Chapter 7

Product integration

7.1 Ordinary differential equations

The product integral first arose as a way of representing the solutions of initial value problems for linear ordinary differential equations. Consider a linear kth order ordinary differential equation

$$\frac{d^k f(t)}{dt^k} + A_{k-1}(t)\frac{d^{k-1} f(t)}{dt^{k-1}} + \ldots + A_0(t) f(t) = B(t), \qquad a < t < b, \tag{7.1}$$

where for the present f and the coefficients A_j and B are real–valued. As is often done in differential equations one can write an equivalent first order linear vector and matrix differential equation

$$dY/dt = M \cdot Y, \tag{7.2}$$

where Y is the $(k+1) \times 1$ column vector $Y = (y_0, y_1, \ldots, y_k)'$ with $y_0 \equiv 1$, $y_1 \equiv f(t)$, $y_j = d^{j-1} f(t)/dt^{j-1}$ for $j = 2, \ldots, k$ and

$$M = \begin{pmatrix} 0 & 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 1 & \ldots & 0 \\ & \cdots & \cdots & \cdots & \cdots & \\ 0 & . & . & . & \ldots\ 0 & 1 \\ B & -A_0 & . & . & \ldots & -A_{k-1} \end{pmatrix}$$

with $B := B(t)$, $A_j := A_j(t)$, $j = 0, 1, \ldots, k-1$. If (7.2) holds at a point t, then

$$Y(t+h) - Y(t) = hM(t) \cdot Y(y) + o(h) \qquad \text{as } h \downarrow 0, \qquad \text{or}$$

$$Y(t+h) = (I + hM(t)) \cdot Y(t) + o(h) \qquad \text{as } h \downarrow 0, \tag{7.3}$$

where I is the $(k+1) \times (k+1)$ identity matrix. Let $N(t) := \int_a^t M(u)\, du$. If (7.2) holds for $a < t < b$ and Y is continuous on $[a, b]$ then informally, one sees by iterating (7.3) and letting $h \downarrow 0$ that for the product integral as defined in Section 3.5,

$$Y(u) = \Big(\prod_t^u (I + dN)\Big)\, Y(t) \qquad \text{for } a \le t \le u \le b. \tag{7.4}$$

More formally, making a stronger regularity assumptions than necessary for simplicity, if M is continuous, then (7.2) implies the integral equation v

$$Y(u) = Y(t) + \fint_t^u dN(x) \cdot Y(x), \qquad a \leq t \leq u \leq b, \tag{7.5}$$

and (7.4) will indeed give the unique solution of (7.5) and thus of (7.1) by Theorem 7.13 below, or by Theorem 5.21 in [29], as follows. These facts give that $H(u) := \prod_t^u(\mathbb{I}+dN)$ is the unique matrix–valued solution of the matrix integral equation

$$H(u) = I + \int_t^u dN(x) \cdot H(x), \qquad t \leq u \leq b. \tag{7.6}$$

Suppose that $Z(\cdot)$ were another vector–valued solution of

$$Z(u) = Y(t) + \int_t^u dN(x) \cdot Z(x).$$

Let $A(u) = (A_{ij}(u))$ be defined by

$$A_{ij}(u) := \begin{cases} (Z_j - Y_j)(u) & \text{for } i = 1 \\ 0 & \text{for } i > 1. \end{cases}$$

Then $H + A$ would give a solution of (7.6) different from H, a contradiction.

The above considerations also apply to systems of first–order linear ordinary differential equations, where M can be a general (continuous) square matrix–valued function instead of having the specific form coming from one kth order equation.

Equation (7.4) gives the solution of an initial–value problem for linear ordinary differential equations in terms of an operator which is clearly linear in the initial values $Y(t)$. It is non–linear in the coefficients. For a kth order linear ordinary differential equation (7.1) the solution will be $y_1(t)$. The operator will be shown to be holomorphic in the coefficients in Section 7.5.

It's easy to check that for two 3×3 matrices A, B of the form of M, so that $k = 2$, we have $AB = BA$ if and only if $A = B$. The same is true for any $k \geq 2$: consider the next–to–last row of the products. Thus commuting matrices will be obtained only for differential equations with constant coefficients if $k \geq 2$.

7.2 Product integrals

We were introduced to probabilistic and statistical uses of the product integral by Gill and Johansen [43], who use an interval function formulation. In our work [29] we used point functions. Now we have come to prefer interval functions, but perhaps there is some advantage in seeing and comparing both formulations.

The book of Dollard and Friedman [17] contains a survey and examples of applications of several types of product integrals. In the case when a function h is real valued, the product integral with respect to h usually can be written explicitly as in Theorem 7.8 below. Therefore we consider product integrals for functions with values in a Banach algebra $\mathbb{B}$ with an identity $\mathbb{I}$. Several results stated in earlier sections for real–valued functions will be used for $\mathbb{B}$–valued functions. Their proofs in the more general case are the same.

In Section 7.5 we apply some ingenious inequalities of T. Lyons [76] to bounding Peano series expansions of product integrals. It has not escaped our attention that these inequalities can have other applications. We present what we could write up in the time available so far.

Let μ be an additive interval function on an interval $[a,b]$ with values in $\mathbb{B}$. As before the class of all such functions is denoted by $\mathcal{I}([a,b];\mathbb{B})$. Also, as in Section 4.5 we use two classes of subintervals of $[a,b]$: $\mathfrak{I}_{os}([a,b]) := \{(c,d),\{c\}\colon a \le c \le d \le b\}$ and $\mathfrak{I}([a,b]) := \{[c,d],(c,d],[c,d),(c,d)\colon a \le c \le d \le b\}$. Let $\mathrm{IP}_{os}([a,b])$ be the set of all partitions of $[a,b]$ by intervals from $\mathfrak{I}_{os}([a,b])$. Recall that the set $\mathrm{IP}([a,b])$ of interval partitions was introduced in Section 2.1. For $J \in \mathfrak{I}([a,b])$ and $\mathfrak{D} = (J_1,\dots,J_N) \in \mathrm{IP}(J)$, let

$$P(\mu;\mathfrak{D}) := P_{\mathfrak{I}(J)}(\mu;\mathfrak{D}) := \prod_{j=1}^{N}[\mathbb{I}+\mu(J_j)] := [\mathbb{I}+\mu(J_N)]\cdot\ldots\cdot[\mathbb{I}+\mu(J_1)].$$

Similarly define $P_{\mathfrak{I}_{os}(J)}(\mu;\mathfrak{D})$. Since $\mathbb{B}$ may not be commutative, the product sign is used below with the prescribed order. Some authors use the reverse order.

Definition 7.1. Let $\mu \in \mathcal{I}([a,b];\mathbb{B})$ be additive and let $J \in \mathfrak{I}([a.b])$. The *product integral with respect to μ over J* is defined to be the limit

$$\prod_{J}(1+d\mu) := \lim_{\mathfrak{D}\uparrow} P_{\mathfrak{I}(J)}(\mu;\mathfrak{D}) \tag{7.7}$$

if it exists under refinements of $\mathfrak{I}$–partitions of J (cf. Definition 6.23).

Lemma 7.2. *Let $\mu \in \mathcal{I}([a,b];\mathbb{B})$ be additive and upper continuous. Then the limit in* (7.7) *exists if and only if it exists under refinements of $\mathfrak{I}_{os}$–partitions of J, and the two are equal.*

Proof. Since $\mathrm{IP}_{os}(J) \subset \mathrm{IP}(J)$ and each $\mathfrak{I}$–partition can be $\mathfrak{I}_{os}$–refined, if (7.7) exists then so does the limit under refinements of $\mathfrak{I}_{os}$-partitions of J, and the two are equal. The proof of the converse implication is similar to the proof of Theorem 6.26 and is omitted. □

As in Dudley and Norvaiša [29, Definition 4.20] given a $\mathbb{B}$-valued function h on $[a,b]$, the product integral

$$\prod_{a}^{b}(\mathbb{I}+dh) \tag{7.8}$$

is defined to exist and equal A iff for every $\epsilon > 0$ there is a $\lambda \in \mathrm{PP}([a,b])$ such that

$$\|P(h;\kappa)-A\| < \epsilon \tag{7.9}$$

for each refinement κ of λ, where $P(h;\kappa) := \prod_{i=1}^{n}[1+h(x_i)-h(x_{i-1})]$ for $\kappa = \{x_i\colon i = 0,\dots,n\} \in \mathrm{PP}([a,b])$.

To relate the product integral (7.7) with (7.8) recall the definition (4.32) of the interval function μ_h corresponding to a regulated function h. Then we have:

Theorem 7.3. *Let $h \in \mathcal{R}([a,b];\mathbb{B})$ be either right-continuous or left-continuous at each point of (a,b) and let $h(a)=0$.*

(1) *The existence of any of the three product integrals*:

$$\prod_{a}^{b}(\mathbb{I}+dh), \qquad \prod_{a+}^{b}(\mathbb{I}+dh), \qquad \prod_{a}^{b-}(\mathbb{I}+dh) \tag{7.10}$$

implies the existence of the corresponding product integral with the same value:

$$\prod_{[a,b]}(\mathbb{I}+d\mu_h), \qquad \prod_{(a,b]}(\mathbb{I}+d\mu_h), \qquad \prod_{[a,b)}(\mathbb{I}+d\mu_h). \tag{7.11}$$

(2) *If, in addition, there is a finite constant $C(h)$ such that*

$$\sup_{\mathfrak{D}\in \mathrm{IP}\,(J)} \|P(\mu_h;\mathfrak{D})\| \leq C(h) \tag{7.12}$$

for each subinterval $J \subset [a,b]$, then the converse to (1) *holds.*

Remark. By Lemma 4.19 of [29], condition (7.12) is satisfied provided $v_p(h;[a,b]) < \infty$ for some $0 < p < 2$.

Proof. We prove the stated relation only for the product integrals $\prod_a^b(\mathbb{I}+dh)$ and $\prod_{[a,b]}(\mathbb{I}+d\mu_h)$ because a proof for the other two pairs in (7.10) and (7.11) is similar. To prove (1) let $\epsilon > 0$. By definition of the product integral with respect to h, there exists $\lambda \in \mathrm{PP}\,([a,b])$ such that (7.9) holds for each refinement κ of λ. Let $\mathfrak{D}_\lambda$ be the $\mathfrak{I}_{os}$-partition of $[a,b]$ such that the corresponding point partition $\wp(\mathfrak{D}_\lambda) = \lambda$ and let $\mathfrak{D}$ be any $\mathfrak{I}_{os}$-refinement of $\mathfrak{D}_\lambda$. Let $\kappa := \{x_i\colon i = 0,\dots,n\} = \wp(\mathfrak{D})$ be the corresponding point partition of $[a,b]$. Let $I_r := \{i = 1,\dots,n-1\colon \Delta^+h(x_i) = 0\}$ and $I_l := \{i = 1,\dots,n-1\} \setminus I_r$. Then $\Delta^-h(x_i) = 0$ for each $i \in I_l$ by assumption. Choose points x_i', $i = 0,\dots,n$ recursively as follows: let $x_0' \in (x_0,x_1)$ and supposing x_{i-1}' to be chosen, then, if $i < n$, take either $x_i' \in (x_{i-1} \vee x_{i-1}', x_i)$ if $i \in I_r$, or $x_i' \in (x_i, x_{i+1})$ if $i \in I_l$, while if $i = n$ take $x_n' \in (x_{n-1} \vee x_{n-1}', x_n)$. Let $D_0' := [\mathbb{I} + h(x_1' \wedge x_1) - h(x_0')][\mathbb{I} + h(x_0')]$, for each $i = 1,\dots,n-1$, let

$$D_i' := \begin{cases} [\mathbb{I} + h(x_{i+1}' \wedge x_{i+1}) - h(x_i)][\mathbb{I} + h(x_i) - h(x_i')] & \text{if } i \in I_r \\ [\mathbb{I} + h(x_{i+1}' \wedge x_{i+1}) - h(x_i')][\mathbb{I} + h(x_i') - h(x_i)] & \text{if } i \in I_l \end{cases}$$

and let $D_n' := \mathbb{I} + h(x_n) - h(x_n')$. Letting $x_i' \downarrow x_i$ if $i \in \{0\} \cup I_l$ and $x_i' \uparrow x_i$ if $i \in I_r \cup \{n\}$, we get that $D_0' \to [\mathbb{I} + h(x_1-) - h(x_0+)][\mathbb{I} + h(x_0+)] = [\mathbb{I} + \mu_h((x_0,x_1))][\mathbb{I} + \mu_h(\{x_0\})]$,

$$D_i' \to \left\{\begin{matrix} [\mathbb{I} + h(x_{i+1}-) - h(x_i)][\mathbb{I} + \Delta^-h(x_i)] & \text{if } i \in I_r \\ [\mathbb{I} + h(x_{i+1}-) - h(x_i+)][\mathbb{I} + \Delta^+h(x_i)] & \text{if } i \in I_l \end{matrix}\right\} = [\mathbb{I} + \mu_h((x_i,x_{i+1}))][\mathbb{I} + \mu_h(\{x_i\})]$$

for each $i = 1,\dots,n-1$, and $D_n' \to \mathbb{I} + \mu_h(\{x_n\})$. Therefore

$$P(h;\kappa') = D_n' \cdots D_1' D_0' \to P_{\mathfrak{I}_{os}}(\mu_h;\mathfrak{D}) \quad \text{as } \kappa' \to \kappa.$$

Since $\lambda \subset \kappa \subset \kappa'$, by (7.9), it follows that $\|\prod_a^b(\mathbb{I}+dh) - P_{\mathfrak{I}_{os}}(\mu_h;\mathfrak{D})\| \leq \epsilon$ for any $\mathfrak{I}_{os}$-refinement $\mathfrak{D}$ of $\mathfrak{D}_\lambda$. Statement (1) now follows from Lemma 7.2.

To prove (2), take $C(h) \geq 1$ and again let $\epsilon > 0$. Since each $\mathfrak{I}$-partition can be $\mathfrak{I}_{os}$-refined, there is an $\mathfrak{I}_{os}$-partition $\mathfrak{D}_\epsilon$ of $[a,b]$ such that

$$\Big\| \prod_{[a,b]} (\mathbb{I} + d\mu_h) - P_{\mathfrak{I}_2}(\mu_h;\mathfrak{D}) \Big\| < \epsilon \tag{7.13}$$

for each $\mathfrak{I}$-partition $\mathfrak{D} \sqsupseteq \mathfrak{D}_\epsilon$. Let $\{z_j\colon j = 0,\dots,m\} := \wp(\mathfrak{D}_\epsilon)$ be the corresponding point partition, and let $\{v_{j-1}, u_j\colon j = 1,\dots,m\}$ be a set of points such that $z_{j-1} < v_{j-1} < u_j < z_j$ for $j = 1,\dots,m$ and

$$\max_{1\leq j\leq m} [\mathrm{Osc}(h;(z_{j-1},v_{j-1}]) \vee \mathrm{Osc}(h;[u_j,z_j))] < \epsilon/[2\|h\|_\infty C(h)^2(m+1)]. \tag{7.14}$$

Let $\lambda := \{z_j, v_{j-1}, u_j\colon j = 0,\dots,m\}$ and let $\kappa = \{x_i\colon i = 0,\dots,n\} \in \mathrm{PP}\,([a,b])$ be a refinement of λ. Define the sets I_r, I_l as in the above proof of statement (1), and for each $j = 0,\dots,m$, let $i(j) \in \{0,1,\dots,n\}$ be such that $x_{i(j)} = z_j$. Let $\mathfrak{D}_\kappa$ be the $\mathfrak{I}$-partition of $[a,b]$ consisting of the intervals: $\{z_0\}, J_1,\dots,J_{i(j)}, \{z_j\}, J_{i(j)+1},\dots,J_n, \{z_m\}$, where, for $j = 0,\dots,m-1$,

$$J_{i(j+1)} := \begin{cases} (x_{i(j+1)-1}, z_{j+1}) & \text{if } i(j+1) \in I_r \\ [x_{i(j+1)-1}, z_{j+1}) & \text{if } i(j+1) \in I_l, \end{cases} \quad \text{and for } i(j)+1 < i < i(j+1)$$

$$J_{i(j)+1} := \begin{cases} (z_j, x_{i(j)+1}] & \text{if } i(j)+1 \in I_r \\ (z_j, x_{i(j)+1}) & \text{if } i(j)+1 \in I_l \end{cases} \qquad J_i := \begin{cases} [x_{i-1}, x_i] & \text{if } i-1 \in I_l,\, i \in I_r \\ (x_{i-1}, x_i] & \text{if } i-1, i \in I_r \\ [x_{i-1}, x_i) & \text{if } i-1, i \in I_l \\ (x_{i-1}, x_i) & \text{if } i-1 \in I_r,\, i \in I_l. \end{cases}$$

Then $\mathfrak{D}_\kappa \in \mathrm{IP}_2([a,b])$ and $\mathfrak{D}_\kappa \sqsupseteq \mathfrak{D}_\epsilon$. Due to the assumption of left- or right-continuity we have $\mu_h(J_{i(j+1)}) = h(z_{j+1}-) - h(x_{i(j+1)-1})$, $\mu_h(J_{i(j)+1}) = h(x_{i(j)+1}) - h(z_j+)$ and $\mu_h(J_i) = h(x_i) - h(x_{i-1})$. Let $B_n := [\mathbb{I} + \mu_h(\{z_m\})][\mathbb{I} + \mu_h(J_n)]$,

$$B_i := \begin{cases} [\mathbb{I} + \mu_h(\{z_j\})][\mathbb{I} + \mu_h(J_{i(j)})] & \text{if } i(j) \in I_r, \\ [\mathbb{I} + \mu_h(J_{i(j)+1})][\mathbb{I} + \mu_h(\{z_j\})] & \text{if } i(j) \in I_l, \\ \mathbb{I} + \mu_h(J_i) & \text{if } i \in \{1, \dots, n-1\} \setminus (I_l \cup I_r) \end{cases}$$

and $B_1 := (\mathbb{I} + \mu_h(J_1))(\mathbb{I} + \mu_h(\{z_0\}))$. Using a telescoping sum representation, we get

$$T := P(h;\kappa) - P_{\mathfrak{J}}(\mu_h; \mathfrak{D}_\kappa) = \prod_{i=1}^{n} (\mathbb{I} + \Delta^i h) - \prod_{i=1}^{n} B_i = [\mathbb{I} + \Delta^n h - B_n] \prod_{i=1}^{n-1} B_i$$

$$+ \sum_{j=1}^{m-1} \Big(\prod_{i=i(j)+1}^{n} (\mathbb{I} + \Delta^i h) \Big) [\mathbb{I} + \Delta^{i(j)} h - B_{i(j)}] \Big(\prod_{i=1}^{i(j)-1} B_i \Big) + \prod_{i=2}^{n} (\mathbb{I} + \Delta^i h) [\mathbb{I} + h(x_1) - B_1].$$

Then using (7.12) and (7.14), we get the bound

$$\|T\| \le C(h) \|\Delta^- h(b)[h(b-) - h(x_{n-1})]\| + \sum_{j=1}^{m-1} C(h)^2 H_j + C(h) \|h(a+)[h(x_1) - h(a+)]\| < \epsilon,$$

where

$$H_j := \begin{cases} \|\Delta^- h(z_j)[h(z_j) - h(x_{i(j)-1})]\| & \text{if } i(j) \in I_r \\ \|[h(x_{i(j)+1}) - h(z_j+)]\Delta^+ h(z_j)\| & \text{if } i(j) \in I_l. \end{cases}$$

By (7.13), it then follows that $\| \prod_{[a,b]} (\mathbb{I} + d\mu_h) - P(h;\kappa)\| < 2\epsilon$ for any point partition $\kappa \supset \lambda$. Since $\epsilon > 0$ is arbitrary, (2) holds. The proof of Theorem 7.3 is complete. $\square$

As the following example shows, the one–sided continuity assumptions of the preceding theorem cannot be dropped.

Example 7.4. For $\xi \in [0,1]$, let $h_\xi(x) := 0$ for $x \in [0,1)$, $h_\xi(1) := \xi$ and $h_\xi(x) := 1$ for $x \in (1,2]$. Then the product integrals with respect to h_ξ and μ_{h_ξ} exist, and

$$\prod_0^2 (1 + dh_\xi) = (1+\xi)(2-\xi) \quad \text{and} \quad \prod_{[0,2]} (1 + d\mu_{h_\xi}) = 2.$$

Therefore the two product integrals have the same value if and only if either $\xi = 0$ or $\xi = 1$; that is if and only if h_ξ is either left-continuous or right-continuous.

Recall that $\mathcal{I}_p([a,b], \mathbb{B})$ is the set of all $\mu \in \mathcal{I}([a,b]; \mathbb{B})$ with bounded p-variation as defined by (4.38). Next are the main facts about the product integral to be used in the subsequent sections.

Corollary 7.5. *Let $\mu \in \mathcal{I}_p([a,b];\mathbb{B})$, $1 \le p < 2$, be additive and upper continuous. Then the product integral with respect to μ is defined over any subinterval of $[a,b]$ and the mapping*

$$\mathfrak{I}([a,b]) \ni J \mapsto \prod_J (\mathbb{I} + d\mu) \in \mathbb{B} \tag{7.15}$$

is a multiplicative and upper continuous interval function on $[a,b]$ with bounded p-variation.

Proof. The functions $R_{\mu,a}$ and $L_{\mu,a}$ defined by (4.29), are respectively right–continuous and left–continuous on (a,b), and both are 0 at a. By Proposition 4.37 and Theorem 4.39, $\mu = \mu_h$ with $h = R_{\mu,a}$, or $h = L_{\mu,a}$, and h has bounded p-variation. Therefore the existence of the product integral with respect to μ over any subinterval of $[a,b]$ follows from Theorem 7.3 above and Theorem 4.23 of [29]. To see that (7.15) is multiplicative let $a < c < b$. Then by Theorem 7.3 above, Corollary 4.24 and Lemma 5.1 of [29], we have

$$\begin{aligned}\prod_{[a,b]} (\mathbb{I} + d\mu) &= \prod_a^c (\mathbb{I} + dR_{\mu,a}) \prod_{c+}^b (\mathbb{I} + dR_{\mu,a}) = \prod_{[a,c]} (\mathbb{I} + d\mu) \prod_{(c,b]} (\mathbb{I} + d\mu) \\ &= \prod_a^{c-} (\mathbb{I} + dL_{\mu,a}) \prod_c^b (\mathbb{I} + dL_{\mu,a}) = \prod_{[a,c)} (\mathbb{I} + d\mu) \prod_{[c,b]} (\mathbb{I} + d\mu).\end{aligned}$$

Similarly using Theorem 7.3 above and Lemma 5.2 of [29] one can show that (7.15) is upper continuous at $\emptyset$. This yields upper continuity by Theorem 4.35. The interval function (7.15) has bounded p-variation by an analogous result for the indefinite product integral with respect to point functions proved in Proposition 5.3 of [29] and by Theorem 4.39 and Lemma 4.13. Thus the conclusion of Corollary 7.5 holds. $\square$

The condition on p-variation in the preceding statement is sufficient but not necessary. In the case of real–valued functions we have necessary and sufficient conditions for existence of a non–zero product integral. The following statement is Theorem 4.4 of Dudley and Norvaiša [29].

Theorem 7.6. *For a real–valued function h on $[a,b]$, the product integral* (7.8) *exists and is non zero if and only if the following two conditions hold:*
(1) $h \in \mathcal{W}_2^*[a,b]$;
(2) $\Delta^- h(x) \neq -1 \neq \Delta^+ h(y)$ *for* $a < x \le b$, $a \le y < b$.

Let g be a $k \times k$ matrix-valued function defined on an interval $[a,b]$ and let $\kappa = \{x_i \colon i = 0, \dots, n\}$ be a partition of $[a,b]$. Volterra [110, Opere, I, p. 235] defined the product integral of the function g over $[a,b]$ to be the limit

$$\lim_{\kappa} \prod_{i=1}^{n} [\mathbb{I} + g(y_i)(x_i - x_{i-1})] \tag{7.16}$$

if it exists, where $\mathbb{I}$ is the identity matrix, y_i are any points from $[x_{i-1}, x_i]$ for $i = 1, \dots, n$ and the limit is taken as the mesh $|\kappa|$ of the partition κ tends to zero. He proved the existence of the limit under conditions analogous to those under which Riemann [94] proved the existence of his integral. Masani [79, Sec. V] proved among other things that for g with values in a Banach algebra the limit (7.16) exists if and only if the function g is Riemann integrable. Next we prove a similar result for the (RYS) integral. Let f be a real-valued function on $[a,b]$

and let $h \in \mathcal{R}[a,b]$. For $\kappa = \{x_i\colon i = 0, \ldots, n\} \in \mathrm{PP}([a,b])$ and open intermediate partition $\sigma = \{y_i\colon i = 1, \ldots, n\}$ of κ, let

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

The subscript Y for P_Y refers to the similarity between this product and the corresponding Young sum.

Definition 7.7. Let f be a function on $[a,b]$ and let $h \in \mathcal{R}[a,b]$. The *product integral* of f with respect to h over $[a,b]$ exists and has value A, if for each $\epsilon > 0$ there exists a partition $\lambda \in \mathrm{PP}([a,b])$ such that

$$|P_Y(f,h;\kappa,\sigma) - A| < \epsilon \tag{7.17}$$

for each refinement κ of λ and open intermediate partition σ of κ. If the product integral is defined then we write

$$\prod_a^b (1 + f\,dh) := A. \tag{7.18}$$

A product $\prod_i A_i$ will be said to *converge absolutely* if $\sum_i |A_i - 1| < \infty$ and $A_i \neq -1$ for all i. Let

$$\chi[z] := (1+z)e^{-z}, \qquad z \in \mathbb{R}. \tag{7.19}$$

For $J \subset [a,b]$ and $\delta > 0$, let $I^{\sharp}(\delta; J) := \{x \in J\colon |\Delta^{\sharp}_{a,b}(f)| > \delta\}$, where $\sharp = -$ and $\sharp = +$. Let

$$\prod_J \chi[f\Delta^{\pm} h] := \lim_{\delta \downarrow 0} \prod_{x \in I^-(\delta;J)} \chi[f\Delta^-_{a,b} h](x) \prod_{x \in I^+(\delta;J)} \chi[f\Delta^+_{a,b} h](x) \tag{7.20}$$

provided the two products converge absolutely.

Theorem 7.8. *Let f be a bounded function on $[a,b]$, $h \in \mathcal{W}_2^*[a,b]$, both real–valued, and*

$$f(y)\Delta^- h(y) \neq -1 \neq f(x)\Delta^+ h(x) \tag{7.21}$$

for $a \le x < y \le b$. Then the integral $(RYS)\int_a^b f\,dh$ exists if and only if the product integral (7.18) *does, and if so they satisfy the relation*

$$\prod_a^b (1 + f\,dh) = \exp\Big\{(RYS)\int_a^b f\,dh\Big\} \prod_{[a,b]} \chi(f\Delta^{\pm} h). \tag{7.22}$$

Remarks. The product over jumps is as in the stochastic case, e.g. when $f \equiv 1$ ([15]). Condition (7.21) does not restrict small jumps of h, that is, (7.21) holds if $|\Delta^- h(y)| \vee |\Delta^+ h(x)| < 1/\|f\|_\infty$ for $a \le x < y \le b$.

Proof. A Taylor series expansion with remainder for the function (7.19) yields

$$\chi(u) = 1 - \theta(u)u^2, \quad \text{where } 1/(4\sqrt{e}) \le \theta(u) \le 3\sqrt{e}/4 \text{ for } |u| \le 1/2.$$

Thus for any finite set $\nu \subset (a,b)$ of points x such that $|[f\Delta^- h](x)| \vee |[f\Delta^+ h](x)| \le 1/2$, we have the bound

$$\sum_{\nu} |1 - \chi(f\Delta^{\pm} h)| \le (3\sqrt{e}/4)\|f\|_\infty^2 \mathfrak{S}_2(h)^2 < \infty.$$

That is, the product $\prod_{[a,b]}\chi(f\Delta^{\pm}h)$ in (7.22) converges absolutely.

By another Taylor series expansion with remainder, it follows that

$$\log(1+u) = u - \theta(u)u^2, \quad \text{ where } 2/9 \le \theta(u) \le 2 \text{ for } |u| \le 1/2. \tag{7.23}$$

Let $\epsilon > 0$. There exists $\lambda = \{z_j\colon j = 0,\dots,m\} \in \mathrm{PP}([a,b])$ satisfying the following four properties:

(1) $D := I^-(1/2;[a,b]) \cup I^+(1/2;[a,b]) \subset \lambda$;

(2) $\mathrm{Osc}(h;(z_{j-1},z_j)) < 1/(2\|f\|_\infty)$ for $j = 1,\dots,m$;

(3) for each finite set $\nu \subset (a,b)$ disjoint from λ,

$$\sum_\nu [(\Delta^- h)^2 + (\Delta^+ h)^2] < \epsilon/(4\|f\|_\infty);$$

(4) for each refinement $\kappa = \{x_i\colon i = 0,\dots,n\}$ of λ,

$$\sum_{i=1}^n [h(x_i-) - h(x_{i-1}+)]^2 < \epsilon/(4\|f\|_\infty).$$

The existence of λ follows from Corollary 4.11, Theorem 4.10, $\mathfrak{S}_2(f) < \infty$ and Lemma 4.18, applied in that order. Let $\kappa = \{x_i\colon i = 0,\dots,n\}$ be a refinement of λ and let $\sigma = \{y_i\colon i = 1,\dots,n\}$ be an open intermediate partition of κ. Then let

$$\Delta_i^\sharp := \Delta_i^\sharp(\kappa) := [f\Delta^\sharp h](x_i) \quad \text{and} \quad \Delta_i := \Delta_i(\kappa,\sigma) := f(y_i)[h(x_i-) - h(x_{i-1}+)],$$

where $\sharp = -$ and $\sharp = +$. Also, for each index $j = 0,\dots,m$ of λ, let $i(j) \in \{0,\dots,n\}$ be the index of κ such that $x_{i(j)} = z_j$ and let $I := \{0,\dots,n\} \setminus \{i(j)\colon j = 0,\dots,m\}$. By the above conditions (1) and (2), we have

$$\begin{aligned}
\log\{P_Y(f,h;\kappa,\sigma)\} &= \log\Big\{\prod_{j=1}^m (1+\Delta^+_{i(j-1)})(1+\Delta^-_{i(j)})\Big\} + \sum_{i\in I}[\log(1+\Delta_i^-) + \log(1+\Delta_i^+)] \\
&\quad + \sum_{i=1}^n \log(1+\Delta_i) \qquad \text{by (7.23) and (7.20)} \\
&= \log\Big\{\prod_\lambda \chi[f\Delta^{\pm}h]\Big\} + S_Y(f,h;\kappa,\sigma) - R(\kappa,\sigma),
\end{aligned} \tag{7.24}$$

where

$$R(\kappa,\sigma) := \sum_{i=1}^n \theta(\Delta_i)(\Delta_i)^2 + \sum_{i\in I}[\theta(\Delta_i^-)(\Delta_i^-)^2 + \theta(\Delta_i^+)(\Delta_i^+)^2].$$

By the above conditions (3) and (4), it then follows that

$$R(\kappa,\sigma) \le 2\|f\|_\infty \sum_{i=1}^n [h(x_i-) - h(x_{i-1}+)]^2 + 2\|f\|_\infty \sum_{i\in I}[(\Delta^+h(x_i))^2 + (\Delta^-h(x_i))^2] < \epsilon$$

for any refinement κ of λ and open intermediate partition σ of κ. Since $\log\{\prod_\lambda \chi[f\Delta^{\pm}h]\}$ in (7.24) does not depend on κ and σ, the (RYS) integral exists if and only if the corresponding product integral (7.18) does, and relation (7.7) follows. The proof of the theorem is complete. $\square$

7.3 The Duhamel formula

We prove the Duhamel formula, which is an integral representation of a difference between two product integrals. Its discrete analog is a simple algebraic identity:

$$\prod_{i=1}^{n} a_j - \prod_{j=1}^{n} b_j = \sum_{i=1}^{n} \Big(\prod_{j=i+1}^{n+1} a_j \Big)(a_i - b_i)\Big(\prod_{j=0}^{i-1} b_j \Big) \tag{7.25}$$

valid by a telescoping sum for any elements $a_1, \dots, a_n$ and $b_1, \dots, b_n$ of a Banach algebra, with $b_0 := a_{n+1} := \mathbb{I}$. An integral $\neq\!\!\!\!\int_{[a,b]} f(y)\,\rho(dy)\,g(y)$ is defined as a limit under refinement of sums $\sum_{j=1}^{n} f(y_j)\rho(I_j)g(y_j)$ over interval partitions $\{I_j\}_{j=1}^{n}$ of $[a,b]$ and $y_j \in I_j$, $j = 1, \dots, n$. This reduces to the previous definition if $\mathbb{B}$ is commutative.

Theorem 7.9 (Duhamel formula). *Let $\mu, \nu \in \mathcal{I}_p([a,b], \mathbb{B})$, $1 \leq p < 2$, be additive and upper continuous. Then*

$$\prod_{[a,b]} (\mathbb{I} + d\mu) - \prod_{[a,b]} (\mathbb{I} + d\nu) = \neq\!\!\!\!\int_{[a,b]} \prod_{(y,b]} (\mathbb{I} + d\mu)(\mu - \nu)(dy) \prod_{[a,y)} (\mathbb{I} + d\nu), \tag{7.26}$$

where a product integral over the empty set is defined as $\mathbb{I}$.

First proof. Let $R := R_{\mu,a}$ and $L := L_{\nu,a}$ be the functions defined by (4.29). By Theorem 4.39, R and L have bounded p-variation. Thus by Theorem 7.3 and Corollary 7.5, the following product integrals are defined and satisfy the relations:

$$\prod_{[a,b]} (\mathbb{I} + d\mu) - \prod_{[a,b]} (\mathbb{I} + d\nu) = \prod_{a}^{b} (\mathbb{I} + dR) - \prod_{a}^{b} (\mathbb{I} + dL), \qquad \prod_{[a,y)} (\mathbb{I} + d\nu) = \prod_{a}^{y-} (\mathbb{I} + dL)$$

for $a < y \leq b$, and

$$\prod_{(y,b]} (\mathbb{I} + d\mu) = \prod_{y+}^{b} (\mathbb{I} + dR) = \Big(\prod_{y}^{b} (\mathbb{I} + dR) \Big)^{(a+)}$$

for $a \leq y < b$, because R is right–continuous only on (a,b). Here $f^{(a+)}$ is a function on $[a,b]$ defined by $f^{(a+)}(x) := f(x)$ for $a < x \leq b$ and $f^{(a+)}(a) := f(a+)$. Since R is right–continuous on (a,b) and L is left–continuous on (a,b), the Duhamel formula of [29, Theorem 5.5] can be written in terms of the (CY) integral as follows:

$$\begin{aligned} \prod_{a}^{b} (\mathbb{I} + dR) - \prod_{a}^{b} (\mathbb{I} + dL) &= (CY) \int_a^b \Big(\prod_{y}^{b} (\mathbb{I} + dR) \Big)^{(a+)} d(R - L) \Big(\prod_{a}^{y} (\mathbb{I} + dL) \Big)_{-}^{(a)} \\ &= (RYS) \int_a^b \prod_{(y,b]} (\mathbb{I} + d\mu)\, d(R - L) \prod_{[a,y)} (\mathbb{I} + d\nu), \end{aligned}$$

where the last equality holds by Proposition 6.22. Since the (RYS) integral does not depend on values of the integrator at jump points, and since $(\mu - \nu)([a,x]) = R(x) - L(x)$ for all $x \in [a,b]$ except at atoms of $\mu - \nu$, the Duhamel formula (7.26) holds. ☐

We give a second proof of the Duhamel formula which in view of what is used from [29] would give a shorter total proof. It shows how the (RYS) integral appears in (7.26) as compared with the Left Young and Right Young integrals used in the Duhamel formula of [29, Theorem 5.5].

Second proof. Let $\mathfrak{D} = (J_1, \dots, J_N)$ be a Young interval partition of $[a,b]$ and let $\wp(\mathfrak{D}) = \{x_0, x_1, \dots, x_n\}$ be the corresponding point partition. For any $I \subset \{1, \dots, N\}$ and an additive interval function γ, let

$$P_\gamma^{\mathfrak{D}}(\cup_{i\in I} J_i) := \prod_{i\in I}[\mathbb{I} + \gamma(J_i)] \quad \text{and} \quad P_\gamma^{\mathfrak{D}}(\emptyset) := \mathbb{I}.$$

Define step functions $G^{\mathfrak{D}}$ and $F^{\mathfrak{D}}$ on $[a,b]$ by

$$G^{\mathfrak{D}}(y) := \begin{cases} P_\mu^{\mathfrak{D}}([x_i, b]) & \text{if } y \in (x_{i-1}, x_i) \text{ for some } i = 1, \dots, n \\ P_\mu^{\mathfrak{D}}((x_i, b]) & \text{if } y = x_i \text{ for some } i = 0, \dots, n \end{cases}$$

and

$$F^{\mathfrak{D}}(y) := \begin{cases} P_\nu^{\mathfrak{D}}([a, x_{i-1}]) & \text{if } y \in (x_{i-1}, x_i) \text{ for some } i = 1, \dots, n \\ P_\nu^{\mathfrak{D}}([a, x_i)) & \text{if } y = x_i \text{ for some } i = 0, \dots, n. \end{cases}$$

Letting $\gamma := \mu - \nu$, by (7.25), we have

$$\begin{aligned} \fint_a^b G^{\mathfrak{D}}\, d\gamma\, F^{\mathfrak{D}} &= \sum_{i=1}^n P_\mu^{\mathfrak{D}}([x_i, b])\gamma((x_{i-1}, x_i))P_\nu^{\mathfrak{D}}([a, x_{i-1}]) + \sum_{i=0}^n P_\mu^{\mathfrak{D}}((x_i, b])\gamma(\{x_i\})P_\nu^{\mathfrak{D}}([a, x_i)) \\ &= P_\mu^{\mathfrak{D}}([a,b]) - P_\nu^{\mathfrak{D}}([a,b]). \end{aligned} \tag{7.27}$$

In the next lemma we construct an increasing sequence $\{\mathfrak{D}_m\colon m \geq 1\}$ of Young partitions of $[a,b]$ such that $D := \cup_m \wp(\mathfrak{D}_m)$ is dense in $[a,b]$, contains all the atoms of γ and, for each $y \in D$,

$$\lim_{m\to\infty} G^{\mathfrak{D}_m}(y) = \prod_{(y,b]} (\mathbb{I} + d\mu), \qquad \lim_{m\to\infty} F^{\mathfrak{D}_m}(y) = \prod_{[a,y)} (\mathbb{I} + d\nu), \tag{7.28}$$

$$\lim_{m\to\infty} P_\mu^{\mathfrak{D}_m}([a,b]) = \prod_{[a,b]} (\mathbb{I} + d\mu) \quad \text{and} \quad \lim_{m\to\infty} P_\nu^{\mathfrak{D}_m}([a,b]) = \prod_{[a,b]} (\mathbb{I} + d\nu).$$

Then (7.26) will follow provided

$$\lim_{m\to\infty} \fint_a^b G^{\mathfrak{D}_m}\, d\gamma\, F^{\mathfrak{D}_m} = \fint_a^b \prod_{(y,b]} (\mathbb{I} + d\mu)\gamma(dy) \prod_{[a,y)} (\mathbb{I} + d\nu). \tag{7.29}$$

Lemma 7.10. *There exists an increasing sequence $\{\mathfrak{D}_m\colon m \geq 1\}$ of Young interval partitions of $[a,b]$ such that $D := \cup_m \wp(\mathfrak{D}_m)$ is dense in $[a,b]$, contains all the atoms of γ,*

$$\lim_{m\to\infty} P_\mu^{\mathfrak{D}_m}([y,b]) = \prod_{[y,b]} (\mathbb{I} + d\mu), \qquad \lim_{m\to\infty} P_\nu^{\mathfrak{D}_m}([a,y]) = \prod_{[a,y]} (\mathbb{I} + d\nu), \tag{7.30}$$

and (7.28) *holds for $y \in D$.*

Proof. Let $\{\xi_m\colon m \geq 1\}$ and $\{\zeta_m\colon m \geq 1\}$ denote the atoms of μ and ν, respectively. For each $m \geq 1$, put $r_m := \{a + l2^{-m}\colon l = 1, \dots, [2^m(b-a)]\}$. We will proceed recursively. By Corollary 7.5, there exists a Young partition $\mathfrak{D}_1$ such that, for each refinement $\mathfrak{D}$ of $\mathfrak{D}_1$,

$$\|P_\mu^{\mathfrak{D}}([a,b]) - \prod_{[a,b]} (\mathbb{I} + d\mu)\| \vee \|P_\nu^{\mathfrak{D}}([a,b]) - \prod_{[a,b]} (\mathbb{I} + d\nu)\| < 1.$$

We may and do assume that ξ_1, ζ_1 and the points of r_1 are endpoints of some elements of $\mathfrak{D}_1$. Assume that $m \geq 2$ and we already have Young partitions $\mathfrak{D}_1, \dots, \mathfrak{D}_{m-1}$ with $\wp(\mathfrak{D}_{m-1}) =$

$\{x_0^{m-1}, \ldots, x_{n(m-1)}^{m-1}\}$. To construct $\mathfrak{D}_m$ using Corollary 7.5, for each $i = 0, \ldots, n(m-1)-1$, one can find a Young partition $\mathfrak{D}_m^{i,\mu}$ of $[x_i^{m-1}, b]$ such that

$$\Big\| P_\mu^{\mathfrak{D}}([x_i^{m-1}, b]) - \prod_{[x_i^{m-1}, b]} (\mathbb{I} + d\mu) \Big\| \vee \Big\| P_\mu^{\mathfrak{D}}((x_i^{m-1}, b]) - \prod_{(x_i^{m-1}, b]} (\mathbb{I} + d\mu) \Big\| < 1/m$$

for all refinements $\mathfrak{D}$ of $\mathfrak{D}_m^{i,\mu}$. Similarly for each $i = 1, \ldots, n(m-1)$, one can find a Young partition $\mathfrak{D}_m^{i,\nu}$ of $[a, x_i^{m-1}]$ such that

$$\Big\| P_\nu^{\mathfrak{D}}([a, x_i^{m-1}]) - \prod_{[a, x_i^{m-1}]} (\mathbb{I} + d\mu) \Big\| \vee \Big\| P_\nu^{\mathfrak{D}}([a, x_i^{m-1})) - \prod_{[a, x_i^{m-1})} (\mathbb{I} + d\mu) \Big\| < 1/m$$

for all refinements $\mathfrak{D}$ of $\mathfrak{D}_m^{i,\nu}$. We may and do assume that among the endpoints of elements of the new partitions $\mathfrak{D}_m^{i,\mu}$ and $\mathfrak{D}_m^{i,\nu}$ are all points $r_m \cap [x_i^{m-1}, b]$ and $r_m \cap [a, x_i^{m-1}]$, respectively. Let $\mathfrak{D}_m$ be the Young partition of $[a, b]$ such that

$$\wp(\mathfrak{D}_m) = \{\xi_m, \zeta_m\} \cup \Big(\cup_{i=0}^{n(m-1)-1} \wp(\mathfrak{D}_m^{i,\mu}) \Big) \cup \Big(\cup_{i=1}^{n(m-1)} \wp(\mathfrak{D}_m^{i,\nu}) \Big).$$

We arrive recursively at a sequence of Young partitions $\{\mathfrak{D}_m\colon m \geq 1\}$. By construction the set $D := \cup_m \wp(\mathfrak{D}_m)$ is dense in $[a, b]$, contains all atoms of γ and relations (7.28), (7.30) hold. □

To prove (7.29) we will use a convergence theorem for (RYS) integrals. The theorem applies to sequences $\{f_m\} = \{f_m\colon m \geq 1\}$ of functions which have uniformly bounded p-variation and converge locally uniformly. That is, we say that $\{f_m\}$ is a *$\mathcal{W}_p$–sequence* if $\sup_m \|f_m\|_{[p]} < \infty$ and $\{f_m\}$ *converges locally uniformly to f at y*, if given $\epsilon > 0$, there is an m_0 and a $\delta > 0$ such that

$$\|f_m(z) - f(z)\| < \epsilon \tag{7.31}$$

for all $m \geq m_0$ and all z distant less than δ from y. To show local uniform convergence in the proof below we first show *uniform convergence on the left at y* meaning that (7.31) holds for $z < y$, and the other conditions are the same. Then we show *uniform convergence on the right at y* meaning that (7.31) holds for $z \geq y$.

Proposition 7.11. *Let $h \in \mathcal{W}_p([a, b]; \mathbb{B})$ and let $\{g_m\}$, $\{f_m\}$ be $\mathcal{W}_q$–sequences of $\mathbb{B}$–valued functions on $[a, b]$, where $p^{-1} + q^{-1} > 1$, $p, q > 0$. Suppose that $\{g_m\}$ and $\{f_m\}$ converge pointwise on a dense set to functions $g, f \in \mathcal{W}_q([a, b]; \mathbb{B})$, respectively, and that both sequences converge locally uniformly to their limits at each discontinuity of h. Then*

$$\lim_{m\to\infty} (RYS) \int_a^b g_m \, dh \, f_m = (RYS) \int_a^b g \, dh \, f. \tag{7.32}$$

Proof. By Proposition 6.22, the (RYS) integrals exist and are equal to the corresponding (CY) integrals. Thus the stated convergence theorem follows from Proposition 3.33 of [29]. □

Continuation of the second proof of Theorem 7.9. We check the hypotheses of Proposition 7.11 with $q = p$ only for $g_m = G^m := G^{\mathfrak{D}_m}$, $m \geq 1$. The hypotheses for $\{F^{\mathfrak{D}_m}\}$ can be verified similarly. For the rest of the proof we use without special mention Theorems 4.39 and 7.3 relating the p-variation and the product integrals for μ, $R_{\mu,a}$ and $L_{\mu,a}$. Let

$$C(\mu) := \sup \{ \|P_\mu^{\mathfrak{D}}(J)\|\colon \mathfrak{D} \in \mathrm{IP}_1(J),\ J \in \mathfrak{I}_{os}([a, b]) \}.$$

Since $v_p(\mu) < \infty$, by Lemma 4.19 of [29], $C(\mu) < \infty$. Suppose for a moment that $\rho := C_p\|\mu\|_{(p)} < 1$ for a suitable constant C_p from Lemma 4.18 of [29]. Then for $\kappa \in \mathrm{PP}\,([a,b])$, by inequality (4.52) of [29], we have

$$\begin{aligned} s_p(G^m;\kappa) &\le C(\mu)^p \sup\Big\{\sum_j \|P_\mu^{\mathfrak{D}_m}(J_j) - \mathbb{I}\|^p \colon \{J_j\} \in \mathrm{IP}\,(\mathfrak{I}_{os}([a,b])),\ \{J_j\} \subset \mathfrak{D}_m\Big\} \\ &\le [C(\mu)/(1-\rho)]^p \sup\Big\{\sum_j v_p(\mu; J_j) \colon \{J_j\} \in \mathrm{IP}\,(\mathfrak{I}_{os}([a,b])),\ \{J_j\} \subset \mathfrak{D}_m\Big\} \\ &\le [C(\mu)/(1-\rho)]^p \|\mu\|_{(p)}^p. \end{aligned}$$

Since κ and m are arbitrary, $\{G^m\}$ is a $\mathcal{W}_p$–sequence in the case when $\|\mu\|_{(p)} < 1/C_p$. In the general case we use Lemma 4.7 above to split the interval $[a,b]$ by a Young partition $\mathfrak{D}$ so that the above bound holds on each open subinterval of $\mathfrak{D}$. This yields that $\{G^m\}$ is a $\mathcal{W}_p$–sequence.

Let y be an atom of $\gamma = \mu - \nu$. Then $y \in D = \cup_m \wp(\mathfrak{D}_m)$. First, we show that if $y > a$,

$$G^m \to \prod_{(\cdot,b]} (\mathbb{I} + d\mu) \quad \text{as } m \to \infty, \text{ uniformly on the left at } y. \tag{7.33}$$

Let $\epsilon > 0$. By Lemma 7.10, there exists m_0 so large that, for all $m \ge m_0$, $y \in \wp(\mathfrak{D}_m)$ and

$$\Big\|G^m(y-) - \prod_{[y,b]} (\mathbb{I} + d\mu)\Big\| < \epsilon. \tag{7.34}$$

By Lemmas 2.18 and 5.2 of [29], there exists $\delta \in (0, y-a)$ such that

$$v_p(\mu; [y-\delta, y)) < \{[\epsilon/(2C(\mu))] \wedge (2C_p)^{-1}\}^p, \tag{7.35}$$

where C_p is as before a suitable constant from Lemma 4.18 of [29], and for all $z \in [y-\delta, y)$,

$$\Big\|\prod_{(z,b]} (\mathbb{I} + d\mu) - \prod_{[y,b]} (\mathbb{I} + d\mu)\Big\| = \Big\|\prod_{z+}^{b} (\mathbb{I} + dL_{\mu,a}) - \prod_{y}^{b} (\mathbb{I} + dL_{\mu,a})\Big\| \le \epsilon \tag{7.36}$$

because $L_{\mu,a}$ is left–continuous at y. Thus by inequality (4.52) of [29] and (7.35) above, we have for $z \in [y-\delta, y)$,

$$\begin{aligned} \|G^m(z) - G^m(y-)\| &\le C(\mu) \max\{\|P_\mu^{\mathfrak{D}_m}(\{x,y)) - \mathbb{I}\| \colon x \in \wp(\mathfrak{D}_m) \cap [y-\delta, y)\} \\ &\le 2C(\mu) v_p(\mu; [y-\delta, y))^{1/p} < \epsilon, \end{aligned}$$

where $\{\cdot, y)$ is an interval right open at y. Thus by (7.34) and (7.36), for each $m \ge m_0$ and $z \in [y-\delta, y)$, we get

$$\begin{aligned} \Big\|G^m(z) - \prod_{(z,b]} (\mathbb{I} + d\mu)\Big\| &\le \|G^m(z) - G^m(y-)\| + \Big\|G^m(y-) - \prod_{[y,b]} (\mathbb{I} + d\mu)\Big\| \\ &\quad + \Big\|\prod_{[y,b]} (\mathbb{I} + d\mu) - \prod_{(z,b]} (\mathbb{I} + d\mu)\Big\| < 3\epsilon. \end{aligned}$$

Therefore (7.33) holds.

Second, we show that if $y < b$,

$$G^m \to \prod_{(\cdot,b]} (\mathbb{I} + d\mu) \quad \text{as } m \to \infty, \text{ uniformly on the right at } y. \tag{7.37}$$

Let $\epsilon > 0$. By Lemma 7.10, there exists m_0 so large that, for all $m \geq m_0$, $y \in \wp(\mathfrak{D}_m)$ and

$$\Big\| G^m(y) - \prod_{(y,b]} (\mathbb{I} + d\mu) \Big\| < \epsilon. \tag{7.38}$$

By Lemmas 2.18 and 5.2 of [29], there exists $\delta \in (0, b - y)$ such that

$$v_p(\mu; (y, y + \delta]) < \{[\epsilon/(2C(\mu))] \wedge (2C_p)^{-1}\}^p, \tag{7.39}$$

where C_p is again a suitable constant from Lemma 4.18 of [29], and for all $z \in (y, y + \delta]$,

$$\Big\| \prod_{(z,b]} (\mathbb{I} + d\mu) - \prod_{(y,b]} (\mathbb{I} + d\mu) \Big\| = \Big\| \prod_{z+}^{b} (\mathbb{I} + dL_{\mu,a}) - \prod_{y+}^{b} (\mathbb{I} + dL_{\mu,a}) \Big\| \leq \epsilon. \tag{7.40}$$

We can and do assume that $y + \delta \in \wp(\mathfrak{D}_m)$ for $m \geq m_0$. Then by inequality (4.52) of [29] and (7.39) above, we have for $z \in (y, y + \delta]$,

$$\begin{aligned} \|G^m(z) - G^m(y)\| &\leq C(\mu) \max\{\|P_\mu^{\mathfrak{D}_m}((y, x\}) - \mathbb{I}\|: x \in \wp(\mathfrak{D}_m) \cap (y, y + \delta)\} \\ &\leq 2C(\mu) v_p(\mu; (y, y + \delta])^{1/p} < \epsilon, \end{aligned}$$

where $(y, \cdot\}$ is an interval left open at y. Thus by (7.38) and (7.40), for each $m \geq m_0$ and $z \in (y, y + \delta]$, we get

$$\begin{aligned} \Big\| G^m(z) - \prod_{(z,b]} (\mathbb{I} + d\mu) \Big\| &\leq \|G^m(z) - G^m(y)\| + \Big\| G^m(y) - \prod_{(y,b]} (\mathbb{I} + d\mu) \Big\| \\ &\quad + \Big\| \prod_{(y,b]} (\mathbb{I} + d\mu) - \prod_{(z,b]} (\mathbb{I} + d\mu) \Big\| < 3\epsilon. \end{aligned}$$

Therefore (7.37) holds. This together with (7.33) yields that $\{G^m\}$ converges locally uniformly at each atom of γ. By Lemma 7.10 it also converges densely. Hence all the hypotheses of Theorem 7.11 are satisfied, and hence (7.29) holds. The second proof of the Duhamel formula (7.26) is complete. $\square$

7.4 Integral equations

First, as an example, let's consider the integral equation

$$G(x) = C + \int_0^x G(y)\, dF(y).$$

Let $0 \leq x \leq 2$ and $F := 1_{[1,\infty)}$. Then for the integrals we are considering other than the (MRS) and (RRS) integrals (which may be undefined), the equation gives

$$G(x) = \begin{cases} C, & \text{if } 0 \leq x < 1; \\ C + G(1), & \text{if } 1 \leq x \leq 2. \end{cases}$$

For $x = 1$, this yields a contradiction (unless $C = 0$). In accordance with some developments in analysis (see below) and stochastic analysis (e.g. Doléans-Dade [15]) it seems desirable that in such an equation the integrand should be left–continuous while the solution is right–continuous. Writing

$$G(x) = G_+^{(2)}(x) = C + \int_0^x G_-^{(0)}(y)\, dF(y)$$

we get

$$G(x) = \begin{cases} C, & \text{if } 0 \le x < 1; \\ 2C, & \text{if } 1 \le x \le 2 \end{cases}$$

and now there is no contradiction. The integral can be any of those we are considering (in Figure 2.1) except for the (MRS) integral but now including the (RRS) integral.

In analysis several approaches have been suggested to solving a linear Stieltjes type integral equation with respect to a possibly discontinuous function F of bounded variation. Using an integral equivalent to the (RYS) integral, Hildebrandt [51, p. 359] proved that if a matrix–valued function F with bounded variation has a finite number of discontinuities on $[a, b]$ then the integral equation

$$G(x) = G(a) + \int_a^x dF(y)\, G(y), \qquad a \le x \le b, \tag{7.41}$$

has a unique solution if and only if the matrices $\mathbb{I} - \Delta^- F(x)$ have inverses for all points of discontinuity of F. Thus a situation similar to the above example is excluded. Hildebrandt considers cases where F may have non–zero jumps $(\Delta^+ F)(u)$ and $(\Delta^- F)(u)$ on the two sides of the same point and seeks an actual solution of (7.41) with $G(y)$ (not $G(y-)$). The solution is a kind of product integral where the factor at a point u where F has jumps on one side or the other is $[\mathbb{I} + \Delta^+ F(u)][\mathbb{I} - \Delta^- F(u)]^{-1}$. MacNerney [77] used the left Cauchy and right Cauchy integrals defined as (RRS) integrals except that tagged partitions $(\{x_i\}, \{y_i\})$ are used such that $y_i = x_{i-1}$ and $y_i = x_i$, respectively. Hönig [53] used the interior (RRS) integral defined as the (RRS) integral except that the tagged partitions $(\{x_i\}, \{y_i\})$ defining the integral satisfy the relation: $x_{i-1} < y_i < x_i$. The book of Schwabik, Tvrdý and Vejvoda [101] treats linear integral equations with respect to functions of bounded variation using the (HK) integral.

The form of the Duhamel formula from the preceding section suggests the following:

Definition 7.12. Let ν be a $\mathbb{B}$–valued additive and upper continuous interval function on $[a, b]$. We say that a $\mathbb{B}$–valued interval function α satisfies the *forward linear integral equation with respect to* ν, if the integral $\not\fint_{[a,b]} \nu(dx)\, \alpha([a, x))$ is defined, and for all $y \in [a, b]$,

$$\alpha([a, y]) = \mathbb{I} + \not\fint_{[a,y]} \nu(dx)\, \alpha([a, x)). \tag{7.42}$$

Similarly, we say that a $\mathbb{B}$–valued interval function α satisfies the *backward linear integral equation with respect to* ν, if the integral $\not\fint_{[a,b]} \alpha((x, b])\, \nu(dx)$ is defined, and for all $y \in [a, b]$,

$$\alpha([y, b]) = \mathbb{I} + \not\fint_{[y,b]} \alpha((x, b])\, \nu(dx). \tag{7.43}$$

Theorem 7.13. *Let $\nu \in \mathcal{I}_p([a, b]; \mathbb{B})$, $1 \le p < 2$ be additive and upper continuous. Then the product integral $\prod(\mathbb{I} + d\nu)$ is defined, is in $\mathcal{I}_p([a, b]; \mathbb{B})$ and is the unique solution in $\mathcal{I}_r([a, b]; \mathbb{B})$, for any $r \ge p$ such that $1/p + 1/r > 1$, of the forward integral equation* (7.42).

Proof. The product integral $\prod(\mathbb{I}+d\nu)$ exists and is in $\mathcal{I}_p([a,b];\mathbb{B})$ by Corollary 7.5. Taking $\mu = 0$ in the Duhamel formula (7.26) we get that $\prod(\mathbb{I}+d\nu)$ satisfies the forward linear integral equation with respect to ν. Let μ be another solution in $\mathcal{I}_p([a,b];\mathbb{B})$ and let

$$F(y) := \prod_{[a,y]} (\mathbb{I}+d\nu) - \mu([a,y]) \quad \text{for } a < y \le b \text{ and } F(a) := \mathbb{I} - \mu(\emptyset).$$

Then for some $h \in [\nu]_a$ (cf. (4.34)) and all $y \in [a,b]$,

$$F(y) = (RYS)\int_a^y dh\, F_-^{(a)}.$$

So $F(a) = 0$. By (6.8), $\Delta^+ F(a) = (\Delta^+ h)(a)F(a) = 0$. Thus $F(a+) = 0$. By the Love–Young inequality for the indefinite (RYS) integral (Corollary 4.28), and since $\|F\|_{[p]}$ is a non-increasing function of p (Lemma 4.5), for $r \ge p$ such that $1/p + 1/r > 1$, it follows that

$$\|F\|_{[r]} \le \|F\|_{[p]} \le D_{p,r}\|h\|_{(p)}\|F\|_{[r]}, \tag{7.44}$$

where $D_{p,r} = 2\zeta(p^{-1}+r^{-1})$. Then (7.44) holds for norms over any subinterval $(a,y]$ of $[a,b]$. Taking y close enough to a we will have $\|h\|_{(p)} < 1/D_{p,r}$ by Lemma 4.8, and thus $F \equiv 0$ on $[a,y]$. There is a largest y for which this holds, with $y \le b$. If $y = b$ then we are done. If $y < b$ then we can repeat the argument with y in place of a and contradict maximality of y. So uniqueness is proved. $\square$

7.5 Lyons' inequalities and series expansions

In this section we prove that the product integral with respect to a regulated additive interval function μ on $[a,b]$ with bounded p-variation $1 \le p < 2$, has the series expansion:

$$\prod_{[a,b]} (\mathbb{I}+d\mu) = \mathbb{I} + \mu([a,b]) + \sum_{k\ge 2} \oint_{[a,b]} \mu(dx_1) \oint_{[a,x_1)} \mu(dx_2)\cdots \oint_{[a,x_{k-1})} \mu(dx_k).$$

Proposition 7.14. *For $1 \le p < 2$, let $\mu \in \mathcal{I}_p([a,b];\mathbb{B})$ be additive and upper continuous, and let $f \in \mathcal{W}_p([a,b];\mathbb{B})$. Suppose that f is right-continuous on $[a,b)$, $f(a) = 0$ and $v_p(f;[u,v]) \le \rho^k(v) - \rho^k(u)$ for $a \le u < v \le b$ and $k \ge 1$, where $\rho(x) := v_p(\mu;[a,x])$ for $x \in [a,b]$. Then for the function H defined on $[a,b]$ by $H(x) := \oint_{[a,x]} d\mu\, f_-^{(a)}$, and any $a \le u < v \le b$,*

$$V_p(H;[u,v]) \le C_p\Big(\frac{\rho^{k+1}(v) - \rho^{k+1}(u)}{k+1}\Big)^{1/p}, \tag{7.45}$$

where $C_p := 1 + 4^{1/p}\zeta(2/p)$, and H is right-continuous on $[a,b)$.

The proof of this proposition will be based on two facts. The first one is a part of Theorem 1.1 of Lyons [76], up to the constant 4.

Lemma 7.15. *Let ρ be a non-decreasing function on $[a,b]$, k be a positive integer and $\kappa = \{x_i\colon i = 0,\dots,n\} \in \mathrm{PP}\,([a,b])$. Then there exists an index $i \in \{1,\dots,n-1\}$ such that*

$$[\rho^k(x_i) - \rho^k(x_{i-1})][\rho(x_{i+1}) - \rho(x_i)] \le \frac{4}{(n-1)^2}\Big(\frac{\rho^{k+1}(b) - \rho^{k+1}(a)}{k+1}\Big).$$

Proof. Denoting $\Delta^i\phi := \phi(x_i) - \phi(x_{i-1})$ for a function ϕ, for each $i = 1, \ldots, n-1$, we have

$$\begin{aligned}
\Delta^i\rho^k\Delta^{i+1}\rho &= k\int_{\rho(x_{i-1})}^{\rho(x_i)} u^{k-1}\,du\,\Delta^{i+1}\rho \le k\int_{\rho(x_{i-1})}^{\rho(x_i)} u^{\frac{k-1}{2}}\,du\,[\rho(x_i)]^{\frac{k-1}{2}}[\rho(x_{i+1}) - \rho(x_i)] \\
&\le k\int_{\rho(x_{i-1})}^{\rho(x_i)} u^{(k-1)/2}\,du\int_{\rho(x_i)}^{\rho(x_{i+1})} u^{(k-1)/2}\,du. \qquad (7.46)
\end{aligned}$$

Then using the inequality $2\sqrt{xy} \le x + y$ for $x, y \ge 0$, we get

$$\begin{aligned}
\min_{1\le i\le k-1}\{\Delta^i\rho^k\Delta^{i+1}\rho\} &\le \Big\{\frac{1}{n-1}\sum_{i=1}^{n-1}\sqrt{\Delta^i\rho^k\Delta^{i+1}\rho}\Big\}^2 \\
\text{by (7.46)} \quad &\le \Big\{\frac{1}{n-1}\frac{\sqrt{k}}{2}\Big(\int_{\rho(a)}^{\rho(x_{n-1})} u^{\frac{k-1}{2}}\,du + \int_{\rho(x_1)}^{\rho(b)} u^{\frac{k-1}{2}}\,du\Big)\Big\}^2 \\
&\le \Big\{\frac{\sqrt{k}}{n-1}\int_{\rho(a)}^{\rho(b)} u^{\frac{k-1}{2}}\,du\Big\}^2 = \Big\{\frac{2\sqrt{k}}{n-1}\Big(\frac{\rho^{(k+1)/2}(b) - \rho^{(k+1)/2}(a)}{k+1}\Big)\Big\}^2 \\
&\le \frac{4}{(n-1)^2}\Big(\frac{\rho^{k+1}(b) - \rho^{k+1}(a)}{k+1}\Big),
\end{aligned}$$

where in the last step, the inequality $(d-c)^2 < d^2 - c^2$ for $0 < c < d$ was used. The proof is complete. $\square$

The second fact to be used in proving Proposition 7.14 is the following:

Lemma 7.16. *Let $f, h \in \mathcal{W}_p([a,b];\mathbb{B})$ for $1 \le p < 2$, let f be right–continuous on $[a,b)$ and let h be right–continuous on (a,b). Then for any $\epsilon > 0$ there is a partition $\{z_j\}_{j=0}^m \in \mathrm{PP}\,([a,b])$ such that*

$$\Big\|(RYS)\int_a^b dh\,f_-^{(a)} - \sum_{j=1}^m [h(z_j) - h(z_{j-1})]f(z_{j-1})\Big\| < \epsilon.$$

Proof. Since $\|f_-^{(a)}\|_{(p)} \le \|f\|_{(p)}$, and $f_-^{(a)}$, h have no common one–sided discontinuities, the (RYS) integral exists and equals an (RRS) integral by Theorems 4.26 and 6.13. Thus there exists a partition $\{z_j\}_{j=0}^m$ such that

$$\Big\|(RRS)\int_a^b dh\,f_-^{(a)} - \sum_{j=1}^m [h(z_j) - h(z_{j-1})]f_-^{(a)}(y_j)\Big\| < \epsilon$$

whenever $z_{j-1} \le y_j \le z_j$ for $j = 1, \ldots, m$. We can let $y_j \downarrow z_{j-1}$ for $j = 1, \ldots, m$, giving $f(z_{j-1})$ by right–continuity of f on $[a,b)$. The conclusion follows. $\square$

Proof of Proposition 7.14. Let $a \le u < v \le b$. Since $R_{\mu,a}$ is right–continuous on (a,b), by Proposition 3.25 and Corollary 3.20 of [29], we have

$$H(u) - H(v) = (RYS)\int_u^v dR_{\mu,a}\,f_-^{(u)}.$$

Let $\epsilon > 0$. By Lemma 7.16, there exists $\lambda = \{x_i\colon i = 0, \ldots, n\} \in \mathrm{PP}\,[u,v]$ such that

$$\Big\|(RYS)\int_u^v dR_{\mu,a}\,f_-^{(u)} - S(\lambda)\Big\| < \epsilon,$$

where $S(\lambda) := \sum_{i=1}^{n}[R_{\mu,a}(x_i)-R_{\mu,a}(x_{i-1})]f(x_{i-1})$. Let $\lambda^i := \lambda\backslash\{x_i\}$ for some $i \in \{1,\ldots,n-1\}$. Then we have

$$\begin{aligned}\|S(\lambda)-S(\lambda^i)\| &\le \|R_{\mu,a}(x_{i+1})-R_{\mu,a}(x_i)\|\,\|f(x_i)-f(x_{i-1})\|\\ &\le \{[\rho(x_{i+1})-\rho(x_i)]\,[\rho^k(x_i)-\rho^k(x_{i-1})]\}^{1/p}.\end{aligned}$$

By Lemma 7.15, there exists an index $i \in \{1,\ldots,n-1\}$ such that

$$\|S(\lambda)-S(\lambda^i)\| \le \Big\{\frac{4}{(n-1)^2}\Big(\frac{\rho^{k+1}(v)-\rho^{k+1}(u)}{k+1}\Big)\Big\}^{1/p}.$$

Deleting further points according to Lemma 7.15, we obtain the bound

$$\begin{aligned}\Big\|(RYS)\int_u^v dR_{\mu,a}\, f_-^{(u)}\Big\| &\le 4^{1/p}\sum_{n=2}^{\infty}\frac{1}{(n-1)^{2/p}}\Big(\frac{\rho^{k+1}(v)-\rho^{k+1}(u)}{k+1}\Big)^{1/p}\\ &\quad+\|[R_{\mu,a}(v)-R_{\mu,a}(u)]f(u)\|+\epsilon.\end{aligned}$$

Since $f(a)=0$, we have

$$\|[R_{\mu,a}(v)-R_{\mu,a}(u)]f(u)\|^p \le [\rho(v)-\rho(u)]\rho^k(u) \le \int_{\rho(u)}^{\rho(v)} t^k\,dt = \frac{\rho^{k+1}(v)-\rho^{k+1}(u)}{k+1}.$$

Letting $\epsilon \downarrow 0$, it follows that

$$\Big\|(RYS)\int_u^v dR_{\mu,a}\, f_-^{(u)}\Big\| \le [1+4^{1/p}\zeta(2/p)]\Big(\frac{\rho^{k+1}(v)-\rho^{k+1}(u)}{k+1}\Big)^{1/p}.$$

Using this inequality on subintervals, (7.45) follows. $\square$

Proposition 7.17. *Let $\mu \in \mathcal{I}_p([a,b];\mathbb{B})$, $1 \le p < 2$, be additive and upper continuous. For each $k = 2,3,\ldots$ the function $I_k(\mu)$ on $[a,b]$ defined by*

$$I_k(\mu)(y) := \mathop{\not\!\!\int}_{[a,y]}\mu(dx_1)\mathop{\not\!\!\int}_{[a,x_1)}\mu(dx_2)\cdots\mathop{\not\!\!\int}_{[a,x_{k-1})}\mu(dx_k),\quad a<y\le b,\qquad I_k(\mu)(a):=0$$

exists, and satisfies the bound

$$\|I_k(\mu)\|_{[p]} \le C_p^{k-1}\Big(\frac{\rho^k(b)-\rho^k(a)}{k!}\Big)^{1/p}, \tag{7.47}$$

where $\rho(y) := v_p(\mu;[a,y])$ and $C_p = 2(1+4^{1/p}\zeta(2/p))$.

Proof. Let $I_1(\mu)(x) := \mu((a,x])$, $x \in [a,b]$. Then $I_k(\mu)(y) = \not\!\!\int_{[a,y]}\mu(dx)(I_{k-1}(\mu))_-^{(a)}$ for each $k = 2,3,\ldots$. It is clear that $I_1(\mu)$ is right–continuous on $[a,b)$ and 0 at a. Because of the relations $a \le x_k < x_{k-1} < \cdots < x_1 \le y$, we have

$$I_k(\mu)(y) = \mathop{\not\!\!\int}_{(a,y]}\mu(dx_1)\mathop{\not\!\!\int}_{(a,x_1)}\mu(dx_2)\cdots\mathop{\not\!\!\int}_{[a,x_{k-1})}\mu(dx_k),$$

so that each $I_k(\mu)$ is right–continuous on $[a,b)$ and 0 at a by the definition. Thus Proposition 7.17 follows from Proposition 7.14 by induction. $\square$

Theorem 7.18. *Let $\mu \in \mathcal{I}_p([a,b];\mathbb{B})$, $1 \le p < 2$, be additive and upper continuous. Then the series $\sum_{k\ge 2} I_k(\mu)$ converges in $\mathcal{W}_p([a,b];\mathbb{B})$, the product integral with respect to μ exists, and they satisfy the relation:*

$$\prod_{[a,y]} (\mathbb{I} + d\mu) = \mathbb{I} + \mu([a,y]) + \sum_{k\ge 2} I_k(\mu)(y), \quad \text{for } y \in [a,b]. \tag{7.48}$$

Proof. The series $\sum_{k\ge 2} I_k(\mu)$ converges in $\mathcal{W}_p$ by Proposition 7.17. The product integral with respect to μ exists by Corollary 7.5. To prove (7.48) let α be the interval function on $[a,b]$ defined by

$$\alpha(J) := \mathbb{I} + \mu(J) + \sum_{k\ge 2} \,⨎_J d\mu\, (I_{k-1}(\mu))_-^{(u)},$$

where u is the left endpoint of $J \subset [a,b]$ and $I_1(\mu)(x) := \mu((a,x])$, $x \in [a,b]$. By Theorem 4.39, α has bounded p-variation. Also for each $y \in [a,b]$, it follows that

$$\begin{aligned} ⨎_{[a,y]} \mu(dx)\alpha([a,x)) &= \mu([a,y]) + ⨎_{[a,y]} \mu(dx)\mu([a,x)) + ⨎_{[a,y]} \mu(dx) \sum_{k\ge 2} ⨎_{[a,x)} d\mu\, (I_{k-1}(\mu))_-^{(a)} \\ &= \mu([a,y]) + ⨎_{[a,y]} \mu(dx_1) ⨎_{[a,x_1)} \mu(dx_2) + \sum_{k\ge 3} ⨎_{[a,y]} d\mu\, (I_k(\mu))_-^{(a)} \\ &= \alpha([a,y]) - \mathbb{I}. \end{aligned}$$

To show that the integrals in the series all exist it is enough to apply the Love–Young inequality. Thus α satisfies the forward linear integral equation with respect to μ. Since the unique solution in $\mathcal{I}_p([a,b];\mathbb{B})$ of this equation is the product integral with respect to μ by Theorem 7.13, (7.48) holds. The proof is complete. ☐

Corollary 7.19. *Let $\mu \in \mathcal{I}_p([a,b];\mathbb{B})$, $1 \le p < 2$, be additive and upper continuous. Then the mapping*

$$\mathcal{I}_p \ni \mu \mapsto \prod(\mathbb{I} + d\mu) \in \mathcal{I}_p \tag{7.49}$$

is analytic with infinite radius of uniform convergence of its Taylor series around any point.

Proof. By Corollary 7.5, the mapping (7.49) is defined. (7.48) gives a power series expansion of the product integral with an infinite radius of uniform convergence at 0 by (7.47). Thus by Theorem 11.11 of [11, p. 163], where we suggest that $\|A_m\|$ be replaced by $\|\widehat{A}_m\|$ in the proof, the same is true at any other element of $\mathcal{I}_p$. ☐

7.6 Comments and related results

Survival and hazard functions Let T be a positive and bounded random variable on a probability space $(\Omega, \Pr)$. The *survival function* S on $[0,\infty)$ is defined by $S(t) := \Pr(\{T > t\})$ for $0 \le t < \infty$. Let $\tau_S := \sup\{t \ge 0\colon S(t) > 0\}$ which is finite because T is bounded. A nonnegative measure Λ on $[0,\infty)$ is a *hazard measure* if, letting $\tau_\Lambda := \inf\{t \ge 0\colon \Lambda((t,\infty)) = 0\}$, we have

(*a*) $0 < \tau_\Lambda < \infty$;
(*b*) for $0 < s < \tau_\Lambda$, $\Lambda(\{s\}) < 1$ and $\Lambda([0,s]) < \infty$;
(*c*) either $\Lambda([0,\tau_\Lambda)) < \infty$ and $\Lambda(\{\tau_\Lambda\}) = 1$ or $\Lambda([0,\tau_\Lambda)) = \infty$ and $\Lambda(\{\tau_\Lambda\}) = 0$.

Theorem 7.20. *There is a $1-1$ correspondence between hazard measures Λ on $[0,\infty)$ and survival functions S on $[0,\infty)$ given by*

$$S(t) = \prod_{[0,t]} (1 - d\Lambda), \quad 0 \le t < \tau_\Lambda, \qquad S(\tau_\Lambda) = 0 \qquad \textit{and}$$

$$\Lambda([0,t]) = -\int_{[0,t]} \frac{dS(u)}{S(u-)}, \quad 0 \le t < \tau_S, \quad \Lambda(\{\tau_S\}) := \begin{cases} 1 & \textit{if } S(\tau_S-) > 0 \\ 0 & \textit{otherwise} \end{cases} \qquad \Lambda((\tau_S,\infty)) := 0,$$

with $\tau_\Lambda = \tau_S$.

This is approximately Theorem 11 of Gill and Johansen [43] (see also [42, Section 5], [1, Theorem II.6.6]).

In survival analysis there is a sample of n individuals each of whom has a lifetime T_i and a "censoring time" C_i. Let T_i be independent identically distributed positive random variables with an unknown law having a survival function S to be estimated, and let $C_1, \dots, C_n$ be independent identically distributed positive random variables with a law having a survival function H. For each i one observes $X_i := \min(T_i, C_i)$ and $D_i = 1_{\{X_i = T_i\}}$, that is whether X_i actually equals T_i or C_i. We assume that for some $\tau_0 < \infty$, $C_i \le \tau_0$ for all i. Let $\tau_n := \max_{1 \le i \le n} X_i \le \tau_0$. For $0 \le t < \tau_n$, let

$$N_n(t) := \frac{1}{n}\#\{i\colon X_i \le t,\, D_i = 1\}, \quad Y_n(t) := \frac{1}{n}\#\{i\colon X_i \ge t\}, \quad \text{and} \quad \Lambda_n([0,t]) := \int_{[0,t]} \frac{dN_n}{Y_n}.$$

The *empirical hazard measure* Λ_n so defined is called the Nelson–Aalen estimator of the hazard measure. Since $\Lambda_n([0,\tau_n))$ is always finite in this case, $\Lambda_n(\{\tau_n\}) := 1$ is needed to fit the above definition of hazard measure. Unlike the point masses at $X_i = T_i < \tau_n$, representing deaths, or other endpoints of studies, $\Lambda_n(\{\tau_n\}) = 1$ simply means that only the distribution of $\min(X_i, \tau_n)$ can be estimated at all. The *empirical survival function* S_n defined by

$$S_n(t) := \prod_{[0,t]} (1 - d\Lambda_n), \qquad 0 \le t \le \tau_n,$$

is called the Kaplan-Meier estimator of the unknown survival function $S(t) = \Pr(\{T_i > t\})$. S_n is the nonparametric maximum likelihood estimator of S on $[0,\tau_S)$ from the given data set, and has other good properties. However, no estimator can be reliable for t large enough so that $Y_n(t)$ is small.

An alternative definition of the product integral Let M_k be the set of $k \times k$ matrices with complex entries, let μ be a nonnegative measure on an interval $[a,b]$ and $f \in \mathcal{L}^1([a,b],\mu,M_k)$. Dollard and Friedman [17, pp. 155-156] define a form of product integral

$$\prod_{(a,b]} \exp\{f\,d\mu\} \quad \text{as a limit of products} \quad \prod_{i=1}^{n} \exp\{\nu((x_{i-1},x_i])\} \tag{7.50}$$

over partitions $a = x_0 < x_1 < \cdots < x_n = b$ as the mesh of partitions goes to 0, where $\nu((c,d]) := \int_{(c,d]} f\,d\mu$. The exponential is defined by its usual power series. If the signed measure ν has no atoms, then

$$\prod_{(a,b]} \exp\{f\,d\mu\} = \prod_{(a,b]} (\mathbb{1} + d\nu) \tag{7.51}$$

(Dollard and Friedman [17, Theorem I.7.1, p. 52]; the continuity assumption there can be avoided because of our different formulation). The function $g(z) = e^z$ or $1+z$ can be replaced by any function analytic in a neighborhood of 0 with $g(0) = g'(0) = 1$. Thus $g(z) = 1+z+O(z^2)$ as $z \to 0$. Beside $1+z$ and e^z, the function $1/(1-z)$ also has been applied, cf. Hildebrandt [51] and Freedman [36]. If ν has atoms the product integral depends on the choice of g and corresponding definitions give different results.

If the values of ν on different intervals commute, e.g. if ν is real–valued, then simply

$$\prod_{(a,b]} \exp\{f\,d\mu\} = \exp\{\nu((a,b])\}, \tag{7.52}$$

which is never 0 since ν is a finite signed measure. If ν is taken as a hazard measure and has atoms, (7.50) does not give the Kaplan–Meier estimator but a different, unwanted estimator.

While the product integral $\prod(\mathbb{I}+d\nu)$ gives solutions of integral equations (7.42) and (7.43), the form

$$P(a,b) := \prod_{(a,b]} \exp\{f\,d\mu\}$$

gives solutions of integral equations

$$P(a,b) = \mathbb{I} + \int_{(a,b]} f_\phi(s)P(a,s)\,\mu(ds),$$

where $f_\phi(s) := \phi(f(s)\mu(\{s\}))f(s)$ and $\phi(x) := (e^z-1)/z$, $z \neq 0$, $\phi(0) := 1$ (Dollard and Friedman [17, pp. 160-161]). For natural integral equations with atoms (jumps), then, it seems that $\prod(\mathbb{I}+d\nu)$ might be preferred. If ν has no atoms, then in light of (7.52) one can see how the form $\prod_{(a,b]}\exp\{f d\mu\}$ might have been preferred, although it is equivalent to $\prod(\mathbb{I}+d\nu)$ by (7.51).

We next extend the definition of the product integral $\Pi\, e^{df}$. Let $\kappa := \{I_j\}_{j=1}^n$ be a partition of an interval J into disjoint subintervals (some of which may be singletons), i.e. $\kappa \in \mathrm{IP}\,(J)$. Then $x < y$ whenever $x \in I_i$, $y \in I_j$, and $1 \leq i < j \leq n$. Let ν be an additive interval function defined on all subintervals of J, with values in a Banach algebra $\mathbb{B}$ with identity $\mathbb{I}$. Recall that for $A \in \mathbb{B}$, $e^A := \exp(A) := \sum_{j=0}^\infty A^j/j!$, where $A^0 := \mathbb{I}$. We define the product $\Pi_\kappa e^{\Delta\nu} := \Pi_{j=1}^n \exp\{\nu(I_j)\}$. We say that the product integral $\Pi_J e^{d\nu}$ exists and equals $A \in \mathbb{B}$ if for every $\epsilon > 0$ there is a $\kappa \in \mathrm{IP}\,(J)$ such that for every refinement λ of κ, $\|A - \Pi_\lambda e^{\Delta\nu}\| < \epsilon$.

If $\mathbb{B}$ is commutative then $e^{A+B} = e^Ae^B$ for any $A, B \in \mathbb{B}$ and the following is immediate:

Theorem 7.21. *If $\mathbb{B}$ is a commutative Banach algebra with identity then for an arbitrary finitely additive interval function ν into $\mathbb{B}$, defined on all subintervals I of an interval J, $\Pi_I e^{d\nu}$ exists and equals $e^{\nu(I)}$. In fact, $\Pi_\kappa e^{\Delta\nu} = e^{\nu(I)}$ for any partition κ of I. If I ranges over subintervals of J then the map $\nu \mapsto (I \mapsto \Pi_I e^{d\nu})$ is holomorphic from the supremum norm $\|\nu\|_\infty := \sup_{I\subset J}\|\nu(I)\|$ to itself.*

In the noncommutative case, it is not hard to show that $\Pi_J e^{d\nu}$ is defined in $\mathbb{B}$ at least for $\nu \in \mathcal{W}_1(J, \mathbb{B})$. Note that such a ν is countably additive. But we have:

Proposition 7.22. *Let M_2 be the Banach algebra of 2×2 real matrices with $I := \left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right)$. Then the maps $\nu \mapsto \Pi_{[0,1]}e^{d\nu}$ and $\nu \mapsto \prod_{[0,1]}(I+d\nu)$ on $\mathcal{W}_1([0,1], M_2)$ are not continuous in the p-variation norm for any $p > 2$.*

Proof. Let $B := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $C := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, and for $0 < \delta < 1/2$ let $X := I + \delta B$ and $Y := I + \delta C$,. Then $B^2 = C^2 = 0$, so $X^{-1} = I - \delta B$, $Y^{-1} = I - \delta C$, $e^{\delta B} = X$ and $e^{\delta C} = Y$. Let $T := YXY^{-1}X^{-1}Y^{-1}X^{-1}YX$. Then straightforward algebra gives $T_{11} = 1 - 2\delta^2 + \delta^4 + \delta^6$, $T_{12} = T_{21} = 2\delta^5 + \delta^7$, and $T_{22} = 1 + 2\delta^2 + 3\delta^4 + 3\delta^6 + \delta^8$. Here T is symmetric (self-adjoint), as can be seen directly from its definition since $X = Y^t$, etc. Thus T can be diagonalized by some rotation. The characteristic equation for its eigenvalues is $\lambda^2 - t\lambda + 1 = 0$ where $t = 2 + 4\delta^4 + 4\delta^6 + \delta^8$, and T has determinant 1 since each factor does. The roots are thus $\lambda = [t \pm (t^2 - 4)^{1/2}]/2$. It follows easily that the larger eigenvalue is $\geq 1 + 2\delta^2 \geq \exp\{\delta^2\}$.

Now for each $n = 1, 2, \cdots$, let κ be the partition of $(0, 1]$ into the $8n$ subintervals $((j - 1)/(8n), j/(8n)]$, $j = 1, \cdots, 8n$. Let the values of ν on the first eight subintervals be successively $\delta C, \delta B, -\delta C, -\delta B, -\delta C, -\delta B, \delta C, \delta B$; and then the cycle repeats n times. If I is one of the $8n$ intervals in κ and $J \subset I$ let $\nu(J) := 8n\lambda(J)\nu(I)$ where λ is Lebesgue measure. The sum of the 8 values in a cycle is 0. It follows by Lemma 4.20 that the p-variation of ν is bounded above by $Kn\delta^p$ for some constant K. Since $p > 2$ we can choose $\delta_n := n^{-2/(2+p)}$ so that $n\delta_n^p \to 0$ but $n\delta_n^2 \to \infty$ as $n \to \infty$. Then $\Pi_{[0,1]} e^{d\nu} = T^n$, via the partition κ and any refinement of it, and $\prod_{(0,1]} \mathbb{I} + d\nu = T^n$ also. Since T^n has an eigenvalue $\geq \exp(n\delta_n^2)$, the conclusion follows. $\square$

In Proposition 7.22, p-variation for $p > 2$ can be replaced by ϕ-variation for any ϕ-function such that $\phi(x) = o(x^2)$ as $x \downarrow 0$.

Comparing product integrals and integral equations There are different possibilities for defining product integrals and integral equations they satisfy, which are actually different in case there are jumps. As we have seen,

Dollard and Friedman [17] defined a product integral Πe^{df} and obtained an unusual integral equation.

We, in [29], defined a product integral $\prod(\mathbb{I} + df)$ and obtained an unusual integral equation in terms of "Left Young" or "Right Young" integrals.

Hildebrandt [51] solved a natural integral equation and found an unusual product integral with factors $(\mathbb{I} + \Delta^+ f)(\mathbb{I} - \Delta^- f)^{-1}$.

Stochastic analysts, some analysts, Gill and Johansen [43], and Gill [42], defined integral equations with left-continuous integrands and right-continuous solutions. Also, results were formulated in terms of interval functions in [43], [42], where the product integral is $\prod(\mathbb{I} + d\mu)$. These formulations, as we have seen in Sections 7.4 and 7.5, seem to fit together well.

Bibliography

[1] P. K. Andersen, Ø. Borgan, R. D. Gill and N. Keiding. *Statistical Models based on Counting Processes*. Springer, New York, 1993.

[2] J. Appell and P. P. Zabrejko. *Nonlinear superposition operators*. Cambridge University Press, 1990.

[3] * R. G. Bartle. Return to the Riemann integral. *Amer. Math. Monthly*, **103** (1996), 625-632.

[4] N. K. Bary. *A treatise on trigonometric series*. Vol. 1. Transl. by M. F. Mullins. MacMillan, New York, 1964.

[5] J. Beirlant and P. Deheuvels. On the approximation of P-P and Q-Q plot processes by Brownian bridges. *Statist. Probab. Letters*, **9** (1990), 241-251.

[6] K. Bichteler. Stochastic integration and L^p-theory of semimartingales. *Ann. Probab.*, **9** (1981), 49-89.

[7] R. M. Blumenthal and R. K. Getoor. Sample functions of stochastic processes with stationary independent increments. *J. Math. and Mech.*, **10** (1961), 493-516.

[8] J. Bretagnolle. p-variation de fonctions aléatoires. In: *Séminaire de Probabilités VI*; Lect. Notes in Math. **258**, 51-71. Springer, Berlin, 1972.

[9] T. Carleman. Über die Fourierkoeffizienten einer stetigen Funktion. *Acta Math.* (Sweden), **41** (1918), 377-384.

[10] L. Carleson. On convergence and growth of partial sums of Fourier series. *Acta Math.* (Sweden) **116** (1966), 135-157.

[11] S. B. Chae. *Holomorphy and calculus in normed spaces*. Dekker, New York, 1985.

[12] M. Csörgő. *Quantile Processes with Statistical Applications*. SIAM, Philadelphia, 1983.

[13] P. J. Daniell. Stieltjes derivatives. *Bull. Amer. Math. Soc.* **26**, 444-448.

[14] C. Dellacherie and P.-A. Meyer. *Probabilities and potential B*. Transl. by J. P. Wilson. North–Holland, Amsterdam, 1982.

[15] C. Doléans–Dade. Quelques applications de la formule de changement de variables pour les semimartingales. *Z. Wahrsch. Verw. Gebiete*, **16** (1970), 181-194.

[16] C. Doléans–Dade and P. A. Meyer. Intégrales stochastiques par rapport aux martingales locales. *Séminaire de Probabilités IV*, Univ. Strasbourg, *Lecture Notes in Math.* (Springer) **124** (1970), 77-107.

[17] J. D. Dollard and C. N. Friedman. *Product Integration with Applications to Differential Equations. Encyclopedia of Math. and its Appl.*, Vol. 10, Addison-Wesley, Reading, Massachusetts, 1979.

[18] J. L. Doob. *Stochastic Processes*. Wiley, New York, 1953.

[19] R. M. Dudley. The size of compact subsets of Hilbert space and continuity of Gaussian processes. *J. Functional Analysis*, **1** (1967), 290-330.

[20] R. M. Dudley. Sample functions of the Gaussian process. *Ann. Probab.*, **1** (1973), 66-103.

[21] R. M. Dudley. Wiener functionals as Itô integrals. *Ann. Probab.*, **5** (1977), 140-141.
[22] R. M. Dudley. *Real analysis and probability.* 2d printing, Chapman & Hall, New York, 1993.
[23] R. M. Dudley. Fréchet differentiability, p-variation and uniform Donsker classes. *Ann. Probab.*, **20** (1992), 1968-1982.
[24] R. M. Dudley. The order of the remainder in derivatives of composition and inverse operators for p-variation norms. *Ann. Statist.*, **22** (1994), 1-20.
[25] R. M. Dudley. Empirical processes and p-variation. In *Festschrift for Lucien Le Cam* (D. Pollard, E. Torgersen and G. L. Yang, eds.) 219-233. Springer, New York, 1997.
[26] R. M. Dudley. *Uniform Central Limit Theorems.* Cambridge University Press, Cambridge, to appear 1999.
[27] R. M. Dudley. Differentiability of the composition and quantile operators for regulated and a.e. continuous functions. Preprint.
[28] R. M. Dudley and R. Norvaiša. A survey on differentiability of six operators in relation to probability and statistics. Preprint.
[29] R. M. Dudley and R. Norvaiša. Product integrals, Young integrals and p-variation. Preprint.
[30] R. M. Dudley and W. Philipp. Invariance principles for sums of Banach space valued random elements and empirical processes. *Z. Wahrscheinlichkeitsth. verw. Geb.*, **62** (1983), 509-552.
[31] A. Dvoretzky and C. A. Rogers. Absolute and unconditional convergence in normed linear spaces. *Proc. Nat. Acad. Sci. U.S.A.*, **36** (1950), 192-197.
[32] W. Fernández de la Vega. On almost sure convergence of quadratic Brownian variation. *Ann. Probab.*, **2** (1974), 551-552.
[33] L. T. Fernholz. *von Mises calculus for statistical functionals. Lect. Notes in Statist.*, **19**. Springer, New York, 1983.
[34] H. Föllmer. Calcul d'Itô sans probabilités. *Séminaire de Probabilités XV*, Eds. J. Azéma and M. Yor; *Lect. Notes in Math.* **850** (1981), 143-150.
[35] M. A. Freedman. Operators of p-variation and the evolution representation problem. *Trans. Amer. Math. Soc.*, **279** (1983), 95-112.
[36] M. A. Freedman. Necessary and sufficient conditions for discontinuous evolutions with applications to Stieltjes integral equations. *J. Integral Equations*, **5** (1983), 237-270.
[37] B. Fristedt. Sample functions of stochastic processes with stationary, independent increments. In: *Advances in probability and related topics*, Vol. 3. Dekker, New York, 1974, pp. 241-396.
[38] B. Fristedt and S. J. Taylor. Strong variation for the sample functions of a stable process. *Duke Math. J.*, **40** (1973), 259-278.
[39] F. W. Gehring. A study of α-variation. I. *Trans. Amer. Math. Soc.*, **76** (1954), 420-443.
[40] I. I. Gikhman and A. V. Skorokhod. *Introduction to the theory of random processes.* W.B. Saunders, Philadelphia, 1969.
[41] R. D. Gill. Non– and semi–parametric maximum likelihood estimators and the von Mises method (Part 1). *Scand. J. Statist.*, **16** (1989), 97-128.
[42] R. D. Gill. Lectures on Survival Analysis. In *Ecole d'été de Probabilités de Saint-Flour* (P. Bernard, ed.). *Lecture Notes in Math.*, **1581** (1994), 115-241. Springer, Berlin.
[43] R. D. Gill and S. Johansen. A survey of product-integration with a view toward application in survival analysis. *Ann. Statist.*, **18** (1990), 1501-1555.
[44] V. I. Glivenko. *The Stieltjes integral.* (In Russian) ONTI, Leningrad, 1936.

[45] R. A. Gordon. *The integrals of Lebesgue, Denjoy, Perron, and Henstock.* Amer. Math. Soc., Providence, 1994.
[46] R. A. Gordon. The use of tagged partitions in elementary real analysis. *Amer. Math. Monthly*, **105** (1998), 107-117.
[47] G. H. Hardy, J. E. Littlewood and G. Pólya. *Inequalities.* Cambridge University Press, 2d ed. 1952, repr. 1967.
[48] R. Henstock. *The General Theory of Integration.* Clarendon Press, Oxford, 1991.
[49] T. H. Hildebrandt. On integrals related to and extensions of the Lebesgue integrals. *Bull. Amer. Math. Soc.*, **24** (1917), 113-144 and 177-202.
[50] T. H. Hildebrandt. Definitions of Stieltjes integrals of the Riemann type. *Amer. Math. Monthly*, **45** (1938), 265-278.
[51] T. H. Hildebrandt. On systems of linear differentio–Stieltjes–integral equations. *Illinois J. Math.*, **3** (1959), 352-373.
[52] T. H. Hildebrandt. *Introduction to the theory of integration.* Academic Press, New York, 1963.
[53] C. S. Hönig. *Volterra Stieltjes-Integral Equations.* In: *Functional Differential Equations and Bifurcation*, (A. F. Izé, ed.). Lect. Notes in Math. (Springer), **799** (1980), 173-216.
[54] Yen-Chin Huang. *Empirical distribution function statistics, speed of convergence, and p-variation.* Ph. D. thesis, Massachusetts Institute of Technology, 1994.
[55] Yen-Chin Huang. Speed of convergence of classical empirical processes in p-variation norm. Preprint, 1998.
[56] R. A. Hunt. On the convergence of Fourier series. In: *Orthogonal Expansions and their Continuous Analogues*, ed. D. T. Haimo, Southern Illinois Univ. Press [Carbondale], (1968), 235-255.
[57] K. Itô. Stochastic processes. *Lecture Notes Series*, **16**. Aarhus Universitet, 1969.
[58] J. Jacod and P. Protter. Asymptotic error distributions for the Euler method for stochastic differential equations. *Ann. Probab.* **26** (1998), 267-307.
[59] N. C. Jain and D. Monrad. Gaussian measures in B_p. *Ann. Probab.*, **11** (1983), 46-57.
[60] T. Kawada and N. Kôno. On the variation of Gaussian processes. In: *Proc. Second Japan-USSR Symp. on Probab. Theory*; (G. Maruyama and Yu. V. Prokhorov, eds.). *Lect. Notes Math.* (Springer), **330** (1973), 176-192.
[61] J. Kiefer. Deviations between the sample quantile process and the sample df. In *Nonparametric Techniques in Statistical Inference* (M. L. Puri, ed.), 299-319. Cambridge University Press, 1970.
[62] A. N. Kolmogorov. Une série de Fourier-Lebesgue divergente partout. *C. R. Acad. Sci. Paris*, **183** (1926), 1327-1328.
[63] A. N. Kolmogorov. Untersuchungen über den Integralbegriff. *Math. Ann.*, **103** (1930), pp. 654-696. Russian transl. by P. Osvald in *A. N. Kolmogorov. Collected works. Mathematics and mechanics*, Moskow, Nauka, 1985, pp. 96-134.
[64] ∗ J. König. A határozott intégralok elmeletéz. *Mat. és. Természettudmányi értesitö*, 1897, 380-384.
[65] G. Kowalewski. Über den zweiten Mittelwertsatz der Infinitesimalrechnung. *Math. Annalen*, **60** (1905), 151-156.
[66] G. L. Krabbe. Integration with respect to operator-valued functions. *Bull. Amer. Math. Soc.*, **67** (1961), 214-218.
[67] G. L. Krabbe. Integration with respect to operator-valued functions. *Acta Sci. Math.* (Szeged), **22** (1961), 301-319.

[68] J. Kurzweil. Generalized ordinary differential equations and continuous dependence on a parameter. *Czechoslovak Math. J.*, **7** (1957), 418-446.
[69] M. Ledoux and M. Talagrand. *Probability in Banach spaces.* Springer, Berlin, 1991.
[70] R. Leśniewicz and W. Orlicz. On generalized variations (II). *Studia Math.* **45** (1973), 71-109.
[71] P. Lévy. Le mouvement brownien plan. *Amer. J. Math.*, **62** (1940), 487-550.
[72] P. Lévy. Random functions: general theory with special reference to Laplacian random functions. *Univ. of Calif. Publs. in Statist.*, **1** no. 12, 1953, 331-388; repr. in *Oeuvres de Paul Lévy* IV, *Processus Stochastiques*, Gauthier–Villars, Paris, 1980, 187-286.
[73] R. Lipschitz. De explicatione per series trigonometricas instituenda functiorum unius variabilis arbitrariarum ... *J. reine angew. Math.*, **63** (1864), 296-308. French. transl. by P. Montel in *Acta Math.*, **36** (1912-13), 281-295.
[74] E. R. Love. The refinement–Ross–Riemann–Stieltjes (R^3S) integral. In: *Analysis, Geometry and Groups: A Riemann Legacy Volume*, (H. M. Srivastava and Th. M. Rassias, eds.), pp. 289-312; Hadronic Press, Palm Harbor, Florida, 1993.
[75] E. R. Love and L. C. Young. On fractional integration by parts. *Proc. London Math. Soc.*, Ser. 2, **44** (1938), 1-35.
[76] T. Lyons. Differential equations driven by rough signals (I): an extension of an inequality of L. C. Young. *Math. Res. Letters*, **1** (1994), 451-464.
[77] J. S. MacNerney. Integral equations and semigroups. *Illinois J. Math.*, **7** (1963), 148-173.
[78] J. Marcinkiewicz. On a class of functions and their Fourier series. *C. R. Soc. Sci. Varsovie*, **26** (1934), 71-77. Reprinted in: *J. Marcinkiewicz. Collected Papers.* (A. Zygmund, ed.) Państwowe Wydawnictwo Naukowe, Warsaw, 1964.
[79] P. R. Masani. Multiplicative Riemann integration in normed rings. *Trans. Amer. Math. Soc.*, **61** (1947), 147-192.
[80] E. J. McShane. *Unified integration.* Academic Press, Orlando, 1983.
[81] F. A. Medvedev. *Development of the Concept of Integral* (in Russian). Moscow, Nauka, 1974.
[82] P.-A. Meyer. *Probability and Potentials.* Blaisdell, Waltham, MA, 1966.
[83] I. Monroe. On the γ-variation of processes with stationary independent increments. *Ann. Math. Statist.*, **43** (1972), 1213-1220.
[84] I. Monroe. On embedding right continuous martingales in Brownian motion. *Ann. Math. Statist.*, **43** (1972), 1293-1311.
[85] E. H. Moore. Definition of limit in general integral analysis. *Proc. Nat. Acad. Sci. U.S.A.*, **1** (1915), 628-632.
[86] E. H. Moore and H. L. Smith. A general theory of limits. *Amer. J. Math.*, **44** (1922), 102-121.
[87] W. Orlicz. Beiträge zur Theorie der Orthogonalentwicklungen (III). *Bull. Acad. Pol. Sci.*, **1932**, 229-238.
[88] R. E. A. C. Paley, N. Wiener and A. Zygmund. Notes on random functions. *Math. Z.*, **37** (1933) 647-668. Reprinted in: *Norbert Wiener. Collected Works with commentaries*, Vol. I (P. Masani, ed.) pp. 536-557; MIT Press, Cambridge, 1976.
[89] Lee Peng–Yee. *Lanzhou Lectures on Henstock integration.* World Scientific, Singapore, 1989.
[90] W. F. Pfeffer. *The Riemann approach to integration.* Cambridge University Press, Cambridge, 1993.

[91] S. Pollard. The Stieltjes' integral and its generalisations. *Quart. J. Pure and Appl. Math.*, **49** (1923), 73-138.
[92] Jinghua Qian. The p-variation of partial sum processes and the empirical process. *Ann. Probab.*, **26** (1998), 1370-1383.
[93] J. A. Reeds III. *On the definition of von Mises functionals.* Ph. D. thesis, Statistics, Harvard University, 1976.
[94] G. F. B. Riemann. Über die Darstellbarkeit einer Function durch eine trigonometrische Reihe. *Abh. Gesell. Wiss. Göttingen*, **13** (1866-67: publ. 1868), math. Kl., 87-132. Reprinted in: *Oeuvres mathématiques de Riemann*, (L. Laugel, ed.), 1968, Paris, 225-279.
[95] K. A. Ross. Another approach to Riemann–Stieltjes integrals. *Amer. Math. Monthly*, **87** (1980), 660-662.
[96] K. A. Ross. *Elementary analysis: The theory of calculus.* Springer, New York, 1980.
[97] T. Runst and W. Sickel. *Sobolev Spaces of Fractional Order, Nemytskij Operators, and Nonlinear Partial Differential Equations.* de Gruyter, Berlin, 1996.
[98] S. Saks. *Theory of the Integral.* 2d ed. English transl. by L. C. Young. Hafner, New York. Repr. corrected Dover, New York, 1964.
[99] R. Salem. La convergence des series de Fourier. *C. R. Acad. Sci. Paris*, **202** (1938), 662-66. Repr. as: Essais sur les séries trigonometriques. Chap. VI, in *Actual. Sci. et Industr.*, **862**. Paris, 1940, and in: *Œuvres Mathématiques*, pp. 156-161. Hermann, Paris, 1967.
[100] Š. Schwabik. On the relation between Young's and Kurzweil's concept of Stieltjes integral. *Časopis Pěst. Mat.*, **98** (1973), 237-251.
[101] Š. Schwabik, M. Tvrdý and O. Vejvoda. *Differential and integral equations.* Reidel, Dordrecht, 1979.
[102] A. V. Skorohod. On a representation of random variables. *Theory Probab. Appl.*, **21** (1976), 628-632.
[103] A. V. Skorohod. On a generalization of the stochastic integral. *Theory Probab. Appl.*, **20** (1975), 219-233.
[104] ∗ T. J. Stieltjes. Recherches sur les fractions continues. *Ann. Fac. Sci. Toulouse*, Sér. 1 **8** (1894), 1-122.
[105] M. Talagrand. Regularity of Gaussian processes. *Acta Math.* (Sweden), **159** (1987), 99-149.
[106] S. J. Taylor. Exact asymptotic estimates of Brownian path variation. *Duke Math.*, **39** (1972), 219-241.
[107] H. Triebel. *Theory of function spaces.* Birkhäuser, Basel, 1983.
[108] A. van der Vaart. Efficiency and Hadamard differentiability. *Scand. J. Statist.*, **18** (1991), 63-75.
[109] A. W. van der Vaart and J. A. Wellner. *Weak Convergence and Empirical Processes, with Applications to Statistics.* Springer, New York, 1996.
[110] V. Volterra. Sui fondamenti della teoria delle equazioni differenziali lineari. *Memorie della Società Italiana delle Scienze (detta dei XL)*, serie III, vol. VI, n. 8 (1887). In: "Opere Matematiche" I, 1954, pp. 209-290. Accad. Nazionale dei Lincei, Roma.
[111] A. J. Ward. The Perron–Stieltjes integral. *Math. Z.*, **41** (1936), 578-604.
[112] N. Wiener. Differential space. *J. Math. Phys.* M.I.T. **2** (1923), 131-174. Also in Wiener (1976), I, pp. 455-498.
[113] N. Wiener. The quadratic variation of a function and its Fourier coefficients. *J. Math. and Phys.* (MIT, Cambridge, Mass), **3** (1924), 72-94. Also in Wiener (1976), II, pp. 36-58.

[114] N. Wiener (1976-1986). *Norbert Wiener: Collected Works with Commentaries.* (P. Masani, ed.) 4 vols. M.I.T. Press, Cambridge, Mass.

[115] Quanhua Xu. Espaces d'interpolation réels entre les espaces V_p: propriétés géométriques et applications probabilistes. In: *Séminaire d'Analyse Fonctionnelle* 1985/ 1986/1987, Paris VI-VII. Paris: Publ. Math. Univ. Paris VII **28** 1988, pp. 77-123.

[116] G. C. Young. On infinite derivatives. An Essay. *Quarterly J. Pure Appl. Math.*, **47** (1916), 127-175.

[117] L. C. Young. An inequality of the Hölder type, connected with Stieltjes integration. *Acta Math.* (Sweden), **67** (1936), 251-282.

[118] L. C. Young. General inequalities for Stieltjes integrals and the convergence of Fourier series. *Math. Ann.*, **115** (1938), 581-612.

[119] R. C. Young. On Riemann integration with respect to a continuous increment. *Math. Z.*, **29** (1929), 217-233.

[120] W. H. Young. On integration with respect to a function of bounded variation. *Proc. London Math. Soc.* (Ser. 2), **13** (1914), 109-150.

[121] A. Zygmund. *Trigonometric series.* Cambridge University Press, 2d ed. 1959, repr. with corrections and some additions 1968, repr. 1988.

$*$ We learned of these references from secondary sources and have not seen them in the original.