SIAM Review
Vol. 59, No. 3, pp. 671–678

Elementary Number Theory and Rader's FFT[*]

Shlomo Engelberg[†]

Abstract. This note provides a self-contained introduction to Rader's fast Fourier transform (FFT). We start by explaining the need for an additional type of FFT. The properties of the multiplicative group of the integers modulo a prime number are then developed and their relevance to the calculation of the discrete Fourier transform is explained. Rader's FFT is then derived, Rader's zero-padding technique is described, and the performance of the unpadded and the zero-padded approaches is examined.

Key words. Rader's FFT, Cooley–Tukey FFT, zero-padding, $\mathbb{Z}_p$, $\mathbb{Z}_p^\times$

AMS subject classifications. 65T50, 11L03

DOI. 10.1137/15M1044990

1. The Cooley–Tukey FFT.

The discrete Fourier transform (DFT), $\{Y_m\}$, of a sequence, $\{y_k\}$, with N elements is defined as

$$Y_m = \sum_{k=0}^{N-1} \exp(-2\pi jkm/N)y_k, \qquad m = 0, \ldots, N-1.$$

A naive estimate of the number of operations necessary to calculate a DFT leads to the conclusion that the DFT should require on the order of N^2 calculations. The Cooley–Tukey fast Fourier transform (FFT) [3] is a divide-and-conquer algorithm for calculating the DFT of a sequence when the number of elements in the sequence, N, is a composite number, $N = I \cdot L$. It is based on rewriting the DFT as

$$\begin{aligned} Y_m = &\sum_{k=0}^{I-1} \exp(-2\pi jkm/I)y_{kL} + \exp(-2\pi jm/N) \sum_{k=0}^{I-1} \exp(-2\pi jkm/I)y_{Lk+1} \\ &+ \cdots + \exp(-2\pi j(L-1)m/N) \sum_{k=0}^{I-1} \exp(-2\pi jkm/I)y_{Lk+(L-1)}, \qquad N = I \cdot L, \end{aligned}$$

noting that each sum is now an I-term DFT that should require on the order of I^2 calculations, and noting that whenever you need terms that go past the "natural end" of a DFT you do not need to perform more calculations: instead, you find the additional terms by making use of the periodicity of the DFT. Calculating the DFT using this strategy takes on the order of $I^2L+(L-1)N$ operations, where the $(L-1)N$

[*] Received by the editors October 22, 2015; accepted for publication (in revised form) November 8, 2016; published electronically August 7, 2017.
http://www.siam.org/journals/sirev/59-3/M104499.html

[†] Department of Electrical and Electronics Engineering, Jerusalem College of Technology, Jerusalem, Israel (shlomoe@jct.ac.il).

operations are performed when adding up each of the smaller DFTs. As $I^2L = N^2/L$ and as $(L-1)N$ is on the order of N, one saves a fair amount of work by making use of this identity once. If I is itself composite, then the calculation of the smaller DFTs can be made more efficient by subdividing the calculation further.

The easiest case to handle is the one for which $N = 2^M$. In this case, by repeatedly splitting the DFTs, you can reduce the work from being of order N^2 to being of order $N\log(N)$—a huge savings. The problem with the Cooley–Tukey FFT is that when N is a prime number, it does not lead to any way to improve the efficiency of the calculation, and this is where Rader's FFT [6] comes in.

When *you* are picking the number of samples to be taken, the number of elements in the sequence, as a rule you should pick a power of two. Even if you did not pick the number of samples and the person who did (perhaps foolishly) chose to take a prime number of samples, all may not be lost. If the reason for calculating the DFT is to estimate the continuous-time Fourier transform, you can zero-pad the samples that have been taken until you have a convenient number of samples [4]. When you *must* calculate the true DFT of a sequence with a prime number of elements, Rader's FFT can be used to perform the calculation more efficiently.

As Rader's FFT makes essential use of some very pretty and elementary number theory, we develop that material now. Since this background material is interesting in its own right, we spend more time developing the background than might otherwise have been expected.

2. Some Elementary Number Theory.

2.1. The Field $\mathbb{Z}_p$. When considering the integers modulo a prime number, p, denoted by $\mathbb{Z}_p$, you find that every number, a, has an additive inverse, $-a$, and every nonzero number, b, has a multiplicative inverse, b^{-1}. For example, considering the integers modulo 7, $\mathbb{Z}_7$, and using $\equiv$ to denote equality modulo p, you find that

$$\begin{aligned}
&0+0=0 \Rightarrow -0 \equiv 0,\\
&1+6=7\equiv 0 \Rightarrow -1\equiv 6, -6\equiv 1,\\
&2+5=7\equiv 0 \Rightarrow -2\equiv 5, -5\equiv 2,\\
&3+4=7\equiv 0 \Rightarrow -3\equiv 4, -4\equiv 3,
\end{aligned}$$

and that

$$\begin{aligned}
&1\cdot 1 = 1 \Rightarrow 1^{-1}\equiv 1,\\
&2\cdot 4 = 8 \equiv 1 \Rightarrow 2^{-1}\equiv 4, 4^{-1}\equiv 2,\\
&3\cdot 5 = 15 \equiv 1 \Rightarrow 3^{-1}\equiv 5, 5^{-1}\equiv 3,\\
&6\cdot 6 = 36 \equiv 1 \Rightarrow 6^{-1}\equiv 6.
\end{aligned}$$

(The existence of additive inverses is clear. The existence of multiplicative inverses follows from the fact that there is no way to factor a prime number into factors that are smaller than the prime number.) The numbers $0,\dots,p-1$ form a field under addition and multiplication modulo p, and the field is denoted by $\mathbb{Z}_p$. The numbers $1,\dots,p-1$ form a group under multiplication modulo p; the group is denoted by $\mathbb{Z}_p^\times$.

2.2. Fermat's Little Theorem. Fermat's little theorem states that every nonzero element, x, of $\mathbb{Z}_p^\times$ satisfies the equation $x^{p-1}\equiv 1$. Proving this identity is simple. As every nonzero element of $\mathbb{Z}_p$, x, has a multiplicative inverse, the numbers $x\cdot 1$,

$x \cdot 2, \dots, x \cdot (p-1)$ must be the numbers $1, \dots, p-1$ in a (possibly) different order. Thus,

$$x^{p-1} 1 \cdots (p-1) \equiv 1 \cdots (p-1).$$

Canceling like terms (which is possible because every nonzero element has a multiplicative inverse) gives us the identity

$$x^{p-1} \equiv 1.$$

Thus, if you consider the sequence $x, x^2, \dots$ modulo p, you find that for all x the sequence has period $p-1$; the sequence may have a shorter period as long as that period divides $p-1$.

2.3. How Many Sequences of Length n Will There Be? Consider a value of n that divides $p-1$. Then $p-1 = nk$. It is easy to check that in this case

$$x^{p-1} - 1 = x^{nk} - 1 = (x^n - 1)\left(1 + x^n + \cdots + x^{(k-1)n}\right).$$

As the polynomial on the left has $p-1$ solutions, so must the product of polynomials on the right. As in any field a polynomial has no more solutions than its order, it is clear that there are n solution of $x^n - 1 = 0$ and $(k-1)n$ solutions of the second polynomial on the right for a total of $kn = p-1$ solutions. This is equivalent to saying that of the $p-1$ elements of $\mathbb{Z}_p^\times$ there are n elements that are associated with sequences of period n (or have periods that themselves divide n).

2.4. A Simple Example. Consider $\mathbb{Z}_7$. The elements of $\mathbb{Z}_7^\times$ are $1, \dots, 6$. All elements of this group satisfy $x^6 \equiv 1$, and from what we have just seen, three must satisfy $x^3 \equiv 1$, two must satisfy $x^2 \equiv 1$, and one must satisfy $x \equiv 1$.

Clearly, 1 satisfies all of the equalities. As $6^2 = 36 \equiv 1$, it satisfies $x^2 \equiv 1$. Thus, the two solutions of $x^2 \equiv 1$ are 1 and 6. As $2^3 = 8 \equiv 1$ and $4^3 = 64 \equiv 1$, the three solutions of $x^3 \equiv 1$ are $1, 2$, and 4. Clearly, 3 and 5 must have period 6 and no smaller periods. In fact, we find that $\{3, 3^2 \equiv 2, 3^3 \equiv 6(\equiv -1), 3^4 \equiv 4, 3^5 \equiv 5, 3^6 \equiv 1\}$ and that $\{5, 5^2 \equiv 4, 5^3 \equiv 6(\equiv -1), 5^4 \equiv 2, 5^5 \equiv 3, 5^6 \equiv 1\}$ have period 6, as they must.

2.5. A Definition, an Example, and a Theorem. A group is said to be *cyclic* if the powers of one or more of its elements give all the members of the group. We now consider the question, "Is the group $\mathbb{Z}_p^\times$ cyclic?" We start by considering an example: the group $\mathbb{Z}_{19}^\times$. We know that every nonzero element is associated with a sequence that is 18 elements long or whose period divides 18. We want to know if any of the elements are associated with a sequence whose period is *truly* 18.

Note that $18 = 18 \cdot 1 = 3 \cdot 3 \cdot 2 = 9 \cdot 2 = 3 \cdot 6$, that is, 18 is divisible by $1, 2, 3, 6$, and 9. Thus, there is one element with period 1, and there are two with period 2, three with period 3, six with period 6, and nine with period 9. There really is one element of order 1; it is $x \equiv 1$ whose associated series is $1, \dots, 1, \dots$. This sequence is 1-periodic, but it is also "any number periodic" and is counted as 2-periodic, 3-periodic, 6-periodic, and 9-periodic. Thus, there is only one sequence that is truly 2-periodic. It is the sequence associated with $-1 \equiv 18$. The 1-periodic sequence is also counted as a 3-periodic sequence, so there are two "genuine" 3-periodic sequences. All the 3-periodic sequences and the 2-periodic sequences are counted as 6-periodic. As there are four distinct sequences of periods 2 and 3, there are two "genuine" 6-periodic sequences. As the 3-periodic sequences are also counted as 9-periodic sequences, there are six "genuine" 9-periodic sequences. All the other sequences, $18 - 1 - 1 - 2 - 3 - 6 = 6$

sequences, must be genuinely 18-periodic. It is not hard to check that the elements associated with 18-periodic sequences are $2, 3, 10, 13, 14$, and 15.

This type of argument can be used to prove the following theorem, though there are easier ways of proving it [1].

THEOREM 1. *$\mathbb{Z}_p^\times$ is a cyclic group.*

That is, for any prime number there is always at least one element whose associated sequence has period $p-1$. This is equivalent to saying that powers of any such element give every element in $\mathbb{Z}_p^\times$. The fact that $\mathbb{Z}_p^\times$ is cyclic is *crucial* to the development of Rader's FFT.

3. Rader's FFT. In 1965, Cooley and Tukey published an article [3] in which they demonstrated that by reorganizing the calculations needed to perform a DFT one could make the computation much more efficient when the number of elements is highly composite. Their method works very well when the number of elements is, for example, a power of two. (Cooley and Tukey were not the first people to develop an FFT [2], but they are the modern proposers of the most commonly used FFT.) In 1968, Rader published an article [6] in which he showed that by reorganizing the computations in a different way, you could also make the calculation of the DFT more efficient when the number of elements is prime.

Following Rader, we make use of the fact that $\mathbb{Z}_p^\times$ is a cyclic group—that the group is generated by an element, a, and that for every value l between 1 and $p-1$ there exists a power, n, between 0 and $p-2$ for which $l \equiv a^n$. Also, note that if the powers of a generate all of the elements of $\mathbb{Z}_p^\times$, so do the powers of a^{-1} (which follows from the fact that $(a^n)^{-1} = (a^{-1})^n$ and inverses are unique).

Considering the definition of the DFT and, recalling that $a^l = a^l \bmod p + n_l p$ for some integer n_l, we find that

$$\begin{aligned}
Y_m &= \sum_{k=0}^{p-1} \exp(-2\pi jkm/p) y_k \\
&= y_0 + \sum_{l=0}^{p-2} \exp(-2\pi j(a^l \bmod p)m/p) y_{a^l \bmod p} \\
&\overset{\exp(-2\pi j n_l pm/p)=1}{=} y_0 + \sum_{l=0}^{p-2} \exp(-2\pi j(a^l \bmod p + n_l p)m/p) y_{a^l \bmod p} \\
&= y_0 + \sum_{l=0}^{p-2} \exp(-2\pi j a^l m/N) y_{a^l \bmod p}.
\end{aligned}$$

The only change here is to the order in which the elements are summed. If we now consider the value of $Y_{a^m \bmod p}$, following the same logic we find that

$$\begin{aligned}
Y_{a^m \bmod p} &= y_0 + \sum_{l=0}^{p-2} \exp(-2\pi j a^l a^m/p) y_{a^l \bmod p} \\
&= y_0 + \sum_{l=0}^{p-2} \exp(-2\pi j a^{l+m}/p) y_{a^{-(-l)} \bmod p}.
\end{aligned}$$

We have seen that the sequence $a^{-k} \bmod p$, whose elements are the inverses of the elements of a^k, sweeps out $\mathbb{Z}_p^\times$ and that the sequence is $p-1$ periodic. Let

$b_k = y_{a^{-k} \bmod p}$, and let $c_k = \exp(-2\pi j a^k/p)$. Then both sequences are $p-1$ periodic. Noting that

$$Y_{a^m \bmod p} - y_0 = \sum_{l=0}^{p-2} c_{l+m} b_{-l}$$

is the cyclic convolution of two sequences and recalling that the DFT of the cyclic convolution of two sequences is the DFT of the first times the DFT of the second [4], we find that

$$Y_{a^m \bmod p} - y_0 = \mathrm{IDFT}(\{C_n B_n\})(m),$$

where $\mathrm{IDFT}(\cdot)(m)$ is the m^{th} element of the inverse DFT of a sequence and where C_n and B_n are the DFTs of c_k and b_k. (Because the calculation of the IDFT is so similar to that of the DFT, an FFT algorithm can be made into an equally effective fast *inverse* DFT algorithm.) As $\{C_n\}$ can be precalculated, calculating the values of $Y_{a^m \bmod p}$ for $m = 0, \ldots, p-2$ takes about as long as calculating two DFTs of length $p-1$. This is worse than one finds with the Cooley–Tukey FFT, but if $p-1$ is highly composite (as is the case, for example, when $p = 257$ and $p-1 = 2^8$), it need not be too much worse. As p is a prime number (and greater than two), it is odd. Consequently, sequences with $p-1$ elements have an even number of elements, and it is possible to perform at least one iteration of the Cooley–Tukey type FFT.

You suffer from one more minor inconvenience when using this method: As no power of a gives 0, Y_0 must be calculated separately. Luckily, it is easy to calculate: $Y_0 = y_0 + \cdots + y_{p-1}$.

4. Trouble in Paradise.

4.1. Sophie Germain Primes and Cunningham Chains. A Sophie Germain prime is a prime, p, for which $2p+1$ is also prime. A Cunningham chain of the first kind is a sequence whose elements are all one greater than twice the previous element and are all prime [9]. Thus, all elements but the last element in a Cunningham chain are Sophie Germain primes. The numbers 2, 5, 11, and 23 are Sophie Germain primes, and the sequence 2, 5, 11, 23, and 47 is a Cunningham chain of the first kind.

Suppose that we would like to calculate the DFT of a sequence with 47 elements. We find that at each stage we must use the ideas behind the Rader transform. When calculating the 47-term DFT we find that we need to rephrase the problem as a cyclic convolution of sequences of length 46. This leads to the necessity of considering sequences of length 23. That leads to considering sequences of length 11, which leads to considering sequences of length 5, which leads to considering sequences of length 2. This is very resource intensive and presents us with a new problem to be overcome [10]. What should we do when the number of elements, the number of samples, is, for example, the last element of a Cunningham chain of the first kind?

5. The Solution: A Different Type of Zero-Padding. We have seen that when N is a prime number, it is possible to represent the calculations needed to calculate the DFT as a cyclic convolution and to use the ideas underlying the FFT to make this computation more efficient. Unfortunately, when dealing with the terminal element of a Cunningham chain, at every stage you must convert the calculation into another cyclic convolution, and only then can you start calculating DFTs. There is a better way [6].

Suppose that the sequences a_k and b_k are of length $p-1$ and that we would like to calculate the cyclic convolution of the sequences. We extend each sequence to a sequence whose length is $M \geq 2p-3$, where M is generally chosen to be a highly

Table 1 *Run times for several values of p. The first column gives the number of elements in the sequence. The following two give the number of factors of $p-1$ and the largest factor of $p-1$. The next two columns give the average run time when using Rader's FFT without zero-padding and when using zero-padding. The final column gives the average time to calculate the FFT when using the MATLAB* `fft` *command.*

p	Factors	Largest	Rader (μs)	Rader w. Z-P (μs)	MATLAB (μs)
3617	6	113	333	319	251
3793	6	79	314	325	245
4057	6	13	228	338	270
4201	7	7	244	551	210
4259	2	2129	479	575	280
4283	2	2141	479	573	280

composite number and perhaps even a power of two. We extend a_k to a sequence, c_k, whose first $p-1$ elements are a_k, whose last $p-2$ elements are also equal to $a_k, k = 1, \ldots, p-2$, and whose middle elements are all zero. We "promote" b_k to a sequence d_k whose first $p-1$ element are b_k and the rest of whose elements are zero. It is easy to see that the first $p-1$ elements of the cyclic convolution of c_k and d_k are identical to the elements of the cyclic convolution of a_k and b_k, and now all the DFTs and inverse DFTs that we need to calculate are of sequences with M elements, where M is highly composite. We are now in a very good position.

6. How Efficient Is It?

When working with a relatively simple set of algorithms to implement FFTs, one might have no choice but to use Rader's FFT with zero-padding, because simple implementations of the FFT might only implement the algorithm for powers of two. Modern implementations of MATLAB use the FFTW ("fastest Fourier transform in the West") package (http://www.fftw.org) [7, 8]. This package provides algorithms that try to determine the best way to calculate an FFT of a given size and then provides code with which to perform the calculation [5]. When working in such an environment, you pass the data to the FFT routine and allow it to decide how to proceed.

In order to understand how using or not using zero-padding affects the speed of computation, we wrote MATLAB code that calculated the FFT of a prime number of samples in several ways. (See section 7 for instructions about how to download the code.) The code takes the samples and uses Rader's FFT both without zero-padding and with zero-padding to perform the calculation. (When using zero-padding, the code zero-pads out to the smallest possible power of two.) The code uses the MATLAB `fft` and `ifft` commands to perform the smaller FFTs and IFFTs that Rader's algorithm requires and uses the `tic` and `toc` commands to see how long each calculation takes. It also calculates the FFT of the prime-length sequence using the MATLAB `fft` command. As Table 1 shows, using just the MATLAB `fft` command generally led to better results. As the MATLAB command is more highly optimized than our code and as it considers many additional ways of calculating the DFT, this is not surprising.

We now consider each line of Table 1 and try to understand when using zero-padding out to the smallest power of two is advantageous. Note that the power of two nearest the numbers of samples in the table is 4096—2^{12}. After zero-padding, the sequences will have either $2^{13} = 8192$ elements or $2^{14} = 16384$ elements.

6.1. $p = 3617$ and $p = 3793$.

When $p = 3617$ we find that $p-1$ has five relatively small prime factors and one relatively large one (113), which means that calculating its FFT should be moderately efficient. On the other hand, because 3617 is less than 4096, even after zero-padding the new sequence will only be a bit more than twice as long as the original one. As the table shows, using Rader's FFT with zero-padding is very moderately more efficient. It seems that the advantages of performing the FFT on a power-of-two number of samples outweigh the disadvantages of performing the FFT on a sequence with 8192 samples. Similar comments apply when $p = 3793$, except that this time not using zero-padding is mildly more efficient.

6.2. $\boldsymbol{p = 4057}$**.** When $p = 4057$, we find that all six of $p-1$'s factors are less than or equal to 13. In this case not using zero-padding should lead to a relatively efficient calculation, and Table 1 bears this out. Rader's method without zero-padding outperforms Rader's method with zero-padding.

6.3. $\boldsymbol{p = 4201}$**.** For $p = 4201$, $p-1$ has seven small prime factors. Thus, even without using zero-padding the FFT should be rather fast. Additionally, as 4201 is larger than 4096, after zero-padding the sequence will be almost four times as long as the original sequence. In this case, we expect not zero-padding to be much more effective than zero-padding, and this is borne out by Table 1.

6.4. $\boldsymbol{p = 4259}$ **and** $\boldsymbol{p = 4283}$**.** For both 4259 and 4283, we find that $p-1$ has only two factors. In such cases, we would expect zero-padding to show to advantage. However, in both cases, zero-padding to the smallest power of two causes the final sequence to have 16,384 samples—nearly four times as many samples as the unpadded sequence. Considering the values in Table 1, we find that not zero-padding works better, but the results when using zero-padding are almost as good despite the fact that the sequence on which the FFT is performed is nearly four times as long as the unpadded sequence.

7. Downloadable Code. When trying to understand an algorithm there is nothing like implementing it, and if time does not permit implementing it yourself and debugging the code, looking at someone else's implementation can be a real help. The MATLAB files we wrote to examine Rader's FFT can be found at

http://www.jct.ac.il/~shlomoe/Public/RaderFiles/DownloadRader.html

They should help you understand precisely what goes into Rader's FFT.

8. Concluding Remarks. We have seen that a little bit of very pretty number theory can go a long way. We made use of the fact that $\mathbb{Z}_p^\times$ is cyclic to develop Rader's FFT and demonstrated how Rader's FFT and a new type of zero-padding (also proposed by Rader) can be used to take what looks like a very hard problem, calculating the DFT of a sequence with a prime number of elements, and convert it to one of our favorite calculations, using the Cooley–Tukey FFT to calculate the DFT of a sequence with a highly composite number of elements.

REFERENCES

[1] M. Artin, 18.701 *Algebra I* (Massachusetts Institute of Technology: MIT OpenCourseWare), http://ocw.mit.edu (accessed July 31, 2015). (Cited on p. 674)

[2] J. Cooley, P. Lewis, and P. Welch, *Historical notes on the fast Fourier transform*, IEEE Trans. Audio Electroacoustics, 15 (1967), pp. 76–79. (Cited on p. 674)

[3] J. W. Cooley and J. W. Tukey, *An algorithm for the machine calculation of complex Fourier series*, Math. Comp., 19 (1965), pp. 297–301. (Cited on pp. 671, 674)

[4] S. Engelberg, *Digital Signal Processing: An Experimental Approach*, Springer, London, 2008. (Cited on pp. 672, 675)

[5] M. Frigo and S. G. Johnson, *The design and implementation of FFTW3*, Proc. IEEE, 93 (2005), pp. 216–231. (Cited on p. 676)

[6] C. M. Rader, *Discrete Fourier transforms when the number of data samples is prime*, Proc. IEEE, 56 (1968), pp. 1107–1108. (Cited on pp. 672, 674, 675)

[7] MathWorks, *Documentation: fft*, http://www.mathworks.com/help/matlab/ref/fft.html, last visited June 30, 2016. (Cited on p. 676)

[8] MathWorks, *Documentation: fftw*, http://www.mathworks.com/help/matlab/ref/fftw.html, last visited June 30, 2016. (Cited on p. 676)

[9] Wikipedia, *Cunningham Chain*, https://en.wikipedia.org/wiki/Cunningham_chain, last visited July 31, 2015. (Cited on p. 675)

[10] Wikipedia, *Rader's FFT Algorithm*, https://en.wikipedia.org/wiki/Rader's_FFT_algorithm, last visited August 18, 2015. (Cited on p. 675)