Bollettino
Unione Matematica Italiana

P. L. Lions

On mathematical finance

Bollettino dell'Unione Matematica Italiana, Serie 8, Vol. **3-B** (2000), n.3, p. 553–572.

Unione Matematica Italiana

<http://www.bdim.eu/item?id=BUMI_2000_8_3B_3_553_0>

Bollettino dell'Unione Matematica Italiana, Unione Matematica Italiana, 2000.

Bollettino U. M. I.
(8) 3-B (2000), 553-572

On Mathematical Finance.

P. L. LIONS (*)

Sunto. – I. Introduzione.
II. Un rapido esame di modelli e strumenti.
III. Simulazioni Monte-Carlo efficienti e calcolo di Malliavin.
IV. Osservazioni parziali e «option pricing».

I. – Introduction.

We present here a brief survey of some issues in Mathematical Finance, concentrating mostly on rather practical problems such as modelling and numerical issues and ignoring some more theoretical issues.

As is well-known, the industry of options (also called derivatives, securities ...) has become a huge one with many «players» consisting not only of financial institutions but also of financial divisions in the major companies worldwide. New products are being designed, almost constantly, and now areas expand rapidly such as credit risk options or options on utilities (gaz, electricity...). Since the pioneering word of F. Black and M. Scholes [4], R. Merton [16], mathematical models are being used and developed by many research groups for option pricing and hedging (defined more precisely in section II below), leading to numerical software on which traders often base their activities. The industry of derivatives is clearly a high technology one and it is thus not surprising to see that models, numerical methods (and computer systems) are becoming more and more sophisticated with the input of thousands of mathematically trained financial engineers. Of course, mathematical models and numerical methods will probably never suffice to yield a complete description of the derivatives market, but their overall efficiency and precision is quite remarkable. A general argument for the mathematical approach can be made once we recall that, by analogy with various other sciences, partial differential equations are often efficient to describe, in an average fashion, the collective dynamical behavior of large numbers of interacting particles. In the financial

(*) Conferenza tenuta a Napoli il 13 settembre 1999 in occasione del XVI Congresso U.M.I.

applications, one can then argue (somewhat vaguely) that the particles are the agents (or contracts) which interact by trading (1).

We describe in section II below the classical theory of option pricing and some more recent models or products together with some of the issues concerning them, such as, for instance, the issue of numerical computations. In section III, we present briefly some work taken from [9], [8] that allows for efficient Monte-Carlo computations of hedges. Finally, in section IV, we illustrate a general approach, developed in collaboration with J.-M. Lasry [12], to the issue of option pricing and hedging with partial observations of informations. And we do so in the context of rather typical financial problems.

II. – A quick tour of models and issues.

Of course, this section is a biased selection of topics and the interested reader should consult various references that present the field in more details and more completely, such as D. Duffie [6], [7], M. Musiela and M. Rutkowski [17], R. Rebonato [19] ...

II.1. *Classical Black-Scholes theory revisited.*

We denote by S_t the value of some assets (equities, interest rates ...) at time t and we assume that S_t solves the following stochastic differential equation for $t \geqslant 0$.

$$dS_t = \sigma(S_t).\, dW_t + b(S_t)\, dt\,, \qquad S_0 = S \in R^N \tag{1}$$

where W_t is a standard Brownian motion in R^m and σ, b are functions of S that we take to be time-independent in order to simplify the presentation. Also, we shall not bother to state the precise regularity and growths conditions on (σ, b) needed in this section. Of course, typical examples of (1) are the *Brownian case* where σ and b are constants that we still denote by σ and b, and $S_t = S_0 + \sigma.\, W_t + bt$, and the *log-normal case* where $\frac{\sigma}{S}$ and $\frac{b}{S}$ are constants and $S_t = S_0 \exp\left\{\sigma W_t + \left(b - \frac{\sigma^2}{2}\right) t\right\}$. Other explicit «gaussian cases» are used in Finance together with models for which (1) cannot be solved «explicitely».

A typical example of an option is a contract between two parties, one of which pays at $t = 0$ some price for the future payment by the other party at some specified time $T > 0$ (the maturity of the option) of $\Phi(S_T)$ where Φ is a specified function of S. Once more, in order to simplify the notation and presenta-

(1) It is rather striking to recall that traders call the most commonly exchanged products liquids products (although no relationship with Fluid Mechanics is to be expected).

tion, we shall take the interest rate to be 0. Then, the obvious questions of interest are: i) what is the *price of this option* i.e. how much should the first party be asked, ii) how should the second party manage this contract i.e. how should he invest (sell or buy) in S_t in order to protect himself from the random fluctuations of $\Phi(S_T)$ (this is the so-called *hedging* question).

Under various natural assumptions (complete market, no arbitrage, continuous time hedging ...) that we do not wish to describe here, these problems were solved in [4], [16]. Various presentations are possible and we follow one which is based upon optimal control theory (and utility functions) since this is basically the only one that remains once we discard one of the aforementioned assumptions. This is why we introduce the wealth process

$$dP_t = \alpha_t \cdot dS_t \quad for\ t \geq 0\,, \quad P_0 = P \in R \tag{2}$$

where P_t stands for the wealth (of the second party), α_t is the hedge or in other words the number of assets owned during the «interval $(t, t+dt)$», and obviously we may choose α_t as we wish among all processes (adapted to the filtration generated by W) with apropriate bounds that we do not specify here. At time T, the wealth will be $P_T - \Phi(S_T)$ and we measure the «quality» of the hedge by the following expected utility

$$E[U(P_T - \Phi(S_T))] \tag{3}$$

where U is an arbitrary utility function that we assume to be, for instance, continuous, concave and increasing on R. (In fact, the result mentioned below is true for much more general functions U but we shall not detail this point here ...). And we are thus interested in the following stochastic control problem

$$V = \operatorname*{Max}_{\alpha_t} E[U(P_T - \Phi(S_T)] \tag{4}$$

and V is obviously a function of (P, S, T): $V = V(P, S, T)$.

Pricing and hedging the option many then be formulated as follows: in (2), P stands for the total wealth at time $t = 0$ which, of course, contains the price paid by the first party. This is why, it is natural to consider an auxiliary control problem, corresponding to the case when no option is bought or sold, namely

$$V_0(P, S, T) = \max_{\alpha_t} E[U(P_T)]\,. \tag{5}$$

Then, the «fair» price of the option is the quantity, denoted by u, such that

$$V_0(P-u, S, T) = V(P, S, T)\,, \tag{6}$$

and the optimal hedge is $\delta_t = \overline{\alpha}_t - \overline{\alpha}_t^0$ where $\overline{\alpha}_t$, $\overline{\alpha}_t^0$ are respectively maximizers of

(4), (5) respectively (provided, of course, all these objects $u, \overline{\alpha}, \overline{\alpha}^0$ exist ...).

It turns out that u is entirely determined by the following expression

$$u(S,\,T) = E[\Phi(S_T^0)]\,, \tag{7}$$

which is thus independent of P and U, and

$$\delta_t = \delta(S,\,t)\,, \qquad \delta = \frac{\partial u}{\partial S}(S,\,t)\,. \tag{8}$$

In the above formula, S_t^0 stands for the solution of (1) with $b=0$ and, infact, we can also write (7) with S_t solving (1) but, then, we need to replace E (i.e. the probability) by a new one according to the classical Girsanov formula. This change of probability is an important notion in Mathematical Finance but we shall not develop this aspect here.

An easy proof of (7)-(8) can be made using the Hamilton-Jacobi-Bellman equations associated to (4) and (5). And we just sketch a formal proof of (7) (which also yields (8) with little more work ...): we recall first that V and V_0 solve the following equation, where we denote by $a = \sigma\sigma^T$

$$\begin{gathered} \frac{\partial W}{\partial T} - \frac{1}{2} Tr\left(a\cdot\frac{\partial^2 W}{\partial S^2}\right) - b\cdot\frac{\partial W}{\partial S} + \frac{1}{2}\left(\frac{\partial^2 W}{\partial P^2}\right)^{-1}\cdot \\ \cdot\left(a\cdot\frac{\partial^2 W}{\partial S\partial P} + b\frac{\partial W}{\partial P}\right)^T\cdot a^{-1}\cdot\left(a\cdot\frac{\partial^2 W}{\partial S\partial P} + b\frac{\partial W}{\partial P}\right) = 0\,, \end{gathered} \tag{9}$$

and, of course, $V|_{T=0} = U(P\cdot\Phi(S))$, $V_0\,|_{T=0} = U(P)$.

Next, a straightforward computation shows that $V_0(P-u(S,\,T),\,S,\,T)$, also solves (9) and satisfies the same initial condition than V if and only if u solves

$$\frac{\partial u}{\partial T} - \frac{1}{2} Tr\left(a\cdot\frac{\partial^2 u}{\partial S^2}\right) = 0\,, \qquad u|_{T=0} = \Phi(S)\,. \tag{10}$$

And (7) is nothing but the probabilistic representation of the solution of (10).

REMARK. – The same facts hold for american options which can be exercised at any (stopping) time τ in $[O,\,T]$. Then, the optimal hedging formula remains the same ((8)) while the price is given by

$$u(S,\,T) = \max_{0\leq\tau\leq T} E[\Phi(S_\tau)] \tag{11}$$

which solves uniquely the following obstacle problem (variational inequality ...)

$$\begin{cases} \operatorname{Min}\left(\dfrac{\partial u}{\partial T} - \dfrac{1}{2}\operatorname{Tr}\left(a \cdot \dfrac{\partial^2 u}{\partial S^2}\right), u - \Phi\right) = 0 \\ u|_{T=0} = \Phi(S). \end{cases} \tag{12}$$

Historically, the first models that have been used were one-dimensional, mostly log-normal models (but also some Gaussian models for interest rates) which make possible to derive *explicit representations* of the price u and of the hedge $\frac{\partial u}{\partial S}$ (also called *delta*).

These formulae are even particularly simple for the most common products such as calls ($\Phi(S) = (S-K)_+$), puts ($\Phi(S) = (K-S)_+$), digitals ($\Phi(S) = 1_{(S \geqslant K)}$) ..., or even bareers which are slightly more elaborate contracts and whose prices essentially solve (10) with appropriate Dirichlet boundary conditions on some line ($S = B$). However, no such formula exists in the case of american options, which were solved numerically by tree methods i.e. explicit finite difference methods (see for instance D. Lamberton and B. Lapeyre [10]).

II.2. *Increasing dimension.*

As we saw above, determining the price of an option amounts to solve a linear parabolic equation (or an obstacle problem for american options). From a numerical viewpoint, this is an easy matter in dimension $N = 1$ and, in addition, for many simple models explicit or semi-explicit formulae are often available. However, it is clear by now that one has to consider situations in higher dimensions ($N \geqslant 2$). One can list a few reasons for such a necessity

i) Higher dimensional models

More accurate models involve at least two dimensions: this is the case for stochastic volatility models such as

$$\begin{cases} dS_t = \sigma_t S_t \, dW_t \\ d\sigma_t = \nu \sigma_t \, dB_t + a(\sigma_t - \overline{\sigma})\, dt \end{cases} \tag{13}$$

where ν, a, $\overline{\sigma}$ are positive constants and B_t is a Brownian motion with a fixed correlation parameter ϱ with W_t. Exactly as before, we obtain a price $u(S, \sigma, T)$ which solves

$$\frac{\partial u}{\partial T} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 u}{\partial S^2} - \frac{\nu^2}{2}\sigma^2 \frac{\partial^2 u}{\partial \sigma^2} - \varrho \nu \sigma^2 S \frac{\partial^2 u}{\partial \sigma \partial S} - a(\sigma - \overline{\sigma})\frac{\partial u}{\partial \sigma} = 0\,, \tag{14}$$

a two dimensional parabolic equation which becomes degenerate (and delicate to solve numerically) if $|\varrho|$ is close to 1. Other examples include two-factors interest rate models ...

ii) Options involving several assets

Typical examples include options on baskets, or worst-of-N. We then have N one-dimensional assets $S_t^1, \ldots, S_t^N$ and the pay off Φ depends of $(S_t^1, \ldots, S_t^N) = S_t$ (for instance of $\frac{1}{N}\sum_{i=1}^{N} a_i S_t^i$, or $\min_{i \leq i \leq N} (a_i S_t^i)$ for some weights $a_1, \ldots, a_N > 0$). In such examples, N may reach 10 or even be larger ...!

iii) Exotic and path-dependent options

Even if the underlying asset S_t is one-dimensional, the form of the pay-off may lead to equations in higher dimensions. This is the case with exotic options such as asian options where $\Phi = \Phi\left(S_T, \int_0^T S_t\,dt\right)$ or look-backs where $\Phi = \Phi\,(S_T, \max_{0 \leq t \leq T} S_t)$ (for instance ...). In those two examples, the price (and the hedge) of the option is determined by the solution of a two-dimensional equation namely $u(S, T) = u(S, I, T)\,|_{I=0}$ and

$$\frac{\partial u}{\partial T} - \frac{1}{2}\sigma^2(S)\frac{\partial^2 u}{\partial S^2} - S\frac{\partial u}{\partial I} = 0\,, \qquad u_{|T=0} = \Phi(S, I) \tag{15}$$

in the case of «asian options», and $u(S, T) = u(S, M, T)\,|_{M=S}$ and

$$\frac{\partial u}{\partial T} - \frac{1}{2}\sigma^2(S)\frac{\partial^2 u}{\partial S^2} = 0 \quad \textit{for } S < M, \ \frac{\partial u}{\partial M} = 0 \ \textit{at } S = M, \ u_{|T=0} = \Phi(S, M)\,, \tag{16}$$

in the case of «look-backs» (see G. Barles, E. Daher and M. Romano [2]).

These two examples show how the addition of one rule (the integral or the maximum) in the contract leads to one more dimension for the associated partial differential equation. This explains why the so called path-dependent options which depend in a non trivial way on the path $(t \mapsto S_t)$ are often untreatable by PDE techniques (this is the case for «structured» products like Mortgage Backed Securities, CMO, IAB ...). Numerical evaluations can then only be made by Monte-Carlo simulations.

II.3. *Nonlinear equations.*

It is reasonable to expect that more and more financial situations will require the case of nonlinear (partial differential) equations. This is in fact already the case for american options (see II.1 above) or related objects, or for «passport options» which lead intrinsically to stochastic control problems and their corresponding Hamilton-Jacobi-Bellman equations. Another reason for the occurence of nonlinear equations is to take into account market imperfections (and their impact on pricing and hedging) such as imperfect volatility calibration (see the worst case model by M. Avellaneda, A. Paras and A. Levy [1], or the partial observation model by J.-M. Lasry and the author [11]), market frictions (transaction costs for instance, see the asymptotic model by G. Barles and M.H. Soner [3]) or discrete hedging in order to take into account the fact that continuous hedging is not realistic (see the model by J.-M. Lasry and P.-L. Lions [11]). For such nonlinear equations, the issues are modelling and numerical ones but not a theoretical one since the theory of *viscosity solutions* provides the desired mathematical framework. It might be worth pointing out that this to be expected since viscosity solutions theory has been designed to take care of all equations enjoying maximum and comparison principle properties. And this is a natural fact in Finance since two options (with the same maturity) with pay-offs $\Phi_1 \geqslant \Phi_2$ are to be priced with the same ordering ($u_1 \geqslant u_2$)!

II.4. *Numerical perspective.*

We shall not discuss here an important practical issue namely the calibration of models i.e. the numerical determination of the various paremeters entering the chosen models in view of observed market prices. This type of problem falls into the classical field of inverse problems which are, in general, quite delicate.

Leaving aside the specific difficulties associated with calibration, the main numerical issues concern, of course, the computation of the prices and of the hedges. At this point, one needs to make a distinction between deterministic methods such as trees (=explicit finite differences) or more generally finite differences or finite elements, and Monte-Carlo methods. For low dimensions ($N = 1$ or 2, possibly 3...), the former are obviously preferred because of their efficiency (speed and precision) and because they allow to treat american options or, more generally, nonlinear equations. On the other hand, they degrade as N increases and, in particular, are useless for «really» path-dependent products. This is where Monte-Carlo methods are used: their advantages being the simplicity of implementation, their intrinsic parallel structure and the possibility of computing in «high dimensions». The drawbacks of Monte-Carlo methods are the oscillating nature of the (slow) convergence to the desired result (prices and hedges) and the fact that they cannot be used, with classical approaches, to compute the solution of nonlinear equations. Various tricks or recipes have been proposed to cure at least part of the first draw back (variance reduction, imputance sampling, control variables, blocks counting ...) and we wish to mention one aspect namely the generation of Brownian paths.

Let us first recall how classical Monte-Carlo simulations are performed (for european options) in order to approximate, say, the price (for instance $u(S, T)$ given by (7)). Given a time discretization Δt, one generates $N \times M$ (with $M = T/\Delta t$) independent centered Gaussian variables with variance Δt in order to approximate the increments of a N-dimensional Brownian motion $(W_{j\Delta t} - W_{(j-1)\Delta t}, 1 \leqslant j \leqslant M)$. Then, one solves (1) by a discretization scheme (many are possible, Euler for instance) and generates $S_{j\Delta t}$ for $1 \leqslant j \leqslant M$. Along this discretized path, one then computes the pay off. Doing this computation for n paths (i.e. generating $n \times N \times M$ independent Gaussian variables) one obtains n values $(\Phi_i)_{1 \leqslant i \leqslant n}$. And, finally, an approximation of the price namely

$$\frac{1}{n} \sum_{i=1}^{n} \Phi_i .$$

The typical rate of convergence (neglecting the time discretization error) of this approximation is $1/\sqrt{n}$ if one use «randomly generated» Gaussian variables and is thus rather slow. On the other hand, one can generate these variables using the so-called low discrepancy sequences which are much more effective in low dimensions, the dimension is in our case typically $n \times N \times M$ or $n \times N$ in some particular cases and is thus in general quite large. Various groups, including ours — see also [5] —, advocate the use of mixed generations of Brownian paths by, for example, generating through low discrepancy sequences intermediate points (midpoints for instance) connected by randomly generated Brownian bridges. All these tricks and recipes are quite useful for practical computations but we shall not attempt to detail them more here. The next section is devoted to a more conceptual improvement of Monte-Carlo simulations for the computation of hedges i.e. the «delta» $\partial u/\partial s$ and more generally all sensitivities of the price with respect to important parameters, that are called in the financial jargon «greeks», like the «gamma» $\partial^2 u/\partial S^2$, «vega» $\partial u/\partial v$...

III. – Efficient Monte-Carlo simulations and Malliavin calculus.

The contents of this section are taken from [9] and [8].

III.1. *Position of the problem.*

As we saw in the preceding section, we are mainly interested in the following quantities

$$u_\lambda = E[\Phi(S(T, \lambda), \lambda)], \qquad \frac{\partial^m u_\lambda}{\partial \lambda^m} = \frac{\partial^m}{\partial \lambda^m} E[\Phi(S(T, \lambda))] \tag{17}$$

where $m \geqslant 1$ and λ is a parameter (say in R) and $S_t = S(t, \lambda)$ solves

$$dS_t = \sigma(S_t, \lambda) \cdot dW_t, \qquad S_0 = S \in R^N. \tag{18}$$

Once more, we do not make precise the regularity and growths assumptions on σ and Φ in S and λ (that we may take as smooth as we wish ...). A typical example is: $\lambda = S$ (if $N = 1$, say), in which case $\partial^m u / \partial \lambda^m$ is the «delta» if $n = 1$ and the «gamma» if $n = 2$.

The computation, by Monte-Carlo simulations, of $\partial u_\lambda / \partial \lambda$ (or $\partial^2 u_\lambda / \partial \lambda^2$) is classically performed by difference quotients i.e. by computing (by a Monte-Carlo procedure) the following expectation (for instance)

$$E\left[\frac{1}{h} \{ \Phi(S(T, \lambda + h)) - \Phi(S(T, \lambda)) \} \right]$$

for an appropriate «small» h (which has to be chosen well ...). Obviously, this difference quotient is a (crude) approximation of the following expression

$$\frac{\partial u_\lambda}{\partial \lambda} = E\left[\frac{\partial \Phi}{\partial \lambda}(S_T, \lambda) + \frac{\partial \Phi}{\partial S}(S_T, \lambda) \frac{\partial S_T}{\partial \lambda} \right] \tag{19}$$

where $\dfrac{\partial S_T}{\partial \lambda}$ solves the following affine stochastic differential equation

$$d\left(\frac{\partial S_t}{\partial \lambda} \right) = \left[\frac{\partial \sigma}{\partial S}(S_t, \lambda) \frac{\partial S_t}{\partial \lambda} + \frac{\partial \sigma}{\partial \lambda}(S_t, \lambda) \right] \cdot dW_t. \tag{20}$$

In fact, (19) makes sense if Φ has some regularity in S (for instance, $\Phi \in C^1$ and $\partial \Phi / \partial S$ bounded, or even Φ Lipschitz if σ is nondegenerate ...). And we also have

$$\frac{\partial^2 u_\lambda}{\partial \lambda^2} = E\left[\frac{\partial^2 \Phi}{\partial \lambda^2} + 2 \frac{\partial^2 \Phi}{\partial S \partial \lambda} \frac{\partial S_T}{\partial \lambda} + \frac{\partial \Phi}{\partial S} \frac{\partial^2 S_T}{\partial \lambda^2} \right] \tag{21}$$

with

$$d\left(\frac{\partial S_t^2}{\partial \lambda^2} \right) = \left[\frac{\partial \sigma}{\partial S} \frac{\partial^2 S_t}{\partial \lambda} + \frac{2 \partial^2 \sigma}{\partial S \partial \lambda} \frac{\partial S_t}{\partial \lambda} + \frac{\partial^2 \sigma}{\partial \lambda^2} \right] \cdot dW_t. \tag{22}$$

The regularity of Φ in S is a serious difficulty since $\Phi(S) = 1_{(S > K)}$ for a digital and $\Phi(S) = (S - K)_+$ for a call (or $(K - S)_+$ for a put) and thus

$$\frac{\partial \Phi}{\partial S} = \delta_K(S) \text{ for a digital}$$

$$\frac{\partial^2 \Phi}{\partial S^2} = \delta_K(S) \text{ for a call, } -\delta_k(S) \text{ for a put.}$$

In particular, for a digital, the approximation by difference quotients of $\partial u_\lambda/\partial\lambda$ amounts to the following quantity, in the simple example where $\lambda = S$, $S_t = S + \sigma W_t$,

$$E\left[\frac{1}{h}1_{(0<S_T-S<h)}\right]$$

which is clearly an unstable quantity to compute by a Monte-Carlo simulation (if h is small, otherwise the expectation is not necessarily close to the desired quantity). This shows why classical Monte-Carlo methods to compute grecks are not efficient and we shall see below that Malliavin calculus provides a cure by allowing to integrate by parts within the expectation.

III.2. *Malliavin calculus and integration by parts.*

We just present the basic facts of Malliavin derivatives and integration by parts and we refer the reader interested in more details to the books by D. Nualart [18] and P. Malliavin [15]. Given a notion of integral, it is natural to expect a notion of derivative. And the notion of Malliavin derivative of random variables (in a Wiener space i.e. in the probability space associated to a standard Brownian motion ...) is naturally associated to the notion of ($It\hat{o}$) stochastic integrals. More precisely, if $N = 1$ in order to simplify notation, the Malliavin derivative D_t, whenever it exists, acts linearly on random variables F, satisfies the chain rule, $(D_tF)_{t\geq 0}$ is an adapted process, $D_tF = 0$ for $t \geq T$ $(T < \infty)$ if F is measurable with respect to $(W_s)_{0\leq s\leq T}$ and D_tF is defined on smooth (dense) random variables of the form $F = \varphi\left(\int_0^\infty h_1(t)\,dW_t, \ldots, \int_0^\infty h_m(t)\,dW_t\right)$ with $\varphi \in C_0^\infty\ (R^m)$, $m \geq 1$, $h_i \in L^2(0,\ \infty)$ $(1 \leq i \leq m)$ by

$$D_tF = \sum_{i=i}^{m} \frac{\partial\varphi}{\partial x_i}(-)\, h_i(t)\,. \tag{23}$$

Denoting by $\mathcal{F}_t$ the σ field generated by $(W_s)_{0\leq s\leq t}$, we have the two following fundamental properties if F is $\mathcal{F}_T$ measurable $(0 < T < \infty)$

$$F = E[F] + \int_0^T E[D_tF \mid \mathcal{F}_t]\,dW_t, \tag{24}$$

and the integration by parts formula

$$E\left[\int_0^T D_tF\alpha_t\,dt\right] = E[F\delta(\alpha)] \tag{25}$$

where $\delta(\alpha)$ is the so-called Skorohod (stochastic) integral that extends to smooth processes the Itô (stochastic) integral. In particular, if α_t is adapted then $\delta(\alpha) = \int_0^T \alpha_t dW_t$.

A simple explanation of (25) is provided by looking at the example when $F = \varphi(W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_T - W_{t_{m-1}})$ where $\varphi \in C_0^\infty(R^m)$, $m \geq 1$, $0 = t_0 < t_1 < t_2 < \ldots < t_{m-1} < t_m = T$, and $\alpha_t = \psi_i(W_{t_1}, \ldots, W_T - W_{t_{m-1}})$ if $t_{i-1} \leq t < t_i$, with ψ_i (say) $\in C_0^\infty(R^m)$. Then, we have

$$E\left[\int_0^T D_t F \alpha_t dt\right] = \sum_{i=1}^m E\left[\frac{\partial\varphi}{\partial x_i}\psi_i\right](t_i - t_{i-1}) =$$

$$\sum_{i=1}^m (t_i - t_{i-1}) \int_{R^m} \frac{\partial\varphi}{\partial y_i}(y)\, \psi_i(y) \exp\left(-\sum_j \frac{y_j^2}{2(t_j - t_{j-1})}\right) \prod_j (2\pi(t_j - t_{j-1}))^{-1/2} dy =$$

$$\sum_{i=1}^m \int_{R^m} \varphi\psi_i y_i \exp\left(-\sum_j \frac{y_j^2}{2(t_j - t_{j-1})}\right) \prod_j (2\pi(t_j - t_{j-1}))^{-1/2} dy -$$

$$\sum_{i=1}^m (t_i - t_{i-1}) \int_{R^m} \varphi(div\,\psi) \exp\left(-\sum_j \frac{y_j^2}{2(t_j - t_{j-1})}\right) \prod_j (2\pi(t_j - t_{j-1}))^{-1/2} dy\,.$$

In particular, if ψ_i depends only on y_k for $k < i$ — which is equivalent to requesting that α_t is adapted —, we have $div\,\psi = 0$ and we obtain

$$E\left[\int_0^T D_t F \alpha_t dt\right] = \sum_{i=1}^m E[\varphi\psi_i(W_{t_i} - W_{t_{i-1}}] = E\left[\varphi\int_0^t \alpha_t dW_t\right].$$

III.3. *Applications to the representation of greeks.*

We may now go back to the expressions (19) (and (21)) and we wish to integrate by parts inside the expectation the term $E\left[\frac{\partial\Psi}{\partial S}(S_T)\frac{\partial S_T}{\partial\lambda}\right]$. The main idea is to find an «integrating factor» α (which is a random process) such that

$$\frac{\partial\Phi}{\partial S}(S_T)\,\frac{\partial S_t}{\partial\lambda} = \int_0^T D_t\{\Phi(S_T)\}\,\alpha_t dt \qquad a.s. \tag{26}$$

If we can find one, we then have

$$E\left[\frac{\partial\Phi}{\partial S}(S_T)\,\frac{\partial S_T}{\partial\lambda}\right] = E[\Phi(S_T)\,\pi] \tag{27}$$

where $\pi = \delta(\alpha)$. In conclusion, we obtain a representation of $\frac{\partial u_\lambda}{\partial\lambda}$ in terms of quantities that are straightforward to compute by Monte-Carlo simulations provided we determine α which solves (26). In order to do so, we first observe that we have

$$\int_0^T D_t\{\Phi(S_T)\}\,\alpha_t\,dt = \frac{\partial\Phi}{\partial S}(S_T)\int_0^T D_t S_T\,\alpha_t\,dt\,,$$

hence, we need to determine α solving

$$\int_0^T D_t S_T\,\alpha_t\,dt = \frac{\partial S_t}{\partial\lambda} \qquad a.s.$$

And if $D_t S_T \not\equiv 0$ a.s., this is certainly possible choosing for instance $\alpha_t = (D_t S_T)\frac{\partial S_t}{\partial\lambda}\left(\int_0^T (D_t S_T)^2\,dt\right)^{-1/2}$. Finally, one can check that $D_t S_T = \sigma(S_t)\,\xi_T\,\xi_t^{-1}$, where ξ_t satisfies: $d\xi_t = \sigma'(S_t)\,\xi_t\,dW_t$, $\xi_0 = 1$. And, if for example when $\lambda = S$ (the delta), σ is nondegenerate, we may check (see [9]) for more details) that we may choose α in such a way that we find

(28) $$\pi = \frac{1}{T}\int_0^T \sigma^{-1}(S_t).\ \xi_t\,dW_t.$$

Let us also observe that there exist many weights π (the question of the «optimal» one with minimal variance is studied in [9]) for which (27) holds and that we also have (by an easy density argument)

(29) $$\frac{\partial u}{\partial S} = E[\Phi(S_T)\,\pi]$$

for all $\Phi \in L^\infty$ (for instance).

Let us give another application to an asian option where we are interested in $\frac{\partial u}{\partial S}$ with $u = E[\Phi(S_T, I_T)]$ and $I_T = \int_0^T S_t\,dt$. Then, we have

$$\frac{\partial u}{\partial S} = E\left[\frac{\partial\Phi}{\partial S}\,\xi_T + \frac{\partial\Phi}{\partial T}\left(\int_0^T \xi_t\,dt\right)\right]$$

$$= E\left[\int_0^T D_t\{\Phi(S_T, I_T)\}\,\alpha_t\,dt\right]$$

if α_t satisfies

$$\int_0^T \alpha_t \sigma(S_t)\, \xi_T \xi_t^{-1}\, dt = \xi_T$$

and

$$\int_0^T dt \int_t^T \alpha_t \sigma(S_t)\, \xi_s \xi_t^{-1}\, ds = \int_0^T \xi_t\, dt\,,$$

and since the linear forms $\sigma(S_t)\, \xi_t^{-1}$ and $\sigma(S_t)\, \xi_t^{-1} \int_t^T \xi_s\, ds$ are obviously linear independent, the existence of such a factor α and thus of a weight π is insured. Examples are given in [9].

III.4. *Localization.*

Numerical illustrations may be found in [9], [8]. We only wish to mention here that the practical implementation of the above approach requires, in order to produce a very efficient numerical method (much more than the classical difference quotients approach), to localize the above integration by parts around the singulaties of Φ. This is explained in detail in [9], [8] but the idea is simple. For instance, in the Brownian case $S_t = S + \sigma W_t$, (28)-(29) yield

$$\frac{\partial u}{\partial S} = E\left[\Phi(S_T)\, \frac{W_T}{\sigma_T}\right].$$

And multiplying Φ by W_T may lead to a Monte-Carlo simulation of a random variable with a rather large variance which is thus quite slow. This is why one has to localize the integration by parts: for instance, if Φ is smooth except at K, we introduce Ψ smooth such that $\Psi \equiv \Phi$ outside $[K_1, K_2]$ (with $K_1 < K < K_2$) and we write

$$\frac{\partial u}{\partial \lambda} = E\left[\Psi'\, \frac{\partial S_T}{\partial \lambda} + (\Phi - \Psi)\, \pi\right].$$

The quantity $(\Phi - \Psi)\, \pi$ is now better behaved since, in the Brownian case, $(\Phi - \Psi)\, W_T$ is now compactly supported ...

III.5. *Conditional expectations.*

Malliavin calculus can also be used for the Monte-Carlo simulations of *conditional expectations*. One class of example (others may be found in [9]) is the following one:

(30) $$u = E[\Phi(S_T)\, |\, S_t = S']$$

where $0 < t < T$ and S' is arbitrary. Of course, in view of the Markov property, this is nothing else than $u(S', T-t)$. In particular, our approach allows to compute, with a *single* set of Monte-Carlo paths emanating from a fixed position S, the solution of a linear second order parabolic equation at any point S' (and positive time) which is a completely new fact in scientific computing. Of course, traditional Monte-Carlo simulations of (30) are hopeless since «almost all» paths generated «randomly» will miss the value S' at time t.

Formally, the idea of our approach is to write

$$u = \frac{E[\Phi(S_T)\, \delta_{S'}(S_t)]}{E[\delta_{S'}(S_t)]} \tag{31}$$

and then only needs to integrate by parts $E[\Phi(S_T)\, \delta_{S'}(S_t)]$, for each Φ, without differentiating $\Phi(S_T)$. This is indeed possible provided we introduce $H(S) = 1/2 \text{ sign } (S - S')$ and we write

$$\begin{aligned} E[\Phi(S_T)\, H'(S_t)] &= E\left[\int_0^T D_s\{\Phi(S_T)\, H(S_t)\}\, \alpha_s\, ds\right] \\ &= E[\Phi(S_T)\, H(S_t)\, \delta(\alpha)] \end{aligned}$$

provided α satisfies

$$\int_0^T (D_s S_T)\, \alpha_s\, ds = 0\,, \qquad \int_0^T (D_s S_t)\, \alpha_s\, ds = 1$$

i.e.

$$\int_0^T \sigma(S_s)\, \xi_s^{-1}\, \alpha_s\, ds = 0\,, \qquad \int_0^t \sigma(\xi_s)\, \xi_s^{-1}\, \alpha_s\, ds = \xi_t^{-1}.$$

We may then choose for instance, $\alpha_s = \frac{1}{\sigma(\xi_s)} \frac{\xi_s}{\xi_t}\left(\frac{1}{t} 1_{[0,\, t)}/s\right) - \frac{1}{T-t} 1_{[t,\, T)}(s))$. And, denoting by $\pi = \delta(\alpha)$, we obtain finally the desidered representation namely

$$u = \frac{E[\Phi(S_T)\, H(S_t)\, \pi]}{E[H(S_t)\, \pi]}\,. \tag{32}$$

And this representation can be computed with the same generation of paths than the one used initially to compute $E[\Phi(S_T)]$...

Let us finally mention that this approach is used in [14] to make a full Monte-Carlo computation of american options, thus opening the road to Monte-Carlo simulations of *all* financial products...

IV. – Partial observations and option pricing.

The work described below is taken from [12] and should be seen only as an example showing, we hope, the usefulness of partial observations models in Finance. Other examples may be found in [12] and will be developed elsewhere.

IV.1. *An industrial problem.*

We begin with a practical financial issue encountered by many industrial groups that need to protect themselves againts currency fluctuations during the negotiation (and execution) of a contract. Typically, we may need an option at time T (which could be thought as an insurance policy against currency fluctuations for instance ...) but the actual need of this option and its precise form depend an a random event which is (essentially) independent from the financial sphere (signing or not signing a contract for instance) and on which we have some informations (the probability of signing the contract for example) or we may obtain (= buy) some informations. The issues are numerous and certainly include the price (and the hedge) of such a product and the understanding of the relationships between the price of this option and the price of informations.

IV.2. *Without observations.*

We consider an asset governed by (1) and a pay off of the form

$$\Phi(S_T, Y)$$

where Y is independent from $(W_t)_{t \geqslant 0}$ and its law is given (or estimated) by a probability measure m. Examples include: *i*) (contract) $Y = 1$ or 0 with probability p, $1 - p$ and $\Phi(S, 1) = \Phi(S)$ (we need the option if $Y = 1$), $\Phi(S, 0) = 0$ (we do not need it if $Y = 0$); *ii*) (sales) $Y = y \in R^k$ with a law $dm(y)$ and $\Phi(S, y) = \Phi(S) \cdot y \left(= \sum_{i=1}^{k} \Phi_i(S)\, y_i \right)$.

We then use the utility function approach described in section 2 and introduce the wealth process P_t satisfying (2). Then, α_t is obviously adapted to the filtration generated by W_t (it is «independent» from Y). And we consider

$$V(P, S, T) = \max_{\alpha_t} \overline{E}[U(P_T - \Phi(S_T)] \tag{33}$$

where $\overline{E}$ is the «total» expectation with respect to (W, Y) and $U(p)$ is a utility function that we take, in order to simplify the presentation, to be

$$U(p) = 1 - e^{-\lambda p} \tag{34}$$

and $\lambda > 0$ is a positive parameter that corresponds to the absolute risk aversion. In other words, we have

$$V(P, S, T) = \max_{\alpha_t} E \int U(P_T - \Phi(S_T, y))\, dm(y) \tag{35}$$

and, as in section 3, we introduce V_0 (given by (5)) and the «fair» price u of the product such that (6) holds.

It is shown in [12] (it is infact a very special case of the results shown therein) that u is given by

$$u(S, T) = E[\Psi(S_T^0)], \qquad \Psi(S) = \frac{1}{\lambda} \log \int e^{\lambda \Phi(S, y)}\, dm(y), \tag{36}$$

and that the «optimal» hedge is still given by (8). Formally, this can be easily understood since V and V_0 still solve the same HJB equation (9) and we have now

$$\begin{aligned} V|_{T=0} &= \int U(P - \Phi(S, y))\, dm(y) \\ &= 1 - e^{-\lambda P}\Big[\int e^{\lambda \Phi(S, y)}\, dm(y)\Big] \\ &= 1 - e^{-\lambda(P - \Psi(S))}. \end{aligned}$$

The meaning of the above result is that we just need to consider an equivalent option whose pay off is Ψ. In addition, for low risk aversion i.e. small λ,

$$\Psi(S) = \int \Phi(S, y)\, dm(y) + \frac{\lambda}{2}\Big[\int \Phi^2(S, y)\, dm(y) - \Big(\int \Phi(S, y)\, dm(y)\Big)^2\Big] + o(\lambda),$$

i.e. the first order approximation is simply to consider the averaged payoff $\int \Phi(S, y)\, dm(y)$ (i.e. Φp for a contract) which is a very natural quantity from a financial view-point, but the next order (that builds up the price of the product) involves the variance in y of the pay-off which, of course, measures the risk induced by fluctuations of Y. Notice finally that the price (u) is increasing with respect to λ i.e. grows if risk aversion grows!

IV.3. *Partial observations and the cost of informations.*

We now consider a pay-off given by $\Phi(S_T, Y_T)$ where Y_t satisfies

$$dY_t = \textstyle\sum(Y_t) \cdot dB_t + \gamma(y_t)\, dt, \qquad Y_0 = y \in R^k, \tag{37}$$

B_t is a n-dimensional Brownian motion independent of W and $\sum$, γ are smooth (with bounds that we do not specify here, see [12]). And we wish to incorporate in our model the following possible actions (controls): on each time interval $(t, t + dt)$, either we choose $\beta_t = 1$, we observe dY_t and we pay cdt, or we choose

β_t and we do not observe dY_t. Then, (2) is replaced by

$$dP_t = \alpha_t \cdot dS_t - c\beta_t dt \tag{38}$$

and α_t (the hedging strategy) is now adapted to the filtration generated by the observations (which are, in some sense, parts of the controls, a feature that makes this type of control problems rather new ...). The cost c (for observation) is t aken to be a positive constant in order to simplify notation.

In order to write precisely the control problem, we introduce, as is customary in optimal stochastic control under partial observations and nonlinear filtering, the conditional low of Y_t given the observations that we denote by ν_t. This conditional low is a random process that depends on y, and which is a.s. a probability measure in y. The above heuristics are then translated in the following stochastic parabolic equation

$$d\nu_t = B\nu_t \cdot \beta_t dB_t + A\nu_t dt \tag{39}$$

where $A\nu = \frac{1}{2}\frac{\partial^2}{\partial y_i \partial y_j}(K_{ij}\nu)$, $K = \sum \cdot \sum^T$, $\beta\nu = -div_y(\sum \nu)$. This equation is nothing else than a variant of the famous Zakai's equation in nonlinear filtering theory. Then, we may write an infinite-dimensional stochastic control with full observations

$$V(P, S, \nu, T) = Max_{\alpha, \beta} E \int d\nu_t(y)\, U(P_T - \Phi(S_T, y)) \tag{40}$$

where ν is the initial probability law on $y = Y_0$, so that $\nu_t \,|_{t=0} = \nu$.

Working with the associated infinite-dimensional Hamilton-Jacobi-Bellman equations, thanks to viscosity solutions theory as developed in P.-L. Lions [13], and making manipulations somewhat related to those introduced in section 2.1 (which can also be justified with the help of infinite-dimensional viscosity solutions theory), one can show the following results that we state somewhat imprecisely (in order to avoid rather unpleasant technicalities...).

THEOREM. – *We have*

$$V(P, S, \nu, T) = V_0(P - u(S, \nu, T), S, T) \tag{41}$$

where u (*the price of the option*) *is the solution of the following infinite-dimensional nonlinear Black-Scholes equation*

$$\left\{ \frac{\partial u}{\partial T} - \frac{1}{2} Tr\left(a \cdot \frac{\partial^2 u}{\partial S^2}\right) - \left\langle \frac{\partial u}{\partial \nu}, A\nu \right\rangle + \left(-\frac{1}{2}\frac{\partial^2 u}{\partial \nu^2}(B\nu, B\nu) + \frac{\lambda}{2}\left\langle \frac{\partial u}{\partial \nu}, B\nu \right\rangle^2 - c \right)_+ \right\} = 0 \tag{42}$$

(*for all* S, *bounded non negative measures* ν, $T \geqslant 0$) *and*

$$u|_{T=0} = \frac{1}{\lambda} \log \int e^{\lambda \Phi(S, y)}\, d\nu(y)\,. \tag{43}$$

In addition, the hedging strategy is still given by (8).

Of course, this infinite-dimensional nonlinear equation is by no means easy to solve and the following two corollaries provide some intuition about the structure of u.

COROLLARY 1. – *As* λ *goes to* 0_+, $u = u^0 + \lambda u^1 + 0(\lambda)$ *where* u^0, u^1 *solve*

$$\frac{\partial w}{\partial T} - \frac{1}{2} Tr\left(a \cdot \frac{\partial^2 w}{\partial S^2}\right) - \left\langle \frac{\partial w}{\partial \nu}, A\nu \right\rangle = 0 \tag{44}$$

with $u^0|_{T=0} = \int \Phi(S, y)\, d\nu(y), u^1|_{T=0} = \int \Phi^2(S, y)\, d\nu(y) - \left(\int \Phi(S, y)\, d\nu(y)\right)^2$, *i.e.*

$$u^0(S, \nu, t) = E\left[\int \Phi(S_T^0, y)\, d\nu_T^0(y)\right], u^1(S, \nu, T) =$$

$$E\left[\int \Phi^2(S_T^0, y)\, d\nu_T(y) - \left(\int \Phi(S_T^0, y)\, d\nu_T^0(y)\right)^2\right],$$

with

$$\frac{\partial v_t^0}{\partial t} = Av_t^0; \qquad \nu_t^0|_{t=0} = \nu\,. \tag{45}$$

In addition, the optimal control β_t *vanishes identically i.e. we never buy information for* $\lambda > 0$ *small enough.*

Next, we consider the special case when Σ and γ are constant (i.e. independent of y) and when $\Phi(S, y) = \Phi(S) \cdot y$; and we introduce

$$\Phi_i^1(S, T) = E[\Phi_i^1(S_T^0)] \tag{46}$$

$$\Phi_{ij}^2(S, T) = E[\Phi_i^1 \Phi_j^1(S_T^0)]\,. \tag{47}$$

COROLLARY 2. – *For any* ν *such that* $\int e^{c|y|}\, d\nu(y) < \infty$ *for all* $c > 0$,

$$
(48) \qquad \begin{cases} \dfrac{\partial u}{\partial T} - \dfrac{1}{2} Tr\left(a \dfrac{\partial^2 u}{\partial S^2}\right) = \dfrac{\lambda}{2} K_{ij} \cdot \Phi^2_{ij} + \gamma_i \Phi^1_i - \left(\dfrac{\lambda}{2} K_{ij} (\Phi^2_{ij} - \Phi^1_i \Phi^1_j) - c\right)_+ \\ u|_{T=0} = \dfrac{1}{\lambda} \log \displaystyle\int e^{\lambda \Phi(S) \cdot y}\, d\nu(y)\,. \end{cases}
$$

And, an optimal feed back control β_t *is given by* $\beta(S, t) = 1$ *if and only if* $\frac{\lambda}{2} K_{ij} (\Phi^2_{ij}(S, t) - (\Phi^1_i \Phi^1_j)(S, t)) > c$.

The equation (48) is a simple modification of the «Black-Scholes» equation and shows various interesting phenomena. First of all, the decision of buying information is independent of the initial guess ν on y. Next, the scaling for the cost c of information is (risk aversion) $\times$ (variance of Y)^2 $\times$ (variance of option prices).

REFERENCES

[1] M. AVALLANEDA - A. PARAY - A. LEVY, *Pricing and hedging derivative securities in marckets with uncertain volatilities*, Appl. Math. Finance, **2** (1995), 73-88.

[2] G. BARLES - E. DAHER - M. ROMANO, preprint.

[3] G. BARLES - M. H. SONER, *Option pricing with transaction costs and a nonlinear Black-Scholes equation*, Finance and Stochastics, to appear.

[4] F. BLACK - M. SCHOLES, *The pricing of options and corporate liabilities*, J. Political Economy, **81** (1973), 637-659.

[5] R. CAFFLISCH - W. MARCKOFF - A. OWAR, *Valuation of mortgage-backed securities using Brownian bridges to reduce effective dimension*, J. Comp. Finance, **1** (1997), 27-46.

[6] D. DUFFIE, *Security Markets: Stochastics Models*, Academic Press, Boston, 1988.

[7] D. DUFFIE, *Dynamic Asset Pricing Theory*, 2nd edit., Princeton Univ. Press, Princeton, 1996.

[8] E. FOURNIÉ - J. LEBUCHAUX - J. M. LASRY - P. L. LIONS - N. TAURI, *Applications of Malliavin calculus to Monte-Carlo methods in Finance*, Finance and Stochastics, **3** (1999).

[9] E. FOURNIÉ - J. LEBUCHAUX - J.-M. LASRY - P.-L. LIONS, *Applications of Malliavin calculus to Monte-Carlo methods in Finance, II*, Finance and Stochastics, to appear.

[10] D. LAMBERTON - B. LAPEYIRE, *Introduction au calcul stochastique appliquè a la Finance*, 1991.

[11] J. M. LASRY - P. L. LIONS, work in preparation.

[12] J.-M. LASRY - P.-L. LIONS, *Control stochastique avec informations partielles et applications a la Finance*, C.R. Acad. Sci. Paris, **328** (1999), 1003-1010.

[13] P.-L. LIONS, *Viscosity solutions of fully nonlinear second order equations and optimal stochastic control in infinite dimension*, Part II: Optimal control of Zakai's equation. In *Stochastic Partial Differential Equations and Applications*, II. Springer L.N. Maths, 1340, Berlin, 1989.
[14] P.-L. LIONS - H. RÉGNIER, work in preparation.
[15] P. MALLIAVIN, *Stochastic Analysis*, Springer, Berlin, 1997.
[16] R. MERTON, *Theory of rational option pricing*, Bell J. Econom. Manag. Sci, 4 (1973), 141-183.
[17] M. MUSIELA - M. RUTKOWSKI, *Mastingale methods in Financial Modelling*, Springer, Berlin, 1997.
[18] D. NUALART, *The Malliavin calculus and related topics*, Springer, Berlin, 1995.
[19] R. ROBONTOS, *Interest rate option models*, Wiley, New-York, 1997.

P. L. Lions: Ceremade UMR CNRS 7534, Université Paris-Dauphine
Place de Lattre de Tassigny, Paris 75775 Cedex 16, France

Pervenuta in Redazione
il 13 settembre 1999