Malliavin’s calculus in insider models: additional utility and free lunches

Peter Imkeller
Institut für Mathematik
Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin
Germany
February 14, 2002

Abstract

We consider simple models of financial markets with regular traders and insiders possessing some extra information hidden in a random variable which is accessible to the regular trader only at the end of the trading interval. The problems we focus on are the calculation of the additional utility of the insider and a study of his free lunch possibilities. The information drift, i.e. the drift to eliminate in order to preserve the martingale property in the insider’s filtration, turns out to be the crucial quantity needed to answer these questions. It is most elegantly described by the logarithmic Malliavin trace of the conditional laws of the insider information with respect to the filtration of the regular trader. Several examples are given to illustrate additional utility and free lunch possibilities. In particular, if the insider has advance knowledge of the maximal stock price process, given by a regular diffusion, arbitrage opportunities exist.

1991 AMS subject classifications: primary 60 G 48, 90 A 09; secondary 60 H 07, 94 A 17.

Key words and phrases: insider trading; enlargement of filtrations; Malliavin’s calculus; free lunch; arbitrage; equivalent martingale measure; Bessel process.

Introduction

The mathematical study of financial markets with economic agents possessing different information levels proceeds via essentially two approaches.

Mostly discrete models originating in economics oriented papers (see O’Hara [34], Kyle [31]) investigate auction trading with three agents: a market maker, a noise trader, and a risk neutral insider, to whom the final value of one stock is known in advance.
Back [4], [5] extends these models in which the insider’s actions may have a backlash on the pricing rules to the time continuous setting. For more information on this and related classes of models see for example the thesis by Wu [38].

The approach from the point of view of martingale theory, which we shall follow in this paper was first taken in Karatzas, Pikovsky [28], and Pikovsky [35]. They study a continuous time model on a Wiener space with two agents: a regular agent whose information coincides with the natural filtration of the price processes, and an insider who possesses some extra information hidden in a random variable $G$ from the beginning of the trading interval. Discussing questions like the insider’s additional utility with respect to special utility functions, and martingale representation properties in the insider’s filtration, they introduce the powerful technique of grossissement de filtrations (see Yor [39], [40], [41], [42], [43], Jeulin [27], Jacod [25]) to this economic context. Subsequent work along this punch line studies admissible strategies for insiders, their additional profits, and possibilities to detect them on stochastic bases of increasing complexity. See Grorud, Pontier [17], Denis, Grorud, Pontier [13], [19].

The discussion of models of this kind in the present paper focuses on two questions:

- how can one calculate the additional utility of an insider?
- does the insider possess opportunities of riskless free lunches or arbitrage?

Work mainly concentrating on the first question performed in Amendinger [1], Amendinger, Imkeller and Schweizer [3], Amendinger, Becherer and Schweizer [2] and [23] underlines that the domain of techniques of enlargement of filtrations is general semimartingale theory. Investigations of the second question in Imkeller, Pontier and Weisz [22] for special interesting insider informations, however, showed that the usual framework of Jacod [25] in which the conditional laws of $G$ are supposed to be absolutely continuous with respect to its law is insufficient. Much more flexibility and transparency is obtained if the techniques of Malliavin’s calculus become available. See [20], [21].

The most important observation in this context is based on a measure valued version of the formula of Clark-Ocone presented in [22], applied to the conditional laws of $G$ with respect to the regular trader’s filtration. It allows in the first place to interpret the information drift, i.e. the drift to eliminate in passing to the insider model in order to keep the Brownian motion a martingale, as a logarithmic Malliavin trace of the conditional law. Secondly, it allows to extend Jacod’s framework of enlargement of filtrations to cover quite general additional informations $G$. The explicit description of the information drift allows to derive verifiable criteria for non-existence of equivalent martingale measures from the perspective of the regular trader, i.e. measures under which the insider does not see Brownian motion as a martingale any more. This provides tools to access the existence of free lunches and arbitrage in simple interesting examples.

This paper which aims at stressing the natural role played by Malliavin’s calculus in insider models, and which partially reviews main results of [3] and [22], is organized as follows. In section 1 we explain our insider model in the simplest possible framework, and provide several both illustrative and analytically accessible examples of insider
information $G$. In section 2 the additional expected logarithmic utility of an insider is computed in terms of the information drift (Theorem 2.1), and linked to the entropy of $G$. In the framework of Jacod’s approach of enlargement of filtrations the information drift is expressed in terms of conditional densities of $G$ (Theorem 2.2). Motivated by the example $G = \sup_{t \in [0,1]} W_t$, section 3 is devoted to the extension of Jacod’s framework by means of Malliavin’s calculus for measure valued martingales on Wiener space. It culminates in a measure valued Clark-Ocone formula (Theorem 3.1) that allows to reinterpret the information drift as the Radon-Nikodym density of the Malliavin trace of the conditional laws of $G$ with respect to themselves (Theorem 3.2). In section 4 the important example $G = \sup_{t \in [0,1]} X_t$, where the insider knows in advance the maximal price of a risky asset, modeled by a regular diffusion $X$, is considered. Properties of the information drift are derived which imply that there exist free lunches with no risk (Theorem 4.1 and Theorem 4.2).

1 A simple model of a market with an insider

Our basic probability space is the 1-dimensional canonical Wiener space $(\Omega, \mathcal{F}, P)$, equipped with the canonical Wiener process $W = (W_t)_{t \geq 0}$. More precisely, $\Omega = C(\mathbb{R}_+; \mathbb{R})$ is the set of continuous functions on $\mathbb{R}_+$ starting at 0, $\mathcal{F}$ the $\sigma$-algebra of Borel sets with respect to uniform convergence on compact subsets of $\mathbb{R}_+$, $P$ Wiener measure and $W$ the coordinate process. The natural filtration $(\mathcal{F}_t)_{t \geq 0}$ of $W$ is assumed to be completed by the sets of $P$-measure 0.

To simplify the exposition of the main aspects of the additional utility a better informed trader may have, or even his opportunities to exercise arbitrage, we stick to the simplest possible setting. So we suppose that the trading interval is given by $[0, 1]$. We let our financial market model $(\alpha, \sigma)$ consist of a progressively measurable mean rate of return process $\alpha$ which satisfies $\int_0^1 |\alpha_t| dt < \infty$ $P$-a.s. and of a progressively measurable volatility process $\sigma$ satisfying $\int_0^1 \sigma_t^2 dt < \infty$, $\sigma^2 > 0$ $P$-a.s. They determine a (stock) price process given by

$$\frac{dX_t}{X_t} = \alpha_t dt + \sigma_t dW_t.$$  

For convenience, we let $X_0 = 1$. Two traders may act on this simple market. The regular trader’s information level corresponds to the natural information flow of the market. So at time $t$ the trader’s knowledge is given by $\mathcal{F}_t$. The insider’s additional information is supposed to be available from the very beginning of the trading interval, and consists in the knowledge of a random variable $G$ which is $\mathcal{F}_1$- or $\mathcal{F}_{1+\epsilon}$-measurable for some (small) $\epsilon > 0$. Hence the insider’s filtration is given by $\mathcal{G} = (\mathcal{G}_t)_{t \in [0,1]}$, where

$$\mathcal{G}_t = \mathcal{F}_t \vee \sigma(G),$$

$t \in [0,1]$. Important, analytically tractable examples of additional information are as follows.
**Examples:**
1. Let \( a, b \in \mathbb{R}, a < b \). Then
   \[
   G = 1_{[a,b]}(X_t), \quad \text{or} \quad G = 1_{[a,b]}(X_{1+\epsilon})
   \]
represents the binary advance knowledge whether the price of the asset \( X \) at time 1 or \( 1 + \epsilon \) is inside or outside of the interval \([a,b]\).

2. Knowing
   \[
   G = X_1 \quad \text{or} \quad G = X_{1+\epsilon}
   \]
means to have advance knowledge of the exact price of the asset \( X \) at time 1 or \( 1 + \epsilon \).

3. Advance knowledge of the maximal stock price in the trading interval is assumed if
   \[
   G = \sup_{t \in [0,1]} X_t.
   \]

4. The insider might also be in possession of the advance knowledge of the random time of the last crossing of some level \( a \) by the stock price process
   \[
   G = \sup \{ t : t \in [0,1], X_t = a \}.
   \]

While for the regular trader a portfolio process will be an \( \mathbb{F} \)-progressively measurable process \( \pi \) such that
\[
\int_0^1 |\pi_t \alpha_t| dt < \infty \quad \text{P-a.s.}
\]
and
\[
\int_0^1 |\pi_t \sigma_t|^2 dt < \infty \quad \text{P-a.s.,}
\]
the insider’s portfolio processes, analogously defined by replacing \( \mathbb{F} \) with \( \mathbb{G} \), will be allowed to base cleverer portfolio decisions on his additional information.

The *value process* \( V \) of a portfolio \( \pi \) is given by the formula
\[
\frac{dV_t}{V_t} = \pi_t \frac{dX_t}{X_t}.	ag{1}
\]

Here, of course, we suppose that \( X \) is a semimartingale in both filtrations, a key property to be discussed at length below. Again for convenience, let us suppose \( V_0 = 1 \).

To motivate the notion of *information drift* which will be of central importance, we shall just consider the logarithmic utility function
\[
U(x) = \ln x, \quad x > 0,
\]
to measure the utility a trader draws from his wealth \( V_t \) at the end of the trading interval. Hence the *expected maximal logarithmic utility* of the regular trader is given by
\[
N_\mathbb{F} = \max_{\pi \in \mathbb{F}-\text{portfolio}} E(\ln V_t), \tag{2}
\]
while the expected maximal logarithmic utility of the insider is expressed by

$$N_G = \max_{\pi_G-portfolio} E(\ln V_t). \quad (3)$$

The additional expected logarithmic utility of the insider due to his information advantage is therefore described by the formula

$$\Delta N = N_G - N_F.$$

2 Additional utility and information drift

Since the regular trader’s filtration is the natural filtration of the Wiener process, his expected logarithmic utility is easy to calculate. First of all, for a given \( \mathbb{F} \)-portfolio \( \pi \) the stochastic differential equation determining \( V \) is a simple linear equation solved by the formula

$$V_t = \exp\left[ \int_0^t \pi_s \sigma_s \, dW_s - \frac{1}{2} \int_0^t \pi_s^2 \sigma_s^2 \, ds \right].$$

Due to the local martingale property of \( \int_0^t \pi_s \sigma_s \, dW_s, t \in [0,1] \), the expected logarithmic utility of the regular trader is deduced from the maximization problem

$$N_F = \max_{\pi_F-portfolio} E\left[ \int_0^1 \pi_s \alpha_s \, ds - \frac{1}{2} \int_0^1 \pi_s^2 \sigma_s^2 \, ds \right]. \quad (4)$$

The maximization of

$$\pi \mapsto \int_0^1 \pi_s \alpha_s \, ds - \frac{1}{2} \int_0^1 \pi_s^2 \sigma_s^2 \, ds$$

for given processes \( \alpha \) and \( \sigma \) is just a more complex version of the one-dimensional maximization problem for the function

$$\pi \mapsto \pi \alpha - \frac{1}{2} \sigma^2 \, \pi^2$$

with \( \alpha, \sigma \in \mathbb{R} \). Its solution is obtained by the critical value \( \pi = \frac{\alpha}{\sigma^2} \) and thus

$$N_F = \frac{1}{2} E\left[ \int_0^1 \frac{\alpha_s^2}{\sigma_s^4} \, ds \right]. \quad (5)$$

To compute the insider’s expected utility, let \( \pi \) be a \( \mathbb{G} \)-portfolio. Now since \( W \) is not a martingale for the filtration \( \mathbb{G} \), \( \int_0^t \pi_s \sigma_s \, dW_s, t \in [0,1] \), is not a martingale any more, and we are led directly into the basic problem of the sophisticated technique of grossissement de filtrations. Suppose that \( \mathbb{G} \) is small enough so that \( W \) is still a semimartingale with respect to this filtration. Below we shall discuss conditions under which this is guaranteed. More precisely, suppose that there is a \( \mathbb{G} \)-progressively measurable process \( \mu^G \) such that

$$\int_0^1 |\mu_s^G| \, ds < \infty \quad \text{P-a.s.,}$$

5
and such that
\[ W = \tilde{W} + \int_0^t \mu_s^G \, ds \] (6)

with a \( \mathcal{G} \)-Brownian motion \( \tilde{W} \). We call \( \mu^G \) the information drift corresponding to the insider information \( G \). Using (6), we may describe the value process of the portfolio from the insider’s perspective by
\[ V_t = \exp \left[ \int_0^t \pi_s \sigma_s \, d\tilde{W}_s - \frac{1}{2} \int_0^t \pi_s^2 \sigma_s^2 \, ds + \int_0^t \pi_s [\alpha_s + \sigma_s \mu_s^G] \, ds \right]. \]

The essential difference between the two descriptions of the value process from the two traders’ perspectives consists in the fact that for the insider the information drift, modulated by the price volatility, adds to the regular drift \( \alpha \). Repeating the arguments for maximizing the resulting expected logarithmic utility of the insider will therefore lead to the formula
\[ N_G = \frac{1}{2} E \left[ \int_0^1 \frac{[\alpha_s + \sigma_s \mu_s^G]^2}{\sigma_s^2} \, ds \right]. \] (7)

The calculation of the additional expected logarithmic utility from (5) and (7) is now rather easy and results in a remarkable formula. Just observe that \( \alpha \) and \( \sigma \) are both \( \mathcal{F} \)- and \( \mathcal{G} \)-adapted. Hence we have
\[ E(\int_0^1 \frac{\alpha_s \mu_s^G}{\sigma_s} \, ds) = E(\int_0^1 \frac{\alpha_s}{\sigma_s} (dW_s - d\tilde{W}_s)) = 0, \]

which implies that the mixed term in the integral expression for \( \Delta N \) has to vanish. Therefore we have proved

**Theorem 2.1** Assume that there is a \( \mathcal{G} \)-progressively measurable process \( \mu^G \) such that
\[ \int_0^1 |\mu_s^G| \, ds < \infty \quad \text{P-a.s.,} \]

and such that
\[ W = \tilde{W} + \int_0^t \mu_s^G \, ds \]

with a \( \mathcal{G} \)-Brownian motion \( \tilde{W} \). Then the additional expected logarithmic utility of the insider is given by
\[ \Delta N = \frac{1}{2} E \left[ \int_0^1 (\mu_s^G)^2 \, ds \right]. \] (8)

Stated differently, Theorem 2.1 says that the additional expected logarithmic utility of the insider knowing \( G \) in advance is given by some type of energy of the information drift. For more details on the subject of calculating an insider’s additional logarithmic utility in a more general semimartingale framework see Amendinger [1] and Amendinger, Imkeller and Schweizer [3]. In this paper it is also shown that the result of Theorem 2.1 allows an intriguing and very simple representation in terms of the entropy of the additional information \( G \).
Example 1:
Let $a, b \in \mathbb{R}$, $a < b$, and

$$G = 1_{[a,b]}(X_1).$$

Let $p_1 = P(G \in [a,b])$. Then the additional expected logarithmic utility of the insider is given by

$$\Delta N = p_1 \ln(p_1) + (1 - p_1) \ln(1 - p_1).$$

Example 2:
Let the law of $G$ be absolutely continuous. Then

$$\Delta N = \infty.$$

This is the case for example for $G = X_{1+\epsilon}$, $\epsilon > 0$, if $X$ is a regular diffusion.

It remains to clarify the relationship between the additional information $G$ and the information drift $\mu^G$. Actually the discussion of this problem will keep us occupied for a while. It will clearly show how Malliavin’s calculus enters the scene in a very natural, powerful and efficient way.

In a first attempt, we shall work under a condition concerning the laws of the additional information $G$ which has been used as a standing assumption in many papers dealing with *grossissement de filtrations*. See Yor [39], [40], [43], Jeulin [27]. The condition was essentially used in particular in the seminal paper by Jacod [25], and in several equivalent forms in Föllmer and Imkeller [16]. To state and exploit it, let us first mention that all stochastic quantities appearing in the sequel, often depending on several parameters, can always be shown to possess measurable versions in all variables, and progressively measurable versions in the time parameter (see Jacod [25], Stricker and Yor [37]).

Denote by $P^G$ the law of $G$, and for $t \in [0,1], \omega \in \Omega$, by $P^G_t(\omega, dl)$ the regular conditional law of $G$ given $\mathcal{F}_t$ at $\omega \in \Omega$. Then the condition, which we will call *Jacod’s condition*, states that

$$P^G_t(\omega, dl) \text{ is absolutely continuous with respect to } P^G(dl) \text{ for } P-a.e. \omega \in \Omega.$$  \hspace{1cm} (9)

Also its reinforcement

$$P^G_t(\omega, dl) \text{ is equivalent to } P^G(dl) \text{ for } P-a.e. \omega \in \Omega,$$  \hspace{1cm} (10)

will be of relevance. We denote the Radon-Nikodym density process of the conditional laws with respect to the law by

$$p_t(\omega, l) = \frac{dP^G_t(\omega, \cdot)}{dP^G}(l), \quad l \in \mathbb{R}, \omega \in \Omega.$$ 

By the very definition, $t \mapsto p_t(\cdot, dl)$ is a local martingale with values in the space of probability measures on the Borel sets of $\mathbb{R}$. This is inherited to $t \mapsto p_t(\cdot, l)$ for (almost) all $l \in \mathbb{R}$. Let the representations of these martingales with respect to the $\mathbb{F}$–Wiener process $W$ be given by

$$p_t(\cdot, l) = p_0(\cdot, l) + \int_0^t k_u^l dW_u, \quad t \in [0, 1]$$

7
with measurable kernels $k$. To calculate the information drift in terms of these kernels, take $s, t \in [0, 1], s \leq t$, and let $A \in \mathcal{F}_s$ and a Borel set $B$ on the real line determine the typical set $A \cap G^{-1}[B]$ in a generator of $\mathcal{G}_s$. Then we may write

$$E([W_t - W_s] 1_A 1_B(G)) = E\left( \int_B 1_A [W_t - W_s] \, P_t^G(\cdot, dl) \right)$$

$$= \int_B E(1_A [W_t - W_s] [p_t - p_s](\cdot, l)) \, P_t^G(dl)$$

$$= \int_B E(1_A \int_s^t k_u^l \, du) \, P_t^G(dl)$$

$$= \int_B E(1_A \int_s^t \frac{k_u^l}{p_u(\cdot, l)} \, p_u(\cdot, l) \, du) \, P_t^G(dl)$$

$$= \int_B E(1_A \int_s^t \frac{k_u^l}{p_u(\cdot, l)} \, p_t(\cdot, l) \, P_t^G(\cdot, dl)$$

$$= E\left( \int_B 1_A k_u^l \, p_u(\cdot, l) \, P_t^G(\cdot, dl) \right)$$

$$= E\left(1_A 1_B(G) \int_s^t \frac{k_u^l}{p_u(\cdot, l)} |_{t=G} \, du \right).$$

The bottom line of this chain of arguments shows that

$$\tilde{W} = W - \int_0^1 \frac{k_u^l}{p_u(\cdot, l)} |_{t=G} \, du$$

is a $\mathcal{G}$–martingale, hence a $\mathcal{G}$–Brownian motion provided that $\int_0^1 \left| \frac{k_u^l}{p_u(\cdot, l)} \right| |_{t=G} \, du < \infty$ $P$–a.s.. This completes the deduction of an explicit formula for the information drift of $G$ in terms of quantities related to the law of $G$ in which we use the common oblique bracket notation to denote the covariation of two martingales (for more details see Jacod [25]).

**Theorem 2.2** Suppose that Jacod’s condition (9) is satisfied, and furthermore that

$$\mu^G_t = \frac{k_t^l}{p_t(\cdot, l)} |_{t=G} = \left\{ \frac{\partial}{\partial l} \langle p(\cdot, l) \rangle, W \rangle, t \right\}_{t=G}, \quad t \in [0, 1],$$

satisfies

$$\int_0^1 |\mu^G_u| \, du < \infty \quad P-a.s.$$ (12)

Then

$$W = \tilde{W} + \int_0^1 \mu^G_u \, ds$$

is a $\mathcal{G}$–semimartingale with a $\mathcal{G}$–Brownian motion $\tilde{W}$.
3 Extension of Jacod’s framework: the information drift in Malliavin’s calculus

To see how restrictive condition (9) may be, let us illustrate it by looking at two of the examples of additional information given in section 1.

Example 1:
Let $\epsilon > 0$ and suppose that the stock price process is a regular diffusion given by a stochastic differential equation with bounded $\sigma$ and $\alpha, \sigma_t = \sigma(X_t), t \in [0,1]$, where $\sigma$ is a smooth function without zeroes. Let $G = X_{1+\epsilon}$. Then in particular $X$ is a time homogeneous Markov process with transition probabilities $P_t(x, dy), x \in \mathbb{R}_+, t \in [0,1]$, which are equivalent with Lebesgue measure on $\mathbb{R}_+$. For $t \in [0,1]$, the regular conditional law of $G$ given $\mathcal{F}_t$ is then given by $P_{1+\epsilon,t}(X_t, dy)$, which is equivalent with the law of $G$. Hence in this case, even the reinforcement of Jacod’s hypothesis (10) is verified.

Example 2:
Let $\sigma = 1, \alpha = \frac{1}{2}$, so that $X_t = \exp(W_t), t \geq 0$. Further, let $G = \sup_{t \in [0,1]} X_t$. Then by strict monotonicity of $\exp$, we might and do quite as well assume that

$$G = \sup_{t \in [0,1]} W_t.$$  

To abbreviate, denote for $t \in [0,1]$

$$G_t = \sup_{0 \leq s \leq t} W_s, \quad \tilde{G}_1-t = \sup_{t \leq s \leq 1} (W_s - W_t).$$

Finally, let $p_{1-t}$ denote the density function of $\tilde{G}_1-t$. Then we may write for every $t \in [0,1]$

$$G = G_t \vee [W_t + \tilde{G}_1-t].$$  

(13)

Now $G_t$ is $\mathcal{F}_t$–measurable, independent of $\tilde{G}_1-t$, and therefore for Borel sets $A$ on the real line we have

$$P_t^G(\cdot, A) = \int_{-\infty}^{G_t-W_t} p_{1-t}(y) dy \cdot \delta_{G_t}(A) + \int_{A \cap [G_t-W_t, \infty]} p_{1-t}(y) dy.$$  

(14)

Note now that the family of Dirac measures in the first term of (14) is supported on the random points $G_t$, and that the law of $G_t$ is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}_+$. Hence there cannot be any common reference measure equivalent with $\delta_{G_t}, P$–a.s.. Therefore in this example Jacod’s condition is violated.

We see that the framework of Jacod’s hypothesis (9) is insufficient to cover interesting examples such as the one just discussed. On the other hand, Amendinger [1] shows that the stronger condition (10) implies the existence of an equivalent martingale measure under which $W$ is a $\mathcal{G}$–martingale. As a consequence, (10) implies according to the fundamental theorem of asset pricing by Delbaen, Schachermayer [9] that there can be no free lunch with vanishing risk (NFLVR). This finally means that under (10) the insider has no possibilities of exercising arbitrage.
So in the situation of Example 1 the informed trader has infinite additional logarithmic utility, without, however, any arbitrage opportunities. What can be said in situations like Example 2? Can the insider exercise arbitrage or not? To answer this question, obviously the technical framework of Jacod’s hypothesis is not sufficient. In order to find a natural extension of this framework, we are therefore motivated to reconsider the approach which finally led to the identification of the information drift in terms of the conditional densities of the additional information in Theorem 2.2. The link to Malliavin’s calculus is provided by the well known formula of Clark-Ocone. We shall start explaining the identification of the information drift in terms of this stochastic calculus of variations with an emphasis on the concepts, and with less details concerning the formalities. For these the reader is referred to [22], or Nualart [32].

It is still imperative to briefly recall some of the basic concepts of Malliavin’s calculus.

Let \( \mathcal{S} \) be the set of smooth random variables on \((\Omega, \mathcal{F}, P)\), i.e. of random variables of the form

\[
F = f(W_{t_1}, ..., W_{t_n}), \quad f \in C_0^\infty(\mathbb{R}^n), \quad t_1, ..., t_n \in [0, 1].
\]

For \( F \in \mathcal{S} \) we may define the Malliavin derivative

\[
(DF)_s = D_s F = \sum_{i=1}^n \frac{\partial}{\partial x_i} f(W_{t_1}, ..., W_{t_n}) \mathbb{1}_{[0,t_i]}(s), \quad s \in [0, 1].
\]

We may regard \( DF \) as a random element with values in \( L^2([0,1]) \), and then define the Malliavin derivative of order \( k \) by \( k \) fold iteration of the above derivation. It will be denoted by \( D^{\otimes_k} F \), and is a random element with values in \( L^2([0,1]^k) \). Its value at \((s_1, ..., s_k) \in [0,1]^k\) is written \( D^{\otimes_k}_{s_1,...,s_k} F \).

If \( p \geq 1 \) and \( k \in \mathbb{N} \), we denote by \( \mathbf{D}_{p,k} \) the Banach space given by the completion of \( \mathcal{S} \) with respect to the norm

\[
\|F\|_{p,k} = \|F\|_p + \sum_{1 \leq j \leq k} E\left( \left| \int_0^1 (D^{\otimes_j}_{s_1,...,s_j} F)^2 ds_1 ... ds_j \right|^\frac{p}{2} \right)^\frac{1}{2}, \quad F \in \mathcal{S}.
\]

More generally, if \( H \) is a Hilbert space and \( \mathcal{S}_H \) the set of linear combinations of tensor products of elements of \( \mathcal{S} \) with elements of \( H, \mathbf{D}_{p,k}(H) \) will denote the closure of \( \mathcal{S}_H \) w.r. to the norm

\[
\|F\|_{p,k} = \|F\|_H + \sum_{1 \leq j \leq k} E\left( \left| \int_0^1 (D^{\otimes_j}_{s_1,...,s_j} F)^2_H ds_1 ... ds_j \right|^\frac{p}{2} \right)^\frac{1}{2}, \quad F \in \mathcal{S}_H,
\]

where the Malliavin derivatives of smooth functions are given in an obvious way, and \(|-|_H\) denotes the norm on \( H \) induced by the scalar product. These definitions are consistent. For example,

\[
\|F\|_p + \|DF\|_{p,k-1} = \|F\|_{p,k}, \quad F \in \mathbf{D}_{p,k},
\]

if \( H = L^2([0,1]) \).
The classical Clark-Ocone formula states that for $F \in \mathbf{D}_{2,1}$ one has

$$F = E(F) + \int_0^1 E(D_t F | \mathcal{F}_t) \, dW_t.$$  

The objects we are interested in representing are conditional densities. As stated earlier, they are martingales in the time parameter. For this reason, the conditional expectation in the stochastic integrand in the formula of Clark-Ocone can be interchanged with the Malliavin derivative, and one can show under suitable regularity conditions on $G$ (see for example [20] for somewhat too restrictive ones) that the following particular version of the representation formula holds

$$p_t(\cdot, l) = p_0(\cdot, l) + \int_0^t D_u p_u(\cdot, l) \, dW_u, \quad t \in [0, 1], l \in \mathbb{R},$$  

(15)

where the appearing Malliavin trace type object has to be understood as

$$D_u p_u(\cdot, l) = \lim_{t \downarrow u} D_u p_t(\cdot, l).$$

Comparing (15) with our earlier representation using the kernel $k$, we therefore find the following formula for the information drift

$$p^G_t = \frac{d}{dt} \frac{\mathbb{E}(p(\cdot, l), W_t)}{p_t(\cdot, l)}|_{t=G} = \frac{D_t p_t(\cdot, l)}{p_t(\cdot, l)}|_{t=G} = D_t \ln p_t(\cdot, l)|_{t=G}, \quad t \in [0, 1].$$  

(16)

In this intriguing representation, the information drift is identified with a logarithmic Malliavin trace of the conditional density.

Now at this point, after a moment’s thought it becomes clear that the nature of the logarithmic variational trace makes it pointless to take reference to a measure like the law of $G$ and operate on densities of the conditional law with respect to the law. Intuitively, our expression for the information drift is overdetermined due to the intervention of a reference measure. Granted some regularity, it should in fact be possible to interchange the Malliavin derivative $D_t$ and the Radon-Nikodym derivation $\frac{d}{dt}$ to formally obtain for $t \in [0, 1], l \in \mathbb{R}$

$$\frac{D_t p_t(\cdot, l)}{p_t(\cdot, l)} = \frac{d}{dt} \frac{ap^G_t(\cdot, l)}{ap^G_t(\cdot, l)} = \frac{d}{dt} \frac{D_t p^G_t(\cdot, dl)}{P^G_t(\cdot, dl)}(l).$$  

(17)

To identify the last two expressions, we need not ask for the absolute continuity or equivalence of conditional laws and law of $G$. Instead, we have to make sense of $D_t p^G_t(\cdot, dl)$ as a measure valued random variable in the first place, and in the second place to ask for absolute continuity or equivalence of this measure with respect to the regular conditional laws of $G$ given $\mathcal{F}_t$ directly. To tackle the first problem just means to extend the formula of Clark-Ocone to measure valued martingales such as our conditional laws. The basic measure valued Malliavin’s calculus needed for this purpose has been established in [22]. Let us just state its main consequences needed here, and refer the reader to [22] for further details.
Let $\mathbf{M}$ be the space of signed measures on the real line equipped with its Borel sets. The variation norm is denoted by $|\mu|$ for $\mu \in \mathbf{M}$. $\mathbf{M}$ is endowed with the Banach space topology induced by this norm. It is convenient to use a weaker topology: the weak* topology which is induced by the space $C_b(\mathbb{R})$ of continuous bounded functions with the supremum norm $\|\cdot\|$. Endowed with the latter topology, $\mathbf{M}$ is a locally convex space which, due to the separability of $C_b(\mathbb{R})$, is separable. For $\mu \in \mathbf{M}$, $f \in C_b(\mathbb{R})$, we denote $\langle \mu, f \rangle = \int_{\mathbb{R}} f \, d\mu$. We may choose a dense sequence $(f_i)_{i \in \mathbb{N}} \subset C_b(\mathbb{R})$, to use the standard embedding of $\mathbf{M}$ into an infinite dimensional metrizable space

$$\Phi : \mathbf{M} \rightarrow \mathbb{R}^N,$$

$$\mu \mapsto (\langle \mu, f_i \rangle)_{i \in \mathbb{N}}.$$ 

Note that $\mu$ is actually mapped into the compact cube $\prod_{i \in \mathbb{N}} [-|\mu||f_i|, |\mu||f_i|]$. We shall use $\Phi$ to define Malliavin derivatives of $\mathbf{M}$-valued objects. For $h \in L^2([0, 1])$, let $W(h) = \int_0^1 h(s) \, dW_s$. We first define the smooth cylinder functions. Let

$$\mathcal{S}(\mathbf{M}) = \{ F : F = g(W(h_1), \cdots, W(h_k), x) \, dx, g \in C_c^\infty(\mathbb{R}^{k+1}),$$

$$h_1, \cdots, h_k \in L^2([0, 1]), k \in \mathbb{N} \}. $$

For $g \in C_c^\infty(\mathbb{R}^k)$ denote by $\partial_i g$ the partial derivative of $g$ in direction $i$, $1 \leq i \leq k$. So we may define the Malliavin derivative for smooth cylinder functions by

$$D_s F = \sum_{i=1}^k \partial_i g(W(h_1), \cdots, W(h_k), x) \, dx \, h_i(s), \quad s \in [0, 1].$$

We consider $DF$ as an element of $L^2(\Omega \times [0, 1], \mathbf{M})$ with respect to the Banach space topology. We obviously have in terms of the Malliavin derivative of real valued functions

$$\langle DF, f \rangle = D(\langle F, f \rangle), \quad f \in C_b(\mathbb{R}),$$

and hence also

$$DF = \Phi^{-1}((D(\langle F, f_i \rangle))_{i \in \mathbb{N}}).$$

We next introduce a norm on $\mathcal{S}(\mathbf{M})$. For $F \in \mathcal{S}(\mathbf{M})$ let

$$||F||_{1,2} = \left[ E(|F|^2)^{\frac{1}{2}} + E(||DF||^2)^{\frac{1}{2}} \right].$$

Note that by definition, we have indeed $||F||_{1,2} < \infty$ for $F \in \mathcal{S}(\mathbf{M})$. Hence the closure $\mathbf{D}_{2,1}(\mathbf{M})$ of $\mathcal{S}(\mathbf{M})$ with respect to $||\cdot||_{1,2}$ is well defined and nontrivial.

In a similar way, we may define $||\cdot||_{1,p}, p \geq 1$, and $\mathbf{D}_{p,1}(\mathbf{M})$ by replacing the 2–norm by the $p$–norm, as well as for higher derivatives the norms $||\cdot||_{k,p}$ and spaces $\mathbf{D}_{p,k}(\mathbf{M})$, $k \in \mathbb{N}, p \geq 1$.

In this setting, a measure valued version of the Clark-Ocone formula holds and is crucial for our extension of Jacob’s framework.
Theorem 3.1 Let $F \in D_{2,1}(M)$. Then we have

$$F = E(F) + \int_0^1 E(D_s F | \mathcal{F}_s) \, dW_s.$$  

As in the scalar case, we have to specialize this formula to measure valued martingales, given by the conditional laws of our random variable $G$. In order to minimize regularity conditions for $G$, we start with smoothing the variable by some Gaussian unit variable $N$ independent of $\mathbb{F}$, and set for $\epsilon > 0$

$$G_\epsilon = G + \sqrt{\epsilon}N.$$

Let $P^\epsilon_t$, be the family of conditional laws of $G_\epsilon$ given $\mathcal{F}_t, t \in [0,1]$, and suppose that $G \in D_{2,1}$. Then it is easy to see ([22], section 1) by an explicit calculation that $t \mapsto D_t P^\epsilon_t(\cdot, dl) \in L^2([0,1]; M)$ and

$$P^\epsilon_t(\cdot, dl) = P^\epsilon_0(\cdot, dl) + \int_0^t D_u P^\epsilon_u(\cdot, dl) \, dW_u, \quad t \in [0,1]. \quad (22)$$

Taking the smoothing parameter to 0, we get the following minimal version of an extension of the Clark-Ocone formula to the conditional laws.

Theorem 3.2 Suppose that there exists an $M$-valued process $k_t(\cdot, dx), t \in [0,1]$, denoted by

$$D_t P^G_t(\cdot, dl) = k_t(\cdot, dl), \quad t \in [0,1],$$

such that for any $t \in [0,1], f \in C_b(\mathbb{R})$ we have

$$E\left( \int_0^t \langle [D_s P^G_s(\cdot, dx) - D_s P^G_s(\cdot, dx)], f \rangle \, ds \right) \to 0 \quad (23)$$

as $\epsilon \to 0$, and

$$\sup_{f \in C_b(\mathbb{R}): \|f\| \leq 1} E\left( \int_0^t \langle D_s P^G_s(\cdot, dx), f \rangle \, ds \right) < \infty, \quad (24)$$

then for any $t \in [0,1]$

$$P^G_t(\cdot, dx) = P^G_0(\cdot, dx) + \int_0^t D_s P^G_s(\cdot, dx) \, dW_s. \quad (25)$$

Suppose now that we are in the situation of Theorem 3.2. We may then finally extend Jacod’s framework of the preceding section by replacing condition (9) by the condition

$$D_t P^G_t(\cdot, dl) \text{ is absolutely continuous with respect to } P^G_t(\cdot, dl) \text{ P-a.s. for } t \in [0,1]. \quad (26)$$

If (24) is satisfied, let

$$g_t(\cdot, l) = \frac{d}{d} D_t P^G_t(\cdot, dl) \left( \frac{d}{d} P^G_t(\cdot, dl) \right)(l), \quad t \in [0,1], l \in \mathbb{R},$$

as usual taken to be measurable in all variables. Then the obvious extension of Theorem 2.2 gives the following formula for the information drift $\mu^G$.
**Theorem 3.3** Suppose that (24) is satisfied, and furthermore that
\[ \mu_t^G = g_t(\cdot, G) \]  \hspace{1cm} (27)
satisfies
\[ \int_0^1 |\mu_u^G| du < \infty \quad P-a.s. \]  \hspace{1cm} (28)
Then
\[ W = \tilde{W} + \int_0^1 \mu_s^G ds \]
is a $\mathcal{G}$–semimartingale with a $\mathcal{G}$–Brownian motion $\tilde{W}$.

How much more general is the situation described by Theorem 3.3? It is proved in [22] that
\[ E(\int_0^1 g_t^2(\cdot, G) ds) < \infty \quad \text{implies} \quad P_t^G(\cdot, dl) \quad \text{is absolutely continuous for} \quad P^G, \]
$P-a.s., t \in [0, 1]$, and moreover
\[ E(\exp(\frac{1}{2} \int_0^1 g_t^2(\cdot, G) ds)) < \infty \quad \text{implies} \quad P_t^G(\cdot, dl) \quad \text{is equivalent with} \quad P^G, \]
$P-a.s., t \in [0, 1]$.

4 **Arbitrage possibilities for the insider**

Let us now return to the question raised in the context of Example 2 in the preceding section. We will show that for quite general regular diffusion processes describing the evolution of the stock price $X$ there are riskless free lunches, and even arbitrage.

**Example 1:**
Let the stock price process be a regular diffusion given by a stochastic differential equation with $\sigma_t = \sigma(X_t), \quad \alpha_t = \alpha(X_t), \quad t \in [0, 1]$, where $\sigma, \alpha$ are bounded, smooth functions on $\mathbb{R}$, $\sigma$ in addition without zeroes. Denote by $P_x$ the law of the diffusion starting at $x \in \mathbb{R}$, by $\nu_y$ the first time to reach $y \in \mathbb{R}$, and by $s$ the strictly increasing scale function. Moreover, let $G = \sup_{t \in [0, 1]} X_t$. For $\epsilon > 0$ and a standard Gaussian variable $N$ independent of $\mathbb{F}$ let
\[ G_\epsilon = G + \sqrt{\epsilon} N. \]
Fix $t \in [0, 1]$. Then the conditional law of $G_\epsilon$ given $\mathcal{F}_t$ possesses a density given by
\[ y \mapsto E(p_\epsilon(y - G)|\mathcal{F}_t), \]
where $p_\epsilon$ is the Gaussian density of $\sqrt{\epsilon} N$. Now fix also $x \in \mathbb{R}$ and abbreviate $\phi = p_\epsilon(x - \cdot)$. Set $G_t = \sup_{0 \leq s \leq t} X_s$, note that the process $((G_t, X_t))_{t \in [0, 1]}$ is a Markov
process, and let $q_t(x, y), x, y \in \mathbb{R}$ be the probability density of $G_t$ given that $X_0 = x$, which exists due to regularity. Then we have

$$E(\phi(G)|\mathcal{F}_t) = E(\phi(G_t) 1_{\{G_t > \sup_{t \leq s \leq 1} X_s\}}|\mathcal{F}_t) + E(\phi(\sup_{t \leq s \leq 1} X_s) 1_{\{G_t \leq \sup_{t \leq s \leq 1} X_s\}}|\mathcal{F}_t)$$

$$= \phi(G_t) P_{X_t}(\nu_y \geq 1 - t)|_{y=G_t} + \int_{G_t}^\infty \phi(y) q_{1-t}(X_t, y) \, dy.$$

Let us next compute the Malliavin trace of the conditional density. For this purpose, note that by our assumptions $X$ has a $P$-a.s. unique maximum at some random time $\tau_t$ on any fixed interval $[0, t], t \in [0, 1]$, the law of which is absolutely continuous with respect to Lebesgue measure there (see [7]). To denote the time of the maximum on the whole interval $[0, 1]$, we also write $\tau$ instead of $\tau_t$. We first remark that by smoothness of $\sigma$ and $\alpha$, $X_t \in \mathbb{D}_{2,1}, t \in [0, 1]$. Also, by Nualart and Vives [33], we know that $G_t \in \mathbb{D}_{2,1}$ and that its Malliavin derivative is supported on the random interval $[0, \tau_t]$. Since $\tau_t < t$ for $t > 0$ $P$-a.s., we deduce that $D_t G_t = 0$. This in turn implies

$$\frac{d}{d y} E(\phi(G)|\mathcal{F}_t)$$

$$= \phi(G_t) \frac{\partial}{\partial x} P_{X_t}(\nu_y \geq 1 - t)|_{y=G_t} D_t X_t \, dy$$

$$+ \int_{G_t}^\infty \phi(y') \frac{\partial}{\partial x} q_{1-t}(X_t, y') \, dy' D_t X_t \, dy.$$

This crucial observation indeed means that as $\epsilon \to 0$ the Malliavin trace remains a measure which is comparable to the conditional density. So the criterion derived in the preceding section applies, and shows that this type of important example becomes tractable in this framework. By integrability properties of $\frac{\partial}{\partial x} q_{1-t}(x, y)$ the convergence of (30) to

$$D_t P_t(\cdot, dx)$$

$$= \delta_{G_t}(dx) \frac{\partial}{\partial x} P_{X_t}(\nu_y \geq 1 - t)|_{y=G_t} D_t X_t$$

$$+ \int_{G_t}^\infty \frac{\partial}{\partial x} q_{1-t}(X_t, x) \, dx \, D_t X_t,$$

is dominated, so that

$$E(\int_0^t \langle [D_s P_s^f(\cdot, dx) - D_s P_s(\cdot, dx)], f \rangle^2 \, ds \to 0$$

for any $f \in C_b(\mathbb{R})$. By similar arguments, we obtain

$$P_t(\cdot, dx)$$

$$= \delta_{G_t}(dx) \, P_{X_t}(\nu_y \geq 1 - t)|_{y=G_t}$$

$$+ \int_{G_t}^\infty q_{1-t}(X_t, x) \, dx.$$
The explicit representation of the density in [7] also proves that

$$\sup_{f \in C_b(\mathbb{R}), \|f\| \leq 1} E \left( \int_0^t \langle D_s P_s(\cdot, dx), f \rangle^2 ds \right) < \infty,$$

(34)

for $0 \leq t < 1$.

(31) and (33) immediately show that (26) is satisfied and that the information drift corresponding to $G$ is given by

$$\mu^G_t = 1_{G_t}(G) \frac{\partial}{\partial x} \ln P_{X_t}(\nu_y \geq 1 - t)|_{y= G D_t X_t}$$

(35)

$$+ \frac{\partial}{\partial x} q_{1-t}(X_t, G) \langle D_t X_t \rangle$$

$$= 1_{[\tau, 1]}(t) \frac{\partial}{\partial x} \ln P_{X_t}(\nu_y \geq 1 - t)|_{y= G D_t X_t}$$

$$+ \frac{\partial}{\partial x} q_{1-t}(X_t, G) \langle D_t X_t \rangle,$$

$0 \leq t < 1$.

Let us now investigate the integrability properties of the information drift on $[0, 1]$. Note that $D_t X_t = \sigma(X_t) X_t$, $t \in [0, 1]$. The essential contribution comes from the first term in the last line of (35) in a small neighborhood of $\tau$.

To estimate this contribution, let $x < y$, and denote by $n^x_\tau$ the Levy measure of the diffusion on level $x$ (see Itô-McKean [24], p. 214). Then as $x \uparrow y$, to first order $P_x(\nu_y \geq 1 - t)$ behaves like

$$n^x_\tau (1 - t) \frac{|s(y) - s(x)|}{s'(x)},$$

alternatively

$$n^x_\tau (1 - t) |y - x|.$$

Hence, to first order, the largest term in $\mu^G_t$ behaves like

$$\frac{\sigma(X_t)}{G - X_t}$$

as $t \downarrow \tau$. An eventual change of time scale therefore teaches us that $\mu^G$ has the same integrability properties as

$$\frac{1}{S - W_t}$$

on $[\rho, 1]$, where $S = \sup_{t \in [0, 1]} W_t$ and $\rho$ the random time at which $W$ reaches its maximum. This takes our investigation into basic excursion theory for Brownian motion. If we denote by $\rho_0$ the first zero of $W$ after $\rho$, Revuz, Yor [36], Proposition VI.3.13, p. 238, prove that between $\rho$ and $\rho_0$, $W$ behaves like a three dimensional Bessel process. The trajectorial properties of Bessel processes, also described in Revuz, Yor [36], then imply (see [22], section 2)
Theorem 4.1 Let $X$ be a regular diffusion given by a stochastic differential equation with $\sigma_t = \sigma(X_t)$, $\alpha_t = \alpha(X_t)$, $t \in [0,1]$, where $\sigma, \alpha$ are bounded, smooth functions on $\mathbb{R}$, $\sigma$ in addition without zeroes. Moreover, let $G = \sup_{t \in [0,1]} X_t$. Then the information drift $\mu^G$ possesses the following properties

$$\int_0^1 |\mu^G_s| \, ds < \infty \quad P-a.s.,$$

(36)

$$\int_0^1 (\mu^G_s)^2 \, ds = \infty \quad \text{with positive probability for each } 0 < t \leq 1.$$  

(37)

Theorem 1 on the one hand shows that for the additional knowledge of the supremum of a regular diffusion the information drift is tame enough to keep $W$ a semimartingale in the big filtration $G$. But to allow for an equivalent martingale measure under which $W$ is a martingale in the big filtration $G$, $\mu^G$ is obviously too irregular. In order to obtain such a measure, we would have to eliminate the whole effective drift given by (see section 1) $\theta = \frac{\sigma}{\sigma} + \mu^G$. But this is excluded by Theorem 4.1.

Theorem 4.2 Let $\theta = \frac{\sigma}{\sigma} + \mu^G$ with $G$ according to Theorem 4.1. Then we have

$$\int_0^t \theta^2_s \, ds = \infty \quad \text{with positive probability for each } 0 < t \leq 1.$$  

(38)

In particular, in the model of the insider the condition (NFLVR) is violated. Consequently, the insider possesses arbitrage opportunities.

Proof: Write $c_t = \frac{\sigma_t}{\sigma_t}$ and suppose for simplicity $t = 1$. Let us assume that

$$\int_0^1 \theta^2_t \, dt < \infty$$

(39)

$P-a.s.$ Note that $c$ is $\mathbb{F}$-adapted. Hence, by conditioning on $\mathcal{F}_t$ we get for $t \in [0,1]$

$$\|c_t\| \leq E(\|\theta_t\| | \mathcal{F}_t) + E(|\mu^G_t| | \mathcal{F}_t).$$

(40)

If we can show

$$\int_0^1 (E(|\mu^G_t| | \mathcal{F}_t))^2 \, dt < \infty,$$

(41)

(40) and Jensen’s inequality imply $\int_0^1 c^2_t \, dt < \infty$, and therefore the contradiction

$$\int_0^1 (\mu^G_t)^2 \, dt \leq 4 \int_0^1 c^2_t \, dt + \int_0^1 \theta^2_t \, dt < \infty \quad P-a.s.$$

Now note that by the very definition of $\mu^G$ we have for $t \in [0,1]$

$$E(|\mu^G_t| | \mathcal{F}_t) = |D_t P^G_t(\cdot, dl)|,$$

so that we obtain the estimate

$$\int_0^1 (E(|\mu^G_t| | \mathcal{F}_t))^2 \, dt \leq \int_0^1 |D_t P^G_t(\cdot, dl)|^2 \, dt.$$

But this is clearly finite $P-a.s.$ due to (31). This proves (39) and thus the proof is complete. □
References


[23] Imkeller, P. *Random times at which insiders can have free lunches.* To appear in ”Stochastics and Stochastics Reports” (2002).


