Real Analysis Exchange
Vol. 29(2), 2003/2004, pp. 905–920

Solomon Leader, Mathematics Department, Rutgers University, 110 Frelinghuysen Rd., Piscataway, NJ 08854-8019, U.S.A.

CHANGE OF VARIABLE IN KURZWEIL-HENSTOCK STIELTJES INTEGRALS

Abstract

We present conditions under which one may substitute the identity function for h in Kurzweil-Henstock integrals of the form $\int (f \circ h)\, d(g \circ h)$ reducing them to equivalent integrals of the form $\int f\, dg$. Our study requires that we also consider reduction of $\int (f \circ g)|d(g \circ h)|$ to $\int Nf|dg|$ where N is the Banach indicatrix of h.

1 Introduction

Our primary objective here was to find conditions on a continuous function h on $K = [a, b]$ and on functions f, g on $h(K)$ that would ensure the validity of the integration formula

$$\int_a^x (f \circ h)(t)\, d(g \circ h)(t) = \int_{h(a)}^{h(x)} f(y)\, dg(y) \tag{1.1}$$

for all x in K. This objective is attained in Theorems 1 and 3 where both conditional and absolute integrability in (1.1) are treated. Theorem 1 treats the case of nondecreasing h while Theorem 3 allows h to be of bounded variation. In order to get (1.1) under the latter condition we had to determine the validity of

$$\int_K (f \circ h)\, |d(g \circ h)| = \int_{h(K)} Nf\, |dg| \tag{1.2}$$

where $N(y)$ is the number of points x in K such that $h(x) = y$. The function N is the indicatrix of h introduced by Banach [1] who proved that $\int_{-\infty}^{\infty} N(y)\, dy$

Key Words: Kurzweil-Henstock Stieltjes integrals, Banach indicatrix, bounded variation, endpoint-tagged partition
Mathematical Reviews subject classification: 26A39, 26A42, 26A45
Received by the editors July 1, 2003
Communicated by: Peter Bullen

equals the total variation of the continuous function h on K. This is (1.2) for $f = 1$ and g the identity.

In Theorem 3 we get (1.2) from a transform in Theorem 2. This transform was introduced in [3], [4] and subsequently treated in [2, § 9.2]. Our treatment here is more general in that the previous treatments apply to (1.1) and (1.2) only for g the identity function.

All integrals here are Kurzweil-Henstock integrals defined by gauge-directed limits of approximating sums on endpoint-tagged partitions. For a detailed exposition of this integration process and the concept of differential based upon it we direct the reader to [2]. But for easy access we offer the following introductory remarks.

A *cell* here is any closed, bounded, nondegenerate interval $H = [r, s]$ in $\mathbb{R}$. So $-\infty < r < s < \infty$. (Actually, all our results remain valid if either $r = -\infty$ or $s = \infty$, or both.) The *interior* H° of H is the open interval (r, s). The *boundary* $H^\bullet$ consists of the two endpoints r, s. Two cells *overlap* if their intersection is a cell. They *abut* if their intersection is a common endpoint. A *figure* is a finite union of cells. (In terms of its components a figure is a finite union of disjoint cells.) A *partition* of a cell (or figure) K is a finite set $\mathbb{K}$ of cells whose union is K with no two members of $\mathbb{K}$ overlapping. A *tagged cell* is a pair (I, t) where I is a cell and t is one of its endpoints. A *division* $\mathcal{K}$ of K is a finite set of tagged cells such that the cells form a partition of K. A *gauge* on K is a function δ on K such that $\delta(t) > 0$ for all t in K. A tagged cell (I, t) is δ*-fine* if the length of I is less than $\delta(t)$. A δ*-division* is a division whose members are δ-fine. A *summant* S on K is any function $S(I, t)$ on the set of all tagged cells in K. Given a summant S on K each division $\mathcal{K}$ of K yields a real number

$$(\Sigma S)(\mathcal{K}) = \sum_{(I,t)\in\mathcal{K}} S(I, t)\,. \tag{1.3}$$

For each gauge δ on K we let $S^{(\delta)}(K)$ be the supremum, and $S_{(\delta)}(K)$ the infimum, of (1.3) over all δ-divisions $\mathcal{K}$ of K. The *upper* and *lower integrals* of S are defined respectively by

$$\overline{\int_K} S = \inf_\delta S^{(\delta)}(K)\,, \qquad \underline{\int_K} S = \sup_\delta S_{(\delta)}(K) \tag{1.4}$$

taken over all gauges δ on K. The *integral* $\int_K S$ is the common value of (1.4) in $[-\infty, \infty]$ whenever the upper and lower integrals are equal. S is *integrable* if its integral exists and is finite. S is *absolutely integrable* if both S and $|S|$

are integrable. S is *conditionally integrable* if S is integrable but $|S|$ is not. For such S we have $\int_K |S| = \infty$.

An equivalence relation $S_1 \sim S_2$ for summants on K is defined by $\int_K |S_1 - S_2| = 0$. Every summant S on K represents a *differential* σ on K defined to be the set $[S]$ of all summants equivalent to S on K. The differentials on K form a Riesz space with $|\sigma| = \left[|S|\right]$, $\sigma^+ = \left[S^+\right]$, and $\sigma^- = \left[S^-\right]$ for $\sigma = [S]$. The definitions $\overline{\int}_K \sigma = \overline{\int}_K S$ and $\underline{\int}_K \sigma = \underline{\int}_K S$ are effective for $\sigma = [S]$. So the concepts of integral, integrability, absolute and conditional integrability extend to differentials. A differential $\sigma = 0$ if and only if $\int_K |\sigma| = 0$, and $\sigma \geq 0$ if and only if $\sigma^- = 0$.

Every function g on K induces a differential $dg = [\Delta g]$ which is integrable with

$$\int_K dg = \Delta g(K) \tag{1.5}$$

since Δg is additive on abutting cells. Conversely, every integrable differential σ on K equals dg for $g(x) = \int_{[a,x]} \sigma$ with $g(a) = 0$. The integral $\int_K |dg|$ exists for every function g on K and defines the *total variation* (finite or infinite) of g.

For every function f on K and differential $\sigma = [S]$ on K the differential $f\sigma = [fS]$ is effectively defined with $(fS)(I,t) = f(t)S(I,t)$. A subset E of K is *σ-null* if $1_E\,\sigma = 0$ where 1_E is the indicator of E. A function g on K is continuous at a point p if and only if p is dg-null. A condition holds *σ-everywhere* on K if it holds on the complement of a σ-null set in K. The definition of the product $f\sigma$ extends to any function f that is defined and finite σ-everywhere on K.

2 Change of Variable for Continuous, Nondecreasing h

The conclusions of the following theorem also hold on any cell contained in K.

Theorem 1. *Let h be continuous on a cell K with $dh \geq 0$. Let g be continuous on $h(K)$. Then for every function f on $h(K)$*

$$\overline{\int_K} (f \circ h)\, d(g \circ h) = \overline{\int_{h(K)}} f\, dg \tag{2.1}$$

and

$$\overline{\int_K} (f \circ h) \left|d(g \circ h)\right| = \overline{\int_{h(K)}} f \left|dg\right| \tag{2.2}$$

which also hold for the lower integrals. If either of the integrals $\int_K (f\circ h)\, d(g\circ h)$ *or* $\int_{h(K)} f\, dg$ *exists then both exist and they are equal. The same holds for the integrals* $\int_K (f\circ h)\, \big|d(g\circ h)\big|$ *and* $\int_{h(K)} f|dg|$. *Absolute integrability of either* $(f\circ h)\, d(g\circ h)$ *on* K *or* $f\, dg$ *on* $h(K)$ *implies absolute integrability of both.*

PROOF. We dismiss the case of constant h for which all the integrals in (2.1) and (2.2) vanish. So $h(K)$ is a cell L rather than a point.

For any tagged cell (I,t) in K the image $h(I)$ is either a point or a cell. In the latter case $\big(h(I),\, h(t)\big)$ is a tagged cell with $h(t)$ the left or right endpoint according to whether t is the left or right endpoint of I. Whatever the case may be we have

$$(f\circ h)(t)\, \Delta(g\circ h)(I) = f\big(h(t)\big)\, \Delta g\big(h(I)\big) \tag{2.3}$$

where both sides vanish if $h(I)$ is a point.

Let β be an arbitrary gauge on L. Continuity of h yields a gauge α on K such that

$$\begin{aligned} &\big(h(I), h(t)\big) \text{ is a } \beta\text{-fine tagged cell in } L \text{ if } (I,t) \\ &\text{is an } \alpha\text{-fine tagged cell in } K \text{ with } \Delta h(I) > 0 \end{aligned} \tag{2.4}$$

If I, J are nonoverlapping cells in K with both $\Delta h(I) > 0$ and $\Delta h(J) > 0$ then the cells $h(I)$, $h(J)$ do not overlap. Thus each α-division $\mathcal{K}$ of K is carried by h into a β-division $\mathcal{L}$ of L (discarding those (I,t) in $\mathcal{K}$ with $\Delta h(I) = 0$) such that

$$\Big(\Sigma(f\circ h)\, \Delta(g\circ h)\Big)(\mathcal{K}) = \Big(\Sigma f\, \Delta g\Big)(\mathcal{L}) \tag{2.5}$$

by (2.3) and (2.4). For the supremum of the right-hand side of (2.5) over all β-divisions $\mathcal{L}$ of L we get

$$\Big(\Sigma(f\circ h)\, \Delta(g\circ h)\Big)(\mathcal{K}) \le \big(f\, \Delta g\big)^{(\beta)}(L) \tag{2.6}$$

for all α-divisions $\mathcal{K}$ of K. Taking the supremum of the left-hand side of (2.6) over all such $\mathcal{K}$ we get

$$\Big((f\circ h)\, \Delta(g\circ h)\Big)^{(\alpha)}(K) \le \big(f\, \Delta g\big)^{(\beta)}(L)\,. \tag{2.7}$$

By (1.4) and (2.7)

$$\overline{\int_K} (f\circ h)\, d(g\circ h) \le (f\, \Delta g)^{(\beta)}(L) \tag{2.8}$$

for all gauges β on L. Hence (1.4) gives

$$\overline{\int_K}(f\circ h)\,d(g\circ h)\le\overline{\int_L}f\,dg\,. \tag{2.9}$$

To reverse the inequality in (2.9) let α be an arbitrary gauge on K. Choose a gauge β on L as follows. Given u in L take an open interval W in $\mathbb{R}$ about u small enough so that both $q-p<\alpha(q)$ and $s-r<\alpha(r)$ for $[q,r]=h^{-1}(u)$ with $q\le r$ and the cell $[p,s]=h^{-1}(\overline{W})$. Such W exist since h is continuous and nondecreasing. Take $\beta(u)$ small enough so that if J is any cell in L with u as an endpoint and length less than $\beta(u)$ then J is contained in W.

Given a β-division $\mathcal{L}$ of L each member (J,u) of $\mathcal{L}$ yields an α-fine tagged cell (I,t) in K where $I=\overline{h^{-1}(J^\circ)}$ and t is the left (right) endpoint of I if u is the left (respectively, right) endpoint of J. In this way $\mathcal{L}$ yields an α-division $\mathcal{H}$ of a figure H in K. Since h is constant on each component of the complementary figure $H'=\overline{(K\setminus H)^\circ}$ the differential $d(g\circ h)=0$ on H'. Take any α-division $\mathcal{H}'$ of H' (with $\mathcal{H}'$ empty if $H=K$) and let $\mathcal{K}=\mathcal{H}\cup\mathcal{H}'$. Then $\mathcal{K}$ is an α-division of K such that (2.5) holds for the given β-division $\mathcal{L}$ of L. By an analogous argument to that which led from (2.5) to (2.7) we get the reversal of the inequality in (2.9). So equality holds in (2.9) which proves (2.1).

The proof of (2.1) gives a proof of (2.2) if we replace $\Delta(g\circ h)$, Δg, $d(g\circ h)$ and dg by their absolute values. To get (2.1) and (2.2) for the lower integrals apply them to $-f$.

The rest of the theorem follows from these results since they also apply to $|f|$. □

3 The Transform Theorem

Our next theorem is needed to treat change of variable for continuous h of bounded variation. Its proof requires two lemmas.

Lemma 1. *Let h be continuous on a cell I. Given a cell J in $h(I)$ there exists a cell H in I such that*

$$h(H^\bullet)=J^\bullet \tag{3.1}$$

and

$$h(H^\circ)=J^\circ\,. \tag{3.2}$$

If J_1, J_2 are nonoverlapping cells in $h(I)$ and H_1, H_2 are cells in I such that (3.2) holds for H_1, J_1 and H_2, J_2 then H_1, H_2 are nonoverlapping. For every *function g on $h(I)$*

$$\left|\Delta(g\circ h)(I)\right| \le \int_{h(I)} |dg| \le \int_I \left|d(g\circ h)\right|. \tag{3.3}$$

So g is of bounded variation on $h(I)$ if $g\circ h$ is of bounded variation on I.

PROOF. Since the endpoints p, q of J are distinct points in $h(I)$ the nonempty sets $h^{-1}(p)$ and $h^{-1}(q)$ are disjoint and compact. So the minimum distance $|x-y|$ between them is attained by some x in $h^{-1}(p)$ and y in $h^{-1}(q)$. Let H be the cell with endpoints x, y. Then (3.1) clearly holds. Since H has minimum length for all cells linking $h^{-1}(p)$ with $h^{-1}(q)$ its interior H° contains no point of these two sets while H links them. So (3.2) follows from the intermediate value theorem for the continuous function h.

Given (3.2) for the pairs H_1, J_1 and H_2, J_2 let $H = H_1 \cap H_2$. Then $H^\circ = H_1^\circ \cap H_2^\circ$ so $h(H^\circ) \subseteq J_1^\circ \cap J_2^\circ$ by (3.2). Hence if $J_1^\circ \cap J_2^\circ$ is empty then so is H°. That is, if J_1, J_2 do not overlap then neither do H_1, H_2.

Given a partition $\mathbb{J} = \{J_1, \ldots, J_n\}$ of $h(I)$ choose for each J_i a cell H_i in I such that (3.1) and (3.2) hold for H_i, J_i. Then (3.1) implies that for every function g on $h(I)$

$$\left|\Delta g(J_i)\right| = \left|\Delta(g\circ h)(H_i)\right| \tag{3.4}$$

for $i = 1, \ldots, n$. Since no two members of $\mathbb{J}$ overlap, the same holds for $\mathbb{H} = \{H_1, \ldots, H_n\}$. So $\mathbb{H}$ is a partition of the figure $H_1 \cup \ldots \cup H_n$ in I. Summation of (3.4) gives

$$\big(\Sigma|\Delta g|\big)(\mathbb{J}) = \big(\Sigma|g\circ h|\big)(\mathbb{H}) \le \int_I \left|d(g\circ h)\right|. \tag{3.5}$$

Since $\mathbb{J}$ is an arbitrary partition of $h(I)$ the second inequality in (3.3) follows from (3.5). The first inequality in (3.3) is obvious. □

Note that (3.2) implies (3.1).

Lemma 2. *Let h be continuous on a cell K. Let g be continuous on $h(K)$ with $g\circ h$ of bounded variation on K. Then*

$$\overline{\int_{h(K)}} 1_{h(D)}|dg| \le \overline{\int_K} 1_D\left|d(g\circ h)\right| \tag{3.6}$$

for every subset D of K. So $h(D)$ is dg-null for every $d(g\circ h)$-null subset D of K.

PROOF. Let U be any relatively open subset of K containing D. Then $h(D) \subseteq h(U) = \cup_{i\in M} h(I_i)$ where the I_i's are the components of U indexed by a finite or countably infinite set M. So

$$\begin{aligned}\overline{\int_{h(K)}} 1_{h(D)}\,|dg| &\le \overline{\int_{h(K)}} 1_{h(U)}\,|dg| \\ &\le \sum_{i\in M}\int_{h(I_i)} |dg| \\ &\le \sum_{i\in M}\int_{I_i} \big|d(g\circ h)\big| \\ &= \int_K 1_U\,\big|d(g\circ h)\big|\end{aligned}$$

by Theorem 2 (§2.7) and Theorem 9 (§4.3) in [2], and (3.3) in Lemma 1. Thus

$$\overline{\int_{h(K)}} 1_{h(D)}\,|dg| \le \int_K 1_U\,\big|d(g\circ h)\big| \tag{3.7}$$

for every open set U containing D. The infimum of the right-hand side of (3.7) over all such U equals the upper integral on the right-hand side of (3.6) by Theorem 3 (§5.1) in [2]. So (3.6) holds. □

We can now confront the transform theorem. The transform (3.8) is determined by the continuous function h of bounded variation on K. It takes ϕ on K to $\hat{\phi}$ on the set Y of all y in $\mathbb{R}$ for which the set $h^{-1}(y)$ is finite. It acts as a positive linear transformation taking the Lebesgue space $L^1\big(|d(g\circ h)|\big)$ on K into $L^1\big(|dg|\big)$ on $h(K)$ for every continuous g on $h(K)$ such that $g\circ h$ (hence also g) is of bounded variation.

Theorem 2. *Let h be a continuous function of bounded variation on a cell K. Let g be continuous on $h(K)$ with $g\circ h$ of bounded variation on K. Given a function ϕ on K such that $\phi d(g\circ h)$ is absolutely integrable on K define the function $\hat{\phi}$ dg-everywhere on $h(K)$ by*

$$\hat{\phi}(y) = \sum_{x\in h^{-1}(y)} \phi(x) \tag{3.8}$$

which is valid since the set $h^{-1}(y)$ is finite for dg-all y in $h(K)$. Then $\hat{\phi}dg$ is absolutely integrable on $h(K)$ and

$$\int_K \phi\big|d(g\circ h)\big| = \int_{h(K)} \hat{\phi}|dg|\,. \tag{3.9}$$

PROOF. Since $g \circ h$ is a continuous function of bounded variation there exists a continuous, nondecreasing function V on K such that

$$dV = \big|d(g \circ h)\big| \quad \text{on } K. \tag{3.10}$$

Moreover, ϕdV is absolutely integrable since both $d(g \circ h)$ and $\phi\, d(g \circ h)$ are absolutely integrable, and $d(g \circ h)$ has a Borel measurable Hahn decomposition of K. So there is a continuous function F of bounded variation on K such that

$$dF = \phi\, dV\,. \tag{3.11}$$

We shall treat three cases of increasing generality beginning with the condition

$$0 \le \phi \le k \text{ for some positive integer } k. \tag{3.12}$$

Define the summant S, which is independent of the tags, by

$$S(I) = \begin{cases} \frac{\Delta F}{\Delta V}(I) & \text{if } \Delta V(I) > 0, \\ 0 & \text{if } \Delta V(I) = 0 \end{cases} \tag{3.13}$$

for all cells I in K. Also independent of the tags is the summant T defined by

$$T(I) = \int_{h(I)} |dg| \tag{3.14}$$

for all cells I in K. By (3.3) in Lemma 1 and (3.10)

$$\big|\Delta(g \circ h)\big| \le T \le \Delta V\,. \tag{3.15}$$

By (3.11) and (3.12) $\quad 0 \le dF \le k dV$ whose integrals give

$$0 \le \Delta F \le k\Delta V\,. \tag{3.16}$$

From (3.16) and (3.13) we get

$$0 \le S \le k \tag{3.17}$$

and

$$S\Delta V = \Delta F\,. \tag{3.18}$$

From (3.18), (3.17) and (3.15) we get

$$\big|\Delta F - ST\big| = S\Delta V - ST = S(\Delta V - T) \le k\big(\Delta V - |\Delta(g \circ h)|\big)\,.$$

That is,

$$\big|\Delta F - ST\big| \le k\big(\Delta V - |\Delta(g \circ h)|\big)\,. \tag{3.19}$$

Take a sequence of partitions $\mathbb{K}_j$ of K such that for $j = 1, 2, \ldots$

$$\mathbb{K}_{j+1} \text{ refines } \mathbb{K}_j \tag{3.20}$$

and

$$\text{each member of } \mathbb{K}_j \text{ has length less than } \frac{1}{j}\,. \tag{3.21}$$

Since $g \circ h$ is continuous (3.20), (3.21) and (3.10) imply that as $j \nearrow \infty$

$$\big(\Sigma|\Delta(g \circ h)|\big)(\mathbb{K}_j) \nearrow \Delta V(K)\,. \tag{3.22}$$

From (3.22) and (3.19) we have the convergence

$$\big(\Sigma ST\big)(\mathbb{K}_j) \to \Delta F(K) \tag{3.23}$$

since ΔV and ΔF are additive on abutting cells.

For each j define ψ_j on $h(K)$ in terms of (3.13) by

$$\psi_j(y) = \sum_{I \in \mathbb{K}_j} S(I) 1_{h(I)}(y)\,. \tag{3.24}$$

So ψ_j is a linear combination of indicators of intervals. By Lemma 1 g is of bounded variation. Thus $\psi_j\,|dg|$ is integrable on $h(K)$ and

$$\int_{h(K)} \psi_j\,|dg| = \big(\Sigma ST\big)(\mathbb{K}_j) \tag{3.25}$$

by (3.24), (3.14), and continuity of g. By (3.23) and (3.25)

$$\int_{h(K)} \psi_j\,|dg| \to \Delta F(K) \tag{3.26}$$

as $j \to \infty$.

Let $N_j(y)$ be the number of members I of $\mathbb{K}_j$ such that y belongs to $h(I)$. That is,

$$N_j(y) = \sum_{I \in \mathbb{K}_j} 1_{h(I)}(y)\,. \tag{3.27}$$

By continuity of g, the second inequality in (3.15), and the definition (3.14) of T we can integrate (3.27) against $|dg|$ to get

$$\int_{h(k)} N_j |dg| = (\Sigma T)(\mathbb{K}_j) \le \Delta V(K)\,. \tag{3.28}$$

By (3.24), (3.17), and (3.27)

$$0 \le \psi_j \le kN_j\,. \tag{3.29}$$

Let E be the countable set of all endpoints of members of the $\mathbb{K}_j$'s. Consider any y in $A = h(K)\backslash h(E)$. For all j (3.20) implies

$$0 \le N_j(y) \le N_{j+1}(y) \le N(y) \quad \text{and} \quad N_j(y) < \infty \tag{3.30}$$

where N is the Banach indicatrix of h on K. Since g is continuous the countable set $h(E)$ is dg-null. That is, $1_A\, dg = dg$. So (3.30) holds at dg-all y in $h(K)$.

Given y, in A consider any finite set $x_1 < \ldots < x_m$ of points in the nonempty set $h^{-1}(y)$. For j large enough so that $\frac{1}{j} < x_{i+1} - x_i$ for $i = 1, \ldots, m-1$ (3.21) implies that each member I of $\mathbb{K}_j$ contains at most one x_i. So

$$m \le N_j(y) \;\text{ ultimately as } j \to \infty\,. \tag{3.31}$$

Thus if $N(y) < \infty$ then for $m = N(y)$ the inequalities (3.30) and (3.31) imply $N_j(y) = N(y)$ ultimately as $j \to \infty$. If $N(y) = \infty$ then (3.31) holds for arbitrarily large m. That is, $N_j(y) \nearrow \infty$ in (3.30). In summary, for all y in A

$$N_j(y) \nearrow N(y) \;\text{ as }\; j \nearrow \infty\,. \tag{3.32}$$

Since $1_A\, dg = dg$ (3.32) holds for dg-all y in $h(K)$.

Let B be the set of y in A such that $N(y) < \infty$. For y in B and j large enough so that each member I of $\mathbb{K}_j$ contains at most one point of $h^{-1}(y) = \{x_1, \ldots, x_m\}$ the definition (3.24) of ψ_j takes the form

$$\psi_j(y) = \sum_{i=1}^{m} S(I_{j,i}) \tag{3.33}$$

where $I_{j,i}$ is the unique member I of $\mathbb{K}_j$ such that I° contains the point x_i of $h^{-1}(y)$.

Let C consist of all y in B such that

$$\frac{dF}{dV}(x) = \phi(x) \tag{3.34}$$

for all x in $h^{-1}(y)$. At such a point x $\Delta V(I) > 0$ for every cell I in K that contains x, and

$$\frac{\Delta F}{\Delta V} \to \phi(x)$$

as the length of I goes to 0 with x in I. For y in C (3.13) and (3.21) imply that for each term of the sum in (3.33)

$$S(I_{j,i}) \to \phi(x_i) \quad \text{as} \quad j \to \infty\,. \tag{3.35}$$

By (3.35), (3.33), and the definition (3.8) of $\hat{\phi}$

$$\psi_j(y) \to \hat{\phi}(y) \quad \text{as} \quad j \to \infty \tag{3.36}$$

for all y in C.

From (3.28) and (3.32) the monotone convergence theorem gives

$$\int_{h(K)} N_j \, |dg| \quad \nearrow \quad \int_{h(K)} N \, |dg| \le \Delta V(K) < \infty \tag{3.37}$$

as $j \nearrow \infty$. This implies that $N < \infty$ dg-everywhere on $h(K)$. That is, $1_B \, dg = dg$ since $1_A \, dg = dg$.

Theorem 2 (§6.3) in [2] applied to (3.11) gives (3.34) at all x in $K \setminus D$ where D is dV-null, that is, $d(g \circ h)$-null by (3.10). So $h(D)$ is dg-null by Lemma 2. Hence $1_C \, dg = dg$ since $1_B \, dg = dg$.

On C we have the convergence (3.36) with the ψ_j's bounded by kN according to (3.29) and (3.32). Thus, since $kN \, |dg|$ is integrable by (3.37), the dominated convergence theorem in [2, Theorem 4, §2.8] gives

$$\int_{h(K)} \psi_j \, |dg| \to \int_{h(K)} \hat{\phi} \, |dg|\,. \tag{3.38}$$

Comparison of (3.38) with (3.26) gives

$$\Delta F(K) = \int_{h(K)} \hat{\phi} \, |dg|\,. \tag{3.39}$$

By (3.10) and (3.11) equation (3.39) is just (3.9). So (3.9) holds for the case (3.12).

We can now prove (3.9) for the case

$$0 \le \phi(x) < \infty \quad \text{for all } x \text{ in } K. \tag{3.40}$$

Let $\phi_k = \phi \wedge k$ for each positive integer k. Since $\phi \, dV$ is integrable so is $\phi_k \, dV$. Since each ϕ_k satisfies (3.12) we have (3.9) in the form

$$\int_K \phi_k \, dV = \int_{h(K)} \widehat{\phi_k} \, |dg| \tag{3.41}$$

for ϕ_k by (3.10) Since $\phi_k \nearrow \phi$ as $k \nearrow \infty$ the monotone convergence theorem in [2, Theorem 3, §2.7] gives

$$\int_K \phi_k \, dV \quad \nearrow \quad \int_K \phi \, dV \,. \tag{3.42}$$

Similarly $\widehat{\phi_k} \nearrow \hat{\phi}$ dg-everywhere under (3.8). So

$$\int_{h(K)} \widehat{\phi_k} \, |dg| \quad \nearrow \quad \int_{h(K)} \hat{\phi} \, |dg| \,. \tag{3.43}$$

The last three displays give (3.9) for the case (3.40).

For general case, $-\infty < \phi(x) < \infty$ for all x in K, we apply the case (3.40) to ϕ^+ and ϕ^-, noting that $\hat{\phi} = \widehat{\phi^+} - \widehat{\phi^-}$, to get

$$\int_K \phi \, dV = \int_K \phi^+ \, dV - \int_K \phi^- \, dV$$

$$= \int_{h(k)} \widehat{\phi^+} \, |dg| - \int_{h(K)} \widehat{\phi^-} \, |dg| = \int_{h(K)} \hat{\phi} \, |dg| \,. \qquad \square$$

For the case $\phi = 1$ (3.8) gives $\hat{1} = N$, the Banach indicatrix of h. So (3.9) then gives Lindner's indicatrix integral for the total variation of $g \circ h$,

$$\int_K |d(g \circ h)| = \int_{h(K)} N \, |dg| \,. \tag{3.44}$$

Lindner [5] proved this for all continuous h, g with both integrals equal to ∞ when $g \circ h$ is of unbounded variation. Banach's indicatrix theorem [1] is the special case of (3.44) for g the identity function,

$$\int_K |dh| = \int_{-\infty}^{\infty} N(y) \, dy \tag{3.45}$$

for any continuous h on K. For g the identity function, Theorem 2 reduces to Theorem 1 (§9.2) in [2].

4 Change of Variable for Continuous h of Bounded Variation

Theorem 3. *Let h be a continuous function of bounded variation on a cell $K = [a, b]$ with Banach indicatrix N. Let g be continuous on $h(K)$ with $g \circ h$ of bounded variation on K. Let f be a function on $h(K)$.*

(i) *If* $(f \circ h)\, d(g \circ h)$ *is absolutely integrable on* K *then*

$$\int_K (f \circ h)\, |d(g \circ h)| = \int_{h(K)} Nf\, |dg| \tag{4.1}$$

and there exists a continuous function F *of bounded variation on* $h(K)$ *such that*

$$dF = f\, dg \quad on \quad h(K) \tag{4.2}$$

and

$$d(F \circ h) = (f \circ h)\, d(g \circ h) \quad on \quad K\,. \tag{4.3}$$

For all x *in* K

$$\int_{h(a)}^{h(x)} f(y)\, dg(y) = \int_a^x (f \circ h)(t)\, d(g \circ h)(t)\,. \tag{4.4}$$

(ii) *If the integrability condition* (4.2) *holds for some* F *on* $h(K)$ *such that* $h^{-1}(E)$ *is* $d(F \circ h)$*-null for every* dF*-null subset* E *of* $h(K)$ *then* (4.3) *and* (4.4) *hold.*

PROOF. To prove (4.1) in (i) apply Theorem 2 with $\phi = f \circ h$. Since ϕ has constant value $f(y)$ on $h^{-1}(y)$ (3.8) gives $\hat{\phi} = Nf$. So (3.9) gives (4.1).

To complete the proof of (i) we collect some results from [2] that we shall need.

Given g continuous and of bounded variation on a cell L we call a point x in L Δg*-positive* (Δg*-negative*) whenever $\Delta g(I) > 0$ (respectively, $\Delta g(I) < 0$) for all sufficiently small cells I in L containing x. For P the set of all Δg-positive points and Q the set of all Δg-negative points in L we have $1_P\, dg = (dg)^+ = 1_P\, |dg|$ and $-1_Q\, dg = (dg)^- = 1_Q\, |dg|$. So

$$1_{P+Q}\, dg = dg\,, \quad (1_P - 1_Q)\, dg = |dg|, \quad \text{and} \quad (1_P - 1_Q)\, |dg| = dg\,. \tag{4.5}$$

(See [2, Theorem 3, §6.3].) We shall apply these results to g, h, and $g \circ h$. So let A be the set of all Δh-positive, and B the set of all Δh-negative, points in K. Let C be the set of all $\Delta(g \circ h)$-positive, and D the set of all $\Delta(g \circ h)$-negative, points in K.

For each subset E of K let N_E be the Banach indicatrix of h on E. That is $N_E(y)$ is the number of points x in E such that $h(x) = y$.

Consider any point x in $h^{-1}(P)$. That is, $h(x)$ is Δg-positive. Such a point x is $\Delta(g \circ h)$-positive if and only if it is Δh-positive. That is, $C \cap h^{-1}(P) = A \cap h^{-1}(P)$. So $1_P N_C = N_{C \cap h^{-1}(P)} = N_{A \cap h^{-1}(P)} = 1_P N_A$. That is,

$$1_P N_C = 1_P N_A\,. \tag{4.6}$$

Now we consider any x in $h^{-1}(Q)$. That is, $h(x)$ is Δg-negative. Such a point x is $\Delta(g \circ h)$-positive if and only if it is Δh-negative. That is, $C \cap h^{-1}(Q) = B \cap h^{-1}(Q)$ which gives

$$1_Q N_C = 1_Q N_B \,. \tag{4.7}$$

Similar arguments give

$$1_P N_D = 1_P N_B \tag{4.8}$$

and

$$1_Q N_D = 1_Q N_A \,. \tag{4.9}$$

The sum of (4.6) and (4.7) minus the sum of (4.8) and (4.9) gives

$$1_{P+Q}(N_C - N_D) = (1_P - 1_Q)(N_A - N_B)\,. \tag{4.10}$$

By (4.5) and (4.10)

$$(N_C - N_D)|dg| = (N_A - N_B)\, dg\,. \tag{4.11}$$

We also have

$$w\,|d(g \circ h)| = d(g \circ h) \quad \text{for} \quad w = 1_C - 1_D \quad \text{on} \quad K\,. \tag{4.12}$$

For $\phi = wf \circ h$ in (3.8) we get $\hat{\phi} = \hat{w} f = (\widehat{1_C} - \widehat{1_D})f = (N_C - N_D)f$ which together with (4.12) and (4.11) converts (3.9) into

$$\int_K (f \circ h)\, d(g \circ h) = \int_{h(K)} (N_A - N_B) f\, dg\,. \tag{4.13}$$

To get (4.4) from (4.13) for the case $x = b$ we need only prove

$$\int_{h(K)} (N_A - N_B) f\, dg = \int_{h(a_)}^{h(b)} f(y)\, dg(y) \tag{4.14}$$

under the convention $\int_p^q = -\int_q^p$.

Since $g \circ h$ is of bounded variation Lindner's indicatrix integral in (3.44) is finite. So $N < \infty$ dg-everywhere and the same holds for N_A and N_B.

Consider any y distinct from $h(a)$ and $h(b)$ with $N(y) < \infty$. As t advances continuously from a to b the point $h(t)$ passes through y in the positive direction $N_A(y)$ times and in the negative direction $N_B(y)$ times. Since h is continuous these positive and negative transits must alternate. Let J be the closed interval with endpoints $h(a)$ and $h(b)$. For y in J° the value of $N_A - N_B$ at y is 1 if $h(a) < h(b)$ and -1 if $h(b) < h(a)$. For y in $\mathbb{R} \setminus J$ the value of

$N_A - N_B$ at y is 0 in both of the cases $h(a) \le h(b)$ and $h(b) \le h(a)$. In summary, since the endpoints of J are dg-null for the continuous function g,

$$N_A - N_B = 1_J \operatorname{sgn}\bigl(h(b) - h(a)\bigr) \tag{4.15}$$

dg-everywhere on $h(K)$. Integration of (4.15) against $f\,dg$ gives (4.14). Since our proof is valid for $[a, x]$ in place of K we get (4.4) for all x in K.

Since $(f \circ h)\,d(g \circ h)$ is absolutely integrable on K, $f\,dg$ is absolutely integrable on $h(K)$ by (4.1) and (4.4). Hence (4.2) holds for some continuous F of bounded variation on $h(K)$. So the left-hand side of (4.4) equals $\Delta(F \circ h)[a, x]$ which under (4.4) gives (4.3). So we have the proof of (i).

Given the hypothesis in (ii) we invoke Theorem 2 (§6.3) in [2] to get $h(K) = D + E$ with $\frac{dF}{dg}(y) = f(y)$ for all y in D, and E dg-null. So for all y in D

$$F(z) - F(y) - f(y)\bigl(g(z) - g(y)\bigr) = o\bigl(|g(z) - g(y)|\bigr) \tag{4.16}$$

as $z \to y$ in $h(K)$. For tagged cells (I, t) in K with t in $h^{-1}(D)$ we can apply (4.16) with $y = h(t)$ and $z = h(s)$ where s is the endpoint of I opposite t. This gives

$$(1_D \circ h)(t)\bigl|\Delta(F \circ h)(I) - (f \circ h)(t)\Delta(g \circ h)(I)\bigr| = o\bigl(|\Delta(g \circ h)(I)|\bigr) \tag{4.17}$$

as the length of I goes to 0 with t an endpoint of I. This convergence is just $s \to t$ which by the continuity of $g \circ h$ implies $\Delta(g \circ h)(I) \to 0$. So (4.17) implies

$$1_{h^{-1}(D)}\,\bigl|d(F \circ h) - (f \circ h)\,d(g \circ h)\bigr| = 0 \tag{4.18}$$

since $g \circ h$ is of bounded variation and $1_D \circ h = 1_{h^{-1}(D)}$. By hypothesis $h^{-1}(E)$ is $d(F \circ h)$-null since E is dg-null. So

$$1_{h^{-1}(D)}\,d(F \circ h) = d(F \circ h)\,. \tag{4.19}$$

By Theorem 3 (§5.1) in [2] there exists a dg-null Borel set A containing the dg-null set E. Since h is continuous, $h^{-1}(A)$ is a Borel set. So $1_{h^{-1}(A)}\,d(g \circ h)$ is absolutely integrable since $g \circ h$ is of bounded variation. Thus Theorem 2 applies with $\phi = 1_{h^{-1}(A)}$ for which $\hat{\phi} = N1_A$. So by (3.9) we conclude that $h^{-1}(A)$, hence also $h^{-1}(E)$, is $d(g \circ h)$-null. Thus

$$1_{h^{-1}(D)}\,d(g \circ h) = d(g \circ h)\,. \tag{4.20}$$

The last three displays give (4.3). Integration of (4.2) and (4.3) gives (4.4) completing the proof of (ii). $\square$

We remark that while (ii) demands integrability of (4.2), (i) demands absolute integrability of (4.3). The case $f = 1$ in (i) reduces (4.1) to (3.44). More generally, for $f = 1_E$ with E a Borel set in $h(K)$ we have absolute integrability of both $1_E\, dg$ on $h(K)$ and $1_{h^{-1}(E)}\, d(g \circ h)$ on K, since $h^{-1}(E)$ is a Borel set. So Theorem 3 gives

$$\int_K 1_{h^{-1}(E)} \left|d(g \circ h)\right| = \int_{h(K)} N1_E\, |dg|$$

for (4.1) and

$$\int_{h(a)}^{h(x)} 1_E(y)\, dg(y) = \int_a^x 1_{h^{-1}(E)}(t)\, d(g \circ h)(t)$$

for (4.4) since $1_{h^{-1}(E)} = 1_E \circ h$.

References

[1] S. Banach, *Sur les Lignes Rectifiables et les Surfaces dont l'Aire est Finie*, Fund. Math., **7** (1925), 225–236.

[2] S. Leader, *The Kurzweil-Henstock Integral and its Differentials*, Marcel Dekker, New York, 2001.

[3] S. Leader, *Conversion Formulas for the Lebesgue-Stieltjes Integral*, Real Anal. Exch., **20(2)** (1994/5), 527–535.

[4] S. Leader, *Transforming Lebesgue-Stieltjes Integrals into Lebesgue Integrals*, Real Anal. Exch., **20(2)** (1994/5), 603–616.

[5] S. Lindner, *Generalization of the Banach Indicatrix Theorem*, Real Anal. Exch., **27(2)** (2001/2), 721–724.