TOPICAL SURVEY Real Analysis Exchange Vol. 8 (1982-83)

B.S. Thomson - Simon Fraser University, Burnaby, B.C., Canada V5A 1S6

DERIVATION BASES ON THE REAL LINE (I)

Preface. The material presented here is a preliminary version of a projected monograph on the subject of differentiation and integration theory on the real line. I would like to thank the editors of the Exchange for their encouragement in this project and their willingness to present this admittedly tentative version. I trust the reader too will be as generous in overlooking the obvious rough edges.

The intention of this study is to present a language and a framework within which a large portion of classical real analysis permits a relatively clear and simple expression. Our unifying concept is that of a derivation basis. Relative to any derivation basis there are three fundamental related concepts, giving rise to a derivation theory, an integration theory and a measure theory. Many of the concerns of analysts over the years can be considered as entirely natural problems that arise within such a setting and by placing them within this setting one acquires a convenient way of expressing the problems, a clearer picture of the many interrelations between problems, and a unified methodology for attacking the problems.

Our language draws on two main sources: the abstract differentiation theory introduced some sixty years ago by R. de Possel and developed since then by numerous authors, and the abstract integration theory (generalized Riemann integration) introduced by R. Henstock some twenty years ago and developed since then mainly by Henstock himself and his students. As we choose to place everything just on the real line we will not require the full apparatus of these two abstract theories nor will we require of the reader any familiarity with them.

The first chapter is only an attempt to motivate the general theory and to give an indication of the form it is to take. One generally needs a good bit of motivation in order to pursue an abstract theory and the ideas here are no exception. The fact that there are by now dozens of distinct derivation and integration processes and hundreds of papers generated in an attempt to sort out their properties certainly provides adequate motivation for a general theory. Indeed numerous authors have responded to this situation (which Professor Garg describes as a "jungle") by putting forth concepts designed to unify and simplify. Of the many unifying attempts in the theory of integration on the real line that of Henstock appears to be the most successful. Here we take this a step further and show that Henstock's ideas also provide a unification of the many derivation processes and that they provide at the same time a better view of the historically close connection between integration and differentiation.

The second chapter introduces the notion of an abstract derivation basis on the real line and develops all of the terminology needed for the remainder of the work. The ideas are really quite simple but they apparently are rather compact and the notation takes a bit of familiarity; for this reason we have liberally given examples (all in italics) to help illustrate the ideas and to ease somewhat the burden on the reader. There is always the danger in the development of an abstract language that the author will create merely his own private fantasy, unshared by others; certainly our subject already has a number of apparently profound but clearly unreadable works. For this reason the terminology has been kept to a bare minimum and mainly suggestive notation and labelling has been used, although of course the risk remains.

The third chapter develops the properties of the variation. This concept is a common generalization of such diverse notions as Peano-Jordan measure, upper Darboux integrals, upper Lebesgue and Lebesgue-Stieltjes integrals, total variation of a function, Burkill integration and Hellinger integration. The variation is the most convenient tool for developing properties of the derivation or the integration and it provides remarkably simple proofs of a variety of well known theorems. In addition there are a number of concepts which arise directly or indirectly from a consideration of the variation and these too will appear in this chapter; thus the measure theory and the theory of the upper integral are given here as well as such notions as a generalized continuity.

Chapters four and five, which give treatment of the abstract differentiation theory and the integration theory, will appear in a later issue of the Exchange.

Proofs are given for all the main results. For the illustrative examples usually only a statement of the result appears with, perhaps, a reference to the literature where one is known. A number of items in the text have appeared as queries; this is meant to indicate only that at the time of writing this question occurred to the writer with no corresponding answer coming to mind. Answers to the queries, suggestions and critical comments would be most welcome.

CHAPTER ONE

INTRODUCTION

§1. Abstract differentiation theory. Our concern throughout is with abstract differentiation theory in its simplest setting, the real line. In that setting much of the usual machinery of that theory becomes empty and most of the traditional problems of the subject meaningless. However, new problems and new machinery arise naturally by reviewing classical analysis in such an abstract setting. It is hoped that such a study will throw some light on a number of problems in analysis and will enrich the general subject of abstract differentiation theory itself.

We begin with a very sketchy review of the abstract theory: Let $(X, \mathcal{M}, \mu)$ be a measure space and select from $\mathcal{M}$ a distinguished class of sets $\mathcal{I} \subset \mathcal{M}$ so that $0 < \mu(I) < +\infty$ for every $I \in \mathcal{I}$. The class $\mathcal{I}$ is to play the role of the "intervals" and so we shall refer to them as "generalized intervals".

DEFINITION 1.1 A differentiation basis B on the measure space $(X, \mathcal{M}, \mu)$ is a filterbase on the product set $\mathcal{I} \times X$.

That is B is a collection of subsets of $\mathcal{I} \times X$, so that each $\beta \in B$ contains numerous pairs (I,x) (I an interval and $x \in X$), and B has the filterbase properties: (a) $\emptyset \notin B$, (b) if β_1 and β_2 belong to B then there is a $\beta_3 \subset \beta_1 \cap \beta_2$ that also belongs to B.

For any function $F : \mathcal{I} \to R$ we may define its derivative at a point $x \in X$ with respect to this differentiation basis B by writing

$$D_B F(x) = \lim F(I)/\mu(I)$$

where the limit is taken in the sense of the filterbase B. In the simplest language $D_B F(x) = c$ if for every $\varepsilon > 0$ there is a $\beta \in B$ so that

$$|F(I)/\mu(I) - c| < \varepsilon \text{ for all } (I,x) \in \beta .$$

In less simple language this can be written by setting

$$D(F,x;\beta) = \{F(I)/\mu(I) : (I,x) \in \beta\}$$

and then $\{D(F,x;\beta) : \beta \in B\}$ is a filterbase on the real line and it converges to the derivative if such exists.

The central problem of abstract differentiation theory can be formulated with a minimum of terminology now; a function f on X is locally integrable if it is integrable on each set $I \in \mathcal{I}$, and if so we may write $F^f(I) = \int_I f\, d\mu \quad (I \in \mathcal{I})$ for its indefinite integral.

BASIC PROBLEM OF ABSTRACT DIFFERENTIATION THEORY. Given a class $\mathcal{F}$ of locally integrable functions, what are the necessary and sufficient conditions that a basis B should have in order for the assertion

$$D_B F^f(x) = f(x) \quad \mu\text{-almost everywhere in } X$$

to hold for every function f in the class $\mathcal{F}$?

The charm and challenge of this problem is in the quite surprising fact that the answer lies not in measure-theoretic considerations nor in topological considerations but in the geometry of the differentiation basis B. For a full explanation of this vague term and the world of Vitali conditions, halo conditions, etc. to which the problem leads the reader might consult Bruckner [8] for an overview and then de Guzmàn [37] and Hayes and Pauc [39] for a serious study of the subject.

If we restrict our attention to the real line with μ as Lebesgue measure then we are naturally drawn to use genuine intervals $[a,b]$ for the class $\mathcal{I}$. Thus, if $\mathcal{I}$ is the collection of all closed bounded nondegenerate intervals on the real line R, a differentiation basis B is a filterbase on $\mathcal{I} \times R$. Such objects are the principal object of our study in the present work.

We are not however interested in the basic problem given above, for in the setting of the real line almost any reasonable differentiation basis would differentiate the integrals of any locally integrable function even if for μ we take Lebesgue-Stieltjes measures μ_G generated by a continuous bounded variation function G. Our program instead is to seek other problems natural to real analysis and to discover the geometry of a differentiation basis that lies at the heart of the problem. This is completely within the spirit of abstract differentiation theory but will lead us to entirely different notions.

In the next three sections we show how the notion of an abstract differentiation basis on the real line arises naturally in the study of three quite different concepts.

§2. Generalized derivations. It is common now, in the study of real functions, to replace the derivative (which may not exist) by the extreme derivates (which do exist). Thus for any function F one introduces the bilateral derivates defined as

$$\overline{D}F(x) = \limsup_{y\to x} \frac{F(y)-F(x)}{y-x}$$

and

$$\underline{D}F(x) = \liminf_{y\to x} \frac{F(y)-F(x)}{y-x}$$

and the four unilateral derivates (also called Dini derivatives)

$$^{+}\overline{D}F(x) = \limsup_{y\to x+} \frac{F(y)-F(x)}{y-x},$$

$$^{+}\underline{D}F(x) = \liminf_{y\to x+} \frac{F(y)-F(x)}{y-x}$$

$$^{-}\overline{D}F(x) = \limsup_{y\to x-} \frac{F(y)-F(x)}{y-x}$$

$$^{-}\underline{D}F(x) = \liminf_{y\to x-} \frac{F(y)-F(x)}{y-x}.$$

These derivation processes are the most natural but they do not exhaust the many other reasonable (and unreasonable) ways in which the derivative has been generalized. For later reference (as well as to impress on the reader the bewildering variety of such inventions) we list those generalized derivatives that are known to us. Since our concern is to present a unified and simplified treatement of these derivation processes we should first confront the confusion into which some order is to be thrown.

(i) uniform derivation. A function f is said to be a uniform derivative of a function F on a set X provided

$$\lim_{h\to 0} [F(x+h) - F(x)]/h = f(x) \quad \text{uniformly for } x \in X .$$

While such a derivation should have limited interest as being far too restrictive this process plays an historically important role and will play an important illustrative role in our general theory. A number of authors have used this concept (e.g., Weinstock [117], Lahiri [74], Bhakta and Mukhopadhyay [6], Manna [80]).

(ii) sharp derivation. For any function F we define the sharp extreme derivates as

$$\overline{D}^{\#} F(x) = \limsup_{\substack{(y,z)\to(x,x)\\ y\neq z}} \frac{F(y) - F(z)}{y - z}$$

and

$$\underline{D}^{\#} F(x) = \liminf_{\substack{(y,z)\to(x,x)\\ y\neq z}} \frac{F(y) - F(z)}{y - z} .$$

These evidently satisfy the inequality

$$\underline{D}^{\#} F(x) \le \underline{D} F(x) \le \overline{D} F(x) \le \overline{D}^{\#} F(x)$$

and so represent not a weakening of the ordinary derivation process but a sharpening of it and hence the name. It would be appropriate to name them after Peano who first introduced them in 1892 (Peano [99]) but his name is now firmly attached to a different process. Some authors have referred to this as "strong" derivation but that term is reserved for quite a different notion in general differentiation theory (*cf.* Saks [105, p. 106]) and would interfere with the terminology needed for vector-valued extensions of this concept. Bruckner [10, p. 69] suggests "unstraddled" but only in passing.

The concept itself might at first sight appear too restrictive to play any serious role. For example, in order for a sharp derivative $D^{\#} F(x)$ to exist in an interval F must be continuously differentiable there. Even so the sharp derivation can be used to clarify the nature of assertions about other derivations. For example the following two classical theorems for the Dini derivatives (the first due to Dini himself and the second to W.H. Young) are much clearer when expressed in terms of the sharp derivation.

THEOREM 1. *If $^{+}\underline{D} F$ is continuous at a point x_0 then F is differentiable at x_0.*

THEOREM 2. *If F is continuous then residually $^{+}\underline{D} F(x) = {}^{-}\underline{D} F(x)$.*

In their sharper versions (the latter of which is due to Bruckner and Goffman [11]) we have

THEOREM 1'. *If $^{+}\underline{D} F$ is continuous at a point x_0 then F has a sharp derivative there.*

THEOREM 2'. *If F is continuous then residually $\underline{D}^{\#} F(x) = {}^{+}\underline{D} F(x) = {}^{-}\underline{D} F(x)$.*

Any study of this derivation process should begin with a reading of Peano's original paper ([99]) which contains a number of basic observations. (For some more recent studies see Esser and Shisha [30], Bhakta and Mukhopadyay [6], Belna, Evans and Humke [3].)

(iii) density derivation. The most profound and useful of the generalizations of the ordinary derivation process was given by Denjoy [24] and by Khintchine [70]. For any function F the derivate $ap-\overline{D}F(x)$ is the infimum of all numbers c for which the set

$$\left\{ y : \frac{F(y)-F(x)}{y-x} > c \right\}$$

has x as a point of dispersion (i.e., of outer density zero).

This, together with its corresponding lower derivate, is the process of bilateral approximate derivation. One sided versions are available merely by considering the appropriate one sided density condition, and we label these as

$$ap-{}^{+}\overline{D}\,F(x)\ ,\ ap-{}^{+}\underline{D}\,F(x)\ ,\ ap-{}^{-}\overline{D}\,f(x)\ ,\ \text{and}\ \ ap-{}^{-}\underline{D}\,F(x)$$

and refer to them as the approximate Dini derivatives.

There is an extensive literature devoted to the study of these derivations and a formidable range of results has been obtained. The survey article of Bruckner and Goffman [12] provides a recent and quite comprehensive review of the subject.

By relaxing the density requirements in the above definition we can obtain a number of generalizations of the approximate derivative. For a pair of number (ρ, λ) chosen from the interval $[0,1)$ define the derivate $ap_{(\rho,\lambda)}-\overline{D}\,F(x)$ to be the infimum of the numbers c for which the set

$$\left\{ y : \frac{F(y)-F(x)}{y-x} > c \right\}$$

has outer right upper density less than $1-\rho$ and outer left upper density less than $1-\lambda$.

Although these derivations are of much less significance than the approximate given above they do play a role in some investigations. Denjoy [24] studied the $ap_{(\frac{1}{2}, \frac{1}{2})}$ - process and some investigations of Denjoy and Khintchine (cf. Saks [105, pp. 295-297]) involve the process $ap_{(0, 0)}$.

In their study of Perron type integrals Sarkhel and De [107] have been led to consider a different type of modification of the above density condition: a set E is sparse on the right at a point x if for every $\varepsilon > 0$ there is a $\delta > 0$ so that every interval $(a,b) \subset (x, x+\delta)$ with $(a-x) < \delta (b-x)$ contains a point y such that $|E \cap I| < |I| \varepsilon$ where $I = (x,y)$. Clearly, a sparse set has lower density zero but examples are given [107, p. 30] to show that its upper density may be arbitrarily close to 1 . Using this notion of sparse in place of density zero we obtain the "proximal derivation" process.

(iv) category based derivation. It is only natural, given the success and importance of the density based derivation process, that one tries to utilize other measures of size. Thus one could consider using the concept "first category in some neighbourhood of x" in place of "density zero at x" in the previous definitions. This leads to the notion of a qualitative derivation introduced by S. Marcus [82]. A number of authors have shown that such derivations share many properties of the ordinary derivative; for example, it is shown in Bruckner, O'Malley and Thomson [13] that a qualitatively differentiable function is in fact differentiable.

(v) selective derivation. A selection is a function $s: R^2 \to R$ that is symmetric (i.e., $s(x,y) = s(y,x)$) and for any $x < y$, has $x < s(x,y) < y$. This concept was used by Neugebauer [94] to characterize those functions which are in the first class of Baire and have the Darboux property, and (solving formally a long standing problem of W.H. Young) to characterize those functions which are everywhere the derivative of some continuous function. These considerations have led O'Malley [95] to define a derivation process relative to a selection and he has used this class of derivations to shed some light on other types of derivatives. (We might mention as well that the notion of a selection has been used to provide a modification of the Riemann-Stieltjes integral; see for example Baker and Shive [2].)

Thus if s is a selection one has defined

$$\overline{D}_s F(x) = \limsup_{y \to x} [F(s(x,y)) - F(x)]/[s(x,y)-x]$$

and

$$\underline{D}_s F(x) = \liminf_{y \to x} [F(s(x,y)) - F(x)]/[s(x,y)-x] .$$

One then studies the properties of selective derivates and selective derivatives relative to general selections or selections having some further properties; since a number of different derivations can be realized as selective derivatives this gives a general approach to studying a certain class of generalized derivations. (See the article of O'Malley [95] for details.)

O'Malley has also defined a related class of "bi-selective derivations".

(vi) parametric derivation. Evans and Humke [29] define a derivation process of the form

$$\underline{D}_\varphi F(x) = \liminf_{h \to 0+} [F(x - \varphi(h)) - F(x - \varphi(h) - h)]/h$$

and

$$\overline{D}_\varphi F(x) = \limsup_{h \to 0+} [F(x - \varphi(h)) - F(x - \varphi(h) - h)]/h$$

where φ is an appropriate monotonic function considered as a "parameter", and they obtain numerous properties of such derivates and derivatives. Other variants on this theme are possible by the choice of similar "parameters". In a sense, in fact, the selective derivative is a certain type of parametric derivative.

(vii) neighbourhood filter derivation. For each $x \in R$ suppose there has been given a filter $N(x)$ converging to x ; then one defines $\overline{D}_N F(x)$ to be the infimum of the numbers c so that

$$\left\{ y : \frac{F(y) - F(x)}{y - x} < c \right\} \cup \{x\} \in N(x) .$$

Similarly $\underline{D}_N F(x)$ is the supremum of the numbers c for which

$$\left\{ y : \frac{F(y) - F(x)}{y - x} > c \right\} \cup \{x\} \in N(x) .$$

If one takes for $N(x)$ the usual neighbourhood filter on the real line at x then D_N reduces to ordinary derivation. If one takes for $N(x)$ the collection of all sets η that have inner density 1 at x then D_N realizes the approximate derivation $ap - D$.

Although this is a most useful and most general manner of expressing a derivation process there appear to have been only two publications investigating the notion: the original work of Swiatkowski [111] (in Polish) in which the idea was introduced and an article of Mastalerz-Wawrznczak [83] obtaining a version of the Goldowski-Tonelli theorem for such derivations.

A special case of the neighbourhood filter derivation has been considered by Császár [22]. Let $\mathcal{S}$ be a σ-ideal of subsets of R , i.e., if $A \in \mathcal{S}$ and $B \subset A$ then $B \in \mathcal{S}$ ($\mathcal{S}$ is hereditary) and $\mathcal{S}$ is closed under countable unions. Define $\mathcal{F}$ to be the filter of sets that are complements of sets in $\mathcal{S}$ and at each point x use $N(x) = \mathcal{F}$. Then the derivation D_N corresponds to that of Császár: because $N(x)$ is the same filter at each point and because it is closed under countable rather than merely finite intersections further properties should be available. See [22] for an account of these.

A similar theory of limits based on a system of filters has been investigated by Jedrzejewski [66].

(viii) path derivation. For a measurable function F that has an approximate derivative $ap-D\,F(x_0)$ at a point there is a particularly convenient expression of that derivative: for some set E measurable and with density 1 at x_0 the function F has a derivative at x_0 relative to E equal to the number $ap-D\,F(x_0)$, i.e., $\lim\,[F(y)-F(x)]/[y-x]$ as $y \to x$ with $y \in E$ is this number. This realization of a generalized derivation as a derivation relative to some set has a number of convenient properties. For this reason the notion of path derivation was introduced in an attempt to unify a variety of themes in differentiation theory.

A system $E = \{E_x : x \in R\}$ is a system of paths if each E_x is a set of real numbers having x as a point of accumulation. Then derivation relative to the paths in E is defined as

$$\overline{D}_E\, F(x) = \limsup_{\substack{y \to x \\ y \in E_x}} \frac{F(y) - F(x)}{y - x}$$

and

$$\underline{D}_E\, F(x) = \liminf_{\substack{y \to x \\ y \in E_x}} \frac{F(y) - F(x)}{y - x} \ .$$

The theory of such derivates and their corresponding derivatives proceeds by investigating properties that arise in the derivates from assumptions about the thickness of the paths and the manner in which pairs of paths E_x and E_y intersect. (See Bruckner, O'Malley and Thomson [13] for some such results.)

A special case is obtained by fixing a set Q that has 0 as a point of accumulation and defining E to be the system $\{Q + x : x \in R\}$. If we denote the corresponding derivation D_E as $[Q]-D$ we have the "congruent derivation" of Sindalovski [109].

A special case in turn of the congruent derivative can be obtained by fixing a sequence $\{h_n\}$ convergent to zero and taking for Q the range of that sequence. If we denote the derivation as $[\{h_n\}]-D$ we have the sequential derivation of Petruska and Lazcovich [73].

Finally, a more elaborate path derivation can be defined relative to a sequence $\{E_n\}$ of sets that cover the line: one writes

$$[\{E_n\}]-\overline{D}\,F(x) \;=\; \sup_n \;\limsup_{\substack{y\to x\\ y\in E_n}} \frac{F(y)-F(x)}{y-x}$$

and

$$[\{E_n\}]-\underline{D}\,F(x) \;=\; \inf_n \;\liminf_{\substack{y\to x\\ y\in E_n}} \frac{F(y)-F(x)}{y-x} \;.$$

This is the "composite path derivation" of O'Malley and Weil [97]. It is motivated by O'Malley's observation in [96] that an approximately differentiable function permits a decomposition of the line into sets $\{E_n\}$ for which $ap-D\,F(x) = [\{E_n\}]-D\,F(x)$ must hold.

This type of derivative was first studied by Ridder [103] and Tolstov [116] in order to provide a Perron type characterization of the general integral of Denjoy.

(ix) <u>symmetric derivation</u>. One of the most familiar and useful of the many generalized derivations is obtained by writing

$$\text{sym-}\underline{D}\,F(x) \;=\; \liminf_{h\to 0+}\; [F(x+h) \;-\; F(x-h)]/2h$$

and

$$\text{sym-}\overline{D}\,F(x) \;=\; \limsup_{h\to 0+}\; [F(x=h) \;-\; F(x-h)]/2h \;.$$

This process has been extensively studied by numerous authors and there are a great many results that have been obtained. A list of references would be too lengthy and almost certainly incomplete; the interested reader should consult Khintchine for the earliest and Preiss and Larson for the most recent of the deep results known for this derivative. Bruckner [10, p. 168] lists some bibliography.

By combining the idea behind the symmetric derivation process with some of the previously mentioned ideas one can arrive at an "approximate symmetric derivation", "preponderant symmetric derivation", "symmetric path derivation", etc.. These ideas have been pursued by a number of authors (e.g., Evans [28], Larson [76], Mukhopadyay [93]).

(x) relative derivation. By altering the difference quotient in each of the preceding derivations one can arrive at the notion of "relative derivation" or derivation relative to a function G. In place of a quotient $[F(y) - F(z)]/(y - z)$ write $[F(y) - F(z)]/[G(y) - G(z)]$. The study of such derivatives goes back quite far; Lebesgue made a number of contributions. Several results are given in Saks [105, pp. 272-277] and (in rather more arcane language) in Kenyon and Morse [68].

A similar alteration was made by Besicovitch [5] who studied the limits of the quotient $[F(y) - F(z)]/[y - z]^s$ for $0 < s < 1$ and obtained results closely related to Hausdorff dimension (as might be expected). A common generalization of this relative derivation and Besicovitch's fractional derivation is obtained by considering simply the quotient $h(I)/g(I)$ with $I = [z,y]$ and where h and g are arbitrary (not necessarily additive) interval functions. This study along with a corresponding integration theory was initiated by Burkill [14], [15] and has been continued by numerous later authors (e.g., Kempisty [67], Henstock [41], [42], Cesari [20]).

All of the examples of derivations given above, from the Dini derivation through to Burkill's general derivates can be expressed vaguely in the form

$$\underline{GD}\, h_g(x) = \liminf_{I \Rightarrow x} h(I)/g(I)$$

and

$$\overline{GD}\, h_g(x) = \limsup_{I \Rightarrow x} h(I)/g(I)$$

under various interpretations as to what $I \Rightarrow x$ (" I shrinks to x ") might mean. Such generalized derivation processes as can be expressed in this manner include a broad spectrum of generalized derivatives and we propose to study just this type of derivative. We must ignore then a number of other ideas such as the introduction of convergence factors (Cesaro derivatives, L_p-derivatives), higher order derivatives (Peano derivatives, second symmetric derivatives), and vector-valued functions, but this still leaves an extensive theory.

To set up such a theory within the context of a derivation basis using the terminology in §1 above let $\mathcal{I}$ denote the family of all nondegenerate compact intervals. Then we shall be studying the following notions.

DEFINITION 1. By a <u>differentiation basis</u> on the real line we mean a filterbase of subsets of $\mathcal{I} \times R$.

DEFINITION 2. If B is a differentiation basis on the real line then by the <u>extreme</u> B-<u>derivates</u> of a function F we mean

$$\overline{D}_B F(x) = \inf_{\beta \in B} \sup_{(I,x) \in \beta} F(I)/|I|$$

and

$$\underline{D}_B F(x) = \sup_{\beta \in B} \inf_{(I,x) \in \beta} F(I)/|I| \quad .$$

This then captures most of the ideas of differentation theory; we in fact will go somewhat further and define a <u>derivation basis</u> B to be any nonempty family of subsets of $\mathcal{I} \times R$. This very general object and the notions that arise from it are the subject of our study and they provide a unified approach to the study of generalized derivations.

§3. Riemann type integrals. We have promised in the previous section that our study is concerned exclusively with derivation bases (i.e., families of subsets of $I \times R$ and especially filterbases on $I \times R$). That this should have even the most remote connection with integrals of Riemann type may appear strange. To see how this connection arises let us set derivation bases and filterbases aside and follow the history of Riemann type integrals.

It was Cauchy [19] who proved that for any continuous function f on an interval $[a,b]$ the limit of the sums $\sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1})$ over partitions of the interval $[a,b]$ with the limit taken in a now familiar sense can be used to compute the integral. It was then only left to Riemann a half century later to take Cauchy's theorem as a definition of "integrability" and to proceed from there. This definition of an integral as a limit of Riemann (Cauchy?) sums has a natural and simple appeal. It also has some decided advantages over more modern integration techniques. For example, (i) a generalization to Stieltjes integrals as $\lim \sum_{i=1}^{n} f(\xi_i)(G(x_i) - G(x_{i-1}))$ is immediate, (ii) generalizations to integrals of arbitrary interval functions as $\lim \sum_{i=1}^{n} h([x_{i-1},x_i])$ giving Burkill or Hellinger integrals is equally immediate, (iii) vector-valued integration presents no barrier - indeed since the process is merely a sum followed by a limit our functions may assume values in a topological semigroup and the formal definition remains unchanged.

This procedure, of defining an integral as a limit of Riemann sums, has been severely discredited by a generation that now considers an integral to be nothing more or less than a countably additive signed measure. Within that viewpoint the Riemann definition of an integral is considered to be intimately joined to a finitely additive measure (Peano-Jordan measure) and hence forever doomed to enjoy no adequate limit theorems; consequently any of the presumptive advantages that such a Riemann type definition possesses should be cheerfully abandoned in favour of the deeper analytic properties available in measure theory. Thus, for example, Luxemberg [79] shows that the convergence properties of the Lebesgue integral arise from the countable additivity of Lebesgue measure. For measure theorists there is no issue here.

However, on the real line there is a certain convenience provided by an integral defined as a limit of Riemann sums; and the Riemann integral lacks no convergence properties - it is merely somewhat short of integrable functions with which to express them. In fact Lebesgue himself showed that his integral was expressible as a limit of Riemann sums (Lebesgue [77]) but the limit proved too intractable to be taken as a definition. The issue is simple: the usual limit operation that is used in the calculus to define the Riemann integral is too coarse and allows too few functions to be integrable, therefore a finer limit operation should be substituted.

If we express this in the correct language the solution will be apparent. By a partition π of the interval $[a,b]$ we mean that $\pi = \{(I_i,x_i) : i = 1,2,\ldots,n\}$ where I_i are nonoverlapping subintervals of $[a,b]$ whose union is $[a,b]$ and x_i is a point of the interval $[a,b]$ (in fact in most applications $x_i \in I_i$). By Π we mean the collection of all such partitions of $[a,b]$. For a function f on $[a,b]$ we write

$$S(f,\pi) = \Sigma_{(I,x)\in\pi} f(x)\ |I|$$

and for any subset Π_0 of Π we write

$$S(f,\Pi_0) = \{S(f,\pi) : \pi\in\Pi_0\} .$$

The usual convergence notion used in the calculus can be described by writing for any $\delta > 0$,

$$\Pi_\delta = \{\pi\in\Pi : \text{ if } (I,x)\in\pi \text{ then } |I| < \delta\} .$$

Then $\{\Pi_\delta : \delta > 0\}$ is a filterbase on Π and convergence of the corresponding filterbase $\{S(f,\Pi_\delta) : \delta > 0\}$ is equivalent to the Riemann integrability of f in the classical sense.

Now we can give a precise formulation of the extension problem for the Riemann integral: find a filterbase F on Π that is finer than $\{\Pi_\delta : \delta > 0\}$ and provides an abundance of functions f for which $S(f,F)$ converges.

Surprisingly this formulation returns us to our starting point. In order to find an appropriate extension of the Riemann integral it would be enough to find a differentiation basis $\mathbf{B}$ such that each element $\beta \in \mathbf{B}$ contains a partition of the interval $[a,b]$. Then if we write for each such β, $\Pi_\beta = \{\pi \in \Pi : \pi \subset \beta\}$ we will have that $\{\Pi_\beta : \beta \in \mathbf{B}\}$ is a filterbase on Π. In fact there are a number of generalized derivations whose differentiation bases have this property that partitions always exist and each such derivation yields an extension of the Riemann integral in the above prescribed manner. The table below tells the story. The differentiation basis that expresses the generalized derivative listed in the first column provides a natural extension of the Riemann integral. The classical name for the extension (where it has been previously named) is listed in the second column.

Except for the Riemann integral which, of course, is defined as a limit of Riemann sums these characterizations are quite recent. Lebesgue did point out ([77]) that his integral could be obtained as a limit of Riemann sums but although a number of authors did pursue the idea the simple characterization here was discovered much later. Independently Kurzweil [72] and Henstock [45] found the characterization of the Denjoy-Perron integral. Henstock pointed out that similar Riemann sums characterizations of the approximate Perron integral (Henstock [47]) and the general Denjoy integral (Henstock [54, p. 222] and the corrected version in [60, pp. 2-3]) were also available. It was McShane [89] and [90] who noted the adjustment needed in order to characterize precisely the Lebesgue integral in this way. The interpretation of these facts within the setting of derivation bases is really very obvious but has received no explicit comment from previous authors.

TABLE 3.1

DIFFERENTIATION BASIS	CORRESPONDING EXTENSION OF THE RIEMANN INTEGRAL
uniform derivation	classical Riemann integral
sharp derivation	classical Lebesgue integral
ordinary derivation	integral of Denjoy and Perron
approximate derivation	approximate Perron integral (of J.C. Burkill)
composite path derivation (under some hypotheses)	general Denjoy integral
selective derivation	(unnamed)
qualitative derivation	(unnamed)

This Table alone should be sufficient motivation to pursue a study of derivation bases. That this collection of integrals which for so long were considered as dramatically different should be so easily unified certainly suggests that the underlying structure is of some interest. But apart from this unification there are other strong reasons for wishing to study these ideas systematically. Numerous other integration procedures on the real line have also been invented and these too can be placed within this setting. Thus the various Stieltjes integrals (the Riemann-Stieltjes, Lebesgue-Stieltjes, Darboux-Stieltjes, Perron-Stieltjes, etc.), the integrals of arbitrary interval functions (Burkill, Hellinger, Burkill-Cesari), and a variety of modifications of these procedures (mean-Stieltjes, modified-Stieltjes, "belated" integrals, Lane integral) can receive a systematic study. As the list of integration procedures on the real line devised to handle specific problems is proving to be endless any attempt to simplify matters should be welcome.

§4. Measure theory. We have seen that the study of differentiation bases leads to a number of solutions to the extension problem for the Riemann integral. There is a parallel problem in measure theory which we discuss here.

Let us begin by sketching the development of measure theory in the nineteenth century. Using the terminology of the previous section we may define the Peano-Jordan outer measure $\overline{m}(E)$ of a set E contained in the interval $[a,b]$ as follows: for any partition π of $[a,b]$ set

$$m(\pi,E) = \Sigma \{|I_i| : (I_i,x_i) \in \pi , x_i \in E\}$$

and for any subset Π_0 of Π define

$$\overline{m}(\Pi_0,E) = \sup \{m(\pi,E) : \pi \in \Pi_0\} .$$

Then for the measure of the set E we take

$$\overline{m}(E) = \lim m(\Pi_\delta, E)$$

in the sense of the filterbase $\{m(\Pi_\delta, E) : \delta > 0\}$. This is essentially the definition of Stolz [ ] although the filter approach does not make this entirely transparent.

This manner of defining a measure has a number of advantages: (i) the generalization to Stieltjes measures $\overline{G}(E)$ is managed merely by replacing $m(\pi, E)$ by $m(\pi, E; G) = \Sigma\{|G(I_i)| : (I_i, x_i) \in \pi, x_i \in E\}$, (ii) a further generalization in the same spirit places the Darboux-Stieltjes upper integral in the same setting - take

$$m(\pi, E; f, G) = \Sigma\{|f(x_i)G(I_i)| : (I_i, x_i) \in \pi, x_i \in E\}$$

and then the resulting measure $(\overline{fG})(E)$ would be the same as the upper Darboux-Stieltjes integral $\overline{\int_a^b} |f(x)| \chi_E(x) d|G|$, (iii) an extension to measures generated by interval functions whose values lie in some abstract structure could be accomplished by replacing $\overline{m}(\Pi_0, E)$ with the set $m(\Pi_0, E) = \{m(\pi, E) : \pi \in \Pi_0\}$ and so producing a set-valued measure.

The Peano-Jordan measure has one useful limit property: if $\{F_n\}$ is a shrinking sequence of closed subsets of $[a,b]$ then $\overline{m}(\cap F_n) = \lim \overline{m}(F_n)$. It also has a property which to a twentieth century eye seems highly undersirable, $\overline{m}(E) = \overline{m}(\overline{E})$ for all $E \subset [a,b]$, but to the mathematicians of the previous century would have seemed a <u>sine qua non</u> of any measure theory. When Lebesgue began searching for a measure theory with which to generalize the integral he was led to an entirely different method of construction, largely because of the influence of Borel. But this method of construction of Stolz can be extended merely by searching for a finer filterbase than $\{\Pi_\delta : \delta > 0\}$.

In fact then there is associated with any derivation basis on the real line a fully developed measure theory and it should be studied along with the corresponding integration theory and differentiation theory. From this viewpoint all the apparatus of Lebesgue's theory of measure arises naturally as part of the study of the ordinary derivative.

Even better is the fact that a number of variational ideas that have played a key role in various studies in differentiation theory and in integration theory are expressible directly in terms of the measures that arise. Thus for a function F there will be a measure F_B corresponding to any derivation basis; in the particular case where B expresses the ordinary derivative this measure F_B carries information as to whether F has bounded variation or is VBG_*, or AC_* or ACG_* on a set. Also because the same structure is used to define the measures, the derivatives, and the integrals, the interrelations are frequently very easy to establish and the role of certain classical hypotheses becomes rather transparent.

§5. The program. Our program in this study is to use the notion of an abstract derivation basis as a unifying concept in the treatment of a variety of ideas in classical real analysis. In particular, by developing in this setting the three basic concepts of the integral, the derivative, and the variation we can give a simpler and more directed account of a great many concerns of real analysts.

A largely suppressed motivation for this study rests on the fact that these ideas should have some considerable impact on the study of differentiation and integration in higher dimensions. Authors such as Mawhin and Pfeffer have obtained some interesting results; Henstock has devoted much time in the development of an abstract theory that will apply in higher dimensional and even infinite dimensional spaces (our bibliography lists his contributions); McShane in [89] and later work has introduced these ideas into the study of stochastic integrals. There are a number of other themes that are being or could be studied (Burkill-Cesari integration for example) in this setting. Although we do not here pursue any of these ideas it is possible that a detailed study of derivation bases on the real line will provide some clues as to the type of results that might be sought in higher dimensions; thus Bruckner [7] has obtained analogues of classical theorems for the Dini derivatives in higher dimensions and one could expect that more of the detailed knowledge of derivatives on the real line could be lifted to higher dimensions.

CHAPTER TWO

DERIVATION BASES

§1. Basic definitions. The focus of attention is on intervals and interval functions. Indeed, as J.C. Burkill pointed out, "almost every process of analysis involves the manipulation of functions of intervals, which are not usually additive" [14, pp. 275-276]. We actually go somewhat further and study what we call "interval-point" functions and shall hazard the proposition that every process of classical real analysis can be realized as the manipulation of such functions. In this section we present the terminology needed for an investigation of these interval-point functions and their related concepts.

(1) [intervals] The collection of all closed bounded intervals is denoted as $\mathcal{I}$. By $\mathcal{I}_+$ we mean all finite unions of intervals.

(2) [interval functions] An interval function is a mapping from $\mathcal{I}$ having real values. We prefer upper case letters F, G, H, etc. for such functions.

If $f : R \to R$ then Δf denotes the interval function defined by $\Delta f([x,y]) = f(y) - f(x)$. For many applications it is best not to distinguish notationally between a function $F : R \to R$ and its corresponding interval function ΔF.

(3) [additive interval functions] An interval function F is additive if $F(I \cup J) = F(I) + F(J)$ for any pair of nonoverlapping intervals I and J for which $I \cup J$ is an interval. Such an F can always be extended to $\mathcal{I}_+$ and this will be done without comment.

(4) [point functions] A function $f : R \to R$ is called a point function, merely to distinguish it from the other types of functions under study.

(5) [interval-point pairs] Our concern in the sequel is with pairs (I,x) where $I \in \mathcal{I}$ and $x \in R$ and with the collection of all such pairs, namely the product set $\mathcal{I} \times R$.

We use lower case greek α, β, γ etc. usually to denote subsets of $\mathcal{I} \times R$.

(6) [partitions] A finite subset π of $\mathcal{I} \times R$ is a partition if

$$\pi = \{(I_i, x_i) : i = 1,2,3,\ldots,n\}$$

has I_i and I_j nonoverlapping for distinct i and j. It is said to be a partition of $\bigcup_{i=1}^{n} I_i$ and where that set (which of course belongs to $\mathcal{I}_+$) is also an interval we have more or less a traditional partition except that the associated points x_i are carried along. Kurzweil [65, p. 515] calls them "pointed partitions", Henstock [55] calls them "divisions", and McLeod [88] calls them "tagged division".

(7) [interval-point functions] A function $h : \mathcal{I} \times \mathcal{R} \to \mathcal{R}$ is called an interval point function. We consider point functions $f : R \to R$ and interval functions $F : \mathcal{I} \to R$ as special cases of interval-point functions by agreeing that $f(I,x) = f(x)$ and $F(I,x) = F(I)$. In this way we even have the product fF defined as an interval-point function, by $(fF)(I,x) = f(x)F(I)$.

The special interval (and hence interval-point) function $I \to |I|$, the length of the interval I is denoted as m so that $m(I) = m(I,x) = |I|$ for all $I \in \mathcal{I}$ and $x \in R$.

(8) [Riemann sums] If h is an interval-point function and π is a partition then the sum

$$\Sigma\ \{h(I,x)\ :\ (I,x) \in \pi\}$$

is our version of a Riemann sum and we use the shorthand $\Sigma_\pi h$, $\Sigma_\pi h(I,x)$, or $\Sigma_{(I,x) \in \pi} h(I,x)$ to denote such sums. In particular the traditional Riemann sum

$$\sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1})$$

assumes the form

$$\Sigma_\pi fm \quad \text{or} \quad \Sigma_\pi f(x)m(I)$$

where π is the partition $\pi = \{I_i, \xi_i) : i = 1,2,\ldots,n\}$ with $I_i = [x_{i-1}, x_i]$.

(9) [derivation basis] A derivation basis is a nonempty collection of subsets of $\mathcal{I} \times R$.

(10) [sections of a derivation basis] For any $\beta \subset \mathcal{I} \times R$ and any $X \subset R$ we write

$$\beta(X) = \{(I,x) \in \beta : I \subset X\} \quad \text{and} \quad \beta[X] = \{(I,x) \in \beta : x \in X\}.$$

Then if B is a derivation basis so too are the objects

$$B(X) = \{\beta(X) : \beta \in B\} \quad \text{and} \quad B[X] = \{\beta[X] : \beta \in B\}$$

which are called sections of the derivation basis B.

The square bracket sections $B[X]$ are used frequently especially in the measure theory. The round bracket sections $B(X)$ will be used only in the case of $X = I$ an interval or $X = G$ an open set. In the latter case the section $B(G)$ would be considered in the language of abstract differentiation theory (cf. Hayes and Pauc [39, p. 12]) a "G-pruning".

(11) [partial order of derivation bases]. The collection of derivation bases is partially ordered in a natural way: if A and B are derivation bases such that given any element $\beta \in B$ there is an element $\alpha \in A$ such that $\alpha \subset \beta$ then we say A is finer than B and write $A \leq B$.

This then induces an equivalence relation: A is said to be equivalent to B and we write $A \cong B$ if $A \leq B$ and $B \leq A$.

(This partial order suggests a natural way of combining two derivation bases B_1 and B_2. One can write $B_1 \vee B_2$ and $B_1 \wedge B_2$ for the two derivation bases

$$B_1 \vee B_2 = \{\beta_1 \cup \beta_2 : \beta_1 \in B_1 , \beta_2 \in B_2\}$$

and

$$B_1 \wedge B_2 = \{\beta_1 \cap \beta_2 : \beta_1 \in B_1 , \beta_2 \in B_2\} .$$

Note that $B_1 \wedge B_2 \leqq B_i \leqq B_1 \vee B_2$ (i=1,2). A number of later ideas permit an expression in this language: for example a derivation basis B that has the property $B \cong B \wedge B$ is said in §4 below to be filtering down. The derivation bases D, RD, and LD that represent the ordinary derivation, right derivation and left derivation (see §3 below) are directly related by the assertion $D \cong RD \vee LD$. We do not intend to use this terminology in the sequel but mention it parenthetically as it may prove useful in some contexts.)

§2. The three fundamental concepts arising from a derivation basis. As has been mentioned in the introduction there are three fundamental objects of study in the theory of derivation bases; loosely these are a differentiation theory, an integration theory, and a measure theory. These concepts dominate the theory.

A. [DIFFERENTIATION THEORY] Let B be a derivation basis and let h and k be interval-point functions.

(a) [exact derivatives] A function f is an exact B-derivative of h relative to k if for every $\varepsilon > 0$ there is a $\beta \in B$ with

$$|h(I,x) - f(x)k(I,x)| \leqq \varepsilon\, |k(I,x)|$$

for all $(I,x) \in \beta$.

In symbols we may write $D_B\, h_k = f$.

In the special case $k = m$ we return to conventional rather than relative derivatives and then f is an exact B-derivative of h if for every $\varepsilon > 0$ there is a $\beta \in B$ with

$$\left| \frac{h(I,x)}{|I|} - f(x) \right| < \varepsilon \quad \text{for all} \quad (I,x) \in \beta$$

and we write $D_B\, h = f$.

(b) [extreme derivates] The upper and lower extreme B-derivates of h relative to k at a point x are

$$\overline{D}_B\, h_k(x) = \inf_{\beta \in B} \sup_{(I,x) \in \beta} h(I,x)/k(I,x)$$

and

$$\underline{D}_B\, h_k(x) = \sup_{\beta \in B} \inf_{(I,x) \in \beta} h(I,x)/k(I,x) ,$$

subject to the interpretation of $0/0$ as 0 and $c/0$ as $+\infty$ or $-\infty$ depending on the sign of c .

In the special case where $k = m$ we return to conventional rather than relative derivates and we do not have to worry about division by zero; the notation will be simply $\overline{D}_B\, h(x)$ and $\underline{D}_B\, h(x)$.

B. [MEASURE THEORY] Let B be a derivation basis and let h be an interval-point function.

(a) [the variation] For any nonempty subset β of $\mathcal{I} \times R$ we write

$$V(h,\beta) = \sup \{\Sigma_\pi |h(I,x)| : \pi \subset \beta , \pi \text{ a partition}\}$$

and refer to $V(h,\beta)$ as the variation of h over β; for $\beta = \emptyset$ we take $V(h,\emptyset) = 0$.

The variation of h over B is defined as

$$V(h,B) = \inf_{\beta \in B} V(h,\beta) .$$

(b) [the variational measure] For any set $X \subset R$ the B-variational measure of h on X is the variation of h over the section B[X] of B , i.e.,

$$h_B (X) = V(h,B[X]) .$$

(c) [the upper integral] For any set $X \subset R$ the B-upper integral of h over X is

$$B\text{-}\overline{\int}_{[X]} d|h| = V(h,B[X]) ,$$

in short, then merely another notation for the measure $h_B (X)$.

C. [INTEGRATION THEORY] Let B be a derivation basis and h an interval-point function.

(a) [the B-integral] The B-integral of h over the interval I_0 is any number c for which, given $\varepsilon > 0$, a $\beta \in B$ can be found so that

$$\left|\Sigma_{(I,x) \in \pi} h(I,x) - c\right| < \varepsilon$$

for every partition π of I_0, $\pi \subset \beta$.

(b) [B-integrable] The function h is B-integrable over I_0 if such a number c exists and is unique; implicit in this is the requirement that every $\beta \in B$ contain a partition of the interval I_0.

If h is B-integrable over an interval I_0 we write this number c as

$$(B) \int_{(I_0)} dh \quad \text{or} \quad \int_{(I_0)} dh$$

if the context is clear.

If h is the product of a point function f and an interval function G then we would prefer the notation

$$\int_{(I_0)} d(fG) = \int_{(I_0)} f\, dG .$$

The use of the round bracket notation in the integral and the square bracket notation in the upper integral is intended to reflect the fact that

$$\int_{(I)} dh$$

arises from the derivation basis B(I) and that

$$\overline{\int}_{[X]} dh$$

arises from the derivation basis B[X] and the connection between the two concepts is only strong in certain examples, but not in general. This really reflects the classical distinctions between integration thought as a "function of sets" or as a "function of intervals" which distinction often causes the student some pain when he thinks simultaneously about the nature of the Riemann-Stieltjes and the Lebesgue-Stieltjes integrals.

§3. Fundamental examples of derivation bases. In order to give some substance to the preceding material, which will certainly strike most readers as forbiddingly abstract, we now present a series of examples to motivate and illustrate the theory. Later examples in the text will make use of this terminology and notation. In each case some brief mention will be given to indicate the nature of the three fundamental concepts as they are realized in that setting.

EXAMPLE 3.1 (The trivial derivation basis) The derivation basis

$$\mathrm{T} = \{\emptyset\}$$

is called the trivial basis. It is finer than every other basis. The variation $V(h,\mathrm{T})$ *clearly vanishes for any interval-point function* h *and so any measure* h_{T} *also vanishes. There is no integration theory available as this basis does not supply any partitions.*

The derivation theory is also satisfyingly trivial in that one always has

$$\underline{D}_{\mathrm{T}}\, h_k(x) = +\infty \quad \text{and} \quad \overline{D}_{\mathrm{T}}\, h_k(x) = -\infty$$

and

$$D_{\mathrm{T}}\, h_k = f$$

is invariably true.

At first sight one might expect to exclude such objects from consideration and restrict the definition of a derivation basis to disallow the inclusion of the empty set; the only reason we do not do so is to preserve the feature that whenever B *is a derivation basis then so too is any section* B[*X*] *or* B(*I*) *. Note however that any basis* B *that contains the empty set is equivalent to* T .

EXAMPLE 3.2 (The uniform derivation basis) For any positive number δ *we write*

$$\beta_\delta = \{(I,x) : I \in \mathcal{I} \text{ , } x \in I, \ |I| < \delta\}$$

and refer to the derivation basis

$$\mathrm{U} = \{\beta_\delta : \delta > 0\}$$

as the uniform basis.

It should be apparent that U *expresses uniform derivatives and the Riemann and Riemann-Stieltjes integration procedures. It is less obvious that* m_{U} *yields the classical Peano-Jordan measure.*

EXAMPLE 3.3 (The refinement basis) Let S *be any set of real numbers that has no point of accumulation and for any such* S *write*

$$\beta_S = \{([a,b],x) : a < b \text{ , } a \le x \le b \text{ , } S \cap (a,b) = \emptyset\}$$

and

$$\mathrm{R} = \{\beta_S : S \subset R \text{ , } S \text{ has no accumulation points}\} \text{ .}$$

This basis has its greatest interest in the integration theory and expresses an integral based on a familiar partition-refinement type limit as has been used in the study of certain versions of the Riemann-Stieltjes integral (e.g., Pollard [100]*, Getchell* [35]*). The differentiation theory is essentially the same as that for* U *and just expresses uniform derivatives.*

Note that R *is finer than* U, *i.e., in symbols* $R \leq U$, *since if a* $\beta \in U$ *is given one can select a set* S *so that* $\beta_S \subset \beta$. *In particular the integration theory provided by* R *is slightly more general than that given by* U *(as is familiar from elementary analysis) although both give precisely the classical Riemann integral when applied to a function of the form* $f(x)m(I)$. *The difference between the two bases emerges for Stieltjes type integrals* $\int fdg$; *in our setting the familiar advantage that the basis* R *has over* U *is expressed by the fact that* R *has the additive property defined in §4 below.*

EXAMPLE 3.4 (The ordinary derivation basis) For any positive function δ *on* R *define the collections* β_δ *and* $\beta_\delta^{\,o}$ *by*

$$\beta_\delta^{\,o} = \{(I,x) : I \in \mathcal{I} ,\ x \in I ,\ |I| < \delta(x)\}$$

and

$$\beta_\delta = \{(I,x) : I \in \mathcal{I} ,\ x \ \text{ an endpoint of } I ,\ |I| < \delta(x)\} .$$

Then the derivation bases

$$D^o = \{\beta_\delta^{\,o} : \delta \ \text{ a positive function on } R\}$$

and

$$D = \{\beta_\delta : \delta \ \text{ a positive function on } R\}$$

are two versions of the ordinary derivation basis. The former will be called a full version and the latter an endpoint tagged version.

It is not difficult to see that the extreme derivates $\underline{D}_D F(x)$ *and* $\overline{D}_D F(x)$ *are just the usual bilateral extreme derivatives of* F , $\underline{D} F(x)$ *and* $\overline{D} F(x)$, *and that the assertion* $D_D F = f$ *just says that* f *is the derivative of* F .

It is far from obvious though that the measure theory here gives m_D *and* m_{D^o} *equivalent to Lebesgue outer measure on the line and that the integration theory yields the Denjoy-Perron and Denjoy-Perron-Stieltjes integrals.*

These two bases treat additive interval functions in the same way and only diverge in their treatment of nonadditive interval functions or arbitrary interval-point functions. Mainly D *is to be preferred as it has some sharper properties. This feature of ordinary derivation is common: there are a number of applications where a choice between a full version or an endpoint tagged version needs to be made differently depending on the theorem desired.*

Note that $D \le D^o \le U$ *and* $D \le R$.

EXAMPLE 3.5 (The sharp derivation basis) In the basis D *we required of the pairs* (I,x) *appearing that* x *be an endpoint of* I . *If we remove the restriction that* x *even belong to the interval* I *and adjust so that* I *is "close" to* x *then we arrive at the sharp derivation process. For a positive function* δ *on* R *write*

$$\beta_\delta^{\#} = \{(I,x) : I \in \mathcal{I} , x \in R , I \subset (x - \delta(x) , x + \delta(x))\}$$

and then

$$D^{\#} = \{\beta_\delta^{\#} : \delta \text{ a positive function on } R\}$$

is called the sharp derivation basis.

It should be clear that the $D^{\#}$ *derivates and the* $D^{\#}$ *derivative is just the "sharp" differentiation we have described in the introduction; it is however surprising and perhaps curious that the integration theory developed by this basis includes a characterization of the Lebesgue integral.*

Note that $D \le D^{\#}$ *and* $D^o \le D^{\#}$ *but that neither* U *nor* R *are comparable with* $D^{\#}$. *It is useful to introduce a sharp version of the uniform basis by writing for any positive number* δ ,

$$\beta_\delta^{\#} = \{(I,x) : I \in \mathcal{I} , x \in R , I \subset (x - \delta, x + \delta)\}$$

and

$$U^{\#} = \{\beta_\delta^{\#} : \delta \text{ a positive number}\} .$$

Then certainly $U \le U^{\#}$ *and* $D^{\#} \le U^{\#}$ *; interestingly, and significantly,* *the integration theory for the* $\mathrm{U}^{\#}$ *basis gives the Riemann integral where one might have expected from these relations that it would only give a restricted version of it.*

EXAMPLE 3.6 (The dual of the ordinary derivation basis) *The basis* D *has a dual in a sense that will be formalized in §8 below.* D^* *denotes the derivation basis that contains all elements* $\beta^* \subset I \times R$ *that have the following property: if* $x \in R$ *and* $\varepsilon > 0$ *are given then there is at least one number* $y \neq x$ *such that* $|y - x| < \varepsilon$ *and* $([x,y],x)$ *(or* $([y,x],x)$*) belongs to* β^* *.*

This basis has the remarkable feature that it reverses the roles of the upper and lower extreme derivates: for a function F *the* D *and* D^* *derivates are related by*

$$\underline{D}_{\mathrm{D}} F(x) = \overline{D}_{\mathrm{D}^*} F(x) \quad \text{and} \quad \overline{D}_{\mathrm{D}} F(x) = \underline{D}_{\mathrm{D}^*} F(x) .$$

Also an assertion $D_{\mathrm{D}^*} F = f$ *is equivalent to the fact that at each point* x *the number* $f(x)$ *is a derived number for the function* F *(so that in particular such* f *need not at all be unique).*

There is no integration theory available as D^* *need not contain enough partitions but the measure theory does play a role in the subsequent theory. Note that* $\mathrm{D}^* \leq \mathrm{D}$ *, a fact that will not hold necessarily for all dual bases but does for most familiar ones.*

EXAMPLE 3.7 (Dini derivation) *In order to express one-sided derivatives and Dini derivatives we need one-sided versions of* D *. For any positive function* δ *on* R *define*

$$R\beta_\delta = \{([x,y],x) : 0 < y - x < \delta(x)\},$$

$$L\beta_\delta = \{([y,x],x) : 0 < x - y < \delta(x)\},$$

$$\mathrm{RD} = \{R\beta_\delta : \delta \text{ a positive function on } R\}$$

and

$$\mathrm{LD} = \{L\beta_\delta : \delta \text{ a positive function on } R\} .$$

We refer to these as the right and left Dini derivation bases. Note that $\mathrm{RD} \le \mathrm{D}$ *and* $\mathrm{LD} \le \mathrm{D}$ *and that* $\mathrm{D}^* \le \mathrm{RD}$ *and* $\mathrm{D}^* \le \mathrm{LD}$ *(in fact* $\mathrm{D}^* \supset \mathrm{RD} \cup \mathrm{LD}$*).*

EXAMPLE 3.8 (Filtered or natural derivation) Suppose that there is given a system $\{N(x) : x \in R\}$ *of filters such that each* $N(x)$ *is a filter converging to* x *. The derivation process generated by such a system of filters can be described by the following derivation basis. A* <u>*choice*</u> *relative to the system* $\{N(x) : x \in R\}$ *is a function* η *on* R *such that each* $\eta_x \in N(x)$ *: corresponding to any choice* η *we write*

$$\beta_\eta = \{([y,z],x) : y = x , z > x , z \in \eta_x \text{ or } z = x , y < x , y \in \eta_x\}$$

and

$$\mathrm{N} = \{\beta_\eta : \eta \text{ a choice}\} .$$

This type of basis is sufficiently general that it can be studied and characterized in our abstract terminology and henceforth when it appears it will not be hidden in italics but will join the main body of the text.

For various choices of the system N *one has the derivation bases* D, RD *or* LD . *If each* $N(x)$ *is defined to be the collection of all sets having inner density* 1 *at* x *then the corresponding derivation basis is written as* A *and called the approximate derivation basis.*

EXAMPLE 3.9 (Composite path derivation) In order to define a derivation basis that captures the notion of a composite path derivative (as given in §2 of Chapter One) we must proceed in a slightly different manner: given a sequence $E = \{E_n\}$ of closed sets such that $\cup E_n = R$ and a positive function δ on $R \times N$ (i.e., for $x \in R$ and for $n = 1,2,3,\ldots$ there is a positive number $\delta(x,n)$ defined) we write

$$\beta_{E,\delta} = \{([x,y],z) : \text{for some } n \in N ,\ x,y,z \in E_n \\ x \le z \le y \text{ and } 0 < y-x < \delta(z,n)\} .$$

Then the basis C_E defined for a fixed sequence $E = \{E_n\}$ is

$$C_E = \{\beta_{E,\delta} : \delta \text{ a positive function on } R \times N\} .$$

The properties of the basis C_E depend usually on properties possessed by the sequence E. In order for $D_{C_E} F = f$ to hold it is necessary and sufficient that $f(x)$ be the derivative of F at x relative to any set E_n that contains x.

For a more general basis write

$$C = \cup \{C_E : E = \{E_n\} \text{ any sequence of closed sets covering } R\} .$$

This basis expresses the general notion of composite path derivation. In particular if there is some sequence $\{P_n\}$ of closed sets covering R for which $f(x)$ is the derivative of F relative to each P_n containing x, then $D_C F = f$. Note that $C \leqq C_E$ for any sequence E.

EXAMPLE 3.10 (The symmetric derivation basis) The symmetric derivative can be expressed naturally by a derivation basis as follows: if δ *is a positive function on* R *denote* β_δ^s *as*

$$\beta_\delta^s = \{([x-h,x+h],x) \ : \ 0 < h < \delta(x)\}$$

and

$$S = \{\beta_\delta^s \ : \ \delta \text{ a positive function on } R\} \ .$$

Except for the obvious relation $S \leq D^o$ *this basis is quite remote both in methods and properties from the other bases defined above.*

§4. Elementary properties. The definition of a derivation basis requires only that one has a nonempty collection of subsets of $I \times R$. As the example of the trivial basis (Example 3.1) shows there can be no theorems at all in this generality. This is similar to the situation in a development of topology where a general topological space has scarcely any genuine theorems; as in topology where the theorems flow from strong assumptions (separation properties for example) we need here a number of assumptions that can be used to develop a theory. In this section we outline just the simplest and most elementary of these and present a few examples to help illustrate. In §5, §6, and §7 below we will investigate the deeper and more powerful assumptions that are needed to give the heavier results of the theory.

DEFINITION 4.1 [FILTERING DOWN] A derivation basis B is said to be filtering down if for every β_1 and β_2 in B there is an element $\beta_3 \in B$ with

$$\beta_3 \subset \beta_1 \cap \beta_2 \ .$$

The main intention of this property is to allow limit operations, for then (unless B is trivial) B is a filterbase and the derivation theory and the integration theory that arise from B can be expressed in terms of such a limit. In Henstock's development of this type of theory he uses the phrase "directed in the sense of divisions" in more or less the same sense (see Henstock [55]) and in his lectures refers to this informally as "B shrinks". If B is nontrivial (i.e., not equivalent to the trivial basis) and is filtering down one can replace B by the filter $\tilde{B}$ on $I \times R$ generated by B and then, since $\tilde{B} \cong B$, the ensuing theory is unchanged.

EXAMPLE 4.2 The derivation bases T , U , R , $D^\#$, $U^\#$, D , D^0 , N , *and* S *are all filtering down. The example* D^* *(the dual of* D*) is not filtering down and this accounts for some of its unorthodox behaviour (as for example the fact that a* D^**-derivative need not be unique).*

DEFINITION 4.3 [STRADDLED, ENDPOINT TAGGED] A derivation basis B is said to be straddled if for every $\beta \in B$ and any pair $(I,x) \in \beta$ one has $x \in I$. A derivation basis is said to be endpoint tagged if moreover for such pairs (I,x) the point x must even be an endpoint of the interval I .

EXAMPLE 4.4 All the bases we have defined in the previous section are straddled with the exception only of the sharp bases $U^\#$ *and* $D^\#$.

Only the bases D , RD , LD , N *have been defined in such a way as to be endpoint tagged. The basis* D^* *has not been defined this way but a moments reflection will show that there is a basis equivalent to* D^* *that is endpoint tagged.*

DEFINITION 4.5 [SEPARATION PROPERTIES] Two subsets β_1 and β_2 are said to be separated if for any pair $(I_1,x_1) \in \beta_1$ and any pair $(I_2,x_2) \in \beta_2$ the intervals I_1 and I_2 are nonoverlapping. A derivation basis B is said to separate two sets $X, Y \subset R$ if there is an element $\beta \in B$ such that $\beta[X]$ and $\beta[Y]$ are separated.

EXAMPLE 4.6 The basis D *separates any two sets* X *and* Y *that are topologically separated* (*i.e.*, *such that there are disjoint open sets* G_1 *and* G_2 *with* $X \subset G_1$ *and* $Y \subset G_2$); *the basis* U *does not have this property.*

DEFINITION 4.7 [ADDITIVE PROPERTY] A derivation basis B is said to be additive if for every interval I, $B \leq B(I) \cup B(R \setminus I^0)$. Equivalently this says that for any pair β_1, β_2 from B and any interval I there is a $\beta \in B$ for which

$$\beta \subset \beta_1(I) \cup \beta_2(R \setminus I^0) .$$

We can say that β splits at I here in the sense that for every pair $(J,x) \in \beta$ either $J \subset I$ or else J and I do not overlap.

EXAMPLE 4.8 The uniform derivation basis U *and the refinement basis* R *help demonstrate the additive property. The former is not additive and the latter is. It is precisely this fact that makes the basis* R *preferable in defining Stieltjes-type integrals. This distinction also shows that the two versions of the ordinary derivation basis* D^0 *and* D *are slightly different in their effect: the former is not additive while the latter, the endpoint tagged version, is additive and again is to be preferred in a development of a Stieltjes-type integral.*

DEFINITION 4.9 [FINER THAN THE TOPOLOGY] A derivation basis B is said to be finer than the topology if for every open set G ,

$$B[G] \leq B(G) .$$

Equivalently this asserts that for every element $\beta_1 \in B$ and every open set G there is an element $\beta_2 \in B$ for which

$$\beta_2[G] \subset \beta_1(G) .$$

EXAMPLE 4.10 Neither of the bases U *nor* R *are finer than the topology although all of the other bases we have defined are.*

DEFINITION 4.11 [IGNORES A POINT] A derivation basis B is said to ignore a point x if there is an element $\beta \in B$ for which there is no pair $(I,x) \in \beta$ for any $I \in \mathcal{I}$. Equivalently B ignores the point x if the section $B[\{x\}]$ is equivalent to the trivial basis.

EXAMPLE 4.12 Usually we will invoke definition 4.11 in a form to exclude the possibility that a basis ignores a point. Most of the bases we have so far defined ignore no point. However, a basis may ignore every point and yet not be trivial (i.e., is not equivalent to the trivial basis). The composite path basis has been defined in such a way that it ignores every point.

We give a further example here of a basis which ignores each point, is nontrivial and might have some interest. From the ordinary derivation basis D *we collect sections as follows:*

$$\mathrm{DD} = \{\beta[R\setminus N] \ : \ \beta \in \mathrm{D} \ \text{and} \ N \subset R \ \text{is a set of measure zero}\} .$$

While this basis does inherit some properties from D *it nonetheless ignores every point. The statement* $D_{\mathrm{DD}} F = f$ *is equivalent to the assertion* $F'(x) = f(x)$ *almost everywhere.*

§5. The partitioning property. One of the most interesting and useful properties to emerge in this study of derivation bases arises from the observation that a number of important derivation bases contain an abundance of partitions. While the properties of the previous section have only relatively minor and predictable consequences in the theory the partitioning property has deep and far ranging consequences.

DEFINITION 5.1 A derivation basis B is said to have the partitioning property if for every $\beta \in B$ and every interval I there is a partition π of I contained in β.

EXAMPLE 5.2 The bases U *(the uniform derivation basis) and* R *(the refinement basis) can be easily seen to have the partitioning property. Of course it is this fact that permits a Riemann integral to be defined in terms of* U*-limits or* R*-limits.*

EXAMPLE 5.3 (The partitioning property for the ordinary derivation basis) The partitioning property for the bases D *and* D^o *follows from an elementary compactness argument; indeed this property is equivalent to compactness and the various characterizations of compactness (Heine-Borel, Bolzano-Weierstrass) can be deduced in turn from the partitioning property (see Example 5.6 below). This has a curious history. It was proved, of course, by Henstock* [45] *and by Kurzweil* [72] *in their development of the generalized Riemann integral that now bears their name. Henstock since has traced the history of the idea back further; in* [50, *p. 124*] *he provides the references Hildebrandt* [63], *W.H. Young and G.C. Young* [118], *and Lusin* [78] *all of whom use an idea of this sort. Recently the history has been pushed back even further to the close of the last century by the references Goursat* [36] *and Cousin* [21] *(we have obtained these latter two references through Professor G. Goodman). Possibly one can go even earlier but for now we should, as does Mawhin* [85], *label this "Cousin's lemma".*

EXAMPLE 5.4 (The partitioning property for the approximate derivation basis) That the approximate derivation basis also possesses this property was first indicated by Henstock [~~47~~ 4?] using a standard category argument. This same type of argument has been formalized by Romanovski (see Kestelman [69, p. 217]) and can be considered an abstract presentation of a partitioning property. Romanovski's lemma can be written as follows:

> *let* ∇ *be a collection of intervals that is hereditary (i.e., if* $I \subset J \in \nabla$ *then* $I \in \nabla$*), that contains every* I *with the property* $J \subset I^0 \Rightarrow J \in \nabla$ *, and that has the property: if* P *is perfect and every interval contiguous to* P *is in* ∇ *then for some* $I \in \nabla$ *,* $I^0 \cap P \neq \emptyset$ *. Then* ∇ *contains a partition of every interval.*

The role that Romanovski's lemma plays in the theory of integration is closely related to the role the partitioning property in general must play.

This partitioning property has many implications for the differentiation theory, the measure theory, and the integration theory associated with a derivation basis. Immediately we see that the precise setting in which the integration theory must take place is in that of a derivation basis that is filtering down and which enjoys the partitioning property. Indeed corresponding to any such derivation basis is a Riemann type integral.

This property has many far reaching effects; we will isolate one particularly useful one. If P is a class of intervals it is said to be additive if whenever $[a,b]$ and $[b,c]$ belong to P so also does $[a,c]$, and P is said to be B-local for a derivation basis B if there is in B at least one element β such that for every $(I,x) \in \beta$ it is the case that $I \in P$. We then have the following elementary, but useful, theorem. We call this the partitioning argument: loosely it asserts that if a property of intervals is additive and holds locally for any appropriate derivation basis, then that property holds globally.

THEOREM 5.5 [the partitioning argument] Let $\mathcal{P}$ be a class of intervals that is additive and B-local for a derivation basis B that has the partitioning property. Then $\mathcal{P}$ contains all intervals.

The following examples illustrate the variety of applications of this principle.

EXAMPLE 5.6 (The Bolzano-Weierstrass theorem) The partitioning property for the derivation basis D *is proved by a compactness argument; conversely the partitioning argument can establish any of the several equivalent expressions of compactness on the real line. For an example we show that any set S that has no accumulation points must be finite in any interval which is the Bolzano-Weierstrass theorem. Say that* $[a,b]$ *belongs to* $\mathcal{P}$ *if and only if* $S \cap [a,b]$ *is finite. Then clearly* $\mathcal{P}$ *is additive and since S has no accumulation points it is easy to see that* $\mathcal{P}$ *is* D*-local. By the partitioning argument it follows that* $\mathcal{P}$ *contains every interval and this proves our claim. Note that the basis* $D^{\#}$ *(which also has the partitioning property) could have been used here.*

Similar arguments can be used to obtain the Heine-Borel theorem.

EXAMPLE 5.7 (A monotonicity theorem) Let F be a real function whose lower bilateral derivate $\underline{D}\,F(x)$ *is everywhere positive. Then a compactness argument would show that F is strictly increasing. The same can be based on a partitioning argument using the derivation basis* D *: let* $[a,b]$ *belong to* $\mathcal{P}$ *if and only if* $F(b) - F(a) > 0$ *. Then* $\mathcal{P}$ *is clearly additive and our assumptions on the derivate of F show that* $\mathcal{P}$ *is* D*-local. Thus every interval* $[a,b]$ *belongs to* $\mathcal{P}$ *and so F is strictly increasing.*

While this is not too surprising an application it is remarkable that the same argument will give a monotonicity theorem for a number of generalized derivates provided only that the associated derivation basis has the partitioning property. Thus one has this same feature for approximate derivates, preponderant derivates, qualitative derivates, and selective derivates with an identical simple proof in each case.

EXAMPLE 5.8 (A Darboux property) If F is an everywhere approximately continuous function then it has the Darboux property. To prove this using a partitioning argument let A *denote the derivation basis that expresses approximate derivation and which is known to have the partitioning property (see §9 below for more details). It is enough if we show that F nonvanishing would require that it be always positive or else always negative. Say that [a,b] belongs to P if and only if F(a)F(b) > 0 : then P is additive and an easy argument shows, since F does not vanish, that P is* A*-local. By the partitioning argument P contains every interval and that proves our claim.*

This same argument can be used for a variety of generalized notions of continuity to show that functions continuous in such a sense must have the Darboux property (e.g., preponderantly continuous, selectively continuous).

The partitioning argument is just a formulation of some very familiar arguments in analysis. The formalization is convenient in that it can suggest methods of proof that might not otherwise come to mind; for example many properties of approximate derivatives and approximately continuous functions can be proved by this argument that have in the past been approached by more complicated methods. Note that these ideas have been given formal treatments in the past: thus there are the "full covering properties" of Thomson [115], the "interval-additive propositions" of Ford [32], the "creeping lemma" of Moss and Roberts [92], and the "local, additive families" of Shanahan [108].

The partitioning property has an important implication for the variation of additive interval functions and nonnegative subadditive interval functions which we express in the following lemma.

LEMMA 5.9 Let B be a derivation basis that has the partitioning property and let h be an additive interval function or a nonnegative subadditive interval function. If V(h,B(I)) = 0 for an interval I then h vanishes on every subinterval of I ; if V(h,B) = 0 then h vanishes identically.

PROOF. Since $|h|$ is a nonnegative subadditive interval function when h is additive it is enough to prove the lemma for the subadditive case. For any $\varepsilon > 0$ if $V(h,B(I)) = 0$ then there is a $\beta \in B$ for which $V(h,\beta(I)) < \varepsilon$. If $J \subset I$ then, by partitioning property, there is a partition π of J with $\pi \subset \beta$. Thus

$$h(J) \le \Sigma_{(I',x') \in \pi} h(I') \le V(h,\beta(I)) < \varepsilon$$

and so $h(J) < \varepsilon$ for all $J \subset I$ and all $\varepsilon > 0$. Hence h vanishes on every subinterval of I as required. The final assertion of the lemma now follows easily since if $V(h,B) = 0$ then $V(h,B(I)) = 0$ for every interval I .

This property of derivation bases that have the partitioning property is important on its own and can be used even when the basis is not partitioning.

DEFINITION 5.10 [H-COMPLETENESS] Let B be a derivation basis and H a family of nonnegative subadditive interval functions. B is said to be H-complete if for any $h \in H$ for which $V(h,B) = 0$ one has $h \equiv 0$.

An important special case occurs if C is the collection of all continuous subadditive interval functions: such an h is continuous if for every $\varepsilon > 0$ there is a $\delta > 0$ so that $h(I) < \varepsilon$ whenever $|I| < \delta$. A basis that is at least C-complete has a number of desirable properties. C-completeness or H-completeness relative to any collection H is, by lemma 5.9, weaker than the partitioning property. The terminology is from Henstock [50].

EXAMPLE 5.11 (The symmetric basis is C-complete) The derivation basis S *that expresses symmetric derivation does not have the partitioning property but is C-complete. This follows from a weak type of partitioning property available for* S *that a student of Henstock's, J.J. McGrotty* [87], *has discovered:*

let $\beta \in S$, *and* $c \in R$ *then there must exist a set* $C \subset (0, +\infty)$ *that is both closed and countable such that* β *contains a partition of every interval of the form* $[c - t, c + t]$ *for* $t \notin C$.

This suggests however a modification of the symmetric derivation basis that could possibly be used to develop an integration theory (in particular that would invert the symmetric derivatives of continuous functions). To each $\beta \in S$ *one can find a countable set* N_β *so that* β *contains a partition of every interval* $[x,y]$ *for which* $x \notin N_\beta$ *and* $y \notin N_\beta$. *Then if one defines*

$$PS = \{\beta \cup \beta' \ : \ \beta \in S, \ \beta' \in D[N_\beta]\}$$

this enlarged derivation basis should have the partitioning property, should be filtering down, etc.. We leave it as a query as to whether this can be pursued to give an interesting integration theory, and especially as to whether it is related to Denjoy's symmetric totalization process (Denjoy [25]).

EXAMPLE 5.12 (The Dini derivation bases are C-complete) It is easy to see that neither of the bases RD *nor* LD *that express the one-sided derivations have the partitioning property. That they are C-complete can be obtained from a form of a partitioning property that is related to a notion of Lebesgue:*

if $\beta \in RD$ *and* $[a,b]$ *is an interval then there is a transfinite increasing sequence* $\{\xi_i\}$ *such that* a *and* b *appear in the sequence and every pair.*

$$([\xi_i, t], \xi_i) \ \text{for} \ \xi_i < t \le \xi_{i+1}$$

belongs to β.

On the basis of the above property of RD *we can see that* RD *is C-complete; for if* h *is a continuous nonnegative subadditive interval function and* $V(h,\beta) < \varepsilon$ *for some* $\beta \in$ RD *then for any interval* $[a,b]$ *there is such a sequence* $\{\xi_i\}$ *linking* a *and* b *and by arguing along the sequence in an obvious fashion one obtains that* $h([a,b]) < \varepsilon$ *, and hence the result. (This can be proved of course much more simply than employing the above "Lebesgue chain" but it seems appropriate here to connect the idea with some kind of partitioning argument.)*

(Perhaps more surprising is the fact that the dual basis RD* *defined in §8 below has almost the same "partitioning property" as that above: for that basis each* $\beta^* \in$ RD* *permits such a transfinite sequence* $\{\xi_i\}$ *for which each term* $([\xi_i , \xi_{i+1}] , \xi_i) \in \beta^*$ *and so it too must be C-complete.)*

EXAMPLE 5.13 (Composite path derivation) For a fixed sequence $\{E_n\}$ *of closed sets covering the real line we have defined the composite derivation basis (relative to* $E = \{E_n\}$*) as* $C_E = \{\beta_{E,\delta} ; \delta\}$ *where*

$$\beta_{E,\delta} = \{([x,y] , z : x,y,z \in E_n \text{ for some } n, x \leq z \leq y , \text{ and } 0 < y - x < \delta(z,n)\} .$$

We can prove the following: <u>*if for each*</u> n <u>*every point of*</u> E_n <u>*is a point of bilateral accumulation of*</u> E_{n+1} <u>*then*</u> C_E <u>*has the partitioning property.*</u>

The general composite derivation basis C *which is essentially just* $\cup \{C_E : \text{all such } E = \{E_n\}\}$ *does not have the partitioning property and as it stands is not quite suitable for a Riemann-type integration theory. Henstock suggests the following modification (Henstock [60, p. 3]): to each sequence* $E = \{E_n\}$ *of closed sets covering* R *let* A_E *denote the points in any* E_n *that are isolated on one side at least in* E_n *; this set* A_E *is evidently countable (*<u>*cf.*</u> *Saks [105, p. 260]). Define*

$$\text{PC} = \{\beta_{E,\delta} \cup \beta' : \beta_{E,\delta} \in C_E , \beta' \in D[A_E] , \text{ all sequences } E = \{E_n\}\} .$$

Then the arguments in Henstock [60, *p.* 3] *can be used to show that* PC *has the partitioning property and Henstock claims that the resulting integral will be, because of Tolstov* [116], *equivalent to the general integral of Denjoy.*

§6. Local character. The derivation basis U which expresses uniform derivation and the derivation basis D which expresses ordinary derivation have a number of features in common: both are filtering down and both have the partitioning property. Indeed $D \leq U$ so that D is a finer filterbase than U . The integration theories that arise in the two theories are of course similar in many ways; but the integration theory for D extends the integration theory for U in some powerful ways. The scope of the limit theorems available in the D setting is truly impressive. It is natural to ask for the abstract property resident in the derivation basis that expresses these limit properties. Some authors have fixed on the fact that U is defined as $\{\beta_\delta : \delta > 0\}$ for fixed positive numbers δ and that D employs positive functions δ . Thus we have the introduction of the function δ as a "gauge" and attention focused on this distinction (e.g., in McLeod [88]). McShane interpreted this distinction in his setting ([89]) by the phrase "pointwise character". The clearest expression of this property is however given in Henstock [55] using the terminology "decomposable" and "fully decomposable". We will use different language so that we can reserve the term "decomposition" for a concept that is a decomposition in the usual sense of that word.

DEFINITION 6.1 [LOCAL CHARACTER] A derivation basis B is said to have local character if

$$B \leqq \bigcup_{x \in R} B[\{x\}] .$$

Less compactly this requirement can be written as follows: if to each $x \in R$ an element $\beta_x \in B$ is given then there is a $\beta \in B$ for which $\beta \subset \bigcup_{x \in R} \beta_x[\{x\}]$ (i.e., for which $\beta[\{x\}] \subset \beta_x$ for all $x \in R$). Loosely put this just asserts that B is completely determined by its sections $B[\{x\}]$ as is the case for the ordinary derivation basis but is not the case for the uniform basis.

Henstock also uses a weaker concept in his investigations. The stronger concept (local character) is a natural one to impose in differentiation theory, but for a great many of the results of his integration theory either in our setting on the real line or in a very abstract setting one only needs a "countable" version of this property.

DEFINITION 6.2 [σ-LOCAL CHARACTER] A derivation basis B is said to have σ-local character if for every sequence of disjointed sets $\{X_n\}$ one has

$$B[\bigcup_{n=1}^{\infty} X_n] \leqq \bigcup_{n=1}^{\infty} B[X_n] .$$

Less compactly written this says that given any such sequence $\{X_n\}$ and any sequence $\{\beta_n\} \subset B$ there is a $\beta \in B$ for which

$$\beta[X_n] \subset \beta_n \quad \text{for all } n .$$

Certainly a basis that has local character must have σ-local character; the converse is not true as the example of the basis DD below (Example 6.5) will show.

In our development we will see that these properties of local character and σ-local character will provide all of the needed convergence results. In particular in their presence the measures h_B become genuine outer measures (i.e., countably subadditive set functions) and the integrals $\int_{(I)} f dh$ assume all of the power of the Lebesgue integral as regards taking limits inside the integral. Indeed, as Henstock has remarked [ ], these properties in our setting do for integration theory what "countable additivity" does in the measure theoretic setting.

The examples indicate which bases possess which properties.

EXAMPLE 6.3 Neither U *nor* R *have local character or σ-local character. They do have a weaker version of these properties however: if* $\{\beta_n\}$ *is a* *finite* *sequence of elements of* U *(or of* R*) and* $\{X_n\}$ *a corresponding sequence of disjointed sets then there is an element* $\beta \in$ U *(or in* R*) such that* $\beta[X_n] \subset \beta_n$ *for each* n .

EXAMPLE 6.4 All of the bases D, D^o, D^*, $D^\#$, N, S, RD, *and* LD *have local character merely because of the manner in which they were defined.*

EXAMPLE 6.5 (A basis with σ-local character but not local character) The basis DD *defined by sectioning the ordinary basis* D *by sets of full measure,* *i.e.*,

$$DD = \{\beta[R\setminus N] \;:\; \beta \in D \text{ and } N \subset R \text{ of measure zero}\} ,$$

does not have local character because if $x \in R$ *is given there is in* DD *an element* β_x *with* $\beta_x[\{x\}] = \emptyset$ *so that local character would imply that* $\emptyset \in$ DD *which is false. It does have σ-local character however. To see this let* $\{X_k\}$ *be a sequence of disjointed sets and* β_k *a sequence from* DD *: for each* β_k *there is an element* $\beta_{k'}$ *from* D *and a set* N_k *of measure zero so that* $\beta_k = \beta_{k'}[R\setminus N_k]$. *Define*

$$N = \bigcup_{k=1}^{\infty} N_k$$

and select a β' *in* D *so that* $\beta'[X_k] \subset \beta_{k'}$; *then the element* β *in* DD *defined as* $\beta = \beta'[R\setminus N]$ *is the one needed to verify the σ-local character property.*

This same observation applies to the modified version of the symmetric basis PS *(Example 5.10). Neither has local character but each can be proved to have* σ*-local character.*

§7. Decomposition properties. We begin by illustrating a most common device in differentiation theory. For example, in the study of the Dini derivative $^{+}\underline{D}\, F(x)$ one might encounter a situation in which everywhere on a set X one has $^{+}\underline{D}\, F(x) > c$. In some sense then one expects the function $F(x) - cx$ to be increasing on X although that cannot precisely be the case. What is available, however, is a decomposition of the set X into a sequence of sets $\{X_n\}$ on each of which the function $F(x) - cx$ is indeed increasing. To see how this may be done, classically, we choose a number $0 < \delta(x) < 1$ at each point $x \in X$ so that for any $0 < y - x < \delta(x)$ one has $F(y) - F(x) > c(y-x)$. This function δ induces a decomposition on X by writing

$$X_{mj} = \{x \in X : 2^{-m-1} \le \delta(x) < 2^{-m}\} \quad \cap \; [j2^{-m-1}, (j+1)2^{-m-1})$$

for $m = 0,1,2,3,\ldots$ and $j = 0, \pm 1, \pm 2, \pm 3, \ldots$. On each set X_{mj} it is easy to see that $F(t) - ct$ is increasing.

This device is much used in differentiation theory in the study of all manner of generalized derivatives but in spite of the fact that it recurs as a theme in a number of instances it has not been singled out and formalized prior to its mention in Bruckner, O'Malley, and Thomson [13] as a "δ-decomposition". This will play a role in our theory as well.

DEFINITION 7.1 Let δ be a positive function defined on a set X. Then by a δ-decomposition of the set X we mean a sequence $\{X_n\}$ which is a relabelling of the double sequence $\{X_{mj}\}$ defined above. Such a decomposition has the following properties:

(i) $\{X_n\}$ is a disjointed sequence whose union is X ,

(ii) if two points x and y belong to the same set X_n then $|y-x| < \min\{\delta(x), \delta(y)\}$, and

(iii) if $x \in X$ is a point of accumulation of some set X_n then every point y in the set $X_n \cap (x-\delta(x), x+\delta(x))$ has $|y-x| < \min\{\delta(x), \delta(y)\}$.

DEFINITION 7.2 Let B be a derivation basis and let $X \subset R$. By a decomposition property of the derivation basis B we mean an assertion that for every $\beta \in B$ there exists a decomposition $\{X_n\}$ of the set X for which each section $\beta[X_n]$ enjoys some stated property.

DEFINITION 7.3 Let B be a derivation basis and let $X \subset R$ be a closed set. By a closed decomposition property of the derivation basis B we mean an assertion that for every $\beta \in B$ there is a decomposition of the set X , $\{X_n\}$, such that sections $\beta[X_n]$ and $\beta[\overline{X}_n]$ enjoy some stated property.

There are a number of decomposition properties that are useful in the development. We shall give several general decomposition properties that are shared by a number of derivation bases as well as find, for a particular example of a derivation basis, a decomposition property that is peculiar to it. Indeed in this theory it is most useful to sort out in advance of the study of some derivation process the decomposition properties that will be available for its corresponding basis. Our first general property shall be labelled as a Y-decomposition property after W.H. Young and G.C. Young who first proved a number of the results for the Dini derivatives that we are able to generalize by utilizing this decomposition property.

DEFINITION 7.4 A derivation basis B will be said to have the Y-decomposition property if for every set $X \subset R$ and every $\beta \in B$ there is a decomposition of the set X into a sequence of sets $\{X_n\}$ with the property that $\beta[X_n]$ contains a partition of any interval with endpoints in X_n. B is said moreover to have the closed Y-decomposition property if for X closed such a decomposition can be found so that moreover if $y \in \overline{X}_n$ and $x \in X_n$ is sufficiently close to y then $\beta[\overline{X}_n]$ contains a partition of the interval $[x,y]$ (or $[y,x]$ if $y < x$).

(In particular note that $\beta[\overline{X}_n]$ would then always contain a partition of any interval $[a,b]$ with $a, b \in \overline{X}_n$ and a and b are not isolated on the right and left, respectively, in $\overline{X}_n$, while if a and b are so isolated one can at least find points $a' \le a$ and $b' \ge b$ arbitrarily close to a and b respectively and so that partitions are available for the intervals $[a',a]$, $[a',b']$, and $[b',b]$; from this one can at least obtain approximate results for $[a,b]$ (e.g., any additive interval function h would have $h([a,b]) = h([a',b']) - h([a',a]) - h([b,b'])$ and information about the three terms on the right of this equality would be available.

EXAMPLE 7.5 (A decompostion property for the uniform basis) The basis U *has both the Y-decomposition property and the closed Y-decomposition property (from 7.7 below and because* $RD \le U$*). It is easy to see that it has a much stronger property: if* $\beta \in U$ *and* X *bounded is given there is a finite decomposition of* X, $\{X_1,X_2,\dots X_n\}$ *such that* $\beta[X_k]$ *contains any interval point pair* $([a,b], c)$ *for which* $c \in [a,b]$ *and* $[a,b]$ *lies inside the bounds of* X_k.

EXAMPLE 7.6 (A decomposition property for the sharp basis) The basis $D^{\#}$ *has a decomposition property almost as strong as that for* U *: if* $\beta \in D^{\#}$ *and* $X \subset R$ *is given then there is a decomposition of* X *into a sequence of sets* $\{X_n\}$ *such that for any interval* $[a,b]$ *that lies within the bounds of* X_n *there is an element* $([a,b], c) \in \beta[X_n]$.

EXAMPLE 7.7 (Decomposition properties for the Dini derivation bases) The bases $\mathbb{RD}$ and $\mathbb{LD}$ have the Y-decomposition property and the closed Y-decomposition property. In fact, given $\beta \in \mathbb{RD}$ and a set X there is a decomposition $\{X_n\}$ of X so that any interval $[a,b]$ whose right endpoint is in X_n and which lies inside the bounds of X_n must have the element $([a,b],a) \in \beta[X_n]$.

To see this let $\delta(x)$ be given so that

$$\beta = \{([x,y],x) \ : \ 0 < y-x < \delta(x)\} ,$$

and let $\{X_n\}$ be a δ-decomposition of X. Then $\{X_n\}$ must have the stated property and this decomposition can be used to verify the conditions needed for the Y-decomposition property. For the closed Y-decomposition property the same with X closed works.

EXAMPLE 7.8 (A decomposition property for the ordinary basis) The basis $\mathbb{D}$ has the following decomposition property that is stronger than the Y-decomposition property: if $\beta \in \mathbb{D}$ and $X \subset R$ is given then there is a decomposition $\{X_n\}$ of X with the property that $\beta[X_n]$ contains a partition π (containing no more than two elements) of any interval $[a,b]$ that lies within the bounds of X_n and intersects X_n.

EXAMPLE 7.9 We shall give some general methods later in §9 to show that the derivation bases that express approximate derivation, selective derivation, qualitative derivation, and preponderant derivation all have the Y-decomposition property and the closed Y-decomposition property.

EXAMPLE 7.10 (Decomposition properties for the symmetric basis) The symmetric derivation basis $\mathbb{S}$ has the following closed decomposition properties which are of some use in establishing a number of properties of symmetric derivatives.

(A) *If* $X \subset R$ *is closed and* $\beta \in S[X]$ *there exists a decomposition* $\{X_n\}$ *of* X *with the property that if* $x \in X$ *is a right hand limit point of* X_n *then for* $y \in X$, $y > x$ *and sufficiently close to* x *there must be a point* $z \in (x,y) \cap X_n$ *for which* (I_x, x), (I_y, y) *and* (I_z, z) *all belong to* X *where*

$$I_x = [x - (y-z),\ x + (y-z)],$$
$$I_y = [y - (z-x),\ y + (z-x)], \quad \text{and}$$
$$I_z = [z - (y-x),\ z + (y-x)]$$

so that $F(I_z) = F(I_x) + F(I_y)$ *whenever* F *is additive.* *(The same assertion holds for* x *a left limit point of* X_n *with appropriate changes of course.)*

(B) *This decomposition* $\{X_n\}$ *can be arranged as well so that whenever* x *is a right hand limit point of* X_n *and* $y > x$ *is a left hand limit point of* X_n *and* I^* *denotes the interval*

$$\left[\frac{x+y}{2} - (y-x),\ \frac{x+y}{2} + (y-x)\right] = \left[\frac{3}{2}x - \frac{1}{2}y,\ \frac{3}{2}y - \frac{1}{2}x\right]$$

then there are $(I_i, x_i) \in \beta[X]$, $i = 0,1,\ldots,5$ *so that*

$$F(I^*) = \Sigma_{i=0}^{5} (-1)^i F(I_i)$$

for additive F.

QUERY 7.11 What are the appropriate decomposition properties for the derivation basis that expresses the approximate symmetric derivation?

For a first application of the Y-decomposition properties we establish a useful criterion for a derivation basis to have the partitioning property.

THEORM 7.12 Let B be a derivation basis that has the closed Y-decomposition property and has the following local property as well: if $\beta \in B$, $x \in R$ and $\varepsilon > 0$ then there are numbers y and z with $x - \varepsilon < y < x < z < x + \varepsilon$ so that β contains a partition of the intervals [y,x] and [x,z] . Then B has the partitioning property.

PROOF. Let $\beta \in B$ be given. We need to prove that β contains a partition of every interval. Let J denote the collection of all intervals [a,b] such that β contains a partition of every subinterval of [a,b] and write

$$G = \bigcup \{(a,b) : [a,b] \in J\} .$$

The set G is open and the theorem is proved if we are able to show that $G = R$ for if $[c,d] \subset G$ then since a finite number of elements J will cover [c,d] we can obtain a partition of [c,d] . To obtain a contradiction suppose that the set $Q = R \setminus G$ is nonempty. We already know that it must be closed: if [c,d] is an interval contiguous to Q then β contains a partition of every subinterval [a,b] with $c < a < b < d$. By the hypothesis of the theorem β must contain a partition of [c,a] and [b,d] for some such a and b so that we now have that β contains a partition of [c,d] and also every subinterval. Consequently Q can have no isolated points.

Using the closed Y-decomposition property we can decompose Q into a sequence of sets $\{Q_n\}$ such that $\beta[Q_n]$ contains a partition of any interval [x,y] with endpoints in Q_n and $\beta[\overline{Q}_n]$ contains a partition of [x,y] provided $x,y \in \overline{Q}_n$ and not isolated on the right and left respectively. By Baire's theorem (see Saks [105, p. 54]) one of these sets (Q_m say) is dense in a nonempty portion of Q ($Q \cap [c,d]$ say). A consequence of this is that β must contain a partition of any subinterval of [c,d] . To see this we only have to subdivide any such interval into subintervals that are contiguous or complementary to Q , and subintervals [x,y] with

$x,y \in Q \cap [c,d] = \overline{Q}_m \cap [c,d]$ for which x is not isolated on the right and y is not isolated on the left. Since β contains a partition of any interval of either of these types we can find a partition of every subinterval of $[c,d]$ as claimed. But then $Q \cap (c,d) = \emptyset$ which is impossible. This contradiction proves the theorem.

§8. The dual basis. Conventional abstract differentiation theory is dominated by the duality between the notions of a full cover and a fine (or Vitali) cover. That general theory commonly proceeds by making some assumption on the nature of these covers, most frequently that some version of the Vitali covering theorem is available (e.g., see Hayes and Pauc [39, Chapter II]). On the real line we require a more delicate approach which we achieve by the notion of a dual basis; in this way theorems which would normally require a Vitali-like assumption merely require the appearance of the dual basis. Thus our viewpoint should be that in the study of a derivation basis B, particularly in the study of the differentiation theory that arises, the dual basis B^* must take a natural appearance.

DEFINITION 8.1 Let B be a derivation basis. A subset β^* of $I \times R$ is said to be B-fine if for every $\beta \in B$ and every $x \in R$ either $\beta[\{x\}] = \emptyset$ or else $\beta^* \cap \beta[\{x\}] \neq \emptyset$. The collection of all B-fine subsets of $I \times R$ is denoted as B^* and referred to as the dual of B .

Usually we shall assume that B ignores no point so that the condition in 8.1 need only read that $\beta^* \cap \beta\, [\{x\}] \neq \emptyset$ for all $\beta \in B$ and all $x \in R$. We choose to call the derivation basis B^* the dual of B because of the fact that under most circumstances (see Theorem 8.6 below) the dual of B^* , $(B^*)^*$, is equivalent to B (i.e., $B \cong B^{**}$). In fact if B is filtering down and ignores no point then B^{**} is the filter generated by the filterbase B .

EXAMPLE 8.2 The dual basis of the trivial basis T *consists of all subsets of* $I \times R$ *and so in particular* $T \cong T^*$.

EXAMPLE 8.3 The dual basis of D *is the basis* D^* *introduced in §3 above.*

EXAMPLE 8.4 The dual basis of D^o *is the same as the dual basis of* U *. From this we see that the second dual of* U *,* U^{**} *, is equivalent to* D^o *.*

There are a number of natural questions that we can now answer for dual bases. Some of these are trivial and some quite surprising. For most derivation bases B (those that are filtering down and that ignore no point) the dual basis B^* behaves in a startling manner with respect to derivation: for interval point functions h and k the bases B and B^* interchange the upper and the lower derivations so that

$$\underline{D}_B\ h_k(x) = \overline{D}_{B^*}\ h_k(x) \quad \text{and} \quad \overline{D}_B\ h_k(x) = \underline{D}_{B^*}\ h_k(x)$$

A referee of an earlier version of this theory has remarked that this might be quite shocking to some readers and recommends that it be emphasized that this interchange is just our abstract expression of a well known fact regarding the computation of a "lim sup" : a lim sup can be viewed as a "sup - inf" or equally well as an "inf - sup" by a familiar device. One has for example

$$\limsup_{y \to x+} f(y) = \inf_{\delta > 0} \sup \{f(y) \ : \ 0 < y-x < \delta\}$$

and

$$\limsup_{y \to x+} f(y) = \sup \{ \inf_n \ f(x+h_n) \ : \ \{h_n\} \text{ a sequence } \downarrow 0 \}$$

The remaining properties should be less "shocking".

LEMMA 8.5 <u>Let</u> A <u>and</u> B <u>be derivation bases that ignore no point.</u> <u>If</u> $A \le B$ <u>then</u> $A^* \subset B^*$; <u>if</u> $A \cong B$ <u>then</u> $A^* = B^*$.

PROOF. If $\beta^* \in A^*$ is given and $A \leq B$ then for any $\beta_2 \in B$ and $x \in R$ there is an element $\beta_1 \in A$ with $\beta_1 \subset \beta_2$ and so

$$\beta^* \cap \beta_2 [\{x\}] \supset \beta^* \cap \beta_1 [\{x\}] \neq \emptyset$$

by the definition of the dual A^*. Consequently $\beta^* \in B^*$ and hence $A^* \subset B^*$ as required. The final assertion follows from this in an obvious manner.

THEOREM 8.6 *Let* B *be a derivation basis and* B^* *its dual. Then the following must hold:*

8.6.1 *if* $\beta^* \in B^*$ *and* $\beta^* \subset \gamma \subset I \times R$ *then necessarily* $\gamma \in B^*$,

8.6.2 B^* *has local character, and*

8.6.3 *if* B *ignores no point then* $B^{**} \supset B$.

If we assume as well that B *is filtering down and ignores no point then the following also hold:*

8.6.4 $B^* \supset B$,

8.6.5 *if* $\beta \in B$ *and* $\beta^* \in B^*$ *then* $\beta \cap \beta^* \in B^*$,

8.6.6 *if* B *has local character then* B^{**} *is the filter generated by the filterbase* B *on* $I \times R$, *and*

8.6.7 *if* B *has local character then* $B^{**} \cong B$.

PROOF. Assertions 8.6.1 and 8.6.2 are obvious. For 8.6.3 if B ignores no point then B^* also can ignore no point; thus if $\beta^* \in B^*$ and $\beta \in B$ we have

$$\beta^* \cap \beta [\{x\}] = \beta^* [\{x\}] \cap \beta \neq \emptyset$$

so that by this symmetry when β is in B it is also in B^{**} .

For 8.6.4 under the additional assumptions here any pair β_1 and β_2 in B have $\beta_1 \cap \beta_2$ $[\{x\}] \neq \emptyset$ (since there is a $\beta_3 \subset \beta_1 \cap \beta_2$ and β_3 $[\{x\}]$ cannot be empty) so that each element of B is in B^*. For 8.6.5 if $\beta \in B$ and β^* in B^* are given and $x \in R$ and $\beta_1 \in B$ are also given we certainly have

$$\beta_1 \cap \beta_2 \cap \beta^* \ [\{x\}] \ = \ \beta_2 \cap \beta^* \ [\{x\}] \neq \emptyset$$

for any $\beta_2 \in B$ with $\beta_2 \subset \beta_1$. This shows that $\beta \cap \beta^*$ is in B^* as required.

Finally we need only show that given an element β^{**} of B^{**} there is a $\beta \in B$ for which $\beta \subset \beta^{**}$ for then our assertions would follow from 8.6.1 and 8.6.3. To this end we fix $x \in R$ and determine a $\beta_x \in B$ so that $\beta_x[\{x\}] \subset \beta^{**}$. If this were not possible i.e., if there is no such β_x, then one can define

$$\beta^* \ = \ \{(I,x) \ : \ I \in \mathcal{I} \ , \ (I,x) \notin \beta^{**}\} \cup \mathcal{I} \times (R \setminus \{x\}) \ .$$

By our assumptions on β^{**} we see that $\beta^* \cup \beta$ $[\{x\}] \neq \emptyset$ for every $y \in R$ and every $\beta \in B$: for $y \neq x$ this is immediate and for $y = x$ this is because β^{**} can contain no $\beta[\{x\}]$ for $\beta \in B$. But this contradicts the fact that $\beta^{**} \in B^{**}$ for this β^* so defined is an element B^* that does not meet β^{**} at the point x.

Hence we have for every $x \in R$ and element $\beta_x \in B$ for which β_x $[\{x\}] \subset \beta^{**}$. Because B has local character we may find an element β of B so that $\beta[\{x\}] \subset \beta_x$ for each x ; this is of course the required β that is a subset of β^{**} and the proof is complete.

EXAMPLE 8.7 The assumption that a basis ignores no point is important in order for the dual basis to have any properties. For example the basis DD *ignores every point and it follows that its dual* DD^* *contains every subset of* $I \times R$ *and in particular* DD^* *is equivalent to the trivial basis.*

Thus even though $DD \le D$ *one does not have* $DD^* \subset D^*$ *as Lemma 8.5 would show, and other properties of the dual that require a basis not to ignore a point also fail.*

THEOREM 8.8 Let B be a derivation basis that is filtering down and does not ignore a point $c \in R$. Then the derivates of an interval-point function h relative to an interval-point function k at c with respect to B and its dual B^* have the following relationship:

$$\underline{D}_B\, h_k(c) = \overline{D}_{B^*}\, h_k(c) \quad \text{and} \quad \overline{D}_B\, h_k(c) = \underline{D}_{B^*}\, h_k(c)\ .$$

PROOF. Suppose that $\overline{D}_B\, h_k(c) < r$; then there is a $\beta \in B$ so that

$$\frac{h(I,c)}{k(I,c)} < r$$

for all $(I,c) \in \beta$. Let β^* belong to B^* so that by definition $\beta^*[\{c\}]$ must meet β : this gives

$$\inf\left\{\frac{h(I,c)}{k(I,c)} \ :\ (I,c) \in \beta^*\right\} < r\ ,$$

and hence $\underline{D}_{B^*}\, h_k(c) \le r$. As this holds for all such r we have established the inequality $\underline{D}_{B^*}\, h_k(c) \le \overline{D}_B\, h_k(c)$.

In the opposite direction if $\overline{D}_B\, h_k(c) > r$ then every $\beta \in B$ must have at least one element (I,c) for which

$$\frac{h(I,c)}{k(I,c)} > r$$

and we can define the collection

$$\beta^* = \{(I,c) : I \in \mathcal{I} , \frac{h(I,c)}{k(I,c)} > r\} \cup \mathcal{I} \times (R \setminus \{c\}) .$$

From the above remarks β^* must belong to B^* and so

$$\inf \left\{ \frac{h(I,c)}{k(I,c)} : (I,c) \in \beta^* \right\} \geq r$$

which proves that $\underline{D}_{B^*} h_k(c) \geq r$, yielding the inequality

$$\underline{D}_{B^*} (h_k(c) \geq \overline{D}_B h_k(c)$$

and the theorem follows.

§9. Natural derivation bases. If we are given a family of filters $\{N(x) : x \in R\}$ such that $N(x)$ is a filter convergent to x then we have defined in §3 above a derivation basis generated by the family $\{N(x) : x \in R\}$ and denoted for convenience N . Not only does this formulation express a great many familiar derivation bases (approximate, selective, path derivation etc.) it is a sufficiently general object on its own that it can be studied abstractly to some advantage. Indeed there is a characterization of such derivations in terms of the simple language we have previously developed.

THEOREM 9.1 *Let B be a derivation basis that is filtering down, endpoint tagged, finer than the topology, and has local character. Then there is a system of filters $\{N(x) : x \in R\}$ with each $N(x)$ a filter converging to the point x such that the derivation basis N generated by the family is equivalent to B . If B does not ignore a point x_0 then the filter $N(x_0)$ is nontrivial (i.e., $\{x_0\} \notin N(x_0)$).*

PROOF. For each $x \in R$ and $\beta \in B$ let $M_\beta(x) = \{y \in R : ([x,y],x)$ or $([y,x],x)$ is in $\beta\}$, and let $N(x)$ be the filter generated by the filter-base $\{M_\beta(x) \cup \{x\} : \beta \in B\}$.

The characteristics of a derivation basis that make it expressible as a filtered derivation process occur frequently enough to warrant a new definition.

DEFINITION 9.2 [NATURAL DERIVATION BASIS] A derivation basis that is filtered down, endpoint tagged, finer than the topology, has local character and ignores no point is said to be a natural derivation basis.

In any discussion of a natural derivation basis we will move freely between the generating filter family and the basis itself; certain properties prove to be more readily expressible in terms of the nature of the sets in the filters $N(x)$ than in the compact language of the derivation basis terminology. Note especially that a natural derivation basis is additive as the next lemma shows. If N is a natural derivation basis then always $N \leq D$ and so $N^* \subset D^*$.

LEMMA 9.3 If a derivation basis B is filtering down, endpoint tagged and finer than the topology it is necessarily additive. In particular every natural derivation basis is additive.

PROOF. Given β_1 , $\beta_2 \in B$ and $I = [a,b]$ define the open sets $G_1 = R\setminus\{a,b\}$ and $G_2 = (a - \frac{b-a}{4}, a + \frac{b-a}{4}) \cup (b - \frac{b-a}{4}, b + \frac{b-a}{4})$. Since B is finer than the topology we may select elements β_3 , $\beta_4 \in B$ so that $\beta_3[G_1] \subset \beta_1(G_1)$ and $\beta_4[G_2] \subset \beta_1(G_2)$. Finally, since B is filtering down we may select an element $\beta \in B$ with $\beta \subset \beta_1 \cap \beta_2 \cap \beta_3 \cap \beta_4$. The lemma is proved now merely by showing that this β splits at I , i.e., that if $(J,y) \in \beta$ then J cannot overlap I , and this follows easily from the construction and the fact that y here must be an endpoint of J .

REMARK 9.4 *Natural derivation is very general. Suppose that a function F is given and some generalized derivation process has yielded the extreme derivates $\underline{GD}\, F(x)$ and $\overline{GD}\, F(x)$ such that each is a derived number of the function F at x and $\underline{GD}\, F(x) \le \overline{GD}\, F(x)$. Then write for any $\varepsilon > 0$*

$$\eta_\varepsilon(x) = (x-\varepsilon, x+\varepsilon) \cap \{y : \underline{GD}\, F(x) - \varepsilon < \frac{F(y)-F(x)}{y-x} < \overline{GD}\, F(x) + \varepsilon\}$$

and let $N(x)$ be the filter generated by the filterbase

$$\{\eta_\varepsilon(x) : \varepsilon > 0\} .$$

Each $N(x)$ is a nontrivial filter converging to x and the corresponding derivation basis N *has $\underline{D}_N F(x) = \underline{GD}\, F(x)$ and $\overline{D}_N F(x) = \overline{GD}\, F(x)$.*

In particular every generalized derivative whose values are restricted to lie at derived numbers (unlike the symmetric derivative but like the approximate) gives rise to a natural derivation process that expresses it, but tailored, it must be noted, to the particular function under investigation.

There are a number of more delicate properties of natural derivation bases that are expressible in terms of the sets that appear in the families $N(x)$. One group of these, called intersection conditions, is particularly useful. These are related to similar notions explored in Bruckner, O'Malley and Thomson [13].

DEFINITION 9.5 Let $\{N(x)\}_{x\in R}$ be a system of filters with each $N(x)$ converging to x. A system $\{\eta_x\}_{x\in R}$ is said to be a choice from N if each $\eta_x \in N(x)$.

DEFINITION 9.6 A system $\{N(x)\}_{x\in R}$ will be said to have the stated intersection condition if corresponding to any choice $\{\eta_x\}_{x\in R}$ from N there is a positive function δ on R so that whenever

$$0 < y - x < \min \{\delta(x), \delta(y)\}$$

the two sets η_x and η_y from the choice intersect in the described manner:

9.6.1 [INTERSECTION CONDITION] $\eta_x \cap \eta_y \cap [x,y] \neq \emptyset$,

9.6.2 [INTERNAL INTERSECTION CONDITION] $\eta_x \cap \eta_y \cap (x,y) \neq \emptyset$,

9.6.3 [EXTERNAL INTERSECTION CONDITION [p] , p > 0]

$$\eta_x \cap \eta_y \cap (x - p(y-x),\ x] \neq \emptyset$$

and

$$\eta_x \cap \eta_y \cap [y\ ,\ y + p(y-x)) \neq \emptyset ,$$

9.6.4 [ONESIDED EXTERNAL INTERSECTION CONDITION [p], p > 0]

as for 9.6.3 but only <u>one</u> of the two intersections need be nonempty.

Our first result shows that in the presence of an intersection condition the derivation basis inherits some strong properties.

THEOREM 9.7 <u>Let</u> N <u>be a natural derivation basis that satisfies the intersection condition (9.6.1)</u>. <u>Then</u> N <u>has the</u> Y-<u>decomposition property and the closed</u> Y-<u>decomposition property</u>. <u>If in addition each set</u> $\eta_x \in N(x)$ <u>is twosided at</u> x (<u>i.e.</u>, <u>the sets</u> $[x,+\infty)$ <u>and</u> $(-\infty, x]$ <u>do not belong to</u> N(x)) <u>then</u> N <u>has the partitioning property</u>.

PROOF. Let $\beta \in N$ be given. Then by the definition of N there is a choice η , $\{\eta(x) : x \in R\}$, so that

$$\beta = \{([y,z],\ x) : y,z \in \eta(x)\ ,\ y \le x \le z\} .$$

Corresponding to this choice η there is a positive function δ on R so that if $0 < y - x < \min \{\delta(x), \delta(y)\}$ then the intersection condition 9.6.1 is met for $\eta(x)$ and $\eta(y)$.

If $X \subset R$ is given then let $\{X_n\}$ be a δ-decomposition of the set X . If $x,y \in X_n$ with $x < y$ then by the nature of the decomposition $\eta(x) \cap \eta(y) \cap [x,y]$ contains a point z say. If $z = y$ then the pair $([x,y],x)$ belongs to $\beta[X_n]$; if $z = x$ then the pair $([x,y],y)$ belongs to $\beta[X_n)$; if $x < z < y$ then both pairs $([x,z], x)$ and $([z,y], y)$ belong to $\beta[X_n]$. Thus we see that $\beta[X_n]$ contains in any case a partition of any interval with endpoints in X_n . This proves the Y-decomposition property.

If X is closed and $\{X_n\}$ denotes the same decomposition then for any point x in the closure of X_n , but not in X_n , we must have $X_n \cap (x - \delta(x), x + \delta(x))$ nonempty and every point y in that intersection will satisfy

$$|x - y| < \min \{\delta(x), \delta(y)\}$$

by the nature of the δ-decomposition. Again then the above arguments supply a partition of $[x,y]$ or $[y,x]$. This proves that the closed Y-decomposition property holds.

Finally to see that the partitioning property holds we need only appeal to Theorem 7.12 since B has the closed Y-decomposition property and the twosided assumptions that are given in the statement of our theorem give precisely the local property needed for 7.12. This then completes the proof.

EXAMPLE 9.8 (Intersection conditions for the approximate derivation) *Let* $\{N(x) : x \in R\}$ *be a system of filters with the following density restrictions: each* $\eta_x \in N(x)$ *has lower (inner) density on the right at* x *exceeding* ρ *and on the left exceeding* λ *. We say then that* N *is of* (ρ,λ)-*<u>density type</u>. Type* $(1,1)$ *is defined by having density* *equal to* 1 *and for* $0 \le \rho, \lambda < 1$ *the "exceeding" is used. The corresponding natural derivation bases can be used to express the standard density derivations:* $(1,1)$*-density type for the approximate derivative,* $(\frac{1}{2}, \frac{1}{2})$ *-density type for the preponderant, and* $(0,0)$*-density type for the weakest density derivation.*

The following density types have the following intersection properties:

$(1,1)$ *-density type*	--	*intersection condition, internal intersection condition, and external intersection condition* [p] *for all* $p > 0$;
(ρ,λ) *-density type* $\rho + \lambda \ge 1$	--	*intersection condition and internal intersection condition;*
(ρ,λ) *-density type,* $\rho > \frac{1}{2}$ *and* $\lambda > \frac{1}{2}$	--	*intersection condition, external intersection condition* p *for some* $p > 0$.

Note as a consequence of these intersection conditions that a natural derivation basis of (ρ,λ)*-density type with* $\rho + \lambda \ge 1$ *(in particular the approximate derivation basis) has the partitioning property.*

EXAMPLE 9.9 (Intersection conditions for paths) A system $\{E_x : x \in R\}$ *where each* E_x *is a set having* x *as a point of accumulation gives rise to a system of filters by taking merely for each* $N(x)$ *the filter generated by the filterbase* $\{E_x \cap (x - \varepsilon, x + \varepsilon) : \varepsilon > 0\}$. *In this case the intersection conditions assume a simpler form; for example this* N *has the internal intersection condition if and only if there exists a positive function* δ *on* R *so that if* $0 < y-x < \min\{\delta(x), \delta(y)\}$ *then* $E_x \cap E_y \cap (x,y) \ne \emptyset$.

Thus our intersection conditions for a natural derivation basis translate into direct assertions about the intersections of the paths for path derivations. (See Bruckner, O'Malley and Thomson [13] *for an account expressed in simpler language.)*

QUERY 9.10 (Selective derivation) The idea behind O'Malley's notion of selective derivation is very closely related to the internal intersection condition. There are two problems here that are worth considering.

#1. If f is the selective derivative of a function F (for a given selection) is there a choice of paths $\{E_x : x \in R\}$ possessing the internal intersection condition so that f is the path derivative of F for the system $\{E_x : x \in R\}$?

#2. If $D_{\mathbf{N}} F = f$ for a natural derivation basis $\mathbf{N}$ that possesses the internal intersection condition is there a selection so that f can be realized as the selective derivative of F relative to that selection?

CHAPTER THREE

THE VARIATION

§1. Elementary properties. Recall that for an arbitrary interval-point function h and any derivation basis B the variation of h over B has been defined by setting $V(h,B) = \inf\{V(h,\beta) : \beta \in B\}$ where $V(h,\emptyset) = 0$ and for a nonempty $\beta \subset \mathcal{I} \times R$ we have defined

$$V(h,\beta) = \sup\{\Sigma_{(I,x)\in\pi} |h(I,x)| : \pi \text{ a partition, } \pi \subset \beta\} .$$

This concept is fundamental to all of our concerns in this study; it yields the measure theory, an upper integral and it permits an expression of most of the principal results in the differentiation theory and in the integration theory. As it plays this fundamental role we need to address the properties of the variation prior to developing any other theory.

In this section we develop the basic computational properties of the variation. In particular we need to focus on the variational expressions

$$V(fh,B[X]) \quad\text{and}\quad V(fh,B(I))$$

where f is a point function, h an interval-point function, B a derivation basis, and $B[X]$ and $B(I)$ sections of B corresponding to subsets $X \subset R$ and $I \in \mathcal{I}$. By considering these expressions separately as functions of f, h, B, X, and I we can see numerous manipulations that require some investigation. For the first of these we observe the relationship between the variations relative to two different but related derivation bases.

LEMMA 1.1 Let h be an interval-point function and suppose that A and B are derivation bases.

1.1.1 if $A \leq B$ then $V(h,A) \leq V(h,B)$, and

1.1.2 if $A \cong B$ then $V(h,A) = V(h,B)$.

PROOF. If $A \leq B$ and $\beta_2 \in B$ then there is a $\beta_1 \in A$ such that $\beta_1 \subset \beta_2$ and hence $V(h,A) \leq V(h,\beta_1) \leq V(h,\beta_2)$. As this is true for all $\beta_2 \in B$ we must have $V(h,A) \leq V(h,B)$. The second assertion of the lemma follows easily from the first.

The most important and frequently used of the elementary properties of $V(h,B)$, considered as a function of h , is the seminorm property: namely that under the additional hypothesis that B is filtering down $V(h + h', B) \leq V(h,B) + V(h',B)$ for any pair of interval-point functions h and h' . This property along with some other simpler and similar ones is our concern in the next pair of lemmas.

LEMMA 1.2 *Let* h *and* h' *be interval-point functions and let* β *and* β' *be subsets of* $I \times R$. *Then*

1.2.1 $0 \leq V(h,\beta) \leq +\infty$,

1.2.2 *if* $\beta \subset \beta'$ *then* $V(h,\beta) \leq V(h,\beta')$,

1.2.3 *for any real number* c *not equal to zero* ,

$$V(ch,\beta) = |c| \; V(h,\beta) \;,$$

1.2.4 *if* $|h| \leq |h'|$ *then*

$$V(h,\beta) \leq V(h',\beta) \;,$$

1.2.5 $V(h + h',\beta) \leq V(h,\beta) + V(h',\beta)$,

1.2.6 *if* β *and* β' *are separated then*

$$V(h, \beta \cup \beta') = V(h,\beta) + V(h,\beta') \;.$$

PROOF. Each of 1.2.1 through 1.2.5 is completely elementary. For 1.2.6 we only have to note that if β and β' are separated then any partition $\pi \subset \beta \cup \beta'$ can be split into two separate partitions π_1 and π_2 with $\pi = \pi_1 \cup \pi_2$ and $\pi_1 \subset \beta$ and $\pi_2 \subset \beta'$, and conversely given any two such partitions π_1 and π_2 the set $\pi = \pi_1 \cup \pi_2 \subset \beta \cup \beta'$ is again a partition.

LEMMA 1.3 Let h and h' be interval-point functions and suppose that B is a derivation basis. Then

1.3.1 $0 \leq V(h,B) \leq +\infty$,

1.3.2 for any real number c not equal to zero ,

$$V(ch,B) = |c| \, V(h,B) ,$$

1.3.3 if $|h| \leq |h'|$ then

$$V(h,B) \leq V(h',B) ,$$

1.3.4 if B is filtering down then

$$V(h + h', B) \leq V(h,B) + V(h',B) .$$

PROOF. Each of 1.3.1, 1.3.2, and 1.3.3 follow from corresponding assertions in the previous lemma. For 1.3.4 note that if β_1 , β_2 and β_3 are given in B with $\beta_3 \subset \beta_1 \cap \beta_2$ then

$$\begin{aligned} V(h + h', B) &\leq V(h + h', \beta_3) \leq V(h, \beta_3) + V(h', \beta_3) \\ &\leq V(h, \beta_1) + V(h', \beta_2) , \end{aligned}$$

by computations above. Since B is filtering down there is such a β_3 for any choice of β_1 and β_2 and now 1.3.4 follows immediately.

EXAMPLE 1.4 It should be noticed here that the seminorm property of the variation (i.e., that $V(h_1 + h_2, B) \leq V(h_1, B) + V(h_2, B)$*) requires that* B *be filtering down, and this property may fail for some derivation bases.* D^* *, the dual of the ordinary derivation basis, does not enjoy this property. Let* $h(I,x) = |I|$ *if* $|I|$ *is rational and zero otherwise, and let* $g(I,x) = |I| - h(I,x)$ *. Then* $V(h,D^*) = V(g,D^*) = 0$ *and yet* $V(h+g, D^*) = +\infty$.

Our next concern is with the expressions $V(h,B[X])$ and $V(h,B(I))$ thought of as functions of the sets X and I .

LEMMA 1.5 *Let* h *be an interval-point function and* B *a derivation basis. Then for sets* $X, Y \subset R$ *and* $I, J \in \mathcal{I}_+$,

1.5.1 *if* B *is filtering down,*

$$V(h,B[X \cup Y]) \leq V(h,B[X]) + V(h,B[Y]) ,$$

1.5.2 *if* B *is filtering down and separates* X *and* Y ,

$$V(h,B[X \cup Y]) = V(h,B[X]) + V(h,B[Y]) ,$$

1.5.3 *if* I *and* J *do not overlap* ,

$$V(h,B(I \cup J)) \geq V(h,B(I)) + V(h,B(J)) ,$$

1.5.4 *if* B *is filtering down and additive, and* I *and* J *do not overlap,*

$$V(h,B(I \cup J)) = V(h,B(I)) + V(h,B(J)) ,$$

1.5.5 *if* B *is filtering down and additive,*

$$V(h,B) = V(h,B(I)) + V(h,B(R \setminus I^0)) .$$

PROOF. Assertion 1.5.1 is a special case of 1.3.4 : set $h_1 = \chi_X h$ and $h_2 = \chi_Y h$. Assertion 1.5.2 follows routinely from 1.2.6.

For 1.5.3 note that if $\beta \in B\ (I \cup J)$ and I and J do not overlap then $\beta(I)$ and $\beta(J)$ are separated and so 1.5.3 can also be made to follow from 1.2.6.

For 1.5.4 if B is also additive as well as filtering down then given any β_1 , $\beta_2 \in B$ there is a $\beta_3 \subset \beta_1 \cap \beta_2$ with $\beta_3 \in B$ and β_3 splits at I and J . Thus

$$\begin{aligned} V(h,B(I \cup J)) &\leq V(h,\beta_3(I \cup J)) = V(h,\beta_3(I)) + V(h,\beta_3(J)) \\ &\leq V(h,\beta_1(I)) + V(h,\beta_2(J)) \end{aligned}$$

and this together with 1.5.3 can be used to establish the desired equality. Similar arguments will prove the final assertion.

QUERY 1.6 (Additivity of the variation) One might wish to address the problem of determining conditions under which the additive formula

$$V(F + G,\mathrm{B}) = V(F,\mathrm{B}) + V(G,\mathrm{B}) \quad \text{might hold}$$

for additive interval functions F and G and for a derivation basis B. For instance such an investigation is given in Cater [18] *with* $\mathrm{B} = \mathrm{D}([a,b])$ *and necessary and sufficient conditions for this formula to hold are presented in terms of the Dini derivatives of the functions F and G. Are there other theorems of this type that are of any interest?*

§2. The fundamental lemma of the variational theory. Almost all of the important results expressible in our language arise from two deep properties of the variation which we present as our fundamental lemma. The hypothesis of local character (as well as the weaker one of σ-local character) plays a key role. In addition part of our lemma needs a uniformity assumption; because of the role it plays in Henstock's theory of integration we refer to this as "property H".

DEFINITION 2.1 An interval-point function h is said to have the property H relative to a derivation basis B if for every positive number ε there is a $\beta \in \mathrm{B}$ with

$$V(h,\beta(I)) \leq V(h,\mathrm{B}(I)) + \varepsilon$$

for all $I \in \mathcal{I}_+$.

The property H plays an important technical role in establishing some of the limit properties of the variation. We address immediately the problem of determining some situations in which that property is available. Note in particular that whenever the derivation basis is additive this property is easily obtained; thus any natural derivation basis allows the application freely of these lemmas.

LEMMA 2.2 *Let* B *be a derivation basis and* h *an interval-point function. Then in order that the function* h *shall have the property* H *relative to* B *either of the following two conditions suffice:*

(a) $V(h,B) < +\infty$ *and* B *is both filtering down and additive,*

(b) $V(h,B) < +\infty$, h *is an additive interval function or a non-negative subadditive interval function, and* B *is filtering down and has the partitioning property.*

PROOF. If B is filtering down and additive with $V(h,B) < +\infty$ then, using Lemma 1.4.5 we must have

$$(*) \qquad V(h,B) = V(h,B(I)) + V(h,B(R\setminus I^0))$$

for any $I \in \mathcal{I}$. Given $\varepsilon > 0$ choose $\beta \in B$ so that

$$V(h,\beta) \le V(h,B) + \varepsilon$$

and observe that

$$\begin{aligned} V(h,\beta(I)) &\le V(h,\beta) - V(h,\beta(R\setminus I^0)) \\ &\le V(h,B) + \varepsilon - V(h,B(R\setminus I^0)) \\ &\le V(h,B(I)) + \varepsilon \, , \end{aligned}$$

using (*) in the final inequality, and this is the property H as required. The same proof works under hypothesis (b) simply by proving that (*) again holds in such a circumstance; we omit the details.

It will be shown below (Chapter Five, §2) that any uniformly integrable interval-point function h has the property H relative to a basis that is filtering down and has the partitioning property. This allows a number of convergence properties of integrals to be proved without an "additive" assumption on the derivation basis. This is the technical reason that Henstock in a recent paper [61] was able to obtain various limit theorems in the setting of what he calls "non-additive division spaces".

Property H yields a useful property of the variation that it is convenient for us to express here.

LEMMA 2.3 Let B be a derivation basis and suppose that an interval-point function h has property H relative to B. Then

$$V(h,B) = \sup \{V(h,B(I)) : I \in \mathcal{I}\} .$$

PROOF. Because we are assuming the property H for the interval-point function h there must be for any $\varepsilon > 0$ an element $\beta \in B$ with

$$V(h,\beta(I)) \leq V(h,B(I)) + \varepsilon/2$$

for all $I \in \mathcal{I}_+$. If c is any number for which $c \leq V(h,B)$ then there must exist at least one partition $\pi \subset \beta$ for which

$$\Sigma_{(J,x) \in \pi} |h(J,x)| > c - \varepsilon/2$$

and then if I is an interval for which $I \supset \sigma(\pi)$ (i.e., each $J \subset I$ if $(J,x) \in \pi$),

$$V(h,B(I)) + \varepsilon/2 \geq V(h,\beta(I)) \geq \Sigma_\pi |h(J,x)| > c - \varepsilon/2 .$$

From this we can conclude that $\sup_{I \in \mathcal{I}} V(h,B(I)) \geq c - \varepsilon$ for every $\varepsilon > 0$. But $\varepsilon > 0$ is arbitrary and $c \leq V(h,B)$ is arbitrary and this gives $V(h,B) \leq \sup_{I \in \mathcal{I}} V(h,B(I))$; since the opposite inequality holds trivially the lemma has been proved.

We can now state our fundamental lemma which is to play an obvious role in the development of the measure theory, and plays a strong technical role in the development of the properties of integrals and derivatives in the next two chapters.

LEMMA 2.4 [FUNDAMENTAL LEMMA OF THE VARIATIONAL THEORY] *Let* B *be a derivation basis that has* σ*-local character and let* h *be an interval-point function.*

2.4.1 *for any sequence of sets* $X, X_1, X_2, X_3, \ldots$ *with*

$$X \subset \bigcup_{i=1}^{\infty} X_i ,$$

$$v(h,B[X]) \le \sum_{i=1}^{\infty} V(h,B[X_i]) ;$$

2.4.2 *assuming in addition that* B *is filtering down, for any sequence of nonnegative point functions* $f, f_1, f_2, f_3, \ldots$ *with* $\{f_i\}$ *nondecreasing,* $0 \le f(t) \le \sup_i f_i(t)$, *and such that each function* $f_i h$ *has the property* H *relative to* B ,

$$v(fh,B) \le \lim_{n \to \infty} V(f_n h,B) .$$

PROOF. We prove the first assertion (cf. Henstock [54, Theorem 44.10, p. 232]). It is clear that whenever $Y \subset Z$, $V(h,B[Y]) \le V(h,B[Z])$ so that there is no loss in generality if we assume that the given sequence $\{X_i\}$ is disjointed. For $\varepsilon > 0$ choose elements $\beta_n \in B$ such that

$$V(h,\beta_n[X_n]) \le V(h,B[X_n]) + \varepsilon\, 2^{-n} .$$

Because B has σ-local character and the sets $\{X_i\}$ are disjointed there must be an element $\beta \in B$ with the property that

$$\beta[X_n] \subset \beta_n$$

for each n .

Let $\pi \subset \beta[X]$ be an arbitrary partition and define $\pi_n = \pi[X_n]$: then

$$\Sigma_{\pi} |h(I,x)| = \sum_{n=1}^{\infty} \Sigma_{\pi_n} |h(I,x)|$$

$$\le \sum_{n=1}^{\infty} V(h,\beta_n[X_n])$$

$$\le \sum_{n=1}^{\infty} (V(h,B[X_n] + \varepsilon\, 2^{-n}) .$$

as this holds for all such partitions $\pi \subset \beta[X]$ we have

$$V(h,B[X]) \le V(h,\beta[X]) \le \sum_{n=1}^{\infty} V(h,B[X_n]) + \varepsilon$$

and since $\varepsilon > 0$ is arbitrary the first part of the theorem is proved.

For the second part of the theorem (cf. Henstock [54, Theorem 44.9, p. 231]) suppose that $\varepsilon > 0$ and $0 < c < 1$ are given. For each $x \in R$ there is a least integer $n(x)$ such that for every $m \ge n(x)$ one has $f_m(x) \ge c\, f(x)$.

Define the sequence of sets $X_n = \{x \in R : n(x) = n\}$ and observe that this is a disjointed sequence whose union is all of R. Select an element $\beta_n \in B$ so that for every $I \in I_+$

$$V(f_n h,\beta_n(I)) \le V(f_n h,B(I)) + \varepsilon\, 2^{-n} ;$$

this uses the fact that each $f_n h$ has the property H. Using the σ-local character of B, as in the first part of the theorem, choose an element $\beta \in B$ so that $\beta[X_n] \subset \beta_n$ for all n.

We now compute $V(fh,B)$: if $\pi \subset \beta$ is a partition write again, $\pi_n = \pi[X_n]$, $I_n = \sigma(\pi_n)$ and let N be the first integer for which $\pi_n = \emptyset$ for $n \ge N+1$ (this is possible because π is finite). Then

$$
\begin{aligned}
\Sigma_{\pi} |f(x)h(I,x)| &= \sum_{n=1}^{N} \Sigma_{\pi_n} |f(x)h(I,x)| \\
&\le c^{-1} \sum_{n=1}^{N} \Sigma_{\pi_n} |f_n(x)h(I,x)| \\
&\le c^{-1} \sum_{n=1}^{N} V(f_n h, \beta_n(I_n)) \\
&= c^{-1} \sum_{n=1}^{N} \{V(f_n h, B(I_n)) + \varepsilon\, 2^{-n}\} \\
&= c^{-1} \left(\sum_{n=1}^{N} V(f_N h, B(I_n)) + \varepsilon \right) \\
&\le c^{-1}\, V(f_N h, B) + c^{-1} \varepsilon .
\end{aligned}
$$

(We have used here a number of the elementary computations from the previous section without comment.)

In this inequality c may be arbitrarily close to 1 and ε to 0 ; the choice of N depends on the pair c and ε but it is clear from this inequality that

$$V(fh,B) \le \sup_i \; V(f_i h,B)$$

and the proof of the theorem is complete.

There are some immediate corollaries that we shall state for future reference.

COROLLARY 2.5 *Let* h *be an interval-point function and* f *a point function. If* B *is a derivation basis that has* σ*-local character then*

2.5.1 *if* $V(h,B) = 0$ *it follows that* $V(fh,B) = 0$, *and conversely*

2.5.2 *if* $V(fh,B) = 0$ *then* $V(h,B[X]) = 0$ *where*

$$X = \{x : f(x) \neq \emptyset\} .$$

PROOF. If $V(h,B) = 0$ then by writing $X_n = \{x \in R : |f(x)| \le n\}$ we have $V(fh,B[Y_n]) \le n\, V(h,B[Y_n]) \le n\, V(h,B) = 0$ by some of our elementary properties. But the fundamental Lemma then gives

$$V(fh,B) \le \sum_{n=1}^{\infty} V(fh,B[Y_n]) = 0$$

as required.

The second part of the corollary is proved in precisely the same way but using the sets $Y_n = \{x \in R : |f(x)| > 1/n\}$. The set X of assertion 2.5.2 is just the union of these sets and so the proof is obtained in the obvious manner.

COROLLARY 2.6 Let B be a derivation basis that has σ-local character and is finer than the topology. Then an interval-point function h has $V(h,B(I)) = 0$ for every interval I if and only if $V(h,B) = 0$.

PROOF. Certainly if $V(h,B) = 0$ then $V(h,B(I)) = 0$ for every interval. For the converse under the additional hypotheses on B note that if $V(h,B([a,b])) = 0$ then $V(h,B[[c,d]]) = 0$ for $a < c < d < b$ since, using the fact that B is finer than the topology, one has $B[[c,d]] \le B([a,b])$.

Then by the fundamental Lemma we have

$$V(h,B) \leqq \sum_{n=1}^{\infty} V(h,B[[-n,n]])$$

and so on the assumption that each of these vanishes so too does $V(h,B)$.

EXAMPLE 2.7 This corollary requires some hypothesis such as σ-local character. For example it is false for the uniform basis U *: let*

$$h(I,x) = \begin{cases} |I| & \text{if } x \text{ is an integer} \\ 0 & \text{otherwise} \end{cases}$$

and it is simple to compute that $V(h,U(I)) = 0$ *for every interval* I *and yet* $V(h,U) = +\infty$.

Note that this same example shows that the fundamental lemma itself requires the hypothesis of σ-local character. For example let

$$X_n = [-n,n] \quad \text{and} \quad f_n(x) = \chi(X_n,x)$$

and then it is clear that $V(h,U[X_n]) = V(f_n h,U) = 0$ *for all* n *but* $V(h,U)$ *is not less than either of*

$$\Sigma\, V(h,U[X_n]) \quad \text{or} \quad \sup V(f_n h,U) .$$

§3. <u>The fundamental lemma of the derivation theory</u>. There is an intimate relationship between the variation and the derivation expressed very loosely by the fact that $D_B\, h_k = f$ is almost equivalent to the assertion $V(h - fk,B) = 0$. Although the lemma is not quite as simple as this the relationship is very strong and has many implications in our subsequent theory. Consequently statements about derivatives translate into variational equations and the tools we have developed in the previous sections can be applied to provide numerous properties of the derivatives.

LEMMA 3.1 [FUNDAMENTAL LEMMA OF THE DERIVATION THEORY] <u>Let</u> h <u>and</u> k <u>be interval-point functions, let</u> f <u>be a point function, and suppose that</u> B <u>is a derivation basis that is filtering down</u>.

3.1.1 <u>If</u> $V(k,B) < +\infty$ <u>and</u> $D_B\, h_k = f$ <u>then</u> $V(h - fk,B) = 0$, <u>and conversely</u>

3.1.2 <u>if</u> $V(h - fk,B) = 0$ <u>then there is a set</u> $N \subset R$ so that

$$D_{B[R\setminus N]}\, h_k = f \quad \underline{\text{and}} \quad V(k,B^*[N]) = 0 .$$

PROOF. To prove the first assertion let $\varepsilon > 0$ be given and choose β_1 and β_2 from B so that

$$|h(I,x) - f(x)k(I,x)| \le \varepsilon\, |k(I,x)| \quad \text{if} \quad (I,x) \in \beta_1$$

and

$$V(k,\beta_2) \le V(k,B) + 1 .$$

Since B has been assumed to be filtering down there is a $\beta_3 \in B$ with $\beta_1 \cap \beta_2 \supset \beta_3$. For this we have evidently

$$V(h - fk, \beta_3) \leq V(\varepsilon k, \beta_3) \leq \varepsilon(V(k,B) + 1)$$

and since $\varepsilon > 0$ is arbitrary and $V(k,B)$ is finite the assertion of the lemma follows.

For the second assertion, remembering that B^* is the dual basis for B, define the sequence of sets $\{Y_n\}$ so that $x \in Y_n$ if and only if for every element $\beta \in B$ either B ignores x or else β contains at least one pair (I,x) for which

$$|h(I,x) - f(x)k(I,x)| \geq \frac{|k(I,x)|}{n} .$$

Define $N = \bigcup_{n=1}^{\infty} Y_n$. We will show that $V(k,B^*[N]) = 0$ and that $D_{B[R \setminus N]} h_k = f$.

To see the first of these let $\varepsilon > 0$ be given and select a $\beta \in B$ so that $V(h-fk,\beta) < \varepsilon$. Let β_{n^*} denote the collection of all pairs $(I,x) \in \beta$ for which

$$|h(I,x) - f(x)k(I,x)| \geq \frac{|k(I,x)|}{n} .$$

By the definition of the set Y_n this collection β_{n^*} must be an element of $B^*[Y_n]$ and hence

$$V(k,B^*[Y_n] \leq V(k,\beta_{n^*}) \leq n\, V(h-fk,\beta) < n\,\varepsilon .$$

As $\varepsilon > 0$ is arbitrary we must have $V(k,B^*[Y_n]) = 0$ and so, using the fact that B^* has local character and that the fundamental lemma of the variation theory is hence available, we have

$$V(k,B^*[N]) \leqq \sum_{n=1}^{\infty} V(k,B^*[Y_n]) = 0$$

as required.

Consequently the lemma is proved if we show that $D_A h_k = f$ for the derivation basis $A = B[R \setminus N]$. Given any integer n there must be an element $\beta \in B$ so that

$$|h(I,x) - f(x)k(I,x)| < \frac{|k(I,x)|}{n}$$

for every $(I,x) \in \beta[R \setminus N]$ merely because of the way in which the sets $Y_n \subset N$ are defined. By definition then this means

$$D_{B[R \setminus N]} h_k = f$$

and we are done.

EXAMPLE 3.2 Applications of this lemma require rather more of the integration and differentiation theory than we have so far developed but we can indicate here an example that shows why the dual basis must enter in. If

$$F(b) - F(a) = \int_a^b f(x)dx$$

holds on every interval for the Riemann integral then this can be seen (see §4 below) to be equivalent to the assertion $V(F - fm, \mathrm{U}([a,b])) = 0$ *for all* $[a,b]$. *The lemma we have just proved then asserts that* f *is the uniform derivative of* F *on* $[a,b] \setminus N$ *where* $N \subset [a,b]$ *is a set of measure zero in the sense that* $V(m, \mathrm{U}^*[N]) = 0$, *and this amounts to saying that* N *has Lebesgue measure zero. Note that for* N *to have measure zero in the sense of* U *rather than in the sense of its dual says that* N *has Peano-Jordan measure zero (i.e., that in fact* $\overline{N}$ *also has measure zero) and this conclusion is not valid here (e.g.,* f *can be taken to be discontinous at every rational point).*

§4. The fundamental lemma of the integration theory. Just as there is an intimate relationship between the variation and the differentiation there is also such a relationship available between the variation and the integration. Loosely this can be expressed by the claim that the assertion

$$H(I) = \int_{(I)} dh \quad (I \in \mathcal{I}_+)$$

is more or less equivalent to the variational equation $V(H - h, B) = 0$. As a consequence a variety of properties of the integral can be made to appear as easy consequences of the properties of the variation.

Throughout this section we assume that B is a derivation basis that is filtering down and that possesses the partitioning property. In that case we have the following defintions:

4.1 An interval-point function h is integrable on a set $I_0 \in \mathcal{I}_+$ if there is a number c such that for every $\varepsilon > 0$ an element $\beta \in B$ can be found such that

$$\left| \Sigma_{(I,x) \in \pi} h(I,x) - c \right| < \varepsilon$$

for any partition π of I_0 with $\pi \subset \beta$.

4.2 Because of the assumptions on B (that it is filtering down and has the partitioning property) such a c if it exists is unique and so if h is integrable on I we write

$$\int_{(I)} dh \quad \text{or} \quad (B) - \int_{(I)} dh$$

for this number.

4.3 An interval-point function is integrable uniformly on $\mathcal{I}_0$ where $\mathcal{I}_0$ is a subset of $\mathcal{I}_+$ if it is integrable on each $I \in \mathcal{I}_0$ and for every $\varepsilon > 0$ there is a $\beta \in B$ so that

$$\left| \int_{(\sigma(\pi))} dh - \Sigma_{(I,x) \in \pi} h(I,x) \right| < \varepsilon$$

for every partition $\pi \subset \beta$ for which $\sigma(\pi) \in \mathcal{I}_0$, where $\sigma(\pi)$ denotes the union of the intervals I for which a pair (I,x) belongs to π.

We can now state our fundamental lemma linking the integration definitions with the variation.

LEMMA 4.4 [FUNDAMENTAL LEMMA OF THE INTEGRATION THEORY] Let h be an interval-point function and $J \in \mathcal{I}_+$. If B is a derivation basis that is filtering down and partitions every subinterval of J then the following assertions are equivalent:

(i) h is integrable on J,

(ii) h is integrable uniformly on $J \cap \mathcal{I}_+$, and

(iii) there is an additive interval function H for which

$$V(H - h, B(J)) = 0 .$$

In this latter case $H(I) = \int_{(I)} dh$ for every subinterval I of J.

PROOF. We begin by making a simple but useful observation. Integrability and uniform integrability can be characterized in terms of a familiar "Cauchy" criterion.

4.4.1 h is integrable on J if and only if for every $\varepsilon > 0$ there is a $\beta \in B$ such that whenever π and π' are partitions of J from β then

$$\left| \Sigma_{(I,x) \in \pi} h(I,x) - \Sigma_{(I',x') \in \pi'} h(I',x') \right| < \varepsilon$$

4.4.2 h is integrable uniformly on $\mathcal{J} \subset \mathcal{I}_+$ if and only if for every $\varepsilon > 0$ there is a $\beta \in B$ so that whenever π and π' are partitions from β with $\sigma(\pi) = \sigma(\pi') \in \mathcal{J}$ then

$$\left| \Sigma_{(i,x) \in \pi} h(I,x) - \Sigma_{(I',x') \in \pi'} h(I',x') \right| < \varepsilon .$$

These will be used in the proof.

We begin with the implication: (iii) implies (ii). If $V(H - h, \mathcal{B}(J)) = 0$ then for any $\varepsilon > 0$ there is a $\beta \in \mathcal{B}(J)$ for which $V(H - h, \beta) < \varepsilon$. If $\pi \subset \beta$ is a partition then $\sigma(\pi) \in J \cap I_+$ and

$$|H(\sigma(\pi)) - \Sigma_\pi h(I,x)| \leq \Sigma_\pi |H(I) - h(I,x)| \leq V(H-h,\beta) < \varepsilon .$$

By definition then h is integrable uniformly on $J \cap I_+$ and $H(I) = \int_{(I)} dh$ there.

Thus we have (iii) ⇒ (ii), and since (ii) ⇒ (i) is obvious we need only show that (i) ⇒ (iii). Given $\varepsilon > 0$ choose $\beta \in \mathcal{B}(J)$ so that $\pi \subset \beta$ is a partition of J then

$$(*) \qquad \left| \int_{(J)} dh - \Sigma_\pi h(I,x) \right| < \varepsilon / 8 .$$

Let E be an element of $J \cap I_+$ and suppose that we are given two partitions π and π' from β such that both π and π' are partitions of E ; then there must exist a partition π'' which adjoined to π or π' yields a partition of J ; thus setting $\pi_1 = \pi \cup \pi''$ and $\pi_2 = \pi' \cup \pi''$ we have

$$|\Sigma_{\pi_1} h(I,x) - \Sigma_{\pi_2} h(I,x)| = |\Sigma_\pi h(I,x) - \Sigma_{\pi'} h(I,x)|$$

which cannot exceed $\varepsilon/4$ because of (*) above.

As such a choice of $\beta \in \mathcal{B}(J)$ can be made for any $\varepsilon > 0$ we have by our preliminary observations, 4.4.1 and 4.4.2, that h is integrable uniformly on $J \cap I_+$ and so we may set

$$H(E) = \int_{(E)} dh$$

for any such E and it will follow that

$$|H(E) - \Sigma_{\pi_1} h(I,x)| < \varepsilon/2$$

for partitions $\pi_1 \subset \beta$ of E .

If E_1 and E_2 are nonoverlapping members of $J \cap I_+$ then a further application of this inequality in the obvious manner gives

$$|H(E_1 \cup E_2) - H(E_1) - H(E_2)| < 3\varepsilon/2$$

from which we readily see that H must be additive.

It now remains only to see that $V(H-h,B(J)) = 0$. Let $\beta \in B(J)$ be as above and let $\pi \subset \beta$ be any partition. Write

$$\pi' = \{(I,x) \in \pi : H(I) - h(I,x) \geq 0\} ,$$

$$\pi'' = \{(I,x) \in \pi : H(I) - h(I,x) < 0\} ,$$

$$E' = \sigma(\pi') , \text{ and}$$

$$E'' = \sigma(\pi'') .$$

The above estimates now give

$$\begin{aligned} \Sigma_{\pi} |H(I) - h(I,x)| &= \Sigma_{\pi'} (H(I) - h(I,x)) + \Sigma_{\pi''} (h(I,x) - H(I)) \\ &= |H(E') - \Sigma_{\pi'} h(I,x)| + |H(E'') - \Sigma_{\pi''} h(I,x)| < \varepsilon . \end{aligned}$$

As this holds for all such partitions we see that

$$V(H-h,B(J)) = V(H-h,\beta) < \varepsilon$$

and the lemma follows.

§5. <u>The upper integral</u>. For a fixed derivation basis B and a fixed interval-point function h the functional $f \to V(fh,B)$ defined for all point functions serves as an upper integral. It is a direct generalization of the classical Darboux upper integral, and includes abstractly a number of other concepts. For convenience we restrict attention to non-negative finite valued real functions.

DEFINITION 5.1 Let h be an interval-point function and B a derivation basis. Then for any nonnegative point function f we write

$$h_B(f) = V(fh,B)$$

and refer to the functional h_B as the upper integral associated with h and B .

Modern notation (cf. A.I. Tulcea and C.I. Tulcea [64, pp. 1-2]) usually requires of an upper integral that it be defined also for extended real-valued functions and have stronger limit properties than are available here without further assumptions. Our next theorem shows that these limit properties are available under some simple natural assumptions; and it would be an easy matter to extend this functional to functions that might assume the value $+\infty$.

THEOREM 5.2 [PROPERTIES OF THE UPPER INTEGRAL] *Let h_B be the upper intergral associated with an interval-point function h and a derivation basis B . The following properties hold:*

5.2.1 $0 \le h_B(f) \le +\infty$,

5.2.2 $h_B(0) = 0$ (*the first* "0" *denoting the zero function*),

5.2.3 *if* $f \le g$ *then* $h_B(f) \le h_B(g)$,

5.2.4 *if* B *is assumed to be filtering down, then*

$$h_B(f+g) \le h_B(f) + h_B(g) ,$$

5.2.5 $h_B(cf) = c\, h_B(f)$ *if* c *is positive*.

On the further assumption that the derivation basis B *is filtering down, has* σ-*local character, and is additive, the stronger properties below also hold:*

5.2.6 *if* $f \le \sum_{n=1}^{\infty} g_n$ *then* $h_B(f) \le \sum_{n=1}^{\infty} h_B(g_n)$, *and*

5.2.7 *if* $\{g_n\}$ *is an increasing sequence and* $f \le \sup_n g_n$ *then* $h_B(f) \le \lim_{n\to\infty} h_B(g_n)$.

PROOF. Each of these assertions is merely a translation into the present notation of a previous statement about the variation that has been proved earlier.

EXAMPLE 5.3 (Classical upper integrals) In the event that the derivation basis is U *(the uniform basis) or* D *(the ordinary basis) and the function* h *is of the form* $h(I,x) = G(I)$ *where* G *is a nondecreasing function on* R *(i.e., considered as an interval function it is additive and nonnegative) then the upper integrals* h_{U} *and* h_{D} *can be written in a more traditional form:*

$$(\mathrm{U})-\overline{\int_a^b} f(x)dG(x) = G_{\mathrm{U}([a,b])}(f)$$

and

$$(\mathrm{D})-\overline{\int_{[X]}} f(x)dG(x) = G_{\mathrm{D}[X]}(f)$$

for $f \geq 0$ *. The former is called the* <u>*upper Darboux-Stieltjes*</u> *integral and the latter the* <u>*upper Lebesgue-Stieltjes*</u> *integral of the nonnegative function* f *with respect to the nondecreasing function* G *.*

These are largely incomparable notions and the notation is meant to reflect that. However if $G(x) = x$ *(i.e., as an interval function* $G = m$*) then we can introduce three classical upper integrals that are closesly related:*

$$(\mathrm{U})-\overline{\int_a^b} f(x)dx = V(fm,\mathrm{U}([a,b])) \qquad \text{(upper Darboux),}$$

$$(\mathrm{D})-\overline{\int_a^b} f(x)dx = V(fm,\mathrm{D}([a,b])) \qquad \text{(upper Lebesgue), and}$$

$$(\mathrm{D}^*)-\overline{\int_a^b} f(x)dx = V(fm,\mathrm{D}^*([a,b])) \qquad \text{(upper dual).}$$

Because of the relation $D^* \leq D \leq U$ *we have immediately for any nonnegative function* f *on* $[a,b]$ *:*

$$(D^*)\overline{\int_a^b} f(x)dx \leq (D)\overline{\int_a^b} f(x)dx \leq (U)\overline{\int_a^b} f(x)dx \,.$$

Note that only the middle integral, based on D*, here permits all of the conclusions of Theorem 5.2 to hold as* U *does not have local character and* D^* *is not filtering down. It is not too difficult to see what the* U *and the* D *integral are doing; the* D^* *integral is rather more mysterious. The following might be conjectured and we leave it as a query:*

QUERY 5.4 *Is it the case that for any nonnegative function* f *on* $[a,b]$

$$(D^*)\overline{\int_a^b} f(x)\, dx = \sup \{(D)\overline{\int_a^b} g(x)dx : g \leq f,\ g \text{ is upper semicontinuous}\}\ ?$$

This seems likely since if f *is measurable then the* (D) *and the* (D^*) *upper integrals coincide.*

§6. <u>The measure theory</u>. For a fixed derivation basis B and a fixed interval-point function h the set function $X \to V(h,B[X])$ defined for all subsets $X \subset R$ serves as an outer measure. It is a direct generalization of the classical Peano-Jordan outer measure that was developed in the last two decades of the nineteenth century; in most cases of interest to us it is a genuine outer measure in the modern sense of that term (<u>e.g.</u>, Munroe [91]).

DEFINITION 6.1 Let h be an interval-point function and B a derivation basis. Then for any set $X \subset R$ we write

$$h_B(X) = V(h,B[X])$$

and refer to the set function h_B as the "<u>outer measure</u>" associated with h and B .

Modern notation requires countable subadditivity of outer measures and our first theorem shows that this is available under natural assumptions on the derivation basis. Without those assumptions the set function h_B will still be called an outer measure but in quotations ("outer measure") to indicate that the deeper properties may not be true. Note that the use of the same terminology for the upper integral and the outer measure ($h_B(X)$ and $h_B(f)$) is justified by the fact that the obvious equation

$$h_B(X) = h_B(\chi_X)$$

is the traditional method of deriving a measure from an upper integral and in our case is provided by Definition 6.1 directly. Also in our setting there is a particularly nice relationship existing for the standard notion of product of function and measure; here that product is in fact a genuine product, *viz.*

$$(fh)_B(X) = h_B(f\,\chi_X) = V(fh, B[X]) .$$

THEOREM 6.2 [PROPERTIES OF THE "OUTER MEASURE"] *Let* h_B *be the "outer measure" associated with an interval-point function* h *and a derivation basis* B . *Then the following properties hold:*

6.2.1 $0 \le h_B(X) \le +\infty \quad (X \subset R)$,

6.2.2 $h_B(\emptyset) = 0$,

6.2.3 *if* $X \subset Y$ *then* $h_B(X) \le h_B(Y)$,

6.2.4 *if* B *is filtering down then*

$$h_B(X \cup Y) \le h_B(X) + h_B(Y) ,$$

6.2.5 *if* B *has* σ*-local character and* $X \subset \bigcup_{n=1}^{\infty} X_n$ *then*

$$h_B(X) \le \sum_{n=1}^{\infty} h_B(X_n) ,$$

6.2.6 if B is filtering down and separates the sets X and Y then

$$h_B(X \cup Y) = h_B(X) + h_B(Y) \; ,$$

6.2.7 if B is filtering down, has σ-local character, and is additive then for any increasing sequences of sets $\{X_n\}$

$$h_B\left(\bigcup_{n=1}^{\infty} X_n\right) = \lim_{n\to\infty} h_B(X_n) \; .$$

PROOF. Each of the stated properties of the set function h_B is just a restatement of an earlier proved property of the variation.

THEOREM 6.3 [PROPERTIES OF THE OUTER MEASURE] Let h_B be the outer measure associated with an interval-point function and a derivation basis B that is assumed to be of σ-local character. Then

6.3.1 h_B is a genuine outer measure,

6.3.2 on the additional assumption that B is filtering down and is finer than the topology h_B is a metric outer measure,

6.3.3 on the additional assumption that B is filtering down and is additive then h_B has the increasing sets property.

PROOF. That h_B is a true outer measure is merely the content of 6.2.2 and 6.2.5. Under the additional assumptions on B in 6.3.2 we see that B separates any two sets X and Y that are topologically separated; thus by an elementary property of the variation we have

$$V(h,B[X \cup Y]) = V(h,B[X]) + V(h,B[Y])$$

and this shows that h_B is additive over topologically separated sets. By definition then (e.g., Munroe [91, Ch. 2]) h_B is a metric outer measure; this is equivalent we should recall to the fact that all Borel sets are h_B-measurable. Finally, the increasing sets property has already been expressed in the previous theorem.

THEOREM 6.4 *Let* A *and* B *be derivation bases and* h *an interval-point function.*

6.4.1 *if* $A \le B$ *then* $h_A \le h_B$,

6.4.2 *if* $A \cong B$ *then* $h_A \equiv h_B$.

PROOF. These are easy consequences of our elementary computations for the variation since if $A \le B$ (or $A \cong B$) then every section $A[X] \le B[X]$ (or $A[X] \cong B[X]$) .

THEOREM 6.5 *Let* B *be a derivation basis that is filtering down. Then for any interval-point function* h , $h_{B*} \le h_B$.

PROOF. From our results in Chapter Two, §8 we know that $B^* \le B$ whenever B is filtering down and so this follows from the previous theorem.

Note that in this theorem B may not have local character or σ-local character so that h_B might not be an outer measure, but h_{B*} is in fact always an outer measure.

THEOREM 6.6 [PRODUCT OF A POINT FUNCTION AND A MEASURE] *Let* B *be a derivation basis that has* σ*-local character and suppose that* f *is a point function and* h *an interval-point function. Then the outer measures* h_B *and* $(fh)_B$ *have the following relations:*

6.6.1 $(fh)_B(X) = 0$ *if and only if* $f(x) = 0$ *for* h_B*-almost every point* x *in* X ,

6.6.2 *if* h_B *is* σ*-finite on a set* X *then so too is* $(fh)_B$, *and*

6.6.3 *if* h_B *vanishes on a set* X *then so too does* $(fh)_B$.

PROOF. The first and last of these assertions are just translations into our present language of statements proved in §2 above for the variation. For the second assertion suppose that

$$X = \bigcup_{n=1}^{\infty} X_n$$

where $h_B(X_n) < +\infty$ and write

$$X_{mn} = \{x \in X_n : |f(x)| \le m\} .$$

Clearly $fh_B(X_{mn}) = V(fh,B[X_{mn}]) \le mV(h,B[X_{mn}]) \le m\, h_B(X_n) < +\infty$ so that $(fh)_B$ is also σ-finite on X as required.

As an easy application of this lemma we can easily connect the vanishing of the upper integral $h_B(f)$ with the vanishing of the function f.

COROLLARY 6.7 Let B be a derivation basis, let h be an interval-point function and f a point function. If B has σ-local character then $h_B(|f|) = 0$ if and only if f vanishes h_B-almost everywhere on R. In general if B is not claimed to have σ-local character but is filtering down and ignores no point then one has at least that $h_B(|f|) = 0$ entails $f(x) = 0$ h_{B^*}-almost everywhere.

PROOF. The first part of the corollary follows directly from the theorem and is only a translation into the present language. The latter part uses only the fact that under these assumptions $h_{B^*}(|f|) \le h_B(|f|)$ and B^* must have local character.

EXAMPLE 6.8 (Classical measure theory) The standard interval function $m(I) = |I|$ *that expresses the length of the interval* I *gives rise to "outer measures"* m_U *and* m_D *relative to the two derivation bases* U *(the uniform basis) and* D *(the ordinary basis) that play historically important roles. We define*

m_U - *Peano-Jordan "outer measure"*

m_D - *Lebesgue outer measure.*

Note the following properties. In particular the properties of m_D *will justify our calling it Lebesgue measure (we have throughout been using the notation* $|X|$ *for the Lebesgue outer measure of* X*).*

(i) m_U *is finitely subadditive, but in general not countably subadditive;*

(ii) m_D *is a metric outer measure that has the increasing sets property;*

(iii) *for any set* $X \subset R$,

$$m_U(X) = \inf \{\Sigma |I_i| : \{I_i\} \text{ a finite sequence of intervals covering } X\},$$

and

$$m_D(X) = \inf \{\Sigma |I_i| : \{I_i\} \text{ a sequence of intervals covering } X\};$$

(iv) *for bounded sets* X , $m_U(X) = m_D(X)$ *if and only if the set* $\overline{X} \setminus X$ *has measure zero (i.e.,* $|\overline{X} \setminus X| = 0$*); in particular then* m_U *and* m_D *agree on compact sets;*

(v) $m_D \leq m_U$.

The two measures are obviously closely related. This relationship can be exhibited in an unusual light as well: the study of the Riemann integral has a somewhat distressing feature to the beginning analysis student in that although the Peano-Jordan measure seems the most natural one in that setting a number of results can only be satisfactorily expressed by using Lebesgue measure which at first sight seems like an intruder. However, an investigation of U *must lead to a consideration of the duals* U^* *and* U^{**}, *and since* $U^{**} \cong D^o$ *the Lebesgue measure* m_{D^o} *enters the scene naturally. Indeed one has* $m_{U^*} = m_{U^{**}} = m_{D^o}$ *and Lebesgue measure is unavoidable.*

EXAMPLE 6.9 (Peano-Jordan measurability) *The "outer measure" m_U is defined on all subsets of R but is better behaved on a special subclass of sets that were isolated and studied by Jordan. We shall say that a bounded set E is* <u>*PJ-measurable*</u> *if it satisfies any of the following equivalent assertions:*

(1) $m_U(I) = m_U(I \cap E) + m_U(I \setminus E)$ <u>*for all intervals*</u> I ;

(2) χ_E <u>*is Riemann integrable on every interval*</u> I;

(3) <u>*the set of boundary points of*</u> E <u>*has (Lebesgue or equivalently Peano-Jordan) measure zero;*</u>

(4) <u>*for every*</u> $\varepsilon > 0$ <u>*there is a*</u> $\beta \in U$ <u>*so that*</u>

$$\Sigma_{(I,x) \in \pi[E]} \Sigma_{(I',x') \in \pi[R \setminus E]} |I \cap I'| < \varepsilon$$

<u>*for all partitions*</u> π <u>*and*</u> π' <u>*of the same interval with both*</u> π <u>*and*</u> π' <u>*contained in*</u> β .

It is now easy to show that m_U is additive (finitely that is) on the class of PJ-measurable sets; this follows from any one of the four characterizations above but perhaps from (2) this is most transparent.

EXAMPLE 6.10 (Lebesgue measurable sets) *Parallel to the characterizations of the class of PJ-measurable sets is a similar characterization of the class of Lebesgue measurable sets. Each of the following is equivalent:*

(1) $m_D(I) = m_D(I \cap E) + m_D(I \setminus E)$ <u>*for all intervals*</u> I ;

(2) $m_D(T) = m_D(T \cap E) + m_D(T \setminus E)$ <u>*for all sets*</u> $T \subset R$;

(3) $\chi_E m$ <u>*is*</u> D-<u>*integrable on every interval*</u> I ;

(4) <u>*there are open sets*</u> G_1 <u>*and*</u> G_2 <u>*with*</u> $m_D(G_1 \cap G_2)$ <u>*arbitrarily small and*</u> $G_1 \supset E$ <u>*and*</u> $G_2 \supset R \setminus E$;

(5) <u>*for every*</u> $\varepsilon > 0$ <u>*there is a*</u> $\beta \in D$ <u>*so that*</u>

$$\Sigma_{(I,x) \in \pi[E]} \Sigma_{(I',x') \in \pi[R \quad E]} |I \cap I'| < \varepsilon$$

<u>*for all partitions*</u> π <u>*and*</u> π' <u>*of the same interval with*</u> π, $\pi' \subset \beta$.

QUERY 6.11 For $0 < p \le 1$ let m^p denote the interval function

$$I \to |I|^p = m^p(I) .$$

Then the measure $m^p{}_D$ *evidently is related to the classical Hausdorff p-dimensional measure. More generally if* h *is a monotonically increasing function on* $[0,\infty), h(0) = 0$, *then* $h \circ m$ *denotes the function* $I \to h(m(I)) = h(|I|)$ *and* $(h \circ m)_D$ *again represents a measure on* R *that should be related to similar ideas in the theory of Hausdorff measures. What is the exact relation here?*

EXAMPLE 6.12 (The Stieltjes measures) For any additive interval function F *the measures* F_D *and* F_{D^*} *provide an expression of the total variation of the function* F. *This will be discussed in greater detail in §9 below. Here we point out some simple computations for these measures.*

(i) F_D and F_{D^*} are metric outer measures,

(ii) $F_{D^*} \le F_D$,

(iii) for any point x,

$$F_D(\{x\}) = \limsup_{h\to 0+} |F(x+h) - F(x)| + \limsup_{h\to 0+} |F(x) - F(x-h)|$$

and

$$F_{D^*}(\{x\}) = \min\{\liminf_{h\to 0+} |F(x+h) - F(x)|, \liminf_{h\to 0+} |F(x) - F(x-h)|\}$$

(iv) for $G = \cup (a_i, b_i)$ open,

$$F_D(G) = \Sigma\, Var(F, (a_i, b_i)) .$$

§7. <u>The measure theory (continued)</u>. Measure theory plays a number of important roles in analysis. For our purposes we can outline these roles as resting within three categories: (i) it provides relative to a given outer measure, μ say, a class of μ-measure zero sets that may serve as the exceptional sets for a certain class of theorems; (ii) it provides a class of finite and σ-finite measure sets and so a way of categorizing certain objects as being not too large; and (iii) it provides a theory of integration. Our largest concern is with the first two of these applications of measure theory; sets of measure zero relative to an outer measure h_B provide the numerous exceptional sets needed in both the differentiation and the integration theory below. Outer measures h_B that are finite or σ-finite on a set X provide an expression of some familiar variational concepts (e.g., VB, VBG_*) that are frequent tools in analysis.

But associated with any outer measure h_B that arises in our setting would be a corresponding measure-theoretic integration theory that we will have no particular use for. We have an upper integral and also a Riemann-type integral to hand in most of our applications and no need for a measure-theoretic integral. Since the connections will not be immediately apparent we shall in this section show the relationship that exists between our upper integral $h_B(f)$ and the measure-theoretic integral

$$\int f(x)\,dh_B(x)$$

that a measure theorist would generate from h_B as an outer measure.

Throughout this section let us suppose that B is a derivation basis that has σ-local character and is filtering down; some of the results can do with less but this setting simplifies matters. For any interval-point function h then the set function h_B is a true outer measure and we can construct a measure-theoretic integral as follows:

(i) a nonnegative point function g is said to be h_B-elementary if it is h_B-measurable and countably valued, i.e., if

(*) $$g(t) = \Sigma\, a_i\, \chi(E_i, t)$$

for a sequence $\{a_i\}$ of positive numbers and a sequence $\{E_i\}$ of h_B-measurable sets (which can be taken as disjoint);

(ii) for g that is h_B-elementary we write

$$\int g\,dh_B = \Sigma_{c \in R}\, c\, h_B(g^{-1}(\{c\}))$$

or, using the expression (*) above,

$$\int g\,dh_B = \Sigma\, a_i h_B(E_i) .$$

(iii) for any nonnegative point function f we write

$$\tilde{\int} f\,dh_B = \inf \{ \int g\,dh_B : g \geq f ,\ g \text{ is } h_B\text{-elementary}\} .$$

We do not pause to give the traditional justifications of the above definitions but proceed immediately to show how this measure theoretic upper integral based on the outer measure h_B relates to our h_B-upper integral.

LEMMA 7.1 Let the derivation basis B be filtering down and suppose that h is an interval-point function and E_1, E_2 a pair of disjoint sets. Then if E_1 is h_B-measurable, and c_1, c_2 are positive real numbers,

$$h_B(c_1\chi_{E_1} + c_2\chi_{E_2}) = c_1h_B(E_1) + c_2h_B(E_2) .$$

PROOF. By elementary arguments of measure theory we have that $h_B(E_1 \cup E_2) = h_B(E_1) + h_B(E_2)$. Given $\varepsilon > 0$ choose an element $\beta \in B$ so that

$$V(h,\beta[E_1]) \leq h_B(E_1) + \varepsilon/2 ,$$

$$V(h,\beta[E_2]) \leq h_B(E_2) + \varepsilon/2 , \text{ and}$$

$$V((c_1\chi(E_1) + c_2\chi(E_2)h,\beta) \leq h_B(c_1\chi(E_1) + c_2\chi(E_2)) + \varepsilon .$$

Since B is filtering down we may choose β in such a way that all three inequalities hold.

We may select a partition $\pi \subset \beta[E_1 \cup E_2]$ so that

$$\Sigma_{(I,x) \in \pi} |h(I,x)| \geq h_B(E_1 \cup E_2) - \varepsilon/2 = h_B(E_1) + h_B(E_2) - \varepsilon/2 .$$

This requires

$$\begin{aligned}\Sigma_{(I,x) \in \pi[E_1]} |h(I,x)| &= \Sigma_\pi |h(I,x)| - \Sigma_{\pi[E_2]} h(I,x) \\ &\geq h_B(E_1) + h_B(E_2) - \varepsilon/2 - \{h_B(E_2) + \varepsilon/2\} \\ &\geq h_B(E_1) - \varepsilon .\end{aligned}$$

An identical computation would yield

$$\Sigma_{(I,x) \in \pi[E_2]} |h(I,x)| \geq h_B(E_2) - \varepsilon .$$

Now using this partition π and the above estimates we obtain

$$\begin{aligned}V((c_1\chi(E_1) + c_2\chi(E_2))h,\beta) &\geq \Sigma_{(I,x) \in \pi} (c_1\chi(E_1,x) + c_2\chi(E_2,x)) |h(I,x)| \\ &\geq c_2 \Sigma_{\pi[E_1]} |h(I,x)| + c_2 \Sigma_{\pi[E_2]} |h(I,x)| \\ &\geq c_1 h_B(E_1) + c_2h_B(E_2) - (c_1 + c_2) \varepsilon .\end{aligned}$$

This inequality evidently holds for all $\beta \in B$ and so, since $\varepsilon > 0$ is arbitrary we have established the inequality

$$h_B(c_1\chi(E_1) + c_2\chi(E_2)) \geq c_1h_B(E_1) + c_2h_B(E_2) .$$

The opposite inequality is available for derivation bases that are filtering down from our elementary theory and so the lemma is proved.

LEMMA 7.2 Let the derivation basis B be filtering down, let h be an interval-point function, let $\{E_i\}$ be a sequence of disjoint h_B-measurable sets and $\{c_i\}$ a sequence of positive numbers. Then

$$h_B(\Sigma_{i=n}^{n} c_i\chi_{E_i}) = \Sigma_{i=n}^{n} c_i\, h_B(E_i) .$$

If further B has σ-local character and g is defined, where

$$g(t) = \Sigma_{n=1}^{\infty} c_i\chi(E_i,t)$$

then

$$h_B(g) = \Sigma_{n=1}^{\infty} c_i\, h_B(E_i) .$$

PROOF. For fixed n write $g_n(t) = \Sigma_{i=1}^{n} c_i\chi(E_i,t)$. Then for the first part of the lemma we wish to prove that

$$h_B(g_n) = \Sigma_{i=1}^{n} c_i h_B(E_i) .$$

We already know this for n=1 and n=2. For $n \geq 3$ and a given $\varepsilon > 0$ we can use the fact that B is filtering down to select a $\beta \in B$ so that each of the following inequalities holds:

$$V(g_n h,\beta) \leq h_B(g) + \varepsilon ,$$

and

$$V(h,\beta[\bigcup_{i\neq m} E_i]) \leq h_B(\bigcup_{i\neq m} E_i) + \varepsilon/2 \quad \text{for } m=1,2,\ldots,n .$$

There must be in $\beta[\bigcup_{i=1}^{n} E_i]$ a partition π so that

$$\Sigma_{(I,x)\in\pi} |h(I,x)| \geq h_B(\bigcup_{i=1}^{n} E_i) - \varepsilon/2 .$$

Putting these together, and remembering that the $\{E_i\}$ are disjoint and h_B-measurable, we obtain for any $m=1,2,\ldots,n$ the following estimate:

$$\Sigma_{\pi[E_m]} |h(I,x)| \geq \Sigma_\pi |h(I,x)| - \Sigma_{\pi[\bigcup_{i=m} E_i]} |h(I,x)|$$

$$\geq \Sigma_{i=1}^{n} h_B(E_i) - \varepsilon/2 - \{\Sigma_{i\neq m} h_B(E_i) + \varepsilon/2\}$$

$$\geq h_B(E_m) - \varepsilon .$$

Now using this partition π and the above estimates we obtain

$$h_B(g_n) + \varepsilon \geq V(g_n h, \beta) \geq \Sigma_{(I,x)\in\pi} g_n(x) |h(I,x)|$$

$$\geq \Sigma_{i=1}^{n} \Sigma_{\pi[E_i]} c_i |h(I,x)| \geq \Sigma_{i=1}^{n} \{c_i h_B(E_i) - c_i \varepsilon\} .$$

Since the ε here is an arbitrary positive number we have proved the inequality

$$h_B(g_n) \geq \Sigma_{i=1}^{n} c_i h_B(E_i) .$$

As B is filtering down the opposite inequality holds as well and thus we have proved the first part of our lemma.

Turning now to the second part and noting that $g \geq g_n$ we have immediately that

$$h_B(g) \geq h_B(g_n) = \Sigma_{i=1}^{n} c_i h_B(E_i)$$

for all n. Also using the fundamental lemma of the variation theory, since now B is taken to have σ-local character, we obtain

$$h_B(g) = V(gh,B) = V(gh,B[\bigcup_{i=1} E_i]) \ \Sigma_{i=1}^{\infty} V(gh,B[E_i]) = \Sigma_{i=1}^{\infty} c_i h_B(E_i) .$$

This proves that for all n,

$$\Sigma_{i=1}^{n} c_i h_B(E_i) \leq h_B(g) \leq \Sigma_{i=1}^{\infty} c_i h_B(E_i)$$

and now the lemma follows.

We may now prove our main theorem.

THEOREM 7.3 Let the derivation basis B be filtering down and have σ-local character, and let h be an interval-point function. Then

7.3.1 for any nonnegative point function f,

$$h_B(f) \leq \tilde{\int} f dh_B ,$$

7.3.2 for any nonnegative, bounded h_B-measurable point function f,

$$h_B(f) = \tilde{\int} f dh_B ,$$

7.3.3 assuming in addition that B is additive, for any nonnegative h_B-measurable point function f,

$$h_B(f) = \tilde{\int} f dh_B .$$

PROOF. For the first part observe that if there is no h_B-elementary function g exceeding f then $\tilde{\int} f dh_B = +\infty$ so there is nothing to prove; if there is such a function $g \geq f$ then

$$h_B(f) \leq h_B(g) = \int g dh_B$$

and, since this holds for all such functions g, again the inequality must hold.

For (.2), with f both bounded and h_B-measurable let $r > 1$ and suppose $f(x) < M$ for all x. Write

$$E_n = \{x \in R : Mr^{-n} \leq f(x) < Mr^{-n+1}\}$$

and define

$$g(t) = \Sigma_{n=1}^{\infty} Mr^{-n+1} (E_n, t) .$$

By our assumptions on f g is h_B-elementary and $f \leq g \leq rf$ everywhere. Thus $h_B(f) \leq h_B(g) = \int g dh_B \leq rh_B(f)$ for all $r > 1$. From this assertion (.2) easily follows.

Finally, assertion (.3) just uses the monotone convergence property of the variation (§3 above), which is available for additive derivation bases, in the traditional way to extend from bounded h_B-measurable functions to unbounded ones. (Note here that the additivity assumption is used only to assure that all the functions gh that appear have the property H and thus permit the fundamental lemma of the variation to be applied. This also would hold under weaker assumptions.)

EXAMPLE 7.4 (The Lebesgue upper integral) Using the basis D *that expresses ordinary derivation we have defined*

$$\overline{\int}_{[X]} f(x)dx = V(fm, \mathrm{D}[X])$$

for any nonnegative function f *and have referred to this as the Lebesgue upper integral (see §5 above). By Theorem 7.3 we see now that this is indeed the same integral as would be generated by using the Lebesgue outer measure* m_{D}. *This same observation applies to Stieltjes versions of these integrals in the obvious way.*

Note in particular that the Riemann-type integral generated by D *can be expressed also as a measure-theoretic integral at least for nonnegative functions: thus for any interval-point function* $h \geqq 0$ *and any point function* $f \geqq 0$ *if* fh *is* D*-integrable on an interval then*

$$\int_{(I)} fdh = V(fh, \mathrm{D}(I)) = \widetilde{\int} fd\psi_{h,I} \quad \text{where} \quad \psi_{h,I}$$

is the outer measure

$$\psi_{h,I}(X) = V(h, \mathrm{D}(I)[X]) .$$

§8. Generalized continuity: B-continuous functions. There are numerous problems in real analysis that can be expressed in terms of the variation relative to some derivation basis. Thus many generalizations of continuity (uniform continuity, right or left continuity, approximate continuity, preponderant continuity, selective continuity, symmetric continuity, etc.) can be expressed by a single definition and studied systematically. In this section we introduce the notion of a B-continuous function relative to a derivation basis B as well as the more restrictive notions of B-null and B-locally constant.

DEFINITION 8.1 Let B be a derivation basis and h an interval-point function. Then

8.1.1 h is said to be B-continuous at a point x if for $\varepsilon > 0$ there is a $\beta \in B$ with $V(h,\beta[\{x\}]) < \varepsilon$;

8.1.2 h is said to be B-continuous if for every $\varepsilon > 0$ there is a $\beta \in B$ with $V(h,\beta[\{x\}]) < \varepsilon$ for all x ;

8.1.3 h is said to be B-null if $V(h,B) = 0$;

8.1.4 h is said to be B-locally constant if there is a $\beta \in B$ such that $h(I,x) = 0$ for all $(I,x) \in \beta$.

We shall be interested in applying these concepts mainly for functions h of the form Δf, i.e., for additive interval functions, although the definitions are useful in general. Note that the concepts increase in generality: a B-locally constant function is necessarily B-null, a B-null function is necessarily B-continuous, and a B-continuous function must be B-continuous at each point. It is also easy to find examples to show that each of the four definitions is distinct.

Note that for additive interval functions F some connections are clear: if B has the partitioning property then only the zero function F can be B-null or B-locally constant; the notion of B-continuous or pointwise B-continuous is clearly related to familiar generalizations of continuity. In particular our next theorem and the examples that follow it should clarify these latter two concepts.

THEOREM 8.2 Let h be an interval-point function and B a derivation basis. Then the following are true:

8.2.1 if h is B-continuous then h is B-continuous at every point,

8.2.2 if h is B-continuous at every point and B has local character then h is B-continuous,

8.2.3 if h is a continuous additive interval function and B is finer than the topology and straddled then h is B-continuous at each point.

PROOF. The first two of these are obvious. For the third, if h is a continuous additive interval function then at any point x and for any $\varepsilon > 0$ there is a $\delta > 0$ so that $|h([y,z])| < \varepsilon/2$ if $x - \delta < y \leq x \leq z < x + \delta$. Using the open set $G = (x-\delta, x+\delta)$ and the fact that B is finer than the topology there must be a $\beta \in B$ with $\beta[\{x\}] \subset \beta(G)$. Since for any $(I,x) \in \beta$ we have $|h(I)| < \varepsilon/2$ and since B is straddled this gives

$$V(h, \beta[\{x\}]) < \varepsilon .$$

This proves that h is B-continuous at x as required.

EXAMPLE 8.3 (Uniform continuity) If U *denotes the uniform derivation basis then a function* F *is* U*-continuous if and only if* F *is uniformly continuous. Since* U *does not have local character it is not true that* U*-continuity at each point is sufficient to ensure* U*-continuity.*

EXAMPLE 8.4 (Ordinary continuity) If D *denotes the ordinary derivation basis then it is clear that* D*-continuity is merely equivalent to pointwise continuity in the usual sense.*

EXAMPLE 8.5 (Sharp continuity) For the sharp derivation basis $D^{\#}$ *the notion of* $D^{\#}$*-continuity has an unusual aspect which is a reflection of the fact that* $D^{\#}$ *is not straddled. A function* F *is* $D^{\#}$*-continuous at a point* x *if and only if* F *is continuous and of bounded variation in some interval containing* x *.*

We prove here two theorems which relate some integrability and derivability results to the notion of B-continuity. These are just general versions of familiar results for integrals and derivatives.

THEOREM 8.6 [CONTINUITY OF THE INTEGRAL] Let the derivation basis B be filtering down and have the partitioning property and suppose that

$$F(I) = \int_{(I)} f\,dh \quad (I \subset J)$$

where f is a point function and h an interval-point function for which fh is integrable on an interval J relative to B . Then $F - fh$ is $B(J)$-continuous. In particular if h is $B(J)$-continuous at a point then so too is F .

PROOF. In fact, by the fundamental lemma of the integration theory (§4 above), we have $V(F - fh, B(J)) = 0$ which is stronger than our assertion here.

THEOREM 8.7 Let B be a derivation basis that is filtering down and let h and k be interval-point functions such that k is B-continuous at a point x . Then if both $\underline{D}_B\, h_k(x)$ and $\overline{D}_B\, h_k(x)$ are finite, h must also be B-continuous at x .

PROOF. Using a fixed point x there must be a positive number M so that for some β_1, β_2 in B we have

$$h(I,x)/k(I,x) < M \quad \text{for} \quad (I,x) \in \beta_1$$

and

$$h(I,x)/k(I,x) > -M \quad \text{for} \quad (I,x) \in \beta_2 ,$$

remembering of course to interpret $0/0$ as 0. Since B is filtering down there is a β_3 in B with $\beta_3 \subset \beta_1 \cap \beta_2$. This gives

$$|h(I,x)| \leq M|k(I,x)|$$

for all $(I,x) \in \beta_3$ and hence for all $\beta \in B$,

$$V(h,B[\{x\}]) \leq V(h,\beta_3 \cap \beta[\{x\}]) \leq M\, V(k,\beta[\{x\}])$$

so that in fact $V(h,B[\{x\}]) \leq M\, V(k,B[\{x\}])$ and the result follows.

In the spirit of the last two results we can also show that B-continuity is closed under a natural limit operation involving the variation.

LEMMA 8.8 <u>Let the derivation basis</u> B <u>be filtering down and let</u> $h,h_1,h_2,h_3,\ldots$ <u>be interval-point functions such that each</u> h_n <u>is</u> B-<u>continuous and so that</u> $\lim V(h_n - h,B) = 0$. <u>Then</u> h <u>too is</u> B-<u>continuous</u>.

PROOF. Given $\varepsilon > 0$ choose N so large that $V(h_N - h,B) < \varepsilon/3$ and choose a $\beta_1 \in B$ so that

$$V(h_N - h,\beta_1) < V(h_N - h,B) + \varepsilon/6 < \varepsilon/2 .$$

Take $\beta_2 \in B$ so that $V(h_N,\beta_2[\{x\}]) < \varepsilon/2$ at each x. Now if $\beta_3 \in B$ is chosen so that $\beta_3 \subset \beta_1 \cap \beta_2$ then it is easy to see that

$$V(h,\beta_3[\{x\}]) \leq V(h_N - h,\beta_1) + V(h_N,\beta_2[\{x\}]) < \varepsilon$$

as required.

The first of our main theorems in this section gives a condition under which a function that is B-continuous will be in the first class of Baire. This theorem generalizes a number of classical results for approximately continuous functions, preponderantly continuous functions, selectively continuous functions, and others. The condition we need is just a more detailed version of the closed Y-decomposition property but one which is enjoyed by all of the examples we have so far shown have the weaker version of that property.

DEFINITION 8.9 A derivation basis B is said to have *Y_m-decomposition property* or the *closed* *Y_m-decomposition property* where $m = 2,3,...$ provided it has the corresponding Y-decomposition property as expressed in Definition 7.4 of Chapter Two but with the additional hypothesis that the partition that is required to exist there contain no more than m elements.

THEOREM 8.10 *Let* B *be a derivation basis that has the closed* Y_m*-decomposition property for some integer* m . *Then if* F *is a real function on* R *that is* B-*continuous,* F *is in the first class of Baire*.

PROOF. To show that F is Baire 1 we obtain a contradiction by a standard device: if F is not Baire 1 then there must exist a perfect set Q such that the oscillation of F restricted to Q exceeds some positive number ε at each of its points. Choose an element $\beta \in B$ so that

$$|F(I)| < \varepsilon/3m \quad \text{for all} \quad (I,x) \in \beta$$

(this just uses the definition of B-continuity) and choose a decomposition $\{Q_n\}$ of Q corresponding to this choice of $\beta \in B$ as in Definition 8.9. By Baire's theorem one of these sets, say Q_k , is dense in a portion of Q , say $Q \cap [c,d]$.

We obtain a contradiction by showing that the oscillation of F on this portion does not exceed ε and the proof is complete. To this end observe that for any points x and y in Q_k $(x < y)$ there is in β a partition π of the interval $[x,y]$ and $\text{card}(\pi) \leq m$. Thus

$$|F(y) - F(x)| \leq \Sigma_\pi |F(I)| < \varepsilon/3 .$$

If x and y are now arbitrary points in $\overline{Q}_k \supset Q \cap [c,d]$ then there must be points x' and y' in Q_k that are sufficiently close to x and y respectively so that again there are partitions from β with no more than m elements of each of the intervals $[x,x']$ (or $[x',x]$, and $[y,y']$ (or $[y',y]$)). This gives again, as before

$$|F(x) - F(x')| < \varepsilon/3 \quad \text{and} \quad |F(y) - F(y')| < \varepsilon/3 .$$

But we already have $|F(x') - F(y')| < \varepsilon/3$ and thus putting these three together in the obvious way yields $|F(x) - F(y)| < \varepsilon/3$, which inequality holds everywhere in this portion of Q. Since this is the desired contradiction the proof is complete.

THEOREM 8.11 Let B be a derivation basis that is endpoint tagged and has the partitioning property. Then any B-continuous function has the Darboux property.

PROOF. Suppose that F never assumes the value c, and define $\mathcal{P}$ to be the collection of intervals $[a,b]$ for which $(F(a) - c)(F(b) - c > 0$. If $[a,b]$, $[b,b']$ are abutting intervals in $\mathcal{P}$ then $[a,b']$ is in $\mathcal{P}$ since

$$\{(F(a) - c)(F(b) - c)\}\{(F(b) - c)(F(b') - c\} = (F(a) - c)(F(b') - c)(F(b) - c)^2 .$$

Since F does not assume the value c at any point x we can write $\varepsilon = |F(x) - c| > 0$ and then there is an element $\beta \in B$ so that if $([x,y],x)$ or $([y,x],x)$ is in β then

$$|F(y) - F(x)| < \varepsilon/2$$

which means that $|F(y) - c| \geq |F(x) - c| - \frac{\varepsilon}{2} > 0$; in particular $F(x) - c)(F(y) - c) > 0$ for every $([x,y],x)$ or $([y,x],x)$ in β . From the partitioning property it now follows that $\mathcal{P}$ includes every interval so that $(F(a) - c)(F(b) - c) > 0$ always. Since this is just an interpretation of the Darboux property (i.e., if $F(a) < c < F(b)$ or $F(b) < c < F(a)$ then F assumes the value c in that interval) the theorem is proved.

THEOREM 8.12 Let N be a natural derivation basis that satisfies the intersection condition. Then any N-continuous function is in the first class of Baire. If in addition N is not onesided (i.e., for no x does $[x,+\infty)$ or $(-\infty, x]$ belong to $N(x)$) then such a function is also Darboux.

PROOF. This is just an application of the previous two theorems for if N satisfies the intersection condition then we have seen (Chapter Two, section 8) that a Y-decomposition and a closed Y-decomposition are available; in fact, checking the size of the partitions that are used there we see that a closed Y_7-decomposition has been obtained. For the Darboux property we need only recall that the non-onesided assumption together with the intersection condition proves the partitioning property.

We shall not give the details but any one of our intersections conditions from Chapter Two, §9 can be used to supply that B-continuous functions are Baire 1, and then again a non-onesided assumption carries this to Darboux Baire 1.

EXAMPLE 8.13 (Approximately continuous functions are Darboux Baire 1) *If* A *denotes the approximate derivation basis then a function* F *is approximately continuous if and only if it is* A*-continuous. We have already seen earlier (Chapter One, §9) that the basis* A *has the partitioning property and that it has the decomposition property of Theorem 8.10 (in fact with* $N=6$ *) so that it follows from our results that approximately continuous functions are necessarily Darboux Baire 1. (This same observation applies to functions that are preponderantly continuous.)*

EXAMPLE 8.14 (Unilaterally continuous functions are Baire 1) The decomposition property of Theorem 8.10 holds for the bases RD *and* LD *that express onesided derivation. The corresponding continuity for these bases is just a unilateral assertion of ordinary continuity; consequently unilaterally continuous functions are (of course) Baire 1. It is easy to see that such functions need not have the Darboux property and this is reflected in the fact that neither basis* RD *nor* LD *has the partitioning property (in the language of Theorem 8.12 each is in fact a onesided natural basis).*

EXAMPLE 8.15 (Selectively continuous functions are Darboux Baire 1) O'Malley [95] *has proved that a function continuous in the selective sense must be Darboux Baire 1 and, indeed, it is a Theorem of Neugebauer* [94] *that this notion can characterize the class of Darboux Baire 1 functions. In our setting these properties follow from Theorems 8.10 and 8.11.*

*EXAMPLE 8.16 (*B*-locally constant/*B*-null/*B*-continuous) For any derivation basis the classes of functions that are* B*-locally constant,* B*-null or* B*-continuous are of some considerable interest and it is worth investing some time to discover characterizations of the classes. In the table below we shall exhibit some of these classes: an asterisk (*) entry indicates that there is further discussion below, while an entry (?) indicates that we do not know of an appropriate characterization of the class of functions indicated.*

TABLE 8.17.1

	DERIVATION BASIS B	B-LOCALLY CONSTANT (1)	B-NULL (2)	B-CONTINUOUS (3)
(1)	T *the trivial basis*	*all functions*	*all functions*	*all functions*
(2)	U *the uniform basis*	*constant functions*	*constant functions*	*uniformly continuous*
(3)	D *the ordinary basis*	*constant functions*	*constant functions*	*continuous functions*
(4)	A *the approximate basis*	*constant functions*	*constant functions*	*approximately continuous*
(5)	D[X] *sections of the ordinary basis*	(?)	(?)	*continuous at each* $x \in X$
(6)	RD *the right Dini derivation basis*	(*)	(?)	*right continuous*
(7)	D^* *the dual of the basis D*	(*) *locally recurrent*	(?)	*each* $f^{-1}((c,d))$ *is dense in itself*
(8)	S *the symmetric derivation basis*	(*) *locally symmetric*	(?)	*symmetrically continuous*
(9)	AS *the approximate symmetric derivation basis*	(?)	(?)	*approximately symmetrically continuous* (?)
(10)	Q *the qualitative derivation basis*	*constant*	*constant*	*continuous*

The entries (2,1), (2,2), (3,1), (3,2), (4,1), (4,2), (10,1), and (10,2) arise from the fact that the derivation bases given have the partitioning property. We should elaborate on the type of problem intended for row (5) that uses the sections D[X] *of* D . *Let* Ω *denote a family of (presumably large) sets, such as for example sets whose complements are countable (or of measure zero, or first category, or* σ*-porous) and then the problem is to characterize the class of functions* f *for which there is some* $O \in \Omega$ *for which* f *is* D[O] - *locally constant (or* D[O]*-null).*

The entry (6,1) defines functions f *with the property that for any* $x \in R$ *there is a* $\delta(x) > 0$ *so that* $f(y) = f(x)$ *for all* $y \in [x, x + \delta(x))$ *(right locally constant functions). One can show that for any right locally constant function* f *there must be a countable closed set* C *such that for every interval* (a,b) *complementary to* C , f *is constant on* $[a,b)$.

The entry (7,1) defines functions f *that are locally recurrent, namely for each* $x \in R$ *and every* $\varepsilon > 0$ *there is at least one point* $y \in (x-\varepsilon, x+\varepsilon)$ *distinct from* x *for which* $f(y) = f(x)$. *This terminology has been used by a number of authors (e.g., Marcus* [82], *Šalát* [106], *Engquist* [27], *Bush* [17] *and Benson* [4]). *Using the basis* RD^* *which is the dual of* RD *one would have the corresponding idea of right locally recurrent. More interesting is to study the problem of characterizing the class of functions that are* $D^*[X]$ *or* $RD^*[X]$ *locally constant for* X *in various classes of sets (as above say). For this perhaps a better viewpoint is to look at the level structure of the functions and to express the problems in this language.*

The entries in row (8) are of some interest. A function f *is locally symmetric if for every* $x \in R$ *there is a* $\delta(x) > 0$ *so that* $f(x+h) = f(x-h)$ *for all* $0 < h < \delta(x)$. *One can prove that for a locally symmetric function* f *there is a countable closed set* C *such that* f *is constant on the complement of* C . *(This and related concerns can be found in Davies* [23], *Foran* [31], *Ruzza* [104], *and Thomson* [115].)

A function f is said to be symmetrically continuous if $f(x+h) - f(x-h) \to 0$ as $h \to 0$. The problem of determining the structure of such functions was posed by Hausdorff [38] *and has been studied by a number of authors, (See,* e.g., *Fried* [33], *Marcus* [81], *Ponomarev* [101], *and Preiss* [102]*).*

Finally, let us mention another direction that these concepts might lead to. Following Heindl and Kohler [40] *we say the function f is* locally increasing *if for every $x \in R$ there is a $\delta(x) > 0$ so that $f(y) \leq f(x) \leq f(z)$ for all $x-\delta(x) < y \leq x \leq z < x+\delta(x)$. It is clear that in our language such a function has the property that the interval function Δf^-, the negative part of Δf, is* D*-locally constant. Since* D *has the partitioning property it is immediate that Δf^- must vanish and so f is in fact increasing (in the loose sense). Of course this is elementary and not very interesting, but it does suggest a number of variants on this theme that could be pursued.*

One such variant has been given by Ornstein [98]. *His theorem can be rewritten in our language as follows:* if F is approximately continuous and ΔF^- is RA_0-locally constant then F is nondecreasing. *Here the derivation basis RA_p for $0 \leq p < 1$ is a right hand basis using sets of density exceeding p on the right.*

§9. The total variation of a function. Given a real function F on R and a set $X \subset R$ there are a number of ways of constructing a measure μ_F so that $\mu_F(X)$ in some way reflects the variation of the function F on the set X. The most familiar of these would seem to be to take $\mu_F(x) = \lim \mu_F^{(n)}(X)$ where

$$\mu_F^{(n)}(X) = \inf \{ \Sigma |F(b_i) - F(a_i)| : \cup [a_i, b_i] \supset X, b_i - a_i < \frac{1}{n} \}.$$

This construction is due to Carathéodory and is popularly known as "Munroe's Method II" after Munroe [91]. This method has the advantage of producing a metric outer measure for any function F (and indeed one could replace

$|F(b_i) - F(a_i)|$ by $|h([a_i, b_i])|$ where h is any interval function and the resulting measure still has these desirable properties). For functions F of bounded variation the measure μ_F does indeed carry the variational information of F ; but for general F the method has serious disadvantages. Ellis and Burry [16] have given an example of a continuous function F on the interval $[0,1]$ for which $\mu_F([0,1]) = 0$ and Bruckner [9] has shown that such behaviour is "typical" of continuous functions ("typical" in the familiar sense of Bruckner [10, Ch. XIII]). For the purposes of discussing total variation then these measures are of limited use beyond the case where F has bounded variation.

In the present theory the most natural candidate for a total variation measure is the "outer measure" F_B associated with a derivation basis B. The choice of B would be dictated by the intended application and so a variety of total variation measures are available. In particular the measures F_U and F_D associated with the uniform and the ordinary derivation bases respectively are very useful. It is immediately clear that the finiteness of these measures should have some relation to the boundedness of the variation of F. Henstock pointed out some time ago that the concepts of VBG_* and ACG_* that play such an important role in certain investigations in integration and differentiation can be expressed directly in terms of the measures F_D. (See Henstock [50, pp. 56-69] and the more recent paper Henstock [60].)

In this section we begin an investigation of the measures F_B for F an additive interval function using the point of view that these measures provide variational information about F. In any study there will be in fact two measures F_B and F_{B^*} where B^* is the dual basis for B and the pair of measures together describes the total variation of F in a useful manner.

Our first results connect the finiteness of the measures with more classical variational concepts.

THEOREM 9.1 Let B be a derivation basis and F an additive interval function.

9.1.1 *if* F *has bounded variation then* F_B *is finite,*

9.1.2 *conversely, assuming* B *has the partitioning property, if* F_B *is finite* F *has bounded variation on* R,

9.1.3 *if* B *is finer than the topology and* B *has the partitioning property then* $F_B(G) < +\infty$ *for an open set* G *if and only if* F *has bounded variation on* G .

PROOF. If F has bounded variation, say $v(F,R) < +\infty$ then for any $\beta \in B$ we certainly have $V(F,\beta) \leq v(F,R) < +\infty$ so that F_B is finite. In the other direction if $\beta \in B$ contains a partition of each of a finite sequence $\{J_i\}$ of nonoverlapping intervals then

$$\Sigma \left| F(J_i) \right| \leq V(F,\beta) .$$

Thus, for 9.1.2, if F_B is finite there is a $\beta \in B$ for which $V(F,\beta) < +\infty$ and from the above observation we see that if B has the partitioning property

$$v(F,R) \leq V(F,\beta) < +\infty$$

and F has bounded variation on R . Similarly the final assertion is obtained by using the fact that $B[G] \leq B(G)$ for a B that is finer than the topology and that B[G] permits partitions of every subinterval of G , if B itself has the partitioning property.

THEOREM 9.2 *Let* B *be a derivation basis,* F *an additive interval function, and suppose that* $F_B(X) < +\infty$ *for a set* $X \subset R$:

9.2.1 *then if* B *has the* Y-*decomposition property,* F *is* VBG *on* X , *and*

9.2.2 *if* B *has the closed* Y-*decomposition property and* X *is closed,* F *is* [VBG] *on* X .

The square bracket notation [VBG] indicates that the sequence of sets on which the function may be taken as VB can be chosen in such a way as to be closed. Similarly [ACG] may be defined. Recall that because of Saks [, p. 229] the notion [VBG*] would be merely equivalent with VBG*. Note that in this theorem we have not asserted the fact that F must be VB on X if $F_B(X) < +\infty$; also we see that under the assumption of 9.2.1 if F_B is σ-finite on a set X then F must be VBG on X .

PROOF. If $F_B(X) < +\infty$ then there is an element $\beta \in B[X]$ with $V(F,\beta) < +\infty$ and hence, using the Y-decomposition property, we may choose a decomposition of X into a sequence $\{X_n\}$ such that $\beta[X_n]$ partitions every interval $[x,y]$ with $x,y \in X_n$: let $\{J_i\}$ be any sequence of non-overlapping intervals with endpoints in X_n. Then $|F(J_i)| \le V(F,\beta(J_i))$ because $\beta(J_i)$ contains a partition of J_i and so

$$\Sigma\,|F(J_i)| \le \Sigma\, V(F,\beta(J_i)) \le V(F,\beta) < +\infty$$

which asserts that F is VB on each set X_n. This exhibits F as VBG on $X = \bigcup_{n=1}^{\infty} X_n$ as required.

For the second part of the theorem suppose that X is closed and that $\{X_n\}$ has been chosen so as to supply a closed Y-decomposition of X. Let $J_i = [a_i,b_i]$ $(i=1,2,\ldots,n)$ be a sequence of intervals with endpoints in $\overline{X}_m$, such that no two of these intervals are abutting. Then there are disjoint open intervals $\{K_i\}$ so that $K_i \supset J_i$ and we can treat each J_i in the following manner: choose $x_i,y_i \in X_m \cap K_i$ sufficiently close to a_i,b_i respectively so that $\beta[\overline{X}_m]$ partitions each of the intervals $[a_i,x_i]$ (or $[x_i,a_i]$) and $[b_i,y_i]$ (or $[y_i,b_i]$) and $[x_i,y_i]$; this is possible because of the nature of the decomposition. This gives

$$|F(J_i)| = |F(b_i) - F(a_i)| \leq |F(x_i) - F(a_i)| + |F(y_i) - F(x_i)|$$
$$+ |F(b_i) - F(y_i)|$$
$$\leq 3\ V(F, \beta(K_i))$$

and so

$$\Sigma\, |F(J_i)| \leq 3\ V(F, \beta[\overline{X}_m]) \leq 3\ V(F, \beta[X]) < +\infty\ .$$

From this inequality we can conclude that F is VB on each $\overline{X}_m$ and so [VBG] on X as required.

EXAMPLE 9.3 (Characterization of the class VBG). Using the ordinary derivation basis* D *the outer measures* F_{D} *carry a great deal of variational information about the function* F *. Certainly because of Theorem 9.1 we know that the finiteness of* F_{D} *on any interval* $[a,b]$ *would require* F *to be of bounded variation there. More generally we can even characterize the class* VBG_* *. We will restrict attention just to continuous functions on an interval* $[a,b]$ *, although more general formulations are possible: the following assertions are equivalent for a continuous function* F *on an interval* $[a,b]$ *and for* $X \subset (a,b)$ *:*

(1) F *is* VBG_* *on* X ,

(2) *[HENSTOCK]* F_{D} *is σ-finite on* X ,

(3) *[WARD] there is a continuous increasing function* G *on* $[a,b]$ *so that* $\overline{D}\, F_G(x)$ *and* $\underline{D}\, F_G(x)$ *are finite at each point of* X ,

(4) *there is an increasing sequence of sets* $\{X_n\}$ *with* $X = \cup X_n$ *and a sequence of continuous functions* $\{G_n\}$ *of bounded variation on* $[a,b]$ *so that*

$$V(F - G_n, \mathrm{D}[X_n] = 0 \quad \text{for all} \quad n\ .$$

QUERY 9.3.1 The class of continuous functions that are VBG_* can be topologized in a way similar to the usual topologization of the class of continuous functions of bounded variation; for the latter class one simply takes the norm $\|F\| = V(F,D)$. If $CVBG_*$ denotes the linear space of all continuous additive interval functions that are VBG_* then write $CVBG_*(\{E_n\})$, where $\{E_n\}$ is an expanding sequence of closed sets covering the real line, for those F for which $F_D(E_n) < +\infty$ for each n and topologize $CVBG_*(\{E_n\})$ with the collection of seminorms $\{p_n\}$ where $p_n(F) = F_D(E_n)$. Finally then the space $CVBG_*$ is just the union of the spaces $CVBG_*(\{E_n\})$ over all such sequences $\{E_n\}$ and it can be topologized in the familiar manner.

Can this construction be used to give any insight into the nature of VBG_* functions? In particular can the type of questions addressed by Garg [34] *for the Banach space of functions of bounded variation be carried over to this setting?*

Our next lemma is due to Henstock [55] and is a key tool in the study of these total variation measures. It represents one of the main consequences of the additivity assumption on a derivation basis and is, perhaps, the reason that in Henstock's abstract versions of this theory he has incorporated the additivity assumption into his definition of a "division space".

LEMMA 9.4 [HENSTOCK] Let B be a derivation basis that is additive, filtering down, and has the partitioning property and suppose that h is a nonnegative subadditive interval function. If the function H defined by $H(I) = V(h,B(I))$ $(I \in I_+)$ is finite then H is an additive interval function and

$$V(H-h,B(I)) = 0$$

for all $I \in I_+$. If moreover B is assumed to have σ-local character and to be finer than the topology then

$$V(H-h,B) = 0$$

and consequently $F_B \equiv h_B$. On the further and final assumption that B ignores no point, $F_{B^*} \equiv h_{B^*}$.

PROOF. For a fixed interval I and an $\varepsilon > 0$ there must be a partition π_0 of I for which

$$H(I) = V(h,B(I)) \le \sum_{i=1}^{n} h(I_i) + \varepsilon$$

where we have written

$$\pi_0 = \{(I_i,x_i) : i=1,2,\dots,n\} .$$

Using the fact that B is additive and filtering down we may choose a $\beta \in B(I)$ that splits at each I_i , i.e., so that

$$\beta \subset \bigcup_{i=1}^{n} \beta(I_i) .$$

Let $\pi \subset \beta$ be any partition of I ; from the subadditivity of h we have

$$H(I) - \varepsilon < \sum_{i=1}^{n} h(I_i) \le \sum_{i=1}^{n} \Sigma_{(J,x) \in \pi(I_i)} h(J) = \Sigma_{(J,x) \in \pi} h(J) .$$

By our elementary computations for the variation we know, since B is additive, that H is an additive interval function and that H exceeds h : hence for any partition $\pi \subset \beta$, which we may always take to be a partition of I , we will have

$$\Sigma_{(J,x) \in \pi} |H(J) - h(J)| = \Sigma_{(J,x) \in \pi} H(J) - h(J) = H(I) - \Sigma_{\pi} h(J) < \varepsilon .$$

Thus $V(H-h,\beta) < \varepsilon$ and so $V(H-h,B(I)) < \varepsilon$. Since $\varepsilon > 0$ and I are arbitrary we have proved the first part of the lemma.

For the second part recall from §2 above that if B is of σ-local character and finer than the topology then $V(H-h,B) = 0$ whenever $V(H-h,B(I)) = 0$ for every interval I . Thus from elementary properties of the variation we obtain

$$\begin{aligned}|H_B(X) - h_B(X)| &= |V(H,B[X]) - V(h,B[X])| \\ &\leq V(H-h,B[X]) \\ &\leq V(H-h,B) = 0\end{aligned}$$

and so the outer measures h_B and H_B agree on each set as stated.

Finally using some properties of the dual basis established in Chapter Two, §8, let $\beta \in B[X]$ and $\beta^* \in B^*[X]$ so that we know $\beta \cap \beta^*$ also belongs to $B^*[X]$ and then we have

$$\begin{aligned}H_{B^*}(X) &= V(H,B^*[X]) \leq V(H,\beta \cap \beta^*) \\ &\leq V(H-h , \beta \cap \beta^*) + V(h,\beta \cap \beta^*) \\ &\leq V(H-h,\beta) + V(h,\beta^*) .\end{aligned}$$

Now letting β and β^* vary we obtain

$$H_{B^*}(X) \leq 0 + V(h,B^*[X]) = h_{B^*}(X) .$$

As the other inequality, $h_{B^*}(X) \leq H_{B^*}(X)$, would follow in the same way the proof of the lemma is complete.

Henstock's lemma is really an abstract version of the classical Jordan decomposition theorem as our next theorem illustrates.

THEOREM 9.5 [JORDAN DECOMPOSITION THEOREM] Let B be a derivation basis that is additive, filtering down, and has the partitioning property and suppose that F is an additive interval function that has bounded variation on every interval. Define each of the interval functions F^+, F^-, P, N, and T by writing for every interval I,

$$F^+(I) = \max\{F(I),0\}, \quad F^-(I) = \max\{-F(I),0\},$$

$$P(I) = V(F^+,B(I)), \quad N(I) = V(F^-,B(I)), \text{ and}$$

$$T(I) = V(F,B(I)).$$

Then

(i) P, N, and T are additive nonnegative interval functions,

(ii) $F = P - N$ and $T = P + N$, and

(iii) $V(P - F^+,B(I)) = V(N - F^-,B(I)) = V(T - F,B(I)) = 0$ for every interval I.

If moreover B has σ-local character and is finer than the topology then

(iv) $V(P - F^+,B) = V(N - F^-,B) = V(T - F,B) = 0$, and

(v) $P_B \equiv F^+_B$, $N_B \equiv F^-_B$, and $T_B \equiv F_B$.

Finally, if B also ignores no point then,

(vi) $P_{B*} \equiv F^+_{B*}$, $N_{B*} \equiv F^-_{B*}$, and $T_{B*} \equiv F_{B*}$.

PROOF. For the most part this follows directly from the preceding lemma. For example, suppose that we wish to prove that $F = P-N$. We have, using that lemma for the subadditive interval functions F^+ and F^-:

$$F = F^+ - F^-$$

$$V(F^+ - P,B(I)) = 0$$

$$V(F^- - N,B(I)) = 0$$

so that

$V(F-(P-n),B(I)) \leq V(F-(F^+-F^-),B(I)) + V(F^+-P,B(I)) + V(F^- -N,B(I)) = 0$

for all intervals I. Since B has the partitioning property and $F-(P-N)$ is additive, $F-(P-N)$ vanishes identically proving that $F = P-N$. The remaining details are similar.

EXAMPLE 9.5.1 (Decomposition theorems for functions of bounded variation) By using in Theorem 9.5 the basis R *or* D *one has the standard Jordan decomposition of a function* F *that is locally of bounded variation. This is rather more interestingly expressed as an integration (using either* R *or* D*):*

a function F *that is locally of bounded variation may be written as the difference of two nonnegative additive interval functions* $F = P-N$ *where*

$$P(I) = \int_{(I)} dF^+ \quad \text{and} \quad N(I) = \int_{(I)} dF^- .$$

Continuing this same theme we can establish in the same simple way some other familiar decomposition theorems for such functions. Here we require the integral based on D *, the ordinary derivation basis.*

[DISCRETE/CONTINUOUS DECOMPOSITION]: a function F *that is locally of bounded variation may be written as the sum of a continuous and a discrete function* $F = F_c + F_d$ *where*

$$F_c(I) = \int_{(I)} \chi(C_F, \cdot)dF \quad \text{and} \quad F_d(I) = \int_{(I)} \chi(C_F', \cdot)dF$$

and

$$C_F = \{x \in R : F \text{ is continuous at } x\} , \quad C_F' = R \setminus C_F .$$

(The integral expressing $F_d(I)$ *can be easily seen to have the value*

$$\Sigma_{c \in I \cap C_F'} \{F(c+) - F(c)\} + \{F(c) - F(c-)\}$$

with an appropriate term deleted from the sum if c *is an endpoint of* I *.)*

[LEBESGUE DECOMPOSITION]: *a function* F *that is locally of bounded variation may be written as the sum of an absolutely continuous and a singular function* $F = F_{ac} + F_s$ *where*

$$F_{ac}(I) = \int_{(I)} \chi(\Delta_F, \cdot)dF \quad \text{and} \quad F_d(I) = \int_{(I)} \chi(\Delta_F', \cdot)dF$$

and

$$\Delta_F = \{x \in R : F'(x) \text{ exists finitely}\} \quad \text{and} \quad \Delta_F' = R \setminus \Delta_F .$$

(The integral expressing $F_{ac}(I)$ *can also be written in the form* $F_{ac}(I) = \int_{(I)} f(x)dx$ *where* $F(x) = F'(x)$ *for* $x \in \Delta_F$ *and* $f(x) = 0$ *otherwise.)*

Given the apparatus we have developed the proofs of the decomposition properties are remarkably simple, and the form of the decomposition may be suggestive of generalizations.

The next theorem plays a key role in a number of investigations in the differentiation theory. Because it is so intimately connected with the classical Vitali theorem we have so labelled it.

THEOREM 9.6 [VITALI THEOREM FOR LEBESGUE-STIELTJES MEASURES] Let B be a derivation basis that is straddled, filtering down, additive, finer than the topology, and has σ-local character and the partitioning property. Suppose that F is a continuous function having bounded variation on every interval and let T be its corresponding total variation function. Then if μ_T denotes the usual Lebesgue-Stieltjes outer measure generated by T we must have

$$F_B \equiv F_{B^*} \equiv T_B \equiv T_{B^*} \equiv \mu_T .$$

PROOF. We already know from the previous theorem, the Jordan decomposition theorem, that in this situation $F_B = T_B$ and $F_{B^*} = T_{B^*}$, so we can focus attention just on T which is essentially a continuous monotonic nondecreasing function on R, and μ_T which is the usual Lebesgue-Stieltjes measure generated by T (by the Carathéodory process outlined at the opening of this section). We will prove that

$$T_B \le \mu_T \le T_{B^*}$$

and since we know already that $T_{B^*} \le T_B$ the theorem will be proved.

To see that $\mu_T \le T_{B^*}$ suppose that $X \subset R$ and that $\beta^* \in B^*[X]$: then applying the Vitali covering theorem relative to the measure μ_T (for a proof of the Vitali covering theorem for Lebesgue-Stieltjes measures see, for example, de Guzman [37, p. 27]) there must be a sequence $\{(I_i, x_i\} \subset \beta^*$ with I_i and I_j nonoverlapping for distinct i and j and such that

$$\mu_T(X \bigcup_{n=1}^{\infty} I_n) = 0 .$$

This gives

$$\mu_T(X) \le \sum_{n=1}^{\infty} \mu_T(I_n) = \sum_{n=1}^{\infty} T(I_n) \le V(T,\beta^*)$$

and hence letting β^* vary we have $\mu_T(X) \le T_{B^*}(X)$ for any X as required to establish the inequality $\mu_T \le T_{B^*}$.

To see that $T_B \le \mu_T$ let $X \subset R$ and $\varepsilon > 0$ be given and then, using a well known property of Lebesgue-Stieltjes measures, there must be an open set $G \supset X$ such that

$$\mu_T(G) \le \mu_T(X) + \varepsilon .$$

Because B is finer than the topology there is a $\beta \in B$ for which $\beta[G] \subset \beta(G)$. Writing $\{I_i\}$ for the sequence of closed intervals whose interiors are the components of the open set G we must have

$$T_B(X) \le V(T,\beta[X]) \le V(T,\beta[G]) \le V(T,\beta(G)) \le \sum_{n=1}^{\infty} T(I_n) = \mu_T(G) \le \mu_T(X) + \varepsilon .$$

As $\varepsilon > 0$ is arbitrary and this holds for any $X \subset R$ we have proved that $T_B \le \mu_T$ as required. The theorem now follows.

The scope of this theorem is severely limited by restricting it to functions of bounded variation on each interval. One method whereby these results can be lifted to a larger class of functions is presented in the next definition. The class of functions so presented form a generalized version of the class of functions that are continuous and VBG_* .

DEFINITION 9.7 Let B be a derivation basis and h an interval-point function. We shall say that h is CVBG(B) on a set $X \subset R$ if there is an expanding sequence of sets $\{X_n\}$ with

$$X = \bigcup_{n=1}^{\infty} X_n$$

and a sequence of functions $\{G_n\}$ that are continuous and of bounded variation on each interval so that

$$V(h - G_n, B[X_n]) = 0$$

for each n .

Note that a function h that is CVBG(B) on a set X is necessarily B-continuous at each point of X under the natural assumption that any function that is continuous is also B-continuous. This is the case with most derivation bases (although it would not be true for the sharp derivation basis $D^{\#}$). Our next theorem takes the Vitali theorem and extends it from Lebesgue-Stieltjes measures to more general measures h_B for functions that are CVBG(B) , under certain hypotheses.

THEOREM 9.8 [VITALI THEOREM FOR CVBG(B) FUNCTIONS] Let B be a derivation basis that is straddled, additive, filtering down, finer than the topology, and has σ-local character and the partitioning property. Then if h is an interval-point function that is CVBG(B) on a set X , 9.8.1 $h_B(Y) = h_{B^*}(Y)$ <u>for all</u> $Y \subset X$,

9.8.2 h_B <u>and</u> h_{B^*} <u>are</u> σ-<u>finite on</u> X ,

9.8.3 $h_B(\{x\}) = h_{B^*}(\{x\}) = 0$ <u>for all</u> $x \in X$ <u>so that in particular</u> h <u>is</u> B-<u>continuous at each point of</u> X .

PROOF. Let $\{X_n\}$ be an expanding sequence of sets covering X and $\{G_n\}$ a sequence of functions continuous and of bounded variation on each interval such that $V(h - G_n, B[X_n]) = 0$ for all n . Then for any $Y \subset X$ we must have by the Vitali theorem and by the increasing sets property,

$$h_B(Y \cap X_n) = (G_n)_B(Y \cap X_n) = (G_n)_{B^*}(Y \cap X_n) = h_{B^*}(Y \cap X_n)$$

and so

$$h_B(Y) = \lim h_B(Y \cap X_n) = \lim h_{B^*}(Y \cap X_n) \le h_{B^*}(Y) .$$

But since $h_{B^*} \le h_B$ always in this situation it follows that for any $Y \subset X$, $h_B(Y) = h_{B^*}(Y)$ as required.

EXAMPLE 9.9 (Characterization of the class ACG_*). Parallel to the characterization in example 9.3 above of the class VBG_* we can provide a similar characterization of the class ACG_* , again using the derivation basis D .

<u>*The following assertions are equivalent for a continuous function*</u> F <u>*on an interval*</u> $[a,b]$ <u>*and a closed set*</u> $X \subset (a,b)$:

(1) F <u>*is*</u> ACG_* <u>*on*</u> X ,

(2) F_D <u>*is*</u> σ-<u>*finite on*</u> X <u>*and*</u> $F_D(N) = 0$ <u>*for every set*</u> $N \subset X$ <u>*for which*</u> $|N| = 0$,

(3) <u>*there is an increasing sequence of sets*</u> $\{X_n\}$ <u>*with*</u> $X = \bigcup X_n$ <u>*and a sequence*</u> $\{G_n\}$ <u>*of absolutely continuous functions on*</u> $[a,b]$ <u>*so that*</u>

$$V(F - G_n, \mathcal{D}[X_n]) = 0 \ \underline{\text{for all}} \ n .$$

QUERY 9.10 *It seems likely that there is a Ward type characterization of the class* ACG_* *to add to example 9.9. Thus,*

(4) *there is an absolutely continuous increasing function* G *on* $[a,b]$ *so that* $\overline{D}\,F_G(x)$ *and* $\underline{D}\,F_G(x)$ *are finite at each point of* X .

Is this equivalent with (1), (2), and (3)? Saks [105, p. 237] *asserts this only in the case that* X *be an open interval.*

EXAMPLE 9.11 *(Zero variation) For an arbitrary function* F *and a set* $X \subset R$ *the expressions* $F_{\mathcal{D}}(X)$, $F_{\mathcal{D}^*}(X)$, *and* $|F[X]|$ *are quite independent. But there are some interesting interrelations, particularly in the event that one or more vanishes. Here we use the notation* $F[X] = \{y : F(x) = y$ *for some* $x \in X\}$ *so that* F *must be interpreted as a point function (although* $|F[X]|$ *can be determined solely from* ΔF *).*

These results are known: let F *be continuous on* R *and let* $X \subset R$.

9.11.1 *if* $|F[X]| = 0$ *then* $F_{\mathcal{D}^*}(X) = 0$, *(Thomson* [113]*)*

9.11.2 $F_{\mathcal{D}}(X) = 0$ *if and only if* F *is* VBG_* *on* X *and* $|F[X]| = 0$ *(Henstock* [60, *Theorem 4*]*)* .

EXAMPLE 9.12 *(Total variation of a typical continuous function) Using the word "typical" in the sense of Bruckner* [10, *Chapter XIII*] *we can show that the measures* $\Delta f_{\mathcal{D}}$, $\Delta f_{\mathcal{D}^*}$, *and* μ_f *(that we know are closely related for continuous* VBG_* *functions) have a certain predictable behaviour. Here* $C[a,b]$ *is the usual Banach space of continuous functions on the interval* $[a,b]$.

A typical continuous function f *in* $C[a,b]$ *(i.e., for a residual set of* f *in that space) has*

$$\mu_f([a,b]) = \Delta f_{D^*}([a,b]) = 0$$

while Δf_D *is non* σ*-finite on any set of positive measure. (Bruckner [9] for* μ_f *and Thomson [113] for* Δf_D *and* Δf_{D^*} *.)*

EXAMPLE 9.13 (Lebesgue differentiation theorem) We indicate here an interesting if unorthodox proof of the Lebesgue differentiation theorem (i.e., the assertion that continuous functions of bounded variation must have finite derivatives almost everywhere). Lebesgue's proof relied heavily on the integration theory he had developed and since then numerous proofs have been given. The proof we give here exploits the properties of the variation and should not be considered elementary since it uses the Vitali theorem. Nonetheless the computations are quite simple:

Let F *be a continuous function that is* VBG_* *on a set* X *. Then* $F'(x)$ *exists finitely almost everywhere in* X *, and* $F'(x)$ *exists finitely or infinitely* F_D*-almost everywhere in* X *.*

The proof (cf. Thomson [114]) follows from three simple steps. The first step requires showing that the finiteness of F_D *on a set* $Y \subset X$ *requires that both* $\underline{D}\, F(x)$ *and* $\overline{D}\, F(x)$ *be finite a.e. in* Y *.*

For the second step note that for a continuous function $F_+'(x)$ *either exists finitely or infinitely or else there are rational numbers* r *and* s *with* $|r| \neq |s|$ *such that both* r *and* s *are derived numbers of* F *at* x *. Let* $X_{rs} = \{x \in X : r$ *and* s *are derived numbers of* F *at* $x\}$ *. Then* $V(F - rm, D^*[X_{rs}]) = V(F - sm, D^*[X_{rs}]) = 0$ *. But the identities* $F_D \equiv F_{D^*}$ *and* $m_D \equiv m_{D^*}$ *on* X *then lead easily to the assertion*

$$|r| m_D(X_{rs}) = F_D(X_{rs}) = |s| m_D(X_{rs}) .$$

From this we conclude that the set of points x *in* X *at which the right derivative* $F_+'(x)$ *[and hence similarly the left derivative* $F_-'(x)$*] does not exist finitely or infinitely has both* m_D *and* F_D *measure zero.*

The final step then needs only a proof that the set of points x *at which* $F_+'(x) = \pm\infty$ *and yet* $F_-'(x) = \mp\infty$ *is countable. With these three steps then the theorem now follows easily.*

INDEX OF NOTATIONS

RD, LD	right and left Dini derivation basis	II,§3(3.7)
PS	modified symmetric basis	II,§5(5.10)
RA_p	$(0 \leq p < 1)$ right density p basis	III,§8(8.16)
S	symmetric derivation basis	II,§3(3.10)
T	trivial basis	II,§3(3.1)
U	uniform basis	II,§3(3.2)
$U^{\#}$	sharp version of uniform basis	II,§3(3.5)

<u>properties of a derivation basis</u>:

additive	II,§4(4.7)
$\mathcal{C}$-complete	II,§5(5.9)
endpoint tagged	II,§4(4.3)
filtering down	II,§4(4.1)
finer than the topology	II,§4(4.9)
decomposition properties	II,§7
$\mathcal{H}$-complete	II,§5(5.9)
ignores no point	II,§4(4.10)
intersection conditions	II,§9(9.6)
local character	II,§6(6.1)
σ-local character	II,§6(6.2)
partitioning property	II,§5(5.1)
separates	II,§4(4.5)
splits	II,§4(4.5)
straddled	II,§4(4.3)

REFERENCES

[1] A.P.Baisnab, On absolute Dini derivatives, Rev. Roum. Math. Pures et Appl. 15 (1970), 1593-1597.

[2] J.D.Baker and R.A.Shive, On the existence of ψ-integrals, Rend. Circ. Mat. Palermo (2) 21 (1972), 293 - 304.

[3] C.L.Belna, M.J.Evans, and P.D.Humke, Symmetric and strong different-i tion, Amer. Math. Monthly, 86 (1979), 121-123.

[4] D.C.Benson, Nonconstant locally recurrent functions, Pacific J. Math. 21 (1967), 437-443.

[5] A.S.Besicovitch, On linear sets of points of fractional dimension, Math. Ann. 101 (1929), 161-193.

[6] P.C.Bhakta and D.K.Mukhopadhyay, On approximate strong and approx-imate uniform differentiability, J.Indian Inst. Sci. 61 (1979), 103-118.

[7] A.M.Bruckner, On the differentiation of integrals in Euclidean space, Fund. Math. 66 (1969/70), 129 - 135.

[8] ------------, Differentiation of integrals, Amer. Math. Monthly 78 (1971),1-51.

[9] ------------, A note on measures determined by continuous functions, Canad. Math. Bull. (15) 2 (1972), 289-291.

[10] ------------, <u>Differentiation of Real Functions</u>, Lecture Notes in Mathematics #659, Springer-Verlag (1978).

[11] A.M.Bruckner and C.Goffman, The boundary behaviour of real functions in the upper half-plane, Rev. Roum. Math. Pures et Appl. XI (1966), 507-518.

[12] --------------------------, Approximate differentiation, Real Analysis Exchange, 6 (1980-81),9-65.

[13] A.M.Bruckner, R.J.O'Malley and B.S.Thomson, Path derivatives: a unified view of certain generalized derivatives (to appear).

[14] J.C.Burkill, Functions of intervals, Proc. London Math. Soc. 2 22 (1924), 275-310.

[15] -----------, The derivates of functions of intervals, Fund. Math. 5 (1924), 321-327.

[16] J.H.Burry and H.W.Ellis, On measures determined by continuous functions that are not of bounded variation, Canad. Math. Bull. (13) 1 (1970), 121-124.

[17] K.A.Bush, Locally recurrent functions, Amer. Math. Monthly 69 (1962) 199-206.

[18] F.S.Cater, When total variation is additive, Proc. Amer. Math.Soc. 84 (1982), 504-508.

[19] O.Cauchy, Résumé des lecons données a l'Ecole Royale Polytechnique, le calcul Infinitésimal, vol. I (Paris), 1823.

[20] L.Cesari, Quasi additive set functions and the concept of integral over a variety, Trans. Amer. Math. Soc. 102 (1962), 94-113.

[21] P.Cousin, Sur les fonctions de n variables complexes, Acta. Math. 19 (1895), 1-62.

[22] A.Császár, Sur une généralisation de la notion de dérivée, Acta. Sci. Math. Szeged 16 (1955), 137-159.

[23] R.O.Davies, Symmetric sets are measurable, Real Analysis Exchange 4 (1978/79), 87-89.

[24] A.Denjoy, Memoire sur la totalisation des nombres dérivées non sommables, Ann. Ecole Norm., 33 (1916), 127-222.

[25] --------, Totalisation des dérivées premieres generalisees I, C.R. Acad. Sci. Paris 241 (1955), 617-620.

[26] --------, Totalisation des dérivées premieres symmetrique II, C.R. Acad. Sci. Paris 241 (1955), 829-832.

[27] M.L.Engquist, On the existence of locally recurrent functions, Amer. Math. Monthly 84 (1977), 191-195.

[28] M.J.Evans, On continuous functions and the approximate symmetric derivatives, Colloq. Math. 31 (1974), 129-136.

[29] M.J.Evans and P.D.Humke, Parametric differentiation, Colloq. Math. 45 (1981), 125-131.

[30] M.Esser and O.Shisha, A modified differentiation, Amer. Math. Monthly 71 (1964), 904-906.

[31] M.Foran, Symmetric functions, Real Analysis Exchange, 1 (1976), 38-40.

[32] L.R.Ford, Interval additive propositions, Amer. Math. Monthly 64 (1957), 106-108.

[33] H.Freid, Über die symmetrische Stetigkeit von Funktionen, Fund. Math. 29 (1937), 134-137.

[34] K.M.Garg, Characterizations of absolutely continuous and singular functions, Proc. of the Conference on the Constructive Theory of Functions (Approximation theory) Budapest, 1969, 183-188.

[35] B.C.Getchell, On the equivalence of two methods of defining Stieltjes integrals, Bull. Amer. Math. Soc., (2) 41 (1935), 413-418.

[36] E.Goursat, Sur la définition générale des fonctions analytiques d'après Cauchy, Trans. Amer. Math. Soc. 1 (1900), 14-16.

[37] M.de Guzmán, Differentiation of Integrals in R^n, Lecture Notes in Math. #481, Springer-Verlag (1975).

[38] F.Hausdorff, Problem 62, Fund. Math. 25 (1935), 578.

[39] C.Hayes and C.Pauc, Derivation and Martingales, Springer, Berlin (1970).

[40] G.Heindl and G.Köhler, Ein Monotoniekriterium, Bayer. Akad. Wiss. Math.-Natur. Kl. S.-B. 1968, Abt. II,(1969), 107-112.

[41] R.Henstock, On interval functions and their integrals, J. London Math. Soc. 21 (1946), 204-209.

[42] ----------, On interval functions and their integrals II, J. London Math. Soc. 23 (1948), 118-128.

[43] ----------, Density integration, Proc. London Math. Soc. (2) 53 (1951), 192-211.

[44] ----------, On Ward's Perron Stieltjes integral, Canad. J. Math. 9 (1957), 96-109.

[45] ----------, A new descriptive definition of the Ward integral, J. London Math. Soc. 35 (1960), 43-48.

[46] ----------, The use of convergence factors in Ward integration, Proc. London Math. Soc. (3) 10 (1960), 107-121.

[47] ----------, The equivalence of generalized forms of the Ward, variational, Denjoy-Stieltjes, and Perron-Stieltjes integrals, Proc. London Math. Soc. (3) 10 (1960), 281-303.

[48] ----------, N-variation and N-variational integrals of set functions, Proc. London Math. SOc. (3) 11 (1961), 109-133.

[49] ----------, Definitions of Riemann-type of the variational integrals, Proc. London Math. SOc. (3) 11 (1961), 402-418.

[50] ----------, Theory of integration, (Butterworths, 1963).

[51] ----------, The integrability of functions of interval functions, J.London Math. Soc. 39 (1964), 589-597.

[52] ----------, Majorants in variational integration, Canad. J. Math. 18 (1966), 49-74.

[53] ----------, A Riemann-type integral of Lebesgue power, Canad. J. Math. 20 (1968), 79-87.

[54] ----------, Linear Analysis, (Butterworths, 1968).

[55] ----------, Generalized integrals of vector-valued functions, Proc. London Math. Soc. (3) 19 (1969), 509-536.

[56] ----------, Integration by parts, Aequationes Math. 9 (1973),1-18.

[57] ----------, Integration in product spaces, including Wiener and Feynmann integration, Proc. London Math. Soc. (3) 27 (1973), 317-344.

[58] ----------, Additivity and the Lebesgue limit theorems, Proc. C. Caratheodory Int. Symposium (Athens, 1973), 223-241. Greek Math. Soc., Athens, 1974.

[59] ----------, Integration, variation, and differentiation in division spaces, Proc. Roy. Irish Acad. Sect. A 78 (1978), 69-85.

[60] ----------, The variation on the real line, Proc. Roy. Irish Acad. Sect. A 79 (1979), 1-10.

[61] ----------, Generalized Riemann integration and an intrinsic topology, Canad. J. Math. 32 (1980), 395-413.

[62] ----------, Division spaces, vector-valued functions and backwards martingales, Proc. Roy. Irish Acad. Sect. A 80 (1980), 217-233.

[63] T.H.Hildebrandt, The Borel theorems, Bull. Amer. Math. Soc. 32 (1926), 423-474.

[64] A.Ionescu Tulcea and C.Ionescu Tulcea, Topics in the theory of lifting, Springer (1969).

[65] K.Jacobs, Measure and integrals, Academic Press, New York, 1978 (with an appendix by J.Kurzweil).

[66] J.M.Jędrzejewski, On limit numbers of real functions, Fund. Math. 83 (1973/74), 269-281.

[67] S.Kempisty, Fonctions d'intervalle non additive, Hermann et Cie, Paris (1939).

[68] H.Kenyon and A.P.Morse, Web derivatives, Memoirs Amer. Math. Soc. 132 (1973).

[69] H.Kestelman, Modern theories of integration, Dover (1960).

[70] A.Khintchine, Sur la dérivation asymptotique, C.R.Acad. Sci. Paris 164 (1917), 142-144.

[71] ------------, Recherches sur la structure des fonctions mesurables, Fund. Math. 9 (1927), 212-279.

[72] J.Kurzweil, Generalized ordinary differential equations and continuous dependence on a parameter, Czech. Math. Journal 82 (1957), 418-449.

[73] M.Laczkovitch and G.Petruska, Remarks on a problem of A.M.Bruckner, Real Analysis Exchange, 6 (1980/81), 120-126.

[74] B.K.Lahiri, On non-uniform differentiability, Amer. Math. Monthly 67 (1960), 649.

[75] L.Larson, The symmetric derivative (to appear).

[76] --------, The Baire class of approximate symmetric derivatives, (to appear).

[77] H.Lebesgue, Sur les intégrales singulieres, Ann. Fac. Sci. Toulouse (1) 3 (1909), 4-117.

[78] N.Lusin, Sur un théorème fondamentale du calcul intégral (in Russian) Trans. Moscow Math. Soc. 28 (1911-12), 270.

[79] W.A.J.Luxemburg, Arzela's dominated convergence theorems for the Riemann integral, Amer. Math. Monthly 78 (1971), 970-979.

[80] N.C.Manna, On uniform Dini derivatives, Colloq. Math. 21 (1970), 95-100.

[81] S.Marcus, Les ensembles F_σ et la continuité symmétrique, Acad. R.P. Romine. Bul. Şti. Secţ. Şti. Mat. Fiz. 7 (1955), 871-886.

[82] --------, La limite approximative qualitative, Com. Acad. R.P. Romane 3 (1953), 9-12.

[83] M.Mastalerz-Wawrzńczak, On a certain condition of the monotonicity of functions, Fund. Math. 97 (1977), 187-198.

[84] J.G.Mauldon, The differentiability of locally recurrent functions, Amer. Math. Monthly 72 (1965), 983-985.

[85] J.Mawhin, Introduction a l'analyse, Cabay, Louvain-la-Neuve (1979).

[86] P.McGill, Properties of the variation, Proc. Roy. Irish Acad. Sect. A 75 (1975), 73-77.

[87] J.McGrotty, A theorem on complete sets, J. London Math. Soc. 37 (1962), 338-340.

[88] R.M.McLeod, The generalized Riemann integral, Carus Math. Monographs, #20, M.A.A. (1980).

[89] E.J.McShane, A Riemann-type integral that includes Lebesgue-Stieltjes, Bochner and stochastic integrals. Memoir Amer. Math. Soc., #88 , (1969).

[90] -----------, A unified theory of integration, Amer. Math. Monthly 80 (1973), 349-359.

[91] M.E.Monroe, Measure and integration, Second Ed., Addison-Wesley (1971).

[92] R.M.F.Moss and G.T.Roberts, A creeping lemma, Amer. Math. Monthly 75 (1968), 649-652.

[93] S.N.Mukhopadhyay, On approximate Schwarz differentiability, Monatsh. Math. 70 (1966), 454-460.

[94] C.Neugebauer, Darboux functions of Baire class one and derivatives, Proc. Amer. Math. Soc. 13 (1962), 838-843.

[95] R.J.O'Malley, Selective derivates, Acta Math. Acad. Sci. Hung. 29 (1977), 77-97.

[96] ------------, Decomposition of approximate derivatives, Proc. Amer. Math. Soc. 69 (1978), 243-247.

[97] R.J.O'Malley and C.E.Weil, Composite derivatives, (to appear).

[98] D.Ornstein, A characterization of monotone functions, Illinois J. Math. 15 (1971), 73-76.

[99] G.Peano, Sur la definition de la derivee, Opere Scelte, v.I, Edizioni Cremonese, Roma (1957), 210-212.

[100] S.Pollard, The Stieltjes integral and its generalizations, Quart. J. Math. Oxford Ser. 49 (1920), 87-94.

[101] S.P.Ponomarev, Symmetrically continuous functions, Mat. Zametki 1 (1967), 385-390.

[102] D.Preiss, A note on symmetrically continuous functions, Časopis Pěst. Mat. 96 (1971), 262-264.

[103] I.Ridder, Das allegemeine Denjoysche Integral, Fund. Math. 21 (1933).

[104] I.Z.Ruzsa, Locally symmetric functions, Real Analysis Exchange 4 (1978-79), 84-86.

[105] S.Saks, Theory ofthe integral (Warsaw, 1937).

[106] T.Šalát, On locally recurrent functions, Amer. Math. Monthly 77 (1970), 384-385.

[107] D.N.Sarkhel and A.K.De, The proximally continuous integral, J. Austral. Math. Soc. (Series A) 31 (1981), 26-45.

[108] P. Shanahan, A unified proof of several basic theorems of real analysis, Amer. Math. Monthly 79 (1972), 895-898.

[109] G.H.Sindalovskiĭ, The derived numbers of continuous functions, Izv. Akad. Nauk SSSR Ser. Mat. 32 (1968), 983-1023.

[110] O.Stolz, Über einen zu einer unendlichen Punktmenge gehorigen Grenzwert, Math. Ann. 23 (1884), 152-156.

[111] T.Świątkowski, On a certain generalization of the notion of the derivative (in Polish), Zeszyty Naukowe Ploitechniki Łódzkiej 149, Matematyka 2.1 (1972), 89-103.

[112] B.S.Thomson, On the total variation of a function, Canad. Math. Bull. (24) 3 (1981), 331 - 340.

[113] -----------, Outer measures and total variation, Canad. Math. Bull. (24) 3 (1981), 341 - 345.

[114] -----------, On the derived numbers of VBG_* functions, Journal London Math. Soc. (2) 22 (1980), 473-485.

[115] -----------, On full covering properties, Real Analysis Exchange 6 (1980/81), 77-93.

[116] G.Tolstov, La méthode de Perron pour l'intégrale de Denjoy, Mat. Sbornik T.8 (50) N.1 (1940), 149-167.

[117] R.Weinstock, Continuous differentiability, Amer. Math. Monthly 64 (1957), 492.

[118] W.H.Young and G.C.Young, On the reduction of sets of intervals, Proc. London Math. Soc. (2) 14 (1915), 111-130.