JOURNAL OF APPROXIMATION THEORY 1, 66–76 (1968)

Efficiency of Polynomials on Sequences

D. J. NEWMAN[1]

Belfer Graduate School of Science, Yeshiva University, New York, New York 10033

Let S be a compact set of real numbers and let P be any n-dimensional space of continuous functions defined on S. The "approximation index" for P may be defined as

$$\max_{f\in\mathscr{S}} \min_{p\in P} \max_{x\in S} |f(x)-p(x)|$$

(where $\mathscr{S}$ is the class of $f(x)$ such that $|f(x)-f(y)| \leqslant |x-y|$ for all x, y).

It is known [*2*] that the smallest possible value for this approximation index over all possible choices of P is given exactly by $\frac{1}{2}\epsilon_n(S)$, where $\epsilon_n(S)$, the "massivity" of S, is defined by

$$\epsilon_n(S) = \max_{y_1, y_2, \ldots y_{n+1}\in S} \; \min_{i\neq j} |y_i - y_j|.$$

This massivity, thus defined in a purely geometric way, gives the absolute lower bound for the degrees of approximation by various function spaces. For example, when S is the unit interval, ϵ_n is identically equal to $1/n$, and thus the lower bound is $1/2n$. Viewed in this light, Jackson's theorem [*1*] takes on a special significance. It says here that for P the polynomials of degree $< n$, the approximation index is $\leqslant 1/n$ and this is, aside from some constant factor, the same as the absolute lower bound $1/2n$! In loose language we can say that on the unit interval polynomials are essentially as useful for approximation as any other space of functions.

DEFINITION. We say that polynomials are efficient on the set S if the approximation index for the polynomials of degree $< n$ is bounded by $A\epsilon_n(S)$ for all n with A independent of n.

Thus Jackson's theorem says that polynomials are efficient on intervals. On what other sets?

In this context sets of positive measure offer no new interest. They are sufficiently like intervals for efficiency to follow directly from Jackson's theorem. We go to the opposite extreme then and consider sets S which are mere sequences. Indeed, our sets will all consist of 0 and a sequence $\{x_n\}$ which decreases to 0. We are able to show that polynomials are efficient for the

[1] Supported by NSF Grant GP 4391.

common garden variety sequences such as $1/n^\alpha$, $\log^\beta(n+1)/n^\alpha$, $1/2^n$, $1/n!$, etc., and indeed, for all sequences with a "regular" growth pattern. We will then show by a counterexample that if a sequence does not have sufficiently "regular" growth, efficiency can fail.

Our main positive result is

THEOREM 1. *Polynomials are efficient on* $\{0;x_n\}$ if $\{x_n\}$ *is logarithmically convex.*

Since this does not include sequences such as $1/n!$, we will also provide a complementary result, namely

THEOREM 2. *Polynomials are efficient on* $\{0;x_n\}$ *if* $x_{n+1}/x_n < c < 1$.

(Together these theorems capture all the nameable sequences!)

It is pleasant to look at special cases. For example, when $x_n = 1/n$ we obtain the following: For $f(x) \in \mathscr{S}$ there exists $P(x)$ of degree $< n$ such that

$$\left|f\left(\frac{1}{k}\right) - P\left(\frac{1}{k}\right)\right| \leqslant \frac{A}{n^2} \qquad \text{for all } k = 1,2,3,\ldots.$$

By the usual trick of preapproximation by polygonal functions, this can be elevated to: For any continuous $f(x)$ there exists $P(x)$ of degree $< n$ such that

$$\left|f\left(\frac{1}{k}\right) - P\left(\frac{1}{k}\right)\right| \leqslant A\omega_f\left(\frac{1}{n^2}\right) \qquad \text{for all } k,$$

(ω_f denoting the modulus of continuity of $f(x)$).

Similarly, we can produce an endless variety of such theorems. In terse form some of these read

$$\left|f\left(\frac{1}{\log k}\right) - P\left(\frac{1}{\log k}\right)\right| \leqslant A\omega_f\left(\frac{1}{n\log n}\right),$$

$$\left|f\left(\frac{1}{2^k}\right) - P\left(\frac{1}{2^k}\right)\right| \leqslant A\omega_f\left(\frac{1}{2^n}\right),$$

$$|f(\exp -k^{1/2}) - P(\exp -k^{1/2})| \leqslant A\omega_f\left(\frac{1}{n^{1/2}}\exp -n^{1/2}\right),$$

$$|f(k^{-\alpha}) - P(k^{-\alpha})| \leqslant A\omega_f(n^{-(1+\alpha)}),$$

$$\left|f\left(\frac{1}{k!}\right) - P\left(\frac{1}{k!}\right)\right| \leqslant A\omega_f\left(\frac{1}{n!}\right),$$

and of course, in all cases these are best possible estimates!

Since the proof of Theorem 1 is a bit long, let us give this preview of it. Our job is to produce a good approximation to a given $f(x) \in \mathscr{S}$. We do this by obtaining an interpolating polynomial at the first few points $x_1, x_2, \ldots, x_m$

of S, and then a Jackson polynomial on the whole remaining interval $[0, x_{m-1}]$. These two polynomials are then fused together to produce the required approximator. The right choice of m and the "fusing" procedure constitute the bulk of the proof.

Before we can accomplish this project, however, we require certain preliminaries. Our first lemma yields a handy way of determining the exact order of magnitude of ϵ_n. (For example, it gives $1/n\log n$ for the sequence $1/\log n$, $1/n^{1/2}\exp-n^{1/2}$ for the sequence $\exp-n^{1/2}$, $1/2^n$ for $1/2^n$, etc.)

Let us define

$$\delta_n = \min_{1\leqslant k\leqslant n} \frac{x_k}{n-k+1}.$$

LEMMA 1. *We always have $\epsilon_n \leqslant \delta_n$. If the sequence x_n is convex, then we have in addition, $\epsilon_n \geqslant \frac{1}{2}\delta_n$.*

Proof: Let there be given any $n+1$ members of S, and let k be any integer in $[1,n]$. Since the interval $(x_k, x_1]$ contains at most $k-1$ of these members, the interval $[0, x_k]$ must contain at least $n-k+2$ of them. By the "pigeon hole principle", two of these must have mutual distance $\leqslant x_k/n-k+1$. It follows that $\epsilon_n \leqslant x_k/n-k+1$, and since k was arbitrary, that $\epsilon_n \leqslant \delta_n$.

Now assume that the sequence x_n is convex. Choose j so that $x_j - x_{j+1} < \frac{1}{2}\delta_n \leqslant x_{j-1} - x_j$. Since $x_j/n-j+1 \geqslant \delta_n$, it follows that all the intervals

$$I_m = [x_j - m\delta_n, x_j - (m-\tfrac{1}{2})\delta_n], \qquad m = 1,2,3,\ldots,n-j+1$$

lie in $[0, x_j]$. Convexity insures that any subinterval of $[0, x_j]$ of length $\frac{1}{2}\delta_n$ contains a member of S. In particular, we conclude that each I_m contains such a member which we call y_m. The points $x_1, x_2, \ldots, x_j; y_1, y_2, \ldots, y_{n-j+1}$ are $n+1$ in number and they clearly have all mutual distances $\geqslant \frac{1}{2}\delta_n$. This proves that $\epsilon_n \geqslant \frac{1}{2}\delta_n$.

(We remark that a logarithmically convex sequence is *a priori* a convex one, so that this lemma is applicable to our case.) Next we need a general result concerning logarithmically convex sequences.

LEMMA. *Let $\{x_n\}$ be logarithmically convex and decreasing. If $x_j \geqslant \exp\pi(\frac{2}{3})^{1/2} x_k$, then*

$$\prod_{\substack{i=1\\ i\neq j}}^{k} \left|1 - \frac{x_j}{x_i}\right| \geqslant 1.$$

If, moreover, $x_j \geqslant e^{\pi} x_k$, then

$$\prod_{\substack{i=1\\ i\neq j}}^{k} \left|1 - \frac{x_j}{x_i}\right| \geqslant \frac{1}{3}\frac{x_j}{x_k} e^{(k-j-1)/2}.$$

Proof. Call $x_j/x_k = e^{\lambda}$ and $k-j=\mu$. Logarithmic convexity insures

$$\frac{x_j}{x_i} \geqslant \left(\frac{x_j}{x_k}\right)^{i-j/k-j} = e^{\lambda/\mu(i-j)} \qquad \text{for } i>j. \tag{1}$$

In particular $x_j/x_{j+1} \geqslant e^{\lambda/\mu}$ and so, again by logarithmic convexity, $x_j/x_{j-1} \leqslant x_{j+1}/x_j \leqslant e^{-\lambda/\mu}$. Yet another application of logarithmic convexity now gives $x_j/x_i \leqslant (x_j/x_{j-1})^{j-i}$ for $i<j$. Combining these two, yields

$$\frac{x_j}{x_i} \leqslant e^{\lambda/\mu(i-j)} \qquad \text{for } i<j. \tag{2}$$

Applying (1) and (2) to the product (call it P) in question, produces the lower bound

$$\begin{aligned}
&\prod_{i=1}^{j-1}(1-e^{\lambda/\mu(i-j)}) \prod_{i=j+1}^{k}(e^{\lambda/\mu(i-j)}-1)\\
&\quad=\prod_{i=1}^{j-1}(1-e^{\lambda/\mu(i-j)}) \prod_{i=j+1}^{k}(1-e^{\lambda/\mu(j-i)}) \prod_{i=j+1}^{k} e^{\lambda/\mu(i-j)}\\
&\quad=\prod_{\nu=1}^{j-1}(1-e^{-\lambda\nu/\mu}) \prod_{\nu=1}^{\mu}(1-e^{-\lambda\nu/\mu})\, e^{\lambda/\mu(1+2+3+\dots+\mu)}\\
&\quad\geqslant\prod_{\nu=1}^{\infty}(1-e^{-\lambda\nu/\mu})^2\, e^{\lambda/2(\mu+1)}.
\end{aligned}$$

Now we borrow a simple estimate from the theory of partitions [3], namely

$$\prod_{\nu=1}^{\infty}(1-e^{-t\nu}) \geqslant e^{-\pi^2/6t}.$$

Applying this to the above gives, as our lower bound

$$P \geqslant e^{\lambda/2(\mu+1)-\mu(\pi^2/3\lambda)}. \tag{3}$$

If $\lambda \geqslant \pi(\frac{2}{3})^{1/2}$, then $\lambda/2 - \pi^2/3\lambda \geqslant 0$, and this is surely $\geqslant 1$.
(We remark that the constant $\exp \pi(\frac{2}{3})^{1/2}$ is best possible for this result.)
If $\lambda \geqslant \pi$, then we have by (3),

$$\begin{aligned}
P &\geqslant e^{\lambda/2(\mu+1)-\mu(\pi^2/3\lambda)} = e^{\lambda-(\pi^2/3\lambda)}\, e^{(\mu-1)(\lambda/2-\pi^2/3\lambda)}\\
&\geqslant e^{\lambda-(\pi/3)}\, e^{(\mu-1)\pi/6} \geqslant \frac{e^{\lambda}}{3} e^{(\mu-1)/2} = \frac{1}{3}\frac{x_j}{x_k} e^{(k-j-1)/2}.
\end{aligned}$$

The proof is complete.

We can now give the proof of Theorem 1. Choose k so that $x_k/(n-k+1) = \delta_n$, and then choose m so that $x_m \geqslant e^{\pi} x_k > x_{m+1}$ (where we adopt the convention $x_0 = \infty$).

Now introduce the Lagrange interpolation polynomials $p_j(x)$ for $j = 1, 2, 3, \ldots, m$. These are given by

$$p_j(x) = \prod_{\substack{i=1\\ i\neq j}}^{k} \frac{x - x_i}{x_j - x_i}, \qquad j \leqslant m. \tag{4}$$

If x lies in $[0, x_k]$, then

$$|p_j(x)| \leqslant |p_j(0)| = 1 \Big/ \prod_{\substack{i=1\\ i\neq j}}^{k} \left|1 - \frac{x_j}{x_i}\right|.$$

Lemma 2 is applicable because of our choice of m, and we obtain

$$|p_j(x)| \leqslant \frac{3x_k}{x_j} e^{(j+1-k)/2}, \qquad \text{for } j \geqslant m \text{ and } 0 \leqslant x \leqslant x_k. \tag{5}$$

Next we introduce modified interpolation polynomials by setting

$$q_j(x) = \left(\frac{x}{x_j}\right)^K p_j(x), \qquad \text{for } j \leqslant m, \text{ where } K = \left[\frac{n-k+1}{2}\right]. \tag{6}$$

The crucial step in our proof is the establishment of the following important inequality for these modified interpolators.

LEMMA 3. *For $x = x_{m+1}, x_{m+2}, \ldots$ we have*

$$\sum_{j=1}^{m} x_j |q_j(x)| \leqslant 8\delta_n$$

Proof. Since each of the q_j vanishes at all the points $x_{m+1}, x_{m+2}, \ldots, x_k$, we need only prove our inequality for $x = x_{k+1}, x_{k+2}, \ldots$. Indeed, we will prove that it holds throughout $[0, x_k]$.

For this we need the following estimate:

$$\left(\frac{x_k}{x_j}\right)^K \leqslant \frac{1}{n-k+1}. \tag{7}$$

Namely

$$\frac{x_k}{x_j} \leqslant \frac{x_k}{x_m} \leqslant e^{-\pi} \quad \text{and} \quad e^{-\pi K} \leqslant \frac{1}{2K+1} \leqslant \frac{1}{n-k+1}.$$

Thus, by (7) and (6), we have for $0 \leqslant x \leqslant x_k, j \leqslant m$,

$$|q_j(x)| \leqslant \left(\frac{x_k}{x_j}\right)^K |p_j(x)| \leqslant \frac{|p_j(x)|}{n-k+1}.$$

Applying (5) to this gives us the estimate

$$|q_j(x)| \leqslant \frac{1}{n-k+1} \cdot \frac{3x_k}{x_j} \cdot e^{(j+1-k)/2} \qquad \text{for } 0 \leqslant x \leqslant x_k, j \leqslant m. \tag{8}$$

Now multiply by x_j and sum over j. The result is that, throughout $[0, x_k]$,

$$\sum_{j=1}^{m} x_j |q_j(x)| \leqslant \frac{3x_k}{n-k+1} \sum_{j=1}^{m} e^{(j+1-k)/2}$$
$$= 3\delta_n \sum_{j=1}^{m} e^{(j+1-k)/2} \leqslant 3\delta_n \sum_{j=1}^{k-1} e^{(j+1-k)/2}$$
$$\leqslant 3\delta_n \sum_{\nu=0}^{\infty} e^{-\nu/2} = \frac{3\delta_n}{1-e^{-1/2}} \leqslant \frac{3\delta_n}{1-(1-\frac{1}{2}+\frac{1}{8})} = 8\delta_n,$$

and the lemma is proved.

We are now in a position to produce our "good" polynomial approximation to the given $f(x) \in \mathscr{S}$. Namely, choose $q(x)$ as its best Tchebychev approximator of degree $[(n-k)/2]$ over the interval $[0, x_{m+1}]$, and write

$$P(x) = q(x) + \sum_{j=1}^{m} (f(x_j) - q(x)) q_j(x). \tag{9}$$

This polynomial is immediately seen to interpolate $f(x)$ exactly at x_1, x_2, $\ldots$, x_m. Also its degree is bounded by

$$\left[\frac{n-k}{2}\right] + \left(\left[\frac{n-k+1}{2}\right] + k - 1\right) = n-1.$$

Now let $s > m$ and obtain, from (9),

$$|P(x_s) - f(x_s)| \leqslant |q(x_s) - f(x_s)| + \sum_{j=1}^{m} |f(x_j) - q(x_s)| |q_j(x_s)|. \tag{10}$$

An application of Jackson's theorem [*1*] yields

$$|q(x) - f(x)| \leqslant \frac{x_{m+1}}{[(n-k)/2]+1} \leqslant \frac{2x_{m+1}}{n-k+1} \text{ throughout } [0, x_{m+1}], \tag{11}$$

and among other things this tells us that

$$|q(x_s) - f(x_j)| \leqslant |q(x_s) - f(x_s)| + |f(x_s) - f(x_j)|$$
$$\leqslant \frac{2x_{m+1}}{n-k+1} + x_j \leqslant 3x_j \qquad \text{for } j \leqslant m < s. \tag{12}$$

Inserting (11) and (12) into (10) gives

$$|P(x_s) - f(x_s)| \leqslant \frac{2x_{m+1}}{n-k+1} + 3 \sum_{j=1}^{m} x_j |q_j(x_s)|.$$

An application of Lemma 3 now yields

$$|P(x_s) - f(x_s)| \leqslant \frac{2x_{m+1}}{n-k+1} + 24\delta_n. \tag{13}$$

Recalling that $x_{m+1} < e^{\pi} x_k < (47/2) x_k$ allows us to conclude from (13) that

$$|P(x_s) - f(x_s)| \leqslant 71\delta_n. \tag{14}$$

Finally, by Lemma 1, (14) becomes $|P(x_s) - f(x_s)| \leqslant 142\epsilon_n$ and the proof is complete.

Note the remarkable fact that we obtain the same constant (142) for all sequences! We not only have efficiency, but "uniform" efficiency! Nothing is more than 284 times as good as polynomials (that is, for approximation on logarithmically convex sequences).

The proof of Theorem 2 is a simplified revision of the proof of Theorem 1.

Choose $P(x)$ as that polynomial of degree $< n$ which interpolates $f(x)$ at the points $x_1, x_2, \ldots, x_{n-1}$; 0 ($f(x)$ a given member of $\mathscr{S}$).

We have, identically,

$$P(x) = x \sum_{j=1}^{n-1} \frac{f(x_j) - f(0)}{x_j} \prod_{\substack{i=1 \\ i \neq j}}^{n-1} \frac{x - x_i}{x_j - x_i} + f(0). \tag{15}$$

Throughout $[0, x_n]$ then, we obtain the estimate

$$|P(x) - f(0)| \leqslant x_n \sum_{j=1}^{n-1} \prod_{\substack{i=1 \\ i \neq j}}^{n-1} \left|1 - \frac{x_j}{x_i}\right|^{-1}. \tag{16}$$

Now note that,

$$\text{for } i < j, \ \frac{x_j}{x_i} \leqslant c^{i-j}, \text{ while for } i > j, \ \frac{x_j}{x_i} \geqslant c^{i-j}, \tag{17}$$

so that we obtain

$$\prod_{\substack{i=1 \\ i \neq j}}^{n-1} \left|1 - \frac{x_j}{x_i}\right| \geqslant \prod_{j < i < n} (1 - c^{i-j}) \prod_{i < j} (c^{i-j} - 1) \tag{18}$$

$$\geqslant \prod_{\nu=1}^{\infty} (1 - c^{\nu})^2 c^{-\binom{n-j}{2}}.$$

If we use (18) in (16), we find

$$|P(x) - f(0)| \leqslant x_n \prod_{\nu=1}^{\infty} (1 - c^{\nu})^{-2} \sum_{j=1}^{n-1} c^{-\binom{n-j}{2}} \leqslant A x_n \tag{19}$$

where

$$A = \prod_{\nu=1}^{\infty} (1 - c^{\nu})^{-2} \sum_{\nu=1}^{\infty} c^{\binom{\nu}{2}}.$$

Thus, again throughout $[0, x_n]$, we have

$$|P(x) - f(x)| \leqslant |P(x) - f(0)| + |f(0) - f(x)| \leqslant A x_n + x_n = (A+1) x_n. \tag{20}$$

Since $P(x)$ interpolates $f(x)$ at $x_1, x_2, \ldots, x_{n-1}$, this inequality persists for all x_s! Theorem 2 follows from this and the simple observation that $\epsilon_n \geqslant \alpha x_n$ where $\alpha = \min((1/c) - 1, 1)$.

We now construct our counterexample. Since the details are quite intricate it might be useful if we indicated the simple idea behind it. This is just the fact that the massivity of a set can be thrown off (made small) for certain n by the existence of *just one* unusually close pair of points. The approximation index for polynomials, however, is not so local and will not be made small by just one such pair. By exploiting this fact then, we can arrange to have the polynomial approximation index much larger than the massivity number on occasion.

Here are the details:

For convenience we will write $c_k = 2^{2^k}$. Set

$$t_m = \frac{1}{c_{k+2}} \qquad \text{for } c_{k-1} \leqslant m < c_k,$$
$$k = 1, 2, 3, \ldots,$$

and let S be the set consisting of 0 and the numbers

$$x_j = \sum_{m>j} t_m.$$

Surely these numbers decrease to 0 (even convexly!) Now choose $n = c_i - 1$ (i large). We claim first that, for this n, we have

$$\epsilon_n \leqslant \frac{2c_{i+1}}{c_{i+3}}.$$

Indeed, it is clear that $\epsilon_n \leqslant x_n$ (by Lemma 1, e.g.). For this particular case then,

$$\epsilon_n \leqslant \sum_{m=c_i}^{\infty} t_m \leqslant \sum_{k=i+1}^{\infty} \frac{c_k - c_{k-1}}{c_{k+2}} \leqslant \frac{c_{i+1}}{c_{i+3}}\left(1 + \frac{1}{2} + \frac{1}{4} + \ldots\right)$$
$$= \frac{2c_{i+1}}{c_{i+3}}.$$

Now define $f(x)$ as follows: For

$$c_{i-1} - 1 \leqslant j < c_i, \quad f(x_j) = \frac{(-1)^{j-1}}{2c_{i+2}},$$

and for

$$c_i \leqslant j < c_i + c_{i-1}, \quad f(x_j) = \frac{1}{2c_{i+2}}.$$

This function satisfies $|f(x) - f(y)| \leqslant |x - y|$, and so can be extended (e.g. linearly) to become a member of $\mathcal{S}$.

Our main effort will be to prove that for $p(x)$ any polynomial of degree $< n$, we must have $|f(x_j) - p(x_j)| \geqslant 1/6c_{i+2}$ for some j, $c_{i-1} - 1 \leqslant j < c_i + c_{i-1}$.

This will surely prove our nonefficiency assertion since $1/6c_{i+2}$ is enormously larger than $2c_{i+1}/c_{i+3}$.

But first let us pretty things up with some normalization, viz. set

$$F(x) = 2c_{i+2} f\left(x_{c_{i-1}} + \frac{x}{c_{i+2}}\right).$$

Then $F(j) = (-1)^j$ for $j = 0, 1, 2, \ldots, c_i - c_{i-1}$, and $F(-k/c_{i+2}) = 1$ for $k = 1, 2, \ldots, c_{i-1}$. Thereby our problem is exactly expressed as the following:

THEOREM 3. *Let $p(x)$ be of degree $< c_i$. Either there is a $j = 0, 1, 2, \ldots, c_i - c_{i-1}$ for which $|(-1)^j - p(j)| \geqslant \frac{1}{3}$, or there is a $k = 1, 2, \ldots, c_{i-1}$ for which $|1 - p(-k/c_{i+2})| \geqslant \frac{1}{3}$.*

Indeed, we will prove a generalization:

THEOREM 4. *Let r, s, N be any positive integers and*

$$\delta = \frac{\int_0^1 [(1+u^N)^r - (1-u^N)^r](1-u)^{s-1}\, du}{\int_0^1 [(1+u^N)^r (1-u)^{s-1} + (1-u^N)^r (1+u)^{s-1}]\, du}.$$

If $p(x)$ is any polynomial of degree $< r + s$ then either $|(-1)^j - p(j)| \geqslant \delta$ for some $j = 0, 1, 2, \ldots, r$, or $|1 - p(-k/N)| \geqslant \delta$ for some $k = 1, 2, \ldots, S$.

To see that Theorem 4 really includes Theorem 3, choose $r = c_i - c_{i-1}$, $s = c_{i-1}$, $N = c_{i+2}$. All we need show is that the resulting δ is $\geqslant \frac{1}{3}$. We note in fact, that for these values

$$2\int_0^1 (1-u^N)^r (1+u)^{s-1}\, du < \int_0^1 (1+u^N)^r (1-u)^{s-1}\, du.$$

Indeed, the left side is bound by 2^s. The right side is

$$> \int_{1-2/N}^{1-1/N} > \frac{1}{N}\left(\frac{11}{10}\right)^r \frac{1}{N^{s-1}} = \frac{1}{N^s}\left(\frac{11}{10}\right)^r$$

and of course $(11/10)^r > (2N)^s$.

LEMMA. *There exists a polynomial $P(x)$ of degree $< r + s$ for which*

$$P(j) = (-1)^j (1-\delta), \qquad \text{for } j = 0, 1, 2, \ldots, r,$$

$$P\left(-\frac{k}{N}\right) = 1 - (-1)^k \delta, \qquad \text{for } k = 1, 2, \ldots, s.$$

Proof. Consider the function $G(k)$, $k = 1, 2, \ldots, s$, given by

$$G(k) = \frac{(-1)^{r+1}}{N(r+1)} \binom{-k/N}{r+1}^{-1} \left[1 - \delta(-1)^k - (1-\delta) \sum_{j=0}^{r} (-2)^j \binom{-k/N}{j}\right].$$

We can express this $G(k)$ as a simple definite integral. Indeed, by the usual Beta-function identities we have, for $x > 0$,

$$\binom{-x}{j} \Big/ \binom{-x}{r+1} = -(r+1)\binom{r}{j} \int_0^1 t^{j+x-1}(t-1)^{r-j}\, dt.$$

In particular, for $j = 0$, this becomes

$$1 \Big/ \binom{-x}{r+1} = (-1)^{r+1}\,(r+1) \int_0^1 t^{x-1}(1-t)^r\, dt.$$

Hence

$$\binom{-x}{r+1}^{-1} \sum_{j=0}^{r} (-2)^j \binom{-x}{j} = (r+1)\,(-1)^{r+1} \int_0^1 \sum_{j=0}^{r} \binom{r}{j} (2t)^j (1-t)^{r-j}\, t^{x-1}\, dt$$

$$= (r+1)\,(-1)^{r+1} \int_0^1 (1+t)^r\, t^{x-1}\, dt,$$

and so,

$$G(k) = \int_0^1 \Big[[1-\delta(-1)^k]\,(1-t)^r - [1-\delta]\,(1+t)^r\Big]\, t^{k/N-1}\, dt$$

or, with $u = t^{1/N}$,

$$G(k) = \int_0^1 \Big[[1-\delta(-1)^k]\,(1-u^N)^r - [1-\delta]\,(1+u^N)^r\Big]\, u^{k-1}\, du.$$

Next we compute the $(s-1)$th difference of $G(k)$, viz.

$$\sum_{k=1}^{s-1} (-1)^{k-1} \binom{s-1}{k-1}\, G(k).$$

This equals

$$\int_0^1 \sum_{k=1}^{s} \binom{s-1}{k-1} \{[(-u)^{k-1} + \delta u^{k-1}]\,(1-u^N)^r - (1-\delta)\,(1+u^N)^r\,(-u)^{k-1}\}\, du$$

$$= \int_0^1 \Big[[(1-u)^{s-1} + \delta(1+u)^{s-1}]\,(1-u^N)^r - (1-\delta)\,(1+u^N)^r\,(1-u)^{s-1}\Big]\, du,$$

and this, by our very choice of δ, is equal to 0.

So the $(s-1)$th difference of $G(k)$ is 0 and therefore $G(k)$ can be extended to be a polynomial of degree $< s-1$.

We need now simply choose, for our $P(x)$

$$P(x) = (1-\delta) \sum_{j=0}^{r} (-2)^j \binom{x}{j} + N(-1)^{r+1}\,(r+1) \binom{x}{r+1}\, G(-Nx).$$

Direct verification shows that this choice satisfies all the requirements of our lemma.

Proof of Theorem 4. Standard sign change counting. If $P(x)$ violated our assumption then we would have to have $p(0) > P(0)$, $p(1) < P(1)$, and, in general, $p(j) > P(j)$ for even j, $p(j) < P(j)$ for odd j, $j = 0, 1, 2, \ldots, r$. Also, we must have

$$p\left(-\frac{1}{N}\right) < P\left(-\frac{1}{N}\right), p\left(-\frac{2}{N}\right) > P\left(-\frac{2}{N}\right),$$

and in general $p(-k/N) > P(-k/N)$ for even k, $p(-k/N) < P(-k/N)$ for odd k; $k - 1, 2, \ldots, s$. This would force $p(x) - P(x)$ to have $r + s + 1$ sign changes (and hence $r + s$ zeroes) and this is impossible for a nontrivial polynomial of degree $<r + s$. Q.E.D.

A close examination of this counterexample, and in fact the simple idea behind it, shows that the deviation between ϵ_n and the polynomial approximation index is large only on rare occasions. Can one produce an S for which this vast deviation occurs for all n? Or must there always be a subsequence on which polynomials are efficient?

References

1. D. Jackson, "The Theory of Approximation". American Mathematical Society Colloquium Publications, Vol. XI, New York, 1930.
2. G. G. Lorentz, Lower bounds for the degree of approximation. *Trans. Am. Math. Soc.* **97** (1960), 25–34.
3. D. J. Newman, Simplified proof of the Partition formula. *Mich. Math. J.* **9** (1962), 283–287.