On certain almost Brownian filtrations

M. Émery

IRMA, 7, rue René Descartes, 67 084 Strasbourg cedex, France

Received 11 September 2003; accepted 13 January 2005

Abstract

A consequence of Vershik’s results on discrete-time filtrations is the existence, in continuous time, of filtrations $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$ which are “Brownian after zero” (that is, for each $\varepsilon > 0$, $\mathcal{F}^\varepsilon = (\mathcal{F}_{\varepsilon + t})_{t \geq 0}$ is generated by $\mathcal{F}_\varepsilon$ and some $\mathcal{F}^\varepsilon$-Brownian motion), but not generated by $\mathcal{F}_0$ and any Brownian motion. Among the filtrations that are Brownian after zero, how are the truly Brownian ones characterized? An answer is given by the self-coupling criterion (ii) of Theorem 1. This criterion is always satisfied when $\mathcal{F}$ is immersible into the filtration of an infinite-dimensional Brownian motion.

© 2005 Elsevier SAS. All rights reserved.

Résumé

Une conséquence de la théorie des filtrations élaborée par Vershik il y a plus de trente ans est l’existence, en temps continu, de filtrations $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$ telles que, pour chaque $\varepsilon > 0$, $(\mathcal{F}_{\varepsilon + t})_{t \geq 0}$ soit engendrée par $\mathcal{F}_\varepsilon$ et par un mouvement brownien (nous les dirons « browniennes après zéro »), mais que $\mathcal{F}$ ne soit pas engendrée par $\mathcal{F}_0$ et un mouvement brownien. Comment caractériser les filtrations browniennes parmi les filtrations browniennes après zéro ? Le théorème 1(ii) répond à cette question par un critère d’auto-couplage ; ce critère est toujours satisfait lorsque $\mathcal{F}$ est immersible dans la filtration du mouvement Brownien à une infinité de dimensions.

© 2005 Elsevier SAS. All rights reserved.
have constructed has two further features: first, it is Brownian on the interval \([\varepsilon, \infty]\) for each \(\varepsilon > 0\) (that is, \(\mathcal{F}^\varepsilon = (\mathcal{F}_{t+\varepsilon})_{t \geq 0}\) is generated by \(\mathcal{F}_t\) and some \(\mathcal{F}^\varepsilon\)-Brownian motion); second, it is not immersible (this term will be defined soon) into “the” filtered probability space generated by an infinite-dimensional Brownian motion. The latter property looks stronger than mere non-Brownianness; but we shall see in Corollary 1 that, for a filtration which is Brownian on \([\varepsilon, \infty]\) for each \(\varepsilon > 0\), this property is in fact equivalent to being non-Brownian. In other words, consider a filtration \(\mathcal{F}\) which is Brownian on \([\varepsilon, \infty]\) for each \(\varepsilon > 0\); Corollary 1 says that \(\mathcal{F}\) is Brownian as soon as it is immersible into some infinite-dimensional Brownian filtration.

Most ideas and arguments are borrowed, some almost verbatim, from the theory of standard filtrations in discrete, negative time, mostly due to Vershik [14]. For instance, condition (iv) of Theorem 1 is a straightforward adaptation of his first-level standardness criterion (Condition 2 in Theorem 3.2 of [14]), and our proof of it just repeats his surprising argument. Similarly, the self-coupling condition (ii) of Theorem 1 is the continuous-time analogue of the discrete-time property called “I-cosiness” in [6], which is equivalent to Vershik’s standardness criterion [14].

But Vershik’s criterion, with its tower of measures on iterated measure spaces, does not lend itself to a restatement in continuous time; so the argument in [6], where Vershik’s criterion is used to show that I-cosiness entails standardness, does not adapt to our setting. Fortunately, another proof, bypassing Vershik’s criterion, has been given by S. Laurent [9]. Laurent’s method can be made to work in continuous time, but some adaptation is required: his approach relies on the explicit knowledge of the extremal points of some finite-dimensional set of joint laws; in continuous time, the set becomes infinite-dimensional, and has too many extremal points. This difficulty can be overcome by replacing the extremality argument with a density property (our Proposition 2).

This proposition says that, if \(\mathcal{F}\) is a filtration and if \(X\) and \(Y\) are two \(\mathcal{F}\)-Brownian motions, their joint law \(\mathcal{L}(X, Y)\) can be approximated by laws of the form \(\mathcal{L}(X', Y')\), where \(X'\) and \(Y'\) are two Brownian motions generating the same filtration. This was unexpected in the case when, for instance, \(Y = X\) on \([0, T]\) and \(dY = -dX\) after \(T\), where \(T\) is a stopping time independent of \(X\) and \(Y\).

Conventions, notation and definitions

The words ‘positive’ and ‘increasing’ are always understood in the broad sense: the null function is positive and increasing. An open interval is written \([s, t)\); similarly \([s, t]\) and \([s, t]\) denote half-open intervals.

By a probability space, we always mean a triple \((\Omega, \mathcal{A}, \mathbb{P})\) where the \(\sigma\)-field \(\mathcal{A}\) is \(\mathbb{P}\)-complete and essentially separable. By a sub-\(\sigma\)-field of \(\mathcal{A}\), we mean an \((\mathcal{A}, \mathbb{P})\)-complete sub-\(\sigma\)-field of \(\mathcal{A}\); it automatically inherits the property of being essentially separable. If \(\mathcal{C}\) is a class of events, or \(X\) a r.v. or a process, \(\sigma(\mathcal{C})\) or \(\sigma(X)\) denotes the \(\sigma\)-field generated by \(\mathcal{C}\) or \(X\), and by all negligible events of \(\mathcal{A}\); similarly, the expression ‘generated by’ means ‘generated by . . . and by the null events’. A random variable is a \(\mathbb{P}\)-equivalence class of measurable maps; \(L^0\) denotes the space of all a.s. finite r.v., endowed with the topology of convergence in probability; the elements of \(L^p\) are also defined up to \(\mathbb{P}\)-equivalence. If \(\mathcal{B}\) and \(\mathcal{C}\) are two sub-\(\sigma\)-fields of \(\mathcal{A}\), \(\mathcal{B} \vee \mathcal{C}\) denotes the \(\sigma\)-field \(\sigma(\mathcal{B} \cup \mathcal{C})\).

An isomorphism between two probability spaces \((\Omega', \mathcal{A}', \mathbb{P}')\) and \((\Omega'', \mathcal{A}'', \mathbb{P}'')\) is a bijection between the quotient \(\sigma\)-fields \(\mathcal{A}'/\mathbb{P}'\) and \(\mathcal{A''}/\mathbb{P}''\) that preserves the \(\sigma\)-field structures and the probabilities; it extends (uniquely) to random variables.

A filtration on \((\Omega, \mathcal{A}, \mathbb{P})\) is an increasing, right-continuous family \(\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}\) of sub-\(\sigma\)-fields of \(\mathcal{A}\); a filtered probability space is a system \((\Omega, \mathcal{A}, \mathbb{P}, \mathcal{F})\) where \(\mathcal{F}\) is a filtration. If \(\mathcal{F}\) is a filtration, \(\mathcal{F}_\infty\) denotes the \(\sigma\)-field \(\bigvee \mathcal{F}_t\). Given two filtrations \(\mathcal{F}\) and \(\mathcal{G}\) on \((\Omega, \mathcal{A}, \mathbb{P})\), we say that \(\mathcal{F}\) is included in \(\mathcal{G}\) if \(\mathcal{F}_t \subset \mathcal{G}_t\) for each \(t\); and we say that \(\mathcal{F}\) is immersed in \(\mathcal{G}\) if furthermore every \(\mathcal{F}\)-martingale is a \(\mathcal{G}\)-martingale. We denote by \(\mathcal{F} \vee \mathcal{G}\) the smallest filtration containing \(\mathcal{F}\) and \(\mathcal{G}\); it is given by \((\mathcal{F} \vee \mathcal{G})_t = \bigcap_{\varepsilon > 0} (\mathcal{F}_{t+\varepsilon} \vee \mathcal{G}_{t+\varepsilon})\). We say that \(\mathcal{F}\) and \(\mathcal{G}\) are jointly immersed if there exists a filtration \(\mathcal{H}\) on the same probability space such that each of \(\mathcal{F}\) and \(\mathcal{G}\) is immersed in \(\mathcal{H}\). In that case, one can always choose \(\mathcal{H} = \mathcal{F} \vee \mathcal{G}\). [Proof: each \(\mathcal{F}\)-martingale is an \(\mathcal{H}\)-martingale adapted to the smaller filtration \(\mathcal{F} \vee \mathcal{G}\), hence also an \(\mathcal{F} \vee \mathcal{G}\)-martingale; similarly for \(\mathcal{G}\)-martingales.]
Given two filtered probability spaces \((\Omega', \mathcal{A}', \mathbb{P}', \mathcal{F}')\) and \((\Omega'', \mathcal{A}'', \mathbb{P}'', \mathcal{F}'')\), we say that \(\mathcal{F}'\) and \(\mathcal{F}''\) are isomorphic if there exists an isomorphism \(\Psi\) between the probability spaces \((\Omega', \mathcal{F}'', \mathbb{P}')\) and \((\Omega'', \mathcal{F}', \mathbb{P}'')\) such that, for each \(t\), \(\mathcal{F}'_t\) and \(\mathcal{F}''_t\) are in correspondence by \(\Psi\). And we say that \(\mathcal{F}'\) is isomorphic to a filtration immersed in \(\mathcal{F}''\).

By convention, a Brownian motion (or, shortly, a BM) is always started at the origin. It may be one-dimensional, or \(d\)-dimensional, where \(d \leq \infty\); the case \(d = \infty\) means that the process consists of countably many independent components, each of them a real BM. We shall abbreviate ‘\(d\)-dimensional BM’ by BM\(^d\). We call Brownian (more precisely: \(d\)-Brownian) any filtration generated by some BM (more precisely: by some BM\(^d\)). An \(\mathcal{F}\)-BM (or \(\mathcal{F}\)-BM\(^d\)) is a \((d\)-dimensional\) Brownian motion for the filtration \(\mathcal{F}\): its increment on \([s, t]\) is independent of \(\mathcal{F}_s\) (equivalently, it is a BM and an \(\mathcal{F}\)-martingale; or it is a BM whose natural filtration is immersed in \(\mathcal{F}\)).

A non-Brownian example

To make the kind of situation we have in mind less abstract, we start with an example. It will not be used nor referred to in the sequel; skipping it is harmless.

**Example 1.** Suppose given a one-dimensional Brownian motion \(B\), a finite set \(A\) (called the alphabet) with at least two elements, and a sequence \(W = (W_n)_{n \geq 0}\) where, for each \(n\), \(W_n\) is a random word of length \(2^n\), that is, a random element of \(A^{2^n}\). Each word \(W_n\) is uniformly distributed on \(A^{2^n}\), and is independent of the whole process \(B\); but \(W_n\) is far from independent of \(B\), for we suppose \(W_n\) to be the first (resp. second) half of the twice longer word \(W_{n+1}\) iff \(B_{2^{-n}}\) is larger (resp. smaller) than \(B_{2^{-(n+1)}}\). It is easy to show the existence of such a pair \((B, W)\): first, taking \(W_n\) independent of \(B\), construct the law of \((B, W_n, W_{n-1}, \ldots, W_0)\), and verify that \(W_{n+1}\) too is uniform and independent of \(B\); then, take a projective limit.

From \(B\) and \(W\), define a filtration \(\mathcal{F}\) by

\[
\mathcal{F}_t = \sigma(B_s, s \in [0, t]) \vee \sigma(W_n, 2^{-n} \leq t).
\]

It is possible to show that \(\mathcal{F}\) is indeed a (right-continuous) filtration, enjoying the following three properties:

1. for \(0 < s \leq t\), \(\mathcal{F}_t = \mathcal{F}_s \vee \sigma(B_u-B_s, u \in [s, t])\).
2. \(\mathcal{F}_0 = (\mathcal{F}_{0+})\) is degenerate: every \(\mathcal{F}_{0+}\)-measurable r.v. is a.s. constant;
3. \(\mathcal{F}\) is not Brownian.

The first property is very easy: at time \(2^{-n}\) the new word \(W_n\) is observed; it is a function of the previously observed word \(W_{n+1}\) and of the increment of \(B\) between \(2^{-(n+1)}\) and \(2^{-n}\), so the increments of \(B\) suffice to generate all information necessary to increment \(\mathcal{F}\). This property is not valid for \(s = 0\), for there is more information in \(\mathcal{F}\) than in \(B\) only, each \(W_n\) being independent of \(B\).

The second property, right-continuity at 0, is less straightforward, but not difficult. The third one is much deeper; it can be deduced from Smorodinsky’s study [13] of the process \(W\) (in inverse, discrete time), or, slightly less directly, from Vershik’s Example 3 in his theory of reversed, discrete-time filtrations [14]. We shall not attempt to give, or even sketch, the proofs of these properties. This example exhibits a pathological behaviour at time 0+, the non-Brownianness of \(\mathcal{F}\) being a germ property at time 0+, even though \(\mathcal{F}_{0+}\) is degenerate.

**Definition.** Fix \(1 \leq d \leq \infty\). A filtration \(\mathcal{F}\) is \(d\)-Brownian after zero if there exists an \(\mathcal{F}\)-BM\(^d\) \(B\) such that, for all \(t \geq s > 0\), \(\mathcal{F}_t\) is generated by \(\mathcal{F}_s\) and by the process \((B_u-B_s, u \in [s, t])\), which is independent of \(\mathcal{F}_s\).
Lemma 1. Let $\mathcal{F}$ be a filtration; call $\mathcal{F}^s = (\mathcal{F}_{s+1})_{s \geq 0}$ the filtration $\mathcal{F}$ shifted by $s$. Suppose that for each $s > 0$, there exists an $\mathcal{F}^s$-BM $B^s$ such that the filtration $\mathcal{F}^s$ is generated by its initial $\sigma$-field $\mathcal{F}_s$ and by the process $B^s$. Then $\mathcal{F}$ is $d$-Brownian after zero.

Proof. Partition $[0, \infty[$ into intervals $I_n = [2^n, 2^{n+1}]$ where $n$ ranges over $\mathbb{Z}$, and define an $\mathcal{F}$-BM $B$ by $B_t = B_{m \alpha^n}$ for $t \in I_n$ (the increments of $B$ on $I_n$ are given by the increments of $B^2$ on $[0, 2^n]$). By induction on $n$, the property by $\mathcal{F}_t = \mathcal{F}_t \lor \sigma(B_u, u \in [0, t])$, holds when $s = 2^n$ and $t = 2^n$ with $m \leq n$; it then extends easily to the general case when $s$ and $t$ are real numbers verifying $0 < s \leq t$. □

The (admittedly artificial) example of this section shows that, for a filtration $\mathcal{F}$ which is Brownian after zero, the degeneracy of $\mathcal{F}_0$ is not sufficient to imply Brownianess. The next theorem bridges this gap and gives a necessary and sufficient condition for Brownianess when the filtration is a priori known to be Brownian after zero.

Main results

Theorem 1. Fix $d \in \mathbb{N}^* \cup \{\infty\}$ and a filtered probability space $(\Omega, \mathcal{A}, \mathbb{P}, \mathcal{F})$; suppose that $\mathcal{F}$ is $d$-Brownian after zero. The following two statements are equivalent:

(i) $\mathcal{F}$ is $d$-Brownian.
(ii) (Self-coupling condition.) For each $R \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ and each $\delta > 0$, there exists a probability space $(\Omega', \mathcal{A}', \mathbb{P})$ endowed with two filtrations $\mathcal{F}'$ and $\mathcal{F}''$ verifying the following four conditions:

(a) $\mathcal{F}'$ and $\mathcal{F}''$ are isomorphic to $\mathcal{F}$; in particular, there are two $\nu$. $R' \in L^1(\mathcal{F}_\infty)$ and $R'' \in L^1(\mathcal{F}'_\infty)$ corresponding to $R$ by these isomorphisms;
(b) $\mathcal{F}'$ and $\mathcal{F}''$ are jointly immersed;
(c) for some $s > 0$, $\mathcal{F}'_s$ and $\mathcal{F}''_s$ are $\mathbb{P}$-independent;
(d) $\|R' - R''\|_{L^1(\Omega)} < \delta$.

At first reading, condition (ii) looks awful, with the appearance of another filtered probability space $(\Omega', \mathcal{A}', \mathbb{P})$ where two filtrations are jointly immersed. One way of understanding it is to consider $(\Omega, \mathcal{A}, \mathbb{P})$ as an “enlargement” of $(\Omega', \mathcal{A}', \mathbb{P})$, in the sense of the sentence ‘at the cost of enlarging the space, we may suppose that…’ The new filtration $\mathcal{F}'$ isomorphic to $\mathcal{F}$ should then be understood as being $\mathcal{F}$ itself, and the meaning of (ii) becomes: “After enlarging $\Omega$ if necessary, there exists another filtration $\mathcal{F}''$, jointly immersed with $\mathcal{F}$, such that…”

Condition (ii) is the analogue, in our continuous-time setting, of the property called ‘$I$-cosiness’ in [6]. It is dubbed ‘self-coupling’ because, if $\mathcal{F}$ is generated by some process $X$, (ii) means that it is possible, in some universe $\Omega'$, to run simultaneously two copies of $X$, in such a way that they are independent up to some time $s > 0$, but, at time infinity, the two values taken by some given functional $R$ of $X$ have become close to each other. This idea should be compared with the classical coupling method used to establish estimates for Markov processes.

The proof of Theorem 1 is rather long; it will be given later. Meanwhile, we shall comment on it, state and prove a corollary, and establish Proposition 2, a crucial density property.

Remarks. (1) Let $D$ be a dense subset of $L^1(\Omega, \mathcal{A}, \mathbb{P})$. If condition (ii) is satisfied for each $R \in D$ and each $\delta > 0$, it holds in full generality. Indeed, for $R \in L^1$ and $\delta > 0$, there exists $S \in D$ such that $\|R - S\| < \delta$; (ii) applied to $S$ gives $\|R' - S''\| < \delta$, and by isomorphism $\|R' - S''\| = \|R' - S''\| = \|R - S\| < \delta$, whence $\|R' - R''\| < 3\delta$, and (ii) holds for $R$ too.

In particular, taking for $D$ the set of all simple, $\mathcal{F}_\infty$-measurable random variables, it suffices to verify (ii) for $R$ taking finitely many values. (I do not know if it suffices to verify it for $R$ taking two values, that is, for indicators of events; already in Vershik’s theory of filtrations in discrete, negative time, the corresponding question is open.)
(2) In condition (ii), the space $L^1$ and its norm can equivalently be replaced by $L^p$ for any $p \in [1, \infty[$, or by $L^0$, with the distance $d_0(X,Y) = \mathbb{E}[1 \wedge |X - Y|]$ for instance. Indeed, as shown by the preceding remark, we may suppose $|R|$ to be bounded by some number $k \geq 1$; so $|R' - R''| \leq 2k$. It then suffices to use the elementary inequality $\mathbb{E}[|R' - R''|^p] \leq (2k)^p \mathbb{E}[1 \wedge |R' - R''|]$. 

(3) (Remark due to S. Laurent [9].) If condition (ii) holds for some $R$ and all $\delta$, it holds also for all $R_1 \in L^1(\sigma(R))$ and all $\delta$; consequently, it suffices to verify it for one $R$ that generates $\mathcal{F}_\infty$. To see this, suppose (ii) to hold for $R$ and consider the set $\Phi$ of all bounded, Borel functions, for, if $f$ is $k$-Lipschitz, $\|f \circ R' - f \circ R''\|_{L^1} \leq k\|R' - R''\|_{L^1}$. Second, the set $\Phi$ is closed under uniformly bounded pointwise limits: if $f_n \in \Phi$ are uniformly bounded and $f_n(x) \to f(x)$ for all $x \in \mathbb{R}$, then $f \in \Phi$. To check this, fix $\delta > 0$. The quantity $\mathbb{E}[|f_n \circ R' - f \circ R'|]$ tends to 0 when $n \to \infty$, so it is smaller than $\delta$ for some $n$ (fixed in the sequel). Since (ii) holds for $f_n \circ R$ and $\delta$, there are $\mathcal{F}'$ and $\mathcal{F}''$ verifying (ii-a), (ii-b), (ii-c) and $\mathbb{E}[|f_n \circ R' - f \circ R'|] < \delta$. By isomorphism, one has also $\mathbb{E}[|f_n \circ R' - f \circ R'|] = \mathbb{E}[|f_n \circ R'' - f \circ R''] < \delta$; so $\mathbb{E}[|f \circ R' - f \circ R''|] < 3\delta$, yielding $f \in \Phi$. These two properties of $\Phi$ entail that $\Phi$ contains all bounded, Borel functions.

For $d \leq \infty$, call $\mathcal{B}^d$ “the” filtration of $d$-dimensional Brownian motion (it is unique up to isomorphism, whence the quotation marks). The next statement, a corollary of Theorem 1, says that the filtration $\mathcal{B}^\infty$ plays in our continuous-time framework the same rôle as the standard, non-atomic, discrete-time filtration in Vershik’s theory.

**Corollary 1.** Fix $d \leq \infty$. Assume that a filtration $\mathcal{F}$ is $d$-Brownian after zero. The following three statements are equivalent:

(i) $\mathcal{F}$ is $d$-Brownian;

(ii) the independent product $\mathcal{F} \otimes \mathcal{B}^\infty$ is equal to $\mathcal{B}^\infty$;

(iii) $\mathcal{F}$ is immersible into $\mathcal{B}^\infty$.

The independent product $\mathcal{F} \otimes \mathcal{B}^\infty$ in (ii) is the filtration (more precisely: the filtered probability space) generated by two independent filtrations, respectively isomorphic to $\mathcal{F}$ and to $\mathcal{B}^\infty$; this product is well defined up to isomorphism only, so ‘equal to’ in (ii) really means ‘isomorphic to’.

**Proof (the theorem is admitted).** (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) are trivial. To prove (iii) $\Rightarrow$ (i), applying Theorem 1 to $\mathcal{B}^\infty$, one sees that $\mathcal{B}^\infty$ satisfies the self-coupling property. By immersion, every subfiltration immersed in $\mathcal{B}^\infty$ also satisfies the self-coupling property; this transfers to $\mathcal{F}$ by isomorphism. Applying now the theorem to $\mathcal{F}$, one sees that $\mathcal{F}$ is $d$-Brownian. $\square$

The interest of the theorem and its corollary is more theoretical than practical. In the special case that the filtration is Brownian after zero, they answer the following general question (see D. Revuz and M. Yor [12, p. 219]): if a filtration has the predictable representation property w.r.t. some BM$^d$, which additional assumption is sufficient to imply that it is $d$-Brownian? In view of Corollary 1, this question can be sharpened: If a filtration has the predictable representation property w.r.t. some BM$^d$ and is immersible into $\mathcal{B}^\infty$, is it necessarily $d$-Brownian?

As observed in the introduction, the corollary also explains why the non-Brownian counterexamples constructed in [4] and [5] have the stronger property of not being immersible into $\mathcal{B}^\infty$.

**Brownian examples**

Putting the corollary at work to establish that a given particular filtration is Brownian, is disappointing: in all instances I know, exhibiting a generating BM is possible, and more informative than merely asserting its existence.
Nevertheless, in some cases, referring to the corollary is shorter and simpler than the constructive proof. Here are two examples; in both of them, a constructive proof of Brownianness is already available.

**Example 2 (Generalized Goswami–Rao filtrations).** Fix an integer \( d < \infty \) and a subgroup \( G \) of \( \text{O}(d) \). Let \((W^d, \mathcal{B}^d, \lambda^d, \mathcal{F}^d)\) denote the canonical filtered probability space of BM\(^d\), and \( B \) the canonical process on \( W^d \). Each orthogonal transformation \( g \in G \) acts on the paths \( \omega \in W^d \) by \((g \cdot \omega)(t) = g \cdot (\omega(t))\); this action is an isomorphism from \((W^d, \mathcal{B}^d, \lambda^d, \mathcal{F}^d)\) to itself. A r.v. \( Y \) is said to be invariant if, for each \( g \in G \), one has \( Y = Y \circ g \) a.s.; call \( I \) the sub-\( \sigma \)-field of invariant events. As the action of \( G \) commutes with conditional expectation w.r.t. \( \mathcal{B}^d \), the subfiltration \( \mathcal{F} \) of \( \mathcal{B}^d \) defined by \( \mathcal{F}_t = \mathcal{B}^d \cap I \) is immersed in \( \mathcal{B}^d \).

Malric has shown in [10] and [11] that the filtration \( \mathcal{F} \) is \( d \)-Brownian. Since \( \mathcal{F} \) is immersed in \( \mathcal{B}^d \), this can also be derived from Corollary 1, provided we show \( \mathcal{F} \) to be \( d \)-Brownian after zero. As we shall now see, in the particular case when \( G \) is finite, this is straightforward. (But Malric’s proof works for all \( G \), finite or not.)

We suppose from now on \( G \) to be finite. There exists a Borel set \( A \subset \mathbb{R}^d \) such that \( A \cap g \cdot A = \emptyset \) for each \( g \in G \setminus \{1\} \), and that \( G \cdot A = \bigcup_{g \in G} g \cdot A \) has a Lebesgue-negligible complementary set. (One can for instance choose a point \( z \in \mathbb{R}^d \) such that \( g \cdot z \neq z \) for all \( g \in G \setminus \{1\} \), and take \( A = \{y \in \mathbb{R}^d : \forall g \in G \setminus \{1\}, \|y - z\| < \|y - g \cdot z\|\} \).) The action of \( G \) on \( A \) is faithful, so for each \( x \in G \cdot A \) there exists a unique \( \gamma^x \in G \) such that \( \gamma^x \cdot x \in A \); moreover, \( x \mapsto \gamma^x \) is a Borel map on \( G \cdot A \). Clearly, \( \gamma^{gx} = \gamma^x \gamma^{-1} \) since each of them maps the point \( g \cdot x \) to an element of \( A \).

For fixed \( s > 0 \), \( B_s \) is a.s. in \( G \cdot A \), so \( \gamma^s B_s \) is a.s. well defined. The process \( \beta_t = \gamma^{B_t} \cdot (B_{s+t} - B_s) \) is a BM\(^d\) for the filtration \((\mathcal{B}^d_{s+}, t \geq 0)\). It is also invariant, for

\[
\beta_t \circ g = \gamma^{g^{-1} B_t} \cdot (g \cdot B_{s+t} - g \cdot B_s) = \gamma^{B_t} g^{-1} g \cdot (B_{s+t} - B_s) = \beta_t.
\]

Hence it is adapted to the smaller filtration \( \mathcal{F}^s \), and it is an \( \mathcal{F}^s \)-BM\(^d\). Using Lemma 1, to establish that \( \mathcal{F} \) is Brownian after zero, it remains to see that \( \mathcal{F}_s \) and \( \beta \) generate \( \mathcal{F}^s \). Putting \( V = (B_{s+} - B_s, v \in [0, t]) \) and \( \tilde{V} = (\beta_v, v \in [0, t]) \), we have to show that, for \( Y = \mathcal{E}_t \), the conditional expectation \( \mathbb{E}[Y | I] \) is measurable for \( \mathcal{F}_s \vee \sigma(\tilde{V}) \). We may take \( Y \) of the form \( U \circ \phi(V) \), where \( U \) is \( \mathcal{B}^d \)-measurable; and it then suffices to write

\[
\mathbb{E}[Y | I] = \frac{1}{|G|} \sum_{g \in G} Y \circ g = \frac{1}{|G|} \sum_{g \in G} (U \circ g) \phi(V \circ g)
\]

\[
= \frac{1}{|G|} \sum_{g \in G} (U \circ g) \gamma^B \phi(V \circ g) = \frac{1}{|G|} \sum_{g \in G} (U \circ g) \gamma^{B_s} \phi(V \circ g)
\]

\[
= \frac{1}{|G|} \sum_{g \in G} (U \circ g) \gamma^{B_s} \phi(\tilde{V}) = \phi(\tilde{V}) \frac{1}{|G|} \sum_{g \in G} (U \circ g) = \phi(\tilde{V}) \mathbb{E}[U | I].
\]

**Example 3 (Stationary Brownian motion on a sphere).** Let \( M \) be a compact, connected, \( d \)-dimensional Riemannian manifold without boundary, and call \((X_t)_{t \in \mathbb{R}}\) the stationary Brownian motion with values in \( M \); for each \( t \in \mathbb{R} \), \( X_t \) is distributed on \( M \) according to the Riemannian measure. Arnaudon [1] has established that the filtration \((\mathcal{F}_t)_{t \in \mathbb{R}}\) generated by \( X \) is \( d \)-Brownian, that is, generated by an \( \mathbb{R}^d \)-valued Ornstein–Uhlenbeck process indexed by \( \mathbb{R} \) (or, equivalently, by a BM\(^d\) which is forced to jump to 0 at each integer instant \( t \in \mathbb{Z} \); equivalently again, the logarithmically time-changed filtration \((\mathcal{F}_{mt})_{t \geq 0}\) is \( d \)-Brownian). It may be interesting to observe here that Arnaudon’s proof relies on a coupling argument. This situation, and other stationary processes indexed by \( \mathbb{R} \), is a typical instance where Theorem 1 or Corollary 1, or their analogues obtained by a logarithmic time-change, can be expected to enter the picture, because the hypothesis that everything goes well on each interval \([s, \infty[ \) is clearly satisfied; the problem is at time \(-\infty \) only.
Corollary 1 gives a new proof of Arnaudon’s result, provided it is established that the process $X$ is immersed in some, possibly higher-dimensional, Brownian motion. As we shall now see, this immersion is quite simply obtained in the particular case when $M$ is a $d$-dimensional sphere.

So we are working in eternal time ($t \in \mathbb{R}$), and $X$ takes values in the unit sphere $S \subset \mathbb{R}^{d+1}$. Denote by $dB_t = (dB_t^1, \ldots, dB_t^{d+1})$ a $(d + 1)$-dimensional Brownian innovation, and consider the SDE on the sphere

$$dX_t = \text{proj}_{X_t^d}(dB_t) - \frac{d}{2}X_t \, dt.$$  

The drift term $-\frac{d}{2}X_t \, dt$ compensates for the extrinsic curvature of the sphere, so $X$ remains on $S$ and is a Brownian motion on $S$. This SDE generates a stochastic flow of diffeomorphisms $\Phi_{st}$ on $S$. A consequence of Corollary 1 gives a new proof of Arnaudon’s result, provided it is established that the process $(\Phi_{st}(x), \Phi_{st}(y))$ has only one characteristic exponent, namely $-d/2$. Since it is strictly negative, one has

$$\sup_{x,y \in S} \mathbb{E}[\text{dist}(\Phi_{st}(x), \Phi_{st}(y))] \to 0 \quad \text{when} \quad t - s \to \infty. \quad (*)$$

(This property can also be obtained directly, by studying the diffusion $\text{dist}(\Phi_{st}(x), \Phi_{st}(y))$, which is the solution of some one-dimensional SDE.) A consequence of $(*)$ is that, for fixed $x \in S$,

$$\mathbb{E}[\text{dist}(\Phi_{st}(x), \Phi_{st}(y))] = \mathbb{E}[\text{dist}(\Phi_{rt}(x), \Phi_{st}(x))] \to 0$$

when $r$ and $s$ tend to $-\infty$ with $r \leq s$. By Cauchy’s criterion, for fixed $x \in S$, the limit $\overline{X}_t = \lim_{s \to -\infty} \Phi_{st}(x)$ exists in probability; the flow property then shows that $\overline{X}_t$ is a solution to the SDE, hence a Brownian motion on $S$. As $\overline{X}_t$ is also measurable w.r.t. the past innovation $\mathbb{F}^{d+1} = \sigma(B_t - B_s, \ s \in ]-\infty, t])$, this realizes an immersion of $X$ in the $(d+1)$-dimensional Brownian filtration $\mathbb{B}^{d+1}$, and Corollary 1(iii) applies.

More generally, the same argument works for any embedded compact manifold such that all characteristic exponents of the gradient Brownian flow (or, for that matter, of some Brownian flow) are strictly negative.

We now start proving Theorem 1. We begin with two sections devoted to establishing technical results that will be useful later.

**Preliminaries:** (1) **Substantial families of $\sigma$-fields**

If $(\Omega, \mathcal{A}, \mathbb{P})$ is a probability space and $F$ a finite set, we denote by $L(\mathcal{A}; F)$ the set of all $F$-valued, $\mathcal{A}$-measurable random variables; $L(\mathcal{A}; F)$ is a metric space when endowed with the distance $(R, S) \mapsto \mathbb{P}[R \neq S]$.

If $(E, d)$ is a separable metric space, $L^1(\mathcal{A}; (E, d))$ (or $L^1(\mathcal{A}; (E, d))$ for short) denotes the set of all $E$-valued, $\mathcal{A}$-measurable random variables $R$ such that $\mathbb{E}[d(R, x)]$ is finite for some ($\Leftrightarrow$ for all) $x \in E$. It is endowed with the distance $(R, S) \mapsto \mathbb{E}[d(R, S)]$. It is well known that the set $L^1(\mathcal{A}; E)$ of all simple, $E$-valued, $\mathcal{A}$-measurable r.v.’s, is a dense subset of $L^1(\mathcal{A}; (E, d))$.

[Proof: Choose a numbering $(x_n)_{n \geq 0}$ of a countable dense subset $D$ of $E$; define $\psi_n : D \to D$ by $\psi_n(x_m) = x_m$ if $m \leq n$ and $\psi_n(x_m) = x_0$ if $m > n$. For $R \in L^1(\mathcal{A}; (D, d))$, $\mathbb{E}[d(\psi_n \circ R, R)]$ tends to 0 by dominated convergence; so $L^1(\mathcal{A}; (D, d))$ is dense in $L^1(\mathcal{A}; (D, d))$. To see that the latter is dense in $L^1(\mathcal{A}; (E, d))$, define $\phi_{\delta} : E \to D$ by $\phi_{\delta}(x) = \text{first } x_n$ such that $d(x_n, x) < \delta$; $\phi_{\delta}$ is measurable and verifies $d(x, \phi_{\delta}(x)) < \delta$, so $\mathbb{E}[d(R, \phi_{\delta}(R))] < \delta.$]

**Definition.** Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $B$ a set of sub-$\sigma$-fields of $\mathcal{A}$. We shall say that $B$ is **substantial in $\mathcal{A}$** if $\bigcup_{B \in B} L^1(\Omega, B, \mathbb{P})$ is dense in $L^1(\Omega, \mathcal{A}, \mathbb{P})$.

For instance, if $(\mathcal{A}_n)_{n \in \mathbb{N}}$ is an increasing sequence of sub-$\sigma$-fields of $\mathcal{A}$, the set $\{\mathcal{A}_n, \ n \in \mathbb{N}\}$ is substantial in $\bigcap_{n \in \mathbb{N}} \mathcal{A}_n$.

\[1\] This argument is due to S. Watanabe; I thank him for allowing me to publish it here.
Observe that if $B$ is substantial in $A$, the class $\bigcup_{\mathcal{B} \in \mathcal{B}} \mathcal{B}$ generates the $\sigma$-field $A$; but the converse is false: for instance, the set $B = \{\sigma(A), A \in A\}$ is in general not substantial in $A$.

**Lemma 2.** Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $B$ a set of sub-$\sigma$-fields of $A$. The following three conditions are equivalent:

(i) $B$ is substantial in $A$;

(ii) for each finite set $F$, $\bigcup_{\mathcal{B} \in \mathcal{B}} \mathcal{B}(F; B)$ is dense in $L(A; F)$;

(iii) for each separable metric space $(E, d)$, $\bigcup_{\mathcal{B} \in \mathcal{B}} \mathcal{B}(E; d)$ is dense in $L^1(A; (E, d))$.

**Proof.** (iii) $\Rightarrow$ (i) is trivial: take $E = \mathbb{R}$.

(i) $\Rightarrow$ (ii). Fix $F$ finite, $R \in L(A; F)$ and $\delta > 0$. Let $\phi : F \to \mathbb{R}$ be a map such that for all $e$ and $f$ in $F$, $e \neq f$ there is $|\phi(e) - \phi(f)| \geq \delta$. Applying hypothesis (i) to $\phi R$, one obtains a $B \in \mathcal{B}$ and a $T \in L^1(\mathcal{B}; \mathbb{R})$ such that $E[|T - \phi R|] < \delta$; so $\mathbb{P}[|T - \phi R| \geq \delta] < \delta$. Define $\psi : \mathbb{R} \to \phi(F)$ by $\psi(x)$ is the point in $\phi(F)$ closest to $x$ (the leftmost such point if there are two). On $|T - \phi R| < 1$, one has $\psi R = \phi R$ by the choice of $\phi$; hence

$$\mathbb{P}[\psi R \neq \phi R] \leq \mathbb{P}[|T - \phi R| \geq \delta] < \delta.$$

It suffices to set $S = \phi^{-1}\circ \psi R$ to have $S \in L(\mathcal{B}; F)$ and $\mathbb{P}[S \neq R] = \mathbb{P}[\phi R \neq \phi R] < \delta$.

(ii) $\Rightarrow$ (iii). Fix $R \in L^1(\mathcal{A}; (E, d))$ and $\delta > 0$. As $L^1(\mathcal{A}; E)$ is dense in $L^1(\mathcal{A}; E)$, there exist some finite subset $F$ of $E$ and some $T \in L(\mathcal{A}; F)$ such that $\mathbb{E}[d(R, T)] < \delta/2$. Call $a$ the diameter of $F$. Hypothesis (ii) yields a $B \in \mathcal{B}$ and a $S \in L(\mathcal{B}; F)$ such that $\mathbb{P}[S \neq T] < \delta/2a$. Since $d(S, T) \leq a\mathbb{P}[S \neq T]$, one has $\mathbb{E}[d(S, T)] \leq a\mathbb{P}[S \neq T] < \delta/2$, and $\mathbb{E}[d(R, S)] \leq \mathbb{E}[d(R, T)] + \mathbb{E}[d(S, T)] < \delta$. □

Lemma 2 will be used only once, in the proof of Lemma 3; what we shall need is only (i) $\Rightarrow$ (iii), in the particular case when the separable metric space $E$ is equal to $\mathbb{R}^k$.

**Lemma 3.** Given $(\Omega, \mathcal{A}, \mathbb{P})$, let $\mathcal{B}$ and $\mathcal{C}$ be sub-$\sigma$-fields of $\mathcal{A}$, and $D$ a set of sub-$\sigma$-fields of $\mathcal{C}$, substantial in $\mathcal{C}$. The set $[\mathcal{B} \vee D, D \in D]$ is substantial in $\mathcal{B} \vee C$.

**Proof.** We have to show that $\bigcup_{D \in D} L^1(\mathcal{B} \vee D)$ is dense in $L^1(\mathcal{B} \vee C)$. But the random variables of the form $\sum \alpha_i I_{B_i C_i}$ (finite sum, $\alpha_i \in \mathbb{R}$, $B_i \in \mathcal{B}$, $C_i \in \mathcal{C}$) are dense in $L^1(\mathcal{B} \vee C)$. Fix such a r.v. $\alpha_1 I_{B_1 C_1} + \cdots + \alpha_k I_{B_k C_k}$. Applying Lemma 2 to the random variable $(\alpha_1 I_{C_1}, \cdots, \alpha_k I_{C_k})$ with values in $(E, d) = (\mathbb{R}^k, \sum |x_i - y_i|)$, one gets the existence of some $D \in D$ and $(D_1, \ldots, D_k) \in L^1(\mathcal{D}, E)$ such that $\mathbb{E}[\sum |\alpha_i I_{C_i} - D_i|] < \delta$, and a fortiori $\mathbb{E}[\sum |\alpha_i I_{B_i C_i} - I_{B_i D_i}|] < \delta$. As the r.v. $\sum I_{B_i D_i}$ belongs to $L^1(\mathcal{B} \vee D)$, the proof is over. □

**Lemma 4.** Fix $s \geq 0$. Let $\mathcal{F}$ be a filtration and $G$ a set such that each $\mathcal{G} \subset G$ is a filtration immersed in $\mathcal{F}$; suppose $[\mathcal{G}_\infty, \mathcal{G} \subset G]$ to be substantial in $\mathcal{F}_\infty$. Then $[\mathcal{G}_s, \mathcal{G} \subset G]$ is substantial in $\mathcal{F}_s$; moreover, if $\mathcal{C}$ is a $\sigma$-field such that $\mathcal{F}_s \subset \mathcal{C} = \mathcal{F}_\infty$, then $[\mathcal{G}_s \subset \mathcal{C}, \mathcal{G} \subset G]$ is substantial in $\mathcal{F}_\infty$.

**Proof.** Since the map $E_s : X \mapsto \mathbb{E}[X \mid \mathcal{F}_s]$ is a contraction from $L^1(\mathcal{F}_\infty)$ onto $L^1(\mathcal{F}_s)$, and since $\bigcup_{s \geq 0} L^1(\mathcal{G}_\infty)$ is dense in $L^1(\mathcal{F}_\infty)$, $\bigcup_{s \geq 0} E_s L^1(\mathcal{G}_\infty)$ is dense in $L^1(\mathcal{F}_s)$. The immersion hypothesis gives $\mathbb{E}_s L^1(\mathcal{G}_\infty) = L^1(\mathcal{G}_s)$; so $[\mathcal{G}_s, \mathcal{G} \subset G]$ is substantial in $\mathcal{F}_s$. The second part of the conclusion follows immediately from this via Lemma 3. □
Preliminaries: (2) Joint laws of Brownian motions

Throughout this section, the dimension $d$ is fixed and finite.

Call $\text{COV}(d)$ the set of all $d \times d$ real matrices $h$ such that, for all real vectors

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_d \end{pmatrix} \quad \text{and} \quad w = \begin{pmatrix} w_1 \\ \vdots \\ w_d \end{pmatrix},$$

one has the inequality

$$t^ivhw = \sum_{i,j} v_i h_{ij} w_j \leq \|v\|\|w\|$$

(where $t^iv$ denotes the transpose of $v$ and $\|v\|$ its Euclidean norm).

The set $\text{COV}(d)$ draws its name from ‘covariance’; for if $X$ and $Y$ are two random vectors in $\mathbb{R}^d$ with unit variance, their covariance $h_{ij} = \text{Cov}[X_i, Y_j]$ belongs to $\text{COV}(d)$; this stems directly from property (iv) in Proposition 1 below. But $\text{COV}(d)$ also means ‘covariation’, for if $X$ and $Y$ are two BM$^d$ for some filtration, the following lemma asserts that their covariation $(d/dt)\langle X_i, Y_j \rangle_t$ takes its values in $\text{COV}(d)$.

**Lemma 5.** If $X$ and $Y$ are two BM$^d$ for some filtration $\mathcal{F}$, there exists a predictable process $H$ with values in $\text{COV}(d)$ such that $(d/dt)\langle X_i, Y_j \rangle(t) = H_{ij}(t)dt$.

**Proof.** The Kunita–Watanabe inequality says that there exists a predictable process $H$, with values in $d\times d$ matrices, defined up to a $d\times d\mathbb{P}$ negligible set, such that $(d/dt)\langle X_i, Y_j \rangle(t) = H_{ij}(t)dt$; moreover, as

$$\langle v \cdot X, v \cdot X \rangle_t = \|v\|^2t, \quad \langle w \cdot Y, w \cdot Y \rangle_t = \|w\|^2t \quad \text{and} \quad \langle v \cdot X, w \cdot Y \rangle_t = \int_0^t \sum_{ij} v_i H_{ij}(s)w_j ds,$$

one has

$$\frac{1}{t-s} \int_s^t \sum_{ij} v_i H_{ij}(u)w_j du \leq \|v\|\|w\|$$

for all $s, t$, almost all $\omega$ and all rational vectors $v$ and $w$. Hence,

$$\frac{1}{t-s} \int_s^t H(u) du \in \text{COV}(d)$$

for all $s, t$ and almost all $\omega$. In the limit, $H_t(\omega)$ belongs to the closed set $\text{COV}(d)$ for almost all $(t, \omega)$. Replacing $H$ by 0 wherever necessary makes it possible to choose a version of $H$ which is identically $\text{COV}(d)$-valued.

**Proposition 1.** (1) For a $d \times d$ matrix $h$, the following are equivalent:

(i) the matrix $h$ belongs to $\text{COV}(d)$;
(ii) the transpose $t^ih$ belongs to $\text{COV}(d)$;
(iii) for all vectors $v$, $\|hv\| \leq \|v\|$;
(iv) the symmetric, $2d \times 2d$ matrix $\begin{pmatrix} I_h & h \\ t^ih & I \end{pmatrix}$ is positive;
(v) the symmetric, $d \times d$ matrix $I - t^ihh$ is positive.
(2) The set \( \text{COV}(d) \) is convex and compact. Its extremal points are the orthogonal matrices \( h \in O(d) \).

(3) Put \( d' = d^2 + 1 \). There exist \( d' \) measurable maps \( r_1, \ldots, r_{d'} \) from \( \text{COV}(d) \) to \( O(d) \) and \( d' \) measurable maps \( \alpha_1, \ldots, \alpha_{d'} \) from \( \text{COV}(d) \) to \([0, 1]\) such that, for each \( h \in \text{COV}(d) \)

\[
\sum_{i=1}^{d'} \alpha_i(h) r_i(h) = h.
\]

**Remark** (not used in the sequel). More generally, if \( G \) is any compact group of real \( d \times d \) matrices, the set of all extremal points of the convex hull of \( G \) is \( G \) itself. Indeed, \( G \) is a subgroup of \( O(d) \), and every point of \( G \) is extremal in \( \text{COV}(d) \) by (2), and a fortiori in the smaller set \( \text{conv} G \).

**Proof of Proposition 1.** (1) Equivalence between (i) and (ii) stems trivially from \( t'hw = t'(t'hw) = t'hw' \). Also, by homogeneity, (i) is equivalent to

\[
\forall w \quad \sup_{v: \|v\|=1} t'hw \leq \|w\|,
\]

or to (iii) since the left-hand side is \( \|hw\| \). In turn, (iii) amounts to \( t'hv \leq t'v \), which says that \( I - t'hh \geq 0 \), that is, (v). Last, (iv) means that \( t'hw = \frac{1}{2}(\|v\|^2 + \|w\|^2) \) for all \( v \) and \( w \). If \( h \) is in \( \text{COV}(d) \), this inequality holds because the right-hand side is minorated by \( \|v\| \|w\| \); conversely, if this inequality holds, one has \( t'hw = \|v\| \|w\| \) for all unit vectors, hence also for all vectors by homogeneity.

(2) Convexity, closedness and boundedness (each entry of \( \text{COV}(d) \) verifies \( |h_{ij}| \leq 1 \)) are obvious on the definition of \( \text{COV}(d) \). The inclusion \( O(d) \subset \text{COV}(d) \) stems for instance from (v).

It is not difficult to see that the extremal subset of \( \text{COV}(d) \) is equal to \( O(d) \); see [8].

(3) The last assertion of Proposition 1 comes from Carathéodory’s theorem (if \( K \) is a compact, convex subset of \( \mathbb{R}^d \), every point of \( K \) is a barycentre of at most \( n + 1 \) extremal points of \( K \), and the measurable section theorem. \( \square \)

Still for finite \( d \), denote by \( W^d = \{ w \in C(\mathbb{R}_+, \mathbb{R}^d): w(0) = 0 \} \) the canonical space of \( d \)-dimensional Brownian motion, and by \( \lambda^d \) the Wiener measure on \( W^d \).

We shall call \( \text{JIB}^d \) (for “jointly immersed Brownian motions”) the set of probability measures on the product \( W^d \times W^d \) defined as follows: a probability \( \mu \) on \( W^d \times W^d \) is an element of \( \text{JIB}^d \) if and only if there exist, on some filtered probability space \( (\Omega, \mathcal{G}, \mathbb{P}, \mathcal{F}) \), two \( \mathcal{F} \)-BM\(^d \) \( X \) and \( Y \) with joint law \( \mathcal{L}[X, Y] = \mu \). Observe that if \( \mu \in \text{JIB}^d \), both marginals of \( \mu \) are equal to \( \lambda^d \); but the converse does not hold. For instance, if \( X \) is a BM\(^d \) and if \( Y_t = X_{t+1} - X_t \), the law of \( (X, Y) \) does not belong to \( \text{JIB}^d \), because \( X_2 - X_1 \) is not independent of \( Y_1 \).

**Lemma 6.** The set \( \text{JIB}^d \) is convex; endowed with the topology of weak convergence (for the bounded, continuous functions on \( W^d \times W^d \), it is compact.

If \( X \) and \( Y \) are two BM\(^d \) for some filtration \( \mathcal{F} \), and if \( \mathcal{C} \) is a sub-\( \sigma \)-field of \( \mathcal{F}_0 \), the conditional joint law \( \mathcal{L}[X, Y | \mathcal{C}] \) a.s. belongs to \( \text{JIB}^d \).

If \( X \) and \( Y \) are two processes, and if \( \mathcal{F}_0 \) is a \( \sigma \)-field such that \( \mathcal{L}[X, Y | \mathcal{F}_0] \) is a.s. in \( \text{JIB}^d \), then \( X \) and \( Y \) are BM\(^d \) for the smallest filtration \( \mathcal{G} \) such that \( \mathcal{G}_0 \supseteq \mathcal{F}_0 \) and that \( X \) and \( Y \) are \( \mathcal{G} \)-adapted.

---

2 I thank an anonymous referee for this reference.
Proof. If $X$ and $Y$ are $\mathcal{F}$-BM$^d$, then, for all $k \geq 1$, all $0 < s_1 < \cdots < s_k < t < u$, and all $\rho, \sigma_1, \ldots, \sigma_k, \tau_1, \ldots, \tau_k$ in the dual of $\mathbb{R}^d$, 

$$E \left[ \left( e^{i \rho (X_u - X_t)} - e^{-\frac{1}{2} \|\rho\|^2 (u-t)} \right) \exp i \sum_{j=1}^k (\sigma_j X_{s_j} + \tau_j Y_{s_j}) \right] = 0;$$

so, putting

$$\chi(x, y) = \left( e^{i \rho (y - x)} - e^{-\frac{1}{2} \|\rho\|^2 (u-t)} \right) \exp i \sum_{j=1}^k (\sigma_j x_{s_j} + \tau_j y_{s_j}),$$

every $\mu \in \text{JIB}^d$ verifies $\mu(\chi) = 0$. Similarly, $\mu(\tilde{\chi}) = 0$, where $\tilde{\chi}(x, y) = \chi(y, x)$.

Conversely, if $\mu(\chi) = \mu(\tilde{\chi}) = 0$ for all $k$ and all dyadic $s_j, t, u, \rho, \sigma_j, \tau_j$, then, calling $\mathcal{F}$ the smallest $\mu$-complete, right-continuous filtration on $\mathbb{W}^d \times \mathbb{W}^d$ to which the canonical process $(x, y)$ is adapted, $x$ and $y$ are $(\mu, \mathcal{F})$-BM$^d$, and so $\mu$ is in $\text{JIB}^d$.

Hence, $\text{JIB}^d$ is characterized as the set of all probabilities $\mu$ on $\mathbb{W}^d \times \mathbb{W}^d$ such that $\mu(f) = 0$ for all $f$ belonging to some countable set of bounded, continuous functions. Consequently, $\text{JIB}^d$ is closed and convex. It is also relatively compact, because it is tight: tightness follows immediately from both marginals of any $\mu \in \text{JIB}^d$ being equal to $\lambda^d$.

If $X$ and $Y$ are two $\mathcal{F}$-BM$^d$, and if $\mathcal{E}$ is a sub-$\sigma$-field of $\mathcal{F}_0$, one has $\mathbb{E}[\chi(X, Y) \mid \mathcal{E}] = \mathbb{E}[\tilde{\chi}(X, Y) \mid \mathcal{E}] = 0$ a.s. for each dyadic $\chi$, showing that $\mathcal{L}[X, Y \mid \mathcal{E}] \in \text{JIB}^d$ a.s.

Conversely, if $X$ and $Y$ are two processes, and if $\mathcal{L}[X, Y \mid \mathcal{F}_0]$ belongs almost surely to $\text{JIB}^d$, one has $\mathcal{L}[X \mid \mathcal{F}_0] = \lambda^d$ a.s., whence $\mathcal{L}[X] = \lambda^d$, and $X$ is a BM$^d$. Moreover, for every $\mathcal{F}_0$-measurable r.v. $F_0$ and every dyadic $\chi$, one has $\mathbb{E}[\chi(X, Y) \exp i F_0] = 0$. Consequently, for all dyadic $s < t < u$ and $\rho$, and all $\tilde{g}_s$-measurable, bounded r.v. $G_s$, one has

$$E\left[ (e^{i \rho (X_u - X_t)} - e^{-\frac{1}{2} \|\rho\|^2 (u-t)}) G_s \right] = 0,$$

showing that $X$ is a $\tilde{g}$-BM$^d$. □

Keeping $d$ finite, call $\text{MAB}^d$ (for “mutually adapted Brownian motions”) the set of all probabilities $\mu$ on $\mathbb{W}^d \times \mathbb{W}^d$ such that there exist two Brownian motions $X$ and $Y$ which generate the same filtration and have joint law $\mathcal{L}[X, Y] = \mu$. Clearly, $\text{MAB}^d$ is a subset of $\text{JIB}^d$.

Remark (not used in the sequel). It is easy to see that each element of $\text{MAB}^d$ is an extremal point of $\text{JIB}^d$, because it is already extremal in the much bigger set consisting of all probabilities on $\mathbb{W}^d \times \mathbb{W}^d$ with first margin $\lambda^d$. The converse does not hold: there are many extremal points of $\text{JIB}^d$ that do not belong to $\text{MAB}^d$. (Consider for instance, if $d = 1$, the joint law of $X$ and $T X$, where $T$ is the Lévy transform.) It would be interesting to characterize all extremal points of $\text{JIB}^d$; this seems to be a difficult problem. (Already, in two dimensions, call $\text{JU}$ (for ‘joint uniform’) the set of all probabilities $\mu$ on the unit square such that both marginals of $\mu$ are the Lebesgue measure; characterizing the extremal points of $\text{JU}$ is, as far as I know, an open question.)

Proposition 2. The set $\text{MAB}^d$ is dense in $\text{JIB}^d$. In other words, if $X$ and $Y$ are two $\mathcal{F}$-BM$^d$, there exist Brownian motions $\xi^n$ and $\eta^n$ such that

- for each fixed $n$, $\xi^n$ and $\eta^n$ generate the same filtration;
- when $n$ tends to infinity, $(\xi^n, \eta^n)$ converges in law to $(X, Y)$.

Remark (not used in the sequel). An analogue of Proposition 2 is the following fact: when $T$ ranges over the set of all bimeasurable, Lebesgue-measure preserving bijections of $[0, 1]$ to itself, the probabilities $\int_0^1 \hat{\delta}(s, T_s) \, dx$ are
understood by keeping in mind two very different particular cases: the case when
\( \langle \\rangle \) Proof of Proposition 2. The goal is to approximate an arbitrary
\( h \) and \( Z \) are two uniform r.v. on \([0, 1]\), there exist \( \xi^n, \eta^n \), also with uniform law on \([0, 1]\), such that

- for each \( n, \xi^n \) and \( \eta^n \) generate the same \( \sigma \)-field;
- when \( n \) tends to infinity, \( (\xi^n, \eta^n) \) converges in law to \( (X, Y) \).

A proof of this remark is given by Gangbo [7].\(^3\) Another, more direct, proof is the following. We shall work with
\([0, 1]\) instead of \([0, 1]\); by ‘interval’, we shall always mean an interval of the form \([a, b]\). Call \( JU \) the subset of \( JU \) which we claim is dense.

Partition \([0, 1]\) into \( n \) subintervals \( I_1, \ldots, I_n \) of length \( 1/n \); put \( J_k = I_k \). (The intervals \( I_k \) will be used on
the first factor, the \( J_k \) on the second factor.) The square \( S = [0, 1] \times [0, 1] \) is partitioned into \( n^2 \) smaller squares
\( S_{k\ell} = I_k \times J_\ell \). Given a probability \( \mu \in JU \), it suffices to exhibit a probability \( v \in JU' \) such that \( v(S_{k\ell}) = \mu(S_{k\ell}) \) for
all \( k \) and \( \ell \); when \( n \) tends to infinity, these \( v \) will converge towards \( \mu \), thus proving the remark.

Put \( m_{k\ell} = \mu(S_{k\ell}) \). The hypothesis \( \mu \in JU \) implies that the \( 2n \) sums \( \sum_{k=1}^n m_{k\ell} \) and \( \sum_{\ell=1}^n m_{k\ell} \) are all equal
to \( 1/n \). This makes it possible to partition the interval \( I_k \) (respectively \( J_\ell \)) into \( n \) subintervals \( I_{k1}, \ldots, I_{kn} \) (respectively
\( J_{\ell1}, \ldots, J_{\ell n} \)) such that \( \text{Leb}(I_{k\ell}) = \text{Leb}(J_{k\ell}) = m_{k\ell} \). (It is not really a partition, because if \( m_{k\ell} = 0 \) the
corresponding subintervals are empty.) The family \( \{I_{k\ell}, 1 \leq k, \ell \leq n\} \) (respectively \( \{J_{k\ell}, 1 \leq k, \ell \leq n\} \)) is a “par-
tition” of \([0, 1]\) into \( n^2 \) subintervals with lengths \( m_{k\ell} \). Since \( \text{Leb}(I_{k\ell}) = \text{Leb}(J_{k\ell}) \), there exists a bi-Borel, measure-

preserving bijection \( T : [0, 1] \rightarrow [0, 1] \) such that \( T(I_{k\ell}) = J_{k\ell} \) for all \( k \) and \( \ell \); the measure \( v = \int_0^1 \delta_{(x,Tx)} \) belongs to \( JU' \). Call \( G \) the graph of \( T \). The subset \( G_{k\ell} \) of \( G \) corresponding to abscissae in \( I_{k\ell} \) and ordinates
in \( J_{k\ell} \) verifies \( G_{k\ell} \subset I_{k\ell} \times J_{k\ell} \subset S_{k\ell} \); as \( \bigcup_{k,\ell} G_{k\ell} = G \) and as the \( S_{k\ell} \) are disjoint, one has \( G_{k\ell} = G \cap S_{k\ell} \). Thus
\( v(S_{k\ell}) = v(G_{k\ell}) = \text{Leb}(I_{k\ell}) = \text{Leb}(J_{k\ell}) \), that is, \( v(S_{k\ell}) = m_{k\ell} = \mu(S_{k\ell}) \). The claim is established.

We now come back to Proposition 2. Before proving it, here is a small lemma, obvious from the point of view
of Rohlin’s theory, that will be needed in the proof.

**Lemma 7.** Let \( Z \) and \( h \) be two r.v. on some probability space \((\Omega, \mathcal{A}, \mathbb{P})\). On another probability space \((\Omega', \mathcal{A}', \mathbb{P}')\),
let \( \zeta \) and \( \gamma \) be two independent r.v., such that \( \zeta \) has the same law as \( Z \) and \( \gamma \) has a diffuse law. There exists on \( \Omega' \) a r.v. \( h' \), measurable for \( \sigma(\zeta, \gamma) \), such that the joint law \( \mathcal{L}(h', \zeta) \) is equal to \( \mathcal{L}(h, Z) \).

**Proof of Lemma 7.** By replacing \( \gamma \) with \( F\gamma \), where \( F(x) = \mathbb{P}[\gamma \leq x] \), we may suppose \( \gamma \) to be unifor-

mously distributed on \([0, 1]\). Call \( \rho_z(dj) \) a regular version of the conditional law \( \mathbb{P}[h \in dj \mid Z = z] \), let \( f_z(r) = \inf\{j : \rho_z(\{1,\infty, j\}) > r\} \) denote the inverse of the distribution function of \( \rho_z \), and set \( h' = f_z(\gamma) \). Owing to the independence of \( \zeta \) and \( \gamma \), one has \( \mathcal{L}(h') = \mathcal{L}(f_z\gamma) = \rho_z \). Calling \( \mu \) the law of \( Z \) and \( \zeta \), this gives

\[
\mathbb{P}(h' \in dj, \zeta \in dz) = \mu(dz)\rho_z(dj) = \mathbb{P}(h \in dj, Z \in dz),
\]
so \( (h', \zeta) \) has the same law as \( (h, Z) \). \( \square \)

We now have all the necessary ingredients to prove Proposition 2. The proof is rather long; it can be best
understood by keeping in mind two very different particular cases: the case when \( (Y, T, X) \), \( h t \), where \( h \) belongs to \( \text{COV}(d) \) by Lemma 5 and is \( \mathcal{F}_0 \)-measurable; and the case when \( dY_t = (I_{[0,T]} - I_{[T,\infty]})(t) \) \( dX_t \), where \( T \) is a stopping time, independent of \( X \) and of \( Y \).

**Proof of Proposition 2.** The goal is to approximate an arbitrary \( \mu \in JIB^d \) by elements of \( MAB^d \). The first steps
of the proof will consist in replacing \( \mu \) by an element of some suitably chosen sequence tending to \( \mu \) (that is, in taking \( \mu \) in some suitable dense subset of \( JIB^d \)).

\(^3\) I thank F. Delbaen and S. Laurent who independently pointed out this reference to me.
So we start with a filtered probability space \((\Omega, \mathcal{A}, \mathbb{P}, \mathcal{F})\) and two \(\mathcal{F}\)-BM\(d\) \(X\) and \(Y\). According to Lemma 5, there exists a \(\text{COV}(d)\)-valued predictable process \(H\) such that

\[
\langle Y_t, X_j \rangle_t = \int_0^t H_{ij}(s) \, ds,
\]

or, in matrix notation, considering \(X\) and \(Y\) as column vectors,

\[
d\langle Y, tX \rangle = H \, dt.
\]

By (v) and (ii) of Proposition 1, \(I - H^t H \geq 0\).

By an independent enlargement of the filtered probability space, we may suppose without loss of generality that there exists an \(\mathcal{F}\)-BM\(d\) \(\beta\) independent of \((X,Y)\). Put \(Y_\theta = Y \cos \theta + \beta \sin \theta\). For each \(\theta\), \(Y_\theta\) tends in law to \((X,Y)\). So we may work with \((X,Y_\theta)\) in place of \((X,Y)\), for some \(\theta \in [0, \frac{\pi}{2}]\) fixed in the sequel. What we have gained from this replacement is that \(H\) has been multiplied by \(\cos \theta< 1\), so we have a stronger inequality \(I \cos^2 \theta - H^t H \geq 0\). Thus, the predictable process \(K = \sqrt{I - H^t H}(\text{symmetric square root of a symmetric, positive matrix})\) has a well-defined inverse \(K^{-1}\) which is bounded (by \(1/\sin \theta\)).

This makes it possible to define a \(d\)-dimensional \(\mathcal{F}\)-local martingale \(U\) by

\[
U = \int K^{-1}(dY - H \, dX).
\]

As

\[
d(U, tU) = K^{-1}(dY - H \, dX)(t \, dY - H \, dX)K^{-1} = K^{-1}(I \, dt - H^t H \, dt - H^t H \, dt + H^t H \, dt)K^{-1} = K^{-1}(I - H^t H)K^{-1} \, dt = I \, dt,
\]

\(U\) is an \(\mathcal{F}\)-BM\(d\); and as

\[
d(U, tX) = K^{-1}(dY - H \, dX)^t \, dX = K^{-1}(H \, dt - H \, dt) = 0,
\]

\(U\) is independent of \(X\). By the very definition of \(U\), one has

\[
Y = \int H \, dX + \int K \, dU.
\]

Now, for \(\delta = 2^{-n}\), define

\[
H^\delta_t = \begin{cases} 0 & \text{if } t \in [0, \delta], \\ \frac{1}{\delta} \int_{(k-1)\delta}^k H_s \, ds & \text{if } t \in [(k+1)\delta, k\delta] \text{ for } k \geq 1. \end{cases}
\]

By Lemma 5, the process \(H\) takes values in the compact, convex set \(\text{COV}(d)\); so does also \(H^\delta\). Hence we may put \(K^\delta = \sqrt{I - H^\delta \, tH^\delta}\) and define a new \(\mathcal{F}\)-local martingale by

\[
Y^\delta = \int H^\delta \, dX + \int K^\delta \, dU.
\]

Since \(d(Y^\delta, tY^\delta) = (H^\delta \, tH^\delta + K^\delta \, t) \, dt = I \, dt\), \(Y^\delta\) is an \(\mathcal{F}\)-BM\(d\). When \(n \to \infty\) (and \(\delta = 2^{-n} \to 0\)), \(H^\delta\) tends to \(H\) in \(L^1([0,t])\) for fixed \(\omega\) and \(t\), and also, by boundedness, in \(L^2([0,t] \times \Omega)\) for each \(t > 0\); similarly for \(K^\delta\). Thus \(\mathbb{E}[\sup_{s \leq t} \|Y^\delta_s - Y_s\|^2] \to 0\), and \((X, Y^\delta)\) tends in law to \((X, Y)\). Working with \((X, Y^\delta)\) instead of \((X, Y)\), and consequently with \(H^\delta\) instead of \(H\), we may henceforth forget about the preceding construction, and simply...
suppose that, for $n = n_0$ (and hence also for all $n \geq n_0$), the process $H$ is constant (with respect to time) on each interval $]k2^{-n}, (k+1)2^{-n}].$

Let $V$ be a BM$^d$ independent of $\mathcal{F}_\infty$. Call $(\mathcal{V}_t)_{t \geq 0}$ the filtration generated by $V$, and, for $\varepsilon = 2^{-n}$ with $n \geq n_0$, define

$$X_\varepsilon^t = \begin{cases} V_t & \text{if } t \leq \varepsilon, \\ V_t + X_{t-\varepsilon} & \text{if } t > \varepsilon, \end{cases} \quad Y_\varepsilon^t = \begin{cases} V_t & \text{if } t \leq \varepsilon, \\ V_t + Y_{t-\varepsilon} & \text{if } t > \varepsilon, \end{cases} \quad \mathcal{G}_\varepsilon^t = \begin{cases} \mathcal{V}_t & \text{if } t \leq \varepsilon, \\ \mathcal{V}_t \vee \mathcal{F}_{t-\varepsilon} & \text{if } t > \varepsilon. \end{cases}$$

Both $X^\varepsilon$ and $Y^\varepsilon$ are $\mathcal{G}^\varepsilon$-Brownian motions; when $n \to \infty$ (and $\varepsilon = 2^{-n} \to 0$), $(X^\varepsilon, Y^\varepsilon)$ tends to $(X, Y)$ pathwise, hence also in law. Call $Z^\varepsilon$ (resp. $Z$) the $2d$-dimensional process $(X^\varepsilon, Y^\varepsilon)$ (resp. $(X, Y)$).

For fixed $\varepsilon$, we shall now construct two Brownian motions $\xi^\varepsilon$ and $\eta^\varepsilon$, generating the same filtration, and such that, putting $\xi^\varepsilon = (\xi^\varepsilon, \eta^\varepsilon)$, one has

$$\mathcal{L}[\xi^\varepsilon_1, \xi^\varepsilon_2, \ldots, \xi^\varepsilon_{k_{\varepsilon}}, \ldots] = \mathcal{L}[Z^\varepsilon_1, Z^\varepsilon_2, \ldots, Z^\varepsilon_{k_{\varepsilon}}, \ldots].$$

The existence of such $\xi^\varepsilon$ will suffice to prove the proposition. Indeed, let $t_1, \ldots, t_p$ be some dyadic instants. For $\varepsilon = 2^{-n}$ small enough, each $t_i$ is a multiple of $\varepsilon$, hence, when $n \to \infty$,

$$\mathcal{L}[\xi^\varepsilon_{t_1}, \ldots, \xi^\varepsilon_{t_p}] = \mathcal{L}[Z^\varepsilon_{t_1}, \ldots, Z^\varepsilon_{t_p}] \to \mathcal{L}[Z_{t_1}, \ldots, Z_{t_p}],$$

that is, the finite-dimensional dyadic marginal laws of the process $\xi^\varepsilon$ converge to those of $Z$. By a well-known criterion for weak convergence (see for instance Billingsley [2]), this implies $\mathcal{L}[\xi^\varepsilon] \to \mathcal{L}[Z]$ provided the sequence of laws $\mathcal{L}[\xi^\varepsilon]$ is tight; but tightness is here a direct consequence of both components $\xi^\varepsilon$ and $\eta^\varepsilon$ of $\xi^\varepsilon$ having the same law $\lambda^d$ independent of $\varepsilon$. So $\mathcal{L}[\xi^\varepsilon] \to \mathcal{L}[Z]$, establishing the proposition.

It only remains to construct, for fixed $\varepsilon$, two mutually adapted Brownian motions $\xi^\varepsilon$ and $\eta^\varepsilon$ such that the sequences $(\xi^\varepsilon_{k_{\varepsilon}})_{k_{\varepsilon} \geq 0}$ and $(Z^\varepsilon_{k_{\varepsilon}})_{k_{\varepsilon} \geq 0}$ have the same law.

Let $\eta^\varepsilon$ be some BM$^d$ (on any probability space) and $\mathcal{F}$ the filtration generated by $\eta^\varepsilon$. We shall construct $\eta^\varepsilon$ stepwise, successively on each interval $[k\varepsilon, (k+1)\varepsilon]$.

First, on the time-interval $[0, \varepsilon]$, set $\eta^\varepsilon = \xi^\varepsilon$. Since, on $[0, \varepsilon]$, $X^\varepsilon = Y^\varepsilon = V$, one has $\mathcal{L}[\xi^\varepsilon] = \mathcal{L}[Z^\varepsilon]$ (both are normal, centred, with covariance $\varepsilon \begin{pmatrix} I & I \\ I & I \end{pmatrix} \varepsilon$). On $[0, \varepsilon]$, $\xi^\varepsilon$ and $\eta^\varepsilon$ generate the same filtration. Fix an independent sequence $(\gamma_k)_{k \geq 0}$ of $\mathcal{F}$-measurable random variables with diffuse laws, starting with $\gamma_0 = \xi^\varepsilon$. (Take for instance $\gamma_k = \int_0^1 \phi_\varepsilon(t) \, d\xi^\varepsilon_t$, where $\phi_\varepsilon$ are orthogonal in $L^2([0, \varepsilon])$ and $\phi_0 = 1$.)

Now suppose that, for some $k \geq 1$, the Brownian motion $\eta^\varepsilon$ has been defined on the interval $[0, k\varepsilon]$, satisfying the following three properties:

(a) the random vector $(\xi^\varepsilon_{k\varepsilon}, \ldots, \xi^\varepsilon_{k_{\varepsilon}})$ has the same law as $(Z^\varepsilon_{k\varepsilon}, \ldots, Z^\varepsilon_{k_{\varepsilon}})$;
(b) on the time interval $[0, k\varepsilon]$, $\eta^\varepsilon$ generates the same filtration $\mathcal{F}$ as $\xi^\varepsilon$;
(c) the process $(\xi^\varepsilon_{t}, \quad t \in [\varepsilon, k\varepsilon])$ is independent of the sequence $(\gamma_{k}, \gamma_{k+1}, \ldots)$.

We shall further extend $\eta^\varepsilon$ to $[k\varepsilon, (k+1)\varepsilon]$ so that these three properties hold with $k + 1$ instead of $k$. As they already hold for $k = 1$, by induction on $k$ this will define the process $\eta^\varepsilon$ for all $t$, in such a way that, by (a), the sequences $(\xi^\varepsilon_{k_{\varepsilon}})_{k_{\varepsilon} \geq 0}$ and $(Z^\varepsilon_{k_{\varepsilon}})_{k_{\varepsilon} \geq 0}$ have the same law, and, by (b), $\eta^\varepsilon$ generates $\mathcal{F}$. The proposition will thus be proved.

Recall the existence of a COV($d$)-valued r.v. $h_\varepsilon$ such that $d(Y, t X) = h_\varepsilon \, dt$ on $[(k-1)\varepsilon, k\varepsilon]$. By right-continuity of the filtration $\mathcal{F}$, $h_\varepsilon$ is $\mathcal{F}_{(k-1)\varepsilon}$-measurable. Thus the conditional law $\mathcal{L}[Z_{(k-1)\varepsilon} \mid \mathcal{F}_{(k-1)\varepsilon}]$ is the normal distribution with mean $Z_{(k-1)\varepsilon}$ and covariance $\varepsilon \begin{pmatrix} I & h_\varepsilon \\ h_\varepsilon & I \end{pmatrix}$. Enlarging $\mathcal{F}_{(k-1)\varepsilon}$ with the $\sigma$-field $\mathcal{V}_\varepsilon$ (which is independent of $\mathcal{F}_\infty$) does not modify this conditional law; so

$$\mathcal{L}[Z_{(k+1)\varepsilon} - Z_{k\varepsilon} \mid \mathcal{G}_{k\varepsilon}] = \mathcal{L}[Z_{(k+1)\varepsilon} - Z_{(k-1)\varepsilon} \mid \mathcal{F}_{(k-1)\varepsilon}] = N \left( 0, \varepsilon \begin{pmatrix} I & h_\varepsilon \\ h_\varepsilon & I \end{pmatrix} \varepsilon \right).$$
Call $\rho_{z_1,...,z_k} (dh)$ a regular version of the conditional law $P[h_t \in dh \mid Z^e_{t'} = z_1, \ldots, Z^e_{t''} = z_k]$. Conditioning both sides of the above equality w.r.t. the sub-$\sigma$-field $\sigma(Z^e_{t'}, \ldots, Z^e_{t''})$ of $g^e_{t''}$ yields

$$\mathcal{L}[Z^e_{(k+1)e} \mid Z^e_{t'}, \ldots, Z^e_{t''}] = \int_{h \in \text{COV}(d)} N \left( \begin{array}{c} Z^e_{t'} \\ \epsilon \left( \begin{array}{c} h \\ -l \end{array} \right) \end{array} \right) \rho_{Z^e_{t'},\ldots,Z^e_{t''}} (dh). \quad (*)$$

By hypothesis (c), $\gamma_k$ is independent of $(\zeta^e_{t'}, \ldots, \zeta^e_{t''})$; so Lemma 7 gives the existence of a COV($d$)-valued r.v. $h^e_k$ of the form

$$h^e_k = f_k(y_k, \zeta^e_{t'}, \ldots, \zeta^e_{t''})$$

and such that $\mathcal{L}[h^e_k, \zeta^e_{t'}, \ldots, \zeta^e_{t''}] = \mathcal{L}[h_k, Z^e_{t'}, \ldots, Z^e_{t''}]$; consequently

$$\mathcal{L}[h^e_k \mid \zeta^e_{t'} = z_1, \ldots, \zeta^e_{t''} = z_k] = \mathcal{L}[h_k \mid Z^e_{t'} = z_1, \ldots, Z^e_{t''} = z_k] = \rho_{z_1,\ldots,z_k}. \quad (***)$$

By Proposition 1, there exist random variables $o_1 = r_1(h^e_k), \ldots, o_d = r_d(h^e_k)$ in $\text{O}(d)$ and $\beta_1 = \alpha_1(h^e_k), \ldots, \beta_d = \alpha_d(h^e_k)$ in $[0, 1]$ such that $\beta_1 + \cdots + \beta_d = 1$ and $\beta_1 o_1 + \cdots + \beta_d o_d = h^e_k$; by definition of $h^e_k$ and by hypothesis (b), all these r.v. are $\mathcal{E}_{t''}$-measurable. This entails that, for $i \in \{0, 1, \ldots, d' \}$, $T_i = k e + (\beta_1 + \cdots + \beta_i) \epsilon$ is an $\mathcal{E}$-stopping time; observe that $T_0 = k e$ and $T_{d'} = (k + 1) \epsilon$. Define now the Brownian motion $\eta^e$ on the interval $[k e, (k + 1) \epsilon]$ by

$$1_{[k e, (k + 1) \epsilon]}(t) \, d\eta^e = \sum_{i=1}^{d'} 1_{[T_i - 1, T_i]}(t) o_i \, d\xi^e_i.$$ 

This formula inverts itself as

$$1_{[k e, (k + 1) \epsilon]}(t) \, d\xi^e_i = \sum_{i=1}^{d'} 1_{[T_i - 1, T_i]}(t) o_i^{-1} \, d\eta^e_i,$$

showing that the mutual adaptedness of $\xi^e$ and $\eta^e$ extends to the interval $[0, (k + 1) \epsilon]$; so (b) holds for $k + 1$.

From the definition of $\eta^e$ on $[k e, (k + 1) \epsilon]$, the conditional law $\mathcal{L}[\zeta^e_{(k+1)e} \mid \mathcal{E}_{t''}]$ is the centred normal distribution with covariance

$$\sum_{i=1}^{d'} (T_i - T_i - 1) \begin{pmatrix} I & o_i \\ o_i & I \end{pmatrix} = \sum_{i=1}^{d'} \beta_i \begin{pmatrix} I & o_i \\ o_i & I \end{pmatrix} = \varepsilon \left( \begin{array}{cc} I & h^e_k \\ h^e_k & I \end{array} \right).$$

Using (**) and further conditioning w.r.t. $(\zeta^e_{t'}, \ldots, \zeta^e_{t''})$ gives

$$\mathcal{L}\left[\begin{array}{c} \zeta^e_{(k+1)e} \\ \zeta^e_{t'}) \end{array} \right] = \int_{h \in \text{COV}(d)} N \left( \begin{array}{c} \zeta^e_{(k+1)e} \\ \epsilon \left( \begin{array}{c} h \\ -l \end{array} \right) \end{array} \right) \rho_{\zeta^e_{t'},\ldots,\zeta^e_{t''}} (dh).$$

Comparing with (*) and recalling that, by induction hypothesis, $\mathcal{L}[\zeta^e_{t'}, \ldots, \zeta^e_{t''}] = \mathcal{L}[Z^e_{t'}, \ldots, Z^e_{t''}]$, one obtains

$$\mathcal{L}[\zeta^e_{t'}, \ldots, \zeta^e_{t''}, \zeta^e_{(k+1)e}] = \mathcal{L}[Z^e_{t'}, \ldots, Z^e_{t''}, Z^e_{(k+1)e}],$$

showing that (a) holds for $k + 1$.

Last, there remains to show that (c) holds for $k + 1$. If $Q$ is any process, denote by $dQ_{[s,t]}$ the process $(Q_u - Q_s, u \in [s,t])$, and by $Q_{[s,t]}$ the process $(Q_u, u \in [s,t])$. By induction hypothesis, we know that the sequence $(\zeta^e_{[s,k e]})$ is $\mathcal{E}_{k e}$-measurable, it is also independent of $d\xi^e_{k e,(k+1) e}$; hence the sequence $(d\xi^e_{k e,(k+1) e}, \zeta^e_{[s,k e],(k+1) e})$ is independent too. So the $\sigma$-field $\mathcal{E} = \sigma(d\xi^e_{k e,(k+1) e}, \zeta^e_{[s,k e],(k+1) e})$ is independent of $(\zeta^e_{[k e,k e+1]}, \ldots)$. Now, $h^e_k$ is $\mathcal{E}$-measurable by definition, so $d\eta^e_{k e,(k+1) e}$ and consequently also $d\xi^e_{k e,(k+1) e}$ are $\mathcal{E}$-measurable, and by concatenation with $\zeta^e_{[s,k e],(k+1) e}$, finally $\zeta^e_{[s,(k+1) e]}$ is $\mathcal{E}$-measurable and therefore independent of $(\zeta^e_{[k e,k e+1]}, \ldots)$. \[\square\]
Restatement of Theorem 1

Theorem 1 asserts the equivalence of two conditions. To prove it, it will be convenient to introduce two more equivalent conditions, thus splitting up the argument into shorter chunks and making the ideas stand out more clearly. That is why we state it again, with four equivalent conditions instead of two; the first half of the statement is copied verbatim from Theorem 1.

Theorem 1'. Fix \( d \in \mathbb{N}^* \cup \{ \infty \} \) and a filtered probability space \((\Omega, \mathcal{A}, \mathbb{P}, \mathcal{F})\); suppose that \( \mathcal{F} \) is \( d \)-Brownian after zero. The following four statements are equivalent:

(i) \( \mathcal{F} \) is \( d \)-Brownian.

(ii) (Self-coupling condition.) For each \( R \in L^1(\mathcal{F}_\infty) \) and each \( \delta > 0 \), there exists a probability space \((\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})\) endowed with two filtrations \( \mathcal{F}' \) and \( \mathcal{F}'' \) verifying the following four conditions:
   (a) \( \mathcal{F}' \) and \( \mathcal{F}'' \) are isomorphic to \( \mathcal{F} \); in particular, there are two r.v. \( R', R'' \in L^1(\mathcal{F}'_\infty) \) corresponding to \( R \) by these isomorphisms;
   (b) \( \mathcal{F}' \) and \( \mathcal{F}'' \) are jointly immersed;
   (c) for some \( s > 0 \), \( \mathcal{F}' \) and \( \mathcal{F}'' \) are \( \mathbb{P} \)-independent;
   (d) \( \|R' - R''\|_{L^1(\bar{\mathcal{F}})} < \delta \).

(iii) (Approximation by a Brownian filtration.) For each \( R \in L^1(\mathcal{F}_\infty) \) and each \( \delta > 0 \), there exists a probability space \((\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})\) with two filtrations \( \mathcal{F}' \) and \( \mathcal{B} \) such that
   (a) \( \mathcal{F}' \) is isomorphic to \( \mathcal{F} \); in particular, there is a r.v. \( R' \in L^1(\mathcal{F}'_\infty) \) corresponding to \( R \) by this isomorphism;
   (b) \( \mathcal{B} \) is \( d \)-Brownian;
   (c) \( \mathcal{F}' \) and \( \mathcal{B} \) are jointly immersed;
   (d) there exists \( R'' \in L^1(\mathcal{B}) \) such that \( \|R' - R''\|_{L^1(\bar{\mathcal{F}})} < \delta \).

(iv) (Vershik’s first-level criterion.) For each \( R \in L^1(\mathcal{F}_\infty) \) and each \( \delta > 0 \), there exists an \( \mathcal{F} \)-BM\(^d\) \( B \) such that
   (a) \( \forall s > 0 \forall t \geq s \mathcal{F}_t = \mathcal{F}_s \vee \sigma(B_u - B_t, u \in [s, t]) \);
   (b) there exists \( R' \in L^1(\sigma(B)) \) such that \( \|R - R'\|_{L^1(\mathcal{F})} < \delta \).

Remarks (extension to (iii) and (iv) of the remarks following Theorem 1). (1) Let \( D \) be a dense subset of \( L^1(\mathcal{F}_\infty) \). In Remark 1 following Theorem 1, we saw that \( \forall R \in L^1(\mathcal{F}_\infty) \) in (ii) can be replaced with \( \forall R \in D \). It is also true in (iii) and (iv), by the same argument.

(2) As in condition (ii), and for the same reason, the space \( L^1 \) and its norm can equivalently be replaced in (iii) or (iv) by \( L^p \) for any \( p \in [1, \infty[ \), or by \( L^0 \), with the distance \( \mathbb{E}[1 \wedge |X - Y|] \). In the case when \( p = 2 \), one can take \( R'' \) in (iii-d) or \( R' \) in (iv-b) to be the best \( L^2 \)-approximation, that is, \( R'' = \mathbb{E}[R' | \mathcal{B}_\infty] \) in (iii-d) or \( R' = \mathbb{E}[R | \sigma(B)] \) in (iv-b).

(3) S. Laurent’s observation that (ii) needs to be verified for one \( R \) only, provided this \( R \) generates \( \mathcal{F}_\infty \), also extends to (iii) and (iv).

Proof of (i) \( \Rightarrow \) (ii) in Theorems 1 and 1’

Suppose \( \mathcal{F} \) to be generated by some BM\(^d\) \( B \), and fix \( R \in L^1(\mathcal{F}_\infty) \) and \( \delta > 0 \). The \( \sigma \)-fields \( \mathcal{A}_n = \sigma(B_t - B_{1/n}, \ t \in [1/n, \infty[) \) form an increasing sequence with limit \( \mathcal{F}_\infty = \mathcal{F}_\infty \); hence, for a suitable integer \( m \), the r.v. \( S = \mathbb{E}[R | \mathcal{A}_m] \) verifies \( \|S - R\|_{L^1(\mathcal{F})} < \delta \).

Let \( B' \) and \( C \) be two independent BM\(^d\) on some filtered probability space \((\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}}, \bar{\mathcal{F}})\); the process \( B'' \) defined by \( B''_t = 0 \) and
\[
dB''_t = 1_{[0,1/m]}(t)\,dC_t + 1_{[1/m,\infty[}(t)\,dB'_t
\]
is also an \( \bar{\mathcal{F}} \)-BM\(^d\). Call \( \mathcal{F}' \) and \( \mathcal{F}'' \) the filtrations on \( \bar{\Omega} \) respectively generated by \( B' \) and \( B'' \).
(b) Since $B'$ and $B''$ are $\mathcal{F}$-Brownian motions, $\mathcal{F}'$ and $\mathcal{F}''$ are immersed in $\mathcal{F}$.

(a) $\mathcal{F}'$ and $\mathcal{F}''$ are isomorphic to $\mathcal{F}$, by the isomorphisms $\Psi'$ and $\Psi''$ such that $B' = \Psi'(B)$, $B'' = \Psi''(B)$. Put $R' = \Psi'(R)$, $R'' = \Psi''(R)$, $S' = \Psi'(S)$ and $S'' = \Psi''(S)$.

(c) $\mathcal{F}'_{1/m} = \sigma(B'_t, t \in [0, 1/m])$ and $\mathcal{F}''_{1/m} = \sigma(C_t, t \in [0, 1/m])$ are independent, because $B'$ and $C$ are.

(d) As $S$ is $\mathcal{A}_m$-measurable, one has $S = h(B'_1 - B''_{1/m}, t \in [1/m, \infty[)$ a.s. for some Borel $h$. By isomorphisms, $S' = h(B'_1 - B''_{1/m}, t \in [1/m, \infty[)$ a.s. and $S'' = h(B''_{1/m} - B''_{1/m}, t \in [1/m, \infty[)$ a.s. By definition of $B''$, this implies $S' = S''$. Now, from $\|R - S\|_{L^1(\Omega)} < \delta$, one gets by isomorphisms $\|R' - S'\|_{L^1(\mathcal{F})} < \delta$ and $\|R'' - S''\|_{L^1(\mathcal{F})} < \delta$; consequently $\|R' - R''\|_{L^1(\mathcal{F})} < 2\delta$.

**Proof of (ii) ⇒ (iii) in Theorem 1’**

We shall give the proof in the (slightly more complicated) case that $d = \infty$, and briefly indicate between brackets the simplifications to be made when $d$ is finite.

Fix $R \in L^1(\mathcal{F}_\infty)$ and $\delta > 0$; fix also $\mathcal{F}'$, $\mathcal{F}''$ and $s$ verifying conditions (ii-a) to (ii-d). Call $\mathcal{F}'$ the filtration $\mathcal{F}$ shifted by $s$. As $\mathcal{F}$ is $d$-Brownian after zero, there exists an $\mathcal{F}'$-$BM^d$ $X'$ such that $\mathcal{F}$ is generated by $\mathcal{F}_s$ and $X'$. Call $\mathcal{A}_n$ the $\sigma$-field generated by $\mathcal{F}_s$ and by the first $n$ components of $X'$; these $\mathcal{A}_n$ form an increasing sequence with limit $\bigvee_n \mathcal{A}_n = \mathcal{F}_s \vee \sigma(X') = \mathcal{F}_\infty$. By martingale convergence, there exists an integer $e < d = \infty$ and a v.f. $S \in L^1(\mathcal{A}_e)$ such that $\|S - R\|_{L^1(\mathcal{A}_e)} < \delta$. Call $X$ the $BM^e$ consisting of the first $e$ components of $X'$, so that $S$ is measurable for $\mathcal{F}_s \vee \sigma(X)$. [In the case when $d < \infty$, the $\mathcal{A}_n$ and the replacement of $d$ by $e$, $R$ by $S$, $X$ by $X$, are not needed.] Calling $S'$ and $S''$ the copies of $S$ in $\mathcal{F}'$ and $\mathcal{F}''$, one has $\|S' - R\|_{L^1(\mathcal{F})} < \delta$ and $\|S'' - R\|_{L^1(\mathcal{F})} < \delta$ by isomorphisms; from $\|R' - R''\|_{L^1(\mathcal{F})} < \delta$, one gets $\|S' - S''\|_{L^1(\mathcal{F})} < 3\delta$.

Let $F'$ denote a v.f. generating the $\sigma$-field $\mathcal{F}_s$, and $\Phi$ the law of $F$. There exists a measurable $f$ such that $S = f(F, X)$; transferring $F$ and $X$ from $\mathcal{F}$ to $\mathcal{F}'$ and $\mathcal{F}''$ yields $F', X', F'', X''$ such that $\mathbb{E}[\|f(F', X') - f(F', X'')\|] < \delta$.

By (ii-b), $\mathcal{F}'$ and $\mathcal{F}''$ are jointly immersed; so are also $(\mathcal{F}_s)'_{1 \geq 0}$ and $(\mathcal{F}''_s)'_{1 \geq 0}$; hence $X'$ and $X''$ are $\mathcal{F}_0$-$BM^d$, for some filtration $\mathcal{F}_0$ on $\Omega$ such that $\mathcal{F}_s \vee \mathcal{F}_0 \subset \mathcal{F}_0$. By Lemma 6, the conditional joint law $\mathcal{L}[X', X'' | \mathcal{F}_s \vee \mathcal{F}_0]$ belongs to JIB$^e$. So, using the independence of $\mathcal{F}_s$ and $\mathcal{F}_0$, given by (ii-c), the joint law of $F'$, $F''$, $X'$, $X''$ can be written

$$\mathbb{P}[F' \in dz', F'' \in dz'', X' \in dx', X'' \in dx''] = \Phi(dx'')\Phi(dx')\mu_{z',z''}(dz', dz'')$$

with $\mu_{z',z''} \in JIB^e$ for $\Phi$-almost all $z'$ and $z''$. The estimate of $\|S' - S''\|$ becomes

$$\int |f(z', x') - f(z'', x'')|\Phi(dx'')\Phi(dx')\mu_{z',z''}(dz', dz'') < 3\delta.$$ 

Consequently, the set of all $z''$ such that

$$\mu_{z',z''} \in JIB^e \text{ for } \Phi\text{-almost all } z'$$

$$\int |f(z', x') - f(z'', x'')|\Phi(dx')\mu_{z',z''}(dz', dx'') < 3\delta$$

is not $\Phi$-negligible and a fortiori non empty. Fix such a $z''$ and put $g(y) = f(z'', y)$ and $\nu_z = \mu_{z',z''}$; one has $\nu_z \in JIB^e$ for almost all $z$ and

$$\int |f(z, x) - g(y)|\Phi(dx)\nu_z(dx, dy) < 3\delta.$$ 

On a suitable $(\mathcal{G}_1, A, \mathbb{P})$, define $\mathcal{F}$, $\mathcal{X}$, $\mathcal{X}', \mathcal{X}''$ with joint law $\Phi(dx)\nu_z(dx, dy)\lambda^\infty(du)\lambda^d(du)$. [When $d$ is finite, the $BM^\infty$ $\mathcal{U}$ is not needed.] Since the first marginal of $\nu_z$ is the Wiener measure $\chi$, $X$ is a $BM^\infty$ independent of $\mathcal{F}$, and the infinite-dimensional process $\mathcal{X}' = (\mathcal{X}, \mathcal{U})$ is a $BM^\infty$ independent of $\mathcal{F}$. [When $d$ is finite, there is no $\mathcal{U}$, and $\mathcal{X}' = \mathcal{X}$ is a $BM^d$.] The pair $(\mathcal{F}, \mathcal{X}')$ has the same law as $(F, X)$; this gives an isomorphism between $(\mathcal{G}_1, \mathcal{F}_\infty, \mathbb{P})$ and $(\mathcal{G}_1, \sigma(\mathcal{F}, \mathcal{X}', \mathbb{P}))$; call $\mathcal{F}$ the filtration on $\mathcal{G}_1$ corresponding to $\mathcal{F}$ by this isomorphism. Similarly,
using the second marginal of \( v_s \), \( \hat{Y} \) is a BM by, and the infinite-dimensional process \( \hat{Y}^s = (\hat{Y}, \hat{U}) \) is a BM in. [When \( d \) is finite, \( \hat{Y}^s = \hat{Y} \) is a BM]. So the process

\[
\hat{B}_t = \begin{cases} 
\hat{V}_t & \text{if } t \leq s, \\
\hat{V}_t + \hat{Y}_{t-s} & \text{if } t > s
\end{cases}
\]

is also a BM in \([a, b] \) if \( d \) is finite. Call \( \hat{B} \) the filtration it generates. To establish (iii), it suffices to check that \( \hat{F} \) and \( \hat{B} \) satisfy conditions (iii-a) to (iii-d); (a) and (b) follow from the definitions of \( \hat{F} \) and \( \hat{B} \).

(d) Since the law of \((\hat{F}, \hat{X}, \hat{Y})\) is \( \phi(dx_1, dx_2, dy) \), one has \( \mathbb{E}[|f(\hat{F}, \hat{X}) - g(\hat{Y})|] < 3\delta \). Now, \( f(\hat{F}, \hat{X}) \) is nothing but \( \delta \), corresponding to \( \hat{F} \) to \( S \) in \( \mathcal{F} \). Hence, \( \|\hat{S} - g(\hat{Y})\|_{L^1} < 3\delta \), and, since \( \|\hat{S} - \hat{H}\|_{L^1} < \delta \), one has \( \|\hat{H} - g(\hat{Y})\|_{L^1} < 4\delta \). As \( \hat{S} \) is \( \hat{B}_s \)-measurable, this gives (d) with 4\( \delta \) instead of \( \delta \) [with \( \delta \) if \( d \) is finite].

It remains to show (c). Since \( v_s \in \text{JIB}^c \) for almost all \( z \), Lemma 6 says that \( \hat{X} \) and \( \hat{Y} \) are BM by for the filtration generated by \( \hat{X}, \hat{Y} \), and the constant process \( \hat{F} \). Performing the independent enlargement of \( \hat{U} \), one obtains that \( \hat{X}^s \) and \( \hat{Y}^s \) are BM in \( \mathcal{F} \), generated by \( \hat{X}, \hat{Y} \), and \( \hat{F} \) is immersed in \( \hat{\mathcal{F}} \), and \( \hat{Y}^s \) is a BM in \( \hat{\mathcal{F}} \). Define \( \hat{g} \) by

\[
\hat{g}_t = \begin{cases} \hat{F}_t & \text{if } t < s, \\
\hat{g}_t & \text{if } t \geq s.
\end{cases}
\]

\( \hat{F} \) is immersed in \( \hat{g} \); this is trivial on \([0, s] \), and holds on \([s, \infty) \) because \( \hat{F}_{t+} = \hat{F}_t \) is immersed in \( \hat{g} \). Similarly, \( \hat{Y}_{t+} = \hat{Y}_t \), \( \hat{g} \) being an \( \mathcal{F} \)-martingale (null on \([0, s] \), Brownian after \( s \)). The independent enlargement of \( \hat{g} \) with the process \( \hat{V} \) gives a filtration \( \hat{H} \) such that \( \hat{F} \) is immersed in \( \hat{H} \), and \( \hat{Y}_{t+} = \hat{Y}_t \) is an \( \hat{H} \)-martingale. Since \( \hat{V} \) is an \( \hat{H} \)-BM, \( \hat{B} \) is an \( \hat{H} \)-BM too, and this means that \( \hat{B} \) is immersed in \( \hat{H} \). So \( \hat{F} \) and \( \hat{B} \) are jointly immersed.

**Proof of (iii) \( \Rightarrow \) (iv) in Theorem 1’**

Fix \( R \in L^1(\mathcal{F}_\omega) \) and \( \delta > 0 \). Assumption (iii) provides us with \( \mathcal{F}^r \) and \( \mathcal{B} \); by isomorphism, instead of proving property (iv) for the r.v. \( R \) in the filtration \( \mathcal{F} \), it suffices to prove it for \( R^r \) in \( \mathcal{F}^r \). So, without loss of generality, we shall suppose that the isomorphism in (iii-a) is identity, \( \mathcal{F}^r = \mathcal{F} = \mathcal{F}^r = R \). Hypothesis (iii) provides us with a filtration \( \mathcal{H} \) such that \( \mathcal{F} \) is immersed in \( \mathcal{H} \), and an \( \mathcal{H} \)-BM \( Y \) such that \( R \) is \( \delta \)-close in \( L^1 \) to a functional of \( Y \).

Approximating \( \sigma(Y) \) with \( \sigma(Y^t = Y^t_1 - Y^t_2, t \in [r, \infty)) \) for a suitable \( \epsilon < \infty \) [equal to \( d \) if \( d \) is finite] and a suitable \( r > 0 \), we obtain that

\[
\|R - g(\hat{Y}^t_1 - \ldots, \hat{Y}^t_\epsilon)\|_{L^1} < \delta
\]

where \( \hat{Y}_t^i = Y^i_{t+} - Y^i_r \). Call \( F \) a r.v. generating \( \mathcal{F}_r \), and \( \phi \) the law of \( F \). Since \( \mathcal{F} \) is \( \mathcal{F} \)-Brownian after zero, there exists an \( \mathcal{F} \)-BM \( X \) generating the increments of \( \mathcal{F} \) after \( r \); by the same argument as above, one has also

\[
\|R - f(X^1, \ldots, X^\epsilon)\|_{L^1} < \delta
\]

for a suitable finite \( \epsilon \geq \epsilon \) [equal to \( d \) and \( d \) if \( d \) is finite]. Adding if necessary some unused arguments to the functional \( g \), we henceforth suppose \( \epsilon = \epsilon \); so both processes \( \hat{X} = (X^1, \ldots, X^\epsilon) \) and \( \hat{Y} = (\hat{Y}^1, \ldots, \hat{Y}^\epsilon) \) are BM by for the filtration \( (\mathcal{H}_t)_{t>0} \).

Now, \( f(F, \hat{X}) \) can be \( L^1 \)-approximated by \( \hat{f}^c(F, \hat{X}) \) with \( f^c \) bounded and continuous on \( \mathbb{R} \times W^c \); similarly, \( g(\hat{Y}) \) is approximated by \( \hat{g}(\hat{Y}) \) with \( \hat{g}^c \) bounded and continuous on \( W^c \). So we have \( \|R - f^c(F, \hat{X})\| < \delta \) and \( \|R - g^c(\hat{Y})\| < \delta \), whence

\[
\mathbb{E}[|f^c(F, \hat{X}) - g^c(\hat{Y})|] < 2\delta.
\]

The function \( h_z(x, y) = |f^c(z, x) - g^c(y)| \) is bounded and continuous in \((z, x, y)\).
As \( F \) generates \( \mathcal{F}_r \) which is included in \( \mathcal{H}_r \), and as \( \overline{X} \) and \( \overline{Y} \) are \( (\mathcal{H}_{r+t})_{t \geq 0} \)-BM\(^c\), Lemma 6 says that the conditional joint law

\[
\mu_z = \mathcal{L}[\overline{X}, \overline{Y} \mid F = z]
\]

belongs to JIB\(^c\) for \( \phi \)-almost all \( z \).

The set MAB\(^c\), with the topology inherited from JIB\(^c\), is separable (because JIB\(^c\) is separable and metrizable). Let \( (m_k)_{k \in \mathbb{N}} \) be a dense sequence in MAB\(^c\); by Proposition 2, this sequence is also dense in JIB\(^c\). Consequently,

\[
n(z) = \inf\{k \in \mathbb{N} \mid |m_k(h_z) - \mu_z(h_z)| < \delta \}
\]

is finite for \( \phi \)-almost all \( z \). Since the map

\[
z \mapsto \mu_z(h_z) = \mathbb{E}[|f^c(z, \overline{X}) - g^c(\overline{Y})| \mid F = z]
\]

is measurable and \( z \mapsto m_k(h_z) \) is continuous, \( z \mapsto n(z) \) is measurable too. Integrating \( z \) out of the inequality \( m_{n(z)}(h_z) < \mu_z(h_z) + \delta \), one gets

\[
\int h_z(x, y)\phi'(dz)m_{n(z)}(dx, dy) < \int h_z(x, y)\phi'(dz)\mu_z(dx, dy) + \delta = \mathbb{E}[h_F(\overline{X}, \overline{Y})] + \delta < 3\delta.
\]

Each \( m_k \) is the joint law of two BM\(^c\) \( \xi \) and \( \eta \) that generate the same filtration; let each \( m_k \) is associated a transformation \( T_k : W^c \to W^c \) which is invertible, bi-measurable, \( \lambda^c \)-preserving, \( 4 \) filtration-preserving, and such that \( \eta = T_k \xi \) a.s. and one has \( m_k(dx, dy) = \lambda^c(dx)\delta_{T_k(y)}(y) \), where \( \delta_x \) denotes the unit mass at \( x \). Putting \( \overline{U} = T_n(F)\overline{X} \), the \( 3\delta \)-estimate becomes

\[
3\delta > \int h_z(x, y)\phi'(dz)m_{n(z)}(dx, dy) = \int h_z(x, T_n(c)x)\phi'(dz)\lambda^c(dx) = \mathbb{E}[h_F(\overline{X}, T_n(F)\overline{X})] = \mathbb{E}[|f^c(F, \overline{X}) - g^c(\overline{U})|].
\]

The events \( A_k = \{n(F) = k\} \) form a countable, \( \mathcal{F}_r \)-measurable partition of \( \Omega \). As \( T_k \) is invertible and filtration-preserving, the filtrations

\[
\overline{X}_k = \mathcal{F}_r \lor \sigma(\overline{X}_s, s \in [0, t]) \quad \text{and} \quad \overline{U}_k = \mathcal{F}_r \lor \sigma(\overline{U}_s, s \in [0, t])
\]

agree on each \( A_k \); as \( A_k \) belongs to both \( \overline{X}_0 \) and \( \overline{U}_0 \), the filtrations \( \overline{X} \) and \( \overline{U} \) are equal. On \( A_k \), \( \overline{U} \) is equal to \( T_k \overline{X} \) which is an \( \overline{X} \)-BM\(^c\); so \( \overline{U} \) is an \( \overline{X} \)-BM\(^c\). The independent enlargement with the components of \( X \) which do not appear in \( \overline{X} \) shows that the process \( U = (\overline{U}^1, \ldots, \overline{U}^r, X^{r+1}, X^{r+2}, \ldots) \) [take \( U = \overline{U} \) if \( d \) is finite] is an \((\mathcal{F}_{r+t})_{t \geq 0}\)-BM\(^d\) such that the filtrations

\[
\overline{X}_t = \mathcal{F}_t \lor \sigma(\overline{X}_s, s \in [0, t]) \quad \text{and} \quad \overline{U}_t = \mathcal{F}_t \lor \sigma(\overline{U}_s, s \in [0, t])
\]

are equal; as \( \overline{X}_t \lor \sigma(\overline{X}_s, s \in [0, t]) \lor \sigma(\overline{X}_s, s \in [0, t]) \lor \sigma(\overline{X}_s, s \in [0, t]) \) is also an \( \mathcal{F} \)-BM\(^d\) with the same generating property; in other words, it satisfies (iv-a).

Recall that \( \mathcal{F}_t = \mathcal{F}_s \lor \sigma(X_u, u \in [s, t]) \) for all \( 0 < s \leq t \). As \( \mathcal{U}_t = \mathcal{F}_{r+t} \), the process

\[
B_t = \begin{cases} X_t & \text{if } t \leq r, \\ X_t + U_{t-r} & \text{if } t > r \end{cases}
\]

is also an \( \mathcal{F} \)-BM\(^d\) with the same generating property; in other words, it satisfies (iv-b) (with \( 4\delta \) instead of \( \delta \)).

---

[4] Recall that \( \lambda^c \) denotes the \( c \)-dimensional Wiener measure.
Proof of (iv) ⇒ (i) in Theorem 1'

Fix a r.v. $R \in L^1(\mathcal{F}_\infty)$ such that $R$ generates the $\sigma$-field $\mathcal{F}_\infty$. By induction, we shall construct a sequence of numbers $s_n > 0$ and a sequence of processes $X^n$ such that

1. $s_n \leq s_{n-1}/2$;
2. $X^n = 0$ on $[0, s_n]$;
3. $(X^n_{s_{n+1}}, t)_{t \geq 0}$ is a BM$^d$ for the filtration $(\mathcal{F}_{s_{n+1}})_{t \geq 0}$;
4. $dX^n = dX^{n-1}$ on $[s_{n-1}, \infty]$;
5. $\mathcal{F}_{s_n} \vee \sigma(X^n) = \mathcal{F}_\infty$;
6. for some $U_n \in L^1(\sigma(X^n))$, $\|R - U_n\|_{L^1} < 1/n$.

Assuming this has been done, the $\mathcal{F}$-BM$^d$ $Y$ defined by

$$dY = \sum_{n=0}^\infty 1_{[s_n, s_{n+1}]}(t) dX^n_t + 1_{[0, s_0]}(t) dX^0_t$$

verifies $X^n = \int 1_{[0, \infty]} dY$ and in particular $\sigma(X^n) \subset \sigma(Y)$; consequently each $U_n$ is $\sigma(Y)$-measurable, and so is $R = L^1$-lim $U_n$ too. The filtration $\mathcal{F}$ generated by $Y$ is immersed in $\mathcal{F}$ (because $Y$ is an $\mathcal{F}$-BM) and verifies $\mathcal{F}_\infty = \mathcal{F}_\infty$ (because $R$ generates $\mathcal{F}_\infty$ and is $\mathcal{F}_\infty$-measurable). So, if $A \in \mathcal{F}_\infty$, the martingale $P[A \mid Y_t]$ is the $\mathcal{F}$-martingale with terminal value $1_A$; in particular, for $A \in \mathcal{F}$, $P[A \mid Y_t] = P[A \mid \mathcal{F}_t] = 1_A$, showing $Y_t = \mathcal{F}_t$; and $\mathcal{F}$ is generated by $Y$, proving (i).

It remains to perform the induction. Call $B$ the set of all $\mathcal{F}$-BM$^d$ $B$ such that, for all $0 < s \leq t$, $\mathcal{F}_t = \mathcal{F}_s \vee \sigma(B_\infty - B_s, u \in [s, t])$. By (iv) we know that the set $[\sigma(B), B \in B]$ is substantial in $\mathcal{F}_\infty$. For $B \in B$, put $B^s_t = B_t - B_t\vee s = \int_0^s 1_{[u \leq s]} dB_u$. Start the induction by choosing some $s_0 > 0$ and some $Z \in B$, and by setting $X^0 = Z|_{\sigma(X^n)}$; $s_0$ and $X^0$ verify properties (2), (3) and (5) with $n = 0$.

Now, for some fixed $n \geq 0$, suppose $s_n$ and $X^n$ have been constructed, verifying (2), (3) and (5); we shall show the existence of $s_{n+1}$ and $X^{n+1}$ verifying the six properties (1)-(6).

Define a set $G^s$ by $\mathcal{G} \in G^s$ iff $\mathcal{G}$ is the natural filtration of some $B^s$, with $B \in B$ and $s \in [0, s_n/2]$. For fixed $B \in B$, when $s \downarrow 0$, the increasing limit of $\sigma(B^s)$ is $\sigma(B)$; hence, the set

$$\{\mathcal{G}\in G^s \mid \mathcal{G} \sigma(B^s), B \in B, s \in [0, s_n/2]\}$$

is substantial in $\mathcal{F}_\infty$. As every $B \in B$ is an $\mathcal{F}$-BM$^d$, each $\mathcal{G}$ is immersed in $\mathcal{F}$. Since $\mathcal{F}_{s_n} \vee \sigma(X^n) = \mathcal{F}_\infty$ by induction hypothesis, $(\mathcal{G}_{s_n} \vee \sigma(X^n)) \in \mathcal{F}_\infty$ is substantial in $\mathcal{F}_\infty$ by Lemma 4. Consequently, there exist $C \in B$ and $s_{n+1} \in [0, s_n/2]$ such that $\|R - U_{s_{n+1}}\| < 1/(n + 1)$ for some $U_{s_{n+1}} \in L^1(\sigma(C_{[s_{n+1}], X^n})$. Define $X^{n+1}$ by

$$dX^{n+1} = 1_{[s_{n+1}, s_n]}(t) dC_t + 1_{[0, s_0]}(t) dX^0_t.$$

Properties (1)-(4) hold for $s_{n+1}$ and $X^{n+1}$ by the very choice of these objects; (6) holds because $\sigma(C_{[s_{n+1}], X^n}) = \sigma(X^{n+1})$; last, (5) holds since, as $C \in B$, one has

$$\mathcal{F}_{s_{n+1}} \vee \sigma(X^{n+1}) = \mathcal{F}_{s_n} \vee \sigma(C_{[s_{n+1}], X^n}) \vee \sigma(X^n) = \mathcal{F}_{s_n} \vee \sigma(X^n) = \mathcal{F}_\infty.$$

References