UNIFORM APPROXIMATION FROM TCHEBYCHEFF SYSTEMS

by

R.G. Brookes

No. 39

January 1987

Contents

UNIFORM APPROXIMATION FROM TCHEBYCHEFF SYSTEMS

By R.G. BROOKES

Mathematics Department, University of Canterbury, Christchurch, New Zealand

Section 1: Introduction

This report is concerned with the study of best uniform approximation to $f \in C[a,b]$ from the linear space generated by some finite subset $U = \{u_0, u_1, \ldots, u_n\}$ of $C[a,b]$. By a best uniform approximation we mean $p^* \in \text{span } U$ such that

$$\max\{|f(x) - p^*(x)| : x \in [a,b]\} = \min\{\max\{|f(x) - p(x)| : x \in [a,b]\} : p \in \text{span } U\} .$$

We explore, firstly, the case $U = \{1, x, \ldots, x^n\}$. It will be shown in Section 4 that in this situation each $f \in C[a,b]$ has a unique best approximation and for this best approximation there is a strong characterisation theorem. It is then natural to ask whether these results are true for a more general $U = \{u_0, u_1, \ldots, u_n\}$. If a strong type of linear independence known as the Haar condition is imposed on U then this will indeed turn out to be the case. We will attempt to develop this condition using an approach more intuitively obvious than those found in many standard texts.

When the Haar condition is not satisfied the problem rapidly becomes complicated and it appears that much work remains to be done in this area. A theorem concerning a particularly simple situation is given in Section 8.

Section 2: Notation

A brief summary of notation which will be used without definition during the course of this report.

will signify that a desired contradiction has been achieved.

$P_n[x]$ will denote the set of real polynomials in x of degree $\leqslant n$.

Let $f \in C[a,b]$. Then $\|f\| = \max\{|f(x)| : x \in [a,b]\}$, the uniform norm of f on $[a,b]$.

Let $f \in C[a,b]$, $p \in \text{span } U$. Then

$$E^+ = \{x \in [a,b] : (f(x) - p(x)) = \|f - p\|\}$$

$$E^- = \{x \in [a,b] : -(f(x) - p(x)) = \|f - p\|\}$$

$$E(f) \text{ (or just } E) = E^+ \cup E^-.$$

Note that $E \subset [a,b]$ is the inverse image of $\{-\|f-p\|, \|f-p\|\}$ under the continuous function $f - p$ so that E is closed and bounded hence compact.

Section 3: Preliminary Theorems

We are interested in finding a best approximation to $f \in C[a,b]$ from P, a finite dimensional subspace of $C[a,b]$. The first question to be asked is, of course, does a best approximation always exist? The answer is clearly affirmative as outlined in the following theorem.

Theorem (3.1) Let $f \in C[a,b]$, P a finite dimensional subspace of $C[a,b]$. Then $\exists p^* \in P$ such that $\|f - p^*\| = \min\{|f - p| : p \in P\}$, i.e. there does exist a best approximation.

Proof

Let $q \in P$, $Q = \{g \in C[a,b] : \|f - g\| \leqslant \|f - q\|\}$. Then $\min\{\|f - p\| : p \in P\} = \min\{\|f - p\| : p \in P \cap Q\}$ and since $\|f - p\|$ is a continuous function on $P \cap Q$ (a compact set) this minimum is attained, i.e. $\exists p^* \in P \cap Q \subset P$ such that

$$\|f - p^*\| = \min\{\|f - p\| : p \in P\}.$$

There is a more useful description of the best approximation known as the Kolmogorov Characterisation Theorem. It is just a formal expression of the idea that $p^* \in P$ is a best approximation to $f \in C[a,b]$ iff we can't add any $p \in P$ to p^* to give a better approximation.

Theorem (3.2) (Kolmogorov Characterisation Theorem). Let P be a linear subspace of $C[a,b]$ and $f \in C[a,b]\setminus P$. Then $p^* \in P$ is a best approximation to f from P iff $\nexists p \in P$ such that

$$\begin{cases} p(x) > 0 & x \in E^+ \\ p(x) < 0 & x \in E^- \end{cases}$$

(Recall that $E^+ = \{x \in [a,b] : f(x) - p^*(x) = \|f - p^*\|\}$,
$E^- = \{x \in [a,b] : f(x) - p^*(x) = -\|f - p^*\|\}$).

Proof

[$\Rightarrow$]

Suppose that p^* is a best approximation and that $\exists p \in P$ satisfying the requirements of the theorem.

Then $\forall x \in E$, $(f(x) - p^*(x))p(x) > 0$.

Since E is compact and $(f(x) - p^*(x))p(x)$ is continuous $\exists \varepsilon > 0$ such that

$$\min\{(f(x) - p^*(x))p(x) : x \in E\} = 2\varepsilon.$$

Also, by continuity

$$G = \{x \in [a,b] : (f(x) - p^*(x))p(x) \in (\varepsilon,\infty)\}$$

is an open set (containing E). For $\lambda \in \mathbb{R}$ such that $0 < \lambda < \frac{2\varepsilon}{\|p\|^2}$ we have,

$$\begin{aligned}\forall x \in G,\ 0 \leqslant (f(x) - (p^*(x)) + \lambda p(x)))^2 &= (f(x) - p^*(x))^2 - 2\lambda(f(x) - p^*(x))p(x) \\ &\quad + \lambda^2(p(x))^2 \\ &\leqslant \|f - p^*\|^2 - 2\varepsilon\lambda + \lambda^2\|p\|^2 \\ &< \|f - p^*\|^2.\end{aligned}$$

Now $F = [a,b]\setminus G$ is compact and $|f(x) - p^*(x)| < \|f - p^*\|$ so $\exists \delta > 0$ such that $\forall x \in F$, $|f(x) - p^*(x)| < \|f - p^*\| - \delta$. For $\lambda \in \mathbb{R}$ such that $0 < \lambda < \frac{\delta}{2\|p\|}$ we have

$$\begin{aligned}\forall x \in F,\ |f(x) - (p^*(x) + \lambda p(x))| &\leqslant |f(x) - p^*(x)| + \lambda|p(x)| \\ &\leqslant \|f - p^*\| - \delta + \lambda\|p\| \\ &< \|f - p^*\|.\end{aligned}$$

Thus for $\lambda \in \mathbb{R}$ such that $0 < \lambda < \min\left\{\frac{2\varepsilon}{\|p\|^2}, \frac{\delta}{2\|p\|}\right\}$ $p^* + \lambda p \in P$ is a better approximation to f than p^*. #

[$\Leftarrow$]

Suppose that $p^* \in P$ is not a best approximation. Then there exists a better approximation $q \in P$.

Clearly, $\forall x \in E^+$, $f(x) - q(x) < f(x) - p^*(x)$

$$\Rightarrow \quad 0 < q(x) - p^*(x)$$

$\forall x \in E^-$, $f(x) - q(x) > f(x) - p^*(x)$

$$\Rightarrow \quad 0 > q(x) - p^*(x)$$

i.e. letting $p = q - p^*$ completes the proof.

This characterisation of p^* is of vital importance in what follows.

Section 4: Approximation from $P_n[x]$

As an introduction to our main topic we consider uniform approximation to $f \in C[a,b]$ from $P_n[x]$. It will be shown that there is a concrete and useful characterisation of best approximation in this case and that we can, in fact, speak of the unique best approximation. We hope to understand the fundamental properties of this problem and so make some sort of meaningful generalisation. It is with this aim in mind that the following approach has been chosen. Although it is neither the briefest nor most elegant it appeals as that which shows most clearly the idea involved. An alternative development in a more general setting will be presented in Section 6.

Firstly, recall the Kolmogorov Characterisation Theorem from the previous section.

Theorem (3.2) Let P be a linear subspace of $C[a,b]$ and $f \in C[a,b]\setminus P$. Then $p^* \in P$ is a best approximation to f from P iff $\nexists p \in P$ such that

$$\begin{cases} p(x) > 0 & x \in E^+ \\ p(x) < 0 & x \in E^- . \end{cases}$$

Now consider the case $P = P_n[x]$. It should be obvious that if E consists of less than $n + 2$ points, then we can construct a $p(x)$ as above simply by interpolation. So $|(f - p^*)(x)|$ attains its maximum at (at least) $n + 2$ points. With a little further thought one realises that given a set of less than $n + 2$ points $x_1 < x_2 < \ldots < x_j$ it is easy to construct $p \in P_n[x]$

such that $p(x)$ has alternating signs from x_i to x_{i+1}. Bearing these ideas in mind the next theorem follows naturally.

<u>Theorem (4.1)</u> Let $f \in C[a,b]\setminus P_n[x]$, $p^* \in P_n[x]$. Then p^* is a best approximation to f from $P_n[x]$ iff $(f-p^*)(x)$ attains its maximum modulus at $n+2$ points of $[a,b]$ with alternating sign.

<u>Proof</u> ([3], pp8-9).

$[\Leftarrow]$

Suppose that p^* is not a best approximation. Then $\exists p \in P_n[x]$ satisfying the conditions of Theorem (3.2). This $p(x)$ is continuous, of degree at most n and has alternate sign at $n+2$ points. Thus it has at least $n+1$ roots and so $p(x) \equiv 0$. #

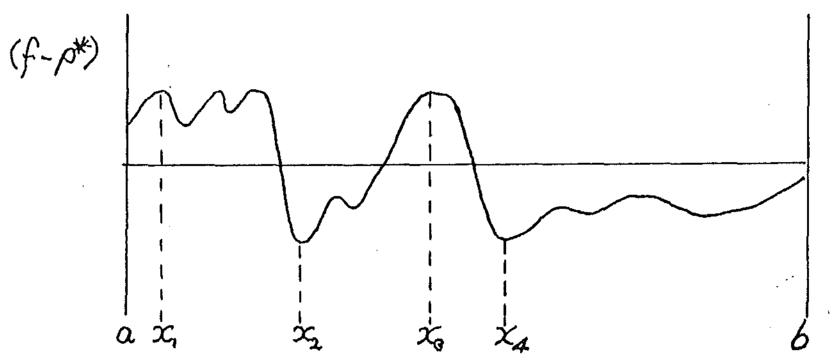
Hence p^* is a best approximation.

$[\Rightarrow]$

(This proof is just an expanded version of that given in Shapiro [3]. The necessary detail tends to obscure the simplicity of the argument, the essential points of which are in italic print, with justification following).

We construct a sequence of points from E *with* x_{i+1} *being the next point from* E *after* x_i *with* $(f-p^*)(x_i) = -(f-p^*)(x_{i+1})$.

eg.



Note that E^+, E^- are compact. Let x_1 be the smallest element of E and suppose without loss of generality that $x_1 \in E^+$. Let x_2 minimise $\{x - x_1 : E^- \cap [x_1,\infty)\}$. ($E^- \cap [x_1,\infty)$ is compact so contains such an x_2) and note that $x_2 > x_1$. Let x_3 minimise $\{x - x_2 : x \in E^+ \cap [x_2,\infty)\}$. We have

$x_3 > x_2$.

Continue in this fashion to obtain a sequence $x_1 < x_2 < \ldots$ such that $(f - p^*)(x)$ attains its maximum modulus on $\{x_i\}$ with alternating sign. The sequence terminates only if $E^{(-1)^j} \cap [x_j, \infty) = \phi$ for some $j \in \mathbb{N}$.

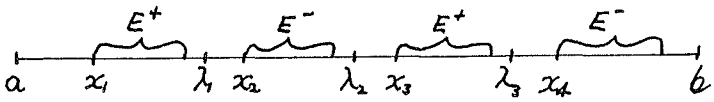
If our set $\{x_i\}$ *has* $n + 2$ *or more members then the proof is complete. So suppose it has only* j *members,* $j < n + 2$.

We can define $\lambda_1, \lambda_2, \ldots, \lambda_{j-1}$ *such that*

(a) $\lambda_i \in (x_i, x_{i+1})$

(b) $E \cap [\lambda_i, x_{i+1}) = \phi$

eg.

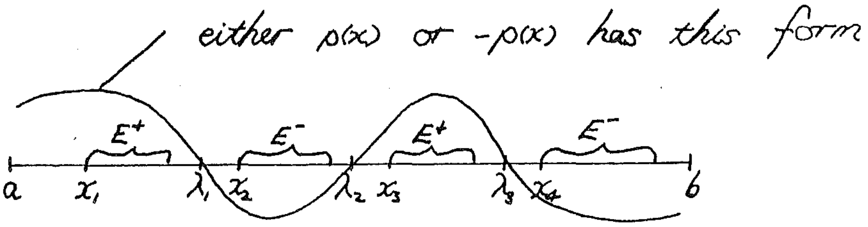


Note that $E^{(-1)^i} \cap (x_i, x_{i+1}) = \phi$ (1)

since otherwise we contradict the minimality definition of x_{i+1}. Let $\mu_i = \max\{x : x \in E^{(-1)^{i+1}} \cap [x_i, x_{i+1}]\}$. Then clearly $\mu_i \geqslant x_i$ and $\mu_i < x_{i+1}$ since $x_{i+1} \in E^{(-1)^i}$. Let $\lambda_i \in (\mu_i, x_{i+1})$. Then $E^{(-1)^{i+1}} \cap [\lambda_i, x_{i+1}) = \phi$ and this, with (1) gives

$$E \cap [\lambda_i, x_{i+1}) = \phi . \tag{2}$$

Form $p(x) = \prod_{i=1}^{j-1} (x - \lambda_i) \in P_n[x]$. *Then either* $p(x)$ *or* $-p(x)$ *satisfies the conditions of Theorem (3.2) and so* p^* *is not a best approximation.* #



$p(x)$ has roots $\{\lambda_1, \lambda_2, \ldots, \lambda_{j-1}\}$ precisely and changes sign at these points. Suppose without loss of generality, that j is odd and that $p(x_1) > 0$. Then $E^+ \cap (x_n, x_{n+1}) = \phi$ from (1) and this with (2) gives $E^+ \cap (\lambda_{2n-1}, \lambda_{2n}) = \phi \quad \forall n \in \{2, 3, \ldots, \frac{j-1}{2}\}$. Since $p(x) > 0$ $\forall x \notin (\lambda_{2n-1}, \lambda_{2n})$ it follows that $p(x)(f(x) - p^*(x)) > 0$, $\forall x \in E^+$.

Similarly $p(x)(f(x) - p^*(x)) < 0$, $\forall x \in E^-$ and so p^* is not a best approximation. #

This completes the proof.

<u>Definition (4.1)</u> Let $f \in C[a,b]$, $p^* \in P_n[x]$ and suppose that $\exists\ x_0, x_1, \ldots, x_{n+1} \in [a,b]$ such that $x_0 < x_1 < \ldots < x_{n+1}$ with

(a) $|f(x_i) - p^*(x_i)| = \|f - p^*\|$, $\forall i \in \{0,1,\ldots, n+1\}$

(b) $f(x_i) - p^*(x_i) = -(f(x_{i+1}) - p^*(x_{i+1}))$.

Then $\{x_0, x_1, \ldots, x_{n+1}\}$ is called an alternating set for $f - p^*$ and p^* is called an alternant of f.

It is important to examine the properties of $P_n[x]$ which are used in this proof. If we require an analogue to the above theorem in some generalisation of polynomial approximation then clearly we will need to preserve these essential properties in some form. The only properties of $P_n[x]$ used are (in the $[\Leftarrow]$ and $[\Rightarrow]$ proofs respectively):

(a) Any $p \in P_n[x]\setminus\{0\}$ has at most n roots

(b) Given $\lambda_1 < \lambda_2 < \ldots < \lambda_j$, $j \leqslant n$, $\exists p \in P_n[x]$ such that $p(x)$ changes sign at each λ_i and $p(x) = 0$ on $\{\lambda_1, \lambda_2, \ldots, \lambda_j\}$ precisely.

If for some finite dimensional subspace P of $C[a,b]$ we have these properties then the above theorem will hold. It will be shown in Section 5 that (b) is, essentially, a consequence of (a) and so this generalised fundamental theorem of algebra will be <u>the essential condition</u> in all that follows.

Given that $p_1^*, p_2^* \in P_n[x]$ are both best approximations to $f \in C[a,b]$ we know that there must exist alternating sets for $f - p_1^*$ and $f - p_2^*$. This is a strong requirement and it is not unreasonable to expect p_1^*, p_2^* to be closely related in some way. In fact they must be equal, as the following theorem shows.

<u>Theorem (4.2)</u> Each $f \in C[a,b]$ has a <u>unique</u> best approximation from $P_n[x]$.

Proof

Let $p_1^*, p_2^* \in P_n[x]$ be best approximations to f. Then $\frac{1}{2}(p_1^* + p_2^*)$ is also a best approximation so $\exists\ x_1, x_2, \ldots, x_{n+2} \in [a,b]$ such that $a \leqslant x_1 < x_2 < \ldots < x_{n+2} \leqslant b$ and

$$f(x_i) - \frac{1}{2}(p_1^* + p_2^*)(x_i) = (-1)^i \varepsilon \qquad \forall i \in \{1, \ldots, n+2\}$$

where $|\varepsilon| = \|f - p_1^*\| = \|f - p_2^*\|$.

i.e. $\frac{1}{2}(f - p_1^*)(x_i) + \frac{1}{2}(f - p_2^*)(x_i) = (-1)^i \varepsilon.$

Since $|(f - p_1^*)(x_i)|, |(f - p_2^*)(x_i)| \leqslant |\varepsilon|$ we have

$$(f - p_1^*)(x_i) = (f - p_2^*)(x_i) = (-1)^i \varepsilon, \qquad \forall i \in \{1, \ldots, n+2\}$$

$\Rightarrow \qquad p(x_i) = p_2(x_i)$

$\Rightarrow \qquad p_1^*(x) = p_2^*(x), \qquad \forall x \in [a,b].$

The above characterisation of p^*, the best approximation to f, is useful in a number of cases, either to actually determine p^* or, given an alternant q of f, to identify q as p^*. Following are two simple examples.

Example (4.1) Find the best uniform approximation to e^x on $[0,1]$ from $P_1[x]$.

By theorems (4.1) and (4.2) it is sufficient to find $p(x) = ax + b$ such that $\exists\ x_1, x_2, x_3 \in [0,1]$, $0 \leqslant x_1 < x_2 < x_3 \leqslant 1$ extreme points of $e(x) = e^x - p(x)$ with $e(x) = -e(x_2) = e(x_3)$. The extreme points of $e(x)$ are either the endpoints or where $e'(x) = 0$. But

$$e'(x) = 0 \Rightarrow e^x - a = 0 \Rightarrow x = \ln a$$

so $\qquad x_1 = 0,\ x_2 = \ln a,\ x_3 = 1.$

Thus $\quad e(0) = -e(x_2) = e(1),$

so $\quad 1 - b = a\ln a + b - a = e - (a+b),$

hence $\quad a = e - 1$ and $b = \frac{1}{2}[(1-e)(\ln(e-1) - 1) + 1]$

i.e. $(e-1)x + \frac{1}{2}[(1-e)(\ln(e-1) - 1) + 1]$ is the best approximation from $P_1[x]$ to e^x on $[0,1]$.

Example (4.2) The monic polynomial of degree n with the least maximum modulus on $[-1,1]$ is $2^{1-n}T_n(x)$, where $T_n(x) = \cos(n \cos^{-1} x)$ is the Tchebycheff polynomial of degree n. (Note that $2^{1-n}\,T_n(x)$ <u>is</u> monic, of degree n).

Let p(x) be monic, of degree n. Then $\exists r(x) \in P_{n-1}[x]$ such that $p(x) = x^n - r(x)$. Clearly p(x) will be the polynomial we seek iff r(x) is the best approximation to x^n from $P_{n-1}[x]$. By theorem (4.1) this is the case iff $p(x) = x^n - r(x)$ has an alternating set (of n + 1 points). But $\left\{x_j = \frac{\cos j\pi}{n} : j = 0,1,\ldots, n\right\}$ is just such a set for $2^{1-n}T_n(x)$ so it follows that $2^{1-n}T_n(x) = p(x)$.

Section 5: Tchebycheff Systems

In the previous section we developed, in theorems (4.1) and (4.2), two powerful properties of best approximations to $f \in C[a,b]$ from the linear space generated by $\{1,x,\ldots, x^n\}$. Our immediate aim is to answer the following question: for which sets of continuous functions $\{u_0,u_1,\ldots, u_n\}$ can we develop similar theorems concerning the best approximations to $f \in C[a,b]$ from the linear space generated by $U = \{u_0,u_1,\ldots, u_n\}$?

<u>Definition (5.1)</u> An element of span U is termed a generalised polynomial (since it is just a linear combination of u_i's).

For reasons outlined after the proof of Theorem (4.1) the next definition is the obvious starting point.

<u>Definition (5.2)</u> Let $U = \{u_0,u_1,\ldots, u_n\} \subset C[a,b]$. Then U is said to satisfy the Haar condition on [a,b] iff each non-trivial generalised polynomial has no more than n zeros on [a,b]. Alternatively, the functions u_i, i = 0,1,..., n are said to form a Tchebycheff system (on [a,b]).

<u>Definition (5.3)</u>

$$D[x_0,x_1,\ldots, x_n] \equiv \begin{vmatrix} u_0(x_0) & u_1(x_0) & \cdots & u_n(x_0) \\ u_0(x_1) & u_1(x_1) & \cdots & u_n(x_1) \\ \vdots & \vdots & & \vdots \\ u_0(x_n) & u_1(x_n) & \cdots & u_n(x_n) \end{vmatrix}.$$

An equivalent formulation of Tchebycheff systems will also be useful.

Lemma (5.1) U forms a Tchebycheff system iff $\forall x_0, x_1, \ldots, x_n \in [a,b]$ such that $a \leqslant x_0 < x_1 < \ldots < x_n \leqslant b$

$$D[x_0, x_1, \ldots, x_n] \neq 0.$$

Proof

Let $A = \begin{bmatrix} u_0(x_0) & u_1(x_0) & \ldots & u_n(x_0) \\ u_0(x_1) & u_1(x_1) & \ldots & u_n(x_1) \\ \vdots & \vdots & & \vdots \\ u_0(x_n) & u_1(x_n) & \ldots & u_n(x_n) \end{bmatrix}$

and let $p(x) = \sum_{i=0}^{n} c_i u_i(x)$ be a non-trivial generalised polynomial.

Then $A\underline{c} = \begin{bmatrix} p(x_0) \\ p(x_1) \\ \vdots \\ p(x_n) \end{bmatrix}$.

So $p(x)$ has $n+1$ zeros iff for some $x_0, x_1, \ldots, x_n \in [a,b]$ $A\underline{c} = \underline{0}$

i.e. iff A is singular

$\Leftrightarrow$ $\det A = 0 = D[x_0, x_1, \ldots, x_n]$.

The result follows.

The above formulation leads easily to a fundamental property of Tchebycheff systems.

Lemma (5.2) Let U form a Tchebycheff system. Then $\forall x_0, x_1, \ldots, x_{n-1} \in [a,b]$ such that $a \leqslant x_0 < x_1 < \ldots < x_{n-1} \leqslant b$ there exists a generalised polynomial $u(x)$ vanishing only at $\{x_0, x_1, \ldots, x_{n-1}\}$ and changing sign at each $x_i \in (a,b)$.

Proof

Put $u(x) = D[x, x_0, x_1, \ldots, x_{n-1}]$. Then clearly $u(x)$ is a generalised polynomial vanishing at $\{x_0, x_1, \ldots, x_{n-1}\}$ and by Lemma (5.1) vanishing only on this set.

It remains only to show the sign changing property. Consider without loss of generality, the case $i = 0$ with $a < x_0$. Let $h \in \mathbb{R}$ such that $a < x_0 - h < x_0 + h < x_1$ and let $\phi(t) = D[x_0 - h + t,\ x_0 + t,\ x_1, \ldots, x_{n-1}]$ (a continuous function of t). By Lemma (5.1) $\phi(t) \neq 0$, $\forall t \in [0,h]$ and so has constant sign on $[0,h]$.

$$\begin{aligned} \text{Now} \quad \phi(0) &= u(x_0 - h) \\ \text{while} \quad \phi(h) &= D[x_0, x_0 + h, x_1, \ldots, x_{n-1}] \\ &= -D[x_0 + h, x_0, x_1, \ldots, x_{n-1}] \\ &= -u(x_0 + h). \end{aligned}$$

Thus u changes sign at x_0.

The point behind our definition of a Tchebycheff system was a desire to develop theorems analogous to Theorem (4.1) and Theorem (4.2) with the linear subspace generated by $U = \{u_0, u_1, \ldots, u_n\}$ (denoted by span U) playing the role of $P_n[x]$. Recalling the proofs of these theorems it should now be apparent that in order to do this we need only the following lemma.

<u>Lemma (5.3)</u> Let U form a Tchebycheff system. Then $\forall x_0, x_1, \ldots, x_k \in (a,b)$ such that $a < x_0 < x_1 < \ldots < x_k < b$ and $k \leqslant n - 1$ there exists a generalised polynomial u vanishing only on $\{x_0, x_1, \ldots, x_k\}$ and changing sign at each x_i.

<u>Remark.</u> Noting that the case $k = n - 1$ has already been proven (i.e. Lemma (5.2)) one would expect this result to follow easily. In fact, the proof is rather long.

<u>Proof of the Lemma</u>

Let $\varepsilon > 0$ such that $a < a + (n-k-1)\varepsilon < x_0$. Then by Lemma (5.2)

$$D[x, a+\varepsilon, a+2\varepsilon, \ldots, a+(n-k-1)\varepsilon, x_0, x_1, \ldots, x_k] = 0$$

only on $S = \{a+\varepsilon, \ldots, x_k\}$ and changes sign at each point in S.

Let $u_\varepsilon(x) = (-1)^t D[x, a+\varepsilon, \ldots, x_k]$ where t is chosen so that $u_\varepsilon(b) > 0$. We can assume that $\|u_\varepsilon(x)\| = 1$ (just multiply $u_\varepsilon(x)$ by a positive constant). Now the sequence defined by putting $\varepsilon = \frac{1}{n}$, $n \in \mathbb{N}$ is a subset of the compact set $\{p \in \text{span } U : \|p\| = 1\}$ so has a non-trivial limit polynomial $u(x)$.

Clearly $\quad u(x) \geqslant 0 \quad x \in [x_k,b] \cup [x_{k-2},x_{k-1}] \cup \ldots$

$\quad u(x) \leqslant 0 \quad x \in [x_{k-1},x_k] \cup [x_{k-3},x_{k-2}] \cup \ldots$

$\quad u(x) = 0 \quad x \in \{x_0,\ldots, x_k\}$.

We now distinguish two cases:

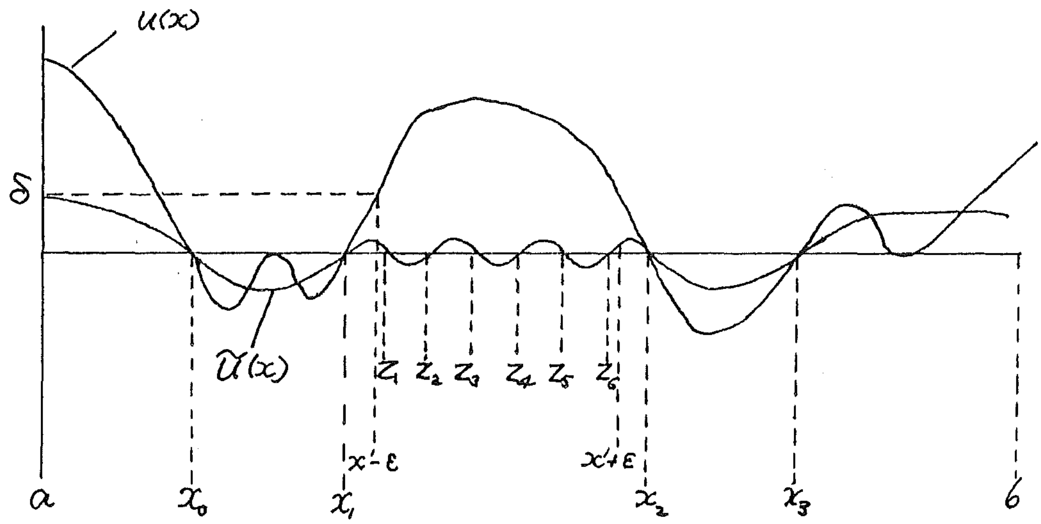
Case 1. $n - k - 1$ even.

Let $x' \in [a,b]$ such that $u(x') \neq 0$. Suppose without loss of generality that $u(x') > 0$. Then $\exists \varepsilon,\delta > 0$ such that $[x'-\varepsilon,x'+\varepsilon] \subset (a,b)$ and $x \in [x'-\varepsilon,x'+\varepsilon] \Rightarrow u(x) > \delta$.

Put $z_i = x' - \varepsilon + \frac{2i\varepsilon}{n-k-1}$, $i = 1,2,\ldots, n-k-1$.

Then by Lemma (5.2) there exists a non-trivial polynomial $\tilde{u}(x)$ changing sign at $\{x_0,\ldots, x_k, z_1,\ldots, z_{n-k-1}\}$ and agreeing in sign with $u(x)$ on $[a,b]\setminus[x'-\varepsilon,x'+\varepsilon]$. Now let $\lambda \in \mathbb{R}$ such that $0 < \lambda < \frac{\delta}{\|\tilde{u}(x)\|}$.

Then $(u + \lambda\tilde{u})(x)$ vanishes only on $\{x_0,\ldots, x_k\}$ and changes sign at each of these points.



Case 2. $n - k - 1$ odd.

Define $x',\varepsilon,\delta,z_1,\ldots, z_{n-k-2}$ as before. Then by Lemma (5.2) there exists a non-trivial polynomial $\tilde{u}(x)$ changing sign on $\{x_0,\ldots, x_k,z_1,\ldots, z_{n-k-2}\}$ and vanishing also at b. Similarly there exists $\tilde{u}(x)$ changing sign at $x_0,\ldots, x_k,z_1,\ldots, z_{n-k-2}$ and vanishing also at a.

It follows that $\exists \lambda, \mu \in \mathbb{R}$ such that $(u + \lambda\tilde{v} + \mu\tilde{u})$ vanishes only on $\{x_0, x_1, \ldots, x_k\}$ and changes sign at each of these points.

Theorem (5.1) Let U form a Tchebycheff system on $[a,b]$ and let $f \in \mathbb{C}[a,b]$. Then

(i) p^* is a best approximation to f from span U iff $f - p^*$ has an alternating set.

(ii) The best approximation to f is unique.

Proof.

These proofs are obvious after consideration of the proofs of theorems (4.1) and (4.2). Note in particular that the $[\Rightarrow]$ direction of (i) requires Lemma (5.3).

The proof of the above theorem was essentially, the goal of this report. An attempt has been made to justify the necessary steps as they were made rather than from hindsight. The surprising fact is that we have made just the right generalisations and definitions in the sense that U must form a Tchebycheff system if each $f \in C[a,b]$ is to have a unique best approximation from span U.

Theorem (5.2) (Haar Unicity Theorem). ([1],p81)
Let $U = \{u_0, u_1, \ldots, u_n\} \subset C[a,b]$. If $\forall f \in C[a,b]$ f has a unique best approximation from span U then U forms a Tchebycheff system.

Proof.

Suppose that U does not form a Tchebycheff system. Then by Lemma (5.1) there exist $x_0, x_1, \ldots, x_n \in [a,b]$ such that $a \leqslant x_0 < x_1 < \ldots < x_n \leqslant b$ and the matrix $[u_j(x_i)]$ is singular. Let $(a_0, \ldots, a_n)$, $(b_0, \ldots, b_n)$ be non-trivial vectors orthogonal to the row and column spaces of $[u_j(x_i)]$ respectively. Then $\sum_{i=0}^{n} a_i u_i(x_j) = 0 = \sum_{i=0}^{n} b_i u_j(x_i)$, $j = 0,1,\ldots, n_n$.
Note that for any $p(x) = \sum_{i=0}^{n} c_i u_i(x)$ we have $\sum_{i=0}^{n} b_j p(x_j) = \sum_{j=0}^{n} b_j \sum_{i=0}^{n} c_i u_i(x_j) \equiv 0$

(1)

Put $q(x) = \sum_{i=0}^{n} a_i u_i(x) \in \text{span } U$. Then $q(x)$ is non-trivial with $q(x_j) = 0$ $j = 0,1,\ldots, n$ and without loss of generality we may suppose that $\|q\| < 1$. Now construct $f \in C[a,b]$ such that $\|f\| = 1$ and $f(x_j) = \text{sgn} b_j$, $j = 0,1,\ldots, n$ (such a function clearly exists). Let $h(x) = f(x)(1 - |q(x)|)$ so that $h(x_j) = f(x_j) = \text{sgn} b_j$, $j = 0,1,\ldots, n$. We claim that for any $p \in \text{span } U$ we have $\|h - p\| \geqslant 1$. If not then $\|h - p\| < 1$ so $\text{sgn} p(x_j) = \text{sgn} f(x_j) = \text{sgn} b_j$, $j = 0,1,\ldots, n$. But then $\sum_{j=0}^{n} b_j p(x_j) > 0$. # (see (1))

Now $\forall \lambda \in [0,1]$, $\forall x \in [a,b]$,

$$\begin{aligned} |h(x) - \lambda q(x)| &\leqslant |h(x)| + \lambda |q(x)| \\ &= |f(x)|(1 - |q(x)|) + \lambda |q(x)| \\ &\leqslant 1 \end{aligned}$$

so $\lambda q(x)$ is a best approximation to $h(x)$ $\forall \lambda \in [0,1]$ thus there is no unique best approximation to f.

Section 6: An alternative approach

We present here a considerably more elegant route to Theorem (5.1). The disadvantage of this approach is that there is a lack of intuitive continuity from one step to the next. Because of this the argument is presented simply as a series of theorems with little attempt to justify their general direction. It is suggested that one first read the statements of all the theorems in this section before considering the individual proofs. (This treatment is exactly that given by Cheney [1] and is included solely for completeness).

Definition (6.1) A finite convex combination of elements from a set A is a linear combination of the form $\sum_{i=0}^{n} \theta_i a_i$ where $n \in \mathbb{N}$, $\sum_{i=0}^{n} \theta_i = 1$ and $\theta_i \geqslant 0$, $a_i \in A$ $i = 0,\ldots, n$.

Definition (6.2) A set A is convex iff each finite convex combination of elements from A is itself in A.

Definition (6.3) Given any set A we define $\mathcal{H}(A)$, the convex hull of A to be the set of all finite convex combinations of elements from A.

Remark. $\mathcal{H}(A)$ is the smallest convex set containing A.

Theorem (6.1) (Carathéodory) ([1],p17).

Let A be a subset of an n dimensional linear space. Then each point of $\mathcal{H}(A)$ is expressible as a convex combination of n + 1 or fewer elements of A.

Proof.

Let $a \in \mathcal{H}(A)$. Then $a = \sum_{i=0}^{k} \theta_i a_i$ for some k, θ_i, a_i such that $k \in \mathbb{N}$, $\theta_i \geq 0$, $a_i \in A$ $i = 0,1,\ldots, k$ and $\sum_{i=0}^{k} \theta_i = 1$.

Assume that k is minimal (i.e. there is no convex combination of fewer elements from A that is equal to a). Then $\theta_i > 0$ $i = 0,\ldots, k$. Let $y_i = a_i - a$. Note that $\sum_{i=0}^{k} \theta_i y_i = 0$.

Suppose that $k > n$. Then $\{y_1,\ldots, y_k\}$ has more than n elements so is a linearly dependent set. Thus $\exists \alpha_i \in \mathbb{R}$ such that $\sum_{i=1}^{k} |\alpha_i| \neq 0$ and $\sum_{i=1}^{k} \alpha_i y_i = 0$. If we define $\alpha_0 = 0$ then $\forall \lambda \in \mathbb{R}$, $\sum_{i=0}^{k} (\theta_i + \lambda\alpha_i) y_i = 0$. Choose λ so that $|\lambda|$ is as small as possible under the condition that $\theta_j + \lambda\alpha_j = 0$ for some $j \in \{0,\ldots, k\}$. The remaining coefficients are all non-negative and do not all vanish since $\theta_0 + \lambda\alpha_0 = \theta_0 > 0$. But now

$$\sum_{\substack{i=0\\ i\neq j}}^{k} (\theta_i + \lambda\alpha_i)(a - a_i) = 0$$

$$\Rightarrow a = \Big\{\sum_{\substack{i=0\\ i\neq j}}^{k} (\theta_i + \lambda\alpha_i)\Big\}^{-1} \sum_{\substack{i=0\\ i\neq j}}^{k} (\theta_i + \lambda\alpha_i) a_i$$

- a finite convex combination contradicting the minimality of k.

Hence we have the result.

Corollary (6.1) ([1],p18).

The convex hull of a compact subset, X, of $\mathbb{R}_n$ is compact.

Proof. (For a more standard proof see [1]).

Let $S = \{(\theta_0, \ldots, \theta_n) : \theta_i \geq 0, \sum_{i=0}^{n} \theta_i = 1\}$. S is a compact subset of $\mathbf{R}_{n+1}$. Then by Theorem (6.1) $\mathcal{H}(X)$ is a continuous image of $S \times \underbrace{X \times \ldots \times X}_{n+1}$ (compact) hence compact.

Theorem (6.2) ([1],p19)

Every closed convex subset, K, of $(\mathbf{R}_n, \| \ \|_2)$ contains a point of minimum norm.

Proof.

Let $(x_n) \in K$ such that $\lim_{n\to\infty} \|x_n\| = d = \inf_{x \in K} \|x\|$

Then $\|x_i - x_j\|^2 = 2\|x_i\|^2 + 2\|x_j\|^2 - 4\|\frac{1}{2}(x_i + x_j)\|^2$ and since K is convex $\frac{1}{2}(x_i + x_j) \in K$ so $\|\frac{1}{2}(x_i + x_j)\| \geq d$.

Thus $\quad \|x_i - x_j\|^2 \leq 2\|x_i\|^2 + 2\|x_j\|^2 - 4d$

$$\to 0 \quad \text{as} \quad i,j \to \infty$$

hence $(x_n) \to x \in K$ and $\|x\| = d$.

Theorem (6.3) ([1],p19).

Let U be a compact subset of $(\mathbf{R}_n, \| \ \|_2)$. Then $\nexists z \in \mathbf{R}_n$ such that $\langle u,z\rangle > 0$, $\forall u \in U$ iff $0 \in \mathcal{H}(U)$.

Proof.

[⇒]

Suppose that $0 \notin \mathcal{H}(U)$. Then by Corollary (6.1) $\mathcal{H}(U)$ is compact and so by Theorem (6.2) $\exists z \in \mathcal{H}(U)$ of minimum norm. Let $u \in U$.

Then $\theta u + (1-\theta)z \in \mathcal{H}(U)$, $\forall \theta \in \mathbf{R}$ such that $0 \leq \theta \leq 1$. Hence

$$0 \leq \|\theta u + (1-\theta)z\|^2 - \|z\|^2 = \theta^2\|u - z\|^2 + 2\theta < u - z, z > .$$

Considering small positive θ this implies that $\langle u,z\rangle \geq \|z\|^2 > 0$ so z is a solution of the inequality.

[⇐]

Suppose that $0 \in \mathcal{H}(U)$. Then $\exists \lambda_i \geq 0$, $u_i \in U$ such that $0 = \sum_{i=0}^{n} \lambda_i u_i$.

So $\forall z \in R_n$, $0 = \sum_{\ell=0}^{0} \lambda_i < u_i, z >$. This equation would be violated if $< u_i, z > > 0$, $i = 0,\ldots,n$.

Theorem (6.4) ([1],p74).

Let $\{u_0,\ldots,u_n\}$ form a Tchebycheff system, $a \leqslant x_0 < x_1 < \ldots < x_{n+1} \leqslant b$ and $\lambda_0,\lambda_1,\ldots,\lambda_{n+1} \in R\setminus\{0\}$. Let $\hat{x} = [u_0(x),u_1(x),\ldots,u_n(x)]$ (so $\hat{x} \in R_{n+1}$). Then $0 \in \mathcal{H}(\{\lambda_0\hat{x}_0,\ldots,\lambda_{n+1}\hat{x}_{n+1}\})$ iff $\lambda_i\lambda_{i+1} < 0$, $i = 0,1,\ldots,n+1$.

Proof.

$D[x_0\ldots,x_n]$ has constant sign independent of the points $a \leqslant x_0 < \ldots < x_n \leqslant 0$.

If not we have $a \leqslant x_0 < \ldots < x_n \leqslant b$, $a \leqslant y_0 < \ldots < y_n \leqslant b$ such that $D[x_0,\ldots,x_n] < 0 < D[y_0,\ldots,y_n]$. Since D is continuous in its $n+1$ variables $\exists\lambda \in (0,1)$ such that $D[\lambda x_0 + (1-\lambda)y_0,\ldots, \lambda x_n + (1-\lambda)y_n] = 0$. By Lemma (5.1) it follows that

$\lambda x_i + (1-\lambda)y_i = \lambda x_j + (1-\lambda)y_i$ for some $i,j,i \neq j$. This is impossible since $\{x_i\},\{y_i\}$ are strictly increasing.

$0 \in \mathcal{H}(\{\lambda_0\hat{x}_0,\ldots,\lambda_{n+1}\hat{x}_{n+1}\})$ iff $\exists\theta_0,\ldots,\theta_{n+1} > 0$ such that $\sum_{i=0}^{n+1} \theta_i\lambda_i\hat{x}_i = 0$.

[⇐]

If there exist such θ_i then dividing by $\sum_{i=0}^{n+1} \theta_i$ gives $0 \in \mathcal{H}(\{\lambda_0\hat{x}_0,\ldots,\lambda_{n+1}\hat{x}_{n+1}\})$.

[⇒]

If $0 \in \mathcal{H}(\{\lambda_0\hat{x}_0,\ldots,\lambda_{n+1}\hat{x}_{n+1}\})$ then $\exists\theta_0,\ldots,\theta_{n+1} \geqslant 0$ with $0 = \sum_{i=0}^{n+1} \theta_i\lambda_i\hat{x}_i$. Suppose without loss of generality that $\theta_0 = 0$. Then $\{\hat{x}_1,\ldots,\hat{x}_{n+1}\}$ is linearly dependent so $D[x_1,\ldots,x_{n+1}] = 0$. #

Hence $0 \in \mathcal{H}(\{\lambda_0\hat{x}_0,\ldots,\lambda_{n+1}\hat{x}_{n+1}\})$ iff $\exists\theta_0,\ldots,\theta_{n+1} > 0$ such that $\hat{x}_0 = \sum_{i=0}^{n+1} - \left(\frac{\theta_i\lambda_i}{\theta_0\lambda_0}\right)\hat{x}_i$, i.e. (by Cramer's rule)

$$-\frac{\theta_i\lambda_i}{\theta_0\lambda_0} = \frac{D[x_1,\ldots,x_{i-1},x_0,x_{i+1},\ldots,x_{n+1}]}{D[x_1,\ldots,x_{n+1}]} .$$

Since it requires $i-1$ interchanges in the numerator determinant to restore the natural order of the arguments it follows that the required θ_i's exist iff the λ_i's alternate in sign.

We are now ready to reprove Theorem (5.1)(i).

Theorem (5.1)(i).

Let $\{u_0,\ldots,u_n\}$ form a Tchebycheff system on $[a,b]$ and let $f \in C[a,b]$. Then p^* is a best approximation to f from span $(\{u_0,\ldots,u_n\})$ iff $f-p^*$ has an alternating set.

Proof.

$[\Leftarrow]$

Obvious.

$[\Rightarrow]$

The case $f \in \text{span}(\{u_0,\ldots,u_n\})$ is obvious so suppose $\|f-p^*\| > 0$.

By Theorem (3.2) $\exists p \in \text{span}(\{u_0,\ldots,u_n\})$ such that

$$(f-p^*)(x)p(x) > 0 \quad \forall x \in E.$$

Writing $p(x) = \sum_{i=0}^{n} d_i u_i(x) = \langle d,\hat{x}\rangle$ we have $\exists d \in R_{n+1}$ such that $\langle d,(f-p^*)(x)\hat{x}\rangle > 0$, $\forall x \in E$. Since $\{(f-p^*)(x)\hat{x} : x \in E\}$ is compact we can apply Theorem (6.3) to give

$$0 \in \mathcal{H}(\{(f-p^*)(x)\hat{x} : x \in E\}) .$$

By Theorem (6.1) $0 = \sum_{i=0}^{k} \lambda_i (f-p^*)(x_i)\hat{x}_i$, $k \leqslant n$, $\lambda_i > 0$ $i = 0,\ldots,k$ and $x_i \in E$, $x_i \neq x_j$ $i \neq j$. Since $\{u_0,\ldots,u_n\}$ forms a Tchebycheff system $\{\hat{x}_i : i = 0,1,\ldots,k\}$ is linearly independent for $k < n$ (by Lemma (5.1)) so it follows that $k = n$. If we now arrange the x_i in increasing order then by Theorem (6.4) the $\lambda_i(f-p^*)(x_i)$ alternate in sign hence so do the $(f-p^*)(x_i)$.

Section 7: Examples

We now present some basic examples and applications of the theorems of Section 5.

Example 7.1. $\{1,x,\ldots x^n\}$ forms a Tchebycheff system on any $[a,b] \subseteq \mathbf{R}$.

Example 7.2. $\{1,x^2,x^4\}$ forms a Tchebycheff system on $[0,1]$ but not on $[-1,1]$.

$[D[x_1,x_2,x_3] = (x_2^2 - x_1^2)(x_3^2 - x_1^2)(x_3^2 - x_2^2)$

(by analogy with the Vandermonde determinant) so

$D[x_1,x_2,x_3] \neq 0 \quad \forall x_1,x_2,x_3 \in [0,1]$

but, for example $D[-1,0,1] = 0]$.

Example 7.3. $\{1,\cos\theta,\cos 2\theta,\ldots,\cos n\theta,\sin\theta,\ldots,\sin n\theta\}$ forms a Tchebycheff system on any $[a,b] \subseteq [0,2\pi)$.

[Any generalised polynomial $\sum_{k=0}^{n} (a_k \cos k\theta + b_k \sin k\theta)$ can be written in form

$$\sum_{k=-n}^{n} \alpha_k e^{ik\theta} = e^{-in\theta} \sum_{k=0}^{2n} \alpha_{k-n} e^{ik\theta}$$

Let $z = e^{i\theta}$ and (by the fundamental theorem of algebra) $\sum_{k=0}^{2n} c_k z^k$ has at most $2n$ roots. Since $[a,b] \subseteq [0,2\pi]$ it follows that $\forall z \in \mathbf{C}$, $z = e^{i\theta}$ has at most one solution $\theta \in [a,b]$ and so $\sum_{k=0}^{n} (a_k \cos k\theta + b_k \sin k\theta)$ has at most $2n$ roots].

Example 7.4. Suppose $\{u_0,u_1,\ldots,u_n\}$ forms a Tchebycheff system and let $r \in C[a,b]$ such that $\forall x \in [a,b]$, $r(x) \neq 0$. Then $\{ru_0,ru_1,\ldots,ru_n\}$ also forms a Tchebycheff system on $[a,b]$.

[Since $\sum_{i=0}^{n} a_i(ru_i)(x) = 0 \iff \sum_{i=0}^{n} a_i u_i(x) = 0]$.

Example 7.5. ([2],pp287-288). This example is sketched only. The analytic details are left to the reader.

We wish to find p^* the best approximation to $\frac{1}{x-c}$ $(c > 1)$ on $[-1,1]$ from $P_n[x]$.

Consider the function

$$\psi(x) = \frac{K}{2}\left[z^n \frac{\beta - z}{1-\beta z} + z^{-n}\frac{1-\beta z}{\beta - z}\right]$$

where $x = \frac{1}{2}(z + z^{-1})$, $c = \frac{1}{2}(\beta + \beta^{-1})$, $K = 4\beta^{n+2}(1-\beta^2)^{-2}$.

We claim that:

(i) $\exists p^* \in P_n[x]$ such that $\psi(x) = \frac{1}{x-c} - p^*(x)$

(ii) $\psi(x)$ has an alternating set of $n+2$ points.

Firstly, since ψ is a rational function of z, $\psi(x) = \frac{p(z)}{q(z)}$ for some polynomials p,q. Also $\psi(x(z)) = \psi(x(z^{-1}))$ so

$$\psi(x) = \frac{p(z)}{q(z)} = \frac{p(z^{-1})}{q(z^{-1})} = \frac{p(z) + p(z^{-1})}{q(z) + q(z^{-1})} .$$

and thus $\psi(x) = \frac{p_1(z+z^{-1})}{q_1(z+z^{-1})}$ for some polynomials p_1, q_1. (this is clear after consideration of the expansion of $(z + z^{-1})^i$). Hence $\psi(x) = \frac{p_1(2x)}{q_1(2x)}$.

Clearly $\psi(x(z))$ is discontinuous only at $z = \{\beta, \frac{1}{\beta}\}$ and in both cases $x = c$.

Since $\lim\limits_{x\to c} (x-c)\psi(x) = \frac{K(1-\beta^2)^2}{4\beta^{n+2}} = 1$, $\psi(x)$ must be of the form

$$\psi(x) = \frac{1}{x-c} - p^*(x)$$

where $p^*(x) \in P_n[x]$.

Now if we put $z = e^{i\theta}$ then $x = \frac{1}{2}(z + \frac{1}{z}) = \frac{e^{i\theta} + e^{-i\theta}}{2} = \cos\theta$ so that as x moves along the interval $[-1,1]$ z traces the curve $\{e^{i\theta} : 0 \leqslant \theta \leqslant \pi\}$.

Note that

$$\arg\left\{z^n \frac{\beta - z}{1-\beta z}\right\} = \begin{cases} 0 \\ \pi \end{cases} \Rightarrow \psi(x) = \begin{cases} K \\ -K \end{cases}$$

and that $|z| = 1 \Rightarrow \left|z^n \frac{\beta - z}{1-\beta z}\right| = 1$

$$\Rightarrow \psi(x) \leqslant K .$$

It follows that as θ moves from 0 to π, $\arg\{z^n \frac{\beta - z}{1-\beta z}\}$ increases from π to $(n+2)\pi$ so $\psi(x)$ has an alternating set and $p^*(x)$ is the best approximation to $\frac{1}{x-c}$. Since $\beta = c - \sqrt{c^2 - 1}$,

$\| \frac{1}{x-c} - p^*(x) \| = K = (c - \sqrt{c^2 - 1})^n / (c^2 - 1)$.

As a trivial extension note that

$$\frac{x^{k+1}}{x-\lambda} = \frac{x^{k+1} - \lambda^{k+1}}{x-\lambda} + \frac{\lambda^{k+1}}{x-\lambda} = x^k + \lambda x^{k-1} + \ldots + \lambda^k + \frac{\lambda^{k+1}}{x-\lambda}, \; k \leqslant n$$

so the best approximation to $\frac{x^{k+1}}{x-\lambda}$ on $[-1,1]$ from $P_n[x]$ is

$$x^k + \lambda x^{k-1} + \ldots + \lambda^k + \lambda^{k+1} p^*(x) .$$

Section 8: Uniform Approximation without the Haar Condition

If the functions $\{u_0, \ldots, u_n\}$ do not form a Tchebycheff system then the question remains, "For which functions $f \in C[a,b]$ does there exist a unique best approximation?". Theorem (5.2) tells us that the set F of such functions is certainly a proper subset of $C[a,b]$ and the following theorem shows us that $F \neq \phi$ in general. Given $\{u_0, \ldots, u_n\}$ can we characterise F in any way? The answer to this question does not appear obvious, even in simple cases.

Theorem (8.1)

Let p^* be a best approximation to $f \in C[a,b]$ from span $\{u_0, \ldots, u_n\}$ on $[a,b]$. If $\exists c,d \in R$ such that $E(f) \subset [c,d] \subset [a,b]$ and $\{u_0, \ldots, u_n\}$ forms a Tchebycheff system on $[c,d]$ then p^* is the unique best approximation to f on $[a,b]$ (and $[c,d]$) and $f-p^*$ has an alternating set (in $[c,d]$).

Proof.

It is clearly sufficient to show that p^* is a best approximation to f on $[c,d]$. If not then by Theorem (3.2) $\exists p \in$ span U such that

$$\begin{cases} p(x) > 0 & x \in E^+ \\ p(x) < 0 & x \in E^- \end{cases}$$

(Note $E(f) \subset [c,d]$)

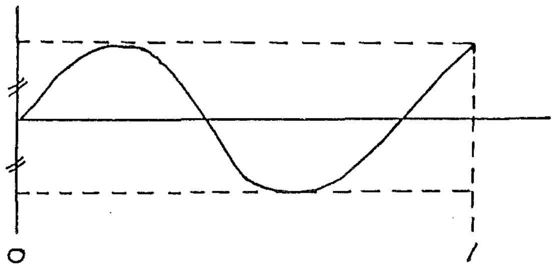
and again by Theorem (3.2) p^* is not a best approximation on $[a,b]$. #

This theorem is demonstrated by the following example.

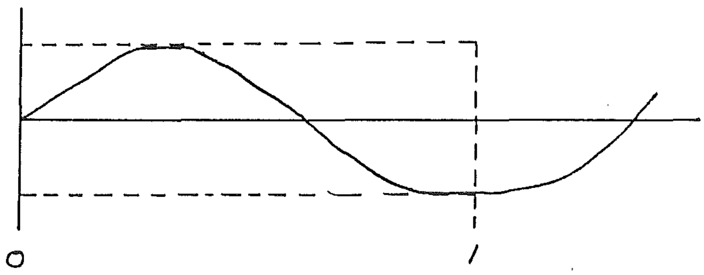
Example (8.1) Any $f \in C[0,1]$ with $f(0) = 0$ has a unique best approximation p^* from $\text{span}\{x, x^2\}$ on $[0,1]$ and $f - p^*$ has an alternating set.

For instance if $f(x) = x^3$ we can construct our unique best approximation as follows.

We require $a, b \in \mathbb{R}$ such that $x^3 - ax^2 - bx$ has the following form.



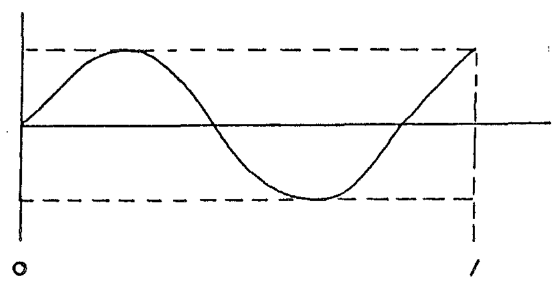
Now the Tchebycheff polynomial of degree three is $x(4x^2 - 3)$ so $(x - \frac{\sqrt{3}}{2})[4(x - \frac{\sqrt{3}}{2})^2 - 3]$ looks like



and thus $\exists! c \in (\sqrt{3}, \infty)$ such that

$$(c - \frac{\sqrt{3}}{2})[4(c - \frac{\sqrt{3}}{2})^2 - 3] = 1$$

with $(cx - \frac{\sqrt{3}}{2})[4(cx - \frac{\sqrt{3}}{2})^2 - 3]$ having the form:



Hence

$x^3 - \frac{1}{4c^3}\{(cx - \frac{\sqrt{3}}{2})[4(cx - \frac{\sqrt{3}}{2})^2 - 3]\}$ is the required best approximation.

Section 9: Conclusion

Originally, I started out by considering the problem outlined in Section (8) (that is - "What is the situation if U is not a Tchebycheff system?") and quickly decided that this question was difficult, if not impossible, to answer. It seemed that as a first step I should familiarise myself with an approach to Theorem (5.1) which lent itself to consideration of this more general problem. The difficulty that I encountered in finding such a treatment in any of the standard approximation theory texts led to the idea of this report.

I hope that at this stage the reader has a good intuitive understanding of the principles and problems of best uniform approximation from finite-dimensional subspaces of C[a,b]. If this is so then my aim has been achieved. As to the original problem, it remains unanswered but Theorem (8.1) is, at least, a step in the right direction.

Section 10: Acknowledgements, References

I would like to thank Dr A.W. McInnes for his guidance and for posing the original problem. Both he and Dr R.K. Beatson suffered numerous discussions of the topic. Much of the material in this report is from standard texts and from Dr Beatson's lectures of paper 480. In particular, the [⇒] proof of Theorem (4.1) follows closely that of Shapiro [3], Theorem (5.2) and Section 6 are drawn from Cheney [1] and Example (7.5) is from Karlin and Studden [2].

References

[1] Cheney, E.W. Introduction to Approximation Theory, McGraw-Hill, New York 1966.

[2] Karlin, S. and Studden, W.J. Tchebycheff Systems with Applications in Analysis and Statistics, John Wiley and Sons, New York 1966.

[3] Shapiro, Harold S. Topics in Approximation Theory, Springer-Verlag, Heidelberg 1971.