PACS numbers: **05.40.–a**, 05.45.Df, 89.65.Gh
DOI: 10.3367/UFNe.0181.201107k.0779

Econophysics and the fractal analysis of financial time series*

M M Dubovikov, N V Starchenko

1. Introduction

The term '*econophysics*' was coined in 1995 by H Eugene Stanley as a common name for research in which methods of statistical physics were applied to analyzing the behavior of financial markets. Such research efforts were stimulated most of all by the revolution in computer technologies, which by that time had led to the creation of huge readily accessible arrays of financial data which had been painstakingly accumulated from the middle of 1980s. Later on, the term began to be used in a wider context, indicating that a paper on economics or another social science had been written by a physicist. Since 2002, such publications have begun to appear regularly in all the major general physical journals, such as *Reviews of Modern Physics*, *Physical Review E*, *Physical Review Letters*, and some others. At this time, a course on econophysics is being taught in the West in the most prestigious universities, and a section dedicated to it has grown to become an integral part of major annual international and national conferences on social sciences [e.g., ESHIA (Economic Science with Heterogeneous Interacting Agents), AKSOE (Arbeitskreis Physik Sozio-Ökonomischer Systeme), and others]. The first All-Russia Congress of Econophysics was convened in June 2009 in Moscow.

The first econophysics work whose popularity far transcended the bounds of any one field of science was the publication by Mantegna and Stanley in *Nature* [1]. In essence, this work developed the rather old idea of Benoit Mandelbrot concerning the *Lévy flight* [2] so as to make it to agree with new empirical data.

This article is devoted to the development in the same field of another of Mandelbrot's seminal ideas, which was also first advanced in paper [2] in the study of financial time series. Subsequently, this idea has been successfully applied in a number of very different fields of physics [3].

Ever since the 1950s, experts have been quite familiar with the proposition that movements of prices of most financial tools over various time and price scales look very similar. An observer cannot identify from the shape of the charts if the data describe weekly, daily, or hourly fluctuations [3]. In today's language, the indicated self-similarity signifies that financial time series are *fractals* [4]. The main characteristic of such structures is, as we know, the fractal dimension D. In the case of chaotic time series, this indicator defines the Hurst index H ($D = 2 - H$), which is a measure of persistence in a time series (the ability to sustain a certain trend). However, an impossibly large representative scale is required for a reliable calculation of D (as well as H), which excludes any chance of using D as an indicator defining the local dynamics of a time series.

In this paper we introduce new fractal parameters: the *dimension of minimum covers* and the related *index of fractality*. It has been rigorously proved in the principal order in δ (here, δ is the minimum scale of partitioning the time series) that for $\delta \to 0$ the dimension of the minimum cover is identical to D. By the example of financial time series, it has been proved that the amount of data contained in the minimum scale required for determining the introduced indicators with an acceptable accuracy is less by two orders of magnitude than the corresponding scale for the determination of the Hurst index H. This makes it possible to consider the index of fractality as a local indicator of stability of the time series. An empirical justification of the concept of *stability* on the financial market has been proposed, based on the index of fractality. An effect of enhancement of large-scale fluctuations and suppression of small-scale oscillations has been revealed; it was used to build an indicator of strong fluctuations in the global financial market.

2. Fractal structures

1. Objects for which the methods of *classical* analysis proved totally unsuitable (such as the Cantor set, the Weierstrass function, and the Peano curve) were found in mathematics for the first time at the end of the 19th century. All of them were built using very simple rules of iterative procedure, and all possessed scalable self-similarity (consisted of parts similar to the whole). By the beginning of the 20th century, the number of such objects became sufficiently large, and to analyze them, Felix Hausdorff offered in 1919 his definition of the dimension of a compact set in an arbitrary metric space [5]. Hausdorff noticed that if these sets are covered by spheres with a radius δ, the minimum number $N(\delta)$ of such spheres will grow with diminishing δ by the power-law dependence

$$N(\delta) \sim \left(\frac{1}{\delta}\right)^D . \qquad (1)$$

Notice that the power exponent D is typically calculable exactly. This was the exponent that Hausdorff called 'dimension'.[1] If we now take the logarithms of both sides of this expression and rewrite them in the form of an equality for D, we obtain the exact definition of the Hausdorff dimension:

$$D = \lim_{\delta \to 0} \left[\frac{\ln N(\delta)}{\ln(1/\delta)}\right] . \qquad (2)$$

For sets that are familiar in classical calculus (e.g., smooth curves or surfaces), the exponent D coincides with the

* Dedicated to the memory of Benoit Mandelbrot (20.11.1924–14.10.2010).

M M Dubovikov, N V Starchenko INTRAST Management Company, Moscow Engineering Physics Institute, Moscow, Russian Federation
E-mail: Dubovikov@intrast.ru

Uspekhi Fizicheskikh Nauk **181** (7) 779 – 786 (2011)
DOI: 10.3367/UFNr.0181.201107k.0779
Translated by V I Kisin; edited by A Radzig

[1] This quantity is sometimes called the ultimate capacity (see, e.g., monograph [6]) perceiving here by the Hausdorff dimension d_H the critical value of the argument of the function

$$m(p) = \sup_{\varepsilon>0} \inf_{\{A_i^\varepsilon\}} \sum (\operatorname{diam} A_i^\varepsilon)^p ,$$

which possesses the following property: $m(p) = \infty$ for $p < d_H$, and $m(p) = 0$ for $p > d_H$. Here A_i^ε is the coverage of the original set by the family of sets A_i^ε with a diameter less than ε. Such a dimension is of more recent origin and extends to an unbounded set. As a rule, $d_H = D$, but counterexamples have been found. One can only state in the general case that $d_H \leqslant D$.

topological dimension D_T equal to the minimum number of coordinates necessary to describe such sets (e.g., one coordinate is sufficient to describe the line, two coordinates to describe the surface, and three coordinates to describe the body [2]). It was found that for the nonclassical sets mentioned above, the Hausdorff dimension (typically a fractional quantity) is always greater than the topological dimension D_T. The last property was later used by Mandelbrot for one of the possible definitions of the fractal, which states that a fractal is a set where $D > D_T$ [3].

It should be noted that if the original set is immersed in an Euclidean space, then other approximations of the set cover by simple shapes (e.g., cells) of size δ can be used instead of covering this set by spheres. In addition, new fractal dimensions (cellular, internal, etc.) appear, along with the original spherical dimension D; they usually coincide as limiting values for $\delta \to 0$. However, the rates of convergence to this limit may vary considerably for these dimensions.

Consider, for instance, the *Sierpiński carpet*, which is constructed as follows. Take a unit square, and in the first step divide it into nine equal squares, of which the middle one is thrown out (Fig. 1a). In the next step, this procedure is repeated with all remaining squares, and so forth. In the limit, the set obtained by an iterative procedure is known as the Sierpiński carpet (it can be shown that $D_T = 1$ for this object). Notice that in constructing model fractals, a set consisting of $N(\delta)$ elementary simplexes of linear size δ usually emerges in the nth iteration step. Mandelbrot called this set the pre-fractal of the nth generation. This set for the Sierpiński carpet consists of $N(\delta) = 8^n$ cells with a side length $\delta = (1/3)^n$. If now we use pre-fractals in definition (2) instead of covering set by spheres, the dimension D can be calculated in a straightforward manner. Indeed, passing to the limit $n \to \infty$ in our case, we obtain from formula (2):

$$D = \lim_{\delta \to 0} \left(\frac{\ln 8^n}{\ln 3^n} \right) = \lim_{n \to \infty} \left(\frac{n \ln 8}{n \ln 3} \right) = \frac{\ln 8}{\ln 3} \ (\approx 1.89) \,.$$

The result will not change if spheres are chosen instead of cells. However, the characteristics of the algorithms of direct calculation of these two dimensions are very different. In order to show this, we construct for each dimension the plot of the function $N(\delta)$ at $\delta = (1/3)^n$ on a double logarithmic scale (Fig. 1b). On this scale, all power-law functions are linear, and the exponent D is defined as the slope of the regression line corresponding to the plot. For cellular covers (pre-fractals), all points of the plot of the function $N(\delta)$ lie on one straight line. This means that the function $N(\delta)$ *rapidly reaches the asymptotic power mode* (1), which allows us to get the value of D already in the first iteration step. If we use spheres instead of cells for calculating D, the corresponding plot becomes closer to power law (1) only asymptotically as $\delta \to 0$. A more profound analysis reveals that the above property of $N(\delta)$ for pre-fractals of the Sierpiński carpet emerges due to the fact that cellular cover is in a sense a minimal cover in each iteration step. Therefore, it is precisely the minimality of the covers which is the reason why the appropriate function determined by the covers and being used to calculate the dimension D grows rapidly and reaches the power-law asymptotic mode. As will be shown in Section 3, this principle allows straightforward generalization to the case of chaotic time series.

2. Objects with a nontrivial Hausdorff dimension were for a long time regarded only as a figment of the sophisticated mathematical intellect. These days, largely through the efforts of Benoit Mandelbrot, we know that fractals are all around us. Some fractals are continually changing, like moving clouds or flickering flame, while others preserve the structure created in the process of evolution, as happened with coastlines, trees, or our vascular systems. The real range of scales in which fractal structures are observed stretches from intermolecular distances in polymers to distances between clusters of galaxies in the Universe.

We need to point to the main features of natural fractals that distinguish them from model ones. *First*, natural fractals are never strictly symmetrical. Self-similarity holds for them only on average. *Second*, calculations of the dimensions of natural fractals inevitably exclude scales that are smaller than a certain minimum scale δ_0 of the structure. This means that power law (1) manifests itself as an 'intermediate asymptotics' (as $\delta \to 0$, the scale considered is much smaller than a certain characteristic scale but greater than the minimum scale δ_0). *Third*, for natural fractals there is no system of pre-fractals. Therefore, the system of approximations by simplexes, required for the construction of the function $N(\delta)$ when $\delta \to 0$, is in the general case fairly arbitrary. Consequently, the computation of the dimension D as the slope of the regression line $N(\delta)$ on the double logarithmic scale needs a large amount of data, since the function $N(\delta)$ usually converges to the power law (1) very slowly.

We will show, nevertheless, in Section 3 that for computing the dimensions of fractal time series it is possible to construct a sequence of minimal covers similar to the sequence of pre-fractals of a Sierpiński carpet.

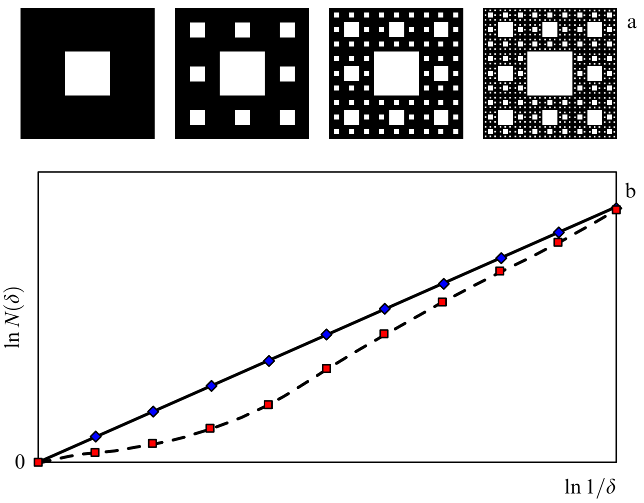

Figure 1. Pre-fractals of four generations for the Sierpiński carpet (a), and the function $N(\delta)$ in double logarithmic scale (b) for cellular (solid line) and Hausdorff (dashed curve) dimensions.

[2] This approach can be generalized (e.g., for an arbitrary compact set) in at least two ways [7]. The first approach is based on the fact that any two closed disjoint subsets of the original set of dimension $n+1$ can be split with a partition of dimension n. The dimension is introduced here by induction. The second method is based on the fact that the minimum multiplicity of the cover of a set of dimension n by closed sets with an arbitrarily small diameter equals $n+1$. Multiplicity is understood here to equal the maximum number of the cover elements having nonempty intersections.

3. The dimension of the minimum cover. The index of fractality

1. *Chaotic time series* form the most important class of natural fractals. Such series with an extremely irregular behavior are found in the observations of various natural, social, and technological processes. Some of these processes are traditional (geophysical, economic, medical), and some were discovered fairly recently (daily variations in crime level or in traffic accidents in an administrative region, fluctuations in the number of hits of certain sites on the Internet, etc.). Such series of data are usually generated by complex nonlinear systems of various natures. However, all of them behave in essentially the same characteristic manner within a certain range of scales. The easiest method of studying the fractal structure of these series is based on calculating the cellular dimension D_c. To find D_c, one divides the plane, on which the diagram of the time series is defined, into cells of size δ. Then, for different δ we plot the function $N(\delta)$ which is equal to the number of cells of size δ that contain at least one point on the diagram. The dimension D_c is found from the slope of the regression line $N(\delta)$ on the double logarithmic scale. It is readily shown that $D_c = D$. The fractal dimension for *chaotic* series happens to be especially important because this indicator is closely related to the Hurst exponent (index) H which is usually calculated using the normalized amplitude range and, as we know, is an indicator of *persistence* (ability to sustain trends) of a time series. Notice that if $H > 0.5$, the series is persistent (it is likely that the movement of the series in a certain direction on an interval will initiate movement in the same direction on the next interval). If $H < 0.5$, the series is *antipersistent* (it is likely that the movement of the series in a certain direction on an interval will initiate movement in the opposite direction on the next interval). Finally, if $H \approx 0.5$, the series has *zero persistence* (the motion of the series on any interval is independent of its motion on the previous interval).

More than ten different algorithms for the calculation of this indicator were created later owing to its importance [8–11]. It seems that the simplest method for calculating the exponent H is based on the formula

$$\langle |f(t+\delta) - f(t)| \rangle \sim \delta^H \quad \text{as} \quad \delta \to 0\,, \tag{3}$$

where angle brackets denote averaging over the time interval, and $f(t)$ is the value assumed by the time series at the instant of time t. The exponent H is found from the corresponding regression line. It is easy to show that for Gaussian random processes $H = 2 - D$. Virtually all experts agree that this relation has a wider range of applicability, since it has been confirmed for all the observed chaotic time series in all those cases in which both indicators are accurately determinable. Also, all difficulties associated with the computation of dimension D are transferred to the algorithms for the calculation of the exponent H. Thus, any reliable determination of H requires a representative scale of several thousand data sets. As a rule, a time series changes the parameters of its behavior on such a long scale many times, which greatly devalues the analysis of time series with the aid of the Hurst exponent H. As we saw for the dimension D, this difficulty stems from the fact that the convergence of the corresponding function to power law (3) for $\delta \to 0$ unfolds extremely slowly. To overcome this obstacle, it is possible to follow the analogy of how this is done in the case of the Sierpiński carpet and to determine the sequence of approximations of a series, which consists of minimal covers for any fixed δ. Indeed, if we multiply both sides of formula (1) by δ^2, the definition of the dimension can be rewritten as a power law for the approximation area $S(\delta)$:

$$S(\delta) \sim \delta^{2-D} \quad \text{for} \quad \delta \to 0\,. \tag{4}$$

Notice that, in contrast to formula (1), this form does not require that the simplexes of which each individual approximation consists be identical. It would be sufficient for them to have one and the same geometric factor δ. It is this circumstance that allows us to use approximations which are minimal covers.

2. Indeed, let a function $y = f(t)$ having not more than a finite number of points of discontinuity of the first kind be defined on a segment $[a, b]$: it is natural to consider precisely such functions as model ones, e.g., for financial time series. We introduce a uniform partition of the segment, $\omega_m = [a = t_0 < t_1 < \ldots < t_m = b]$, where $t_i - t_{i-1} = \delta = (b-a)/m$, $(i = 1, 2, \ldots, m)$. We cover the graph of this function with rectangles in such a way that this cover is the minimum area in the class of covers by rectangles with base δ (Fig. 2). Then the height of the rectangle on the segment $[t_{i-1}, t_i]$ equals the amplitude $A_i(\delta)$ which is the difference between the maximum and minimum values of function $f(t)$ on this segment. We now introduce a quantity

$$V_f(\delta) \equiv \sum_{i-1}^{m} A_i(\delta)\,. \tag{5}$$

The total area $S_\mu(\delta)$ of the minimum cover can then be written as $S_\mu(\delta) = V_f(\delta)\,\delta$. Consequently, formula (4) implies that

$$V_f(\delta) \sim \delta^{-\mu} \quad \text{for} \quad \delta \to 0\,, \tag{6}$$

where $\mu = D_\mu - 1$. We call the dimension D_μ *the dimension of the minimal covers*. To appreciate the differences between D_μ and other dimensions, especially the cellular dimension D_c, we construct the cellular partition of the plane of the graph of the function $f(t)$ as shown in Fig. 2. Let $N_i(\delta)$ be the number of cells that cover the plot of $f(t)$ on the segment $[t_{i-1}, t_i]$. It is seen from the figure that

$$0 < N_i(\delta)\,\delta^2 - A_i(\delta)\,\delta < 2\delta^2\,. \tag{7}$$

Divide this inequality by δ and sum up over i, taking into account Eqn (5). As a result, we have

$$0 < N(\delta)\,\delta - V_f(\delta) < 2(b-a)\,, \tag{8}$$

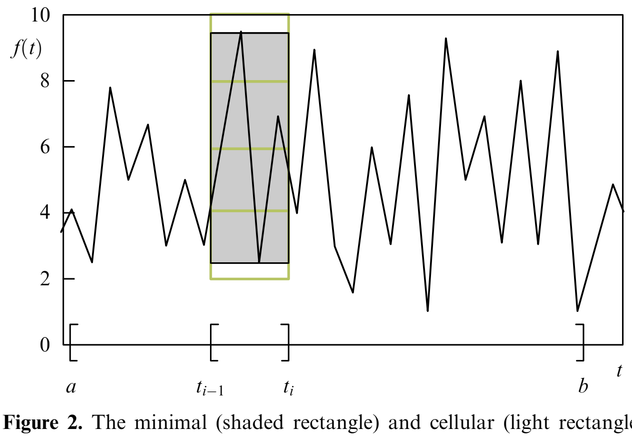

Figure 2. The minimal (shaded rectangle) and cellular (light rectangle) covers of the function $f(t)$ on the interval $[t_{i-1}, t_i]$ of length δ.

where $N(\delta) = \sum N_i(\delta)$ is the total number of cells of size δ that cover the plot of function $f(t)$ on the segment $[a, b]$. Passing to the limit $\delta \to 0$ and taking formula (6) into account, we obtain

$$N(\delta)\,\delta \sim V_f(\delta) \sim \delta^{-\mu} = \delta^{1-D_\mu}\,. \tag{9}$$

On the other hand, it follows according to formula (4) that

$$N(\delta)\,\delta = S_c(\delta)\,\delta^{-1} \sim \delta^{1-D_c}\,. \tag{10}$$

Hence, $D_c = D_\mu$. Note, however, that despite this equality, the minimal and cellular covers for real fractal functions may provide different convergences of the quantity $S(\delta)$ to the asymptotic mode (4), and this difference may be quite large. Next, since $D_c = D_\mu = D$, $\mu = D_\mu - 1$ and since $D_T = 1$ for the one-dimensional function, we have $\mu = D - D_T$. In this case, therefore, the index μ can naturally be called the *index of fractality*. In what follows, we will analyze financial time series and regard this index as the main fractal indicator.

4. Financial time series. Problems of identification and prediction

1. The most popular representatives of fractal time functions are financial time series (first and foremost, series of stock prices and currency rates). There is reliable numerical evidence of the fractal structure of such series [12, 13]. Theoretically, fractality is usually linked to the fact that investors with different investment horizons (from several hours to several years) must be active in the market for sustaining its stability. This is the factor that produces scaling invariance (absence of a singled out scale) of price series over the corresponding time interval [14, 15].

As an example, a database was investigated which included price series for shares of thirty companies included in the Dow Jones Industrial Index (DJII) from 1970 to 2002. Each series contained about eight thousand records. Each record corresponded to a running day of trading and included four values: the lowest and highest prices, and the opening and closing prices. In the literature, financial series are usually represented using the *Japanese candles chart*. A fragment of such a series for the Coca-Cola Co. is displayed in Fig. 3a. To simplify the analysis, only the last $2^{12} = 4096$ records for each company were considered. To compute the variation index μ, the sequence of m nested divisions ω_m for $m = 2^n$ $(n = 0, 1, 2, \ldots, 12)$ were used. Each division consisted of 2^n intervals containing 2^{12-n} trading days. For each division ω_m, the value of $V_f(\delta)$ was calculated by formula (6). Here, $A_i(\delta)$ equals the difference between the highest and lowest prices over the interval $[t_{i-1}, t_i]$ (thus, if $\delta = \delta_0$, $A_i(\delta)$ equals the difference between the highest and lowest prices over one day). A typical example of the behavior of $V_f(\delta)$ on the double logarithmic scale for the Microsoft company is illustrated in Fig. 4. We see that the data lie with amazing precision on a straight line, except for the last two points, which deviate from the linear mode and exhibit a 'break'. To find the value of μ from these data should exclude the last two points and determine the regression line. At the confidence level

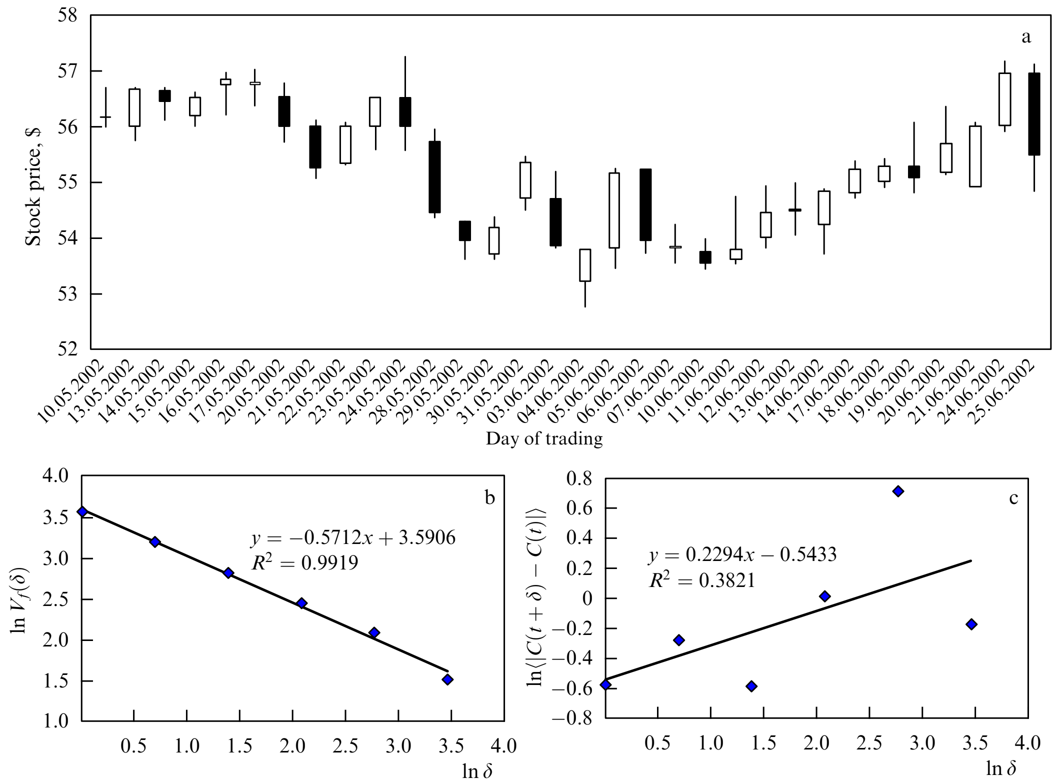

Figure 3. (a) Typical *Japanese-candles* financial series in the interval of 32 days (the graph of Coca-Cola share prices was used). Each of the rectangles (known as the candle body) with two vertical bars above and below (known as candle shadows) symbolizes price fluctuations during one day of trading. The top point of the upper shadow indicates the highest day price, and the bottom of the lower shadow is the lowest day price. The upper and lower boundaries of the candle body show the opening price and closing price on the trading day. The white (black) color of the candle body indicates that the closing price was above (below) the opening price. (b) The result of calculation of $V_f(\delta)$ on the double logarithmic scale for the presented time series. The relation $y = ax + b$ was calculated by the method of least squares; here $\mu = -a$. (c) The result of calculations of $\langle|C(t+\delta) - C(t)|\rangle$ for the same series, and the corresponding function $y = ax + b$, for $H = a$.

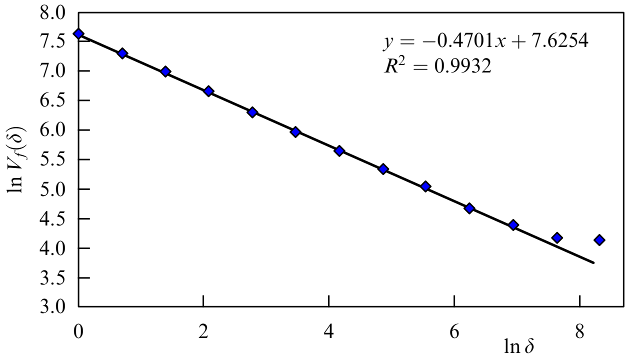

Figure 4. The result of calculation of $V_f(\delta)$ and the corresponding relation $y = ax + b$ for the time series of Microsoft stock prices in the interval of 4096 days.

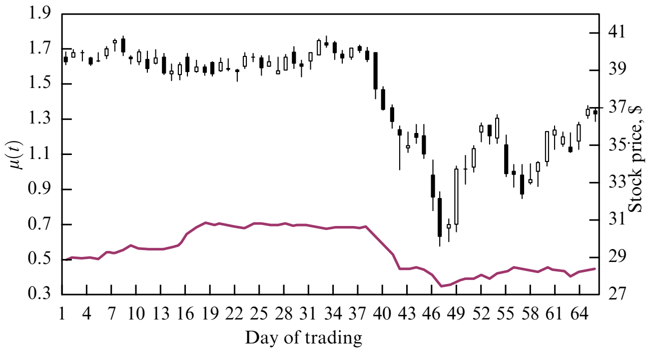

Figure 5. Daily prices of Exxon Mobil Corporation shares (Japanese candles, right-hand scale) and plot of the function $\mu(t)$ (solid curve, left-hand scale).

$\alpha = 0.95$, we have in the above example $\mu = 0.472 \pm 0.008$ and $R^2 = 0.999$. Here, R^2 is the determination index for the regression line. A comparison of this algorithm of computation of D ($D = \mu + 1$) and, correspondingly, H ($H = 2 - D = 1 - \mu$) with standard algorithms for computing these indices shows that the results are consistent with acceptable accuracy. However, the values of $V_f(\delta)$ on double logarithmic scale fall appreciably more accurately onto a straight line (except for the last two points) than those values corresponding to other algorithms, which also allows us to determine the characteristic scale on which the break of the linear mode occurs.

Now we need to point out that for each of the 30 companies the plot of $V_f(\delta)$ on the double logarithmic scale fits the straight line, almost as accurately, on all shorter representative intervals too, down to 32 days, and sometimes even down to 16 days. Note that on intervals shorter than 500 days the break on the linear part of the plot, as a rule, disappears.

A typical example of the behavior of the function $V_f(\delta)$ on a segment of financial time series 32 days long (Fig. 3a) is given in Fig. 3b. If $\alpha = 0.95$, we obtain $\mu = 0.571 \pm 0.071$, $R^2 = 0.992$. For comparison, Fig. 3c displays an example of the behavior of $\langle |C(t+\delta) - C(t)| \rangle$ on the same segment (we use here $32 + 1 = 33$ close prices $C(t)$ and averaging is carried out in nonintersecting intervals $\delta = 2^n$ in length, where $n = 0$, 1, 2, 3, 4, 5). In this case, $H = 0.229 \pm 0.405$, $R^2_{\mathrm{H}} = 0.382$. It immediately becomes evident that the calculation of the index H over this interval is simply meaningless.

The conclusion is that the fact that the quantity $V_f(\delta)$ rapidly reaches the asymptotic power-law mode makes it possible to reliably calculate the index of fractality μ over short intervals, as well. Further analysis showed that the power law for the function $V_f(\delta)$ fits the results with remarkable precision in the range of scales from several minutes to several years. It was understood that this property helps achieve significant progress in solving the two main problems of time series analysis: identification and prediction.

2. The problem of identification usually consists in determining the state of the system (the macrostate of the time series) on the basis of the observed values of the series in some local range. Specialists identify three types of local states for financial time series: trend (movement directed upward or downward), flat (relatively stable state), and random walk (intermediate state between trend and flat). In order to correlate the value of μ with the states of the financial time series, we introduce a function $\mu(t)$ as such value of μ that can still be calculated at an acceptable accuracy on a minimal interval τ_μ foregoing time t. If the argument t is continuous, we could choose for such an interval an arbitrarily small one. However, since in practical cases the time series always has a minimum scale (in our case, it spans one day), τ_μ is of finite length (in our case, we take $\tau_\mu = 32$ days). Such a function $\mu(t)$ was constructed for each of the companies in the Dow Jones index.

Figure 5 shows a typical fragment of the price series of one of these companies together with the function $\mu(t)$ calculated for this fragment. Suffice it to throw a quick glance at Fig. 5 to understand that the index μ has a direct bearing on the states of the time series. Indeed, $\mu(t) > 0.5$ in the interval between the 1st and the 39th days, where prices are relatively stable (flat). Further, simultaneously with the unfolding of the trend state in the price chart, $\mu(t)$ drops sharply to values below 0.5, and finally, after the 56th day when the prices are in an intermediate state between trend and flat, $\mu(t)$ returns to a value of $\mu \approx 0.5$. The original series thus becomes more stable as μ increases. Also, if $\mu > 0.5$, the flat state is observed, and if $\mu < 0.5$, the trend state is observed. Finally, if $\mu \approx 0.5$, then the series resides in the random walk state, which is intermediate between the trend and a flat states. Such a correlation between the value of μ and the behavior characteristics of the original time series was observed for all investigated series. A theoretical basis for this correlation can be found, for instance, in paper [16]. We will show below how the function $\mu(t)$ can be used to justify the classical theory of finance.

3. The basic model of financial time series is the random walk model.[3] Rethinking this model led to the concept of the *effective* market (Effective Market Hypothesis, EMH) on which the price fully reflects all available information. For such a market to exist, it is sufficient to assume that it has a large number of fully informed, rational agents with uniform preferences, which instantly adjust the prices and bring them into equilibrium. It is natural that the basic model of such a

[3] The first random walk model [17] was constructed by Luis Bachelier in 1900 (five years before Albert Einstein proposed his model of Brownian motion), who used it for describing the behavior of stock prices on the Paris Stock Exchange. Many of the results linked to this model, which were later obtained by other authors (Chapman–Kolmogorov equation, martingale theory, Black–Scholes equation), were already implied in Bachelier's paper.

market is the random walk model. It should be noted that all the main results of the classical theory of finance [portfolio theory, CAPM (Capital Asset Pricing Model), Black–Scholes model, etc.] have been obtained within the framework of precisely this approach. At present the "the concept of effective market continues to play a dominant role both in financial theory and in financial business" [20].

However, by the beginning of the 1960s some empirical studies showed that large fluctuations of yield series occur much more often than could be expected on the basis of the normal distribution (the problem of 'fat tails'), plus these large changes usually followed one another (effect of volatility clusterization). Mandelbrot [2] was one of the first to severely criticize the above concept. Indeed, if we calculate the value of the exponent H ($H = 1 - \mu$ in our case) for some share, then in all likelihood (see the beginning of this section), this value will differ from $H = 0.5$, which corresponds to the random walk model.

The reader will recall that two postulates lie at the base of this model. First, the price increments [4] in any time interval have a normal (Gaussian) distribution; this follows from the central limit theorem and is obtained as a result of summation of a sufficiently large number of independent random variables with finite variance. Second, these increments are statistically independent in disjoint intervals. It was the rejection of the first postulate, while maintaining the second one, that led Mandelbrot to consider a random process which he called the Lévy flight [2]. The rejection of the second postulate, while maintaining the first, led him to introducing the concept of *generalized Brownian motion* (Fractional Brownian Motion) [21].

The behavior of a time series for which $H \neq 0.5$ can be described using any of these processes. For the ideological base, people typically use *the concept of fractal market* (Fractal Market Hypothesis, FMH), which is usually considered as an alternative to EMH. This concept assumes that the market comprises a wide range of agents with different investment horizons and, therefore, with different preferences. These horizons vary from one minute for *intraday* traders to several years for banks and corporations. The stable equilibrium in this market is the regime for which the mean yield is independent of scale, except for a multiplication by the appropriate scale coefficient [2]. Since this coefficient has an undefined power exponent, we are actually dealing with a whole class of regimes, each of which is determined by its specific value of the index H. Consequently, the value of $H = 0.5$ is fully equivalent to any other value ($0 < H < 1$). Similar arguments caused serious doubts about the reality of equilibrium on the stock market (see, e.g., Refs [20–24]) and, hence, about the validity of the modern theory of finance.

Investigation of the function $\mu(t)$ using the initial basis (see the beginning of the section), as well as of Russian (included in the MMVB index) and American (included in the DJII index) companies, together with the corresponding indices for the last ten years, makes it possible to clearly show that the value $H = 0.5$ is a distinguished one.

Figure 6 displays typical probability distributions of the values of index μ for a time series of one of the shares included in the DJII index on the intervals of different lengths (from 8 to 256 days). All distributions are asymmetric. This means

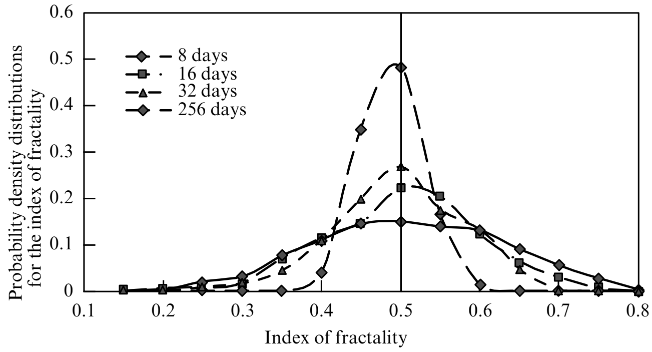

Figure 6. Probability density distributions for the values of the index μ for a series of daily changes in Ford share prices over a period from 03.01.2000 to 30.11.2010 (2745 records all-in-all, of trading days, each storing information on opening and closing prices, as well as daily records of the highest and lowest prices). The function $\mu(t)$ was calculated from the foregoing intervals containing from 8 to 256 values, and then empirical probability densities were constructed in each case from the values of μ.

that the average value of the index of fractality for this stock differs from $\mu = 0.5$ in appropriate intervals. However, all these distributions have the *principal mode* precisely at this value of μ.

To a first approximation, the following general pattern is observed in all series. The function $\mu(t)$ performs quasiperiodic oscillations around the position $\mu = 0.5$ between the values of $\mu < 0.5$ and $\mu > 0.5$. The mode of the time series is continuously changing, from the trend state via the random walk to flat and then back. For each series, the states with relatively stable values of μ (see Fig. 5) emerge time and again, then disappear. Among these states, the mode $\mu = 0.5$ occupies an obviously privileged position. For each time series, it is the longest in all intervals containing 8 or more points.

It should be emphasized that the agent-oriented interpretation of the price fluctuations may vary greatly on different scales. Thus, for example, agents' intraday behavior is apparently very close to rational behavior on a small scale when more than 50% of transactions are concluded (on U.S. markets) by trading robots. Unlike this, an essential role on scales from several days to several months is played by the social psychology, which always involves an irrational element. Incidentally, the unchanging nature of these fluctuations is reproduced on all scales, beginning with the shortest. This last remark points to a conjecture that some common mechanism of retardation, accompanying most decision-making processes, constitutes the nature of such fluctuations. But the principal state is, nevertheless, the random walk, which remains to act as the main attraction regime on all scales.

4. Generally, the prediction problem seeks to determine certain qualitative or quantitative parameters of the future behavior of a time series on the basis of the entire array of historical data. The most interesting in this situation is the problem of determining the earliest precursors of the critical behavior of a time series. We shall consider one approach to solving this problem. Starting with Eqn (5), we introduce the average amplitude $A(\delta)$ using the formula

$$A(\delta) \equiv \langle A_i(\delta) \rangle = m^{-1} V_f(\delta)\,. \tag{11}$$

[4] Various modifications of the Bachelier model [18, 19] typically operate with logarithms of price increments instead of increments themselves. This difference is not significant for us in this context.

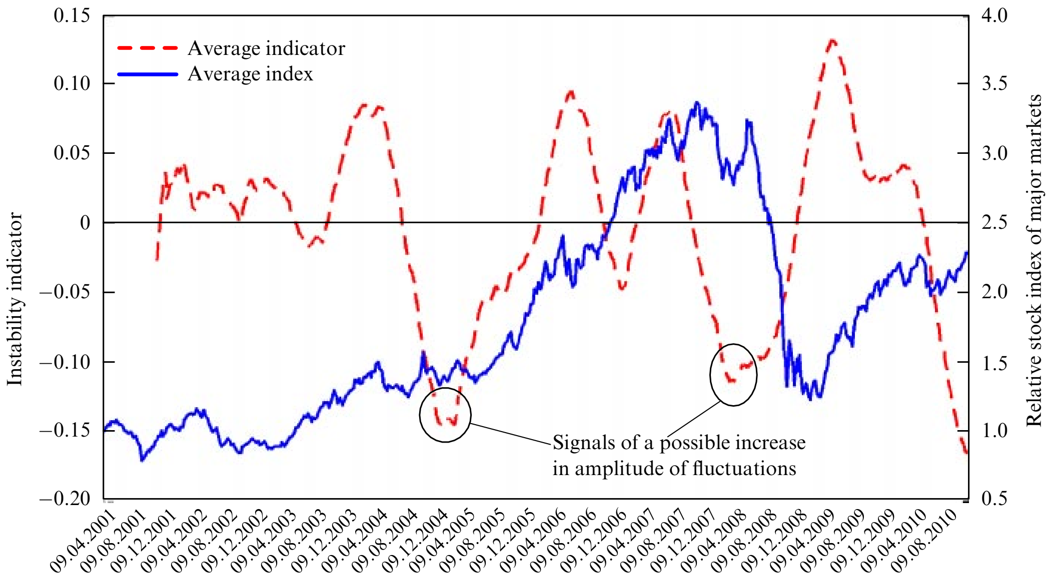

Figure 7. The aggregated index in relative units (solid curve, right-hand scale) and the appropriate indicator of instability centered about the mean value (dashed line, left-hand scale).

Multiplying Eqn (5) by $m^{-1} \sim \delta$ and substituting into Eqn (6), we obtain

$$A(\delta) \sim \delta^{H_\mu} \quad \text{for} \quad \delta \to 0\,, \qquad (12)$$

where $H_\mu \equiv 1 - \mu$. Comparing H_μ with Fig. 5, we obtain a visual confirmation that this index is a measure of persistence for a series (see item 3 of this section) and a direct generalization of the Hurst exponent H to small intervals. In item 3 of this section we have seen that the index μ essentially imposes a certain method of functional integration of the time series, and that it is then clear that the series corresponding to the random walk acquires the maximum weight. It then becomes possible to build a number of distributions, including conditional ones, over μ. It is becoming clear that the effect of enhancement of large-scale fluctuations, and at the same time a certain weakening of small-scale ones, plays a special role here. This effect manifests itself because, first, the power law for the function $A(\delta)$ (as well as for the function $V_f(\delta)$) holds in an immense range of scales: from several minutes to several years. Second, the power-law function possesses an important property: the slower it decays (in comparison with a function with a different power exponent) for $\delta \to 0$, the faster it grows for $\delta \to \infty$. This fact implies that a change in the regime of the system caused by an abrupt drop in μ (increase in the index H_μ) leads to a further suppression of small-scale fluctuations and at the same time to intensification of large-scale fluctuations in the series. This means that an abrupt drop in small-scale fluctuations at present may, under certain conditions, be a precursor to strong large-scale fluctuations in the future. Testing over the entire database mentioned above has demonstrated that this effect manifests itself at a probability of about 70–80%. Note that this percentage grows even higher in those cases in which the impact of external factors can be reduced to a minimum.

Figure 7 plots the indicator (dashed curve) which was constructed on the basis of this effect in the INTRAST Management Company in 2007. The solid curve represents the initial series, namely, the aggregated index including the stock indices of both developed and developing countries (one from each country).[5] This approach excludes the factor of mutual influences exerted by stock markets of different countries against each other and produced by intercountry flows of capital on the global financial market. Figure 7 gives evidence that twice after 2001 the indicator revealed an abrupt drop in small-scale fluctuations. The first time it happened was in December 2004, and it was followed by steep growth of all indices half a year later; the growth lasted for about two years. The second time it happened was in April 2008 after which — also about half a year later — the crisis triggered a sharp drop in all indices. Moreover, we see from this figure that a new signal is being actively formed now (08.11.2010), which is a precursor of intense fluctuations of the stock market in the mid-term (from six months to one year).[6] Even though the indicator says nothing about the direction of this strong movement, the information received may prove to be sufficient, e.g., for building a successful asset management strategy on the stock market.

To complete this section, several words are in order about the effect of intensification of large-scale fluctuations as the small-scale ones decrease. In essence, this effect signifies that the trends in complex systems (natural, social, technological), which form very slowly and imperceptibly but show more-than-average implacability, often grow globally with time and dictate the main vector of evolution of such systems. Note that the well-known *calmness effect* (i.e., suppression of the high-frequency noise component) which usually precedes natural disasters (e.g., earthquakes) is a particular manifestation of this effect. Therefore, in their evolution, many global trends do resemble *the mustard seed* (of the Gospel parable): "...Although it is the smallest of all seeds, when it is fully grown it is larger than the garden plants and becomes a tree, and the birds in the sky come and nest in its branches" (Matthew, 13:32).

[5] USA, Germany, France, Japan, Russia, Brazil, China, Korea.

[6] This prediction is in clear contradiction to the general expectation of 'a slow emergence out of recession'.

5. Conclusion

To recapitulate, new fractal indices have been proposed for a one-dimensional fractal function $f(t)$: its dimension D_μ and the related index μ. The limiting value of dimension D_μ for $\delta \to 0$ coincides with the usual fractal dimension D. Numerical calculations carried out for stock price series have shown that the application of minimal covers leads to rapid convergence to a power-law asymptotic behavior of the function $V_f(\delta)$ with respect to δ. This is the reason why the representative scale required to determine these parameters with acceptable accuracy contains less data by two orders of magnitude than, for example, the scale determining the Hurst exponent H. This makes it possible to treat the index μ as a local characteristic and to introduce the function $\mu(t)$ which is an indicator of local stability of the time series: the greater μ, the more stable the series. It has been shown, using a very rich empirical data array, that the index of fractality, in essence, defines a natural way of integrating over all possible price trajectories (starting from the shortest). It turns out that trajectories corresponding to random walk have the greatest weight. This fact can well be regarded as a justification for the modern theory of finance. Finally, an early precursor of strong fluctuations in stock markets was built, based on the effect of enhancement of large-scale fluctuations accompanied with suppression of small-scale ones.

Benoit Mandelbrot, who should rightly be considered one of the main predecessors of econophysics, had the notoriety of 'iconoclast, a dynamiter of foundations' and earned complete rejection by some among the economics community. He was one of the most ardent critics of the modern theory of finance based on the concept of general equilibrium from its conception, and was seeking an acceptable alternative to it until the last days of his life. Econophysics tries to propose an alternative to the concept of general equilibrium in a similar manner, but now in the framework of the entire theory of economics. Nevertheless, it was indeed Mandelbrot who was able to work out a system of concepts which allows, as we have shown, not only generating an efficient prognosis but also proposing, after appropriate modifications, what at the moment appears to be the only empirical justification of the *classical* theory of finance.

In conclusion, the authors take this opportunity to express their profound gratitude to V M Polterovich for unflinching support, helpful advice, and valuable comments.

References

1. Mantegna R N, Stanley H E *Nature* **376** 46 (1995)
2. Mandelbrot B *J. Business* **36** 394 (1963)
3. Mandelbrot B B *The Fractal Geometry of Nature* (San Francisco: W.H. Freeman, 1982)
4. Mandelbrot B *Les Objets Fractals* (Paris: Flammarion, 1975) [Translated into English: *Fractals* (San Francisco: W.H. Freeman, 1977)]
5. Hausdorff F *Math. Ann.* **79** 157 (1919)
6. Malinetskii G G, Potapov A B, Podlazov A V *Nelineinaya Dinamika: Podkhody, Rezul'taty, Nadezhdy* (Nonlinear Dynamics. Approaches, Results, Expectations) (Moscow: URSS, 2006)
7. Aleksandrov P S, Pasynkov B A *Vvedenie v Teoriyu Razmernosti* (Introduction to Dimension Theory) (Moscow: Nauka, 1973)
8. Brown J et al., Working Papers in Economics 08/13 (Christchurch, New Zealand: Univ. of Canterbury, Department of Economics and Finance, 2008)
9. Clegg R *Int. J. Simulation: Syst., Sci. Technol.* **7** (2) 3 (2006)
10. Mishura Y *Stochastic Calculus for Fractional Brownian Motion and Related Processes* (New York: Springer, 2007)
11. Brodu N, nlin/0511041
12. Kantz H, Schreiber T *Nonlinear Time Series Analysis* (Cambridge: Cambridge Univ. Press,1997)
13. Abarbanel H D I *Analysis of Observed Chaotic Data* (New York: Springer, 1996)
14. Peters E E *Fractal Market Analysis: Applying Chaos Theory to Investment and Economics* (New York: J. Wiley, 1994) [Translated into Russian (Moscow: Internet-Trading, 2004)]
15. Peters E E *Chaos and Order in the Capital Markets* (New York: Wiley, 1996) [Translated into Russian (Moscow: Mir, 2000)]
16. Dubovikov M M, Starchenko N S, Dubovikov M S *Physica A* **339** 591 (2004)
17. Bachelier L, in *The Random Character of Stock Market Prices* (Ed. P H Cootner) (Cambridge, Mass.: M.I.T. Press, 1964) p. 17
18. Kendall M G *J. R. Statistical Soc.* **96** 11 (1953)
19. Samuelson P A *Industrial Management Rev.* **6** 13 (1965)
20. Shiryaev A N *Osnovy Stokhasticheskoi Finansovoi Matematiki* (Essentials of Stochastic Financial Mathematics) Vol. 1 (Moscow: Fazis, 1998)
21. Mandelbrot B B, Van Ness J W *SIAM Rev.* **10** 422 (1968)
22. Shiryaev A N *Teor. Veroyatn. Primenenie* **39** 5 (1994) [*Theor. Probab. Appl.* **39** (1) 1 (1994)]
23. Shiryaev A N *Obozrenie Priklad. Promyshl. Mat.* **1** 780 (1994)
24. Polterovich V M *Ekonom. Nauka Sovremennoi Rossii* (1) 46 (1998)