

# ON $n$ -PARAMETER FAMILIES OF FUNCTIONS AND ASSOCIATED CONVEX FUNCTIONS

BY

LEONARD TORNHEIM

**Introduction.** Let  $f(x)$  be a real-valued function continuous on the interval  $a \leq x \leq b$ . Then  $f(x)$  is said to be strictly convex if and only if the graph of any linear function for  $a \leq x \leq b$  meets the graph of  $f(x)$  in at most two points. In this situation, one may consider the linear functions on  $a \leq x \leq b$  as a two-parameter family—for each pair of points  $(x_1, y_1)$  and  $(x_2, y_2)$ ,  $x_1 \neq x_2$ , there is exactly one linear graph through these points—and the strictly convex functions as “associated” with the linear functions.

Beckenbach and Bing [1, 3]<sup>(1)</sup> generalized this situation by replacing the linear functions by a more general 2-parameter family, that is, a family of continuous functions such that for each pair of points  $(x_1, y_1)$  and  $(x_2, y_2)$ ,  $x_1 \neq x_2$ , there is one and only one member of the family through these points; then in a natural way they have introduced the associated convex functions. These authors have shown that many properties of the class of linear functions and convex functions hold for 2-parameter families and their associated convex functions. One surprising result was the observation that a 2-parameter family need not be topologically equivalent to the family of linear functions on the interval  $0 \leq x \leq 1$ .

T. Popoviciu [9] has given the definition for  $n$ -parameter families, but stated no properties. We obtain results here for such families of functions and their associated convex functions which are in part generalizations of those obtained by Beckenbach and Bing. We also obtain results related to the work of T. Popoviciu [7, 8] on convex functions associated with linear families, to that of M. M. Peixoto [6] on the derivatives of generalized convex functions, and to that on approximation discussed by S. Bernstein [4] and C. J. de la Vallee Poussin [5].

## 1. Definitions and elementary properties.

**DEFINITION 1.** An  $n$ -parameter family is a set of single-valued, real, continuous functions  $f(x)$  on an interval  $a \leq x \leq b$  such that for every set of points  $(x_i, y_i)$  ( $i = 1, \dots, n$ ) with  $a \leq x_1 < x_2 < \dots < x_n \leq b$  there is exactly one  $f(x)$  with  $f(x_i) = y_i$ .

Hereafter all functions are assumed to be single-valued, real, and continuous on  $a \leq x \leq b$ . We designate an  $n$ -parameter family by  $F$ .

Simple examples of  $n$ -parameter families are the set of all polynomials of

---

Presented to the Society, September 5, 1947; received by the editors March 19, 1949 and, in revised form, May 1, 1950.

(<sup>1</sup>) Numbers in bracket refer to the bibliography at the end of the paper.

degree not greater than  $n-1$  on a fixed closed interval and the set of all linear combinations of either  $1, \sin x, \sin 2x, \dots, \sin (n-1)x$  on  $a \leq x \leq b$  where  $0 < a < b < \pi$  or  $1, \sin x, \cos x, \sin 2x, \cos 2x, \dots$ , and so on, with  $n$  such functions and  $0 < a < b < 2\pi$ . These families are in particular *linear* because there exist  $n$  functions  $f_1, \dots, f_n$  in terms of which every  $f$  may be expressed as a linear combination:  $f = a_1 f_1 + \dots + a_n f_n$ . Popoviciu [8] has discussed them.

**THEOREM 1.** *If an  $n$ -parameter family  $F$  is closed under addition, then it is linear.*

**Proof.** Choose  $n$  vertical axes  $x = x_1, \dots, x = x_n$  where  $x_1, \dots, x_n$  are distinct values on  $(a, b)$ . Any member  $f$  of  $F$  is uniquely described by the  $n$  values  $(f(x_1), \dots, f(x_n))$ ; we represent  $f$  by this vector. This correspondence preserves addition. Since the set of  $n$ -dimensional vectors forms a group under addition and has a basis of  $n$  vectors, by the isomorphism the same must be true of  $F$ .

The example given by Beckenbach and Bing of a 2-parameter family which is not topologically equivalent to the family of linear polynomials may be modified slightly so that all the functions are analytic; for example, for the member  $f(x)$  determined by  $(x_1, y_1), (x_2, y_2)$  take

$$\begin{aligned} f(x) &= cx + d & \text{if } (y_1 - y_2)/(x_1 - x_2) \leq 0, \\ f(x) &= e^{x+c} + d & \text{otherwise,} \end{aligned}$$

with suitably chosen  $c, d$ .

However 1-parameter families are topologically equivalent to the set of constant functions  $y = c$ . A 1-1 correspondence is attained by setting  $c = f(a)$  for a given  $f$  in  $F$ . An application of Theorem 5 below completes the proof.

The next definition is reminiscent of an unsatisfactory definition of being tangent.

**DEFINITION 2.** A function  $g(x)$  is said to *graze*  $h(x)$  at  $x = x_0$  if (i)  $g(x_0) = h(x_0)$  and (ii) there is a positive  $\epsilon$  such that for  $a \leq x_0 - \epsilon < x < x_0 + \epsilon \leq b$ ,  $g(x) - h(x)$  does not change sign.

**DEFINITION 3.** A function  $g(x)$  is said to be *convex* with respect to a given  $n$ -parameter family  $F$  if  $g(x)$  is real-valued and continuous on  $(a, b)$  and intersects no member of  $F$  more than  $n$  times, multiplicities *not* being reckoned.

**THEOREM 2.** *If  $g$  is convex and has  $n$  intersections with an  $f$  of  $F$ , then  $g$  does not graze  $f$  anywhere for  $a < x < b$ .*

This result was proved by Beckenbach for  $n=2$ . We shall prove it for  $n=1$  and then reduce the general case to this one. Suppose  $g$  grazes  $f$  at  $x'$ . Then  $g-f$  always has the same sign and is zero only for  $x = x'$ ; assume  $g-f \geq 0$ . Furthermore  $x' \neq a$  or  $b$  by Definition 2. Let  $f_1$  and  $f_2$  be members of  $F$ ,  $f_1$  intersecting  $g$  at  $x = a$ , and  $f_2$  intersecting  $g$  at  $x = b$ . Now  $f$  and  $f_1$ , being

members of a 1-parameter family, never intersect. Thus  $f_1 - f$  is always of the same sign, which is seen to be positive by taking  $x = a$ . Now  $g$  and  $f_1$ , having a point in common, cannot intersect again. Since  $f_1 = g$  at an end point,  $f_1 - g$  always has the same sign, which is indeed positive by taking  $x = x'$ . Similarly  $f_2 - g \geq 0$ . Now  $f_1 - f_2 = (f_1 - g) - (f_2 - g)$  and is negative for  $x = a$  and positive for  $x = b$ . Being continuous, it must be zero somewhere on  $(a, b)$ . But this is a contradiction to the fact that  $f_1$ , being distinct from  $f_2$ , cannot intersect it.

We next consider the general case. Suppose  $g$  grazes  $f$  at  $x = x_1$ . Let  $(a', b')$  be a closed interval with  $x_1$  in its interior and containing abscissas of none of the other  $n - 1$  points of intersection. If we take only those functions  $f$  in  $F$  which pass through the other  $n - 1$  points of intersection and restrict them to  $(a', b')$ , we obtain a 1-parameter family  $F'$  in which  $g$  restricted to  $(a', b')$  is a convex function  $g'$ . But then  $g'$  cannot graze any function of  $F'$ ; hence  $g$  does not graze  $f$  at  $x_1$ .

**COROLLARY.** *If  $g$  is convex and has  $n - 1$  intersections with an  $f$  of  $F$  but not at  $x = a$  or  $b$ , and if  $\text{sign } \{g(a) - f(a)\} = (-1)^{n+1} \text{sign } \{g(b) - f(b)\}$ , then  $g$  and  $f$  intersect exactly  $n - 1$  times.*

**Proof.** There could be at most one more intersection and if this occurred  $g - f$  would change sign at each intersection according to the theorem so that  $\text{sign } \{g(a) - f(a)\} = (-1)^n \text{sign } \{g(b) - f(b)\}$ . This is a contradiction to the hypothesis of the corollary.

**THEOREM 3.** *If  $f_1$  and  $f_2$  are distinct members of  $F$  and intersect  $n - 1$  times, then  $f_1$  does not graze  $f_2$  anywhere on  $a < x < b$ .*

This is a consequence of the previous theorem. Suppose  $f_1$  grazes  $f_2$  at  $x_1$ . Let  $(a', b')$  be a closed interval having  $x_1$  in its interior and containing none of the abscissas of the other  $n - 2$  points of intersection. Choose  $x'$  not in  $(a', b')$  and with  $f_1(x') \neq f_2(x')$ . Let  $F'$  be the 1-parameter family of those functions of  $F$  passing through the other  $n - 2$  points of intersection and also through  $(x', f_2(x'))$ , restricted to the interval  $(a', b')$ . Then  $f_1$  restricted to  $(a', b')$  is convex with respect to  $F'$  and intersects  $f_2$  on  $(a', b')$  at one point. By the previous theorem  $f_1$  cannot graze  $f_2$  at  $x_1$ .

**COROLLARY.** *If  $f_1$  and  $f_2$  are in  $F$ , and intersect at least  $n - 2$  times but not for  $x = a$  or  $b$ , and if  $\text{sign } \{f_1(a) - f_2(a)\} = (-1)^n \text{sign } \{f_1(b) - f_2(b)\}$ , then  $f_1$  intersects  $f_2$  exactly  $n - 2$  times.*

A proof of this uses this same reasoning as was used for the corollary of the preceding theorem.

We have inserted the following theorem here, even though its proof depends on Theorem 5 below, because it is closely related to the material of this section.

**THEOREM 4.** *If  $g$  is convex and if  $f$  in  $F$  intersects  $g$  in  $n$  points  $x = x_1, \dots, x_n$*

$(x_1 < \cdots < x_n)$ , the sign of  $f(x) - g(x)$  on  $x_1 < x < x_2$  is independent of the choice of  $f$ .

If  $n=1$  the theorem is to be interpreted as stating that the sign of  $f-g$  for  $x_1 < x \leq b$  is independent of the choice of  $f$ .

**Proof.** Let  $x_1, \dots, x_n$  vary continuously in  $(a, b)$ , only maintaining that  $x_1 < \cdots < x_n$ , and let  $f(x)$  vary accordingly so that  $f(x) = g(x)$  at  $x_1, \dots, x_n$ . Since  $f(x)$  maintains its maximum number of variations in sign (Theorem 2), by continuity (Theorem 5) these variations must remain in the same order.

As a consequence of Theorem 4 it follows that convex functions as defined here are concave or convex in the sense of Popoviciu [8]. The converse is obvious.

**2. Convergence.** The set  $F$  may be easily metrized by choosing  $n$  distinct  $x_i$  ( $i=1, \dots, n$ ) on the interval  $(a, b)$  and defining the distance from  $f_1$  to  $f_2$  as

$$(f_1, f_2) = \max_i |f_1(x_i) - f_2(x_i)|.$$

That  $(f_1, f_2) \geq 0$  is obvious from its definition; also  $(f_1, f_2) = (f_2, f_1)$ . That  $(f_1, f_2) = 0$  if and only if  $f_1 = f_2$  follows immediately from the definition of  $F$ . Finally

$$(f_1, f_3) \leq (f_1, f_2) + (f_2, f_3).$$

It is easily seen that the space  $F$  is complete with respect to this metric. Furthermore the functions converge point-wise. Indeed the convergence is uniform as a consequence of the next theorem.

**THEOREM 5.** If  $2n$  sequences  $(x_{1k}), (y_{1k}), \dots, (x_{nk}), (y_{nk})$  converge to  $x_1, y_1, \dots, x_n, y_n$  respectively, where all the  $x$ 's are on the interval  $(a, b)$  and if  $x_{ik} \neq x_{jk}$ ,  $x_i \neq x_j$  ( $i \neq j$ ), then the sequence of functions  $f_k$  in  $F$  determined by  $(x_{ik}, y_{ik})$  ( $i=1, \dots, n$ ) converges uniformly to the function  $f$  in  $F$  passing through  $(x_1, y_1), \dots, (x_n, y_n)$ .

The proof is indirect. If the sequence  $\{f_i\}$  does not converge uniformly to  $f$ , then there is a positive  $\epsilon$ , an infinite subsequence  $\{f_{k_i}\}$ , and a sequence  $\{\xi_{k_i}\}$  for which  $|f_{k_i}(\xi_{k_i}) - f(\xi_{k_i})| > \epsilon$ . Since the  $\xi_{k_i}$  are bounded, they have an accumulation point  $\xi$  to which a subsequence  $\{\xi_{k_j}\}$  of  $\xi_{k_i}$  converges and we may also require that  $f_{k_j}(\xi_{k_j}) - f(\xi_{k_j})$  always have the same sign. Let  $\lambda$  be 1 or  $-1$  taking that same sign. Assume that  $x_1 < x_2 < \cdots < x_n$ , and for simplicity of notation we denote the sequence  $\{f_{k_j}\}$  by  $\{f_j\}$  since we no longer refer to the original sequence  $\{f_i\}$ . Similarly the notation  $\{\xi_{k_j}\}$  will be replaced by  $\{\xi_j\}$ .

Case:  $x_r < \xi < x_{r+1}$  for some  $r$  such that  $1 < r < n-1$ , and  $x_n \neq b$ . Set  $x'_j = (x_j + x_{j+1})/2$  ( $j \neq r, n$ ),  $x'_r = \xi$ ,  $x'_n = (x_n + b)/2$ . Let  $f'$  in  $F$  be determined by the  $n$  points

$$(x'_i, f(x'_i)) \quad (i \neq r), \quad (x'_r, f(x'_r) + \lambda\epsilon/2).$$

We shall find an  $f_j$  distinct from  $f'$  having  $n$  intersections with  $f'$ , and this is a contradiction.

Now  $f'$  intersects  $f$  at the  $n-1$  points where  $x=x'_i$  ( $i \neq r$ ), and since  $x'_i < x_{i+1} < x'_{i+1}$ , by Theorem 2,  $f'(x_i) - f(x_i)$  must alternate in sign as  $i$  successively takes the values from 1 to  $r$  and from  $r+1$  to  $n$ , but it must have the same sign for  $i$  equal to  $r$  and  $r+1$ , the sign being the same as at  $x'_r$ , that is, the same as the sign of  $\lambda$ .

Suppose

$$(1) \quad \eta = \min_i |f'(x_i) - f(x_i)|;$$

then  $\eta > 0$  since  $f$  does not intersect  $f'$  at any  $x_i$ . Because of the continuity of  $f'$  there is a positive  $\delta$  such that

$$(2) \quad |f'(x) - f'(x_i)| < \eta/2 \quad \text{if} \quad |x - x_i| < \delta.$$

Under the hypothesis of the theorem for sufficiently large  $m$

$$(3) \quad |x_i - x_{im}| < \delta$$

and

$$(4) \quad |f(x_i) - f_m(x_{im})| < \eta/2.$$

Suppose  $f'(x_i) - f(x_i)$  is positive; if negative a similar discussion occurs. By (1), (2), and (3),  $f'(x_{im}) > [f'(x_i) + f(x_i)]/2$ ; whereas by (1) and (4),  $f_m(x_{im}) < [f'(x_i) + f(x_i)]/2$ . Hence  $f'(x_{im}) - f_m(x_{im})$  has the same sign as  $f'(x_i) - f(x_i)$ , the conclusion which is also true if  $f'(x_i) - f(x_i)$  were negative. Because this last quantity alternates in sign for  $i=1, \dots, r$  and also for  $i=r+1, \dots, n$ , there will be the same alternations in sign for  $f'(x_{im}) - f_m(x_{im})$  and hence at least  $n-2$  intersections of  $f'$  and  $f_m$  inasmuch as the functions are continuous.

Finally we shall prove the existence of two more intersections on the interval  $x_r < x < x_{r+1}$ . Since  $f'$  and  $f$  are continuous there is a positive  $\delta$  such that for  $x$  in the interval  $x_r < \xi - \delta < x < \xi + \delta < x_{r+1}$ ,  $0 < |f(x) - f'(x)| < 3\epsilon/4$ , the value of  $|f(\xi) - f'(\xi)|$  being  $\epsilon/2$ . For  $m$  sufficiently large,  $|\xi_m - \xi| < \delta$ . Because  $|f(\xi_m) - f'(\xi_m)| < |f(\xi_m) - f_m(\xi_m)|$ ,  $f(\xi_m)$  and  $f'(\xi_m)$  are both larger or both smaller than  $f_m(\xi_m)$  or, in other words,  $f'(\xi_m) - f_m(\xi_m)$  has the same sign as  $f(\xi_m) - f_m(\xi_m)$ , this sign being the same as that of  $-\lambda$  from the definition of  $\lambda$ . But it was shown in the preceding paragraph that  $f'(x_{im}) - f_m(x_{im})$  has the same sign as  $f'(x_i) - f(x_i)$ . This sign is the same as that of  $\lambda$ , as was observed at the beginning of the proof of this case. Thus the sign of  $f'(x) - f_m(x)$  alternates as  $x$  goes from  $x_{rm}$  to  $\xi_m$  to  $x_{r+1,m}$ . For sufficiently small  $\delta$ ,  $x_{rm} < \xi_m < x_{r+1,m}$ , and  $f'$  and  $f_m$  have at least two more intersections.

For the other cases slight modifications of this proof are used. If instead of  $x_n \neq b$  we had  $x_1 \neq a$ , the same proof applies after first reflecting with re-

spect to the  $y$ -axis. If however  $x_1 = a$  and  $x_n = b$  we proceed much as before. But we first introduce two auxiliary functions in  $F$ :  $f'_1$  determined the same as the former  $f'$ , and  $f'_2$  determined by the same  $n$  points except that  $(a, f(a))$  is used instead of  $(x'_n, f(x'_n))$ . Notice that  $x'_n = b$ . Since these two functions intersect in  $n-1$  points,  $f'_1 - f'_2$  must alternate in sign from one intersection to the next by Theorem 3. Another function  $f'_3$  in  $F$  determined by the  $n-1$  intersections of  $f'_1$  and  $f'_2$  and the point  $(a, [f'_1(a) + f'_2(a)]/2)$  will always lie between  $f'_1$  and  $f'_2$  also by Theorem 3. This  $f'_3$  has properties that the function  $f'$  had which are needed for the proof in the first case; namely that  $f'(x_i) - f(x_i)$  alternates sign except for  $i = r$  and  $r+1$ . This can be seen by considering  $f'_1(x_i) - f(x_i)$  and  $f'_2(x_i) - f(x_i)$ . These have the same sign (or are zero) for fixed  $i$  and alternate sign for  $i = 1, \dots, r$  and again for  $i = r+1, \dots, n$ . But  $f'_3(x_i) - f(x_i)$  must do likewise because  $f'_3$  always lies between  $f'_1$  and  $f'_2$ .

If originally  $\xi > x_n$ , the proof is exactly the same as in the first case but with  $r = n$ . Then  $n-1$  intersections are obtained on the intervals  $x_i < x < x_{i+1}$  ( $i = 1, \dots, n-1$ ) and another intersection for  $x_n < x < b$ .

If  $\xi < x_1$ , a reflection with respect to the  $y$ -axis transforms this case into that of the paragraph above.

We next consider the changes in the original proof if  $\xi$  equals  $x_r$ .

If  $\xi \neq x_n$  or  $x_1$ , then we may take the sequence  $\xi_j$  such that  $\xi_j - \xi$  always has the same sign; assume it to be positive. Let  $f'$  in  $F$  be determined by  $(x_i, f(x_i) \pm \lambda\epsilon/2)$  ( $i = 1, \dots, n$ ) where the ambiguous sign is chosen  $+$  for  $i = r$  and alternates sign for  $i \leq r$ , while it is taken  $+$  for  $i = r+1$  and alternates sign for  $i \geq r+1$ . Then the same type of reasoning as used originally shows there is at least one intersection of  $f'$  and  $f_m$  (for sufficiently large  $m$ ) on each of the  $n-2$  intervals  $x_i < x < x_{i+1}$  ( $i \neq r$ ). Two more intersections are obtained much as before on the interval  $x_r < x < x_{r+1}$  by the use of the smaller interval  $x_r < x < x_r + \delta$ .

If  $\xi = x_n$ , we proceed as above providing we may take the sequence  $\{\xi_j\}$  such that  $\xi_j < x_n$ . Otherwise we model the proof after the one for the case  $\xi > x_n$ . Determine  $f'$  in  $F$  by the points  $(x_i, f(x_i) + (-1)^{n-i}\lambda\epsilon/2)$  ( $i = 1, \dots, n$ ). The alternations in sign from one  $x_i$  to the next will provide  $n-1$  instead of  $n-2$  intersections with the  $f_m$  as in the case  $\xi < x_n$ . Another intersection is obtained because  $f_m(\xi_m) - f'(\xi_m)$  and  $f_m(x_n) - f'(x_n)$  have opposite signs.

**3. Derivatives of convex functions.** Peixoto [6] has given several results about the existence of derivatives for 2-parameter families. We shall prove the following theorem.

**THEOREM 6.** *If all the functions in an  $n$ -parameter family  $F$  ( $n \geq 3$ ) have derivatives on  $a < x < b$ , then so does any function  $g$  convex with respect to  $F$ .*

Choose  $x_0$  on  $a < x < b$ . Then there is a positive  $h$  such that  $a < x_0 - h < x_0 + 2h < b$ . We reduce to the situation of a 3-parameter family by first choosing

$n-3$  points  $x_4, \dots, x_n$  on the open interval  $(x_0+2h, b)$ . Those members of  $F$  passing through  $(x_i, g(x_i))$  ( $i=4, \dots, n$ ) form a 3-parameter family  $F'$  on  $a \leq x \leq x_0+2h$  and  $g$  restricted to that interval is a function  $g'$  convex with respect to  $F'$ .

Start with  $f'_0$  in  $F'$  determined by having it intersect  $g'$  at  $x=x_0-h, x_0, x_0+h$ . It is essentially no restriction to assume  $f'_0 - g' > 0$  on the interval  $(x_0-h, x_0)$ . Consider all  $f'_1$  in  $F'$  which intersect  $g'$  at  $x=x_0, x_0+h$ , and a value  $x'_0$  between  $x_0$  and  $x_0+h$ . Then by Theorem 4,  $f'_1 - g' < 0$  for  $x < x_0$ ; and consequently  $f'_1 - f'_0 < 0$  at  $x=x_0-h$ . Let  $y_1$  be the least upper bound of  $f'_1(x_0-h)$ . This bound is not attained, for if  $f'_1$ , corresponding to such an upper bound, intersected  $g$  at  $x=x_0, x'_0, x_0+h$ , then the function in  $F'$  determined by intersections with  $g$  at  $x=x_0, x_0+x'_0/2, x_0+h$  lies below  $f'_1$  for  $x_0 < x < x'_0$  and must lie above  $f'_1$  for  $x < x_0$  by Theorem 3 so that  $y_1$  would not have been an upper bound. Let  $f_1$  be that function in  $F'$  determined by

$$(x_0-h, y_1), \quad (x_0, g(x_0)), \quad (x_0+h, g(x_0+h)).$$

It does not intersect  $g$  for  $x_0 < x < x_0+h$  by the argument above, nor for  $x > x_0+h$  because there it is above any  $f'_1$  by Theorem 3. It does not intersect  $g$  for  $x < x_0$  or for  $x > x_0+h$  by Theorem 5 since  $f_1$  is the limit of functions which do not intersect  $g$  there.

Also  $f_1 \leq g$  when  $x_0-h < x < x_0+h$ . For,  $f_1 < g$  at  $x=x_0-h$  where  $g=f'_0$ , and hence  $f_1 < g$  when  $x < x_0$ . Next  $f_1 > f'_1$  when  $x < x_0$ , this being true at  $x_0-h$ ; by Theorem 3,  $f_1 < f'_1$  when  $x_0 < x < x_0+h$ . But  $f'_1(x'_0) = g(x'_0)$ ; hence  $f_1 < g$  at  $x=x'_0$  and consequently for  $x_0 < x < x_0+h$ .

By the same process we can find an  $f_2$  in  $F'$  intersecting  $g$  just at  $x=x_0-h$  and  $x_0$  and such that  $f_2 \geq g$  for  $x_0-h < x < x_0+h$ . Because the derivatives of  $f_1$  and  $f_2$  exist and  $f_2$  grazes  $f_1$  at  $x_0$ , we conclude  $df_2/dx = df_1/dx$  at  $x=x_0$ . Finally  $f_2(x') - f_2(x_0) \geq g(x') - g(x_0) \geq f_1(x') - f'_1(x_0)$  for  $x_0-h < x' < x_0+h$ . On dividing by  $x'-x_0$  and letting  $x' \rightarrow x_0$ , we finish the proof of the theorem.

**4. Approximation of functions.** Within the set  $F$  we have a very simple metric whose determination in a particular case is a finite process. But for the topic of approximation we introduce the metric commonly used for continuous functions; namely, if  $g_1, g_2$  are continuous functions on  $a \leq x \leq b$ , the distance between  $g_1$  and  $g_2$  is

$$\|g_1 - g_2\| = \max_x |g_1(x) - g_2(x)|.$$

In  $F$  the two metrics are equivalent by Theorem 5.

**DEFINITION 4.** A best approximant in  $F$  of a continuous function  $g$  is a function  $f'$  in  $F$  for which

$$\|f' - g\| = \text{g.l.b.}_{f' \in F} \|f - g\|$$

and this value is called the modulus of approximation of  $g$  in  $F$ .

THEOREM 7. *For any function  $g$ , a best approximant in  $F$  exists.*

For let  $k$  be the modulus of approximation. Then there is a sequence  $f_m$  such that  $\lim_{m \rightarrow \infty} \|f_m - g\| = k$ . Choose  $n$  distinct values  $x_1, \dots, x_n$  on the interval  $(a, b)$ . There is a one-to-one correspondence between the functions  $f$  in  $F$  and the sets of values  $y_1, \dots, y_n$  taken by the functions  $f$  at  $x = x_1, \dots, x_n$ . Let  $f_m(x_i) = y_{mi}$ . Then there is a subsequence  $f_{m_j}$  of  $f_m$  for which  $y_{m_j i}$  is convergent ( $i = 1, \dots, n$ ), converging to  $y'_i$ . Let  $f'$  correspond to the set  $y'_1, \dots, y'_n$ ; by Theorem 5,  $\|f' - g\| = k$ .

THEOREM 8. *If  $f$  is a best approximant in  $F$  of a continuous function  $g$ , then  $f - g$  takes the value  $\pm \|f - g\|$  at a set of  $n + 1$  points and with opposite signs for consecutive points of the set.*

Let  $k = \|f - g\|$ . Let  $\xi_1$  be the smallest value of  $x$  for which  $|f - g| = k$ , a smallest value existing because  $f - g$  is continuous. This  $\xi_1$  is the beginning of an increasing sequence of values  $\xi_1, \xi_2, \dots$  such that  $\xi_{i+1}$  is the smallest  $x$  larger than  $\xi_i$  for which  $f(x) - g(x) = -[f(\xi_i) - g(\xi_i)]$ . Suppose that the conclusion of the theorem is not satisfied so that the sequence ends with  $\xi_m$ ,  $m \leq n$ . Let  $x_i$  ( $i = 1, \dots, m - 1$ ) be the largest value of  $x$  smaller than  $\xi_{i+1}$  at which  $f - g = 0$ . These values also exist because  $f - g$  is continuous. Also  $\xi_i < x_i$  because at  $\xi_i$  the sign of  $f - g$  is opposite to its sign at  $\xi_{i+1}$ . Let  $x_0 = a$ ,  $x_m = b$ . Suppose  $(-1)^i [f(\xi_i) - g(\xi_i)] = k$ ; a similar argument applies if the signs are all reversed. Let  $k'_i$  ( $i = 1, \dots, m - 1$ ) be the maximum of  $(-1)^{i+1}(f - g)$  on the interval  $(x_{i-1}, x_i)$  and  $k' = \max k'_i$ . By our construction  $0 \leq k' < k$ , for the smallest value of  $x$  larger than  $x_{i-1}$  at which  $(-1)^{i+1}(f - g) = k$  is  $\xi_{i+1}$ , and  $\xi_{i+1} > x_i$ . Let  $\epsilon = (k - k')/2$ . Since  $f(x_1) - g(x_1) = 0$ , there is an interval  $x_1 < x < x_1 + \delta < \xi_2$  on which  $|f(x) - g(x)| < k'$ .

If  $n - m$  is even, choose distinct points  $x'_1, \dots, x'_{n-m}$  on the open interval  $(x_1, x_1 + \delta)$ , and finally take  $x'' = x_1 + \delta$ . Determine  $f'$  in  $F$  so that it intersects  $f$  at  $x = x_1, \dots, x_{m-1}, x'_1, \dots, x'_{n-m}$  and passes through  $(x'', f(x'') - \eta)$ , where  $\eta$  is positive and chosen sufficiently small by Theorem 5 so that  $\|f - f'\| < \epsilon$ . Since  $f$  and  $f'$  have  $n - 1$  intersections, they cross at each intersection and nowhere else. By construction the sign of  $f(\xi_2) - g(\xi_2)$  is positive while that of  $f'(x'') - f(x'')$  is negative and  $\xi_2$  is on the interval  $(x'', x_2)$ . Thus the sign of  $f(\xi_i) - g(\xi_i)$  is opposite to that of  $f' - f$  on  $(x_{i-1}, x_i)$  ( $i = 3, \dots, m$ ).

For a moment let us discuss values on the interval  $(x_2, x_3)$ ; here  $f < f'$  since  $f(\xi_3) - g(\xi_3) < 0$ . At a point where  $f < f' < g$ ,  $|f' - g| < |f - g| \leq k$ . If  $f < g < f'$ , then  $|f' - g| < |f' - f| < \epsilon < k$ . If  $g < f$ , then  $|f' - g| < |f' - f| + |f - g| < \epsilon + k' < k$ . Thus on  $(x_1, x_2)$  always  $|f' - g| < k$ . Much the same argument proves that  $|f' - g| < k$  on the intervals  $(x_{i-1}, x_i)$  ( $i = 4, \dots, m$ ) and on  $(x'', x_2)$ , and also on  $(x_0, x_1)$  since there are an odd number of intersections of  $f$  and  $f'$  from  $x = x_2$  to  $x = x_1$ . On the interval  $(x_1, x'')$ ,  $|f' - g| < |f' - f| + |f - g| < \epsilon + k' < k$ .



Thus everywhere  $|f' - g| < |f - g|$ ;  $\|f' - g\| < \|f - g\|$  since all these functions are continuous; and  $f$  cannot be a best approximant.

If however  $n - m$  is odd then choose just  $n - m - 1$  points  $x'_1, \dots, x'_{n-m-1}$  as before. If  $|f(a) - g(a)| \neq \|f - g\|$ , let  $x'_{n-m-1} = a$ ; or if  $|f(b) - g(b)| \neq \|f - g\|$  we may take  $x'_{n-m-1} = b$ . Then choose  $x''$  as before and the proof goes through much the same. But if  $a = \xi_1$ ,  $b = \xi_m$ , then let  $y_a = f(a) - \eta[f(a) - g(a)]$ ,  $y_b = f(b) - \eta[f(b) - g(b)]$ , where  $\eta > 0$  and sufficiently small that  $\|f - f'\| < \epsilon$ . Here  $f'$  is determined by the points of  $f$  at  $x = x_1, \dots, x_{m-1}, x'_1, \dots, x'_{n-m-1}$ , and by  $(a, y_a)$ ,  $(b, y_b)$ . Thus  $f'$  has  $n - 2$  intersections with  $f$ . It has no more by the corollary of Theorem 3. Since the intersections are precisely  $x_1, \dots, x_{m-1}$  and an even number  $x'_1, \dots, x'_{n-m-1}$ , the former proof may now be carried out.

**COROLLARY 1.** *A necessary and sufficient condition that  $f$  in  $F$  be a best approximant of  $g$  is that  $f - g = \pm \|f - g\|$  at  $n + 1$  values of  $x = x_1, \dots, x_{n+1}$  and with opposite signs for consecutive  $x_i$ .*

Let  $f_1$  be a best approximant, and assume  $f$  is not. Then at the values of  $x$  where  $f - g = \pm \|f - g\|$ ,  $f - f_1$  is not zero and takes the sign of  $f - g$ . Thus  $f - f_1$  changes sign  $n$  times and consequently  $f$  and  $f_1$  intersect  $n$  times. This is a contradiction to  $f$  and  $f_1$  being distinct.

It is not true that  $\pm \|f - g\|$  must be attained at  $a$  and at  $b$ ; let  $F$  be the 3-parameter family of parabolas and straight lines and let  $g$  be a sine curve having several cycles between  $a$  and  $b$  and not taking an extreme value at  $a$  or at  $b$ .

**COROLLARY 2.** *If  $f - g$  equals  $m_1, \dots, m_{n+1}$  at the successive values  $x_1, \dots, x_{n+1}$  and if successive  $m_i$  have opposite signs, then the modulus of approximation is not less than the minimum of the  $m_i$  nor more than their maximum.*

For the proof much the same reasoning applies as was used to justify Corollary 1.

**THEOREM 9.** *If  $g$  is continuous there is exactly one best approximant in  $F$  of  $g$ .*

Let  $f$  and  $h$  be best approximants in  $F$  of  $g$ . Let  $f - g$  take the values  $\pm \|f - g\|$  at points with  $x = x_1, \dots, x_{n+1}$ , where the sign alternates for adjacent points and  $x_1 < \dots < x_{n+1}$ ; similarly for  $h$  and points at  $x = x'_1, \dots, x'_{n+1}$ . We may assume that if  $x_i$  is on an interval where  $|f - g| = \|f - g\|$ , then  $x_i$  is the least value of  $x$  on that interval. Arrange the  $x_i, x'_i$  ( $i = 1, \dots, n + 1$ ) in order, obtaining a set  $\xi_1, \dots, \xi_{2n+2}$ . By a slight deformation which does not change the position of any intersections or introduce additional intersections, we may assume  $\xi_i < \xi_{i+1}$ . We consider the possibility of an intersection of  $f$  and  $h$  on the interval  $\xi_i < x \leq \xi_{i+1}$  and we would like to average two values of  $\xi$  for every intersection in order to obtain  $n$  intersections. The

following cases are typical of all possible cases; here  $f^+$  means  $f-g=||f-g||$  and  $f^-$  that  $f-g=-||f-g||$ .

Case	I	II	III
$\xi_i$	$f^+$	$f^+$	$f^+$
$\xi_{i+1}$	$f^-$	$h^+$	$h^-$
min. no. of intersections	0	1	0

We need consider in detail only those cases where the number of intersections is 0, namely, I and III. Now I gives 0 intersections only if  $h(\xi_i)=f(\xi_i)$  and  $h>f$  for  $\xi_i<x\leq\xi_{i+1}$ . At  $\xi_{i+2}$  we may have  $f^+$ ,  $h^-$ , or  $h^+$ . In the first two then there would be an intersection for  $\xi_{i+1}<x\leq\xi_{i+2}$ , and the last by itself is not possible because since we have  $h^+$  at  $\xi_i$  and not on  $\xi_i<x\leq\xi_{i+2}$ , we cannot have  $h^+$  at  $\xi_{i+2}$ . Thus we have used only two  $\xi$ 's and obtained an intersection.

For case III at  $\xi_{i+2}$  we may have  $h^+$  or  $f^-$  and both of these give an intersection.

In Theorem 9 the hypothesis that  $g$  be continuous is not superfluous; for example, let  $F$  be the 2-parameter family of linear functions on the interval  $(-1, 1)$  and let  $g$  be  $-1$  when  $-1\leq x<0$  and  $+1$  when  $0\leq x\leq 1$ . Then all the lines  $y=mx$ ,  $0\leq m\leq 2$ , are best approximants. Theorem 7 is true if  $g$  is bounded even if not continuous, after redefining  $||f-g||$  of course as l.u.b.  $|f(x)-g(x)|$ , the proof being as before. Theorems 8 and 9 are still true if  $g$  is merely bounded when  $F$  is a 1-parameter family. To prove the former, if  $f$  is in  $F$ , let l.u.b.  $(f-g)=k$ , g.l.b.  $(f-g)=-k'$ , and  $k\neq k'$ . Assume  $k-k'$  is positive and let it equal  $2\epsilon$ . By Theorem 5 there is an  $f'$  in  $F$  such that  $0<f-f'<\epsilon$ . Since  $f$  and  $f'$  are continuous and do not intersect,  $f-f'>\eta$  where  $\eta>0$ . Then  $f'-g<f-g-\eta\leq k-\eta$ , while  $f'-g=(f'-f)+(f-g)\leq -\epsilon-k'$  so that  $f'$  approximates  $g$  better than  $f$ . Next to prove Theorem 9 without the assumption that  $g$  is continuous, we let  $f$  and  $f'$  be two best approximants in  $F$  and we may assume  $f'-f>\epsilon>0$  on  $(a, b)$  since  $f$  and  $f'$  are continuous and do not intersect. Choose  $x_1$  such that  $f(x_1)-g(x_1)>||f-g||-\epsilon$ . Now  $f'(x_1)-g(x_1)>f(x_1)-g(x_1)+\epsilon>||f-g||=||f'-g||$ , and this is impossible by the definition of  $||f'-g||$ .

#### BIBLIOGRAPHY

1. E. F. Beckenbach, *Generalized convex functions*, Bull. Amer. Math. Soc. vol. 43 (1937) pp. 363-371.
2. ———, *Convex functions*, Bull. Amer. Math. Soc. vol. 54 (1948) pp. 439-460, especially pp. 444-445. It contains additional references.
3. E. F. Beckenbach and R. H. Bing, *On generalized convex functions*, Trans. Amer. Math. Soc. vol. 58 (1945) pp. 220-230.
4. S. Bernstein, *Leçons sur les propriétés extrémales et la meilleure approximation des fonctions analytiques d'une variables réelle*, 1926.

5. C. J. de la Vallée Poussin, *Leçons sur l'approximations des fonctions d'une variable réelle*, 1919.
6. M. M. Peixoto, *On the existence of derivative of generalized convex functions*, Summa Brasiliensis Mathematicae vol. 2 (1948) pp. 35-42.
7. T. Popoviciu, *Sur quelques propriétés des fonctions d'une ou de deux variables réelles*, Mathematica vol. 8 (1934) pp. 1-85.
8. ———, *Notes sur les fonctions convexes d'ordre supérieur* (1), Mathematica vol. 12 (1936) pp. 81-92.
9. ———, *Les fonctions convexes*, 1945.

UNIVERSITY OF MICHIGAN,  
ANN ARBOR, MICH.