Ljiljana Petrović; Sladjana Dimitrijević
Statistical causality and adapted distribution

Czechoslovak Mathematical Journal, Vol. 61 (2011), No. 3, 827–843

Persistent URL: http://dml.cz/dmlcz/141642

Czechoslovak Mathematical Journal, 61 (136) (2011), 827–843

STATISTICAL CAUSALITY AND ADAPTED DISTRIBUTION

Ljiljana Petrović, Beograd, Sladjana Dimitrijević, Kragujevac

(Received June 18, 2010)

Abstract. In the paper *D. Hoover, J. Keisler*: Adapted probability distributions, Trans. Amer. Math. Soc. *286* (1984), 159–201 the notion of adapted distribution of two stochastic processes was introduced, which in a way represents the notion of equivalence of those processes. This very important property is hard to prove directly, so we continue the work of Keisler and Hoover in finding sufficient conditions for two stochastic processes to have the same adapted distribution. For this purpose we use the concept of causality between stochastic processes, which is based on Granger's definition of causality. Also, we provide applications of our results to solutions of some stochastic differential equations.

Keywords: filtration, causality, adapted distribution, weak solution of stochastic differential equation

MSC 2010: 60G07, 03C98, 60H10

1. Introduction

Following the introduction, in the second section of this paper some preliminary definitions are illustrated. Then various concepts of causality relationship between flows of information (represented by filtrations) are considered. Also, a generalization of a causality relationship "**G** entirely causes **H** within **F**" is included which (in terms of σ-algebras) was first presented in [13] and is initially based on Granger's definition of causality (the nonlinear version) given in [6].

In the third section we present the different notions of equivalence of two stochastic processes outlined by Aldous, Keisler, Hoover and Fajardo (in [1], [8], [7], [9], [10], [11], [3]). We point out some connections between them and the given causality concepts from the second part of this paper.

Our main results are stated in the fourth section. We prove several results which link the given definition of causality with the concept of adapted distribution.

This research was supported by Science Fund of Serbia.

Two processes with the same adapted distribution share many probabilistic properties (being adapted, being a martingale, having the Markov property, being a local martingale and semimartingale, as was proved in [11]). We now prove that such processes share causality properties, too. Also, in [11] it was proved that for Markov processes to have the same adapted distribution is sufficient to have the same finite dimensional distribution. In this paper we prove sufficient conditions for having the same adapted distribution for a wider class of stochastic processes, which is defined by properties of causality.

The last section contains of several results connected with causality, the weak solutions of stochastic differential equations and the adapted distribution, which are corollaries of results from the previous section.

2. Causality and related concepts

Let $(\Omega, \mathscr{F}, P)$ be an arbitrary probability space and let $\mathbf{F} = \{\mathscr{F}_t,\ t \in I,\ I \subseteq \mathbb{R}^+\}$ be a family of sub-σ-algebras of $\mathscr{F}$. The sub-σ-algebra $\mathscr{F}_t$ can be interpreted as a set of events observed up to time t. Whether or not $\sup I = +\infty$ is true we define $\mathscr{F}_\infty$ as the smallest σ-algebra containing all the $\mathscr{F}_t$ (even when $\sup I < +\infty$). So, we have $\mathscr{F}_\infty = \bigvee_{t \in I} \mathscr{F}_t$. The filtration $\mathbf{F} = \{\mathscr{F}_t,\ t \in I\}$ is a nondecreasing family of σ-subalgebras of $\mathscr{F}$, that is

$$\mathscr{F}_s \subseteq \mathscr{F}_t, s \leqslant t.$$

A probabilistic model for a time-dependent system is described by $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$, where $(\Omega, \mathscr{F}, P)$ is a probability space and $\{\mathscr{F}_t,\ t \in I\}$ is a "framework" filtration, that is, $\mathscr{F}_t$ are all events in the model up to and including time t and $\mathscr{F}_t$ is a sub-σ-algebra of $\mathscr{F}$. We suppose that the filtration $(\mathscr{F}_t)$ satisfies the usual conditions‰, which means that $(\mathscr{F}_t)$ is right continuous and each $\mathscr{F}_t$ is complete.

Analogous notation will be used for filtrations $\mathbf{H} = \{\mathscr{H}_t,\ t \in I,\ I \subseteq \mathbb{R}^+\}$ and $\mathbf{G} = \{\mathscr{G}_t,\ t \in I,\ I \subseteq \mathbb{R}^+\}$.

The family of σ-algebras induced by a stochastic process $\mathbf{X} = \{X_t, t \in I, I \subseteq \mathbb{R}^+\}$ is given by $\mathbf{F}^X = \{\mathscr{F}_t^X, t \in I\}$, where

$$\mathscr{F}_t^X = \sigma\{X_u, u \in I,\ u \leqslant t\},$$

being the smallest σ-algebra with respect to which the random variables X_u, $u \leqslant t$, are measurable. The filtration $\mathbf{F}^X = \{\mathscr{F}_t^X, t \in I\}$ is called a natural filtration of the process $\mathbf{X} = \{X_t\}$.

The process $\mathbf{X} = \{X_t\}$ is $(\mathscr{F}_t)$-adapted (or adapted to the filtration $\mathbf{F}$) if all X_u, $u \leqslant t$, are $\mathscr{F}_t$-measurable, that is, if $\mathscr{F}_t^X \subseteq \mathscr{F}_t$ for each t. The notation $(X_t, \mathscr{F}_t)$ means that the process $\mathbf{X} = \{X_t\}$ is $(\mathscr{F}_t)$-adapted.

We now give some concepts of causality relationship between the flows of information (represented by filtrations) and between the stochastic processes. Also we provide some basic results that will be of use later.

Using the notion of conditional independence we introduce a statistical concept of causality which unifies the nonlinear Granger-causality (see [6]) with some related concepts given by Mykland (see [13]).

The intuitively plausible notion of causality formulated in terms of Hilbert spaces, is given in [15]. We shall use an analogous notion of causality in terms of filtrations (see [17]). Let **F**, **G** and **H** be arbitrary filtrations on the same probability space. We can say that "**G** entirely causes **H** within **F**" if $\mathscr{H}_\infty$ and $\mathscr{F}_t$ are conditionally independent when $\mathscr{G}_t$ is given, which we write in the following manner:

$$\mathscr{H}_\infty \perp \mathscr{F}_t \mid \mathscr{G}_t, \tag{2.1}$$

because the essence of (2.1) is that all information about $\mathscr{H}_\infty$ which $\mathscr{F}_t$ provides comes via $\mathscr{G}_t$ for an arbitrary t. In other words, $\mathscr{G}_t$ contains all information from the $\mathscr{F}_t$ needed for predicting $\mathscr{H}_\infty$. Relation (2.1) is equivalent to $\mathscr{H}_\infty \perp \mathscr{F}_t \vee \mathscr{G}_t \mid \mathscr{G}_t$. This relation means that the condition $\mathbf{G} \subseteq \mathbf{F}$ does not represent an essential restriction. Thus, it is natural to introduce the following definition of causality between filtrations.

Definition 2.1 (see [15]). Let **F**, **G** and **H** be arbitrary filtrations on the same probability space. It is said that **G** entirely causes (or briefly said "causes") **H** within **F** relative to P (written as $\mathbf{H} \mid\!< \mathbf{G}; \mathbf{F}; P$) if $\mathscr{H}_\infty \subseteq \mathscr{F}_\infty$, $\mathbf{G} \subseteq \mathbf{F}$ and if $\mathscr{H}_\infty$ is conditionally independent of $\mathscr{F}_t$ given $\mathscr{G}_t$ for each t, that is, if

$$\mathscr{H}_\infty \perp \mathscr{F}_t \mid \mathscr{G}_t \text{ for each } t, \tag{2.2}$$

or

$$(\forall A \in \mathscr{H}_\infty)\, P(A|\mathscr{F}_t) = P(A|\mathscr{G}_t).$$

If there is no doubt about P, we omit "relative to P".

Remark 2.2. The condition (2.2) is equivalent to $\mathscr{H}_u \perp \mathscr{F}_t \mid \mathscr{G}_t$ for each t and each u.

Intuitively, $\mathbf{H} \mid\!< \mathbf{G}; \mathbf{F}$ means that, for arbitrary t, information about $\mathscr{H}_\infty$ provided by $\mathscr{F}_t$ is not "bigger" than that provided by $\mathscr{G}_t$, or that it is possible to reduce the available information from $\mathscr{F}_t$ to $\mathscr{G}_t$ in order to predict $\mathscr{H}_\infty$.

A definition similar to Definition 2.1 was first given in [13]: "It is said that **G** entirely causes **H** within **F** relative to P (denoted as $\mathbf{H} \mid\!< \mathbf{G}; \mathbf{F}; P$) if $\mathbf{H} \subseteq \mathbf{F}$, $\mathbf{G} \subseteq \mathbf{F}$ and if $\mathscr{H}_\infty \perp \mathscr{F}_t \mid \mathscr{G}_t$ for each t". However, this definition (from [13]) contains the condition $\mathbf{H} \subseteq \mathbf{F}$, or equivalently $\mathscr{H}_t \subseteq \mathscr{F}_t$ for each t (instead of $\mathscr{H}_\infty \subseteq \mathscr{F}_\infty$) which does not have intuitive justification. Since Definition 2.1 is more general than the definition given in [13], all results resulting from Definition 2.1 will be true for the definition from [13], when we add the condition $\mathbf{H} \subseteq \mathbf{F}$ to them.

It should be mentioned that the definition of causality from [13] is equivalent to the definition of strong global noncausality as given in [4]. So, Definition 2.1 is a generalization of the notion of strong global noncausality.

If filtrations $\mathbf{G}$ and $\mathbf{F}$ are such that $\mathbf{G} \mid\!< \mathbf{G}; \mathbf{G} \vee \mathbf{F}$ (where $\mathbf{G} \vee \mathbf{F}$ is the family determined by $(\mathscr{G} \vee \mathscr{F})_t = \mathscr{G}_t \vee \mathscr{F}_t$), we shall say that $\mathbf{F}$ does not cause $\mathbf{G}$. Obviously the interpretation of Granger's causality is that $\mathbf{F}$ does not cause $\mathbf{G}$ if $\mathbf{G} \mid\!< \mathbf{G}; \mathbf{G} \vee \mathbf{F}$, because this relation means that we are not able to predict more precisely $\mathbf{G}$ using the additional information from $\mathbf{F}$ which does not exist in $\mathbf{G}$.

It can be shown, without difficulty, that this term and the term "$\mathbf{F}$ does not anticipate $\mathbf{G}$" (as introduced in [20]) are identical.

Definition 2.3 (see [13]). When filtrations $\mathbf{G}$ and $\mathbf{F}$ are such that $\mathbf{G} \mid\!< \mathbf{G}; \mathbf{F}$, we say that $\mathbf{G}$ is its own cause within $\mathbf{F}$.

Note that the notion of subordination (as introduced in [19]) is equivalent to the notion of being one's own cause, as defined here. Also "$\mathbf{G}$ is its own cause" sometimes occurs as a useful assumption in the theory of martingales and stochastic integration (see [2], [21], [17] and [22]).

These definitions can be applied to stochastic processes if we are talking about the corresponding induced filtrations. For example, an $(\mathscr{F}_t)$-adapted stochastic process $\mathbf{X} = \{X_t\}$ is its own cause if $(\mathscr{F}_t^X)$ is its own cause within $(\mathscr{F}_t)$, that is if

$$\mathbf{F}^X \mid\!< \mathbf{F}^X; \mathbf{F}; P.$$

The extensions of the definitions to vector processes are usually straightforward.

The following result shows that a process $\mathbf{X}$ which is its own cause is completely described by its behavior relative to $\mathbf{F}^X$.

Proposition 2.4. *A process $\mathbf{X} = \{X_t, t \in [0,T]\}$ is a Markov process relative to the filtration $\mathbf{F} = \{\mathscr{F}_t, t \in [0,T]\}$ on a filtered probability space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ if and only if $\mathbf{X}$ is a Markov process relative to the filtration $\mathbf{F}^X$ and the process $\mathbf{X}$ is its own cause within $\mathbf{F}$ relative to P.*

Proof. Let $\mathbf{X}$ be a Markov process relative to $\mathbf{F}$, that is, let

$$P(A \cap B | X_t) = P(A|X_t) P(B|X_t) \text{ a.s.}$$

for each $t \in [0,T]$, $A \in \mathscr{F}_t$, $B \in \mathscr{F}_{[t,\infty)} = \sigma\{X_s, s \geqslant t\}$, hold. Then we have

$$(\forall t)\,(\forall A \in \mathscr{F}_\infty^X),\ E(\chi_A|\mathscr{F}_t) = E(\chi_A|X_t) \text{ a.s.}$$

Now, from $\mathbf{F}^X \subseteq \mathbf{F}$ we have

$$(\forall t)\,(\forall A \in \mathscr{F}_\infty^X),\ P(A|\mathscr{F}_t) = P(A|\mathscr{F}_t^X) \text{ a.s.}$$

Therefore, we conclude that

$$\mathbf{F}^X \mid\!< \mathbf{F}^X; \mathbf{F}; P\,,$$

and it is clear that $\mathbf{X}$ is a Markov process relative to $\mathbf{F}^X$.

It is easy to see that the converse is true. □

Corollary 2.5. *The Brownian motion $\mathbf{W} = \{W_t, \mathscr{F}_t, t \in [0,T]\}$ on a filtered probability space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ is its own cause within the filtration $\mathbf{F} = \{\mathscr{F}_t, t \in [0,T]\}$ relative to the probability P.*

We shall give a few properties of the causality relationship from Definition 2.1 which we will need later.

Proposition 2.6 (see [16]). *From $\mathbf{H} \mid\!< \mathbf{G}; \mathbf{F}$ and $\mathbf{H} \subseteq \mathbf{F}$ it follows that $\mathbf{H} \subseteq \mathbf{G}$.*

It follows from the following result that the relationship "being one's own cause" is a transitive relationship.

Proposition 2.7 (compare with [5] and [16]). *From $\mathbf{H} \mid\!< \mathbf{H}; \mathbf{G}$ and $\mathbf{G} \mid\!< \mathbf{G}; \mathbf{F}$ it follows that $\mathbf{H} \mid\!< \mathbf{H}; \mathbf{F}$.*

The following result gives the invariance under convergence for the causality relationship from Definition 2.1.

Proposition 2.8 ([13]). *Let $\mathbf{F}$ and $\mathbf{G}$ be filtrations on the probability space $(\Omega, \mathscr{F}, P)$. Let $\{X^{(n)}\}$ be a sequence of stochastic processes satisfying*

$$X_t^{(n)} \xrightarrow{P} X_t, \text{ when } n \to +\infty, \quad \text{for every } t \in I \subset \mathbb{R}$$

and

$$\mathbf{F}^{X^{(n)}} \mid\!< \mathbf{G};\ \mathbf{F}, \quad \text{for every } n.$$

Then the process X satisfies

$$\mathbf{F}^{\mathbf{X}} \mid\!< \mathbf{G};\ \mathbf{F}.$$

3. Adapted distribution

In this section we consider several notions of equivalence of two stochastic processes originally given by Aldous, Keisler, Hoover and Fajardo (in [1], [8], [7], [9], [11], [10], [3]). After that we link these notions to the causality relationship which we defined earlier.

If two stochastic variables have the same distribution then that is a very strong notion of equivalence of those two variables. But, when two processes have the same distribution we know much less about them, because we do not have any information about those properties of processes which take into account the relations of the processes to their underlying filtrations (for instance being adapted, being a martingale). Therefore it is natural to make an attempt to discover more general notions of equivalence for processes which involve filtrations.

Aldous introduced in [1] the weakest of those notions—the synonymity of two processes.

Definition 3.1 ([8]). Let $(X_t, \mathscr{F}_t)_{t\in\mathbb{R}^+}$ be a stochastic process on a stochastic base Ω and let $(Y_t, \mathscr{G}_t)_{t\in\mathbb{R}^+}$ be another process on a possibly different base. We say $\mathbf{X} = \{X_t\}$ and $\mathbf{Y} = \{Y_t\}$ are synonymous, and write $\mathbf{X} \equiv_1 \mathbf{Y}$, if and only if for any $n \in \mathbb{N}$, any $t_1, \dots, t_n, u_1, \dots, u_n \geqslant 0$, and any bounded Borel functions $\varphi, \varphi_1, \dots, \varphi_n$: $\mathbb{R}^n \to \mathbb{R}$ we have

$$E[\varphi(E[\varphi_1(X_{u_1}, \dots, X_{u_n})|\mathscr{F}_{t_1}], \dots, E[\varphi_n(X_{u_1}, \dots, X_{u_n})|\mathscr{F}_{t_n}])]$$
$$= E[\varphi(E[\varphi_1(Y_{u_1}, \dots, Y_{u_n})|\mathscr{G}_{t_1}], \dots, E[\varphi_n(Y_{u_1}, \dots, Y_{u_n})|\mathscr{G}_{t_n}])].$$

In [1] it is shown that several important properties of stochastic processes (being adapted, being a martingale, and having the Markov property) are preserved under the synonymity relation $\mathbf{X} \equiv_1 \mathbf{Y}$. However, the same cannot be concluded for any property of stochastic processes.

Keisler and Hoover introduced in [11] a stronger notion—two processes have the same adapted distribution or have the same adapted law. Their thesis was that two processes with the same adapted distribution share the same probabilistic properties. They proved that fact for all the most interesting probabilistic properties (being adapted, being a martingale, having the Markov property, being a local martingale and semimartingale). The most powerful property of $\equiv_{\mathrm{AD}}$ which they proved is the existence of spaces with a saturation property, which fails to be the case with synonymity.

Some preliminary definitions are given before the definition of the adapted distribution. Let M be a Polish space (complete separable metrizable topological space) which remains fixed throughout our discussion.

Definition 3.2 ([11]). For each n, each bounded continuous function $\Phi\colon M^n \to \mathbb{R}$ and each stochastic process $\mathbf{X}$ with values in M, $\hat{\Phi}\mathbf{X}$ is the n-fold stochastic process

$$\hat{\Phi}\mathbf{X}(t_1, \ldots, t_n) = \Phi(X_{t_1}, \ldots, X_{t_n}).$$

Remark 3.3. Two stochastic processes $\mathbf{X}$ and $\mathbf{Y}$ have the same finite dimensional distribution if and only if

$$E[\hat{\Phi}\mathbf{X}(t_1, \ldots, t_n)] = E[\hat{\Phi}\mathbf{Y}(t_1, \ldots, t_n)]$$

for all Φ and all $t_1, \ldots, t_n$.

The finite dimensional distribution of a process $\mathbf{X}$ depends only on $(\Omega, P, \mathbf{X})$ and not on the filtration $(\mathscr{F}_t)$. The next notion depends strongly on the filtration $(\mathscr{F}_t)$.

The class CP is a family of functions f, called conditional processes, which associate with each stochastic process $\mathbf{X}$ on Ω an n-fold stochastic process $f\mathbf{X}$ on Ω.

Definition 3.4 ([11]). The class CP of conditional processes (in M) is defined inductively as follows:

(i) (Basis) For each n and bounded continuous $\Phi\colon M^n \to \mathbb{R}$, $\hat{\Phi} \in$ CP.

(ii) (Composition) If $f_1, \ldots, f_n \in$ CP and $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ is a bounded continuous function, then $\varphi(f_1, \ldots, f_n) \in$ CP, where $\varphi(f_1, \ldots, f_n)\mathbf{X} = \varphi(f_1\mathbf{X}, \ldots, f_n\mathbf{X})$.

(iii) (Conditional Expectation) If f is an n-fold conditional process $f\mathbf{X}(t_1, \ldots, t_n)$ then $E[f|t]$ is an $(n+1)$-fold conditional process, where $E[f|t]\mathbf{X}(t, t_1, \ldots, t_n)$ is a version of $E[f\mathbf{X}(t_1, \ldots, t_n)|\mathscr{F}_t]$.

The number of iterations of the unexpected value operation in f is called the rank of f.

Definition 3.5 ([11]). The rank of conditional processes is defined by:

(i) for each n and Φ, the conditional process $\hat{\Phi}$ has rank zero;

(ii) the rank of the composition $\varphi(f_1, \ldots, f_n)$ is the maximum of the ranks of the conditional processes $f_1, \ldots, f_n$;

(iii) if f is a conditional process of rank r, then $E[f|t]$ is a conditional process of rank $r + 1$.

Now we are ready to introduce the main notion.

Definition 3.6 ([11]). Two stochastic processes $\mathbf{X}$ and $\mathbf{Y}$ (on perhaps different adapted spaces) have the same adapted distribution (or adapted law), in symbols $\mathbf{X} \equiv_{AD} \mathbf{Y}$, if

$$E[f\mathbf{X}(t_1, \ldots, t_n)] = E[f\mathbf{Y}(t_1, \ldots, t_n)] \tag{3.1}$$

holds for every n-fold conditional process f and all $(t_1, \ldots, t_n) \in (\mathbb{R}^+)^n$.

Definition 3.7 ([11]). Processes $\mathbf{X}$ and $\mathbf{Y}$ have the same adapted distribution up to rank r, in symbols $\mathbf{X} \equiv_r \mathbf{Y}$, if (3.1) holds for every f of rank at most r and all $(t_1, \dots, t_n) \in (\mathbb{R}^+)^n$.

Remark 3.8. Notice that $\mathbf{X} \equiv_0 \mathbf{Y}$ means that the processes $\mathbf{X}$ and $\mathbf{Y}$ have the same finite dimensional distribution.

Of course, $\mathbf{X} \equiv_{\mathrm{AD}} \mathbf{Y}$ implies $\mathbf{X} \equiv_r \mathbf{Y}$ for every r, and the reverse is not true.

Several very interesting notions concerning stochastic processes and filtrations given by Hoover in [10], have many similarities with our notion of causality from Definition 2.1.

Definition 3.9 ([10]). A subfiltration $\mathbf{G}$ of a filtration $\mathbf{F}$ is self-contained in $\mathbf{F}$ ($\mathbf{F}$ is an extension of $\mathbf{G}$) if for each $t \in \mathbb{R}^+$ the σ-algebras $\mathscr{G}_\infty$ and $\mathscr{F}_t$ are conditionally independent given $\mathscr{G}_t$, that is, if

$$(\forall A \in \mathscr{G}_\infty)\, P(A|\mathscr{F}_t) = P(A|\mathscr{G}_t).$$

In terms of causality, the notion *self-contained* is analogous to the notion *be its own cause.*

We have already mentioned that this relation arises in many parts of the stochastic theory with different names applied by different authors.

In the paper [10] the notion *intrinsic filtration of* $\mathbf{X}$ is defined.

If $\mathbf{X} = (X, \mathbf{F})$ is a random variable with filtration, then there exists a smallest self-contained subfiltration $\mathbf{I}(\mathbf{X})$ of $\mathbf{F}$ such that X is $\mathscr{I}(\mathbf{X})_\infty$-measurable. That filtration $\mathbf{I}(\mathbf{X})$ we call the intrinsic filtration of $\mathbf{X}$.

In terms of causality, the intrinsic filtration of a process $\mathbf{X}$, defined on $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$, means the smallest filtration which is its own cause and which entirely causes the natural filtration $\mathbf{F}^X$ of the process $\mathbf{X}$ within $(\mathscr{F}_t)$.

In the paper [10] (Theorem 2.3) it was proved that processes $\mathbf{X}$ and $\mathbf{Y}$ have the same adapted distribution if and only if there is a filtration isomorphism $h\colon \mathbf{I}(\mathbf{X}) \to \mathbf{I}(\mathbf{Y})$ such that $h(X) = Y$.

Often, especially when we are solving same stochastic differential equation and trying to find a weak solution, we must enlarge the probability space. In [10] Hoover gave one very useful definition of extension of a probability space with filtration $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$.

Definition 3.10 ([10]). An extension of a probability space with filtration $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ is a space $(\dot{\Omega}, \mathscr{G}, \mathscr{G}_t, Q)$ satisfying:

(i) $\dot{\Omega} = \Lambda^1 \times \Omega \times \Lambda^2$ for some sets Λ^1, Λ^2, and $\mathscr{G} = \mathscr{H}^1 \times \Omega \times \mathscr{H}^2$, $\mathbf{G} = (\mathscr{G}_t)$ is the smallest filtration such that for each t, $\mathscr{G}_i \supseteq \mathscr{H}_t^1 \times F_t \times \mathscr{H}_t^2$, for some $\mathscr{H}_t^i$ σ-algebras and $\mathbf{H}^i$ filtrations on Λ^i, $i = 1, 2$;

(ii) for each $F \in \mathscr{F}$, $\dot{F} = \Lambda^1 \times F \times \Lambda^2 \in \mathscr{G}$ and $Q(\dot{F}) = P(F)$;
(iii) for all s, $\dot{\mathscr{F}}_\infty$ and $\mathscr{G}_s$ are conditionally independent given $\dot{\mathscr{F}}_s$.

The purpose of this definition is to make sure that an induced process $\dot{\mathbf{X}}$ on $\dot{\Omega}$, given by

$$\dot{X}(\lambda_1, \omega, \lambda_2) = X(\omega),$$

has the same properties relative to the filtration $\mathbf{G} = (\mathscr{G}_t)$ as $\mathbf{X}$ has relative to the filtration $\mathbf{F} = (\mathscr{F}_t)$, that is, $(\mathbf{X}, \mathbf{F}) \equiv_{\text{AD}} (\dot{\mathbf{X}}, \mathbf{G})$.

The essence of this definition is condition (iii) which means $\dot{\mathbf{F}} \models\!< \dot{\mathbf{F}}; \mathbf{G}; Q$, ($\dot{\mathbf{F}}$ is its own cause within $\mathbf{G}$), and we explained earlier the effects of this property.

4. Causality and adapted distribution

We have already pointed out similarity in defining the adapted distribution and the causality. Now we give some connections between these two notions.

Theorem 4.1. *Let $\mathbf{X} = \{X_t\}$ be a process on $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ and let $\mathbf{G} = (\mathscr{G}_t)$ be subfiltration of $\mathbf{F} = (\mathscr{F}_t)$. Then*

$$\mathbf{F}^X \models\!< \mathbf{G}; \mathbf{F} \text{ if and only if } (\mathbf{X}, \mathbf{G}) \equiv_{\text{AD}} (\mathbf{X}, \mathbf{F}).$$

Proof. From $(\mathbf{X}, \mathbf{G}) \equiv_{\text{AD}} (\mathbf{X}, \mathbf{F})$ it follows that

$$(\forall A \in \mathscr{F}^X_\infty)\, E(I_A|\mathscr{F}_t) = E(I_A|\mathscr{G}_t),$$

that is,

$$(\forall A \in \mathscr{F}^X_\infty)\, P(A|\mathscr{F}_t) = P(A|\mathscr{G}_t),$$

or simply

$$\mathbf{F}^X \models\!< \mathbf{G}; \mathbf{F}.$$

On the other hand, from $\mathbf{F}^X \models\!< \mathbf{G}; \mathbf{F}$ we have

$$(\forall A \in \mathscr{F}^X_\infty)\, P(A|\mathscr{F}_t) = P(A|\mathscr{G}_t),$$

that is,

$$(\forall A \in \mathscr{F}^X_\infty)\, E(I_A|\mathscr{F}_t) = E(I_A|\mathscr{G}_t).$$

Then for every simple (step) function $f_N(\omega) = \sum_{n=1}^{N} c_n I_{A_n}(\omega)$ we also have

$$E(f_N|\mathscr{F}_t) = E(f_N|\mathscr{G}_t) < E(f_N) < +\infty.$$

For every bounded continuous function f there exists a sequence of simple (step) functions (f_N) which converges almost surely to it. Now, when $N \to +\infty$, by Lebesgue's dominated convergence theorem, for every bounded continuous function f we have

$$E(f|\mathscr{F}_t) = E(f|\mathscr{G}_t). \tag{4.1}$$

From the equality (4.1) and the definition of the n-fold conditional process f it follows that (4.1) is true for every n-fold conditional process f, as well. □

Corollary 4.2. *Let* $\mathbf{X}$ *be a process on* $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ *and let* $\mathbf{G}$ *and* $\mathbf{H}$ *be subfiltrations of* $\mathbf{F}$ *such that* $\mathbf{G} \subseteq \mathbf{H} \subseteq \mathbf{F}$. *Then from* $\mathbf{F}^X \mid< \mathbf{G}; \mathbf{F}$ *it follows that* $(\mathbf{X}, \mathbf{G}) \equiv_{\mathrm{AD}} (\mathbf{X}, \mathbf{H})$.

Theorem 4.3. *Let* $\mathbf{X}$ *be a process on* $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$, *let the probability space* $(\dot{\Omega}, \mathscr{G}, \mathscr{G}_t, Q)$ *be an extension of the space* $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ *in the sense of Definition* 3.10 *and let* $\dot{\mathbf{X}}$ *be the induced process (on* $(\dot{\Omega}, \mathscr{G}, \mathscr{G}_t, Q)$ *) of the process* $\mathbf{X}$. *Then*

$$\mathbf{F}^X \mid< \mathbf{F}^X; \mathbf{F}; P \quad \textit{implies} \quad \mathbf{F}^{\dot{X}} \mid< \mathbf{F}^{\dot{X}}; \mathbf{G}; Q.$$

P r o o f. Because of $\mathbf{F}^X \mid< \mathbf{F}^X; \mathbf{F}; P$ we have $\mathbf{I}(\mathbf{X}) = \mathbf{F}^X$. Since the probability space $(\dot{\Omega}, \mathscr{G}, \mathscr{G}_t, Q)$ is an extension of the space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ we have $(\mathbf{X}, \mathbf{F}) \equiv_{\mathrm{AD}} (\dot{\mathbf{X}}, \mathbf{G})$. So, there is a filtration isomorphism $h\colon \mathbf{F}^X \to \mathbf{I}(\dot{\mathbf{X}})$. Further, have $h(\mathbf{F}^X) = \mathbf{F}^{\dot{X}}$. This implies $\mathbf{I}(\dot{\mathbf{X}}) = \mathbf{F}^{\dot{X}}$, or equivalently, $\mathbf{F}^{\dot{X}} \mid< \mathbf{F}^{\dot{X}}; \mathbf{G}; Q$. □

It has already been proved in [10] and [11] that many properties of stochastic processes (probably all of some significance) are preserved under the relation $\equiv_{\mathrm{AD}}$. We prove that the properties of causality are preserved, too.

Theorem 4.4. *Let* $\mathbf{X}$ *be a process on* $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$; $\mathbf{Y}$ *a process on* $(\bar{\Omega}, \mathscr{G}, \mathscr{G}_t, Q)$, *and* $(\mathbf{X}, \mathbf{F}) \equiv_{\mathrm{AD}} (\mathbf{Y}, \mathbf{G})$. *Then*

$$\mathbf{F}^X \mid< \mathbf{F}^X; \mathbf{F}; P \quad \textit{implies} \quad \mathbf{F}^Y \mid< \mathbf{F}^Y; \mathbf{G}; Q.$$

P r o o f. From $(\mathbf{X}, \mathbf{F}) \equiv_{\mathrm{AD}} (\mathbf{Y}, \mathbf{G})$ we have that there is a filtration isomorphism $h\colon \mathbf{I}(\mathbf{X}) \to \mathbf{I}(\mathbf{Y})$. From $\mathbf{F}^X \mid< \mathbf{F}^X; \mathbf{F}; P$ it follows that $\mathbf{I}(\mathbf{X}) = \mathbf{F}^X$. Then, according to the Amalgamation Theorem (Theorem 3.2 in [10]), there exists a common extension $(\widetilde{\Omega}, \widetilde{\mathscr{H}}, \widetilde{\mathscr{H}_t}, \widetilde{P})$ of spaces $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ and $(\bar{\Omega}, \mathscr{G}, \mathscr{G}_t, Q)$. On that extension there are processes $h_1(\mathbf{X}) = \dot{\mathbf{X}} = (\dot{X}, \widetilde{\mathbf{H}})$ and $h_2(\mathbf{Y}) = \dot{\mathbf{Y}} = (\dot{Y}, \widetilde{\mathbf{H}})$ such that

$$(\dot{\mathbf{X}}, \widetilde{\mathbf{H}}) \equiv_{\mathrm{AD}} (\mathbf{X}, \mathbf{F}) \quad \text{and} \quad (\dot{\mathbf{Y}}, \widetilde{\mathbf{H}}) \equiv_{\mathrm{AD}} (\mathbf{Y}, \mathbf{G}).$$

From the last two relations and $(\mathbf{X},\mathbf{F}) \equiv_{\mathrm{AD}} (\mathbf{Y},\mathbf{G})$ we have that $(\dot{\mathbf{X}},\widetilde{\mathbf{H}}) \equiv_{\mathrm{AD}} (\dot{\mathbf{Y}},\widetilde{\mathbf{H}})$ and there is a filtration isomorphism $\dot{h}\colon\ \mathbf{I}(\dot{\mathbf{X}}) \to \mathbf{I}(\dot{\mathbf{Y}})$. Now, for every set $A_Y \in \mathscr{F}^Y_\infty$ we have

$$\begin{aligned} Q(A_Y|\mathscr{G}_t) &= \widetilde{P}(\dot{A}_Y|\widetilde{\mathscr{H}_t}) = \widetilde{P}(\dot{A}_X|\widetilde{\mathscr{H}_t}) \\ &= P(A_X|\mathscr{F}_t) = P(A_X|\mathscr{F}^X_t) \\ &= Q(h(A_X)|\mathscr{F}^Y_t) = Q(A_Y|\mathscr{F}^Y_t), \end{aligned}$$

and we conclude that $\mathbf{F}^Y \mid\!< \mathbf{F}^Y;\mathbf{G};Q$. □

The next theorem follows directly from the Adjunction Theorem (Theorem 3.3) in [10].

Theorem 4.5. *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be processes defined on a filtered probability space* $(\Omega,\mathscr{F},\mathscr{F}_t,P)$ *and let* $\mathbf{F}^X \mid\!< \mathbf{F}^Y;\mathbf{F};P$ *hold. Let* $\overline{\mathbf{X}}$ *be a process defined on* $(\overline{\Omega},\overline{\mathscr{F}},\overline{\mathscr{F}_t},\overline{P})$ *and* $(\mathbf{X},\mathbf{F}) \equiv_{\mathrm{AD}} (\overline{\mathbf{X}},\overline{\mathbf{F}})$*. Then there exists an extension* $(\dot{\Omega},\dot{\mathscr{F}},\dot{\mathscr{F}}_t,\dot{P})$ *of the space* $(\overline{\Omega},\overline{\mathscr{F}},\overline{\mathscr{F}_t},\overline{P})$ *and processes* $\dot{\mathbf{X}}$ *and* $\dot{\mathbf{Y}}$ *defined on it such that* $\mathbf{F}^{\dot{X}} \mid\!< \mathbf{F}^{\dot{Y}};\dot{\mathbf{G}};\dot{Q}$ *and* $(\mathbf{X},\mathbf{Y},\mathbf{F}) \equiv_{\mathrm{AD}} (\dot{\mathbf{X}},\dot{\mathbf{Y}},\dot{\mathbf{G}})$.

One of the aims of the paper [11] was to find classes of processes for which we can get the adapted distribution under a weaker condition. We have proved one result of that type (Theorem 4.7), too. That is our main result. But first we shall prove the next lemma, which is needed for proving Theorem 4.7.

Lemma 4.6. *Let* $\mathbf{X} = (X_t)$ *and* $\mathbf{Y} = (Y_t)$ *be stochastic processes. If for every bounded continuous function* $\overline{h}\colon\ M^n \to \mathbb{R}$*, for every* $n \in \mathbb{N}$ *and for every* $(t_1,t_2,\ldots,t_n) \in (\mathbb{R}^+)^n$ *the equality*

$$E[\overline{h}(X_{t_1},X_{t_2},\ldots,X_{t_n})] = E[\overline{h}(Y_{t_1},Y_{t_2},\ldots,Y_{t_n})]$$

holds, then

$$E[h(X_{t_1},X_{t_2},\ldots,X_{t_n},\ldots)] = E[h(Y_{t_1},Y_{t_2},\ldots,Y_{t_n},\ldots)]$$

holds for every bounded continuous function $h\colon\ M^{\mathbb{N}} \to \mathbb{R}$ *and for every sequence* $(t_1,t_2,\ldots,t_n,\ldots) \in (\mathbb{R}^+)^{\mathbb{N}}$.

P r o o f. We shall use a shorter notation

$$h^X = h(X_{t_1},X_{t_2},\ldots,X_{t_n},\ldots), \quad h^Y = h(Y_{t_1},Y_{t_2},\ldots,Y_{t_n},\ldots)$$

and

$$h^X_n = h(X_{t_1},X_{t_2},\ldots,X_{t_n},0,0,\ldots), \quad h^Y_n = h(Y_{t_1},Y_{t_2},\ldots,Y_{t_n},0,0,\ldots).$$

Because the function h is bounded, there is a constant M such that for every n

$$h_n^X \leqslant M \quad \text{and} \quad h_n^Y \leqslant M.$$

Also, we have almost sure convergence of the sequence $(X_{t_1}, X_{t_2}, \ldots, X_{t_n}, 0, 0, \ldots)$, that is,

$$x_n = (X_{t_1}, X_{t_2}, \ldots, X_{t_n}, 0, 0, \ldots) \xrightarrow{\text{a.s.}} x = (X_{t_1}, X_{t_2}, \ldots, X_{t_n}, \ldots), \quad n \to +\infty.$$

For the process $\mathbf{Y}$ the same holds. Since the function h is continuous we have, also, the almost sure convergence of the sequences (h_n^X) and (h_n^Y), that is,

$$h_n^X \xrightarrow{\text{a.s.}} h^X, \quad n \to +\infty,$$

and

$$h_n^Y \xrightarrow{\text{a.s.}} h^Y, \quad n \to +\infty.$$

Now, by Lebesgue's dominated convergence theorem, it follows that

$$\lim_{n\to+\infty} E[h_n^X] = E[h^X] \quad \text{and} \quad \lim_{n\to+\infty} E[h_n^Y] = E[h^Y]. \tag{4.2}$$

We define functions $\overline{h}_n^X : M^n \to \mathbb{R}$, $\overline{h}_n^Y : M^n \to \mathbb{R}$ by

$$\begin{aligned} \overline{h}_n^X(X_{t_1}, X_{t_2}, \ldots, X_{t_n}) &= h_n^X = h(X_{t_1}, X_{t_2}, \ldots, X_{t_n}, 0, 0, \ldots), \\ \overline{h}_n^Y(Y_{t_1}, Y_{t_2}, \ldots, Y_{t_n}) &= h_n^Y = h(Y_{t_1}, Y_{t_2}, \ldots, Y_{t_n}, 0, 0, \ldots). \end{aligned}$$

Now, because of $\mathbf{X} \equiv_0 \mathbf{Y}$, for every n we have

$$E[h_n^X] = E[\overline{h}_n^X(X_{t_1}, X_{t_2}, \ldots, X_{t_n})] = \overline{h}_n^Y(Y_{t_1}, Y_{t_2}, \ldots, Y_{t_n})] = E[h_n^Y].$$

From this equality and (4.2) we get

$$E[h^X] = E[h^Y].$$

□

Theorem 4.7. *Let a process* $\mathbf{X}$ *be defined on a probability space* $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ *and let* $\mathbf{X}$ *be its own cause within the filtration* $\mathbf{F} = (\mathscr{F}_t)$, *that is* $\mathbf{F}^X \mid\!< \mathbf{F}^X; \mathbf{F}; P$, *and let a process* $\mathbf{Y}$ *be defined on a probability space* $(\overline{\Omega}, \mathscr{G}, \mathscr{G}_t, Q)$ *and let* $\mathbf{Y}$ *be its own cause within the filtration* $\mathbf{G} = (\mathscr{G}_t)$, *that is,* $\mathbf{F}^Y \mid\!< \mathbf{F}^Y; \mathbf{G}; Q$. *Then* $\mathbf{X} \equiv_{\mathrm{AD}} \mathbf{Y}$ *if and only if* $\mathbf{X} \equiv_1 \mathbf{Y}$.

P r o o f. Implication $\mathbf{X} \equiv_{\mathrm{AD}} \mathbf{Y} \Rightarrow \mathbf{X} \equiv_1 \mathbf{Y}$ is obvious.

To prove the converse we shall show, by induction, that for every n-fold conditional process f and every $\vec{t} \in (\mathbb{R}^+)^{\mathbb{N}}$ there is a bounded Borel function $\psi_{f,\vec{t}}\colon M^{\mathbb{N}} \to \mathbb{R}$ such that

$$f\mathbf{X}(\vec{t}) = \psi_{f,\vec{t}}(X_{t_1}, \ldots, X_{t_n}, \ldots) \ \text{ a.s.} \tag{4.3}$$

and

$$f\mathbf{Y}(\vec{t}) = \psi_{f,\vec{t}}(Y_{t_1}, \ldots, Y_{t_n}, \ldots) \ \text{ a.s.} \tag{4.4}$$

For each bounded continuous function $\Phi\colon M^{\mathbb{N}} \to \mathbb{R}$, the function $\psi_{\hat{\Phi},\vec{t}} = \Phi$ has the required properties (4.3) and (4.4). If $f \in \mathrm{CP}$ is of the form $f = \varphi(f_1, \ldots, f_m)$, we take

$$\psi_{f,\vec{t}} = \varphi(\psi_{f_1,\vec{t}}, \ldots, \psi_{f_m,\vec{t}}).$$

This takes care of the basis step and the composition step in the induction. For the conditional expectation step, let $g = E[f|s]$ where f is an n-fold conditional process and suppose $s \in \mathbb{R}^+$, $\vec{t} \in (\mathbb{R}^+)^{\mathbb{N}}$ and $\psi_{f,\vec{t}}$ satisfies (4.3). Since the process $\mathbf{X}$ is its own cause within the filtration $\mathbf{F}$ and $\psi_{f,\vec{t}}$ is bounded, we have

$$E[\psi_{f,\vec{t}}(X_{t_1}, \ldots, X_{t_n}, \ldots)|\mathscr{F}_s] = E[\psi_{f,\vec{t}}(X_{t_1}, \ldots, X_{t_n}, \ldots)|\mathscr{F}_s^X].$$

Because $E[\psi_{f,\vec{t}}(X_{t_1}, \ldots, X_{t_n}, \ldots)|\mathscr{F}_s^X]$ is measurable with respect to $\mathscr{F}_s^X$, we have

$$E[\psi_{f,\vec{t}}(X_{t_1}, \ldots, X_{t_n}, \ldots)|\mathscr{F}_s^X] = h(X_{s_1}, \ldots, X_{s_n}, \ldots) \ \text{ a.s.}$$

where $s_1, s_2, \ldots, s_n, \ldots \in [0, s]$ and h is a $\mathscr{B}_s$-measurable function (see [12], p. 22). Therefore there is a bounded Borel function $\psi_{g,\vec{s}}$ such that

$$g\mathbf{X}(s\vec{t}) = E[\psi_{f,\vec{t}}(X_{t_1}, \ldots, X_{t_n}, \ldots)|\mathscr{F}_s] = \psi_{g,\vec{s}}(X_{s_1}, \ldots, X_{s_n}, \ldots). \tag{4.5}$$

From $\mathbf{X} \equiv_1 \mathbf{Y}$ and the causality relationships $\mathbf{F}^X \mid\!< \mathbf{F}^X; \mathbf{F}; P$ and $\mathbf{F}^Y \mid\!< \mathbf{F}^Y; \mathbf{G}; Q$ it follows that

$$E[h(X_{t_1}, \ldots, X_{t_n}, \ldots)|\mathscr{F}_s^X] = E[h(Y_{t_1}, \ldots, Y_{t_n}, \ldots)|\mathscr{F}_s^Y].$$

Due to the previous equality ($\mathbf{X}$ and $\mathbf{Y}$ have the same transition function), equalities (4.4) and (4.5) hold for $\mathbf{Y}$ with the same choice of $\psi_{g,\vec{s}}$. This completes the induction. Finally, using Lemma 4.6, we get that

$$E[f\mathbf{X}(\vec{t})] = E[\psi_{f,\vec{t}}(X_{t_1},\ldots,X_{t_n},\ldots)] = E[\psi_{f,\vec{t}}(Y_{t_1},\ldots,Y_{t_n},\ldots)] = E[f\mathbf{Y}(\vec{t})]$$

is true for all $f \in \mathrm{CP}$ and all $\vec{t}$, that is, $\mathbf{X} \equiv_{\mathrm{AD}} \mathbf{Y}$. □

In [11] it was proved that for Markov processes $\mathbf{X}$ and $\mathbf{Y}$, from $\mathbf{X} \equiv_0 \mathbf{Y}$ it follows that $\mathbf{X} \equiv_{\mathrm{AD}} \mathbf{Y}$. The previous theorem gives us a similar conclusion for a wider class of stochastic processes, processes which are their own cause in their spaces. For such two processes $\mathbf{X}$ and $\mathbf{Y}$, $\mathbf{X} \equiv_1 \mathbf{Y}$ implies $\mathbf{X} \equiv_{\mathrm{AD}} \mathbf{Y}$.

5. Applications to stochastic differential equations

Now, we will consider some stochastic differential equations, and give a few results about their weak solutions and the adapted distribution.

In general, when we solve some equation on a specifically given filtered probability space with a known driving process we try to find a strong solution of that equation. But if for some equation we have only prescribed nonanticipative functionals and we try to find the filtered probability space in which the solution process $\mathbf{X}$ and the driving process exist which satisfy the equation, we speak about a weak solution.

Let $\mathbf{B}^H = \{B_t^H,\, t \in [0,T]\}$ be a fractional Brownian motion with Hurst parameter $H \in (0,1)$ defined on a probability space $(\Omega, \mathscr{F}, P)$. That is, $\mathbf{B}^H$ is a centered Gaussian process with covariance

$$R_H(t,s) = E(B_t^H B_s^H) = \frac{1}{2}\left\{|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right\}.$$

If $H = \frac{1}{2}$ the process $\mathbf{B}^H$ is the standard Brownian motion.

For each $t \in [0,T]$ we denote by $\mathscr{F}_t^{B^H}$ the σ-algebra generated by random variables B_s^H, $s \in [0,t]$ and the sets of probability zero. So, $\mathbf{F}^{B^H} = \{\mathscr{F}_t^{B^H}, t \in [0,T]\}$ is the natural filtration of the fractional Brownian motion $\mathbf{B}^H = \{B_t^H,\, t \in [0,T]\}$.

Consider the stochastic differential equation (see [14])

$$X_t = x + B_t^H + \int_0^t b(s, X_s)\,\mathrm{d}s,\ \ t \in [0,T] \tag{5.1}$$

where b is a Borel function on $[0,T] \times \mathbb{R}$.

By a weak solution of equation (5.1) we mean a couple of adapted continuous processes $(\mathbf{B}^H, \mathbf{X})$ on a filtered probability space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$, such that:

(i) $\mathbf{B}^H$ is a $(\mathscr{F}_t)$-fractional Brownian motion,

(ii) $\mathbf{X}$ and $\mathbf{B}^H$ satisfy equation (5.1).

Also, sometimes, we say that the process $\mathbf{X}$ is a weak solution of equation (5.1).

Theorem 5.1. *Suppose that the process* $(\mathbf{B}^H, \mathbf{X})$ *on a filtered probability space* $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ *and the process* $(\overline{\mathbf{B}^H}, \mathbf{Y})$ *on a filtered probability space* $(\overline{\Omega}, \overline{\mathscr{F}}, \overline{\mathscr{F}}_t, \overline{P})$ *are weak solutions of equation* (5.1). *If* $(\mathbf{B}^H, \mathbf{X}) \equiv_1 (\overline{\mathbf{B}^H}, \mathbf{Y})$ *then* $\mathbf{X} \equiv_{\mathrm{AD}} \mathbf{Y}$ *and* $\mathbf{B}^H \equiv_{\mathrm{AD}} \overline{\mathbf{B}^H}$.

Proof. In [18] (Theorem 2.1) it is proved that every process of fractional Brownian motion is its own cause and that every weak solution (process $\mathbf{X}$) of equation (5.1) is its own cause (Theorem 1.3). Now, from $\mathbf{B}^H \equiv_1 \overline{\mathbf{B}^H}$ together with the fact that both the processes $\mathbf{B}^H$ and $\overline{\mathbf{B}^H}$ are their own cause, using Theorem 4.7 of the present paper, we conclude that $\mathbf{B}^H \equiv_{\mathrm{AD}} \overline{\mathbf{B}^H}$. In a similar way, we have that $\mathbf{X} \equiv_{\mathrm{AD}} \mathbf{Y}$. □

Remark 5.2. The essence of Theorem 5.1 is that if weak solutions of some equation (5.1) are unique up to synonymity then they are unique up to the same adapted distribution (or adapted law).

Corollary 5.3. *When every two weak solutions of equation* (5.1) *are synonymous, then if a strong solution of that equation exists, it also has the same adapted distribution (or the adapted law).*

Proof. The result follows from the fact that every strong solution is a weak solution, too. □

For similar stochastic differential equations, but with a Wiener process, we have an even better result.

For a given interval $[0, T]$ let $\alpha\colon [0, T] \times \mathbb{R} \to \mathbb{R}$ be a measurable nonanticipative function and $\mathbf{W} = \{W_t, \mathscr{F}_t, 0 \leqslant t \leqslant T\}$ a Wiener process.

We say that a stochastic differential equation

$$\mathrm{d}X_t = \alpha(t, X)\,\mathrm{d}t + \mathrm{d}W_t \tag{5.2}$$

with an initial condition η, having the prescribed distribution function $F(x)$ has a weak solution if there exist:

(i) a probability space $(\Omega, \mathscr{F}, P)$,

(ii) a nondecreasing family of sub-σ-algebras $(\mathscr{F}_t)$, $0 \leqslant t \leqslant T$,

(iii) a continuous random process $\mathbf{X} = (X_t, \mathscr{F}_t)$,

(iv) a Wiener process $\mathbf{W} = (W_t, \mathscr{F}_t)$ such that

$$P\Big(\int_0^T |\alpha(t,X)|\,\mathrm{d}t < \infty = 1\Big)$$

and

(v) with probability 1 for each t, $0 \leqslant t \leqslant T$,

$$X_t = \eta + \int_0^t \alpha(s,X)\,\mathrm{d}s + W_t.$$

The process $\mathbf{X}$ given above will be called weak a solution of equation (5.2).

Proposition 5.4. *If for every two weak solutions* $\mathbf{X}$ *and* $\mathbf{Y}$ *(defined on a possibly different probability spaces) of equation* (5.2) *we have* $\mathbf{X} \equiv_0 \mathbf{Y}$, *then every weak solution of that equation has the same adapted distribution (adapted law).*

P r o o f. Every weak solution of equation (5.2) is a Wiener process (a consequence of the Girsanov theorem), so it has the Markov property. Therefore, $\mathbf{X}$ and $\mathbf{Y}$ are Markov processes with the property $\mathbf{X} \equiv_0 \mathbf{Y}$. From Proposition 2.4 it follows that any Markov process on $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ is its own cause within the filtration $(\mathscr{F}_t)$. Now, using Theorem 2.8. in [11], we can conclude that $\mathbf{X} \equiv_{\mathrm{AD}} \mathbf{Y}$. □

Remark 5.5. The essence of this proposition is that weakly unique weak solutions (having the same distribution) of an equation (5.2) are unique in the AD sense (have the same adapted distribution or adapted law).

Corollary 5.6. *If for every two weak solutions* $\mathbf{X}$ *and* $\mathbf{Y}$ *of equation* (5.2) *we have* $\mathbf{X} \equiv_0 \mathbf{Y}$, *then a strong solution of that equation, if it exists, has the same adapted distribution (or the adapted law).*

P r o o f. Since every strong solution is a weak solution, the statement holds. □

References

[1] *D. Aldous*: Weak convergence and the general theory of proccesses, preprint (1981).

[2] *P. Bremaud, M. Yor*: Changes of filtrations and of probability measures. Z. Wahrscheinlichkeitstheor. Verw. Geb. *45* (1978), 269–295.

[3] *S. Fajardo, H. J. Keisler*: Model Theory of Stochastic Processes, Lecture Notes in Logic vol. 14, Urbana, 2002, Association for Symbolic Logic.

[4] *J. P. Florens, D. Fougères*: Noncausality in continuous time. Econometrica *64* (1996), 1195–1212.

[5] *J. B. Gill, L. Petrović*: Causality and stochastic dynamic systems. SIAM J. Appl. Math. *47* (1987), 1361–1366.

[6] *C. W. J. Granger*: Investigating causal relations by econometric models and cross spectral methods. Econometrica *37* (1969), 424–438.

[7] *D. N. Hoover*: Adapted distribution, Probability theory and applications. Proc. World Congr. Bernoulli Soc., Tashkent/USSR 1986 *1* (1987), 201–204.

[8] *D. N. Hoover*: Synonymity, generalized martingales, and subfiltrations. Ann. Probab. *12* (1984), 703–713.

[9] *D. N. Hoover*: A characterization of adapted distribution. Ann. Probab. *15* (1987), 1600–1611.

[10] *D. N. Hoover*: Extending probability spaces and adapted distribution, Séminare de probabilitités XXVI. Lect. Notes Math. *1526* (1992), 560–574.

[11] *D. N. Hoover, H. J. Keisler*: Adapted probability distributions. Trans. Am. Math. Soc. *286* (1984), 159–201.

[12] *R. S. Lipster, A. N. Shiryaev*: Statistics of Random Processes I, General theory. Applications of Mathematics vol. 5, Springer-Verlag, New York-Heidelberg-Berlin, 1977.

[13] *P. A. Mykland*: Statistical causality. Statistical Report no. 14, Dept. of Mathematics, University of Bergen, 1986.

[14] *D. Nualart, Y. Ouknine*: Regularization of differential equations by fractional noise. Stochastic Processes. Appl. *102* (2002), 103–116.

[15] *L. Petrović*: Causality and Markovian representations. Stat. Probab. Lett. *29* (1996), 221–227.

[16] *L. Petrović*: Causality and stochastic realization. Int. J. Math. Math. Sci. (2005), 349–356.

[17] *L. Petrović, D. Stanojević*: Statistical Causality, Extremal Measures and Weak Solutions of Stochastic Differential Equations with Driving Semimartingales, vol. 9. JMMA, 2010, pp. 113–128.

[18] *L. Petrović, S. Dimitrijević*: Some models of causality and stochastic differential equations driven by fractional Brownian motion. Facta Univ., Ser. Math. Inf. *20* (2005), 113–122.

[19] *Y. A. Rozanov*: Innovation Processes, Scripta Series in Mathematics. Washington, V. H. Winston and Sons, New York, 1977.

[20] *Y. A. Rozanov*: Markov Random Fields. Springer-Verlag, Berlin-New York-Heidelberg, 1982.

[21] *D. W. Strook, M. Yor*: On extremal solutions of martingale problems. Ann. Sci. Éc. Norm. Supér. (4) *13* (1980), 95–164.

[22] *M. Yor*: Sur L'Étude des Martingales Continues Extremales. Stochastics *2* (1979), 191–196.

Authors' addresses: Ljiljana Petrović, University of Belgrade, Faculty of Economics, Kamenička 6, 11000 Belgrade, Serbia, e-mail: petrovl@ekof.bg.ac.rs; Sladjana Dimitrijević, University of Kragujevac, Faculty of Science, Radoja Domanovića 12, 34000 Kragujevac, Serbia, e-mail: sladjana_dimitrijevic@kg.ac.rs.