PRICING EUROPEAN OPTIONS WITHOUT PROBABILITY*

V.G. Vovk

June 1995

It is well known that in the case where the stock price S_t is governed by the equation $dS_t/S_t = \mu dt + \sigma dW_t$, any European option satisfying weak regularity conditions has a fair price (the Black—Scholes formula and its generalizations). We consider the case where no probabilistic assumptions are made about S_t; instead, we assume that the derivative security D which pays a dividend of $(dS_t/S_t)^2$ (the squared relative increase in the price of S_t) each instant dt is traded in the market. We prove that the "regular" European options have fair prices provided that both S_t and D_t (the price process of D) are continuous and the fractal dimensions of the graphs of S_t and D_t satisfy certain inequalities. Intuitively our assumptions are much weaker than the usual assumption $dS_t/S_t = \mu dt + \sigma dW_t$.

Key Words: Black—Scholes formula, fractal dimension, pathwise stochastic integral, nonstandard analysis

*The research described in this publication was made possible in part by Grant No. MRS300 from the International Science Foundation and Russian Government. It was finished while the author was a Fellow at the Center for Advanced Study in the Behavioral Sciences (Stanford, CA), with financial support provided by National Science Foundation (#SES-9022192). I am grateful to Glenn Shafer for supplying information about Bick and Willinger (1994) and to Glenn, Phil Dawid, Steffen Lauritzen, and Shashi Murthy for useful comments.

1 MAIN RESULT

We will use some simplest notions of nonstandard analysis. Exact definitions are given in Appendix A, but for understanding our results (both statements and proofs) an intuitive idea of "infinitely large" and "infinitely small" (or "infinitesimal") will suffice. All numbers considered are ordinary integer or real (i.e., "standard") numbers unless explicitly stated otherwise.

Our framework is intermediate between continuous-time and discrete-time: our time interval is $[0, T]$ but it is divided into Ω equal infinitely short subintervals. Here Ω is an infinitely large positive integer fixed throughout the paper; the time horizon $T > 0$ is also fixed throughout the paper. We put

$$dt := \frac{T}{\Omega},\ \mathcal{T} := \{idt \mid 0 \le i \le \Omega\},$$

where i ranges over the nonnegative integers, maybe infinitely large.

We consider a security market where two primary securities, bond B and stock S, are traded. Both securities pay no dividends; the price for a unit of S (resp. B) at time $t \in [0, T]$ is denoted S_t (resp. B_t). The market is frictionless and there are no limitations on short selling. We will assume $B_t \equiv 1$ (taking B as a new numeraire); therefore, the bond being available for trade amounts to the possibility of interest-free borrowing and lending.

Let $U\colon \mathbb{R} \to \mathbb{R}$ be some "regular" nonnegative "payoff function". We are interested in the fair price of the European option with maturity date T and payoff $U(S_T)$. First we briefly recall the Black—Scholes approach to this problem (for details see, e.g., Shiryaev (1994), Shiryaev et al. (1994a, 1994b), Harrison and Pliska (1981)).

The stock price S_t is assumed to be governed by the stochastic differential equation

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t, \tag{1.1}$$

where W_t is a standard Wiener process, $\mu \in \mathbb{R}$ is the "appreciation rate", and $\sigma > 0$ is the "volatility coefficient"; the initial value $S_0 > 0$ is known. Under this assumption, the fair price for the European option $U(S_T)$ exists and equals

$$\int_{\mathbb{R}} U\left(S_0 e^w\right) \mathcal{N}_{-T\sigma^2/2,\, T\sigma^2}(dw), \tag{1.2}$$

where $\mathcal{N}_{a,b}$ stands for the normal distribution with mean a and variance b (see Shiryaev et al. (1994b), Theorem 4.1, or Harrison and Pliska (1981), (1.15) and Section 5).

Assumption (1.1) is very strong: it suffices to recall that

$$\frac{1}{\sigma}\left(\ln\frac{S_t}{S_0}-\left(\mu-\frac{\sigma^2}{2}\right)t\right)$$

is a standard Wiener process (Shiryaev et al. (1994b), (1.4)), and so its paths satisfy, with probability 1, many very special properties such as the local law of the iterated logarithm. Besides, it is widely believed that the empirical data contradict assumption (1.1) (Mandelbrot (1963, 1982), Fama (1965), Rachev and Rüschendorf (1994)). The usual remedy is to replace (1.1) by other probabilistic assumptions, e.g., to replace the normal distribution by a distribution with "heavier tails". Our approach is, in some sense, more radical: we will not impose any probabilistic assumptions on S_t; instead, we will make an assumption of more "financial" nature. (Though the fair price for $U(S_T)$ will still be given by (1.2), unlike many other alternatives to the Black—Scholes approach; cf. discussion in Section 3 below.)

Formula (1.2) explicitly contains $T\sigma^2$, the total volatility of $\ln S_t$ over the whole time interval $[0,T]$. What can we replace this value with? Our choice is to measure the volatility of S by some derivative security D. It will turn out that adding only one derivative security D will enable us to price infinitely many European options $U(S_T)$. (It may seem that the assumption that some extra security is traded in the market is very restrictive; however, we will later see that our assumption is intuitively weaker than (1.1).)

We assume that the following derivative security D is traded in our security market: at each time $t\in\mathcal{T}\setminus\{0\}$, D pays a dividend of $\left(\frac{S_t-S_{t-dt}}{S_{t-dt}}\right)^2$ (that is, the squared relative increase in the price of S during $[t-dt,t]$). Let D_t be the price for a unit of D at time t. We will replace $T\sigma^2$ by D_0 in (1.2). To do so, we will have to impose some restrictions on the paths $t\mapsto S_t$ and $t\mapsto D_t$.

First of all, we assume that the paths S_t and D_t are continuous. Our next assumption will concern the "oscillation rate" of S_t and D_t. Let us give formal definitions.

Consider a continuous function $f\colon[0,T]\to\mathbb{R}$. For each constant $c\in$

$[1, \infty[$, we define the *c-variation* $\operatorname{var}_c f \geq 0$ to be the sum

$$\sum_{t \in \mathcal{T} \setminus \{T\}} |f(t + dt) - f(t)|^c. \tag{1.3}$$

(For $c = 2$, this may be called the *quadratic variation*—cf. Theorem I.4.47 in Jacod and Shiryaev (1987).)

Lemma 1.1 *For each continuous function* $f : [0, T] \to \mathbb{R}$ *there exists a unique number* $\dim f \in [1, \infty]$ *(the* oscillating dimension *of* f*) such that*

- $\operatorname{var}_c f$ *is infinitely large when* $1 \leq c < \dim f$*;*
- $\operatorname{var}_c f$ *is infinitely small when* $c > \dim f$*.*

To understand the intuition behind this notion, suppose that the increments $f(t + dt) - f(t)$ of f have the order of magnitude $(dt)^H$. (For the "usual" deterministic functions like $\sin t$ we have $H = 1$; for the typical paths of "regular" diffusion processes with positive diffusion coefficient, such as (1.1), we have $H = \frac{1}{2}$.) In this case, $\operatorname{var}_c f$ has the order of magnitude $(dt)^{cH-1}$, and we obtain $\dim f = \frac{1}{H}$. Reversing this argument, we define the *Hölder exponent* of f to be $H(f) := \frac{1}{\dim f}$.

Remark 1.1 We can regard $\dim f$ as the fractal dimension of the graph of f (Mandelbrot (1986), Part II, 2.3), and there are even more convincing reasons (Mandelbrot, 1986) to regard $2 - H(f)$ as the fractal dimension of the graph of f.

Now we formally define the notion of fair price. A *trading strategy* is a pair (Δ, Σ) of functions $\Delta : \mathbb{R}^2 \to \mathbb{R}$ and $\Sigma : \mathbb{R}^2 \to \mathbb{R}$. At each time $t \in \mathcal{T}$, the pair $(\Delta(D_t, S_t), \Sigma(D_t, S_t))$ is interpreted as the portfolio recommended by the strategy: $\Delta(D_t, S_t)$ is the number of units of D in the portfolio, $\Sigma(D_t, S_t)$ is the number of units of S, and the amount of bond B is uniquely determined by the requirement that the strategy be self-financing.

Remark 1.2 This is a narrow definition: e.g., for more general trading strategies, the amounts of D and S recommended at time t would be allowed to depend not only on D_t and S_t but also on t and (D_s, S_s) for $s < t$. However, our simple definition will be sufficient for our purpose.

With the strategy (Δ, Σ) and the paths $D: t \mapsto D_t$, $S: t \mapsto S_t$ we associate the *wealth increment* $X(D, S, \Delta, \Sigma)$ defined by

$$X(D,S,\Delta,\Sigma) := \sum_{t \in \mathcal{T} \setminus \{T\}} \Bigg(\Delta(D_t, S_t) \left(D_{t+dt} - D_t + \left(\tfrac{S_{t+dt}-S_t}{S_t}\right)^2 \right) + \Sigma(D_t, S_t)\,(S_{t+dt} - S_t) \Bigg) \tag{1.4}$$

(this formula takes into account both the capital gains from holding D and S and the dividends paid by D). If our initial wealth is c and we follow (Δ, Σ), our final (i.e., time T) wealth will be $c + X(D, S, \Delta, \Sigma)$.

Let $\gamma(D, S)$ be some property of the paths D_t and S_t (such as $\frac{1}{\dim D} + \frac{1}{\dim S} > 1$). We say that $c \in \mathbb{R}$ is the *fair price* for a European option $U(S_T)$ (where $U: \mathbb{R} \to \mathbb{R}$ is the payoff function) *provided* $\gamma(D, S)$ if for each $\epsilon > 0$ there exists a trading strategy (Δ, Σ) such that, for all possible paths $D, S : [0, T] \to \mathbb{R}$,

$$\gamma(D, S) \Longrightarrow |c + X(D, S, \Delta, \Sigma) - U(S_T)| < \epsilon.$$

We call a function $U: \mathbb{R} \to \mathbb{R}$ *Lipschitzian* if, for some constant $c > 0$,

$$|U(x) - U(y)| \le c|x - y|, \ \forall x, y \in \mathbb{R}$$

(cf. condition (4.2) in Shiryaev et al. (1994b)).

Theorem 1.1 *Let $U: \mathbb{R} \to \mathbb{R}$ be Lipschitzian and $\delta, \sigma > 0$. The fair price for a European option $U(S_T)$ is*

$$\int_{\mathbb{R}} U(\sigma e^w)\, \mathcal{N}_{-\delta/2,\,\delta}(dw) \tag{1.5}$$

provided that $D_0 = \delta$, $S_0 = \sigma$, D and S are continuous,

$$\forall t \in [0, T]\colon\ S_t > 0,\ D_t \ge 0,\ D_t = 0 \Leftrightarrow t = T,$$

and

$$\dim D < 2,\ \dim S < 3,\ \frac{1}{\dim D} + \frac{1}{\dim S} > 1. \tag{1.6}$$

Intuitively, this theorem is much stronger than its conventional counterpart. Suppose the stock price is governed by (1.1). In this case (see Lemma B.2 of Appendix B) $\dim S = 2$ a.s., and conditions (1.6) boil down to $\dim D < 2$. (The case of dimension 2 is typical for the paths of stochastic processes, cf. Remark B.1 in Appendix B, so D is required to be non-stochastic in the least degree.) If the derivative security D is actually traded in an arbitrage-free market, then $D_t = \sigma^2(T-t)$ a.s. (see Appendix B) and, therefore, $\dim D = 1 < 2$ a.s. However, this security, under the diffusion model (1.1), is essentially useless, and so it does not matter whether it is available for trade. In Appendix B (Lemma B.3) we prove that, with probability 1, between any two times $t_1, t_2 \in [0,T]$, $t_1 < t_2$, D pays $\sigma^2(t_2 - t_1)$ (to within an infinitely small amount). Thus the investor who holds one unit of D in his portfolio from t_1 to t_2 achieves nothing by doing it: at time t_1 he pays $D_{t_1} = \sigma^2(T-t_1)$, during $[t_1, t_2]$ he receives $\sigma^2(t_2 - t_1)$ in dividends, and at t_2 he sells D for $D_{t_2} = \sigma^2(T-t_2)$; therefore, his total profit is

$$-\sigma^2(T-t_1) + \sigma^2(t_2-t_1) + \sigma^2(T-t_2) = 0.$$

On the other hand, the financial intermediaries will be willing to sell and buy such security D at time t for about $\sigma^2(T-t)$. So the assumption that D is available for trade is simultaneously superfluous and innocuous under the diffusion model.

There are two important special cases of (1.6). The first case was already mentioned: $\dim D < 2, \dim S \le 2$. The second case is $\dim D \le \frac{3}{2}, \dim S < 3$. The assumptions $\dim D < 2$ and $\dim D \le \frac{3}{2}$ can be regarded as different explications of the assumption that the past observations carry little information about the future rates of return $\frac{S_{t+dt} - S_t}{S_t}$.

Remark 1.3 Our result is very close in spirit to the Black—Scholes formula of Bick and Willinger (1994, Proposition 1). However, Bick and Willinger do not go beyond the quadratic variation; essentially, they consider the usual case where $\dim S = 2$.

2 PROOFS

Proof of Lemma 1.1

Let $c_1 < c_2$; it suffices to prove that the ratio

$$\sum_t |df(t)|^{c_2} \Big/ \sum_t |df(t)|^{c_1} \qquad (2.1)$$

(where t ranges over $\mathcal{T} \setminus \{T\}$ and $df(t) := f(t+dt) - f(t)$) is infinitely small. (We assume, without loss of generality, that f is not constant, so the denominator is positive.) Fix any $\epsilon > 0$; we are required to prove that (2.1) is less than ϵ. Let $\epsilon_1 > 0$ be so small that $\epsilon_1^{c_2 - c_1} < \epsilon$. Since f is continuous on the closed interval $[0, T]$, it is uniformly continuous on it and, therefore, $|df(t)| \le \epsilon_1$ for all t. So we have:

$$\sum_t |df(t)|^{c_2} = \sum_t |df(t)|^{c_2 - c_1} |df(t)|^{c_1} \le \epsilon_1^{c_2 - c_1} \sum_t |df(t)|^{c_1} < \epsilon \sum_t |df(t)|^{c_1}.$$

Proof of Theorem 1.1

First we will assume that U is a smooth function such that the derivatives $U^{(1)}$, $U^{(2)}$, $U^{(3)}$, and $U^{(4)}$ are bounded. In this case, we will find a trading strategy (Δ, Σ) such that

$$\left.\begin{array}{r} \dim D < 2 \\ \dim S < 3 \\ \frac{1}{\dim D} + \frac{1}{\dim S} > 1 \end{array}\right\} \Longrightarrow \begin{array}{l} \int_{\mathbb{R}} U(S_0 e^w)\, \mathcal{N}_{-D_0/2,\, D_0}(dw) \\ +X(D, S, \Delta, \Sigma) \approx U(S_T), \end{array} \tag{2.2}$$

where $a \approx b$ means that $|a - b|$ is infinitely small (i.e., a and b are *infinitely close*).

Put, for $D \ge 0$ and $S > 0$,

$$\overline{U}(D, S) := \int_{\mathbb{R}} U(S e^w)\, \mathcal{N}_{-D/2,\, D}(dw). \tag{2.3}$$

Note that the equality in (2.2) can be rewritten as

$$\overline{U}(D_0, S_0) + X(D, S, \Delta, \Sigma) \approx \overline{U}(D_T, S_T).$$

Directly differentiating (2.3), we can see that $\overline{U}$ satisfies (for $D > 0$) the well-known equation

$$\frac{\partial \overline{U}}{\partial D} = \frac{1}{2} S^2 \frac{\partial^2 \overline{U}}{\partial S^2}. \tag{2.4}$$

Our trading strategy is simply

$$\Delta := \frac{\partial \overline{U}}{\partial D}, \; \Sigma := \frac{\partial \overline{U}}{\partial S}.$$

Using Taylor's formula and the notation $df_t := f_{t+dt} - f_t$, we find for $t \in \mathcal{T} \setminus \{T\}$:

$$
\begin{aligned}
d\overline{U}(D_t, S_t) &= \tfrac{\partial \overline{U}}{\partial D}(D_t, S_t)dD_t + \tfrac{\partial \overline{U}}{\partial S}(D_t, S_t)dS_t \\
&+\tfrac{1}{2}\tfrac{\partial^2 \overline{U}}{\partial D^2}(D_t + \theta dD_t, S_t + \theta dS_t)(dD_t)^2 \\
&+\tfrac{1}{2}\tfrac{\partial^2 \overline{U}}{\partial S^2}(D_t + \theta dD_t, S_t + \theta dS_t)(dS_t)^2 \\
&+\tfrac{1}{2}\tfrac{\partial^2 \overline{U}}{\partial D \partial S}(D_t + \theta dD_t, S_t + \theta dS_t)dD_t dS_t,
\end{aligned}
\tag{2.5}
$$

where $\theta \in]0,1[$. Applying Taylor's formula to $\frac{\partial^2 \overline{U}}{\partial S^2}$, we transform the penultimate addend this way:

$$
\begin{aligned}
\tfrac{1}{2}\tfrac{\partial^2 \overline{U}}{\partial S^2}(D_t + \theta dD_t, S_t + \theta dS_t)(dS_t)^2 &= \tfrac{1}{2}\tfrac{\partial^2 \overline{U}}{\partial S^2}(D_t, S_t)(dS_t)^2 \\
&+\tfrac{1}{2}\tfrac{\partial^3 \overline{U}}{\partial D \partial S^2}(D_t + \theta_1 dD_t, S_t + \theta_1 dS_t)\theta dD_t (dS_t)^2 \\
&+\tfrac{1}{2}\tfrac{\partial^3 \overline{U}}{\partial S^3}(D_t + \theta_1 dD_t, S_t + \theta_1 dS_t)\theta (dS_t)^3,
\end{aligned}
$$

where $\theta_1 \in]0, \theta[$. Plugging this and (2.4) into (2.5), we find:

$$
\begin{aligned}
d\overline{U}(D_t, S_t) &= \tfrac{\partial \overline{U}}{\partial D}(D_t, S_t)\left(dD_t + \left(\tfrac{dS_t}{S_t}\right)^2\right) + \tfrac{\partial \overline{U}}{\partial S}(D_t, S_t)dS_t \\
&+\tfrac{1}{2}\tfrac{\partial^2 \overline{U}}{\partial D^2}(D_t + \theta dD_t, S_t + \theta dS_t)(dD_t)^2 \\
&+\tfrac{1}{2}\tfrac{\partial^3 \overline{U}}{\partial D \partial S^2}(D_t + \theta_1 dD_t, S_t + \theta_1 dS_t)\theta dD_t (dS_t)^2 \\
&+\tfrac{1}{2}\tfrac{\partial^3 \overline{U}}{\partial S^3}(D_t + \theta_1 dD_t, S_t + \theta_1 dS_t)\theta (dS_t)^3 \\
&+\tfrac{1}{2}\tfrac{\partial^2 \overline{U}}{\partial D \partial S}(D_t + \theta dD_t, S_t + \theta dS_t)dD_t dS_t.
\end{aligned}
$$

Summing over t, we further find:

$$
\begin{aligned}
&\left|\left(\overline{U}(D_T, S_T) - \overline{U}(D_0, S_0)\right) - X(D, S, \Delta, \Sigma)\right| \\
&\le \tfrac{1}{2}\sup\left|\tfrac{\partial^2 \overline{U}}{\partial D^2}\right| \operatorname{var}_2 D + \tfrac{1}{2}\sup\left|\tfrac{\partial^3 \overline{U}}{\partial D \partial S^2}\right| \theta \textstyle\sum_t |dD_t||dS_t|^2 \\
&+\tfrac{1}{2}\sup\left|\tfrac{\partial^3 \overline{U}}{\partial S^3}\right| \theta \operatorname{var}_3 S + \tfrac{1}{2}\sup\left|\tfrac{\partial^2 \overline{U}}{\partial D \partial S}\right| \textstyle\sum_t |dD_t||dS_t|,
\end{aligned}
\tag{2.6}
$$

all suprema being over the convex hull of $\{(D_t, S_t) \mid 0 < t < T\}$.

We are required to show that the right-hand side of (2.6) is infinitely small. First we will show that

$$
\sum_t |dD_t||dS_t|^2, \ \sum_t |dD_t||dS_t|
$$

are infinitely small. This is easy: Hölder's inequality and $\frac{1}{\dim D} + \frac{1}{\dim S} > 1$ give

$$\sum_t |dD_t||dS_t| \le \left(\sum_t |dD_t|^p\right)^{1/p} \left(\sum_t |dS_t|^q\right)^{1/q} \approx 0,$$

where $p, q \in \,]1, \infty[$ satisfy $\dim D < p$, $\dim S < q$, and $\frac{1}{p} + \frac{1}{q} = 1$; and the uniform continuity of S on $[0, T]$ gives $\sup_t |dS_t| \le 1$ and, thus,

$$\sum_t |dD_t||dS_t|^2 \le \sum_t |dD_t||dS_t| \approx 0.$$

It remains to prove that all suprema in (2.6) are finite. Using (2.4), we find:

$$\begin{aligned}
\frac{\partial^2 \overline{U}}{\partial D^2} &= \frac{1}{2}\frac{\partial}{\partial D}\left(S^2 \frac{\partial^2 \overline{U}}{\partial S^2}\right) = \frac{1}{2} S^2 \frac{\partial^3 \overline{U}}{\partial D \partial S^2}, \\
\frac{\partial^3 \overline{U}}{\partial D \partial S^2} = \frac{\partial^3 \overline{U}}{\partial S^2 \partial D} &= \frac{1}{2}\frac{\partial^2}{\partial S^2}\left(S^2 \frac{\partial^2 \overline{U}}{\partial S^2}\right) = \frac{\partial^2 \overline{U}}{\partial S^2} + 2S\frac{\partial^3 \overline{U}}{\partial S^3} + \frac{1}{2} S^2 \frac{\partial^4 \overline{U}}{\partial S^4}, \\
\frac{\partial^2 \overline{U}}{\partial D \partial S} &= \frac{1}{2}\frac{\partial}{\partial S}\left(S^2 \frac{\partial^2 \overline{U}}{\partial S^2}\right) = S\frac{\partial^2 \overline{U}}{\partial S^2} + \frac{1}{2} S^2 \frac{\partial^3 \overline{U}}{\partial S^3}.
\end{aligned}$$

To upper bound the partial derivatives $\frac{\partial^n \overline{U}}{\partial S^n}$, $n = 2, 3, 4$, we first note that $\mathbf{E}\, e^{x\xi} = e^{x^2/2}$ for $x \in \mathbb{R}$ and a standard normal variable ξ. Therefore ($\|U^{(n)}\|$ standing for $\sup |U^{(n)}|$),

$$\begin{aligned}
\frac{\partial^n \overline{U}}{\partial S^n} &= \textstyle\int_{\mathbb{R}} U^{(n)}(S e^w) e^{nw} \mathcal{N}_{-D/2,\, D}(dw) \\
&\le \|U^{(n)}\| \textstyle\int_{\mathbb{R}} e^{nw} \mathcal{N}_{-D/2,\, D}(dw) = \|U^{(n)}\|\, \mathbf{E}\, e^{n(-D/2+\sqrt{D}\xi)} \\
&= \|U^{(n)}\| e^{-nD/2} e^{n^2 D/2} = \|U^{(n)}\| e^{n(n-1)D/2}
\end{aligned}$$

(we were allowed to differentiate under the integral sign due to, e.g., Proposition II.3.7 in Bourbaki (1958); see also Remark 1 after that proposition). We can see that the right-hand side of (2.6) is infinitely small, which completes the proof for the case of smooth U with bounded $U^{(1)}$–$U^{(4)}$.

Now we drop the assumption that U is smooth with bounded $U^{(1)}$–$U^{(4)}$. Let $\sigma > 0$ be a small parameter. Put

$$V(S) := \int_{\mathbb{R}} U(S + w)\, \mathcal{N}_{0,\sigma^2}(dw)$$

(i.e., V is a smoothed version of U); we are going to apply (2.2) to V.

First we find (C_1, C_2,... denote various positive constants and c is the constant from the definition of U being Lipschitzian):

$$\begin{aligned}|V(S) - U(S)| &= \left|\textstyle\int_{\mathbb{R}} U(S+w) - U(S)\,\mathcal{N}_{0,\sigma^2}(dw)\right| \\ &\le \textstyle\int_{\mathbb{R}} |U(S+w) - U(S)|\,\mathcal{N}_{0,\sigma^2}(dw) \le c\textstyle\int_{\mathbb{R}} |w|\,\mathcal{N}_{0,\sigma^2}(dw) \\ &= c\sigma \textstyle\int_{\mathbb{R}} |w|\,\mathcal{N}_{0,1}(dw) \le C_1\sigma\end{aligned} \tag{2.7}$$

and ($\overline{V}$ is defined in the same way as $\overline{U}$)

$$\begin{aligned}|\overline{V}(D_0,S_0) - \overline{U}(D_0,S_0)| &= \left|\textstyle\int_{\mathbb{R}} V(S_0e^w) - U(S_0e^w)\,\mathcal{N}_{-D_0/2,\,D_0}(dw)\right| \\ &\le \textstyle\int_{\mathbb{R}} |V(S_0e^w) - U(S_0e^w)|\,\mathcal{N}_{-D_0/2,\,D_0}(dw) \le C_1\sigma\end{aligned} \tag{2.8}$$

(the last inequality follows from (2.7)). Assuming that V is a smooth function with bounded $V^{(1)}$–$V^{(4)}$, we deduce from (2.7) and (2.8) the existence of a trading strategy (Δ,Σ) that ensures

$$\begin{aligned}&|U(S_T) - \overline{U}(D_0,S_0) - X(D,S,\Delta,\Sigma)| \\ &\le |U(S_T) - V(S_T)| + |\overline{U}(D_0,S_0) - \overline{V}(D_0,S_0)| \\ &\quad + |V(S_T) - \overline{V}(D_0,S_0) - X(D,S,\Delta,\Sigma)| \\ &< 3C_1\sigma.\end{aligned}$$

Since σ can be taken arbitrarily small, the fair price for $U(S_T)$ is $\overline{U}(D_0,S_0)$.

It remains to prove that $V^{(1)}$–$V^{(4)}$ are bounded. It suffices to consider the case $\sigma = 1$. From

$$V(S) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-(x-S)^2/2} U(x)\,dx,$$

we find, for $n = 0, 1, \ldots$,

$$V^{(n)}(S) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-(x-S)^2/2} H_n(x-S) U(x)\,dx,$$

where H_n are Hermite's polynomials (see, e.g., Shiryaev (1989), Example II.11.1). Assuming, without loss of generality, $U(S) = 0$, we obtain for $n = 1, \ldots, 4$:

$$\begin{aligned}|V^{(n)}(S)| &\le C_2 \textstyle\int_{\mathbb{R}} e^{-(x-S)^2/2} |H_n(x-S)| c|x-S|\,dx \\ &= C_2 \textstyle\int_{\mathbb{R}} e^{-x^2/2} |H_n(x)| c|x|\,dx \le C_3\end{aligned}$$

(c is the constant from the definition of U being Lipschitzian). This completes the proof of the theorem.

3 DISCUSSION

In this paper we considered a security market with two primary securities, bond B and stock S. Since stock is a risky security, people often want to hedge against unfavorable changes in its price in the future, which creates demand for various derivative securities such as European options. The number of derivative securities which might turn out useful for the hedgers is enormous, and only few of them can be offered by the market (because of the low demand for most of these derivative securities). An ideal situation is where there are several "basic" derivative securities from which we can "construct" the other derivative securities that we may need.

We showed that a wide class of European options can be "constructed" from a single derivative security D. Our assumptions were that $\dim S < 3$, $\dim D < 2$, and $\frac{1}{\dim S} + \frac{1}{\dim D} > 1$. This calls for an empirical investigation of the oscillating dimensions for the primary and various derivative securities.

It is not obvious that our security D is the best choice. Another possible choice is security E paying a dividend of $(dS_t)^2$ (the squared absolute increase in the price of S_t) each instant dt. (This security is implicit in Bachelier's approach; see Shiryaev (1994), 3.1.) As noted by Samuelson (of course, in different terms; see Shiryaev (1994), 3.1), this choice is much worse than D. This can be seen from the fact that the valuation formula (1.5) will transform into $\int_{\mathbb{R}} U(S_0 + w)\mathcal{N}_{0,E_0}(dw)$ and so the fair price for $U(S_T)$ will depend on the impossible values $U(x)$, $x < 0$. Perhaps more reasonable choice would be a security intermediate between D and E (cf. the end of Section 4 of Shiryaev (1994)).

This paper is part of a general program of providing more parsimonious foundations for probability theory and its applications. In the most complete form this program is described in Shafer (1996); see also Dawid (1984, 1985) and Vovk (1993a, 1993b).

APPENDIX A: BASICS OF NONSTANDARD ANALYSIS

In this appendix we give a very simple construction of nonstandard numbers (for more details see, e.g., Davis (1977)).

A *nontrivial ultrafilter* in the set $\mathbb{N}$ of positive integers is a family $\mathcal{U} \subseteq 2^{\mathbb{N}}$ of numeric sets such that:

$$\left.\begin{array}{r} A \in \mathcal{U} \\ A \subseteq B \subseteq \mathbb{N} \end{array}\right\} \Longrightarrow B \in \mathcal{U},$$
$$A, B \in \mathcal{U} \Longrightarrow A \cap B \in \mathcal{U},$$
$$\forall A \subseteq \mathbb{N}\colon A \in \mathcal{U} \text{ or } \mathbb{N} \setminus A \in \mathcal{U},$$
$$\forall n \in \mathbb{N}\colon \{n\} \notin \mathcal{U}.$$

Let us fix such $\mathcal{U}$ (its existence follows from the axiom of choice). A set $A \subseteq \mathbb{N}$ is *big* if $A \in \mathcal{U}$.

A *nonstandard* (real) *number* a is a sequence $(a_1 a_2 \ldots)$ of real numbers. The operations such as $+$, $-$, $\times$, $/$, $|\cdot|$ over nonstandard numbers are defined term-wise: e.g.,

$$(a_1 a_2 \ldots) + (b_1 b_2 \ldots) := ((a_1 + b_1)(a_2 + b_2) \ldots);$$

the relations such as $=$, $<$, $\leq$, $>$, $\geq$ are extended to the nonstandard numbers through "voting": e.g., $(a_1 a_2 \ldots) \leq (b_1 b_2 \ldots)$ means that the set $\{n \,|\, a_n \leq b_n\}$ is big. As usual, we do not distinguish nonstandard numbers a and b such that $a = b$.

For each $A \subseteq \mathbb{R}$ we denote by ${}^*A$ the set of all nonstandard numbers $(a_1 a_2 \ldots)$ with $a_n \in A$, $\forall n$. The elements of ${}^*\mathbb{N}$ are called *nonstandard positive integers*. We embed $\mathbb{R}$ into ${}^*\mathbb{R}$ identifying each $t \in \mathbb{R}$ with $(tt\ldots) \in {}^*\mathbb{R}$.

It is easy to see that for all $a, b \in {}^*\mathbb{R}$ one and only one of the following three possibilities holds: $a < b$, $a = b$, or $a > b$. We say that $a \in {}^*\mathbb{R}$, $a \geq 0$, is *infinitely small* if $a < \epsilon$ for each real $\epsilon > 0$ and *infinitely large* if $a > N$ for each positive integer N.

In Section 1, we introduced an infinitely large positive integer Ω. Such numbers exist in ${}^*\mathbb{N}$: e.g., take $\Omega := (1, 2, \ldots)$. The ratio $dt = \frac{T}{\Omega}$ will be infinitely small. Sum (1.3) is interpreted term-wise: if $\Omega = (\Omega_n)_{n=1}^{\infty}$, then (1.3) is the nonstandard number

$$\left(\sum_{i=0}^{\Omega_n - 1} \left| f\left((i+1)T/\Omega_n\right) - f\left(iT/\Omega_n\right) \right|^c \right)_{n=1}^{\infty}.$$

The sum in (1.4) is also interpreted term-wise; namely, as the nonstandard number

$$\left(\sum_{i=0}^{\Omega_n-1}\left(\Delta(D_{iT/\Omega_n}, S_{iT/\Omega_n})\left(D_{(i+1)T/\Omega_n} - D_{iT/\Omega_n} + \left(\frac{S_{(i+1)T/\Omega_n}-S_{iT/\Omega_n}}{S_{iT/\Omega_n}}\right)^2\right)\right.\right.$$
$$\left.\left.+\Sigma(D_{iT/\Omega_n}, S_{iT/\Omega_n})\left(S_{(i+1)T/\Omega_n} - S_{iT/\Omega_n}\right)\right)\right)_{n=1}^{\infty}.$$

The reader will have no difficulties in interpreting the proofs of Section 2 in this rigorous framework. In the rest of the paper we will consider assertions whose proofs require the exact definition of nonstandard numbers.

Remark A.1 An important open problem in the foundations of our approach to option pricing is to find out for which properties $\gamma(D, S)$ the fair price for $U(S_T)$ provided $\gamma(D, S)$ is unique; despite our use of the definite article, it may not be unique if γ is too restrictive. (Notice that the non-uniqueness of fair price implies that there exist $\epsilon > 0$ and a trading strategy (Δ, Σ) such that, for all possible paths D and S,

$$\gamma(D, S) \Longrightarrow X(D, S, \Delta, \Sigma) \geq \epsilon.)$$

The uniqueness is almost obvious in the context of Theorem 1.1 (though even in this case a rigorous proof would need to overcome some technical difficulties); however, we can also ask if the uniqueness will remain when, say, $\dim S < 3$ in (1.6) is replaced by $2 < \dim S < 3$ (if uniqueness is lost, this will mean that the "superstochastic" case $\dim S > 2$ is impossible: it allows the investors to pump money out of the market).

APPENDIX B: CONNECTIONS WITH THE DIFFUSION MODEL

The first two results of this appendix are explications, within our mathematical framework, of the "$\sqrt{\Delta t}$-effect" discovered by Bachelier (Shiryaev (1994), 3.1; Shiryaev et al. (1994b), 1.1). Intuitively they are obvious because of the similarity between the definition of 2-variation $\operatorname{var}_2 S$ and the usual definition of quadratic variation $[S, S]_T$; however, their proofs are not quite trivial.

Lemma B.1 *The path* $W\colon [0,T] \to \mathrm{I\!R}$ *of a standard Wiener process satisfies* $\dim W = 2$ *a.s. Moreover,* $\operatorname{var}_2 W \approx T$ *a.s.*

Proof. Let $\epsilon > 0$ be arbitrarily small. For each $N = 1, 2, \ldots$, we have

$$\text{(B.1)} \qquad \operatorname{prob}\left\{\sum_{i=0}^{N-1}(W_{(i+1)T/N} - W_{iT/N})^2 \geq T + \epsilon\right\} \leq e^{-Nc(\epsilon)},$$

$$\text{(B.2)} \qquad \operatorname{prob}\left\{\sum_{i=0}^{N-1}(W_{(i+1)T/N} - W_{iT/N})^2 \leq T - \epsilon\right\} \leq e^{-Nc(\epsilon)},$$

where $c(\epsilon)$ is a positive constant. These inequalities follow from

$$\mathbf{E}(W_{(i+1)T/N} - W_{iT/N})^2 = T/N$$

and the standard large-deviation results (see, e.g., Shiryaev (1989), Section IV.5, (11) and (12)). Combining (B.1) and (B.2) with the Borel—Cantelli lemma (Shiryaev (1989), II.10.3), we obtain that, with probability 1,

$$\left|\sum_{i=0}^{N-1}(W_{(i+1)T/N} - W_{iT/N})^2 - T\right| \geq \epsilon$$

only for finitely many N; therefore, $|\operatorname{var}_2 W - T| < \epsilon$ a.s. Since ϵ can be taken arbitrarily small, $\operatorname{var}_2 W \approx T$ a.s. □

Lemma B.2 *The path* $S\colon [0,T] \to \mathrm{I\!R}$ *of the diffusion process governed by (1.1) satisfies* $\dim S = 2$ *a.s.*

Proof. Since

$$\text{(B.3)} \qquad S_t = S_0 e^{(\mu-\sigma^2/2)t+\sigma W_t}$$

for a standard Wiener process W (see Shiryaev et al. (1994b), (1.4)), we have:

$$\begin{aligned} dS_t &= S_t e^{(\mu-\sigma^2/2)dt+\sigma dW_t} - S_t \\ &= S_t e^{\theta_t((\mu-\sigma^2/2)dt+\sigma dW_t)}((\mu-\sigma^2/2)dt + \sigma dW_t) \asymp (\mu-\sigma^2/2)dt + \sigma dW_t, \end{aligned}$$

where $\theta_t \in\,]0,1[$ and $a_t \asymp b_t$ means that $|a_t| \leq C|b_t|$ and $|b_t| \leq C|a_t|$ for some positive constant C (maybe, dependent on the path S_t). Therefore,

$$\begin{aligned} \operatorname{var}_2 S &\asymp \textstyle\sum_t((\mu-\sigma^2/2)dt + \sigma dW_t)^2 \\ &= (\mu-\sigma^2/2)^2 \textstyle\sum_t (dt)^2 + 2(\mu-\sigma^2/2)\sigma \sum_t dt dW_t + \sigma^2 \sum_t (dW_t)^2 \\ &\approx \sigma^2 \operatorname{var}_2 W, \end{aligned}$$

t ranging over $\mathcal{T}\setminus\{T\}$, and Lemma B.1 shows that, with probability 1, $\mathrm{var}_2\, S$ is neither infinitely small nor infinitely large. □

Remark B.1 It is natural to expect that for a "regular" continuous semimartingale S we will have, for almost all paths $S : [0,T] \to \mathbb{R}$, $\dim S \in \{1,2\}$ ($\dim S = 1$ corresponding to a trivial "martingale component") and

$$\mathrm{var}_2\, S = [S,S]_T = \langle S,S\rangle_T$$

(cf. Jacod and Shiryaev (1987), Theorems I.4.47 and I.4.52). However, the proof may not be easy and may require that our basic infinitely large number Ω be sufficiently large, e.g., $\Omega := (1,2,4,\ldots)$ (dependence on the choice of Ω is analogous to dependence on the choice of the partition sequence of $[0,T]$ in Bick and Willinger (1994)).

The next lemma can be interpreted as asserting the uselessness of the derivative security D in the diffusion model (1.1).

Lemma B.3 *Let derivative security D pay a dividend of $\left(\frac{S_{t+dt}-S_t}{S_t}\right)^2$ at the end of each interval $[t,t+dt]$, $t \in \mathcal{T}\setminus\{T\}$, where S_t is governed by (1.1). With probability 1, the total amount of dividends paid by D during every time interval $[t_1,t_2]$, $0 \le t_1 < t_2 \le T$, is infinitely close to $\sigma^2(t_2-t_1)$.*

Proof. First we consider fixed t_1 and t_2. It is easy to see that we can assume that the dividend paid by D at the end of each $[t,t+dt]$ is $(\sigma dW_t)^2$ rather than $\left(\frac{S_{t+dt}-S_t}{S_t}\right)^2$, where W_t is the standard Wiener process satisfying (B.3). Analogously to (B.1) and (B.2), we have, for any $\epsilon > 0$ and from some N on,

$$\mathrm{prob}\left\{\sum_{i=\lceil Nt_1/T\rceil-1}^{\lfloor Nt_2/T\rfloor-1}\left(\sigma W_{(i+1)T/N}-\sigma W_{iT/N}\right)^2 \ge \sigma^2(t_2-t_1)+\epsilon\right\} \le e^{-Nc(\epsilon)},$$

$$\mathrm{prob}\left\{\sum_{i=\lceil Nt_1/T\rceil-1}^{\lfloor Nt_2/T\rfloor-1}\left(\sigma W_{(i+1)T/N}-\sigma W_{iT/N}\right)^2 \le \sigma^2(t_2-t_1)-\epsilon\right\} \le e^{-Nc(\epsilon)}.$$

Again applying the Borel—Cantelli lemma, we obtain that the total amount of dividends paid during $[t_1,t_2]$ is $\sigma^2(t_2-t_1)$ to within ϵ, a.s.; since ϵ can be arbitrarily small, it is infinitely close to $\sigma^2(t_2-t_1)$, a.s.

Now it is clear that for almost all paths S_t the total amount of dividends paid by D during $[t_1, t_2]$ is infinitely close to $\sigma^2(t_2 - t_1)$ for all rational $t_1, t_2 \in [0, T]$, $t_1 < t_2$. Consider a path S_t satisfying this property; let $t_1, t_2 \in [0, T]$, $t_1 < t_2$, be not necessarily rational. Since t_1 and t_2 can be arbitrarily accurately approximated from below and from above by rational numbers, the total amount of dividends paid by D during $[t_1, t_2]$ is infinitely close to $\sigma^2(t_2 - t_1)$. $\Box$

Lemma B.3 immediately implies that in the case where the derivative security D is traded in an arbitrage-free market its price D_t (as usual in probability theory, we assume that D_t is continuous from the right) satisfies $D_t = \sigma^2(T - t)$ a.s. and, therefore, $\dim D = 1$ a.s. To show it, it suffices to show that $D_t = \sigma^2(T - t)$ a.s. for any fixed $t \in [0, T]$. Fix $t \in [0, T]$ and suppose that, on the contrary, $D_t \neq \sigma^2(T - t)$ with positive probability. Then we will make a profit with positive probability and lose nothing with probability 1 if, at time t, we

- buy a unit of D if $D_t < \sigma^2(T - t)$,
- sell a unit of D if $D_t > \sigma^2(T - t)$.

REFERENCES

Bick, A., and W. Willinger (1994): "Dynamic Spanning without Probabilities," *Stochastic Process. Appl.*, 50, 349–374.

Bourbaki, N. (1958): *Éléments de Mathématique, Livre IV, Fonctions d'une Variable Réelle (Théorie Élémentaire)*, 2nd ed. Paris: Hermann.

Davis, M. (1977): *Applied Nonstandard Analysis.* New York: Wiley.

Dawid, A.P. (1984): "Statistical Theory. The Prequential Approach" (with Discussion), *J. R. Statist. Soc.* A, 147, 278–292.

Dawid, A.P. (1985): "Calibration-Based Empirical Probability" (with Discussion), *Ann. Statist.*, 13, 1251–1273.

Fama, E.F. (1965): "The Behavior of Stock-Market Prices," *J. Business*, 38, 34–105.

Harrison, J.M., and S.R. Pliska (1981): "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stochastic Process. Appl.*, 11, 215–260.

Jacod, J., and A.N. Shiryaev (1987): *Limit Theorems for Stochastic Processes*. Berlin: Springer.

Mandelbrot, B.B. (1963): "New Methods in Statistical Economics," *J. Political Economy*, 71, 421–440.

Mandelbrot, B.B. (1982): *The Fractal Geometry of Nature*. New York: Freeman.

Mandelbrot, B.B. (1986): "Self-Affine Fractal Sets." *In* L. Pietronero and E. Tosatti (eds.), *Fractals in Physics*. Amsterdam: North-Holland, pp. 3–28.

Rachev, S.T., and L. Rüschendorf (1994): "Models for Option Prices," *Theory Probab. Appl.*, 39, 120–152.

Shafer, G. (1996): *The Art of Causal Conjecture*. MIT Press, to be published.

Shiryaev, A.N. (1989): *Veroyatnost'*, 2nd ed. Moscow: Nauka. English translation of the 1st ed.: *Probability*. Berlin: Springer (1984).

Shiryaev, A.N. (1994): "On Some Concepts and Stochastic Models in Financial Mathematics," *Theory Probab. Appl.*, 39, 1–13.

Shiryaev, A.N., Yu.M. Kabanov, D.O. Kramkov, and A.V. Mel'nikov (1994a): "Toward a Theory of Pricing Options of European and American Types. I. Discrete Time," *Theory Probab. Appl.*, 39, 14–60.

Shiryaev, A.N., Yu.M. Kabanov, D.O. Kramkov, and A.V. Mel'nikov (1994b): "Toward a Theory of Pricing Options of European and American Types. II. Continuous Time," *Theory Probab. Appl.*, 39, 61–102.

Vovk, V.G. (1993a): "A logic of Probability, with Application to the Foundations of Statistics" (with Discussion), *J. R. Statist. Soc.* B, 55, 317–351.

Vovk, V.G. (1993b): "Forecasting Point and Continuous Processes: Prequential Analysis," *Test*, 2, 189–217.