Lithuanian Journal of Physics, 2001, **41**, No. 4-6, 551-555.

Modelling share volume traded in financial markets

V. Gontis

Institute of Theoretical Physics and Astronomy,
A. Goštauto 12, 2600 Vilnius, Lithuania
E-mail: gontis@ktl.mii.lt

Abstract. A simple analytically solvable model exhibiting a $1/f$ spectrum in an arbitrarily wide frequency range was recently proposed by Kaulakys and Meškauskas (KM). Signals consisting of a sequence of pulses show that inherent origin of the $1/f$ noise is Brownian fluctuations of the average intervent time between subsequent pulses of the pulse sequence. We generalize the KM model to reproduce the variety of self-affine time series exhibiting power spectral density $S(f)$ scaled as power of their frequency f. Numerical calculations with the generalized discrete model (GDM) reproduce power spectral density $S(f)$ scaled as power of frequency $1/f^{\beta}$ for various values of β, including $\beta = 1/2$ for applications in financial markets. The particular applications of the model proposed are related with financial time series of share volume traded.

1 Introduction

Physicists have recently begun doing research in finance with wide application of models earlier introduce in physics [1, 2] . The statistical properties of financial time series are attracting experts of statistical physics, fluctuations analysis, dynamical chaos and others. Papers on finance are appearing with some frequency in physics journals. A new movement called *econophysics* has been established [2]. On the other hand, several empirical studies have determined scale-invariant behavior of both the long range correlations of price volatility and share volume traded in financial markets [3, 4, 5, 6]. Mandelbrot introduced the concept of fractals in terms of statistical self similarity [7] and using the context of self-affine time series extended the concept to time series [8]. The basic definition of self-affine time series is that the power spectral density of the time series has a power-law dependence on frequency $S(f) = f^{-\beta}$. Universality of $1/f$ noise, when $\beta = 1$, has led to speculations that there might exist some generic mechanism underlying production of so general statistical properties. Such concepts as fractional Brownian motion provide a formal procedure how to produce self-affine time series, but can't serve as generic mechanism. We will generalize the simple model introduced by Kaulakys and Meškauskas (KM) [9] to generate time series in the range $0 \leq \beta \leq 2$ with particular interest in financial time series.

Long range correlations in time series $I(t)$ are quantified by autocovariance (autocorrelation) function $C(s)$:

$$C(s) = C(-s) = \left\langle \frac{1}{T} \int_0^{T-s} I(t) I(t+s) dt \right\rangle, \quad (1)$$

with Wiwner-Khintchine relation to power spectral density $S(f)$ defined as:

$$S(f) = \lim_{T \to \infty} \left\langle \frac{2}{T} \left| \int_0^T I(t) e^{-i2\pi f t} dt \right|^2 \right\rangle = 4 \lim_{T \to \infty} \int_0^T C(s) \cos(2\pi f s) ds, \quad (2)$$

where T denotes the considered time interval of time series $I(t)$. The KM model is based on the time series generated as:

$$I(t) = \sum_k q_k \delta(t - t_k) \quad (3)$$

where q_k is a contribution to the signal of one pulse and noise is due to the correlations between the occurrence times t_k of δ type pulses. This model corresponds to the flow of point objects: photons, electrons, cars, trades in financial markets and so on. The simplest version of the model consist of sequence of transit times t_k described by recurrence equations:

$$\begin{aligned} t_k &= t_{k-1} + \tau_k, \\ \tau_k &= \tau_{k-1} - \gamma(\tau_{k-1} - \overline{\tau}) + \sigma\varepsilon_k. \end{aligned} \tag{4}$$

Here the recurrence time $\tau_k = t_k - t_{k-1}$ fluctuates due to the external random perturbation of the system by sequence of uncorrelated normally distributed random variables $\{\varepsilon_k\}$ with zero expectation and unit variance, where σ denotes the standard deviation of the white noise, $\gamma \ll 1$ is the recurrence time τ relacsation rate to the some average value $\overline{\tau}$. Note that τ follows an autoregressive process AR(1).

This model containing only one relaaxation time γ^{-1} can for sufficiently small parameters γ produce an exact $1/f$ - like spectrum in wide range of frequence [9]:

$$S(f) = \frac{\alpha_H}{\overline{\overline{\tau}}^2 f}, \quad f_1 < f < \min(f_2, f_{\overline{\tau}}), \tag{5}$$

where α_H is a dimensionless constant — the Hooge parameter:

$$\alpha_H = \frac{2}{\sqrt{\pi}} K e^{-K^2}, \qquad K = \frac{\overline{\tau}\sqrt{\gamma}}{\sigma}. \tag{6}$$

and $f_1 = \gamma^{3/2}/\pi\sigma$, $f_2 = 2\gamma^{1/2}/\pi\sigma$, $f_\tau = (2\pi\overline{\tau})^{-1}$.

Here we present generalized KM model with particular interest to reproduce the statistical properties of share volume traded in financial markets. We will generate discrete share volume time series with recurrence time of distinct trades described as AR(2) process. This particular application of the model proposed by B. Kaulakys and T. Meškauskas will exhibit univesality of the mechanism underlying production of an $f^{-\beta}$ noise.

2 Model definition

In the numerical data analysis we usually deal with discrete sets of data. There is the first need to modify the model to the discrete one. Let us introduce a new conventional time scale defined with a time interval τ_d. Integrating continuous signal $I(t)$ in the subsequent intervals of length τ_d

we will get discrete time series (let us call it volume V_r) :

$$V_r = \int_{t_r}^{t_r+\tau} I(t)dt = \sum_k q_k, \quad t_r = r\tau_d. \tag{7}$$

Consequently, the discrete power spectral density can be expressed as:

$$S(f_s) = \left\langle \frac{2}{\tau_d n} \left| \sum_{r=1}^{n} V_r e^{-i2\pi(s-1)(r-1)/n} \right|^2 \right\rangle, \tag{8}$$

where $f_s = \frac{s-1}{T}$, $T = \tau_d n$.

New discrete series are equivalent to the initial sequence of δ functions, when $\tau_d \gtrsim \bar{\tau}$ whit the same values of γ and σ. Numerical results of $S(f_s)$ calculated with various parameters are shown in Fig. 1. Multiple numerical calculations confirm full correspondence of the discrete model (DM) defined here with earlier introduced by B. Kaulakys and T. Meškauskas [9].

The model can produce the $1/f$- like spectrum in an arbitrarily wide range of frequencies $f_1 < f < f_2, f_\tau$ and is free from unphysical divergence of the spectrum at $f \to 0$; for $f < f_1$:

$$S(f) = \bar{\tau}^{-2} \frac{2\sigma^2/\bar{\tau}\gamma^2}{1 + \sigma^4/4\bar{\tau}^2\gamma^4}. \tag{9}$$

Due to the long memory random process, defining transit time sequence t_k, the model describes long time correlations quantified in power spectral density $S(f) \backsim 1/f$. This model may be also generalized for non-Gaussian distribution of the periods τ_k. Then

$$S(f) = 2\bar{\tau}^{-1}\psi(0)/f, \tag{10}$$

where $\psi(\tau)$ is the distribution density of periods τ_k. This makes the model applicable to the wide variety of stochastic processes, which have well defined distribution function $\psi(\tau)$ in the vicinity of $\tau_k = 0$. Numerical calculations confirm that, when $\psi(0) = 0$, the dependance of the power spectral density on frequency appears as $S(f) \backsim 1/f^{3/2}$ [10].

One more generalization of the model is needed for applications in financial markets and other self-affine time series with a power-law dependence on frequency $S(f) = f^{-\beta}$ with $\beta \simeq 1/2$. We will strengthen high frequencies

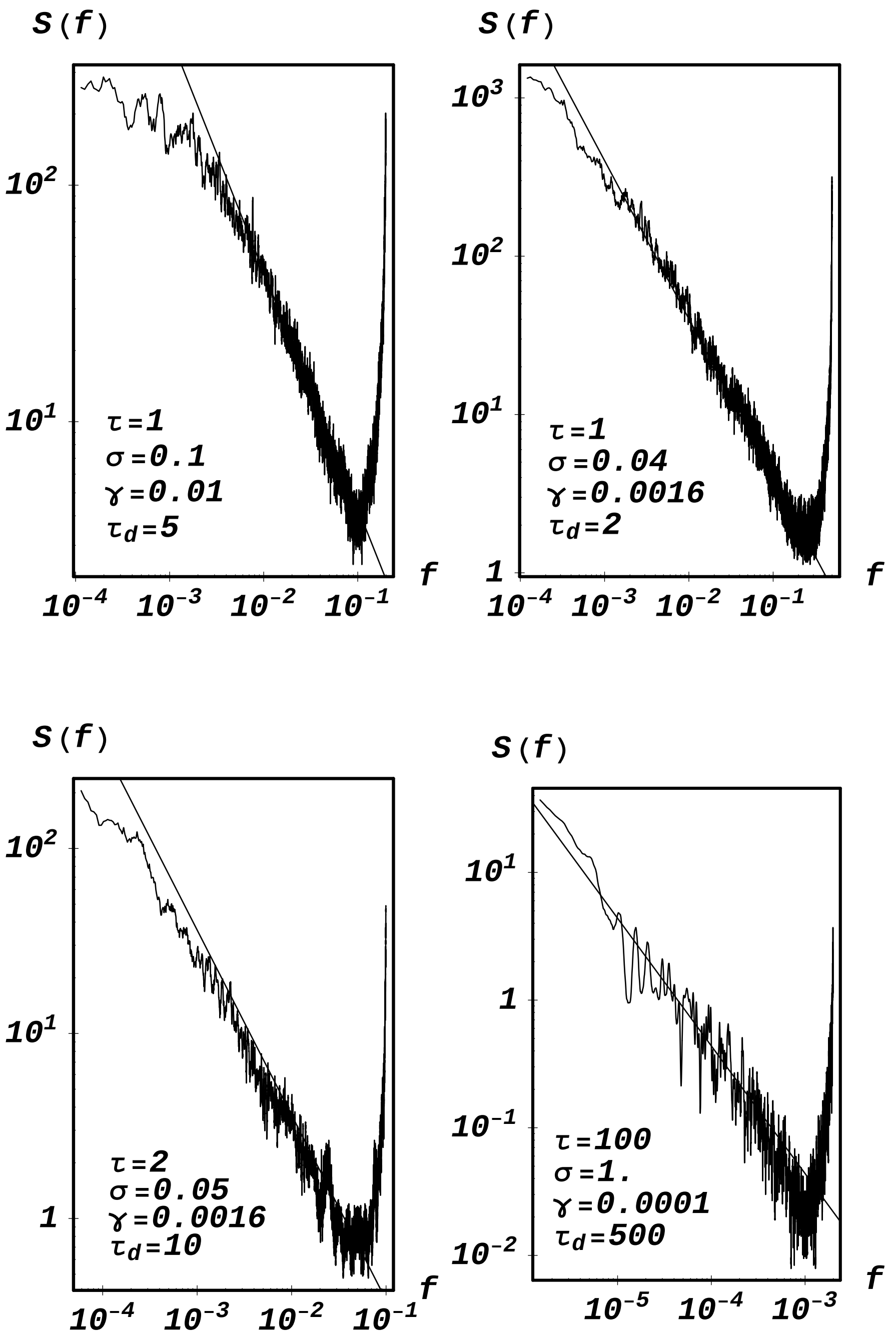

Figure 1: Power spectral density versus frequency calculated from the model described by Eqs. (3), (4), (7), (8). The main parameters defining the model are: $\overline{\tau}, \sigma, \gamma$. Time scale τ_d determines the highest frequency $f \leq \frac{1}{\tau_d}$ under consideration. The sinuous curves represent the results of numerical simulations averaged over five realizations, and the straight lines represent the analytical spectra described by Eqs. (5) and (6).

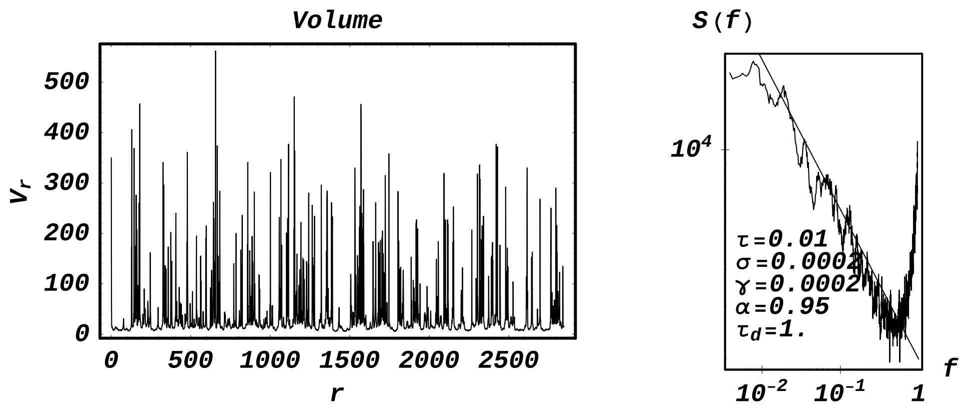

Figure 2: Volume V_r versus the number r of time interval τ_d and power spectral density $S(f)$ versus frequency f, calculated from the generalized discrete model (GDM) with $\sigma = 0.0002$, $\overline{\tau} = 0.01$, $\alpha = 0.95, \gamma = 0.0002, \tau_d = 1$. S is averaged over five realizations. The straight lines represent the fractional power law $S(f) = 2 * 10^3/f^{1/2}$.

and will account for positive playback of τ_k increment by adding the term $\alpha\Delta\tau_{k-1} = \alpha(\tau_{k-1} - \tau_{k-2})$ to the τ_k recurrent expression:

$$\tau_k = \tau_{k-1} + \alpha(\tau_{k-1} - \tau_{k-2}) - \gamma(\tau_{k-1} - \overline{\tau}) + \sigma\varepsilon_k. \tag{11}$$

Note that this new term changes autoregressive process AR(1) to the higher one AR(2). Multiple numerical calculations with the generalized discrete model (GDM) exhibit dependance of β on α and other parameters of the model: $\sigma, \overline{\tau}, \gamma$. We demonstrate an example of numerical calculation with GDM in Fig. 2, which exhibits clear fractional power law with $\beta = 1/2$ of the power spectral density.

3 Application to the financial market

An important quantity that characterizes the dynamics of price movement in financial markets is the number of shares V_r (share volume) traded in a time interval $r\tau_d < t < (r+1)\tau_d$. The statement "It takes volume to move stock prices" accumulates very general idea that statistical properties of financial markets are enclosed in time series of share volume. Very direct confirmation of this statement and quantitative investigation of the largest 1000 stocks in three major US stock markets was recently presented in [6]. This work provides an evidence that long range correlations in share volume and price volatility are largely due to those of the number of trades N_r in time interval τ_d. Close correlation between N_r and V_r is imposed by the relation:

$$V_r \equiv \sum_{i=1}^{N_r} q_i \tag{12}$$

and very weak correlation in time sequence of share volume per transaction q_i. These results suggest us to apply GDM as a model for the time series of share volume traded in financial markets , with a simple assumption that in the first approach the average $\langle q_i \rangle$ can be included into the normalization factor. This mean that simple relation $V_r = N_r \langle q_i \rangle$ enables us to make comparison of GDM to the variety of real market data. In Fig. 3. we demonstrate comparison of Lithuanian Stock Market data with numerical results from GDM. The volume of shares included in the index LITIN is normalized to average of 100 trades per $\tau_d = 1$day . Note that despite the model simplicity it serves as market data generator and reproduces the main statistical property of the system, i.e., the power law dependence ($\beta = 1/2$) on the frequency of the power spectral density.

4 Conclusion

The empirical evidence provided by Gopikrishnan *et al* [6], that the number of transactions N_r in the subsequent time intervals τ_d define long range correlations of share volume traded, enabled us to apply simple model of a $1/f$ noise [9] and to reproduce long-range correlations of the share volume traded in financial markets. We generalized the KM model by integrating the sequence of pulses in a conventional time interval τ_d and replacing the recurrent time τ_k stochastic process from AR(1) to the AR(2) one. Numerical calculations with the generalized discrete model (GDM) reproduce power spectral density $S(f)$ scaled as power of frequency $1/f^\beta$ for various values of β, including $\beta = 1/2$ for applications in financial markets. Further investigation of the model with its possible applications in financial time series is in progress.

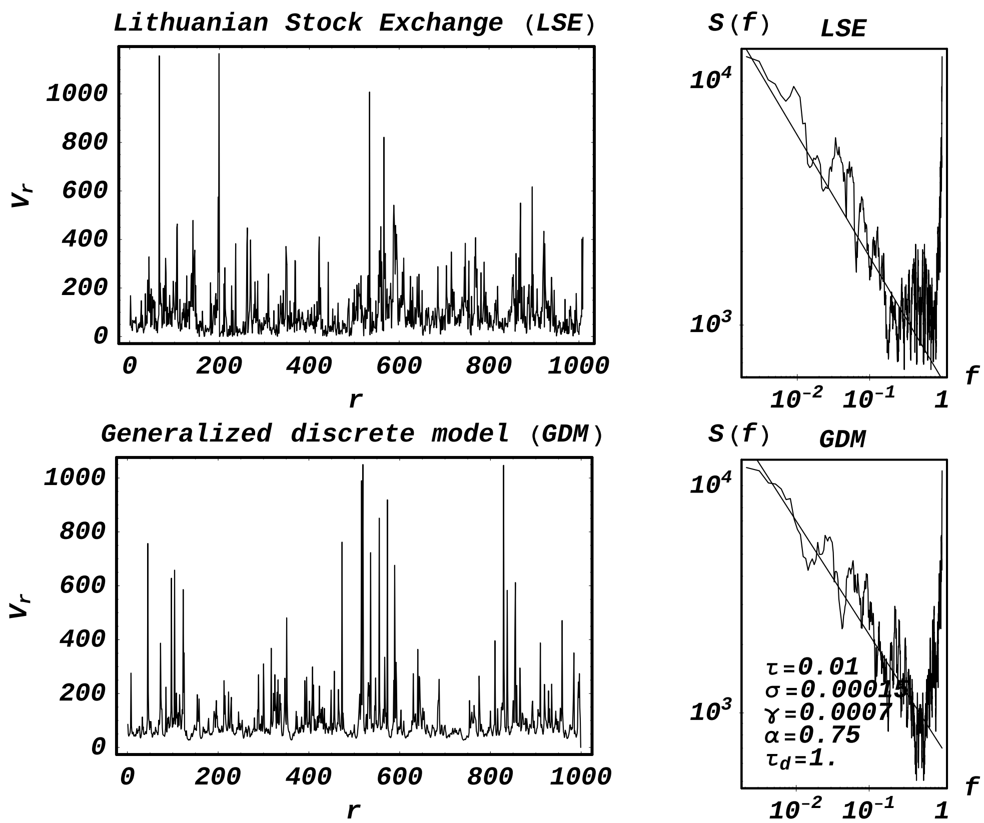

Figure 3: **(LSE)** The volume of shares V_r , included in the index LITIN of Lithuanian Stock Exchange, normalized to an average of 100 trades per $\tau_d = 1day$ versus the number of day traded and corresponding power spectral density $S(f)$ versus frequency f calculated from Fast Fourier Transform of discrete data. **(GDM)** Volume V_r versus the number r of the time interval $\tau_d = 1$ and power spectral density $S(f)$ versus frequency f, calculated from the generalized discrete model (GDM) with $\sigma = 0.00015$, $\overline{\tau} = 0.01$, $\alpha = 0.75, \gamma = 0.0007, \tau_d = 1$. The straight lines approximating the power spectral density curves represent the fractional power law $S(f) = 700/f^{1/2}$.

Acknowledgment

The author gratefully acknowledges numerous and long discussions with Bronislovas Kaulakys regarding inherent origin of $1/f$ noise and applicability of KM model.

References

[1] **J. D. Farmer,** Computing in Science & Engineering, V. 1, p. 26 (1999).

[2] **R. N. Mantegna and H. E. Stanley,** *Introduction to Econophysics: Correlations and Complexity in Finance*, (UK,Cambridge, Cambridge Univ. Press., 1999).

[3] P. **Gopikrishnan et al.** Phys. Rev. E, V. 60, p. 5305 (1999).

[4] **V. Plerou et al.,** Phys. Rev. E, V. 60, p. 6519 (1999)

[5] **Y. Liu et al.,** Phys. Rev. E, V. 60, p. 1390 (1999).

[6] P. **Gopikrishnan et al.,** Phys. Rev. E, V. 62 (4), p. R4493-R4496 (2000).

[7] **B. B. Mandelbrot,** Science, V. 156, p. 636-638 (1967).

[8] **B. B. Mandelbrot, J. W. van Ness,** SIAM Rev., V. 10, p. 422-437 (1968).

[9] **B. Kaulakys and T. Meškauskas,** Phys. Rev. E, V 58, p.7013 (1998).

[10] **B. Kaulakys and T. Meškauskas,** Nonlinear Analysis: Modelling and Control, Vilnius, IMI, 1999, No 4