JOURNAL OF APPROXIMATION THEORY 3, 243–260 (1970)

Approximation of Norms

JAAK PEETRE

Department of Mathematics, Lund Institute of Technology, Lund, Sweden

Communicated by E. W. Cheney

Received June 2, 1969

INTRODUCTION

In many questions of analysis we encounter the problem of approximating a given norm by "simpler" ones. For example, if

$$\| a \| = \sup_{0 \leqslant x \leqslant 1} | a(x) |,$$

a suitable candidate for an approximate norm is

$$\| a \|_m = \max_{0 \leqslant k \leqslant m-1} \left| a\left(\frac{k}{m}\right) \right|,$$

with m large. More generally,

$$\| a \| = \left(\int_0^1 | a(x) |^p \, dx \right)^{1/p} \qquad (p \geqslant 1)$$

can be approximated by

$$\| a \|_m = \left(\sum_{k=0}^{m-1} \left| a\left(\frac{k}{m}\right) \right|^p \frac{1}{m} \right)^{1/p}.$$

These are examples of so-called *discrete* approximations. In other problems, one would like to approximate

$$\| a \| = \sup | a(x) |$$

by

$$\| a \|_m = \left(\int_0^1 | a(x) |^{p_m} \, dx \right)^{1/p_m},$$

where $p_m \to \infty$.

We believe that it is worth while to put the above on a more formal basis. The present paper is a first, modest attempt in this direction. In particular, our general point of view allows us to treat, in a more systematic fashion, the convergence of a number of algorithms in approximation theory (discretization of continuous Tschebycheff-approximation, Pólya algorithm etc.) We also give a new look at the little (or Hausdorff) moment problem. These general ideas are also very useful in connection with the theory of *interpolation spaces* but this we shall treat elsewhere.

A notion of convergence of normed or, more generally, metric linear spaces has been studied in a paper by Semadeni [7] but there seems to be hardly any connection with the present work. (More close to our viewpoint comes a paper by Kripke [3].)

The plan of the present paper is as follows. In Section 1, the general definitions are given. In Section 2, we briefly review some known facts in linear approximation theory. In Section 3, we give a general theorem on the convergence of algorithms, along the lines of a theorem by Cheney [1] dealing with the concrete case of approximation by algebraic polynomials. This case is studied here in Section 4. Finally, in Section 5, we use our ideas in connection with the moment problem, mentioned above.

1. Basic Definitions

Let A be any vector space over $\mathbf{R}$.

Definition 1.1. By a norm $\| \; \|$, we mean a mapping $A \to \mathbf{R}^+ : a \to \| a \|$ such that

$$\| a + b \| \leqslant \| a \| + \| b \|,$$
$$\| \lambda a \| = | \lambda | \, \| a \|.$$

Consider the linear subspace $N = \{a \mid \| a \| = 0\}$. If $N = 0$, we speak of a *proper* norm. If N is of finite codimension, we speak of a *discrete* norm.

We have, thus, departed slightly from the usual terminology.

Our terminology	Usual terminology
Norm	Seminorm
Proper norm	Norm

Definition 1.2. By a *normed space* we mean a vector space A with a fixed (usually proper) norm $\| \; \| = \| \; \|_A$.

Next, we consider sequences of norms $\| \;\|_m$ on A (where, usually, $m = 1, 2, \ldots$).

DEFINITION 1.3. $\| \;\|_m$ is an *approximation* of $\| \;\|$ if

$$\lim_{m\to\infty} \| a \|_m = \| a \| \qquad (\forall a \in A).$$

If all norms $\| \;\|_m$ are discrete, we speak of a discrete approximation.

DEFINITION 1.4. $\| \;\|_m$ is a null-sequence if

$$\lim_{m\to\infty} \| a \|_m = 0 \qquad (\forall a \in A).$$

The following result is obvious.

PROPOSITION 1.1. *$\| \;\|_m$ is an approximation of $\| \;\|$ if there exist null-sequences $\| \;\|_m^*$ and $\| \;\|_m^{**}$ such that*

$$\| a \| \leqslant \| a \|_m + \| a \|_m^* , \tag{1.1}$$

$$\| a \|_m \leqslant \| a \| + \| a \|_m^{**}. \tag{1.2}$$

DEFINITION 1.5. $\| \;\|_m^*$ is called a *majorant* of the approximation and $\| \;\|_m^{**}$ a *minorant.*

The following is a general way of constructing approximations: Let A_m be a sequence of normed spaces. Let there be given, for each m, linear mappings $Q_m : A \to A_m$ and $P_m : A_m \to A$ such that

$$U_m \to I \qquad \text{(pointwise convergence: } \| U_m a - a \| \to 0)$$

where we have set $U_m = P_m Q_m$. Then we may take

$$\| a \|_m = \| Q_m a \|_{A_m} .$$

If

$$\| P_m a \|_A = \| a \|_{A_m} ,$$

$$\lim_{m\to\infty} \| Q_m a \|_{A_m} \leqslant \| a \|,$$

then clearly (cf. Theorem 5.1)

$$\| a \|_m \to \| a \| = \| a \|_A .$$

If U_m is of finite rank then this is a discrete approximation. Question: Does any (separable) normed space admit a discretization of this type? The answer is trivially positive if the space posseses a Schauder basis.

2. The Fundamental Problem of Linear Approximation Theory

Let B be a given subspace of a normed space A. Let us set

$$E(a) = E(a, A, B) = \inf_{b \in B} \| a - b \|.$$

Clearly

$$E(a) \leqslant \| a - b \| \qquad (\forall b \in B)$$

and in particular (take $b = 0$),

$$E(a) \leqslant \| a \|.$$

The fundamental problem of linear approximation theory consists of finding $b \in B$ such that

$$E(a) = \| a - b \|.$$

We say that b is a solution. It is a classical fact that a solution always *exists* in the following two cases:

(a) B is finite dimensional (see [1], p. 20).

(b) A is uniformly convex, B is complete (see [1], p. 22).

(For more recent results in this direction, see also Cheney and Wulbert [2].) Concerning *uniqueness*, we list two typical cases where it holds:

(α) A strictly convex, B finite dimensional (see [1], p. 23).

(β) $\| a \|$ a Tschebycheff-norm (maximum-norm), B a Haar subspace (see [1], p. 80).

(For more recent results, see, e.g., Phelps [6], Singer [8].) If uniqueness holds, we denote the unique solution by Ta (Tschebycheff-operator). Clearly,

$$E(a) = \| a - Ta \| \leqslant \| a - b \| \qquad (\forall b \in B).$$

We say that we have *strong uniqueness* if

$$\gamma \| b - Ta \| + \| a - Ta \| \leqslant \| a - b \| \qquad (\forall b \in B)$$

where $\gamma > 0$ depends on a only. Strong uniqueness is known to hold in case

(β) above (see [1], p. 80). In case (α), strong uniqueness does not hold. However, we have the following substitute

$$E(a)\,\delta_1^{-1}\left(\frac{\| b - Ta \|}{E(a)}\right) + \| a - Ta \| \leqslant \| a - b \|,$$

where

$$\delta_1(\epsilon) = (1 + \epsilon)\,\delta\left(\frac{\epsilon}{1 + \epsilon}\right),$$

$\delta(\epsilon)$ denoting the modulus of convexity.

DEFINITION 2.1. By an *algorithm* for T we mean a sequence of Tschebycheff-operators T_m corresponding to an approximation $\| \cdot \|_m$ of $\| \cdot \|$.

3. CONVERGENCE OF ALGORITHMS (GENERAL CASE)

We consider the following situation: T_m is an algorithm for the Tschebycheff-operator T. We assume (strong uniqueness) that

$$\phi(\| b - Ta \|) + \| a - Ta \| \leqslant \| a - b \| \qquad (\forall b \in B), \tag{3.1}$$

$$\phi_m(\| b - T_m a \|_m) + \| a - T_m a \|_m \leqslant \| a - b \|_m \qquad (\forall b \in B), \tag{3.2}$$

where ϕ and ϕ_m are positive functions depending on a. We also assume that there are given a majorant $\| \;\|_m^*$ for $\| \;\|_m$ such that

$$\| b \|_m^* \leqslant N_m^* \| b \| \qquad (\forall b \in B) \tag{3.3}$$

and a minorant $\| \;\|_m^{**}$ such that

$$\| b \|_m^{**} \leqslant N_m^{**} \| b \| \qquad (\forall b \in B) \tag{3.4}$$

for some constants $N_m{}^*$ and N_m^{**}. If B is finite dimensional, the existence of such constants is automatically guaranteed. Also $N_m{}^* \to 0$ and $N_m^{**} \to 0$ as $m \to \infty$. First we prove:

LEMMA 3.1. If (3.3) holds and $N_m{}^* < 1$, then

$$\| b \| \leqslant \frac{1}{1 - N_m{}^*} \| b \|_m \qquad (\forall b \in B). \tag{3.5}$$

Proof. From (1.1) and (3.3)

$$\| b \| \leqslant \| b \|_m + \| b \|_m^* \leqslant \| b \|_m + N_m{}^* \| b \|.$$

Hence

$$(1 - N_m{}^*) \| b \| \leqslant \| b \|_m .$$

Our main result is

THEOREM 3.1. *Assume that* (3.1), (3.2), (3.3), *and* (3.4) *hold and that* $N_m{}^* < 1$. *Then*

$$\phi(\| Ta - T_m a \|) + \phi_m(\| Ta - T_m a \|_m)$$
$$\leqslant \| a \|_m^* + \| a \|_m^{**} + \frac{2N_m{}^*}{1 - N_m{}^*} \| a \|_m + 2N_m^{**} \| a \| \qquad (\forall a \in A). \quad (3.6)$$

Proof. Using (3.1) and (3.2), we get

$$\phi(\| Ta - T_m a \|) + \phi_m(\| Ta - T_m a \|_m)$$
$$\leqslant (\| a - T_m a \| - \| a - Ta \|) + (\| a - Ta \|_m - \| a - T_m a \|_m).$$

But, by (1.1) and (1.2), we have

$$\| a - T_m a \| - \| a - Ta \| \leqslant \| a - T_m a \|_m - \| a - Ta \| + \| a - T_m a \|_m^* ,$$
$$\| a - Ta \|_m - \| a - T_m a \|_m \leqslant \| a - Ta \| - \| a - T_m a \|_m + \| a - T_m a \|_m^{**}.$$

Adding up, we arrive at

$$\phi(\| Ta - T_m a \|) + \phi_m(\| Ta - T_m a \|_m) \leqslant \| a - T_m a \|_m^* + \| a - Ta \|_m^{**}. \quad (3.7)$$

We estimate each term separately. Using (3.3) and (3.5), we get

$$\| a - T_m a \|_m^* \leqslant \| a \|_m^* + N_m{}^* \| T_m a \| \leqslant \| a \|_m^* + \frac{N_m{}^*}{1 - N_m{}^*} \| T_m a \|_m .$$

But

$$\| T_m a \|_m \leqslant \| a \|_m + \| a - T_m a \|_m \leqslant 2 \| a \|_m .$$

Hence

$$\| a - T_m a \|_m^* \leqslant \| a \|_m^* + \frac{2N_m{}^*}{1 - N_m{}^*} \| a \|_m . \quad (3.8)$$

Next, using (3.4), we get

$$\| a - Ta \|_m^{**} \leqslant \| a \|_m^{**} + N_m^{**} \| Ta \|.$$

But, again,

$$\| Ta \| \leqslant \| a \| + \| a - Ta \| \leqslant 2 \| a \|.$$

Hence,

$$\| a - Ta \|_m^{**} \leqslant \| a \|_m^{**} + 2N_m^{**} \| a \|. \tag{3.9}$$

Inserting inequalities (3.8) and (3.9) into inequality (3.7), we end up with (3.6). Q.E.D.

Two special cases deserve special mention.

COROLLARY 3.1. *If* $\| \ \|_m^{**} = 0$, *so that* $\| a \|_m \leqslant \| a \|$ $(\forall a \in A)$, *and if* $N_m{}^* < 1$, *then*

$$\phi(\| Ta - T_m a \|) + \phi_m(\| Ta - T_m a \|_m) \leqslant \| a \|_m^* + \frac{2N_m{}^*}{1 - N_m{}^*} \| a \|. \tag{3.10}$$

COROLLARY 3.2. *If* $\| \ \|_m^{**} = \| \ \|_m^*$, $N_m^{**} = N_m{}^* < 1$, *then*

$$\begin{aligned} &\phi(\| Ta - T_m a \|) + \phi_m(\| Ta - T_m a \|_m) \\ &\qquad \leqslant 2 \| a \|_m^* + \frac{2N_m{}^*}{1 - N_m{}^*} \| a \|_m + 2N_m{}^* \| a \|. \end{aligned} \tag{3.11}$$

4. CONVERGENCE OF ALGORITHMS (CONCRETE CASES)

We now turn to concrete applications of the results of Section 3.

EXAMPLE 4.1. We take

$A = C^1 =$ the set of continuously differentiable functions on $I = [0, 1]$,
$B =$ the set of algebraic polynomials of degree $\leqslant n$,

$$\| a \| = \max_{x \in I} | a(x) |,$$

$$\| a \|_m = \max_{x \in I_m} | a(x) |,$$

where I_m is a finite subset of I consisting of points x_k, which we call *nodes*. $\| a \|_m$ is a discrete norm. Moreover,

$$\| a \|_m \leqslant \| a \|,$$

so we can take (as in Corollary 3.1)

$$\| a \|_m^{**} = 0.$$

Choose

$$x_k = \frac{k}{m} \qquad (k = 0, 1, \ldots, m - 1).$$

(A slightly more advantageous choice would have been

$$x_k = \frac{k + (1/2)}{m} \qquad (k = 0, 1, \ldots, m - 1);$$

cf. [1], p. 93). Let $x_k \leqslant x \leqslant x_k + 1/m$. Since

$$a(x) = a(x_k) + \int_{x_k}^{x} a'(\xi)\, d\xi,$$

$$| a(x)| \leqslant | a(x_k)| + (x - x_k) \max | a'(\xi)| \leqslant \| a \|_m + \frac{1}{m} \| a' \|.$$

Hence we get

$$\| a \| \leqslant \| a \|_m + \frac{1}{m} \| a' \|,$$

so that we may choose

$$\| a \|_m^* = \frac{1}{m} \| a' \|.$$

If $b \in B$ (i.e., b is a polynomial of degree $\leqslant n$), we have by Markoff's inequality (cf. [1], p. 91)

$$\| b \|_m^* = \frac{1}{m} \| b' \| \leqslant \frac{2n^2}{m} \| b \|.$$

Thus, we may take

$$N_m{}^* = \frac{2n^2}{m}.$$

Applying Corollary 3.1, we now get with $\phi(\sigma) = \gamma\sigma$ (dropping the term involving ϕ_m):

$$\gamma \| Ta - T_m a \| \leqslant \frac{1}{m} \| a' \| + \frac{2(2n^2/m)}{1 - (2n^2/m)} \| a \| \qquad (m > 2n^2), \quad (4.1)$$

an inequality essentially contained in Cheney [1], p. 92. We are here particularly interested in the behavior of the left hand side of (4.1) as $m \to \infty$. Obviously, (4.1) implies

$$\| Ta - T_m a \| = O\left(\frac{1}{m}\right). \qquad (4.2)$$

To improve this estimate we have to put further restrictions upon the functions a. [Cheney's result in [1], p. 92, is in the opposite direction. He requires just continuity of a and gets a weaker estimate in terms of the modulus of continuity $\omega(t, a)$.]

EXAMPLE 4.2. Take

$A = C^2 =$ the set of twice continuously differentiable functions on $I = [0, 1]$, B, $\| a \|$, $\| a \|_m$ as in Example 4.1. We augment, however, I_m, by adding to it $x_m = m/m = 1$. Thus, we have

$$x_k = \frac{k}{m} \qquad (k = 0, 1, \ldots, m).$$

If $x_k \leqslant x \leqslant x_k + 1/m = x_{k+1}$, we use the formula

$$a(x) = a(x_k)\, m(x_{k+1} - x) + a(x_{k+1})\, m(x - x_k) + \int_0^1 K(x, \xi)\, a''(\xi)\, d\xi,$$

with

$$K(x, \xi) = \begin{cases} (x - x_k)(\xi - x_{k+1}) & \text{if } x \leqslant \xi, \\ (\xi - x_k)(x - x_{k+1}) & \text{if } x \geqslant \xi, \end{cases}$$

and deduce

$$\begin{aligned} | a(x) | &\leqslant \max(| a(x_k) |, | a(x_{k+1}) |) + \tfrac{1}{2}(x - x_k)(x_{k+1} - x) \max | a''(\xi) | \\ &\leqslant \| a \|_m + \frac{1}{8m^2} \| a'' \|. \end{aligned}$$

Hence

$$\| a \| \leqslant \| a \|_m + \frac{1}{8m^2} \| a'' \|$$

and we are lead to take

$$\| a \|_m^* = \frac{1}{8m^2} \| a'' \|.$$

If $b \in B$ we get, again, by Markoff's inequality (iterated)

$$\| b \|_m^* = \frac{1}{8m^2} \| b'' \| \leqslant \frac{n^4}{2m^2} \| b \|$$

and subsequently

$$N_m{}^* = \frac{n^4}{2m^2}\,.$$

Corresponding to (4.1), we thus have

$$\gamma \| Ta - T_m a \| \leqslant \frac{1}{m^2} \| a'' \| + \frac{2(n^4/2m^2)}{1 - (n^4/2m^2)} \| a \| \qquad (m > 1/\sqrt{2}\, n^2) \tag{4.3}$$

and corresponding to (4.2),

$$\| Ta - T_m a \| = O\left(\frac{1}{m^2}\right). \tag{4.4}$$

We have improved the previous $O(1/m)$ to $O(1/m^2)$. It does not seem likely that this can be easily improved further; $O(1/m^2)$ is about the optimum which we can hope for.

Remark 4.1. A *theoretical* possibility of improving the estimate is, however, the following. For each m, we consider the mapping (function to sequence)

$$Q_m : a \to (a(x_0), \ldots, a(x_{r_m}))$$

and the mapping (sequence to function)

$$P_m : a \to \sum_{k=0}^{r_m} h_{mk}(x)\, a_k ,$$

where $h_{mk}(x)$ are given functions and $r_m = 1 + \operatorname{card} I_m$. (In Example 4.1 we had

$$h_{mk}(x) = \begin{cases} 1 & \text{if} \quad x \in \left[x_k , x_k + \dfrac{1}{m}\right], \\ 0 & \text{elsewhere,} \end{cases} \qquad r_m = m,$$

and in Example 4.2,

$$h_{mk}(x) = \begin{cases} m(x - x_{k-1}) & \text{if} \quad x \in [x_{k-1} , x_k], \\ m(x_{k+1} - x) & \text{if} \quad x \in [x_k , x_{k+1}], \\ 0 \quad \text{elsewhere} \end{cases} \qquad r_m = m + 1.)$$

The basic assumption is that

$$U_m \to I \qquad \text{as} \qquad m \to \infty,$$

where $U_m = P_m Q_m$. The corresponding assumptions on $h_{mk}(x)$ are well-known. Indeed, under suitable assumptions on $h_{mk}(x)$, it is even possible to prove a much stronger result, namely,

$$\| a - U_m a \| \leqslant \frac{C}{m^N} \| a^{(N)} \|.$$

For example, it suffices to assume that

$$\sum_{k=1}^{r_m} h_{mk}(x) = 1,$$

$$\sum_{k=1}^{r_m} (x - x_k)\, h_{mk}(x) = 0,$$

$$\cdots$$

$$\sum_{k=1}^{r_m} (x - x_k)^{N-1} h_{mk}(x) = 0,$$

$$\sup_m \sum_{k=1}^{r_m} | x - x_k |^N | h_{mk}(x) | < \infty.$$

With

$$A = C^N = \text{the set of } N \text{ times continuously differentiable functions on}$$
$$I = [0, 1], \qquad \| a \|_m = \| U_m a \|,$$

and B, $\| a \|$ as before (Examples 4.1 and 4.2), we are, then, lead to

$$\| Ta - T_m a \| = O\left(\frac{1}{m^N}\right). \tag{4.5}$$

The problem is that $\| a \|_m$ is quite a complicated norm, in general not of the Tschebycheff type, so in concrete cases the computation of $T_m a$ might cause difficulties and the fact that we have a better estimate will be of little actual help.

EXAMPLE 4.3. We take A and B as in Example 4.1; $A = C^1$, $B =$ the set of polynomials of degree $\leqslant n$, but change the norms, namely,

$$\| a \| = \left(\int_I | a(x)|^p \, dx\right)^{1/p},$$

and

$$\| a \|_m = \left(\sum_{I_m} | a(x_k)|^p \frac{1}{m}\right)^{1/p} \qquad \left(\text{with } x_k = \frac{k}{m}\right).$$

Here $1 \leqslant p < \infty$. We have

$$\begin{aligned}
\| a \|^p - \| a \|_m^p &= \int_0^1 | a(x)|^p \, dx - \sum_{k=0}^{m-1} | a(x_k)|^p \frac{1}{m} \\
&= \sum_{k=0}^{m-1} \int_{x_k}^{x_k+1/m} (| a(x)|^p - | a(x_k)|^p) \, dx \\
&= \sum_{k=0}^{m-1} \int_{x_k}^{x_k+1/m} \left(\int_{x_k}^{x} p \, | a(\xi)|^{p-2} a(\xi) \, a'(\xi) \, d\xi\right) dx.
\end{aligned}$$

In view of Hölder's inequality, the last integral is bounded above by

$$p \left(\int_{x_k}^{x_k+1/m} | a(\xi)|^p \, d\xi\right)^{1-1/p} \left(\int_{x_k}^{x_k+1/m} | a'(\xi)|^p \, d\xi\right)^{1/p},$$

and, thus, the corresponding summand is bounded above by the last expression multiplied by $1/m$. Thus, using once more Hölder's inequality (for sums), we get

$$| \| a \|^p - \| a \|_m^p | \leqslant \frac{p}{m} \left(\sum_{k=0}^{m-1} \int_{x_k}^{x_k+1/m} | a(\xi)|^p \, d\xi \right)^{1-1/p} \left(\sum_{k=0}^{m-1} \int_{x_k}^{x_k+1/m} | a'(\xi)|^p \, d\xi \right)^{1/p}$$

$$= \frac{p}{m} \| a \|^{p-1} \| a' \|.$$

On the other hand,

$$| \| a \|^p - \| a \|_m^p | \geqslant \| a \|^{p-1} | \| a \| - \| a \|_m |,$$

and so,

$$| \| a \| - \| a \|_m | \leqslant \frac{p}{m} \| a' \|.$$

Thus, we can take

$$\| a \|_m^* = \| a \|_m^{**} = \frac{p}{m} \| a' \|.$$

If $b \in B$, we may apply the L_p-version of Markoff's inequality (cf. Stein [9]) and get

$$\| b \|_m^* = \| b \|_m^{**} \leqslant \frac{pA_p n^2}{m} \| b' \|,$$

where A_p is a constant depending on p only. Thus, we may take

$$N_m{}^* = N_m^{**} = \frac{B_p n^2}{m} \qquad \text{(with } B_p = pA_p\text{)}.$$

Applying Corollary 3.2, we now get with $\phi(u) = \gamma u^q$, $1/q = \max(\frac{1}{2}, 1 - 1/p)$,

$$\gamma \| Ta - T_m a \|^q \leqslant \frac{A_p}{m} \| a' \| + \frac{2(B_p n^2/m)}{1 - (B_p n^2/m)} \| a \| + \frac{2B_p n^2}{m} \| a \|_m \qquad (m > B_p n^2), \tag{4.6}$$

which, in particular, implies

$$\| Ta - T_m a \| = O\left(\frac{1}{m^{1/q}}\right). \tag{4.7}$$

EXAMPLE 4.4. [Pólya (or de la Vallée–Poussin) algorithm.] Take

$A = C^0 =$ the set of continuous functions on $I = [0, 1]$,

$B =$ the set of algebraic polynomials of degree $\leqslant n$,

$$\| a \| = \max_I | a(x)|,$$

$$\| a \|_m = \left(\int_I | a(x)|^{p_m} \, dx \right)^{1/p_m},$$

where $p_m \to \infty$ as $m \to \infty$. From the inequality

$$| a(x)| \leqslant | a(\xi)| + | x - \xi | \, \| a' \|,$$

we obtain by virtue of the L_p-triangle inequality for $p \geqslant 1$,

$$| a(x)| \left(\int_{J_h} 1^p \, d\xi\right)^{1/p} \leqslant \left(\int_{J_h} | a(\xi)|^p \, d\xi\right)^{1/p} + \left(\int_{J_h} | x - \xi |^p \, d\xi\right)^{1/p} \| a' \|,$$

where J_h is any subinterval of I of length $h \leqslant \frac{1}{2}$, containing x. This yields

$$\begin{aligned} \| a \| &\leqslant h^{-1/p_m} \| a \|_m + h \| a' \| \\ &= \| a \|_m + ((e^{(\log(1/h)/p_m)} - 1) \| a \|_m + h \| a' \|). \end{aligned}$$

On the other hand, it is trivial that

$$\| a \|_m \leqslant \| a \|.$$

We now choose

$$h = h_m = \frac{\log p_m}{p_m}.$$

(This is about the best choice.) We then end up with

$$\| a \|_m \leqslant \| a \| \leqslant \| a \|_m + \frac{C \log p_m}{p_m} (\| a \| + \| a' \|),$$

with C independent of m. Accordingly, we choose

$$\| a \|_m^* = \frac{C \log p_m}{p_m} (\| a \| + \| a' \|), \qquad \| a \|_m^{**} = 0.$$

If $b \in B$, Markoff's inequality is again available. We do not include the details and content ourselves with the estimate

$$\| Ta - T_m a \| = O\left(\frac{\log p_m}{p_m}\right), \qquad m \to \infty. \tag{4.8}$$

We feel that it is unlikely that this can be improved upon very much.

5. An Abstract Moment Problem

Let A be a normed space and let A_m be a sequence of such spaces. Let $P_m : A_m \to A$ and $Q_m : A \to A_m$ be as in Section 1, with

$$U_m \to I, \tag{5.1}$$

where $U_m = P_m Q_m$. If ϕ is a continuous linear functional on A, its norm is by definition given by

$$\| \phi \|_{A'} = \sup | \phi(a)| / \| a \|_A .$$

Let us define, for each m, a continuous linear functional ϕ_m on A_m, by setting

$$\phi_m(a) = \phi(P_m a).$$

The corresponding norm is

$$\| \phi_m \|_{A_m'} = \sup | \phi_m(a)| / \| a \|_{A_m} = \sup | \phi(P_m a)| / \| a \|_{A_m} .$$

What we term as a moment problem is to relate the norms $\| \phi \|_{A'}$ and $\| \phi_m \|_{A_m'}$. To this end, we prove

THEOREM 5.1. *Assume that besides* (5.1), *we have*

$$\| P_m a \|_A \leqslant \| a \|_{A_m}, \tag{5.2}$$

$$\overline{\lim_{m\to\infty}} \| Q_m a \|_{A_m} \leqslant \| a \|. \tag{5.3}$$

Then

$$\lim_{m\to\infty} \| \phi_m \|_{A_m'} = \| \phi \|_{A'} .$$

Proof. We have by (4.2)

$$| \phi_m(a)| = |\phi(P_m a)| \leqslant \| \phi \|_{A'} \| P_m a \|_A \leqslant \| \phi \|_{A'} \| a \|_{A_m} .$$

Therefore

$$\overline{\lim_{m\to\infty}} \| \phi_m \|_{A_m'} \leqslant \| \phi \|_{A'} .$$

To prove an inequality in the opposite sense, choose, for $\epsilon > 0$, an a such that

$$| \phi(a)| \geqslant \| \phi \|_{A'}(1 - \epsilon), \qquad \| a \|_A = 1,$$

and m so that

$$\| a - U_m a \| < \epsilon;$$

we have used (5.1). It follows that

$$| \phi(U_m a)| \geqslant \| \phi \|_{A'}(1 - 2\epsilon).$$

But

$$\phi(U_m a) = \phi(P_m Q_m a) = \phi_m(Q_m a),$$

so that

$$\overline{\lim_{m\to\infty}} \mid \phi(U_m a)\mid \leqslant [\overline{\lim_{m\to\infty}} \| \phi_m \|_{A_m'}] \lim_{m\to\infty} \| U_m a \|$$

$$= [\overline{\lim_{m\to\infty}} \| \phi_m \|_{A_m'}] \| a \| = \overline{\lim_{m\to\infty}} \| \phi_m \|_{A_m'} .$$

Therefore

$$\overline{\lim_{m\to\infty}} \| \phi_m \|_{A_m'} \geqslant \| \phi \|_{A'}(1 - 2\epsilon),$$

and, $\epsilon > 0$ being arbitrary,

$$\overline{\lim_{m\to\infty}} \| \phi_m \|_{A_m'} \geqslant \| \phi \|_{A'} . \qquad \text{Q.E.D.}$$

We now give a concrete application (corresponding to the classical little moment problem; cf., e.g. [10], Chap. III).

EXAMPLE 5.1. We take

$$A = C^0 = \text{the set of continuous functions on } I = [0, 1],$$

$$\| a \|_A = \| a \|_p = \left(\int_I \mid a(x)\mid^p dx\right)^{1/p} \qquad \text{with} \quad 1 \leqslant p \leqslant \infty$$

$$\text{(interpreted as } \max_{x\in I} \mid a(x)\mid \text{ if } p = \infty),$$

$$A_m = \text{Euclidean } (m + 1)\text{-space},$$

$$\| a \|_{A_m} = \| a \|_p = \left(\sum_{k=0}^{m} \mid a_k \mid^p \frac{1}{m}\right)^{1/p}$$

$$\text{(interpreted as } \max_{1\leqslant k\leqslant m} \mid a_k \mid \text{ if } p = \infty),$$

$$(P_m a)(x) = \sum_{k=0}^{m} \beta_{km}(x)\, a_k , \qquad \text{with} \quad \beta_{km}(x) = \binom{m}{k} x^k(1 - x)^{m-k},$$

$$Q_m a = \left(a(0), a\left(\frac{1}{m}\right), a\left(\frac{2}{m}\right),..., a(1)\right).$$

We check the validity of (5.1), (5.2), and (5.3). That (5.1) holds is the classical theorem of Bernstein (cf. [1], p. 66–69). Note that $U_m = P_m Q_m$ is the Bernstein operator. That (5.3) holds is obvious (existence of Riemann integral, cf. also Example 4.3). There remains thus (5.2), i.e., the inequality

$$\| P_m a \|_p \leqslant \| a \|_p \qquad \text{for} \quad 1 \leqslant p \leqslant \infty. \tag{5.4}$$

In view of the Schur interpolation theorem (a special case of the M. Riesz interpolation theorem), it suffices to prove (5.4) in the extremal cases $p = 1$ and $p = \infty$. We obtain

$$\| P_m a \|_1 \leqslant \sum_{k=0}^{m} \| \beta_{km} \|_1 \, | a_k | \leqslant m[\max_{0 \leqslant k \leqslant m} \| \beta_{km} \|_1] \, \| a \|_1 ,$$

$$\| P_m a \|_\infty \leqslant \Big\| \sum_{k=0}^{m} \beta_{km} \Big\|_\infty \| a \|_\infty = \| a \|_\infty .$$

Using the Euler integrals, we have

$$\begin{aligned} \| \beta_{km} \|_1 &= \binom{m}{k} \int_0^1 x^k (1-x)^{m-k} \, dx \\ &= \binom{m}{k} B(k+1, m-k+1) = \binom{m}{k} \frac{\Gamma(k+1)\,\Gamma(mk+1)}{\Gamma(m+2)} \\ &= \binom{m}{k} \frac{k!\,(m-k)!}{(m+1)!} = \frac{1}{m+1} \leqslant \frac{1}{m}, \end{aligned}$$

so that $\| P_m a \|_1 \leqslant \| a \|_1$. Let ϕ be any continuous linear functional on A. We now have

$$\phi_m(a) = \phi(P_m a) = \sum_{k=0}^{m} \phi(\beta_{km})\, a_m = \sum_{k=0}^{m} \lambda_{km} a_m \frac{1}{m}, \tag{5.5}$$

with $\lambda_{km} = m\phi(\beta_{km})$. Therefore we find

$$\| \phi_m \|_{A_m'} = \left(\sum_{k=0}^{m} | \lambda_{km} |^{p'} \frac{1}{m} \right)^{1/p'} \qquad \left(\frac{1}{p} + \frac{1}{p'} = 1 \right).$$

Application of Theorem 5.1 thus yields

$$\| \phi \|_{A'} = \lim_{m \to \infty} \left(\sum_{k=0}^{m} | \lambda_{km} |^{p'} \frac{1}{m} \right)^{1/p'}. \tag{5.6}$$

This should be compared with the classical results (cf., notably, [8]).

Remark 5.1. Using a more general interpolation theorem we can cover the case of an arbitrary rearrangement invariant norm (in place of the L_p norm $\| \ \|_p$); cf. [3], p. 80. This is the widest range for Schur interpolation.

EXAMPLE 5.2. We take

$A = C^1 =$ the set of continuously differentiable functions on $I = [0, 1]$,

$$\| a \|_A = \max_{0 \leqslant x \leqslant 1} | a'(x)|,$$

$A_m =$ Euclidean $(m + 1)$-space

$$\| a \|_{A_m} = \max_{0 \leqslant k \leqslant m-1} | a_{k+1} - a_k |,$$

$$P_m \text{ and } Q_m \qquad \text{as in Example 5.1.}$$

The following formula is of interest:

$$DP_m = mP_{m-1}\Delta. \tag{5.7}$$

Here D and Δ denote differentiation and the difference operators, respectively, i.e.,

$$Da(x) = a'(x), \qquad \Delta a_k = a_{k+1} - a_k .$$

Using (5.7), it is not hard to see that

$$\max_{0 \leqslant x \leqslant 1} | DU_m a(x) - Da(x)| \to 0 \text{ as } m \to \infty.$$

In other words, (5.1) holds in this case, too. Also (5.2) and (5.3) can be readily verified. If ϕ is any continuous linear functional on A, the corresponding ϕ_m is again given by (5.5) and we have

$$\| \phi_m \|_{A_m'} = \sum_{k=0}^{m-1} | \Lambda_{km} | \frac{1}{m},$$

where Λ_{km} is defined by

$$\begin{aligned} \Lambda_{k-1,m} - \Lambda_{k,m} &= \lambda_{k,m} \qquad (k = 1,..., m - 1), \\ -\Lambda_{0,m} &= \lambda_{0,m}, \qquad \Lambda_{m,m} = \lambda_{m,m} . \end{aligned} \tag{5.8}$$

Note that Λ_{km} is well-defined since the compatibility condition

$$\sum_{k=0}^{m} \lambda_{k,m} = 0$$

for the solvability of (5.8) is obviously fulfilled. Application of Theorem 5.1 now yields

$$\| \phi \|_{A'} = \lim_{m \to \infty} \sum_{k=0}^{m-1} | \Lambda_{km} | \frac{1}{m}. \tag{5.9}$$

Remark 5.2. The result (5.9) can be extended in several direction. For example, we can treat the case of Lipschitz norms (i.e.,

$$\| a \|_A = \sup | a(x) - a(y)|/| x - y |^\alpha$$

with $0 < \alpha < 1$). Here it is advantageous to use the theory of interpolation spaces. Note that these spaces were used by Löfström [4] to solve a dual problem.

References

1. E. W. Cheney, "Introduction to Approximation Theory," McGraw-Hill, New York, 1966.
2. E. W. Cheney and D. E. Wulbert, Existence and unicity of best approximations, *Math. Scand.* **24** (1969), 113–140.
3. B. Kripke, Best approximation with respect to nearly norms, *Numer. Math.* **6** (1964), 103–105.
4. J. Löfström, On the moments of functions satisfying a Lipschitz condition, *Math. Scand.* **23** (1968), 177–187.
5. G. G. Lorentz, "Bernstein Polynomials," Toronto Univ. Press, Toronto, Canada, 1953.
6. R. R. Phelps, Uniqueness of Hahn–Banach extensions and unique best approximation, *Trans. Amer. Math. Soc.* **95** (1960), 238–255.
7. Z. Semadeni, Limit properties of ordered families of linear metric spaces, *Studia Math.* **20** (1961), 245–270.
8. I. Singer, Caractérisation des éléments de meilleure approximation dans un espace de Banach quelconque, *Acta Sci. Math.* (*Szeged*) **17** (1956), 181–189.
9. E. M. Stein, Interpolation in polynomial classes and Markoff's inequality, *Duke Math. J.* **24** (1957), 467–476.
10. D. V. Widder, "The Laplace Transform," Princeton Univ. Press, Princeton, N. J., 1946.