ON THE RELATION BETWEEN THE NELSON STOCHASTIC DERIVATIVE AND THE FEYNMAN-KAC THEOREM

Francesco Guida
Department of Computer Science
University of Verona
Strada le Grazie 15, 37134 Verona, ITALY

Abstract: This paper proposes a different point of view to look at the Nelson’s derivatives. In particular we relate the Nelson’s forward derivative of a function of a diffusion, with certain properties, to the partial differential equation coming from the Feynman-Kac’s theorem. We will see that whenever the Nelson forward derivative of functions of a diffusion process, that satisfies a certain stochastic differential equations, is computed at the initial point of the process and whenever we put it equal to zero, we obtain a partial differential equation. The latter is the same equation coming from the Feynman-Kac’s theorem. Finally we will see how we could reduce the stochastic derivatives to the forward Nelson’s derivative of a certain process by supposing the existence of a new complex processes related to the backward filtration.

1. Nelson Stochastic Calculus

1.1. Introduction

Nelson’s derivatives arises naturally when we are searching for a stochastic counterpart of the usual classical derivatives.
The research for this stochastic counterpart obviously meet as a first problem: the non-differentiability of trajectories of a Brownian motion. To overcome that problem we have to be able to capture the kinematic of our process using the information available. If $X(t)$ is a stochastic process, then its kinematic is something captured by a quantity like this

$$\frac{X(t + \Delta t) - X(t)}{\Delta t}$$

but considering also the information available at time $t$, because we are far from every deterministic space. These arguments suggest a precise description of kinematic of a stochastic process. That description will lead us to a characterization of martingales. In particular if we apply that characterization to the right martingale we’ll obtain a characterization of Feynman-Kac’s theorem via Nelson’s forward derivative.

1.2. The Nelson’s Derivatives

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, where $\mathcal{A}$ is the $\sigma$-algebra of all measurable events and $\mathbb{P}$ is a probability measure defined on $\mathcal{A}$. Let $\mathcal{P} = \{\mathcal{P}_t\}$ and $\mathcal{F} = \{\mathcal{F}_t\}$ be an increasing and a decreasing family of sub-$\sigma$-algebras, respectively.

**Definition 1.** A $d$-dimensional random process $X(\cdot)$ defined on $I \times \Omega$, where $I$ is an open interval of $\mathbb{R}$, is an SO-process if: $X(\cdot)$ has continuous sample paths, $X(\cdot)$ is $\mathcal{P}$ and $\mathcal{F}$ adapted, for all $t \in I$, $X(t) \in L^2(\Omega)$, the mapping $t \mapsto X(t)$ from $I$ to $L^2(\Omega)$ is continuous.

**Definition 2.** A $d$-dimensional random process $X(\cdot)$ is an S1-process if it is an SO-process such that:

1.

$$DX(t) = \lim_{h \to 0^+} \mathbb{E} \left[ \frac{X(t + h) - X(t)}{h} | \mathcal{P}_t \right]$$

2.

$$D* X(t) = \lim_{h \to 0^+} \mathbb{E} \left[ \frac{X(t) - X(t-h)}{h} | \mathcal{F}_t \right]$$

exist in $L^2(\Omega)$ for $t \in I$, and the mappings $t \mapsto DX(t)$ and $t \mapsto D* X(t)$ are both continuous from $I$ to $L^2(\Omega)$. 
We prove now a theorem that can be found in [16] that explains the behaviour of the Nelson’s forward derivative. This behaviour lead to a characterization of martingales via Nelson’s forward derivative.

**Theorem 1.** Let $X$ be an $S^1$ process, and let $a \leq b, a, b \in I$. Then

$$
\mathbb{E}[X(b) - X(a) | \mathcal{F}_a] = \mathbb{E}
\left[
\int_a^b DX(s) ds | \mathcal{F}_a
\right]
$$

Notice that the integral exists since $s \rightarrow DX(s)$ is continuous.

**Proof.** Let $\varepsilon > 0$ and let $J$ be the set of all $t$ in $[a, b]$ s.t.

$$
||\mathbb{E}[X(s) - X(a) | \mathcal{F}_a] - \mathbb{E}
\left[
\int_a^s DX(r) dr | \mathcal{F}_a
\right]||_1 \leq \varepsilon(s - a)
$$

for all $a \leq s \leq t$, where $|| \cdot ||_1$ denotes the $L^1$ norm. Clearly, $a$ is in $J$, and $J$ is a closed subinterval of $[a, b]$. Let $t$ be the right end-point of $J$, and suppose that $t < b$. By the definition of $DX(t)$, there is a $\delta > 0$ s.t.

$$
||\mathbb{E}[X(t + \Delta t) - X(t) | \mathcal{F}_t] - DX(t) \Delta t||_1 \leq \frac{\varepsilon}{2} \Delta t
$$

for $0 \leq \Delta t \leq \delta$. Since conditional expectations reduce the $L^1$ norm and since $\mathcal{F}_t \cap \mathcal{F}_a = \mathcal{F}_a$,

$$
||\mathbb{E}[X(t + \Delta t) - X(t) | \mathcal{F}_a] - \mathbb{E}[DX(t) \Delta t | \mathcal{F}_a]||_1 \leq \frac{\varepsilon}{2} \Delta t
$$

for $0 \leq \Delta t \leq \delta$. By reducing $\delta$ if necessary, we find:

$$
||DX(t) \Delta t - \int_t^{t+\Delta t} DX(s) ds||_1 \leq \frac{\varepsilon}{2} \Delta t
$$

for $0 \leq \Delta t \leq \delta$, since $s \mapsto DX(s)$ is continuous in $L^1$. Therefore,

$$
||\mathbb{E}[DX(t) \Delta t | \mathcal{F}_a] - \mathbb{E}
\left[
\int_t^{t+\Delta t} DX(s) ds | \mathcal{F}_a
\right]||_1 \leq \frac{\varepsilon}{2} \Delta t
$$

for $0 \leq \Delta t \leq \delta$. From (1) for $s = t$, (2), and (3), it follows that (1) holds for all $t + \Delta t$ with $0 \leq \Delta t \leq \delta$. This contradicts the assumption that $t$ is the end-point of $J$, so we must have $t = b$. Since $\varepsilon$ is arbitrary, the statement holds. \qed
Theorem 2. An $S_1$-process is a martingale if and only if $DX(t) = 0$, $t \in I$.

This theorem is an immediate consequence of the previous theorem and of the definition of martingale.

1.3. Good Diffusion Processes

We introduce a class of functions for which we can explicitly compute Nelson’s derivatives.

Definition 3. We denote by $\Lambda^1$ the set of diffusion processes $X$ satisfying the following conditions:

i. $X$ solves a SDE:

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))dW(t), \quad X(0) = X_0$$

where $X_0 \in L^2(\Omega)$, $W$ is a $\mathcal{F}$-Brownian motion, i.e., a Brownian motion w.r.t. the increasing filtration, $b : I \times \mathbb{R}^d \to \mathbb{R}^d$ and $b : I \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ are Borel measurable functions satisfying the linear growth and Lipschitz conditions:

$$\sup_t (|\sigma(t, x) - \sigma(t, y)| + |b(t, x) - b(t, y)|) \leq K|x - y|$$

$$\sup_t (|\sigma(t, x)| + |b(t, x)|) \leq (1 + |x|)$$

ii. For any $t \in I, X(t)$ has a density $p_t(x)$ at a point $x$.

iii. Setting $a_{ij} = (\sigma\sigma^*)_{ij}$, for any $i \in \{1, \ldots, n\}$, for any $t_0 \in I$,

$$\int_{t_0}^1 \int_{\mathbb{R}^d} |\partial_j(a_{ij}(t, x)p_t(x))|dxdt < +\infty$$

iv. $X(\cdot)$ is also a $\mathcal{F}$-Brownian diffusion, i.e., a Brownian diffusion w.r.t. the decreasing filtration.

Definition 4. We denote by $\Lambda^2$ the subset of $\Lambda^1$ whose diffusion processes $X$ are such that their drift $b$ and the function

$$(t, x) \to \frac{\partial_j(a_{ij}(t, x)p_t(x))}{p_t(x)}$$

are bounded, belong to $C^{1,2}(I \times \mathbb{R}^d)$, and have all its first and second order derivatives bounded.
1.4. The Nelson’s Derivatives for Good Diffusion Processes

A useful property of good diffusion processes is that their Nelson’s derivatives can be explicitly computed. The proof of the following theorem can be found in [9]. That theorem is very useful to understand the deep relation between generators of stochastic processes and their Nelson’s derivatives. We would like to underline that such relation can be used to better study a wide range of models as those arise in the mathematical finance scenario, particularly in the case of delayed forward-backward S(P)DEs, see, e.g., [5].

**Theorem 3.** Let $X \in \Lambda^1$ which writes $dX(t) = b(t, X(t))dt + \sigma(t, X(t))dW(t)$. Then $X$ is Markov diffusion w.r.t. to the increasing filtration $\mathcal{F} = \{\mathcal{F}_t\}$ and the decreasing filtration $\mathcal{G} = \{\mathcal{G}_t\}$. Moreover, $DX$ and $D_*X$ exists w.r.t. these filtrations and:

$$DX(t) = b(t, X(t))$$

$$D_*X(t) = b_*(t, X(t)) \text{ where } x \to p_t(x) \text{ denotes the density of } X(t) \text{ at } x$$

and

$$b_*(t, x) = b^j(t, x) - \frac{\partial_j(a_{ij}(t, x)p_t(x))}{p_t(x)}$$

with the convention that the term involving $\frac{1}{p_t(x)}$ is 0 if $p_t(x) = 0$.

**Proof.** Let $X \in \Lambda^1$. Then $X$ is a Markov diffusion w.r.t. the increasing filtration $\mathcal{F} = \{\mathcal{F}_t\}$, so:

$$\mathbb{E}\left[\frac{X(t+h) - X(t)}{h} \middle| \mathcal{F}_t\right] = \mathbb{E}\left[\frac{1}{h} \int_t^{t+h} b(s, X(s))ds \middle| \mathcal{F}_t\right],$$

and

$$\mathbb{E}\left[\left|\frac{X(t+h) - X(t)}{h} \middle| \mathcal{F}_t\right| - b(t, X(t))\right| \leq \frac{1}{h} \mathbb{E}\int_t^{t+h} |b(s, X(s)) - b(t, X(t))|ds$$

$$= \frac{1}{h} \int_t^{t+h} \mathbb{E}|b(s, X(s)) - b(t, X(t))|ds.$$  

Now $b$ is Lipschitz and $t \to \mathbb{E}[X_t]$ is locally integrable. We can conclude that

$$\frac{1}{h} \int_t^{t+h} \mathbb{E}|b(s, X(s)) - b(t, X(t))|ds \to 0 \text{ a.s.}$$
as \( h \to 0 \). Therefore \( DX \) exists and \( DX(t) = b(t, X(t)) \). We now prove \( D_* X(t) = b_*(t, X(t)) \).

As \( X \in \Lambda^1 \) we have that \( \overline{X}(t) = X(1 - t) \) is a diffusion process w.r.t. the increasing filtration \( (\mathcal{F}_{1-t}) \) and whose generator reads

\[
\mathcal{L}_t f = \overline{b}^i \partial_i f + \frac{1}{2} \overline{a}^{ij} \partial_{ij} f
\]

with \( \overline{a}^{ij}(1 - t, x) = a^{ij}(t, x) \) and

\[
\overline{b}^i(1 - t, x) = -b^i(t, x) + \frac{\partial_j(a_{ij}(t, x)p_t(x))}{p_t(x)}
\]

We have:

\[
E \left[ \frac{X(t) - X(t - h)}{h} \bigg| \mathcal{F}_t \right] = E \left[ \frac{\overline{X}(1 - t) - \overline{X}(1 - t + h)}{h} \bigg| \mathcal{F}_{1-t} \right]
\]

\[
= E \left[ \frac{1}{h} \int_{1-t}^{1-t+h} b(s, \overline{X}(s))ds \bigg| \mathcal{F}_{1-t} \right]
\]

Using the same calculations as above since hypothesis iii) implies that

\[
t \to E \left[ \frac{\partial_j(a_{ij}(t, x)p_t(x))}{p_t(x)} \right]
\]

is locally integrable, we obtain that \( D_* X(t) \) exists and is equal to

\[
-\overline{b}(1 - t, \overline{X}(1 - t)).
\]

\[ \square \]

In the following we compute Nelson’s derivatives of \( f(t, X(t)) \) where \( X(t) \) is a diffusion process and \( f \) is a smooth function.

**Theorem 4.** Let \( X \in \Lambda^1 \) and \( f \in C^{1,2}(I \times \mathbb{R}^d) \). Then we have:

\[
Df(t, X(t)) = \left[ \partial_t f + DX(t) \cdot \nabla f + \frac{1}{2} a^{ij} \partial_{ij} f \right](t, X(t)),
\]

\[
D_* f(t, X(t)) = \left[ \partial_t f + D_* X(t) \cdot \nabla f - \frac{1}{2} a^{ij} \partial_{ij} f \right](t, X(t)).
\]
Proof. The proof can be found in [9]. Let \( X \in \Lambda^1 \) and \( f \in C^{1,2}(I \times \mathbb{R}^d) \). Thus \( f \) belongs to the domain of the generators \( L_t \) and \( \overline{L}_t \) of the diffusions \( X(t) \) and \( \overline{X}(t) \). Moreover these regularity assumptions allow us to write:

\[
Df(t, X(t)) = \partial_t f(t, X(t)) + L_t(f(t, \cdot))(X(t)) = \\
= \left[ \partial_t f + b^i \partial_i f + \frac{1}{2} a^{ij} \partial_{ij} f \right](t, X(t)) = \\
= \partial_t f + DX(t) \cdot \nabla f + \frac{1}{2} a^{ij} \partial_{ij} f \right](t, X(t))
\]

and

\[
Df(t, X(t)) = \partial_t f(t, X(t)) + \overline{L}_{1-t}(f(t, \cdot))(X(t)) = \\
= \left[ \partial_t f + D_\ast X(t) \cdot \nabla f - \frac{1}{2} a^{ij} \partial_{ij} f \right](t, X(t)).
\]

\[\Box\]

2. The Feynman-Kac’s Theorem

The idea behind this section is that we want to apply the theorem 2 to the martingale that we usually find when we derive the partial differential equation related to the underlied stochastic differential equation. Indeed, the general principle behind the proof of the Feynman-Kac’s theorem is to find the right martingale, take the differential and set the drift term equal to 0. In the following we stop at the first step and we proceed via Nelson’s derivatives theory. We keep in mind the description in [20] in its chapter 6 of how to find that martingale.

Let \( X \in \Lambda^1 \) and \( 0 \leq t \leq T \) be given, and let \( h(y) \) be a Borel-measurable function. Denote by

\[
f(t, x) = \mathbb{E}^{t,x} h(X(T)) \tag{4}
\]

the expectation of \( h(X(T)) \), where \( X(T) \) is the solution of the underlied stochastic differential equation with initial condition \( X(t) = x \).
Theorem 5. Let \( X(t), t \geq 0 \) be a process in \( \Lambda^1 \) with initial condition given at time 0. Then, for \( 0 \leq t \leq T \),

\[
E[h(X(T))|\mathcal{F}(t)] = f(t, X(t))
\]

where here the filtration is the one generated by the Brownian motion.

This theorem underlies the Feynman-Kac’s theorem. The latter relates stochastic differential equations and partial differential equations.

Theorem 6. Consider \( X \in \Lambda^1 \) and \( f(t, x) \) defined as in (4). Then

\[
Df(t, X(t)) = 0
\]

so that

\[
f_t(t, x) + b(t, x)f_x(t, x) + \frac{1}{2}\sigma^2(t, x)f_{xx}(t, x) = 0 \tag{5}
\]

and we have also the terminal condition:

\[
f(T, x) = h(x) \quad \text{for all } x \tag{6}
\]

The proof of that theorem resides on theorems (4) and (2) and on the following.

Theorem 7. Let \( X \in \Lambda^1 \). Let \( h(y) \) be a Borel-measurable function, fix \( T > 0 \), and let \( f(t, x) \) be given by (4). Then the stochastic process

\[
f(t, X(t)), \quad 0 \leq t \leq T,
\]

is a martingale.

3. Reducing the Stochastic Derivative

In this section we define a natural extension of the classical derivative on real stochastic processes. This stochastic derivative turns out to be necessarily complex valued. We proceed as in [9]. As we will see, the stochastic derivative will be a certain combination of both Nelson’s derivatives. We will also represent the stochastic derivative in terms of the only forward Nelson’s derivative. But
in this way we will have to study deeper geometric properties of underlying stochastic processes.

The best extension proposed in [9] which still maintain good properties, in particular it is the usual classical derivative if restricted to deterministic processes, and it is $\mathbb{R}$-linear, is represented by:

**Definition 5.** We denote by $D_\mu$, the operators defined by

$$D_\mu = \frac{D + D_*}{2} + i\mu\frac{D - D_*}{2} \quad \mu = \pm 1$$

We note that these stochastic derivatives are complex. We need to extend stochastic derivatives to complex processes. There are two principal ways to do that: an algebraic way and an analytical one. We will extend it in the analytical way. We first have that $D$ and $D_*$ have a natural extension to complex processes. Indeed, let $X = X_1 + iX_2$ be a complex process where $X_1, X_2$ are real processes. Then the natural extension of Nelson’s derivatives is:

$$D(X_1 + iX_2) = D(X_1) + iD(X_2) \quad \text{and} \quad D_*(X_1 + iX_2) = D_*(X_1) + iD_*(X_2)$$

As a consequence, we can naturally extend stochastic derivatives simply by:

$$D_\mu = \frac{D + D_*}{2} + i\mu\frac{D - D_*}{2} \quad \mu = \pm 1$$

with the natural extension of $D$ and $D_*$. Turn back for a moment to the derivatives of functions of good diffusion processes. We have seen that the Feynman-Kac’s theorem counterpart in the Nelson’s theory simply reduced to the formula:

$$Df(t, X(t) = x) = 0$$

If we consider the same equation but with the $D_*$ operator, what we are asking for?

$$D_*f(t, X(t) = x) = 0 \iff \left[ \partial_t f + D_* X(t) \cdot \nabla f - \frac{1}{2} a^{ij} \partial_{ij} f \right](t, X(t) = x) = 0$$

Remind now that for good diffusion processes by the theorem (3) we have that:
\[ D_*X(t) = b_*(t, X(t)) \] where \[ b^i_*(t, x) = b^i(t, x) - \frac{\partial_j(a_{ij}(t, x)p_t(x))}{p_t(x)} \]

In a certain sense we are looking for a complex process \( \hat{X} \) that has the following dynamic:

\[ d\hat{X}(t) = b_*(t, X(t)) + i\sigma(t, X(t))dW(t) \]

for that process we also have: \( D\hat{X}(t) = b_*(t, X(t)) = D_*X(t) \).

So considering now the stochastic derivative we have:

\[ D_\mu X = \frac{DX + D_*X}{2} + i\mu \frac{DX - D_*X}{2} = \frac{DX + D\hat{X} + i\mu(DX + D\hat{X})}{2} \]

that is:

\[ D_\mu X = \frac{1}{2}D[(1 - i\mu)(X + \hat{X})] \quad \mu = \pm 1 \]

We have found that computation of backward Nelson’s derivative, could imply the existence of a complex process that is related to the backward filtration. This might suggests that when we consider the effects of a filtration bringing information from the future, we intrinsically suppose the existence of a complex process that has its randomness in its imaginary part. We have also found that supposing that process really exists and has good properties the computation of the stochastic derivatives reduces to computing the forward Nelson’s derivative of the sum of the real processes \( X \) and the complex "counterpart" \( \hat{X} \).

4. Conclusion and Perspectives

The main aim of the present paper is to underline a particular class of mathematical links as in the case of the connection between the random variable \( Df(t, X(t)) \) and the discussed PDE. This relation could give us a natural way to study properties such as symmetries of partial differential equation directly studying what change in the underlying stochastic process whenever we apply a symmetric transformation. It is worth to mention that such an approach has been extensively used during last decades by a number of different points of view, see, e.g., [13] concerning the analysis of nonlinear PDEs, [3, 4, 8] to what concerns applications in mathematical finance, also in the light of the following
papers [6, 7] and references therein, and also to simplify certain type of computational tasks that arise in the field of multivariate statistics when huge quantities of data have to be taken into account, see, e.g., [10, 11] and references therein. It follows that the development of a general theory that relates Lie theory of symmetries for PDEs with the transformation of the stochastic process that the symmetry could bring to it, could potentially bring new methods to better analyse relevant models characterizing heterogeneous fields, as in the case, e.g., of meteorology, finance, biology, etc.

References


