arXiv:1201.1996v2 [math.PR] 8 Jun 2012

A SIMPLE PROOF OF THE BICHTELER–DELLACHERIE THEOREM

MATHIAS BEIGLBÖCK AND PIETRO SIORPAES

Abstract. We give a short and rather elementary proof of the celebrated Bichteler–Dellacherie Theorem, which states that a process S is a good integrator if and only if it is the sum of a local martingale and a finite-variation process. As a corollary, we obtain a characterization of semimartingales along the lines of classical Riemann integrability.

1. Introduction

In this paper we provide a short proof of the Bichteler–Dellacherie theorem, which basically asserts that one can integrate with respect to a process S iff S is a semimartingale, i.e., the sum of a local martingale and a finite-variation process.

The Doob-Meyer decomposition theorem leads to the following reformulation: a bounded process allows for a good integration theory iff it is (locally) the difference of two submartingales. This is analogous to the deterministic case, where one can integrate with respect to a function f iff f can be written as a difference of two increasing functions. We find that this analogy is sound, as the simple proof in the deterministic set-up can be reinterpreted to establish the Bichteler–Dellacherie theorem in full generality.

As a corollary, we obtain that semimartingales can be characterized by Riemann-sums in the following way: a càdlàg adapted process $(S_t)_{t\in[0,1]}$ is a semimartingale iff for every bounded adapted continuous process H the sequence of Riemann-sums

$$\sum_{i=0}^{2^n-1} H_{\frac{i}{2^n}}(S_{\frac{i+1}{2^n}} - S_{\frac{i}{2^n}}) \tag{1}$$

converges in probability. This observation is perhaps new and emphasizes the viewpoint that semimartingales are the stochastic equivalent of processes of finite variation.

2. Definitions, assumptions and main statement

Throughout this article we consider a finite time horizon T, which wlog we take to be equal to 1, and a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$. We assume that the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,1]}$ satisfies the usual conditions of right continuity and saturatedness. A *simple integrand* is a stochastic process $H = (H_t)_{t\in(0,1]}$ of the form

$$H = \textstyle\sum_{i=1}^k H^i \mathbb{1}_{(\tau_i,\tau_{i+1}]}, \tag{2}$$

where k is a finite number, $0 \le \tau_1 \le \ldots \le \tau_{k+1} \le 1$ are stopping times, and H^i are bounded $\mathcal{F}_{\tau_i}$-measurable random variables. The vector space of simple integrands will be denoted by $\mathcal{S}$, and will be endowed with the sup norm

$$\|H\|_\infty := \| \sup_{t\in[0,1]} |H_t| \, \|_{L^\infty(\mathbb{P})}. \tag{3}$$

Given an adapted (real-valued) process $S = (S_t)_{t\in[0,1]}$ and a simple integrand H as in (2), it is natural to define the (Itô) integral $\mathcal{I}_S(H)$ of $H \in \mathcal{S}$ with respect to S as the random variable

$$\mathcal{I}_S(H) := \textstyle\sum_{i=1}^k H^i(S_{\tau_{i+1}} - S_{\tau_i}). \tag{4}$$

This defines the integral as a linear operator $\mathcal{I}_S$ from the normed space $\mathcal{S}$ to the topological vector space $L^0(\mathbb{P})$ (the space of all random variables, with the metrizable topology of convergence in probability). A process S is then called a *good integrator* if $\mathcal{I}_S : \mathcal{S} \to L^0(\mathbb{P})$ is continuous, i.e. if $H^n \in \mathcal{S}$, $\|H^n\|_\infty \to 0$ implies that $\mathcal{I}_S(H^n)$ goes to 0 in probability as $n \to \infty$.

It is easy to show that (locally) square integrable martingales and processes of finite variation are good integrators. It is also true that any (local) martingale is a good integrator, although this requires a little more work; we refer to [Edw90] for an elementary proof of this result which does not make use of the structure of local martingales in continuous time.

The converse result is of key importance to stochastic analysis, as it characterizes the processes S for which one can build a powerful integration theory. This is the object of the following well known theorem, commonly known as the Bichteler-Dellacherie Theorem.

Theorem BD. [Bic79, Bic81, Del80] *Let $(S_t)_{0\le t\le 1}$ be a càdlàg adapted process. If $\mathcal{I}_S : \mathcal{S} \to L^0(\mathbb{P})$ is continuous then S can be written as a sum of a càdlàg local martingale and a càdlàg adapted process of finite variation.*

Theorem BD evolved in the Strasbourg school of P.A. Meyer and was originally published in [Del80] and, independently, [Bic79]. Mokobodski deserves particular credit (see for instance the discussion in [DM88]); however since the result is usually baptized after Bichteler and Dellacherie, we stick to this name.

Standard accounts of the result employ functional analytic machinery and change of measure techniques, see [DM88, Pro05]. A relatively direct proof is given in [BSV11a], where the desired semimartingale decomposition is obtained from discrete-time Doob decompositions. However, the limiting procedure involves some rather delicate estimates. Another proof, based on an orthogonal decomposition, is given by Lowther in [Low11].

We emphasize that the definition of good integrators requires that the integrands are adapted. Simply dropping this assumption would amount to considering all simple processes that are adapted to the constant filtration $\mathcal{G}_t := \mathcal{F}_1$, $0 \le t \le 1$. Since $(\mathcal{G}_t)$-local martingales are constant, Theorem BD implies that every $(\mathcal{G}_t)$-good integrator has paths of finite variation. So, if one chooses to consider integrands which are not necessarily adapted (predictable) one is left with an unreasonably small class of integrators.

Since submartingales provide a filtration-dependent stochastic equivalent of increasing functions, we believe that the following reformulation of Theorem BD is quite intuitive.

Theorem 2.1. *Let $S = (S_t)_{0\le t\le 1}$ be a bounded càdlàg adapted process. If S is a good integrator then it is locally the difference of two càdlàg submartingales*

We recall that a process defined on $[0, 1]$ satisfies a property *locally* if, for each $\varepsilon > 0$, there exist a $[0, 1] \cup \{\infty\}$-valued stopping time ϱ such that S^ϱ satisfies that property and $\mathbb{P}(\varrho = \infty) \ge 1 - \varepsilon$.

Our main contribution consists in a simple proof of Theorem 2.1. Its equivalence with Theorem BD easily follows from the Doob–Meyer decomposition theorem, of which in recent years simple and elementary proofs have been obtained: we refer the reader to [Bas96, Jak05] resp. [BSV11b].

The paper is organized as follows. After recalling Rao's Theorem in the next section, we provide the proof of Theorem 2.1 in Section 4. The fact that Theorem 2.1 implies Theorem BD is shown in detail in Section 5. In Section 6 we include Stricker's simple proof of Rao's Theorem. Finally, in Section 7 we discuss certain ramifications of the Theorem BD (including the characterization (1)) and, en passant, give an elementary derivation of some closure properties of the space of semimartingales.

We conclude this section with some definitions that will be used throughout the paper. As it is customary, we will denote by X^+ (X^-) the positive (negative) part of a random variable X, and by D_n the n-th dyadic partition of $[0,1]$, i.e. $D_n = \{0, 1/2^n, 2/2^n, \ldots, 1\}$. We will not be picky about the difference between functions and their equivalence classes. Given a simple integrand H, $H \cdot S$ denotes the process given by $(H \cdot S)_t := \mathcal{I}_{S^t}(H)$. Recall that a family $F \subseteq L^0(\mathbb{P})$ is bounded if for every $\varepsilon > 0$ there exists a constant C such that $\mathbb{P}(|X| \geq C) \leq \varepsilon$ for every $X \in F$. By the usual proof, a linear operator from a normed space to $L^0(\mathbb{P})$ is continuous iff it is bounded, i.e., it maps bounded sets into bounded sets. We will use this fact without further mention. We will say that a process $(S_t)_t$ is right continuous in probability if S_{t_n} converges in probability to S_t whenever $t_n \to t$, $t_n > t$.

3. Quasimartingales

To prove that a given function $f = f(t)$ can be written as a difference of two increasing functions, one would typically show that f has finite variation. This has an analogue in the stochastic world; to state it, we recall the notion of quasimartingale.

Let $S = (S_t)_{0 \leq t \leq 1}$ be an adapted process such that $S_t \in L^1$ for all $t \in [0,1]$. Given a partition $\pi = \{0 = t_0 < t_1 < \ldots < t_n = 1\}$ of $[0,1]$, the *mean variation* of S *along* π is defined as

$$\mathrm{MV}(S, \pi) = \mathbb{E} \textstyle\sum_{t_i \in \pi} \left|\mathbb{E}[S_{t_i} - S_{t_{i+1}} | \mathcal{F}_{t_i}]\right|.$$

Note that the mean variation along π is an increasing function of π, i.e. we have $\mathrm{MV}(S, \pi) \leq \mathrm{MV}(S, \pi')$, whenever π' is a partition refining π: this follows from the conditional Jensen inequality $|\mathbb{E}(X|\mathcal{G})| \leq \mathbb{E}(|X|\,|\mathcal{G})$.

By definition, S is a *quasimartingale*[1] if it is adapted, $S_t \in L_1, t \in [0,1]$ and the *mean variation*

$$\mathrm{MV}(S) := \sup\nolimits_\pi \mathrm{MV}(S, \pi)$$

of S is finite. We will use that if S is bounded and càdlàg then trivially $\mathrm{MV}(S) = \lim_n \mathrm{MV}(S, D_n)$.

The stochastic analogue of the fact that a function has bounded variation if and only if it can be written as a difference of two increasing functions is then provided by the following characterization of quasimartingales, usually known as Rao's theorem ([Rao69]; see also [Pro05, Chapter 3, Theorem 17], [RW00, Chapter 6, Theorem 41.3]).

Theorem 3.1. *A càdlàg process S is a quasimartingale if and only it has a decomposition $S = Y - Z$ as the difference of two càdlàg submartingales Y and Z.*

Following [Str77], we provide a simple proof of Theorem 3.1 in Section 6.

In dealing with the mean variation of stopped processes the following lemma is useful.

[1] The study of quasimartingales goes back to Fisk [Fis65], Orey [Ore67] and Rao [Rao69].

Lemma 3.2. *Let S be a bounded process. Given a partition π and a stopping time ϱ define $\varrho+ := \inf\{t \in \pi : t \geq \varrho\}$. Then*

$$\mathrm{MV}(S^{\varrho+}, \pi) = \mathbb{E} \textstyle\sum_{t_i \in \pi} \mathbb{1}_{\{t_i < \varrho\}} \big|\mathbb{E}[S_{t_{i+1}} - S_{t_i} | \mathcal{F}_{t_i}]\big| \tag{5}$$

and $|\,\mathrm{MV}(S^{\varrho+}, \pi) - \mathrm{MV}(S^{\varrho}, \pi)| \leq 2\|S\|_\infty$.

Proof. To obtain (5), observe that for each $t_i \in \pi$

$$\mathbb{E}[S^{\varrho+}_{t_{i+1}} - S^{\varrho+}_{t_i} | \mathcal{F}_{t_i}] = \mathbb{E}[(S_{t_{i+1}} - S_{t_i})\mathbb{1}_{\{t_i<\varrho\}} | \mathcal{F}_{t_i}] = \mathbb{1}_{\{t_i<\varrho\}} \mathbb{E}[(S_{t_{i+1}} - S_{t_i}) | \mathcal{F}_{t_i}].$$

Given processes S', S'' the conditional Jensen inequality implies

$$|\,\mathrm{MV}(S', \pi) - \mathrm{MV}(S'', \pi)| \leq \mathbb{E} \textstyle\sum_{t_i \in \pi} |(S'_{t_{i+1}} - S'_{t_i}) - (S''_{t_{i+1}} - S''_{t_i})|.$$

Applying this to $S' = S^\varrho, S'' = S^{\varrho+}$ concludes the proof, as the only (possibly) non-zero term in the above sum is the one for which $\varrho \in [t_i, t_{i+1})$. □

4. The technical core

The aim of this section is to establish Theorem 2.1. To motivate our approach, assume that a continuous function $f : [0, 1] \to \mathbb{R}$ gives rise to a Riemann-Stieltjes integral

$$h \mapsto \textstyle\int h(t)\, df(t)$$

which is continuous on the space of piecewise constant functions $h : [0, 1] \to \mathbb{R}$, endowed with the sup norm. Then f has finite total variation; indeed the sequence of piecewise constant functions

$$h^n := \textstyle\sum_{t_i \in D_n} \mathbb{1}_{(t_i, t_{i+1}]} \operatorname{sign}\big(f(t_{i+1}) - f(t_i)\big)$$

is bounded uniformly and

$$\textstyle\int_0^1 h^n\, df = \sum_{t_i \in D_n} |f(t_{i+1}) - f(t_i)|$$

converges to the total variation of f. The subsequent proof is merely a translation of this standard argument to the stochastic setting, where the integrands are assumed to be adapted.

Lemma 4.1. *Let $S = (S_t)_{0\leq t\leq 1}$ be a càdlàg bounded adapted good integrator. Then for every $\varepsilon > 0$ there exist a constant C and a sequence of $[0, 1] \cup \{\infty\}$-valued stopping times $(\varrho_n)_n$ such that $\mathbb{P}(\varrho_n = \infty) \geq 1 - \varepsilon$ and $\mathrm{MV}(S^{\varrho_n}, D_n) \leq C$.*

Proof. Since S is a good integrator, given $\varepsilon > 0$ there exists $C > 0$ so that for all simple processes H with $\|H\|_\infty \leq 1$ we have $\mathbb{P}((H \cdot S)_1 \geq C - 2\|S\|_\infty) \leq \varepsilon$. For each n we define the simple process H^n and the stopping time ϱ_n as

$$H^n := \textstyle\sum_{t_i \in D_n} \mathbb{1}_{(t_i, t_{i+1}]} \operatorname{sign}\big(\mathbb{E}[S_{t_{i+1}} - S_{t_i} | \mathcal{F}_{t_i}]\big),$$
$$\varrho_n := \inf\{t \in D_n : (H^n \cdot S)_t \geq C - 2\|S\|_\infty\}.$$

Notice that, on the set $\{\varrho_n < \infty\}$,

$$(H^n 1_{(0,\varrho_n]}) \cdot S = (H^n \cdot S)^{\varrho_n} \text{ satisfies } (H^n \cdot S)^{\varrho_n}_1 \geq C - 2\|S\|_\infty,$$

and thus $\mathbb{P}(\varrho_n = \infty) \geq 1 - \varepsilon$. Moreover, since the jumps of S are bounded by $2\|S\|_\infty$, $C \geq (H^n \cdot S)^{\varrho_n}_1$ holds, so we find, with the help of lemma 3.2 ,

$$C \geq \mathbb{E}(H^n \cdot S)^{\varrho_n}_1 = \mathbb{E} \sum_{t_i \in D_n} \mathbb{1}_{\{t_i<\varrho_n\}} \operatorname{sign}\big(\mathbb{E}[S_{t_{i+1}} - S_{t_i} | \mathcal{F}_{t_i}]\big)(S_{t_{i+1}} - S_{t_i}) =$$
$$= \mathbb{E} \sum_{t_i \in D_n} \mathbb{1}_{\{t_i<\varrho_n\}} \Big|\mathbb{E}[(S_{t_{i+1}} - S_{t_i}) | \mathcal{F}_{t_i}]\Big| = \mathrm{MV}(S^{\varrho_n}, D_n). \quad \square$$

Given that $\mathrm{MV}(S^{\varrho_n}, D_k) \le C$ for every $k \le n$, it is desirable to define an "accumulation stopping time" ϱ of the stopping times $(\varrho_n)_n$, so that $\mathrm{MV}(S^{\varrho}, D_k) \le C$ will hold for every k, proving that S^ϱ is a quasimartingale. A similar technique is also used in [BSV11a, Proposition 3.6].

Lemma 4.2. *Assume that $(\varrho_n)_n$ is a sequence of $[0,1]\cup\{\infty\}$-valued stopping times such that $\mathbb{P}(\varrho_n = \infty) \ge 1-\varepsilon$, $n \ge 1$ for some $\varepsilon > 0$. Then there exists a stopping time ϱ and for each $n \ge 1$ convex weights $\mu_n^n, \dots, \mu_{N_n}^n$ such that*[2] *$\mathbb{P}(\varrho = \infty) \ge 1-3\varepsilon$ and for all $n \ge 1$*

$$\mathbb{1}_{[0,\varrho]} \le 2 \textstyle\sum_{k=n}^{N_n} \mu_k^n \mathbb{1}_{[0,\varrho_k]}. \tag{6}$$

Proof of Lemma 4.2. Recall the following classical result by Mazur: if $(f_n)_n$ is a bounded sequence in a Hilbert space then there exists vectors $g_n \in \mathrm{conv}(f_n, f_{n+1}, \dots)$, $n \ge 1$ such that $(g_n)_n$ converges in Norm.[3] We apply this to the random variables

$X_n = \mathbb{1}_{\{\varrho_n = \infty\}} \in L^2(\mathbb{P}), n \ge 1$ to obtain for each n convex weigths $\mu_n^n, \dots, \mu_{N_n}^n$ such that

$$Y_n := \mu_n^n X_n + \ldots + \mu_n^{N_n} X_{N_n}$$

converges to some random variable X in $L^2(\mathbb{P})$. Relabeling sequences if necessary, we assume that the convergence holds also almost surely.

From $X \le 1$ and $\mathbb{E}[X] \ge 1-\varepsilon$ we deduce that $\mathbb{P}(X < 2/3) < 3\varepsilon$. Since $\mathbb{P}(\lim_m Y_m \ge 2/3) > 1-3\varepsilon$, by Egoroff's theorem we deduce that there exists a set A with $\mathbb{P}(A) \ge 1-3\varepsilon$ such that $Y_n \ge 1/2$ on the set A, for all n greater or equal than some $n_0 \in \mathbb{N}$, which we can assume to be equal to 1.

We now define the desired stopping time ϱ by

$$\varrho = \inf_{n\ge1} \inf\{t : \mu_n^n \mathbb{1}_{[0,\varrho_n]}(t) + \ldots + \mu_n^{N_n} \mathbb{1}_{[0,\varrho_{N_n}]}(t) < 1/2\}.$$

Then clearly (6) holds, and from $A \subseteq \{\varrho = \infty\}$ we obtain $\mathbb{P}(\varrho = \infty) \ge 1-3\varepsilon$. □

We are now in the position to complete the proof of Theorem 2.1

Proof of Theorem 2.1. Given $\varepsilon > 0$, pick C, $(\varrho_n)_n$ and ϱ according to Lemma 4.1 resp. Lemma 4.2. Fixing $n \ge 1$ we obtain from (6) that

$$\mathbb{E} \sum_{t_i \in D_n} \mathbb{1}_{\{t_i < \varrho\}} \Big| \mathbb{E}[S_{t_{i+1}} - S_{t_i} \| \mathcal{F}_{t_i}] \Big| \le 2\mathbb{E} \sum_{t_i \in D_n} \sum_{k=n}^{N_n} \mu_k^n \mathbb{1}_{\{t_i < \varrho_k\}} \Big| \mathbb{E}[S_{t_{i+1}} - S_{t_i} | \mathcal{F}_{t_i}] \Big|. \tag{7}$$

By Lemma 3.2, $\mathrm{MV}(S^\varrho, D_n)$ differs from the left side of (7) by at most $2\|S\|_\infty$. Applying Lemma 3.2 once more, the right side of (7) is bounded by

$$2 \textstyle\sum_{k=n}^{N_n} \mu_k^n (\mathrm{MV}(S^{\varrho_k}, D_n) + 2\|S\|_\infty) \le 2C + 4\|S\|_\infty.$$

Combining these facts and letting $n \to \infty$ we conclude $\mathrm{MV}(S^\varrho) \le 2C + 6\|S\|_\infty$. By Rao's theorem 3.1 this yields Theorem 2.1. □

[2] We note that the constant 2 in (6) can be replaced by $1+\delta$, for $\delta > 0$ in which case one is only guaranteed to find ϱ satisfying $\mathbb{P}(\varrho = \infty) \ge 1 - \eta\varepsilon$ for $\eta > (1-(1+\delta)^{-1})^{-1}$. But we do not need this.

[3] This can be seen as a consequence of weak compactness combined with the fact that weak and strong closure coincide for convex sets. Alternatively one may simply pick the elements g_n to have (asymptotically) minimal norm in $\mathrm{conv}(f_n, f_{n+1}, \dots)$, $n \ge 1$.

5. Every good integrator is a Semimartingale

In this section we show in detail how Theorem BD follows from Theorem 2.1. All arguments are however quite standard.

Lemma 5.1. *Let a process S be locally a semimartingale. Then S is a semimartingale.*

Proof. By definition, S is locally a semi martingale then there exists a sequence $(\varrho_n)_n$ of stopping times such that $\mathbb{P}(\varrho_n < \infty) \leq 2^{-n}$ and for each n a Martingale M_n and a process A_n of finite variation such that $S^{\varrho_n} = M_n + A_n$. Set $\sigma_k := \inf_{n \geq k} \varrho_n$. Then

$$\begin{aligned} S &= S^{\sigma_1} + (S^{\sigma_2} - S^{\sigma_1}) + (S^{\sigma_3} - S^{\sigma_2}) + \ldots \\ &= [M_1^{\sigma_1} + (M_2^{\sigma_2} - M_1^{\sigma_1}) + \ldots] + [A_1^{\sigma_1} + (A_2^{\sigma_2} - A_1^{\sigma_1}) + \ldots], \end{aligned}$$

showing that S is the sum of a local martingale and a process of finite total variation. □

Recall that a process X is of class D if the family $\{X_\sigma : \sigma \text{ stopping time}\}$ is uniformly integrable.

Lemma 5.2. *Let $S = (S_t)_{0 \leq t \leq 1}$ be a càdlàg submartingale. Then S is locally of class D.*

Proof. Define the stopping time

$$T_n := \inf\{t \in [0,1] : |S_t| \geq n\},$$

then if σ is an arbitrary stopping time we have that $|S_\sigma^{T_n}| \leq n + |S_{1 \wedge T_n}|$. By the optional sampling theorem[4] $S_{1 \wedge T_n}$ is integrable, showing that $\{S_\sigma^{T_n} : \sigma \text{ stopping time}\}$ is uniformly integrable. □

Proof of Theorem BD. We note that S can be written as the sum two adapted processes, one of finite variation and one locally bounded: indeed, since S is càdlàg, $\Delta S_t := S_t - S_{t-}$ and $J_t := \sum_{0<s\leq t} \Delta S_t \mathbb{1}_{\{|\Delta S_t| \geq 1\}}$ are well defined (the sum defining $J_t(\omega)$ is finite for each t, ω). Since J has finite variation and is adapted, and $S - J$ has bounded jumps, $S = J + (S - J)$ is a decomposition as required. Notice that J is a càdlàg good integrator (since it has finite variation), and so such is $S - J$. Thus, we may assume without loss of generality that there exists a localizing sequence $(\varrho_n)_n$ such that each $S^{\varrho_n}, n \geq 1$ is bounded. By Theorem 2.1 it follows that S^{ϱ_n} is locally the difference of two càdlàg submartingales. By Lemma 5.2 and the Doob-Meyer decomposition theorem S^{ϱ_n} is locally a semimartingale, and thus applying twice Lemma 5.1 we obtain that S is a semimartingale. □

6. Proof of Rao's Theorem

Below we provide the classical proof of Rao's Theorem 3.1. We record the following consequence of the proof, since it will be useful in the next section.

Remark 6.1. *If a quasimartingale is right continuous in probability, then it has a càdlàg modification.*

Proof of Theorem 3.1. We will prove here the statement with submartingales replaced by supermartingales, which is obviously the same. Trivally the difference of two supermartingales is a quasimartingale, so let us prove the converse.

[4] For a proof see [KS91, Theorem 3.22] or [RW00, Theorem II.77.1].

For any n, we define positive supermartingales ${}^{\pm}Y^n$ on the dyadic times $s \in D := \bigcup_k D_k$ by

$$^{\pm}Y^n_s := \mathbb{E}\Big[\sum_{t_i \in D_n \cap [s,1]} \mathbb{E}[S_{t_i} - S_{t_{i+1}} | \mathcal{F}_{t_i}]^{\pm} \Big| F_s\Big],$$

so that ${}^{+}Y^n_s - {}^{-}Y^n_s$ equals $S_s - E[S_1|F_s]$ for n big enough that $s \in D_n$.

Since

$$\mathbb{E}[X|G]^+ \leq \mathbb{E}[X^+|G] \text{ implies that } 0 \leq {}^{\pm}Y^n_s \leq {}^{\pm}Y^{n+1}_s,$$

the monotone convergence theorem and $\mathbb{E}[{}^{\pm}Y^n_s] \leq \mathrm{MV}(S, D)$ yield that, for any $s \in D$, $({}^{\pm}Y^n_s)_n$ converges almost surely and in L^1. Thus, the processes indexed by the dyadic times

$$^{\pm}\hat{Y}_s := \mathbb{E}[S^{\pm}_1|F_s] + \lim_n {}^{\pm}Y^n_s$$

are positive supermartingales such that almost surely ${}^{+}\hat{Y}_s - {}^{-}\hat{Y}_s = S_s$ for any $s \in D$. By Doob's regularity theorem (for a proof see [KS91, Prop 1.3.14] or [RY99, Chapter II, Prop 2.4]), and since the filtration is right continuous, the positive processes defined by

$$^{\pm}Y_t := \lim_{s \downarrow t, s \in D} {}^{\pm}\hat{Y}_s$$

whenever the limit exists, and defined to be 0 otherwise, are supermartingales which are càdlàg a.s. From the right continuity in probability of S it follows that, for every $s \in [0, 1]$, ${}^{+}Y_s - {}^{-}Y_s = S_s$ holds a.s.; is particular ${}^{+}Y - {}^{-}Y$ is an a.s. càdlàg modification of S. Moreover, if S has also right continuous paths, it is indistinguishable from ${}^{+}Y - {}^{-}Y$. Since the filtration is saturated, one can modify ${}^{\pm}Y$ on a null set, making them càdlàg and such that $S = {}^{+}Y - {}^{-}Y$. □

We remark that the above proof of Theorem 3.1 shows that the two submartingales Y, Z of Theorem 2.1 can be chosen to be negative.

7. Ramifications of the Bichteler-Dellacherie theorem

In this section we prove that Riemann integrators are good integrators, and somewhat strengthen Theorem BD.

It is well known that, in the definition of good integrators, the space $\mathcal{S}$ can be replaced by the subset of *elementary integrands*, which consists of all processes H of the form

$$H = \sum_{i=1}^{k} H^i \mathbb{1}_{(t_i, t_{i+1}]}, \tag{8}$$

where t_i are *deterministic* times such that $0 \leq t_1 < \ldots < t_{k+1} = 1$, and each H_i is bounded $\mathcal{F}_{t_i}$-measurable. In Remark 7.1 we prove this fact in a slightly stronger form, which will be useful in what follows.

Let $\mathcal{E}_{D_n}$ be the space of all processes H of the form

$$H = \sum_{i=0}^{2^n - 1} H^i \mathbb{1}_{(\frac{i}{2^n}, \frac{i+1}{2^n}]}, \tag{9}$$

where, for each $i = 1, ..., 2^n - 1$, H^i is bounded and $\mathcal{F}_{\frac{i-1}{2^n}}$-measurable (not only $\mathcal{F}_{\frac{i}{2^n}}$-measurable), and $H^0 = 0$; then, define $\mathcal{E}_D := \bigcup_{n \geq 1} \mathcal{E}_{D_n}$.

Lemma 7.1. *Let S be an adapted process which is right continuous in probability. Then $\mathcal{I}_S : \mathcal{S} \to L^0(\mathbb{P})$ is a continuous operator if and only if its restriction to $\mathcal{E}_D$ is continuous.*

Proof. We have to show that if $\mathcal{I}_S$ is a bounded operator on $\mathcal{E}_D$, then it is also a bounded operator on $\mathcal{S}$. Given $\varepsilon > 0$, pick $C > 0$ such that $\mathbb{P}(|\mathcal{I}_S(K)| > C) < \varepsilon$ for every process $K \in \mathcal{E}_D$ satisfying $\|K\|_\infty \leq 1$. Let H be a simple integrand as in (2) and satisfying $\|H\|_\infty \leq 1$, and define the stopping times

$$\sigma^n_i := 1 \wedge (i + 2)/2^n \text{ on } \{i/2^n < \tau_i \leq (i + 1)/2^n\}.$$

Then the process

$$K^n := \sum_{i=1}^k H^i \mathbb{1}_{(\sigma_i^n, \sigma_{i+1}^n]}$$

is actually in $\mathcal{E}_{D_n}$: this follows from the fact that the stopping times $(\sigma_i^n)_i$ have values in D_n and satisfy $\tau_i + 1/2^n \leq \sigma_i^n$, while H^i is $\mathcal{F}_{\tau_i}$-measurable. Moreover $\mathcal{I}_S(K^n)$ converges to $\mathcal{I}_S(H)$ in probability (since S is right continuous) and so, taking n big enough, it follows that

$$\mathbb{P}(|\mathcal{I}_S(H)| > C) \leq \mathbb{P}(|\mathcal{I}_S(H) - \mathcal{I}_S(K^n)| > C) + \mathbb{P}(|\mathcal{I}_S(K^n)| > C) < 2\varepsilon.$$

Since the choice of C was independent of $H \in \mathcal{S}$, this proves that $\mathcal{I}_S$ is bounded on $\mathcal{S}$. □

As a corollary of Theorem BD, we obtain that semimartingales can be characterized by Riemann-sums. Indeed, if S is a semimartingale, the stochastic dominated convergence theorem implies that, for every left-continuous adapted process H, the random variables

$$\sum_{\tau_i \in \pi_n} H_{\tau_i}(S_{\tau_{i+1}} - S_{\tau_i})$$

converge in probability (to $\mathcal{I}_S(H)$) as $n \to \infty$, for any sequence $(\pi_n)_n$ of random partitions whose mesh is going to 0. Conversely, we find that this property characterizes semimartingales. Indeed, define a càdlàg adapted process $(S_t)_{0 \leq t \leq 1}$ to be a *Riemann integrator* if for every bounded adapted continuous process H the sequence of random variables

$$\sum_{i=0}^{2^n-1} H_{\frac{i}{2^n}}(S_{\frac{i+1}{2^n}} - S_{\frac{i}{2^n}})$$

converges in probability as $n \to \infty$. Then, the following holds:

Corollary 7.2. *Every Riemann integrator is a semimartingale.*

To prove Corollary 7.2 we need some additional definitions. Consider the Banach space $L^\infty(\Omega; C^0([0,1]))$ of all bounded continuous processes $(H_t)_{t \in [0,1]}$, endowed with the sup norm (3). Let X be the subspace constituted by the processes which are adapted; this is a closed subspace, and hence a Banach space with the induced norm. Finally, define the linear continuous operator $\mathcal{I}_S^n : X \to L^0(\mathbb{P})$ by

$$\mathcal{I}_S^n(H) := \mathcal{I}_S(H^{D_n}), \text{ where } H^{D_n} := \sum_{i=0}^{2^n-1} H_{\frac{i}{2^n}} \mathbb{1}_{(\frac{i}{2^n}, \frac{i+1}{2^n}]}.$$

By definition, S is a Riemann integrator if, for every $H \in X$, $\mathcal{I}_S^n(H)$ converges in probability as $n \to \infty$. The Banach-Steinhaus theorem[5] [Rud91, Theorem 2.6] then yields the following:

Lemma 7.3. *If S is a Riemann good integrator, then for every $\varepsilon > 0$ there is some $C > 0$ such that $\mathbb{P}(\mathcal{I}_S^n(H) \geq C) \leq \varepsilon$ for all $n \geq 1$ and all continuous adapted processes H such that $\|H\|_\infty \leq 1$.*

It is now fairly straightforward to establish that every Riemann integrator is a good integrator.

Proof of Theorem 7.2. Let $H \in \mathcal{E}_{D_n}$ be as in (9) and satisfy $\|H\|_\infty \leq 1$. Define a process K by declaring it equal to H^i at time $t = i/2^n$, for $0 \leq i \leq 2^n - 1$, and equal to zero at time 1, and extending it to $t \in [0,1]$ by affine interpolation. Then K is a continuous *adapted* process such that $\|K\|_\infty \leq 1$ and $\mathcal{I}_S(H) = \mathcal{I}_S^n(K)$. Thus, by Lemma 7.3, $\mathcal{I}_S$ is bounded on $\mathcal{E}_D$. Lemma 7.1 then shows that S is a good integrator, and so Theorem BD implies that S is a semimartingale. □

[5]Which is also commonly called "the uniform boundedness principle".

If we want to define the stochastic integral with respect to a process S in a way that the stochastic dominated convergence theorem holds, we must require the process S to be right continuous in probability.[6] The following well known[7] theorem shows that is minimal continuity assumption is actually strong enough for the purpose.

Theorem 7.4. *If a good integrator is right continuous in probability, then it has a càdlàg modification.*

Conveniently enough, the proof of Theorem 2.1 yields this result *automatically* if the good integrator is bounded (this follows from Remark 6.1). With just a little more work we can remove the assumption of boundedness, by proving first the invariance of good integrators under convex transformations, which is interesting in itself.

Lemma 7.5. *Assume that S is a good integrator and let $f:\mathbb{R}\to\mathbb{R}$ be a convex map whose left derivative f' is bounded. Then $f(S)$ is a good integrator.*

Proof. Given a partition π of $[0,1]$, define the processes

$$G^\pi := \textstyle\sum_{t_i\in\pi} f'(S_{t_i})\mathbb{1}_{(t_i,t_{i+1}]}, \quad A^\pi := f(S)-f(S_0)-G^\pi\cdot S\,.$$

Then G^π is a simple integrand and, for $t\in\pi$, the representation of A^π_t as a telescopic sum

$$A^\pi_t = \textstyle\sum_{t_j<t, t_j\in\pi} f(S_{t_{j+1}})-f(S_{t_j})-f'(S_{t_j})(S_{t_{j+1}}-S_{t_j}),$$

shows that $(A_t)_{t\in\pi}$ is increasing. Thus, given a simple integrand H of the form (2) with $t_i\in\pi$ for every i, we obtain the estimate

$$|(H\cdot f(S))_1| = |(H\cdot A^\pi)_1 + (H\cdot(G^\pi\cdot S))_1| \le \|H\|_\infty |A^\pi_1| + |((HG^\pi)\cdot S)_1|.$$

Since S is a good integrator, A^π_1 as well as $((HG^\pi)\cdot S)_1$ are bounded in probability, uniformly over all partitions π and all simple integrands H satisfying $\|H\|_\infty\le 1$, which concludes the proof of the lemma. □

Proof of Theorem 7.4. By the previous lemma, $\tan^{-1}(S)$ is a good integrator, as $\tan^{-1}$ is the difference of two convex functions with bounded derivatives. The arguments in Section 4 remain valid if the assumption of càdlàg paths is relaxed to right continuity in probability, thus proving that $\tan^{-1}(S)$ has locally finite mean variation. Remark 6.1 then yields that (locally and thus globally) there exists a càdlàg modification of $\tan^{-1}(S)$; so S has a càdlàg modification. □

Remark 7.6. Lemma 7.5 and its proof admit a trivial generalization to the multi-dimensional case: one just needs to replace the left derivative of f with a Borel-measurable selection of the sub-differential of f.

Moreover it clearly sufficient for f to be defined (and have bounded "derivative") on an open convex set on which S takes its values; in particular applying Lemma 7.5 to the functions $f(x)=|x|$ and $f(x)=x^2$ and pre-localizing S (to make it bounded) it follows that the space of good integrators is a lattice and an algebra.

---

[6] Indeed, since $\mathbb{1}_{(t,t+\frac{1}{n}]}$ converges to zero, $\mathcal{I}_S(\mathbb{1}_{(t,t+\frac{1}{n}]}) = S_{t+\frac{1}{n}} - S_t$ must converge to zero in probability.

[7] See, for instance, [DM88, Chapter 8, Paragraph 81] or [Bic81, Theorem 2.3.4]

References

[Bas96] R. F. Bass. The Doob-Meyer decomposition revisited. *Canad. Math. Bull.*, 39(2):138–150, 1996.

[Bic79] K. Bichteler. Stochastic integrators. *Bull. Amer. Math. Soc. (N.S.)*, 1(5):761–765, 1979.

[Bic81] K. Bichteler. Stochastic integration and L^p-theory of semimartingales. *Ann. Probab.*, 9(1):49–89, 1981.

[BSV11a] M. Beiglböck, W. Schachermayer, and B. Veliyev. A direct proof of the Bichteler-Dellacherie theorem and connections to arbitrage. *Ann. Probab.*, 2011. to appear.

[BSV11b] M. Beiglböck, W. Schachermayer, and B. Veliyev. A short proof of the Doob-Meyer theorem. *Stochastic Process. Appl.*, 2011. to appear.

[Del80] C. Dellacherie. Un survol de la théorie de l'intégrale stochastique. *Stochastic Process. Appl.*, 10(2):115–144, 1980.

[DM88] C. Dellacherie and P.-A. Meyer. *Probabilities and potential. C*, volume 151 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1988. Potential theory for discrete and continuous semigroups, Translated from the French by J. Norris.

[Edw90] D. A. Edwards. A note on stochastic integrators. *Math. Proc. Cambridge Philos. Soc.*, 107(2):395–400, 1990.

[Fis65] Donald L. Fisk. Quasi-martingales. *Trans. Amer. Math. Soc.*, 120:369–389, 1965.

[Jak05] A. Jakubowski. An almost sure approximation for the predictable process in the Doob-Meyer decomposition theorem. In *Séminaire de Probabilités XXXVIII*, volume 1857 of *Lecture Notes in Math.*, pages 158–164. Springer, Berlin, 2005.

[KS91] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[Low11] G. Lowther. Almost sure. *http://almostsure.wordpress.com*, 2011.

[Ore67] Steven Orey. F-processes. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), Vol, II: Contributions to Probability Theory, Part 1*, pages 301–313. Univ. California Press, Berkeley, Calif., 1967.

[Pro05] P. E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2005. Second edition. Version 2.1, Corrected third printing.

[Rao69] K. Murali Rao. Quasi-martingales. *Math. Scand.*, 24:79–92, 1969.

[Rud91] Walter Rudin. *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill Inc., New York, second edition, 1991.

[RW00] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 2*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Itô calculus, Reprint of the second (1994) edition.

[RY99] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[Str77] Christophe Stricker. Quasimartingales, martingales locales, semimartingales et filtration naturelle. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 39(1):55–63, 1977.