STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

Report S 356 (VP 26)

ON MONOTONICITY OF HIGHER TYPE

Preliminary Report

by

W. Molenaar

december 1965

Printed at the Mathematical Centre at Amsterdam,49,2nd Boerhaavestraat.
The Netherlands.

The Mathematical Centre, founded the 11th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications, and is sponsored by the Netherlands Government through the Netherlands Organization for Pure Scientific Research (Z.W.O.) and the Central National Council for Applied Scientific Research in the Netherlands (T.N.O.), by the Municipality of Amsterdam and by several industries.

1. Summary

A real-valued function f is called I_k if its k-th order differences are nonnegative regardless of the choice of the choice of the k+1 equidistant points involved. It is shown that measurability, boundedness except near endpoints, and continuity are equivalent properties for I_k functions ($k \geq 2$) defined on an interval. Moreover, continuity and the I_k property together are equivalent to convexity of order k as defined by KARLIN. The results are well known for k=2 (convex functions).

This report is the first of two preliminary reports discussing properties of I_k functions that will be needed in research now in progress at the Mathematisch Centrum. The second one [5] will appear shortly.

2. Definitions

Throughout this paper, f will be a function with finite real values defined on an open interval (a,b), where $a = -\infty$ and $b = \infty$ are allowed. The letter h will always denote a positive real number, and k a positive integer. "Measurable" stands for Lebesgue measurable.

Definition 1. For all $x \in (a,b-kh)$

(2.1) $$\Delta_h^k \; f(x) \stackrel{def}{=} \Delta_h^{k-1} \; f(x+h) - \Delta_h^{k-1} \; f(x)$$

and for all $x \in (a,b)$ $\Delta_h^0 \; f(x) \stackrel{def}{=} f(x)$.

Clearly this implies

(2.2.) $$\Delta_h^k \; f(x) = \sum_{j=0}^{k} (-)^{k-j} \binom{k}{j} \; f(x + jh).$$

This asymmetric definition of the k-th order difference, where x denotes the lowest argument and not the middle one, is more convenient for our purposes. We note that $\Delta_h^k \, . \, (x)$ is a linear operator on the space of all real valued finite functions.

Definition 2. The function f is called I_k (increasing of type k) if, for all $h > 0$ and all $x \in (a, b - kh)$, we have $\Delta_h^k f(x) \geq 0$. It is called D_k (decreasing of type k) if the reversed inequality holds for all h and x, and M_k (monotone of type k) if it is either I_k or D_k.

Of course I_1 means nondecreasing, while I_2 means convex, in the sense

$$(2.3) \qquad f\left(\frac{x+y}{2}\right) \leq \frac{f(x) + f(y)}{2} .$$

If f is k times differentiable, I_k is equivalent to the nonnegativity of the k-th derivative. It is clear that an I_k function is not necessarily I_j for $j < k$.

It is well known that there exists convex functions which are discontinuous and unbounded on every interval. From any Hamel basis one can construct many functions f such that $f(x+y) = f(x) + f(y)$. Inserting this equality in (2.2), one finds that these functions are not only convex but also I_k for higher k. Unless f is linear they are not I_1, as $\Delta_h^1 f(x) = f(x+h) - f(x) = f(h)$, and f has sign changes when it is not linear (and possibly also when it is linear).

To exclude such pathological examples, one could define convexity by

$$(2.4) \qquad f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \quad \text{for } 0 < \lambda < 1.$$

It is known (see e.g. [2] p. 116-117) that an I_2 function is bounded and continuous, and satisfies (2.4), as soon as it is assumed to be measurable. In section 3 a similar result for I_k functions with $k > 2$ is derived by an adaptation of the proof for k=2. We shall see in section 4 that the continuous (or measurable) I_k functions are precisely KARLIN's convex functions of order k (see definition 4 below).

Definition 3. A function f has k sign changes, if k+1 is the maximal number of points $x_1 < x_2 < \ldots < x_{k+1}$ such that either $f(x_i)$ is positive for odd k-i and negative for even k-i [we call this: k sign changes with a plus ending] or $f(x_i)$ is positive for even k-i and negative for odd k-i [with a minus ending].

Definition 4. A function f is C_k (convex of order k), if for each polynomial p(x) of degree at most k-1, the function f-p has at most k sign changes, and if exactly k then with a plus ending.

Definition 3 is essentially KARLIN's definition of the number $V^-(f)$ of (strong) sign changes. Definition 4 is given by KARLIN and PROSCHAN [3] with a "plus beginning", i.e. $f(x_1)$ positive in our notation; so f is C_k in our sense iff $(-)^k f$ is C_k in the sense of [3] p. 732. In [4] p. 344 one finds definition 4 with the restricition that the leading coefficient of p should be positive. Consideration of the cases k=1,2 (nondecreasing and convex functions) makes it desirable to remove this restriction.

3. Measurability, boundedness and continuity

Theorem 1. A measurable M_k function defined on (a,b) is bounded on every closed subinterval of (a,b).

Proof. As the case k=1 is trivial we shall assume $k \geq 2$. Suppose f were unbounded in a neighbourhood of $x_0 \in (a,b)$; there would exist for $\varepsilon = \min(b-x_0, x_0-a, 1)$ and for every N a point $x_N \in (x_0-\frac{1}{2}\varepsilon, x_0+\frac{1}{2}\varepsilon)$ with $|f(x_N)| > 2^k N$. We shall show that for each $h \leq \frac{\varepsilon}{2k}$, the inequality $|f(x_N-jh)| > N$ holds either for at least one index $j \in \{1,2,\ldots,k\}$ (case A) or for at least one index $j \in \{-1,1,2,\ldots,k-1\}$ (case B).

If f is I_k and $f(x_N) < -2^k N$, it is obvious from

$$0 \leq \Delta_h^k f(x_N-kh) = f(x_N) + \sum_{j=1}^{k} (-)^j \binom{k}{j} f(x_N-jh) <$$

$$(3.1) \qquad < -2^k N + \sum_{j=1}^{k} \binom{k}{j} |f(x_N-jh)|,$$

that we are in case A. The same conclusion is seen to hold if f is D_k

and $f(x_N) > 2^k N$. In the two remaining cases [f is I_k and $f(x_N) > 2^k N$, or f is D_k and $f(x_N) < -2^k N$] we consider $\Delta_h^k f(x_N - (k-1)h)$ and we find in a similar way that we are in case B.

For each y in $J \overset{\text{def}}{=} [x_N - \varepsilon/2, x_N)$ we define

(3.2)
$$Q_0(y) \overset{\text{def}}{=} x_N + k(x_N - y),$$
$$Q_j(y) \overset{\text{def}}{=} x_N - \frac{k}{j}(x_N - y) \qquad (j=1,2,\ldots,k).$$

It is obvious that these mappings preserve the measurability of sets; the relations

(3.3)
$$\mu(Q_0(E)) = k\,\mu(E),$$
$$\mu(Q_j(E)) = \frac{k}{j}\,\mu(E) \qquad (j=1,2,\ldots,k),$$

where μ denotes Lebesgue measure, hold for all intervals E and by the uniqueness of the extension for all measurable sets E.

In case A, if we put $A_N \overset{\text{def}}{=} J \cap \{x \mid |f(x)| > N\}$, then A_N is measurable, and we shall show that

(3.4) $$\bigcup_{j=1}^{k} Q_j(A_N) \supseteq J.$$

In fact for $x \in J$ there exists $j \in \{1,2,\ldots,k\}$ such that $x_N - j(x_N - x)/k \in A_N$, and this implies by (3.2)

(3.5) $$x = Q_j(x_N - j(x_N - x)/k) \in Q_j(A_N).$$

From (3.4) and $\mu(J) = \frac{1}{2}\varepsilon$ follows that $\mu(Q_j(A_N)) \geq \varepsilon/(2k)$ for at least one $j \in \{1,2,\ldots,k\}$, so by (3.3)

(3.6) $$\mu(A_N) \geq \frac{j}{k} \cdot \frac{\varepsilon}{2k} \geq \frac{\varepsilon}{2k^2}.$$

In case B we put $B_N \overset{\text{def}}{=} [x_N - \varepsilon/2,\ x_N + \varepsilon/(2k)] \cap \{x \,|\, |f(x)| > N\}$. Then B_N is measurable and

$$(3.7) \qquad \bigcup_{j=0}^{k-1} Q_j\ (B)_N \supseteq J;$$

for if there is a $j \in \{1, 2, \ldots, k-1\}$ with $x_N - j(x_N - x)/k \in B_N$ then $x \in Q_j(B_N)$, and if $x_N - (-1)(x_N - x)/k \in B_N$ then $x \in Q_0(B_N)$.
As before we find $\mu(B_N) \geq \varepsilon/(2k^2)$.

Now in both cases we have found a set measure at least $\varepsilon/(2k^2)$ which is contained in $C_N \overset{\text{def}}{=} (x_0 - \varepsilon,\ x_0 + \varepsilon) \cap \{x \,|\, |f(x)| > N\}$.
The decreasing sequence $\{C_N\}$, with $\mu(C_N) \geq \varepsilon/(2k^2)$ for all N, converges towards $C \overset{\text{def}}{=} (x_0 - \varepsilon,\ x_0 + \varepsilon) \cap \{x \,|\, |f(x)| = \infty\}$. This leads to a contradiction as f was assumed to have finite values on (a,b).

<u>Theorem 2.</u> If f is bounded on every closed subinterval of (a,b) and f is M_k for some $k \geq 2$, then f is continuous on (a,b).

<u>Proof.</u> From (2.1) we find, by introduction on n, for all positive integer j:

$$(3.8) \qquad \Delta_h^n f(x+jh) = \Delta_h^n f(x) + \sum_{i=0}^{j-1} \Delta_h^{n+1} f(x+ih);$$

$$(3.9) \qquad \Delta_h^n f(x-jh) = \Delta_h^n f(x) - \sum_{i=1}^{j} {}^{n+1}_{h} f(x-ih).$$

The proof is now based on the following lemma.

<u>Lemma.</u> If f is bounded on every closed subinterval of (a,b), and S_m means: $\lim_{h \downarrow 0} \Delta_h^m f(x+ph) = 0$ for all integer $p \geq -1$ and all $x \in (a,b)$, then for $m \geq 1$ S_{m+1} implies S_m.

<u>Proof of the lemma.</u> Let S_{m+1} hold for some $m \geq 1$ and suppose that

$$(3.10) \qquad \limsup_{h \downarrow 0} \Delta_h^m f(y+qh) = \varepsilon > 0$$

for some $\varepsilon > 0$, some $y \in (a,b)$ and some $q \geq -1$.

Choose $c > a$ and $d < b$ such that $y \in (c,d)$, then there exists a constant M such that $|f(x)| < M$ for $x \in [c,d]$. Consequently we have, for all positive integer n, all $h > 0$ and all $x \in [c, d-nh]$

(3.11) $$|\Delta_h^n f(x)| < 2^n M.$$

Choose an integer $N > \max\{\frac{1}{\varepsilon}, 2^{m+1} M\}$ and select

(3.12) $$h < \min \{y-c, \frac{d-y}{N^2+q+m+1}\}$$

such that simultaneously

(3.13) $$\Delta_h^m f(y+qh) > \frac{1}{N}$$

and

(3.14) $$\Delta_h^{m+1} f(y+qh+jh) > -\frac{1}{2N^3} \text{ for } j = 0,1,\ldots,N^2.$$

This is possible because of (3.10) and S_{m+1}. With (3.8) we find for $j=0,1,\ldots,N^2$:

(3.15) $$\Delta_h^m f(y+qh+jh) > \frac{1}{N} - \frac{N^2}{2N^3} = \frac{1}{2N},$$

and with (3.8) and (3.11)

(3.16) $$\Delta_h^{m-1} f(y+qh+N^2h) > -2^{m-1}M + \frac{N^2}{2N} > 2^{m-1} M,$$

which contradicts (3.11). We next derive a contradiction from

(3.17)
$$\liminf_{h\downarrow 0} \Delta_h^m f(y+qh) = -\varepsilon < 0$$

for some $\varepsilon > 0$, some $y \in (a,b)$ and some $q \geq -1$, by rewriting the preceding proof with from (3.13) onward each inequality "$> C$" replaced "$< -C$". As both (3.10) and (3.17) lead to a contradiction, S_m must hold and the lemma is proved.

To prove the theorem, we observe that S_{k-1} follows from the boundedness and the M_k property. For if f is I_k, then (3.10) for $m=k-1$ leads to a contradiction just as in the proof of the lemma: we have only used S_{m+1} in (3.14), and the nonnegativity of the k-th differences is even stronger. If f is D_k and (3.10) would hold for $m=k-1$, then select N as before and $h < (y-c)/(N^2+1)$ such that (3.13) holds. By (3.9), (3.13) and the D_k property we find

$$(3.18) \qquad \Delta_h^{k-1} f(y+qh-jh) > \frac{1}{N} \text{ for } j=0,1,\ldots,N^2,$$

and from (3.9), (3.11) and (3.18) follows

$$(3.19) \qquad \Delta_h^{k-2} f(y+qh-N^2h) > -2^{k-2}M + \frac{N^2}{N} > 2^{k-2}M,$$

again contradicting (3.11). A contradiction is derived from (3.17) for D_k f as in the lemma, and for I_k f because we can derive then (3.19) with $< -2^{k-2}M$ instead of $> 2^{k-2}M$. Thus S_{k-1} must hold.

By repeated application of the lemma we arrive at S_1, i.e. for $p=0$ and $p=-1$

$$(3.20) \qquad \lim_{h\downarrow 0} \Delta_h^1 f(x) = 0 \quad \text{and} \quad \lim_{h\downarrow 0} \Delta_h^1 f(x-h) = 0,$$

which means continuity from the right and from the left.

From theorems 1 and 2 and the fact that continuous functions are measurable, we find:

<u>Theorem 3</u>. For M_k functions defined on an interval, measurability, boundedness on every closed subinterval and continuity are equivalent provided $k \geq 2$.

4. Convexity of order k

Theorem 4. Any C_k function is I_k.

Proof. Let f be C_k (definition 4) and suppose that

$$(4.1) \qquad \Delta_h^k f(c) = -\varepsilon < 0,$$

for some $\varepsilon > 0$, some $h > 0$ and some $c \in (a,b-kh)$. In the polynomial of degree at most k-1

$$(4.2) \qquad \begin{aligned} p(x) &\stackrel{\text{def}}{=} p_0 + \sum_{j=0}^{k-2} p_{j+1} \prod_{m=0}^{j} (x-c-mh) = \\ &= p_0 + p_1(x-c) + p_2(x-c)(x-c-h)+\ldots+p_{k-1}(x-c)\ldots(x-c(k-2)h) \end{aligned}$$

we define the coefficients p_j successively for $j=0,1,\ldots,k-1$ by requiring

$$(4.3) \qquad f(c+jh) - p(c+jh) = (-)^{k-j-1}\, 2^{-k}\, \varepsilon \qquad (j=0,1,\ldots,k-1).$$

For any polynomial $p(x) = \sum_{m=0}^{k-1} a_m x^m$ it is clear that

$$(4.4) \qquad \begin{aligned} \Delta_h^k p(x) &= \sum_{j=0}^{k} (-)^{k-j} \binom{k}{j} \sum_{m=0}^{k-1} a_m (x+jh)^m = \\ &= \sum_{m=0}^{k-1} a_m \sum_{i=0}^{m} \binom{m}{i} x^{m-i} h^i \sum_{j=0}^{k} (-)^{k-j} \binom{k}{j} j^i = 0; \end{aligned}$$

there are several ways to show that the last sum over j is zero for $i=0,1,\ldots,k-1$; see e.g. [1], II. 12 problem 16.

So $-\varepsilon = \Delta_h^k f(c) = \Delta_h^k \{f(c) - p(c)\}$, and we use (2.2) and (4.3) to find

$$(4.5) \qquad f(c+kh) - p(c+kh) = -\varepsilon + \sum_{j=0}^{k-1} \binom{k}{j} 2^{-k} \varepsilon = -2^{-k}\varepsilon.$$

Now f-p has exactly k sign changes, with a minus ending. As this contradicts the C_k property, (4.1) was incorrect and f is I_k.

Theorem 5. A C_k function is continuous if $k \geq 2$.

Proof. Let y be a discontinuity point of a C_k function f, but suppose f is continuous in $(y-c, y)$ and $(y, y+c)$ for some $c > 0$. As C_k implies I_k, this means by theorem 3 that for $0 < \delta < c/2$ there is a number M such that $|f(x)| < M$ if $x \in [y-c+\delta, y-\delta]$ or $x \in [y+\delta, y+c-\delta]$, while f is unbounded on $(y-\delta, y+\delta)$.

We now choose $\delta < c/(2k)$, and suppose first there is an $x \in (y-\delta, y+\delta)$ with $f(x) < -2^k M$. Then by (2.2)

$$(4.6) \qquad \Delta^k_{2\delta} f(x-2k\delta) < -2^k M + (2^k-1)M < 0,$$

as $x - 2j\delta \in [y-c+\delta, y-\delta]$ for $j=1,2,\ldots,k$. Suppose next there is an $x \in (y-\delta, y+\delta)$ with $f(x) > 2^k M$. In a similar way one finds

$$(4.7) \qquad \Delta^k_{2\delta} f(x-2k\delta+2\delta) < -k\, 2^k M + (2^k-k)M < 0,$$

as $x + 2\delta \in [y+\delta, y+c-\delta]$ and $x + 2\delta - 2j\delta \in [y-c+\delta, y-\delta]$ for $j=2,3,\ldots,k$; hence f is bounded on $(y-\delta, y+\delta)$.

So the supposition that y is an isolated discontinuity point of the I_k function f leads to a contradiction: in each neighbourhood of each discontinuity point of an I_k function there is a new discontinuity point.

But for a finite valued function f which is unbounded in any neigbourhood of an infinite number of discontinuity points, it is clear that even subtraction of a constant will lead to more than k sign changes (as soon as we have 4k discontinuities this can easily be shown explicitly). So the initial supposition of the existence of a discontinuity point y was incorrect, and every C_k function is continuous.

Theorem 6. If for some $k \geq 2$ f is continuous and I_k, then $\Delta^1_h f(.)$ is continuous and I_{k-1}, for all $h > 0$.

Proof. Continuity is trivial. For I_{k-1} we use that

(4.8) $$\Delta^1_{h_2} \quad \Delta^1_{h_2} \ldots \Delta^1_{h_k} \quad f(x)$$

is nonnegative for all positive h_i, as soon as f is I_k and continuous. This will be shown by VAN ZWET in [5]: if all h_i are integer multiples of a fixed $h > 0$, (4.8) can be rewritten as a sum of Δ^k_h -differences, and the general case follows by continuity. We find by choosing $h_1=h_2=\ldots=h_{k-1}$ that $\Delta^1_h f$ is I_{k-1}.

Theorem 7. If f is continuous and I_k then it is C_k.

Proof. The theorem is trivial for k=1; suppose it is true for k-1. Choose a polynomial p of degree at most k-1 and put $g(x) \overset{def}{=} f(x) - p(x)$. Then g is continuous, so we can divide the domain of definition (a,b) in "plusintervals" (where g is nonnegative, zero at the endpoints and positive in some interior point) and "minusintervals" between them (g nonpositive, zero at endpoints, negative somewhere). On the first and last interval we drop the condition g=0 for the endpoint a or b.

For sufficiently small h, Δ^1_h g (.) changes sign at least once on every plus- or minusinterval, except perhaps for the intervals at both ends of (a,b). As g is continuous and I_k, we know by theorem 6 and the induction assumption that $\Delta^1_h g$ has at most k-1 sign changes, and if exactly k-1 then with a plus ending. So there are at most k-1 plus- or minusintervals (the ones with endpoints a or b again excepted), and if exactly k-1 then the last one is a minusinterval (where Δ^1_h g changes from - to +). This proves that g has at most k sign changes, and if k then with a plus ending.

Theorems 4, 5 and 7 are now summarized:

Theorem 8. For $k \geq 2$ the I_k property plus continuity (or measurability) is equivalent to the C_k property.

References:

[1] W. FELLER, An introduction to probability theory and its applications Vol. 1, Second edition, Wiley 1957.

[2] H. HAHN and A. ROSENTHAL, Set functions, University of New Mexico Press 1948.

[3] S. KARLIN and F. PROSCHAN, Pólya type distributions of convolutions, Ann. Math. Stat. 31(1960), 721-736.

[4] S. KARLIN, Total positivity and convexity preserving transformations, Proc. Symp. Pure Math. 7, 329-347, Amer. Math. Soc., 1963.

[5] W.R. Van ZWET Report S 357 (VP 27), Statistische Afdeling, Mathematisch Centrum (to appear).