Chapter 5

Stochastic processes and p-variation

It will be seen that several classes of stochastic processes have sample functions of bounded p-variation for some p. But we did not find any results on p-variation for general Markov processes, which do have regulated and right–continuous paths under mild conditions.

5.1 Stochastic processes with regulated sample functions

As usual, a *stochastic process* $X = \{X(t): t \geq 0\}$ is a family of random variables $X(t) = X(t, \cdot)$ defined on a complete probability space $(\Omega, \mathcal{F}, \Pr)$. For each $\omega \in \Omega$, the function $X(\cdot) = X(\cdot, \omega)$ is called a *sample function* of X. A stochastic process $X = \{X(t): t \geq 0\}$ has regulated sample functions if $X(\cdot, \omega) \in \mathcal{R}[0, \infty)$ for almost all $\omega \in \Omega$. In this case we write $X \in \mathcal{R}[0, \infty)$.

Let $X \in \mathcal{R}[0, \infty)$ and let $\Omega_X := \{\omega \in \Omega: X(\cdot, \omega) \in \mathcal{R}[0, \infty)\}$. Due to completeness of the underlying probability space, $\Omega_X \in \mathcal{F}$ and $\Pr(\Omega_X) = 1$. For each $t \in [0, \infty)$, let

$$X_+(t) := X_+(t, \omega) := \begin{cases} X(t+, \omega) & \text{if } \omega \in \Omega_X \\ X(t, \omega) & \text{otherwise.} \end{cases}$$

Similarly define $X_-(t)$ for each $t \in (0, \infty)$ and let $X_-(0) := X(0)$. For each $t \in [0, \infty)$, $X_+(t)$ is a random variable because it is the limit for $\omega \in \Omega_X$ of the random variables $X(r_n)$ for r_n rational, $r_n \downarrow t$, and equals $X(t)$ otherwise. Likewise, $X(t-)$ is a random variable. Therefore $X_+ = \{X_+(t): t \geq 0\}$ and $X_- = \{X_-(t): t \geq 0\}$ are stochastic processes on the same probability space as X.

Recall that $\Delta^+ X(t) := X^+(t) - X(t)$ and $\Delta^- X(t) := X(t) - X^-(t)$. For each $t \in [0, \infty)$, let

$$\Omega_d(t) := \{\omega \in \Omega: \text{ either } \Delta^- X(t) \neq 0 \text{ or } \Delta^+ X(t) \neq 0\}.$$

A point $t \in [0, \infty)$ is called a point of *fixed discontinuity* if $\Pr(\Omega_d(t)) > 0$. If t is not a point of fixed discontinuity then

$$\lim_{s \to t} X(s) = X(t) \qquad \text{almost surely.} \tag{5.1}$$

Indeed, for each $t \in (0, \infty)$ and $\omega \in \Omega_X \setminus \Omega_d(t)$, we have

$$\lim_{s \uparrow t} X(s, \omega) = X(t-, \omega) = X(t, \omega) = X(t+, \omega) = \lim_{s \downarrow t} X(s, \omega).$$

Since $\Pr(\Omega_X \setminus \Omega_d(t)) = 1$, (5.1) holds when $t \in (0,\infty)$. The same argument yields (5.1) when $t = 0$. If the stochastic process X has no points of fixed discontinuity then the three processes X_-, X_+ and X are modifications of each other, that is

$$\Pr(\{X(t) = X(t-)\}) = \Pr(\{X(t) = X(t+)\}) = 1$$

for each $t \in [0,\infty)$.

A point $t \in [0,\infty)$ is a point of *stochastic continuity* of X if $X(s) \to X(t)$ in probability as $s \to t$.

Theorem 5.1. *Let $X \in \mathcal{R}[0,\infty)$. The set of points of fixed discontinuity of X is at most countable and coincides with the set of points of stochastic discontinuity of X.*

Proof. Since the limit of a sequence convergent in probability is unique almost surely, one can show that X is not stochastically continuous at some $t \in [0,\infty)$ if t is a point of fixed discontinuity. Also, X is stochastically continuous at t whenever t is not a point of fixed discontinuity. This yields the second part of the claim. The first part follows from Theorem 11.1, Ch. VII, of Doob [18]. □

By the preceding theorem, the set of fixed discontinuities of any stochastic process with regulated sample functions is at most countable. However, almost every sample function of such a process may have a non-fixed discontinuity, e.g. $X_t(\omega) = 1_{t \geq \omega}$, where ω has a uniform distribution in $[0,1]$.

5.2 Martingales

Let $\psi_1(u) := u^2/\log\log(1/u)$ for $0 < u \leq e^{-e}$. Then $\psi_1(u)$ can be defined for $u > e^{-e}$ so that ψ_1 is a strictly increasing continuous convex function from $[0,\infty)$ onto itself. For Brownian motion $B(t) := B_t$ and $0 < T < \infty$, Taylor [106] showed that $v_{\psi_1}(B;[0,T]) < \infty$ almost surely while if $\psi_1(t) = o(\phi(t))$ as $t \downarrow 0$, $v_\phi(B;[0,T]) = +\infty$ almost surely. Thus almost surely

$$v_p(B;[0,T]) \begin{cases} < +\infty & \text{if } p > 2 \\ = +\infty & \text{if } p \leq 2. \end{cases}$$

Let $(\Omega, \mathcal{A}, \Pr)$ be a probability space. Then a *filtration* $\mathbb{F} = \{\mathcal{F}_t \colon t \geq 0\}$ is a family of sub–σ–algebras of $\mathcal{A}$ such that $\mathcal{F}_s \subset \mathcal{F}_t$ for $0 \leq s \leq t$. The filtration $\mathbb{F}$ is *right–continuous* iff $\mathcal{F}_{t+} := \cap_{s>t}\mathcal{F}_s = \mathcal{F}_t$ for all $t \geq 0$. A real–valued stochastic process $X = \{X_t \colon t \geq 0\}$, is *adapted* to $\mathbb{F}$ iff X_t is $\mathcal{F}_t$ measurable for each t. A pair $(X,\mathbb{F})$ is a *martingale* iff $E(X_t|\mathcal{F}_s) = X_s$ almost surely for $0 \leq s \leq t$. Thus X is automatically adapted. It is a *supermartingale* if it is adapted and $E(X_t|\mathcal{F}_s) \leq X_s$ almost surely for $0 \leq s \leq t$. A stochastic process $Y = \{Y_t \colon t \geq 0\}$ is a *modification* of X iff for each $t \geq 0$, $X_t = Y_t$ almost surely.

It is known that if $\mathbb{F}$ is right–continuous, then a supermartingale $(X,\mathbb{F})$ has a modification whose paths $t \mapsto Y_t(\omega)$ are regulated and right–continuous if and only if $t \mapsto EX_t$ is right–continuous (Meyer, [82, p. 95]). For a martingale EX_t is constant, so every martingale $(X,\mathbb{F})$ with $\mathbb{F}$ right–continuous has a right–continuous modification $M = \{M_t \colon t \geq 0\}$. For any martingale $(X,\mathbb{F})$ and almost all $\omega \in \Omega$, the function $t \mapsto X_t(\omega)$ has left and right limits at each point through the countable dense set of rational numbers. Thus since $t \mapsto M_t(\omega)$ is right–continuous, it is automatically also regulated (has left limits).

Theorem 5.2. *Let $p > 2$, $0 < T < \infty$ and let $M = \{M_t\colon 0 \le t \le T\}$ be a right–continuous martingale. Then M has almost all paths of bounded p-variation.*

Proof. Monroe [84] showed that in distribution, $M_t = B_{\tau(t)}$, $0 \le t \le T$, where $B = \{B_u\colon u \ge 0\}$ is a Brownian motion and $\tau(t)$ are stopping times with $\tau(s) \le \tau(t)$, $0 \le s \le t \le T$. It follows that $v_p(M;[0,T]) \le v_p(B;[0,\tau(T)]) < +\infty$ almost surely. □

While a local martingale plus a process locally of bounded variation yields a semimartingale (Doléans–Dade and Meyer [16] and Dellacherie and Meyer [14]) Theorem 5.2 allows one to add a process of bounded p-variation for $p > 2$ and get a sum with the same p-variation property.

If $EM_t = 0$, $0 \le t \le T$, and $EM_T^2 < \infty$ then $E\tau(T) = EM_T^2$ (Monroe, [84, Theorems 5 and 11]). Thus for any set of martingales $M_i = \{M_{i,t}\colon 0 \le t \le T\}$, if $\sup_i EM_{i,T}^2 < +\infty$ then $v_p(M_i;[0,T])$ is bounded in probability uniformly in i.

5.3 Gaussian stochastic processes

Let $X = \{X(t)\colon t \ge 0\}$ be a Gaussian stochastic process, and let $\sigma_X(s,t) := E|X(s) - X(t)|$. For $0 < T < \infty$, let

$$v_p(\sigma_X;[0,T]) := \sup\Big\{\sum_{i=1}^n \sigma_X(t_{i-1},t_i)^p\colon \{t_i\colon i = 0,\dots,n\} \in \mathrm{PP}\,([0,T])\Big\}$$

For a function f on $[a,b]$, define the *index of p-variation* by

$$\upsilon(f;[a,b]) := \begin{cases} \inf\{p > 0\colon v_p(f) < \infty\} & \text{if the set is nonempty} \\ +\infty & \text{otherwise.} \end{cases}$$

Using the p-variation index the result of Jain and Monrad [59, Theorem 3.2] can be stated as follows:

Theorem 5.3. *Let X be a separable mean zero Gaussian process. Then with probability 1*

$$\upsilon(X;[0,T]) = \inf\{p > 0\colon v_p(\sigma_X;[0,T]) < \infty\}.$$

Fractional Brownian motion A fractional Brownian motion $B_H = \{B_H(t)\colon t \ge 0\}$ with the Hurst index $H \in (0,1)$ is a mean zero Gaussian process with the covariance function

$$E[B_H(t)B_H(s)] = \frac{1}{2}\Big[t^{2H} + s^{2H} - |t-s|^{2H}\Big] \qquad \text{for } t,s \ge 0 \tag{5.2}$$

and $B_H(0) = 0$ almost surely. Since the right side of (5.2) is equal to $t \wedge s$ for $H = 1/2$, B_H is a Brownian motion in this case. It follows from (5.2) that the incremental variance σ_H of B_H is given by

$$\sigma_H(u)^2 := E[B_H(t+u) - B_H(t)]^2 = u^{2H} \qquad \text{for } t,u \ge 0,$$

and that B_H has stationary increments. For $0 < T < \infty$ and for each $\epsilon > 0$, let $N_H([0,T];\epsilon)$ be the smallest integer n such that intervals $\{[a_i,b_i]\colon i = 1,\dots,n\}$ with $\sigma_H(b_i - a_i) \le 2\epsilon$ for $i = 1,\dots,n$, cover $[0,T]$. Then there is a finite constant C such that $N_H([0,T];\epsilon) \le C\epsilon^{-1/H}$ for all $\epsilon \in (0,1]$. By Theorem 1.1 of Dudley [20], there is a version of B_H with continuous sample functions on $[0,T]$. We assume throughout that B_H is such a sample-continuous version. A stronger statement follows from Corollary 2.3 of Dudley [20]. Namely, the function $\rho_H(u) := u^H\sqrt{|\log u|}$ for $u \geq 0$ is a sample modulus for B_H. That is for almost all $\omega \in \Omega$ there is a $K_\omega < \infty$ such that

$$|B_H(t,\omega) - B_H(s,\omega)| \leq K_\omega \rho_H(|t-s|) \qquad \text{for } t,s \in [0,T]. \tag{5.3}$$

Let $\kappa = \{t_i\colon i = 0,\dots,n\}$ be a partition of the interval $[0,T]$, and let

$$\varrho_H(u) := \left[u/\sqrt{|\log u|}\right]^{1/H} \quad \text{if } u > 0 \text{ and } \varrho_H(0) := 0.$$

For each $\omega \in \Omega$ satisfying (5.3), we then have

$$\sum_{i=1}^n \varrho_H(|B_H(t_i,\omega) - B_H(t_{i-1},\omega)|) \leq \sum_{i=1}^n \varrho_H(K_\omega \rho_H(t_i - t_{i-1})) \leq C_\omega \sum_{i=1}^n (t_i - t_{i-1}) = TC_\omega$$

for some finite constant C_ω which does not depend on κ. Therefore with probability 1,

$$v_\phi(B_H;[0,T]) = \sup\{v_\phi(B_H;\kappa)\colon \kappa \in \mathrm{PP}\,([0,T])\} < \infty \tag{5.4}$$

whenever $\phi = \varrho_H$. Since this function ϕ is bigger as $u \downarrow 0$ than u^p for any $p > 1/H$, it follows that the p-variation index $\upsilon(B_H) \leq 1/H$ for almost all sample functions of B_H.

However (5.4) also holds for the function $\phi = \phi_H$ defined on $[0,+\infty)$ by

$$\phi_H(t) := \left[t/\sqrt{2LL\,(1/t)}\right]^{1/H} \quad \text{for } t > 0 \text{ and } \quad \phi_H(0) := 0.$$

Define ψ_H on $[0,+\infty)$ by

$$\psi_H(t) := t^H\sqrt{2LL\,(1/t)} \quad \text{for } t > 0 \text{ and } \quad \psi_H(0) := 0,$$

where $L(u) := 1 \vee \log u$. The function ψ_H is continuous and stritly increasing. Let ψ_H^{-1} be its inverse function. Then ϕ_H is *asymptotic* to ψ_H^{-1} near the origin, that is $\phi_H(u)/\psi_H^{-1}(u) \to 1$ as $u \downarrow 0$. The following statement is a special case of the result of Kawada and Kôno [60]. In the case when $H = 1/2$, that is for a Brownian motion, it was proved by Taylor [106].

Theorem 5.4. *Let $1/2 \leq H < 1$ and let ϕ_H be regularly varying and asymptotic to ψ_H^{-1} near 0. For $0 < T < \infty$, with probability 1,*

$$\lim_{\delta\downarrow 0} \sup\{s_{\phi_H}(B_H;\kappa)\colon \kappa \in \mathrm{PP}\,([0,T]),\ |\kappa| \leq \delta\} = T. \tag{5.5}$$

By the same result of Kawada and Kôno [60], the relation (5.5) with "=" replaced by "$\leq$" also holds when the Hurst index $H \in (0,1/2)$.

It follows from the preceding theorem that if $1/2 \leq H < 1$ and $\phi_H(u) = o(\phi(u))$ as $u \downarrow 0$, then (5.4) fails. Thus the p-variation index of the fractional Brownian motion B_H, $1/2 \leq H < 1$, is given by

$$\upsilon(B_H) = 1/H \quad \text{with probability 1.}$$

5.4 Lévy processes

Homogeneous Lévy processes Let $X = \{X(t)\colon t \geq 0\}$ be a continuous in probability stochastic process with independent increments. If almost all sample functions of X are right continuous with left limits and $X(0) = 0$ almost surely then X is called a *Lévy process* (cf. Itô [57, Section 1.3]). A Lévy process X is called *homogeneous* if the distribution of $X(t+s)-X(t)$ with $t, s \geq 0$ does not depend on t. Given a real number a, a number $b \geq 0$, and a measure L on $\mathbb{R} \setminus \{0\}$ such $\int_{\mathbb{R}\setminus\{0\}}(1 \wedge x^2)\, L(dx) < \infty$, let

$$\eta(u) := \eta_{a,b,L}(u) := iau - bu^2 + \int_{\mathbb{R}\setminus\{0\}} (e^{iux} - 1 - iuh(x))L(dx) \tag{5.6}$$

for $u \in \mathbb{R}$, where $h(x) := x/(1 + x^2)$. The essential properties of h are that it is bounded on $\mathbb{R}$ and is $x + O(x^2)$ as $x \to 0$. Then η is called a *characteristic exponent* and L is called a *Lévy measure*. The characteristic function of a homogeneous Lévy process X is given by $E\exp\{iuX(t)\} = \exp\{t\eta(u)\}$ for each $t \geq 0$ and $u \in \mathbb{R}$.

It is well–known that sample functions of a homogeneous Lévy process X with the characteristic exponent (5.6) are of bounded variation if and only if $b = 0$ and

$$\int_{\mathbb{R}\setminus\{0\}} (1 \wedge |x|)\, L(dx) < \infty \tag{5.7}$$

(cf. e.g. Gikhman and Skorokhod [40, Theorem 3, p. 279]). The following result of Bretagnolle [8, Théorème III b] is less well–known. It sharpened an earlier result of Blumenthal and Getoor [7] and Monroe [83].

Theorem 5.5 (J. Bretagnolle). *Let $1 < p < 2$ and let $X = \{X(t)\colon t \geq 0\}$ be a mean zero homogeneous Lévy process with the characteristic exponent (5.6) such that $b \equiv 0$. Then $v_p(X; [0, 1]) < \infty$ with probability 1 if and only if*

$$\int_{\mathbb{R}\setminus\{0\}} (1 \wedge |x|^p)\, L(dx) < \infty. \tag{5.8}$$

If (5.8) fails then $v_p(X; [0, 1]) = +\infty$ almost surely.

α-stable Lévy motion A homogeneous Lévy process X is an *α-stable Lévy motion* of index α if its characteristic exponent (5.6) is given by $b \equiv 0$ and the Lévy measure $L = L_\alpha = L_{\alpha,r,q}$, where

$$L_{\alpha,r,q}(dx) := \begin{cases} rx^{-1-\alpha}dx & \text{if } x > 0 \\ q(-x)^{-1-\alpha}dx & \text{if } x < 0 \end{cases}$$

for $\alpha \in (0, 2)$ and $r, q \geq 0$ with $r + q > 0$. It is easy to see that (5.8) with $L = L_\alpha$ holds if and only if $p > \alpha$. Therefore $v_p(X_\alpha; [0, 1]) < \infty$ with probability 1 for each $p > \alpha$. Below we give a more exact result for α-stable Lévy motion.

Recall that $h(x) := x/(1 + x^2)$. If $\alpha < 1$ then (5.7) holds and $\int h\, dL < \infty$. In this case, it is said that an α-stable Lévy motion has no drift if $a + \int h\, dL = 0$. The following result is due to Fristedt and Taylor [38, Theorem 2].

Theorem 5.6 (B. Fristedt and S. J. Taylor). *Let $X_\alpha = \{X_\alpha(t)\colon t \in [0,1]\}$ be an α-stable Lévy motion with $\alpha \in (0,2)$ having no drift if $\alpha < 1$ and with $r = q$ if $\alpha = 1$. For an increasing function $\psi\colon [0,\infty) \to [0,\infty)$, with probability* 1

$$\lim_{|\kappa|\to 0} \sum_{i=1}^{n} \psi(|X_\alpha(t_i) - X_\alpha(t_{i-1})|) = \sum_{(0,1]} \psi(|\Delta^- X_\alpha|). \tag{5.9}$$

By [38, Proposition 3] the right side of (5.9) is finite almost surely if and only if

$$\int_0^1 \psi(u) u^{-1-\alpha}\, du < \infty.$$

Xu [115, Section 3] established necessary and/or sufficient conditions for the boundedness of the p-variation of a symmetric α-stable processes with possibly dependent increments. More information about p-variation of stable processes can be found in Fristedt [37].

5.5 Empirical processes

Let P be a probability distribution on $\mathbb{R}$ with a continuous distribution function F. Let $X_1, X_2, \ldots, X_n$ be independent with law P and let let F_n be the empirical distribution function $F_n(t) := n^{-1}\sum_{j=1}^{n} 1_{\{X_j \le t\}}$ for $t \in \mathbb{R}$. Let U be the $U[0,1]$ distribution function $U(t) := \max(0, \min(1,t))$, and U_n an empirical distribution function for it. Let $\alpha_n := n^{1/2}(F_n - F)$, the classical empirical process. Clearly $U \circ F \equiv F$ and take $F_n \equiv U_n \circ F$. Then $\alpha_n = n^{1/2}(U_n - U)\circ F$. Almost surely U_n is continuous at 0 and at 1. Then the p-variation of α_n on $\mathbb{R}$ equals that of $\beta_n := n^{1/2}(U_n - U)$ on $[0,1]$. In studying the p-variation of α_n we can thus assume $F \equiv U$ and $\alpha_n = \beta_n$. Qian [92, Theorems 3.2 and 4.2] proves the following

Theorem 5.7 (J. Qian). *Let $1 \le p < 2$. Then there are constants $C_p, \lambda_p < \infty$ depending only on p such that $Ev_p(\alpha_n) \le C_p n^{1-p/2}$, and $\limsup_{n\to\infty} v_p(\alpha_n)/n^{1-p/2} \le \lambda_p$ almost surely.*

Conversely, since the X_i are almost surely distinct, we have $v_p(\alpha_n) \ge n^{1-p/2}$, so the above theorem is sharp up to the constants C_p and λ_p.

For $p = 2$ we have:

Theorem 5.8. (a) *$v_2(\alpha_n) = O_P(\log\log n)$ as $n \to \infty$.*
(b) *If $0 < c < 1/12$, then $P(v_2(\alpha_n) > c \log\log n) \to 1$ as $n \to \infty$.*

Proof. Dudley [25] proved (a) and Qian [92, Theorem 3.1] proved (b). □

Let b_t, $0 \le t \le 1$, be the Brownian bridge process, a Gaussian stochastic process with mean 0 and $Eb_t b_u = t(1-u)$ for $0 \le t \le s \le 1$, and such that $t \mapsto b_t(\omega)$ is continuous for almost all ω. As for Brownian motion $v_p(t \mapsto b_t(\omega)) < \infty$ almost surely for $p > 2$ and $= +\infty$ almost surely for $p \le 2$. Y.–C. Huang [54], [55] proved the following:

Theorem 5.9 (Y.–C. Huang). *Let $2 < p < \infty$. On some probability space there exist empirical processes α_n and Brownian bridges $b^{(n)}$ such that $E\|\alpha_n - b^{(n)}\|_{[p]} = O(n^{\frac{1}{p}-\frac{1}{2}})$.*

The rate of convergence is best possible up to a constant multiple since $b^{(n)}$ is continuous: if the X_i are distinct, as they are almost surely, $\|\alpha_n - b^{(n)}\|_{(p)} \ge n^{\frac{1}{p}-\frac{1}{2}}$.

5.6 Differentiability of operators on processes

Some of the operators described in Chapter 3, e.g. the product integral operator (treated in Chapter 7 below), are only Fréchet differentiable with respect to p-variation norms for $p < 2$. In the Love–Young inequality and an inequality for convolutions (Gehring [39], see Appendix A in [28] for a detailed proof) at least one of the two functions must be in $\mathcal{W}_p$ for $p < 2$. On the other hand paths of Brownian motion and certain other processes only have bounded p–variation for $p > 2$, and the empirical process only converges in p–variation norm to the Brownian bridge for $p > 2$. The question arises how the differentiability can still be useful. We have examined this question for the empirical process.

Let T be an operator on distribution functions, Fréchet differentiable with respect to the p–variation norm at a distribution function F with a derivative linear operator DT and a remainder bound of order $\gamma > 1$. Then for empirical distribution functions F_n,

$$T(F_n) = T(F) + (DT)(F_n - F) + O(\|F_n - F\|_{[p]}^{\gamma}).$$

It follows from Theorem 5.7 that

$$\sqrt{n}(T(F_n) - T(F)) = (DT)(\sqrt{n}(F_n - F)) + O_{\mathrm{Pr}}(n^{[1-\gamma(p-1)]/2}).$$

For $\gamma = 1$ and $p \leq 2$, the remainder term would not approach 0 as $n \to \infty$. But if for a fixed $\gamma > 1$ we take $p < 2$ close enough to 2, then $1 - \gamma(p-1) < 0$ and the remainder becomes small as $n \to \infty$. For the bilinear operators and the product integral we have $\gamma = 2$, so that the remainder becomes $O_{\mathrm{Pr}}(n^{\epsilon - 1/2})$ for any $\epsilon > 0$. For the quantile operator, one has $\gamma = (p+1)/p \uparrow 3/2$ as $p \uparrow 2$ by Theorem 3.7, so that the remainder is $O_{\mathrm{Pr}}(n^{\epsilon-1/4})$ for any $\epsilon > 0$, where $-1/4$ is the best possible exponent by Bahadur–Kiefer theorems, cf. Section 3.4 above. Also for the composition operator (Theorem 3.3), γ does not approach 1 as $p \uparrow 2$.

Whenever $1 - \gamma(p-1) < 0$, if $(DT)(\sqrt{n}(F_n - F))$ converges in distribution to $(DT)(b \circ F)$ for a Brownian bridge b, then so does $\sqrt{n}(T(F_n) - T(F))$.

Chapter 6

Integration

A whole world, created by the Young dynasty, with its own ideas, methods and language.

Transl. from F. A. Medvedev. *Development of the Concept of Integral* (in Russian) [81, p. 378].

6.1 The refinement–Riemann–Stieltjes integral

In this section we consider an extension of the mesh–Riemann-Stieltjes integral suggested by in S. Pollard in 1923. The idea of the extension is to replace the convergence as the mesh tends to zero used to define the (MRS) integral by a type of convergence now called net convergence or, from another angle, filter convergence. Net convergence had previously been introduced by Moore [85], Moore and Smith [86]. In his paper [91] Pollard shows that the new integral extends and improves the mesh–Riemann-Stieltjes integral in some ways.

Throughout the chapter f and h are real-valued functions on a closed interval $[a,b]$. Let $\kappa = \{x_i\colon i=1,\ldots,n\}$ be a partition of $[a,b]$ and let $\sigma = \{y_i\colon i=1,\ldots,n\}$ be an intermediate partition of κ, that is $x_{i-1} \le y_i \le x_i$ for $i=1,\ldots,n$. As before the set of all partitions κ of $[a,b]$ is denoted by $\mathrm{PP}([a,b])$ and the Riemann-Stieltjes sum based on (κ,σ) is defined by (4.22). Recall that the refinement–Riemann–Stieltjes integral or (RRS) integral of f with respect to h exists and equals the number A, if for each $\epsilon > 0$ there is a partition λ of $[a,b]$ such that

$$\Big|S_{RS}(f,h;\kappa,\sigma) - A\Big| < \epsilon \tag{6.1}$$

holds for all refinements κ of λ and all intermediate partitions σ of κ. If the refinement–Riemann–Stieltjes integral is defined then we write $(RRS)\int_a^b f\,dh := A$.

Proposition 6.1. *The (RRS) integral is an extension of the (MRS) integral.*

Proof. Suppose $(MRS)\int_a^b f\,dh$ exists with a value A. Then given any positive number $\epsilon > 0$, one can find a $\delta > 0$ such that (6.1) holds whenever the mesh $|\kappa| < \delta$. Let $\lambda := \{a, a+\delta/2, a+\delta, a+3\delta/2, \ldots, a+n\delta, b\}$ for the minimal integer n such that $a+n\delta+\delta/2 > b$. Then (6.1) also holds for each refinement κ of λ. Thus $(RRS)\int_a^b f\,dh$ exists and has the same value A. □

Example 6.2. Let f be a function on $[a,b]$ and let $\ell_c := 1_{[c,b]}$ be the indicator function of $[c,b]$ for some $a < c < b$. For any $\kappa = \{x_i \colon i = 1,\dots,n\}$ and intermediate partition $\sigma = \{y_i \colon i = 1,\dots,n\}$ of κ, we have $S_{RS}(f, \ell_c; \kappa, \sigma) = f(y_i)$ for $x_{i-1} < c \le x_i$ and $y_i \in [x_{i-1}, x_i]$. The integral $(MRS) \int_a^b f\,d\ell_c$ exists and equals $f(c)$ if and only if f is continuous at c. Taking $c \in \kappa$ it follows that the integral $(RRS) \int_a^b f\,d\ell_c$ exists and equals $f(c)$ if and only if f is left-continuous at c. Therefore $(RRS) \int_a^b 1_{[a,c]}\,d\ell_c$ exists and equals 1 while the same integral in the mesh–Riemann-Stieltjes sense doesn't exist. However the other integral $(RRS) \int_a^b \ell_c\,d\ell_c$ also doesn't exist.

The following is proved in Hildebrandt (1963, Theorem II.10.9).

Theorem 6.3. *Let f and h be bounded functions on $[a,b]$. The integral $(MRS) \int_a^b f\,dh$ exists if and only if both $(RRS) \int_a^b f\,dh$ exists and the two functions f, h have no common discontinuities on $[a,b]$.*

Properties 6.4. *Consider the following properties of an integral $\int_a^b f\,dh$:*

I. $\int_a^b 1_{[a,b]}\,dh = h(b) - h(a)$.

II. For $u_1, u_2 \in \mathbb{R}$ and $f_1, f_2 \colon [a,b] \mapsto \mathbb{R}$,

$$\int_a^b (u_1 f_1 + u_2 f_2)\,dh = u_1 \int_a^b f_1\,dh + u_2 \int_a^b f_2\,dh$$

where the left side exists provided the right side does.

III. For $v_1, v_2 \in \mathbb{R}$ and $h_1, h_2 \colon [a,b] \mapsto \mathbb{R}$,

$$\int_a^b f\,d(v_1 h_1 + v_2 h_2) = v_1 \int_a^b f\,dh_1 + v_2 \int_a^b f\,dh_2$$

where again the left side exists provided the right side does.

IV. For $a < c < b$, $\int_a^b f\,dh$ exists if and only if both $\int_a^c f\,dh$ and $\int_c^b f\,dh$ exist, and then

$$\int_a^b f\,dh = \int_a^c f\,dh + \int_c^b f\,dh.$$

V. Let $\int_a^b f\,dh$ exists and let $F(y) := \int_a^y f\,dh$ for $y \in [a,b]$. Then for each $x \in [a,b]$,

$$\lim_{y \to x} \Big\{ F(y) - F(x) - f(x)[h(y) - h(x)] \Big\} = 0. \tag{6.2}$$

Pollard [91, p. 80] showed that the "if" part of IV does not hold for the (MRS) integral. Indeed let $\bar{f} := 1_{[1,2]}$ and $\bar{h} := 1_{(1,2]}$ be indicator functions on $[0,2]$. Then $(MRS) \int_0^1 \bar{f}\,d\bar{h} = 0$ and $(MRS) \int_1^2 \bar{f}\,d\bar{h} = h(2) - h(1) = 1$. For any $\kappa = \{x_i\} \in \mathrm{PP}([a,b])$ such that $1 \notin \kappa$ and an intermediate partition $\sigma = \{y_i\}$ of κ, we have

$$S_{RS}(\bar{f}, \bar{h}; \kappa, \sigma) = \bar{f}(y_i)[\bar{h}(x_i) - \bar{h}(x_{i-1})] = \begin{cases} 1 & \text{if } y_i = x_i \\ 0 & \text{if } y_i = x_{i-1}, \end{cases}$$

where i is such that $x_{i-1} < 1 < x_i$. Clearly κ can have in addition arbitrarily small mesh. Thus $(MRS) \int_0^2 \bar{f}\,d\bar{h}$ doesn't exist. A sufficient condition for the (MRS) integral to satisfy condition IV is given in Theorem II.4.4N of Hildebrandt [52, p. 34].

Theorem 6.5. *The* (RRS) *integral satisfies the above properties* I–V.

Proof. Pollard [91] proved the first four properties. One can also find proofs in Hildebrandt [52, Section II.11]. Here we prove the fifth property for the (RRS) integral.

Suppose the integral $(RRS)\int_a^b f\,dh$ exists. Let $\epsilon > 0$ and $x \in (a,b]$. There exists $\lambda \in \mathrm{PP}\,([a,b])$ such that (6.1) holds for all refinements κ of λ and all intermediate partitions σ of κ. We can and do assume that $x = x_i \in \lambda$ for some $0 \le i < \mathrm{card}\lambda$. Let $y \in (x, x_{i+1}]$. By property IV for the (RRS) integral, there exists a partition $\kappa_{x,y}$ of $[x,y]$ such that the Riemann–Stieltjes sum $S_{RS}(\kappa_{x,y}, \sigma_{x,y})$ differs from $F(y) - F(x)$ by less than ϵ. Let κ_1 and κ_2 be two refinements of λ which coincide with $\kappa_{x,y}$ and $\{x,y\}$ when restricted to $[x,y]$. By (6.1), it then follows that

$$\begin{aligned}\Big|F(y) - F(x) - f(x)[h(y) - h(x)]\Big| &\le \Big|(RRS)\int_x^y f\,dh - S_{RS}(\kappa_{x,y}, \sigma_{x,y})\Big| \\ &\quad + |S_{RS}(\kappa_1, \sigma_1) - S_{RS}(\kappa_2, \sigma_2)| \le 3\epsilon,\end{aligned}$$

where the intermediate partition σ_1 of κ_1, when restricted to $[x,y]$, coincides with the intermediate partition $\sigma_{x,y}$ of $\kappa_{x,y}$, and the intermediate partition σ_2 of κ_2 has an element x corresponding to the interval $[x,y]$. Thus (6.2) holds for $y \downarrow x$ instead of $y \to x$. Similarly, given $x \in (a,b]$, one can show that (6.2) also holds when $y \uparrow x$. The proof of property V for the (RRS) integral is complete. $\square$

Proposition 6.6 (Cauchy criterion). *The integral* $(RRS)\int_a^b f\,dh$ *exists if and only if, for each* $\epsilon > 0$, *there exists* $\lambda \in \mathrm{PP}\,([a,b])$ *such that*

$$|S_{RS}(f,h;\kappa_1,\sigma_1) - S_{RS}(f,h;\kappa_2,\sigma_2)| < \epsilon \tag{6.3}$$

for any refinements κ_1, κ_2 *of* λ *and any intermediate partitions* σ_1, σ_2.

Proof. The "only if" implication is clear. To prove the "if" part let $\epsilon_m \downarrow 0$. For each $m \ge 1$, let λ_m be a partition λ existing by the hypothesis for $\epsilon = \epsilon_m$. One can and does assume that $\lambda_m \subset \lambda_{m+1}$ for each m. Let $\{\sigma_m \colon m \ge 1\}$ be any sequence of intermediate partitions of λ_m, $m \ge 1$. Then $\{S_{RS}(\lambda_m;\sigma_m) \colon m \ge 1\}$ is a Cauchy sequence of real numbers. Let A be its limit. Given $m \ge 1$, let κ be a refinement of λ_m and let σ be an intermediate partition of κ. By (6.3), we then have

$$|S_{RS}(\kappa,\sigma) - A| \le |S_{RS}(\kappa,\sigma) - S_{RS}(\lambda_{m+n},\sigma_{m+n})| + |S_{RS}(\lambda_{m+n},\sigma_{m+n}) - A| < 2\epsilon_m$$

for large enough integers $n \ge 1$. Therefore $(RRS)\int_a^b f\,dh$ exists and has value A. $\square$

Proposition 6.7. *If* $(RRS)\int_a^b f\,dh$ *exists then both* f *and* h *can't have discontinuities on the same side at same point; that is, for* $x \in (a,b]$, *either*

$$\lim_{y\uparrow x} f(y) = f(x) \qquad or \qquad \lim_{y\uparrow x} h(y) = h(x), \tag{6.4}$$

and the same for $x \in [a,b)$ *with* $y \uparrow x$ *replaced by* $y \downarrow x$.

Proof. Let $x \in (a,b]$ and $\epsilon > 0$. Then there exists $\lambda \in \mathrm{PP}\,([a,b])$ such that (6.1) holds for each refinement κ of λ. We can and do suppose that $x \in \lambda$. Subtract two Riemann-Stieltjes sums based on the same κ and two intermediate partitions of κ having all points equal except two of them in the interval with the right endpoint x. This gives the bound

$$|f(y') - f(y'')||h(x') - h(x)| < 2\epsilon$$

for any $x' \le y' < y'' \le x$ and x' close enough to x. Thus (6.4) holds. A symmetric argument yields the second part of the Proposition 6.7. $\square$

Naturally the preceding necessary condition on discontinuities is stronger in the case when the integral exists in the mesh–Riemann-Stieltjes sense.

Proposition 6.8. *If $(MRS) \int_a^b f\,dh$ exists then both f and h can't have a discontinuity at the same point; that is, for $x \in [a,b]$, either $\lim_{y\to x} f(y) = f(x)$ or $\lim_{y\to x} h(y) = h(x)$.*

Proof. Let $x \in [a,b]$ and $\epsilon > 0$. Then there exists $\delta > 0$ such that (6.1) holds for each partition κ with mesh $|\kappa| < \delta$. Subtract two Riemann-Stieltjes sums based on the same κ and two intermediate partitions of κ having all points equal except two of them in the interval containing the point x. This gives the bound

$$|f(y') - f(y'')||h(x') - h(x'')| < 2\epsilon$$

for any $x', x'' \in [a \vee (x-\delta), b \wedge (x+\delta)]$ and $y', y'' \in [x', x'']$. Thus Proposition 6.8 holds. $\square$

6.2 The refinement–Young–Stieltjes integral

Recall the (RYS) integral as defined in Section 2.1. Let f be any function on $[a,b]$, and let $h \in \mathcal{R}[a,b]$. Given a partition $\kappa = \{x_i\colon i = 0,\dots,n\} \in \mathrm{PP}\,([a,b])$, we say that an intermediate partition $\sigma = \{y_i\colon i = 1,\dots,n\}$ of κ is *open* if $x_{i-1} < y_i < x_i$ for $i = 1,\dots,n$, or *left–open* if $x_{i-1} < y_i \le x_i$ for $i = 1,\dots,n$. Then the Young sum $S_Y(f,h;\kappa,\sigma)$ based on (κ,σ) is defined by

$$\begin{aligned} S_Y(f,h;\kappa,\sigma) &:= \sum_{i=1}^{n}\{[f\Delta^{+}h](x_{i-1}) + f(y_i)[h(x_i-) - h(x_{i-1}+)] + [f\Delta^{-}h](x_i)\} \qquad (6.5)\\ &= [f\Delta^{+}h](a) + [f\Delta^{-}h](b) + \sum_{i=1}^{n} f(y_i)[h(x_i-) - h(x_{i-1}+)] + \sum_{i=1}^{n-1}[f\Delta^{\pm}h](x_i). \end{aligned}$$

Example 6.9. As in Example 6.2, for $a < c < b$, let ℓ_c be the indicator function of $[c,b]$ and f be any function, both on $[a,b]$. For each partition κ containing c and an open intermediate partition σ of κ, we have $S_Y(f,\ell_c;\kappa,\sigma) = f(c)$. Therefore the integral $(RYS) \int_a^b f\,d\ell_c$ exists and has value $f(c)$. In particular, $(RYS) \int_a^b \ell_c\,d\ell_c$ exists and has value 1, while the (RRS) integral doesn't exist as shown in Example 6.2. More generally, for any real number r, let $\ell_c^r(x) := \ell_c(x)$ for $x \ne c$ and $\ell_c^r(c) := r$, so that $\ell_c^1 \equiv \ell_c$. Then $(RYS) \int_a^b f\,d\ell_c^r$ exists and has value $f(c)$ for any r.

In fact the (RYS) integral $\int_a^b f\,dh$ does not depend on values of the integrator h at jump points in (a,b). This is because the value of the Young sum (6.5) does not depend on values of h at jump points of the open interval (a,b). The same is not true for the (RRS) integral. Indeed, by changing the value of the integrator h at a point x one can make h either right-continuous or left-continuous at x and so destroy the necessary condition of Proposition 6.7. Unlike the (RRS) integral, the (MRS) integral shares the same property of independence of values of the integrator at discontinuity points with the Young integral. Indeed, by Proposition 6.8, if $(MRS)\int_a^b f\,dh$ exists and h has a jump at $x \in (a,b)$ then f must be continuous at x. For $\xi \in \mathbb{R}$, let $h_\xi := h$ on $[a,x) \cup (x,b]$ and $h_\xi(x) := \xi$. Then for the two Riemann-Stieltjes sums we have

$$S_{RS}(f,h;\kappa,\sigma) - S_{RS}(f,h_\xi;\kappa,\sigma) = \begin{cases} [f(y') - f(y'')][h(x) - \xi] & \text{if } x \in \kappa \\ 0 & \text{if } x \notin \kappa, \end{cases}$$

where $y' \in [x-\delta, x]$ and $y'' \in [x, x+\delta]$ for $\delta := |\kappa|$. Due to continuity of f at x the preceding difference can be made arbitrarily small if the mesh $|\kappa|$ is small enough.

Properties of the (RYS) integral.

Proposition 6.10 (Cauchy criterion). *The integral $(RYS)\int_a^b f\,dh$ exists if and only if, for each $\epsilon > 0$, there exists $\lambda \in \mathrm{PP}([a,b])$ such that*

$$|S_Y(f,h;\kappa_1,\sigma_1) - S_Y(f,h;\kappa_2,\sigma_2)| < \epsilon$$

for any refinements κ_1, κ_2 of λ and any open intermediate partitions σ_1, σ_2.

The proof of the Cauchy criterion for the (RYS) integral is the same as for the (RRS) integral given by Proposition 6.6 above.

Theorem 6.11. *The (RYS) integral satisfies statement IV among Properties 6.4, that is, for $a < c < b$, $(RYS)\int_a^b f\,dh$ exists if and only if both $(RYS)\int_a^c f\,dh$ and $(RYS)\int_c^b f\,dh$ exist, and then*

$$(RYS)\int_a^b f\,dh = (RYS)\int_a^c f\,dh + (RYS)\int_c^b f\,dh. \tag{6.6}$$

Proof. To see that the "if" part holds notice that the additivity relation

$$S_Y(f,h;\kappa,\sigma) = S_Y(f,h;\kappa_1,\sigma_1) + S_Y(f,h;\kappa_2,\sigma_2) \tag{6.7}$$

holds for each $\kappa \in \mathrm{PP}([a,b])$ containing c, $\kappa_1 \in \mathrm{PP}([a,c])$, $\kappa_2 \in \mathrm{PP}([c,b])$, $\kappa_1 \cup \kappa_2 = \kappa$ (as a set), and for corresponding intermediate partitions such that $\sigma_1 \cup \sigma_2 = \sigma$.

For the converse suppose that the integral exists over $[a,b]$. Then the Cauchy condition of Proposition 6.10 holds. We can and do assume that $c \in \lambda \in \mathrm{PP}([a,b])$. Then using this condition when intermediate partitions have the same points from the interval $[c,b]$, one gets that the Cauchy condition holds for the integral over $[a,c]$. A symmetric argument yields that the integral over $[c,b]$ also exists. The additivity relation (6.6) then follows from (6.7). The proof of Theorem 6.11 is complete. $\square$

Let F be an indefinite (RYS) integral defined for a constant C by

$$F(x) := C + (RYS)\int_a^x f\,dh \qquad \text{for } x \in [a,b].$$

According to the following theorem, F is a regulated function, and for $a \le y < x \le b$,

$$(\Delta^- F)(x) = [f\Delta^- h](x) \quad \text{and} \quad (\Delta^+ F)(y) = [f\Delta^+ h](y). \tag{6.8}$$

Theorem 6.12. *The* (RYS) *integral satisfies statement V among Properties 6.4, that is, if* $(RYS)\int_a^b f\,dh$ *exists then, for each* $x \in [a,b]$,

$$\lim_{y\to x}\Big\{F(x)-F(y)-f(x)[h(x)-h(y)]\Big\}=0.$$

Proof. Suppose $(RYS)\int_a^b f\,dh$ exists. Then the indefinite integral F exists by Theorem 6.11. We prove that, for $a < x \le b$,

$$\lim_{y\uparrow x}\Big\{F(x)-F(y)-f(x)[h(x)-h(y)]\Big\}=0. \tag{6.9}$$

The proof when $y \downarrow x$ for $a \le x < b$ is symmetric and is omited. Taking $u_y \in (y,x)$, we have

$$\lim_{y\uparrow x}\Big\{S_Y(f,h;\{y,x\},\{u_y\})-f(x)[h(x)-h(y)]\Big\}$$

$$=\lim_{y\uparrow x}\Big\{[f(u_y)-f(x)][h(x-)-h(y+)]+[f(y)-f(x)]\Delta^+h(y)\Big\}=0.$$

Therefore it is enough to prove

$$\lim_{y\uparrow x}\Big\{F(x)-F(y)-S_Y(f,h;\{y,x\},\{u_y\})\Big\}=0$$

for $u_y \in (y,x)$. Let $a < x \le b$ and $\epsilon > 0$. By definition of the (RYS) integral, there exists $\lambda = \{z_j\colon j=0,\dots,m\} \in \mathrm{PP}\,([a,x])$ such that

$$|S_Y(f,h;\kappa,\sigma)-(RYS)\int_a^x f dh|<\epsilon$$

for each refinement κ of λ and intermediate partition σ of κ. Let $y \in [z_{m-1},x)$. Within ϵ of $(RYS)\int_y^x f\,dh$, choose a Riemann–Stieltjes sum $S_{y,x}$ based on a partition $\kappa_{y,x}$ of $[y,x]$. Let κ_1 and κ_2 be two refinements of λ which coincide with $\kappa_{y,x}$ and $\{y,x\}$, respectively, when restricted to $[y,x]$. Then taking $u_y \in (y,x)$, we have

$$|F(x)-F(y)-S_Y(f,h;\{y,x\},\{u_y\}]| \le \Big|(RYS)\int_y^x f\,dh - S_{y,x}\Big|$$

$$+|S_Y(f,h;\kappa_1,\sigma_1)-S_Y(f,h;\kappa_2,\sigma_2)| \le 3\epsilon$$

Therefore (6.9) holds, and the proof is complete. □

Connection with the (RRS) **integral.**

Theorem 6.13. *Let* f *be a function on* $[a,b]$, *and let* h *be regulated on* $[a,b]$. *If* $(RRS)\int_a^b f\,dh$ *exists then so does* $(RYS)\int_a^b f\,dh$, *and both have the same value.*

Proof. Let $\kappa = \{x_i\colon i=0,\dots,n\}$ be a partition of $[a,b]$ and let $\sigma = \{y_i\colon i=1,\dots,n\}$ be an open intermediate partition of κ. Let $\mu = \{u_{i-1},v_i\colon i=1,\dots,n\}$ be a set of points in (a,b) such that $x_0 < u_0 < y_1 < v_1 < x_1 < \cdots < x_{n-1} < u_{n-1} < y_n < v_n < x_n$. Then $\kappa' := \kappa \cup \mu$

is a partition of $[a,b]$ and $\sigma' := \{x_0, y_1, x_1, x_1, y_2, \dots, x_{n-1}, x_{n-1}, y_n, x_n\}$ is an intermediate partition of κ'. Letting $u_{i-1} \downarrow x_{i-1}$ and $v_i \uparrow x_i$ for $i = 1, \dots, n$, we get

$$\begin{aligned} S_{RS}(f,h;\kappa',\sigma') &= \sum_{i=1}^{n} f(y_i)[h(v_i) - h(u_{i-1})] + f(x_0)[h(u_0) - h(x_0)] \\ &+ \sum_{i=1}^{n-1} f(x_i)[h(u_i) - h(v_i)] + f(x_n)[h(x_n) - h(v_n)] \to S_Y(f,h;\kappa,\sigma). \end{aligned}$$

Therefore for each $\kappa \in \mathrm{PP}\,([a,b])$ and open intermediate partition σ of κ, the Young sums based on κ and σ can be approximated arbitrarily closely by Riemann–Stieltjes sums based on refinements of κ, and the theorem follows. ☐

Lemma 6.14. *Let $\phi, \psi \in \mathcal{R}[a,b]$ be such that ϕ is left–continuous on $(a,b]$ and ψ is right–continuous on $[a,b)$, or vice versa. If (RYS) $\int_a^b \phi\, d\psi$ exists then so does (RRS) $\int_a^b \phi\, d\psi$, and the two integrals have the same value.*

Proof. Let $f, h \in \mathcal{R}[a,b]$. The proof is given only when $\phi = f_-^{(a)}$ and $\psi = h_+^{(b)}$. A proof for the other case is similar. Let $\kappa = \{x_i \colon i = 0, \dots, n\} \in \mathrm{PP}\,([a,b])$ and let $\sigma = \{y_i \colon i = 1, \dots, n\}$ be an open intermediate partition of κ, that is $x_{i-1} < y_i < x_i$ for $i = 1, \dots, n$. Then we have

$$S_{RS}(f_-^{(a)}, h_+^{(b)}; \kappa, \sigma) - S_Y(f_-^{(a)}, h_+^{(b)}; \kappa, \sigma) = \sum_{i=1}^{n} f_-^{(a)}(y_i)[h_+^{(b)}(x_i) - h_+^{(b)}(x_{i-1})]$$

$$-\sum_{i=1}^{n} \Big\{ f(y_i-)[h(x_i-) - h(x_{i-1}+)] + f(x_i-)[h_+^{(b)}(x_i) - h(x_i-)] \Big\}$$

$$= \sum_{i=1}^{n} [f(y_i-) - f(x_i-)][h_+^{(b)}(x_i) - h(x_{i-1}+)] - \sum_{i=1}^{n} [f(y_i-) - f(x_i-)][h(x_i-) - h(x_{i-1}+)]. \quad (6.10)$$

Let $\epsilon > 0$. There exists $\lambda = \{z_j \colon j = 0, \dots, m\} \in \mathrm{PP}\,([a,b])$ such that

$$\Big| S_Y(f_-^{(a)}, h_+^{(b)}; \kappa, \sigma) - (Y) \int_a^b f_-^{(a)}\, dh_+^{(b)} \Big| < \epsilon \quad (6.11)$$

for all refinements $\kappa \in \mathrm{PP}\,([a,b])$ of λ and all open intermediate partitions σ of κ. Choose $\nu = \{v_{j-1}, u_j \colon j = 1, \dots, m\} \subset [a,b]$ such that $z_{j-1} < v_{j-1} < u_j < z_j$ for each $j = 1, \dots, m$,

$$\mathrm{Osc}(f; [u_j, z_j-]) \le \epsilon/(2m\|h\|_\infty) \quad \text{and} \quad \mathrm{Osc}(h; [z_{j-1}+, v_{j-1}+]) \le \epsilon/(2m\|f\|_\infty). \quad (6.12)$$

Let $\kappa = \{x_i \colon i = 0, \dots, n\} \in \mathrm{PP}\,([a,b])$ be a refinement of $\lambda \cup \nu \subset$ and let $\sigma = \{y_i \colon i = 1, \dots, n\}$ be an open intermediate partition of κ. We show that the absolute value of the right side of (6.10) is small. Let $\sigma' = \{y_i' \colon i = 1, \dots, n\}$ be another open intermediate partition of κ. Using (6.11) for two Young sums based on (κ, σ) and (κ, σ'), we get

$$\Big| \sum_{i=1}^{n} [f(y_i-) - f(y_i'-)][h(x_i-) - h(x_{i-1}+)] \Big| < 2\epsilon.$$

Then letting $y_i' \uparrow x_i$ for $i = 1, \dots, n$, it follows that the second sum on the right side of (6.10) has the bound

$$\Big|\sum_{i=1}^{n} [f(y_i-) - f(x_i-)][h(x_i-) - h(x_{i-1}+)]\Big| \le 2\epsilon. \tag{6.13}$$

To bound the first sum on the right side of (6.10) we invoke an auxiliary partition $\kappa' = \{x_i' : i = 0, \dots, n\} \in \mathrm{PP}([a, b])$. For each $j = 0, \dots, m$, let $i(j) \in \{0, \dots, n\}$ be such that $x_{i(j)} = z_j$. Let $J := \{i(j-1)+1, i(j) : j = 1, \dots, m\}$ and let $I := \{0, \dots, n\} \setminus J$. Define κ' by $x_i' := x_i$ for $i \in J$ and $x_i' \in (x_i, y_{i+1})$ for $i \in I$. Notice that $\kappa' \supset \lambda$ and $\sigma = \{y_i : i = 1, \dots, n\}$ is an open intermediate partition of κ'. Let $\sigma' = \{y_i' : i = 1, \dots, n\}$ be another open intermediate partition for κ' defined by $y_i' := y_i$ for $i \in J$ and $y_i' := x_i$ for $i \in I$. Then using (6.11) for two Young sums corresponding to (κ', σ) and (κ', σ') we get the bound

$$\Big|\sum_{i \in I} [f(y_i-) - f(x_i-)][h(x_i'-) - h(x_{i-1}'+)]\Big| < 2\epsilon.$$

Letting $x_i' \downarrow x_i$ for $i \in I$ it follows that

$$\Big|\sum_{i \in I} [f(y_i-) - f(x_i-)][h(x_i+) - h(x_{i-1}+)]\Big| \le 2\epsilon.$$

Since $u_j \le x_{i(j)-1} < y_{i(j)} < z_j$ and $x_{i(j-1)+1} \le v_{j-1}$ for each $j = 1, \dots, m$, by (6.12), we get

$$\Big|\sum_{i=1}^{n} [f(y_i-) - f(x_i-)][h_+^{(b)}(x_i) - h(x_{i-1}+)]\Big| \le 2\|h\|_\infty \sum_{j=1}^{m} \mathrm{Osc}(f; [u_j, z_j-])$$

$$+2\|f\|_\infty \sum_{j=0}^{m-1} \mathrm{Osc}(h; [z_j+, v_j+]) + \Big|\sum_{i \in I} [f(y_i-) - f(x_i-)][h(x_i+) - h(x_{i-1}+)]\Big| \le 4\epsilon.$$

This in conjunction with (6.13) yields that the absolute value of the right side of (6.10) does not exceed 6ϵ. Therefore, by (6.11), we have the bound

$$\Big|S_{RS}(f_-^{(a)}, h_+^{(b)}; \kappa, \sigma) - (Y)\int_a^b f_-^{(a)}\, dh_+^{(b)}\Big| \le 7\epsilon \tag{6.14}$$

for all $\kappa \in \mathrm{PP}([a, b])$ with $\kappa \supset \lambda \cup \nu$ and all open intermediate partitions σ of κ. For any such $\kappa = \{x_i : i = 0, \dots, n\}$, since $f_-^{(a)}$ is left–continuous, (6.14) also holds for all *left-open* intermediate partitions $\sigma = \{y_i : i = 1, \dots, n\}$ of κ, that is $x_{i-1} < y_i \le x_i$ for $i = 1, \dots, n$.

To finish the proof we need to extend the above bound to an arbitrary intermediate partition σ of κ. Given $\epsilon > 0$, let λ be the partition $\{z_j : j = 0, \dots, m\}$ of $[a, b]$ satisfying (6.11), and let ν be the set of points $\{v_{j-1}, u_j : j = 1, \dots, m\}$ such that $z_{j-1} < v_{j-1} < u_j < z_j$ for $j = 1, \dots, m$ and (6.12) holds. By the above proof, (6.14) then holds for any refinement κ of $\lambda \cup \nu'$ for some $\nu' = \{v_{j-1}', u_j' : j = 1, \dots, m\}$ with $z_{j-1} < v_{j-1}' \le v_{j-1} < u_j \le u_j' < z_j$, $j = 1, \dots, m$, and left-open intermediate partition σ of κ. Let $\kappa = \{x_i : i = 0, \dots, n\} \in \mathrm{PP}([a, b])$ be such that $\kappa \supset \lambda \cup \nu \cup \{(z_j + v_j)/2 : j = 0, \dots, m-1\}$ and let $\sigma = \{y_i : i = 1, \dots, n\}$ be an intermediate partition of κ, that is $x_{i-1} \le y_i \le x_i$ for $i = 1, \dots, n$. Let σ_e be the same as σ except that each $y_i \in \sigma$ with i even is replaced by x_i. Similarly, let σ_o be the same as σ except that each $y_i \in \sigma$ with i odd is replaced by x_i. Also, let $\sigma_m := \{x_i : i = 1, \dots, n\}$ and $S(\sigma) := S_{RS}(f_-^{(a)}, h_+^{(b)}; \kappa, \sigma)$. Then we have $S(\sigma) + S(\sigma_m) = S(\sigma_e) + S(\sigma_o)$. Hence

$$|S(\sigma) - A| \le |S(\sigma_m) - A| + |S(\sigma_e) - A| + |S(\sigma_o) - A| \tag{6.15}$$

with $A = (Y)\int_a^b f_-^{(a)}\,dh_+^{(b)}$. The first term on the right side of (6.15) does not exceed 7ϵ by (6.14) because σ_m is a left-open intermediate partition of κ. To bound the second term on the right side of (6.15), let I_σ be the set of indices $i \in \{1,\ldots,n\} \setminus J$ such that $y_i \in \sigma_e$ and $y_i = x_{i-1}$, where $J = \{i(j-1)+1, i(j)\colon j = 1,\ldots,m\}$. So, if $i \in I_\sigma$ then $i \in \{2,\ldots,n-1\}$ is odd and

$$f(y_i-)[h(x_i+) - h(x_{i-1}+)] + f(y_{i-1}-)[h(x_{i-1}+) - h(x_{i-2}+)] = f(x_{i-1}-)[h(x_i+) - h(x_{i-2}+)]$$

because $y_i = y_{i-1} = x_{i-1}$ by the construction of σ_e. Let $\kappa' := \kappa \setminus \{x_{i-1}\colon i \in I_\sigma\}$. Notice that $\kappa' \supset \lambda \cup \{v'_{j-1}, u_j\colon j = 1,\ldots,m\}$ with $v'_{j-1} = x_{i(j-1)+1} < v_{j-1}$ or $v'_{j-1} = x_{i(j-1)+2} \le v_{j-1}$ for $j = 1,\ldots,m$. Also, let σ'_e be the same as σ_e except that each y_i, $i \in J$, such that $y_i = x_{i-1}$, is replaced by x_i, and one of the two points x_{i-1}, $i \in I_\sigma$, is removed. Then σ'_e is a left-open intermediate partition of κ'. By (6.14) and (6.12), it then follows that

$$|S(\sigma_e) - A| \le |S_{RS}(f_-^{(a)}, h_+^{(b)}; \kappa', \sigma'_e) - A| + 2\|f\|_\infty \sum_{j=0}^{m-1} \mathrm{Osc}(h; [z_j+, v_j+])$$

$$+2\|h\|_\infty \sum_{j=1}^{m} \mathrm{Osc}(f; [u_j, z_j-]) \le 9\epsilon.$$

The same reasoning yields the bound $|S(\sigma_o) - A| \le 9\epsilon$. Therefore, by (6.15), $|S(\sigma) - A| \le 25\epsilon$ for an arbitrary intermediate partition σ of κ. The proof of Lemma 6.14 is complete. □

6.3 The central Young integral

L. C. Young [117] defined (and Dudley [23] specified the endpoint terms of) the integral which we call the Central Young integral, or the (CY) integral. The idea of the (CY) integral is to use the (RRS) integral, avoiding its lack of definition when f and h have common one–sided discontinuities by taking a right–continuous version of f and left–continuous version of h, or vice versa, and adding sums of jump terms to restore the desired value of $\int f\,dh$. In general the (CY) integral extends the (RYS) integral. However, if the integrand and integrator have bounded p– and p'–variation, respectively, with $1/p + 1/p' = 1$, then the (CY) integral is defined only if the (RYS) integral is (see Proposition 6.22 below).

Recall the definitions (2.2) of functions $f_+^{(b)}$ and $f_-^{(a)}$, as well as unconditional convergence as defined in Section 2.1. The (CY) integral has two equivalent forms defined next:

Definition 6.15. Let $f, h \in \mathcal{R}([a,b])$. If $a < b$ define the (Y_1) integral

$$(Y_1)\int_a^b f\,dh := (RRS)\int_a^b f_+^{(b)}\,dh_-^{(a)} - [\Delta^+ f \Delta^+ h](a) + [f\Delta^- h](b) - \sum_{(a,b)} \Delta^+ f \Delta^{\pm} h \qquad (6.16)$$

if the (RRS) integral exists and the sum converges unconditionally. If $a = b$ define the (Y_1) integral as 0. Similarly, if $a < b$ define the (Y_2) integral

$$(Y_2)\int_a^b f\,dh := (RRS)\int_a^b f_-^{(a)}\,dh_+^{(b)} + [f\Delta^+ h](a) + [\Delta^- f \Delta^- h](b) + \sum_{(a,b)} \Delta^- f \Delta^{\pm} h \qquad (6.17)$$

if the (RRS) integral exists and the sum converges unconditionally. If $a = b$ define the (Y_2) integral as 0.

By Theorem 6.19 below, the two integrals both exist and have the same value when either exists. To prove this theorem we replace the (RRS) integral in (6.16) and (6.17) by the (RYS) integral. Let

$$f_-^{(a,b)}(x) := \begin{cases} f(a) & \text{for } x = a \\ f(x-) & \text{for } a < x < b \\ f(b) & \text{for } x = b \end{cases} \quad \text{and} \quad f_+^{(a,b)}(x) := \begin{cases} f(a) & \text{for } x = a \\ f(x+) & \text{for } a < x < b \\ f(b) & \text{for } x = b \end{cases}$$

Lemma 6.16. *Let $f, h \in \mathcal{R}([a,b])$. The integral $(RRS)\int_a^b f_+^{(b)}\,dh_-^{(a)}$ exists if and only if $(RYS)\int_a^b f_+^{(a,b)}\,dh$ does, and if they exist, they satisfy*

$$(RYS)\int_a^b f_+^{(a,b)}\,dh = (RRS)\int_a^b f_+^{(b)}\,dh_-^{(a)} - [\Delta^+ f\Delta^+ h](a) + [f\Delta^- h](b). \tag{6.18}$$

The integral $(RRS)\int_a^b f_-^{(a)}\,dh_+^{(b)}$ exists if and only if $(RYS)\int_a^b f_-^{(a,b)}\,dh$ does, and if they exist, then

$$(RYS)\int_a^b f_-^{(a,b)}\,dh = (RRS)\int_a^b f_-^{(a)}\,dh_+^{(b)} + [f\Delta^+ h](a) + [\Delta^- f\Delta^- h](b). \tag{6.19}$$

Proof. By Theorem 6.13 and Lemma 6.14, the integral $(RRS)\int_a^b f_+^{(b)}\,dh_-^{(a)}$ exists if and only if $(RYS)\int_a^b f_+^{(b)}\,dh_-^{(a)}$ does, and then both have the same value. Let $\kappa = \{x_i\colon i = 0,\dots,n\} \in$ PP $([a,b])$ and let $\sigma = \{y_i\colon i = 1,\dots,n\}$ be an open intermediate partition of κ. Then

$$\begin{aligned} S_Y(f_+^{(b)}, h_-^{(a)}; \kappa,\sigma) &= \sum_{i=1}^n f(y_i+)[h(x_i-) - h(x_{i-1}+)] + [f_+\Delta^+ h](a) + \sum_{i=1}^{n-1}[f_+\Delta^\pm h](x_i) \\ &= S_Y(f_+^{(a,b)}, h; \kappa,\sigma) + [\Delta^+ f\Delta^+ h](a) - [f\Delta^- h](b). \end{aligned}$$

Therefore, $(RYS)\int_a^b f_+^{(b)}\,dh_-^{(a)}$ exists if and only if $(RYS)\int_a^b f_+^{(a,b)}\,dh$ does, and (6.18) holds. This proves the first statement. The proof of the second one is similar and is omitted. □

The following is an easy consequence of the preceding lemma.

Proposition 6.17. *For $f, h \in \mathcal{R}([a,b])$, the relations*

$$(Y_1)\int_a^b f\,dh = (RYS)\int_a^b f_+^{(a,b)}\,dh - \sum_{(a,b)} \Delta^+ f\Delta^\pm h \tag{6.20}$$

and

$$(Y_2)\int_a^b f\,dh = (RYS)\int_a^b f_-^{(a,b)}\,dh + \sum_{(a,b)} \Delta^- f\Delta^\pm h \tag{6.21}$$

hold; that is, the integral on the left side exists if and only if both the integral on the right side exists and the sum converges unconditionally.

For the proofs of Theorems 6.19 and 6.20, as well as for later use, we prove the following fact:

Lemma 6.18. *Let $f, h \in \mathcal{R}([a,b])$. The following three statements are equivalent:*
(1). *$(RYS)\int_a^b f\,dh$ exists;*
(2). *$(RYS)\int_a^b f_+^{(a,b)}\,dh$ and $(RYS)\int_a^b \Delta_{a,b}^+ f\,dh$ exist;*
(3). *$(RYS)\int_a^b f_-^{(a,b)}\,dh$ and $(RYS)\int_a^b \Delta_{a,b}^- f\,dh$ exist.*
If any one of the preceding three statements holds then

$$(RYS)\int_a^b f\,dh = (RYS)\int_a^b f_+^{(a,b)}\,dh - \sum_{(a,b)} \Delta^+ f \Delta^\pm h = (RYS)\int_a^b f_-^{(a,b)}\,dh + \sum_{(a,b)} \Delta^- f \Delta^\pm h,$$

where the two sums converge unconditionally.

Proof. We prove only the implication (1) $\Rightarrow$ (2), because the converse follows from linearity of the (RYS) integral and the proof of (2) $\Leftrightarrow$ (3) is similar. Suppose (1) holds. First we show that the sum

$$\sum_{(a,b)} \Delta^+ f \Delta^\pm h \qquad \text{converges unconditionally.} \tag{6.22}$$

Let $\epsilon > 0$. By hypothesis, there exists $\lambda \in \mathrm{PP}([a,b])$ such that

$$\Big| S_Y(f,h;\kappa,\sigma) - (RYS)\int_a^b f\,dh \Big| < \epsilon \tag{6.23}$$

for any refinement κ of λ and open intermediate partition σ of κ. Let $\nu = \{u_l : l = 1,\dots,k\}$ be a finite set of points in (a,b) disjoint from λ. Choose a refinement $\kappa = \{x_i : i = 0,\dots,n\}$ of λ such that for each $l = 1,\dots,k-1$, and some $i(l)$, $l = 1,\dots,k$, $x_{i(l)-1} < u_l < x_{i(l)} < \dots < x_{i(l+1)-1} < u_{l+1} < x_{i(l+1)}$, where $x_{i(l)-1}, x_{i(l)} \notin \lambda$. Let $\sigma' = \{y_i' : i = 1,\dots,n\}$ and $\sigma'' = \{y_i'' : i = 1,\dots,n\}$ be two open intermediate partitions of κ such that $y_{i(l)}' > y_{i(l)}'' = u_l$ for $l = 1,\dots,k$ and $y_i' = y_i'' \in (x_{i-1}, x_i)$ for each $i \notin \{i(1),\dots,i(k)\}$. By (6.23), it follows that

$$|S_Y(f,h;\kappa,\sigma') - S_Y,h;\kappa,\sigma'')| = \Big| \sum_{l=1}^k [f(y_{i(l)}') - f(u_l)][h(x_{i(l)}-) - h(x_{i(l)-1}+)] \Big| < 2\epsilon.$$

Letting $y_{i(l)}' \downarrow u_l$, $x_{i(l)-1} \uparrow u_l$ and $x_{i(l)} \downarrow u_l$, as is possible, one gets $|\sum_{l=1}^k \Delta^+ f(u_l)\Delta^\pm h(u_l)| \le 2\epsilon$ for any finite set $\nu = \{u_l : l = 1,\dots,k\} \subset (a,b)$ which is disjoint from λ. Since $\epsilon > 0$ is arbitrary one can conclude that (6.22) holds.

Next we show that $\Delta_{a,b}^+ f$ is (RYS) integrable with respect to h and

$$(RYS)\int_a^b \Delta_{a,b}^+ f\,dh = \sum_{(a,b)} \Delta^+ f \Delta^\pm h. \tag{6.24}$$

For each partition $\kappa = \{x_i : i = 0,1,\dots,n\} \in \mathrm{PP}([a,b])$ and open intermediate partition $\sigma = \{y_i : i = 1,\dots,n\}$ of κ, we have

$$\begin{aligned} \Big| S_Y(\Delta_{a,b}^+ f, h;\kappa,\sigma) - \sum_{(a,b)} \Delta^+ f \Delta^\pm h \Big| &\le \Big| \sum_{i=1}^n \Delta^+ f(y_i)[h(x_i-) - h(x_{i-1}+)] \Big| \\ + \quad \Big| \sum_{(a,b)} \Delta^+ f \Delta^\pm h - \sum_{i=1}^{n-1} [\Delta^+ f \Delta^\pm h](x_i) \Big| &=: T_1(\kappa,\sigma) + T_2(\kappa). \end{aligned} \tag{6.25}$$

Given $\epsilon > 0$ choose $\lambda \in \mathrm{PP}([a,b])$ such that (6.23) holds and $T_2(\kappa) < \epsilon$ for all refinements $\kappa = \{x_i\colon i = 0,1,\ldots,n\}$ of λ and all open intermediate partitions $\sigma = \{y_i\colon i = 1,\ldots,n\}$ of κ. Then

$$\Big|\sum_{i=1}^{n}[f(y_i') - f(y_i)][h(x_i-) - h(x_{i-1}+)]\Big| < 2\epsilon$$

whenever $y_i' \in (y_i, x_i)$ for $i = 1,\ldots,n$. Letting $y_i' \downarrow y_i$ for each $i = 1,\ldots,n$, yields the bound $T_1(\kappa,\sigma) \le 2\epsilon$. Therefore the right side of (6.25) does not exceed 3ϵ for each refinement κ of λ and each open intermediate partition σ of κ. Since $\epsilon > 0$ is arbitrary (6.24) holds. By linearity of the (RYS) integral, (2) follows and the last conclusion holds. $\square$

Theorem 6.19. *Let $f, h \in \mathcal{R}([a,b])$. The integral $(Y_1)\int_a^b f\,dh$ exists if and only if $(Y_2)\int_a^b f\,dh$ does, and then they are equal.*

Proof. Suppose the (Y_2) integral exists. By Proposition 6.17, $(RYS)\int_a^b f_-^{(a,b)}\,dh$ exists and (6.21) holds. By Lemma 6.18 applied to $f = f_-^{(a,b)}$, $(RYS)\int_a^b f_+^{(a,b)}\,dh$ exists and

$$(RYS)\int_a^b f_+^{(a,b)}\,dh = (RYS)\int_a^b f_-^{(a,b)}\,dh - \sum_{(a,b)}\Delta^{\pm} f\Delta^{\pm} h,$$

where the sum converges unconditionally. Also by linearity, the sum

$$\sum_{(a,b)}\Delta^{+} f\Delta^{\pm} h = \sum_{(a,b)}\Delta^{\pm} f\Delta^{\pm} h - \sum_{(a,b)}\Delta^{-} f\Delta^{\pm} h$$

converges unconditionally. Since the right side of (6.20) is defined, the (Y_1) integral exists and

$$\begin{aligned}(Y_2)\int_a^b f\,dh &= (RYS)\int_a^b f_-^{(a,b)}\,dh + \sum_{(a,b)}\Delta^{-} f\Delta^{\pm} h\\ &= (RYS)\int_a^b f_+^{(a,b)}\,dh - \sum_{(a,b)}\Delta^{\pm} f\Delta^{\pm} h + \sum_{(a,b)}\Delta^{-} f\Delta^{\pm} h\\ &= (RYS)\int_a^b f_+^{(a,b)}\,dh - \sum_{(a,b)}\Delta^{+} f\Delta^{\pm} h = (Y_1)\int_a^b f\,dh.\end{aligned}$$

The proof of the converse implication is symmetric and we omit it. $\square$

Theorem 6.20. *For $f, h \in \mathcal{R}([a,b])$, the integral $(RYS)\int_a^b f\,dh$ exists if and only if the two integrals $(Y_2)\int_a^b f\,dh$ and $(RYS)\int_a^b \Delta_{a,b}^{-} f\,dh$ exist. Also, the relations*

$$(RYS)\int_a^b f\,dh = (Y_2)\int_a^b f\,dh \quad \textit{and} \quad (RYS)\int_a^b \Delta_{a,b}^{-} f\,dh = \sum_{(a,b)}\Delta^{-} f\Delta^{\pm} h \tag{6.26}$$

hold with the sum converging unconditionally, whenever the three integrals exist.

Proof. Suppose that $(Y_2)\int_a^b f\,dh$ and $(RYS)\int_a^b \Delta_{a,b}^{-} f\,dh$ exist. Then by Proposition 6.17, $(RYS)\int_a^b f_-^{(a,b)}\,dh$ exists. Thus $(RYS)\int_a^b f\,dh$ exists by linearity of the (RYS) integral. The relations (6.26) then follow from (6.21).

Suppose that $(RYS) \int_a^b f\,dh$ exists. By statement (3) of Lemma 6.18, $(RYS) \int_a^b \Delta_{a,b}^- f\,dh$ and the right side of (6.21) exist. Thus $(Y_2) \int_a^b f\,dh$ exists and equals $(RYS) \int_a^b f\,dh$ by Proposition 6.17. The proof of Theorem 6.20 is complete. □

Notation. For a set E of real numbers, let $c_0(E)$ be the set of all functions $f\colon E \mapsto \mathbb{R}$ for which there exists a set $\{\xi_i\colon i \geq 1\} \subset E$ of distinct numbers such that $f(\xi_i) \to 0$ as $i \to \infty$ and $f(x) = 0$ if $x \in E \setminus \{\xi_i\colon i \geq 1\}$. Then

$$c_0(E) = \Big\{ \sum_i f_i 1_{\{\xi_i\}}\colon\ \xi_i \neq \xi_j \text{ for } i \neq j,\, \xi_i \in E,\, f_i \to 0 \text{ as } i \to \infty \Big\}. \tag{6.27}$$

If $f \in c_0(J)$ for an open interval J then $f \in \mathcal{R}(J)$ with $f_+ \equiv f_- \equiv 0$. If $f \in c_0((a,b))$ then $f_+(a) = f_-(b) = 0$. Thus defining $f(a) = f(b) := 0$, we have $f_-^{(a)} \equiv f_+^{(b)} \equiv 0$ on $[a,b]$.

Proposition 6.21. *There exist functions f and h on $[0,1]$, where $f \in c_0((0,1))$ and h is continuous, for which $(Y_2) \int_0^1 f\,dh$ exists, while $(RYS) \int_0^1 \Delta_{0,1}^- f\,dh$ and hence $(RYS) \int_a^b f\,dh$ don't exist.*

Proof. Let $f(x) := k^{-1/2}$ if $x = 1/(3k)$ for $k = 1, 2, \ldots$, and $f(x) := 0$ otherwise. Let $h(1/(3k+1)) := 0$ and $h(1/(3k-1)) := k^{-1/2}$ for $k = 1, 2, \ldots$. Let $h(0) = h(1) := 0$ and let h be "linear in between", i.e. on each closed interval where h is so far defined only at the endpoints. Since h is continuous and $f_-^{(0)} \equiv f_+^{(1)} \equiv 0$, $(Y_2) \int_0^1 f\,dh$ exists and is 0. Since $f_-^{(0,1)} \equiv 0$ it is enough to show the non-existence of the integral $(RYS) \int_0^1 f\,dh$. Let $\lambda = \{z_j\colon j = 0, \ldots, m\}$ be any partition of $[0,1]$. Take the smallest m such that $x_m := 1/(3m+1) \leq z_1$. If $\kappa_1 := \lambda \cup \{x_m\}$ then the contribution to any Young sum based on κ_1 coming from $[0, x_m]$ is 0. For $n > m$, consider partitions

$$\kappa_n := \lambda \cup \{1/(3k+1), 1/(3k-1)\colon k = m+1, \ldots, n\}.$$

We form a Young sum based on κ_n by letting it be the same as one for κ_1 on $[x_m, 1]$, and by evaluating f at $1/(3k)$ for $k = m+1, \ldots, n$. Thus the part of our Young sum for κ_n coming from $[0, x_m]$ is

$$\sum_{k=m+1}^{n} f\Big(\frac{1}{3k}\Big)\Big[h\Big(\frac{1}{3k-1}\Big) - h\Big(\frac{1}{3k+1}\Big)\Big] + 0 \cdot \Big[h\Big(\frac{1}{3k-2}\Big) - h\Big(\frac{1}{3k+1}\Big)\Big] = \sum_{k=m+1}^{n} k^{-1/2} k^{-1/2} \to \infty$$

as $n \to \infty$. Thus two Young sums for $(RYS) \int_0^1 f\,dh$, both based on refinements of λ, differ by an arbitrarily large amount. So $(RYS) \int_0^1 f\,dh$ doesn't exist. □

For a set E and $0 < p < \infty$, let

$$\ell_p(E) := \Big\{ f = \sum_i f_i 1_{\{\xi_i\}} \in c_0(E)\colon \sum_i |f_i|^p < \infty \Big\}.$$

Proposition 6.22. *Let $f, h \in \mathcal{R}([a,b])$ be such that $\Delta^- f \in \ell_p((a,b))$ and $h \in \mathcal{W}_{p'}([a,b])$ for some p, p' with $1/p + 1/p' = 1$, $1 \leq p \leq \infty$. Then $(RYS) \int_a^b f\,dh$ exists if and only if $(Y_2) \int_a^b f\,dh$ does, and the two integrals are equal.*

Proof. The proof is given only for the case $p < \infty$ because a proof for the case $p = \infty$ is similar. By Theorem 6.20, it is enough to prove that $(RYS) \int_a^b \Delta_{a,b}^- f\,dh$ exists. By Hölder's inequality, the sum $\sum_{(a,b)} \Delta^- f \Delta^\pm h$ converges unconditionally. Let $\epsilon > 0$. Then there exists a finite set $\lambda_1 \subset (a,b)$ such that

$$\Big| \sum_{x\in\mu} [\Delta^- f \Delta^\pm h](x) - \sum_{(a,b)} \Delta^- f \Delta^\pm h \Big| < \epsilon$$

for any set $\mu \subset (a,b)$ containing λ_1. Since $\Delta^- f \in \ell_p((a,b))$ and $h \in \mathcal{W}_{p'}([a,b])$, there exists a finite set $\lambda_2 \subset (a,b)$ such that

$$\sum_{y\in\nu} |\Delta^- f(y)|^p < \Big(\epsilon / V_{p'}(h;[a,b])\Big)^p$$

for any set $\nu \subset (a,b)$ disjoint from λ_2. Let $\lambda := \{a,b\} \cup \lambda_1 \cup \lambda_2$. Then $\lambda \in \mathrm{PP}\,([a,b])$. For any refinement $\kappa = \{x_i\colon i = 0,\dots,n\}$ of λ and open intermediate partition $\sigma = \{y_i\colon i = 1,\dots,n\}$ of κ, we then have

$$\Big| S_Y(\Delta^-_{a,b} f, h; \kappa, \sigma) - \sum_{(a,b)} \Delta^- f \Delta^\pm h \Big| \le \Big| \sum_{i=1}^{n-1} [\Delta^- f \Delta^\pm h](x_i) - \sum_{(a,b)} \Delta^- f \Delta^\pm h \Big|$$

$$+ \Big| \sum_{i=1}^{n} \Delta^- f(y_i)[h(x_i-) - h(x_{i-1}+)] \Big| \le \epsilon + \Big(\sum_{i=1}^{n} |\Delta^- f(y_i)|^p \Big)^{1/p} V_{p'}(h;[a,b]) < 2\epsilon.$$

Therefore the integral $(RYS) \int_a^b \Delta^-_{a,b} f\, dh$ exists, and the proof is complete. □

6.4 The $(\mathfrak{z})$–integral of A. N. Kolmogorov

Kolmogorov [63] suggested a unified approach to a class of integrals with respect to interval or set functions. The Kolmogorov integral includes the Lebesgue, Lebesgue-Stieltjes and several other integrals. Some of them are considered below. We recall and use Kolmogorov's definition (Definition 6.24 below) in the case when only finite partitions are used.

Let $\mathfrak{M}$ be a multiplicative system of sets; that is, $J_1 \cap J_2 \in \mathfrak{M}$ whenever $J_1, J_2 \in \mathfrak{M}$. For $J \in \mathfrak{M}$, a vector $\mathfrak{D} = (J_1,\dots,J_N)$ of pairwise disjoint sets in $\mathfrak{M}$ will be called an *$\mathfrak{M}$-partition of J* if $J = \cup_{i=1}^N J_i$. The sets $J_1,\dots,J_N$ are called *elements of the $\mathfrak{M}$-partition* $\mathfrak{D}$ and written $J_i \in \mathfrak{D}$, $i = 1,\dots,N$. Let $\mathrm{IP}\,(\mathfrak{M})$ be the set of all $\mathfrak{M}$-partitions $\mathfrak{D}$ of J. For example, if $\mathfrak{M}$ consists of a single set $J \ne \emptyset$ then $\mathrm{IP}\,(\mathfrak{M})$ consists of the single partition of J into itself. An $\mathfrak{M}$-partition $\mathfrak{D}''$ is an *$\mathfrak{M}$-refinement* of an $\mathfrak{M}$-partition $\mathfrak{D}'$ if, for each element J'' of $\mathfrak{D}''$, $J'' \subset J'$ for some element J' of $\mathfrak{D}'$. In this case we write $\mathfrak{D}' \sqsubseteq \mathfrak{D}''$. Then $(\mathrm{IP}\,(\mathfrak{M}), \sqsubseteq)$ is a partially ordered set (a poset). For $\mathfrak{D}', \mathfrak{D}'' \in \mathrm{IP}\,(\mathfrak{M})$ let $\mathfrak{D} := \{J' \cap J''\colon J' \in \mathfrak{D}', J'' \in \mathfrak{D}''\}$. Then $\mathfrak{D}' \sqsubseteq \mathfrak{D}$, $\mathfrak{D}'' \sqsubseteq \mathfrak{D}$ and $\mathfrak{D} \in \mathrm{IP}\,(\mathfrak{M})$ because $\mathfrak{M}$ is multiplicative. Therefore the poset $(\mathrm{IP}\,(\mathfrak{M}), \sqsubseteq)$ is upwards directed.

Definition 6.23. Let S be a function on the set of $\mathfrak{M}$-partitions of $J \in \mathfrak{M}$ whose values are sets of real numbers. We say that a real number A is the limit of S *under refinements of $\mathfrak{M}$-partitions of J* and write $A = \lim_{\mathfrak{D}\uparrow} S(\mathfrak{D})$, if for each $\epsilon > 0$ there is an $\mathfrak{M}$-partition $\mathfrak{D}$ of J such that for all $\mathfrak{D}' \sqsupseteq \mathfrak{D}$, we have $\sup\{|x - A|\colon x \in S(\mathfrak{D}')\} < \epsilon$.

For $J \in \mathfrak{M}$, let $\mathfrak{M}(J)$ be the collection of elements of all partitions $\mathfrak{D}$ of J. Then $\mathfrak{M}(J)$ is a multiplicative system. For a function ν defined on all elements of an $\mathfrak{M}$-partition $\mathfrak{D} = (J_1, \ldots, J_n)$, let

$$S_{\mathfrak{M}}(\nu; \mathfrak{D}) := \sum_{i=1}^{n} \nu(J_i).$$

Definition 6.24. Let ν be a function on $\mathfrak{M}(J)$ for some $J \in \mathfrak{M}$ whose values are sets of real numbers. The $(\mathfrak{M})$*-integral of* ν *on* J will be defined by

$$(\mathfrak{M}) \int_J d\nu := \lim_{\mathfrak{D}\uparrow} S_{\mathfrak{M}}(\nu; \mathfrak{D})$$

provided the limit of $S_{\mathfrak{M}}(\nu; \cdot)$ exists under refinements of $\mathfrak{M}$-partitions of J.

We refer to Kolmogorov [63] for elementary properties of the $(\mathfrak{M})$-integral.

Recall that in Section 4.5 we considered two classes of subintervals of $[a,b]$: $\mathfrak{I}_{os}([a,b]) := \{(c,d), \{c\} \colon a \le c \le d \le b\}$ and $\mathfrak{I}([a,b]) := \{[c,d], (c,d], [c,d), (c,d) \colon a \le c \le d \le b\}$. Next we shall consider the $(\mathfrak{M})$-integral when $\mathfrak{M}(J) = \mathfrak{I}_{os}([a,b])$ or $\mathfrak{M}(J) = \mathfrak{I}([a,b])$, and ν is a point function f multiplied by an additive interval function μ.

Let $\mathfrak{I}$ be either of the two classes $\mathfrak{I}_{os}([a,b])$ or $\mathfrak{I}([a,b])$. These classes possess the following decomposability property (used in Chapter 3, §2 by Kolmogorov [63]): if $J', J \in \mathfrak{I}$ and $J' \subset J$ then there is an $\mathfrak{I}$-partition $\mathfrak{D}$ of J such that J' is an element of $\mathfrak{D}$. Let $f \colon [a,b] \mapsto \mathbb{R}$ be a point function and let $\mu \in \mathcal{I}[a,b]$ be additive. For each $\mathfrak{D} = (J_1, \ldots, J_N) \in \mathrm{IP}\,(\mathfrak{I})$ and $\theta = (y_1, \ldots, y_N) \in J_1 \times \cdots \times J_N$, let

$$S_{\mathfrak{I}}(f, \mu; \mathfrak{D})(\theta) := \sum_{j=1}^{N} f(y_j)\mu(J_j).$$

According to Definition 6.24, the $(\mathfrak{I})$–integral of f with respect to an additive interval function μ on $[a,b]$ is defined by

$$(\mathfrak{I}) \int_{[a,b]} f\, d\mu := \lim_{\mathfrak{D}\uparrow} S_{\mathfrak{I}}(f, \mu; \mathfrak{D})$$

provided the limit exists under refinements of $\mathfrak{I}$–partitions of $[a,b]$. Thus the $(\mathfrak{I})$-integral with $\mathfrak{I} = \mathfrak{I}([a,b])$ is the same as the (RYS) integral and the class of partitions $\mathrm{IP}\,(\mathfrak{I}([a,b]))$ is the same as $\mathrm{IP}\,([a,b])$, both defined in Section 2.1.

Next we describe a relation between $\mathfrak{I}_{os}$–partitions from $\mathrm{IP}_{os}([a,b]) := \mathrm{IP}\,(\mathfrak{I}_{os}([a,b]))$ and corresponding point partitions of $[a,b]$. Let $a = x_0 < x_1 < \cdots < x_n = b$. Then $\kappa = \{x_i \colon i = 0, 1, \ldots, n\}$ is a partition of $[a,b]$ by points and $\mathrm{PP}\,([a,b])$ is the set of all point partitions κ of $[a,b]$. For each $\mathfrak{D} = (J_1, \ldots, J_N) \in \mathrm{IP}_{os}([a,b])$, we have $J_1 = \{a\}$, $J_2 = (a, x_1)$, $J_3 = \{x_1\}$, $J_4 = (x_1, x_2)$, $\ldots$, $J_{2n} = (x_{n-1}, b)$, $J_{2n+1} = \{b\}$ and $N = 2n+1$. Let $\wp$ be the mapping from $\mathrm{IP}_{os}([a,b])$ into $\mathrm{PP}\,([a,b])$ defined by $\wp(\mathfrak{D}) := \{x_0, x_1, \ldots, x_n\} \in \mathrm{PP}\,([a,b])$. Then $\wp$ is an order isomorphism between $(\mathrm{IP}_{os}([a,b]), \sqsubseteq)$ and $(\mathrm{PP}\,([a,b]), \subseteq)$, that is, for each $\mathfrak{D}_1, \mathfrak{D}_2 \in \mathrm{IP}_{os}([a,b])$, $\mathfrak{D}_1 \sqsubseteq \mathfrak{D}_2$ if and only if $\wp(\mathfrak{D}_1) \subseteq \wp(\mathfrak{D}_2)$.

Let $h \in \mathcal{R}[a,b]$, $\kappa = \{x_i \colon i = 0, \ldots, n\} \in \mathrm{PP}\,([a,b])$ and let $\sigma = \{y_i \colon i = 1, \ldots, n\}$ be an open intermediate partition of κ, meaning that $x_{i-1} < y_i < x_i$ for $i = 1, \ldots, n$. Recall that the Young sum is defined by

$$S_Y(f, h; \kappa, \sigma) := \sum_{i=1}^{n} \{[f\Delta^+ h](x_{i-1}) + f(y_i)[h(x_i-) - h(x_{i-1}+)] + [f\Delta^- h](x_i)\}.$$

Notice that the value of the Young sum is the same for each function $h \in [\mu_h]_a$ (cf. (4.34) above); that is, the value of the integral $(RYS) \int_a^b f\,dh$ does not depend on the values of h at its jumps. Also, the existence of $(RYS) \int_a^b f\,dh$ implies the existence of $(\mathfrak{I}_{os}) \int_{[a,b]} f\,d\mu_h$. This follows by comparing the Young sum $S_Y(f, h; \kappa, \sigma)$ with the sum $S_{\mathfrak{I}_{os}}(f, \mu_h; \mathfrak{D})(\theta)$, where $\wp(\mathfrak{D}) = \kappa$ and the coordinates of θ corresponding to the open intervals of $\mathfrak{D}$ are the coordinates of σ. We show next that the converse implication holds whenever an additive interval function μ is upper continuous (see Definition 4.32).

Theorem 6.25. *Let $\mu \in \mathcal{I}[a, b]$ be additive. The following statements about μ are equivalent:*

(*a*) *μ is upper continuous;*

(*b*) *for any $f\colon [a, b] \to \mathbb{R}$, if $(\mathfrak{I}_{os}) \int_{[a,b]} f\,d\mu$ exists then so does $(RYS) \int_a^b f\,dh$ for each $h \in [\mu]_a$, and the two integrals are equal;*

(*c*) *there exists $h \in \mathcal{R}[a, b]$ with $h(a) = 0$ such that, for any $f\colon [a, b] \to \mathbb{R}$, $(RYS) \int_a^b f\,dh$ exists whenever $(\mathfrak{I}_{os}) \int_{[a,b]} f\,d\mu$ does, and the two integrals are equal.*

Proof. (*a*) $\Rightarrow$ (*b*). Let $\mathfrak{D} = (J_1, \ldots, J_N) \in \mathrm{IP}_1([a, b])$ and $J_{2i} = (x_{i-1}, x_i)$ for $i = 1, \ldots, n$ with $N = 2n + 1$. Then $\wp\colon \mathfrak{D} \mapsto \wp(\mathfrak{D}) = \{x_i\colon i = 0, 1, \ldots, n\} =: \kappa \in \mathrm{PP}([a, b])$ is an order isomorphism between $\mathrm{IP}_{os}([a, b])$ and $\mathrm{PP}([a, b])$. Let $\theta = \{x_0, y_1, x_1, \ldots, y_n, x_n\}$, where $y_i \in (x_{i-1}, x_i)$ for $i = 1, \ldots, n$. Since μ is upper continuous and $h \in [\mu]_a$, by Theorem 4.36, we have $\mu((x_{i-1}, x_i)) = \mu_h((x_{i-1}, x_i)) = h(x_i-) - h(x_{i-1}+)$ for $i = 1, \ldots, n$, $\mu(\{a\}) = h(a+)$, $\mu(\{x_i\}) = \mu_h(\{x_i\}) = h(x_i+) - h(x_i-)$ for $i = 1, \ldots, n-1$ and $\mu(\{b\}) = h(b) - h(b-)$. Then we have

$$S_{\mathfrak{I}_1}(f, \mu; \mathfrak{D})(\theta) = \sum_{i=1}^{n} f(y_i)\mu((x_{i-1}, x_i)) + \sum_{i=0}^{n} f(x_i)\mu(\{x_i\}) = S_Y(f, h; \kappa, \sigma),$$

where $\sigma = (y_1, \ldots, y_n)$. Therefore $(RYS) \int_a^b f\,dh$ exists whenever $(\mathfrak{I}_{os}) \int_{[a,b]} f\,d\mu$ does, and the two integrals are equal.

(*b*) $\Rightarrow$ (*c*). This implication is clear because $h \in \mathcal{R}[a, b]$ and $h(a) = 0$ for each $h \in [\mu]_a$ by definition (4.34).

(*c*) $\Rightarrow$ (*a*). Let $a \le u < v \le b$. Then we have

$$\mu_h((u, v)) = h(v-) - h(u+) = (RYS) \int_a^b 1_{(u,v)}\,dh = (\mathfrak{I}_{os}) \int_{[a,b]} 1_{(u,v)}\,d\mu = \mu((u, v)).$$

Similarly, it follows that $\mu_h(\{w\}) = \mu(\{w\})$ for $w \in [a, b]$. The implication then follows from Theorem 4.36. The proof is complete. $\square$

Since $\mathrm{IP}_{os}([a, b]) \subset \mathrm{IP}([a, b])$ and each $\mathfrak{I}$-partition can be $\mathfrak{I}_{os}$-refined, if $(\mathfrak{I}) \int_{[a,b]} f\,d\mu$ exists then so does $(\mathfrak{I}_{os}) \int_{[a,b]} f\,d\mu$, and the two integrals are equal. We show next that the converse holds for upper continuous additive interval functions.

Theorem 6.26. *Let $\mu \in \mathcal{I}[a, b]$ be additive and upper continuous, and let $f\colon [a, b] \to \mathbb{R}$. If $(\mathfrak{I}_{os}) \int_{[a,b]} f\,d\mu$ exists then so does $(\mathfrak{I}) \int_{[a,b]} f\,d\mu$, and the two are equal.*

Proof. Let $\epsilon > 0$. There exists $\mathfrak{D}_0 \in \mathrm{IP}_{os}([a, b]) \subset \mathrm{IP}([a, b])$ such that

$$|S_{\mathfrak{I}_{os}}(f, \mu; \mathfrak{D})(\theta) - (\mathfrak{I}_{os}) \int_{[a,b]} f\,d\mu| \le \epsilon \tag{6.28}$$

for all $\mathfrak{I}_{os}$-refinements $\mathfrak{D} = \{J_1, \ldots, J_N\}$ of $\mathfrak{D}_0$ and all $\theta \in J_1 \times \cdots \times J_N$. It will be shown that (6.28) also holds for $\mathfrak{I}$ in place of $\mathfrak{I}_{os}$, with the integral having the same value.

Let $\mathfrak{D}_1 = (J_1, \ldots, J_N)$ be an $\mathfrak{I}$-refinement of $\mathfrak{D}_0$, $\theta = (y_1, \ldots, y_N) \in J_1 \times \cdots \times J_N$ and let $\lambda := \wp(\mathfrak{D}_0) = \{z_0, \ldots, z_L\} \in \mathrm{PP}\,([a,b])$. Let $\kappa := \{x_i \colon i = 0, \ldots, m\}$ be the set of endpoints of intervals in $\mathfrak{D}_1$. So $\kappa \in \mathrm{PP}\,([a,b])$. For each $i = 0, 1, \ldots, m$, define a sequence $\{x_{in} \colon n \geq 1\}$ as follows. If $\{x_i\}$ is a singleton in $\mathfrak{D}_1$ then let $x_{in} := x_i$ for all $n \geq 1$. If an interval $\{\cdot, x_i] \in \mathfrak{D}_1$ let $x_{in} \downarrow x_i$ or if some $[x_i, \cdot\} \in \mathfrak{D}_1$ let $x_{in} \uparrow x_i$, where in either case x_{in} are not atoms of μ. This can be done by statement (e) of Theorem 4.35. For n large enough, $x_{0n} = a < x_{1n} < \cdots < x_{mn} = b$. Then there is a unique $\mathfrak{I}_{os}$-partition $\mathfrak{D}_{2,n}$ having $\wp(\mathfrak{D}_{2,n}) = \{x_{in} \colon i = 0, \ldots, m\} \in \mathrm{PP}\,([a,b])$. Since each singleton $\{z_j\} = \{x_i\} \in \mathfrak{D}_1$ for some i, $\mathfrak{D}_{2,n}$ is a refinement of $\mathfrak{D}_0$. The intervals in $\mathfrak{D}_{2,n}$ converge to the intervals in $\mathfrak{D}_1$ as $n \to \infty$, except for singletons in $\mathfrak{D}_{2,n}$ with $\mu = 0$ which do not contribute to sums. Each point y_i is eventually in the interval of $\mathfrak{D}_{2,n}$ corresponding to the interval of $\mathfrak{D}_1$. Thus there are sums $S_{\mathfrak{I}_{os}}(f, \mu; \mathfrak{D}_{2,n})(\theta_n)$ converging to $S_{\mathfrak{I}}(f, \mu; \mathfrak{D}_1)(\theta)$, where the coordinates of θ_n are those of θ, and possibly some y_{in} with $\{y_{in}\} \in \mathfrak{D}_{2,n}$ and $\mu(\{y_{in}\}) = 0$. Theorem 6.26 then follows. □

6.5 The Ward–Perron–Stieltjes and gauge integrals

The Henstock integral is designed to integrate highly oscillatory functions which the Lebesgue integral fails to do. It is known as nonabsolute integration and is a powerful tool.

– Lee Peng–Yee [89, p. vii].

If we think of the Lebesgue integral as God sent, then differentiable functions whose derivatives are not Lebesgue integrable may appear evil. The challenge is to resolve the conflict ...

– unattributed aphorism in Pfeffer [90, p. xii].

Ward (1936) defined a Perron–Stieltjes integral which includes both the Lebesgue–Stieltjes and Moore–Pollard–Stieltjes (RRS) integrals. Given two real-valued functions f, g on $[a,b]$, say M is a *major function* of f with respect to g if $M(a) = 0$, M has finite values on $[a,b]$, and for each $x \in [a,b]$ there exists $\delta(x) > 0$ such that

$$\begin{cases} M(z) \geq M(x) + f(x)[g(z) - g(x)] & \text{if} \quad x \leq z \leq \min\{b, x + \delta(x)\}, \\ M(z) \leq M(x) + f(x)[g(z) - g(x)] & \text{if} \quad \max\{a, x - \delta(x)\} \leq z \leq x. \end{cases} \tag{6.29}$$

Thus for each major function M there exists a positive function $\delta_M(\cdot) = \delta(\cdot)$ on $[a,b]$ which satisfies (6.29). Let $\mathcal{U}(f,g)$ be the class of all major functions of f with respect to g, and let

$$U(f,g) := \begin{cases} \inf\{M(b) \colon M \in \mathcal{U}(f,g)\} & \text{if } \mathcal{U}(f,g) \neq \emptyset, \\ +\infty & \text{if } \mathcal{U}(f,g) = \emptyset. \end{cases}$$

A function m is a *minor function* of f with respect to g if $-m \in \mathcal{U}(-f, g)$. Let $\mathcal{L}(f,g)$ be the class of all minor functions of f with respect to g. Thus $m \in \mathcal{L}(f,g)$ provided $m(a) = 0$, m has finite values on $[a,b]$, and there exists a positive function $\delta(\cdot) = \delta_m(\cdot)$ on $[a,b]$ such that

$$\begin{cases} m(z) \leq m(x) + f(x)[g(z) - g(x)] & \text{if} \quad x \leq z \leq \min\{b, x + \delta(x)\}, \\ m(z) \geq m(x) + f(x)[g(z) - g(x)] & \text{if} \quad \max\{a, x - \delta(x)\} \leq z \leq x. \end{cases} \tag{6.30}$$

Let

$$L(f,g) := \begin{cases} \sup\{m(b)\colon m \in \mathcal{L}(f,g)\} & \text{if } \mathcal{L}(f,g) \neq \emptyset, \\ -\infty & \text{if } \mathcal{L}(f,g) = \emptyset. \end{cases}$$

Then the following is true:

Lemma 6.27. *$L(f,g) \leq U(f,g)$.*

Proof. The statement is true if at least one of the sets $\mathcal{L}(f,g)$ or $\mathcal{U}(f,g)$ is empty. So suppose that $\mathcal{L}(f,g)$ and $\mathcal{U}(f,g)$ are not empty. Let $m \in \mathcal{L}(f,g)$ and $M \in \mathcal{U}(f,g)$. Also, let $\delta := \delta_m \wedge \delta_M$ and $w(x) := M(x) - m(x)$ for $x \in [a,b]$. By (6.29) and (6.30), for each $x \in (a,b]$, it follows that

$$\begin{cases} w(z) \geq w(x) & \text{if} \quad x \leq z \leq \min\{b, x+\delta(x)\}, \\ w(z) \leq w(x) & \text{if} \quad \max\{a, x-\delta(x)\} \leq z \leq x. \end{cases}$$

Therefore $\inf\{M(x)\colon M \in \mathcal{U}(f,g)\} - \sup\{m(x)\colon m \in \mathcal{L}(f,g)\}$ is a nondecreasing function of x. Then the statement of the lemma holds because $m(a) = M(a) = 0$. □

If $U(f,g) = L(f,g)$ is finite then denote the common value by $(WPS)\int_a^b f\,dg$ and call it the *Ward–Perron–Stieltjes integral*, or the (WPS) integral. The following is Theorem 5 of Ward [111]. Note that in it, f and g are not necessarily regulated.

Theorem 6.28. *If $(RRS)\int_a^b f\,dg$ exists then so does $(WPS)\int_a^b f\,dg$, and has the same value.*

Proof. For $a \leq u < v \leq b$, let $M(u,v)$ be the least upper bound for Riemann–Stieltjes sums $S_{RS}(f,g;\kappa,\sigma)$ based on partitions $\kappa \in \mathrm{PP}([u,v])$ and intermediate partitions σ of κ. Then

$$\begin{cases} M(u,y) \geq M(u,x) + f(x)[g(y)-g(x)] & \text{if} \quad u < x < y, \\ M(u,y) \leq M(u,x) + f(x)[g(y)-g(x)] & \text{if} \quad u < y < x. \end{cases} \tag{6.31}$$

Suppose $(RRS)\int_a^b f\,dg$ exists and $\epsilon > 0$. Then exists $\lambda = \{z_j\colon j = 0,\dots,m\} \in \mathrm{PP}([a,b])$ such that

$$(RRS)\int_a^b f\,dg - \epsilon < S_{RS}(f,g;\kappa,\sigma) < (RRS)\int_a^b f\,dg + \epsilon \tag{6.32}$$

for each refinement κ of λ and intermediate partition σ of κ. For each $x \in (a,b]$, let $j(x)$ be the greatest integer such that $z_{j(x)} \leq x$. Let $M(a) := 0$, and for $x \in (a,b]$ let

$$M(x) := \sum_{j=1}^{j(x)} M(z_{j-1}, z_j) + M(z_{j(x)}, x),$$

where the sum over the empty set is 0. Define a positive function $\delta(\cdot)$ on $[a,b]$ by $\delta(x) := \min\{x - z_{j(x)}, z_{j(x)+1} - x\}$ if $x \in (z_{j(x)}, z_{j(x)+1})$, $\delta(a) := z_1 - a$, $\delta(b) := b - z_{m-1}$ and $\delta(z_j) := \min\{z_j - z_{j-1}, z_{j+1} - z_j\}$ for $j = 1,\dots,m-1$. By (6.31), the function M so defined satisfies (6.29). Hence it is a major function of f with respect to g. By (6.32), it follows that $M(b) \leq (RRS)\int_a^b f\,dg + \epsilon$. Similarly one can define a minor function m such that $m(b) \geq (RRS)\int_a^b f\,dg - \epsilon$. Since ϵ is arbitrarily small this proves the theorem. □

Kurzweil [68, Section 1.2] suggested an equivalent definition of the (WPS) integral based on an extension of the limit as the mesh of partitions tends to zero in the definition of the RS integral. For a partition $a = x_0 < x_1 < \cdots < x_n = b$, points $y_i \in [x_{i-1}, x_i]$, $i = 1,\dots,n$, are called *tags* and the pair of sets $\tau = (\{x_i\colon i = 0, \dots, n\}, \{y_i\colon i = 1, \dots, n\})$ is called a *tagged partition* of $[a, b]$. A *gauge function* is any function with strictly positive values. Given a gauge function $\delta(\cdot)$ on $[a, b]$, a tagged partition $\tau = (\{x_i\colon i = 0, \dots, n\}, \{y_i\colon i = 1, \dots, n\})$ is *δ-fine* if $y_i - \delta(y_i) \le x_{i-1} \le y_i \le x_i \le y_i + \delta(y_i)$ for $i = 1, \dots, n$. Let $TP(\delta, [a, b])$ be the set of all δ–fine tagged partitions of $[a, b]$.

Lemma 6.29. *For any gauge function δ, $TP(\delta, [a, b])$ is nonempty.*

Proof. Let $\delta(\cdot)$ be a gauge function on $[a, b]$. Then the system of open intervals $\{(y - \delta(y), y + \delta(y))\colon y \in [a, b]\}$ is an open cover of $[a, b]$. Since $[a, b]$ is compact there is a finite subcover $\{J_i := (y_i - \delta(y_i), y_i + \delta(y_i)\colon i = 1, \dots, n\}$ of $[a, b]$ such that any $n - 1$ of J_i do not cover $[a, b]$. Since y_i, $i = 1, \dots, n$ are different we can assume that $y_1 < \cdots < y_n$. Also, $J_i \cap J_{i+1} \ne \emptyset$ for $i = 1, \dots, n-1$, $J_i \cap J_j = \emptyset$ for $|i - j| > 1$, $i, j = 1, \dots, n$, $a \in J_1$ and $b \in J_n$. Hence one can find numbers $a = x_0 < x_1 < \ldots < x_n = b$ such that $x_i \in J_i \cap J_{i+1}$ for $i = 1, \dots, n-1$ and $y_i \in [x_{i-1}, x_i]$ for $i = 1, \dots, n$. Thus $(\{x_i\colon i = 0, \dots, n\}, \{y_i\colon i = 1, \dots, n\}) \in TP(\delta, [a, b])$. $\square$

The sum

$$S_\delta(\tau) := S_\delta(f, g; \tau) := \sum_{i=1}^{n} f(y_i)[g(x_i) - g(x_{i-1})]$$

based on a δ–fine tagged partition $\tau = (\{x_i\colon i = 0, \dots, n\}, \{y_i\colon i = 1, \dots, n\})$ will be called the *gauge sum* based on τ. The *gauge* or *Henstock-Kurzweil* integral $(HK) \int_a^b f\,dg$ is defined as the number A, whenever it exists, such that for each $\epsilon > 0$ there is a gauge function $\delta(\cdot)$ on $[a, b]$ such that $|S_\delta(f, g; \tau) - A| < \epsilon$ for each δ-fine tagged partition τ of $[a, b]$. Kurzweil [68, Theorem 1.2.1] proved the following statement:

Theorem 6.30. *The integral $(WPS) \int_a^b f\,dg$ exists if and only if $(HK) \int_a^b f\,dg$ exists, and then their values are equal.*

Proof. Suppose $(WPS) \int_a^b f\,dg$ is defined. Then given $\epsilon > 0$, there exist a major function $M \in \mathcal{U}(f, g)$ and a minor function $m \in \mathcal{L}(f, g)$ such that

$$M(b) - \epsilon < (WPS) \int_a^b f\,dg < m(b) + \epsilon. \tag{6.33}$$

Let $\delta := \delta_M \wedge \delta_m$ and let $\tau = (\{x_i\colon i = 0, \dots, n\}, \{y_i\colon i = 1, \dots, n\})$ be any δ–fine tagged partition of $[a, b]$, which exists by Lemma 6.29. By (6.29), we have

$$M(x_i) - M(y_i) \ge f(y_i)[g(x_i) - g(y_i)] \quad \text{and} \quad M(y_i) - M(x_{i-1}) \ge f(y_i)[g(y_i) - g(x_{i-1})] \tag{6.34}$$

for each $i = 1, \dots, n$. Adding up $2n$ inequalities (6.34) we get the upper bound $M(b) \ge S_\delta(\tau)$ for the gauge sum $S_\delta(\tau)$. Similarly, using (6.30), we get the lower bound $S_\delta(\tau) \ge m(b)$. By (6.33), it then follows that

$$(WPS) \int_a^b f\,dg - \epsilon < S_\delta(\tau) < (WPS) \int_a^b f\,dg + \epsilon.$$

Thus $(HK) \int_a^b f\,dg$ exists with the same value as the (WPS) integral.

Now suppose $(HK) \int_a^b f\,dg$ is defined. Given $\epsilon > 0$ there exists a gauge function $\delta(\cdot)$ such that for each δ–fine tagged partition τ,

$$(HK) \int_a^b f\,dg - \epsilon < S_\delta(\tau) < (HK) \int_a^b f\,dg + \epsilon. \tag{6.35}$$

For each $x \in (a, b]$, let δ_x be the gauge function δ restricted to the interval $[a, x]$. Then for each $x \in (a, b]$, let

$$m(x) := \inf \{S_{\delta_x}(\tau): \tau \in TP(\delta_x, [a, x])\}, \qquad M(x) := \sup \{S_{\delta_x}(\tau): \tau \in TP(\delta_x, [a, x])\}$$

and $m(a) = M(a) := 0$. Then $m(x)$ and $M(x)$ have finite values for each $x \in (a, b]$. Let $x \in (a, b)$ and $z \in (x, x + \delta(x) \wedge b]$. Then for each $\tau = (\{a = x_0, \ldots, x_m = x\}, \{y_1, \ldots, y_m\}) \in TP(\delta_x, [a, x])$, $(\{x_0, \ldots, x_m, z\}, \{y_1, \ldots, y_m, x\}) \in TP(\delta_z, [a, z])$. Thus

$$S_{\delta_x}(\tau) + f(x)[g(z) - g(x)] \leq M(z)$$

for any $\tau \in TP(\delta_x, [a, x])$. Thus the first inequality in (6.29) holds. Similarly one can show that the second one and (6.30) also hold. Therefore m and M are minor and major functions respectively. By Lemma 6.27 and (6.35), it then follows that

$$(WPS) \int_a^b f\, dg - \epsilon \leq m(b) \leq L(f, g) \leq U(f, g) \leq M(b) \leq (WPS) \int_a^b f\, dg + \epsilon.$$

Since ϵ is arbitrary, $L(f, g) = U(f, g)$ and equals the (HK) integral. The proof of Theorem 6.30 is complete. □

Schwabik [100, Theorem 3.1] proved the following for g of bounded variation, where f is not necessarily regulated, if (a) f is bounded, or (b) f is arbitrary and for any t such that $g(t-) = g(t+)$, we have $g(t) = g(t-)$ also.

Theorem 6.31. *Suppose f and g are regulated on $[a, b]$ and g is right–continuous on $(a, b]$. If $\not\pm_a^b f\, dg$ exists then $(HK) \int_a^b f\, dg$ exists and has the same value.*

Proof. Given $\epsilon > 0$, take a partition $\lambda = \{z_j: j = 0, \ldots, m\}$ of $[a, b]$ such that for any refinement κ of λ and open intermediate partition σ of κ,

$$\Big| S_Y(f, g; \kappa, \sigma) - \not\pm_a^b f\, dg \Big| < \epsilon. \tag{6.36}$$

By Lemma 6.18 we can assume that (6.36) holds also for f replaced by $f_+^{(a,b)}$ or by $\Delta_{a,b}^+ f$, taking λ as a common refinement $\lambda = \lambda' \cup \lambda''$ of partitions λ' for $f_+^{(a,b)}$ and $\epsilon/2$, and λ'' for $\Delta_{a,b}^+ f$ and $\epsilon/2$. It suffices to prove the Theorem for $f_+^{(a,b)}$ and for $\Delta_{a,b}^+ f$.

Define a gauge function $\delta(\cdot)$ on $[a, b]$ by: if $x \notin \lambda$, let $\delta(x) := \min_j |x - z_j|/2$. Then any $\delta(\cdot)$–fine tagged partition must contain each z_j as a tag. For $j = 0, \ldots, m$, define $\delta(z_j)$ such that: $\delta(z_j) \leq \min_{i \neq j} |z_i - z_j|/3$ and

$$\|f\|_\infty \sum_{j=1}^m \Big\{ \mathrm{Osc}(g; [z_j - \delta(z_j), z_j)) + \mathrm{Osc}(g; (z_{j-1}, z_{j-1} + \delta(z_{j-1})]) \Big\} < \epsilon. \tag{6.37}$$

Let $\tau = (\{u_i: i = 0, \ldots, n\}, \{v_i: i = 1, \ldots, n\})$ be any $\delta(\cdot)$–fine tagged partition of $[a, b]$ and let

$$S_\delta(\tau) := \sum_{i=1}^n f(v_i)[g(u_i) - g(u_{i-1})] \tag{6.38}$$

be the gauge sum based on τ. We can assume that $v_i = u_i = v_{i+1}$ never occurs since if it did, the ith and $(i+1)$st term in S could be replaced by $f(v_i)[g(u_{i+1}) - g(u_{i-1})]$, and we would still have a gauge sum based on a $\delta(\cdot)$–fine tagged partition equal to $S_\delta(\tau)$. Since the tags $\{v_i: i=1,\dots,n\}$ must contain the points of λ, for each index $j\in\{0,\dots,m\}$ there is an index $i(j)\in\{1,\dots,n\}$ such that $z_j=v_{i(j)}$. Let $\mu:=\{i(j): j=0,\dots,m\}\subset\{1,\dots,n\}$. For $j=1,\dots,m-1$, by definition of $\delta(\cdot)$, we must have

$$u_{i(j)-1} < v_{i(j)} < u_{i(j)}. \tag{6.39}$$

We also have $i(0)=1$, $v_1=z_0=u_0=a$ and $i(m)=n$, $v_n=z_m=u_n=b$.

First suppose that f is right–continuous on (a,b); that is the case when $f=f_+^{(a,b)}$. We will show that one can find a Young sum based on a refinement of λ, which is arbitrarily close to the gauge sum S. To this aim we will replace the values $g(u_i)$ in S, $i=1,\dots,n-1$, by values $g(x_i)$ at continuity points x_i of g, close to u_i. Then the Young sum terms $f(x_i)\Delta^{\pm}g(x_i)$ will be 0. For each $i=1,\dots,n-1$ such that $u_i<v_{i+1}$, replacing u_i in S by a slightly larger $x_i>u_i$ which is a continuity point of g gives a sum T with $|T-S|<\epsilon$. If $v_{i+1}=u_i$, then let U be a sum equal to T except that for each such i, we replace v_{i+1} by y_{i+1} and u_i by x_i with $u_i<x_i<y_{i+1}<u_{i+1}$, where x_i is a continuity point of g and let $y_{i+1}-u_i$ be as small as desired. Since f is right–continuous on (a,b), we can make $|U-T|<\epsilon$. In either case $v_i\le u_i<x_i$. Let $x_0:=a$ and $x_n:=b$. Let $y_i:=v_i$ for each $i=2,\dots,n-1$ for which y_i was not previously defined. Thus $y_{i(j)}=v_{i(j)}$ for $j=1,\dots,m-1$ by (6.39). Choose any $y_1\in(a,u_1)$ and $y_n\in(x_{n-1},b)$. Let $V:=S_Y(f,g;\zeta,\rho)$ be the Young sum based on $\zeta:=\{x_i: i=0,\dots,n\}$ and $\rho:=\{y_i: i=1,\dots,n\}$. Then

$$\begin{aligned}|V-U| &= |f(a)\Delta^+g(a)+f(y_1)[g(x_1)-g(a+)]-f(a)[g(x_1)-g(a)]\\ &\quad +f(b)\Delta^-g(b)+f(y_n)[g(b-)-g(x_{n-1})]-f(b)[g(b)-g(x_{n-1})]|\\ &= |[f(y_1)-f(a)][g(x_1)-g(a+)]+[f(y_n)-f(b)][g(b-)-g(x_{n-1})]|<4\epsilon\end{aligned}$$

by (6.37). Now let $\kappa:=\lambda\cup\zeta=\{s_l: l=0,\dots,k\}$, a partition of $[a,b]$ with $k=m+n-2$. Each interval (x_{i-1},x_i) for $i=2,\dots,n-1$ contains z_j if $i=i(j)\in\mu$, otherwise contains no z_j and becomes an interval (s_{l-1},s_l) for some l. In the latter case let $t_l:=y_i$. In the former case we have for some l, $s_{l-2}=x_{i-1}<z_j=s_{l-1}<x_i=s_l$. Choose any $t_{l-1}:=w_j$ with $s_{l-2}<t_{l-1}<s_{l-1}$ and $t_l:=r_j$ with $s_{l-1}<t_l<s_l$. Let $W:=S_Y(f,g;\kappa,\sigma)$ be the Young sum based on κ and its open intermediate partition $\sigma:=\{t_l: l=1,\dots,k\}$. Then

$$\begin{aligned}|W-V| &= \Big|\sum_{j=1}^{m-1}\Big\{f(w_j)[g(z_j-)-g(x_{i(j)-1})]+f(z_j)\Delta^-g(z_j)\\ &\quad +f(r_j)[g(x_{i(j)}-)-g(z_j)]-f(z_j)[g(x_{i(j)}-)-g(x_{i(j)-1})]\Big\}\Big|\\ &\le \sum_{j=1}^{m-1}|[f(w_j)-f(z_j)][g(z_j-)-g(x_{i(j)-1})]|\\ &\quad +\sum_{j=1}^{m-1}|[f(r_j)-f(z_j)][g(x_{i(j)}-)-g(z_j)]|<2\epsilon\end{aligned}$$

by (6.37). Here W is a Young sum over a refinement of λ, so $|W-{\neq}\!\!\!\int_a^b f\,dg|<\epsilon$. Thus $|S-{\neq}\!\!\!\int_a^b f\,dg|<9\epsilon$. The conclusion follows when f is right–continuous on (a,b).

Now suppose that $f\in c_0((a,b))$ (cf. notation (6.27)); that is the case when $f=\Delta_{a,b}^{+}f$. There is an open intermediate partition $\{w_j: j=1,\dots,m\}$ of $\lambda=\{z_j: j=0,\dots,m\}$ such that $f(w_j) = 0$ for all j. Thus by (6.36)

$$\Big|\sum_{j=1}^{m-1} f(z_j)\Delta^- g(z_j) - \underset{a}{\overset{b}{\not\!\int}} f\,dg\Big| < \epsilon. \tag{6.40}$$

For any set $\nu \subset \{1,\ldots,n-1\}$, consider partitions $\kappa = \{s_l\colon l = 0,\ldots,k\}$, $k = m+n-2$, consisting of λ, u_i for $i \in \nu$, and for each $i \in \{1,\ldots,n-1\} \setminus \nu$, a continuity point of g close to u_i. Let $\sigma = \{t_l\colon l = 1,\ldots,k\}$ be an open intermediate partition of κ with $f(t_l) = 0$ for all l. Then from (6.40) and (6.36) we obtain

$$\Big|\sum_{i\in\nu} f(u_i)\Delta^- g(u_i)\Big| < 2\epsilon. \tag{6.41}$$

Consider also partitions (κ,σ) defined in the same way except that for $\xi \subset \{1,\ldots,n-1\}$, if $i \in \xi \setminus \mu$ and i is even, we take $s_l = s_{l(i)}$ to be, instead of u_i, a continuity point of g a little larger than u_i, while $s_{l(i)-1}$ is a continuity point of g a little smaller than u_{i-1} and $t_{l(i)} = u_{i-1}$. For each i there will be an $s_{l(i)}$, a continuity point of g near u_i. Let $f(t_l) = 0$ for other l as before. Letting $s_{l(i)} \downarrow u_i$ and $s_{l(i)-1} \uparrow u_{i-1}$ it follows that

$$\Big|\sum \{f(u_{i-1})[g(u_i) - g(u_{i-1}-)]\colon i \in \xi \setminus \mu,\ i \text{ even}\}\Big| \le 2\epsilon.$$

The same holds likewise for i odd. Thus

$$\Big|\sum_{i\in\xi\setminus\mu} f(u_{i-1})[g(u_i) - g(u_{i-1}-)]\Big| \le 4\epsilon, \quad \text{or} \quad \Big|\sum_{i+1\in\xi\setminus\mu} f(u_i)[g(u_{i+1}) - g(u_i-)]\Big| \le 4\epsilon.$$

With $\nu = \{i\colon i+1 \in \xi \setminus \mu\}$ in (6.41) this gives

$$\Big|\sum_{i+1\in\xi\setminus\mu} f(u_i)[g(u_{i+1}) - g(u_i)]\Big| \le 6\epsilon.$$

If $v_{i+1} = u_i$ then $i+1 \notin \mu$ by (6.39). Thus

$$\Big|\sum \{f(u_i)[g(u_{i+1}) - g(u_i)]\colon u_i = v_{i+1},\ \ i = 1,\ldots,n-2\}\Big| \le 6\epsilon.$$

Here $f(u_i)$ could be replaced by $f(u_i) - f(y_{i+1})$, where $u_i < y_{i+1} < u_{i+1}$ and $f(y_{i+1}) = 0$. So the gauge sum (6.38) differs by at most 6ϵ from the sum

$$S' := \sum_{i=1}^{n} f(w_i)[g(u_i) - g(u_{i-1})],$$

where $w_i := y_i$ if $u_{i-1} = v_i$ and $i = 2,\ldots,n-1$, $w_i := v_i$ otherwise. Then $u_{i-1} < w_i \le u_i$ for $i = 2,\ldots,n-1$. The rest of the proof follows as in the case where f is right–continuous except that now there is no need to consider the sum U. The proof of Theorem 6.31 is complete. $\square$

Ward [111] stated and Saks [98, Theorem VI.8.1] gave a proof of the fact that $(WPS)\int_a^b f\,dg$ is defined provided the corresponding Lebesgue–Stieltjes integral $(LS)\int_a^b f\,dg$ is defined, and then they are equal. By Theorem 6.30, the gauge integral is in the same relation with the (LS) integral. The following gives conditions when the converse holds.

Theorem 6.32. *Let f be non-negative on $[a, b]$, and let g be non-decreasing on $[a, b]$. Then $(HK) \int_a^b f\,dg$ exists if and only if $(LS) \int_a^b f\,dg$ does, and then the two are equal.*

Proof. Recall that the McShane integral is defined as the (HK) integral except that in a tagged partition $(\{x_i: i = 0, \ldots, n\}, \{y_i: i = 1, \ldots, n\})$ the tags y_i need not be in the corresponding intervals $[x_{i-1}, x_i]$. By Corollary 6.3.5 of Pfeffer [90, p. 113], under the present hypotheses $(HK) \int_a^b f\,dg$ exists if and only if the corresponding McShane integral exists, and then the two are equal. The equivalence between the McShane and Lebesgue–Stieltjes integral was proved by McShane [80, pp. 552,553]. Also, it follows from Theorem 4.4.7, Proposition 3.6.14 and Theorem 2.3.4 of Pfeffer [90]. The proof of the theorem is complete. □

Suppose a continuous function $f: [a, b] \mapsto \mathbb{R}$ is differentiable everywhere on (a, b). It may be that f' is not Lebesgue integrable, e.g. if $[a, b] = [0, 1]$, $f(x) = x^2 \sin(\pi/x^2)$, $0 < x < 1$, $f(x) = 0$ elsewhere. The Denjoy and Perron integrals were defined so that for every such f', $\int_a^x f'(t)\,dt = f(x) - f(a)$, $a \le x \le b$. These integrals and the later–invented Henstock–Kurzweil integral turned out to be equivalent for integrals $\int_a^b g(x)\,dx$: Gordon [45, Chapter 11]. The following holds:

Theorem 6.33. *Let f be a continuous function: $[a, b) \mapsto \mathbb{R}$ having a derivative $f'(x)$ for $a < x < b$ except for at most countably many x. Then $(HK) \int_a^x f'(t)\,dt$ exists and equals $f(x) - f(a)$ for $a \le x < b$.*

Proof. Apply Theorems 7.2 and 11.1 of Gordon [45]. □

For another example, $f(x) := x^2 \sin(e^{1/x})$, $x \neq 0$, $f(0) := 0$, satisfies the conditions of Theorem 6.33 although f' has very wild oscillations.

While there has been relatively little literature about the Young (RYS) and (CY) integrals, there has been much more about Henstock–Kurzweil (gauge) integrals, e.g. Lee Peng–Yee [89]. A 1991 book by Henstock [48] has a reference list of more than 1200 papers and books, mainly on the theme of "non–absolute integration", if not necessarily about the (HK) integral itself. In fact, Henstock [48] treats integration over general spaces ("division spaces").

6.6 Comments and related results

Integration by parts In stochastic analysis, and for some integral equations as will be seen in Section 7.4, there are advantages in integrating left-continuous integrands with respect to right-continuous functions. In that case we have the following:

Theorem 6.34 (integration by parts). *Let f and g be right-continuous and of bounded variation on $[a, b]$. Then for Lebesgue-Stieltjes integrals,*

$$\int_a^b f_-^{(a)}\,dg + \int_a^b g_-^{(a)}\,df \;=\; fg|_a^b - \sum_{a<x\le b} (\Delta^- f \Delta^- g)(x). \tag{6.42}$$

Proof. By (2.4), $\int_a^b f_-^{(a)} dg = f_-^{(a)} g|_a^b - \int_a^b g\,df_-^{(a)}$. By formulas for the (CY) integral, $\int_a^b g\,df_-^{(a)} = \int_a^b g_-^{(a)}\,df_-^{(a)} + \sum_{(a,b)} \Delta^- g \Delta^- f$, and $\int_a^b g_-^{(a)} df_-^{(a)} = \int_a^b g_-^{(a)}\,df - (g_- \Delta^- f)(b)$. Collecting terms, the result follows. □

The integration by parts formula (6.42) also holds for any of the integrals following from the (LS) integral in Diagram 2.1 (the (RYS), (HK), or (CY) integral). Note that the sum $\sum \Delta^- f \Delta^- g$ appears in (6.42) and in (2.4) with opposite signs. Naturally, Δ^+ terms as in (2.4) don't occur in (6.42) because of right continuity. In stochastic analysis, there are integration by parts formulas with the terms of (6.42) and/or further continuous terms (see e.g. pp. 227-235 and 325-330 in [14]).

W. H. Young's integrals Let h be a monotonic function on $[a,b]$. Suppose f is a step function on $[a,b]$, that is f is constant on each open interval (x_{i-1}, x_i) of a partition $\kappa = \{x_i \colon i = 0, \ldots, n\} \in \mathrm{PP}([a,b])$. Then let

$$\int_a^b f\,dh := \sum_{i=1}^n \Big\{[f\Delta^+ h](x_{i-1}) + f(x_{i-1}+)[h(x_i-) - h(x_{i-1}+)] + [f\Delta^- h](x_i)\Big\}. \tag{6.43}$$

Starting with such integrals for step functions f, W. H. Young (1914) defined integrals for more general functions f by his method of monotone sequences, later adopted by Daniell [13] in defining the "Daniell integral." Thus Young obtained an integral equivalent to the Lebesgue–Stieltjes integral, (see Hildebrandt [49, p. 193]), which Lebesgue had previously defined by a change of variables. If $h(x) = x$, then the right side of (6.43) is a Riemann sum for f because $f(x_{i-1}+) = f(y_i)$ for each $y_i \in (x_{i-1}, x_i)$, $i = 1, \ldots, n$. Also, the right side of (6.43) is equal to the sum $S_Y(f, h; \kappa, \sigma)$ defined by (6.5) for any open intermediate partition σ of κ. Therefore it was natural to define an integral following the definition of the Riemann-Stieltjes integral except that Riemann-Stieltjes sums are replaced by Young sums. However it was known as stated by R. C. Young (1929, p. 221) that such an integral would coincide with the Riemann-Stieltjes integral for monotonic functions h. Namely the following is true:

Proposition 6.35. *Let f be a bounded function on $[a,b]$, let h be of bounded variation on $[a,b]$, and let h be continuous at each $x \in (a,b)$ such that $h(x-) = h(x+)$. Suppose that a number A exists with the property: for each $\epsilon > 0$ there exists $\delta > 0$ such that $|S_Y(f, h; \kappa, \sigma) - A| < \epsilon$ for each $\kappa \in \mathrm{PP}([a,b])$ with mesh $|\kappa| < \delta$ and for each open intermediate partition σ of κ. Then $(MRS) \int_a^b f\,dh$ exists and has value A.*

Remark. Any monotonic function h satisfies the above hypothesis. A different proof of the preceding statement is indicated by Hildebrandt ([52, II.19.3.12], cf. also [50, 8.11]). The following proof also extends to the case when $h \in \mathcal{W}_p$, $f \in \mathcal{W}_q$, $p > 1$, $q > 1$ and $1/p + 1/q = 1$ in place of the assumptions f bounded and $h \in \mathcal{W}_1$.

Proof. Let $\epsilon > 0$ and let $x \in (a,b)$. By hypothesis there exists $\delta > 0$ satisfying the stated property. Let $u < x < v$, $v - u < \delta$ and $y \in (u,v)$. Consider a partition $\kappa_{u,v}$ of $[a,b]$ containing u, v and no other points in between. Let σ_x and σ_y be two open intermediate partitions of κ with the same points except the points x and y. Subtracting two Young sums based on the same partition $\kappa_{u,v}$ and on two open intermediate partitions σ_x and σ_y, we get the bound

$$|[f(x) - f(y)][h(v-) - h(u+)]| < 2\epsilon.$$

Since $\epsilon > 0$ and y are arbitrary, it follows that f is continuous at x whenever $\Delta^{\pm} h(x) = h(x+) - h(x-) \neq 0$. Therefore at each $x \in (a,b)$ either f or h must be continuous. A similar argument applied for $x = a$ and $x = b$ yields that at each endpoint one of the two functions must be one-sidedly continuous. Due to this continuity implication, to show the Riemann-Stieltjes integrability of f with respect to h it is enough to consider Riemann-Stieltjes sums for open intermediate partitions. Let $\kappa = \{x_i : i = 0, \ldots, n\} \in \mathrm{PP}([a, b])$ and let $\sigma = \{y_i : i = 1, \ldots, n\}$ be an open intermediate partition of κ. Then we have

$$\begin{aligned} R(\kappa, \sigma) &:= S_{RS}(f, h; \kappa, \sigma) - S_Y(f, h; \kappa, \sigma) \\ &= \sum_{i=1}^{n} \{[f(y_i) - f(x_{i-1})]\Delta^+ h(x_{i-1}) + [f(y_i) - f(x_i)]\Delta^- h(x_i)\}. \end{aligned} \tag{6.44}$$

Again fix $\epsilon > 0$. Since h has bounded variation, there exists a finite set μ of jump points of h of cardinality m such that the sum of the absolute values of jumps at other points is less than $\epsilon/(2\|f\|_\infty)$. Then choose $\delta > 0$ such that the oscillation of f around each point of μ is less than $\epsilon/(2m\|h\|_\infty)$. By (6.44), it follows that $|R(\kappa, \sigma)| < 2\epsilon$ whenever the mesh $|\kappa| < \delta$. The statement of the proposition follows at once. $\square$

The preceding statement shows that replacing Riemann–Stieltjes sums by Young sums in the definition of the mesh–Riemann–Stieltjes integral essentially gives nothing new. In Section 4.3 we saw that this is not so if the same change is made in the definition of the (RRS) integral. It led to the new integral called the refinement–Young–Stieltjes integral, or just the (RYS) integral. Full-fledged uses of this integral appeared in the works of L. C. Young [118] and Gehring [39]. Hildebrandt discussed the (RYS) integral in his survey paper [50, p. 275] and in his book [52, pp. 88-96].

The early history of Stieltjes integrals Medvedev [81] gives a history of concepts of integration. He describes how Stieltjes (1894) defined an integral $\int_a^b f(u)\,dg(u)$ for f continuous and g nondecreasing. König (1897), in Hungarian, gave the full definition of mesh Riemann-Stieltjes integral, and showed that the classical integration by parts formula holds with endpoint corrections

$$\int_a^b f\,dg + \int_a^b g\,df - (fg)|_a^b = (\Delta^+ f \Delta^+ g)(a) - (\Delta^- f \Delta^- g)(b),$$

which actually must be 0 if the integrals are defined (in the mesh or the later refinement Riemann–Stieltjes sense). Kowalewski (1905) rediscovered the definition and now proved that

$$\int_a^b f\,dg + \int_a^b g\,df = (fg)|_a^b$$

provided that both integrals exist. According to Medvedev, F. Riesz in 1909 made known the existence of König's paper to the world outside of Hungary, but without characterizing it.