Self-similar approach to market analysis

V.I. Yukalov[a]

Bogolubov Laboratory of Theoretical Physics, Joint Institute for Nuclear Research, Dubna 141980, Russia

Abstract. A novel approach to analyzing time series generated by complex systems, such as markets, is presented. The basic idea of the approach is the *Law of Self-Similar Evolution*, according to which any complex system develops self-similarly. There always exist some internal laws governing the evolution of a system, say of a market, so that each of such systems possesses its own character regulating its behaviour. The problem is how to discover these hidden internal laws defining the system character. This problem can be solved by employing the *self-similar approximation theory*, which supplies the mathematical foundation for the law of self-similar evolution. In this report, the theoretical basis of the new approach to analyzing time series is formulated, with an accurate explanation of its principal points.

PACS. 01.75.+m Science and society – 02.30.Lt Sequences, series, and summability – 02.50.-r Probability theory, stochastic processes, and statistics

1 Law of self-similar evolution

There are two sides of the approach I am going to present here: *philosophical* and *mathematical*. The first side gives the general idea of *why* should we think in these or those terms, and the second side is supposed to answer *how* concretely could we realize the idea. It is natural to start with the formulation of the general idea.

The idea to be formulated is general since it may concern any complex system, for instance, a market, a firm, a society, a nation, species, a person, a brain, a heart, and so on. Shortly speaking, a *complex system* is such that cannot, at least at the present time, be completely described by any finite number of given equations. A complex system exists as an empirical organism, with its specific character and behaviour.

To understand the character of a system means to notice some rules governing its behaviour and, therefore, to be able to predict the latter. If the system is complex, such rules cannot be described by fixed equations. But one may always characterize the behaviour in terms of self-similarity, implying that there exist some specific features that, to some extent, are repeated during the evolution of the considered system. Thus, characterizing a person, we name the corresponding features that give us impression of what can be expected from him or her.

The evolution of each complex system is always governed by some laws. This is why there is a similarity in the behavioural facts of this system. Such a behavioural self-similarity is easily noticeable in the life of particular persons as well as in the history of societies, nations, and biological species. This observation suggests that there exists the *law of self-similar evolution*, according to which the main features of a complex system are preserved in the course of its evolution. And if one has grasped these features, he or she should have the feeling of what could be expected in future.

When the behavioural facts, related to the evolution of a complex system, can be associated with some quantities, one gets time series. These can be market prices or indices, measurable signals from a brain or from a heart, some data characterizing societies or species etc. The examples of time series are numerous, their consideration being usually based on statistical analysis and the construction of dynamic regression models [1–5]. These models provide a rather reasonable description of sufficiently stable dynamics with repeated events, like seasonal variations, but they fail in treating such unstable systems as markets. Financial markets are very difficult to predict. Perhaps the most central question in finance is under what circumstances is prediction possible at all?

The widespread opinion is that market prices fluctuate absolutely randomly and, thus, are unforecastable. This point of view is what is called the Efficient Markets Hypothesis, which can be traced back to Samuelson [6]. According to this hypothesis, in an informationally efficient market, properly anticipated prices fluctuate randomly. This includes the assumption that prices fully incorporate the expectations and information of all market participants, that all market traders are identical in the sense that all of them possess the same and the whole information on a market, all traders having the same

---
[a] *Present address*: International Institute of Theoretical and Applied Physics, Iowa State University, Ames, Iowa 50011-3022, USA
e-mail: `yukalov@thsun1.jinr.ru`

physical and mental abilities. Such a hypothesis that all investors are fully rational agents that instantaneously and correctly process all available information is clearly unrealistic. Moreover, people, thank God, are never identical neither in their abilities nor in their wishes. They are not elementary quantum particles, but each of them is a very complex system by its own. There is also increasing empirical evidence that even the most competitive markets are not strictly efficient [7–9].

A word of caution is called for with respect to the meaning of the term "random", which is often confused with "chaotic". These two notions are rather different. A realistic process can be chaotic, at the same time, possessing properties of both regularity and randomness in different proportions, so that it is, in principle, possible to distinguish chaos from randomness, and even control or erase chaos and make predictions [10,11].

Dynamics of all realistic complex systems always exhibit some part of randomness, either due to internal reasons, specific for nonlinear dynamical systems, or caused by external stochastic noise. Therefore randomness is inavoidable to this or that extent and, for some time intervals, it can mask the existence of underlying tendencies and persisting trends. Nevertheless, such trends are to be noticeable on average, whether this concerns the evolution of financial markets or biological species. Persisting features retained during this evolution imply that the complex system evolves self-similarly. It is worth emphasizing that the law of self-similar evolution does not mean and in no sense requires that there should be displayed some rigidly fixed properties of the considered system, but it rather tells that there must exist some trends in evolution. The *evolutional self-similarity* is not a static notion requiring the occurrence of stationary fixed properties but it is a *dynamic concept* stating the persistence of underlying trends.

In this section, the law of self-similar evolution has been described as a philosophical category, in rather general and perhaps vague words. The corresponding mathematical realization will be presented in the following sections. The suggested approach is mathematically based on the self-similar approximation theory whose application to the analysis of asymptotic series is expounded in Section 2. The self-similar extrapolation of asymptotic series can be directly reformulated as a self-similar forecasting for time series, which is explained in Section 3. A method for evaluating the probabilities of self-similar patterns is developed in Section 4. These sections present, for the first time, the complete mathematical foundation of the self-similar approach to analyzing arbitrary time series. Specifications related to market time series are also discussed. Principal points of the approach are summarized in Section 5.

2 Self-similarity in asymptotic series

Assume that we are interested in finding a function $f(x)$ of a real variable x. Let this function be defined by so complicated equations that we are able to extract from them only asymptotic expansions in the vicinity of some

point x_0. Without the loss of generality, the expansion point may be taken as zero, $x_0 = 0$. Let us have several such expansions,

$$f(x) \simeq f_k(x) \qquad (x \to 0) , \qquad (1)$$

enumerated by the index $k = 0, 1, 2, \ldots$ The basic problem is what can be said about the value of $f(x)$ at finite x if all we know are asymptotic expansions $f_k(x)$ in the vicinity of $x \to 0$? This problem is constantly met in physics and applied mathematics, where it is often called the problem of function reconstruction or the problem of summation of asymptotic series [12].

An original general approach, named the *self-similar approximation theory*, for reconstructing functions from a set of their approximate expressions has been developed and successfully applied to various problems [13–23]. The name comes from the basic idea of the approach to present the passage between subsequent approximations as a self-similar transformation. More precisely, it was shown [15–19] that for an approximation sequence $\{f_k(x)\}$ it is possible to construct a cascade, that is a dynamical system with discrete time, whose trajectory is bijective to the sequence $\{f_k\}$, so that the sought function $f(x)$ corresponds to a fixed point of the cascade. If we treat the given complicated equations, together with a calculational algorithm, as a complex system generating a sequence $\{f_k\}$, then the latter is nothing but a prototype of a time series, the approximation number k playing the role of time. Hence the self-similar approximation theory is the mathematical realization of the law of self-similar evolution.

Asymptotic expansions are usually presented as power-law series

$$f_k(x) = \sum_{n=0}^{k} a_n \, x^{\alpha_n} , \qquad (2)$$

where α_n are arbitrary real numbers arranged in ascending order,

$$\alpha_n < \alpha_{n+1} \qquad (n = 0, 1, \ldots , k) . \qquad (3)$$

For the purpose of self-similar analysis, the presentation of the series (2) has to satisfy several general properties:

(i) Analysis should not depend on the choice of units for the variable. Hence the latter is to be taken in a dimensionless form.

(ii) Asymptotic expansions are to be reduced to a scale-invariant form. To this end, it is always possible to factor out the term

$$f_0(x) = a_0 \, x^{\alpha_0} \qquad (a_0 \neq 0) \qquad (4)$$

and to introduce the scale-invariant function

$$\varphi_k(x) \equiv \frac{f_k(x)}{f_0(x)} , \qquad (5)$$

where x is assumed to be dimensionless. Evidently, function (5) does not depend on the change of scales for f_k.

(iii) If the variable x pertains to a finite interval, the latter is to be normalized to the unitary interval. So that everywhere in what follows, it is assumed that $x \in [0,1]$.

Note that if the variable pertains to an infinite interval, one may pursue different ways depending on what additional information on the behaviour of $f(x)$ at $x \to \infty$ is available. If there is no such information, it is possible to transform the infinite interval to the unitary one by means of the change of variables $x' = x/(1+x)$, $x = x'/(1-x')$, so that $x' \in [0,1]$.

For the scale-invariant function (5), we have

$$\varphi_k(x) = \sum_{n=0}^{k} b_n \, x^{\beta_n} \, , \qquad (6)$$

where $x \in [0,1]$ and

$$b_n \equiv \frac{a_n}{a_0} \, , \qquad \beta_n \equiv \alpha_n - \alpha_0 \geq \beta_0 = 0 \, . \qquad (7)$$

Clearly, $\varphi_0(x) = 1$. Also, all powers β_n are positive, even if some α_n are negative, which follows from the ascending order (3). Recall that, by construction, the series (6) are assumed to be asymptotic, having sense only for $x \to 0$, while for any finite x sequence $\{\varphi_k(x)\}_{k=0}^{\infty}$ diverges. It is possible to say that the latter sequence converges just at one point $x = 0$.

To proceed further, we have to transform divergent series to a form that would have sense for finite $x \in [0,1)$. This can be done with the help of control functions [13], which can be introduced in different ways [13–23]. Dealing with asymptotic series, it is convenient to invoke the multiplicative power-law transformation [20–22] defined as

$$\Phi_k(x,s) \equiv x^s \, \varphi_k(x) \, , \qquad (8)$$

with the inverse transformation

$$\varphi_k(x) = x^{-s} \, \Phi_k(x,s) \, . \qquad (9)$$

Since power laws are common in describing fractal objects, equation (8) may be called the *fractal transformation*. The transform (8), according to equation (6), is

$$\Phi_k(x,s) = \sum_{n=0}^{k} b_n \, x^{s+\beta_n} \, . \qquad (10)$$

Here $s = s(x)$ is a control function, whose role is to make the series (10) meaningful for finite x. These series can be considered as an expansion in powers of the new variable x^s. As is evident, such series are asymptotic with respect to $x^s \to 0$. The latter limit can be achieved if, instead of forcing x to zero, we keep $|x| < 1$ and setting $s \to \infty$. Thus, the series (10) can be treated as asymptotic with respect to $s \to \infty$ for all $|x| < 1$. Now we may say that the sequence $\{\Phi_k(x,s)\}_{k=0}^{\infty}$ converges for all $|x| < 1$, provided that $s \to \infty$. In this way, we come to the natural choice of the control function $s \to \infty$. Recall that we are considering the case when no additional constraints are imposed on the behaviour of the sought function and all we know

are its asymptotic expansions (1). In the intermediate expressions, the value of s is assumed to be asymptotically large, and the actual limit $s \to \infty$ is to be taken after the inverse transformation (9).

Since our consideration here concerns functions, we need to define the property of *functional self-similarity*, which should not be confused with geometric self-similarity describing fractals. The notion of geometric self-similarity [24] is connected with the scaling of a variable, which is only a particular kind of the more general notion of the functional or *group self-similarity* [15–19]. To correctly define the latter, we need to introduce some notation. We define the expansion function $x(\varphi, s)$ by the equation

$$\Phi_0(x,s) = \varphi \, , \qquad x = x(\varphi, s) \, . \qquad (11)$$

With the form (10), this gives $x(\varphi, s) = \varphi^{1/s}$ and $\varphi = x^s$. Introduce the mapping

$$y_k(\varphi, s) \equiv \Phi_k(x(\varphi, s), s) \, . \qquad (12)$$

Let $\Phi_k(x,s)$ be real for all $k = 0, 1, 2, \ldots$ and all $x \in [0,1]$. From $\Phi_k(x,s) \in \mathbf{R}$ it follows that $y_k(\varphi, s) \in \mathbf{R}$ for all $k = 0, 1, 2, \ldots$ and $\varphi \in \mathbf{R}$. Therefore the mapping (12) is an endomorphism on $\mathbf{R}$. It is this endomorphism that allows us to formulate the property of group self-similarity we need.

Our aim is to present the change of the endomorphism (12), when varying the approximation number k, as the evolution of y_k with respect to the discrete time k. From the point of view of group theory, self-similarity is nothing but a semigroup property. The latter, for the evolution of y_k with respect to k, reads $y_{k+p} = y_k \, y_p$. As follows from the definition (11), the unit element is y_0, since $y_0(\varphi, s) = \varphi$. The family of endomorphisms, $\{y_k| \, k = 0, 1, 2, \ldots\}$, with the semigroup property forms a dynamical system in discrete time, called the cascade. The semigroup property, in terms of the notation (12), takes the form

$$y_{k+p}(\varphi, s) = y_k(y_p(\varphi, s), s) \, . \qquad (13)$$

Since in the accepted interpretation, we treat the sequence $\{y_k(\varphi, s)\}$ as a trajectory resulting from the evolution of the cascade $\{y_k| \, k = 0, 1, 2, \ldots\}$, the relation (13) may be called the *evolutional self-similarity*. As far as the corresponding semigroup property is natural for dynamical systems, equation (13) may also be termed *dynamic self-similarity*. This equation (13) is a necessary condition for the fastest-convergence criterion [16,17]; the cascade fixed point representing the sought function.

Following the general theory [14–19], the cascade $\{y_k| \, k = 0, 1, 2, \ldots\}$ can be embedded into a flow $\{y_\tau| \, \tau \geq 0\}$, which is a dynamical system in continuous time. For the latter, one may write the Lie equation which is a differential equation of motion. The flow velocity field is defined, by means of the Euler discretization, as the cascade velocity

$$v_n(\varphi, s) \equiv y_n(\varphi, s) - y_{n-1}(\varphi, s) \qquad (n = 1, 2, \ldots, k) \, . \qquad (14)$$

From equations (10–12), it follows that

$$y_k(\varphi, s) = \sum_{n=0}^{k} b_n \, \varphi^{1+\beta_n/s} \, , \qquad (15)$$

which results in

$$v_n(\varphi, s) = b_n \, \varphi^{1+\beta_n/s} \, . \qquad (16)$$

The differential equation of motion can be integrated. The integration over the effective time goes from $\tau = n$ to $\tau = n + \tau_n$, with τ_n being the effective time required for reaching a fixed point after the n-th step. In this way,

$$\int_{n}^{n+\tau_n} \mathrm{d}t = \tau_n \, . \qquad (17)$$

Therefore, the *evolution integral* [15–19] acquires the form

$$\int_{y_{n-1}}^{y_n^*} \frac{\mathrm{d}\varphi}{v_n(\varphi, s)} = \tau_n \, . \qquad (18)$$

This, with the cascade velocity (16), yields

$$\Phi_n^* = \left(\Phi_{n-1}^{-\beta_n/s} - \frac{\beta_n}{s} \, b_n \tau_n \right)^{-s/\beta_n} \, ,$$

where $\Phi_n^* = \Phi_n^*(x, s) \equiv y_n^*(x^s, s)$ and $\Phi_n = \Phi_n(x, s)$. In particular,

$$\Phi_1^*(x, s) = x^s \left(1 - \frac{\beta_1}{s} \, b_1 \tau_1 \, x^{\beta_1} \right)^{-s/\beta_1} \, .$$

Returning back to the function $\varphi_k^*(x)$ by means of the inverse transformation (9), we have to take the limit $s \to \infty$. For example,

$$\varphi_1^*(x) \equiv \lim_{s \to \infty} x^{-s} \, \Phi_1^*(x, s) \, ,$$

which results in the first-order self-similar approximation

$$\varphi_1^*(x) = \exp \left(b_1 \tau_1 \, x^{\beta_1} \right) \, .$$

To obtain higher approximations, we can accomplish such a renormalization procedure $2k$ times for each $\Phi_k^*(x, s)$, which was called [21,22] *self-similar bootstrap*. However the same result can be reached twice faster by accomplishing k renormalizations, in the following way. We may present the function (6) as

$$\varphi_k(x) =$$
$$1 + b_1 \, x^{\beta_1} \left(1 + \frac{b_2}{b_1} \, x^{\beta_2 - \beta_1} \left(1 + \frac{b_3}{b_2} \, x^{\beta_3 - \beta_2}(1 + ...) \right) ... \right) .$$

The latter can be written in the form $\varphi_k(x) = 1 + x_1$, in which x_1 is expressed through x_2, and x_2 through x_3, and so on according to the rule

$$x_n = \frac{b_n}{b_{n-1}} \, x^{\beta_n - \beta_{n-1}} (1 + x_{n+1}) \, ,$$

where $n = 1, 2, \ldots, k$. Considering each x_n as a small parameter, we need to accomplish k times the first-order renormalization procedure described above. As a result, introducing the notation

$$c_n \equiv \frac{a_n}{a_{n-1}} \, \tau_n \, , \qquad \nu_n \equiv \alpha_n - \alpha_{n-1} \qquad (n = 1, 2, \ldots, k) \, , \qquad (19)$$

we come to the k-order *self-similar exponential approximation*

$$\varphi_k^*(x) = \exp \left(c_1 x^{\nu_1} \, \exp \left(c_2 x^{\nu_2} \ldots \exp \left(c_k x^{\nu_k} \right) \right) \ldots \right) \, . \qquad (20)$$

This, for short, can also be named the k-order *superexponential*.

It is worth emphasizing that although the form (20) reminds us of the Euler nested exponentials [25,26], it is principally different from the latter. Firstly, the Euler superexponentials are defined only for integer powers α_n. Secondly, when one tries to sum power series by means of such continued exponentials, one fits the coefficients in the latter in order to reproduce those in the power series; as a result the constructed superexponentials have the same radius of convergence as the related power series on the real axis [27,28].

In our case, the self-similar exponential (20) contains the coefficients c_n given by equation (19), where the effective renormalization time τ_n is yet undefined. The latter plays the role of a control function that is to be determined from additional conditions. One method could be to find τ_n from the fixed-point conditions [22] having the form of the minimal-difference criterion [13]. However, such fixed-point equations do not always possess solutions. Here we suggest another, and more general way of defining the control functions τ_n. This method is commonly employed for defining control functions in the optimal control theory. The idea is to construct a cost functional whose minimization yields the control functions of interest.

The quantity τ_n appears in the evolution integral (18), where it has the meaning of an effective time required for reaching, after the n-th step, a fixed point representing the sought function. One would clearly like to reach the answer as fast as possible. The minimal number of renormalization steps is obviously one. Therefore one would like that the total effective time $n\tau_n$ be also close to one. Hence, one should look for τ_n being close to $1/n$. At the same time, how fast one reaches the fixed point depends on the distance of the latter from the starting point. The distance that is passed during the time τ_n, with a velocity v_n, can be evaluated as $v_n \tau_n$. Thus, we need to find a minimal time τ_n, being close to $1/n$, corresponding to the minimal distance $v_n \tau_n$. This suggests us to construct the *time-distance cost functional*

$$F = \frac{1}{2} \sum_n \left[\left(\tau_n - \frac{1}{n} \right)^2 + (v_n \tau_n)^2 \right] \, , \qquad (21)$$

whose minimization with respect to the control function τ_n yields

$$\tau_n = \frac{1}{n(1 + v_n^2)} \ . \qquad (22)$$

The velocity $v_n = v_n(x)$ is to be understood as the image, in the domain of x, of the cascade velocity (16), while taking account of the inverse transformation (9), which gives $v_n(x) \equiv x^{-s} v_n(x^s, s)$. From here, one has

$$v_n(x) = \varphi_n(x) - \varphi_{n-1}(x) = b_n \, x^{\beta_n} \ , \qquad (23)$$

which defines the control function $\tau_n = \tau_n(x)$ according to the expression (22).

In this way, we find the *controllers*

$$c_n(x) = \frac{a_n}{a_{n-1}} \, \tau_n(x) \ , \qquad \nu_n = \alpha_n - \alpha_{n-1} \ ,$$

$$\tau_n(x) = \frac{1}{n[1 + v_n^2(x)]} \ , \qquad v_n(x) = \frac{a_n}{a_0} \, x^{\alpha_n - \alpha_0} \ . \qquad (24)$$

Combining equations (5) and (20), we obtain the k-order self-similar exponential approximation

$$f_k^*(x) = f_0(x) \exp\left(c_1 x^{\nu_1} \, \exp\left(c_2 x^{\nu_2} \ldots \exp\left(c_k x^{\nu_k}\right)\right) \ldots\right) \ , \qquad (25)$$

in which $c_n = c_n(x)$, and $f_0(x)$ is given by formula (4). The approximant (25) extrapolates the asymptotic series (2), valid only for $x \to 0$, to the region of finite $x \in [0, 1)$. The value of $f(x)$ at the point $x = 1$ can be defined as the limit from the left, as $x \to 1 - 0$. Therefore the superexponential (25) extrapolates the sought function $f(x)$ from asymptotically small $x \to 0$ to the whole unitary interval $0 \leq x \leq 1$.

It is worth recalling that in the asymptotic series (2) the powers α_n were assumed to be arbitrary real numbers, with the sole requirement that they are arranged in ascending order (3). Some, or even all, of these powers could be negative. If so, the initial term (4), with $\alpha_0 < 0$, has the power-law form that has been so much discussed in literature. The asymptotic existence of power laws is well known in critical phenomena. The relevance of power laws has repeatedly been claimed to describe many natural phenomena, ranging from earthquakes [24,29,30] to different economic and financial distributions [31,32]. Since such power laws are practically always asymptotic, their more general form should include corrections leading to the power-law series (2). Extrapolating these asymptotic series in the described self-similar way, one should come to the self-similar exponentials. The first-order approximation then results in a kind of a stretched exponential. The stretched exponential distributions describe many phenomena in nature and economy either not worse or even better than power-law distribution functions [33]. More generally, the extrapolation of power laws should lead to the self-similar nested exponentials (25), whose structure evidently demonstrates the existence of many scales.

3 Self-similarity in time series

The technique of self-similar extrapolation for asymptotic series can be reformulated as the method of forecasting for time series. It is necessary to call attention to the existence of several principal points in this reformulation. Overlooking these points would essentially restrict the applicability of the method. The correct general way of self-similar forecasting for time series is listed below.

First of all, the same basic requirements that were imposed on asymptotic series are compulsory for time series:

(i) The measured quantity is to be presented in a scale-invariant form. This can be easily done by normalizing the given data to the value f_0 of the measured quantity at the initial time, which is analogous to introducing the scale-invariant function (5).

(ii) The time variable has to be normalized to a dimensionless form, such that the prediction time would pertain to the unitary interval $[0, 1]$. This normalization eliminates the ambiguity in defining the power of the power-law transformation (8), requiring that $s \to 0$, which results in self-similar superexponentials (20).

In addition to these requirements, common both for asymptotic as well as for time series, there arises a question that specific for time series: how should the latter be presented, as a backward or forward recursion? The answer to this question follows from the comparison of asymptotic series with time series. In order that the extrapolation of asymptotic series could be directly extended to forecasting for time series, the prediction horizon is to be a unitary interval and the available information from the past has to provide approximate forecasts for asymptotically small time $t \to +0$. This requirement can be strictly accomplished only for backward recursion. Hence, we have one more restriction, specific for time series:

(iii) For the correct usage of the self-similar analysis, time series are to be arranged as a *backward recursion*. This means that, if we are given a set $\{f_n\}$ of data f_n corresponding to a quantity of interest, measured at the times t_n, where $n = 0, 1, 2, \ldots, k$, the moments of time are to be ordered so that

$$t_{n+1} < t_n \qquad (n = 0, 1, 2, \ldots, k) \ , \qquad (26)$$

with the initial time $t_0 = 0$. The past-history *data base* of k-th order is the set

$$\mathbf{D}_k = \{f_k, f_{k-1}, \ldots, f_0 | \, t_k < t_{k-1} < \ldots < 0\} \ . \qquad (27)$$

To start the procedure of self-similar forecasting for future times $t \in [0, 1]$, we need to possess a sequence of approximations valid for asymptotically small $t \to +0$. The role of such asymptotic forecasts can naturally be played by functions $f_k(t)$ interpolating the given data base (27) for the past time horizon $t_k \leq t \leq 0$, so that

$$f_k(t_n) = f_n \qquad (n = 0, 1, \ldots, k) \ . \qquad (28)$$

This interpolation can be uniquely defined by the Lagrange interpolation formula [34] presenting the interpolation function

$$f_k(t) = \sum_{n=0}^{k} f_n \, l_n^k(t) \qquad (k \geq 1) \tag{29}$$

as a series over the Lagrange polynomials

$$l_n^k(t) \equiv \prod_{m(\neq n)}^{k} \frac{t - t_m}{t_n - t_m} \qquad (n \leq k) \, . \tag{30}$$

Because of the property $l_n^k(t_m) = \delta_{mn}$ of the Lagrange polynomials, condition (28) is automatically satisfied. Note that the interpolation form (29) is evidently scale invariant with respect to time and the ratio $f_k(t)/f_0$ is scale invariant with respect to units of f_k. Also, one may remark that the time moments t_n are not necessarily equidistant, but can be chosen arbitrarily.

The interpolation formula (29) can be rewritten as the algebraic polynomial

$$f_k(t) = \sum_{n=0}^{k} a_n \, t^n \qquad (a_0 = f_0) \, , \tag{31}$$

with the coefficients $a_n \equiv a_{nk}$ immediately following from equations (29) and (30). The values of these coefficients depend, of course, on the given data base (27), but for the simplicity of notation, we shall omit in what follows the additional index k. Consider the subsequence $\{f_n(t)\}_{n=0}^{k}$ of the terms

$$f_n(t) \equiv \sum_{m=0}^{n} a_m \, t^m \, , \tag{32}$$

which, for a given k, tends to the polynomial (31) as $n \to k$ and for which

$$f_n(0) = a_0 = f_0 \qquad (n = 0, 1, \ldots, k) \, . \tag{33}$$

Then the subsequence $\{f_n(t)\}_{n=0}^{k}$ can be treated as a sequence of approximations asymptotically valid for $t \to +0$. Employing the self-similar extrapolation, described in the previous section, we may construct a forecast for the time interval $0 \leq t \leq 1$. The result is given by formulae (24) and (25), where we need to change the variable x by the time t and to put $\alpha_n = n$, $\nu_n = 1$. Instead of equations (24), we have the *controllers*

$$c_n(t) = \frac{a_n}{a_{n-1}} \, \tau_n(t) \qquad (n = 1, 2, \ldots, k)$$

$$\tau_n = \frac{1}{n[1 + v_n^2(t)]} \, , \qquad v_n(t) = \frac{a_n}{f_0} \, t^n \, . \tag{34}$$

And the self-similar exponential (25) takes the form

$$f_k^*(t) = f_0 \exp\left(c_1 t \, \exp\left(c_2 t \ldots \exp\left(c_k t\right)\right)\ldots\right) \, , \tag{35}$$

in which $c_n = c_n(t)$. Expression (35) is the *self-similar forecast* for the future time interval $0 \leq t \leq 1$, predicted on the grounds of the past-history data base (27).

4 Ensemble of possible scenarios

For each given data base (27), one may construct the self-similar forecast (35). But one can take several different data bases by varying either the data-base order k, or by changing the time intervals between the moments of time t_n, where $n = 0, 1, \ldots, k$, or by varying both the data-base order as well as the data-base scale. So that, in general, one can consider an ensemble of different data bases. Each of the latter has to be labelled by two indices,

$$\mathbf{D}_k(j) = \left\{ f_k^{(j)}, f_{k-1}^{(j)}, \ldots, f_0 \big| \, t_k^{(j)} < t_{k-1}^{(j)} < \ldots < 0 \right\} \, , \tag{36}$$

one index, k, defining the data base order, and another, j, specifying the chosen time scale of the past. For each data base (36), one obtains a forecast $f_k^*(j, t)$ according to the rule (35), but with different values of $c_n = c_{nk}$. Hence there exists an ensemble $\{f_k^*(j, t)\}$ of possible forecasts, or admissible scenarios. Which of these possible forecasts should one trust?

Life teaches us that in the majority of cases nothing can be trusted a hundred percent. But, when there can happen to be several different events, they can be classified by estimating their probabilities. Hence we need to define a probability measure on the ensemble of scenarios $\{f_k^*(j, t)\}$.

This problem is analogous to the problem of pattern selection occurring for nonlinear differential equations in partial derivatives. Such equations sometimes possess a set of solutions corresponding to different spatio-temporal structures, or patterns [35]. A general approach for treating the problem of pattern selection has been suggested [36]. This approach can be directly applied for weighting possible scenarios from the given ensemble of self-similar forecasts. For this purpose, we may consider the passage from $f_k^*(j, t)$ to $f_{k+1}^*(j, t)$ as the motion with respect to k. Then the map $\{f_k^*(j, t)| \, k = 1, 2 \ldots\}$ can be treated as the image of a dynamical system with discrete time k. The probability of a scenario $f_k^*(j, t)$ can be defined as

$$p_k(j, t) = \frac{1}{Z_k(t)} \, \exp\left\{-\Delta S_k(j, t)\right\} \, , \tag{37}$$

where $Z_k(t)$ is a normalization factor, being the sum

$$Z_k(t) \equiv \sum_j \exp\left\{-\Delta S_k(j, t)\right\}$$

over the pattern indices j, and the entropy variation

$$\Delta S_k(j, t) \equiv S_k(j, t) - S_1(j, t) \tag{38}$$

shows the change of entropy with respect to the effective time k. The entropy of a dynamical system may be defined, by analogy with statistical systems, as the logarithm of an elementary phase volume [36], the latter, in our case, being $|\delta f_k^*(j, t)|$. Thus the dynamical entropy is

$$S_k(j, t) \equiv \ln |\delta f_k^*(j, t)| \, . \tag{39}$$

Then the entropy variation (38) becomes

$$\Delta S_k(j,t) = \ln \left| \frac{\delta f_k^*(j,t)}{\delta f_1^*(j,t)} \right| . \tag{40}$$

With the notation for the mapping multiplier

$$m_k(j,t) \equiv \frac{\delta f_k^*(j,t)}{\delta f_1^*(j,t)} = \frac{\partial f_k^*(j,t)/\partial t}{\partial f_1^*(j,t)/\partial t} , \tag{41}$$

the entropy variation (40) reduces to

$$\Delta S_k(j,t) = \ln |m_k(j,t)| . \tag{42}$$

It is convenient to introduce the average multiplier $\overline{m}_k(t)$ by the relation

$$\frac{1}{|\overline{m}_k(t)|} \equiv \sum_j \frac{1}{|m_k(j,t)|} . \tag{43}$$

Using equations (42) and (43), for the scenario probability (37), we have

$$p_k(j,t) = \left| \frac{\overline{m}_k(t)}{m_k(j,t)} \right| . \tag{44}$$

Note that the *scenario probability* (44) is normalized with respect to the summation over the pattern indices j corresponding to different data-base scales. In a particular case of just one fixed scale, one could accomplish the normalization with respect to the summation over k, which would define the probability weights for a restricted data-base ensemble, as was postulated in reference [37]. The derived scenario probability (44) concerns the general case of an arbitrary ensemble $\{\mathbf{D}_k(j)\}$ of the data bases (36). Being general, the approach of the present paper makes it possible to answer several principal questions.

One important question concerns the choice of the data-base order. More in detail, the problem is as follows. For a fixed time scale, labelled by j, we may analyze different data bases $\mathbf{D}_k(j)$, with varying $k = 1, 2, 3, \ldots$. Then, how many terms $f_k^{(j)}$ should we take? That is, when should we stop increasing k? The answer is straightforward. The data-base order k has to be increased till we reach numerical convergence. More precisely, this means the following. Let us be satisfied by the results of an error ε. Then we need to increase k up to the *saturation number* $N_j = N_j(\varepsilon)$ such that

$$|f_{k+n}^*(j,t) - f_k^*(j,t)| < \varepsilon \tag{45}$$

for $k \geq N_j$, all $n \geq 0$, and $t \in [0,1]$. For this $k = N_j$, we have the *saturated data base* $\mathbf{D}(j) \equiv \mathbf{D}_{N_j}(j)$ and the related *saturated forecast*

$$f^*(j,t) \equiv f_{N_j}^*(j,t) \tag{46}$$

characterized by the *saturated scenario probability*

$$p(j,t) \equiv p_{N_j}(j,t) . \tag{47}$$

Thus, varying the data-base scales, labelled by j, we get the *saturated ensemble* $\{f^*(j,t)\}$ of scenarios (46), with the probability measure (47).

The *most probable scenario* from the ensemble $\{f^*(j,t)\}$ is the forecast $f^*(j_0,t)$ having the largest probability, such that

$$\max_j p(j,t) = p(j_0,t) . \tag{48}$$

The latter, because of the form (44), is equivalent to the condition of the minimum for the absolute value of the multiplier $m(j,t) \equiv m_{N_j}(j,t)$, so that

$$\min_j |m(j,t)| = |m(j_0,t)| . \tag{49}$$

Having the probability measure (47), it is possible to define the average forecast, or the *expected forecast*

$$\langle f(t) \rangle = \sum_j p(j,t) \, f^*(j,t) . \tag{50}$$

The dispersion

$$\sigma^2(t) \equiv \langle f^2(t) \rangle - \langle f(t) \rangle^2 ,$$

in the case of a market, describes the market volatility. The latter can also be characterized by the variance coefficient $\kappa(t) \equiv \sigma(t)/\langle f(t) \rangle$.

The last question to be answered in order to have a completely self-consistent theory is how to choose the data-base time scale. In general, the moments of time $t_n^{(j)}$ should not be compulsory equidistant. However, for practical purpose, it looks more convenient to take them as such, defining the time step as $\Delta_j \equiv t_n^{(j)} - t_{n+1}^{(j)}$. It seems natural to start with the time scale $\Delta_0 = 1$, which equals the prediction horizon. Then, one may decrease as well as increase the time step, for instance, according to the rule $\Delta_{2j} = 2^{-j}$, $\Delta_{2j+1} = 2^j$, with $j = 0, 1, 2, \ldots$ One has to stop decreasing and increasing the data-base time scale at such $j = j_{\max}$, when numerical convergence is reached. This implies that, for a given error ε, one gets the inequality

$$|f^*(j+m,t) - f^*(j,t)| < \varepsilon$$

for $j \geq j_{\max}$, all $m \geq 0$, and $t \in [0,1]$.

5 Summary of main ideas

In this paper, a novel approach to analyzing time series has been presented. There are several principal points distinguishing this approach from the standard one. The aim of this paper has been to clearly describe these principal points forming the body of a self-consistent theory. Not yet all parts of this theory have been exploited in full for practical applications; the calculational work is in process. But some simplified versions of the approach have been illustrated by a number of examples for market time

series [37,38]. In this concluding section, I would like to emphasize again the main ideas the approach is based on.

To better stress the principal difference of the present approach from the standard way of analyzing time series, let us recall the basic idea of the latter: For a given set of data $\{f_n|\ n = 0, 1, 2, \dots, k\}$ one tries to invent a relation $f_t = f(f_0, f_1, \dots, f_k, t, \xi)$ connecting the value f_t at the moment of time t with the past data. This relation can be in the form of an explicit function or in the form of a difference or differential equation, including a stochastic term ξ modelling noise [1–5]. There are, to my mind, two principal deficiencies of such an approach. First, I think that no explicit equations, no matter how elaborated they are, can grasp all peculiarities of a realistic complex system, such as a market. Second, any given relations reflect only the past history, providing an *interpolation* for the learning historical period, while for predicting future one needs an *extrapolation*.

The self-similar approach is based not on attempts to invent a relation between the historical points but it tries to discover *dynamic trends* resulting in these points. For this purpose, instead of studying relations between points, it is necessary to analyze *relations between sequences*. Interpolative formulae are used here only as a starting step. The comparison of different interpolative expressions makes it possible to produce an extrapolation that is forecasting.

The basic philosophical idea of the self-similar analysis of complex systems is the *law of self-similar evolution*, formulated in Section 1. And the mathematical foundation is provided by the self-similar extrapolation of asymptotic series, described in Section 2. This extrapolation can be reformulated as forecasting for time series, as is done in Section 3. Since any forecasting can only be probabilistic, the way of defining the related probability measure is explained in Section 4. Some details of the self-similar analysis can be changed. For example, the number of points in a given data base can be reduced by replacing several neighbouring values f_n by either arithmetic averages or by fitting the values on a large time interval with the help of simple splines [34]. However, in the process of these kinds of averaging, some information on the considered time series will be inavoidably lost.

Another change could concern the definition of the effective control time τ_n. The latter is to be defined from the minimization of a cost functional. As an example, the time-distance cost functional (21) was considered. But, in general, one could opt for another cost functional, depending on the available information and imposed constraints.

Also, it would be possible to deal not with the initially given data base but with some transforms of it. For instance, one could consider the set $\{\ln f_n\}$ instead of $\{f_n\}$. Or one could keep in mind a more elaborated transform, like the wavelet transforms, often employed for analyzing time series [39].

It is worth stressing once more that the notion of self-similarity, exploited throughout the paper, is understood here as the *group self-similarity*, which is a more general notion than the trivial geometric self-similarity one usually talks about in connection with fractals. In the latter case one assumes the existence of the scaling relation $f(\lambda x) = \lambda^\alpha f(x)$ for the considered function $f(x)$. Such a relation, with a given boundary, or initial, condition $f(x_0) = f_0$, immediately results in the power-law function $f(x) = f_0(x/x_0)^\alpha$. One could consider a slightly more complicated scaling relation as $f(\lambda x) = u(\lambda)f(x)$, with a known function $u(\lambda)$. However again, with the given boundary condition, this immediately gives the answer $f(x) = f_0/u(x_0/x)$. All such scaling relations produce the considered function in an explicit form. Whereas the group self-similarity (13) provides an equation that is yet to be solved.

An important point is that the group self-similarity employed here, and which is the basis of the *self-similar approximation theory* [13–23], has to do not with a scaling of a variable but with the motion with respect to the effective time whose role is played by the approximation number. Here it is the *motion on the manifold of approximants*.

Finally, if we may digress from mathematical foundations, in order to conclude why group self-similarity does work for extrapolating asymptotic series and forecasting time series, we should return back to the *law of self-similar evolution*; this tells us that all complex systems develop self-similarly, preserving their generic features in the course of their evolution. Such features may be not noticeable at first and hidden in their genes. The group self-similarity is a kind of *genetic self-similarity*. Remember also that the Lord created the man in a self-similar way [40].

References

1. D.C. Montgomery, L.A. Johnson, *Forecasting and Time Series Analysis* (McGraw-Hill, New York, 1976).
2. C.W. Granger, P. Newbold, *Forecasting Economic Time Series* (Academic, Orlando, 1986).
3. P.J. Brockwell, R.A. Davis, *Time Series: Theory and Methods* (Springer, New York, 1991).
4. R.C. Pindyck, D.L. Rubinfeld, *Econometric Models and Economic Forecasts* (McGraw-Hill, New York, 1991).
5. A. Pankratz, *Forecasting with Dynamic Regression Models* (Wiley, New York, 1994).
6. P.S. Samuelson, *Collective Scientific Papers* (MIT, Cambridge, 1972).
7. J.D. Farmer, *Market Force, Ecology and Evolution*, Santa Fe Inst., preprint 98–12–117 (1998).
8. J.D. Farmer, A.W. Lo, *Frontiers of Finance: Evolution and Efficient Markets*, Santa Fe Inst., preprint 99–06–039 (1999).
9. A. Lo, C. MacKinlay, *A Non-Random Walk Down Wall Street* (Princeton University, Princeton, 1999).
10. G.M. Zaslavsky, Chaos **4**, 25 (1994).
11. G.M. Zaslavsky, Phys. Today N **8**, 39 (1999).

12. J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena* (Clarendon, Oxford, 1996).
13. V.I. Yukalov, Mosc. Univ. Phys. Bull. **31**, 10 (1976).
14. V.I. Yukalov, Int. J. Mod. Phys. B **3**, 1691 (1989).
15. V.I. Yukalov, Physica A **167**, 833 (1990).
16. V.I. Yukalov, J. Math. Phys. **32**, 1235 (1991).
17. V.I. Yukalov, J. Math. Phys. **33**, 3994 (1992).
18. V.I. Yukalov, E.P. Yukalova, Physica A **206**, 533 (1994).
19. V.I. Yukalov, E.P. Yukalova, Physica A **225**, 336 (1996).
20. S. Gluzman, V.I. Yukalov, Phys. Rev. E **55**, 3983 (1997).
21. V.I. Yukalov, S. Gluzman, Phys. Rev. E **55**, 6552 (1997).
22. V.I. Yukalov, S. Gluzman, Phys. Rev. E **58**, 1359 (1998).
23. V.I. Yukalov, E.P. Yukalova, Ann. Phys. **227**, 219 (1999).
24. B.B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, New York, 1983).
25. L. Euler, Acta Acad. Petropolitanae **1**, 38 (1777).
26. R.A. Knobel, Am. Math. Monthly **88**, 235 (1981).
27. C.M. Bender, J.P. Vinson, J. Math. Phys. **37**, 4103 (1996).
28. D. Poland, Physica A **250**, 394 (1998).
29. P. Bak, C. Tang, J. Geophys. Res. **94**, 15635 (1989).
30. J.R. Grasso, D. Sornette, J. Geophys. Res. **103**, 29965 (1998).
31. V. Pareto, *Manuel d'Économie Politique* (Giard, Paris, 1927).
32. B.B. Mandelbrot, *Fractals and Scaling in Finance* (Springer, New York, 1997).
33. J. Laherrère, D. Sornette, Eur. Phys. J. B **2**, 529 (1998).
34. G.G. Lorentz, *Approximation of Functions* (Holt, New York, 1966).
35. M.C. Gross, P.C. Hohenberg, Rev. Mod. Phys. **65**, 851 (1993).
36. V.I. Yukalov, E.P. Yukalova, Phys. Part. Nucl. **31**, 1128 (2000).
37. V.I. Yukalov, S. Gluzman, Int. J. Mod. Phys. B **13**, 1463 (1999).
38. J.V. Andersen, S. Gluzman, D. Sornette, Eur. Phys. J. B **14**, 579 (2000).
39. D.B. Percival, A.T. Walden, *Wavelet Methods for Time Series Analysis* (Cambridge University, Cambridge, 2000).
40. *Holy Bible, Old Testament, Pentateuch, Genesis.*