MONOTONICITY THEOREMS

B. S. THOMSON

Abstract. A generalization of the extreme derivate's of a function is given and used to prove several monotonicity theorems.

1. Introduction. There is now an extensive literature devoted to the investigation of conditions sufficient to ensure that a function be monotonic on an interval. An excellent introduction to this topic can be found in A. M. Bruckner's monograph [1] and in the expository article [2]. The multiplicity of results available in this area suggests an abstract and unified treatment might be useful. Here we introduce the notion of an abstract derivation basis and interpret several monotonicity theorems in terms of the "geometry" of the derivation basis. This permits us to derive a number of classical and recent monotonicity results within a single general setting.

2. General theory. Throughout \([a, b]\) is a fixed interval and \(\mathcal{I}\) denotes the collection of all closed (nondegenerate) subintervals of \([a, b]\). Our structure is built on the product set \(\mathcal{I} \times [a, b]\). Whenever \(S\) is a subset of \(\mathcal{I} \times [a, b]\) and \(X\) is a subset of \([a, b]\) we will write \(S[X] = \{(I, x) \in S: x \in X\}\) and \(S(X) = \{(I, x) \in S: I \subseteq X\}\). If \(\mathcal{B}\) is a family of subsets of \(\mathcal{I} \times [a, b]\) then we write similarly \(\mathcal{B}[X] = \{S(X): S \in \mathcal{B}\}\) and \(\mathcal{B}(X) = \{S(X): S \in \mathcal{B}\}\). A finite subset \(P = \{(I_i, x_i): i = 1, 2, \ldots, n\}\) of \(\mathcal{I} \times [a, b]\) is said to be a (pointed) partition if \(I_i\) and \(I_j\) do not overlap when \(i \neq j\). If \(P\) is a partition as above and \(E = \bigcup \{I_i: i = 1, 2, \ldots, n\}\) then we will say moreover that \(P\) is a partition of \(E\). A subset \(S\) of \(\mathcal{I} \times [a, b]\) is said to partition \(E\) if \(S\) contains a partition of \(E\). A family \(\mathcal{B}\) of subsets of \(\mathcal{I} \times [a, b]\) is said to partition \(E\) if every \(S, S \in \mathcal{B}\), partitions \(E\).

Our main definitions follow.

(1) A nonempty family \(\mathcal{B}\) of subsets of \(\mathcal{I} \times [a, b]\) is said to be a derivation scheme if for every family \(\{S_x: x \in [a, b]\} \subseteq \mathcal{B}\) there is an \(S \in \mathcal{B}\) with \(S(\{x\}) \subseteq S_x\) for each \(x \in [a, b]\).

(2) A derivation scheme \(\mathcal{B}\) is said to be a derivation basis if, moreover, for every set \(G\) open in \([a, b]\) and every \(S_0 \in \mathcal{B}\) there is an \(S \in \mathcal{B}\) for which \(S(G) \subseteq S_0(G)\).

(3) For any derivation scheme \(\mathcal{B}\) and any real-valued interval function \(h\) (\(h: \mathcal{I} \to R\)) the extreme \(\mathcal{B}\)-derivates of \(h\) are defined to be

\[
\overline{D}_\mathcal{B} h(x) = \inf_{S \in \mathcal{B}} \sup \{h(I)/|I|: (I, x) \in S\}
\]
and
\[ D_\delta h(x) = \sup_{S \in \mathcal{G}} \inf \{ h(I)/|I| : (I, x) \in S \} \]

at any point \( x \) in \([a, b]\) where \(|I|\) denotes the length of the interval \( I \) and empty suprema and infima follow the usual conventions. (Note that without further properties assumed for \( \mathcal{G} \) there is no suggestion that \( D_\delta h(x) \) must exceed \( D_{\mathbb{R}} h(x) \).)

If \( f \) is a real-valued function on \([a, b]\) then we will write as \( \Delta f \) the usual interval function defined from \( f \) by setting \( \Delta f(\alpha, \beta) = f(\beta) - f(\alpha) \). Thus our main object of study, the generalized extreme derivatives of \( f \), are written as \( D_\delta \Delta f(x) \) and \( D_{\mathbb{R}} \Delta f(x) \).

(4) For any derivation scheme \( \mathcal{G} \) and any real-valued interval function \( h \) we write
\[ V(h, S) = \sup \{ \Sigma_{(I, x) \in P} h(I) : P \subseteq S, P \text{ a partition} \}, \]
\[ V(h, \mathcal{G}) = \inf \{ V(h, S) : S \in \mathcal{G} \}, \]
and \( h_\mathbb{R}(X) = V(h, \mathcal{G}[X]) \) \((X \subseteq [a, b])\). We will interpret \( V(h, \emptyset) = 0 \).

(5) If \( \mathcal{K} \) is any family of nonnegative subadditive interval functions then a derivation scheme \( \mathcal{G} \) is said to be \( \mathcal{K} \)-complete (cf. \([4]\)) if \( h \in \mathcal{K} \) and \( V(h, \mathcal{G}) = 0 \) together imply that \( h = 0 \). In particular \( \mathcal{G} \) is \( \mathcal{C} \)-complete if this holds for every continuous, nonnegative, subadditive interval function.

To illustrate these definitions and for later reference we give a number of concrete realizations of these ideas. Readers familiar with \([4]\) and \([5]\) should recognize the following examples.

(i) The ordinary extreme (bilateral) derivatives. If \( \delta \) is any positive function defined on \([a, b]\) we will write \( D_\delta \) for the collection of all interval-point pairs \( ([x, y], z) \) with \( x, y, z \in [a, b], z = x \) or \( z = y \), and \( 0 < y - x < \delta(z) \). Then \( \mathcal{D} \) is the family of all such sets \( D_\delta \) for arbitrary positive functions \( \delta \). It is straightforward to verify that \( \mathcal{D} \) is a derivation basis and that \( \mathcal{D} \) partitions every interval in \( \mathcal{G} \) (cf. \([5, p. 221]\)). For any function \( f \) on \([a, b]\) the derivatives \( D_\delta \Delta f(x) \) and \( D_{\mathbb{R}} \Delta f(x) \) are exactly the lower and upper extreme bilateral derivatives of \( f \) at \( x \); if \( f \) is continuous and of bounded variation on \([a, b]\) it is possible to prove that the set function \( \Delta f_{\mathbb{R}} \) gives the usual Lebesgue-Stieltjes outer measure on \([a, b]\).

(ii) The approximate extreme derivatives. Let \((\mu, \lambda)\) be any pair of numbers in the interval \([0, 1]\). Suppose that for each \( x \in [a, b] \) there is given a measurable set \( M(x) \), with \( x \in M(x) \), such that \( M(x) \) has right density at \( x \) exceeding \( \mu \) (or \( \mu = 1 \) then right density equal to 1) and \( M(x) \) has left density at \( x \) exceeding \( \lambda \) (or \( \lambda = 1 \) then left density equal to 1). For such a function \( M(\cdot) \) write \( A_M = \{([x, y], z) : x, y, z \in [a, b], z = x \) or \( z = y, y < x, \) and \( x, y \in M(z) \} \). Then \( \mathcal{A}^{(\mu, \lambda)} \) will denote the family of all such collections \( A_M \) where \( M \) is any function satisfying the above stated properties for a fixed pair \((\mu, \lambda)\).

The family \( \mathcal{A}^{(\mu, \lambda)} \) can be seen to be a derivation basis; Henstock \([4]\) has shown that, provided \( \mu + \lambda > 1 \), \( \mathcal{A}^{(\mu, \lambda)} \) partitions every interval in \( \mathcal{G} \). For measurable functions \( f \) the extreme derivates \( D_{\mathbb{R}} \Delta f(x) \) and \( D_{\mathbb{R}} \Delta f(x) \) with \( \mathcal{A} = \mathcal{A}^{(1,1)} \) are exactly the approximate extreme derivates; the derivation basis \( \mathcal{A}^{(1/2, 1/2)} \) yields similarly the preponderant extreme derivates of a measurable function.

(iii) The symmetric extreme derivatives. For any positive function \( \delta \) on \((a, b)\) write \( S_\delta \) as the collection of all interval-point pairs \( ([x - t, x + t], x) \) in \( \mathcal{G} \times [a, b] \) for
which $0 < t < \delta(x)$, and let $\mathcal{G}$ denote the family of all such $S_\delta$ for arbitrary positive functions $\delta$ on $(a, b)$. Again $\mathcal{G}$ is a derivation basis and the extreme $\mathcal{G}$-derivates are the usual extreme symmetric derivates. $\mathcal{G}$ fails to partition every interval in $\mathcal{G}$ but it follows from a result of McGrotty [6] that $\mathcal{G}$ is $C$-complete, and this fact permits a number of monotonicity theorems to hold for the symmetric derivates in the presence of a continuity hypothesis.

We now proceed to our monotonicity theorems. If $h$ is an interval function then we write $h^+$ and $h^-$ for the interval functions $h^+(I) = \max\{h(I), 0\}$ and $h^-(I) = \max\{-h(I), 0\}$. To prove that a real-valued function $f$ on $[a, b]$ is nondecreasing is then equivalent, obviously, to showing that the interval function $\Delta f^-$ vanishes identically, and this in turn can frequently be established by showing that the associated set function $\Delta f^-$ vanishes. All our theorems reduce essentially to the fact that some information about the $\mathcal{B}$-extreme derivates of $\Delta f$ can be used to prove that $\Delta f^- \equiv 0$. In using this result the two lemmas which follow are basic; proofs can be found in [5].

**Lemma 1.** If $\mathcal{B}$ is a derivation scheme and $h$ a real-valued interval function then $h_\mathcal{B}$ is an outer measure on $[a, b]$. If further $\mathcal{B}$ is a derivation basis then $h_\mathcal{B}$ is a metric outer measure (i.e. all Borel sets are $h_\mathcal{B}$-measurable).

**Lemma 2.** Let $\mathcal{B}$ be a derivation scheme that partitions every interval in $\mathcal{G}$. Then $\mathcal{B}$ is necessarily $\mathcal{K}$-complete for any choice of $\mathcal{K}$, and in particular $\mathcal{B}$ is $C$-complete.

**Theorem 1.** Let $\mathcal{B}$ be a derivation scheme, $X$ a subset of $[a, b]$, and $h$ a real-valued interval function. If $D^-\mathcal{B}(h)(x) > 0$ for $h^-\mathcal{B}$-almost every $x$ in $X$ then $h^-\mathcal{B}(X) = 0$. If $D^-\mathcal{B}(h)(x) > 0$ a.e. in $X$ and $D^-\mathcal{B}(h)(x) > -\infty$ for $h^-\mathcal{B}$-almost every $x$ in $X$ then $h^-\mathcal{B}(X) = 0$, provided $\mathcal{B}$ is also a derivation basis.

**Corollary.** Let $\mathcal{B}$ be a derivation scheme that is $\mathcal{K}$-complete. Then in order that a function $f$ on $[a, b]$ be nondecreasing it is sufficient that $\Delta f^- \in \mathcal{K}$ and either

(a) $D^-\mathcal{B}\Delta f(x) > 0$ for $\Delta f^-\mathcal{B}$-almost every $x$ in $[a, b]$, or

(b) $D^-\mathcal{B}\Delta f(x) > 0$ a.e. in $[a, b]$ and $D^-\mathcal{B}\Delta f(x) > -\infty$ for $\Delta f^-\mathcal{B}$-almost every $x$ in $[a, b]$, with $\mathcal{B}$ being moreover a derivation basis.

**Proof.** The corollary follows immediately from the theorem because of definition (5). To prove the first part of the theorem we need only establish that $h^-\mathcal{B}(X_\delta) = 0$ where $X_\delta = \{x \in X: D^-\mathcal{B} h(x) > 0\}$. For any $\varepsilon > 0$ there must be (by (1) and (3)) an $S \in \mathcal{B}$ with $h(I) > -\varepsilon|I|$ for every $(I, x) \in S[X_\delta]$. This gives $h^-(I) < \varepsilon|I|$ for such $(I, x)$ and so $h^-\mathcal{B}(X_\delta) = V(h^-, S[X_\delta]) < V(h^-, S[X_\delta]) < \varepsilon(b - a)$. As $\varepsilon > 0$ is arbitrary it follows that $h^-\mathcal{B}(X_\delta) = 0$ proving the first part.

For the second part of the theorem we add in the hypothesis that $\mathcal{B}$ is a derivation basis (satisfies (2)). Let us write $Y = \{x \in X: -\infty < D^-\mathcal{B} h(x) < 0\}$; the theorem then follows from the first part if we are able to establish that $h^-\mathcal{B}(Y) = 0$. For each integer $n = 1, 2, 3, \ldots$ write $Y_n = \{x \in X: -(n - 1/2) < D^-\mathcal{B} h(x) < 0\}$. By our assumptions each $Y_n$ has measure zero and so can be included in an open set $G_n$ of measure less than any given $\eta > 0$. Select an $S_n \in \mathcal{B}$ so that $h(I) > -\eta|I|$ whenever $(I, x) \in S_n[Y_n]$ and then use (2) to find an $S'_n \in \mathcal{B}$ with $S'_n[G_n] \subseteq S_n(G_n)$.
From this we see that every \((I, x) \in S_n[x]\) has \(I \subseteq G_n\) and also \(h(I) > -n|I|\). This proves that \(h_\infty(Y_n) = V(h_\infty, S_n[Y_n]) < V(h_\infty, S_\infty[Y_n]) \leq m_\infty\). As \(\eta > 0\) is arbitrary this gives \(h_\infty(Y_n) = 0\) for each \(n\) and so \(h_\infty(Y) < \sum_{n=1}^{\infty} h_\infty(Y_n) = 0\) as required proving the theorem.

With a minimum of structure we have included in this theorem a great many known monotonicity theorems (as is outlined below). The celebrated theorem of Goldowski and Tonelli (see e.g. Saks [8, p. 206]) requires further properties of the derivation basis. We shall need that the family \(\mathcal{B}\) be filtering downwards and that its members permit a decomposition similar to that available for ordinary derivatives (e.g. as in the proof of G. C. Young’s theorem in [1, p. 63]). Specifically we write the properties as (a) for every \(S_1 \in \mathcal{B}\) and \(S_2 \in \mathcal{B}\) there is an \(S_3 \in \mathcal{B}\) with \(S_1 \subseteq S_3 \cap S_2\); and (b) for every \(S \in \mathcal{B}\) and every \(X \subseteq [a, b]\) there is a sequence of sets \(\{X_n\}\) with \(X = \bigcup X_n\) such that \(S[X_n]\) partitions every interval with endpoints in \(X_n\).

We may now state and prove an abstract version of the Goldowski-Tonelli theorem.

**Theorem 2.** Let \(\mathcal{B}\) be a derivation basis having properties (a) and (b) and which is \(C\)-complete. If \(f\) is a continuous function on \([a, b]\) for which (i) \(D_{\mathcal{B}} f(x) > 0\) a.e. and (ii) \(D_{\mathcal{B}} f(x) = D_{\mathcal{B}} f(x)\) everywhere except possibly on a countable set, then \(f\) is nondecreasing on \([a, b]\).

**Proof.** The proof follows very closely the lines of the original as it appears in Saks [8, p. 206]. Let \(G\) be the collection of all points \(x \in [a, b]\) for which \(\Delta f_{\mathcal{B}}(\alpha, \beta)\) vanishes on some interval \((\alpha, \beta)\) containing \(x\). It is clear that \(G\) is open in \([a, b]\) and that \(\Delta f_{\mathcal{B}}(G) = 0\). The conclusion of the theorem is obtained if we prove that \(G = [a, b]\). To this end let \(E = [a, b] \setminus G\); \(E\) is evidently closed and it is to be proved that it is empty.

Write \(X_1 = \{x \in [a, b]: D_{\mathcal{B}} f(x) < -1\}\), \(X_2 = \{x \in [a, b]: D_{\mathcal{B}} f(x) > -2\}\), and observe that because of (ii) these sets exhaust \([a, b]\) except possibly for a countable set. Choose an \(S \in \mathcal{B}\) so that \(\Delta f(S) < -1\) whenever \((I, x) \in S[X_1]\) and \(\Delta f(I)/|I| > -2\) whenever \((I, x) \in S[X_2]\). (This uses property (a).) By property (b) there are sequences \(\{E_n^1\}\) and \(\{E_n^2\}\) with \(X_1 = \bigcup E_n^1\) and \(X_2 = \bigcup E_n^2\) so that \(S[E_n^i]\) partitions every interval with endpoints in \(E_n^i\) (i = 1, 2).

Suppose that \(E\) is nonempty: then by Baire’s theorem [8, p. 54] there must be a nonempty portion \(E \cap (\alpha, \beta)\) in which either (1°) some set \(E_n^1\) is dense, (2°) some set \(E_n^2\) is dense, or (3°) that reduces to a single point. The theorem is proved by establishing that each of these is impossible.

As \(f\) is continuous it is easy to see that \(\Delta f_{\mathcal{B}}(x) = 0\) for every \(x \in [a, b]\). Thus if (3°) is the case, i.e. \(E \cap (\alpha, \beta)\) is a single point, then \(\Delta f_{\mathcal{B}}(\alpha, \beta) \leq \Delta f_{\mathcal{B}}(G) + \Delta f_{\mathcal{B}}(E \cap (\alpha, \beta)) = 0\) which contradicts the definition of \(E\).

If (1°) is the case and \(E_n^1\) is dense in \(E \cap (\alpha, \beta)\) then any interval \(I\) with endpoints in \(E \cap (\alpha, \beta)\) must have \(\Delta f(I)/|I| < -1\). To see this observe that any interval \([x, y]\) with both \(x\) and \(y\) in \(E_n^1\) is partitioned by \(S[E_n^1]\) and so must have \(f(y) - f(x)/(y - x) < -1\); by continuity of \(f\) this obviously extends to the closure
of $E^1_n$ proving the claim. Accordingly $E$ must be nowhere dense in $(\alpha, \beta)$, otherwise we would have $\overline{\Delta}_g f(x) < -1$ everywhere on some subinterval which would violate (i) of the theorem. On the other hand if $(\xi, \eta)$ is a component subinterval of $(\alpha, \beta) \setminus E$ then $\Delta_f G((\xi, \eta)) < \Delta_f G(G) + \Delta_f G((\xi)) + \Delta_f G(\{\eta\}) = 0$ so that $\Delta(\xi, \eta) = 0$ and this contradicts the fact that $\Delta(\xi, \eta) < -1$ as proved above for such intervals.

There now remains only case (2°). Suppose that $E^2_n$ is dense in $E \cap (\alpha, \beta)$. By arguments similar to those above we can prove that $\Delta_f f(I)/|I| > -2$ for any interval $I$ with endpoints in $E \cap (\alpha, \beta)$; in fact this is true for any interval $I \subset (\alpha, \beta)$. For example if $[x, y] \subset (\alpha, \beta)$ with $x \in E$ and $y \in G$ then there is a component $(\xi, \eta)$ of $G$ to which $y$ belongs and, as before, $\Delta_f G((\xi, \eta)) = 0$ so that $f(y) - f(x) > 0$. This gives $f(y) - f(x) = f(y) - f(\xi) + f(\xi) - f(x) > -2$ since $[x, \xi]$ has endpoints in $E \cap (\alpha, \beta)$. The other cases are similar. But this inequality holding for all subintervals of $(\alpha, \beta)$ requires that $\overline{\Delta}_g f(x) > -2$ at every point of $(\alpha, \beta)$; this together with (i) of the theorem and Theorem 1 gives $\Delta_f G(\alpha, \beta) = 0$ which again contradicts the definition of $E$ and completes the proof of the theorem.

The assumption of continuity in this theorem can be replaced by the requirement that $f$ be at least in Baire class 1 and continuous in the Darboux sense. For a proof we need only appeal to a very general reduction theorem of Bruckner [1, p. 181] that permits such extensions under very mild (and frequently encountered) conditions. Here the only condition that needs to be checked is supplied by the following lemma.

**Lemma 3.** Let $\mathcal{B}$ be a derivation basis that has property (b) above. If everywhere in a set $X$ with at most countably many exceptions one of the inequalities $\overline{\Delta}_g f(x) > -\infty$ or $\overline{\Delta}_g f(x) < +\infty$ holds then $f$ must be VBG on $X$.

**Proof.** It is evidently sufficient to show that $f$ is VBG on the set $X = \{ x \in [a, b]: \overline{\Delta}_g f(x) < +\infty \}$ and if we write, for each natural number $n$, $X_n = \{ x \in X: \overline{\Delta}_g f(x) < n \}$ then it is enough to prove that $f$ is VBG on each $X_n$ and it will follow that $f$ is VBG on the set $X = \bigcup X_n$. To this end choose $S \in \mathcal{B}$ so that $\Delta f(I) < n|I|$ for every $(I, x) \in S[X_n]$ and use property (b) to produce a sequence of sets $\{E_m\}$ covering $X_n$ such that $S[E_m]$ partitions every interval with endpoints in $E_m$. If $J$ is any such interval it is evident that $\Delta f(J) < n|J|$. This shows that $f(x) - nx$ is decreasing on $E_m$ and so $f$ is of bounded variation on $E_m$. Since $X_n = \bigcup E_m$ it follows that $f$ is VBG on $X_n$ as required.

We conclude with some remarks illustrating possible applications of the theorems presented here. If $\mathcal{D}$ denotes the derivation basis associated with ordinary differentiation (example (i) above) and the lower extreme derivate of a function $f$ is written in its usual notation $f'(x)$ (rather than $\overline{\Delta}_g f(x)$) then Theorem 1 proves the following monotonicity result: in order for a function $f$ to be nondecreasing on an interval $[a, b]$ it is enough that $f'(x) > 0$ except possibly in a set $N$, or that $f'(x) > 0$ a.e. and $f'(x) > -\infty$ except possibly in a set $N$ with $\Delta f^{-} (N) = 0$. Here $N$ may be taken as empty with no restrictions on $f$, $N$ could be taken as countable on
the assumption that \( f \) is continuous at each point of \( N \), \( N \) could be taken as a set of measure zero provided we know that \( f \) is ACG* or \( N \) could be such that \( |f(N)| = 0 \) provided \( f \) is continuous and VBG*, since in each of these cases it is possible to verify that \( \Delta f_N^-(N) = 0 \). Theorem 2 in this setting is exactly the classical Goldowski-Tonelli theorem.

With \( \mathbb{R} = \mathbb{R}^{(1,1)} \) of example (ii) and writing \( f_{sp}(x) \) in place of \( D_{sp}^A f(x) \) for the lower approximate extreme derivate we have the theorem: in order for a function \( f \) to be nondecreasing on an interval \([a, b]\) it is enough that \( f_{sp}(x) > 0 \) except possibly in a set \( N \), or \( f_{sp}(x) > 0 \) a.e. and \( f_{sp}(x) > -\infty \) except possibly in a set \( N \) where \( f \) is measurable and where \( \Delta f_N^{-}(N) = 0 \). Using \( N \) empty here gives a result of Goffman and Neugebauer [3] and a recent result of O'Malley [7]. Again there are a number of variants possible by varying the scope of the exceptional set \( N \) and restricting \( f \) in such a way that \( \Delta f_N^{-}(N) = 0 \). In this setting Theorem 2 goes back to Tolstoff and Bruckner (see [1, p. 175]).

There are a number of other derivation bases that share the properties of \( \mathcal{D} \) and \( \mathbb{R} \). We shall not give the details but it is possible to investigate qualitative derivatives and selective derivatives in this setting and to prove that the corresponding derivation bases have all the properties needed to apply Theorems 1 and 2. Note that the symmetric derivation basis does not have property (b) and so while Theorem 1 applies to it, Theorem 2 does not.

REFERENCES


DEPARTMENT OF MATHEMATICS, SIMON FRASER UNIVERSITY, BURNABY, BRITISH COLUMBIA, CANADA