INTERPOSITION AND APPROXIMATION

Bernard Kripke and Richard Holmes

Let $\mathscr{B}(X)$ be the space of all bounded real-valued functions on a set X, with the norm $||f|| = \sup\{|f(x)|: x \in X\}$, and let K be any nonempty subset of $\mathscr{B}(X)$. The question whether an element f of $\mathscr{B}(X)$ has a best approximation g in K (such that $||f - g|| = \delta(f) = \inf\{||f - h||: h \in K\}$) can be formulated as the problem of interposing a function g in K between two functions, $L(\cdot, f)$ and $U(\cdot, f)$, which are constructed out of K by certain lattice operations. If K is closed with respect to these lattice operations, or has a certain interposition property, the best approximation will always exist.

For example, X might be a bounded subset of a Banach space E and K might be the set of restrictions to X of the continuous linear functionals in E^* [2, 6]. $U(\cdot, f)$ is then constructed in two stages: first the suprema of bounded subsets of K are formed, and then $U(\cdot, f)$ is obtained as a decreasing sequential limit of such suprema. In two other typical cases, K consists of the bounded continuous functions on a paracompact space [5], or the distance decreasing functions on a metric space. These two share the property of *translational invariance*:

$$(1) \qquad \text{if } f \in K \text{ and } c \text{ is a constant, then } (f + c) \in K,$$

which permits $U(\cdot, f)$ to be constructed by forming suprema alone, without the intervention of decreasing sequential limits. In the last of these sample cases, it actually turns out that $U(\cdot, f)$ is itself in K, and is thus the *largest* of the best approximators to f in K.

1. Mere existence. For every $\rho > 0$, there is a $g \in K$ such that $||f - g|| < \delta(f) + \rho$, or in other words $f - \delta(f) - \rho < g < f + \delta(f) + \rho$. Therefore, $U_\rho(x, f) = \sup\{g(x): g \in K, g \leq f + \delta(f) + \rho\}$ lies between $f - \delta(f) - \rho$ and $f + \delta(f) + \rho$, and dominates

$$L_\rho(x, f) = \inf\{g(x): g \in K, g \geq f - \delta(f) - \rho\}.$$

L_ρ and U_ρ are, respectively, monotonically increasing and decreasing functions of ρ, so that

$$f - \delta(f) \leq L(\cdot, f) = \lim_{n\to\infty} L_{1/n}(\cdot, f) \leq \lim_{n\to\infty} U_{1/n}(\cdot, f)$$

$$= U(\cdot, f) \leq f + \delta(f).$$

PROPOSITION 1.   If $f \in \mathscr{B}(X)$ and $g \in K$, then $\| f - g \| = \delta(f)$ if, only if, $L(\cdot, f) \leqq g \leqq U(\cdot, f)$.

Proof.

$$\| f - g \| = \delta(f) \Longleftrightarrow f - \delta(f) \leqq g \leqq f$$
$$+ \delta(f) \Longleftrightarrow f - \delta(f) - \rho \leqq g \leqq f + \delta(f) + \rho$$

for each $\rho > 0 \Leftrightarrow L_\rho(\cdot, f) \leqq g \leqq U_\rho(\cdot, f)$ for each $\rho < 0 \Leftrightarrow L(\cdot, f) \leqq g \leqq U(\cdot, f)$.

The significance of this proposition is that f has a best approximation in K if, and only if, a function g in K can be interposed between two functions $L(\cdot, f)$ and $U(\cdot, f)$ which are respectively the increasing sequential limit of infima and the decreasing sequential limit of suprema of functions in K.

DEFINITION. K has the *interposition property* (Ip) if for each $f \in \mathscr{B}(X)$, there exists a $g \in K$ such that $L(\cdot, f) \leqq g \leqq U(\cdot, f)$. K has the *upper [lower] interposition property* (Uip [Lip]) if for each $f \in \mathscr{B}(X)$, $U(\cdot, f)$ $[L(\cdot, f)]$ is in K.

COROLLARY 2.   *K has the* Ip *if, and only if, for each $f \in \mathscr{B}(X)$, there is a $g \in K$ such that $\| f - g \| = \delta(f)$.   If K has the* Uip [Lip], *$U(\cdot, f)$ $[L(\cdot, f)]$ is the largest [smallest] function $g \in K$ such that $\| f - g \| = \delta(f)$.*

PROPOSITION 3.   Of the following statements, (2) $\Rightarrow$ (3) $\Rightarrow$ (4).
(2)   For each $x \in X$, there is a family $S(x)$ of functions from $\mathscr{B}(X)$ to R (the real numbers) such that
    (a) $f, g \in \mathscr{B}(X)$, $x \in X$, $\lambda \in S(x)$, $f \leqq g \Rightarrow \lambda(f) \leqq \lambda(g)$;
    (b)   if $\{f_n\}$ is a monotonically decreasing sequence in $\mathscr{B}(X)$ converging pointwise to f, $x \in X$, and $\lambda \in S(x)$, then $\lambda(f) = \lim_{n \to \infty} \lambda(f_n)$;
    (c) $f \in K \Leftrightarrow f(x) \leqq \lambda(f)$ for each $x \in X$ and $\lambda \in S(x)$.
(3)(a)   If F is a nonempty family of functions in K which is uniformly bounded above, its supremum $\varphi(\cdot, F)$ is also in K.
    (b)   If $\{f_n\}$ is a monotonically decreasing sequence of functions in K converging pointwise to a function $f \in \mathscr{B}(X)$, then $f \in K$.
(4) K has the Uip.

Proof.   (2a) & (2c) $\Rightarrow$ (3a).   For each $f \in F$, $f \leqq \varphi(\cdot, F)$.   Thus for each $x \in X$ and $\lambda \in S(x)$, $f(x) \leqq \lambda(f) \leqq \lambda(\varphi(\cdot, F))$.   Therefore, $\varphi(x, F) = \sup \{f(x) : f \in F\} \leqq \lambda(\varphi(\cdot, F))$.   According to (2c), $\varphi(\cdot, F) \in K$. (2b) & (2c) $\Rightarrow$ (3b).   Trivial.   (3) $\Rightarrow$ (4).   Likewise.

2. **Wherein the assumption of translational invariance improves the theory.** *Throughout this section, it shall be assumed that K has the property* (1) *of translational invariance.* It then makes sense to introduce the *upper and lower K-envelopes* of a bounded function, $\alpha(x, f) = \inf \lfloor g(x)\colon g \in K, g \geqq f\rfloor$ and

$$\beta(x, f) = \sup \{g(x)\colon g \in K, g \leqq f\} \ .$$

The quantity $\Delta(f) = (1/2) \| \alpha(\cdot, f) - \beta(\cdot, f) \|$ can be used to estimate $\delta(f)$, while $u(\cdot, f) = \beta(\cdot, f) + \Delta(f)$ and $l(\cdot, f) = \alpha(\cdot, f) - \Delta(\cdot, f)$ can be used to approximate $U(\cdot, f)$ and $L(\cdot, f)$. More precisely, these new constructs are related to U, L, and δ as follows.

PROPOSITION 4. $\delta(f) \geqq \Delta(f)$. $U(\cdot, f) = \beta(\cdot, f) + \delta(f) \geqq u(\cdot, f) \geqq l(\cdot, f) \geqq \alpha(\cdot, f) - \delta(f) = L(\cdot, f)$.

Proof.

$$u(\cdot, f) - l(\cdot, f) = \beta(\cdot, f) - \alpha(\cdot, f) + 2\Delta(f)$$
$$= \| \alpha(\cdot, f) - \beta(\cdot, f) \| - (\alpha(\cdot, f) - \beta(\cdot, f)) \geqq 0 \ .$$

In the presence of (1), it is clear that $U_\rho(\cdot, f) = \beta(\cdot, f) + \delta(f) + \rho$ and $L_\rho(\cdot, f) = \alpha(\cdot, f) - \delta(f) - \rho$. We know that for every $\rho > 0$, there is a $g \in K$ such that

$$\alpha(\cdot, f) - \delta(f) - \rho = L_\rho(\cdot, f) \leqq g \leqq U_\rho(\cdot, f) = \beta(\cdot, f) + \delta(f) + \rho \ ,$$

whence $\alpha(\cdot, f) - \beta(\cdot, f) \leqq 2(\delta(f) + \rho)$. Since $\alpha(\cdot, f) - \beta(\cdot, f)$ is evidently nonnegative, and since ρ can be arbitrarily small, it follows that $2\Delta(f) = \| \alpha(\cdot, f) - \beta(\cdot, f) \| \leqq 2\delta(f)$. Moreover, letting ρ tend to 0, we find that $U(\cdot, f) = \lim_{\rho \to 0} U_\rho(\cdot, f) = \beta(\cdot, f) + \delta(f)$, and likewise $L(\cdot, f) = \alpha(\cdot, f) - \delta(f)$. The rest is trivial.

COROLLARY 5. *In order that an element $g \in K$ satisfy $\| f - g \| = \Delta(f)$, it is necessary and sufficient that $l(\cdot, f) \leqq g \leqq u(\cdot, f)$, in which case $\Delta(f) = \delta(f)$, $U(\cdot, f) = u(\cdot, f)$, and $L(\cdot, f) = l(\cdot, f)$.*

Proof. If $g \in K$ and $l(\cdot, f) \leqq g \leqq u(\cdot, f)$, then

$$f - \Delta(f) \leqq \alpha(\cdot, f) - \Delta(f) = l(\cdot, f) \leqq g \leqq u(\cdot, f)$$
$$= \beta(\cdot, f) + \Delta(f) \leqq f + \Delta(f) \ .$$

Thus

$$\delta(f) \leqq \| f - g \| \leqq \Delta(f) \leqq \delta(f) \ .$$

Conversely, if $g \in K$ and $\| f - g \| = \Delta(f)$, then $\delta(f) \leqq \| f - g \| =$

$\varDelta(f) \leqq \delta(f)$. Therefore, $\delta(f) = \varDelta(f)$, $U(\cdot, f) = u(\cdot, f)$, and $L(\cdot, f) = l(\cdot, f)$. Proposition 1 then shows that $l(\cdot, f) = L(\cdot, f) \leqq g \leqq U(\cdot, f) = u(\cdot, f)$.

Observe that $u(\cdot, f)$ and $l(\cdot, f)$ are respectively the supremum and the infimum of families of functions in K. The preceding results show that when a function in K can be interposed between $l(\cdot, f)$ and $u(\cdot, f)$, the distance from K to f not only is attained, but can also be calculated directly from the upper and lower K-envelopes of f.

DEFINITION. K has the *strong interposition property* (Sip) if it is translationally invariant and for each $f \in \mathscr{B}(X)$, there exists a $g \in K$ such that $l(\cdot, f) \leqq g \leqq u(\cdot, f)$. K has the *strong upper [lower] interposition property* (Suip [Slip]) if it is translationally invariant and for each $f \in \mathscr{B}(X), u(\cdot, f) [l(\cdot, f)]$ is in K.

PROPOSITION 6. *If K is translationally invariant, the following three statements are equivalent.*

(5) K has the Suip.

(6) If F is a nonempty family of functions in K which is uniformly bounded above, its supremum $\varphi(\cdot, F)$ also is in K.

(7) For each $x \in X$, there is a family $S(x)$ of functions from $\mathscr{B}(X)$ into R such that

   (a) $f, g \in \mathscr{B}(X), x \in X, \lambda \in S(x), f \leqq g \Rightarrow \lambda(f) \leqq \lambda(g)$, and

   (b) $f \in K \Rightarrow f(x) \leqq \lambda(f)$ for each $x \in X$ and $\lambda \in S(x)$.

Proof. The demonstration that $(7) \Rightarrow (6)$ was part of the proof of Proposition (3), and the proof that $(6) \Rightarrow (5)$ follows trivially from the definition of $u(\cdot, f)$ as the supremum of the functions in K which are bounded above by $f + \varDelta(f)$. To show that $(5) \Rightarrow (7)$, take $S(x)$ for each x to consist of the single function $f \rightarrow \beta(x, f)$. It is clear at once that $\beta(x, \cdot)$ satisfies the monotonicity condition (7a), and that $f \in K \Rightarrow \beta(x, f) = f(x)$. If, conversely, $f \leqq \beta(\cdot, f)$, then f must equal $\beta(x, f)$ because $\beta(\cdot, f)$ is evidently no larger than f. But then the Suip and (1) imply that $f = \beta(\cdot, f) = u(\cdot, f) - \varDelta(f)$, which is in K.

3. **Applications.** The reader can apply Proposition 6 to his favorite of the families of bounded functions which are closed with respect to taking suprema. For example, if X is a convex subset of a vector space V, let $S(x)$, for each $x \in X$, consist of all mappings of the form $f \rightarrow af(y) + bf(z)$, where $y, z \in X, a, b \geqq 0, a + b = 1$, and $x = ay + bz$. Then f is *convex* if, and only if, $f(x) \leqq \lambda(f)$ for each $x \in X$ and $\lambda \in S(x)$. That is to say, the convex functions are charac-

terized by conditions of the type (7); and it is worth noting that condition (2b) is satisfied in this case as well. One then concludes from Proposition 6 that the family of bounded convex functions on X has the Suip, or from Proposition 3 and a moment's reflection that the family of bounded convex functions on X which vanish on a prescribed subset of X has the Uip. Both of these families of functions are thus *proximinal*—every bounded function has a best approximator in each of them. Indeed, every bounded function has a largest (but not necessarily a smallest) best approximator in each of these families. The second of them is not translationally invariant. Similar arguments can be applied to the families of bounded lower semi-continuous functions on a topological space and of bounded sub-harmonic functions on a region in the plane.

In an earlier paper [5], the authors have announced the result of applying Corollary 5 to the case in which K consists of the bounded continuous functions on a paracompact space. In this situation, $u(\cdot, f)$ and $l(\cdot, f)$ are respectively lower and upper semi-continuous, and the fact that K has the Sip is a consequence of the Interposition Theorem of Dieudonné [3, p. 75]. Similar reasoning was used by S. Mazur to compute the distance from a continuous function f on a compact topological space X to the set K of functions of the form $g \circ \varphi$, where φ is a fixed continuous mapping from X onto a topological space Y, and g is continuous on Y. The answer is, of course, that K is translationally invariant and has the Sip, so that the distance from f to K is $\varDelta(f)$. (The authors are indebted to Professor R. R. Phelps for informing them of Mazur's work. The distance formula was announced without proof in Pelczynski's paper [8], while Mazur's proof can be found in Semadeni's notes [9].)

Cheney and Goldstein [2] proved the following theorem on approximation by linear functionals.

THEOREM. *Let E be a Banach space and X be a subset of E. In order that each bounded function from X to **R** have a best approximation in the set of continuous linear functionals on E which are bounded on X, it is sufficient that 0 be interior to the closed convex hull of $X \cup (-X)$ relative to its linear span.*

In case X is bounded, it was shown by Kripke and Rockafellar [6] that this sufficient condition for the class of continuous linear functionals to be proximal is necessary as well. Using Corollary 2, the result of Cheney and Goldstein can be strengthened as follows. Let V be any locally convex linear topological space, and let X be a subset of V such that the closed convex hull Y of $X \cup (-X)$ has interior relative to the subspace W of V which it spans. Let K be

the set of continuous linear functionals on V which are bounded on X. It will be shown that K has the Ip.

For each $y \in Y$, put $U_\rho(y, f) = \sup \{g(y): g \in K, g \leqq f + \delta(f) + \rho\}$ and $U(y, f) = \lim_{\rho \to 0} U_\rho(y, f)$. These definitions extend to all of Y functions which previously had been defined only on X. U_ρ, being the supremum of a uniformly bounded family of restrictions of linear functionals to Y, is convex; and U, being the *decreasing* limit of convex functions, is convex as well. This much was noted above. In like manner, $L(\cdot, f)$ is a concave function on Y. $\{(y, c): c \geqq U(y, f)\}$ and $\{(y, c): c < L(y, f)\}$ are thus disjoint convex subsets of the product space $W \times \boldsymbol{R}$, both of which have interior. They can be separated by a closed hyperplane H in $W \times \boldsymbol{R}$ (Bourbaki [1]) which is the graph of a continuous affine functional (a linear functional plus a constant) h on W interposed between $L(\cdot, f)$ and $U(\cdot, f)$. But $L(\cdot, f)$ and $U(\cdot, f)$ clearly vanish at 0, so h is actually linear. The Hahn-Banach theorem (Bourbaki [1]) permits h to be extended to a functional $g \in K$.

The last of the typical applications mentioned in the introductory paragraph requires that X be supplied with a pseudo-metric d. K now is to consist of those bounded functions g which satisfy the Lipschitz condition $|g(x) - g(y)| \leqq d(x, y)$ for all $x, y \in X$. A function $g \in \mathscr{B}(X)$ belongs to K if, and only if, it satisfies either of the following equivalent conditions for each $x \in X$:

   (i) $g(x) \leqq g(y) + d(x, y)$ for every $y \in X$,
   (ii) $g(x) \geqq g(y) - d(x, y)$ for every $y \in X$.

The condition (i) is precisely of the type (7) which characterizes families of functions possessing the Suip, while (ii) is characteristic of families possessing the Slip. Since K thus satisfies both the Suip and the Slip, each bounded function f has a largest and a smallest best approximator in K, $u(\cdot, f)$ and $l(\cdot, f)$, and the distance from f to K is precisely $\Delta(f)$.

It is not hard to calculate $\alpha(\cdot, f)$ and $\beta(\cdot, f)$ explicitly. If $g \in K$ and $g \geqq f$, then $g(x) \geqq g(y) - d(x, y) \geqq f(y) - d(x, y)$, so that

$$\alpha(x, f) \geqq A(x, f) = \sup \{f(y) - d(x, y): y \in X\}.$$

On the other hand, the triangle inequality shows that for each y, the function $x \to f(y) - d(x, y)$ is an element of K. $A(\cdot, f)$, being the supremum of such functions, is also in K. Since

$$A(x, f) \geqq f(x) - d(x, x) = f(x), A(\cdot, f) \geqq f.$$

But then $\alpha(\cdot, f)$ can be no larger than $A(\cdot, f)$. That is, it has been shown that $\alpha(x, f) = \sup_y (f(y) - d(x, y))$.

This formula can be used to evaluate $\Delta(f)$.

$$\Delta(f) = \frac{1}{2} \sup_{x} [\alpha(x, f) - \beta(x, f)]$$

$$= \frac{1}{2} \sup_{x} [\sup_{y} (f(y) - d(x, y)) - \inf_{z} (f(z) + d(x, z))]$$

$$= \frac{1}{2} \sup_{x,y,z} [f(y) - d(x, y) - f(z) - d(x, z)]$$

$$\leqq \frac{1}{2} \sup_{y,z} [f(y) - f(z) - d(y, z)] = \frac{1}{2} \sup_{y} [f(y) - \beta(y, f)]$$

$$\leqq \frac{1}{2} \sup_{y} [\alpha(y, f) - \beta(y, f)] = \Delta(f) .$$

To summarize

(8) $$\delta(f) = \Delta(f) = \sup_{y,z} \frac{1}{2} [f(y) - f(z) - d(y, z)] ,$$

$$U(x, f) = \inf_{z} (f(z) + d(x, z)) + \Delta(f) ,$$

and

$$L(x, f) = \sup_{y} (f(y) - d(x, y) - \Delta(f) .$$

The derivation of formula (8) was discovered by Mr. Masao Kishore, a student in a seminar for undergraduates given at the University of California, Berkeley, in 1965.

4. **Brief comment.** As the authors indicated in their earlier paper [5], much of what has just been transacted could be played out in a measure-theoretic setting, in which the essential supremum plays the role of the supremum. It also seems possible to extend some of these results, notably the fact that the space of bounded continuous functions is a proximinal subset of $\mathscr{B}(X)$ when X is paracompact, to the case of complex-valued functions, or more generally functions with values in a reflexive Banach space E. With $\delta(f)$ defined as before, one introduces for each $x \in X$ and $\rho > 0$ the set $T_\rho(x)$, which is the closed convex hull in E of

$$\{g(x): g \in K, \| f - g \| \leqq \delta(f) + \rho\} ,$$

and then the set $T(x) = \bigcap_{\rho > 0} T_\rho(x)$. At least in the case that K consists of the bounded continuous functions on a paracompact space, a theorem of E. A. Michael [7] is available to guarantee that an element g of K can be interposed through the sets $T(x)$—$g(x) \in T(x)$ for each $x \in X$—and this shows by arguments like those that have gone before that f has a best approximation in K. An entirely different line of reasoning can be applied to the set of E-valued

functions on a *compact* metric space which satisfy a Lipschitz condition. One observes that best approximations need only be sought in a *bounded* subset of K, and that if E is given its weak topology, such a subset of K is compact in the topology of uniform convergence on X according to Ascoli's theorem (Kelley [4]). Since $\| f - g \|$ is a lower semicontinuous function of g when E has the weak topology, it attains its infimum. However, Kishore's formula for the distance, as well as certain of its plausible generalizations, fails even for complex-valued functions, as is shown by counterexamples due to another student in the Berkeley undergraduate seminar, Mr. Daniel Chester.

REFERENCES

1. N. Bourbaki, *Espaces Vectorielles Topologiques*, Hermann, Paris, 1953.
2. E. W. Cheney and A. A. Goldstein, *Tchebycheff approximation and related extremal problems*, J. Math. Mech. **14** (1965), 87-98.
3. J. Dieudonné, *Une généralisation des espaces compacts*, J. Math. Pures Appl. **23** (1944), 65-76.
4. J. L. Kelley, *General Topology*, Van Nostrand, Princeton, 1955.
5. B. R. Kripke and R. B. Holmes, *Approximation of bounded functions by continuous functions*, Bull. Amer. Math. Soc. **71** (1965), 896-897.
6. B. R. Kripke and R. T. Rockafellar, *A necessary condition for the existence of best approximations*, J. Math. Mech. **13** (1964), 1037-1038.
7. E. Michael, *Continuous selections*, Ann. of Math. (2) **63** (1956), 361-382.
8. A. Pelczynski, *A generalization of Stone's theorem on approximation*, Bull. Acad. Polon. Sci. Cl. III, **5** (1957), 105-107.
9. Z. Semadeni, *Simultaneous extensions and projections in spaces of continuous functions*, notes, Aarhus Universitet, Denmark, 1965.

UNIVERSITY OF CALIFORNIA, BERKELEY, AND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO