Vol 16 No 4, April 2007
1009-1963/2007/16(04)/0975-09

Chinese Physics

© 2007 Chin. Phys. Soc.
and IOP Publishing Ltd

Description of dynamics of stock prices by a Langevin approach*

Huang Zi-Gang(黄子罡)[a], Chen Yong(陈 勇)[a],
Zhang Yong(张 勇)[b], and Wang Ying-Hai(汪映海)[a]

[a] *Institute of Theoretical Physics, Lanzhou University, Lanzhou* 730000, *China*
[b] *Department of Physics, Center for Nonlinear Studies, and The Beijing-Hong Kong-Singapore Joint Center for Nonlinear and Complex Systems (Hong Kong), Hong Kong Baptist University, Kowloon Tong, Hong Kong, China*

(Received 27 July 2006; revised manuscript received 18 August 2006)

We have studied the Langevin description of stochastic dynamics of financial time series. A sliding-window algorithm is used for our analysis. We find that the fluctuation of stock prices can be understood from the view of a time-dependent drift force corresponding to the drift parameter in Langevin equation. It is revealed that the statistical results of the drift force estimated from financial time series can be approximately considered as a linear restoring force. We investigate the significance of this linear restoring force to the prices evolution from its two coefficients, the equilibrium position and the slope coefficient. The daily log-returns of S&P 500 index from 1950 to 1999 are especially analysed. The new simple form of the restoring force obtained both from mathematical and numerical analyses suggests that the Langevin approach can effectively present not only the macroscopical but also the detailed properties of the price evolution.

Keywords: financial time series, Langevin approach, drift parameter, autocorrelation
PACC: 0580, 0540, 6120L

1. Introduction

The analysis of financial data by methods developed for physical systems has a long tradition,[1−5] and has attracted the interest of physicists. One of the most motivated reasons is that it is a great scientific challenge to understand the dynamics of a strongly fluctuating complex system with a large number of interacting elements. In addition, it may be possible that the experience gained in the process of studying complex physical systems might yield new results in economics.

There are many observables generated from financial markets, and one central issue of the research on the dynamics of financial markets is the statistics of price changes which determine losses and gains. The price changes of a time series of quotations $x(t)$ are commonly measured by returns: $r := x(t+\tau)/x(t)$, log-returns, or increments: $\Delta x := x(t+\tau) - x(t)$ at a time scale τ. In 1900, Bachelier proposed the first model for the stochastic process of returns—an uncorrelated random walk with independent identically distributed Gaussian random variables.[1] However, prices do not follow a single random walk process.[6−10] For example, the daily correlation has been known as the daily log-returns correlated with themselves in such a way that positive returns are followed by positive returns as well.[11,12] Many considerations have been aroused by this effect recently and meanwhile the related research has been reported, not only for daily data[13] but also for high-frequency data.[14,15]

The Langevin equation (LE) has been used to deal with many problems in physics such as Brownian motion.[16−19] It is also one of the methods widely used to investigate phenomena in financial markets, which distinguishes the evolution of the system into the deterministic and random terms. Recently, LE classification has been employed to describe the dynamics of mean-reverting financial variables.[20] In particular, some traditional processes for modelling volatilities[21] or, mainly, squared volatilities are the Hull–White[22] and the Heston[23] models, the latter also known either as the Cox–Ingersoll–Ross[24] or Feller process.[25] The arithmetic and geometric Ornstein–Ulhenbeck processes[26] are particular cases too. On the other

*Project supported by the National Natural Science Foundation of China (Grant No 10305005), the Fundamental Research Fund for Physics and Mathematics of Lanzhou University (Grant No Lzu05008).

hand, using the Langevin approach, Friedrich *et al*[27] and Ivanova *et al*[28] used the statistical way to distinguish and quantify the deterministic and the random influence on the hierarchical structure of the financial time series.

Different from the former studies, we reconstruct the Langevin equation directly from the log-return series. That is, the numerical results of the LE's drift parameter $A(z)$ and diffusion parameter $B(z)$ are directly estimated from log-return series. Based on a sliding-window algorithm, we analyse the time-dependent properties of $A(z)$. The relations between autocorrelation C, average return $\langle z\rangle$ and $A(z)$, from which the practical significance of $A(z)$ is exhibited, are obtained both from mathematical and numerical analyses. It is valuable that our Langevin description contains, as a particular case with flat-$A(z)$, the effect of daily correlation in log-returns. The research is mainly carried out from the samples of the daily log-returns $z(t)|_{\tau=1\mathrm{d}}$ of S&P 500 index from 1950 to 1999, containing 12583 days, thus covering a wide time range with many different economic and political situations.

This paper is organized as follows. In Section 2, taking the daily log-returns as an example, we generally discuss the application of Langevin approach to log-returns series. In Section 3, we show the results and discussions. Finally, the summary and the outlook of this paper are given in Section 4.

2. The Langevin approach to log-returns series

For a time series of prices or market index values $x(t)$, the log-return $z(t) \equiv z(t)_\tau$ over a time scale τ is defined as the forward change in the logarithm of $x(t)$,

$$z(t)_\tau \equiv \ln x(t+\tau) - \ln x(t). \tag{1}$$

The behaviour of daily log-return $z(t)_{1d}$ as a stochastic variable is described by the following LE:

$$\mathrm{d}z = A(z)\mathrm{d}t + B(z)\mathrm{d}w, \tag{2}$$

where the drift parameter $A(z)$ and diffusion parameter $B(z)$ describe respectively the deterministic and the random influences on the time process of log-returns, and $\mathrm{d}w$ denotes the increment of a standard Wiener process. It is assumed that within each sampling window the parameters may depend on the log-returns, but not explicitly on time (stationary). Thus, the drift and diffusion parameters of the sampled period can be extracted simply using the definition,[29]

$$A(z) = \lim_{\Delta t\to 0}\frac{1}{\Delta t}\langle Z(t+\Delta t) - z\rangle|_{Z(t)=z}, \tag{3}$$

$$B(z)^2 = \lim_{\Delta t\to 0}\frac{1}{\Delta t}\langle (Z(t+\Delta t) - z)^2\rangle|_{Z(t)=z}. \tag{4}$$

Here $\langle\ldots\rangle$ denotes the averaging operator and $Z(t+\Delta t)$ is a realization of the LE (2). From Eq.(3), it is obvious that the drift parameter $A(z)$ is the average increment (per unit time) under the condition $Z(t) = z$, and $B(z)$ is the deviation of $A(z)$ which exhibits the random influences. It has been known that the autocorrelation of the log-returns decays very fast,[6] usually characterized by a correlation time much shorter than a trading day.[7] When the time increment Δt is larger than 1 day, the daily log-returns can be considered as the results of many uncorrelated 'shocks'. Thus, in this paper, Δt is mainly set as 1 day. Compared with the length of time window T we will adopt that $\Delta t = 1\mathrm{d}$ approximately accords with the limit in Eqs.(3) and (4).

Sampling all over the long log-return series, the statistical results of $A(z)$ and $B(z)^2$ respectively estimated by Eqs.(3) and (4), denoted as $A'(z)$ and $B'(z)^2$, can be written in general forms. That is, $A'(z)$ is close to a linear form as

$$A(z) = az + b, \tag{5}$$

and $B'(z)^2$ is close to a parabolic form as

$$B(z)^2 = a'z^2 + b'z + c'. \tag{6}$$

As an example, Fig.1 presents the statistical results $A'(z)$ and $B'(z)^2$ (circles) estimated from the daily log-returns series of S&P 500 index from 28 May 1962 to 10 June 1970 (window length $T = 2000\mathrm{d}$). There have been systematical discussions about the linear or nonlinear drift parameter, and additive or multiplicative diffusion parameter.[30−32] However, here, we will just consider the simple forms presented in Eqs.(5) and (6), where $A(z)$ corresponds to a linear drift term, and $B(z)$ corresponds to a state-dependent linear multiplicative noise term $B(z)\mathrm{d}w$ in Eq.(2). The linear multiplicative noise term, which is required by the Langevin description to describe the variability of the log-return series, can be interpreted as the variability of log-returns that increases with log-returns itself.

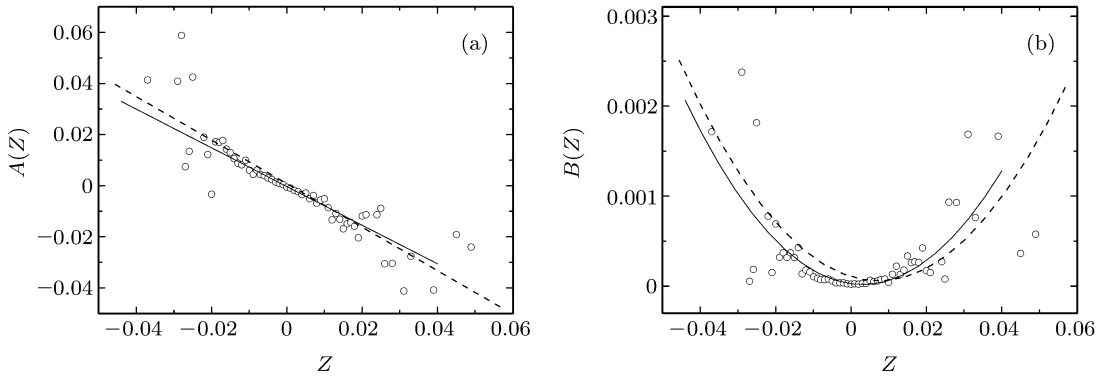


Fig.1. The statistical results of $A(z)$ and $B(z)$ (circles) estimated from the daily log-returns of S&P 500 index from 28 May 1962 to 10 June 1970, a window with length $T = 2000$ days. Their fitting results with equal-weight (dashed lines) and weight factor $P(z_i)$ (solid lines) correspond respectively to the fitting errors $\sigma_A = 2.564\times10^{-4}$ and 8.650×10^{-3} and $\sigma_B = 1.336 \times 10^{-5}$ and 5.144×10^{-4}. The expressions for solid lines are $A(z) = -0.7569z - 3.24305 \times 10^{-4}$ and $B(z)^2 = 0.92478z^2 - 0.00562z + 2.48829 \times 10^{-5}$.

The statistical results for large z are more noisy and uncertain than the points near the origin (see Fig.1), because these border points are visited rarely by the log-returns series. On the contrary, as viewed from statistics, the more the given value z_i ($i = 1, 2, ..., n$, n is the number of regions of z) is visited, the more times the averaging operator $\langle \ldots \rangle|_{Z(t)=z_i}$ in Eqs.(3) and (4) works, which would produce more accurate and reasonable $A'(z_i)$ and $B'(z_i)^2$. Thus, while fitting $A'(z_i)$ ($i = 1, 2, ..., n$) with a linear curve and $B'(z_i)^2$ ($i = 1, 2, ..., n$) with a parabolic curve by least-squares fit, each value z_i should correspond to its own weight. Here, we define the weight factor of z_i proportional to its frequency,

$$P(z_i) = N(z_i)/T, \tag{7}$$

where the $N(z_i)$ is the frequency of z_i within the time window of length T. Figure 1 shows the fitting results to $A'(z)$ and $B'(z)^2$ with weight factor $P(z_i)$ (solid lines) and with equal-weight (dashed lines). The values of the fitting error σ_A and σ_B (i.e., the mean standard deviations of linear and parabolic fitting results from $A'(z)$ and $B'(z)^2$) with weight factor $P(z_i)$ are visibly lower than those with equal-weight. Therefore, while fitting the statistical results, the point at z_i should take the weight $P(z_i)$.

The frequency $N(z_i)$, which influences the accuracy of $A'(z)$ and $B'(z)^2$, also depends on the window length T. Generally, $N(z_i)$ increases with T and thus $A'(z)$ and $B'(z)^2$ become more accurate for large T. We investigate the relations between T and the fitting errors σ_A and σ_B. In order to estimate the error purely made by window length T, instead of by information the time series embodied, we randomly rearrange the daily log-returns series. Then, the fitting errors σ_A, σ_B vs. T (100d − 13000d) are calculated (Fig.2). We find that the σ_A and σ_B of the fitting results with weight factor $P(z_i)$ decline quickly with a perfect power-law behaviour, $\sigma_A = 8.178 \times 10^{-3} \cdot T^{-0.49022}$ and $\sigma_B = 1.249 \times 10^{-4} \cdot T^{-0.31122}$. This behaviour suggests that the more accurate statistical results of $A'(z)$ (or $B'(z)^2$) are more feasible to be fitted with linear (or parabolic) form. That is to say, the linear (or parabolic) fitting line is suitable for the expression of $A'(z)$ (or $B'(z)^2$). Note, incidentally, that only the results with the same window length T could be compared since different T corresponds to different fitting errors.

The algorithm for the detection of the time-dependence of drift term $A(z)$ is as follows: sample the long log-returns series with a sliding window of short length T, and estimate $A'(z)$ for each sampling, whose linear fitting curve presents the characters of financial markets in the corresponding periods. This algorithm is more sensitive than merely studying transient behaviour. The comment for the selection of window length T is to choose T long enough so that the averages in Eqs.(3) and (4) are statistically meaningful, but not so long as to lose the temporal resolution. In the results presented below, 250d and 2000d window lengths and 5d overlapping (window shift by 5 days per time) are used. The corresponding fitting errors are, $\sigma_{A,250\mathrm{d}} = 5.46\times10^{-4}$, $\sigma_{B,250\mathrm{d}} = 2.24\times10^{-5}$, and $\sigma_{A,2000\mathrm{d}} = 1.97 \times 10^{-4}$, $\sigma_{B,2000\mathrm{d}} = 1.17 \times 10^{-5}$. It is expected that the variation of $A(z)$ as a function of time can accurately indicate interesting dynamical changes in financial process.

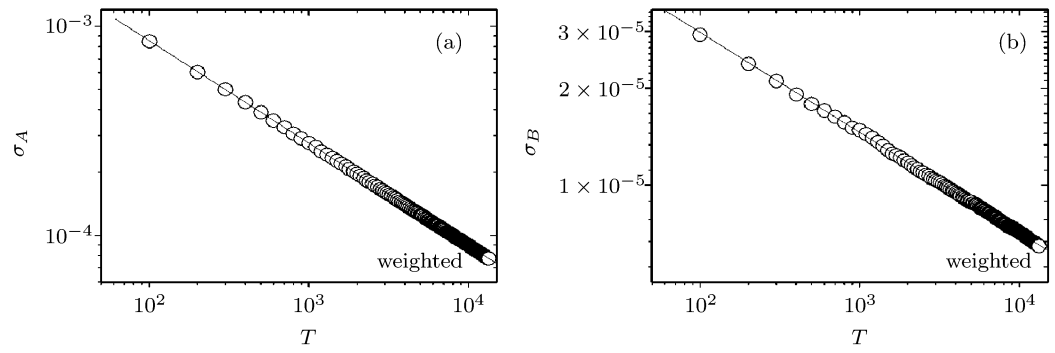


Fig.2. Log-log plots of weighted fitting errors σ_A and σ_B vs. T, with T = 100d–13000d. They exhibit perfect power-law behaviour. The solid lines are power-law regression fits over the entire range, which give estimates of power-law exponents, $\eta_A = 0.49022$ and $\eta_B = 0.31122$.

3. Results and discussions

In the LE, the negative-sloped drift parameter $A(z)$ with the form of Eq.(5) behaves like a restoring force. We call the position $z_0 = -b/a$, at which the force $A(z_0) = 0$, the equilibrium position. And $A(z)$ corresponds to a symmetrical confining potential with $A(z) = -\nabla V(z)$. In the following, the time-dependence of restoring force $A(z)$ will be discussed, including the equilibrium position z_0 and the slope coefficient a. In addition, discussions on the error analysis will be given.

3.1. Equilibrium position z_0

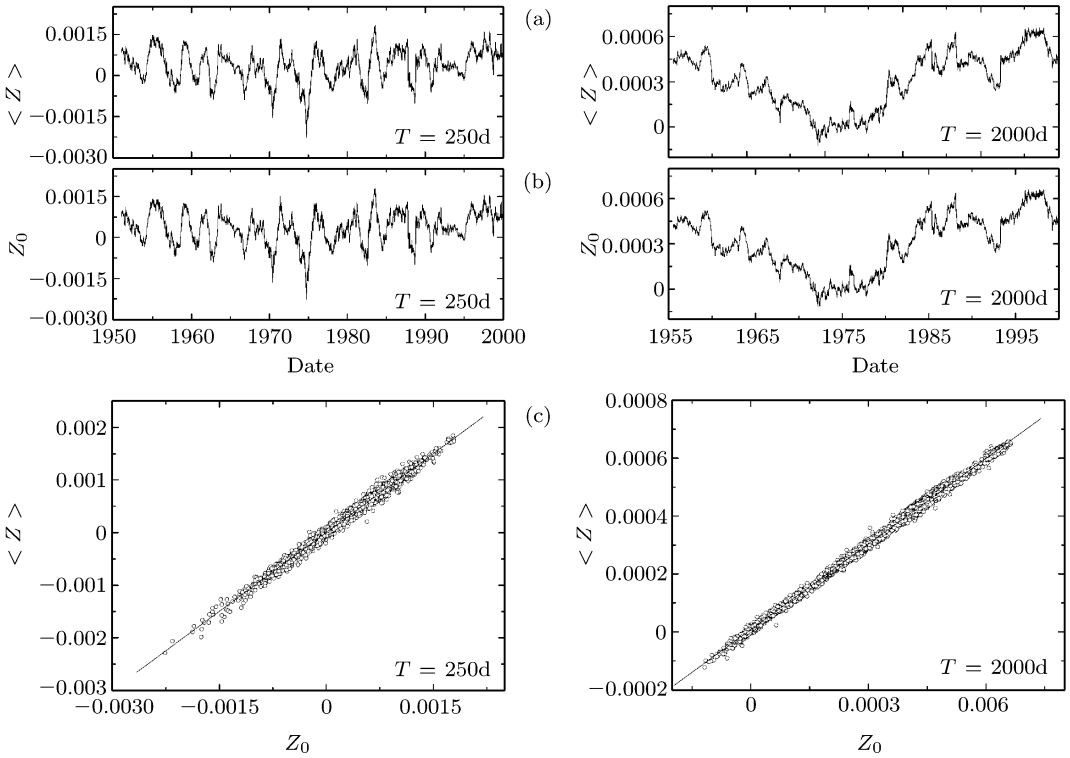


Fig.3. Sliding window analysis of the time-dependence of (a) the average log-return $\langle z \rangle$, and (b) the equilibrium position z_0. The lengths of sliding windows are T = 250d (left) and T = 2000d (right), and windows shift by 5 days per time. (c) The correlation function between z_0 and $\langle z \rangle$. Their linear fitting lines (solid lines) are $\langle z \rangle = 0.9985 z_0 + 3.4060 \times 10^{-6}$ (left), and $\langle z \rangle = 0.9856 z_0 + 7.8319 \times 10^{-6}$ (right).

From the Langevin equation,[26] the so-called equilibrium position z_0, with zero value of force $A(z)$, corresponds to the minimum of the potential well. Also, the average displacement $\langle z\rangle$ of an oscillating particle in the potential well should be the minimum of the potential well, $(-b/a, 0)$. Thus, we have

$$z_0 \simeq \langle z\rangle. \tag{8}$$

The average displacement $\langle z\rangle$ is directly obtained from the log-returns series, and the equilibrium position z_0 is calculated from Eq.(5). The time-dependence of z_0 and $\langle z\rangle$ coincide with each other very well all over the ranges (see Figs.3(a) and (b)), and the plot of z_0 as a function of $\langle z\rangle$ in Fig.3(c) is in excellent agreement with Eq.(8).

In the language of finance, the average log-return $\langle z\rangle$ describes the macroscopical trend of the price movement: $\langle z\rangle > 0$ indicates going up, and $\langle z\rangle < 0$ indicates going down. Therefore, based on Eq.(8), the rising trend of prices could be estimated if the statistical result $z_0 = -a/b > 0$; otherwise, the falling trend could be estimated if $z_0 = -a/b < 0$. Thus, the equilibrium position z_0 of restoring force $A(z)$ can be regarded as the 'trend index' of stock prices. Furthermore, the stock prices and their log-returns are macroeconomic indicators which are widely used because of the strong correlation between financial markets and economic development. In this case, the equilibrium position z_0, which is derived from the LE description of financial time series, would be another important indicator of macroeconomics.

3.2. Slope coefficient a

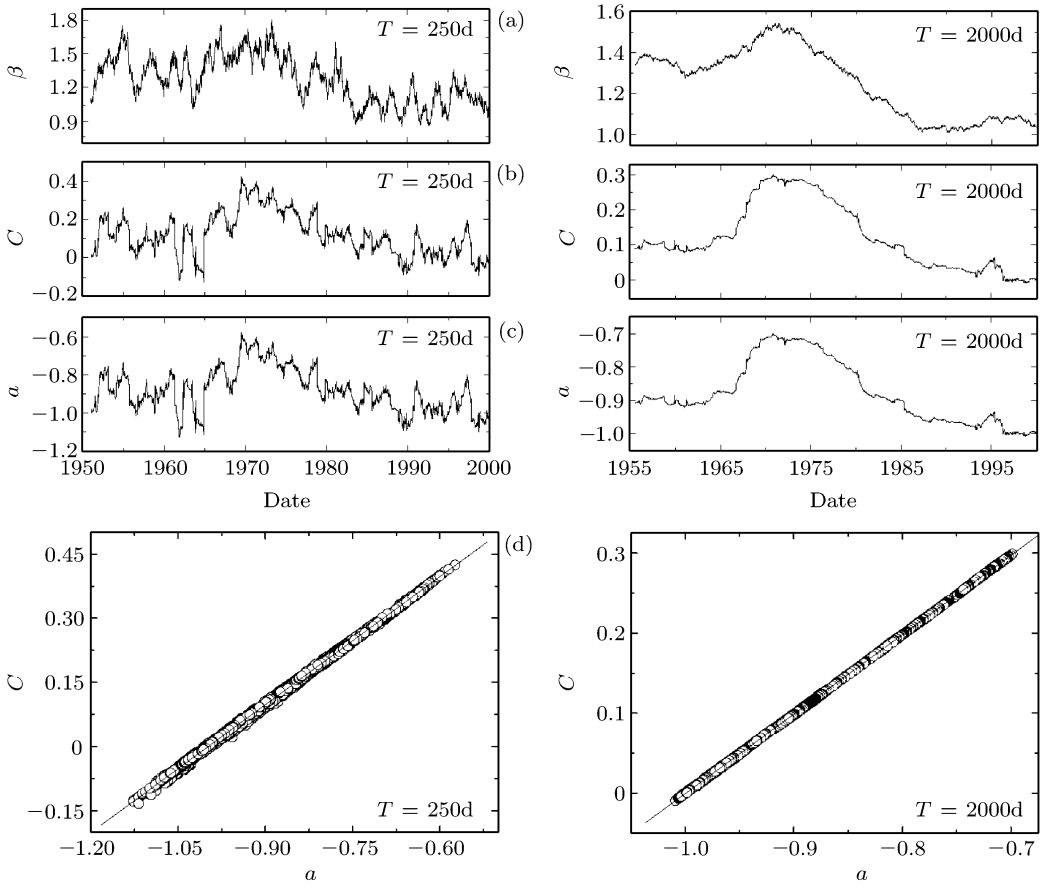


Fig.4. Sliding window analysis of the time-dependence of (a) the slope coefficient a with $\Delta t = 1$d, (b) the autocorrelation function $C_{\Delta t=1}$, and (c) the proportion of sign-converted cases and sign-sustained cases β. The lengths of sliding windows are $T = 250$d (left) and $T = 2000d$ (right), and windows shift by 5 days per time. (d) The correlation function between a and C. Their linear fitting lines (solid lines) are $C = 0.9942 + 0.9958a$ (left), and $C = 0.9932 + 0.9938a$ (right).

In Fig.4(a), the time-dependence of the slope coefficient a calculated from the daily log-returns of S&P 500 with $T = 250$d and 2000d are plotted. The ranges of the slope coefficient a with $T = 250$d and 2000d are $[-1.127, -0.574]$ and $[-1.009, -0.699]$ respectively, both of which are close to -1. We know that in a certain given position z_i, steeper (flatter) slope of $A(z)$ corresponds to larger (smaller) restoring force $A(z_i)$. Thus, one can imagine the mechanism of our model: it would take less time for larger forces (slope: $a < -1$) to draw particles from one side of the equilibrium position z_0 to the other side, which we call 'z_0-crossing' action for the moment, and more time for smaller forces (slope: $-1 < a < 0$). To discuss this mechanism more accurately, we use the normalized daily log-returns,

$$g = \frac{z - \langle z \rangle_T}{\nu}, \nu = \sqrt{\langle z^2 \rangle_T - \langle z \rangle_T^2}, \tag{9}$$

which has zero mean value, $\langle g \rangle = 0$. Here the standard deviation ν of log-returns stands for the time-averaged volatility,[10] and $\langle \ldots \rangle_T$ denotes an average over the entire length of the series within time window T. From Eq.(8), g has the equilibrium position z_0 equal to zero, so that the z_0-crossing action could be reduced to the sign conversion of g. Thus the mechanism can be described as follows: The sign of g changes frequently while the slope is quite steep; on the contrary, the same signs congregate together and sequences of consecutive '+' or '−' appear when the slope is flat.

The sign series of daily log-returns has been considered for the study of the daily correlation in log-returns.[13] Reference [13] obtained the conditional dynamics from the sequence of consecutive '+' and '−'. In this paper we will compare the same-sign sequence with the converted sign one of two neighbouring days to show the relationship between a and z_0-crossing.

The sign-cases of a given day and its previous day are '++', '+−', '−+' and '−−'. One can define '++' and '−−' as the sign-sustained cases, and '+−' and '−+' as the sign-converted cases. Incidentally, the contribution of '++' to the sign-sustained cases is shown to be a little more than '−−' on average over the whole series. Then, the time series of signs of g is investigated by counting the frequencies of sign-sustained cases and sign-converted cases (denoted as m and n, respectively). We calculate the time-dependence of the proportion $\beta = m/n$ from the daily log-returns of S&P 500 with $T = 250$d and 2000d (Fig.4(c)). The similarity of β and a proves their relations and the existence of the correlation between neighbouring daily log-returns. In detail, in one sampling period, while the slope of the restoring force is flat, the given day's sign of normalized log-return is more likely to be the same as its previous day; in contrast, while the slope is steep, the given day's sign is more likely to be different from its previous day. Thus, the mechanism of the daily correlation determined by the sign of log-returns is qualitatively explained by the effect of restoring force.

It is easy to see from Figs.4(a) and (c) that for most of the periods over the whole series, a are larger than -1, and β are larger than 1, with mean values $\bar{a}|_{T=250\text{d}} = -0.883$, $\bar{a}|_{T=2000\text{d}} = -0.874$, and $\bar{\beta}|_{T=250\text{d}} = 1.242$, $\bar{\beta}|_{T=2000\text{d}} = 1.223$. These imply that, in practice, the flat-slope restoring force is more prevalent than the steep-slope one, and the sign-sustained case appears more frequently than the sign-converted one. Consequently, from the general appearance of sign-sustained cases, we arrive at the same conclusion as that of Ref.[13]: the return of the price during a given day can be correlated with the previous day, in particular with the sign of the previous day. Furthermore, it is notable that the analysis with the Langevin approach is general because it contains the positive daily correlation as the particular case with flat-slope restoring force.

On the other hand, the autocorrelation function $C(\Delta t')$, a typical important statistics of stochastic processes, is always used to investigate pairwise correlation of the log-returns of a financial asset. We know that

$$C(\Delta t') = \frac{\langle z(t) \cdot z(t + \Delta t') \rangle_T - \langle z(t) \rangle_T^2}{\langle z(t)^2 \rangle_T - \langle z(t) \rangle_T^2}, \tag{10}$$

where $\langle \ldots \rangle_T$ denotes time averaging over all the trading days within the sampling local period with length T, and $\Delta t'$ the time increment. In the following, we compare the autocorrelation function of the slope coefficient a both from mathematical and numerical analyses.

For the weighted linear least-squares fitting we have used to approximate the statistical results $A'(z)$, the weighted objective function[33,34] can be written as

$$q = \sum_{i=1}^{n} W_i [A(z_i) - (az_i + b)]^2, \tag{11}$$

where W_i is the weight of the point $(z_i, A(z_i))$. It is known that the best values of a and b correspond to the minimal value of function q. From Eq.(11), the solution of slope coefficient a in Eq.(5) is

$$a = \frac{\sum_i W_i \sum_i z_i A(z_i) W_i - \sum_i z_i W_i \sum_i A(z_i) W_i}{\sum_i W_i \sum_i z_i^2 W_i - (\sum_i z_i W_i)^2}. \tag{12}$$

In the above discussion, the weight of z_i is defined as its probability $P(z_i)$ (Eq.(7)). Thus one obtains

$$a = \frac{\sum_i P(z_i) \sum_i z_i A(z_i) P(z_i) - \sum_i z_i P(z_i) \sum_i A(z_i) P(z_i)}{\sum_i P(z_i) \sum_i z_i^2 P(z_i) - (\sum_i z_i P(z_i))^2}. \tag{13}$$

In fact, numerical results of a are calculated by this statistic formula. Substituting Eqs.(3) and (7) into Eq.(13), we have

$$a = \frac{\langle z \cdot A(z) \rangle_T - \langle z \rangle_T \langle A(z) \rangle_T}{\langle z^2 \rangle_T - \langle z \rangle_T^2} = \frac{1}{\Delta t} \left[\frac{\langle z(t) \cdot z(t+\Delta t) \rangle_T - \langle z(t) \rangle_T \langle z(t+\Delta t) \rangle_T}{\langle z(t)^2 \rangle_T - \langle z(t) \rangle_T^2} - 1 \right]. \tag{14}$$

When T, compared to Δt, is sufficiently long, the stationary assumption of financial time series is $\langle z(t) \rangle_T \approx \langle z(t+\Delta t) \rangle_T$. Then, comparing Eq.(14) with Eq.(10), we can obtain the relation between a and C,

$$a(\Delta t) \approx \frac{1}{\Delta t}[C(\Delta t) - 1], \tag{15}$$

which is valid for any value of Δt because there is no limit on Δt during the derivation. However, since the log-returns are lack of correlation for $\Delta t > 1$d, only the case with $\Delta t = 1$d is analysed. We also compare the numerical results of $C_{\Delta t = 1\mathrm{d}}$ and a. The correlation function between them (see Fig.4) shows the good effectiveness of Eq.(15). The analytical results indicate that the sign-cases, slope coefficient a, and autocorrelation function $C_{\Delta t = 1\mathrm{d}}$ reflect the similar properties of time series. Since we know that the correlations observed in financial time series show the incompleteness of the efficient market hypothesis,[13] the three coefficients C, a, and β may probably indicate the degree of market efficiency. From the same tendency of the three coefficients shown in Fig.4 with $T = 2000$d (right), we obtain the following two points: (i) the market was lack of efficiency from 1961 to 1976 relatively, since the values are much larger than the remaining 25 years; (ii) the market tended to be increasingly efficient from 1968 to 1999 because of the decreasing trend of the values.

From the preceding analysis of the equilibrium position z_0 and the slope coefficient a, we finally obtain the form of restoring force $A(z)$ by substituting Eqs.(8) and (15) into Eq.(5),

$$A(z) = \frac{1}{\Delta t}(C-1)(z - \langle z \rangle). \tag{16}$$

This new form as a function of the traditional qualities C and $\langle z \rangle$, is a direct way to understand the dynamical behaviour of time series. As to the financial data, the information given by $A(z)$ includes features of the macroscopical properties together with the detailed properties of the price evolution; that is, the macroscopical trend of prices is presented by z_0, and the detailed correlation between two neighbouring days is expressed by a.

Equation (16) can be informatively rewritten as $A(z) = -(z - z_0)/t_0$, where $t_0 = -\Delta t/a$ is the characteristic relaxation time; t_0 is more visualized than the slope coefficient a, both of which reflect similar properties of the log-returns. The maximum values of t_0 calculated from the sliding windows of various lengths are all less than 2 days, and the average values of t_0 with $T = 250$d and $T = 2000$d are $\bar{t_0}|_{250\mathrm{d}} = 1.153$d and $\bar{t_0}|_{2000\mathrm{d}} = 1.157$d. Therefore, the effect of the restoring force $A(z)$ of one given day decays with the characteristic relaxation time less than 2 days.

3.3. Fitting error σ

In addition, we also investigate the weighted fitting error of $A(z)$ and $B(z)$. Figure 5(b) shows the time-dependence of σ_A and σ_B calculated from the sliding window with $T = 250$d down the log-returns series. The time-averaged volatility of log-returns, a measurement of how liable the market is to fluctuate, is calculated from the same 250d sliding window (see Fig.5(c)). Compared with the original daily log-returns series (Fig.5(a)), one can easily find that the variations of σ_A and σ_B (Fig.5(b)), as well as that of the time-averaged volatility (Fig.5(c)), show sudden jumps when very volatile periods enter or leave the time window. For example, as pointed out by arrows, the jumps at $T = 9500$d and 3110d are caused by the crashes in May 1962 and on the 'black Monday', 19 Oct. 1987. Therefore, one can conclude that the fitting errors, σ_A and σ_B, indeed respond to the volatility of the financial markets: i.e., the information of historical events in real market would influence the accuracy of the fitting results presented in this paper.

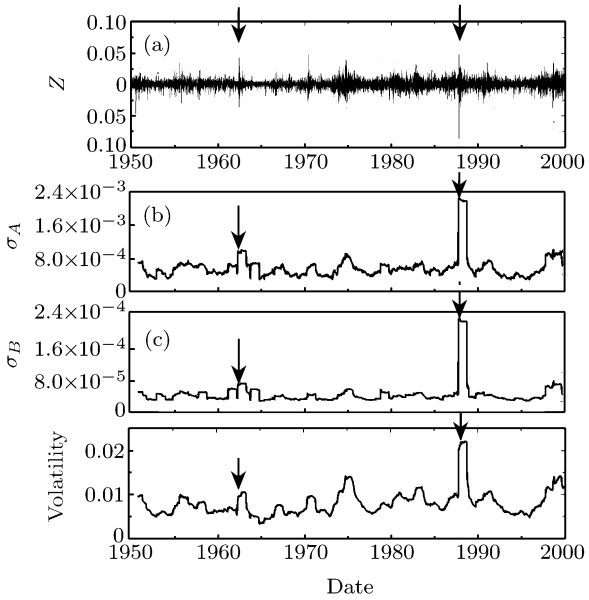


Fig.5. (a) Daily log-returns of S&P 500 index from 1950 to 1999. (b) The time-dependence of fitting errors σ_A and σ_B with $T = 250$d. (c) The time-averaged volatility of the same series with $T = 250$d.

3.4. The results of the log-return series $z(t)|_{\tau\neq 1\mathrm{d}}$

The preceding analysis concentrates on the daily log-return $z(t)|_{1\mathrm{d}}$. In this section, we also investigate $z(t)|_\tau$ with $\tau = 2\mathrm{d}, 3\mathrm{d}, 4\mathrm{d}, 5\mathrm{d}$ in the light of our Langevin approach. It can be concluded that all the forms of $A(z)$ and $B(z)$ (Eqs.(5) and (6)), the correlation function between z_0 and $\langle z\rangle$ (Eq.(8)), the correlation function between a and C (Eq.(15)), and the new form of restoring force $A(z)$ (Eq.(16)) are still effective for these series. Table 1 lists the results from the log-return series of the S&P 500 index on the time scale τ from 1d to 5d.

Table 1. The results from the log-return series of the S&P 500 index on the time scale τ from 1d to 5d. The average slope $\bar{a}$, the average value of autocorrelation $\bar{C}$, the average proportion of the sign-sustained and sign-converted cases $\bar{\beta}$, the average characteristic relaxation time $\bar{t_0}$ with $T = 250$d and 2000d are listed in this table.

τ(d)	$\bar{a}$		$\bar{C}$		$\bar{\beta}$		$\bar{t_0}$(d)	
	250d	2000d	250d	2000d	250d	2000d	250d	2000d
1	-0.883	-0.874	0.115	0.125	1.242	1.223	1.153	1.157
2	-0.467	-0.462	0.533	0.539	2.273	2.252	2.178	2.187
3	-0.303	-0.299	0.697	0.702	3.205	3.137	3.395	3.414
4	-0.224	-0.222	0.776	0.779	4.136	4.059	4.416	4.633
5	-0.182	-0.177	0.818	0.825	4.986	4.886	5.759	5.822

4. Summary and outlook

In this paper, we present a time-dependent Langevin description of the dynamics of stock prices, which is verified to be effective by the results obtained from analysing the S&P 500 index. The time-dependence of the LE's drift parameter $A(z)$ considered as the restoring force is investigated by the sliding window algorithm. Significantly, when choosing the right weight factor (Eq.(7)) to fit the statistical results of $A(z)$, the linear fitting results can reflect both the macroscopical and the detailed properties of the price evolution, and the new form of the restoring force, Eq.(16), is achieved from analytical methods. The mechanism of our model is also discussed by analysing the sign series of log-returns, and the

daily correlation in log-return is exhibited by the flat-slope $A(z)$. Therefore, from the restoring force $A(z)$ in Langevin approach, one can obtain the properties of experimental data, including the financial time series we analysed in this work. Furthermore, it must be pointed out that the diffusion parameter $B(z)$ also plays an important role in the dynamics of financial markets, which will be addressed further in our future work.

Acknowledgment

We would like to thank Professor Zhao Hong and Dr Xu Xin-Jian for helpful discussions.

References

[1] Bachelier L 1900 *Ann. Sci. École Norm. Suppl.* **3** 21
[2] Pareto V 1927 *Manuel d'Economie Politique* (Paris: Marcel Giard)
[3] Levy P 1937 *Théorie de l'Addition des Variables Aléatoires* (Paris: Gauthier-Villars)
[4] Mandelbrot B B 1963 *J. Business* **36** 294
[5] Mao X M, Sun K and Ouyang Q 2002 *Chin. Phys.* **11** 1106
[6] Qiu T, Zheng B, Ren F and Trimper S 2006 *Phys. Rev.* E **73** 065103
[7] Matengna R N and Stanley H E 1971 *An Introduction to Econophysics* (New York: Oxford University Press)
[8] Zhang Y C 1999 *Physica* A **269** 30
[9] Liu Y, Gopikrishnan P, Cizeau P, Meyer M, Peng C K and Stanley H E 1999 *Phys. Rev.* E **60** 1390
[10] Gopikrishnan P, Plerou V, NunesAmaral L A, Meyer M and Stanley H E 1999 *Phys. Rev.* E **60** 5305
[11] Fama E 1970 *J. Finance* **25** 383
[12] LeBaron B 1992 *J. Business* **65** 199
[13] Boguna M and Masoliver J 2004 *Eur. Phys. J.* B **40** 347
[14] Ohira T, Sazuka N, Marumo K, Shimizu T, Takayasu M and Takayasu H 2002 *Physica* A **308** 368
[15] Sazuka N, Ohira T, Marumo K, Shimizu T, Takayasu M and Takayasu H 2003 *Physica* A **324** 366
[16] Zhou G Z, Su Z B, Hao B L and Yu L 1980 *Acta Phys. Sin.* **29** 961 (in Chinese)
[17] Yang H J, Zhuo Y Z and Wu X Z 1994 *Acta Phys. Sin.* **43** 539 (in Chinese)
[18] Zhan Y, Zhao T J, Yu H and Song Y L 2002 *Chin. Phys.* **11** 624
[19] Wang X Y and Meng Q Y 2004 *Acta Phys. Sin.* **53** 388 (in Chinese)
[20] Hull J and White A 1990 *Rev. Financ. Stud.* **3** 573
Cox J C, Ingersoll J E and Ross S A 1985 *Econometrica* **53** 385
[21] Fouque J P, Papanicolaou G and Sircar K R 2000 *Derivatives in Financial Markets with Stochastic Volatility* (Cambridge: Cambridge University Press)
[22] Hull J and White A 1987 *J. Financ.* **42** 281
[23] Heston S L 1993 *Rev. Financ. Stud.* **6** 327
[24] Cox J C, Ingersoll J E and Ross S A 1985 *Econometrica* **53** 363
[25] Feller W 1951 *Ann. Math.* **54** 173
[26] Risken H 1984 *The Fokker–Planck Equation* (Berlin: Springer)
[27] Friedrich R, Peinke J and Renner Ch 2000 *Phys. Rev. Lett.* **84** 5224
[28] Ivanova K, Ausloos M and Takayasu H cond-mat/0301268
[29] Sura P and Barsugli J 2002 *Phys. Lett.* A. **305** 304
[30] Richmond P 2001 *Eur. Phys. J.* B **20** 523
[31] Anteneodo C and Riera R 2005 *Phys. Rev.* E **72** 026106
[32] Wang C J, Chen S B and Mei D C 2006 *Chin. Phys.* **15** 1435
[33] Neter J, Wasserman W and Kutner M H 1983 *Applied Linear Regression Models* (Irwin, IL: McGraw-Hill)
[34] Seber G A F 1977 *Linear Regression Analysis* (New York: Wiley)