Turbulent hydrodynamics and the Stock market: the art of creating a Fokker-Planck equation from a system with high kurtosis

Diego Zocco
Department of Physics, University of California, San Diego, La Jolla, CA 92093, USA
(Dated: December 4, 2007)

The evolution of the probability distributions of different stock markets is described by means of a Fokker-Plank equation. An analogy between turbulent fluid and stock market price changes dynamics is also presented. Finally, a Langevin approach is used to explain the origin of the long tails of the distribution functions for short trading periods of time by using the concept of multiplicative noise.

The availability of high frequency data from financial markets has made possible to study market dynamics on time scales of less than a day. For foreign exchange rates Müller *et al.* [1] have shown that there is a net flow of information from long to short time scales: the behavior of long-term traders influences the behavior of short-time traders who watch the market continuously. Motivated by this feature, it has been explained in a previous work [2] an analogy between these market dynamics and hydrodynamic turbulence. More specifically, the relationship between the probability density function (PDF) of stocks prices Δx and the time delay Δt (fig. 1a) is much the same as the relationship between the probability density of the velocity differences Δv of two points in a turbulent flow and their spacial separation Δr (fig. 1b). A flow of energy from large to small scales is one of the main characteristics of fully developed homogeneous isotropic turbulence in three dimensions. It provides a mechanism for dissipating large amounts of energy in a viscous fluid: energy is pumped into the system at large scales of the order of meters (*e.g.* by a moving car), transferred to smaller scales through eddies of decreasing size, and dissipated at the smallest scale (of the order of millimeters in the previous example). This cascade of kinetic energy extending over several orders of magnitude generates a scaling behavior of the eddies and manifests itself in a scaling of the moments $\langle(\Delta v)^n\rangle$ as $(\Delta r)^{\zeta n}$. In comparison, the information acquired by observing the market after a time Δt can be also related with the scaling of Δx with Δt. Table 1 summarizes the most important concepts of the analogy between turbulence and stock prices. It has also been observed that the *kurtosis*, which is the ratio of the fourth moment and the squared variance, decreases with increasing time delay, that is, the tails of the probability densities loose weight, which is a common feature of the turbulent hydrodynamics and the stock market dynamics. Figure 2 shows the variation of the second, fourth and sixth moments for different time intervals Δt for empirical exchange market data. This is a clear evidence that it is more expected to have Gaussian PDFs for long time intervals Δt rather than for short ones, for which fluctuations of the prices are more important. Several models have been proposed to explain

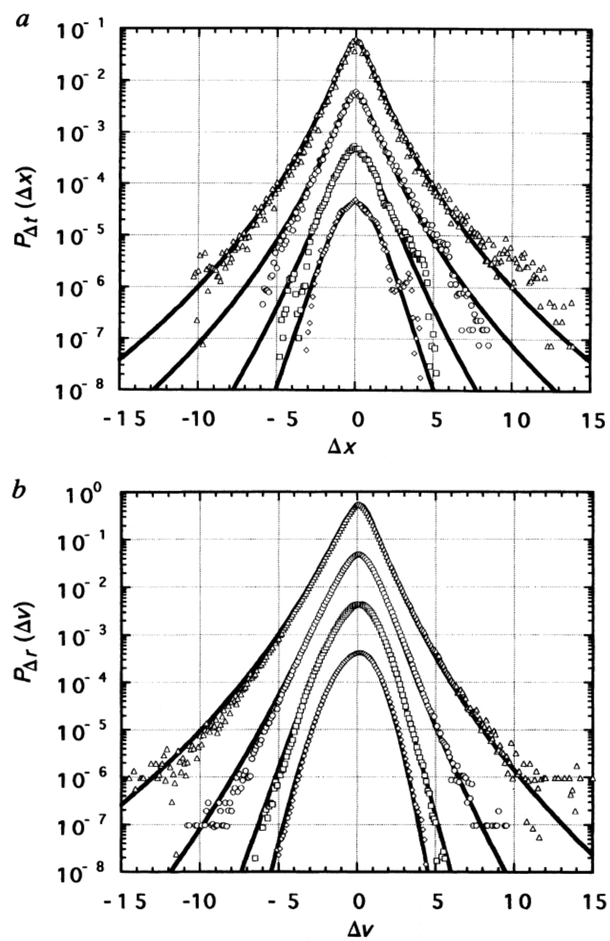

FIG. 1: a) Probability density functions of price changes of stocks Δx for different time delays Δt of 640s, 5120s, 40960s and 163840s (from top to bottom). b) Probability density functions for the turbulent flow case. Fits to these real data will be explained later in the text.

TABLE 1 Correspondence between fully developed three-dimensional turbulence and FX markets

Hydrodynamic turbulence	FX markets
Energy	Information
Spatial distance	Time delay
Laminar periods interrupted by turbulent bursts (intermittency)	Clusters of low and high volatility
Energy cascade in space hierarchy	Information cascade in time hierarchy

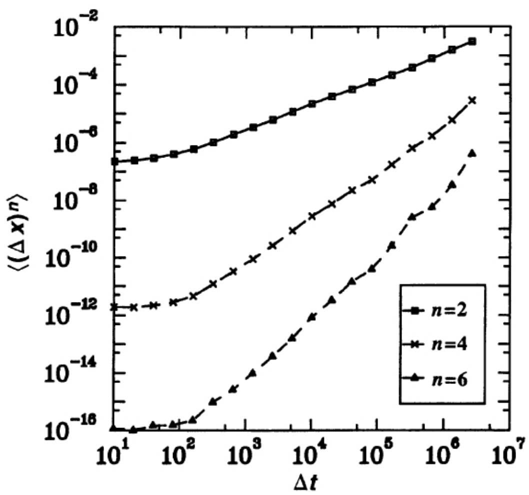

FIG. 2: Second, fourth and sixth moments for the change in stock prices.

the behavior of this systems, including Navier-Stokes like equations also for the case of stock market dynamics, motivated by the similarity of the two phenomena, which is accounted for by the existence of a cascade in both cases. In spite of all the efforts spent, some of the most basic questions concerning the statistics of financial assets were not answered with these models, in particular, the mechanism leading to the fat-tailed (*leptokurtic*) PDFs of fluctuations on small time scales. Quantifying the risks of such large fluctuations is of utmost importance for the pricing of options.

In the following sections, I will explain the approach to this problem developed by Friedrich *et al.* [3], in which they treat the dynamics of the market as Markovian and find a Fokker-Planck equation that describes the evolution of those prices for a large range of time scales Δt.

MATHEMATICS OF A MARKOV PROCESS AND FOKKER-PLANCK EQUATION

Friedrich *et al.* focused on the analysis of price changes measured as increments $\Delta x = x(t + \Delta t) - x(t)$, which are given in units of the standard deviation of Δx at $\Delta t = 40960s$. For this work they have used price quotes from several world markets. Figure 3 shows two of these data sets, corresponding to the NIKKEI 225 and the NASDAQ.

What kind of statistical process underlies the price changes over a series of time delays Δt_i of decreasing duration? In order to characterize the statistics of these price changes, joint probabilities density functions are defined as [4]

$$p^N(\Delta x_1, \Delta t_1; \Delta x_2, \Delta t_2, \ldots, \Delta x_N, \Delta t_N) \tag{1}$$

If the N Δx_i were statistically independent, the joint

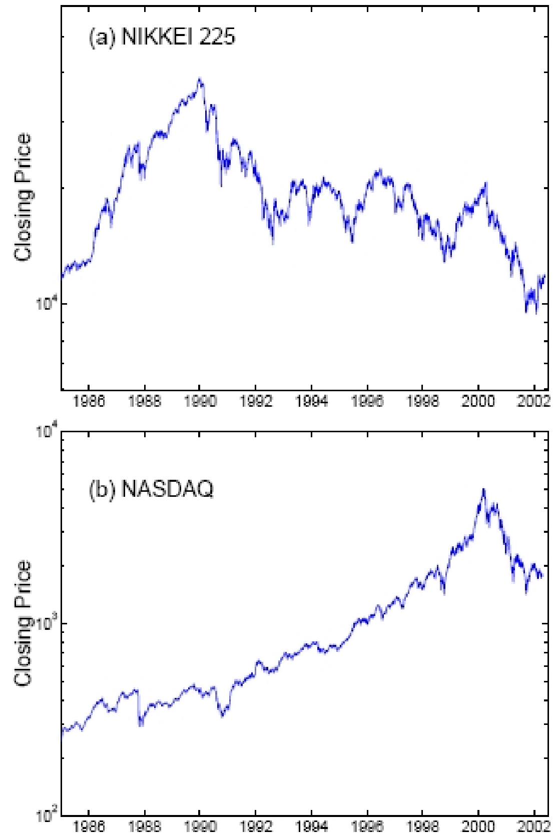

FIG. 3: Daily closing price of a) NIKKEI 225 and b) NASDAQ exchange rates for the period from January 01, 1985 to May 31, 2002.

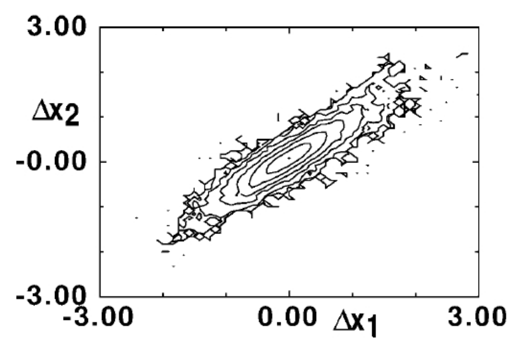

FIG. 4: Contour plot of the joint probability $p^2(\Delta x_1, \Delta t_1; \Delta x_2, \Delta t_2)$.

PDF (1) would be factorizable into a product of N PDFs $p(\Delta x_1, \Delta t_1)p(\Delta x_2, \Delta t_2) \ldots p(\Delta x_N, \Delta t_N)$. This is not the case of the stock market price changes. A plot of a two-price changes joint probability $p^2(\Delta x_1, \Delta t_1; \Delta x_2, \Delta t_2)$ is shown in Figure 4. The tilted form of the joint probability density clearly shows that such factorization does not hold and that the two price changes are statistically dependent. Intuitively, if market values went up over a certain period Δt_1, then it is more likely that, on a shorter period Δt_2 within the larger one, the market went up instead of down. In the case of a Markov process, an important simplification arises: the N-point PDF p^N is generated by a product of the conditional probabilities

$$p(\Delta x_{i+1}, \Delta t_{i+1} | \Delta x_i, \Delta t_i) = \frac{p(\Delta x_{i+1}, \Delta t_{i+1}; \Delta x_i, \Delta t_i)}{p(\Delta x_i, \Delta t_i)} \tag{2}$$

for $i = 1, \ldots, N-1$. As a necessary condition, the Chapman-Kolmogorov equation

$$p(\Delta x_2, \Delta t_2 | \Delta x_1, \Delta t_1) = \int d(\Delta x_i) p(\Delta x_2, \Delta t_2 | \Delta x_i, \Delta t_i) \times \times p(\Delta x_i, \Delta t_i | \Delta x_i, \Delta t_i) \tag{3}$$

should hold for any value of Δt_i, with $\Delta t_2 < \Delta t_i < \Delta t_1$. It is convenient to consider a logarithmic time scale $\tau = ln(40960s/\Delta t)$. Then, the limiting case $\Delta t \to 0$ corresponds to $\tau \to \infty$. The Chapman-Kolmogorov equation formulated in differential form yields a master equation, which can take the form of a Fokker-Plack equation:

$$\frac{d}{d\tau} p(\Delta x, \tau) = [-\frac{\partial}{\partial \Delta x} D^{(1)}(\Delta x, \tau) + + \frac{\partial^2}{\partial \Delta x^2} D^{(2)}(\Delta x, \tau)] p(\Delta x, \tau) \tag{4}$$

The drift and diffusion coefficients $D^{(1)}(\Delta x, \tau)$ and $D^{(2)}(\Delta x, \tau)$ can be estimated directly from the data as moments $M^{(k)}$ of the conditional probability distributions:

$$D^{(k)}(\Delta x, \tau) = \frac{1}{k!} \lim_{\Delta \tau \to 0} M^{(k)} \tag{5}$$

$$M^{(k)} = \frac{1}{\Delta \tau} \int d\Delta x' (\Delta x' - \Delta x)^k p(\Delta x', \tau + \Delta \tau | \Delta x, \tau). \tag{6}$$

The coefficient $M^{(1)}$ shows a linear dependence on Δx, while $M^{(2)}$ can be approximated by a polynomial of degree 2 in Δx. This behavior was found for all scales τ and $\Delta \tau$ (Fig. 5). Therefore the drift term $D^{(1)}$ is well approximated by a linear function of Δx, whereas the diffusion term $D^{(2)}$ follows a function quadratic in Δx:

$$D^{(1)} = -0.44 \Delta x, \tag{7}$$

$$D^{(2)} = 0.003 exp(-\tau/2) + 0.019(\Delta x + 0.04)^2 \tag{8}$$

Since it has been consider a finite data set, it is not possible to prove rigorously that higher orders of $D^{(k)}$ are zero, but the good fits shown in Figure 5 give a hint for the validity of the assumption that the conditional probability density obeys a Fokker-Planck equation (corresponding to only $D^{(1)}$ and $D^{(2)}$ not equal to zero).

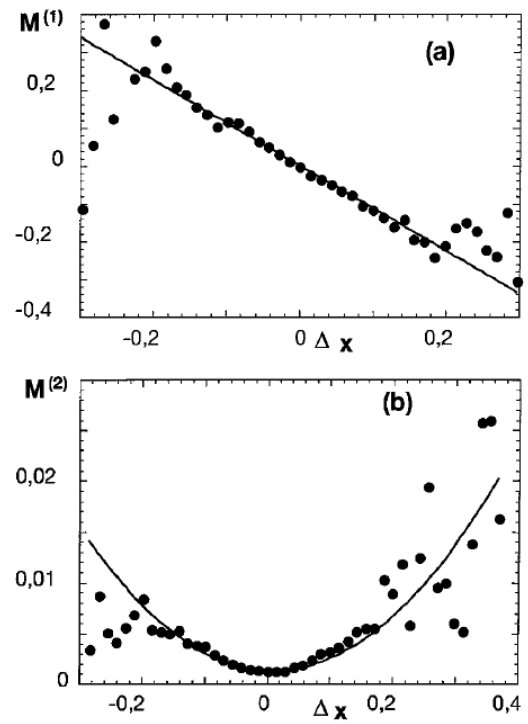

FIG. 5: Kramers-Moyal coefficients $M^{(1)}$ and $M^{(2)}$ estimated from the linear and quadratic fits of the conditional PDFs.

In order to test the solution, these coefficients were used for a numerical solution of the Fokker-Planck equation. Figure 1 shows the fits to the empirical PDFs, for long (Gaussian) and short (long-tail) time intervals. Now it becomes clear that one must not require stationary probability distributions for price differences for different Δt; on the contrary, the coupling between different scales Δt via a Markov process is essential to effectively fit the limit $\tau \to \infty$ ($\Delta t \to 0$).

THE LANGEVIN APPROACH: ADDITIVE VS. MULTIPLICATIVE NOISE

Notes on additive and multiplicative noises [4], [5]. As an example, consider the logistic system for the population n:

$$\frac{dn}{dt} = (k - n)n \tag{9}$$

The term kn represents the growth of the population n, with k being the growth frequency. The term n^2 represents competition for the same species, and sets a limit to the growth. There are two fixed points: $n = 0$ is unstable (unlimited growth of the population) and $n = k$, the stable one (system towards constant population).

a) *Additive noise*: add to eq. (9) a fluctuating term F(t), which is our delta correlated random force,

$\langle F(t_1)F(t_2)\rangle = f_0^2\tau\delta(t_1-t_2)$. This system has no memory on scales $k<1/\tau$. Iterating the continuity equation for the PDF f and averaging over the $F(t)$ time scale τ (f is slowly varying on τ), we can get to a Fokker-Planck equation

$$\frac{\partial\langle f(n,t)\rangle}{\partial t} = -\frac{\partial}{\partial n}[(k-n)n\langle f(n,t)\rangle - D\frac{\partial\langle f(n,t)\rangle}{\partial n}]$$
$$D \equiv \iint F(t)F(t)dtdt' \tag{10}$$

The stationary solution to (10) ($\partial\langle f\rangle/\partial t = 0$)leads to a solution of the form

$$\langle f\rangle \equiv f_{(n)}^{stat,add} = f_0 e^{\frac{1}{D}(\frac{kn^2}{2}-\frac{n^3}{3})} \tag{11}$$

$f_{(n)}^{stat,add}\to 0$ when $n\to\infty$, and $f_{(n)}^{stat,add}\to f_0<\infty$ when $n\to 0$.

In the case of additive noise, the distribution function is a Gaussian that decays (very fast) as e^{-n^3}, with n=k being the most probable value at equilibrium.

b) *Multiplicative noise*: now, instead of considering and additive $F(t)$ noise, let's take a noise of the form $k = k_0+k(t)$, with $k(t)$ being delta correlated: $\langle k(t_1)k(t_2)\rangle = k_0^2\tau\delta(k_1-k_2)$. The population equation is written now as

$$\frac{dn}{dt} = [(k_0+k_{(t)})-n]n = (k_0+k_{(t)})n - n^2 \tag{12}$$

in which is easy to see the multiplicative behavior of the noise. Repeating the same procedure as before, but replacing $F(t)\leftrightarrow k(t)n$, we get the new Fokker-Planck equation:

$$\frac{\partial\langle f(n,t)\rangle}{\partial t} = -\frac{\partial}{\partial n}[(k_0-n)n\langle f(n,t)\rangle - D\frac{\partial n^2\langle f(n,t)\rangle}{\partial n}] \tag{13}$$

The stationary solution in this case is of the form

$$f_{(n)}^{stat,mult} = f_0 e^{-\frac{2n}{D}} n^{(\frac{2k_0}{D}-2)} \tag{14}$$

$$\int_0^\infty f_{(n)}^{stat,mult}dn < \infty \tag{15}$$

At $n\to\infty$, $e^{-\frac{2n}{D}}$ dominates and the integral is finite. In the limit $n\to 0$, $n^{(\frac{2k_0}{D}-2)}$ dominates and it is integrable if $k_0>\frac{D}{2}$. As opposed to the additive noise case where the distribution function decays very fast and has a very short tail, the multiplicative noise problem has a PDF that at equilibrium decays much slower, with a longer tail, and, specially, we have to impose a condition to the diffusion coefficient in order to have an integrable $f_{(n)}^{stat,mult}$.

* * * * **

We know that a Fokker-Planck equation is equivalent to a Langevin equation of the form:

$$\frac{d}{d\tau}\Delta x(\tau) = D^{(1)}(\Delta x(\tau),\tau) + \sqrt{D^{(2)}(\Delta x(\tau),\tau)}F(\tau) \tag{16}$$

where $F(\tau)$ is a fluctuating δ-correlated force with Gaussian statistics: $\langle F(\tau)F(\tau')\rangle = 2\delta(\tau-\tau')$. In the approximation (6) for large τ (small Δt) the stochastic process is very close to a linear stochastic process with multiplicative noise:

$$\frac{d}{d\tau}\Delta x(\tau) = -0.44\Delta x(\tau) + \sqrt{0.19}\Delta x(\tau)F(\tau) \tag{17}$$

(Note constant and linear perturbation terms in $D^{(2)}$ are neglected, because they affect only very small values of Δx). For small τ (large Δt) the influence of the additive term in $D^{(2)}$ (*i.e.* additive noise) becomes more important, which explains that for these τ values the form of the PDFs is closer to a Gaussian shape. For large τ (small Δt) the multiplicative noise dominates and the PDFs become more and more heavy tailed. In the limit $\tau\to\infty$ ($\Delta t\to 0$), the PDF is given by the stationary solution of the Fokker-Planck equation, which exhibits extremely non-Gaussian behavior. However, using the numerical values given in equation (6), it is easily verified that the PDFs in Figure 1 are still far away from the stationary state.

CONCLUSIONS

The knowledge acquired over many years of research on turbulent fluid dynamics has permitted us to establish a better understanding of current stock market prices dynamics. In this paper we can see the main points of this parallelism, and extend it by using a newer model of analysis of stock prices changes. Furthermore, we have seen that the smooth evolution of the PDFs down along the cascades towards smaller time delays is caused by a Markov process with multiplicative noise.

[1] U. A. Müller, M. M. Dacorognaa, R. D. Dava, R. B. Olsen, O. V. Picteta and J. E. von Weizscker, Journal of Empirical Finance **4**, 213 (1997)

[2] A. N. J. Kolmogorov, J. Fluid Mech. **13**, 82-85 (1962)

[3] R. Friedrich, J. Peinke and Ch. Renner, Phys. Rev. Let. **84**, 5224 (2000)

[4] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, Second Edition (1992)

[5] C. W. Gardiner, *Handbook of Stochastic Methods*, Second Edition (1997)