Chebyshev Approximation by Exponentials on Finite Subsets

By Dietrich Braess

Abstract. This paper is concerned with Chebyshev approximation by exponentials on finite subsets. We take into account that varisolvency does not hold for exponentials in general. A bound for the derivatives of exponentials is established and convergence of the solutions for the discrete problems is proved in the topology of compact convergence on the open interval.

1. Introduction. In a recent note, Rosman [9] studied the convergence of best exponential Chebyshev approximation on finite subsets. Unfortunately, his investigations heavily depend on results of Rice from 1962 [6], [8], and he assumed that the family of exponentials

$$ V_n = \left\{ E(x) = \sum_{i=1}^{n} \sum_{j=0}^{m_i} p_{ij} x^{i+j}; p_{ij}, t_i \in \mathbb{R}, \sum_{i=1}^{n} (1 + m_i) = n \right\} $$

has the varisolvency property. But, as was shown by the author in 1967 [1], [3], varisolvency holds only for the special exponentials of the form

$$ \sum_{i=1}^{n} \alpha_i e^{i t}. $$

Moreover, there are two different definitions of varisolvency in the literature. The exponentials of the form (2) are varisolvent in the sense of Rice's papers [6], [7], [8], but not in the sense of Hobby and Rice [5]. For the study of Rosman's proof, this difference cannot be neglected.

In this note, we will present a different proof, using ideas in Werner's [11] and Schmidt's [10] proof for an existence theorem. At first, we establish an estimation of the derivatives of exponentials similar to Bernstein's inequality for polynomials. Computational methods are not considered here; for this, we refer to [2], [8], [12].

2. Estimation of Derivatives. The main result of this section is an estimation of the derivatives of exponentials mentioned in [4]. But the major part will be concerned with the lemmas preparing the convergence theorem in the next section.

**Lemma 1.** Let $x_0 < x_1 < \cdots < x_n$. If $f \in C^r[x_0, x_n]$, and if

$$ |f(x_i)| \leq M, \quad i = 0, 1, \ldots, n, $$

Received October 1, 1971.


Key words and phrases. Chebyshev approximation, exponentials, violation of varisolvency, estimation of derivatives, convergence of discrete approximation.

Copyright © 1973, American Mathematical Society
holds, then there exists a point \( z \in [x_0, x_n] \), such that

\[
|f^{(n)}(z)| \leq M \cdot n! \cdot \prod_{i=0}^{n} \frac{1}{|x_i - x|}.
\]

**Proof.** Consider the polynomial \( p(x) \) of degree \( n \) which interpolates \( f(x) \) at \( x_0, x_1, \ldots, x_n \). Since \( f - p \) has \( n + 1 \) zeros, there is at least one zero of \( (f - p)^{(n)} \). Observe that the right-hand side of (4) represents an upper bound of the \( n \)th derivative of the Lagrangian interpolating polynomial. \( \square \)

By a special choice of points, we get an estimation for functions defined on an interval. For each compact set \( Y \), we define the Chebyshev norm on \( C(Y) \) as

\[
||f||_Y = \sup_{x \in Y} |f(x)|.
\]

**Corollary 2.** Let \( X = [\alpha, \beta] \) and \( d = \beta - \alpha \). For each \( f \in C^n(X) \), there exists a point \( z \in X \), satisfying

\[
|f^{(n)}(z)| \leq 2^{2n-1} n! \cdot d^n \cdot ||f||_X.
\]

**Proof.** Let \( x_i = \frac{1}{2}(\alpha + \beta) - (d/2) \cos(i\pi/n) \) for \( i = 0, 1, \ldots, n \). By applying Lemma 1 to the transformed Chebyshev polynomial \( f(\frac{1}{2}(\alpha + \beta) + dx/2) = T_n(x) \), we obtain the equal sign in (4). From \( T_n^{(n)}(x) = 2^{n-1} n! \), we determine the factor of \( M \) in (4). This yields the theorem. \( \square \)

Now, we have established a priori estimates which are necessary for the application of the main lemma that generalizes a theorem of Schmidt [10]. Notice that derivatives of exponentials are exponentials, too. Thus, they have at most \( n - 1 \) zeros or vanish identically.

To each (finite) sequence of distances \( d_1, d_2, \ldots, d_n \) and to the corresponding a priori constants \( M_1, M_2, \ldots, M_n \), there are associated \( n + 1 \) numbers by a recursive process

\[
K_{n+1} = 0,
\]

\[
K_v = M_v + d_v K_{v+1}, \quad v = n, n - 1, \ldots, 1.
\]

**Lemma 3.** Let \( d_1, d_2, \ldots, d_n \) be positive numbers, satisfying

\[
d_1 + d_2 + \cdots + d_n < d.
\]

Let \( X \supset [x_0 - d, x_0 + d] \) and \( f \in C^{n+1}(X) \). Suppose that \( f^{(n+1)} \) has at most \( n - 1 \) zeros or vanishes identically in \( X \), and, moreover, that in each subinterval of length \( d_v \) (\( v = 1, 2, \ldots, n \)), there is a point \( z \) such that

\[
|f^{(v)}(z)| \leq M_v.
\]

Then

\[
|f'(x_0)| \leq K_1
\]

holds with \( K_1 \) defined by the recursion relation (6).

**Proof.** Suppose to the contrary that (9) is violated and, say, \( f'(x_0) > K_1 \) holds. By an inductive proof we will show that, for \( v = 1, 2, \ldots, n \), there are points \( \xi_v, \eta_v \) such that

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
(10) \[ x_0 - \sum_{\nu=1}^r d_\nu \leq \xi_0 \leq \eta_0 \leq x_0 + \sum_{\nu=1}^r d_\nu, \]

(11) \[ f^{(r+1)}(\xi_0) > K_{r+1}, \quad (-1)^r f^{(r+1)}(\eta_0) > K_{r+1}, \]
and \( f^{(r+1)} \) has \( \nu \) distinct zeros in \([\xi_0, \eta_0] \).

Let \( \nu = 1 \). By assumption, \( |f'(z)| \leq M_1 \) holds for a point \( z_1 \in [x_0 - d_1, x_0] \).

From Rolle's theorem, we obtain a point \( \xi_1 \in [z_1, x_0] \), satisfying

\[ \frac{f''(\xi_1)}{f'(\xi_1)} = \frac{f'(z_1) - f'(z_1)}{z_1 - z_1} = \frac{K_1 - M_1}{d_1} = K_2. \]

A corresponding construction yields \( \eta_1 \in [x_0, x_0 + d_1] \) with the postulated properties.

By virtue of (11), there is a zero of \( f''(x) \) in \((\xi_1, \eta_1)\).

Assume that the statement holds for \( \nu - 1 \leq n - 1 \). Denote by \( x_1, x_2, \ldots, x_{\nu-1} \), the zeros of \( f^{(\nu)} \). By assumption, we have \( |f^{(\nu)}(\xi)| \leq M_\nu \) for a point \( \xi \in [\xi_{\nu-1} - d_\nu, \xi_{\nu-1}] \). Let \( f^{(\nu)} \) attain its maximum in \([\xi, x_1]\) at \( x_1 \). From \( f^{(\nu)}(\xi_{\nu-1}) \geq f^{(\nu)}(\xi_{\nu-1}) > K_\nu \geq M_\nu \geq f^{(\nu)}(\xi_{\nu-1}) \) and \( f^{(\nu)}(x_1) = 0 \), we conclude that \( z_1 \in (z_1, x_1) \) and \( f^{(\nu+1)}(z_1) = 0 \). Set \( z_2 = \min(z_1, \xi_{\nu-1}) \). By virtue of Rolle's theorem, there exists \( \xi_2 \), satisfying

\[ f^{(\nu+1)}(\xi_2) = \frac{f^{(\nu)}(z_2) - f^{(\nu)}(z_1)}{z_2 - z_1} \geq \frac{K_\nu - M_\nu}{d_\nu} = K_{\nu+1}. \]

Construct \( \eta_2 \in [\eta_{\nu-1}, \eta_{\nu-1} + d_\nu] \) by an analogous procedure. Hence, \( f^{(\nu+1)} \) has at least \( \nu - 2 \) zeros between \( x_1 \) and \( x_{\nu-1} \). Moreover, two zeros are determined in \((\xi_1, x_1)\) and \((x_{\nu-1}, \eta_1)\), respectively, and the induction is complete. As a consequence, for \( \nu = n \), there is a contradiction to the assumption on the zeros of \( f^{(n+1)} \). \( \square \)

Now we are ready to prove the desired estimation.

**Theorem 4.** Let \( X = [\alpha, \beta] \) and \( 2d \leq \beta - \alpha \). There exists a constant \( c = c_\nu \), such that, for each exponential \( E \) of degree \( \leq n \),

\[ |E'(x)| \leq (c_\nu/d) \cdot ||E||_{\nu+1} \quad \text{for } x \in [\alpha + d, \beta - d]. \]

**Proof.** It is sufficient to prove the theorem for \( E(x) \neq 0 \). Given \( x_0 \in [\alpha + d, \beta - d] \), set

\[ f(x) = (1/||E||_{\nu+1}) \cdot E(x_0 + dx), \quad -1 \leq x \leq +1. \]

Obviously, \( f(x) \) is an exponential and \( ||f||_{\nu+1} \leq 1 \) holds. Let \( d_1, d_2, \ldots, d_n \) be positive numbers, the sum of which is 1. Set

\[ K_{n+1} = 0, \]

\[ K_\nu = 2^{\nu-1} \cdot d_{\nu-1} + d_\nu K_{\nu+1}, \quad \nu = n, n-1, \ldots, 1, \]

\[ c_\nu = K_1. \]

By virtue of Corollary 2 and Lemma 3 we obtain \( |f'(0)| \leq c_\nu \). From this, the inequality (12) is evident. \( \square \)

3. **Approximation on Finite Subsets.** Let \( X \) be a compact interval on the real line and let \( X_r \) be a set of \( r \) distinct points in \( X \). Then the density of \( X_r \), in \( X \) is measured by
\[ \Delta_r = \max_{x \in X} \min_{v \in X_r} |x - v|. \]

We consider a sequence of subsets \( \{X_r\} \), satisfying \( \Delta_r \to 0 \) as \( r \) tends to infinity. Since \( X \) is compact, this is equivalent to the assumption in [9] that, given \( x \in X \), there is an \( x_r \in X_r \) such that \( x_r \to x \).

As usual, \( E_r \) is called a best approximation to \( f \) on \( X_r \), if the functional \( ||f - E||_{X_r} \) attains its minimum on \( V_n \) at \( E_r \). It is known that best approximations need not exist on finite point sets [8] and unicity of the solution cannot be ensured [1]. From the computational point of view, it is reasonable to assume existence anyway. However, for a rigorous proof of the convergence theorem, we avoid this difficulty by the definition of nearly best approximations. Let \( E^* \) be a best approximation to \( f(x) \) on \( X \). \( E_r \) is called a nearly best approximation to \( f(x) \) on \( X_r \), if

\[ ||f - E_r||_{X_r} \leq ||f - E^*||_{X_r}. \]

Obviously, each best approximation on \( X_\ast \) is a nearly best approximation.

**Theorem 5.** Let \( X = [\alpha, \beta] \), and let \( X_r \) be a sequence of finite subsets, such that \( \Delta_r \to 0 \). Then, each sequence of nearly best approximations \( \{E_r\} \) contains a subsequence that converges to a best approximation \( E^* \) on \( X \) uniformly on each compact subinterval of \( (\alpha, \beta) \). If \( E^* \) has the maximal degree \( n \), then convergence is uniform on the total interval \( X \).

**Proof.** Let \( Y = [\alpha_1, \beta_1] \) be a compact subinterval of \( (\alpha, \beta) \). Set \( Y_1 = [\frac{1}{2}(\alpha + \alpha_1), \frac{1}{2}(\beta + \beta_1)] \). From Corollary 2, we know that, in any interval \( I \) of length \( d \), we can find \( n + 1 \) points \( x_0, x_1, \ldots, x_n \) such that the sum in inequality (4) has the value \( 2^{2n-1}d^n \). Since \( \Delta_r \) tends to zero, for sufficiently large \( r \), we may choose \( n + 1 \) points in \( I \subset X \), such that the sum can be bounded by \( 2^{2n_1}d^n \). By virtue of Lemma 3 and by

\[ ||E_r||_{X_r} \leq ||f - E_r||_{X_r} + ||f||_{X_r} \leq 2 ||f||_{X_r}, \]

there is a constant \( c \) such that, for sufficiently large \( r \),

\[ |E_r(x)| \leq c \cdot ||E_r||_{X_r} \leq 2c \cdot ||f||_{X_r} \quad \text{for} \ x \in Y_1. \]

Since for each \( x \in X \) there is a point in \( X \), with a distance not greater than \( \Delta_r \), we obtain

\[ ||E_r||_{X_\ast} \leq (2 + 2c \cdot \Delta_r) \cdot ||f||_{X}. \]

Hence, \( \{E_r\} \) is bounded on \( Y_1 \). By Corollary 1 in [10], there exists a subsequence converging uniformly on \( Y \subset Y_1 \) to an exponential \( E^* \). By standard arguments, we conclude that this subsequence converges uniformly to \( E^* \) on each compact subset of \( (\alpha, \beta) \). Obviously, \( E^* \) is a best approximation to \( f \) on \( X \). Moreover, from Theorem 4 in [10], we obtain uniform convergence on \( X \), if \( E^* \) has maximal degree. \( \square \)

Institut für Numerische und Instrumentelle Mathematik
Westfälische Wilhelms-Universität
44 Münster, Germany


