JOURNAL OF APPROXIMATION THEORY 18, 50–56 (1976)

Limits of Generalized Polynomials with Nonnegative Coefficients

ROBERT WHITLEY

University of California at Irvine, Irvine, California 92664

Communicated by E. W. Cheney

Received October 18, 1974

INTRODUCTION

The starting point of this paper is a startling theorem, based on results of S. N. Bernstein, and presented by Widder in [10, Theorem 9b]: If f on $[0, 1]$ is the uniform limit of polynomials with nonnegative real coefficients, then f is the restriction of a function analytic in the unit disk. We will consider generalized polynomials in given real functions $\{f_n\}$, namely, finite sums $\sum a_n f_n$ with non-negative real coefficients a_n, and will show that with some conditions on the functions $\{f_n\}$ we can obtain results similar to those in the case $f_n(x) = x^n$. The generalization we obtain requires only simple conditions on the functions $\{f_n\}$, and the Bernstein–Widder results follow directly. It is surprising that this can be done rather easily.

DEFINITION. Let S be a topological space. We will consider a sequence $\{f_n\}$ of functions on S with the following properties.

1. There is a constant c with $c \leqslant f_n(s) \leqslant 1$, for all s in S and $n = 0, 1, 2, \ldots$.
2. There is a nonvoid subset F of S with empty interior, such that $f_n(s) = 1$ for each s in F and $n = 0, 1, 2, \ldots$.
3. For s not in $F, f_n(s) \to 0$.

A (generalized) polynomial in $\{f_n\}$ with (real) coefficients $\{b_j\}$ has the form $\sum b_n f_n$, where only for a finite number of n's the coefficient b_n is nonzero.

THEOREM 1. *Let S be a topological space and $\{f_n\}$ a sequence of functions on S satisfying the conditions of the definition. Assume that f, a real function on S, is the pointwise limit of polynomials $\{P_n\}$ in the functions $\{f_n\}$ with non-negative coefficients. Then*:

I. *There is a convergent series of positive terms* $\sum a_n$, *with*

$$f(s) = \sum a_n f_n(s), \tag{*}$$

uniformly in S–F.

II. *If, in addition,* f *is continuous in* F, *then* (*) *holds uniformly in* S. *Moreover, there is a subsequence of* $\{P_n\}$, *which converges to* f *uniformly in* S.

Proof. We suppose that f is the pointwise limit of a sequence $\{P_n\}$, where each P_n has the form $P_n(s) = \sum a_j^n f_j(s)$, $a_0^n, a_1^n, a_2^n, \ldots$, a sequence of non-negative numbers with only a finite number of nonzero terms. For s in F, $P_n(s) = \sum a_j^n \to f(s)$, so there is a constant M with $\sum a_j^n \leqslant M$ for $n = 0, 1, 2, \ldots$. The sequences $x_n = (a_0^n, a_1^n, a_2^n, \ldots)$, as elements of the Banach space l^1, are uniformly bounded in norm. Regarding l^1 as the conjugate of the space c_0 there is a subsequence converging to $(a_0, a_1, a_2, \ldots)$ in the weak* topology. For simplicity we will also call this subsequence $\{(a_0^n, a_1^n, \ldots)\}$. For s not in F, $\{f_n(s)\}$ belongs to c_0, so $\sum a_j^n f_j(s) \to \sum a_j f_j(s)$; thus $f(s) = \sum a_j f_j(s)$. We remark that $\sum a_j \leqslant M$; consequently the convergence of $\sum_1^n a_j f_j(s)$ to $f(s)$ is uniform.

So far very few properties of the functions $\{f_n\}$ have been used. It would have sufficed to have the sequence uniformly bounded in absolute value, $f_n(s) \to 0$ for s not in F, and $f_n(s_0) \geqslant a > 0$, $n = 0, 1, 2, \ldots$, for some s_0 in F.

Now suppose that f is continuous at a point s in F. Given this s and $\epsilon > 0$, there is a point t in S–F with $|f(s) - f(t)| < \epsilon$. First,

$$\sum a_j \geqslant \sum a_j f_j(t) = f(t) \geqslant f(s) - \epsilon.$$

Second, given k, for large m,

$$\epsilon + f(s) \geqslant P_m(s) = \sum a_j^m \geqslant \sum_0^k a_j^m \to \sum_0^k a_j ;$$

as this holds for all k, $f(s) \geqslant \sum a_j$. Hence $f(s) = \sum a_j$. We will now show that the (subsequence) $\{P_n\}$ converges uniformly to f. Given $\epsilon > 0$, choose k with $\sum_0^k a_j \geqslant \sum_0^\infty a_j - \epsilon$. Then choose n so large that $|a_j^n - a_j| < \epsilon/(k+1)$ for $0 \leqslant j \leqslant k$; it follows that $\sum_0^k a_j^n \geqslant \sum_0^\infty a_j - 2\epsilon$. Because $\sum a_j^n \to \sum a_j$, $\sum_{k+1}^\infty a_j^n < 3\epsilon$ for large n. Hence, for any s in S,

$$\begin{aligned} |P_n(s) - f(s)| &\leqslant \sum |a_j^n - a_j| \\ &\leqslant \sum_0^k |a_j^n - a_j| + \sum_{k+1}^\infty a_j^n + \sum_{k+1}^\infty a_j \\ &< 5\epsilon \end{aligned}$$

for n large. Q.E.D.

Note that in the second part of the proof we establish the result that a sequence $\{x_n\}$ in l^1 which converges in the c_0-topology to x with $\{\| x_n \|\}$ converging to $\| x \|$ must converge to x in norm. Also, for S compact, if we take $\{f_n\}$ and f to be in the Banach space $C(S)$, then Theorem 1 shows that when $\{P_n\}$, a sequence from the positive cone spanned by $\{f_n\}$, converges weakly to f, a subsequence must converge in norm.

In passing, note that some extra condition on f is necessary in Part II of Theorem 1. Consider $S = [0, 1]$, $F = \{1\}$, $f_n(x) = x^n$, $n = 0, 1, 2,\ldots$. Let $g(x) = \sum_0^\infty (1/2^n)\, x^n$, $Q_m(x) = \sum_0^m (1/2^n)\, x^n$. Then define $f(x) = g(x)$, $0 \leqslant x < 1$, and $f(1) = g(1) + 1$; also $P_n(x) = Q_n(x) + x^n$. We have $P_n(x) \to f(x)$ on $[0, 1]$, but $f(x) = \sum_0^\infty (1/2^n)\, x^n$ holds only on $[0, 1)$.

THEOREM 2. (Bernstein [1, Sect. IV]; Widder [10, Chap. IV]. *Let f be a function continuous on* $[0, 1]$. *The following are equivalent.*

(i) *The function f is C^∞ with $f^k(x) \geqslant 0$ for $0 < x < 1$, $k = 0, 1, 2,\ldots$.*

(ii) *The* kth *difference $\Delta_h{}^k f(x)$ is nonnegative for $k = 0, 1, 2,\ldots$, and all positive h.*

(iii) *The function f is a pointwise limit of polynomials with nonnegative coefficients.*

(iv) *There is a convergent series, $\sum a_n$ of nonnegative terms with $f(x) = \sum a_n x^n$ holding uniformly on* $[0, 1]$.

Proof. We show that (i) implies (ii). Recall that the differences of f with increment h are defined by $\Delta_h{}^0 f(x) = f(x)$; $\Delta_h{}^1 f(x) = f(x + h) - f(x)$, $0 \leqslant x \leqslant x + h \leqslant 1$; $\Delta_h^{k+1} f(x) = \Delta_h{}^1 \Delta_h{}^k f(x) = f(x + kh) - \binom{k}{1} f(x + (k - 1)\, h) + \cdots + (-1)^k f(x)$, $0 \leqslant x \leqslant x + kh \leqslant 1$. By repeated applications of the mean value theorem we see that $\Delta_h{}^k f(x) = f^k(c)\, h^k$ with $0 < c < 1$.

We show that (ii) implies (iii). For f continuous on $[0, 1]$, the Bernstein polynomial

$$B_n(x) = \sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1 - x)^{n-k}$$

converges uniformly to f on $[0, 1]$ [3, 8, 9]. Following [8, pp. 12–13] or [10, p. 155], it is easy to see that

$$B(x) = \sum_{k=0}^{n} \binom{n}{k} \Delta_{1/n}^{k} f(0)\, x^k .$$

Thus f is the (uniform) limit of polynomials with nonnegative coefficients.

That (iii) implies (iv) is an immediate consequence of Theorem 1; (i) follows from (iv). Q.E.D.

A simple change of variable establishes the result for the interval $[a, b]$, with $f(x) = \sum b_n(x - a)^n$, f being the pointwise limit of polynomials in powers of $x - a$ with nonnegative coefficients. (Note that $S = [a, b]$, $f_n(x) = ((x - a)/(b - a))^n$ satisfies the hypotheses of Theorem 1.) We can conclude that f is the restriction of a function analytic in a circle of radius $b - a$ centered at a. For example if $\Delta_h{}^k f(x) \geqslant 0$ for x in $[0, \infty)$, $k = 0, 1, 2,\ldots, h > 0$, then for $S = [0, m]$, f is the restriction of a function analytic in a circle of radius m centered at 0. Hence f is the restriction to $[0, \infty)$ of an entire function.

Let S_n be a topological space and $\{f_i{}^n\}$ a sequence of functions on S_n satisfying the conditions of the definition, and in addition suppose that the functions $\{f_j{}^n\}, j, n = 0, 1, 2,\ldots,$ are uniformly bounded below. Define g on ΠS_n by $g_j{}^n(s_0, s_1, \ldots) = f_j{}^n(s_n)$, and let $\{g_n\}$ be some indexing of the collection of all finite products of the functions $g_j{}^n$. Then ΠS_n, $\{g_n\}$ also satisfy the conditions of the definition. A special case of this, on $S = [0, 1] \times [0, 1]$, is to let $g_n(x, y)$ be some indexing of the functions $x^n y^m$, $n, m = 0, 1, 2,\ldots.$ With the aid of Theorem 1 we can easily generalize Theorem 2 to two (or several) variables.

THEOREM 3. *Let f be continuous on* $[0, 1] \times [0, 1]$. *The following are equivalent.*

(i) *The function f has continuous, mixed partial derivative of all orders, with* $(\partial^{n+m}/\partial x^n \partial y^m) f(x, y) \geqslant 0$ *on* $(0, 1) \times (0, 1)$.

(ii) *The differences* $(\Delta_x)_{h_1}^n (\Delta_y)_{h_2}^m f(x, y)$ *are nonnegative for* $n, m = 0, 1, 2,\ldots, h_1 > 0, h_2 > 0$.

(iii) *The function f is a pointwise limit of polynomials in the two variables x and y with nonnegative coefficients.*

(iv) *There is a convergent series* $\sum\sum a_{nm}$ *of nonnegative terms with* $f(x, y) = \sum\sum a_{nm} x^n y^m$ *holding uniformly on* $[0, 1] \times [0, 1]$.

Proof. The proof parallels the proof of Theorem 1. We need the following information.

The partial differences of f are determined as for one variable:

$$\underset{x}{\Delta}_h{}^0 f(x, y) = f(x, y),$$

$$\underset{x}{\Delta}_h{}^1 f(x, y) = f(x + h, y) - f(x, y),$$

$$\underset{x}{\Delta}_h^{k+1} f(x, y) = \underset{x}{\Delta}_h{}^1 \underset{x}{\Delta}_h{}^k f(x, y),$$

and similarly for the y differences. Also, if f has continuous partials of orders less than or equal to $n + m$, then it can be shown inductively that

$$\underset{x}{\Delta}{}^{n}_{h_1}\underset{y}{\Delta}{}^{m}_{h_2} f(0, 0) = h_1{}^n h_2{}^m \frac{\partial^{n+m}}{\partial x^n \, \partial y^m} f(c_1, c_2),$$

where $0 < c_1 < 1$ and $0 < c_2 < 1$.

The Bernstein polynomials

$$B_{n,m}(x, y) = \sum_{k=0}^{n} \sum_{l=0}^{m} \binom{n}{k}\binom{m}{l} f\left(\frac{k}{n}, \frac{l}{m}\right) x^k(1 - x)^{n-k}(y^l)(1 - y)^{m-l}$$

converge uniformly to f on $[0, 1] \times [0, 1]$ [9, p. 10]. Exactly as for one variable, in fact using those results, we obtain

$$B_{n,m}(x, y) = \sum_{k=0}^{n} \sum_{l=0}^{m} \binom{n}{k}\binom{m}{l} \underset{x}{\Delta}{}^{k}_{1/n}\underset{y}{\Delta}{}^{l}_{1/m} f(0, 0)\, x^k y^l.$$

For more about these matters see [2,7].

This is all we need to do the proof. Q.E.D.

A couple of examples of the failure of the conclusions of Theorem 1 to hold will be instructive.

Example 1. On $[0, 1)$, define $f_n(x) = (1 - x)\, x^n$, $n = 0, 1, 2,\ldots$. The function $f(x) = 1$ is the pointwise limit of $\sum_0^m f_n(x)$. We cannot have $f(x) = \sum a_n f_n(x)$, with $\sum a_n$ a convergent series of nonnegative terms, for

$$1 = a_0 + (a_1 - a_0)\, x + (a_2 - a_1)\, x^2 + \cdots$$

shows that $a_j = 1, j = 0, 1, 2,\ldots$.

Example 2. On $(0, 2\pi)$, $f_n(x) = \sin nx$, $n = 1, 2,\ldots$. The function $g(x) = (\pi - x)/2$ is the pointwise limit of the partial sums $\sum_0^m (1/n) \sin(nx)$ of its Fourier series. We cannot have $g(x) = \sum a_n \sin(nx)$ with $\sum a_n$ a convergent series of nonnegative terms, for then $1/m = (1/\pi) \int_0^{2\pi} g(x) \sin(mx)\, dx = a_m$ and $\sum a_m$ diverges.

It is easy to construct examples of functions satisfying the conditions of the definition and for which the conclusions of Theorem 1 are interesting; for example, $f_n(x) = \operatorname{sech}(nx)$, $S = (-\infty, \infty)$, $F = \{0\}$; $f_n(x, y) = (x + y)^n$, $S = \{(x, y): x \geqslant 0, y \geqslant 0, x + y \leqslant 1\}$, F the hypotenuse of the boundary of S. However, stronger examples can be obtained from eigenvalue problems for ordinary and partial differential equations, where expanding a function in a series of eigenfunctions is a classical problem of great interest.

EXAMPLE 3. On $S' = (-1, 1]$, let $P_n(x)$ be the nth Legendre polynomial. Recall that $-1 \leqslant P_n(x) \leqslant 1$, and $P_n(1) = 1$. Using the Laplace integral representation for $P_n(x)$ [5, p. 58],

$$P_n(x) = (1/\pi) \int_{-\pi}^{\pi} [x + i(1 - x^2)^{1/2}(\cos t)]^n \, dt$$

and the inequality $| x + i(1 - x^2)^{1/2} \cos t | < 1$ for $-1 < x < 1$ and $0 < t < \pi$, we see that $P_n(x) \to 0$ for $-1 < x < 1$. (Note that $P_n(-1) = (-1)^n$, so for the sequence $\{P_{2n}\}$ we could take $S = [-1, 1]$ and $F = \{-1, 1\}$.)

EXAMPLE 4. In solving a problem of the diffusion of heat in an infinite cylinder we find that we must expand the solution $u(r, t)$ in a series of the functions $f_n(r, t) = J_0(\lambda_n r) \exp(-a^2 \lambda nt)$, as well as expanding a given function, representing an initial heat distribution, in a series of the functions $J_0(\lambda_n r)$. Here J_0 is the Bessel function of order 0, r is the distance from the axis of the cylinder, t is the time, λ_n are the positive roots of $J_0(\lambda c) = 0$, c is the radius of the cylinder, and a^2 is the thermal conductivity of the cylinder. We take $S = [0, c] \times [0, \infty)$. To see that the conditions of the definition are satisfied for the functions $f_n(r, t)$ we need to know that $\lambda_n \to \infty$, $-1 \leqslant J_0(x) \leqslant 1$, $J_0(0) = 1$, and that $J_0(x) \to 0$ as $x \to \infty$; see [4, Chap. VIII]. We remark that the functions $J_0(\lambda_n r)$ on $[0, c]$ and $\exp(-a^2 \lambda nt)$ on $[0, \infty)$ satisfy the conditions of the definition, so the product, as we remarked following theorem 1, satisfies the conditions on the product space.

There is an interesting point to be brought out here. For certain eigenvalue problems we have the strong results of Theorem 1 holding, results which for analytic $\{f_n\}$ yield (or require) analyticity of the expanded function, while this is not the case for other problems. It would be enlightening if this could be explained in terms of the physics of the underlying physical problem and, conversely, if the conclusions of Theorem 1 could be interpreted in a physical manner.

REFERENCES

1. S. BERNSTEIN, Leçons sur les propriétés extrémales et la meilleure approximation des fonctions analytiques d'une variable réelle; reprinted in "L'Approximation," Chelsea, New York, 1970.
2. P. BUTZER, On two-dimensional Bernstein polynomials, *Canad. J. Math.* **5** (1953), 107–113.
3. E. CHENEY, "Introduction to Approximation Theory," McGraw–Hill, New York, 1966.
4. R. CHURCHILL, "Fourier Series and Boundary Value Problems," McGraw–Hill, New York, 1941.

5. D. Jackson, "Fourier Series and Orthogonal Polynomials," Carus Monograph No. 6, Mathematical Association of America, Washington, D.C., 1941.
6. W. Jurkat and G. Lorentz, Uniform approximation by polynomials with positive coefficients, *Duke Math. J.* **28** (1961), 463–474.
7. E. Kingsleys, Bernstein polynomials for functions of two variables of class $C^{(k)}$, *Proc. Amer. Math. Soc.* **2** (1951), 64–71.
8. G. Lorentz, "Bernstein Polynomials," Univ. of Toronto Press, Toronto, 1953.
9. G. Lorentz, "Approximation of Functions," Holt, Rinehart, and Winston, New York, 1966.
10. D. Widder, "The Laplace Transform," Princeton Univ. Press, Princeton, N.J., 1946.