

FOURIER AND WAVELET REPRESENTATIONS OF FUNCTIONS

NICHOLAS G. ROLAND

ABSTRACT. Representations of functions are compared using the traditional technique of Fourier series with a more modern technique using wavelets. Under certain conditions, a function can be represented with a sum of sine and cosine functions. Such a representation is called a Fourier series. This classical method is used in applications such as storage of sound waves and visual images on a computer. One problem with this sum is that it is infinite. In use, only a finite number of terms can be used. More accuracy requires more terms in the series, but more terms require more time to compute and more space to store. A new type of sum called a wavelet series was first introduced in the 1980's. With these new series the same accuracy often takes fewer terms. Since wavelet representations can be more accurate and take less computer time, they are often more useful.

1. BACKGROUND IN FOURIER SERIES

Jean Baptise Joseph Fourier (1768–1830) was the inventor of Fourier series in the late 1700's. Fourier was a mathematical physicist who developed a way to express a function by combining an infinite number of sine and cosine terms. For example a tuning fork produces a sound wave when it is struck. The sound wave that we hear is a pure tone with one frequency and can be represented by a single sine function. When we hear a piano key struck, we do not hear just one frequency but rather a mixture of frequencies. This mixture contains a fundamental and then a number of overtones (harmonics) of frequency $2, 3, 4, \dots$ multiplied by the fundamental frequency. The combination of the fundamental and the harmonics can be represented by a sum of sine functions. The series which describes this combination is called a Fourier series. Expanding a function in a Fourier series breaks the function down into its various harmonics. More generally, a $2l$ -periodic function $f(x)$, can be represented by its Fourier series,

$$f(x) \sim \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi x}{l}\right) + \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi x}{l}\right),$$

Received by the editors May 9, 2000.

2000 *Mathematics Subject Classification.* 42C40, 42A16.

Key words and phrases. Fourier series, Wavelets.

This paper was written while the author was an undergraduate student at the University of Wisconsin - La Crosse. The paper is also published in the University of Wisconsin - La Crosse Journal of Undergraduate Research.

where the a_n 's and b_n 's are the **Fourier Series coefficients** defined by

$$a_n = \frac{1}{l} \int_{-l}^l f(x) \cos\left(\frac{n\pi x}{l}\right) dx \quad \text{for } n = 0, 1, 2, \dots$$

and

$$b_n = \frac{1}{l} \int_{-l}^l f(x) \sin\left(\frac{n\pi x}{l}\right) dx \quad \text{for } n = 0, 1, 2, \dots \quad .$$

There are some problems with representing functions with this sort of series. The main problem is that it takes an infinite number of terms to represent a function and in practice one can only use a finite number of terms. In applications this Fourier series must be truncated, and this truncation will introduce error since the whole sum is no longer being used. Thus, one must try to find a balance between the number of terms one uses and how much error one is willing to accept. In order to achieve more accuracy, a greater number of terms are needed and this will take up more computer time and storage space.

A second problem with Fourier series is that although it represents the frequency of a function well, it does not do as good a job representing that functions localized properties. For example, a function may contain a high peak on an interval while it is small elsewhere. This function could represent a wave packet, which is just a peak traveling from one point to another in a straight line. Before and after the peak the amplitude is almost zero, as shown in Figure 1. A Fourier series will not do as well when representing this function because the sine and cosine functions, that make up the Fourier series, are all periodic and thus it is hard to focus in on the local behavior of this wave packet. We will see how this is dealt with later.

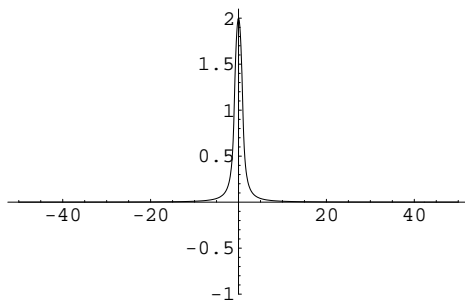


FIGURE 1. Example of wave traveling from left to right.

These were problems that mathematicians as well as physicists had to deal with until the 1980's when a new type of series called wavelet series was introduced. These new series were developed to get around some of the problems associated with Fourier series. For example the reason it take Fourier series many terms to represent localized properties is because they depend on a single basis (sines and cosines) which represent frequencies well but whose support is not localized. Wavelet series give us an infinite number of new basis to choose from so we can choose the "best basis" for a function. Also, wavelets are designed to "zoom in" on localized behavior. Thus wavelet representations may take fewer terms to represent the same function to the same accuracy as the Fourier series.

2. APPLICATIONS FOR WAVELETS

The need to represent functions with fewer terms is one that is growing since more technology is based on moving information from one place to another. The cheaper and faster information can be moved while preserving the accuracy, the better the technology works. This motivated mathematicians to search for new bases. Starting in the 1980's, wavelets were developed to give us new bases. Their structure, see Section 3, also allows them to better represent local behavior of functions. Since wavelets often allow us to reconstruct functions with fewer terms, saving time and storage space, while increasing the accuracy of the reconstruction, they have many applications.

Since wavelets can focus on local behavior, they are often used in medical image processing when looking for potential tumors. They are also used in radar and sonar imaging for oil prospecting to identify the boundary of an oil pocket, or in archaeology to locate artifacts.

Wavelets could also be used in video image analysis. Video telephones are not in high demand now because a high-quality sequence of video images cannot be transmitted along a phone line in real time. This is because such a sequence of images exceeds the phone line's capacity. If video images could be compressed or represented with smaller data sets, without serious destruction to the image, video telephones may become practical. Wavelets could be the answer to this application [Fr].

Wavelets help compress video images for the following reasons. For video images used in the television, the different frames usually only differ slightly from one another. For example, the frames of the background of a television show often do not change. Maybe only a person's arm is moving from frame to frame, so instead of transmitting the entire new frame, only the differences in the frames must be transmitted. If wavelets were used, the wavelet coefficients where the frame stays the same would not change and thus would not have to be updated. The only updating would be for a small number of coefficients near the arm. The updating could be done with fewer data bits so there would be a higher compression rate. Wavelets can do this since they can "zoom in" on specific areas of the image. For more details on the above applications see [Fr].

Recently, wavelets were used to solve another data compression problem, the digital fingerprint. FBI finger prints should be computerized so that local police could have access to them. If there was a way to store fingerprints digitally, then authorities could search for a match electronically. In order to store one person's set of finger prints, that is 10 rolled fingerprints, plus 2 unrolled thumb impressions and 2 impressions of all 5 fingers on a hand, it would require about 10 megabytes of data. This does not sound so bad at first, but then one would have to multiply that by 200 million in order to account for all the people the FBI has fingerprints for. This would require about 2,000 terabytes (a terabyte is 10^{12} bytes). One solution to this is to restrict the number of fingerprints to just the current criminals or approximately 29 million, it would still take 60,000 3-gigabyte hard drives to store. If that amount could be cut by a factor of 20, these same 29 million fingerprint cards could be stored on approximately 3,000 3-gigabyte hard drives. This is a much more realistic number. A group from Los Alamos National Laboratory was given a contract to do just this. This group compressed this data by a factor of 20 by using wavelets. A picture of an original fingerprint along with its wavelet

reconstruction is shown in Figure 2. For more details on this application, see [Br1] and [Br2].

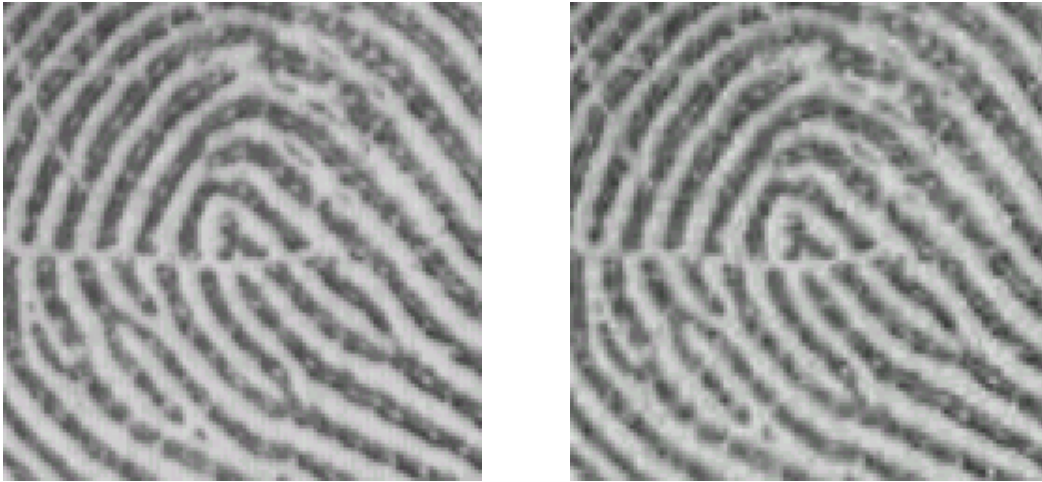


FIGURE 2. Left: Original Fingerprint. Right: Wavelet reconstruction. Images courtesy of C. Brislawn, Los Alamos National Lab.

3. WAVELETS AND THEIR STRUCTURE

Some of the examples in the previous section dealt with two-dimensional images. For this work we focused on functions in only one variable. In order to understand how wavelets can represent localized properties better, we will examine wavelets defined in $L^2(\mathbf{R})$, $L^2(\mathbf{R}) = \{f : \int_{-\infty}^{\infty} |(f(x))|^2 dx < \infty\}$. For simplicity, we can view this set as a set of functions $f(x)$ which get small in magnitude quite rapidly as x goes to plus and minus infinity.

A general structure for wavelets in $L^2(\mathbf{R})$ is called a Multiresolution Analysis [Ma]. This structure is a bit technical, but it will be illustrated afterwards with an example which should help clarify the ideas.

We begin with a family of closed “nested” subspaces

$$\dots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \dots$$

in $L^2(\mathbf{R})$, where

$$\bigcap_{j \in \mathbf{Z}} V_j = \{0\}, \quad \overline{\bigcup_{j \in \mathbf{Z}} V_j} = L^2(\mathbf{R})$$

and

$$f \in V_j \iff f(2 \cdot) \in V_{j+1}.$$

If these conditions are met, then there exists a function $\phi \in V_0$ such that $\{\phi_{j,k}\}_{k \in \mathbf{Z}}$ is an orthonormal basis of V_j , where

$$\phi_{j,k}(x) = 2^{j/2} \phi(2^j x - k).$$

In other words, the function ϕ , called the **father function**, will generate an orthonormal basis for each V_j space.

We then define W_j such that $V_{j+1} = V_j \oplus W_j$. Recall that $V_j \subset V_{j+1}$. This says that W_j is the space of functions in V_{j+1} but not in V_j . Thus, $L^2(\mathbf{R}) = \sum \oplus W_j$. Then there exists a function $\psi \in W_0$ such that $\{\psi_{j,k}\}_{k \in \mathbf{Z}}$ is an orthonormal basis of W_j , and $\{\psi_{j,k}\}_{j,k \in \mathbf{Z}}$ is a wavelet basis of $L^2(\mathbf{R})$, where

$$\psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k).$$

The function ψ is called the **mother function**, and the $\psi_{j,k}$'s are the wavelet functions.

For any function $f \in L^2(\mathbf{R})$, a projection map of $L^2(\mathbf{R})$ onto V_j ,

$$P : L^2(\mathbf{R}) \rightarrow V_j$$

is defined by

$$(1) \quad Pf(x) = \sum_{j=-\infty}^{m-1} \sum_{k=-\infty}^{\infty} \langle f, \psi_{j,k} \rangle \psi_{j,k}(x)$$

$$(2) \quad = \sum_{k \in \mathbf{Z}} \langle f, \phi_{m,k} \rangle \phi_{m,k}(x),$$

where the $\langle f, \psi_{j,k} \rangle = \int_{-\infty}^{\infty} f(x) \psi_{j,k}(x) dx$ are the wavelet coefficient and the $\langle f, \phi_{m,k} \rangle = \int_{-\infty}^{\infty} f(x) \phi_{m,k}(x) dx$ are the scaling coefficient. The sum in (1) is a truncated wavelet series. If j were allowed to go to infinity, we would have the full wavelet sum. Series (2) gives an equivalent sum in terms of the scaling functions $\phi_{m,k}$. As we use a higher m value, which means more terms are used, the truncated series representation of our function improves. In other words, there will be less error if we have a higher m value in the sum. We will come back to these sums when we compare wavelet series with Fourier series in Section 4.

To illustrate the multiresolution structure we will next look at the Haar wavelets. For these wavelets the space V_j is the set of all $L^2(\mathbf{R})$ functions which are constant on each interval $[2^{-j}n, 2^{-j}(n+1))$ for all integers n . Then

$$\phi(x) = \begin{cases} 1, & \text{if } 0 \leq x < 1 \\ 0, & \text{else} \end{cases}$$

and

$$\psi(x) = \begin{cases} 1, & \text{if } 0 \leq x < \frac{1}{2} \\ -1, & \text{if } \frac{1}{2} \leq x < 1 \\ 0, & \text{else} \end{cases}.$$

See Figure 3.

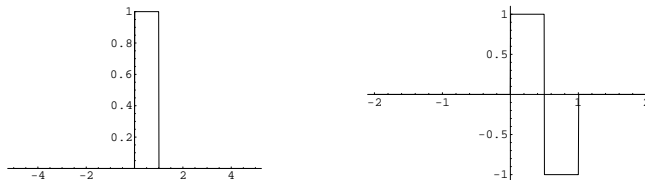


FIGURE 3. Left function: ϕ . Right function: ψ .

The unique thing about using wavelets as opposed to Fourier series is that the wavelets can be moved, stretched, or compressed to accurately represent a functions local properties. This is done by using different j and k values. We will illustrate this with the scaling functions ϕ_{jk} . Recall that $\phi_{jk} = 2^{\frac{j}{2}}\phi(2^jx - k)$, so the ϕ_{jk} is nonzero only on the interval $\frac{k}{2^j} \leq x < \frac{k+1}{2^j}$. The j value gives the function a height of $2^{\frac{j}{2}}$ and it either stretches or compresses (dilates) the function with the $2^{\frac{j}{2}}$ factor. The k value moves (translates) the function a distance from the origin. This is demonstrated in Figure 4. In Figure 4 the ϕ function was used, however the ψ function is translated and dilated in the same fashion.

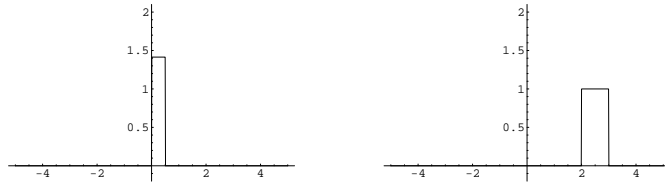


FIGURE 4. Left function: ϕ_{10} . Right function: ϕ_{02} .

4. FOURIER SERIES AND WAVELET REPRESENTATIONS OF FUNCTIONS

The main goal of our work was to compare how well functions are represented using Fourier series and wavelets. To make these comparisons, we compared original functions with finite Fourier series and with finite wavelet sums. We measured errors between the original function and the representation.

The first error examined was a maximum error, defined to be the maximum difference between the original function and the finite series representation divided by the average value of the original function. That is,

$$\text{maximum error} = \frac{\text{maximum difference}}{\text{average value of function}} \cdot 100\%.$$

Here we divided by the average value of the function so that we could compare the size of the difference to the size of the function.

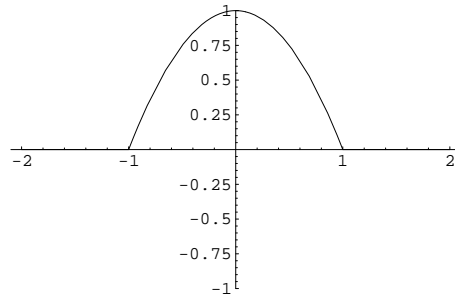
The other error that was studied was an average error, defined to be the area of the difference between the two functions divided by the average value of the original function. That is,

$$\text{average error} = \frac{\text{average value of difference}}{\text{average value of function}} \cdot 100\%.$$

The first function that was investigated was the function shown in Figure 5 and defined as

$$(3) \quad f(x) = \begin{cases} -x^2 + 1, & \text{if } -1 \leq x \leq 1 \\ 0, & 1 < |x| < 10 \end{cases}.$$

This function was chosen for investigation because it exhibits localized behavior, something that we were interested in. For the Fourier series, we looked at a $2(10)$ -periodic extension of this function. For the wavelet expansion we let $f(x) = 0$ elsewhere. We compared the two functions on the interval $[-10, 10]$.

FIGURE 5. $f(x)$ defined in Equation (3).

Fourier series representation of this function were first examined. Several finite sums of Fourier series were computed and graphed with the original function. In Figure 6, fifty nonzero terms from the Fourier series were used to reconstruct the function. Notice how the representation has a “tail” on both ends. That is, the reconstruction has decreasing magnitude where the original function was zero. Since Fourier series use a combination of sine and cosine terms which are periodic, and hence do not decrease in magnitude as they approach infinity, such a series takes many terms to be able to clearly represent some local features in functions. The maximum error when the first 50 nonzero terms were used was 486.31%, and the average error was 98.84%. These errors appear large since we are scaling them with the average value of the function defined over the interval $[-10, 10]$.

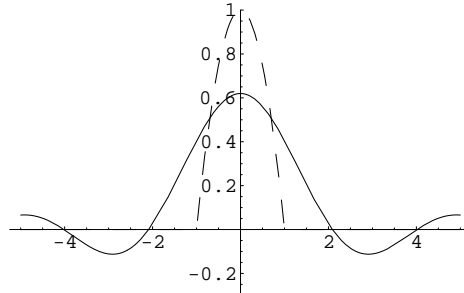


FIGURE 6. Original Function (dashed) and Fourier Series Representation with 50 nonzero terms

In order to decrease the size of the “tail” for the Fourier Series representation and in general get closer to the original function, more nonzero terms must be added. In the Figure 7 the Fourier Series representations for the first 75 and 100 nonzero terms respectively, along with the original function were graphed together. The maximum errors are, 389.95% and 240.75% respectively, and the average errors are 64.60% and 34.12% respectively.

As discussed above, as more nonzero terms are added, the “tail” is being decreased and the overall reconstruction improves. In order to achieve a representation with a small error, many nonzero terms are needed. When more terms are added, the computer time it takes to generate these approximations increases.

The same function was next represented with Haar wavelets, defined in the Section 3. The sums presented in this paper are sums of scaling functions as given

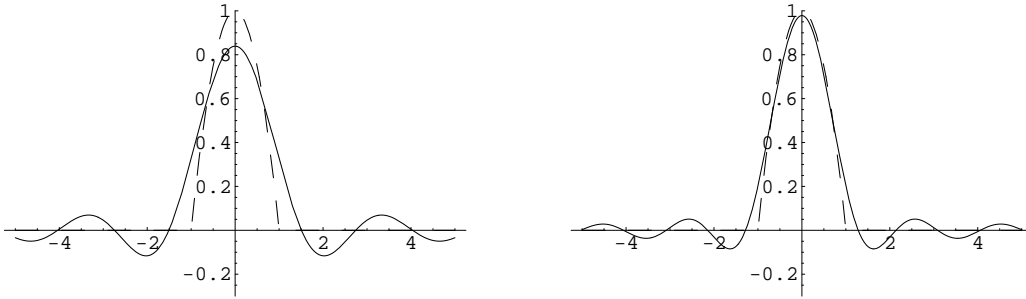


FIGURE 7. Left: 75 nonzero terms. Right: 100 nonzero terms.

in Equation (2). We used Equation (2) for simplicity because we could fix the m value and then sum over the k values instead of summing over both the j values and k values, as given in Equation (1). Recall that the scaling functions for the Haar wavelets at a fixed level are simply rectangles with a constant width.

For the first representation, shown in Figure 8, the m value was set to be three. In examining the interval $[-10, 10]$, this produced 16 nonzero terms where the k value went from $-2^m \leq k \leq 2^m - 1$. The maximum error was 18.08% and the average error was 2.37%.

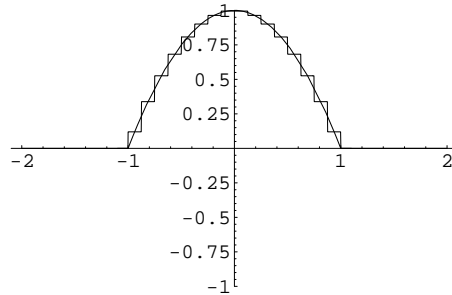


FIGURE 8. 16 nonzero terms, m value fixed at 3.

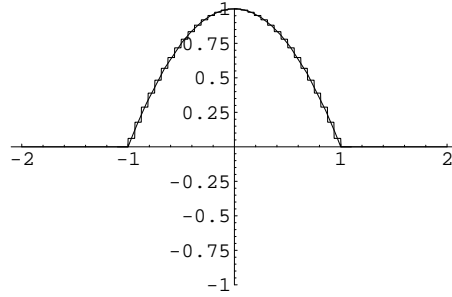
When 50 nonzero terms were used with Fourier series, the maximum error was 486.31% and the average error was 98.84%. As you can see, the scaling functions are already doing a much better job with significantly fewer terms.

In the second representation, shown in Figure 9, the m value was set to be four. This produced 32 nonzero terms. The maximum error was 9.18% and the average error was 1.19%. The Fourier series representation took 4000 terms to get a maximum error of 6.08%, which is comparable to this representations error.

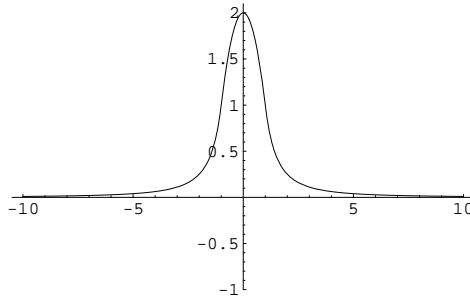
A second function,

$$(4) \quad f(x) = \begin{cases} -x^2 + 2 & \text{if } -1 \leq x \leq 1 \\ \frac{1}{x^2} & \text{if } 1 < |x| < 10 \end{cases},$$

was next examined. Again, for the Fourier series, we looked at a $2(10)$ -periodic extension of this function. For the wavelet expansion we let the function $f(x) = \frac{1}{x^2}$ elsewhere. We compared the two functions on the interval $[-10, 10]$. This function

FIGURE 9. 32 nonzero terms, m value fixed at 4.

is similar to the previous function but outside the interval of $[-1,1]$, it decays like $\frac{1}{x^2}$ instead of simply going to zero. A picture of this function is shown in Figure 10.

FIGURE 10. $f(x)$ defined in Equation(4).

Fourier series representation of this function were first examined. Several finite sums of Fourier series were computed and graphed with the original function. In Figure 11, 50 nonzero terms from the Fourier series were used to reconstruct the function. The maximum error when the first 50 nonzero terms were used was 133.55%, and the average error was 19.20%.

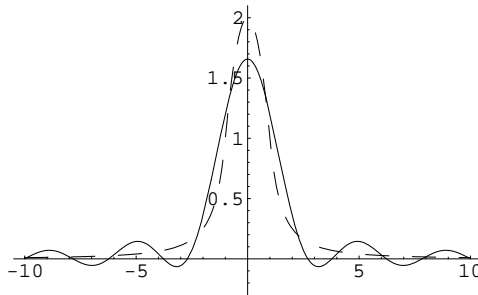


FIGURE 11. Fourier Series representation along with original function.

With the scaling function, shown in Figure 12, the m value was set to be zero. This produced 20 nonzero terms when the k value went from $-10 \cdot 2^m \leq k \leq 10 \cdot 2^m - 1$. The maximum error was 40.00% and the average error was 22.88%.

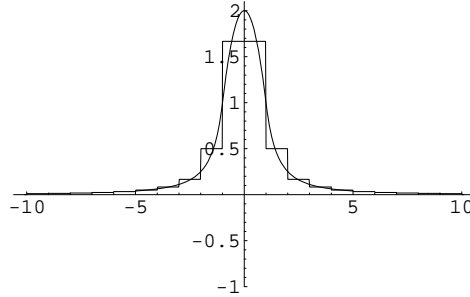


FIGURE 12. 20 nonzero terms, m value fixed at 0.

A second representation using the scaling function where the m value was set to be 2 is shown in Figure 13. This produced 80 nonzero terms. The maximum error was 13.75% and the average error was 9.42%.

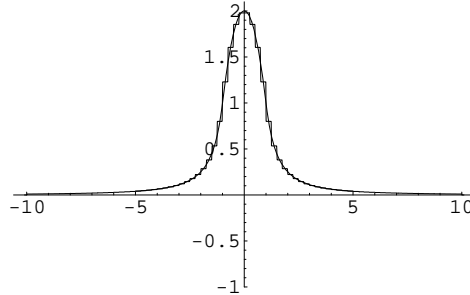


FIGURE 13. 80 nonzero terms, m value fixed at 2.

A second type of wavelet was also used to investigate this last function. This was the D4, four coefficients, wavelet of I. Daubechies. This wavelet was the next step in the natural progression for examining wavelets since it has compact support like the Haar wavelet, but it is also continuous. A graphical approximation of the father function of Daubechies D4 wavelet is shown in Figure 14. This function has compact support and it is continuous. The next wavelet of interest would be a Daubechies D8 wavelet, which not only has compact support and is continuous but is also differentiable. The time scale for this work only allowed time to work with the D4 wavelet.

For our work, this new father function was approximated by using the function illustrated in Figure 15. This is a crude approximation to the D4 function, but it was used to save computer time since an iterative method was needed to generate it (see [Da]).

This father function was used to make representations of functions just as with the Haar wavelets. The m value was set to be zero. In examining the interval

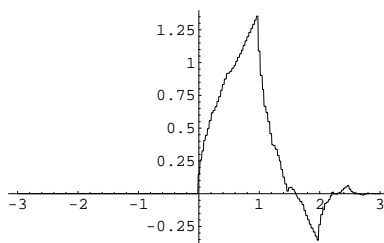


FIGURE 14. The D4 father function.

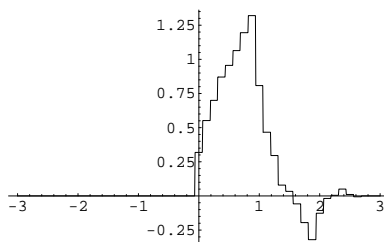


FIGURE 15. Our approximation to the D4 father function.

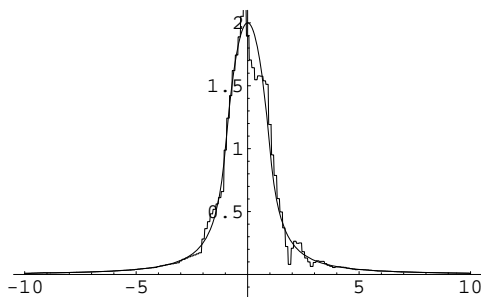


FIGURE 16. Our approximation using D2 with 20 nonzero terms along with the original function.

$[-10, 10]$, this produced 20 nonzero terms when the k value went from $-10 \cdot 2^m - 2 < k < 10 \cdot 2^m - 1$. The maximum error was 137.85% and the average error was 19.48%.

Recall that when 50 nonzero terms were used with the Fourier Series, the maximum error was 133.55%, and when 20 nonzero terms were used with the father function (Haar wavelet), the maximum error was 40.00% and the average error was 22.88%. One possible explanation for why the D4 wavelet has a higher error than the Haar system is that its support is larger, it takes more terms to “zoom in” on the local properties. The second possible explanation is that we used a crude approximation of the D4 function.

5. CONCLUSION

Two techniques were used to make representations of functions. The first was Fourier series which combines an infinite number of sine and cosine functions. This type of series is good for representing the frequency of a function. The major

problem is that it may take many terms to make an accurate approximation to functions that exhibit local properties. The solution and second type of technique is wavelets. Wavelets do a good job representing local properties of functions. Two types of wavelets were examined, Haar and Daubechies D4 wavelets. The Haar wavelets have compact support, and the Daubechies D4 wavelets have compact support and are continuous. The wavelet father functions were used in representing functions for both wavelets since they give a simpler sum to examine. The examples given in this paper illustrate the “zooming in” property of wavelets which make them very useful in representing local properties of functions.

REFERENCES

- [Br1] C.W. Brislawn, *Fingerprints go digital*, Notices of the American Math Soc., 42 (1995), 1278-1283.
- [Br2] C.M. Brislawn, *The FBI Fingerprint Image Compression Standard*, www.c3.lanl.gov/~brislawn/FBI/FBI.html.
- [Da] I. Daubechies, *Orthonormal bases of compactly supported wavelets*, Comm. on Pure and Appl. Math., **XLI** (1988), 909-996.
- [Fr] M. W. Frazier, *An Introduction to Wavelets through Linear Algebra*, Springer-Verlag New York, Inc., (1999), 1-6, 160-169, 380-383.
- [Ke] S. Kelly, *Pointwise Convergence for Wavelet Expansions*, PhD thesis, 1992.
- [Ma] S. Mallat, *Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathbf{R})$* , Trans. of the Am. Math. Soc., **315** (1989), no. 1, 69-87.
- [Wa] J. Walker, *Fourier Analysis*, Oxford University Press, (1988).

MATHEMATICS DEPARTMENT, UW - LA CROSSE, LA CROSSE, WISCONSIN, USA
E-mail address: roland.nich@students.uwlax.edu

SPONSOR: SUSAN E. KELLY, UW - LA CROSSE, LA CROSSE, WISCONSIN, USA
E-mail address: kelly.susa@uwlax.edu