JOURNAL OF APPROXIMATION THEORY 28, 273–292 (1980)

On the Foundations of Best Approximation Theory

DOUGLAS S. BRIDGES*

Department of Mathematical Sciences, New Mexico State University, La Cruces, New Mexico 88003

Communicated by Oved Shisha

Received October 4, 1977

In this paper we excavate the foundations of best-approximation theory with the tools of Bishop's constructive analysis. We prove a general theorem on existence (computability) of best approximations from a given finite-dimensional linear subspace of a normed space E, and illustrate this with the case where E is uniformly convex. The second part of the paper deals with the characterisation and existence of minimax polynomial approximations to elements of $C[0, 1]$, and with the pointwise continuity of the minimax approximation mapping on this space. In particular, the main application of our general existence theorem answers affirmatively the long-open question: Is there a constructive proof of the existence of minimax polynomial approximations?

1. INTRODUCTION

As should be familiar to every advanced undergraduate in mathematics, the theoretical foundation of the numerical analyst's interest in approximation theory is provided by the theorem

a finite-dimensional linear subspace X of a normed space E over R is proximinal in E—that is, (*)

$$\forall a \in E\, \exists \xi \in X \quad \text{dist}(a, X) = \| a - \xi \|.$$

What is remarkable about this theorem is that, although it supports a vital branch of computational mathematics, it admits of no known constructive proof: to be exact, not only do the classical proofs of (*) beg the question of the computability of dist(a, X), but also they deduce the "existence" of the best approximation ξ to a in X from the essentially nonconstructive proposition that a continuous, real-valued function on a compact space attains its infimum [8, 8.3.2].

In this paper, we investigate the problem of existence of best approximations with the techniques of Bishop's constructive analysis. (For general background to constructive mathematics, we refer the reader to [2]: a wider,

* Current address: University College, Buckingham MK18 1EG, England.

0021-9045/80/030273–20$02.00/0

but less up-to-date, coverage of the subject is found in [1].) It is our belief that constructive mathematics, with its insistence on numerical content and computational method, may have considerable importance in the development of numerical analysis, at least in theory. We certainly hope that the questions raised in this work will lead to further investigations in the subject of computability of best approximations, and the constructive approach to numerical analysis in general (cf. [5, 6]).

To return to (*), it is fortunate that the computability of dist(a, X) can be demonstrated as a simple consequence of a result of Bishop [1, Chap. 4, Proposition 13]. Moreover, as we shall show below (2.1), this can be derived also in an elementary manner by an adaptation of a well-known classical proof of (*) [7, pp. 78–80]. It follows that the constructive content of this classical proof is precisely the existence of dist(a, X), and not its attainment at some point ξ of X; indeed, we are tempted to believe that the existence of such ξ is an essentially nonconstructive proposition.

To reassure any numerical analyst who may be distressed by this last possibility, we point out that there are commonly occurring situations in which the existence of best approximations can be established by constructive means. In particular, one corollary of our main general result (2.2) is that, if each finite-dimensional linear subspace of E contains at most one approximation to a given element of E, then all finite-dimensional subspaces of E are proximinal. Moreover, in the general case, the computability of dist(a, X) means that we can compute an approximation to a in X which is as close to a best approximation as we require (for example, to the highest degree of accuracy of any available computer). At the same time, we have no guarantee as yet that this approximation does not jump discontinuously as we try to improve its accuracy.

2. Existence of Best Approximations

Throughout this paper, all normed linear spaces are over the real-number field $\mathbb{R}$. We shall say that a normed space X is *finite dimensional* if there exist finitely many elements $e_1, ..., e_\nu$ of X and linear functionals $\phi_1, ..., \phi_\nu$ on X, such that

$$x = \sum_{k=1}^{\nu} \phi_k(x)\, e_k \qquad (x \in E),$$

$$\phi_j(e_k) = 0 \qquad (1 \leqslant j, k \leqslant \nu, j \neq k),$$

and each ϕ_k is *bounded*—that is, we can compute $c > 0$ such that $|\phi_k(x)| \leqslant c\|x\|$ for each x in X. The number ν of elements of the *basis* $\{e_1, ..., e_\nu\}$ of X is then called the *dimension* of X, and is independent of the basis in question.

We write $X = \operatorname{span}\{e_1, \ldots, e_\nu\}$ when there is no likelihood of confusion as to the linear functionals ϕ_k.

By a *compact space* we mean a metric space that is totally bounded and complete. The closed unit ball of a finite-dimensional normed space is compact, as is its boundary. If f is a uniformly continuous mapping of a compact space K into $\mathbb{R}$, then $\sup f$ and $\inf f$ are computable, although not necessarily attained; for all but countably many real numbers $\alpha > \inf f$, the set $\{x \in K : f(x) \leqslant \alpha\}$ is then compact.

A subset S of a metric space E is *located* in E if $\operatorname{dist}(x, S)$ is computable for each x in E. Every compact subset of a metric space is located.

2.1. Proposition. *A finite-dimensional linear subspace X of a normed linear space E is located.*

Proof. Let $\{e_1, \ldots, e_\nu\}$ be a basis of unit vectors of X, $a \in E$ and define

$$\left\| \sum_{k=1}^{\nu} \lambda_k e_k \right\|_0 \equiv \|\underline{\lambda}\| = \left(\sum_{k=1}^{\nu} |\lambda_k|^2 \right)^{1/2}$$

for each $\underline{\lambda} = (\lambda_1, \ldots, \lambda_\nu)$ in $\mathbb{R}^\nu$. Then $\|\cdot\|_0$ is a norm on X; so that, by the equivalence of all norms on a finite-dimensional linear space, there exists $\mu' > 0$ such that $\mu' \|x\|_0 \leqslant \|x\|$ for each x in X. Thus

$$0 < \mu' \leqslant \mu \equiv \inf \left\{ \left\| \sum_{k=1}^{\nu} \lambda_k e_k \right\| : \underline{\lambda} \in \mathbb{R}^\nu, \|\underline{\lambda}\| = 1 \right\}.$$

For each $r > 0$, define

$$d_r \equiv \inf \left\{ \left\| a - \sum_{k=1}^{\nu} \lambda_k e_k \right\| : \underline{\lambda} \in \mathbb{R}^\nu, \|\underline{\lambda}\| \leqslant r \right\}.$$

Then, choosing in turn $r > 1$ so that $d_1 < r\mu - \|a\|$ and $\underline{\lambda}$ in $\mathbb{R}^\nu$ with $\|\underline{\lambda}\| > r$, we have

$$\begin{aligned} \left\| a - \sum_{k=1}^{\nu} \lambda_k e_k \right\| &\geqslant \|\underline{\lambda}\| \left\| \sum_{k=1}^{\nu} \|\underline{\lambda}\|^{-1} \lambda_k e_k \right\| - \|a\| \\ &\geqslant r\mu - \|a\| \\ &> d_1 \\ &\geqslant d_r . \end{aligned}$$

It is now clear that $\operatorname{dist}(a, X)$ is computable, and equals d_r. ∎

In this last proof, a more natural classical approach to the positivity of μ uses the proposition that a uniformly continuous mapping of a compact space into the positive reals has positive infimum. The constructive validity of this proposition remains an open problem (cf. [3, Sect. 4; 4]).

Another point of divergence between classical and constructive mathematics arises in connection with the comparison of real numbers: as the propositions

$$\forall x \in \mathbb{R}\, (x < 0 \vee x > 0 \vee x = 0)$$

and

$$\forall x \in \mathbb{R}\, \forall y \in \mathbb{R}\, (x < y \vee x = y \vee x > y)$$

are both essentially nonconstructive, constructive analysis must be done using such acceptable substitutes as

$$\forall x \in \mathbb{R}\, ((x > 0 \Rightarrow 0 = 1) \Rightarrow x \leqslant 0)$$

and

$$\forall x \in \mathbb{R}\, \forall y \in \mathbb{R}\, \forall z \in \mathbb{R}\, (x < y \Rightarrow x < z \vee z < y)$$

(see Chapter 1 of [2] for details). In particular, we cannot assert that

$$\forall a \in E\, \forall x \in X\, (\| a - x \| > \operatorname{dist}(a, X) \vee \| a - x \| = \operatorname{dist}(a, X)).$$

To get round this obstacle, we say that $a \in E$ has *at most one best approximation* in the finite-dimensional subspace X of E if

$$\max\,(\| a - x \|, \| a - x' \|) > \operatorname{dist}(a, X)$$

whenever $x \in X$, $x' \in X$ and $\| x - x' \| > 0$.

With this definition, we come to our main general result

2.2. THEOREM. *Let* $\{e_1, ..., e_\nu\}$ *be a basis of the finite-dimensional linear subspace* X *of the normed space* E *over* $\mathbb{R}$. *Suppose that, for each* $k \in \{1, ..., \nu\}$, *each* x *in* E *has at most one best approximation in* span $\{e_1, ..., e_k\}$. *Then* X *is proximinal in* E.

Proof. We proceed by induction on k. Let $a \in E$, $d = \operatorname{dist}(a, \operatorname{span}\{e_1\})$ and $\phi(\lambda) \equiv \| a - \lambda e_1 \|$ $(\lambda \in \mathbb{R})$. We first observe that, if $t_1 > t_2 > 0$ and

$$S_k \equiv \{\lambda \in \mathbb{R} : \phi(\lambda) \leqslant d + t_k\}$$

is compact for $k = 1, 2$, then (as ϕ is uniformly continuous)

$$\phi(\inf S_k) = d + t_k = \phi(\sup S_k) \qquad (k = 1, 2).$$

Hence

$$\inf S_1 < \inf S_2 \leqslant \sup S_2 < \sup S_1\,.$$

Also, $S_2 \subset S_1$; and, as S_k is convex, $S_k = [\inf S_k, \sup S_k]$ $(k = 1, 2)$.

We now construct a sequence $(\alpha_n)_{n\geqslant 1}$ of positive numbers converging to 0, such that, for each n,

$$\begin{aligned} A(n) &\equiv \{\lambda \in \mathbb{R} : \phi(\lambda) \leqslant d + \alpha_n\} \\ &= \{\lambda \in \mathbb{R} : |\lambda| \leqslant \| e_1 \|^{-1}(\| a \| + d + \alpha_n), \phi(\lambda) \leqslant d + \alpha_n\} \end{aligned}$$

is compact. Then

$$A(n + 1) \subset A(n) = [\inf A(n), \sup A(n)]$$

and

$$\inf A(n) < \inf A(n + 1) < \sup A(n + 1) < \sup A(n).$$

Classically, we could now argue that the decreasing, minorized sequence $(\sup A(n))_{n\geqslant 1}$ converges to its infimum M, and hence that $\| a - Me_1 \| = \phi(M) = d$. Constructively, we cannot use the Least Upper-Bound Principle [1, pp. 4–5], and so we adopt the following argument.

We construct a strictly increasing sequence $(\nu_k)_{k\geqslant 1}$ of positive integers such that

$$\sup A(\nu_{k+1}) - \inf A(\nu_{k+1}) \leqslant (\tfrac{2}{3})^k (\sup A(1) - \inf A(1)) \qquad (k \geqslant 1).$$

Having found $\nu_1 \equiv 1,\ldots, \nu_k$, we set

$$m_k \equiv \inf A(\nu_k), \qquad M_k \equiv \sup A(\nu_k),$$

and compute in turn ξ, ν_{k+1} so that

$$\tfrac{1}{2}(m_k + M_k) < \xi < m_k + \tfrac{2}{3}(M_k - m_k),$$

$\nu_{k+1} > \nu_k$ and $\phi(\xi) > d + \alpha_{\nu_{k+1}}$. (These computations are possible as a has at most one best approximation in span$\{e_1\}$.) We then have either

$$\xi < \inf A(\nu_{k+1}) < \sup A(\nu_{k+1}) < M_k$$

or

$$m_k < \inf A(\nu_{k+1}) < \sup A(\nu_{k+1}) < \xi;$$

whence, in either case,

$$\begin{aligned} \sup A(\nu_{k+1}) - \inf A(\nu_{k+1}) &< \tfrac{2}{3}(M_k - m_k) \\ &\leqslant (\tfrac{2}{3})^k (\sup A(1) - \inf A(1)). \end{aligned}$$

This completes our inductive construction.

It now follows that there exists ζ with $\inf A(\nu_k) \leqslant \zeta \leqslant \sup A(\nu_k)$, and therefore $\phi(\zeta) \leqslant d + \alpha_{\nu_k}$, for each $k \geqslant 1$. Hence $\phi(\zeta) \leqslant d$, and so

$$\| a - \zeta e_1 \| = \phi(\zeta) = d = \operatorname{dist}(a, \operatorname{span}\{e_1\}).$$

Now let $1 \leqslant k < \nu - 1$, and suppose that we have proved $Y = \text{span}\{e_1, ..., e_k\}$ proximinal. Defining a new norm and equality on E by

$$\|x\|_1 = \text{dist}(x, Y)$$

and

$$x =_1 x' \Leftrightarrow \|x - x'\|_1 = 0,$$

we note that

$$\begin{aligned} \inf_{\lambda\in\mathbb{R}} \|a - \lambda e_{k+1}\|_1 &= \inf_{\lambda\in\mathbb{R}} \inf_{y\in Y} \|a - \lambda e_{k+1} - y\| \\ &= \text{dist}(a, \text{span}\{e_1, ..., e_{k+1}\}). \end{aligned}$$

Let λ_1, λ_2 belong to $\mathbb{R}$ with $|\lambda_1 - \lambda_2| > 0$, and choose y_1, y_2 in Y so that

$$\|a - \lambda_j e_{k+1} - y_j\| = \text{dist}(a - \lambda_j e_{k+1}, Y) \qquad (j = 1, 2).$$

Then

$$\begin{aligned} \|(\lambda_1 e_{k+1} + y_1) - (\lambda_2 e_{k+1} + y_2)\| &\geqslant \text{dist}((\lambda_1 - \lambda_2)\, e_{k+1}, Y) \\ &= |\lambda_1 - \lambda_2| \, \text{dist}(e_{k+1}, Y) \\ &> 0, \end{aligned}$$

whence

$$\begin{aligned} \max_{j=1,2} \|a - \lambda_j e_{k+1}\|_1 &= \max_{j=1,2} \|a - \lambda_j e_{k+1} - y_j\| \\ &> \text{dist}(a, \text{span}\{e_1, ..., e_{k+1}\}) \\ &= \inf_{\lambda\in\mathbb{R}} \|a - \lambda e_{k+1}\|_1 . \end{aligned}$$

Thus a has at most one best approximation in the one-dimensional subspace $\text{span}\{e_{k+1}\}$ of $(E, \|\cdot\|_1)$. As $a \in E$ is arbitrary, it follows from the first part of the proof that there exists ζ in $\mathbb{R}$ with

$$\begin{aligned} \text{dist}(a - \zeta e_{k+1}, Y) &= \|a - \zeta e_{k+1}\|_1 \\ &= \text{dist}(a, \text{span}\{e_1, ..., e_{k+1}\}). \end{aligned}$$

By our inductive hypothesis, there exists ℓ in Y such that

$$\|a - \zeta e_{k+1} - \ell\| = \text{dist}(a - \zeta e_{k+1}, Y).$$

Clearly, $\ell + \zeta e_{k+1}$ is a best approximation to a in $\text{span}\{e_1, ..., e_{k+1}\}$, and our induction is complete. Taking $k = \nu - 1$, we immediately obtain the proximinality of X. ∎

Note that the best approximation to a in X in 2.2 is unique.

3. First Applications of the Existence Theorem

In order to apply 2.2, we define a normed space E to be *uniformly convex* if, for each $\epsilon > 0$, there exists $\delta > 0$ such that $\|\frac{1}{2}(x + y)\| \leqslant \delta$ whenever $\|x\| = \|y\| = 1$ and $\|x - y\| \geqslant \epsilon$.

3.1. Theorem. *A finite-dimensional subspace of a uniformly convex normed space E is proximinal.*

Proof. In view of 2.2, it will suffice to prove that each element a of E has at most one best approximation in a given finite-dimensional subspace X of E. Let x, x' belong to X, with $0 < \alpha \equiv \|x - x'\|$. With $d \equiv \operatorname{dist}(a, X)$, choose r in $]0, 1[$ so that $\|\frac{1}{2}(s - t)\| \leqslant r$ whenever $\|s\| = \|t\| = 1$ and $\|s - t\| \geqslant \alpha/3(1 + d)$. Suppose that

$$\max(\|a - x\|, \|a - x'\|) < \beta \equiv \min(d + \alpha/6, \alpha/6r).$$

Were $d < \alpha/6$, we would have

$$\begin{aligned} \|x - x'\| &\leqslant \|a - x\| + \|a - x'\| \\ &< 2(d + \alpha/6) \\ &< \alpha, \end{aligned}$$

a contradiction. Thus $d \geqslant \alpha/6 > 0$. With

$$\gamma \equiv \|a - x\|^{-1} + \|a - x'\|^{-1} > 12r/\alpha,$$

we now have

$$\begin{aligned} &\|\, \|a - x\|^{-1}(a - x) + \|a - x'\|^{-1}(a - x')\| \\ &\quad = \gamma \|a - (\gamma^{-1}\|a - x\|^{-1} x + \gamma^{-1}\|a - x'\|^{-1} x')\| \\ &\quad \geqslant \gamma d \\ &\quad \geqslant \gamma\alpha/6 \\ &\quad > 2r \end{aligned}$$

whence

$$\|\, \|a - x\|^{-1}(a - x) - \|a - x'\|^{-1}(a - x')\| \leqslant \alpha/3(1 + d).$$

It follows that

$$\begin{aligned} \|x - x'\| &\leqslant \|x - (a - d\|a - x\|^{-1}(a - x))\| \\ &\quad + d\,\|\, \|a - x\|^{-1}(a - x) - \|a - x'\|^{-1}(a - x')\| \\ &\quad + \|x' - (a - d\|a - x'\|^{-1}(a - x'))\| \\ &\leqslant \|(d\|a - x\|^{-1} - 1)(a - x)\| + d\alpha/3(1 + d) \\ &\quad + \|(d\|a - x'\|^{-1} - 1)(a - x')\| \end{aligned}$$

$$\begin{aligned}(\|a - x\| - d) + \lambda/3 + (\|a - x'\| - d) \\ < \lambda/6 + \lambda/3 + \lambda/6 \\ = \lambda.\end{aligned}$$

This again contradicts the definition of λ. Hence

$$\max(\|a - x\|, \|a - x'\|) \geq \beta > 0,$$

and a has at most one best approximation in X. ∎

Particular cases of interest are those where E is a Hilbert space or an L^p-space $(1 < p < \infty)$: the uniform convexity of the former is comparatively trivial to establish, while that of the latter is proved in the Corollary to Theorem 1, Chapter 9 of [1].

We should point out that the proximinality of a finite-dimensional subspace X of a uniformly convex normed space E can be proved without appeal to 2.2, by an argument akin to that used in 2.3 to prove that an element a of E has at most one best approximation in X (cf. [1, Chap. 9, Exercise 5]).

4. Characterization of Minimax Polynomial Approximations

Perhaps the most interesting example of a best-approximation problem in which E is not uniformly convex, but the conditions of 2.2 are satisfied classically, is that of *minimax approximation by polynomials*, in which $E = C[0, 1]$ (with the usual "sup norm") and $X = \text{span}\{1, x, \dots, x^n\}$, the space of polynomials of degree at most n. The application of 2.2 to this situation appears to require a detailed analysis of the constructive content of the classical characterization of minimax polynomials, obtained by Borel and discussed in Chapter 3 of [7]. Incidentally, it is easy to see that the classical characterization is essentially nonconstructive, even in the simplest case $n = 0$, as it entails that any element of $C[0, 1]$ attains its supremum and infimum.

Throughout the remaining sections of this paper, a will be an element of $C[0, 1]$, ν a nonnegative integer and, for each integer $n \geq 0$, X_n the linear subspace $\text{span}\{1, \dots, x^n\}$ of $C[0, 1]$.

Let $p \in X_\nu$ and $\epsilon > 0$. By an *ϵ-alternant* of a and p, we mean an ordered pair comprising an integer $j \in \{0, 1\}$ and a strictly increasing sequence $(\eta_1, \dots, \eta_{\nu+2})$ of $\nu + 2$ points of $[0, 1]$ such that

$$(-1)^{k+j}(a - p)(\eta_k) > \|a - p\| - \epsilon \qquad (k = 1, \dots, \nu + 2).$$

If also $0 < \epsilon < \|a - p\|$ and $n \in \{0, \dots, \nu\}$, we define an *(n, ϵ)-prealternant* of a and p to be an ordered pair comprising an integer $j \in \{0, 1\}$ and a strictly increasing sequence $0 = x_1 < x_2 < \cdots < x_{2n-4} = 1$ of $2n-4$ points of $[0, 1]$ such that

$$(-1)^j (a-p)(x_2) > \|a-p\| - \epsilon,$$
$$(-1)^{n-1-j} (a-p)(x_{2n+3}) > \|a-p\| - \epsilon,$$
$$(-1)^{k-j} (a-p)(x_r) \geq \|a-p\| - \epsilon \quad (r = 2k-1, 2k-2;\ k = 1, \dots, n)$$

and

$$\sup\{|(a-p)(x)| : x_{2k} \leqslant x \leqslant x_{2k-1}\} < \|a-p\| \qquad (k = 1, \dots, n-1)$$

4.1. LEMMA. *Let $p \in X_\nu$ and $0 < \epsilon < \|a-p\|$. Then either $\|a-p\| > \operatorname{dist}(a, X_\nu)$ or there exists a $(0, \epsilon)$-prealternant of a and p.*

Proof. Let M, m be respectively the sup, inf of $a-p$ over $[0, 1]$. Either $\|a-p\| > \min(-m, M)$ or $\min(-m, M) \geq \|a-p\| - \epsilon$. In the former case, we choose α so that

$$0 < \alpha < \tfrac{1}{2}(\|a-p\| - \min(-m, M)).$$

Then, if $\|a-p\| > -m$(when $\|a-p\| = M$), we set $q = p - \alpha \in X_\nu$; so that, for each x in $[0, 1]$,

$$\begin{aligned}(a-q)(x) &\leqslant \|a-p\| - \alpha,\\ (q-a)(x) &\leqslant \alpha - \sup\{(p-a)(x) : x \in [0, 1]\}\\ &= \alpha - m\\ &< \|a-p\| - \alpha,\end{aligned}$$

and therefore

$$\|a-p\| \geq \|a-q\| + \alpha > \operatorname{dist}(a, X_\nu).$$

We obtain the same inequality in the case $\|a-p\| > M$ by taking $q = p - \alpha$.

On the other hand, if $\min(-m, M) \geq \|a-p\| - \epsilon$, we choose ξ, η in $[0, 1]$ so that

$$(a-p)(\xi) > \|a-p\| - \epsilon$$

and

$$(p-a)(\eta) > \|a-p\| - \epsilon.$$

As $a-p$ is uniformly continuous, we may assume that $\xi < \eta$. We now compute α_1 so that

$$\|a-p\| - \epsilon < \alpha_1 < (a-p)(\xi)$$

and

$$K_1 = \{x \in [0, \eta] : (a-p)(x) \geq \alpha_1\}$$

is compact. With $x_1 = 0$, $x_2 = \sup K_1$, we then compute λ_2 so that

$$\|a - p\| - \epsilon < \lambda_2 < (p - a)(\eta)$$

and

$$K_2 = \{x \in [x_2, \eta] : (p - a)(x) \geq \lambda_2\}$$

is compact. To complete the construction of a $(0, \epsilon)$-prealternant of a and p, it only remains to set $j = 0$, $x_3 = \inf K_2$ and $x_4 = 1$. ∎

4.2. LEMMA. *Let $m \in \{0, \ldots, \nu - 1\}$, and $0 < \epsilon < \|a - p\|$. Suppose that there exists an (m, ϵ)-prealternant of a and p. Then either $\|a - p\| > \operatorname{dist}(a, X_\nu)$ or there exists an $(m - 1, \epsilon)$-prealternant of a and p.*

Proof. Let $(j, (t_1, \ldots, t_{2m-1}))$ be an (m, ϵ)-prealternant of a and p, and define

$$\mu = \max_{k=1,\ldots,m-2} \sup\{(-1)^{k-j} (a - p)(x) : t_{2k-1} \leq x \leq t_{2k}\}.$$

Either $\|a - p\| > \mu$ or $\mu > \|a - p\| - \epsilon$. In the former case, choosing $\lambda > 0$ so that

$$\max(\mu, \max_{k=1,\ldots,m-1} \sup\{|(a - p)(x)| : t_{2k} \leq x \leq t_{2k-1}\}) < \|a - p\| - 2\lambda,$$

we set

$$\beta = 2^{-m-1} \lambda \prod_{k=1}^{m-1} (t_{2k-1} - t_{2k}),$$

$$z_k = \tfrac{1}{2}(t_{2k} + t_{2k+1}) \qquad (k = 1, \ldots, m + 1)$$

and

$$q(x) = p(x) - (-1)^j \lambda \prod_{k=1}^{m-1} (z_k - x) \qquad (x \in [0, 1]).$$

Then $\beta > 0$ and $q \in X_{m-1} \subset X_\nu$. Supposing that $\|a - q\| > \|a - p\| - \beta$, we choose ζ in $[0, 1]$ so that

$$|(a - q)(\zeta)| > \|a - p\| - \beta.$$

Then

$$\begin{aligned} |(a - p)(\zeta)| &\geq |(a - q)(\zeta)| - |(p - q)(\zeta)| \\ &> \|a - p\| - \beta - \lambda \prod_{k=1}^{m-1} |z_k - \zeta| \\ &\geq \|a - p\| - 2\lambda \\ &> \max_{k=1,\ldots,m-1} \sup\{|(a - p)(x)| : t_{2k} \leq x \leq t_{2k-1}\}. \end{aligned}$$

From this and the uniform continuity of $a - p$ on [0, 1], it follows that there exists $i \in \{1,..., m + 2\}$ with $t_{2i-1} \leqslant \zeta \leqslant t_{2i}$. Noting that $(-1)^{i-1} \prod_{k=1}^{m+1} (z_k - \zeta) > 0$, we have

$$\begin{aligned}(-1)^{i-j}(q - a)(\zeta) &= (-1)^{i-j}(q - p)(\zeta) + (-1)^{i-j}(p - a)(\zeta) \\ &\leqslant (-1)^i \alpha \prod_{k=1}^{m+1} (z_k - \zeta) + \| a - p \| \\ &\leqslant -\alpha \prod_{k=1}^{m+1} \tfrac{1}{2}(t_{2k+1} - t_{2k}) + \| a - p \| \\ &= \| a - p \| - \beta.\end{aligned}$$

Hence (by our choice of ζ)

$$(-1)^{i-j} (a - q)(\zeta) > \| a - p \| - \beta,$$

and so

$$\begin{aligned}(-1)^{i-j}(a - p)(\zeta) &= (-1)^{i-j}(a - q)(\zeta) + (-1)^{i-j}(q - p)(\zeta) \\ &> \| a - p \| - \beta + (-1)^i \alpha \prod_{k=1}^{m+1} (z_k - \zeta) \\ &\geqslant \| a - p \| - 2\alpha \\ &> \mu.\end{aligned}$$

This contradicts the definition of μ. Hence we must have

$$\| a - q \| \leqslant \| a - p \| - \beta < \| a - p \| ,$$

and therefore $\| a - p \| > \operatorname{dist}(a, X_\nu)$.

On the other hand, if $\mu > \| a - p \| - \epsilon$, we choose in turn k, α_1 so that

$$\| a - p \| - \epsilon < \alpha_1 < \sup\{(-1)^{k-j} (a - p)(x) : t_{2k-1} \leqslant x \leqslant t_{2k}\}$$

and

$$K_1 \equiv \{x \in [t_{2k-1}, t_{2k}] : (-1)^{k-j} (a - p)(x) \geqslant \alpha_1\}$$

is compact. If $2 \leqslant k \leqslant m + 1$, we set $y_2 \equiv \inf K_1$, $y_3 \equiv \sup K_1$, and (using the properties of an (m, ϵ)-prealternant, and the uniform continuity of $a - p$) observe that $t_{2k-1} < y_2 < y_3 < t_{2k}$ and

$$(-1)^{k-j} (a - p)(y_2) = (-1)^{k-j} (a - p)(y_3) = \alpha_1 > \| a - p \| - \epsilon.$$

Now choose α_2 so that

$$\| a - p \| - \epsilon < \alpha_2 < \min((-1)^{k-j-1} (a - p)(t_{2k-1}), (-1)^{k-j-1} (a - p)(t_{2k}))$$

and the sets

$$K_2 = \{x \in [t_{2k-1}, y_2] : (-1)^{k-j-1}(a-p)(x) \geq \alpha_2\},$$
$$K_3 = \{x \in [y_3, t_{2k}] : (-1)^{k-j-1}(a-p)(x) \geq \alpha_2\}$$

are both compact. With $y_1 = \sup K_2$, $y_4 = \inf K_3$, the uniform continuity of $a - p$ ensures that

$$t_{2k-1} < y_1 < y_2 < y_3 < y_4 < t_{2k}$$

and

$$(-1)^{k-j-1}(a-p)(y_1) = (-1)^{k-j-1}(a-p)(y_4) = \alpha_2 > \|a-p\| - \epsilon.$$

To complete the construction of an $(m-1, \epsilon)$-prealternant $(j, (x_1, \ldots, x_{2m-6}))$ of a and p, it only remains to define

$$x_r = t_r \qquad (r = 1, \ldots, 2k-1),$$
$$x_{2k-1+s} = y_s \qquad (s = 1, 2, 3, 4),$$
$$x_{2m-6} = 1,$$

and, if $k < m-1$,

$$x_{2k+3+s} = t_{2k+1+s} \qquad (s = 1, \ldots, 2m-2k-3).$$

If $k = 1$, we set $x_2 = \sup K_1$, note that $x_2 < t_2$, and choose α_2 so that

$$\|a-p\| - \epsilon < \alpha_2 < (-1)^j (a-p)(t_2)$$

and

$$A = \{x \in [x_2, t_2] : (-1)^j (a-p)(x) \geq \alpha_2\}$$

is compact. Then, setting $x_1 = 1$, $x_3 = \inf A$ and $x_{3+s} = t_{s+1}$ $(s = 1, \ldots, 2m-3)$, we easily show that $(j, (x_1, \ldots, x_{2m-6}))$ is an $(m-1, \epsilon)$-prealternant of a and p.

The case $k = m-2$ is handled in a similar manner. ∎

4.3. Proposition. *If $p \in X_\nu$ and $0 < \epsilon < \|a-p\|$, then either $\|a-p\| > \operatorname{dist}(a, X_\nu)$ or there exists an ϵ-alternant of a and p.*

Proof. Applying 4.1, and then 4.2 repeatedly, we see that either $\|a-p\| > \operatorname{dist}(a, X_\nu)$ or there exists a (ν, ϵ)-prealternant $(j, (t_1, \ldots, t_{2\nu-4}))$ of a and p. In the latter case, choosing points $\eta_k \in [t_{2k-1}, t_{2k}]$ so that

$$(-1)^{k-1-j}(a-p)(\eta_k) > \|a-p\| - \epsilon \qquad (k = 1, \ldots, \nu-2),$$

we obtain an ϵ-alternant $(1-j, (\eta_1, \ldots, \eta_{\nu-2}))$ of a and p. ∎

We are now able to derive the constructive analog of the classical characterization of minimax polynomial approximations.

4.4. THEOREM. *A necessary and sufficient condition that $b \in X_\nu$ be a minimax approximation to a in X_ν is that, for each $\epsilon > 0$, there exists an ϵ-alternant of a and b.*

Proof. Given $\epsilon > 0$, we have either $\frac{1}{2}\epsilon > \|a - b\|$ or $\|a - b\| > 0$. In the former case, as

$$\|a - b\| - \epsilon < -\tfrac{1}{2}\epsilon < -\|a - b\| \leqslant (a - b)(x) \qquad (x \in [0, 1]),$$

if $(\eta_1, \ldots, \eta_{\nu+2})$ is any strictly increasing sequence of $\nu + 2$ points of $[0, 1]$, $(0, (\eta_1, \ldots, \eta_{\nu+2}))$ is an ϵ-alternant of a and b. On the other hand, if $0 < \|a - b\| = \operatorname{dist}(a, X_\nu)$, we see from 4.3 that there exists a $\min(\frac{1}{2}\|a - b\|, \epsilon)$-alternant, which is also clearly an ϵ-alternant, of a and b.

Now suppose the given condition holds, and also that $\|a - b\| > \operatorname{dist}(a, X_\nu)$. Choosing $p \in X_\nu$ so that $\|a - b\| > \|a - p\|$, we set

$$\alpha = \tfrac{1}{2}(\|a - b\| - \|a - p\|)$$

and construct an α-alternant $(j, (\eta_1, \ldots, \eta_{\nu+2}))$ of a and b. Then, for each $k \in \{1, \ldots, \nu + 2\}$,

$$\begin{aligned}(-1)^{k-j}(p - b)(\eta_k) &= (-1)^{k-j}(p - a)(\eta_k) + (-1)^{k-j}(a - b)(\eta_k)\\ &> -\|a - p\| + \|a - b\| - \alpha\\ &= \alpha\\ &> 0.\end{aligned}$$

It follows that the polynomial $p - b$, of degree at most ν, has at least $\nu + 1$ changes of sign. Thus $p = b$, and we obtain the contradiction $\|a - p\| = \|a - b\|$. Hence, in fact, $\|a - b\| = \operatorname{dist}(a, X_\nu)$. ∎

5. EXISTENCE OF MINIMAX POLYNOMIAL APPROXIMATIONS

Having characterized minimax polynomial approximations, we now show how they can be constructed. In order to apply 2.2, we need a lemma on the location of roots of a polynomial.

5.1. LEMMA. *Let n be a positive integer, $c \in \mathbb{R}$, $|c| > 0$. Let $\xi_1, \ldots, \xi_n$ be complex numbers such that $\prod_{r=1}^{n} (x - \xi_r) \in \mathbb{R}$ for each $x \in \mathbb{R}$. Let $(\eta_1, \ldots, \eta_{n+3})$ be a strictly increasing sequence of $n + 3$ points of $\mathbb{R}$, and suppose that $|\eta_j - \operatorname{Re} \xi_k| > 0$ whenever $j \in \{1, \ldots, n + 3\}$ and $k \in \{1, \ldots, n\}$.*

Then

> *there exists* $s \in \{1,\ldots, n-2\}$ *such that* $(-1)^s c \prod_{r=1}^{n}(x - \xi_r) > 0$ *for each* $x \in [\eta_s, \eta_{s+1}]$ *and* $\mathrm{Re}\ \xi_r \notin [\eta_s, \eta_{s+1}]$ *for each* $r \in \{1,\ldots, n\}$. $(*)$

Proof. For convenience, let

$$p(x) = c \prod_{r=1}^{n} (x - \xi_r) \qquad (x \in \mathbb{C}).$$

There are two main steps in the proof.

5.1.1. *Let* $i \in \{0,\ldots, n\}$, $k \in \{0,\ldots, n-i\}$. *Suppose that* $]\eta_k, \eta_{k+i+2}[$ *contains* $\mathrm{Re}\ \xi_r$ *for exactly i distinct values of r; and that, in the case* $i \geqslant 1$, $\mathrm{Re}\ \xi_r \in\ ]\eta_{k+r}, \eta_{k+r+1}[$ *for* $r = 1,\ldots, i$. *Then* $(*)$ *obtains.*

Indeed, as $[\eta_k, \eta_{k+1}]$ is at positive distance from each $\mathrm{Re}\ \xi_r$, there exists $j \in \{0, 1\}$ such that $(-1)^j p(x) > 0$ for each x in $[\eta_k, \eta_{k+1}]$. In the case $i \geqslant 1$, as the numbers $\mathrm{Re}\ \xi_r$ $(r = 1,\ldots, i)$ are distinct, and there are no other roots of p in $[\eta_k, \eta_{k+i+2}]$, we see that $\xi_1, \ldots, \xi_i$ are distinct real roots of p, and that the sign changes of p in $[\eta_k, \eta_{k+i+2}]$ occur precisely at these roots. Thus

$$(-1)^{i+j} p(x) > 0 \qquad (\eta_{k+i+1} \leqslant x \leqslant \eta_{k+i+2}).$$

It is clear that this inequality also holds for $i = 0$. It only remains to set

$$\begin{aligned} s &= k & &\text{if } k + j \text{ is even,} \\ s &= k + i + 1 & &\text{if } k + j \text{ is odd.} \end{aligned}$$

5.1.2. *Let* i, k, λ *be integers with* $0 \leqslant \lambda \leqslant i \leqslant n$, $0 \leqslant k \leqslant n - i$, *and suppose that* $\mathrm{Re}\ \xi_r \in [\eta_k, \eta_{k+i+2}]$ *for exactly* λ *distinct values of r. Then* $(*)$ *obtains.*

As the case $\lambda = 0$ follows from 5.1.1, we may assume that $\lambda \geqslant 1$. Let $m \in \{1,\ldots, n\}$, suppose we have proved 5.1.2 for $i = 0,\ldots, m-1$, and consider the case $i = m$. If any of the $\mathrm{Re}\ \xi_r$ belongs to $[\eta_k, \eta_{k+1}]$, then there are at most $m - 1$ distinct values of r with $\mathrm{Re}\ \xi_r \in [\eta_{k+1}, \eta_{k+m+2}]$, and so $(*)$ obtains. With $s \in \{0,\ldots, \lambda - 1\}$, suppose that $[\eta_k, \eta_{k+s+1}]$ contains $\mathrm{Re}\ \xi_r$ for exactly s distinct values of r; and that, if $s \geqslant 1$, each of the intervals $[\eta_{k+t}, \eta_{k+t+1}]$ $(1 \leqslant t \leqslant s)$ contains $\mathrm{Re}\ \xi_r$ for exactly one value of r. If $[\eta_{k+s+1}, \eta_{k+s+2}]$ contains none of the $\mathrm{Re}\ \xi_r$, we have exactly s distinct values of r with $\mathrm{Re}\ \xi_r \in [\eta_k, \eta_{k+s+2}]$, from which $(*)$ follows. If $\mathrm{Re}\ \xi_r \in [\eta_{k+s+1}, \eta_{k+s+2}]$ for more than one value of r, then there are at most $\lambda - s - 2$ distinct values of r with $\mathrm{Re}\ \xi_r \in [\eta_{k+s+2}, \eta_{k+\lambda+2}]$, and we again have $(*)$. It now follows by induction on s that either $(*)$ holds or, after suitable reindexing of the ξ_r,

$$\mathrm{Re}\ \xi_r \in [\eta_{k+r}, \eta_{k+r+1}] \qquad (r = 1,\ldots, \lambda).$$

From this, we immediately obtain (*) by an application of 5.1.1. This completes the inductive proof of 5.1.2. That of 5.1 is now completed by taking $i = n$ in 5.1.2. ∎

We now reach the end of the search for a constructive proof of the existence of minimax polynomial approximations.

5.2. THEOREM. *Each element a of $C[0, 1]$ has a minimax approximation b in X_ν that is unique in the sense that, if $p \in X_\nu$ and $\| p - b \| > 0$, then $\| a - p \| > \| a - b \|$.*

Proof. Let p, q belong to X_ν, with $\| p - q \| > 0$. In view of 2.2, it will suffice to prove that $\max(\| a - p \|, \| a - q \|) > \operatorname{dist}(a, X_\nu)$. We proceed by induction on ν. If $\nu = 0$, let M, m be respectively, the sup, inf of a over $[0, 1]$. Without loss of generality, we may take $p > q$. Then either $p > \frac{1}{2}(M + m)$, in which case

$$\| a - p \| > p - m > \tfrac{1}{2}(M - m);$$

or $\frac{1}{2}(M + m) > q$, when

$$\| a - q \| > M - q > \tfrac{1}{2}(M - m).$$

As $\frac{1}{2}(M + m) \in X_\nu$ and $\| a - \frac{1}{2}(M + m) \| = \frac{1}{2}(M - m)$, it follows that

$$\max(\| a - p \|, \| a - q \|) > \tfrac{1}{2}(M - m) = \operatorname{dist}(a, X_\nu).$$

Now let n be a positive integer, suppose we have proved the Proposition for $\nu = 0, \ldots, n - 1$, and consider the case $\nu = n$. As

$$\| a - p \| + \| a - q \| \geqslant \| p - q \| > 0,$$

we may assume that $\| a - q \| > 0$. If $p(x) = \sum_{k=0}^{\nu} p_k x^k$ and $q(x) = \sum_{k=0}^{\nu} q_k x^k$, then

$$\| p_\nu - q_\nu \| + \| \sum_{k=0}^{\nu-1} (p_k - q_k)\, x^k \| \geqslant \| p - q \| > 0;$$

so that either $| p_\nu - q_\nu | > 0$ or $\| \sum_{k=0}^{\nu-1} (p_k - q_k)\, x^k \| > 0$. In the latter case, by our induction hypothesis, we have either

$$\begin{aligned} \| a - q \| &= \| a - q_\nu x^\nu - \sum_{k=0}^{\nu-1} q_k x^k \| \\ &> \operatorname{dist}(a - q_\nu x^\nu, X_{\nu-1}) \\ &\geqslant \operatorname{dist}(a, X_\nu); \end{aligned}$$

or, as we may suppose,

$$\|a - q_\nu x^\nu - \sum_{k=0}^{\nu-1} p_k x^k\| < \mathrm{dist}(a - q_\nu x^\nu, X_{\nu-1}).$$

Then either $|p_\nu - q_\nu| > 0$; or

$$|p_\nu - q_\nu| < \|a - q_\nu x^\nu - \sum_{k=0}^{\nu-1} p_k x^k\| - \mathrm{dist}(a - q_\nu x^\nu, X_{\nu-1}),$$

in which case

$$\begin{aligned} \|a - p\| &\geq \|a - q_\nu x^\nu - \sum_{k=0}^{\nu-1} p_k x^k\| - |p_\nu - q_\nu| \\ &> \mathrm{dist}(a - q_\nu x^\nu, X_{\nu-1}) \\ &\geq \mathrm{dist}(a, X_\nu). \end{aligned}$$

It is now clear that we may assume that $|p_\nu - q_\nu| > 0$.

With δ a modulus of uniform continuity for $a - q$ on [0, 1], we set

$$\begin{aligned} \beta &= \tfrac{1}{4}\delta(\|a - q\|), \\ \mu &= \nu\,|p_\nu - q_\nu|\,\beta^{\nu-1}, \\ \epsilon &= \min(\tfrac{1}{2}\|a - q\|, \mu\delta(\|a - q\|)/6). \end{aligned}$$

Then $\beta > 0$, $\mu > 0$, $\epsilon > 0$. By 4.3, either $\|a - q\| > \mathrm{dist}(a, X_\nu)$ or, as we may assume, there exists an ϵ-alternant $(j, (\eta_1, \ldots, \eta_{\nu+2}))$ of a and q.

We now observe that it will suffice to find $k \in \{1, \ldots, \nu + 2\}$ such that $(-1)^{k-j}(p - q)(\eta_k) < -\epsilon$. For then

$$\begin{aligned} \|a - p\| &\geq (-1)^{k-j}(a - p)(\eta_k) \\ &= (-1)^{k-j}(a - q)(\eta_k) + (-1)^{k-j}(q - p)(\eta_k) \\ &> \|a - q\| - \epsilon + \epsilon \\ &= \|a - q\|, \end{aligned}$$

and therefore $\|a - p\| > \mathrm{dist}(a, X_\nu)$.

As either

$$\min_{k=1,\ldots,\nu+2}(-1)^{k-j}(p - q)(\eta_k) < -\epsilon$$

or

$$\min_{k=1,\ldots,\nu+2}(-1)^{k-j}(p - q)(\eta_k) > -2\epsilon,$$

we clearly may assume the latter. For convenience, we also take $j = 0$, the case $j = 1$ being similar. If $\nu = 1$, and $\alpha \in \{0, 1\}$ is chosen so that $|p_1 - q_1| = (-1)^\alpha (p_1 - q_1)$, then

$$\begin{aligned}(-1)^{\alpha+1}(p-q)(\eta_{\alpha+1}) &= (-1)^{\alpha+1}(p-q)(\eta_\alpha) + (-1)^{\alpha+1}(p_1-q_1)(\eta_{\alpha+1}-\eta_\alpha)\\ &< 2\epsilon - |p_1-q_1|\,\delta(\|a-q\|)\\ &< -\epsilon,\end{aligned}$$

and our proof is complete.

We now take $\nu > 1$, and observe that, by [1, Chap. 5, Theorem 8], there exist complex numbers $\xi_1, \dots, \xi_{\nu-1}$ such that

$$(p-q)'(x) = \nu(p_\nu - q_\nu)\prod_{r=1}^{\nu-1}(x-\xi_r) \qquad (x \in [0, 1]).$$

Note also that $|(p-q)'(x)| \geqslant \mu$ whenever $\min_{r=1,\dots,\nu-1} |x - \xi_r| \geqslant \beta$. As $a - q$ is uniformly continuous, we may assume that $|\eta_j - \operatorname{Re}\xi_k| > 0$ whenever $j \in \{1,\dots,\nu+2\}$ and $k \in \{1,\dots,\nu-1\}$. By 5.1, there exists $s \in \{1,\dots,\nu+1\}$ such that $(-1)^s(p-q)'(x) > 0$ for each $x \in [\eta_s, \eta_{s+1}]$ and $\operatorname{Re}\xi_r \notin [\eta_s, \eta_{s+1}]$ for each $r \in \{1,\dots,\nu-1\}$. For each such r and each x in $[\eta_s + \beta, \eta_{s+1} - \beta]$, we then have $|x - \xi_r| \geqslant |x - \operatorname{Re}\xi_r| \geqslant \beta$. Thus

$$\begin{aligned}(-1)^{s+1}(p-q)(\eta_{s+1}) &= (-1)^{s+1}(p-q)(\eta_s)\\ &\quad - \left(\int_{\eta_s}^{\eta_s+\beta} + \int_{\eta_s+\beta}^{\eta_{s+1}-\beta} + \int_{\eta_{s+1}-\beta}^{\eta_{s+1}}\right)(-1)^s(p-q)'(x)\,dx\\ &< 2\epsilon - \int_{\eta_s+\beta}^{\eta_{s+1}-\beta}\mu\,dx\\ &= 2\epsilon - \mu(\eta_{s+1} - \eta_s - 2\beta)\\ &\leqslant 2\epsilon - \tfrac{1}{2}\mu\delta(\|a-q\|)\\ &= -\epsilon.\end{aligned}$$

This completes the proof. ∎

6. Lipschitz Conditions on the Minimax Approximation Mapping

Let P_ν be the mapping which carries an element ψ of $C[0, 1]$ to the unique element $P_\nu\psi$ of X_ν such that $\|\psi - P_\nu\psi\| = \operatorname{dist}(\psi, X_\nu)$. Our aim is to prove that P_ν is locally Lipschitzian on $C[0, 1] - X_\nu$ (6.3 below).

6.1. Lemma. *Let $p \in X_\nu$, $\epsilon > 0$ and $\alpha > 0$. Let $x_1, \dots, x_{\nu-1}$ be points of*

[0, 1] *with* $\min_{k=1,\dots,\nu}(x_{k+1} - x_k) \geqslant \alpha$, *and suppose that* $|p(x_k)| \leqslant \epsilon$ *for each* k *in* $\{1,\dots, \nu + 1\}$. *Then*

$$\|p\| \leqslant \alpha^{-\nu}\left(\sum_{k=1}^{\nu+1} 1/(k-1)!(\nu-k+1)!\right)\epsilon.$$

Proof. For each x in [0, 1], we have

$$p(x) = \sum_{k=1}^{\nu+1} L_k(x)\, p(x_k),$$

where

$$L_k(x) = \left(\prod_{j=1, j\neq k}^{\nu+1} (x - x_j)\right)\Big/ \prod_{j=1, j\neq k}^{\nu+1} (x_k - x_j) \qquad (k = 1,\dots, \nu + 1).$$

The lemma follows from this and the inequality

$$|L_k(x)| \leqslant 1\Big/ \prod_{j=1, j\neq k}^{\nu+1} \alpha\, |k - j| \leqslant 1/\alpha^{\nu}\,(k-1)!(\nu - k + 1)!\,,$$

valid for $k = 1,\dots, \nu + 1$. ∎

6.2. LEMMA. *Let* $\alpha > 0$, *and let* $\eta_1,\dots, \eta_{\nu+2}$ *be points of* [0, 1] *such that* $\eta_{k+1} - \eta_k \geqslant \alpha$ *for each* k *in* $\{1,\dots, \nu + 1\}$. *Then there exists* $c > 0$ *such that* $\|p\| < c\epsilon$ *whenever* $\epsilon > 0$, $p \in X_\nu$ *and* $(-1)^k p(\eta_k) > -\epsilon$ *for each* k *in* $\{1,\dots, \nu + 2\}$.

Proof. Let $\epsilon > 0$, $p \in X_\nu$ and $(-1)^k p(\eta_k) > -\epsilon$ for each k in $\{1,\dots, \nu + 2\}$. If $\nu = 0$, then

$$-\epsilon < p(\eta_1) = p(\eta_2) < \epsilon;$$

so that $\|p\| < \epsilon$, and we can take $c = 1$.

Now let n be a positive integer, suppose we have proved 6.2 for $\nu = 0,\dots, n - 1$, and consider the case $\nu = n$. Writing $p(x) = \sum_{r=0}^{\nu} p_r x^r$, we have either $|p_\nu| > 0$ or $|p_\nu| < \epsilon$. In the latter case, for each $k \in \{1,\dots, \nu + 1\}$,

$$(-1)^k \sum_{r=0}^{\nu-1} p_r \eta_k{}^r = (-1)^k (p(\eta_k) - p_\nu \eta_k{}^{\nu})$$
$$> -\epsilon - |p_\nu|.$$

By our induction hypothesis, there exists $c' > 0$ (c' independent of p and ϵ) such that $\|\sum_{r=0}^{\nu-1} p_r x^r\| \leqslant c'(2\epsilon)$; whence

$$\|p\| \leqslant \left\|\sum_{r=0}^{\nu-1} p_r x^r\right\| + |p_\nu| \leqslant (2c' + 1)\epsilon.$$

It is now clear that we may assume that $|p_\nu| > 0$.

If $\nu = 1$, and $j \in \{0, 1\}$ is chosen so that $|p_1| = (-1)^j p_1$, then $(-1)^j p$ is increasing in $[0, 1]$; so that

$$-\epsilon < (-1)^j p(\eta_j) < (-1)^j p(\eta_{j+1}) < \epsilon.$$

Thus $|p(\eta_j)| < \epsilon$, $|p(\eta_{j+1})| < \epsilon$. The result in this case now follows from 6.1.

We now take $\nu > 1$ and compute $\xi_1, \ldots, \xi_{\nu-1}$ in $\mathbb{C}$ so that

$$p'(x) = \nu p_\nu \prod_{r=1}^{\nu-1} (x - \xi_r) \qquad (x \in [0, 1]).$$

As p is uniformly continuous on $[0, 1]$, we may assume that $|\eta_j - \operatorname{Re} \xi_k| > 0$ whenever $j \in \{1, \ldots, \nu + 2\}$ and $k \in \{1, \ldots, \nu - 1\}$. By 5.1, there exists $s \in \{1, \ldots, \nu + 1\}$ such that $(-1)^s p'(x) > 0$ for each x in $[\eta_s, \eta_{s+1}]$. Thus, for each such x,

$$-\epsilon < (-1)^s p(\eta_s) \leqslant (-1)^s p(x) \leqslant (-1)^s p(\eta_{s+1}) < \epsilon,$$

and therefore $|p(x)| < \epsilon$. Applying 6.1 to the points $\eta_s + k\nu^{-1}(\eta_{s+1} - \eta_s)$ $(k = 0, \ldots, \nu)$, we obtain $\|p\| \leqslant c\epsilon$ with

$$c = \alpha^{-\nu}\nu^\nu \left(\sum_{k=1}^{\nu-1} 1/(k-1)!(\nu-k-1)! \right). \quad \blacksquare$$

6.3. THEOREM. *Let $a \in C[0, 1]$, with $\|a - P_\nu a\| > 0$. Then there exists $c > 0$ such that*

$$\|P_\nu a' - P_\nu a\| \leqslant c \|a' - a\|$$

for each $a' \in C[0, 1]$.

Proof. Given $\alpha \in]0, \|a - P_\nu a\|[$, we construct an α-alternant $(j, (\eta_1, \ldots, \eta_{\nu+2}))$ of a and $P_\nu a$ (4.4), and observe that, for each $k \in \{1, \ldots, \nu + 2\}$,

$$\begin{aligned}
&(-1)^{k-j} (P_\nu a' - P_\nu a)(\eta_k) \\
&\quad = (-1)^{k-j} (a - P_\nu a)(\eta_k) + (-1)^{k-j} (P_\nu a' - a')(\eta_k) - (-1)^{k-j} (a' - a)(\eta_k) \\
&\quad > \operatorname{dist}(a, X_\nu) - \alpha - \operatorname{dist}(a', X_\nu) - \|a' - a\| \\
&\quad \geqslant -2\|a' - a\| - \alpha.
\end{aligned}$$

If δ is a modulus of uniform continuity for $a - P_\nu a$ on $[0, 1]$, our choice of α ensures that

$$\eta_{k+1} - \eta_k \geqslant \delta(\|a - P_\nu a\|) \qquad (k = 1, \ldots, \nu + 1).$$

Thus (6.2) there exists $c > 0$ (c depending on a, but independent of λ and a') such that

$$\|P_v a' - P_v a\| \leq \tfrac{1}{2}c(2\|a' - a\| + \lambda).$$

As $\lambda \in]0, \|a - P_v a\|[$ is arbitrary, we have

$$\|P_v a' - P_v a\| \leq c\|a' - a\| \qquad (a' \in C[0, 1]). \quad \blacksquare$$

Remark. From 6.3, we obtain a particularly simple proof of the pointwise continuity of P_v on $C[0, 1]$. Given $a \in C[0, 1]$ and $\epsilon > 0$, we have either $\|a - P_v a\| < \frac{1}{4}\epsilon$ or $0 < \|a - P_v a\|$. In the former case, if $a' \in C[0, 1]$ and $\|a - a'\| < \frac{1}{4}\epsilon$, then

$$\begin{aligned}\|a' - P_v a'\| &= \mathrm{dist}(a', X_v)\\ &\leq \|a - a'\| + \mathrm{dist}(a, X_v)\\ &< \tfrac{1}{2}\epsilon,\end{aligned}$$

whence

$$\|P_v a' - P_v a\| \leq \|P_v a' - a'\| + \|a' - a\| + \|a - P_v a\| < \epsilon.$$

On the other hand, if $0 < \|a - P_v a\|$, then, computing $c > 0$ as in 6.3, we have $\|P_v a' - P_v a\| < \epsilon$ whenever $a' \in C[0, 1]$ and $\|a - a'\| < c^{-1}\epsilon$.

REFERENCES

1. E. Bishop, "Foundations of Constructive Analysis," McGraw-Hill, New York, 1967
2. D. S. Bridges, "Constructive Functional Analysis," Research Notes in Mathematics No. 28, Pitman Advanced Publishing Programme, London, 1979.
3. D. S. Bridges, "On continuous mappings between locally compact metric spaces," *Bull. London Math. Soc.* **10** (1978), 201–208.
4. D. S. Bridges, "A criterion for compactness in metric spaces?," *Z. Math. Logik Grundlagen Math.* **25** (1979), 97–98.
5. D. S. Bridges, "A constructive proximinality property of finite dimensional linear subspaces," *Rocky Mountain J. Math.*, in press.
6. D. S. Bridges, "A constructive development of Chebyshev approximation theory," manuscript.
7. J. Todd, "Introduction to the Constructive Theory of Functions," Birkhauser-Verlag, Basel, 1963.
8. A. S. Troelstra, "Principles of Intuitionism," Lecture Notes in Mathematics No. 95, Springer-Verlag, Berlin, 1969.

Printed in Belgium