How the FFT Gained Acceptance

JAMES W. COOLEY

The fast Fourier transform (FFT) has had a fascinating history, filled with ironies and enigmas. Even more appropriate for this meeting and its sponsoring professional society,[1] it speaks not only of numerical analysis but also of the importance of the functions performed by professional societies.

The Role of Richard Garwin

My involvement with the FFT algorithm, or algorithms, as we should probably say, started when Dick Garwin[2] came to the computing center of the new IBM Watson Research Center sometime in 1963 with a few lines of notes he made while he was with John Tukey at a meeting of President Kennedy's Scientific Advisory Committee, of which they were both members. John Tukey showed that if N, the number of terms in a Fourier series, is a composite, $N = ab$, then the series can be expressed as an a-term series of subseries of b terms each. If one were computing all values of the series, this would reduce the number of operations from N^2 to $N(a + b)$. Tukey also said that if this were iterated, the number of operations could be reduced from N^2 to $N \log N$. Garwin not only had the insight to see the importance of this idea, but also had the drive to pursue its development and publication.

[1]See Acknowledgment, p.13.

[2]At that time, Garwin was a staff member of the Watson Scientific Laboratory at Columbia University. He is presently at IBM Watson Research Center, Yorktown Hts., NY.

Dick told me that he had an important problem of determining the periodicities of the spin orientations in a 3-D crystal of He^3. I found out later that he was also trying to find ways of improving the ability to do remote seismic monitoring in order to facilitate agreement with Russia on a nuclear test ban and to improve our capability for long-range acoustic detection of submarines. Like many others, I did not see the significance in this improvement and put the job on a back burner while I continued some research I considered more important. However, I was told of Dick Garwin's reputation, and, prodded by his occasional telephone calls (some of them to my manager), I produced a three-dimensional FFT program. I put some effort into designing the algorithm so as to save storage and addressing by overwriting data, and I spent some time working out a three-dimensional indexing scheme that was combined with the indexing within the algorithm.

The Decision to Publish

Garwin publicized the program at first by personal contacts, producing a small but increasing stream of requests for copies. I did a write-up and a version for a program library, but did not plan to publish right away. I gave a talk on the algorithm in one of a series of seminars in our mathematics department. Ken Iverson and Adin Falkoff, the developers of APL, participated, and Howard Smith, a member of the APL group, put the algorithm in APL when it was only a language for defining processes and before it was implemented on any machine. This gave the algorithm a thorough working-over at the seminar.

Another participant was Frank Thomas, a mathematically inclined patent attorney, who kept good contacts in the mathematics department. He suggested that there were patent possibilities and a meeting was called to decide what to do with it. It was decided that the algorithm should be put in the public domain and that this should be done by having Sam Winograd and Ray Miller design a device that could carry out the computation. My part of the strategy was to publish a paper with a footnote mentioning Miller and Winograd and their device. I sent my draft copy to John Tukey, asking him to be coauthor. He made some changes and emendations and added a few references to F. Yates, G. E. P. Box, and I. J. Good. Next came the task of getting it published as quickly as possible. I offered it to *Mathematics of Computation* by sending it to Eugene Isaacson at the Courant Institute of Mathematical Sciences, where I had worked before coming to IBM. I do not know how important my acquaintance with Eugene was or what effect it had on getting the paper published quickly. In any case, it appeared eight months after submission, in the April 1965 issue [1].

I found out later about an excellent paper by Gordon Sande, a bright statistics student of Tukey, who was exposed to the factorization idea in one of Tukey's courses. He carried the subject further, showing how it could be used to reduce computation in covariance calculations. After hearing about our paper going out to *Mathematics of Computation*, he did not publish his in its original form. However, he published several other excellent papers [2], one of which showed that the new algorithm was not only faster but more accurate than older techniques. His form of the FFT is now known as the Sande-Tukey algorithm.

Another result of Dick Garwin's efforts was a seminar run at the IBM Watson Research Center to publicize the algorithm and familiarize IBMers with it. For this, two capable statisticians, Peter D. Welch and Peter A. W. Lewis, joined me in writing a thick research report describing the algorithm and developing some theory and applications. The three of us then published a series of papers on applications of the FFT. These papers elaborated on the theory of the discrete Fourier transform and showed how standard numerical methods should be revised as a result of the economy in the use of the FFT. These included methods for digital filtering and spectral analysis [3].

The IEEE ASSP Digital Signal Processing Committee

The next level of activity came with contact with the speech and signal processing people at MIT — notably Thomas Stockham, Charles Rader, Alan Oppenheim, Charles Rabiner — all of whom have gone on to become highly renowned people in digital signal processing. They had developed digital methods for processing speech, music, and images. The great obstacle to making their methods feasible was the amount of computing required. This was the first really impressive evidence to me of the importance of the FFT. I was invited to join them and others on the Digital Signal Processing Committee of the IEEE Acoustics Speech and Signal Processing Society.

This committee ran the now famous Arden House Workshops on the FFT in 1968 [4] and in 1970 [5]. These were unique in several respects. One was that they collected people from many different disciplines: There were heart surgeons, statisticians, geologists, university professors, and oceanographers, just to name a few. The common interest was in the use of the FFT algorithms and every one of the approximately 100 attending had something useful to say in his or her presentation. People got together to formulate and work out solutions to problems. As an example, Norman Brenner, then of MIT, designed a program that computed the FFT of a sequence of interferometer data of 512,000 elements, which was larger than available high-speed storage. He did this for Mme. Connes of the University of Paris, who returned home to perform a monumental calculation of the infrared spectra of the planets, which has become a standard reference book [6]. Others worked out algorithms for data with special symmetries.

Meanwhile, back at the research center, I started learning the history of the FFT. Dick Garwin questioned his colleague, Professor L. H. Thomas of the Watson Scientific Laboratory of Columbia University, who has an office next to his. Thomas responded by showing a paper he published in 1963 [7]. His paper describes a large Fourier series calculation he did in 1948 on IBM punched card machines; a tabulator and a multiplying punch. He said that he simply went to the library and looked up a method. He found a book by Karl Stumpff [8] that was a cookbook of methods for Fourier transforms of various sizes. Most of these used the symmetries and trigonometric function identities to reduce computations by a constant factor. In a few places, Stumpff showed how to obtain larger transforms from smaller ones, and then left it to the reader to generalize. Thomas made a generalization that used mutually prime factors and got an efficient algorithm for his calculation.

The previously mentioned algorithms of Good and Thomas have some favorable properties, but the constraint that the factors are mutually prime does not give a number of operations proportional to or as low as $N \log N$. Tukey's form of the algorithm, with repeated factors, has the great advantage that a computer program need only contain instructions for the algorithm for the common factor. Indexed loops repeat this basic calculation and permit one to iterate up to an arbitrarily high N, limited only by time and storage.

The credit for what I would consider the first FFT — a computer program implementing this iterative procedure and really giving the $N \log N$ timing — should go to Philip Rudnick of the Scripps Institution of Oceanography in San Diego, California. He wrote to me right after the publication of the 1965 paper to say that he had programmed the radix 2 algorithm using a method published by Danielson and Lanczos in 1942 in the *Journal of the Franklin Institute* [9], a journal of great repute that publishes articles in all areas of science, but which did not enjoy a wide circulation among numerical analysts. Rudnick published some improvements in the algorithm [10] in 1966. I had the pleasure of meeting him and asked why he did not publish sooner. He said that his field was not numerical analysis and that he was only interested in getting a computer program to do his data analysis. Thus, we see another communication failure and lost opportunity, the primary point of Dick Garwin's 1969 Arden House keynote address [11].

Before continuing further with the discussion of the old literature on the FFT, I would like to point out two important concepts in numerical algorithms that had been stated long ago but did not have very much impact until they were demonstrated by the implementation of the FFT on electronic computers. The first is the divide and conquer approach. If a large N-size problem requires effort that increases like N^2, then it pays to break the problem into smaller pieces of the same structure. The second important concept is the asymptotic behavior of the number of operations. Obviously this was not significant for small N, and by habit of thought, people failed to see the importance of early forms of the FFT algorithms even where they would have been useful.

I can illustrate this point by going back to the Danielson and Lanczos paper [9]. The authors describe the numerical problem of computing Fourier coefficients from a set of equally spaced samples of a continuous function. One is faced not only with a long, laborious calculation, but also with the problem of verifying accuracy. Errors can arise from mistakes in computing or from undersampling the data. Lanczos pointed out that although his use of the symmetries of the trigonometric functions, as described by Runge, reduced computation by a significant factor, one still had an N^2 algorithm. In a previous reading of this paper, I obtained and published [12] the mistaken notion that Lanczos got the doubling idea from Runge. In fact, he only attributes the use of symmetries to Runge, citing papers published in 1903 and 1905 that I could not find. The Stumpff paper [8] gave a reference to Runge and König [13] that does contain the doubling algorithm and that appears to have been a standard textbook in numerical analysis. Thus, it appears that Lanczos independently discovered the clever doubling algorithm and used it to solve the problems of computational economy and error control. He says, in the introduction to [9] on page 366, "We shall show that, by a certain transformation process, it is possible to double the number of ordinates with only slightly more than double the labor." He goes on to say [9]:

"In the technique of numerical analysis the following improvements suggested by Lanczos were used: (1) a simple matrix scheme for any even number of ordinates can be used in place of available standard forms; (2) a transposition of odd ordinates into even ordinates reduces an analysis for $2n$ coefficients to two analyses for n coefficients; (3) by using intermediate ordinates it is possible to estimate, before calculating any coefficients, the probable accuracy of the analysis; (4) any intermediate value of the Fourier integral can be determined from the calculated coefficients by interpolation. The first two improvements reduce the time spent in calculation and the probability of making errors, the third tests the accuracy of the analysis, and the fourth improvement allows the transform curve to be constructed with arbitrary exactness. Adopting these improvements the approximation times for Fourier analyses are 10 minutes for 8 coefficients, 25 minutes for 16 coefficients, 60 minutes for 32 coefficients, and 140 minutes for 64 coefficients."

The matrix scheme in (1) in the preceding quotation reduces the data to even and odd components so that real cosine and sine transforms are computed. The rest of the process makes use of the symmetries of the sines and cosines, similar to the methods of Runge. After this, Lanczos uses the doubling algorithm. In step (2) he uses what we have been calling the twiddle factor multiplication, and in step (3) he does the butterfly calculation but observes accuracy by comparing the two inputs: the Fourier coefficients of the sub-series. Thus, it appears that Lanczos had the FFT algorithm; and if he had had an electronic computer, he would have been ready to write a program permitting him to go to arbitrarily high

N. It may seem strange to us, then, to see his remark on page 376, "If desired, this reduction process can be applied twice or three times."

This is an outstanding example of the difference in point of view between different generations of numerical analysts. Here was the doubling algorithm, capable of doing Fourier transforms in *N* log *N* operations, described in detail. It seems to be appreciated as much as a method for checking accuracy as for reducing computing. The authors did not foresee the possibility of automating the procedure. In fact, in the beginning of the Danielson and Lanczos paper, it is presented as an economical way of doing the computation without using a mechanical analyzer available at the time. Then they published it in the *Journal of the Franklin Institute*, where it was unnoticed until Philip Rudnick, who was not a numerical analyst, revived it but ignored the opportunity to show it to the world. Lanczos later published his *Applied Analysis* in 1956 [14] with only a few words and a footnote (page 239) referring to the Danielson and Lanczos paper. I find no references to it at all in his later books, including his 1966 book, *Discourse on Fourier Series* [15].

Gauss and the FFT

After learning of the above early papers, I wrote what I thought to be the very early history of the FFT algorithm [12], going back to Runge and König. Some years later, while working on his book [16], Herman Goldstine told me of a paper by Gauss [17] that contained the FFT algorithm. I got a copy of the paper, which was in a neoclassic Latin that I could not read. The formulas and a slight recognition of parts of words indicated he was doing a kind of Lagrangian interpolation that leads to the basic FFT algorithm. I put this aside as an interesting post-retirement activity.

A few years later, some old signal processing friends, Don Johnson and Sidney Burrus at Rice University, told me that they put a bright, energetic graduate student, Michael Heideman, on the trail of Gauss and the FFT. He not only translated the Gauss article but found and described many others who wrote of FFT methods after Gauss and before my early references [18].

Conclusion

This story of the FFT can be used to help us appreciate the important functions of professional societies such as the ACM and SIAM. The following are some recommendations:

- Prompt publication of significant achievements is essential.
- Reviews of old literature can be rewarding.
- Communication among mathematicians, numerical analysts, and workers in a wide range of applications can be fruitful.
- Do not publish papers in neoclassic Latin.

Acknowledgment

This article is reprinted by permission from *Proceedings of ACM Conference on the History of Scientific and Numeric Computation*, May 1987, and from *A History of Scientific Computing*, ACM Press/Addison-Wesley.

References

[1] J. W. Cooley and J. W. Tukey, "An Algorithm for the Machine Calculation of Complex Fourier Series," *Math. Comp.* 19 (165): 297.

[2] W. M. Gentleman and G. Sande, "Fast Fourier Transforms for Fun and Profit," in *1966 Fall Joint Computer Conf.*, AFIPS Proc. Vol. 29 (Washington, DC; Spartan Press, 1966): 563-78.

[3] J. W. Cooley, P.A.W. Lewis, and P.D. Welch, "The Application of the Fast Fourier Transform Algorithm to the Estimation of Spectra and Cross-Spectra," *J. Sound Vibration* 12(3) (1970): 339-52.

[4] Various authors, "Special Issue on Fast Fourier Transform and Its Application to Digital Filtering and Spectral Analysis," *IEEE Trans. Audio and Electroacoustics* AU-15 (1967): 43-117.

[5] Various authors, "Special Issue on Fast Fourier Transform," *IEEE Trans. Audio and Electroacoustics* AU-17 (1969): 65-186.

[6] J. Connes, P. Connes, and J. P. Maillard, "Atlas des Spectres dans le Proch Infrarouge de Vênus, Mars, Jupiter et Saturne," *Éditions du Centre de la Recherche Scientifique* 15, quai Anatole France Paris VII (1969).

[7] L. H. Thomas, "Using a Computer to Solve Problems in Physics," in *Applications of Digital Computers*, W. F. Freiberger and W. Prager (eds.) (Boston: Ginn and Company, 1963).

[8] Karl Stumpff, *Grundlagen und Methoden der Periodenforschung* (Berlin: Springer, 1937); *Tafeln und Aufgaben zur Harmonischen Analyse und Periodogrammerchnung* (Berlin: Springer, 1939).

[9] G. C. Danielson and C. Lanczos, "Some Improvements in Practical Fourier Analysis and Their Application to X-ray Scattering From Liquids," *J. Franklin Inst.* 233, Pergamon Journals, Ltd. (1942): 365-80, 435-52.

[10] Philip Rudnick, "Note on the Calculation of Fourier Series," *Math. Comp.* 20 (1966): 429-30.

[11] R. L. Garwin, "The Fast Fourier Transform as an Example of the Difficulty in Gaining Wide Use for a New Technique," Special Issue on Fast Fourier Transform, *IEEE Trans. Audio and Electroacoustics*, AU-17 (1969): 69-72.

[12] J. W. Cooley, P. A. W. Lewis, and P. D. Welch, "Historical Notes on the Fast Fourier Transform," *IEEE Trans. Audio and Electroacoustics*, AU-15 (1967): 76-79.

[13] C. Runge and H. König, "Band XI, Vorlesungen Über Numerisches Rechnen," *Grundlehren Math. Wiss.* (Berlin: Verlag von Julius Springer, 1924).

[14] C. Lanczos, *Applied Analysis* (Englewood Cliffs, NJ: Prentice Hall, 1956).

[15] C. Lanczos, *Discourse on Fourier Series* (Edinburgh and London: Oliver and Boyd, 1966).

[16] H. H. Goldstine, *A History of Numerical Analysis from the 16th Through the 19th Century* (New York: Springer-Verlag, 1977), 249-53.

[17] C. F. Gauss, "Nachlass: Theoria interpolationis methodo novo tractata," in *Carl Friedrich Gauss, Werke*, Band 3 (Gottingen; Koniglichen Gesellschaft der Wissenschaften, 1866), 265-303.

[18] M. T. Heideman, D. H. Johnson, and C. S. Burrus, "Gauss and the History of the Fast Fourier Transform," *The IEEE ASSP Magazine* I (1984): 14-21.