Development, Incorporated

Prepared for:

Naval Undersea Center

2 May 1975

DWD

Approved for public release;
distribution unlimited.

DATA/WARE DEVELOPMENT, INC.

DATA/WARE DEVELOPMENT INC.
11585 SORRENTO VALLEY ROAD
SUITE 108
SAN DIEGO, CA 92121

(714) 453-7660

T-55-703

DWD

COMPUTER SYSTEMS, APPLICATIONS, IMPLEMENTATION

(1)

EXACT CONVOLUTIONS BY NUMBER-THEORETIC TRANSFORMS

by

Dr. Philip J. Erdelsky

Contract No. N66001-75-M-C112

Consultant:
Professor Solomon W. Golomb
University of Southern California

Prepared for
Naval Undersea Center
Code 60
May 2, 1975

CONTENTS

I INTRODUCTION

Because of their importance in every engineering discipline, the calculation of convolutions deserves close study. With the advent of the Fast Fourier Transform it was possible to speed up this calculation substantially -- this despite the fact that the FFT itself involves a large number of trigonometric table look-ups and multiplications. This raised the question as to whether some other transform might be more efficient.

Interestingly enough there are many such transforms available to speed up convolutions. Collectively they are referred to as Number Theoretic Transforms. Important examples are the Mersenne and Fermat Transforms discussed by Rader. These transforms, unlike the Discrete Fourier Transform, are defined over finite fields. Thus there are marked similarities between working with these transforms and calculating by means of a typical computer using one's complement notation. A significant difference is that in the finite field there is no roundoff error so that in the words of Rader, "the convolution is obtained with perfect accuracy." A second advantage is that certain of these transforms involve no multiplications at all other than by powers of 2, which are shifts. Higher speeds and simpler hardware implementations result.

In this report the Mersenne and Fermat Transforms are studied in some detail. Both results available in the literature and new results are presented. In addition to Part II, Practical Considerations, and Part III, Theoretical Considerations, a number of Appendices are collected as Part IV and are intended to summarize some basic material needed in the body of the report.

Results on multidimensional transforms to extend the convolution block length and thus overcome an inherent problem are given. Also, there are extensions to complex numbers in which the transform coefficients remain real. It is interesting to note that the Discrete Fourier Transform can itself be expressed as a convolution and computed by these methods.

II PRACTICAL CONSIDERATIONS

1. Required Properties

1.1 Convolutional Transforms Defined

The discrete Fourier transform of a finite sequence $\{a_0, a_1, \ldots, a_{N-1}\}$ is given by

$$A_k = \sum_{j=0}^{N-1} a_j w^{jk}, \quad k = 0, 1, \ldots, N-1, \qquad (1)$$

where $w = \exp(-2\pi i/N)$. It has an inverse given by

$$a_j = N^{-1} \sum_{k=0}^{N-1} A_k w^{-jk}, \quad j=0, 1, \ldots, N-1. \qquad (2)$$

Moreover, if $\{A_k\}$ and $\{B_k\}$ are the transforms of $\{a_j\}$ and $\{b_j\}$, respectively, then the inverse transform of the term-by-term product $\{A_k B_k\}$ is the circular convolution $\{C_m\}$ given by

$$C_m = \sum_{j=0}^{m} a_j b_{m-j} + \sum_{j=m+1}^{N-1} a_j b_{N+m-j},$$

$$m = 0, 1, \ldots, N-1 \qquad (3)$$

This is called the convolution property. Any transform of the form of (1) and (2) with this property will be called a convolutional transform.

In many cases it is more efficient to compute convolutions by this method than to compute them directly from (3). However, a major disadvantage lies in the fact that w is a complex number, and the transforms must be done in complex arithmetic, even when the sequences to be convolved are real. Also, the components of w are irrational for most N, and hence they cannot be represented exactly in machine computations.

1.2 Convolutional Transforms with Modular Arithmetic

There is a large class of convolutional transforms which involve only modular arithmetic on integers, which can be done without round off error in machine computations. (See Appendix A on congruencies and modular arithmetic.) In many of them, w or -w

will be a power of 2, so that multiplications by powers of w (even negative powers, as we shall see in subsection 1.3) can be done by mere shifts and sign changes. We have also examined many kinds of transforms in which $w \neq \pm 2^k$, but they do not appear to be practical. The transforms described in this report include those of Rader (6) and also many others. All of them will permit straightforward extension to complex numbers. (See section 4.)

1.3 Required Properties

Let M be the modulus of the arithmetic. The following conditions on M, w and N are necessary and sufficient for (1) and (2) to be convolutional. (See Theorem 5 in Part III for the proof):

$$w^N \equiv 1 \pmod{M}$$

$$(N,M) = 1$$

$(w^d-1, M) = 1$ for every positive integer d such that N/d is prime.

The notation (m,n) stands for the greatest common division of m and n. (see appendix B). Since $w^{-1} \equiv w^{N-1} \pmod{M}$, negative powers of w are congruent to appropriate positive powers.

The choice of suitable transforms is limited by the additional requirement that N divide p-1 for every prime factor of M (See Theorem 6 in part III for the proof). In particular, there are no suitable transforms where M is of the form 2^m, which is unfortunate because arithmetic modulo M would be easy to implement.

If M is of the form 2^m-1, then arithmetic is only slightly harder to implement. It is simply m-bit ones complement arithmetic with end-around carry. The only transforms in these cases with $w=\pm 2^k$ are Rader's "Mersenne Transforms" and "double length Mersenne transforms" (6), which are described more fully in Section 3. (See Theorems 9 and 10 in Part III for the proofs).

If N is highly composite, then (1) and (2) will have decompositions analogous to the fast Fourier transform. See Section 2 for details. Unfortunately, N is either prime or of the form 2p, where p is prime, in the Mersenne transforms of Rader (6). However, there is a large class of useful transforms with $N=2^n$, but in these cases M is not of the form 2^{m-1}, and modular arithmetic is more difficult to implement, but it is still practical. (See Section 7 for details).

2. Fast Fourier Decompositions and Operation Counts

2.1 The Method

Suppose N in Subsection 1.1 is composite, and write N = QR. For best results, Q should be prime, but this is not necessary. Then we can write

$$j = j_1 Q + j_2, \quad j_1 = 0, 1, \ldots, R-1, \quad j_2 = 0, 1, \ldots, Q-1,$$

$$k = k_1 R + k_2, \quad k_1 = 0, 1, \ldots, Q-1, \quad k_2 = 0, 1, \ldots, R-1.$$

Then

$$A_k = \sum_{j_2=0}^{Q-1} \sum_{j_1=0}^{R-1} a_{j1Q+j2}\, w^{(j1Q+j2)k}$$

$$= \sum_{j2=0}^{Q-1} w^{j2k} \sum_{j1=0}^{R-1} a_{j1Q+j2}\, (w^N)^{j1k1}\, (w^Q)^{j1k2}.$$

Since $w^N = 1$,

$$A_k = \sum_{j2=0}^{Q-1} w^{j2k} \left\{ \sum_{j1=0}^{R-1} a_{j1Q+j2}\, (w^Q)^{j1k2} \right\}$$

For each value of j_2, the expressions in braces are an R- element transform described in Subsection 6.2.

2.2 Operation Count

The number of additions required to compute a transform by this method is

N(Q-1) + Q x (number required for R- element transform)

If R is also composite, the process can be repeated, until finally a prime block length is encountered, which must be computed directly from the definition. The process is most efficient if Q is prime each time. The total number of additions required can be obtained by a simple recursion and is found to be $N(p_1+p_2+\ldots+p_n-n)$, where $N=p_1p_2\ldots p_n$ is the prime factorization of N. H is generally minimized when the prime factors of N are small, especially when $N=2^n$.

The number of shifts (multiplications by powers of w) is the same as the number of additions.

3. Specific Methods

3.1 Generalized Fermat and Mersenne Transforms

One class of methods is given by

$$w = \pm 2^k, \quad N=p^n, \quad M=(T^p-1)/(T-1)$$

where $T = w^{p^{n-1}}$, p is prime, n and k are positive integers, $w \not\equiv 1 \pmod{p}$ and M is the modulus. (See Theorem 7 in part III for the proof.)

3.2 Fermat Transform

The special case p = w = 2 in Subsection 3.1 is the "Fermat transform" of Rader (6). In this case $N^{-1} = 2^{2^n-n}$, so multiplication by N^{-1} is a mere shift. See Subsection 7.5 and 7.6 for implementation of arithmetic modulo M. Since N is highly composite, there is an efficient fast Fourier decomposition, as described in Section 2. The number of additions is $n2^n$.

To extend this method to block length 2N, simply use $v = 2^{2^n-3}(2^{2^n-1}-1)$ and L = 2 in the method described in Subsection 6.3.

3.3 Mersenne Transform

The special case n = 1, and w = 2 in Subsection 3.1 is the "Mersenne Transform" of Rader (6). In this case $M = 2^p-1$, so

arithmetic modulo M is simply p-bit ones-complement arithmetic, and $N^{-1} = -(2^p-2)/p$. The number of additions is p(p-1).

3.4 Generalized Double-Length Mersenne Transform

Another class of possibilities is given by

$$w = \pm 2^k, \quad N = 2p^n, \quad M = (T^p+1)/(T+1),$$

where $T = w^{p^n-1}$, p is an odd prime, n and k are positive integers, $w \not\equiv -1 \pmod{p}$ and M is the modulus. See Theorem 8 in part III for proof.

3.5 Double-Length Mersenne Transform

The special case n = 1 and w = -2 in subsection 3.4 is the "double-length Mersenne Transform" of Rader (6). In this case $M = 2^p-1$, so arithmetic modulo M is simply p-bit ones complement arithmetic, and $N^{-1} = -(2^{p-1}-1)/p$. The number of additions is 2p(p-1).

3.6 Maximal Moduli and Block Lengths

For any given N and w, there is a maximal modulus M which makes the transform convolutional, and all other such moduli are precisely the divisors of M. For the methods described in Subsections 3.1 and 3.4, the given M are maximal. For some other values of N and w, the maximal M (factored into prime factors) are given below:

N	w	M
12	±2	13
"	±4	241
"	±8	37.109
"	±16	97.673
"	±32	13.16.1321
15	2	151
"	4	151.331
"	8	63.23311
"	-2	151.331
"	-4	11.61.151.331.1321
"	-8	331.18837001

For any given modulus M, there is a maximal block length N that admits a convolutional transform (for some w). All other such block lengths are precisely the divisors of N. The maximal block lengths for some M are given below:

M	N
2^m	1
$2^{2m}-1$	2
2^3-1	6
2^5-1	30
2^7-1	126
2^9-1	6
$2^{11}-1$	22
$2^{13}-1$	8190
$2^{15}-1$	6
$2^{17}-1$	131070
$2^{19}-1$	524286
$2^{21}-1$	6
$2^{23}-1$	46
$2^{25}-1$	30
$2^{27}-1$	6
$2^{29}-1$	58
$2^{31}-1$	2147483646

3.7 Finite Fields

If M is prime, then N can be taken to be any divisor of M-1, but w may not be of the form $\pm 2^k$. If M is an odd prime and M-1 is not divisible by 4, then convolutional transforms for the complex numbers whose real and imaginary parts are integers modulo M exist for any N that divides M^2-1, but w may be complex, and even when it is real it may not be simple. See Subsection 4.2 for specific methods in this class.

4. Complex Convolutions and Other Extensions

4.1 Extension to Complex Numbers

All of the foregoing methods also have the desired convolution property when $\{a_j\}$ and $\{b_j\}$ are sequences of complex numbers whose real and imaginary parts are integers modulo M. See Theorem 4 in part III for the proof. The coefficients w^{jk} remain real, and therefore the real and imaginary parts of $\{a_j\}$ can be transformed separately to obtain the real and imaginary parts of $\{A_k\}$. Similar remarks apply to the other transforms. Actual complex arithmetic is used only to compute the products A_kB_k.

4.2 Complex Mersenne Transforms

Reed and Truong (7) have defined a class of methods for complex numbers in which M is a prime number of the form 2^m-1. The arithmetic is easy to implement, and N can be taken to be any divisor of M^2-1. Since M^2-1 is divisible by 2^m, the methods have an efficient fast Fourier decomposition. The disadvantage is in the fact that w is usually a complex number, and not a very simple one at that.

4.3 Other Extensions

The transforms of Subsection 1.1 may also have the convolution property when $\{a_j\}$ and $\{b_j\}$ are other kinds of generalized numbers whose components are integers modulo M. Some possibilities are:

(i) Square matrices (all of the same size)
(ii) Quaternions
(iii) Polynomials
(iv) L-element sequences that are multiplied by convolution (treated more fully in the next section).

In each case, the coefficients w^{jk} remain "scalars", and the generalized multiplication (of matrices, polynomials, etc.) is used only to compute the products A_kB_k.

5. Multidimensional Convolutions and Transforms

5.1 Two-Dimensional Convolutions Defined

The N x L matrix (a_{ij}) can be considered as a sequence $\{A_0, A_1, \ldots, A_{N-1}\}$ of its rows. (Notice the slightly nonstandard numbering of rows). If (b_{ij}) is a similar matrix, we define the two-dimensional circular convolution of (a_{ij}) and (b_{ij}) to be the N x L matrix (c_{ij}) whose rows are given by

$$C_m = \sum_{j=0}^{m} A_j * B_{m-j} + \sum_{j=m+1}^{N-1} A_j * B_{N+m-j},$$

$$m = 0, 1, \ldots, N-1, \qquad (4)$$

where * represents the circular convolution of L-element row vectors, and the usual vector addition is used. It is actually a "convolution of convolutions".

5.2 Transform Techniques

Our transforms are also convolutional when $\{a_j\}$ and $\{b_j\}$ are sequences of L-element row vectors which are multiplied by convolving them. Other vector operations are defined in the usual way.

The components of the vectors may be integers modulo M or complex numbers whose real and imaginary parts are integers modulo M. Hence the transforms can be used to calculate (4) in the following manner. The transforms of $\{A_0, A_1, \ldots, A_{N-1}\}$ and $\{B_0, B_1, \ldots, B_{N-1}\}$ are found by (1). The transforms are convolved, term by term, perhaps by using another transform, and the appropriate inverse transform is applied to the result.

In transforming $\{A_0, A_1, \ldots, A_{N-1}\}$ we notice that the k-th row of the resulting matrix $(\hat{a}_{kl})$ is given by

$$\hat{A}_k = \sum_{j=0}^{N-1} A_j w^{jk},$$

and in particular the element $\hat{a}_{kl}$ in this row is given by

$$\hat{a}_{kl} = \sum_{j=0}^{N-1} a_{jl} w^{jk}.$$

Hence the l-th column of $(\hat{a}_{kl})$ is the ordinary transform of the l-th column of (a_{ij}). Therefore, the transform is done column-by-column. A similar remark applies to the other transforms.

If a transform is also used to convolve rows, the entire process can be described as follows: Transform each column of (a_{ij}). Then transform each row. (The result is often called a two-dimensional transform.) Do the same to (b_{ij}). Then multiply the resulting matrices element-by-element. Then perform an inverse transform on each row. Then perform an inverse transform on each column. The method is possible only if the row and column transforms use the same modulus.

5.3 Higher Dimensions

If * in (4) represents a two-dimensional convolution, then the entire system of equations represents a three-dimensional convolution. The process can be repeated indefinitely to define convolutions of any dimension, but we see no practical use for convolutions of dimensions higher than 2. Higher dimensional convolutions can also be computed by transform techniques.

6. Convolutions of Different Lengths and Nonperiodic Convolution

6.1 General

Our methods generally place restrictions on the block length N. The techniques described in this section can be used to get around these restrictions to some extent.

6.2 Shortening by Interlacing

If $N = QR$, where Q and R are positive integers, then the transform

$$A_k = \sum_{j=0}^{R-1} a_j (w^Q)^{jk}, \quad k = 0, 1, \ldots, R-1,$$

for sequences of length R, is also convolutional. The same modulus is used. (See Theorem 2 in part III for the proof.)

6.3 Lengthening by Roots

If $V^L = w$ and L is not divisible by any prime number that does not also divide N, then the transform

$$A_k = \sum_{j=0}^{NL-1} a_j V^{jk}, \quad k = 0, 1, \ldots, NL-1,$$

for sequences of length NL, is also convolutional. (See Theorem 3 in part III for the proof.) If the original transform was over a finite field (as in the case where the modulus M is prime and real arithmetic is being used, as in Reed and Truong's methods described in Subsection 4.2), then such a V exists whenever the number of nonzero elements in the field is divisible by NL.

6.4 Shortening by Zero Fill

If we want to convolve two sequences $\{a_j\}$ and $\{b_j\}$ of length n, where $2n-1 \leq N$, to obtain $\{c_j\}$, we can simply convolve the N-element sequences $\{a_0, a_1, \ldots, a_{n-1}, 0, 0, \ldots, 0\}$ and $\{b_0, b_1, \ldots, b_{n-1}, 0, 0, \ldots, 0, b_1, b_2, \ldots, b_{n-1}\}$ to obtain $\{c_0, c_1, \ldots, c_{n-1}, x_1, x_2, \ldots, x_{N-n}\}$, where $x_1, x_2, \ldots, x_{N-n}$ are values that are of no interest and need not even be computed.

6.5 Nonperiodic Convolution by Zero Fill

Similarly, the nonperiodic convolution given by

$$c_m = \sum_{j=0}^{m} a_j b_{m-j}, \quad m = 0, 1, \ldots, n-1$$

can be computed by convolving the N-element sequences $\{a_0, a_1, \ldots, a_{n-1}, 0, 0, \ldots, 0\}$ and $\{b_0, b_1, \ldots, b_{n-1}, 0, 0, \ldots, 0\}$ to give $\{c_0, c_1, \ldots, c_{n-1}, y_1, y_2, \ldots, y_{N-n}\}$, where $y_1, y_2, \ldots, y_{N-n}$ need not be computed.

6.6 Lengthening by Two-Dimensional Convolution

If we wish to convolve (periodically) the two sequences $\{a_j\}$ and $\{b_j\}$ of length LN to obtain $\{c_j\}$, we can perform the N x E two-dimensional convolution (see Section 5 for the definition) on

$$A = \begin{bmatrix} a_0 & a_1 & \cdots & a_{L-1} & 0 & 0 & \cdots & 0 \\ a_L & a_{L+1} & \cdots & a_{2L-1} & 0 & 0 & \cdots & 0 \\ & & & * \ * \ * & & & & \\ a_{LN-L} & a_{LN-L+1} & \cdots & a_{LN-1} & 0 & 0 & \cdots & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} b_0 & b_1 & \cdots & b_{L-1} & z & b_{LN-L+1} & b_{LN-L+2} & \cdots & b_{LN-1} \\ b_L & b_{L+1} & \cdots & b_{2L-1} & z & b_1 & b_2 & \cdots & b_{L-1} \\ & & & * \ * \ * & & & & & \\ b_{LN-L} & b_{LN-L+1} & \cdots & b_{LN-1} & z & b_{LN-2L+1} & b_{LN-2L+2} & \cdots & b_{LN-L-1} \end{bmatrix}$$

where $E \geq 2L-1$ and z represents 2L-E-1 zeros, to obtain a matrix C whose first L columns are of the form

$$\begin{bmatrix} C_0 & C_1 & \cdots & C_{L-1} \\ C_L & C_{L+1} & \cdots & C_{2L-1} \\ & * \ * \ * & & \\ C_{LN-L} & C_{LN-L+1} & \cdots & C_{LN-1} \end{bmatrix}$$

The other columns of C are of no interest and need not be computed.

The convolution can be computed as described in the Section 5. In some cases it may be desirable to choose E larger than the minimum value of 2L-1, since convolutions of length E may be easier to compute than those of length 2L-1.

Notice that the transforms of A and B are to be computed column-by-column. The transforms of the last E-L columns of A are obviously zero, and the transforms of the last E-L columns of B are related to those of other columns by property (b) in Appendix E.

7. Accuracy and Arithmetic

7.1 General

The methods described in this report produce exact results modulo M. If $\{a_j\}$ and $\{b_j\}$ are integer sequences which are to be convolved to give $\{c_j\}$, then our methods do not necessarily give the value of each c_j that would have resulted from ordinary arithmetic, but may instead give another value that differs from it by a multiple of M. If it is known in advance that $|c_j| < \frac{1}{2}M$, then this information is sufficient to identify c_j. In fact, the actual value can be obtained by adding an appropriate multiple of M to the computed value.

7.2 Scaling the Data

One way to guarantee that the actual value of c_j is in the desired range is to require that $|a_j|, |b_j| < \frac{1}{2}M/N$. Hence the a_j and b_j must be scaled and rounded to integers in this range. In the case of complex numbers, the same condition will ensure that the real and imaginary parts of the answers are identifiable.

7.3 Use of the Chinese Remainder Theorem

If two methods with relatively prime moduli M_1 and M_2 are used to compute the same convolution, then the answers modulo M_1M_2 can be computed by using the Chinese remainder theorem (see appendix C). However, it is difficult to find transforms with

relatively prime moduli and equal block lengths, and in many such cases it is better to use a transform with modulus M_1M_2. In particular, if w is the same in both transforms, then the transform also applies modulo M_1M_2.

7.4 Choice of Register Length

It can be shown that when M is odd, $2^m \equiv 1 \pmod{M}$ for some positive integer m, and M is odd in all our transforms. Hence arithmetic modulo M can be done in m-bit registers with ones complement arithmetic and end-around carry. For an end-around carry corresponds to replacing 2^m by 1, and these are congruent. Similarly, the ones complement of x is $2^m-1-x \equiv -x \pmod{M}$. Of course, we want m to be as small as possible. Since M divides 2^m-1, this puts a lower limit on m which is reached only for the Mersenne transform.

In methods with $w = 2^k$, we can take $m = kN$, and if $w = -2^k$, we can take $m = 2kN$, but there are smaller m in many cases.

7.5 Reduction Mod M

After the final answers c_j are computed modulo M they must be reduced modulo M to the range $|c_j| < \frac{1}{2}M$. One way to do this is to divide by M, take the remainder, and subtract M from it if it is greater than $\frac{1}{2}M$. If complex arithmetic is being used, both real and imaginary parts must be reduced.

We can avoid division in many cases. In the methods described in Subsections 3.1 and 3.2 in which $w = 2$ and $w = -2$, respectively, $M = T^{p-1}+T^{p-2}+\cdots+1$, where $T = 2^{p^{n-1}}$, and we can take $m = p^n$. Any m-bit answer may be written $GT^{p-1}+H$, where G continues the p^{n-1} most significant bits and H contains the rest. Then

$$GT^{p-1}+H \equiv GT^{p-1}+H-GM \pmod{M}$$

$$= H-(GT^{p-2}+GT^{p-3}+\ldots+G).$$

The expression in parentheses is simply a concatenation of p-1 copies of G. We now have a number in the interval $(-T^{p-1}, T^{p-1})$, and one further addition or subtraction of M will put it into the interval $(-\frac{1}{2}M, \frac{1}{2}M)$.

8. The Discrete Fourier Transform as a Convolution

8.1 General

Our transform methods can be used to compute convolutions. The discrete Fourier transform is not a convolution as it stands, but it can be made into a convolution in the ways described in this section. Transform methods can then be used to compute the convolution. This indirect method is sometimes better than the fast Fourier transform, especially in cases where multiplication is an expensive operation and it is desirable to minimize round-off error.

8.2 Chirp-Z Algorithm

One way to convert the discrete Fourier transform is by the so-called "chirp-Z" algorithm. It is easy to show that (1) can be written as

$$A_k = w^{\frac{1}{2}k^2} \sum_{j=0}^{N-1} (a_j w^{\frac{1}{2}j^2}) w^{-\frac{1}{2}(k-j)^2}, \qquad k = 0, 1, \ldots, N-1.$$

Moreover, when N is even,

$$w^{-\frac{1}{2}(N+k-j)^2} = (w^N)^{-(\frac{1}{2}N+k+j)} \, w^{-\frac{1}{2}(k-j)^2} = w^{-\frac{1}{2}(k-j)^2},$$

since $w^N = 1$. Hence

$$A_k = w^{\frac{1}{2}k^2} \left[\sum_{j=0}^{k} (a_j w^{\frac{1}{2}j^2}) \, w^{-\frac{1}{2}(k-j)^2} + \sum_{j=k+1}^{N-1} (a_j \, w^{\frac{1}{2}j^2}) \, w^{-\frac{1}{2}(N+k-j)^2} \right]$$

The expression in brackets is a circular convolution of $\{ a_j \, w^{\frac{1}{2}j^2} \}$ and $\{ w^{-\frac{1}{2}j^2} \}$, which can be computed by transform methods. This method requires 3N multiplications, compared to $N\log_2 N$ for the fast Fourier transform, so it may be faster when $N > 8$ and multiplication is much slower than other operations.

8.3 Prime Number Algorithm

If N is prime, there is a more efficient way. Let $\langle u \rangle$ represent the integer such that $0 \leq \langle u \rangle \leq N-1$ and $\langle n \rangle \equiv u \pmod N$. Since $w^N = 1$, $w^n = w^{\langle n \rangle}$. Also, $\langle mu \rangle = \langle\langle m \rangle \langle u \rangle\rangle$. (See Appendix A on modular arithmetic for details.) It is known that there is at least one integer g such that $\langle g^0 \rangle, \langle g^1 \rangle, \ldots, \langle g^{N-2} \rangle$ is some permutation of 1, 2, ..., N-1, and $\langle g^{N-1} \rangle = 1$. Write the discrete Fourier transform as

$$A_0 = \sum_{j=0}^{N-1} a_j$$

$$A_k = a_0 + \sum_{j=i}^{N-1} a_j \, w^{jk}, \quad k=1, 2, \ldots, N-1.$$

Then we can write $k = \langle g^m \rangle$ for $m = 0, 1, \ldots, N-2$ and $j = \langle g^{N-2-l} \rangle$, $l = 0, 1, \ldots, N-2$, in the last N-1 equations to obtain

$$A_{\langle g^m \rangle} = a_0 + \sum_{l=0}^{N-2} a_{\langle g^{N-2-l} \rangle} \, w^{\langle g^{N-2-l} \rangle \langle g^m \rangle}$$

$$= a_0 + \sum_{l=0}^{N-2} a_{\langle g^{N-2-l} \rangle} \, w^{\langle g^{N-2+m-l} \rangle}$$

Now $w^{\langle g^{N-1+s} \rangle} = w^{\langle g^{N-1} \rangle \langle g^s \rangle} = w^{\langle g^s \rangle}$, so we can write

$$A_{\langle g^m \rangle} = a_0 + \left\{ \sum_{l=0}^{m} a_{\langle g^{N-2-l} \rangle} \, w^{\langle g^{N-2+m-l} \rangle} + \sum_{l=m+1}^{N-2} a_{\langle g^{N-2-l} \rangle} \, w^{\langle g^{N-1+N-2+m-l} \rangle} \right\}$$

$$m = 0, 1, \ldots, N-2$$

The expressions in braces are the convolution of $\left\{a_{\langle g^{N-2-l}\rangle}\right\}_{l=0}^{N-2}$ and $\left\{w^{\langle g^{N-2+t}\rangle}\right\}_{t=0}^{N-2}$ and can be calculated by transform techniques.

III THEORETICAL CONSIDERATIONS

9. General Theory

In this section, R will be a ring with unit u, N will be a positive integer, and w and N' will be numbers in the center of R. (See appendix D on rings.) For convenience, a symbol such as $\sum i$ will represent summation over i = 0, 1, ..., N-1 and integers will be construed as natural multiples of u (eg., 3 is u+u+u), where the context so requires.

We define the transform of the finite sequence $\{a_0, a_1, \ldots, a_{N-1}\}$ of elements of R to be the sequence $\{A_k\}$ given by

$$A_k = \sum_i a_i w^{ki}, \quad k = 0, 1, \ldots, N-1, \tag{5}$$

and we define the inverse transform of $\{C_k\}$ to be the sequence $\{c_i\}$ given by

$$c_i = N' \sum_k C_k \, w^{N^2-ki}, \quad i = 0, 1, \ldots, N-1. \tag{6}$$

Notice that we have used exponents of the form N^2-ki in (6) instead of the usual $-ki$, since we have not yet shown that w^{-1} exists.

We say that (5) and (6) have the convolution property if the inverse transform $\{c_i\}$ of the term-by-term product $\{A_0B_0, A_1B_1, \ldots, A_{N-1}B_{N-1}\}$ of the transforms $\{A_k\}$ and $\{B_k\}$ of any two sequences $\{a_i\}$ and $\{b_i\}$ is the circular convolution given by

$$C_i = \sum_{j=0}^{i} a_j b_{i-j} + \sum_{j=i}^{N-1} a_j b_{N+i-j}, \quad i = 0, 1, \ldots, N-1. \tag{7}$$

By setting $b_0 = u$ and $b_i = 0$ for $i \geq 1$, we can easily show that if (5) and (6) have the convolution property, they are inverses.

Theorem 1. Of the following conditions, (c1), (c2) and (c3) imply the convolution property, and the convolution property implies (c1), (c2), (c3) and (c4):

(c1) $w^N = u$

(c2) $N'N = u$

(c3) $\sum_k w^{DK} = 0$ for every positive integer D such that N/D is prime

(c4) $\sum_k w^{dk} = 0$ for $d = 1, 2, \ldots, N-1$.

Proof of necessity. The convolution property implies that (5) and (6) are inverses. Let $a_{N-1} = b_1 = c_0 = u$, and let all other a_i, b_i and c_i be zero. Then $\{c_i\}$ is the convolution of $\{a_i\}$ and $\{b_i\}$. Then transforms have $C_1 = u$, $A_1 = w^{N-1}$ and $B_1 = w$. The convolution property and the fact that (5) and (6) are inverses imply that $A_1B_1 = C_1$ or $w^N = u$, which proves (c1).

Now $B_k = w^k$, so the inverse transform of $\{w^k\}$ must be $\{b_i\}$, or for i = 1

$$u = N' \sum_k w^k w^{N^2-k} = N'N,$$

which proves (c2).

For $i = N+1-d$, where $2 \leq d \leq N-1$,

$$0 = N' \sum_k w^k w^{N^2-(N+1-d)k} = N' \sum_k w^{dk}$$

Then multiply by N to obtain $0 = \sum_k w^{dk}$. If d = 1, similarly let i = 0. This proves (c3) and (c4).

To prove sufficiency we need the following Lemma.

Lemma A. Conditions (c1), (c2) and (c3) imply (c4).

Proof. Let r = (N,d) and write N/r = pq, where p is prime. Then (c3) applies with D = qr, and hence $\sum_k w^{qrk} = 0$.

Write $k = gp+h$, where $g = 0, 1, \ldots, qr-1$ and $h = 0, 1, \ldots, p-1$. Then

$$\sum_k w^{qrk} = \sum_{g=0}^{qr-1} \sum_{h=0}^{p-1} w^{qr(gp+h)} = 0.$$

Since $w^{qrp} = w^N = u$, This simplifies to

$$\sum_{g=0}^{qr-1} \sum_{h=0}^{p-1} w^{qrh} = qr \sum_{h=0}^{p-1} w^{qrh} = 0.$$

Multiply by $N' p w^{rf}$, insert the factor $w^{rpqg} = w^{Ng} = u$ into the summand, simplify and sum over $g = 0, 1, \ldots, r-1$, $f = 0, 1, \ldots, q-1$ to obtain

$$\sum_{g=0}^{r-1} \sum_{f=0}^{q-1} \sum_{h=0}^{p-1} w^{r(pqg+qh+f)} = 0.$$

Since k can also be written $k = pqg+qh+f$, where $g = 0, 1, \ldots, r-1$, $h = 0, 1, \ldots, p-1$, and $f = 0, 1, \ldots, q-1$, we have $\sum_k w^{rk} = 0$.

Now let $s = d/r$. Since $(s,N) = 1$ and $(w^r)^N = u$, $(w^r)^{sk}$ for $k = 0, 1, \ldots, N-1$ are $u, w, w^r, (w^r)^2, \ldots, (w^r)^{N-1}$ in some order, and

$$0 = \sum_k w^{rk} = \sum_k w^{rsk} = \sum_k w^{dk}.$$ ∎

Proof of sufficiency. By the definitions (5) and (6), we have

$$C_m = N' \sum_k \left(\sum_i a_i w^{ik}\right) \left(\sum_j b_j w^{jk}\right) w^{N^2-mk}$$

$$= \sum_i \sum_j a_i b_j N'B, \qquad (8)$$

where $B = \sum_k w^{N^2+k(i+j-m)}$.

Let t be the integer such that $0 \le i+j-m+tN<N$. Then since $w^N = u$,

$$B = \sum_k w^{k(i+j-m+tN)}$$

If $i+j-m+tN \neq 0$, then $B = 0$ by the Lemma; otherwise $B = N$. Hence (8) can be written $C_m = \sum_i \sum_j a_i b_j$, where the summation runs only over values of i and j such that $i+j-m+tN = 0$, which is another way of writing (7). ∎

Theorem 2. If the convolution property holds for R, N, w and N', and N is divisible by m, then it also holds for the new system R, N/m, w^m nad mN', respectively.

Proof. Conditions (c1) and (c2) are readily verified for the new system, and (c4) for the old system implies (c3) for the new system. ∎

Theorem 3. If the convolution property holds for R, N, w and N', $v^m = w$, and m is a positive integer not divisible by any prime numbers that do not also divide N, then the convolution property also holds for the new system R, mN, v and $N'm^{-1}$, respectively.

Proof. No generality is lost by assuming m is prime, since the theorem can be applied repeatedly to give the desired result.

Since $(N/m)N'm = u$, we can take $m^{-1} = (N/m)N'$.

Conditions (c1) and (c2) are readily verified for the new system.

In condition (c3), if mN/D is prime, then so is N/(D/m), and

$$\sum_k v^{Dk} = \sum_k (v^m)^{(D/m)k} = \sum_k w^{(D/m)k} = 0. \quad ∎$$

Theorem 4. If the convolution property holds for R, N, w and N', it also holds in any ring that contains R (or an isomorphic copy) as a subring, and it holds in any subring of R that contains w, N' and u.

Proof. This is obvious, since (c1), (c2) and (c3) involve only elements of the subring generated by w, N' and u. ∎

The most useful ring extensions are-

(i) The formal complex numbers x+yi, where x, y ∈ R and operations are defined in the usual ways.

(ii) The n x n matrices with elements in R.

(iii) The n-dimensional vectors with components in R, where vector multiplication is by a convolution.

(iv) The polynomials with coefficients in R.

(v) The formal quaternions with components in R.

The property actually holds over more general structures. For example, in (iii) scalar multiplication of vectors could be used.

10. Modular Arithmetic

In this section, we shall consider only the cases where R is the ring of integers modulo M, where $M \geq 2$. (The results also apply, with a few changes, to the Gaussian integers.)

Theorem 5. If $w \equiv 0 \pmod{M}$, the convolution property does not hold. If $w \equiv 1 \pmod{M}$, the convolution property holds only for N = 1. If $w \equiv -1 \pmod{M}$ the convolution property holds only when N = 2, or when N = 1 and M = 2. In all other cases, (c5), (c6) and (c7) imply the convolution property; and the convolution property implies (c5), (c6), (c7), (c8) and (c9):

(c5) $w^N \equiv 1 \pmod{M}$

(c6) $(N,M) = 1$

(c7) $(w^D-1,M) = 1$ for every positive integer D such that N/D is prime

(c8) $\sum_k w^{dk} \equiv 0 \pmod{M}$ for d = 1, 2, ..., N-1

(c9) $(w^d-1,M) = 1$ for every integer d between 1 and N-1, inclusive, that divides N.

Proof. If $w \equiv 0 \pmod M$ then (c1) is false. If $w \equiv 1 \pmod M$ and $N \geq 2$, then (c2) and (c3) are inconsistent. If $w \equiv -1 \pmod M$ and $M \geq 3$, then (c1) excludes $N = 1$ and (c4) excludes all $N \geq 3$.

In other cases, first assume (c5), (c6) and (c7). Since $(N,M) = 1$ the Euclidean algorithm (see appendix A) can be used to find an N' such that $N'N \equiv 1 \pmod M$. This satisfies (c2). Obviously (c5) implies (c1).

To prove (c3), let D be such that N/D is prime. Then

$$(w^D-1) \sum_k w^{Dk} = w^N - 1 \equiv 0 \pmod M$$

by (c5), that is, the left number is a multiple of M. Since $(w^D-1,M) = 1$, $\sum_k w^{Dk}$ must be a multiple of M, and hence congruent to 0. The desired conclusion follows from Theorem 1.

Conversely, assume the convolution property. Then Theorem 1 gives (c1), (c2), (c3) and (c4). Obviously (c1) implies (c5). From (c2), $N'N \equiv 1 \pmod M$, that is, M divides $N'N-1$. Hence (M,N) divides $N'N-1$ also. Since it divides $N'N$, it must also divide -1; hence (c6) holds. Obviously, (c4) implies (c8).

Now let d be any integer described in (c9). Then by (c8)

$$\sum_k w^{dk} = \frac{w^{Nd}-1}{w^d-1} = \frac{w^{Nd}-1}{w^N-1} \quad \frac{w^{N-1}}{w^d-1}$$

$$= \sum_{k=0}^{d-1} w^{Nk} \quad \frac{w^{N-1}}{w^d-1} \equiv 0 \pmod M.$$

Since $w^N \equiv 1 \pmod M$ by (c5), the expression in parenthesis is congruent to d, and

$$d \frac{w^N-1}{w^d-1} \equiv 0 \pmod M.$$

Since d divides N and (N,M) = 1 by (c6), (d,M) = 1 also. Hence d^{-1} exists in arithmetic modulo M and the above congruence implies that $(w^N-1)/(w^d-1) \equiv 0 \pmod{M}$.

For convenience, define $Q = w^d$ and $e = N/d$. It is easily shown that

$$(Q^{e-2} + 2Q^{e-3} + 3Q^{e-4} + \cdots + e-1)(Q-1) + e$$
$$= Q^{e-1} + Q^{e-2} + \cdots + 1 = \frac{w^N-1}{w^d-1} \equiv 0 \pmod{M}.$$

Since (Q-1,M) divides both Q-1 and right number, it must also divide e. Since e divides N, (Q-1,M) also divides N. Hence (Q-1,M) divides (N,M). Since (N,M) = 1 by (c6), (Q-1,M) = 1, which proves (c9) and (c7). ∎

Theorem 6. [2,3] The convolution property holds for some w if and only if N divides p-1 for every prime factor p of M.

Proof Suppose p is a prime factor of M. There is a generator g such that $g^{p-1} \equiv 1 \pmod{p}$, but $g^k \not\equiv 1 \pmod{p}$ for $k = 1, 2, \ldots, p-2$.

Now we wish to show by induction on k that $g^{p^k(p-1)} \equiv 1 \pmod{p^{k+1}}$ for all nonnegative integers k. For k = 0 it has been proven already. Write $T^p-1 = (T-1)(T^{p-1} + T^{p-2} + \cdots + 1)$ where $T = g^{p^k(p-1)}$. Then p^{k+1} divides T-1 by inductive hypothesis, and since $T \equiv 1 \pmod{p}$, $T^{p-1} + T^{p-2} + \cdots + 1 \equiv 0 \pmod{p}$. Hence p^{k+2} divides T^p-1.

Now let $v = g^s$, where $s = p^{e-1}(p-1)/N$ and p^e is the highest power of p that divides M. Then $v^N \equiv 1 \pmod{p^e}$.

Repeat this process and find such a v for each such p. Then use the Chinese remainder theorem (appendix C) to find a w with $w \equiv v \pmod{p^e}$ for every such p. Then $w^N \equiv 1 \pmod{M}$ and (c5) is established.

In (c7), $w^D-1 \equiv v^D-1 \pmod{p^e}$ for every such prime p, and $w^D-1 \equiv v^D-1 \pmod{p}$ also. Now $v^D = g^t$, where $t = p^{e-1}(p-1)/(N/D)$. Since p-1 does not divide t, $v^D \not\equiv 1 \pmod{p}$ and hence $w^D-1 \not\equiv 0 \pmod{p}$. Since this holds for every prime factor of M, $(w^D-1, M) = 1$, which establishes (c7).

Now (c6) is obviously true, since otherwise for some prime p, p would divide N, which would divide p-1. The desired result follows from Theorem 5.

To prove the converse, we first note that (c5), (c6) and (c8) are also true with M replaced by p. In (c8), if $w^d \equiv 1 \pmod{p}$, then $N \equiv 0 \pmod{p}$, which is impossible since $(N,p) = 1$. Hence $w^d \not\equiv 1 \pmod{p}$ for $d = 1, 2, \ldots, N-1$. By Fermat's theorem, $w^{p-1} \equiv 1 \pmod{p}$. By the Euclidean algorithm there are integers A and B such that $AN + B(p-1) = (N,p-1)$. Hence

$$w^{(N,p-1)} = w^{AN+B(p-1)} = (w^N)^A(w^{p-1})^B \equiv 1 \pmod{p}$$

and $(N,p-1) = N$. Therefore, N divides p-1. ∎

Corollary. If $M = 2^m$ for some positive integer m, then the convolution property holds only when N = 1.

Theorem 7. If p is a prime, n and w are integers, $n \geq 1$, $|w| \geq 2$, $w \not\equiv 1 \pmod{p}$, $N = p^n$, $M = \frac{T^p-1}{T-1}$, $T = w^{p^{n-1}}$,

then the convolution property holds.

Proof. Since M divides T^p-1 and $T^p-1 = w^N-1$, (c5) is established.

By repeated application of Fermat's theorem, one can obtain $T^p \equiv T \equiv w \not\equiv 1 \pmod{p}$. Hence $M \equiv 1 \pmod{p}$ and (c6) is established.

To establish (c7), we note that N/D is prime only when $D = p^{n-1}$, and that $w^D - 1 = T - 1$ in this case. It is easily shown that

$$(T-1)(T^{p-2} + 2T^{p-3} + \cdots + (p-1)) + p$$

$$= T^{p-1} + T^{p-2} + \cdots + 1 = M.$$

Hence $(T-1, M) = (T-1, p) = 1$. The desired result follows from Theorem 5. ∎

Theorem 8. If p is an odd prime, n and w are integers, $n \geq 1$, $w \geq 2$, $w \not\equiv -1 \pmod{p}$

$$N = 2p^n \quad M = \frac{T^p+1}{T+1}, \quad T = w^{p^{n-1}},$$

then the convolution property holds.

Proof. Since M divides T^p+1, and $T^p+1 = w^{p^n}+1$, $w^{p^n} \equiv -1 \pmod{M}$. Hence $w^{2p^n} \equiv 1 \pmod{M}$, and (c5) is established.

By repeated applications of Fermat's theorem one can obtain $T^p \equiv T \equiv w \not\equiv -1 \pmod{p}$. Hence $M \equiv 1 \pmod{p}$. Since $M = T^{p-1} - T^{p-2} + T^{p-3} - \cdots + 1$ is easily seen to be odd, whether T is odd or even, $(M, 2p^n) = 1$ also, and (c6) is established.

To establish (c7), first consider $D = 2p^{n-1}$. It is easily shown that

$$-(T+1)(T^{p-2} - 2T^{p-3} + 3T^{p-4} - \cdots - (p-1)) + p$$

$$= T^{p-1} - T^{p-2} + T^{p-3} - \cdots + 1 = M.$$

Hence $(T+1, M) = (T+1, p) = 1$. Similarly,

$$(T-1)(T^{p-2} + T^{p-4} + \cdots + T) + 1 = T^{p-1} - T^{p-2} + T^{p-3} - \cdots + 1 = M$$

and hence $(T-1, M) = 1$. Therefore $1 = ((T-1)(T+1), M) = (w^D - 1, M)$.

Now consider the other possibility $D = p^n$. In this case $w^D-1 = T^p-1 = M(T+1)+2$, so $(w^D-1, M) = (M,2) = 1$ since M is odd. The desired result follows from Theorem 5. ∎

Lemma B. If m and n are positive integers, then $(2^m-1, 2^n-1) = 1$ if and only if $(m,n) = 1$.

Proof. If $(m,n) > 1$, then $2^{(m,n)}-1$ divides both 2^m-1 and 2^n-1 and $(2^m-1, 2^n-1) > 1$. Now assume $(m,n) = 1$. No generality is lost by assuming $m<n$. Since $2^n-1 = (2^m-1)2^{n-m}+2^{n-m}-1$, $(2^m-1, 2^n-1) = (2^m-1, 2^{n-m}-1)$, where $(m,n-m) = 1$ also. If we continue this process long enough, one exponent will be reduced to 1 and the result will be obvious. ∎

Theorem 9. If $w = 2^k$ and $M = 2^m-1$ for some positive integers k and m, where $m \geq 2$, then the convolution property holds (for some k) if and only if N is prime and $m = N$.

Proof. If N is prime and $m = N$, then Theorem 7 shows that the convolution property holds for $k = 1$.

Conversely, Theorem 5 shows that $2^{kN}-1 \equiv 0 \pmod{2^m-1}$ and $(2^{kd}-1, 2^m-1) = 1$ for every $d < N$ that divides N. By Lemma B, $(kN,m) \neq 1$ and $(kd,m) = 1$. The special case $d = 1$ shows that $(k,m) = 1$ and hence $(N,m) \neq 1$. Now $(N,m) \neq N$, since otherwise $d = (N,m)$ would give a contradiction. Hence m divides N. Since $(kd,m) = 1$ for any $d<N$ that divides N, N must be prime and $m = N$.

Theorem 10 If $w = -2^k$ and $M = 2^m-1$ for some positive integers k and m, where $m \geq 2$, then the convolution property holds (for some k) if and only if $N = 2m$ and m is an odd prime.

Proof Theorem 8 shows that the convolution property holds if $N = 2m$, m is an odd prime, and $k = 1$.

To prove the converse, we first consider the case where N is odd. Then Theorem 5 requires that $-2^{kN} \equiv 1 \pmod{2^m-1}$, or $2^{kN} \equiv -1 \pmod{2^m-1}$ but since $2^m \equiv 1 \pmod{2^m-1}$, the powers of 2 repeat $\pmod{2^m-1}$ and none is ever congruent to -1. Hence N must be even.

Theorem 5 requires that $2^{Nk}-1 \equiv 0 \pmod{2^m-1}$, and $(2^{kd}-1, 2^m-1) = 1$ when $d < N$, d is even and d divides N. By Lemma B, $(Nk,m) \neq 1$ and $(kd,m) = 1$, which implies that $(N,m) \neq 1$ and $(d,m) = 1$ for every appropriate d. This is impossible unless $N = 2m$ and m is an odd prime. ∎

11. Fields

Theorem 12. If R is a field, and w is of order N then the convolution property holds.

Proof To prove (c3), we notice that $(w^D-u)\sum_k w^{Dk} = w^{ND}-u = 0$ Since $w^D-u \neq 0$, this implies that $\sum_k w^{Dk} = 0$.

If the characteristic p of R is either infinite or does not divide N, then $N \times u \neq 0$ and $N' = (N \times u)^{-1}$ exists, which proves (c2). In other cases, consider the subfield generated by u and w. The characteristic is also p, and it is finite, since the fact that $w^N = u$ limits the number of distinct expressions in w. Hence its multiplicative group is of order p^n-1 for some positive integer n. Then N divides p^n-1 and p divides N, an obvious impossibility.

Since (c1) holds by hypothesis, the desired result follows from Theorem 1. ∎

Theorem 13 The ring of formal complex numbers with real and imaginary parts that are integers modulo M is a field if and only if M is an odd prime and M-1 is not divisible by 4.

Proof If M is composite, inverses do not always exist even for nonzero "real" numbers and the ring cannot be a field. If M is prime, then the ring is a field if and only if $\sqrt{-1}$ does not exist modulo M (see appendix D). If M = 2, then $\sqrt{-1} = \sqrt{1} = 1$. If $M \geq 3$ and M-1 is divisible by 4, then there is a number n with $n^4 \equiv 1 \pmod{M}$ but $n, n^2, n^3 \not\equiv 1 \pmod{M}$. Hence $n^2 = -1$ and $\sqrt{-1}$ exists. If $\sqrt{-1}$ exists, then it is of order 4 and M-1 is divisible by 4. ∎

Corollary The ring of formal complex numbers with real and imaginary parts that are integers modulo 2^m-1, where 2^m-1 is prime, are a field.

IV APPENDICES

APPENDIX A

Congruencies and Modular Arithmetic

The notation $a \equiv b \pmod{M}$, which is read "a is congruent to b, modulo M", means that b-a is either zero or a positive or negative multiple of M, which is called the modulus. Congruencies with the same modulus may be treated in many ways as though they were equations. For example, if $a \equiv b \pmod{M}$ and $c \equiv d \pmod{M}$, then $a+c \equiv b+d \pmod{M}$, $a-c \equiv b-d \pmod{M}$ and $ac \equiv bd \pmod{M}$. Any number may be substituted for one that is congruent to it, and numbers congruent to the same number are congruent to each other.

Modular arithmetic is simply arithmetic in which congruence is used instead of equality, and it is usually done only on integers. Division is not always possible in modular arithmetic when the modulus is composite. For example, $2x \equiv 5 \pmod{6}$ has no solution, and $2x \equiv 4 \pmod{6}$ has two distinct (noncongruent) solutions $x = 2$ and $x = 5$. But division is sometimes possible. For example $2 \cdot 5 \equiv 1 \pmod{9}$, so we can divide by 2 in this arithmetic by multiplying by 5. The solution to $2x-1 \equiv 5 \pmod{9}$ is found as in ordinary algebra

$$2x-1 \equiv 5 \pmod{9}$$
$$2x \equiv 6 \pmod{9}$$
$$5 \cdot 2x \equiv 5 \cdot 6 \pmod{9}$$
$$x \equiv 30 \pmod{9}$$

Of course, any other number congruent to 30 (such as 3) is an equally good solution, since the arithmetic does ont distinguish between them.

A general theorem states that one can divide by n in arithmetic modulo M is and only if the greatest common divisor of n and M, often written(n,m), is 1, that is, if n and M have no common divisor greater than 1. By convention (0,M) = M, so we cannot

divide by zero even in modular arithmetic. If M is prime, we can divide by any number that is not (congruent to) zero, and the modular arithmetic is as rich as that of the rational numbers, in a certain sense. A reciprocal in modular arithmetic, if it exists, can be found by the Euclidian algorithm, described in Appendix B.

If M is odd, then any integer is congruent modulo M to exactly one of the integers from -(M-1)/2 to (M-1)/2, inclusive. In fact, if the numbers always remain in this range, there is no difference between modular arithmetic and ordinary arithmetic on integers. Integer arithmetic is usually implimented this way on ones-complement machines. In this case $M = 2^m-1$, where m is the number of bits. Notice that the two representations of zero are congruent modulo M.

In most of our problems, the answers will be reduced to the range -(M-1)/2, (M-1)/2 by adding or subtracting appropriate multiples of M.

APPENDIX B

Greatest Common Divisors and the Euclidean Algorithm

The greatest common divisor (m,n) of two integers m and n, which are not both zero, is the largest positive integer that divides m and n exactly. For example, (3,0) = 3, (12,9) = 3, (5,-5) = 5, (2,7) = 1.

The Euclidean algorithm is a systematic way of finding (m,n). Since (-m,n) = (m,n) = (m,-n) = (-m,-n), we can assume both m and n are positive and arranged so that $n \geq m$. (Note that (o,n)=n.) Then divide n by m and let q_1 and r_1 be the quotient and remainder, respectively. Then $n = mq_1 + r_1$. Now any positive integer that divides m and n exactly also divides r_1 exactly, and any positive integer that divides m and r_1 exactly also divides n exactly. Hence $(m,n) = (r_1,m)$ and $m \geq r_1$. We repeat the process for (r_1,m). We obtain

$$
\begin{aligned}
n &= mq_1 + r_1, && r_1 < m,\\
m &= r_1q_2 + r_2, && r_2 < r_1,\\
r_1 &= r_2q_3 + r_3, && r_3 < r_2,\\
&\vdots\\
r_{n-2} &= r_{n-1}q_n + r_n, && r_n < r_{n-1},\\
r_{n-1} &= r_nq_{n+1}.
\end{aligned}
$$

Since the remainders keep getting smaller, eventually we get a zero remainder and the algorithm stops. Then $(m,n) = (r_1,m) = (r_2,r_1) = (r_3,r_2) = \ldots = (r_n,r_{n-1}) = (0,r_n) = r_n$. Also, if we write

$$
\begin{aligned}
r_1 &= n - mq_1,\\
r_2 &= m - r_1q_2,\\
r_3 &= r_1 - r_2q_3,\\
&\vdots\\
r_n &= r_{n-2} - r_{n-1}q_n,
\end{aligned}
$$

and substitute each equation into the following one, we obtain an equation of the form $(m,n) = r_n = am+bn$. In the special case where $(m,n) = 1$, $am \equiv 1 \pmod{n}$ and a is the reciprocal of m in modular arithmetic.

APPENDIX C

The Chinese Remainder Theorem

Suppose $M_1, M_2, \ldots, M_n$ are relatively prime, that is, no positive integer other than 1 divides more than one of them, and let $M = M_1 M_2 \ldots M_n$. Since $(M/M_i, M_i) = 1$, by appendix A there is an integer R_i, which can be found by the Euclidian algorithm, such that $R_i M/M_i \equiv 1 \pmod{M_i}$.

Now suppose we know that $N \equiv N_i \pmod{M_i}$. Then

$$N \equiv \sum_{i=1}^{n} (R_i M/M_i) N_i \pmod{M}.$$

To prove this, first note that $R_i M/M_i \equiv 0 \pmod{M_k}$ when $k \neq i$, and have

$$\sum_{i=1}^{n} (R_i M/M_i) N_i \equiv (R_k M/M_k) N_k \equiv N_k \equiv N \pmod{M_k}$$

That is, M_k divides $N - \sum_{i=1}^{n} (R_i M/M_i) N_i$ for every k. Since all the M_k are relatively prime, M also divides this expression, and the desired result follows.

APPENDIX D

Groups, Rings and Fields

A group is a set of elements, which are its "numbers" in a generalized sense, together with an operation on those elements, with the following properties. The result of the operation on x and y can be written as xy, which is the "multiplicative" notation, or x+y, which is the "additive" motation. (The latter is usually used for abelian groups.) The required properties are as follows. The unit u or 0 may not depend on x. The properties hold for all x, y, z in the group.

	multiplicative notation	additive notation
associativity	$(xy)z = x(yz)$	$(x+y)+z = x+(y+z)$
unit	$ux = xu = x$	$0+x = x+0 = x$
inverse	$x^{-1}x = xx^{-1} = u$	$-x+x = x+(-x) = 0$

The order of a group is the number of elements. The order of an element x is the smallest positive integer k such that $x^k = u$, where x^k is the product (under the group operative) of k x's. It can be shown that if the order of a group is finite, it is divisible by the order of each of its elements. If the order of an element is the same as that of the group, the element is called a generator.

An abelian group (or commutative group) is a group that obeys the commutative law xy = yx. Additive notation is often used for abelian groups.

A ring is a set of at least two elements, together with two operations, with the following properties. Additive notation is used for one operation and multiplicative notation is used for the other. The required properties are as follows. They must hold for all x, y, and z in the ring.

The ring is an abelian group under addition:

$(x+y)+z = x+(y+z)$

$0+x = x+0 = x$

$-x+x = x+(-x) = 0$

$x+y = y+x$

Distributive laws:

$x(y+z) = xy+xz$

$(x+y)z = xz+yz$.

Multiplicative associative law:

$(xy)z = x(yz)$

If there is an element u in the ring with $ux = xu = x$ for all x in the ring, we call u a unit. The integers modulo M are a ring with unit . (See appendix A).

If n is a positive integer, the n-th natural multiple of an element y in a ring is $n \times y = y+y+\ldots+y$, where the sum contains n terms. For other integers, we define $0 \times y = 0$ and $(-n) \times y = -(n \times y)$. If there is a unit, the set of all its natural multiples is a ring which is isomorphic to the integers or to the integers modulo M. In the latter case, M is said to be the characteristic of the ring. (In the former case the characteristic is said to be infinite) In fact, the characteristic is the order of u in the additive group.

The center of a ring is the set of all elements x such that $yx = xy$ for all y in the ring. If there is a unit, it is in the center. The center is itself a ring.

A ring in which $xy = yx$ for all x and y in the ring is called a commutative ring.

A commutative ring with unit u in which every nonzero element x has a reciprocal x^{-1} with $x^{-1}x = u$ is called a field. The most familiar fields are the fields of real numbers and the field of complex numbers, but there are others. The characteristic of a field, if it is finite, is a prime number.

If a field contains a finite number of elements, it is called a finite field or a Galois field. If the characteristic is p, then it can be shown that the field contains p^n elements for some positive integer n. It can also be shown that for every prime p and positive integer n, there is exactly one field (up to isomorphism) with p^n elements.

Every nonzero element x of a finite field with p^n elements is also an element of the multiplicative group consisting of the p^n-1 nonzero elements of the field. The order of x in this group, that is, the smallest positive integer k for which $x^k = u$, must divide p^n-1. The multiplicative group has a generator.

In the field of integers modulo the prime p, every nonzero element x has an order k that divides p-1. Hence

$$x^{p-1} = (x^k)^{(p-1)/k} = u^{(p-1)/k} = u$$

This gives Fermat's theorem: $x^{p-1} \equiv 1 \pmod{p}$ whenever $x \not\equiv 0 \pmod{p}$, or $x^p \equiv x \pmod{p}$ for any integer x and prime p.

If F is a field, then the formal complex numbers a+bi, where a and b are in F with the usual definitions of addition and multiplication are used, are clearly a commutative ring with unit. It is also a field if, and only if, $\sqrt{-1}$ did not exist in F.

APPENDIX E

Basic Properties of Convolutional Transforms

If a transform is convolutional and the transforms of $\{a_j\}$ and $\{b_j\}$ are $\{A_k\}$ and $\{B_k\}$, respectively, then--

(a) (Linearity) The transform of $\{sa_j + tb_j\}$, is $\{sA_k + tB_k\}$.

(b) (Rotation) The transform of $\{a_{N-1}, a_0, a_1, \ldots, a_{N-2}\}$ is $\{w^k A_k\}$.

(c) The transform of $\{w^j a_j\}$ is $\{A_1, A_2, \ldots, A_{N-1}, A_0\}$.

(d) (Evenness) If $a_{N-j} = a_j$, then $A_{N-k} = A_k$.

(e) (Oddness) If $a_{N-j} = -a_j$, then $A_{N-k} = -A_k$.

(f) If $N = QR$ and $a_j = 0$ except when j is 0 or a multiple of Q, then $A_{k+R} = A_{k+R-N} = A_k$.

(g) If $N = QR$ and $a_{j+R} = a_{j+R-N} = a_j$, then $A_k = 0$ except when k is 0 or a multiple of Q.

(h) (Parseval's equation)

$$N \sum_{j=0}^{N-1} a_j^2 = A_0^2 + \sum_{k=1}^{N-1} A_k A_{N-k},$$

$$N \left(a_0^2 + \sum_{j=1}^{N-1} a_j a_{N-j}\right) = \sum_{k=0}^{N-1} A_k^2 .$$

REFERENCES

1. R. C. Agarwal and C. S. Burrus, "Fast One-Dimensional Digital Convolution by Multidimensional Techniques", IEEE Trans. on Acoustics, Speech and Signal Processing, Vol. ASSP-22, No. 1, February 1974.

2. R. C. Agarwal and C. S. Burrus, "Fast Convolution Using Fermat Number Transforms with Applications to Digital Filtering", IEEE Trans. on Acoustics, Speech and Signal Processing, Vol. ASSP-22, No. 2, April 1974.

3. R. C. Agarwal and C. S. Burrus, "Number Theoretic Transforms to Implement Fast Digital Convolutions", Proc. of the IEEE, Vol. 63, No. 4, April 1975.

4. J. D. Brule, "Fast Convolution with Finite Field Fast Transforms", IEEE Trans. on Acoustics, Speech and Signal Processing, April 1975.

5. S. W. Golomb, I. S. Reed and T. K. Truong, "Integer Convolutions over a Finite Field", unpublished manuscript.

6. C. M. Rader, "Discrete Convolution via Mersenne Transforms", IEEE Trans. on Computers, Vol. C-21, No. 12, December 1972.

7. I. S. Reed and T. K. Truong, "The Use of Finite Fields to Compute Convolutions", IEEE Transactions on Information Theory, Vol IT-21, No. 2, March 1975.

8. I. S. Reed and T. K. Truong, "Complex Integer Convolutions over a Direct Sum of Galois Fields", unpublished manuscript.