

THE SHAPE OF THE ERROR IN WAVELET APPROXIMATION AND PIECEWISE LINEAR INTERPOLATION

ROBERT S. STRICHARTZ*

ABSTRACT. The graph of the error in wavelet approximation, when properly rescaled, is shown to converge in the Hausdorff metric to a limit set Γ . The limit set Γ is not a graph of a function, but rather a region bounded by the graphs of multiples of a derivative of the function, depending on the first non-vanishing moment of the wavelet. A similar result is shown for piecewise linear interpolation. Higher dimensional analogs are discussed.

1. Introduction

What is the size of the error? That is perhaps the most important question in any approximation method. In this paper we consider a different question: *what is the shape of the error?* We will show that for wavelet approximation this question has a precise answer that actually predicts what you will see if you graph the error on an appropriate scale. It is likely that this answer can serve as a paradigm for what happens using other approximation methods, and we give some illustrations of this. The key idea is to avoid making statements about the pointwise behavior of the error, and instead consider the behavior of the graph of the error as a closed set.

So let us consider a continuous real-valued function f defined on a compact set X , and some approximation method that produces a sequence $\{f_n\}$ of approximations to f . Typically we will have some estimate $f - f_n = O(\varepsilon_n)$ for the error, and so by rescaling we obtain a sequence $\varepsilon_n^{-1}(f - f_n)$ of functions of comparable size. Typically we expect a lot of oscillation in the rescaled error, so we would not see any convergence of $\varepsilon_n^{-1}(f - f_n)$ as functions. But we might see convergence of the graphs Γ_n of $\varepsilon_n^{-1}(f - f_n)$ in the Hausdorff metric on the space of nonempty compact sets, $\Gamma_n \rightarrow \Gamma$. (Recall the definition: the distance $d(A, B)$ is the infimum over ε such that A is contained in an ε -neighborhood of B and B is contained in an ε -neighborhood of A ; informally, if you suffer from ε -myopia then A and B look the same.) The limit set Γ would not be

Received March 1, 2000.

*Research supported in part by the National Science Foundation, grant DMS 9970337.

expected to be the graph of a function, but rather a compact subset of $X \times \mathbb{R}$ with nonempty interior. The graph Γ_n for large n would appear as a “scribbled in” approximation to Γ .

This really happens. We can demonstrate this easily in two simple examples. In the first example we approximate f on the unit interval by the piecewise constant functions f_n that arise by taking the mean value of f on each of the 2^n dyadic subintervals of length 2^{-n} . This is, of course, exactly the Haar wavelet approximation method. We need to assume that f is C^1 . Then on the dyadic interval $I_k = [k2^{-n}, (k+1)2^{-n}]$ we may write

$$(1.1) \quad f(x) = f((k+1/2)2^{-n}) + f'((k+1/2)2^{-n})(x - (k+1/2)2^{-n}) + o(2^{-n}).$$

The mean value is then $f((k+1/2)2^{-n}) + o(2^{-n})$. We take $\varepsilon_n = 2^{-n}$ and find

$$(1.2) \quad \varepsilon_n^{-1}(f - f_n) = 2^n f'((k+1/2)2^{-n})(x - (k+1/2)2^{-n}) + o(1) \text{ on } I_k.$$

Ignoring the error term, the graph of this function over I_k is just the slant line rising from $-\frac{1}{2}f'((k+1/2)2^{-n})$ at the left endpoint to $\frac{1}{2}f'((k+1/2)2^{-n})$ at the right endpoint (assuming $f'((k+1/2)2^{-n}) > 0$, with a falling line if $f'((k+1/2)2^{-n}) < 0$). It is clear that Γ_n converges to the region bounded above by the graph of $\frac{1}{2}|f'(x)|$ and below by the graph of $-\frac{1}{2}|f'(x)|$. Moreover, the rate of convergence in the Hausdorff metric will be $O(2^{-n})$ if we assume f is C^2 (the slightly weaker condition that f' is Lipschitz will also suffice).

The next example we consider is piecewise linear interpolation for a dyadic subdivision of the unit interval (other subdivision schemes could also be used). This is the simplest example of a spline approximation method. In this case we need to assume that f is C^2 . On each interval I_k the error $f - f_n$ vanishes at the endpoints and has the same second derivative as f , so

$$(1.3) \quad f - f_n = \frac{1}{2}f''\left(\left(k + \frac{1}{2}\right)2^{-n}\right)(x - k2^{-n})(x - (k+1)2^{-n}) + o(2^{-2n}) \text{ on } I_k.$$

Thus we take $\varepsilon_n = 2^{-2n}$. Ignoring the error term, it is easy to see that when $f''((k + \frac{1}{2})2^{-n}) > 0$, the function $\varepsilon_n^{-1}(f - f_n)$ on I_k takes its minimum value $-\frac{1}{8}f''((k + \frac{1}{2})2^{-n})$ at $x = (k + \frac{1}{2})2^{-n}$ and its maximum value 0 at the endpoints, and similarly varies between $-\frac{1}{8}f''((k + \frac{1}{2})2^{-n})$ and 0 when $f''((k + \frac{1}{2})2^{-n}) < 0$. Thus the limit set Γ is just the region between the graph of $-\frac{1}{8}f''(x)$ and the x -axis. We also have a rate of convergence $O(2^{-n})$ under the hypothesis that f is C^3 .

In Section 2 we will discuss results analogous to the first example for general wavelets in one dimension. It turns out that the limit set Γ is a region bounded by graphs of multiples of $f^{(m)}(x)$, where m is the first non-vanishing moment of the wavelet, under the assumption that f is C^m , with $\varepsilon_n = 2^{-nm}$. This result is also easily deduced from the work of Sweldens and Piessens [SP1,2], who prove that the error is asymptotically expressible as $\varepsilon_n f^{(m)}(x) \tau_0(2^n x)$ for a specific

function τ_0 that is called a *monowavelet*. Moreover, this is only the first term of a longer asymptotic expansion, under the assumption of greater smoothness for f . These results are much more precise than ours, but they might not have as wide a scope. There are graphical illustrations in [SP1,2] of the result. Further development of those ideas may be found in [BU1,2] and [U]. In Section 3 we discuss the situation for wavelets in higher dimensions. In Section 4 we cover the case of piecewise linear interpolation in the plane.

In addition to the intrinsic interest of these results, there may also be practical applications. We note that it is not really necessary to know what f is to see the predicted effects. If we know only $\{f_n\}$ for $n \leq N$, then the graphs of $\varepsilon_n^{-1}(f_N - f_n)$ will show a similar behavior (depending on the specifics, we may have to stop before $n = N - 1$). Moreover, even if we do not know a priori what ε_n should be, we can always make an a posteriori choice to rescale the graphs to be of comparable size. The convergence, or lack of convergence, of the graphs should then be visually apparent. This might be useful as a diagnostic test for the reliability of the approximation method. In cases where it is not known whether or not the function satisfies the requisite smoothness hypothesis, this might give some evidence, and even yield a better idea of what the higher derivative $f^{(m)}(x)$ is like than would be obtained by numerically differentiating the approximations to f .

It is not difficult to give generalizations of the second example to other types of spline interpolation. For example, using cubic splines that interpolate both the value of a function and its first derivative at the nodes, one obtains a region bounded by the x -axis and the graph of multiples of $f^{(4)}(x)$ under the assumption that f is C^4 . It would be interesting to understand what happens for other types of spline approximation, especially those used in the finite element method, and also what happens in higher dimensions.

For convolutions with an approximate identity the situation is much simpler, with $t^{-m}(f(x) - \varphi_t * f)$ converging to a multiple of $f^{(m)}(x)$ where m is the first nonzero moment of φ , under the assumption that f is C^m . This includes summability methods for Fourier series, but the situation for partial sums of Fourier series is not clear.

Another interesting question is what happens if the function has the requisite number of derivatives, but only in a weak sense.

The original inspiration for this paper came from an entirely different approximation method involving sampling, with a rescaling in the x -variable. In this context the results are only experimental, and will be discussed in [ST].

We are grateful to M. Unser for bringing to our attention the earlier work on wavelet approximation.

2. One-dimensional wavelets

For simplicity we will consider compactly supported real-valued orthonormal wavelets on \mathbb{R} . The results are easily extended to bi-orthogonal wavelets, multi-wavelets, or wavelets with sufficiently rapid decay, but we leave the details to the

reader (see also [Bu1,2], [SP1,2] and [U]). Thus we have a compactly supported wavelet generator $\psi(x)$ such that $\{\psi_{jk}\} = \{2^{j/2}\psi(2^jx-k)\}$ forms an orthonormal basis of $L^2(\mathbb{R})$ for $j, k \in \mathbb{Z}$. We will assume that ψ is continuous, although the results are still valid for the Haar wavelet as described in the Introduction. No smoothness is required of ψ , even when we do require smoothness of the function f being approximated. Instead we are interested in the vanishing of moments of ψ . Let

$$(2.1) \quad A_k = \int x^k \psi(x) dx,$$

and let m be the smallest positive integer such that $A_m \neq 0$. We always have $A_0 = 0$, and there are wavelets for which all moments vanish. We will not be able to say anything in such cases, but we also note that such wavelets tend to have very slow decay (this seems to be some kind of uncertainty principle).

We form the auxiliary function

$$(2.2) \quad g(x) = \sum_{k=-\infty}^{\infty} \sum_{j=1}^{\infty} 2^{-jm} \psi(2^jx - k).$$

Note that there is a universal upper bound (roughly the length of the support of ψ) on the number of non-zero k terms for each x and j , so the sum is obviously convergent, and defines a continuous periodic function of period $1/2$. Let g_{\min} and g_{\max} denote the minimum and maximum values of g . Since g has mean value zero, g_{\min} is negative and g_{\max} is positive. The next result is also an easy consequence of results in [SP1,2].

Theorem 2.1. *Let f be a compactly supported real-valued function of class C^m , and let*

$$(2.3) \quad f_n(x) = \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^n \left(\int \psi_{jk}(y) f(y) dy \right) \psi_{jk}(x)$$

be the wavelet approximation to f of order n . Then for $\varepsilon_n = 2^{-mn}$, the graphs of $\varepsilon_n^{-1}(f - f_n)$ over a fixed interval I converge to the set Γ equal to the region bounded by the graphs of $\frac{A_m}{m!} g_{\min} f^{(m)}(x)$ and $\frac{A_m}{m!} g_{\max} f^{(m)}(x)$ over I . Moreover, if f is of class C^{m+1} then the rate of convergence in the Hausdorff metric is $O(2^{-n})$.

Proof. Since the wavelet series converges uniformly, we have

$$(2.4) \quad f(x) - f_n(x) = \sum_{k=-\infty}^{\infty} \sum_{j=n+1}^{\infty} \left(\int \psi_{jk}(y) f(y) dy \right) \psi_{jk}(x).$$

We substitute the Taylor expansion

$$(2.5) \quad f(y) = \sum_{\ell=0}^m \frac{f^{(\ell)}(x)}{\ell!} (y-x)^\ell + \mathcal{R}_m(x, y)$$

into (2.4). Note that all the terms corresponding to $\ell < m$ vanish because the moments A_ℓ are zero. Moreover

$$\int \psi(2^j y - k)(y-x)^m dy = \int \psi(2^j y) y^m dy = 2^{-j(m+1)} A_m$$

because the lower moments vanish. Thus

$$(f(x) - f_n(x)) = \sum_{k=-\infty}^{\infty} \sum_{j=n+1}^{\infty} \left(\frac{f^{(m)}(x)}{m!} 2^{-jm} A_m \psi(2^j x - k) + \int \psi_{jk}(y) \mathcal{R}_m(x, y) dy \psi_{jk}(x) \right),$$

hence

$$(2.6) \quad 2^{nm}(f(x) - f_n(x)) = \frac{A_m}{m!} f^{(m)}(x) g(2^n x) + \mathcal{R}_n(x)$$

where

$$(2.7) \quad \mathcal{R}_n(x) = 2^{nm} \sum_{k=-\infty}^{\infty} \sum_{j=n+1}^{\infty} \int \psi_{jk}(y) \mathcal{R}_m(x, y) dy \psi_{jk}(x).$$

It is easy to show that $\mathcal{R}_n(x) \rightarrow 0$ uniformly over I , and $\mathcal{R}_n(x) = O(2^{-n})$ if f is C^{m+1} . The result now follows from (2.6), since the function $g(2^n x)$ fills in all values between g_{\min} and g_{\max} on each interval of length 2^{-n-1} . \square

Note that if $g_{\min} \neq -g_{\max}$ then Γ determines $f^{(m)}(x)$, and the approximations Γ_n yield approximations to $f^{(m)}(x)$. However, if $g_{\min} = -g_{\max}$, then we can only obtain information about $|f^{(m)}(x)|$ in this manner. Note that the condition $g_{\min} \neq -g_{\max}$ is a kind of asymmetry condition, and might in some way be related to the well-known asymmetry of compactly supported wavelets [D]. See [Bu1,2], [SP1,2] and [U] for numerical computation of the constants.

If we only have access to the wavelet approximations f_n , say for $n \leq N$, then $f_N(x) - f_n(x)$ may be used as a substitute for $f(x) - f_n(x)$ for $n \ll N$, but it is perhaps better to consider the behavior of the graphs of $f_n(x) - f_{n-1}(x)$. Then exactly the same result holds if we replace the function $g(x)$ by the function

$$(2.8) \quad \tilde{g}(x) = \sum_{k=-\infty}^{\infty} \psi(x - k).$$

The situation is slightly different if we allow complex valued functions f , or complex valued wavelets. In the first case Γ will be a 2-dimensional surface in $I \times \mathbb{C}$ foliated by line segments in \mathbb{C} joining the points $\frac{A_m}{m} g_{\min} f^{(m)}(x)$ and $\frac{A_m}{m!} g_{\max} f^{(m)}(x)$ over each point $x \in I$. In the second case we let γ denote the closed (not necessarily simple) curve in \mathbb{C} parametrized by $g(x)$ for $0 \leq x \leq \frac{1}{2}$. In this case $\Gamma = \{(x, \frac{A_m}{m!} f^{(m)}(x)\gamma) : x \in I\}$ is foliated by the curves $\frac{A_m}{m!} f^{(m)}(x)\gamma$ which are dilated and rotated images of γ .

3. Multidimensional wavelets

We consider first the analog of the Haar wavelet approximation method on \mathbb{R}^d . On each dyadic cube of side length 2^{-n} we approximate f by its mean value; this is the approximation f_n . We again take $\varepsilon_n = 2^{-n}$, and assume that f is C^1 . If $k \in \mathbb{Z}^d$ and

$$I_k = [k_1 2^{-n}, (k_1 + 1) 2^{-n}] \times \cdots \times [k_d 2^{-n}, (k_d + 1) 2^{-n}]$$

denotes the associated dyadic cube, then the analog of (1.2) is

$$\varepsilon_n^{-1}(f - f_n) = 2^n \sum_{i=1}^d \frac{\partial f}{\partial x_i} \left(\left(k + \frac{1}{2}\right) 2^{-n} \right) \left(x_i - \left(k_i + \frac{1}{2}\right) 2^{-n} \right) + o(1) \text{ on } I_k.$$

Ignoring the error term, we see that the maximum value is attained at the corner of the cube where the signs of $x_i - (k_i + \frac{1}{2}) 2^{-n}$ match the signs of $\frac{\partial f}{\partial x_i}$, and equals

$$\frac{1}{2} \sum_{i=1}^d \left| \frac{\partial f}{\partial x_i} \left(\left(k + \frac{1}{2}\right) 2^{-n} \right) \right|,$$

and similarly the minimum is attained at the opposite corner. Thus the limit set Γ is bounded above and below by the graphs of

$$\pm \frac{1}{2} \sum_{i=1}^d \left| \frac{\partial f}{\partial x_i} \right|.$$

Next we consider a general wavelet system in \mathbb{R}^d generated by a finite family $\{\psi^{(\ell)}(x)\}$ of $1 \leq \ell \leq L$ of continuous, compactly supported wavelets with respect to a general dilation matrix R . R is assumed to be an integer valued expanding matrix of determinant r (with $|r| > 1$), and $\{\psi_{jk}^{(\ell)}\} = \{r^{j/2} \psi^{(\ell)}(R^j x - k)\}$ is an orthonormal basis of $L^2(\mathbb{R}^d)$ for $1 \leq \ell \leq L$, $j \in \mathbb{Z}$ and $k \in \mathbb{Z}^d$. We also assume that the wavelet series converges uniformly for continuous functions of compact support. It is well-known how to construct such systems with $R = 2I$ by a tensor product method (so-called “separable” wavelets), but other systems are known [BW]. The construction is easier if one allows biorthogonal systems

([CD], [CS]) or drops compact support [S], and our results could be modified to handle these variations.

We let $\alpha = (\alpha_1, \dots, \alpha_d)$ denote a multi-index of nonnegative integers, and consider the moments

$$(3.1) \quad A_\alpha^{(\ell)} = \int_{\mathbb{R}^d} \psi^{(\ell)}(x) x^\alpha dx.$$

We set m equal to the smallest value of $|\alpha|$ such that there exists α and ℓ with $A_\alpha^{(\ell)} \neq 0$, so $A_\alpha^{(\ell)} = 0$ for all $|\alpha| < m$ and all ℓ . We define a family of auxiliary functions

$$(3.2) \quad g^{(\ell)}(x) = \sum_{k \in \mathbb{Z}^d} \sum_{j=1}^{\infty} r^{-jm} \psi^{(\ell)}(R^j x - k).$$

Just as in the one-dimensional case, it is easy to see that the series converges uniformly to a continuous periodic function.

Theorem 3.1. *Let f be a compactly supported real-valued function of class C^m , and let*

$$(3.3) \quad f_n(x) = \sum_{\ell=1}^L \sum_{k \in \mathbb{Z}^d} \sum_{j=-\infty}^n \left(\int \psi_{jk}^{(\ell)}(y) f(y) dy \right) \psi_{jk}^{(\ell)}(x)$$

be the wavelet approximation to f of order n . Then for $\varepsilon_n = r^{-mn}$, the graphs of $\varepsilon_n^{-1}(f - f_n)$ over a fixed cube I converge to the set Γ equal to the region bounded by the graphs of

$$(3.4) \quad \begin{cases} \min_y \sum_{|\alpha|=m} \left(\sum_{\ell=1}^L \frac{A_\alpha^{(\ell)}}{\alpha!} g^{(\ell)}(y) \right) \left(\frac{\partial}{\partial x} \right)^\alpha f(x) & \text{and} \\ \max_y \sum_{|\alpha|=m} \left(\sum_{\ell=1}^L \frac{A_\alpha^{(\ell)}}{\alpha!} g^{(\ell)}(y) \right) \left(\frac{\partial}{\partial x} \right)^\alpha f(x) \end{cases}$$

over I . Moreover, if f is of class C^{m+1} then the rate of convergence in the Hausdorff metric is $O(r^{-n})$.

Proof. The proof follows the same outline as the proof of Theorem 2.1, so we only indicate the differences. In place of (2.6) we will have

$$(3.5) \quad r^{nm}(f(x) - f_n(x)) = \sum_{|\alpha|=m} \sum_{\ell=1}^L \frac{A_\alpha^{(\ell)}}{\alpha!} \left(\frac{\partial}{\partial x} \right)^\alpha f(x) g^{(\ell)}(R^n x) + \mathcal{R}_n(x)$$

with the same estimates on the remainder. For large n the values of $\{g^{(\ell)}(R^n x)\}$ will attain all values $\{g^{(\ell)}(y)\}$ as x varies over a small region, and this explains why (3.4) appears as the boundary of the region Γ . \square

4. Piecewise linear interpolation in \mathbb{R}^2

In this section we consider piecewise linear interpolation on a region in \mathbb{R}^2 with respect to a sequence of triangulations of the region. If we concentrate on a fixed small triangle, and replace the function by its second order Taylor approximation, we are led to the elementary problem of finding the maximum and minimum over the triangle of a quadratic polynomial that vanishes at the vertices. To describe the solution of this problem we introduce a definition.

Definition 4.1: A triple (a, b, c) of positive numbers is called *triangular* if there exists a Euclidean triangle with sides of length a, b, c . Equivalently, the three triangle inequalities

$$(4.1) \quad a < b + c, \quad b < c + a, \quad c < a + b$$

hold.

Lemma 4.2. Let T be a triangle in the plane with sides s_1, s_2, s_3 regarded as vectors in \mathbb{R}^2 . Let $D_j f = s_j \cdot Df$ denote the derivative with respect to the vector s_j . Suppose f is a real quadratic function that vanishes on the vertices of T .

(i) If all the numbers $D_j^2 f$ are positive and $(D_1^2 f, D_2^2 f, D_3^2 f)$ is triangular, then the maximum value of f over T is 0 and the minimum value is

$$(4.2) \quad -\frac{D_1^2 f D_2^2 f D_3^2 f}{32(\det D^2 f)(\text{Area } T)^2}.$$

(ii) If all the numbers $D_j^2 f$ are negative and $(-D_1^2 f, -D_2^2 f, -D_3^2 f)$ is triangular, then (4.2) is the maximum value and 0 is the minimum value of f over T .

(iii) In all other cases, the minimum value of f over T is

$$(4.3) \quad \min \left\{ 0, -\frac{1}{8}D_1^2 f, -\frac{1}{8}D_2^2 f, -\frac{1}{8}D_3^2 f \right\}$$

and the maximum value is

$$(4.4) \quad \max \left\{ 0, -\frac{1}{8}D_1^2 f, -\frac{1}{8}D_2^2 f, -\frac{1}{8}D_3^2 f \right\}.$$

Proof. The extremal values must be assumed either on the boundary or at an interior critical point. Along each side, the extrema occur either at the boundary (0) or at the midpoint $(-\frac{1}{8}D_j^2 f)$. Thus (4.3) and (4.4) give the correct extrema in the case when no interior critical point exists.

To simplify the argument we assume that T is the triangle with vertices $(0, 0)$, $(0, 1)$, $(1, 0)$, since the general triangle can be mapped to this one by an affine transformation, and the result is affine invariant. (More precisely, if

$g(x) = Bx + d$ is the affine map that sends the triangle \tilde{T} to T , and $\tilde{f} = f \circ g$, then $\tilde{D}_j^2 \tilde{f} = D_j^2 f$, where $\tilde{D}_j = \tilde{s}_j \cdot D$, and $D^2 \tilde{f} = B^* D^2 B f$, showing the invariance of the denominator of (4.2) because $\text{Area}(T) = |\det B| \text{Area}(\tilde{T})$. Then

$$(4.5) \quad f(x, y) = \frac{1}{2}a(x^2 - x) + bxy + \frac{1}{2}c(y^2 - y)$$

for some a, b, c . Note that

$$(4.6) \quad (D_1^2 f, D_2^2 f, D_3^2 f) = (a, c, a + c - 2b).$$

In the nondegenerate case $\det D^2 f = ac - b^2 \neq 0$ there is a unique critical point in \mathbb{R}^2 ,

$$(4.7) \quad (x, y) = \left(\frac{(a-b)c}{2(ac-b^2)}, \frac{(c-b)a}{2(ac-b^2)} \right).$$

For this to be an interior point of the triangle we must have

$$(4.8) \quad \begin{cases} 0 < \frac{(a-b)c}{ac-b^2} < 2 \\ 0 < \frac{(c-b)a}{ac-b^2} < 2 \\ 0 < \frac{2ac-ab-bc}{ac-b^2} < 2. \end{cases}$$

We need only consider the elliptic case $ac - b^2 > 0$, since in the hyperbolic case the critical point is a saddle point. If a and c are positive then (4.8) is equivalent to the conditions $b > 0$, $a > b$, $c > b$, and this is exactly the condition that (4.6) is triangular. Thus we are in the case (i). The critical point is the global minimum, and the value at that point is

$$-\frac{ac(a+c-2b)}{8(ac-b^2)}.$$

This is the same as (4.2) since $\text{Area } T = 1/2$ in this case.

If a and c are both negative (still in the elliptic case) we are in case (ii), and the computations are similar. It is not hard to see that the degenerate case $b^2 = ac$ falls under (iii), since if there is a critical point there must be a whole line of critical points along which the value of f is constant, so the maximum and minimum value will be attained on the boundary of the triangle. \square

Note that in the elliptic case one of the extrema is always 0 (the maximum for upward concavity, the minimum for downward concavity), and this is obvious by a convexity argument. In the hyperbolic case it is still possible to have one of the extrema equal to 0, depending on the orientation of the triangle.

In the elliptic case one could choose a different affine transformation to make $f(x, y) = \pm(x^2 + y^2 - r^2)$, and the triangle becomes inscribed into the circle of radius r about the origin. Case (iii) corresponds to an acute triangle, for then the center of the circle lies in the interior of the inscribed triangle, and the other cases correspond to an obtuse or right triangle.

For the next theorem we deal with regular triangulations of nice regions in \mathbb{R}^2 . It should be possible to prove related results for other types of triangulations and more general regions. We fix a pair of triangles T and T^* that together form a parallelogram that tiles the plane with a lattice Λ . If Ω denotes a compact region bounded by a C^1 curve, we consider a sequence of triangulations \mathcal{T}_n of Ω consisting of all triangles of the form $2^{-n}(T + \lambda)$ and $2^{-n}(T^* + \lambda)$ for $\lambda \in \Lambda$ that are contained in Ω . Under our assumptions on Ω , there exists a universal constant c such that every point in Ω is within a distance of $c2^{-n}$ from one of the triangles in \mathcal{T}_n . For any continuous function f on Ω we define the piecewise linear interpolant f_n with domain equal to the union of the triangles of \mathcal{T}_n to be the piecewise linear function with the same values as f at the vertices of those triangles.

Note that both T and T^* have the same sides (up to a sign). We continue to use the notation of Lemma 4.2.

Theorem 4.3. *Assume f is a real valued C^2 function on Ω . Let $\varepsilon_n = 2^{-2n}$, let $\varepsilon_n^{-1}(f - f_n)$ be interpreted to have the same domain as f_n , and denote by Γ_n its graph. Then Γ_n converges in the Hausdorff metric to a set Γ (at rate $O(2^{-n})$ if f is assumed to be C^3) that may be described as the region in $\Omega \times \mathbb{R}$ bounded by the graphs of functions F_{\max} and F_{\min} defined as follows. Partition Ω into $\Omega_{(i)}$, $\Omega_{(ii)}$ and $\Omega_{(iii)}$ according to whether the numbers $D_j^2 f(x)$ satisfy conditions (i), (ii) or (iii) of Lemma 4.2. For $x \in \Omega_{(i)}$, $F_{\max}(x) = 0$ and $F_{\min}(x)$ is given by (4.2). For $x \in \Omega_{(ii)}$, $F_{\min}(x) = 0$ and $F_{\max}(x)$ is given by (4.2). For $x \in \Omega_{(iii)}$, $F_{\min}(x)$ is given by (4.3) and $F_{\max}(x)$ is given by (4.4).*

Proof. On each triangle in \mathcal{T}_n we replace f by its second order Taylor expansion about a point x_0 in the triangle, committing an error of the order $o(2^{-2n})$ that remains small after multiplication by ε_n^{-1} . The sides of the triangles in \mathcal{T}_n are reduced by a factor of 2^{-n} from those of T , and so all the quantities (4.2), (4.3) and (4.4) are reduced by a factor of ε_n . Thus Lemma 4.2 says that if we ignore the error term then the function $\varepsilon_n^{-1}(f - f_n)$ assumes all values between $F_{\min}(x_0)$ and $F_{\max}(x_0)$ on the triangle. The rest of the argument is routine. \square

References

- [BW] E. Belogay and Y. Wang, *Arbitrarily smooth orthogonal nonseparable wavelets in \mathbb{R}^2* , SIAM J. Math. Anal. **30** (1999), 678–697.
- [BU1] T. Blu and M. Unser, *Approximation error for quasi-interpolators and (multi-) wavelet expansions*, Appl. Comput. Harmon. Anal. **6** (1999), 219–251.
- [BU2] ———, *Quantitative Fourier analysis of approximation techniques: Part II - wavelets*, IEEE Trans. Signal Process. **47** (1999), 2796–2806.

- [CD] A. Cohen and I. Daubechies, *Non-separable bidimensional wavelet bases*, Rev. Mat. Iberoamericana **9** (1993), 51–137.
- [CS] A. Cohen and J.-M. Schlenker, *Compactly supported bidimensional wavelet bases with hexagonal symmetry*, Constr. Approx. **9** (1993), 209–236.
- [D] I. Daubechies, *Ten lectures on wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, 61, SIAM, Philadelphia, PA, 1992.
- [S] R.S. Strichartz, *Wavelets and self-affine tilings*, Constr. Approx. **9** (1993), 327–346.
- [ST] R.S. Strichartz and N. Tillman, *Sampling theory for functions with fractal spectrum*, in preparation.
- [SP1] W. Sweldens and R. Piessens, *Quadrature formulae and asymptotic error expansions for wavelet approximations of smooth functions*, SIAM J. Numer. Anal. **31** (1994), 1240–1264.
- [SP2] ———, *Asymptotic error expansions of wavelet approximations of smooth functions II*, Numer. Math. **68** (1994), 377–401.
- [U] M. Unser, *Approximation power of biorthogonal wavelet expansions*, IEEE Trans. Signal Process. **44** (1996), 519–527.

MATHEMATICS DEPARTMENT, MALOTT HALL, CORNELL UNIVERSITY, ITHACA, NY 14853
E-mail address: `str@math.cornell.edu`