AN INTRODUCTION TO p-VARIATION AND YOUNG INTEGRALS.

With emphasis on sample functions of stochastic processes

R. M. Dudley and R. Norvaiša

December 2, 1998

Preface

These lecture notes for a course in Aarhus, January 1999, are a further development of our work on differentiability, p-variation, extended Stieltjes integrals, product integrals, and related topics.

In this survey and exposition, we have not aimed mainly at presenting our own research results, although some are included. More often we have mined the literature, finding under–appreciated gems from up to 85 years ago. We have proved some facts where we saw a need for linking existing results and concepts. We give earlier references where we know them. Our literature searches, although rather extensive in some directions, did not reach out so far in others, and in any case have not (yet) been primarily directed at determining the possible originality of our own results, some of which are recent at this writing. Thus, on any point, the absence of references to previous work, or references only to our own, should not be interpreted as a claim of originality.

We thank Ole Barndorff–Nielsen very much for the opportunity to give these lectures. We also thank Richard Gill and Terry Lyons for stimulating and helpful discussions.

Richard Dudley
Rimas Norvaiša
Cambridge, Mass., November 30, 1998

Contents

Chapter 1

Introduction

Let J be an interval in $\mathbb{R}$, which may be bounded or unbounded, and open or closed at either end. We will be concerned with *regulated* functions f on an interval J, having right limits $f(x+) := \lim_{y\downarrow x} f(y)$ with right jumps $\Delta^+ f(x) := f(x+) - f(x)$ for $x \in J$, not equal to the right endpoint of J, and left limits $f(x-) := \lim_{y\uparrow x} f(y)$ with left jumps $\Delta^-(f)(x) := f(x) - f(x-)$ for $x \in J$, not equal to the left endpoint of J. The functions will at first be real-valued but later may take values in a Banach algebra.

Here are two questions we will address.

1. How should integrals $\int_a^b f\,dg$ be defined?
2. For differentiability of nonlinear operators on subspaces of the space of regulated functions, what are good modes of differentiability (Fréchet, compact, etc.) and for what norms?

We have found that p-variation norms are useful. They are defined as follows. For any interval $J \subset \mathbb{R}$ let $\mathrm{PP}\,(J)$ denote the set of all *point partitions* of J, namely finite sequences $\{x_j\}_{j=0}^n \subset J$, where $x_0 < x_1 < \cdots < x_n$, and if $J = [a,b]$, a closed bounded interval, then $x_0 = a$ and $x_n = b$. Let $\phi\colon [0,\infty) \to [0,\infty)$ be continuous, unbounded, increasing and 0 at 0, and let $f\colon J \to \mathbb{R}$. For each $\kappa = \{x_i\colon i = 0,1,\dots,n\} \in \mathrm{PP}(J)$, let $s_\phi(f;\kappa) := \sum_{i=1}^n \phi(|f(x_i) - f(x_{i-1})|)$. Then the *$\phi$-variation* of f is defined by

$$v_\phi(f;J) := \sup\{s_\phi(f;\kappa)\colon \kappa \in \mathrm{PP}(J)\}.$$

In the special case of $\phi(u) \equiv u^p$, $u \geq 0$, for some $1 \leq p < \infty$, we write $s_p := s_\phi$, $v_p := v_\phi$ and call $v_p(f;J)$ the *p-variation* of f. Let $\mathcal{W}_p(J)$ and $\mathcal{W}_\phi(J)$ be the sets of all functions $f\colon J \mapsto \mathbb{R}$ for which $v_p(f) < \infty$ and $v_\phi(f) < \infty$, respectively. If ϕ is convex let $\|f\|_{(\phi)} := \inf\{c > 0\colon v_\phi(f/c) \leq 1\}$. Thus for $\phi(u) := u^p$, $1 \leq p < \infty$, $\|f\|_{(p)} := \|f\|_{(\phi)} = v_p(f)^{1/p}$. Then $\|\cdot\|_{(p)}$ is a seminorm on $\mathcal{W}_p(J)$, which we call the p-variation seminorm. For $p = \infty$ let $\mathcal{W}_\infty(J)$ be the set of all regulated functions f on J, with $\|f\|_{(\infty)} := \sup f - \inf f$. Let $\|f\|_\infty := \sup_{x\in J} |f(x)|$. For $1 \leq p \leq \infty$ let $\|f\|_{[p]} := \|f\|_{(p)} + \|f\|_\infty$. Then $\|\cdot\|_{[p]}$ is a norm on $\mathcal{W}_p(J)$, called the p-variation norm.

The p-variation spaces and norms have the following invariance property. Let G be any strictly increasing, continuous function from J onto another interval M. Then $f \in \mathcal{W}_p(M)$ if and only if $f \circ G \in \mathcal{W}_p(J)$, with $\|f\|_{[p]} \equiv \|f \circ G\|_{[p]}$. Of the many norms defined on spaces of functions on intervals (e.g. Triebel [107]) we do not know any beside ϕ-variation norms and the supremum norm which share such wide invariance.

Chapter 2

How should integrals of Stieltjes type be defined?

2.1 Definitions of integrals

We begin with several definitions of integrals $\int_J f\,dg$ we will consider, where J is an interval in the real line $\mathbb{R}$, f is a function called the *integrand* and g is a function called the *integrator*. The integrals will be treated in more detail in Chapter 6.

Let J be a closed interval $[a,b]$ with $-\infty < a < b < \infty$. The *mesh* of a point partition $\xi := \{x_j\}_{j=0}^n$ is $\max_{1\le j\le n}(x_j - x_{j-1})$. A *tagged partition* will be a pair (ξ,η) where $\eta = \{y_j\}_{j=1}^n$ and $x_{j-1} \le y_j \le x_j$ for $j = 1,\dots,n$. If (ξ,η) is a tagged partition of $[a,b]$ and f, g are two functions: $[a,b] \mapsto \mathbb{R}$, the *Riemann–Stieltjes sum* $S_{RS}(f,g,\xi,\eta)$ is defined as $\sum_{j=1}^n f(y_j)[g(x_j) - g(x_{j-1})]$. The *mesh–Riemann–Stieltjes integral* $(MRS)\int_a^b f\,dg$ is defined and equal to C iff (if and only if) for every $\epsilon > 0$ there is a $\delta > 0$ such that for every tagged partition (ξ,η) with $\mathrm{mesh}(\xi) < \delta$ we have

$$\Big|S_{RS}(f,g,\xi,\eta) - C\Big| < \epsilon. \tag{2.1}$$

A point partition $\{u_i\}_{i=0}^m$ is a *refinement* of $\{x_j\}_{j=0}^n$ iff each $x_j = u_{i(j)}$ for some $i(j)$. The *refinement–Riemann–Stieltjes* integral $(RRS)\int_a^b f\,dg$ is defined and equal to C iff for every $\epsilon > 0$ there exists a partition $\tau \in \mathrm{PP}\,[a,b]$ such that for every tagged partition (ξ,η) with ξ a refinement of τ, (2.1) holds. To give historical credit where it is due, Hildebrandt [50], [52] calls the (RRS) integral the Moore–Pollard–Stieltjes integral. Clearly, if $(MRS)\int_a^b f\,dg$ exists then so does $(RRS)\int_a^b f\,dg$, with the same value. Some texts define the “Riemann–Stieltjes” integral as the (MRS) integral, others as the (RRS) integral.

A *gauge function* $\delta(\cdot)$: $[a,b] \mapsto \mathbb{R}$ is any strictly positive function. The tagged partition (ξ,η) is called $\delta(\cdot)$*–fine* if $y_j - \delta(y_j) \le x_{j-1} \le x_j \le y_j + \delta(y_j)$ for $j = 1,\dots,n$. A *McShane partition* is a pair (ξ,η) where now $a \le y_1 < y_2 < \cdots < y_n \le b$ and we do not require $y_j \in [x_{j-1},x_j]$.

The *Henstock–Kurzweil* or *gauge* integral $(HK)\int_a^b f\,dg$ is defined to exist and equal C iff for every $\epsilon > 0$ there is a gauge function $\delta(\cdot)$ such that for every $\delta(\cdot)$–fine tagged partition (ξ,η), (2.1) holds. The *McShane integral* is defined in the same way with McShane partitions in place of tagged partitions.

An *interval function* will mean a function defined on intervals which may be open or closed at either end and may be singletons. An interval function ν will be called *finitely* *additive* if whenever $M_1, \cdots, M_k$ are disjoint intervals whose union is an interval M, we have $\nu(M) = \sum_{j=1}^{k} \nu(M_j)$.

Let $\mathcal{R}(J)$ denote the set of all regulated functions on J. If J is a closed interval $[a,b]$ we write $\mathcal{R}[a,b] := \mathcal{R}(J)$. Let g be a regulated function defined on a closed interval $[a,b]$. Let $\nu_g(\{x\}) := g(x+) - g(x-)$ for $a < x < b$, $\nu_g(\{a\}) := g(a+) - g(a)$, $\nu_g(\{b\}) := g(b) - g(b-)$, and $\nu_g((c,d)) := g(d-) - g(c+)$ for $a \le c < d \le b$. Then ν_g extends uniquely to a finitely additive interval function defined on all subintervals of J. If g is of bounded variation, then ν_g extends to a countably additive signed measure on the Borel subsets of J, and the *Lebesgue-Stieltjes integral* $\int_a^b f\,dg$ is defined here as $\int f\,d\nu_g$. (Some other definitions in the literature take different terms at the endpoints.)

An *interval partition* of J will be a collection $\{I_j\}_{j=1}^{n}$ of disjoint non-empty subintervals of J whose union is J, some of which may be singletons, and such that $x < y$ whenever $x \in I_i$, $y \in I_j$, and $1 \le i < j \le n$. Let $\mathrm{IP}\,(J)$ be the set of all interval partitions of J. For $\kappa, \lambda \in \mathrm{IP}\,(J)$, κ will be said to be a *refinement* of λ if each interval in λ is a union of intervals in κ. Given a real-valued function f on J, a finitely additive interval function ν defined on all subintervals of J and $\lambda = \{I_j\}_{j=0}^{n} \in \mathrm{IP}(J)$, a *Kolmogorov sum* for f and λ will mean any sum $\sum_{j=0}^{n} f(y_j)\nu(I_j)$ where $y_j \in I_j$ for each j. The *refinement-Young-Stieltjes integral* (*RYS*) $\int_a^b f\,d\nu$ is defined to exist and equal C if and only if for every $\epsilon > 0$ there is some $\kappa \in \mathrm{IP}\,(J)$ such that for any refinement λ of κ and any Kolmogorov sum S for f and λ we have $|S - C| < \epsilon$. If g is any regulated function on J then we define $(RYS) \int_a^b f\,dg$ as $(RYS) \int_a^b f\,d\nu_g$ if and only if the latter exists. We will also write $\ne\!\!\!\!\int_a^b f\,d\nu := (RYS) \int_a^b f\,d\nu$ iff the latter exists, and $\ne\!\!\!\!\int_a^b f\,dg := (RYS) \int_a^b f\,d\nu_g$ iff the latter exists and g is right–continuous on (a,b). A *Young partition* will be an interval partition consisting of singletons and open intervals. A *Young sum* will be a Kolmogorov sum for which the partition is a Young partition. Any interval partition has a Young partition as a refinement. On the equivalence of integrals defined from Kolmogorov sums or Young sums, see Theorems 4.35, 6.25 and 6.26 below.

A sum $\sum_{i \in I} x_i$ is said to converge *unconditionally* to S iff for every $\epsilon > 0$ there is a finite set $F \subset I$ such that for every finite $G \supset F$, $|S - \sum_{i \in G} x_i| < \epsilon$. A sum $\sum_{i \in I} x_i$ with values in $\mathbb{R}$ or a finite–dimensional normed space converges unconditionally (to some S) iff the sum converges absolutely, $\sum_{i \in I} |x_i| < \infty$. In infinite–dimensional Banach spaces, absolute convergence of a sum implies unconditional convergence but not conversely (Dvoretzky and Rogers, [31]). We will give some definitions in terms of unconditional convergence with a view to eventual extensions to the Banach–valued case.

For $f \in \mathcal{R}[a,b]$ recall that

$$(\Delta^+ f)(x) := f(x+) - f(x), \quad a \le x < b \quad \text{and} \quad (\Delta^- f)(x) := f(x) - f(x-), \quad a < x \le b.$$

Also let $(\Delta^{\pm} f)(x) := (\Delta^+ f)(x) + (\Delta^- f)(x) = f(x+) - f(x-)$ if $a < x < b$. Define functions $f_+^{(b)}$ and $f_-^{(a)}$ on $[a,b]$ by

$$\begin{cases} f_+^{(b)}(x) := f_+(x) := f(x+) := \lim_{z \downarrow x} f(z) & \text{if} \quad a \le x < b, \quad \text{and} \quad f_+^{(b)}(b) := f(b), \\ f_-^{(a)}(x) := f_-(x) := f(x-) := \lim_{z \uparrow x} f(z) & \text{if} \quad a < x \le b, \quad \text{and} \quad f_-^{(a)}(a) := f(a). \end{cases} \tag{2.2}$$

Then the *L. C. Young* or *central Young integral* is defined for $f, h \in \mathcal{R}[a,b]$ by

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

if the (RRS) integral exists and the sum converges unconditionally.

2.2 Choices between integrals

Now, which of the integrals is to be preferred in what situations?

The Lebesgue-Stieltjes integral $(LS)\int_a^b f\,dg$ is at hand if g is of bounded variation (and right-continuous except possibly at a), but what if it is not? The other best-known integrals are the mesh–Riemann-Stieltjes integral $(MRS)\int_a^b fdg$ and the refinement–Riemann–Stieltjes integral $(RRS)\int_a^b fdg$. But $(MRS)\int_a^b fdg$ is not defined if f and g have a common discontinuity and $(RRS)\int_a^b f\,dg$ is not defined if f and g have a discontinuity on the same side of the same point. Thus if $(RRS)\int_a^b f\,df$ is defined, f must be continuous.

Consider, specifically, the following, which we will call the *simple step example*. Let $F = 1_{[1,\infty)}$. Then $\int_0^2 F\,dF$ is not defined for the (MRS) or (RRS) integral. It is defined and equals 1 for the Lebesgue-Stieltjes integral, and all the other integrals defined in the previous section. But the classical integration by parts formula then fails: $2 = \int_0^2 F\,dF + \int_0^2 F\,dF \neq F^2(2) - F^2(0) = 1$. An alternate simple step example is given by $H := 1_{(1,\infty)}$ where again $\int_0^2 H\,dH$ is not defined for the (MRS) or (RRS) integral. The integral is defined and equals 0 for all the other integrals in Figure 2.1 below. Again, classical integration by parts fails since $\int_0^2 H\,dH + \int_0^2 H\,dH = 0 \neq H^2(2) - H^2(0) = 1$.

One useful application of the Lebesgue integral is the Hölder inequality. An analogous application of the less classical integrals defined above is the *Love-Young inequality*, as follows.

Theorem 2.1 (Love-Young inequality). *Let $f \in \mathcal{W}_p[a,b]$ and $g \in \mathcal{W}_q[a,b]$ where $p \geq 1,\ q \geq 1$, and $p^{-1} + q^{-1} > 1$. Then*

$$|\int_a^b f\,dg| \ \leq\ C_{p,q}\|f\|_{[p]}\|g\|_{(q)} \tag{2.3}$$

where $C_{p,q}$ is a constant depending only on p and q, and the integral is defined as a Riemann-Stieltjes integral if f and g have no common discontinuities, as an (RRS) integral if they have no common one-sided discontinuities, and as an (RYS) integral always.

A proof of the Love–Young inequality with $C_{p,q} = \zeta(p^{-1} + q^{-1})$, where $\zeta(s) := \sum_{n=1}^{\infty} n^{-s}$ (the Riemann zeta function), is given in Theorem 4.26.

By making seemingly slight changes in the definition of (RRS) integral – requiring intervals to be disjoint and allowing them to be singletons – integrals $(RYS)\int_a^b f\,dg$ can be defined despite common one-sided discontinuities of f and g. (Just taking disjoint intervals, e.g. left open, right closed intervals $(a,b]$, is not sufficient.) Also, for regulated g, whenever $(RRS)\int_a^b fdg$ exists, so does $(RYS)\int_a^b fdg$ and has the same value by Theorem 6.13 below.

Why has the superiority of the (RYS) integral over the (MRS) and (RRS) integrals only been noticed by a small number of analysts? We venture the following historical remarks, each to the best of our knowledge.

W. H. Young (1914), in a paper written in detail by his wife G. C. Young, first defined what we call Young sums but only for step functions. Their daughter R. C. Young (1928) defined the sums more generally but considered only the mesh $\to 0$ case, where the integral obtained does not go beyond the Riemann-Stieltjes integral (Proposition 6.35). Kolmogorov (1930) defined a refinement integral in general spaces. The (RYS) integral as we defined it is the special case of Kolmogorov's integral for interval partitions. The relation with the Youngs' work was noticed only later. Glivenko (1936, pp. 87-117) wrote a book on the Stieltjes integral where he defined and pointed out the superiority of the integral $(RYS)\int_a^b f\,dg$ over the (RRS) integral for g of bounded variation. Glivenko's work seems to have passed almost unnoticed.

L. C. Young (1936), son of W. H. and G. C. Young, first published a proof of the Love–Young inequality, noting that E. R. Love (unpublished) had proved a form for finite sums earlier. In 1936 Young invented and used the (CY) integral. L. C. Young [118, Section 5] adopted the (RYS) integral. But the Love–Young inequality itself, and thus the need for a suitable integral in it, also received relatively little notice. Hildebrandt (1938) formulated the (RYS) integral. Gehring (1954) proved a sharp inequality for convolutions $\int f(x-y)dg(y)$ in terms of the (RYS) integral. Hildebrandt [52, pp. 92-96] devotes several pages to the (RYS) integral. But Hildebrandt does not cite L. C. Young or Gehring, thus the value of the (RYS) integral in analysis was not made evident. Hildebrandt [52, Section II.19.3.13] states and proves the integration by parts formula, when f and g are of bounded variation,

$$\int_a^b f\,dg + \int_a^b g\,df = fg|_a^b - \sum_{a\le x<b} \Delta^+ f \Delta^+ g + \sum_{a<x\le b} \Delta^- f \Delta^- g \tag{2.4}$$

for the (RYS) integral only. Perhaps some analysts have been deterred from accepting the (RYS) integral by this formula. But (2.4) also holds for the Lebesgue-Stieltjes integral when f and g are both right-continuous or both left-continuous. In that sense (2.4) is correct. In the simple step examples, it gives: $2 = \int_0^2 F\,dF + \int_0^2 F\,dF = F^2(2) - F^2(0) + (\Delta^- F(1))^2 = 2$, $0 = \int_0^2 H\,dH + \int_0^2 H\,dH = H^2(2) - H^2(0) - (\Delta^+ H(1))^2 = 0$. The integration by parts formula (2.4) is not so much a helpful device for evaluating $\int_a^b f\,dg$ but a fact of life. For other forms of integration by parts see [28, Theorem 4.8].

Advanced textbooks of real analysis, including one by one of us (Dudley [22]) usually adopt the Lebesgue integral with perhaps some mention of the (MRS) and/or (RRS) integrals. Intermediate-level texts generally treat the three integrals with little if any treatment of others. The (MRS) and (RRS) integrals are treated as if to some degree they were the integrals previously known by the student and their defects are perhaps taken mainly to show the advantages of the Lebesgue integral. Thus, it is not noticed that integrals $\int_a^b f\,dg$ where g is not of bounded variation could be useful.

It seems that in recent years more and more authors are recognizing the inadequacy of the Riemann–Stieltjes integrals, and are presenting in detail other integrals, beside the Lebesgue–Stieltjes integral, at not too advanced levels of study. Specifically Bartle [3] and Gordon [45], [46] propose or give expositions of the gauge integral, while Ross [95], [96] and Love [74] present the (rediscovered) refinement–Young–Stieltjes integral. The form of gauge integral given by McShane [80, pp. 552-553] turns out to be equivalent to the Lebesgue–Stieltjes integral for g of bounded variation.

What are the implications between existence of different integrals and p-variation conditions? In Figure 2.1, $\longrightarrow$ means that existence of the integral to the left of it implies that of the integral to the right of it, with the same value. For "$\longleftarrow$" left and right are interchanged. The marking "$f \in \mathcal{R}$" or "$g \in \mathcal{R}$" means that the implication holds for regulated f or g, respectively. "$\Longrightarrow$" means that the condition to the left of it implies existence of the integral to the right of it. The condition $\frac{1}{p}+\frac{1}{q}=1$ on the right has no arrows from (or to) it, signifying that there exist $f \in \mathcal{W}_p$ and $g \in \mathcal{W}_q$ with $\frac{1}{p}+\frac{1}{q}=1$, $1<p,q<\infty$ such that $\int_a^b f\,dg$ does not exist for any of the definitions given.

It seems to us that the (MRS) and (RRS) integrals should not be preferred because of their weakness in case of common discontinuities. When g is of bounded variation the Lebesgue–Stieltjes integral is available, as is its extension the (HK) integral. For more general g we still have a choice between a few integrals. If f is not necessarily regulated then the gauge (HK)

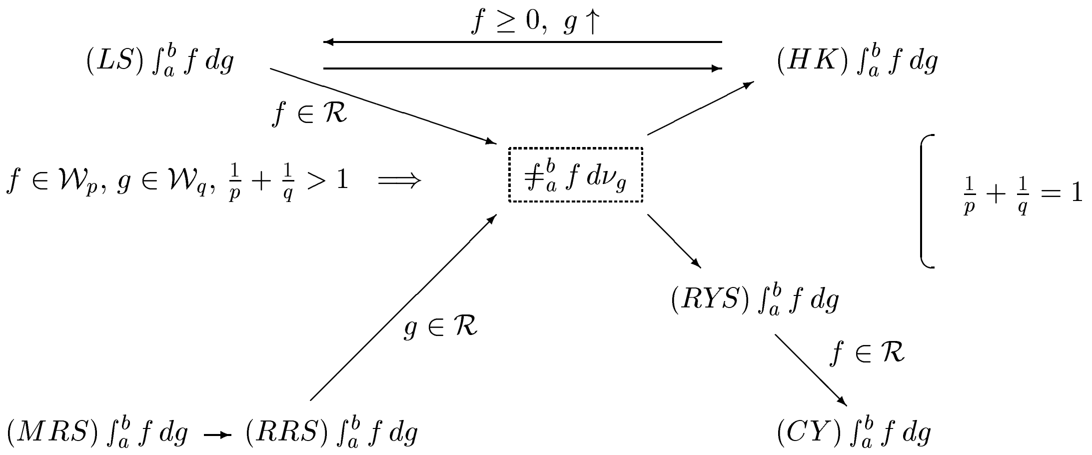

Figure 2.1: Implications for integrals

Table 2.1: References to proofs of the implications shown in Figure 2.1:

$(MRS)\int_a^b f\,dg \longrightarrow (RRS)\int_a^b f\,dg$	Proposition 6.1
$f \in \mathcal{W}_p,\ g \in \mathcal{W}_q,\ \frac{1}{p}+\frac{1}{q} > 1 \Longrightarrow ⨎_a^b f\,d\nu_g$	Theorem 4.26
$(RRS)\int_a^b f\,dg \longrightarrow ⨎_a^b f\,d\nu_g, \quad g \in \mathcal{R}$	Theorem 6.13
$(LS)\int_a^b f\,dg \longrightarrow ⨎_a^b f\,d\nu_g, \quad f \in \mathcal{R}$	Lemma 3.3 of [23] and Proposition 6.22
$(LS)\int_a^b f\,dg \longrightarrow (HK)\int_a^b f\,dg$	Theorem VI.8.1 in [98] and Theorem 6.30
$(HK)\int_a^b f\,dg \longrightarrow (LS)\int_a^b f\,dg, \quad f \geq 0,\ g\uparrow$	Theorem 6.32
$⨎_a^b f\,d\nu_g \longrightarrow (HK)\int_a^b f\,dg$	Theorem 6.31
$⨎_a^b f\,d\nu_g \longrightarrow (RYS)\int_a^b f\,dg$	immediate
$(RYS)\int_a^b f\,dg \longrightarrow (CY)\int_a^b f\,dg$	Theorem 6.20
$f \in \mathcal{W}_p,\ g \in \mathcal{W}_q,\ \frac{1}{p}+\frac{1}{q} = 1,\ 1 < p < \infty$ does not imply any of the above integrals exists	Theorem 4.29

integral seems preferable among the integrals defined here. We are concerned with f and g regulated. Let's compare integrals in that case.

The (RYS) integral, while "stronger", is virtually the same as the $⨎$ integral. In $⨎_a^b f\,dg$, g is assumed right–continuous. The existence and value of $(RYS)\int_a^b f\,dg$ only depend on $g(x-)$ and $g(x+)$ for $a < x < b$, not on $g(x)$ if it happens to differ from both limits. Thus one can add to g a function in $c_0((a,b)) := \{\sum c_n 1_{\{x_n\}} \colon a < x_n < b,\ c_n \to 0 \text{ as } n \to \infty\}$, without changing the value of the integral. Such an addition can cause $(HK)\int_a^b f\,dg$ to be undefined (Example 2.1 in Schwabik [100], attributed to Vrkoč, and Proposition F.6 of [28]), but this in itself does not show any real superiority of the (RYS) over the (HK) integral. Likewise, in our examples where $(CY)\int_a^b f\,dg$ exists but $(RYS)\int_a^b f\,dg$ does not (Proposition 6.21), $f \in c_0((a,b))$, so it is not clear whether the (CY) integral has any real advantage. Its definition as an integral plus a sum is relatively complex.

We are then left with the comparison between the $⨎$ and (HK) integrals. The latter is more inclusive, but in what directions? For $f \geq 0$ and g non–decreasing the (HK) integral is equivalent to the Lebesgue–Stieltjes integral and thus, if f is regulated, to the $\not\equiv$ integral.

For a set of examples of functions f, g with $f \in \mathcal{W}_p$, $g \in \mathcal{W}_q$, $p^{-1} + q^{-1} = 1$, $1 < p, q < \infty$, all the integrals (including the (HK)) are undefined (Theorem 4.29). For $p^{-1} + q^{-1} > 1$, $1 < p, q < \infty$, the $\not\equiv$ integral and thus all to the right of it in Figure 2.1 are defined. For almost all paths B_t of Brownian motion, $\int_0^1 B_t \, dB_t$ does not exist for any of the integrals in Figure 2.1 (Proposition 2.2 below).

Perhaps the best–known feature of Perron–type integrals, including the (HK) integral, which goes beyond other integrals, is the following: if f is a continuous function from an interval $[a, b]$ into $\mathbb{R}$, everywhere differentiable on (a, b) except possibly on a countable set, then $(HK) \int_a^x f'(t) \, dt$ exists and equals $f(x) - f(a)$ for $a \leq x \leq b$ (Theorem 6.33). If h is any regulated function on $[a, b]$, it is bounded and measurable and has an indefinite integral $f(x) := \int_a^x h(t) \, dt$. Then $f'(x) = h(x)$ at any x where h is continuous, and so, except on the at most countable set of jumps of h. Since h is continuous almost everywhere for Lebesgue measure and bounded, it satisfies a well known criterion for Riemann integrability (Riemann [94]). So $\int_a^x h(t) \, dt$ can be any of the integrals in Figure 2.1, and the (HK) integral does not go beyond the others for derivatives, if they are regulated. If f' exists everywhere on (a, b) and is regulated (having limits also at a, b), then clearly f' is continuous on (a, b) and extends to a continuous function on $[a, b]$.

Summing up, if f and g are regulated, as they are if f is in $\mathcal{W}_p$ and g is in $\mathcal{W}_q$, we do not know any examples where $(HK) \int_a^b f \, dg$ exists but $\not\equiv_a^b f \, dg$ does not. Since the definition of $\not\equiv_a^b f \, dg$ is simpler, we prefer it for regulated functions, at least for the present.

2.3 Stochastic integrals

We recall first a classical stochastic integral.

The Itô integral Let $B = \{B(t) \colon t \geq 0\}$ be a Brownian motion on a probability space $(\Omega, \mathcal{A}, \Pr)$. For $t \geq 0$, let $\mathcal{F}_t$ be the σ–algebra generated by the random variables $\{B(r) \colon 0 \leq r \leq t\}$. Then for $0 \leq t < r < \infty$ we have $\{\emptyset, \Omega\} := \mathcal{F}_0 \subseteq \mathcal{F}_t \subseteq \mathcal{F}_r \subseteq \mathcal{A}$. For $0 < T < \infty$ consider the class $\mathcal{H}[0, T]$ of stochastic processes $H = \{H(t) \colon 0 \leq t \leq T\}$ satisfying the following properties:

(1) $(t, \omega) \mapsto H(t, \omega)$ is $\mathcal{B} \times \mathcal{A}$–measurable, where $\mathcal{B}$ is the Borel σ–algebra on $[0, T]$;

(2) H is $\mathcal{F}_t$–adapted, i.e. $H(t, \cdot)$ is $\mathcal{F}_t$–measurable for all $0 \leq t \leq T$;

(3) $E[\int_0^T H^2(t) \, dt] < \infty$.

We now define the Itô integral $(I) \int_0^T H \, dB$ for each $H \in \mathcal{H}([0, T])$. First suppose that H is elementary, that is

$$H(t, \omega) = H_0(\omega) 1_{\{0\}}(t) + \sum_{j=1}^{n} H_j(\omega) 1_{(t_{j-1}, t_j]}(t)$$

for some $0 = t_0 < t_1 < \cdots < t_n = T$, where H_0 is $\mathcal{F}_0$–measurable, each H_j, $j = 1, \ldots, n$, is $\mathcal{F}_{t_{j-1}}$–measurable and $EH_j^2 < \infty$. For such H define the Itô integral by

$$(I) \int_0^T H \, dB := \sum_{j=1}^{n} H_j [B(t_j) - B(t_{j-1})].$$

For an arbitrary $H \in \mathcal{H}[0,T]$, there exists a sequence $\{H_m \colon m \geq 1\}$ of elementary processes H_m such that

$$\lim_{m\to\infty} E\Big[\int_0^T (H(t) - H_m(t))^2\, dt\Big] = 0.$$

Then the Itô integral for H is defined as the limit

$$(I)\int_0^T H\, dB := \lim_{m\to\infty} (I)\int_0^T H_m\, dB \quad \text{in } L_2(\Omega, \Pr),$$

which does not depend on the actual sequence $\{H_m \colon m \geq 1\}$ chosen. The proof of these statements follows from the Itô isometry

$$E\Big[(I)\int_0^T H\, dB\Big]^2 = E\Big[\int_0^T H^2(t)\, dt\Big]$$

valid for each $H \in \mathcal{H}[0,T]$. This construction can be extended in several ways (see e.g. [14, Chapter VIII]). Condition (3) can be relaxed to $\int_0^T H^2(t)\, dt < \infty$ almost surely. Then $(I)\int_0^T H\, dB$ need no longer be square–integrable but can be an arbitrary $\mathcal{F}_T$–measurable random variable (see [21]). Researchers in stochastic analysis have some decades ago defined $\int H(t)\, dX_t$ rather generally, specifically when X_t is a semimartingale (for example, a right-continuous martingale) and $H(t)$ is a locally bounded predictable process (for example, an adapted process with regulated, left-continuous paths): see Doléans–Dade and Meyer [16], Dellacherie and Meyer [14].

To what extent can stochastic integrals be replaced or supplemented by integrals defined pathwise, for fixed ω?

The Love–Young inequality for $f \in \mathcal{W}_p$ and $g \in \mathcal{W}_q$ with $p^{-1} + q^{-1} > 1$, and the Young [118, Section 5] inequality for $f \in \mathcal{W}_\Phi$ and $g \in \mathcal{W}_\Psi$ with $\sum_{n=1}^\infty \Phi^{-1}(n^{-1})\Psi^{-1}(n^{-1}) < \infty$, give conditions for existence of $\neq\!\!\!\!\int_a^b f\, dg$ and bounds for it. Since many stochastic processes on intervals have sample paths in $\mathcal{W}_q$ for some q (see Chapter 5), one can then treat integrands f and integrators g in suitable $\mathcal{W}_p$, $\mathcal{W}_q$.

However, we have:

Proposition 2.2. *For almost all paths $B(t) := B_t$ of standard Brownian motion, $(A)\int_0^1 B_t\, dB_t$ is not defined for $(A) = (MRS)$, (RRS), (LS), (RYS), (CY) or (HK).*

Proof. It suffices to consider (CY) and (HK) by the implications shown in Figure 2.1. If the (CY) integral existed it would equal the (MRS) integral by sample continuity and the definition of the integral, so $\neq\!\!\!\!\int_0^1 B_t\, dB_t$ would exist. Brownian paths almost surely have unbounded 2-variation on every interval $[a,b]$, $a < b$ (Lévy [71, Théorème 9], Taylor [106]). By continuity, there are arbitrarily large sums

$$\sum_{j=1}^n [B(t_j) - B(t_{j-1})][B(v_j) - B(u_j)]$$

with $a = t_0 < \cdots < t_{j-1} < u_j < v_j < t_j < \cdots < t_n = b$. Thus the sums $\sum_{j=1}^n [B(t_j) - B(t_{j-1})]B(w_j)$ with $t_{j-1} < w_j < t_j$ for each j are unbounded. Since these are Young sums, $\neq\!\!\!\!\int_0^1 B_t\, dB_t$ doesn't exist, nor does $(HK)\int_0^1 B_t\, dB_t$ by Proposition F.3 of Dudley and Norvaiša [28]. $\square$

Here is a sense in which at least the particular integral $(I)\int_0^1 B_t\,dB_t$ can be done pathwise. It is a special case of the construction suggested by Föllmer [34] for more general processes. Let $\{\lambda(m)\colon m \geq 1\}$ be a sequence of partitions $\lambda(m) = \{0 = t_0^m < t_1^m < \cdots < t_m^m = 1\}$ of $[0,1]$. Since $B_0 = 0$, for every $m \geq 1$, we have

$$S_L(\lambda(m)) := \sum_{i=1}^{m} B(t_{i-1}^m)[B(t_i^m) - B(t_{i-1}^m)] = \frac{1}{2}[B(1)^2 - s_2(B;\lambda(m))] \tag{2.5}$$

and

$$\sum_{i=1}^{m} B(t_i^m)[B(t_i^m) - B(t_{i-1}^m)] = \frac{1}{2}[B(1)^2 + s_2(B;\lambda(m))]. \tag{2.6}$$

The right sides of (2.5) and (2.6) have limits provided the limit $\lim_{m\to\infty} s_2(B;\lambda(m))$, called the quadratic variation, exists. It is so almost surely in the two cases:

(a) the sequence $\{\lambda(m)\colon m \geq 1\}$ is nested and $\cup_m \lambda(m)$ is dense in $[0,1]$ (Lévy, [71, Section 4, Théorème 5]) or

(b) the mesh of $\lambda(m)$ is of order $o(1/\log m)$ as $m \to \infty$ (Dudley, [20, Theorem 4.5]), and this condition is sharp (Fernández de la Vega, [32]).

If (a) or (b) holds then $\lim_{m\to\infty} s_2(B;\lambda(m)) = 1$ almost surely. From the definition of Itô integral, if $S_L(\lambda(m))$ converge almost surely, they converge to $(I)\int_0^1 B_t\,dB_t$, since they converge to it in L^2. One can take $\lambda(m) := \lambda_m = \{j/m\}_{j=0}^m$ by (b) or its subsequence $\lambda(2^m)$ by (a) or (b). On the other hand, for any partition λ, the supremum of $s_2(B;\kappa)$ over refinements κ of λ is $+\infty$ almost surely, so $\int_0^1 B_t\,dB_t$ is not a left Cauchy integral, meaning that Riemann-Stieltjes sums with $y_j = x_{j-1}$ do not converge as the mesh goes to 0 or as partitions are refined. Also, to deal with jumps of a process X_t, in defining an integral $\int_0^1 X_t\,dX_t$, jump points would need to be included in $\cup_m \lambda(m)$.

Recall that "cadlag" means regulated and right–continuous. Dellacherie and Meyer [14, Chapter VIII, Section 15] consider a rather general stochastic integral

$$S(t) := \int_0^t U_{s-}\,dX_s, \tag{2.7}$$

where U and X are cadlag processes and X is a semimartingale. They show that the integral is a limit in probability of left Cauchy sums

$$S_n(t) := \sum_{j=1}^{2^n} U((j-1)t/2^n)[X(jt/2^n) - X((j-1)t/2^n)].$$

Taking a subsequence if necessary, $S_{n_k}(t)$ will converge almost surely to $S(t)$ for all rational $t > 0$. By a different choice of partitions, letting $[x]$ be the largest integer $\leq x$, define

$$T_n(t) := U([2^n t]/2^n)[X(t) - X([2^n t]/2^n)] + \sum_{i=1}^{[2^n t]} U((i-1)/2^n)[X(i/2^n) - X((i-1)/2^n)].$$

Then for some subsequence, $T_{n_k}(t) \to S(t)$ almost surely for all $t \geq 0$. In this sense the stochastic integral can be done pathwise and, more generally, stochastic differential equations can be solved constructively – we claim no expertise in this direction, but give as references Bichteler [6], Jacod and Protter [58].

We will just mention that there are also ways to define stochastic integrals with anticipating integrands: Skorohod [103].

We would also suggest that integrals $\int Y_t \, dX_t$ can be defined pathwise via the Love–Young inequality if the functions $t \mapsto Y_t$ have bounded p–variation of order $p < 2$, or for processes X_t that have bounded q–variation, $q < 2$, as do e.g. suitable stable processes of order $\alpha < 2$ and some Gaussian processes such as fractional Brownian motion with the Hurst index $1/2 < H < 1$ (see sections 5.3 and 5.4). On the other hand, when the Love–Young inequality does apply, it does not require any probabilistic conditions such as (semi)martingale properties of X_t or predictability of Y_t. Still, when we come to consider integral equations and the product integral (Chapter 7) we will in fact use left–continuous integrands.

Paul Lévy [72] took note of L. C. Young's inequalities [117], [118] and the possibility of applying them to "Stieltjes–Young" integrals $\int Y_t \, dX_t$ for stochastic processes X_t, Y_t.

Non-random integrands and the Wiener process Historically, after Wiener (1923) defined the Brownian motion process B_t and showed its sample continuity, Paley, Wiener and Zygmund (1933) defined a stochastic integral $L(f) = \int_0^1 f(t) \, dB_t(\omega)$ for any non-random $f \in H = L^2[0,1]$ by an extension process. It turns out that L is the isonormal process on the Hilbert space H, a Gaussian process with mean 0 and covariance $E(L(x)L(y)) = (x, y)$, the inner product in the Hilbert space. Any Gaussian process has its finite-dimensional joint distributions completely determined by its means and covariances. Thus the isonormal process is determined by the abstract geometric Hilbert space structure, which has a great deal of rotational invariance. For example, a unitary linear transformation of $L^2[0,1]$ onto itself can take any element with norm 1 (say, a very smooth function) into any other such element (which may be quite non-smooth).

The isonormal process is a linear isometry, preserving inner products, of H into L^2 of a probability space (Ω, P). If one takes $L^2(\Omega, P)$ in turn as the Hilbert space and defines an isonormal process M on it, then $M \circ L$ has all the properties of L, i. e. it is also an isonormal process on H. Moreover, if $(X_t, t \in T)$ is any Gaussian process with mean 0 on any parameter space T, then $t \mapsto M(X_t)$ is another version of the process X_t, with the same covariances and joint distributions. Thus sample continuity and boundedness of Gaussian processes can be studied by way of the isonormal process, in other words, by putting the "intrinsic metric" $d(s,t) = [E(X_s - X_t)^2]^{1/2}$ on T. In fact, if (T, d) is a compact metric space and $t \mapsto X_t$ is continuous in probability, then X_t has continuous paths on T if and only if M does on $\{X_t(\cdot) \colon t \in T\} \subset L^2(\Omega, P)$ (Dudley [26, Theorem 2.8.2]). The focus on the isonormal process, as "the" Gaussian process, advocated by one of us (Dudley [19], [20]) has been well accepted, at least by use of the intrinsic metric (e. g. Talagrand [105]). One can first define the isonormal process L, specifically on $L^2[0, \infty)$, then define Brownian motion simply by $B_t = L(1_{[0,t]})$, e. g. Dudley [22, p. 352]. Thus it is unnecessary to define a stochastic integral $\int f(t) \, dB_t$ per se for non-random $f \in L^2$.