

RIEMANN-INTEGRATION AND A NEW PROOF OF THE BICHTELER–DELLACHERIE THEOREM

M. BEIGLBÖCK[†] AND P. SIORPAES[‡]

ABSTRACT. We give a new 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.

Keywords: Bichteler–Dellacherie Theorem, semimartingale decomposition, good integrators.

Mathematics Subject Classification (2010): 60G05

1. INTRODUCTION

The Bichteler–Dellacherie theorem 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; in this paper we provide a new proof of this celebrated result, together with a new characterization of semimartingales.

The Doob-Meyer decomposition theorem leads to the following reformulation of the Bichteler–Dellacherie theorem: 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

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

[†] University of Vienna, Faculty of Mathematics, *email:* mathias.beiglboeck@univie.ac.at, *Corresponding author*

[‡] University of Vienna, Faculty of Mathematics, *email:* pietro.siorpaes@univie.ac.at.
The first author thanks the Austrian science fund FWF for support through project p21209. Both authors thank Josef Teichmann and Pianta Giacomo for useful comments.

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

We notice that this is in remarkable contrast to a fact which Meyer [Mey81] attributes to Jeulin: there are continuous processes which are not semimartingales for which (1) holds for all integrands H of the type $H_t = f(t, S_t)$, where f is a bounded continuous function.

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

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

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

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

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

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

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 $\mathcal{I}_S : \mathcal{S} \rightarrow L^0(\mathbb{P})$ is continuous, i.e. if $H^n \in \mathcal{S}$, $\|H^n\|_\infty \rightarrow 0$ implies that $\mathcal{I}_S(H^n)$ goes to 0 in probability as $n \rightarrow \infty$.

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. *Let $(S_t)_{0 \leq t \leq 1}$ be a càdlàg adapted process. If $\mathcal{I}_S : S \rightarrow L^0(\mathbb{P})$ 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 has a long history, tracing back to the Rennes school of Metivier and Pellaumail (see for example [Pel73, MP77]), and then evolving in the Strasbourg school of Meyer; it was first published in its present form in [Mey79] and, independently, [Bic79, Bic81]. Mokobodzki deserves particular credit (see for instance the discussion in [DM82]); however since the result is usually baptized after Bichteler and Dellacherie, we stick to this name.

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 $\mathcal{G}_t := \mathcal{F}_1$, $0 \leq t \leq 1$. 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 $S = (S_t)_{0 \leq t \leq 1}$ 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 $\mathbb{P}(\varrho = \infty) \geq 1 - \varepsilon$ (by definition $\mathcal{F}_\infty := \mathcal{F}_1$).

Our main contribution consists in a new 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, BSV11b].

The most popular accounts on Theorem BD employ functional analytic machinery and change of measure techniques, as in Dellacherie–Meyer [DM82]. A modern version of this argument is given in Protter [Pro05]; specifically, one applies a variant of the Hahn–Banach separation theorem (due to Yan [Yan80]) to construct an equivalent measure \mathbb{Q} under which the good integrator S is a quasimartingale. Using the Theorems of Rao and Doob–Meyer it follows that S is a \mathbb{Q} -semimartingale; finally, Girsanov’s Theorem implies that S is also a \mathbb{P} -semimartingale.

Some accounts use the notion of a good integrator (or a similar concept) as starting point to develop the theory of stochastic integration (see e.g. [Mét77,

MP80, Mét82, Pro05]). This greatly simplifies many proofs; however, the useful link to the classical approach based on semimartingales is only obtained a posteriori by Theorem BD. Since the proof of Theorem BD given in the present article does not rely on stochastic calculus nor Girsanov's theorem the equivalence of the two approaches could be established from the start, enabling further simplifications.

For an alternative approach to Theorem BD, based on an orthogonal decomposition, see Lowther [Low11]. A different idea is developed in [BSV11b] which, like the present paper, has an elementary proof based on discrete time arguments, and does not use change of measure techniques; it thus seems interesting to describe this approach more closely.

In [BSV11b], Theorem BD is obtained as a corollary of the fact that every bounded process S satisfying a certain weak No-Arbitrage condition is a semimartingale ([BSV11b, Theorem 1.6]). To prove the latter result, the authors take the discrete time Doob-decomposition $S = M^n + A^n$ of S restricted to the dyadic times of n^{th} -generation and, repeatedly applying the No-Arbitrage property, show that the sequences $(M^n)_n, (A^n)_n$ can be controlled on suitably chosen (random) intervals $[0, \tau_n]$. Using carefully chosen convex combinations it is then possible to pass to limits and obtain processes M, A on $[0, \tau]$ such that M is a martingale, A has finite variation (but is not necessarily predictable), and τ is an arbitrarily “big” stopping time. This comes with necessity to develop quite intricate estimates on the approximations as well as a somewhat complex limiting procedure which takes into account the approximating processes $(M^n)_n, (A^n)_n$ and the intervals $[0, \tau_n]$ *simultaneously*.

This 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. Finally, in Section 6 we discuss certain ramifications of the Theorem BD (including the characterization (1)).

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. $D_n := \{0, 1/2^n, 2/2^n, \dots, 1\}$. 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 $X \in F$. A simple proof, analogous to the one for normed spaces, shows that 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.

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 $\pi = \{0 = t_0 < t_1 < \dots < t_n = 1\}$ of $[0, 1]$, the *mean variation* of S along π is defined as

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

Note that the mean variation along π is an increasing function of π , i.e. we have $\text{MV}(S, \pi) \leq \text{MV}(S, \pi')$, 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*

$$\text{MV}(S) := \sup_{\pi} \text{MV}(S, \pi)$$

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

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; its standard proof, found in most stochastic calculus textbooks (e.g. [Pro05, Chapter 3, Theorem 17], [RW00, Chapter 6, Theorem 41.3], and [DM82, Chapter 6, Theorem 40]), is both short and elementary.

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

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

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*

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

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

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

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

¹ The study of quasimartingales goes back to Fisk [Fis65], Orey [Ore67], Rao [Rao69], and Stricker [Str77].

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

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

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

4. THE TECHNICAL CORE

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

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

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

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

is bounded uniformly and

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

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 $\text{MV}(S^{\varrho_n}, D_n) \leq C$.*

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

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

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

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

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

$$\begin{aligned} C &\geq \mathbb{E}(H^n \cdot S)_1^{\varrho_n} = \mathbb{E} \sum_{t_i \in D_n} \mathbb{1}_{\{t_i < \varrho_n\}} \text{sign} \left(\mathbb{E}[S_{t_{i+1}} - S_{t_i} | \mathcal{F}_{t_i}] \right) (S_{t_{i+1}} - S_{t_i}) = \\ &= \mathbb{E} \sum_{t_i \in D_n} \mathbb{1}_{\{t_i < \varrho_n\}} \left| \mathbb{E}[(S_{t_{i+1}} - S_{t_i}) | \mathcal{F}_{t_i}] \right| = \text{MV}(S^{\varrho_n}, D_n). \quad \square \end{aligned}$$

Given that $\text{MV}(S^{\varrho_n}, D_k) \leq C$ for every $k \leq n$, it is desirable to define an “accumulation stopping time” ϱ of the stopping times $(\varrho_n)_n$, so that $\text{MV}(S^\varrho, D_k) \leq C$ will hold for every k , proving that S^ϱ is a quasimartingale. Ideally, we would want ϱ to be “as big as the ϱ_n ”, and yet such that $\varrho \leq \varrho_{n_k}$ holds for some subsequence n_k . This is not quite possible; however, after rephrasing the previous inequality as $\mathbb{1}_{[0, \varrho]} \leq \mathbb{1}_{[0, \varrho_{n_k}]}$, we can soften this requirement (passing to forward convex combinations instead of subsequences), thus making it compatible with ϱ being “big”. 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 $\mathbb{P}(\varrho_n = \infty) \geq 1 - \varepsilon$, $n \geq 1$ for some $\varepsilon > 0$. Then there exists a stopping time ϱ and for each $n \geq 1$ convex weights $\mu_n^n, \dots, \mu_{N_n}^n$ such that² $\mathbb{P}(\varrho = \infty) \geq 1 - 3\varepsilon$ and for all $n \geq 1$*

$$(6) \quad \mathbb{1}_{[0, \varrho]} \leq 2 \sum_{k=n}^{N_n} \mu_k^n \mathbb{1}_{[0, \varrho_k]}.$$

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 exist vectors $g_n \in \text{conv}(f_n, f_{n+1}, \dots)$, $n \geq 1$ such that $(g_n)_n$ converges in Norm.³ We apply this to the random variables

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

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

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 $X \leq 1$ and $\mathbb{E}[X] \geq 1 - \varepsilon$ we deduce that $\mathbb{P}(X < 2/3) < 3\varepsilon$. Since $\mathbb{P}(\lim_m Y_m \geq 2/3) > 1 - 3\varepsilon$, by Egoroff’s theorem we deduce that there exists a set A with $\mathbb{P}(A) \geq 1 - 3\varepsilon$ such that $Y_n \geq 1/2$ 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.

²We 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 $\mathbb{P}(\varrho = \infty) \geq 1 - \eta\varepsilon$ for $\eta > (1 - (1 + \delta)^{-1})^{-1}$. But we do not need this.

³This 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 $\text{conv}(f_n, f_{n+1}, \dots)$, $n \geq 1$.

We now define the desired stopping time ϱ by

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

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

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 $n \geq 1$ we obtain from (6) that

$$(7) \quad \mathbb{E} \sum_{t_i \in D_n} \mathbb{1}_{\{t_i < \varrho\}} \left| \mathbb{E}[S_{t_{i+1}} - S_{t_i} | \mathcal{F}_{t_i}] \right| \leq 2\mathbb{E} \sum_{t_i \in D_n} \sum_{k=n}^{N_n} \mu_k^n \mathbb{1}_{\{t_i < \varrho_k\}} \left| \mathbb{E}[S_{t_{i+1}} - S_{t_i} | \mathcal{F}_{t_i}] \right|.$$

By Lemma 3.2, $MV(S^\varrho, D_n)$ 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

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

Combining these facts and letting $n \rightarrow \infty$ we conclude $MV(S^\varrho) \leq 2C + 6\|S\|_\infty$. By Rao's theorem 3.1 this yields Theorem 2.1. \square

5. EVERY GOOD INTEGRATOR IS A SEMIMARTINGALE

In this section, for the convenience of the reader, we show in detail how Theorem BD follows from Theorem 2.1; all arguments are however quite standard. For a proof of the following lemma one can also consult [JS03, Proposition 4.25(b), Chapter 1].

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

Proof. If $S = (S_t)_{t \in [0,1]}$ is locally a semimartingale there exists a sequence $(\sigma_n)_n$ of stopping times such that $\mathbb{P}(\sigma_n \leq 1) \rightarrow 0$ and, for each n , a local martingale M_n and a process A_n of finite variation such that $S^{\sigma_n} = M_n + A_n$. By passing to a subsequence n_i s.t. $\mathbb{P}(\sigma_{n_i} \leq 1) \leq 2^{-i}$ and then replacing σ_k with $\varrho_k := \inf_{i \geq k} \sigma_{n_i}$ we can assume that $(\sigma_n)_n$ is increasing (indeed $(\varrho_k)_k$ is increasing and $\mathbb{P}(\varrho_k \leq 1) \leq 2^{-(k-1)} \rightarrow 0$). Since $S^{\sigma_n} = M_n^{\sigma_n} + A_n^{\sigma_n}$ equals $(S^{\sigma_{n+1}})^{\sigma_n} = M_{n+1}^{\sigma_n} + A_{n+1}^{\sigma_n}$,

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

where for each (t, ω) only one term in each sum is non-zero. Since A is of finite variation and M is locally a local martingale, and thus a local martingale (see [JS03, Lemma 1.35(a), Chapter 1]), S is a semimartingale. \square

Notice that the previous lemma implies that Theorem 2 applies also when the time index $[0, 1]$ is replaced with $[0, \infty)$. 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 theorem⁴ $S_{1 \wedge T_n}$ is integrable, showing that $\{S_\sigma^{T_n} : \sigma \text{ stopping time}\}$ is uniformly integrable. \square

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 $J_t := \sum_{0 < s \leq t} \Delta S_s \mathbb{1}_{\{|\Delta S_s| \geq 1\}}$ 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 by localizing and using Lemma 5.1 we may assume without loss of generality that S is bounded. By Theorem 2.1 it follows that S is locally the difference of two càdlàg submartingales. By Lemma 5.2 and the Doob–Meyer decomposition theorem S is locally a local semimartingale, and thus applying twice Lemma 5.1 we obtain that S is a semimartingale. \square

6. 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

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

where t_i are *deterministic* times such that $0 \leq t_1 < \dots < t_{k+1} = 1$, and each H^i is bounded \mathcal{F}_{t_i} -measurable. In Lemma 6.1 we prove this fact in a slightly stronger form, which will be useful in proving Corollary 6.2.

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

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

where, for each $i = 1, \dots, 2^n - 1$, 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 $\mathcal{E}_D := \bigcup_{n \geq 1} \mathcal{E}_{D_n}$.

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

Lemma 6.1. *Let S be an adapted process which is right continuous in probability. Then $\mathcal{I}_S : \mathcal{S} \rightarrow L^0(\mathbb{P})$ 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

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

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}_{τ_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 C was chosen independent of $H \in \mathcal{S}$, this proves that \mathcal{I}_S is bounded on \mathcal{S} . \square

The previous lemma could be reformulated as follows: a cadlag adapted process S is a good integrator iff $\mathcal{I}_S(H^n) \rightarrow 0$ in probability whenever $\|H^n\|_\infty \rightarrow 0$ and $H^n \in \mathcal{E}_{D_n}$ for all n .

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 (resp. cadlag) 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 $\mathcal{I}_S(H)$, resp. $\mathcal{I}_S(H_-)$) as $n \rightarrow \infty$, 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 process⁵ H the sequence of random variables

$$(10) \quad \sum_{i=0}^{2^n-1} H_{\frac{i}{2^n}} (S_{\frac{i+1}{2^n}} - S_{\frac{i}{2^n}})$$

converges⁶ in probability as $n \rightarrow \infty$. Then, the following holds:

⁵As mentioned in the introduction, there are continuous processes which are not semimartingales for which (10) holds for all integrands H of the type $H_t = f(t, S_t)$, where f is a bounded continuous function. On the other hand, every continuous deterministic S for which (10) holds for all integrands H of the type $H_t = f(S_t)$ (with f bounded continuous) is a function of finite variation (see [Str81, Prop. 4.1]).

⁶In fact, to obtain Corollary 6.2 it would be sufficient to require that the sequence in (10) is bounded in L^0 (by the same proof).

Corollary 6.2. *Every Riemann integrator is a semimartingale.*

To prove Corollary 6.2 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 $\mathcal{I}_S^n : X \rightarrow L^0(\mathbb{P})$ by

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

By definition, S is a Riemann integrator if, for every $K \in X$, $\mathcal{I}_S^n(K)$ converges in probability as $n \rightarrow \infty$. The Banach-Steinhaus theorem⁷ [Rud91, Theorem 2.6] then yields the following:

Lemma 6.3. *If S is a Riemann good integrator, then for every $\varepsilon > 0$ there is some $C > 0$ such that $\mathbb{P}(\mathcal{I}_S^n(K) \geq C) \leq \varepsilon$ for all $n \geq 1$ and all continuous adapted processes K such that $\|K\|_\infty \leq 1$.*

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

Proof of Corollary 6.2. 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 $K^{D_n} = H$. Since n was arbitrary and $\mathcal{I}_S(H) = \mathcal{I}_S(K^{D_n}) = \mathcal{I}_S^n(K)$, \mathcal{I}_S is bounded on $\mathcal{E}_D = \bigcup_{n \geq 1} \mathcal{E}_{D_n}$ by Lemma 6.3. Then Lemma 6.1 shows that S is a good integrator, and so Theorem BD implies that S is a semimartingale. \square

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. Veliev. A direct proof of the Bichteler-Dellacherie theorem and connections to arbitrage. *Ann. Probab.*, 2011. to appear.
- [BSV11b] M. Beiglböck, W. Schachermayer, and B. Veliev. A short proof of the Doob-Meyer theorem. *Stochastic Process. Appl.*, 2011. to appear.
- [DM82] C. Dellacherie and P.-A. Meyer. *Probabilities and potential. B*, volume 72 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1982.
- [Edw90] D. A. Edwards. A note on stochastic integrators. *Math. Proc. Cambridge Philos. Soc.*, 107(2):395–400, 1990.

⁷Which is also commonly called “the uniform boundedness principle”.

- [Fis65] D. 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] I. Karatzas and S. E. E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.
- [JS03] Jean Jacod and Albert N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2003.
- [Low11] G. Lowther. Almost sure. <http://almostsure.wordpress.com>, 2011.
- [Mét77] Michel Métivier. *Reelle und vektorwertige Quasimartingale und die Theorie der stochastischen Integration*, volume 607. Springer, 1977.
- [Mét82] Michel Métivier. *Semimartingales: a course on stochastic processes*, volume 2. Walter de Gruyter, 1982.
- [Mey79] Paul-André Meyer. Characterisation des semimartingales, d’après Dellacherie. In *Séminaire de Probabilités XIII*, pages 620–623. Springer, 1979.
- [Mey81] P Meyer. Une question de théorie des processus. *Séminaire de Probabilités XV 1979/80*, pages 142–142, 1981.
- [MP77] M Metivier and J Pellaumail. Mesures stochastiques à valeurs dans des espaces l 0. *Probability Theory and Related Fields*, 40(2):101–114, 1977.
- [MP80] Michel Métivier and Jean Pellaumail. *Stochastic integration*, volume 168. Academic Press New York, 1980.
- [Ore67] S. 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.
- [Pel73] Jean Pellaumail. *Sur l’intégrale stochastique et la décomposition de Doob-Meyer*. Société mathématique de France, 1973.
- [Pro05] P. E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, second edition, 2005.
- [Rao69] K. M. Rao. Quasi-martingales. *Math. Scand.*, 24:79–92, 1969.
- [Rud91] W. 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.
- [Str77] C. Stricker. Quasimartingales, martingales locales, semimartingales et filtration naturelle. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 39(1):55–63, 1977.
- [Str81] Christophe Stricker. Quasimartingales et variations. *Séminaire de Probabilités XV 1979/80*, pages 493–498, 1981.
- [Yan80] J.A. Yan. Caractérisation d’ensembles convexes de L^1 ou H^1 . volume 784 of *Lecture Notes in Math.*, pages 220–222. Springer, 1980.