The Annals of Statistics
1998, Vol. 26, No. 6, 2206–2217

ON GLOBAL PERFORMANCE OF APPROXIMATIONS TO SMOOTH CURVES USING GRIDDED DATA

BY PETER HALL AND MARC RAIMONDO

Australian National University

Approximating boundaries using data recorded on a regular grid induces discrete rounding errors in both vertical and horizontal directions. In cases where grid points exhibit at least some degree of randomness, an extensive theory has been developed for local-polynomial boundary estimators. It is inapplicable to regular grids, however. In this paper we impose strict regularity of the grid and describe the performance of local linear estimators in this context. Unlike the case of classical curve estimation problems, pointwise convergence rates vary erratically along the boundary, depending on number-theoretic properties of the boundary's slope. However, *average* convergence rates, expressed in the L^1 metric, are much less susceptible to fluctuation. We derive theoretical bounds to performance, coming within no more than a logarithmic factor of the optimal convergence rate.

1. Introduction.

1.1. *Information in digitized data.* Consider the problem of digitizing a curve from the trace that it leaves on a grid. Suppose the region above the curve, represented by the equation $y = g(x)$, is colored black and the area below, white; and that the vertices of the grid preserve these colors, but provide no other information. From the color pattern we wish to approximate the curve itself. Even without noise, no precise observations of g are available, unless the boundary should happen to pass directly through a vertex.

In the case of a regular grid, such as a square lattice (used in most imaging equipment) or hexagonal lattice (employed in at least one proprietary image analyser), relatively little theory exists for describing the performance of estimators computed from vertex color data. Curiously, the case of a random array of grid points is simpler and has received considerable attention. A random array might, for example, arise through a Poisson distribution of vertices in the plane and has the attractive property that there are no special orientations for which it is likely to produce better approximations to boundaries. It offers a continuum of possibilities for approximation. This is particularly advantageous for theoretical analysis, and so various "jittered" grids (with vertices perturbed randomly within small regions centered at vertices of a regular lattice) have been treated as approximations to regular grids.

A detailed account of different grid types, including jittered grids, is provided by the monograph of Korostelev and Tsybakov (1993). These authors

Received August 1996; revised June 1998.

AMS 1991 *subject classification*. Primary 62G20.

Key words and phrases. Digital image, image analysis, lattice, local linear smoothing, local polynomial smoothing, L^1 metric, metric number theory, numerical analysis, pixel grid.

devote particular attention to local polynomial approximations to boundaries, using vertex color data. They survey the literature, noting that the regular grid case has received little attention and pointing out the elegance and detail of results available for the continuum or "random grid" case. In the context of engineering, the Freeman code [Freeman (1970)] is a particularly well-known practical method for tracking a boundary using discrete data on vertex colors. Worring and Smeulders (1995) describe related techniques and recent applications and survey the literature.

1.2. *Description of main results.* In this paper we devote attention to regular grids. There, even the noiseless case requires attention, although statistical methods for noisy data are helpful to its analysis; see Section 1.3. It is sufficient to treat the case of a square grid, since almost identical results for its more sophisticated competitors, such as hexagonal arrays of vertices, are readily obtained by noting that square subgrids are embedded within them and that they are in turn embedded within overlapping square grids.

We derive upper bounds to the error of approximation by local linear smoothers and their interpolants. It is shown that a twice-differentiable curve inscribed on a square grid of edge width n^{-1} is approximable at rate $r_n = (n^{-1} \log n)^{4/3}$, in an L^1 sense, with or without noise, and that the rate is valid uniformly in a special class of twice-differentiable functions. It may be proved that in a minimax sense, for randomly tilted and shifted curves with two nonvanishing bounded derivatives, observed on a regular grid, the rate $n^{-4/3}$ is a lower bound.

A convergence rate in the case of regular grids has been given by Korostelev and Tsybakov [(1993), pages 139 and 140], for their "preliminary estimator," although it is only $n^{-1}(\log n)^{1/2}$. This rate is for the Hausdorff metric and in a pointwise sense. The methods of Korostelev and Tsybakov for improving the performance of this estimator require irregularly arranged grid vertices and so are not applicable to the contexts treated in the present paper.

1.3. *Role of number-theoretic properties of slope.* The L^1 measure provides only an average description of convergence. Unlike the classical case of nonparametric curve estimation, the L^1 rate is not available at each fixed point. Indeed, using local linear methods to estimate $y = g(x)$ at a given point $x = t$, where $g'(t)$ is rational, the best rate of convergence that can be expected is $O(n^{-1})$—substantially inferior to the rate r_n. Superior rates can be achieved at other places; for example, it may be shown that the rate of convergence of an appropriately smoothed local linear estimator at points t such that $g'(t)$ is a quadratic irrational [i.e., $g'(t)$ is the solution of a quadratic equation with rational coefficients] is $n^{-4/3}$.

More generally, points t at which g has a rational or irrational slope fall into different categories, and the pointwise rate of convergence to $g(t)$ is driven by properties of a continued fraction expansion of the real number $g'(t)$. That expansion is of infinite extent if and only if $g'(t)$ is irrational.

Convergence rates in L^1, describing "average" performance of an estimator, are arguably of more practical relevance than pointwise rates, which in the present problem have been treated by Hall and Raimondo (1997). In both instances the methods rely on techniques from number theory. Indeed, the fact that the convergence rate r_n is available in an L^1 sense, even though it is not valid for a dense set of fixed points (e.g., the rationals), is a consequence of metric number theory: if the value of $g'(t)$ should be chosen randomly in the interval $[0, 1]$, then with probability 1 the pointwise approximation to g at t would be at least as good as r_n. Versions of the results in this paper for high-order polynomial smoothers are restricted by the difficulty of developing high orders of Diophantine approximation [e.g., Baker (1986), Chapter 1].

One portion of the logarithmic factor in r_n derives from the fact that if $g'(t)$ is chosen randomly, then the ratio of any two consecutive partial denominators in its continued fraction expansion has a Cauchy-like distribution, with the tail of the distribution decreasing like x^{-1}. Thus, even in the absence of noise, the arguments leading to our results rely on statistical analysis.

2. Approximation using noiseless, gridded data. Our basic result, Theorem 2.1, provides a bound for a direct approximation, $y = \bar{g}(x)$, to a boundary $y = g(x)$ at a general point $x = t$, using vertex color data on a square grid. One term in the bound depends on the accuracy of a rational approximation to the slope of g at t and requires a little elaboration. To this end, Theorem 2.2 provides a bound to the integral average of that term. Theorem 2.3 applies that result to describe performance of $\bar{g}$ in the L^1 metric.

We begin by noting number-theoretic results which are important to our analysis. Let $u > 0$ be an irrational number, and consider an expansion of u as a continued fraction:

$$u = [a_0; a_1, a_2, \ldots] = a_0 + \cfrac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots}}}. \tag{2.1}$$

Here, $a_0 \geq 0$, $a_n \geq 1$ for $n \geq 1$ and each a_n is an integer. The numbers

$$\frac{p_0}{q_0} = a_0, \qquad \frac{p_1}{q_1} = a_0 + \frac{1}{a_1}, \qquad \frac{p_2}{q_2} = a_0 + \frac{1}{a_1 + \frac{1}{a_2}}, \ldots$$

are called the *convergents* of u. For $n \geq 1$, p_n and q_n are positive and relatively prime, $q_0 = 1$, the sequence $\{q_n(u), n \geq 1\}$ is strictly increasing,

$$\{q_n(q_n + q_{n+1})\}^{-1} < |u - (p_n/q_n)| < (q_n q_{n+1})^{-1}, \tag{2.2}$$

$$\inf_{p,\, 1 \leq q \leq q_n} |u - (p/q)| = |u - (p_n/q_n)|, \tag{2.3}$$

$$p_{2n}/q_{2n} \uparrow u, \qquad p_{2n+1}/q_{2n+1} \downarrow u. \tag{2.4}$$

See, for example, Chapter 9 of Leveque (1956). In the case of rational u, the sequence $\{q_n(u)\}$ is strictly increasing until, for some finite $n_0 = n_0(u)$, we have $q_n(u) = q_{n_0}(u)$ and $p_n(u)/q_n(u) = u$ for all $n \geq n_0$. [This is a definition; other conventions argue that $q_n(u)$ is not defined for such u and n.] Results (2.2)–(2.4) hold for rational u provided that $1 \leq n \leq n_0 - 1$.

In developing our results, it is convenient to assume initially that the curve is monotone, since this avoids awkward notation connected with either positive or negative slopes. However, our main result in this section, Theorem 2.3, does not require monotonicity.

Let $\mathscr{G}(B)$ be the class of all functions g that are strictly increasing and differentiable on $\mathscr{I} = [0, 1]$ and whose first derivative satisfies the Lipschitz condition

$$\left|g'(x_1) - g'(x_2)\right| \leq B\,|x_1 - x_2| \quad \text{for all } x_1, x_2 \in \mathscr{I}.$$

Let $\mathscr{C}$ denote the curve defined by $y = g(x)$, for $x \in \mathscr{I}$, and consider the square grid whose vertices are at points (in^{-1}, jn^{-1}) for pairs of integers i, j with $0 \leq i \leq n$. Let $\mathscr{S} = \mathscr{S}(n)$ denote the set of all such vertices. Color black each vertex that lies above $\mathscr{C}$ and white each vertex that lies below. (Vertices lying on the curve may be regarded as "colorless," or alternatively, as having either color, specified either randomly or deterministically.) We wish to estimate g from the pattern $\mathscr{P} = \mathscr{P}(\mathscr{C}, \mathscr{I})$ of vertex colors.

Suppose we wish to estimate g at a point $t \in \mathscr{I}$. Let $\mathscr{S}_t$ denote the set of all vertices (in^{-1}, jn^{-1}) such that $in^{-1} \in [t-h, t+h] \cap \mathscr{I}$, where $h > 0$ is a bandwidth, and consider the color pattern $\mathscr{P}_t$ among these vertices. If there exists a straight line $\mathscr{L}$ which is consistent with $\mathscr{P}_t$, let $\bar{g}(t)$ be the y coordinate of the point where this line cuts the vertical line $x = t$. (Any line $\mathscr{L}$ consistent with the color pattern in the thin strip $\mathscr{S}_t$ may be chosen; naturally, we do not ask that it be consistent with the pattern in all of $\mathscr{S}$.) If no line is consistent with the pattern, let $\mathscr{L}$ denote any line which minimizes the errors that arise if it is considered to have produced the pattern $\mathscr{P}_t$, with any positive assignment of weights (bounded away from zero and infinity) to a vertex on the wrong side of $\mathscr{L}$ (e.g., a black vertex below $\mathscr{L}$), and zero weight to a vertex on the correct side; and let $\bar{g}(t)$ be the point at which $\mathscr{L}$ cuts the vertical line $x = t$.

In either of these cases, the straight line $\mathscr{L}$ defining $\bar{g}(t)$ might be regarded as a local "line of best fit" to the vertex color pattern in $\mathscr{S}_t$. Note particularly that our definition permits us to define $\bar{g}(t)$ right up to the ends of $\mathscr{I}$—edge effects do not require us to restrict t to the interval $[h, 1-h]$.

THEOREM 2.1. *Let u denote $g'(t)$. There exist positive absolute constants A_1, A_2 and A_3 such that, if $N = N(s)$ is the largest integer for which $q_N(s) \leq A_1 nh$, then*

$$|\bar{g}(t) - g(t)| < A_2 h\left\{q_{N(u)}(u)\, q_{N(u)+1}(u)\right\}^{-1} + Bh^2 \tag{2.5}$$

for all $g \in \mathscr{G}(B)$, $t \in \mathscr{I}$, $0 < h < \frac{1}{2}$ and $nh \geq A_3$.

REMARK 2.1. *Improving the second term on the right in* (2.5). If g has two bounded derivatives on $\mathscr{I}$ then of course $g \in \mathscr{G}(B)$, with $B = \sup|g''|$. A minor modification to the proof of Theorem 2.1 shows that in that case, (2.5) holds true with the term Bh^2 replaced by $\frac{1}{2}B(h,t)\,h^2$, where

$$B(h,t) = \sup_{s\in\mathscr{I}:\,|s-t|\le h} \left|g''(s)\right|.$$

REMARK 2.2. Origins of the two terms in (2.5). The two terms on the right-hand side of (2.5) represent, respectively, the error due to discretization and the error that arises from the fact that $\bar{g}$ is a linear approximation rather than something more sophisticated. In classical statistical formulations of related problems, the second term would be referred to simply as "bias."

An order h^2 term for bias is, of course, typical in the context of approximating twice-differentiable functions using second-order methods, such as local linear ones. See for example Wand and Jones (1995), Chapter 5. The first term on the right-hand side of (2.5) is also close to being best possible. It arises directly from the upper bound given by (2.2). The lower bound there is always greater than half the upper bound, indicating that the upper bound is accurate.

Next we present a bound for the integral of the first term on the right-hand side of (2.5). Here we need to be a little more restrictive about g. Let $\mathscr{G}_1$ be the class of all functions g that are strictly increasing and have two bounded derivatives on $\mathscr{I} = [0,1]$ and satisfy $\inf_{x\in\mathscr{I}} g''(x) > 0$. Generally, let $N(u)$ denote the largest integer such that $q_{N(u)}(u) \le r$; in Theorem 2.1 we considered the special case $r = A_1 nh$.

THEOREM 2.2. *There exists a positive absolute constant* A_4 *such that*

$$\begin{aligned}\int_{\mathscr{I}} &\left[q_{N\{g'(t)\}}\{g'(t)\}\,q_{N\{g'(t)\}+1}\{g'(t)\}\right]^{-1} dt \\ &< A_4 \left|g'(0) - g'(1)\right| \left(\inf_{\mathscr{I}} \left|g''\right|\right)^{-1} r^{-2} (\log r)^2 \end{aligned} \tag{2.6}$$

for all $r \ge 2$ *and all* $g \in \mathscr{G}_1$.

REMARK 2.3. *Method of proof*. A statistical argument is used to derive Theorem 2.2. It involves writing the left-hand side as

$$E\left[q_{N\{g'(T)\}}\{g'(T)\}\,q_{N\{g'(T)\}+1}\{g'(T)\}\right]^{-1},$$

where T has the uniform distribution on $\mathscr{I}$, and bounding the probabilities that the random argument of the expectation operator takes values of specific size.

REMARK 2.4. *The need to avoid points of inflection*. (Here and below we define a point of inflection of f to be any point where f'' vanishes.) Formula (2.6) suggests that the greatest difficulty is caused by points of inflection, where (perhaps only instantaneously) the derivative of g is not changing. An examination of the derivations of Theorems 2.1 and 2.2 shows that this difficulty is real, not an artifact of the method of proof. Indeed, the "$(\inf |g''|)^{-1}$" term on the right-hand side of (2.6) arises through the necessarily large size of $\{q_{N(u)}(u)\, q_{N(u)+1}(u)\}^{-1}$ if u is changing very slowly; see (3.3).

To better appreciate this point, consider the case where g is purely linear. Then g' is constant and g'' vanishes. If the slope of the straight line $y = g(x)$ is rational, then the best possible rate of approximation to g, using all the information in the full vertex color pattern $\mathscr{P}$, is generally only $O(n^{-1})$ (unless the line happens to pass through vertices). Therefore, the convergence rate of virtually $O(n^{-4/3})$ that is achievable from local linear approximation in other circumstances (see Remark 2.6) is not valid here, and so the constant coefficient associated with that rate must be infinite.

REMARK 2.5. *The case of a turning point*. A turning point is much less of a problem than a point of inflection. Although a turning point is specifically excluded by the conditions of Theorem 2.2, once we have derived that result we may take limits as any turning point is approached, without affecting the validity of (2.6). Indeed, combining Theorems 2.1 and 2.2, we may establish a bound that is applicable to approximations of functions in intervals where there do not exist points of inflection, as follows.

Let $\mathscr{G}_2$ be the class of all functions g that have two bounded derivatives on $\mathscr{I}$, and satisfy $\inf_{x\in\mathscr{I}} |g''(x)| > 0$. Define $B(h,t)$ as in Remark 2.1.

THEOREM 2.3. *Let* $A_1, \dots, A_4$ *be as in Theorems* 2.1 *and* 2.2, *and put* $r = A_1 nh$. *Then*,

$$\int_{\mathscr{I}} |\bar{g}(t) - g(t)|\, dt < A_2 A_4\, h\, |g'(0) - g'(1)| \Big(\inf_{\mathscr{I}} |g''|\Big)^{-1} r^{-2} (\log r)^2 + \tfrac{1}{2} h^2 \int_{\mathscr{I}} B(h,t)\, dt \tag{2.7}$$

for all $g \in \mathscr{G}_2$, $0 < h < \frac{1}{2}$ *and* $nh \geq \max(2/A_1, A_3)$.

Theorem 2.3 is a corollary of Theorems 2.1 and 2.2, noting the symmetry that exists between cases where $g' < 0$ and $g' > 0$ and the fact that $g' = 0$ is allowed by taking limits.

REMARK 2.6. *Convergence rate*. The second term on the right-hand side of (2.7) does not depend on n, and is asymptotic to $\frac{1}{2}h^2 \int_{\mathscr{I}} |g''|$ as $h \to 0$. The first term is asymptotic to $C(n^2 h)^{-1} \{\log(nh)\}^2$ as $h \to 0$ and $nh \to \infty$, where

$$C = A_1^{-2} A_2 A_4\, |g'(0) - g'(1)| \Big(\inf_{\mathscr{I}} |g''|\Big)^{-1}.$$

Equating the orders of magnitude of these two terms, and choosing h to solve this identity, we see that if h is taken equal to a constant multiple of $(n^{-1}\log n)^{2/3}$ then

$$\int_{\mathscr{I}} |\bar{g}(t) - g(t)|\, dt = O\big\{(n^{-1}\log n)^{4/3}\big\}. \tag{2.8}$$

This rate is attained uniformly in functions $g \in \mathscr{G}_2$ for which inf $|g''|$ and sup $|g''|$ are bounded away from zero and infinity, respectively.

This convergence rate is not available uniformly in t or even at individual points. For example, when t is a turning point of g, the bandwidth within which the approximation is conducted needs to be of size at least $h = n^{-1/2}$ in order for vertices on both sides of the curve to be contained within a horizontal strip of width h^2. Since h^2 represents the extent to which bias interferes with local linear approximations, it provides a lower bound, as well as an upper bound, to accuracy; and so $h = n^{-1/2}$ produces an error of n^{-1}.

REMARK 2.7. *Performance in other L^p metrics*. Versions of all these results may be derived in other L^p metrics, although for $p > 1$ there are qualifications. When $0 < p < 1$, the best L^p convergence rate achievable by our methods is actually slightly better, by a logarithmic factor, than the pth power of the rate $O\{(n^{-1}\log n)^{4/3}\}$ noted in Remark 2.6. [Recall that for $0 < p < 1$, the L^p norm is defined by

$$\|g_1 - g_2\| = \int_{\mathscr{I}} |g_1(t) - g_2(t)|^p \, dt,$$

and so is not standardized for p.] However, for $p > 1$ the best convergence rate that we are presently able to derive is inferior to $n^{-4/3}$ by a factor n^{α}, where $\alpha = \alpha(p)$ increases with p.

REMARK 2.8. *Local polynomial estimators*. As an alternative to $\breve{g}$, one may define a version $\bar{g}_\gamma$ of the approximant $\bar{g} = \bar{g}_2$ in which the boundary between black and white vertices is approximated by a polynomial of degree $\gamma - 1$. We might expect an analogue of (2.7) to hold for $\bar{g}_\gamma$; for example,

$$\int_{\mathscr{I}} |\bar{g}_\gamma(t) - g(t)| \, dt = O\{h\,(nh)^{-\gamma+\varepsilon} + h^2\} \tag{2.9}$$

for all $\varepsilon > 0$. This would allow the rate of convergence in (2.8) to be improved to

$$\int_{\mathscr{I}} |\bar{g}_\gamma(t) - g(t)| \, dt = O\big(n^{\varepsilon - 2\gamma/(\gamma+1)}\big)$$

for all $\varepsilon > 0$.

Such results seem particularly difficult to derive, however. Just as our proof of (2.7) relies on bounds to integer approximations to linear functions of integers [see (3.2), for example], so (2.9) seems to require bounds on integer approximations to $(\gamma - 1)$th degree polynomials in integers. Our efforts at deriving such results have not enabled us to achieve the convergence rate described by (2.9). Hence, they will not be presented here. Related work on such approximations in more classical contexts includes those described by Schmidt (1980) and Baker (1986).

REMARK 2.9. *Effect of additive noise*. These results may be extended to cases where the observed color is subject to both systematic and stochastic error. For example, suppose the intensity of the noiseless signal may be expressed as

$$f(x, y) = f_1(x, y) + f_2(x, y)\, I\{y \le g(x)\},$$

where f_1, f_2 and g have two bounded derivatives, f_2 is bounded away from zero and g does not have points of inflection. Suppose we observe f on a grid, contaminated by additive noise ε_{ij}, so that the data $Y(i/n, j/n)$ are generated as $Y(i/n, j/n) = f(i/n, j/n) + \varepsilon_{ij}$ for $1 \le i,\ j \le n$. Assume that the ε_{ij}'s are independent and identically distributed random variables with zero mean and finite moment generating function within some neighborhood of the origin. Under these conditions it is possible to construct a local linear estimator $\hat{g}$, for example, based on a wavelet or least-squares diagnostic, that attains the convergence rate at (2.8):

$$E\left\{ \int_{\mathscr{J}} |\hat{g}(t) - g(t)|\, dt \right\} = O\big\{ (n^{-1} \log n)^{4/3} \big\}.$$

An example of the wavelet approach is given in an unpublished technical report obtainable from the authors, while the least-squares method has been treated by Hall and Raimondo (1997).

3. Technical arguments.

PROOF OF THEOREM 2.1. First we treat the case of a fixed straight line on a fixed square grid whose vertices are at integer pairs (as distinct from integer multiples of n^{-1}; thus, we are in effect considering $n = 1$). Let $0 < c_1 < c_2 < \infty$ be constants and temporarily redefine $\mathscr{S}$ as the set of all vertices whose x coordinates lie in $\mathscr{J} = [c_1 m, c_2 m]$. Let $\mathscr{L}$ denote the line with equation $y = ux + v$ and let $\mathscr{K}$ be the set of all candidates $\hat{\mathscr{L}}$ for $\mathscr{L}$ which produce the same vertex color pattern in $\mathscr{S}$ as $\mathscr{L}$. Let $d_m(\mathscr{J}, \mathscr{L}; x_0, \hat{\mathscr{L}})$ denote the absolute value of the difference between the y-coordinates of the places where $\mathscr{L}$ and $\hat{\mathscr{L}}$ intersect the line $x = x_0$, for $x_0 \in \mathscr{J}$, and define

$$d_m(\mathscr{J}, \mathscr{L}) = \sup_{x \in \mathscr{J};\, \hat{\mathscr{L}} \in \mathscr{K}} d_m(\mathscr{J}, \mathscr{L}; x, \hat{\mathscr{L}}).$$

Let $K = K(u)$ be the largest integer such that $q_K(u) \le \frac{1}{4} c_1 m$. [By convention, $q_0 = 1$, and so $K(u)$ is well defined if $m \ge 4/c_1$.]

LEMMA 3.1. *If $c_2/c_1 \ge 16$ then $d_m(\mathscr{J}, \mathscr{L}) < 18\, c_2 m\, \{q_{K(u)}(u)\, q_{K(u)+1}(u)\}^{-1}$, uniformly in choices of v.*

PROOF. Let V_1 be the black vertex in $\mathscr{S}$ that is nearest to $\mathscr{L}$ among all such vertices (with distance measured vertically, here and below) and let $\mathscr{L}_1$ be the line parallel to $\mathscr{L}$ that passes through V_1. Let V_2 be the white vertex in $\mathscr{S}$ that is nearest to $\mathscr{L}_1$ among all such vertices and let $\mathscr{L}_2$ be the line that is parallel to $\mathscr{L}$ and passes through V_2. Then $\mathscr{L}$ lies between $\mathscr{L}_1$ and $\mathscr{L}_2$. Let d_{1m} denote the vertical separation of $\mathscr{L}_1$ and $\mathscr{L}_2$ and let $L = L(u)$ be the largest integer such that $q_L(u) \le \frac{1}{4}(c_2 - 3c_1)m$. Defining $d_{2m} = c_2 m\,(q_L\, q_{L+1})^{-1}$, we prove that

$$d_{1m} < d_{2m}. \tag{3.1}$$

It follows that both V_1 and V_2 are distant less than d_{2m} from $\mathscr{L}$ in the vertical direction. At this point in the proof we require only $c_2/c_1 > 3$, rather than the stronger assumption $c_2/c_1 > 16$.

We may assume without loss of generality that $L = 2k$ is even, in which case we first develop an approximation to $\mathscr{L}$ using white vertices in the vertical direction. For odd L, this particular approximation could be via black vertices, again vertically.

Let v_1 be chosen such that $\mathscr{L}_1$ has equation $y = ux + v_1$ and write $\langle z\rangle$ for the integer part of the positive number z. We may write $v_1 = i_1 - uj_1$ for integers i_1 and j_1. Then

$$d_{1m} = \min_{c_1 m \le j \le c_2 m} (uj + v_1 - \langle uj + v_1\rangle) = \min_{c_1 m \le j \le c_2 m} \{u(j - j_1) - \langle u(j - j_1)\rangle\},$$

which is the vertical distance of V_2 from $\mathscr{L}_1$. The interval $[c_1 m - j_1, c_2 m - j_1]$ contains an interval of length at least $\frac{1}{2}(c_2 - c_1)m$ which does not have the origin as an interior point. Without loss of generality, the interval is entirely nonnegative. Then, with c_3 denoting any positive number strictly less than $c_4 = \frac{1}{2}(c_2 - c_1)$, there exists an integer $j_2 = j_2(u, v) \in [0, \frac{1}{2}(c_1 + c_2)m]$ such that

$$d_{1m} \le \min_{c_3 m + j_2 \le j \le c_4 m + j_2} (uj - \langle uj\rangle) \le c_2 m \min_{c_3 m + j_2 \le j \le c_4 m + j_2} \big(u - j^{-1}\langle uj\rangle\big). \tag{3.2}$$

Since $c_1 < \frac{1}{3}c_2$, then with $c_3 = \frac{1}{4}(c_2 - 3c_1)$ we have $0 < c_3 < \frac{1}{2}(c_2 - 3c_1)$. In this case, $m/q_{2k} \ge 4/(c_2 - 3c_1) > (c_4 - c_3)^{-1}$, implying that the length of the interval $((c_3 m + j_2)/q_{2k}, (c_4 m + j_2)/q_{2k})$ is strictly greater than 1. Therefore, the interval contains an integer l, say, for which $c_3 m + j_2 \le lq_{2k} \le c_4 m + j_2$. By properties (2.2) and (2.4) of convergents, $0 < u - (lp_{2k}/lq_{2k}) < (q_{2k}q_{2k+1})^{-1}$. Now, $r = \langle ulq_{2k}\rangle$ is the largest integer such that $r/lq_{2k} < u$, and so is the integer such that $u - (r/lq_{2k})$ is as small as possible, subject to being positive. Therefore, by properties (2.3) and (2.4) of convergents, $r = lp_{2k}$. Hence, with $j = lq_{2k}$ we have $u - j^{-1}\langle uj\rangle < (q_{2k}q_{2k+1})^{-1}$. Result (3.1) now follows from (3.2).

Now divide the interval $\mathscr{J} = [c_1 m, c_2 m]$ into three (slightly overlapping) parts, $\mathscr{J}_1 = [c_1 m, c_5 m]$, $[c_5 m, c_6 m]$ and $\mathscr{J}_2 = [c_6 m, c_2 m]$, where $c_5 = 4c_1$ and $c_6 = 5c_1$. Let $\mathscr{S}_i$ be the set of all vertices whose x coordinates lie in $\mathscr{J}_i$. Apply the argument above to $\mathscr{J}_1$ and $\mathscr{J}_2$ instead of $\mathscr{J}$ (and so $[c_1 m, c_5 m]$ and $[c_6 m, c_2 m]$ instead of $[c_1 m, c_2 m]$), establishing the existence of vertices $V_1^{(\mathrm{b})}, V_1^{(\mathrm{w})}, V_2^{(\mathrm{b})}$ and $V_2^{(\mathrm{w})}$ such that (1) $V_i^{(\mathrm{b})}$ and $V_i^{(\mathrm{w})}$ both lie in $\mathscr{S}_i$; (2) vertices $V_1^{(\mathrm{b})}$ and $V_2^{(\mathrm{b})}$ are black, while both the other two are white and (3) each of $V_1^{(\mathrm{b})}, V_1^{(\mathrm{w})}, V_2^{(\mathrm{b})}$ and $V_2^{(\mathrm{w})}$ is distant no more than d_{3m} from $\mathscr{L}$, where $d_{3m} =$

$c_2 m(q_K q_{K+1})^{-1}$. [It is at this point that we need the condition $c_2/c_1 > 16$, which implies that $\frac{1}{4}(c_2 - 3c_6) > \frac{1}{4}c_1$ and $c_2/c_6 > 3$. Note too that $c_5/c_1 > 3$. The quantity d_{3m} is an upper bound to the versions of d_{2m} obtained when $\mathscr{J}$ is replaced by $\mathscr{J}_1$ or $\mathscr{J}_2$ in (3.1).]

Let $\mathscr{L}^{(\mathrm{b})}$ [respectively, $\mathscr{L}^{(\mathrm{w})}$] be the line segment joining $V_1^{(\mathrm{b})}$ to $V_2^{(\mathrm{b})}$ [respectively, $V_1^{(\mathrm{w})}$ to $V_2^{(\mathrm{w})}$], and let $\mathscr{K}_1$ denote the set of all lines which pass between $\mathscr{L}^{(\mathrm{b})}$ and $\mathscr{L}^{(\mathrm{w})}$ without cutting either segment. Given an element $\mathscr{M}$ of $\mathscr{K}_1$, let $d_{4m}(x_0)$ be the absolute value of the difference between the y-coordinates of the places where $\mathscr{L}$ and $\mathscr{M}$ intersect the line $x = x_0$. It may be proved after a little geometry that for all $x_0 \in \mathscr{J}$, $d_{4m}(x) \le 18\, d_{3m}$.

The lemma follows from this result, since any candidate $\hat{\mathscr{L}}$ for $\mathscr{L}$ is necessarily an element of $\mathscr{K}_1$, when restricted to having x coordinates in $\mathscr{J}$. □

Now we return to the proof of Theorem 2.1. Recall that $\mathscr{C}$ is the curve whose equation is $y = g(x)$, drawn across a square grid with vertices at pairs of integer multiples of n^{-1}. Let $t \in \mathscr{I}$, and consider the vertex color pattern produced within the thin vertical strip $\mathscr{S}_t$ of width between h and $2h$, consisting of all vertices whose x coordinate lies in the interval $[t-h, t+h] \cap \mathscr{I}$. If $g \in \mathscr{G}(B)$, then for $|x| \le h$ and $t + x \in \mathscr{I}$, $g(t+x) = ux + v + r(t, x)$, where $u = g'(t)$, $v = g(t)$ and $|r(t, x)| \le Bh^2$. Let $\mathscr{L}$ denote the straight line with equation $y = ux + v$ and consider the following upper and lower bounds to $\mathscr{L}$ within $\mathscr{S}_t$: the line $\mathscr{L}^{\mathrm{up}}$, with equation $y = ux + v + Bh^2$, and the line $\mathscr{L}^{\mathrm{low}}$, with equation $y = ux + v - Bh^2$. Within $\mathscr{S}_t$, the curve $\mathscr{C}$ lies between $\mathscr{L}^{\mathrm{up}}$ and $\mathscr{L}^{\mathrm{low}}$.

Next we apply Lemma 3.1 to deduce the accuracy with which either $\mathscr{L}^{\mathrm{up}}$ or $\mathscr{L}^{\mathrm{low}}$ (denoted generically below by $\mathscr{L}^0$) could be approximated from the vertex color pattern which would arise in $\mathscr{S}_t$ if it were present instead of $\mathscr{C}$. We take $c_1 = 1$, $c_2 = 16$ and $m = m(n) = \langle nh/16 \rangle$. (If $t \in [h, 1-h]$ then a more appropriate choice would be $m(n) = \langle 2nh/16 \rangle = \langle nh/8 \rangle$, reflecting the fact that $\mathscr{S}_t$ is of width $2h$ there, and giving a slightly improved bound.) In view of the lemma, the vertical displacement between $\mathscr{L}^0$ and any one of the possible candidates for it, consistent with the vertex color pattern produced by $\mathscr{L}^0$, would not exceed

$$\delta = n^{-1} 18\, c_2 m(n) \left\{ q_{K(u)}(u)\, q_{K(u)+1}(u) \right\}^{-1} < 2h \left\{ q_{K(u)}(u)\, q_{K(u)+1}(u) \right\}^{-1}$$

at each point in $\mathscr{I}_t = [t-h, t+h]$, provided $nh \ge 16$. Here, K may be taken to equal the largest integer such that $q_K(u) \le (nh/64) - 1$. Of course, the vertical displacement between $\mathscr{L}^{\mathrm{up}}$ and $\mathscr{L}^{\mathrm{low}}$ is $2Bh^2$ and both lines are parallel to $\mathscr{L}$.

Let $\mathscr{P}_t$ denote the vertex color pattern defined in $\mathscr{S}_t$ by (1) all vertices below or on $\mathscr{L}^{\mathrm{low}}$ are white, (2) all vertices above or on $\mathscr{L}^{\mathrm{up}}$ are black, and (3) all vertices strictly between $\mathscr{L}^{\mathrm{low}}$ and $\mathscr{L}^{\mathrm{up}}$ are colorless. We shall say that a line does not conflict with $\mathscr{P}_t$ if the vertex color pattern that it produces in $\mathscr{S}_t$ differs from $\mathscr{P}_t$ only at colorless vertices of the latter. Any line which does not conflict with $\mathscr{P}_t$ has its vertical displacement from $\mathscr{L}$ equal to at most $\delta + Bh^2$ at each point in $\mathscr{I}_t$. The "line of best fit" defined in Section 2 does not conflict with $\mathscr{P}_t$ and so lies no further than $\delta + Bh^2$ from $\mathscr{L}$ throughout $\mathscr{I}_t$. In particular, it is no more than this distance from $\mathscr{L}$ at $t \in \mathscr{I}_t$. This completes the proof of the theorem, since (for appropriate choices of the constants A_1, A_2 and A_3), $\delta + Bh^2$ is an upper bound to the right-hand side of (2.5).

PROOF OF THEOREM 2.2. Since the range of integration is bounded away from points where $g'' = 0$, then g' is monotone in the range. Without loss of generality, it is increasing there, and so $f_1 = (g')^{[-1]}$ (the function inverse of g') is also increasing, and $f_1' = \{g''(f_1)\}^{-1} > 0$. Let T have the uniform distribution on $\mathscr{I}$, let U have the uniform distribution on $\mathscr{J} = [g'(0), g'(1)]$ and let V have the distribution on $\mathscr{J}$ with density f proportional to f_1'. Then, with $R(u) = \{q_{N(u)}(u)\, q_{N(u)+1}(u)\}^{-1}$, we have

$$E\big[R\{g'(T)\}\big] = E\{R(V)\} \le \Big(\sup_{\mathscr{J}} f\Big)\big\{g'(1) - g'(0)\big\}\, E\{R(U)\}. \tag{3.3}$$

Next we bound the function R. Define $b_n = q_{n+1}/q_n > 1$. (Here and below we suppress the argument u.) Trivially, $R \le r^{-2} b_N$. Furthermore, if $b_N > r^2$ then, since $q_N \ge 1$, $q_{N+1}^{-1} < r^{-2}$, and so $R < r^{-2}$. Hence, $R < r^{-2}\big\{1 + b_N\, I\big(b_N \le r^2\big)\big\}$. It is known that $q_i \ge 2^{(i-1)/2}$ [e.g., Khintchine (1963), page 18], and so $N(u) \le C_1 \log r + 1$, where $C_1 = 2\,(\log 2)^{-1}$. Furthermore, $b_i = a_{i+1} + b_{i-1}^{-1} < 2a_{i+1}$ [see Khintchine (1963), page 12, for the equality], where $a_1, a_2, \dots$ are the strictly positive integers in the unique continued fraction expansion of u at (2.1). Therefore,

$$R < r^{-2}\left\{1 + 2 \sum_{i=1}^{C_1 \log r + 2} a_i\, I\big(a_i \le r^2\big)\right\}. \tag{3.4}$$

The result $\sup_{i \ge 1} P\{a_i(U) = j\} \le 2\, j^{-2}$, for all $j \ge 1$, may be deduced from formula (57) of Khintchine (1963), page 68. Therefore,

$$E\big[a_i(U)\, I\big\{a_i(U) \le r^2\big\}\big] \le C_2 \log r$$

for all $i \ge 1$ and $r \ge 2$, where C_2 and C_3 denote absolute constants. Substituting into (3.4) we see that $E\{R(U)\} < C_3\, r^{-2}\, (\log r)^2$ for all $r \ge 2$. Substituting into (3.3) we deduce that

$$\int_{\mathscr{I}} R\{g'(t)\}\, dt < C_3\, \{g'(1) - g'(0)\} \Big(\inf_{\mathscr{I}} g''\Big)^{-1} r^{-2}\, (\log r)^2.$$

The theorem follows from this result. □

REFERENCES

BAKER, R. C. (1986). *Diophantine Inequalities*. Oxford Univ. Press.

FREEMAN, H. (1970). Boundary encoding and processing. In *Picture Processing and Psychopictorics* (B. S. Lipkin and A. Rosenfeld, eds.) Academic Press, New York.

HALL, P. and RAIMONDO, M. (1997). Approximating a line thrown at random onto a grid. *Ann. Appl. Probab.* **7** 648–665.

KHINTCHINE, A. YA. (1963). *Continued Fractions*. Noordhoff, Groningen.

KOROSTELEV, A. P. and TSYBAKOV, A. B. (1993). *Minimax Theory of Image Reconstruction. Lecture Notes in Statist.* **82**. Springer, Berlin.

LEVEQUE, W. J. (1956). *Topics in Number Theory* I. Addison-Wesley, Reading, MA.

SCHMIDT, W. M. (1980). *Diophantine Approximation.* Springer, Berlin.

WAND, M. P. and JONES, M. C. (1995). *Kernel Smoothing.* Chapman and Hall, London.

WORRING, M. and SMEULDERS, W. M. (1995). Digitized circular arcs: characterization and parameter estimation. *IEEE Trans. Pattern Anal. Machine Intell.* **17** 587–598.

CENTRE FOR MATHEMATICS
AND ITS APPLICATIONS
AUSTRALIAN NATIONAL UNIVERSITY
CANBERRA, ACT 0200
AUSTRALIA
E-MAIL: halpstat@pretty.anu.edu.au