The Henryk Niewodniczański
Institute of Nuclear Physics
Polish Academy of Sciences
Department of Complex Systems Theory

Multiscale characteristics of linear and nonlinear cross-correlations in financial markets

Marcin Forczek

This thesis was done under the supervision of
dr hab. inż. Paweł Oświęcimka

Kraków 2016

Acknowledgements

I am profoundly indebted to my supervisor, dr hab. inż. Paweł Oświęcimka, who was very generous with his time and knowledge and assisted me in each step to complete the thesis. I would also like to thank dr hab. Jarosław Kwapień for his remarks and guidance in writing the thesis. Last but not least, I would like to thank prof. dr hab. Stanisław Drożdż, Head of the Department of Complex Systems Theory, for his continued and friendly support throughout my studies.

1 Introduction

With his work "An Inquiry into the Nature and Causes of the Wealth of Nations" published in 1776 [1], the Scottish philosopher and thinker Adam Smith laid foundations for economics as a separate branch of science. In the above-mentioned work, he used the term "invisible hand" meaning a mechanism that directs individual consumers so that their actions are the most beneficial to the whole of society. Although being commonly known and widely covered in the literature, this problem has not been to date put into mathematical formulas in a manner that would enable the market as a whole to be described precisely. This effect is a classic example of an emergent phenomenon, i.e. such a characteristic or property of a system, which has occurred in the system under examination, and which was never previously observed at the level of the individual system components. Phenomena of this type are characteristic of complex systems, where, due to a large number of nonlinear interactions between the elements of the system, evolving structures appear, which cannot be derived from the components in a simple manner. Looking from this perspective on financial markets in a broad sense, one can notice that they perfectly fit into the concept of complex systems. Starting from individual investors in the financial markets or individuals in their households, whose behavior can be described at best in a collective manner, it cannot be determined exactly in which direction the share prices or the domestic economy will, respectively, evolve in the future.

Due to the fact that financial markets have in recent years been supplying a vast amount of information that is readily available at a relatively low cost, more and more researchers from branches of science other than economics are being concerned with market analysis. Since 1973, when currency trading started, the financial institutions have been active round the clock [2]. Until 1995, there had been a nearly 80-fold increase in the volume of transactions, and even more dynamic growth of trade had taken place in the market of derivative financial instruments. A breakthrough in the development of financial market research occurred in 1980 when the capability to make transactions electronically was introduced, which, in turn, allowed the collection of huge amounts of data that became the subject of various analyses. To get an idea of how big and complex system are the financial markets in terms of the number of components, it is sufficient to mention that only in 2013 the two world's largest stock exchanges (NYSE, NASDAQ) traded nearly $12 trillion worth of shares, while the

daily number of NYSE stock exchange transactions amounts to almost 2.5 million orders[1]. Bloomberg, the world's largest supplier of information systems in the area of finances, has nearly 320.000 information terminals in 73 countries, which are used each day by millions of users. Roughly from the mid-1990s, the research activity of physicists in the field of economics began to be increasingly clearly visible. This activity made a bridge between financial mathematics, regarded by many as a kind of technical analysis, and the traditional approach used by financiers. The observation and recording of movements in the financial markets has enabled the construction of theoretical models needed for their description and practical verification.

The first attempts to describe the financial markets, known in the literature, date from the beginning of the last century, when Louis Bachelier, a French mathematician, proposed the random walk as the first model of the dynamics of the share price [3] , based on the assumption that price fluctuations are subject to Gauss distribution. His doctoral thesis entitled "The Theory of Speculation" [3] laid foundations for the development of financial mathematics, and the model described therein has still been used in economics, even though it does not completely correctly describe the dynamics of share prices. It is now known that the probability density distributions of financial fluctuations in the range of large financial events disappear very slowly (fat tails) [4; 5], which in practice means that the probability of rare events is greater than it would appear from the normal distribution. One of the first, who drew attention to this fact, was Benoit Mandelbrot. He formulated the thesis that the distributions of returns should be described with the stable Levy distribution [6]. As has turned out, however, this model does not fully reflect reality, because the tails of actual financial data distribution scale following a power-law relationship, with the scaling factor being outside the range of stable Levy distributions. This means that they do not have an infinite variance. The authors of study [7] have confirmed that the central part of the empirical data distribution perfectly fits into the Levy distribution, while at the distribution ends, the variations are fairly significant, and therefore they have proposed the truncated Levy flight model as more realistic. On the other hand, Plerou et al in their study [8] by analyzing data for the 1000 largest US companies, found that the cumulative distribution of share price fluctuations behaved at its tails as a power-law distribution with an exponent of $\alpha = 3$. It was then that the formulated inverse cubic law of scaling

[1] http://www.nyxdata.com/Data-Products/ access on: 16.03.2014

appeared for the first time in research. However, the concept of power-law distributions had already been known in science since the 19th century, when V. Pareto used them for the first time to describe the sociological phenomena [9]. We deal with scaling when the graph of a function, as represented on a double logarithmic scale, is a straight line, in which case the slope of the line is the exponent of a power-law function. It is worth noting that scaling is, by the way, a natural feature of fractals, and it is fractal studies initiated by Mandelbrot that led to the creation of financial models relying on the self-similarity of financial data, such as *Multifractal Random Walk* [10; 11], and *Multifractal Model of Asset Returns* [12], as well as to the development of techniques enabling the study of fractal structures. Real structures exhibit self-similarity (they scale themselves) in a statistical sense. For example, looking at the stock exchange quotations of any company, we are not able to distinguish whether a given signal comes from minutes' or daily quotations. Exactly the same is with fractals, whose core feature is that, when looking at the fractal structure at different levels of complexity, we see an object similar to the whole. The fractal formalism turned out so interesting to scientists that attempts were made to use it, for instance, in fluid dynamics [13]. The very concept of self-similar structures in fluid dynamics appeared already in Kolmogorov's works in 1941 [14], and was then developed in the works of Mandelbrot [15], Frisch and Nelkin [16; 17] and Meneveau and Sreenivasan [13; 18; 19; 20]. A breakthrough, in the context of finances, was the work by Ghashghaie [21], published in Nature in 1996, which showed a similarity between turbulence and foreign exchange (FX) fluctuations. It has turned out that the flow of information between different time scales in FX markets is hierarchical and there exists its analogy to the turbulent flow of fluid, where the energy is transferred from large to small scales. Study [21] has further demonstrated that the probability density distribution of exchange rate fluctuations within a certain time period is very much similar to the distribution of velocity changes between two points in a turbulent flow with a defined spatial dimension. Despite the fact that another work [22] verified these findings to be too far-reaching, or even indicated significant difference (rather than similarities) in the dynamics of the both systems [23], this did not prevent the development of further studies in this direction [24; 25].

Closely related with turbulence is the notion of cascade processes. In such a process, a large vortex with the energy ε on the scale l_0 transforms or divides itself into smaller vortices, and these in turn divide themselves into even smaller ones, and so no, down to the Kolmogorov scale η, below which the vortex energy is converted into heat.

In each step, the vortices are scaled down with a factor β. Figure 1.1 shows a graphic idea of the cascade process [26].

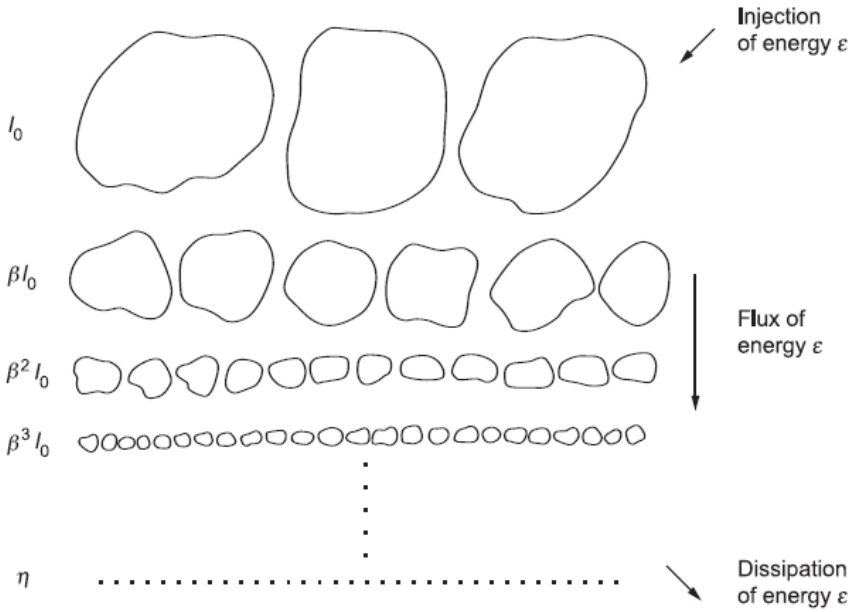

Figure 1.1 A graphic idea of the cascade process.

Processes of the cascade type have also found application in financial modelling in the form of the Markov Switching Multifractal (MSM) model. This is an iterative model that is able to replicate the actual structure of financial data (including variation clustering, which is one of the stylized facts[2]), and to ensure the fractality of the model.

The aim of this work is to perform the fractal analysis of financial data with the emphasis on the interaction between the individual elements of the financial market. The study used data from the FOREX currency market in the form of 9 transaction foreign exchange rates from the period from 17.01.2011 to 21.01.2011; 29 companies included in the DJIA index[3] quoted every minute during the period from 02.01.2008 to 30.09.2009; and for the network analysis of the markets, the 100 largest and smallest US companies quoted in 5 minutes' intervals in the period from 1.12.1997 to 31.12.1999.

[2] Stylized facts – the distinguished, typical characteristics of a given system. In the case of the time series of financial data, in addition variation clustering, stylized facts include fat tails, returns distribution asymmetry, short system memory at the price change level and long at the volatility level and the leverage effect in the case of change in returns variance [118].
[3] The DJIA (Dow Jones Industrial Average) index is a stock market index made up of the 30 largest US companies listed on the New York Stock Exchange (NYSE and Nasdaq). It has been in use since 1896. www.investopedia.com: access on: 15.05.2014.

The thesis consists of six chapters. Chapter Two includes an analysis of the properties of statistical time series, such as the autocorrelation function and the cross correlation function, as well as the distribution of probability density on different time scales. In the Chapter Three, general concepts and fractal characteristics necessary for further understanding of the study are introduced. Chapter Four is a description of modern fractal analysis algorithms and their application on the example of financial data. The focus is primarily on the Multifractal Detrended Fluctuation Analysis (MF-DFA) autocorrelation method, with the obtained results being supplemented with analysis using the Wavelet Leaders method. A new method of power-law Multifractal Cross-Correlation Analysis (MFCCA) has been introduced, while indicating its great advantage over other multifractal cross-correlation analysis methods. The MFCCA algorithm enables the correct identification and quantitative description of the multifractal cross-correlation between two time series and, unlike other multifractal analysis methods, is free from restrictions typical of other methods. In addition, a higher-order cross-correlation coefficient, ρ^q, has been introduced for detrended functions. Chapter Five contains an analysis of financial markets in a network representation, which allows the market structure to be looked at from a completely different perspective. Chapter Six contains summary and conclusions.

2 Statistical properties of financial data

2.1 Models of financial markets

The first financial market evolution model, presented by Louis Bachelier at the beginning of the last century, assumed that the assets price was subject to random walking, and the return had the Gaussian behaviour, so they can be described within a random game [3]. In some financial problems, this model is continued to be used owing to its simplicity, though it does not completely correspond to reality. It is sufficient to look at Figure 2.1 showing the actual financial data (the middle graph) to notice that values occurs in the signals, which have such a large amplitude that practically does not happen in Gaussian distributions. The red line marks the range $\pm 3\sigma$ and for the Gaussian variable 99.7% of the values are contained in this range. At first glance, it can be noticed that for the actual data it is quite different. For this particular reason it is generally accepted, and is confirmed by studies, that the distributions of actual financial data have an elongated central distribution part and fat tails. So, these are leptokurtic distributions. The explanation of this phenomenon might be the fact that financial markets experience intermittent problems with liquidity. In the case of high market liquidity, no large price spikes should appear.

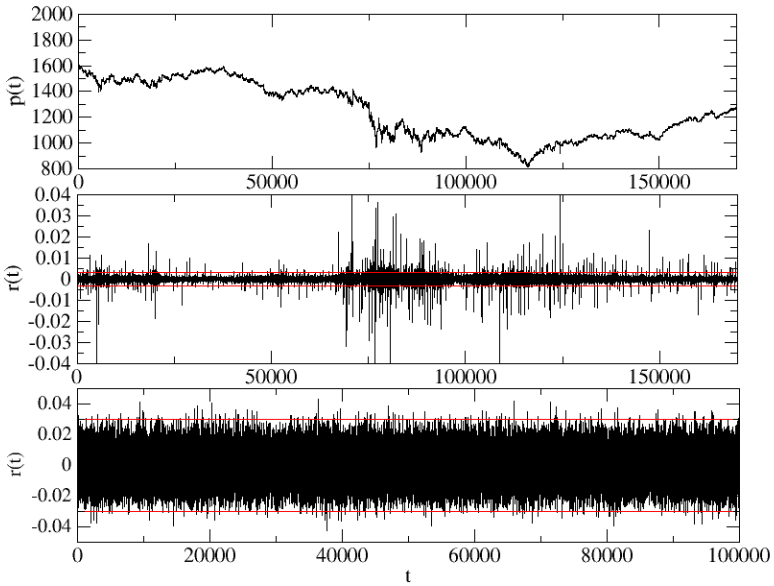

Figure 2.1 A diagram of one minute's change in DJIA index price (the upper figure), the returns of this index (the middle figure) and the Gaussian process. The range $\pm 3\sigma$ is marked in red.

In 1968, Mandelbrot and Van Ness [27] introduced the fractional Brownian motion $B_H(t)$ with the mean value equal to zero and covariance given by the following formula to describe the financial data:

$$E[B_H(t)B_H(s)] = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H}). \quad (1)$$

Unlike the classic Brownian movement, this is the process with memory, which is dependent on the parameter $H \in (0;1)$ and its increments do not need to be independent. In addition, this is a self-affine process, which means that it satisfies the scaling relationship:

$$B_H(ct) \sim |c|^H B_H(t). \quad (2)$$

If $H > 1/2$ we have a persistent time series, in which a positive correlation occurs between successive changes in the series, and the process itself is referred to as a process "with memory". This refers to a long-term memory of the process, which means that the existing change in the series is the resultant of all previous changes, even those very distant, which may only have a fragmentary contribution to the present situations. Moreover, what is happening at present has an effect on distant events in the future. In series of this type, a trend is visible. When $H < 1/2$ the next changes in the series will be correlated negatively, and such a series is called anti-persistent. Consecutive changes between the successive terms of the series are alternating, i.e., if, at a given moment, the series takes on positive values, then it is more likely that it will assume a negative rather than positive value at the next moment. If $H = 1/2$, the process is a random walk and the successive changes in this process will be completely uncorrelated. The parameter *H* defines, therefore, the volatility of the time series. A more detailed description of the *H* is provided in section 4.1.

In the following years, there appeared also geometric and arithmetic Brownian motion models, which relied on the process described by fractional Brownian motion, but which did not have a specific limitation, namely that values assumed by the process might not be negative. They gave better results, though still far from being ideal, just because of the fat tails of the probability distributions of actual data, impossible to be reproduced in the case of the aforementioned processes. Mandelbrot also introduced Levy processes characterized by the probability distribution [7]:

$$L(x) = \frac{1}{\pi}\int_0^{+\infty} e^{-\gamma k^{\alpha_L}} \cos(kx)\, dq \quad (3)$$

for the description of the dynamic of financial instrument prices, where $\gamma > 0$ is a scaling factor and the parameter $0 < \alpha_L \leq 2$ is a stability index. Levi processes have an infinite variance in the above-mentioned stability range, therefore they do not obey the *Central Limit Theorem* [4]. The way of overcoming the infinite variance problem in the context of financial data was the introduction of the Truncated Levy Flights [7], which are given by the following formula:

$$TL(x) = \begin{cases} 0 & x > l \\ cL(x) & -l \leq x \leq l, \\ 0 & x > l \end{cases} \tag{4}$$

where l is the length of truncation, c is a normalization constant and $L(x)$, the Levy distribution. Distributions of this type have a finite variance, exponentially decaying tails, and in the central part they are described by the Levy distribution.

As has been previously mentioned, the tails of actual financial data distributions scale themselves following the power law according to the formula:

$$P(r > x) \sim x^{-\alpha}, \tag{5}$$

for which $\alpha \approx 3$. Scaling of this type can be observed in the case of the indices developed share markets, emerging markets and commodity markets, and is called the *inverse cubic law*. Distribution of this type are not stable in the Levy sense[5] (they have a finite variance), therefore, for sufficiently large time scales, they should converge to the Gaussian distribution in accordance with the *CLT*. For share markets, the convergence to the Gaussian distribution would be relatively slow with increasing scale [7], though recent studies show [4] that for contemporary markets this convergence is noticeable for scales $\Delta t > 1$ min.

Another interesting approach to the description of financial markets has been the formalism of nonextensive statistical mechanics proposed by Tsallis in his wok [28]. It is based on the concept of nonextensive entropy which is one of the possible generalizations of Boltzmann-Gibbs entropy. If a given dynamic system is an ergodic system, then this means that its averaged behaviour over time is the same as the behaviour averaged over the set of all available phase space states. This, in turn, means that all microstates are equally likely in the long run. In this sense, the system is

[4] In accordance with the *CLT*, the distribution of a random variable which is a superposition of independent random variables having a finite variance and the expected value equal to, has an asymptotically normal distribution.

[5] Process X is stable, if the sum of independent processes X_1 and X_2 with a distribution S, is the process with the distribution S: $aX_1 + bX_2 = cX + d$. A feature of stable distributions is that they retain their statistical properties after summing and scaling. For stable financial time series, minute and hourly returns should have the same variance and mean value.

considered in the context of classical Boltzmann-Gibbs thermodynamics which is based on the assumption that all system components are statistically independent and the entropy is given by the formula:

$$S_{BG} = -k\sum_{i=1}^{W} p_i \, lnp_i. \tag{6}$$

The entropy S_{BG} for the entire system is the sum of the entropies of the subsystems (it is extensive).

As will be shown in the next sections, the financial market is a system in which all components are interrelated, therefore the use of distributions based on classical thermodynamics is not necessarily the best choice. The approach based on generalized entropy [29]:

$$S_q = k\frac{1-\sum_{i=1}^{W} p_i^q}{q-1} \ (q \in \mathcal{R}), \tag{7}$$

gives better results when applied to empirical data. The parameter q in the formula above defines the statistics, and for $q = 1$ we have simple entropy S_{BG}. *q-Gaussian* distributions that maximize the entropy (**7**) on the assumption that the generalized mean value satisfies the relationship $\overline{\mu_q} = \int x \frac{[p(x)]^q}{\int [p(x)]^q dx} dx$ and the generalized variance satisfies the relationship $\sigma_q^2 = \int (x-\mu_q)^2 \frac{[p(x)]^q}{\int [p(x)]^q} dx$, are given by the formula:

$$p_q(x) = N_q e_q^{-B_q(x-\mu_q)^2}, \tag{8}$$

where $\mathrm{B}_q = ((3-q)\sigma_q^2)^{-1}$, the *q-exponent* is defined by the formula $e_q^x = [1 + (1-q)x]^{\frac{1}{1-q}}$ while the normalizing factor is:

$$N_q = \begin{cases} \dfrac{\Gamma\left(\frac{5-3q}{2(1-q)}\right)}{\Gamma\left(\frac{2-q}{1-q}\right)}\sqrt{\dfrac{1-q}{\pi}}B_q & q < 1 \\ \dfrac{1}{\sqrt{q}} & q = 1 \\ \dfrac{\Gamma\left(\frac{1}{q-1}\right)}{\Gamma\left(\frac{3-q}{2(q-1)}\right)}\sqrt{\dfrac{q-1}{\pi}} & 1 < q < 3 \end{cases}. \tag{9}$$

What is important in the case of the *q-Gaussian distribution* is the fact that it asymptotically takes on the form of a power-law distribution, as follows:

$$p_q(x) \sim x^{\frac{2}{1-q}}, \tag{10}$$

while if $q = 1$, it assumes the form of a normal Gaussian distribution. For $1 < q < 5/3$ the sum of *q-Gaussians* will be convergent to the Gaussian distribution, while for $5/3 < q < 3$, it will be convergent to the Levy distribution, on account of the Generalized Limit Theorem [30]. It is important, insomuch as using a single parameter q, it is possible to model a time series being either a monofractal (with the Gaussian distribution) or a multifractal (with the Levy distribution), as will be discussed later in this study.

As shown in reference [31], *q-Gaussians* can be successfully used for describing the probability distributions of foreign exchange series. Below, some examples of fitting empirical data with these distributions will be shown. For this purpose, the cumulated form of Formula (**8**) presented in study [31] will be used:

$$P_{q^{\pm}}(x) = N_q \left(\frac{\sqrt{\pi}(\frac{1}{2}(3-q)\beta}{2\Gamma(\beta)\sqrt{\frac{B_q}{\beta}}} \pm (x - \mu_q \right) {}_2F_1(\alpha, \beta; \gamma; \delta). \tag{11}$$

In the above formula, $\alpha = \frac{1}{2}, \beta = \frac{1}{q-1}, \gamma = \frac{3}{2}, \delta = B_q(q-1)\left(\mu_q - x\right)^2$, while ${}_2F_1(\alpha, \beta; \gamma; \delta) = \sum_k^{\infty} \frac{\delta^k (\alpha)_k (\beta)_k}{k! (\gamma)_k}$ is a hypergeometric Gaussian function.

2.2 Probability distributions of returns

A basic characteristic in examining the statistical properties of financial fluctuations are the probability distributions of returns. These distributions make it possible to construct appropriate theoretical models to describe processes, as well as to look inside them. In the calculations, the logarithmic rate of return, R (*return)*, was used, which, for the time series $p(t)$ representing the value of the signal (price) p in time t, is defined by the following formula:

$$R \equiv R(t, \Delta t) = \ln\left(p(t + \Delta t)\right) - \ln\left(p(t)\right), \tag{12}$$

where $t = 1,2, \dots, N$ is the time instant, and Δt, the scale under analysis. Then we normalize the returns:

$$r \equiv r(t, \Delta t) = \frac{R - u}{\nu}, \tag{13}$$

where v is the standard deviation of returns in time T while u is the average value after the time T. Owing to the normalization, calculations performed for different financial data are comparable. The analysis of fluctuations in returns is best carried out on a log-log scale, because on this scale, power-law relationships take on the form of a straight line, which facilitates the scaling identifications. The distribution of fluctuations is approximately a symmetrical distribution 2.1 2.3, so the absolute values of $|r|$ will be analyzed.

There are a number of studies in the literature concerning the distribution of returns for shares [8; 32], foreign exchanges [31] or commodities [33], however, the statistical analysis presented in this paper, based on high-frequency data in the form of transaction foreign exchange rates, to the author's knowledge, has not been developed yet. For the financial data, a systematic analysis of the statistical properties was started from examining the distributions of the DJIA index returns.

2.2.1 The foreign exchange market (FX)

Forex (FX) is a foreign exchange market, in which various financial institutions, banks, corporations and governments carry out foreign exchange operations. It is an *OTC* (*Over the Counter*) market, where brokers or dealers negotiate currency prices directly between each other, and there is no central clearing house. This is the largest market in the world with an estimated daily turnover of 5.3 trillion US dollars[6]. For comparison, the forward exchange market, which is the world's second largest market in terms of turnover, is estimated at less than 440 billion US dollars of daily turnover. Foreign exchange trade takes place round the clock 5 days a week here, excluding weekends, with a huge number of transactions being made; so, this is the most liquid market in the world.

The world's largest foreign exchange broker in 2014 was CITI, whose market share amounted to 16.04%[7] and the currency most often traded in 2013 was the American dollar with a market share of 84%[8]. Movement on the FX has impact on other financial markets, because all other goods are expressed in different currencies, and therefore this market is the world's most important market. From a physicist's point of view, the FX is a complex system with extremely complicated time relationships, as there are many

[6] http://www.bis.org/press/p130905.htm access on: 24.09.2014
[7] http://www.bloomberg.com/news/2014-05-08/deutsche-bank-currency-crown-lost-to-citigroup-on-volatility-1-.html access:24.09.2014
[8] http://www.bis.org/publ/rpfx13fx.pdf access:24.09.2014

factors influencing the formation of individual foreign exchange rates. Each of the currency pairs under analysis is characterized by its own dynamics, being dependent of the national economy, central banks (interest rates/inflation) or, as will be demonstrated later on in this study, events taking place in other markets and in different countries.

The performed analysis covered the transaction exchange rates of 9 currency pairs (Figure 2.2) quoted in the period from 17.01.2011 to 21.01.2011, originating from the Deutsche Bank broker, which was the world's second largest broker in 2014:

- GBPUSD-1’157’571 records,
- USDJPY-485’169,
- GBPJPY – 855’921,
- EURUSD-928’925,
- EURJPY-1’153’303,
- EURGBP-1’359’974,
- EURAUD-1’685’455,
- AUDUSD-775’698,
- AUDJPY-1’204’731.

The data were then converted to the form of equally time-spaced time series with pre-defined intervals of Δt : 5-sec. - 86’400 records; 30-sec. - 14’400; 1-min. - 7’200; 2-min. - 3’600; 4-min. - 1’800; 10-min. - 720, covering the entire trading day from 00:00 to 24:00 hours.

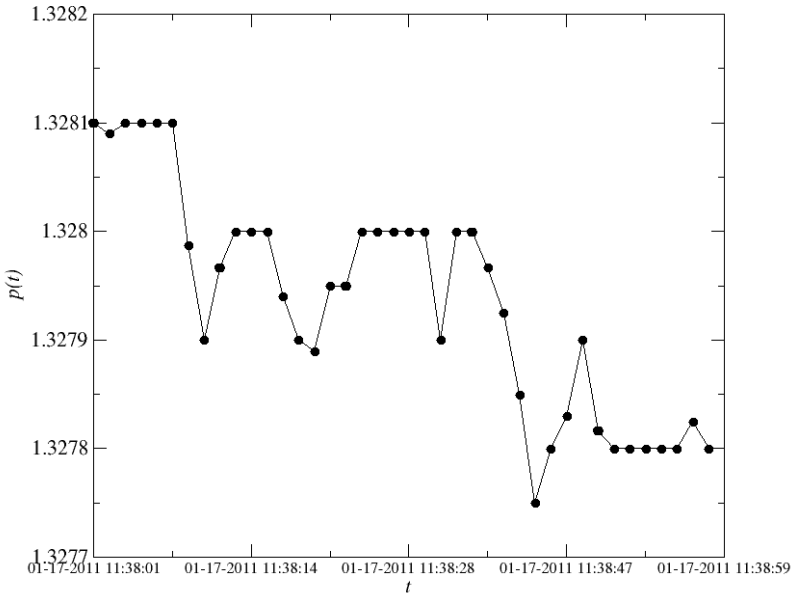

Figure 2.2 A minute of EURUSD currency pair trading, covering 40 transactions.

Subjected to analysis were also the so-called triangular relation deviations, which can be described with the following formula:

$$r^{\Delta}(t_i, \Delta t) = r_A^B(t_i, \Delta t) + r_B^C(t_i, \Delta t) + r_C^A(t_i, \Delta t) \quad (14)$$

and which, assuming the ideal market efficiency [9], should satisfy this equation:

$$r^{\Delta}(t_i, \Delta t) = 0, \quad (15)$$

where $r_A^B = -r_B^A$. In formula (**14**), denotes the logarithmic rate of return for the rate of exchange of currency X for currency Y, in time t_i and on the scale Δt. Deviations from (**15**) offer an excellent opportunity for generating profits at no risk, and are referred to as the triangular arbitrage. On contemporary financial markets, such a behaviour of data rates is actually found on very short time scales. Because of this, it is dedicated software programs being at financial institutions' service that close possible arbitrage items by using so-called *high frequency trading* (HFT)[10]. Considering the fact that triangular relation deviations occur at the 6th or a further decimal place, then in order to achieve a measurable profit, it is necessary to invest big money, which essentially precludes the involvement of individual investors.

Using a signal given by formula (**14**) for analysis, we reduce the number of the system's degrees of freedom, thus going down from three different signal to one. While some exchange rates can be a result of completion of specific investment strategies carried out by traders/institutions, others may be a result of the market movement to eliminate that strategy, or the liquidation of possible arbitrage. When a player buying euros for dollars exchanges dollars for pounds and, for the exchanged pounds, he buys back euros at one time moment, then this transaction will bring in a profit to the player from this transaction, if relationship (**15**) is greater than zero. The occurrence of such a situation in the market causes the activity of other market participants, as a result of which $r^{\Delta}(t_i, \Delta t)$ will be convergent to 0.

For examining the triangular relations, the following solid triangles were used:

- AUD-EUR-JPY: $\frac{AUD}{EUR} * \frac{EUR}{JPY} * \frac{JPY}{AUD} \neq 0$,
- AUD-EUR-USD: $\frac{AUD}{EUR} * \frac{EUR}{USD} * \frac{USD}{AUD} \neq 0$,
- AUD-USD-JPY: $\frac{AUD}{USD} * \frac{USD}{JPY} * \frac{JPY}{AUD} \neq 0$,

[9] The efficient market hypothesis propounds that, at any time moment, the price of a financial asset fully reflects all the information related to it.

[10] HFT is a technique that a large number of transactions to be made automatically using special decision-making algorithms in a very short time (of the order of seconds). www.investopedia.com: access on:19.06.2015

- EUR-GBP-JPY: $\frac{EUR}{GBP} * \frac{GBP}{JPY} * \frac{JPY}{EUR} \neq 0$,
- EUR-GBP-USD: $\frac{EUR}{GBP} * \frac{GBP}{USD} * \frac{USD}{EUR} \neq 0$,
- EUR-USD-JPY: $\frac{EUR}{USD} * \frac{USD}{JPY} * \frac{JPY}{EUR} \neq 0$,
- GBP/USD/JPY: $\frac{GBP}{USD} * \frac{USD}{JPY} * \frac{JPY}{GBP} \neq 0$,

where $\frac{AUD}{EUR}, \frac{EUR}{JPY}$, etc., are the logarithmic rates of return. Figure 2.3 illustrates how the EURUSD exchange rate changes in the period under examination and the returns, r, for $\Delta t = 5\ sec$. The daily activity of the market is clearly visible in the form of five large clusters. The maximum activity falls, more or less, on the middle of the trading day.

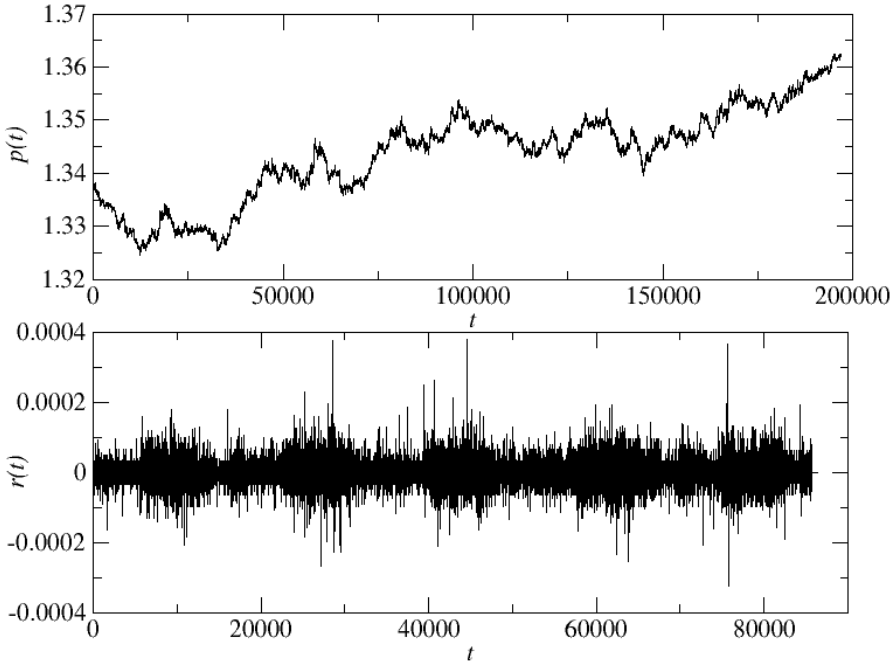

Figure 2.3 The EURUSD exchange rate (top) and the normalized five-second returns calculated for it (bottom).

Figure 2.4 represents the cumulative distributions of returns modules for four selected currencies, though the appearance and behaviour are typical of all the exchange rates examined. Moreover, table 2.1 shows the power-law fitting exponents, using the least squares method, of function given by formula (**5**) to the distribution tails. It should be expected, similarly as in study [4], that a stronger decline of the distribution tails would be noticed with increasing time scale. However, for the analyzed data, it is difficult to observe such a phenomenon as a rule. For exchange rates based on the Australian Dollar

and the GBPJPY currency pair, this is so up to the 2 min. scale, followed by fluctuation on the remaining three scales. For the remaining currency pairs, it is hard to observe any regularity in tail decline; however, it can be assumed that, barring few exceptions, these values are contained in the range between $\alpha = 3$ and $\alpha = 5$. This is likely to be influenced by the length of the time series, because of times scales greater than 5 sec., the distribution end scaling region becomes ever shorter and it is difficult to find an exact fit. The analyzed data indicates also that British pound-based exchange rates have the fattest tails, and they do not change very much with increasing time scale. This finding is consistent with previous observations from the period 2004-2008, made in study [31] .

Table 2.1 Power-law exponents for all currency pairs sampled from 5 seconds to 10 minutes.

	AUDUSD	AUDJPY	EURAUD	EURGBP	EURJPY	EURUSD	GBPJPY	GBPUSD	USDJPY
5s	3.99	3.78	4.06	3.76	4.49	4.34	3.51	3.51	4.75
30s	4.83	3.90	4.10	3.91	4.42	4.01	3.62	3.66	3.05
1m	4.91	4.40	4.53	3.43	3.59	4.28	3.81	3.16	3.66
2m	3.67	5.05	3.99	4.09	3.07	3.71	3.44	3.44	3.65
4m	4.18	4.33	4.14	3.66	3.09	3.51	3.13	2.90	3.18
10m	5.03	3.56	3.75	3.79	3.63	3.81	3.44	4.08	3.33

This may also be indicative of the occurrence of correlations in the examined signals, because their convergence to the normal distribution is slower than for uncorrelated signals. As will be demonstrated later on in this study, these are nonlinear correlations being behind the fractal nature of the examined processes.

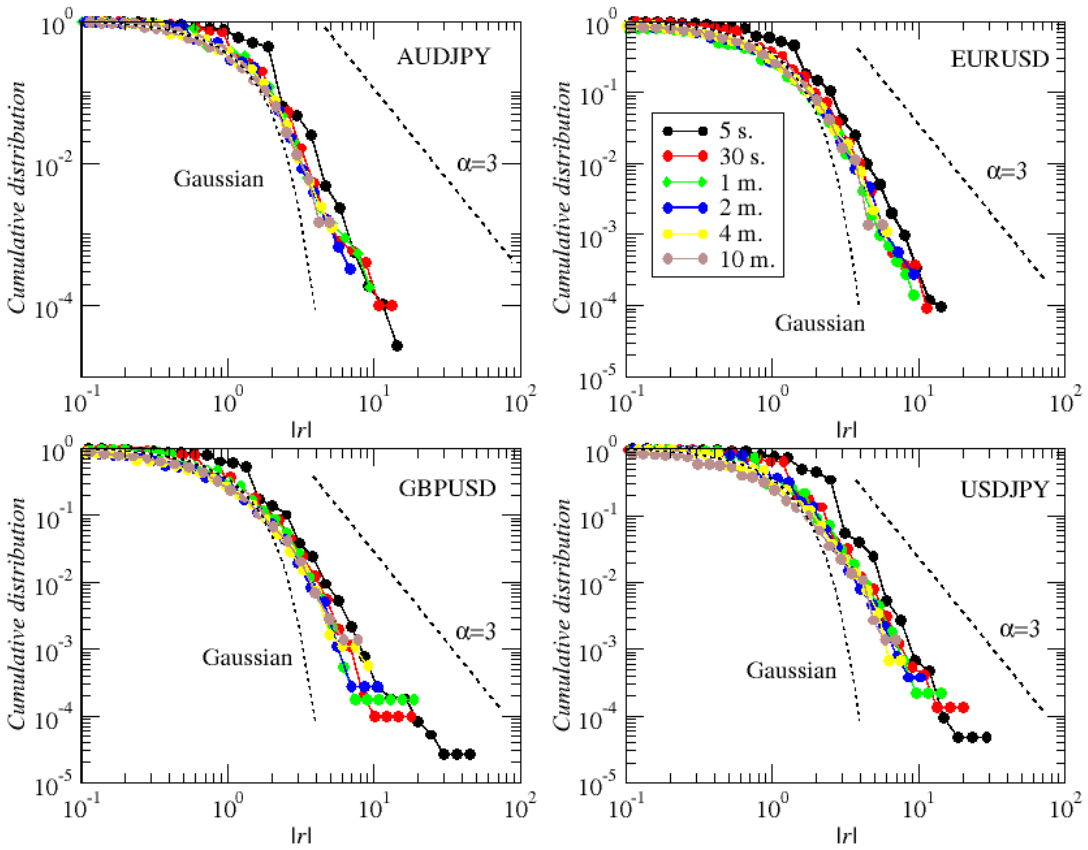

Figure 2.4 Cumulative distributions of currency returns, determined for time scales from 5 sec. to 10 min.

The situation is slightly different for the distributions of triangular relation deviations. Already at the returns level, a difference between r^{Δ} and its component signals can be seen. While the fluctuation magnitudes are the same, the structure of the series r^{Δ} itself is more uniform, with a small number of large-amplitude events. This is illustrated in Figure 2.5 for the AUD-EUR-USD pair.

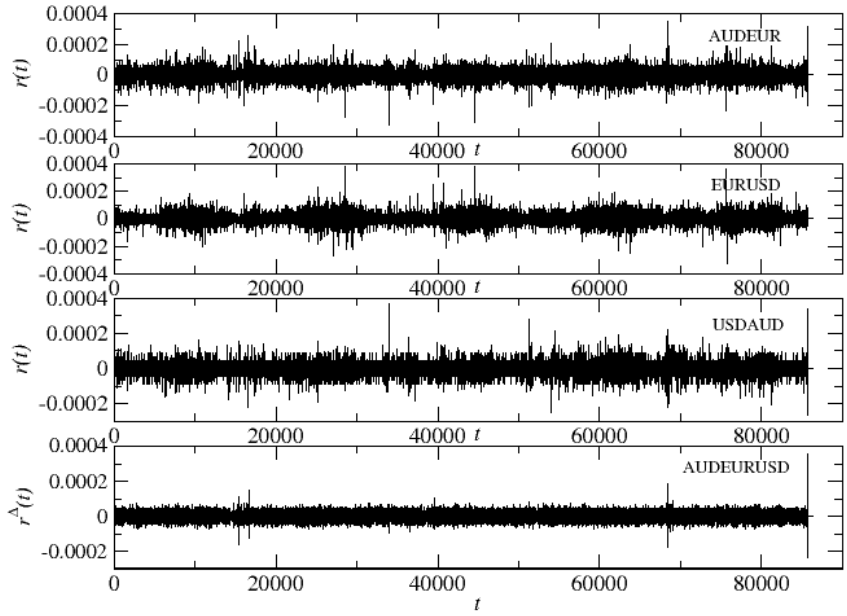

Figure 2.5 The times series of three exchange rates making up the AUD-USD-EUR triangle and the triangle itself (the series at the very bottom).

In contrast to the returns distributions, starting from the 5-second scale, the characteristic inverse cubic scaling law was not observed. All obtained scaling exponents assume values $\alpha > 3$, and the distribution tails are practically of the Gaussian type. The only exception is the EUR-GBP-JPY triangle, where, except for the 2-minute scale, the scaling exponents lie in the range of $\alpha \in (3; 4)$.

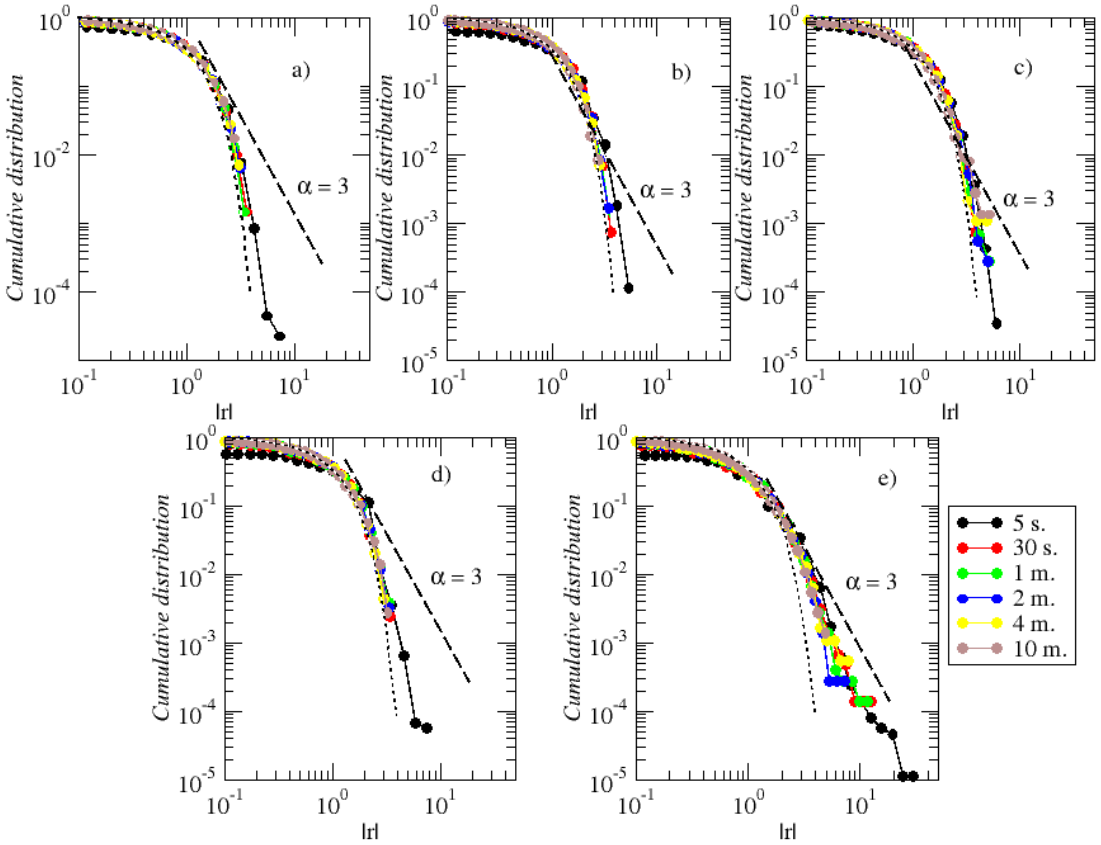

Figure 2.6 The cumulative distribution of triangular relation deviation returns, determined for the time scale from 5 sec. to 10 min. Panel a) shows AUD-EUR-JPY triangular relation deviations, b) EUR-USD-JPY, c) EUR-GBP-USD, d) AUD-USD-JPY, e) EUR-GBP-JPY.

A complete list of power exponents fitted to the distributions of triangular relation deviations is given in Table 2.2 .

Table 2.2 Power exponents for all triangular relation deviations sampled from 5 sec. to 10 min.

	AUD-JPY-USD	AUD-EUR-USD	AUD-EUR-JPY	GBP-USD-JPY	EUR-GBP-USD	EUR-GBP-JPY	EUR-USD-JPY
5s	5.77	4.36	7.28	7.44	8.71	4.07	6.66
30s	5.90	5.82	7.35	8.40	6.75	3.74	7.49
1m	5.67	5.20	8.14	5.53	8.12	3.91	6.91
2m	5.74	5.22	6.61	6.73	5.60	4.98	6.89
4m	7.01	4.31	7.63	7.11	6.60	3.90	5.44
10m	5.66	4.58	5.39	6.63	4.46	3.21	5.61

As has been mentioned above, *q-Gaussians* prove themselves very well in describing real financial data. For this reason, their usefulness for the description of Forex market data was verified. Due to the length of the series, the distributions were fitted only to 5-

second time series. Figure 2.7 shows a cumulative distribution for the AUD-EUR-USD triangle and the AUDUSD exchange rate with the *q-Gaussian* function fitted.

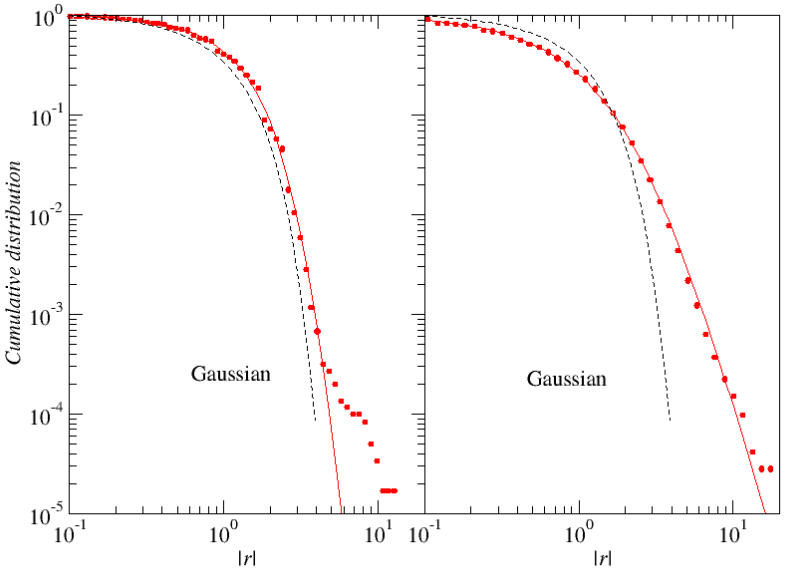

Figure 2.7 The distribution of fluctuations in 5-second AUD-EUR-USD triangular relation deviations (the left-had panel) and the AUDEUR exchange rate (the right-hand panel) along with the *q-Gaussian* function fitted to the distribution tail (the solid line). The dashed line indicates the Gaussian distribution.

These results are representative of all the examined triangles and currency pairs. It is clearly seen that, for the exchange rates, the distribution tails coincide very well with the value of the fitted function. In the case of the triangles, these distributions fairly well coincide with one another in the central part, whereas the tails are modelled not very accurately, and the empirical data are overrepresented with respect to the fitted distribution, more for the signal r^{Δ} than for the returns themselves. For EURAUD $q = 1.44$, which, according to formulae (**5**) and (**10**), means that the distribution tail scales itself approximately with the exponent $\alpha \sim 4.5$.

2.2.2 DJIA

In a similar way, 29 DJIA index companies and the index formed from those companies were analyzed. As has been shown in study [34], the index and its individual components are processes subject to the same probability distribution. The quotations covered the period 02.01.2008 –29.07.2011 and these were one-minute data. The index was made in a manner specific to this type of data, i.e. the share prices at a given time moment were added together, and then the returns were calculated from that signal.

Records formed through so-called share splits [11] likely to cause artificial fluctuations in data analysis, were removed from the time series. Figure 2.8 below represents the cumulative distributions of the one-minute normalized returns of all companies.

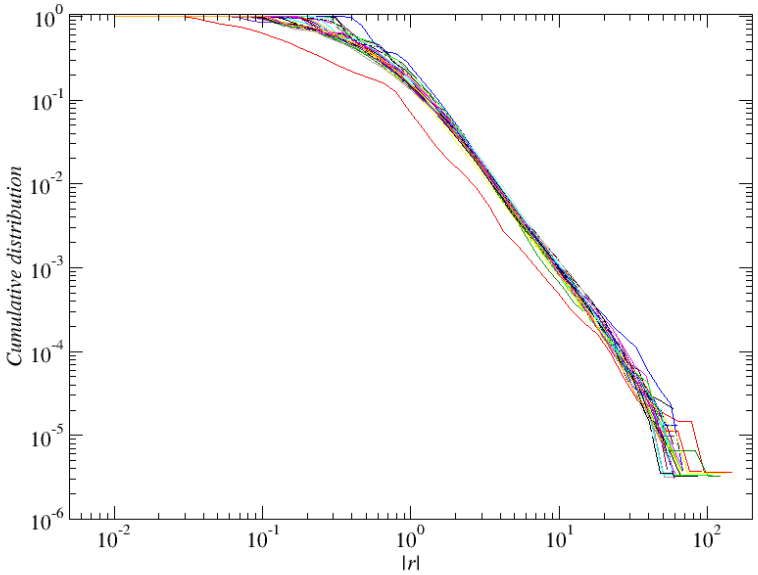

Figure 2.8 The normalized cumulative distributions of 1-minute returns of 29 DJIA index companies.

For the entire set, a very similar behaviour and power-law tail scaling can be noticed. The power-law behaviour of the tails is exhibited by distributions that have a finite variance and are either unstable distributions or stable distributions with the scaling exponent assuming values $\alpha \leq 2$. Table 2.4 summarizes all power exponents fitted to the distribution tails.

Table 2.3 The power exponents of power-law distribution fitting to the 1-minute cumulative distributions of DJIA index companies.

AA	2.52	CSCO	2.51	INTC	2.67	MMM	2.55	T	2.8
AIG	2.15	CVX	2.47	JNJ	2.59	MRK	2.45	UTX	2.52
AXP	2.62	DD	2.56	JPM	2.47	MSFT	2.53	VZ	2.48
BAC	2.22	GE	2.53	KFT	2.37	PFE	2.49	WMT	2.62
BA	2.59	HD	2.57	KO	2.67	PG	2.49	XOM	2.64
CAT	2.53	IBM	2.44	MCD	2.64	TRV	2.67		

Tail scaling along with the increase in the index time scale is illustrated in Figure 2.9. As can be noticed, the exponents take on values $\alpha \sim 2.5$ and lie beyond the stable Lèvy

[11] A share split is a reduction in the value of a given share, with the current level of the company's share capital being maintained. Some companies, wishing to maintain the prices of their shares in a range chosen by them, either emit new shares or split the share price, while retaining the share capital.

region and, with increasing times scale, they come ever closer to the Gaussian distribution, which is consistent with the Central Limit Theorem.

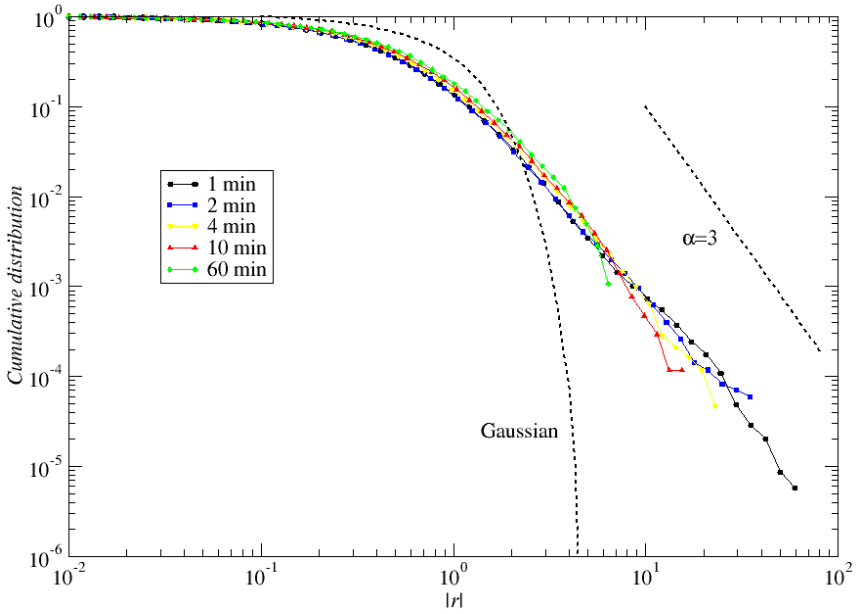

Figure 2.9 The cumulative distribution of the DJIA index for five time scales.

Similarly as for the FX market, an attempt was made to fit the *q-Gaussians* to the distribution of empirical data depending on the parameter q. As the representative set, four companies from different industries were chosen: AA – American Airlines (the aircraft industry), GE – General Electric (conglomerate), MSFT – Microsoft (IT), VZ – Verizon (telecommunications). The fitting was prepared for three time scales $\Delta t \in \{1,10,60\}$ minutes, and its result is shown in Figure 2.10. The obtained result seems to be very good, as the central part of the distribution and a considerable part of its tail coincide perfectly with the theoretical values. Only very rare events occurring at the very end of the distribution diverge slightly. In all cases, with the increase in the time scale, the values of parameter q decrease. This could have been expected from the asymptotic form of the *q-Gaussian* (**10**) and the previously observed fact of the convergence of the distribution tails to the Gaussian distribution with the increase in Δt.

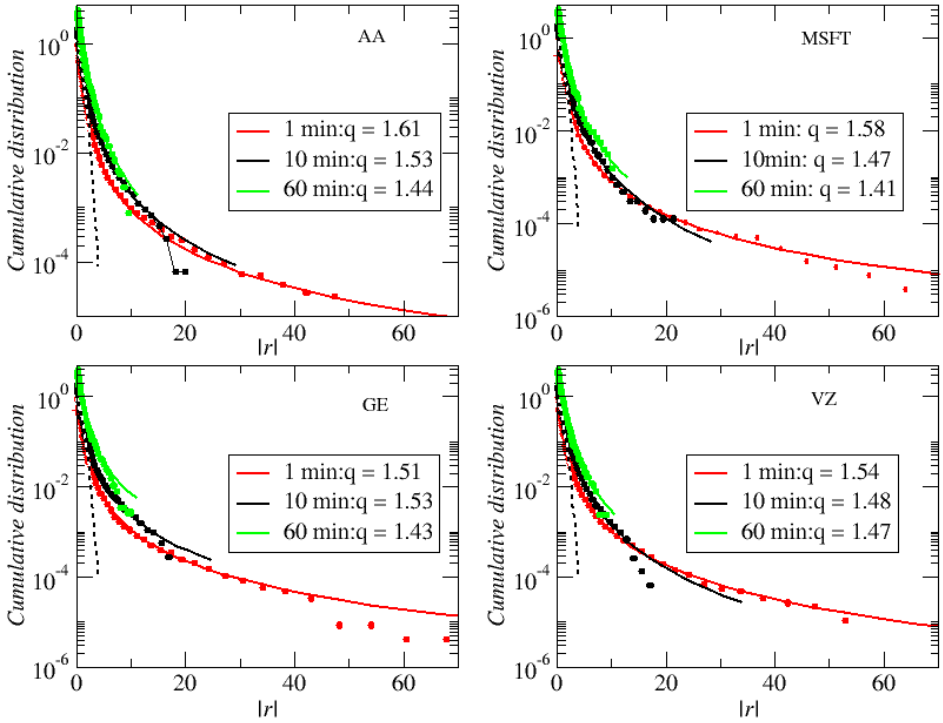

Figure 2.10 The cumulative distributions of the returns of four companies with the *q-Gaussian* function fitted to the distribution tail.

2.3 Autocorrelations and time cross-correlations

Cross-correlations occurring in financial data provide an extremely important element of the analysis of their dynamics. They are part of stylized facts which, in this case, make the function of autocorrelation of the financial time series drop to zero after several minutes [8; 35], while for the volatility (the module of the returns under examination*)* this function remains positive over a considerable period of several weeks [36]. The autocorrelation (self-correlation) function $\rho_x(\tau)$ for function $f(t)$ and shift τ is defined as follows:

$$\rho_x(\tau) = \langle f(t+\tau) f(t) \rangle. \tag{16}$$

With the use of this function, we can determine the effect of the current signal value on the signal value in the future. The function may assume either positive values (when the changes between analyzed values in the series follow one another in the same direction) or negative values (the changes between series values follow one another in the opposite directions) and the maximum falls on the zero-shift point. For random signals, the function value falls immediately to zero. This function lends itself perfectly to the examination of linear relationships in an examined series.

The next stage in the investigation of FX market dynamics will be examining the interrelation between financial data. This is a very important analysis that is widely used in finances, because it leads to the optimization of the investment portfolio and assists in managing the risk. In the analysis of time series and processing signals, the cross-correlation (intercorrelation) between two series is described by normalized covariance that means the degree in which the two series follow one another, assuming that one of these series is shifted by τ relative to the other. For two series X_t and Y_t with a length T, this relationship is defined by a cross correlation function which is given by this formula:

$$\rho_{xy}(\tau) = \frac{\sum_{t=1}^{T}(X(t) - \mu_X)(Y(t+\tau) - \mu_Y)}{\sigma_X \sigma_Y}, \tag{17}$$

where σ_X and σ_Y is the standard deviation of the respective series, while τ is the shift of the series Y relative to X. In the case, where $\rho_{xy}(\tau = 0)$, this is the Pearson correlation coefficient that assumes values in the interval $-1 \leq \rho_{xy} \leq 1$. In the case, when $\rho_{xy} = 1$, we have a perfect correlation between the signal, where $\rho_{xy} = -1$ means anti-correlation and $\rho_{xy} = 0$ lack of correlation between the signals. Using ρ_{xy}, the level of linear relationship between the signals X_t and Y_t is determined. From this point on in this study, the terms cross-correlation and intercorrelation will be used interchangeably. Returns, as well as their modules, were used for analysis in view of the fact that the autocorrelation of volatility may exhibit a different nature.

2.3.1 DJIA

The results obtained for the returns of the DJIA index companies discussed in the previous section are shown in Figure 2.11. They clearly indicate that no long-range correlations occur in financial signals, and only for $\tau = 1$ min. can we speak of a weak anti-correlation of the examined time series. A possible cause of this phenomenon is the occurrence of zeros in the values of the returns on short time scales, being comparable to, or shorter than the frequency of making transactions. Negative autocorrelation is therefore a computational artefact. The behaviour of the entire index is similar, except that in the first minute the autocorrelation level is less negative compared to all component companies and takes on a value close to the noise value. Looking at the process governing the share price from this perspective, one might arrive at the conclusion that it is of the Brownian type (has no memory).

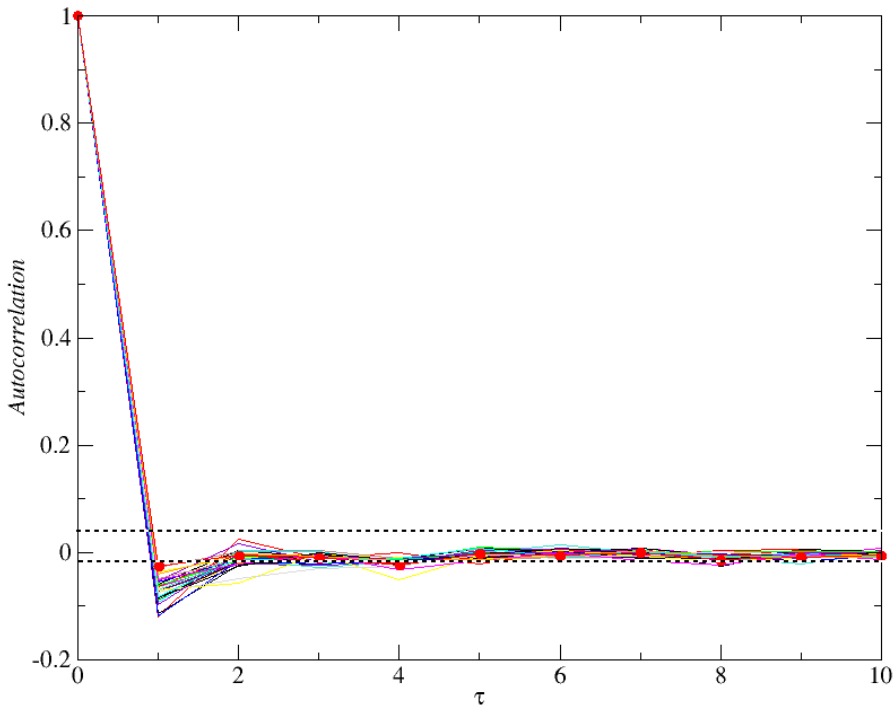

Figure 2.11 Autocorrelation function determined for 1-minute returns of 29 DJIA index companies. The index autocorrelation function is marked with the red line. The dashed lines denote the noise level.

In addition to examining the autocorrelation of the returns, their modules, i.e. *volatility*, was subjected to analysis. This quantity directly takes advantage of the amount of information coming in to the stock exchange. Panel a) in Figure 2.12 represents this value on a linear scale for four companies: AIG, GE, MSFT and CSCO. We can clearly see a daily trend in the form of the regularity of a function, in which characteristic peaks occur exactly every 390 minutes. Panel b) in the above-mentioned figure shows the same data, presented on the log-log scale, along with a power-law function fitted to them. It can be seen that the autocorrelation function of volatility is long-range in character, in contrast to the examined returns. The power-law function fitted to the empirical data (given by formula (5)) has an exponent $\alpha = 0.16$ and seems to adequately describe the character of the fading of correlation in the signal. The same behaviour is visible in Panel c), which shows the intercorrelation function determined for those companies and fitted to the data with a power-law function with power exponent $\alpha = 0.11$.

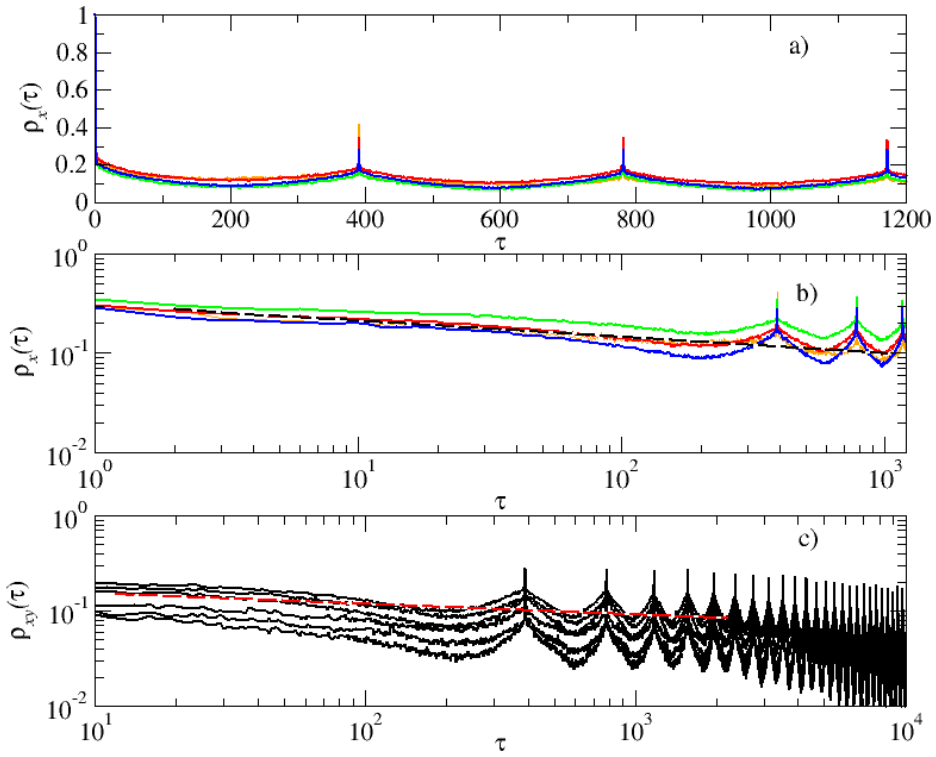

Figure 2.12 The autocorrelation of volatility. Panels a) and b) show the function of autocorrelation of the volatility of four companies. The company AIG is marked in orange colour, GE in red, CSCO in green and MSFT in blue. Panel c) represents the cross-correlation determined among these companies. The dashed line denotes the power-law fit. A daily trend with characteristic peaks every 390 minutes (1 trading day) is clearly visible.

Using formula (**17**), the Pearson coefficient among all companies returns has been determined, which yields 406 possible combinations. All of the examined companies are positively correlated with one another. Two thresholds of correlation values are very well visible, i.e. at $\rho_{xy}(0) \approx 0.4$ and $\rho_{xy}(0) \approx 0.5$, in the vicinity of which more than half of the obtained values lie. This is indicative of a moderate strength of those correlations. In the case of 41 combinations of companies, there is no cross-correlation between signals, and two of them are correlated with each other at a level of $\rho_{xy}(0) = 0.7$. The mean value of cross-correlation in the examined set is $\overline{\rho_{xy}(0)} = 0.35$. When comparing the cross-correlation of companies from different sectors in Table 2.4, it can be noticed that its mean value is the greatest for Materials sector companies, while the least for Services sector companies.

Table 2.4 The mean value of returns cross-correlation $\rho_{xy}(0)$, given by sectors.

Basic Materials	$\overline{\rho_{xy}(0)} \approx 0.38 \pm 0.04$

Consumer Goods	$\overline{\rho_{xy}(0)} \approx 0.36 \pm 0.02$
Financial	$\overline{\rho_{xy}(0)} \approx 0.32 \pm 0.07$
Healthcare	$\overline{\rho_{xy}(0)} \approx 0.34 \pm 0.02$
Industrial Goods	$\overline{\rho_{xy}(0)} \approx 0.39 \pm 0.01$
Services	$\overline{\rho_{xy}(0)} \approx 0.28 \pm 0.06$
Technology	$\overline{\rho_{xy}(0)} \approx 0.35 \pm 0.08$

2.3.2 **FX market**

A similar analysis was made on data from the foreign exchange market, where the strength of correlation between each of the 9 currency pairs was determined. Figure 2.13 represents autocorrelation functions, as dependent on the delay τ, for all currency pairs (the top diagram), as well as triangular relation deviations (the bottom diagram). Just like for financial markets, for FX data this function will assume values at the zero level, practically from $\tau = 2$. Similarly as in study [31], a characteristic feature of almost all examined currency pairs is the existence of a correlation low of negative correlations for $\tau = 1$, being indicative of an (apparent) anti-persistence of the series for short time scales. The only currency pair that has no low for $\tau = 1$ is GPBUSD, which, apart from the EURUSD pair (the minimum low), is the most frequently traded currency pair. Therefore, the series of their returns have the fewest zeros, and it is the presence of zeros that determines whether the occurring persistence is apparent or not. If price-changing transactions occur rarely, then returns different from zero seldom happen in the signal, as isolated peaks surrounded by zeros on both sides. That is, effectively for the algorithm, the large (real) rate of return is followed by a small one (or zero). This resembles exactly anti-persistence, but is not it. Therefore, the negative anti-correlation determined for the FX data is a computational artefact.

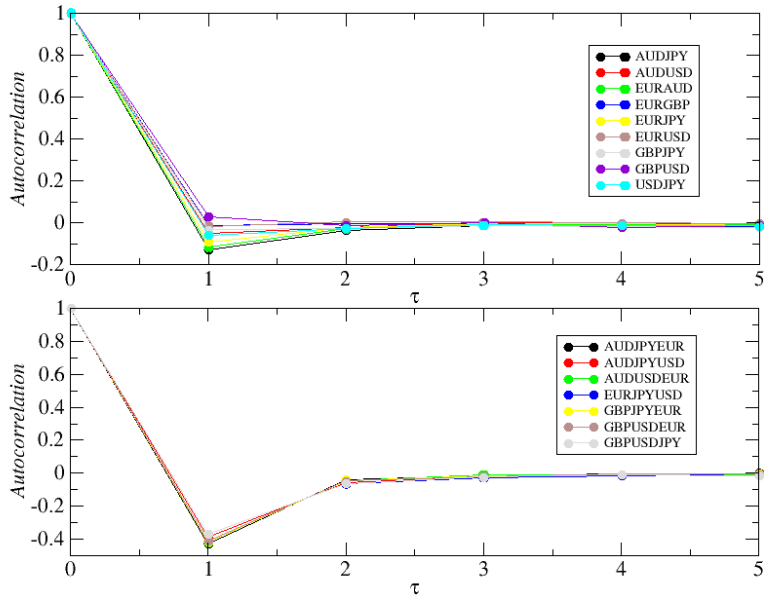

Figure 2.13 The autocorrelation function of 5-second returns corresponding to all currency pairs (the top panel) and triangular relation deviations (the bottom panel).

This implies that the process memory is very short, which is characteristic of Brownian processes.

The volatility autocorrelation functions can be seen in Figures 2.14 and 2.15.

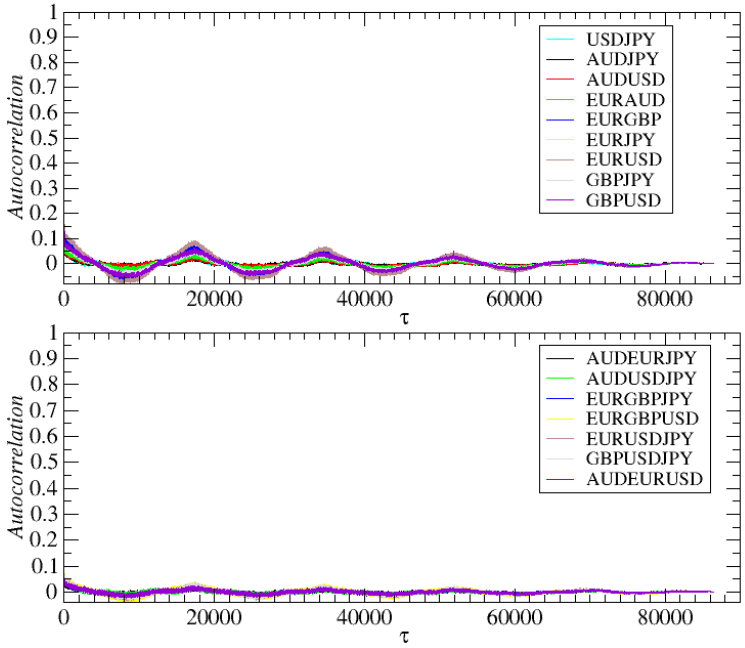

Figure 2.14 The top panel: volatility autocorrelation of 5-second data. Peaks constituting a daily trend are clearly visible. The bottom panel: the volatility of triangular relation deviations for 5-second data.

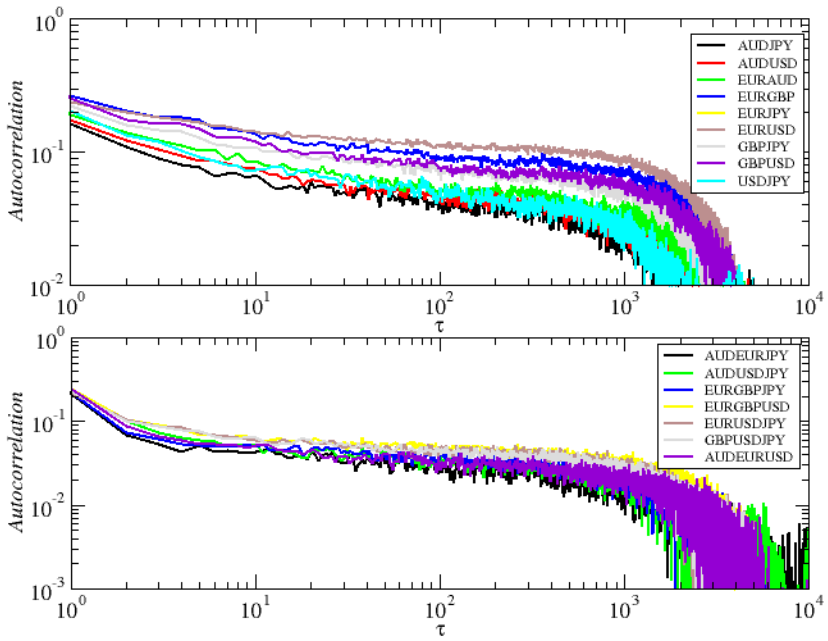

Figure 2.15 The volatility autocorrelation function on the log-log scale of 5-second time series corresponding to all currency pairs (the top panel) and triangular relation deviations (the bottom panel).

When comparing the volatility with the returns it can be seen that the autocorrelation has a long decay time. In approximation, this is a power-law relationship $c(\tau) \sim x^{-\alpha}$, and the power exponent assumes the values $\alpha = 0.32$ and $\alpha = 0.15$, respectively, for returns and the triangular relation deviations.

Using the converse exchange rates ($\frac{EUR}{USD} = \left(\frac{USD}{EUR}\right)^{-1}$), it was possible to examine 24 currency pairs "in the triangle" and 12 currency pairs "beyond the triangle" for cross-correlations. A complete list of combinations is provided in Table 2.5.

Table 2.5 24 exchange rate pairs "in the triangle" (black colour) and 12 exchange rate pairs "beyond the triangle" (red colour).

AUDJPY/EURAUD	**AUDJPY/JPYUSD**	**AUDJPY/USDAUD**
AUDUSD/USDEUR	**AUDUSD/USDJPY**	**EURAUD/GBPEUR**
EURAUD/AUDUSD	**EURAUD/USDEUR**	**AUDJPY/GBPJPY**
EURJPY/AUDEUR	**EURJPY/JPYAUD**	**GBPUSD/USDJPY**
GBPUSD/EURGBP	**GBPUSD/JPYGBP**	**GBPUSD/USDEUR**
GBPJPY/JPYEUR	**GBPJPY/JPYUSD**	**GBPJPY/EURGBP**
EURJPY/USDEUR	**EURJPY/JPYUSD**	**EURGBP/JPYEUR**
EURUSD/GBPEUR	**EURUSD/USDJPY**	**AUDUSD/USDGBP**
AUDJPY/EURGBP	**AUDJPY/GBPUSD**	**AUDUSD/EURGBP**
EURJPY/GBPUSD	**EURJPY/AUDUSD**	**EURUSD/AUDJPY**
GBPUSD/EURAUD	**GBPJPY/AUDUSD**	**EURAUD/GBPJPY**

EURAUD/USDJPY EURGBP/USDJPY EURUSD/GBPJPY

The employed denotation "currencies inside the triangle" means a simplified version of triangular arbitrage, where one single currency will occur between the examined exchange rates. This name will be used later on in this study. In the case of exchange rates "outside the triangle" we have four independent currencies. Moreover, it can be expected that the condition $\frac{\cancel{AUD}}{JPY} * \frac{EUR}{\cancel{AUD}} = \frac{EUR}{JPY}$ will be satisfied for currencies inside the triangle, which enables, using two different exchange rates $\frac{AUD}{JPY}$ and $\frac{EUR}{AUD}$ having a common currency AUD, the determination of the exchange *cross-rate* $\frac{EUR}{JPY}$. For currencies outside the triangle, as in the case of $\frac{AUD}{JPY} * \frac{EUR}{GBP}$, it is not possible to calculate the exchange cross-rate, because there is no common currency between the exchange rates.

The relation between exchange rates was examined using formula (**17**). Except for the AUDJPY/GBPJPY pair, all currencies inside the triangle are negatively correlated, while those outside the triangle, correlated positively (except for the pairs EURAUD/USDJPY and EURGBP/USDJPY). The most negatively correlated is the EURUSD/GBPEUR pair with a correlation value of $\rho_{xy}(0) = -0.63$, AUDJPY/GBPJPY is positively correlated with $\rho_{xy}(0) = 0.37$, while the GBPJPY/AUDUSD pair is almost uncorrelated, with $\rho_{xy}(0) = 0.06$ (the noise level is 0.01). All the results for $\rho_{xy}(0)$ are shown in the inner panel in Figure 2.16. The mean value of the correlation coefficient for the pairs inside the triangle is $\overline{\rho_{xy}(0)} = -0.36$, which means a negative intercorrelation, while for the pairs outside the triangle it is $\overline{\rho_{xy}}(0) = 0.19$, which would suggest virtually no or very little intercorrelation. This value is disturbed by the aforementioned pairs: EURAUD/USDJPY with the coefficient $\overline{\rho_{xy}}(0) = -0.08$ and EURGBP/USDJPY with $\overline{\rho_{xy}}(0) = -0.13$. The EURUSD pair is the one which is strongest linked with the remaining exchange rates and has the mean correlation coefficient equal to $\overline{\rho_{xy}(0)} = -0.31$. This result could have been expected, as this pair is the world's most often traded currency pair with a 25% market share[12]. In turn, an exchange rate least connected with other exchange rates is the AUDJPY with

[12] www.reuters.com: access on 15.09.2015

$\overline{\rho_{xy}(0)} = -0.13$ being, at the same time, a pair with the smallest quota market share (approx. 0.5 %) among the most popular exchange rates.

In the next step, the cross-correlation between the volatility of all currency pairs from Table 2.5 was examined. This is represented in Figure 2.16 where the currencies inside the triangle are shown in black, while the currencies outside the triangle, in red.

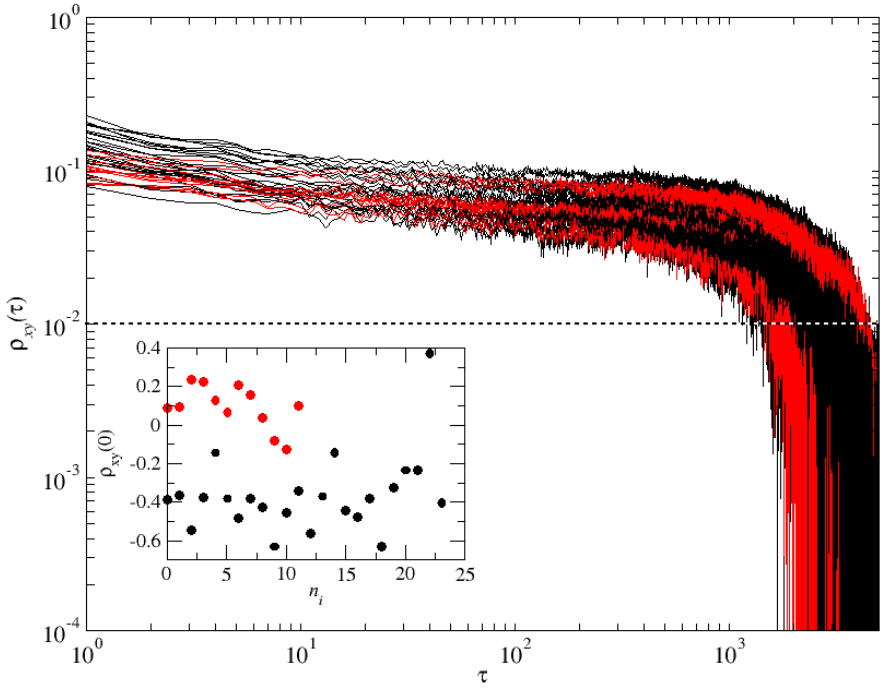

Figure 2.16 The cross-correlation function determined for the volatility of 5-second time series. The currencies inside the triangle are indicated with black colour, while the currencies outside the triangle, with red. The dashed line denotes the noise level. The inset shows the Pearson correlation coefficient $\rho_{xy}(0)$, determined for the returns.

At the volatility level, all the examined exchanges rates are cross-correlated for events distant as much as by $5 * 10^3$ (which gives almost 7 trading hours). For a delay of about 50 seconds, a slightly lower level of cross-correlation can be observed in the currencies outside the triangle, compared to the currencies inside the triangle. In the inset, the cross-correlation level for the returns is also indicated. Cross-correlation clustering on the negative value side for currencies inside the triangle is clearly seen. The only exception is the AUDJPY/GBPJPY currency pair.

Figure 2.17 illustrates cross-correlations between triangular relation deviations. Deviation pairs, for which 2 common currencies exist, such as AUD-USD-JPY/AUD-USD-EUR (with the common currencies AUD and USD), are marked in red colour, while those that have one common currency, e.g. AUD-USD-JPY/GBP-JPY-EUR (the

common currency being JPY), are marked in black. It is clearly visible that, in the case of two common currencies between deviations, there is a cross-correlation for a delay of up to $\tau = 3$ (equal to 15 sec.), thereupon it assumes the level of two uncorrelated noises. In the case of one common currency, the cross-correlation between two series is immediately zero. This results from the fact that for currency triangle pairs, two information transferring currencies are needed in order to be able to determine the cross-correlation between them.

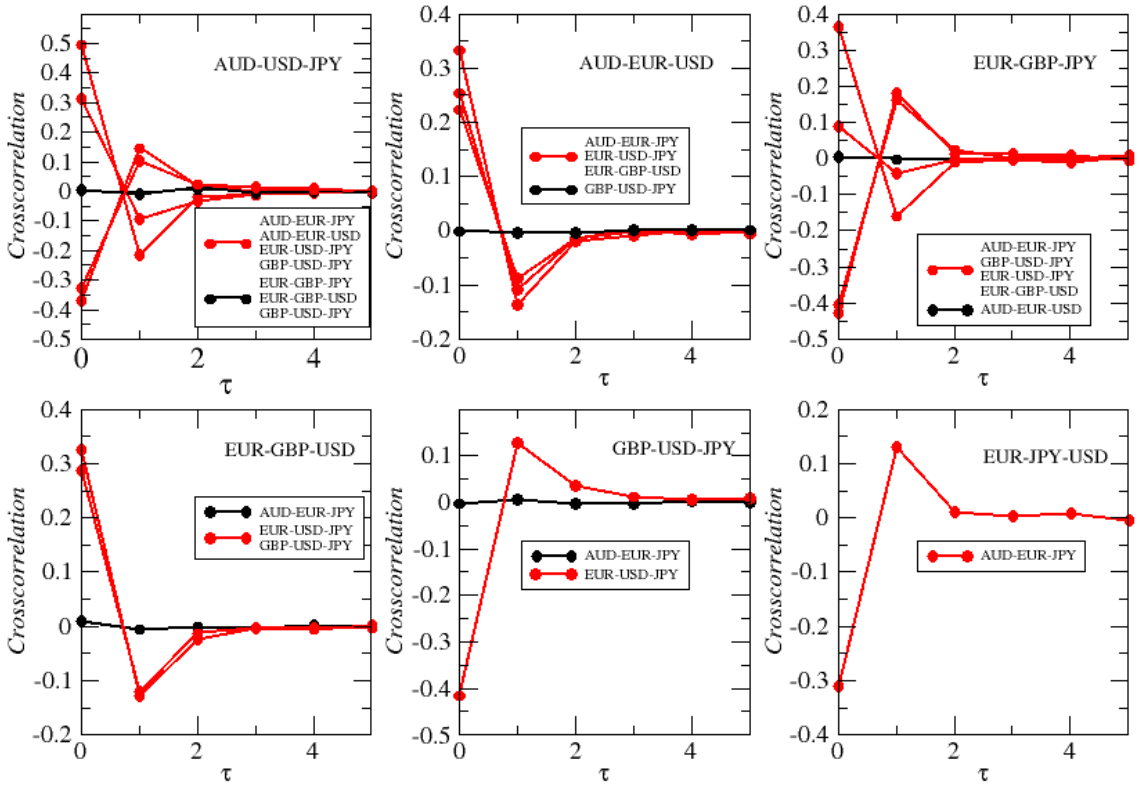

Figure 2.17 The function of cross-correlation between triangular relation deviations. Correlations between triangles having 2 common currencies are marked in red colour.

3 Fractal and multifractal formalism

In the classical Euclidean geometry, we generally deal with regular objects, that are the ones free from roughness (such as curves, polygon sides, spheres or polyhedron walls). As can be easily noticed, the world around us is not so regular and simple in description. The attempt to describe the shape of some real objects (e.g. clouds) using traditional geometry is sometimes a very tedious, if not impossible job. In the 1970s, Benoit Mandelbrot made some groundbreaking discoveries in the study of irregular objects and created fractal geometry [37; 38]. From that moment, fractals have changed the way we view the world. The word 'fractal' itself is derived from Latin and means 'fractional' or 'broken'. When examining financial data charts, Mandelbrot noticed that they are similar to each other in different time scales. This is a feature that is characteristic of fractals and is called self-similarity. An object having this feature is similar to itself regardless of the scale, looking the same when observed either from close up or from a distance. Taking a part of an object and respectively scaling it, we will get an object similar to the whole. Using fractal geometry, we can examine systems, from the complex and irregular structure of which, an ordered and hierarchical picture will emerge. Fractals are present almost everywhere. The shapes of a snow flake, shoreline, trees, human lungs or broccoli – these are just a few examples of natural, random fractals, in which self-similarity is statistical in character. One flake may differ in details from the other, but in terms of general characteristics, they are the same. And it is this type of systems that can be most fully described using fractal geometry. Apart from them, there are plenty of mathematical (deterministic) fractals generated by recurrent formulas, which, unlike natural fractals, make up ideal structures. At each iteration, scaling takes place while preserving the characteristic scale, which is not a rule for random fractals. Very commonly, random fractals are used in video games and computer graphics, where, with the help of simple mathematical rules, complex structures are created, including most often some landscape elements, such as trees or mountains, at a much lower computing cost. Examples of such fractals are shown in Figure 3.1.

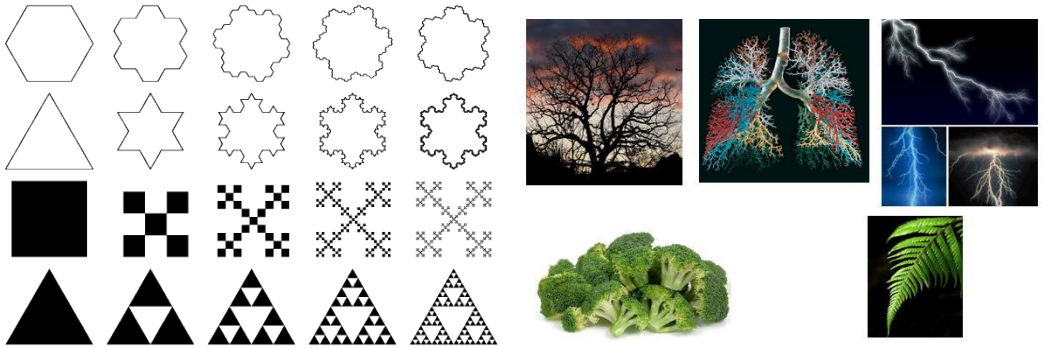

Figure 3.1 Examples of simple mathematical fractals (on the left). The second top structure is a Koch curve, while the first bottom one, a Sierpinski triangle. On the right, there are examples of natural fractals.

Also the time series of financial data have a fractal nature. Looking at the time series of data recorded every second or every minute, it is very difficult to distinguish which series is on what scale (Figure 3.2).

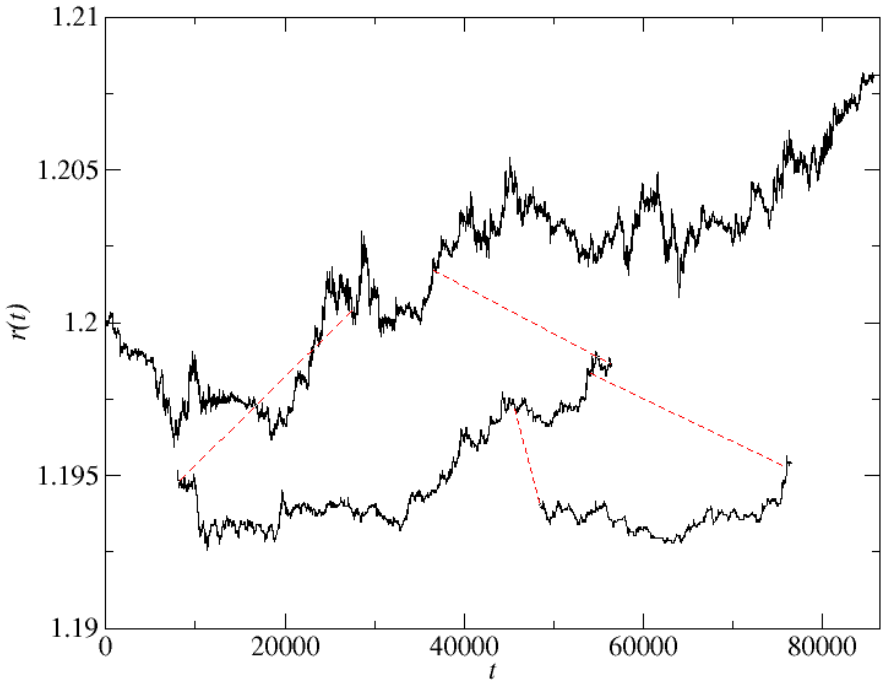

Figure 3.2 Self-similarity of a time series (EUR/USD).

Moreover, the record of share price variations (assuming its continuity) is not one-dimensional (it is not an ordinary curve) or two-dimensional (its surface has a zero area), which, according to the definition below, is sufficient to refer it as a fractal. In addition, unlike classic fractal forms, which are spatially self-similar, time series are self-similar with respect to time.

There is no single definition of the fractal. One of the definitions says that the fractal is a set that exhibits a self-similarity property, at least in an approximate sense, if not exactly. According to another definition, fractals are objects having a non-trivial structure, which cannot be described using the Euclidean geometry. Or, the fractal is an object whose dimension is not an integer.

We assume that the set F is called fractal, if it shows the following features [39]:

1) F has a structure such, that its details can be seen on an arbitrarily small scale;

2) F is too irregular to be described using traditional geometry;

3) F is self-similar (in a statistical sense);

4) fractal dimension of F is usually greater than its topological dimension;

5) in the case of mathematical fractals, it is defined simply, most fully in the form of a recurrent procedure.

Below, basic fractal geometry concepts necessary for the description of fractals will be presented.

3.1 General properties of fractal sets

Fractals are scale free objects, i.e. they do not have a single, distinguished scale. They exhibit so complex structure that no direct and accurate measurement of their length, surface area or volume is possible. During the measurement of these values, it is normally indicated how fast they increase, because when made with increasingly great accuracy, the measurement will lead to divergent results. It should be noted that not every self-similar object is a fractal. Objects not being fractals include, e.g., an interval, a square or a cube. If the linear dimensions of those objects are scaled by an arbitrary number, then their respective surface area and volume will be changed proportionally. In the case of fractals, there is no such proportionality.

Scaling is, therefore, a superposition of several transformations in Euclidean geometry. The transformation $S\colon \mathbb{R}^n \to \mathbb{R}^n$ is called *isometric,* if it retains the distance, that is $|S(x) - S(y)| = |x - y|$ for every $x, y \in \mathbb{R}^n$. An example of isometric transformation is *translation*, or the parallel shift *S(x)* by a vector $a\colon S(x) = x + a$, *rotation* about a point a, $|S(x) - a| = |x - a|$ and *reflection*. *Similarity* transformation by a value c is a transformation, where the relationship $|S(x) - S(y)| = c|x - y|$ occurs for every $c > 0$ and $x, y \in \mathbb{R}^n$, whereby a geometric figure subjected to this transformation turns into the identical figure, but with all its dimensions multiplied by

the value *c*. This is an isotropic transformation, i.e. the transformed object is scaled by the same value in all directions.

A special case of self-similarity is *self-affinity*. The affine transformation S is a transformation in the form of:

$$S(x) = T(x) + a, \tag{18}$$

where a is a point in $\mathbb{R}^n$ while T is singular linear transformation. A transformation T is non-singular, if $T(x) = 0 \Leftrightarrow x = 0$ and T if the following relationship occurs:

$$T(x + y) = T(x) + T(y) \text{ i } T(\lambda x) = \lambda T(x), \tag{19}$$

where $x, y \in \mathbb{R}$ and $\lambda \in \mathbb{R}$.

Affine transformation can be understood as a superposition of the operations of scaling, translation, rotation and reflection, not necessarily by the same value in all directions. So, this is an anisotropic transformation, in which scaling depends on the orientation in space. An example of affine transformation is shown in Figure 3.3 below.

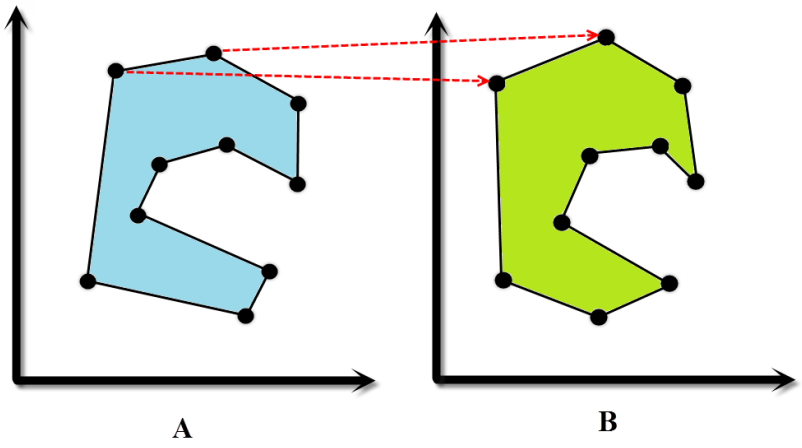

Figure 3.3 A schematic diagram of affine transformation.

Self-affine are the time series of the returns (Figure 3.2), and for the description of the degree of irregularity of function $f(x)$ that can represent such a time series, the Hurst exponent H is used. In the case, where $f(x)$ is self-affine, it satisfies the following relationships:

$$f(x) \cong \lambda^H f(\lambda x). \tag{20}$$

It follows from formula (**20**) that for $H = 1$, $f(x)$ is self-similar. When the exponent H is dependent on time *t*, then the Hölder exponent (function) [40] is used to define the pointwise scaling. A function f is called the Hölder function with an exponent h, if the following relationship is satisfied:

$$|f(x) - P_n(x - x_0)| \leq c|x - x_0|^h, \tag{21}$$

for a constant $c > 0$ and a polynomial P_n of order n. Assuming that $n < h$, the Hölder exponent $h(x_0)$ is defined as the greatest value in the set h which satisfies the relationship (**21**): $h(x_0) = \sup\{s: f \in c^s(x_0)\}$. More information about the Hurst exponent is provided later on in this study.

If one quantity depends on the other in a power-law manner, then we call this scaling. In the case of time series, scaling results directly from self-similarity forming the basis for fractal geometry. Self-similarity, or the lack of a characteristic scale, is closely related to power laws. Systems that can be described using power-law similarity distributions:

$$P(x) = x^{-\alpha}, \tag{22}$$

are scale free objects, such as fractals. By multiplying the argument of function (**22**) by a constant factor *c*, we proportionally scale the value of the function itself, i.e.:

$$P(cx) = (cx)^{-\alpha} = c^{-\alpha}P(x). \tag{23}$$

In this case, scaling involves simply multiplying the function by the coefficients $c^{-\alpha}$. Power-law distributions on a logarithmic-logarithmic scale have the form of a straight line (which is very important in the case of the methods of determining higher-order correlations: sections 4.1 and 4.2). If we find the logarithm of the both sides of $P(x)$ in formula (**23**), then we will get: $\ln P(x) = -\alpha \ln x + \ln c$. Representing this is the coordinates $x' = \ln x'$ and $y' = \ln P(x)$, we obtain the linear relationship $y' = ax' + b$, where $a = -\alpha$ and $b = \ln c$.

It is common practice to use the testing of a function for linearity on the logarithmic scale as the criterion for judging whether a given distribution is a power-law distribution or not. Unfortunately, this may lead to erroneous results, because it often turns out in practice that the size of a test sample strongly influences the occurrence of rare events in that sample (by the very nature of power-law distributions). To circumvent this limitation, the probability distribution for actual data is made on logarithmic cells. Thus, for small but frequent events, we have a large number of cells to be populated, while for large events we will get a small number of cells. Moreover, by being presented on the log-log scale, the power-law distribution diagram will gain in legibility owing to the fact that the logarithmic cells will be equally distant from one another. Power-law distributions are commonly used and observed in physics e.g. noise $1/f$, phase transitions [41], percolation [41], the Insing model [42], biology [43] , economics (Pareto's law [9]) or social sciences (Zipf's law [44]).

Fractality is also a feature of many complex systems, which are understood as systems with a very large number of elements interacting nonlinearly with one another, and in the dynamics of which emergent phenomena can be observed. As has already been mentioned, the feature of such phenomena is that they cannot be accurately foreseen at the macroscopic level based on the knowledge of the structure and interactions among the microscopic components of the system. A slight change in one component has far-reaching consequences. Therefore, the complete description of a system requires all its characteristics to be considered at all organization levels [45; 46]. Complex systems can be observed in many fields of science from medicine, through physics and chemistry to economics and social sciences. Even though those areas are sometimes distant from one another, the behaviour of these systems can be described using the same principles, and the systems themselves exhibit a number of similar features, such as the aforementioned emergence, self-organization or hierarchy. As will be explained later in this study, all of these features are observed in the case of financial markets, where, additionally, the time series of share prices and foreign exchange rates exhibit fractal properties.

3.2 Fractal dimensions

For the quantitative description of scaling and self-affinity, the notion of fractal dimension applies, which provides information on how an object fills the space. Classically, the dimension is defined as the number of degrees of freedom needed for describing an arbitrary point inside that object. In that case, it is the topological dimension D_t. The topological dimension of an interval is one, because only one number is needed to describe the position of an arbitrary point on it. In turn, for the description of a point in a plane, two points are needed, from which it is inferred that the topological dimension of a plane is $D_t = 2$, while for the description of the position of a point in the three-dimensional space, three numbers are required, hence $D_t = 3$. The topological dimension of fractal objects is also an integer; so, it is not the appropriate measure for the description of those objects.

A dimension most often used for fractals is the self-similarity dimension D_s defined by the formula:

$$D_s = \frac{\log(N)}{\log(\frac{1}{s})}, \tag{24}$$

where s is the scale factor, and N is the number of parts into which the object is divided. For geometric objects, such as an interval, a square or cube, the scale factor can

be chosen arbitrarily; but for fractals, it is specific for a given fractal. For example, the Sierpinski triangle scaled with $s = 1/2$ has a fractal dimension of $D_s = \frac{\log(3)}{\log(2)} = 1.5849$, which means that, in the conventional meaning of the Euclidean dimension, this is an object between a straight line and a plane, though it has a topological dimension of 2. An equally common fractal dimension [39] is the box-counting dimension D_b which is defined by the formula:

$$D_b = \lim_{\delta \to 0} \frac{\log N_\delta(F)}{-\log \delta}, \tag{25}$$

where F is a set in $\mathbb{R}^n$ and $N_\delta(F)$ is the smallest number of sets with the diameter δ covering the set F. To calculate the box-counting dimension of a two-dimensional set F, the set on a grid, whose elements have the size δ and count how many elements are needed to completely cover the set F.

For example, the Koch curve is formed by dividing an interval into 3 parts and substituting the median with two intervals, each with a length equal to 1/3 of the initial length, inclined at 60 degrees to one another. Therefore, at the first iteration, we cover the curve with three squares $N_\delta(F) = 3$, each with a side length of $\delta = \frac{1}{3}$.

At the second iteration, we repeat the operation, reducing the square side length to $\delta = \frac{1}{9}$, as a result of which we will have $N_\delta(F) = 12$ squares covering the curve. At the *n*-th iteration, we have $N_\delta(F) = 3 * 4^{n-1}$ and $\delta = \frac{1}{3^n}$,which, after substituting in the formula, will give $D_b \approx 1.26186$. In practice, using the box-counting dimension will be burdened with a large error and may lead to fairly significant distortions as to the actual dimension of the fractal [47]. However, the oldest and, at the same time, the most important dimension is *Hausdorff dimension* [39], while the fractal definitions quoted above are only its simplification. To be able to define it, we must start from the definition of *Hausdorff measure*. Let us assume that U is a non-empty subset of an *n*-dimensional space $\mathbb{R}^n$ with a diameter defined as $|U| = \sup\{|x - y|: x, y \in U\}$ (the greatest distance between any pair of points in U) not greater than δ. If $\{U_i\}$ is a countable set of sets U_i with a diameter of at most δ, which covers the F: $F \subset \bigcup_{i=1}^{\infty} U_i$ i $0 \leq |U_i| \leq \delta$, then we say that $\{U_i\}$ is a δ–cover of F. If F is a subset of $\mathbb{R}^n$ and s is a non-negative number, then for $\delta > 0, \mathcal{H}_\delta^s = \inf\{\sum_{i=1}^{\infty} |U_i|^s: \{U_i\}\}$ is the δ −cover of the set *F*.

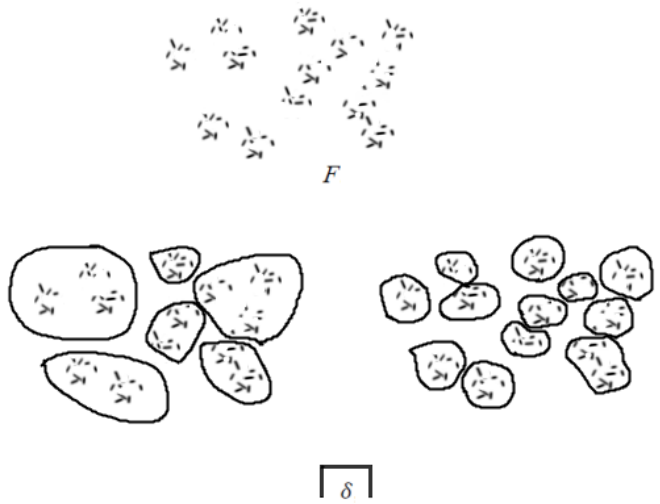

Figure 3.4 Covering of the set with subsets F of a different diameter δ.

We take all coverings of the set F by sets of a diameter δ and search for the one with the smallest sum of the s-degree powers of those diameters. The idea of covering a set with different subsets is shown graphically in Figure 3.4. As δ decreases, the diameters of the sets of each of the covers decrease, whereby we can see more and more details of the set F. At the limit $\delta \to 0$, the measure $\mathcal{H}^s(F)$ is referred to as the s-dimensional Hausdorff measure of the set:

$$\mathcal{H}^s(F) = \lim_{\delta \to 0} \mathcal{H}_\delta^s(F). \tag{26}$$

Scaling of a length, surface or volume is known very well. By scaling them up by a constant λ, we obtain a curve λ two times as long, a surface λ^2 times as large, and a volume λ^3 times as great; the s-dimensional Hausdorff measure scales itself as λ^s. Such scaling properties provide a foundation of the fractal theory.

There is a critical value s, above or below which $\mathcal{H}^s(F)$ is 0 or ∞, and which is called the Hausdorff dimension D_H [39]:

$$\mathcal{H}^s(F) = \begin{cases} \infty & 0 \le s < D_H \\ 0 & s > D_H \end{cases}. \tag{27}$$

The relationship between the introduced dimensions is as follows: $D_S \le D_H \le D_b$. In practice, it is not always the case that dimensions calculated by different methods are the same.

To get the idea of what the fractality of time series is all about, we should have a look at Figure 2.1. We can see that individual returns are not uniform, but there are places where the volatility is greater or less. This reflects the correlation between

market participants that determine the price. How the space is filled by an object depends on the mechanism that forms it. In the case of returns, such driving forces are investors who price the value of assets based on micro- and macroeconomic factors. The time series of financial assets are not exactly random contrary to a time series of the Brownian motion, although the box-counting dimension of these two time series is comparable. The box-counting dimension for the time series in Figure 2.1 is $D_b = 1.43$.

3.3 Multifractals and the singularity spectrum

The above-mentioned measures lend themselves perfectly to the determination of the dimensions of single fractals, but for multifractals (being a combination of individual fractal structures with their own fractal dimension and local scaling properties), other characteristics are needed, which will make it possible to describe the set [48]. Therefore, we introduce a new dimension, D. For a fractal object with a measure μ, the dimension D shows what the increase in mass $\mu(B_x(\epsilon))$ will be with the increase in ϵ:

$$\mu(B_x(\epsilon)) = \int_{B_x(\epsilon)} d\,\mu(y) \sim \epsilon^D, \tag{28}$$

where $B_x(\epsilon)$ is a sphere with a radius ϵ and with the centre x. Generally, fractal measures exhibit multifractal properties, if they scale diversely, depending on the position. This lead to local scaling:

$$\mu(B_x(\epsilon)) \sim \epsilon^{\alpha(x)}, \tag{29}$$

where $\alpha(x)$ is the strength of singularity (*singularity exponent*) of the measure μ at point x. The greater the value $\alpha(x)$, the more uniform the measure is in the vicinity of x and the weaker the singularity. The *singularity spectrum* $f(\alpha)$ describes the distribution of the exponents $\alpha(x)$. If the measure μ is covered with spheres with a size ϵ, then the number of such spheres scales itself, for a given α, in the following manner:

$$N_\alpha(\epsilon) \sim \epsilon^{-f(\alpha)}. \tag{30}$$

At the limit $\epsilon \to 0^+$, $f(\alpha)$ is the Hausdorff dimension of the set of points x such, that $\alpha(x) = \alpha$: $f(\alpha) = d_H(\{x \in \text{Supp}\,\mu, \alpha(x) = \alpha\})$, or the set of the Hausdorff dimensions of points described by the same Hölder exponent. Now we have a distinction between two classes of singular measures. Homogeneous (monofractal) measures are characterized by a singularity spectrum in the form of a single point $(\alpha_0, f(\alpha_0))$, which means that only one type of singularity is present. Fractal measures

refer to different-type singularities in such a manner that the function $f(\alpha)$ has the shape of an inverted parabola extended between the points: $\alpha_{\min}$ (the strongest singularity) and $\alpha_{\max}$ (the weakest singularity).

The *generalized fractal dimension* D_q corresponds to the scaling exponent of the moment q of measure μ. If we cover the measure μ with spheres $B_i(\epsilon)$ with a size ϵ, then we will be able to define the *generalized scaling exponents* $\tau(q)$ using the so-called *partition function*:

$$Z(q,\epsilon) = \sum_{i=1}^{N(\epsilon)} \mu_i^q(\epsilon), \tag{31}$$

where $\mu_i = \mu(B_i(\epsilon))$, $q \epsilon \mathbb{R}$. At the limit $\epsilon \to 0^+, Z(q,\epsilon)$ shows a power-law behaviour:

$$Z(q,\epsilon) \sim \epsilon^{\tau(q)}. \tag{32}$$

The generalized fractal dimension is related to the generalized scaling exponent with the following relationship:

$$D_q = \frac{\tau(q)}{q-1}. \tag{33}$$

For different values of q, D_q takes on some characteristic values. For $q = 0$, the value of function $f(\alpha)$ attains a maximum, while D_0 is then the exponent of the support of measure μ. In the case $q = 1$, the equality $\alpha = f(\alpha)$, occurs, with the corresponding value D_1 being the information dimension. The information dimension tells us how the information needed for describing the point on a fractal scales itself with the increase in dimension ϵ:

$$D_1 = \lim_{\epsilon \to 0} \sum_{i=1}^{N} \frac{p_i(\epsilon)\ln(p_i(\epsilon))}{ln(\epsilon)}, \tag{34}$$

where $p_i(\epsilon)$ is the probability of a sphere i containing the point. The quantity D_1 is related to the information entropy of the system. All values $q \geq 2$ are referred to as the correlation dimension [49] of degree q. The above dimension measures the probability of two randomly points chosen from the set being away from each other by a certain distance. Change in the correlation dimension entails a change in the measure distribution. An example of the singularity spectrum and its characteristic values are shown in Figure 3.5.

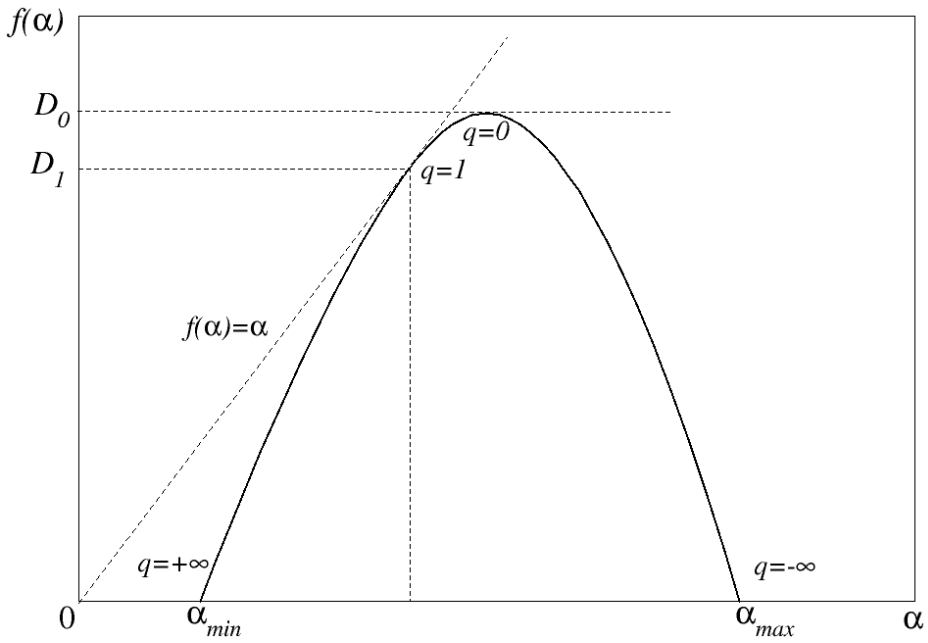

Figure 3.5 A sample multifractal spectrum $f(\alpha)$ of the determined binomial cascade.

If we assume that, for the adopted scale ϵ, the distribution of the values of singularity exponent α has the form of $\varrho(\alpha)\epsilon^{-f(\alpha)}$, then by substituting this expression in Equation (**31**)**,** we will obtain the relationship [48]:

$$Z(q,\epsilon)\sim \int \rho(\alpha)\epsilon^{q\alpha-f(\alpha)}d\alpha. \tag{35}$$

At the limit $\epsilon \to 0^+$, the main contribution to the integral comes from the expression $\epsilon^{\min_\alpha(q\alpha-f(\alpha))}$, from which it follows that $\tau(q)$ assumes the following form:

$$\tau(q) = \min_\alpha\big(q\alpha - f(\alpha)\big). \tag{36}$$

From the above formula, using the inverse Legendre transform, we get the formula for the spectrum $f(\alpha)$:

$$f(\alpha) = \min_q\big(q\alpha - \tau(q)\big). \tag{37}$$

In the case where $f(\alpha)$ is a smooth function, then, following formulas (**36**) and (**37**), we obtain the following relationship [48]:

$$\begin{cases} q = df/d\alpha \\ \tau(q) = q\alpha - f(\alpha). \end{cases} \tag{38}$$

Referring to thermodynamics, we can identify the value q with Boltzmann temperature $\beta = 1/kT$, while $\alpha_i = -\ln(\mu_i)/\ln(1/\epsilon)$, with energy E_i per unit volume of microstate [50].

4 Multifractal analysis of time series

The correlations and memory of time series can be examined in various ways. The analysis of spectral density or autocorrelation function provides, as a rule, information on the linear relationships in the system under analysis. To supplement the system information in terms of non-linear relationships, fractal methods are used. The analysis a time series, in which non-linear long-range correlations occur, is often reduced to the analysis of its multifractal structure. Currently existing algorithms used for such analysis enable the time series fractal dimensions and the scaling exponents to be determined based on the statistic properties of the time series. As these are methods relatively simple and easy to implement, they have found application in many branches of science, such as physics [51], biology [52; 53], chemistry [54], medicine [55], music [56], meteorology [57] or economics [58; 59; 60]. The examination of the autocorrelation and cross-correlation functions enables one to look inside the natural system and observe its dynamics. One of the most commonly known methods of examining the fractal properties of time series is the DFA (*Detrended Fluctuation Analysis*) algorithm. This algorithm was first mentioned in study [61] concerning the DNA analysis. It allows long-range power-law correlations to be determined for non-stationary processes. A generalization of this method on a multifractal case is the MF-DFA (*Multifractal Detrended Fluctuation Analysis*) method that enables correlations other than linear to be identified [62] .

For the proper detection of a correlation in a signal, it is necessary to identify the trend present in the data. Trends may lead to the incorrect identification of the correlation [62] due to the non-stationary nature of the signal. By eliminating the trend, we will obtain a stationary signal. In the above-mentioned methods, for each examined time scale s, the trend represented by the polynomial of degree m is removed. The choice of the polynomial degree is of key importance, as in the case where it is too big, small fluctuations, incorrectly identified as a trend, might be cut out from the signal [63]. If the degree m is too small, the trend may not be completely removed from the signal. Based on the results of study [63], a polynomial of the second degree was used for subsequent calculations, as it best describes the trend in financial data. To supplement the autocorrelation analysis of fluctuations, the data were also analyzed using the *Wavelet Leaders* method [64], which relies on the wavelet transform that decomposes the signal in the domain of time and frequency. To examine the power-law

correlation between signals, the DCCA (*Detrended Cross-Correlation Analysis*) method [65] was employed, which is a generalization of the DFA method on the case of two series. One of the multifractal variations of DCCA is the MF-DXA (*Multifractal Detrended Cross-Correlation Analysis*) method [66]. As is shown below, this method may lead to incorrect results. Therefore, the MFCCA *(Multifractal Cross-Correlation Analysis*) method [67] is proposed, as the one that is able to correctly identify the fractal cross-correlation. The method under consideration enables the determination of the spectrum exponent λ_q that characterizes the multifractal properties of the time series.

4.1 The MF-DFA (Multifractal Detrended Fluctuation Analysis) multifractal method

The British hydrologist Edwin Hurst, during his stay in Egypt from 1906, where he studied the water level in the Nile river basin, was one of the first to observe power laws in real time series [68]. He measured water level fluctuations in the water reservoir behind the dam around its average level. Had it been, as expected, an exactly random relationship, then the range of fluctuations would have varied in time following the classic Brownian motion[13] as $t^{1/2}$. However, Hurst noticed that the results could be best described using the relationship $R/S \sim n^H$, where n is the time scale, H is the Hurst exponent (not necessarily being equal to $1/2$), R is the range under analysis, and S, the standard deviation of the measurement, R. This is the so-called *Rescaled Range Analysis* (R/S), whose advantage is that it enables different types of time series to be compared. The Hurst exponent is dimensionless and is associated with the type of linear correlations occurring in a process. It can be regarded as a characteristic of the volatility of a fractal process, and in a wider context of financial markets, it can be considered a measure of investment risk. In the case of $H < 1/2$, the time series is correlated negatively, for $H > 1/2$ it is correlated positively, while for $H = 1/2$ it is uncorrelated. For the flooding of the Nile, the H value estimated by Hurst is ~ 0.77 [69]. Hurst discovered also that the majority of natural systems are not subject to random walking, but are in some way slightly correlated.

[13] In accordance with the classic Brownian motion law, a distance R covered by a particle is proportional to the square root of time $R \sim T^{1/2}$. This principle is commonly used in investment strategies, where, having the monthly standard deviation T and multiplying it by the root of 12, we obtain the yearly volatility of R [119].

The DFA method [70] has been introduced, because other conventional signal fluctuation description methods, such as R/S, cannot be used for describing non-stationary data. DFA enables the detection of long-range power-law correlations in signals containing trends (most often of the polynomial type), which make the identification of correlations in signal fluctuations difficult.

A generalization of the DFA method on the multifractal case is the MF-DFA method [71; 72], which works as follows: given a time series x_i of the length $N(i = 1,2, \dots N)$, we define the so called profile function according to the formula

$$X(j) = \sum_{i=1}^{j} [x_i - \bar{x}], \tag{39}$$

where $\bar{x}$ is the mean value calculated for the whole signal. Then, the signal X is divided into $\mathrm{M}_s = \frac{N}{s}$ intervals, each of the length s. This operation is repeated two times, from the beginning and end of the series, with the aim of avoiding the situation, where the relevant events occurring at either end of the series are omitted. As a result, we obtain $2M_s$ intervals v, in which determine the trend $P_{X,v}^{(m)}$ being a polynomial of degree m and the variance

$$F_x^2(v,s) = \frac{1}{s}\sum_{k=1}^{s} \left\{\left(X\big((v-1)s + k\big) - P_{X,v}^{(m)}(k)\right)\right\}^2, \tag{40}$$

where $v = 1 \dots M_s$. Then, the variance is averaged over all intervals and the fluctuation function is determined for the parameter $q \neq 0$:

$$F_x(q,s) = \left\{\frac{1}{2M_s}\sum_{v=1}^{2M_s} [F_x^2(v,s)]^{q/2}\right\}^{1/q}. \tag{41}$$

The core of the method is the analysis and respective interpretation of the fluctuation function. If, for different time scales s, the function $F_x(q,s)$ has a power-law behaviour, i.e. the following relationship is satisfied:

$$F_x(q,s) \sim s^{h(q)}, \tag{42}$$

then we can speak of the fractal properties x_i characterized by the generalized Hurst exponent $h(q)$, which for $q = 2$ is the classic Hurst exponent defining linear correlations. The method allows processes of a multifractal, monofractal or a totally non-fractal nature to be distinguished. For multifractal series, $h(q)$ changes its value, depending on the parameter q, while for monofractal series, h is constant and equal to the classic Hurst exponent $h(q) = H$. Having the $h(q)$ determined, using the

relationship given by (**37**) and the relationship $\tau(q) = qh(q) - 1$ [48] we can determine the singularity spectrum. So, for random time series, the singularity spectrum should be reduced to a point located at $\alpha \approx 0.5$ and $f(\alpha) \approx 1$. When examining real monofractal date it often turns out that the determined spectrum is an inverted parabola rather than a point. This is the effect of the finite length of the examined series. Figure 4.1 shows three examples of fractal spectra determined for the fractional Brownian motion[14] (*Fractional Brownian Motion*) [27], given by three different Hurst exponents. In each case, the maximum of the multifractal spectrum $f(\alpha)$ falls almost exactly on α being equal to the given Hurst exponent. A certain narrow spectrum width can also be observed.

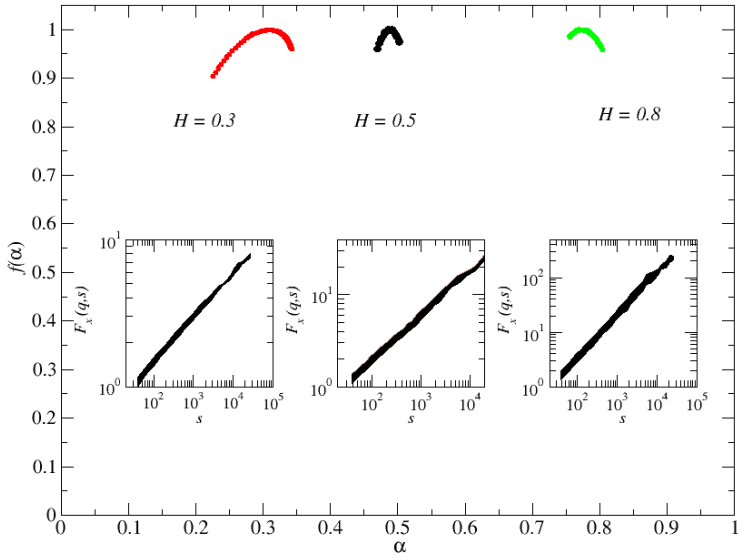

Figure 4.1 Fractal spectra for FBM processes with a different Hurst exponent. The process H = 0.3 is marked in red colour, H = 0.5 in black, and H = 0.8 in green. Inset plots show respective fluctuation functions.

An example of a fractal process is the *Markov-Switching Multifractal Model* (MSM) first introduced in study [73]. This is an iterative model, which is able to replicate the hierarchic structure, providing it a multifractal nature. Its important feature is the fact that it can be used to model the process variation clustering (large fluctuation groups intermingled with groups of small fluctuations), therefore MSM is widely used in finances. The random variable r_t of the MSM type is given by the formula:

[14] The fractional Brownian motion $B_H(t)$, which is a generalization of the classic Brownian motion and is controlled by the Hurst exponent H. Its average value is 0, and the covariance function is given by the $Cov\big(B_H(t_1); B_H(t_2)\big) = \frac{\sigma^2}{2}(t_1^{2H} - |t_1 - t_2|^{2H} + t_2^{2H})$.

$$r_t = \sigma_t u_t, \tag{43}$$

where u_t is the Gaussian random variable, while

$$\sigma^2{}_t = \sigma^2 \prod_{i=1}^{k} M_t^{(i)} \tag{44}$$

is the process volatility determined by k multipliers $M_t^{(1)}, M_t^{(2)}$ ... $M_t^{(i)}$ and a constant factor σ. The model assumes that the multipliers $M_t^{(i)}$ are sampled from a binomial or log-normal distribution. The process volatility is adapted in each step t with the probability γ_i, depending on the rank, and remains unchanged with the probability $1 - \gamma_i$. The probability of change is given by the formula:

$$\gamma_i = 1 - (1 - \gamma_k)^{(b^{i-k})}, \tag{45}$$

where $\gamma_k \in (0; 1)$ and $b \in (1, \infty)$.

Figure 4.2 shows a sample realization of the MSM process generated from the binomial distribution with the weight $m_0 = 1.2$ and its fractal characteristics. The binomial distribution means that the variable $M_t^{(i)}$ can assume the values m_0 and $2 - m_0$ with equal probability.

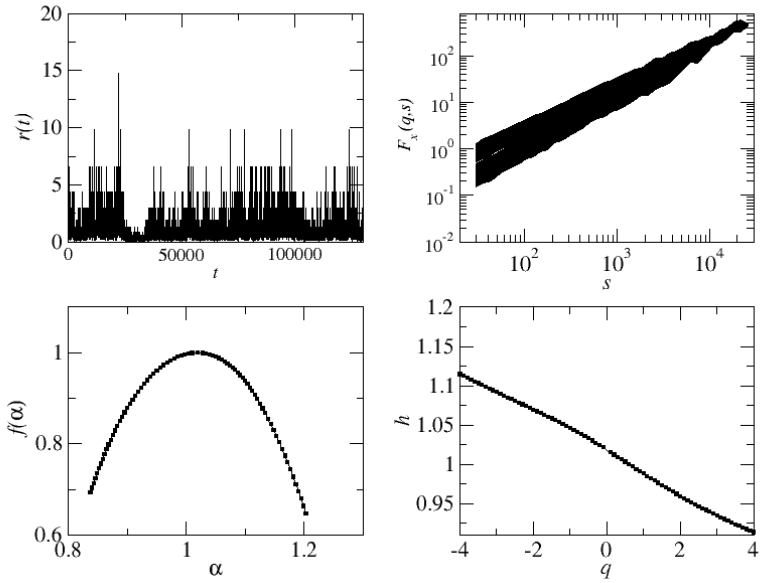

Figure 4.2 The MSM process generated for the weight $m_0 = 1.2$. The panel a) represents the realization of the process, b) the fluctuation function determined for it, c) the multifractal spectrum, while d) the generalized Hurst exponent $h(q)$ determined for each of the moments q.

Unlike the monofractal process, the fluctuation functions $F_x(q, s)$ are not parallel to each other, and for every value of q they assume different slopes. The determined

multifractal spectrum $f(\alpha)$ takes on the shape of an inverted parabola, is $(\Delta\alpha \approx 0.4)$ wide and its maximum falls on the point $\alpha = 1$. The generalized Hurst exponent $h(q)$ is a decreasing function q.

Another example of a multifractal process is the multiplicative cascade which is a generalization of the Cantor set. It is constructed in the following manner: The measure is laid down on a unit segment. The segment is divided into 2 portions, and part of the mass m_0 and m_1 is assigned to each of the portions, where $m_0 + m_1 = 1$ is the total mass. In the next step, each of the newly formed segments is divided again into 2 portions, while assigning part of the mass m_0 to the left-hand segment and m_1 to the right-hand segment. This procedure is repeated for n steps in such a manner that the total mass is the product of the masses of all preceding steps. If the probability distribution of these random variables does not depend on the iteration, then we have a self-similar structure with the retained relation of the scaling moments and the probability distribution. An example of fluctuation function determined for the cascade and Gaussian processes is shown in Figure 4.3.

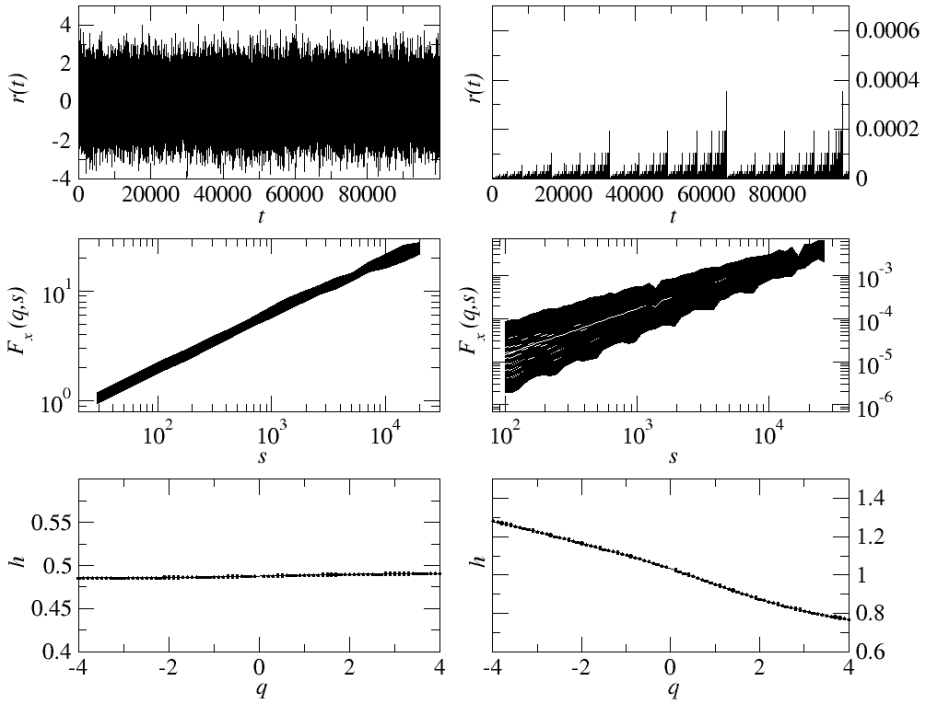

Figure 4.3 Starting from the top: the time series, the fluctuation function $F_x(q,s)$ and the Hurst exponent for, respectively, the Gaussian process (the left-hand panel) and the cascade process (the right-hand panel).

A difference between the two types of processes is clearly seen, on the level of both the fluctuation function $F_x(q,s)$, and the relationship $h(q)$. In the case of the Gaussian process, the values of the function $F_x(q,s)$ for a given scale s and moment q do not practically differ from one another and are parallel to each other. By contrast, for the cascade process, the dependence of the slope of the fluctuation function on the parameter q is clearly seen. Moreover, 'steps' in the graph of the function $F_x(q,s)$, can be observed; its values take on different values, depending on q, and with increasing scale the functions become closer to each other. The generalized Hurst exponent $h(q)$ is constant with respect to q for the Gaussian process, whereas for the cascade process it is not a constant function.

4.1.1 Analysis of foreign exchange market data

The multifractal analysis of the financial data from Section 3 using the MF-DFA method will be presented here.

4.1.1.1 Base currencies

As has already been mentioned, the fluctuation function $F_x(q,s)$ determined using the MFDFA algorithm is the basic characteristic informing about a possible fractal nature of the time series. Its behaviour with the change of the time scale enables one to answer the question whether the dependence of the function on the scale follows the power law, and if so, for what scales. Figure 4.4 represents such a function for 5 seconds' returns of the EURUSD currency pair for different values of the parameter q.

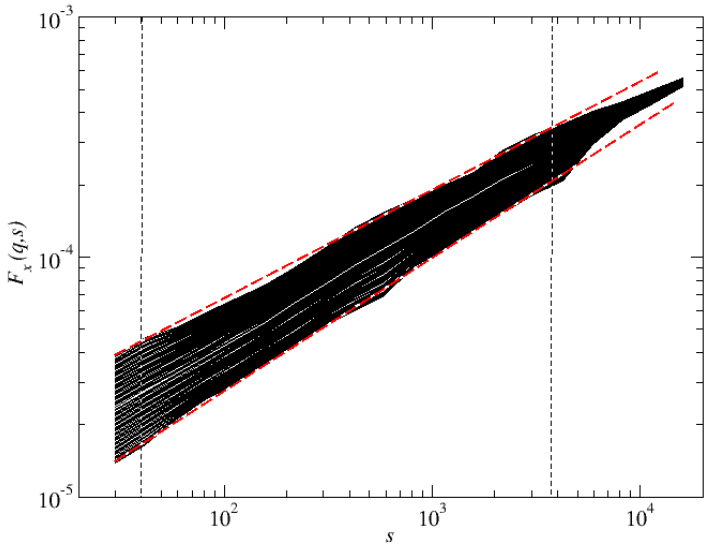

Figure 4.4 Fluctuation function $F_x(q,s)$ for the EURUSD currency pair. The black line indicates the scaling interval that extends from the scale from $s = 40$ to $s = 3500$. The red dashed line denotes the power-law matching to the extreme fluctuation function values.

The function $F_x(q,s)$ has a power-law behaviour in the scaling interval from $s = 40$ (slightly over 3 trading minutes) to $s = 3500$ (about five trading hours), which indicates a fractal nature of the autocorrelation. For the largest scales, the function F_x starts fluctuating, and therefore no such intervals are taken into account when determining the Hurst exponent. In this case, the good scaling quality allows the multifractal spectrum to be determined for the time series under examination. As has been demonstrated in section 2.2.2, the tails of the distributions of the returns and volatilities are leptokurtic (Figure 2.4). For this reason, the multifractal spectra were determined for q from the interval $q \in [-4,4]$, while for detrending, a polynomial of degree 2 was used, which, as has been shown in study [63], well reflects a trend in financial data. The limited selection of the values of q is due to the fact that the inverse cubic law essentially precludes the moments $q > 4$ which, for greater q values, are divergent [74]. The degree of the multifractal complexity of the examined signal can be determined from the relationship:

$$\Delta\alpha = \alpha_{max} - \alpha_{min} = \alpha(q_{min}) - \alpha(q_{\max}). \qquad (46)$$

The greater the value of $\Delta\alpha$, the greater the diversity of individual fractals that make up the examined multifractal. As the criterion of the multifractality of the examined series, $\Delta\alpha \geq 0.15$ is adopted [63].

The obtained spectra for all exchange rates are shown in Figure 4.5 (indicated with the black line). They are all multifractal, and in the case of the returns, the EURJPY pair has the least width of $\Delta\alpha = 0.15$, while the maximum of this spectrum falls on $\alpha_0 \approx 0.46$. The widest spectrum of $\Delta\alpha = 0.29$ is exhibited by the GBPUSD currency pair, which may be indicative of the complex dynamics of this pair, with the maximum spectrum being at point $\alpha \approx 0.5$. All the examined spectra have their maxima in the range of $\alpha_0 \in (0.45; 0.51)$, which indicates a lack of linear correlation in the analyzed signals or the presence of very weak anti-persistence.

For volatility, the situation is different (the red line in Figure 4.5). Except for the British pound-based currency pairs (the last row in the above figure), the widths of all spectra are comparable to those of the spectra of the returns. The narrowest spectrum width $\Delta\alpha = 0.15$ is shown by the USDJPY pair, with the maximum falling on point $\alpha \approx 0.64$. The widest spectra are exhibited by the British pound-based currency pairs, with widths (and a position) of, respectively, $\Delta\alpha = 0.41$ ($\alpha_0 = 0.68$) for the GBPUSD pair, $\Delta\alpha = 0.41$ ($\alpha_0 = 0.68$) for the GBPJPY pair, and $\Delta\alpha = 0.56$ ($\alpha_0 = 0.71$) for the EURGBP pair. Interestingly enough, the spectra $f(\alpha)$ attain negative values, which is the result of the fact that the function $F_x(q, s)$ changes from the increasing function to a decreasing one for $q > 0$ and small time scales s. For the first time, this effect was observed at the returns level in study [31], and is regarded as an anomaly [75]. All volatility spectra are distinctly located for $\alpha > 0.5$, which evidences a positive autocorrelation. Moreover, all of them are, to a greater or lesser extent, left-side asymmetric. This suggest the existence of a more complex organization of the series at the level of large fluctuations (for $q > 0$), whereas the second portion of the spectrum ($q < 0$) is generated by smaller fluctuations, most likely by a kind of "noise" which is more monofractal (in the shape of Gaussian noise). To determine the degree of asymmetry, the following relationship [76]:

$$A_\alpha = \frac{\Delta\alpha_L - \Delta\alpha_R}{\Delta\alpha_L + \Delta\alpha_R}, \tag{47}$$

where $\Delta\alpha_L = \alpha_0 - \alpha_{min}$ and $\Delta\alpha_R = \alpha_{max} - \alpha_0$ denote the distance of the spectrum maximum position α_0 from, respectively, the least and greatest determined value of α.

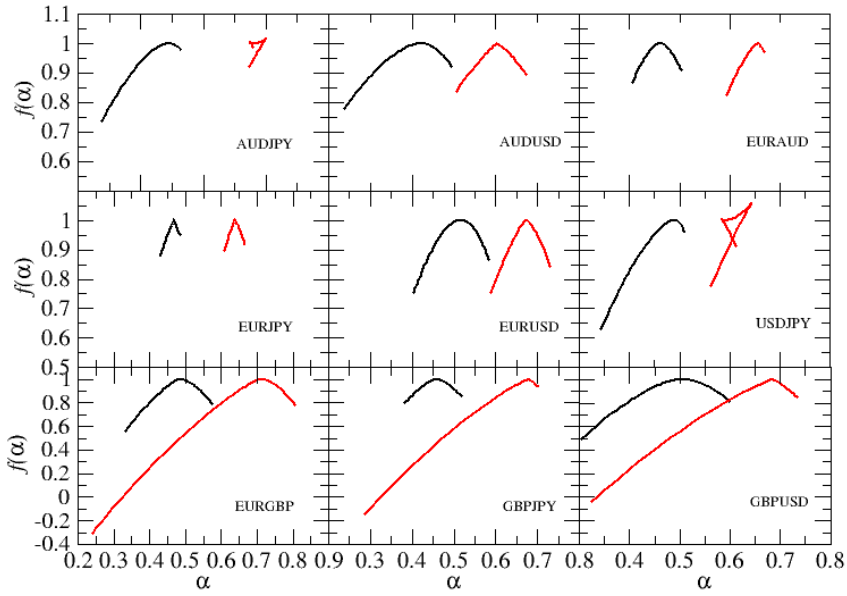

Figure 4.5 Multifractal spectra for the returns (black colour) and volatility (red colour).

The positive value of A_α means a left-hand asymmetry, while the negative value, a right-hand asymmetry. For the returns, the highest degree of asymmetry is shown by the USDJPY currency pair, where $A_\alpha = 0.76$, while the lowest, by GBPJPY with $A_\alpha = 0.07$. In the case of volatility, these are, respectively, GBPJPY with $A_\alpha = 0.85$ and EURJPY with $A_\alpha = 0.01$. Spectra of this type (left-asymmetric) were first described in study [77], and for financial data, literary texts and inter-transaction times, were also comprehensively analyzed in study [76]. Using artificial time series generated with the use of symmetric binomial cascades (Figure 4.2) and setting them appropriately, the authors constructed a multifractal process having left-hand spectra.

To sum up, the analyzed FX data have the nature of processes that are linearly uncorrelated at the returns level and correlated at the volatility level. The majority of spectra are asymmetric in shape, of which the British pound-related exchange rates EURGBP, GBPUSD and GBPJPY have the left-hand arm most developed, which indicates the greatest contribution of large fluctuations to the multifractality of these signals.

As has been mentioned, multifractality is a non-linear phenomenon, therefore non-linear rather than linear correlations can only be its source. The $f(\alpha)$ spectra determined for Gaussian processes exhibit a monofractal character, they are narrow and approximately punctual, whereas the spectra determined for processes with the leptokurtic distribution of fluctuations (e.g. truncated Levy flights [78]) assume the shape of an inverted parabola. This might suggest that the data distribution could affect

the multifractal structure. As have been demonstrated previously, FX data are distinguished by the fat tails of fluctuation distributions and time correlations, therefore the both factors can potentially be a source of signal fractality [79; 80; 81; 82]. It is worth noting, however, that probability distributions can lead only to an apparent fractality, as shown in study [30]. By using uncorrelated *q-Gaussian* series, the authors have demonstrated that for unstable time series (i.e. those convergent to the Gaussian distribution in accordance with the Central Limit Theorem) and $1.55 < q < 1.65$, two areas of scaling of the fluctuation function $F_x(q,s)$ are visible. For small scales s, the process looks multifractal, while for larger scales, monofractal. This effect is the more visible, the higher the value of q is (the thicker the distribution tails are). In that case, a larger range of scales is needed to reconstruct the monofractal process, to which a convolution of independent *q-Gaussians* should be convergent.

In order to confirm the reliability of the obtained results, all linear correlations in the signal were destroyed by randomly mixing the data. Thus, only the character of the probability distributions was preserved. When comparing the spectra of the mixed time series (Figure 4.6, the right-hand panel) with the original data, a difference is clearly seen at the level of $f(\alpha)$ spectrum width and $\Delta\alpha$ width. The average spectrum width for the entire set of 9 exchange rates is $\overline{\Delta\alpha} = 0.27$, and for the mixed data, $\overline{\Delta\alpha} = 0.09$. Moreover, in each of the analyzed cases, the maximum of the mixed series spectrum is located at point $\alpha_0 \sim 0.5$. This confirms the statement that it is correlations that underlay the fractal structure of data.

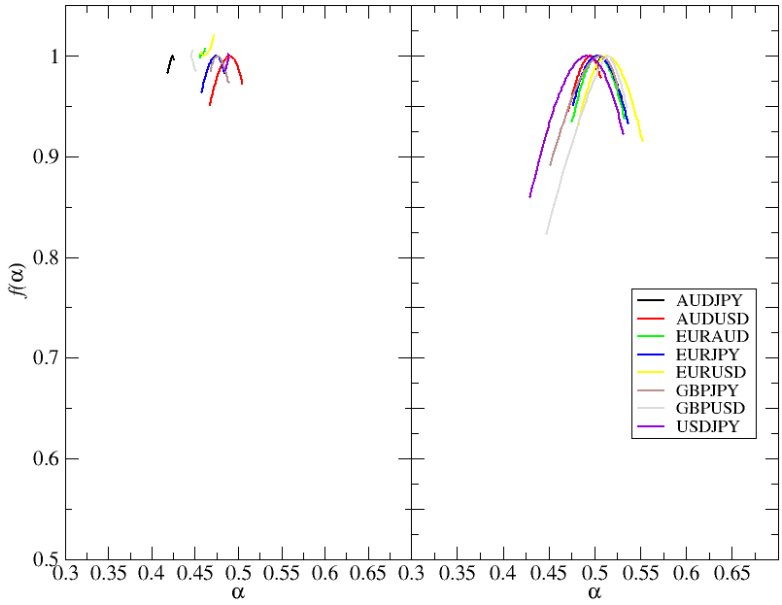

Figure 4.6 The multifractal spectra of mixed data (the right-hand panel) and Fourier surrogates (the left-hand panel).

Taking a step forward, non-linear correlations were removed from all-time series by the Fourier substitute method [15], leaving linear correlations intact. This operation leads to a reduction of the $f(\alpha)$ spectrum width, while retaining its maximum at the same point as for the original data spectrum. The average spectrum determined for all currencies is $\overline{\Delta\alpha} = 0.01$, which indicates the monofractal nature of the substitutes. This implies that the source of multifractality in the examined signals are non-linear correlations, as demonstrated also in reference [80].

4.1.1.2 Deviations from the triangular relation

For the same $q \in [-4,4]$ parameter values as before, deviations from the triangular relation satisfying Equation (**14**) were analyzed. Figure 4.7 shows the fluctuation function $F_x(q,s)$ determined for 7 time series of triangular relation deviations (5-second data). It can be clearly seen that, in contrast to the analysis made for exchange rates in the preceding section, in this case the fluctuation function behaves differently. No scaling practically exists in the scale interval under consideration. The result is surprising, insomuch as the authors of study [31] when examining 1-minute exchange rates, observed a multifractal structure in them. The spectra determined in the

[15] In creating surrogates, the Fourier transformation of the time series is made, the amplitude is retained and the phase is randomly mixed. From so modified series, by means of the inverse Fourier transform, a surrogate signal is obtained.

study quoted above were asymmetric with the developed left-hand side and negative values of the scaling exponent and the multifractal spectrum.

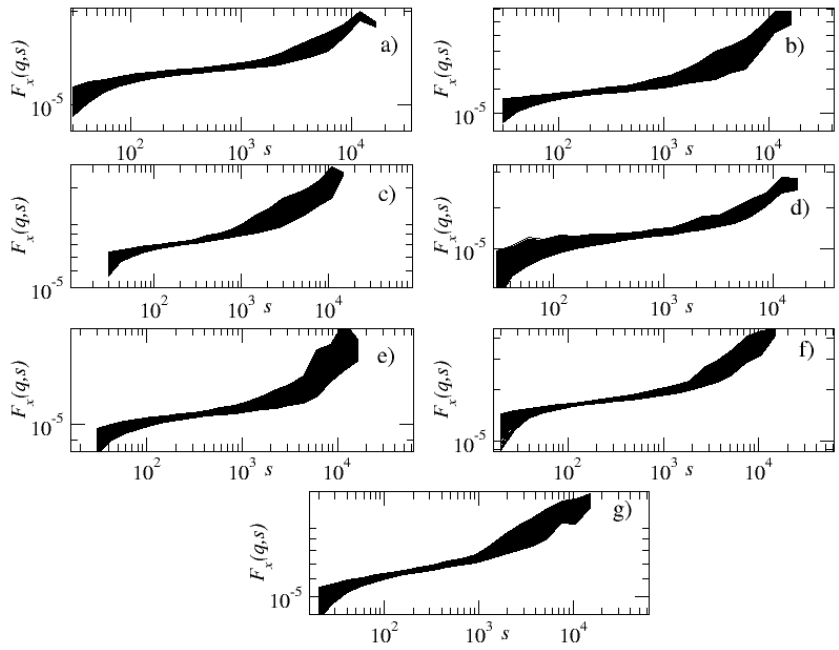

Figure 4.7 The fluctuation functions $F_x(q,s)$ determined for all deviations from the triangular relations: a) AUD-EUR-JPY, b) AUD-EUR-USD, c) AUD-USD-JPY, d) EUR-GBP-JPY, e) EUR-GBP-USD, f) EUR-USD-JPY, g) GBP-USD-JPY.

So, the question arises if possible linear correlations manifest themselves in this case only on larger time scales, starting from the 1 min. scale. To this end, triangular relation deviations were also analyzed for the 1-minute series. Similarly as for the 5-second data, the fluctuation function did not show suitable scaling, which made the multifractal spectrum estimation impossible. The discrepancy in results compared to study [31] referred to above might be due to the different periods from which the data were derived. More recent data, as those considered in this study, may reflect faster processing of information on the exchange markets, which translates into higher market efficiency and, as a result, leads to a shortening of the arbitrage scale. An additional explanation of the discrepancy in the results might be the fact that the proper identification of the fractality of such time series is affected by the length of the series under analysis. In the present study, series of a length of slightly above 86000 points were used for analysis, while in study [31] series above 1.7 million points in length were examined. A shorter series simply means that there are no rare events in the signal, which did not have enough time to occur within the time frame under consideration.

4.1.1.3 Cross exchange rates– the effect of the number of currencies on the f(α) spectrum width

As has been shown in section 2.3.2, all exchange rates examined in this study are more or less cross-correlated. Making use of this fact, the behaviour of exchange rates calculated by an indirect method, i.e. using cross-exchange rates, was also analyzed. Each of the exchange rates analyzed above can be calculated, e.g., using other exchange rates by passing through, respectively, two, three or four such exchange rates according to the following formula:

$$\frac{A}{B}\frac{B}{C}\frac{C}{D}\cdots\frac{Y}{Z}=\frac{A}{Z},$$

where $\frac{X}{Y}$ is the exchange rate of respective currencies. Figure 4.8 shows sample multifractal spectra and fluctuation functions $F_x(q,s)$ for the GBPUSD pair (black colour), as dependent on the number of exchange rates used for exchange rate calculations:

- 2 exchange rates = $\frac{GBP}{JPY}\frac{JPY}{USD}=\frac{GBP}{USD}$ (red colour)
- 3 exchange rates = $\frac{GBP}{JPY}\frac{JPY}{EUR}\frac{EUR}{USD}=\frac{GBP}{USD}$ (green colour)
- 4 exchange rates = $\frac{GBP}{JPY}\frac{JPY}{AUD}\frac{AUD}{EUR}\frac{EUR}{USD}=\frac{GBP}{USD}$ (blue colour).

The function $F_x(q,s)$ was determined for all examined scales and it exhibits scaling from, more or less, the scale $s=200$ to the scale $s=5000$. All determined singularity spectra indicate a multifractal nature of the signal under analysis.

Figure 4.9 presents the widths of the spectra and their maximum, as dependent on the number of exchange rates used. In the top panel, where the location of the spectrum maximum is shown, it can be seen that with the increase in the number of exchange rates used for the calculation of the proper exchange rate, the maximum shifts towards the anti-persistent series. This is not a strong effect, though it is characteristic of each of the exchange rates examined.

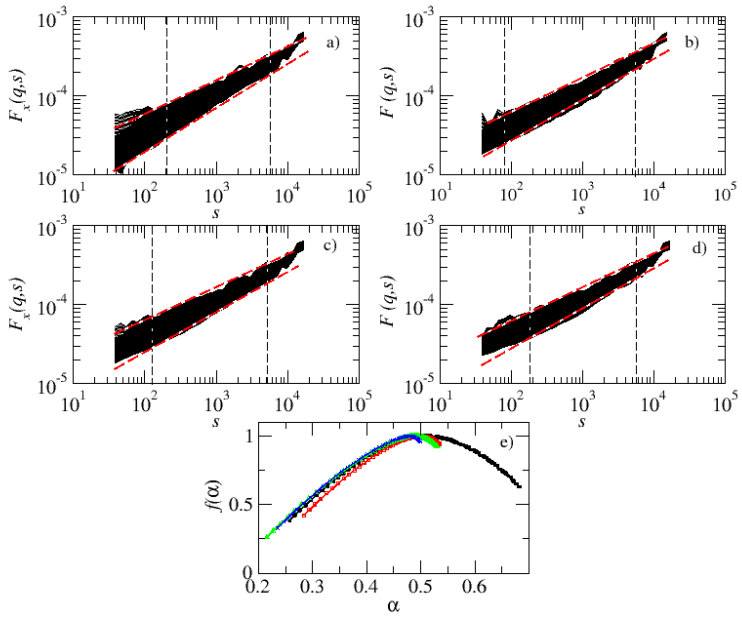

Figure 4.8 Panels (a-d) present the returns fluctuation function for the exchange rates a) $\frac{GBP}{JPY}$, b) $\frac{JPY}{AUD}$,c)$\frac{AUD}{EUR}$,d) $\frac{EUR}{USD}$, used for determining the GBP/USD exchange rate. The black dashed line denotes the scaling interval, while the red dashed line, the matching of the power-law function to the extreme values of the fluctuation function. Panel e) represents the multifractal spectra for the GBPUSD pair, determined for a different number of transition exchange rates. The red colour denotes two, green three, and blue four transition exchange rates. The original exchange rate is marked in black colour.

One can speculate here on the possible cause of this behaviour. What seems the most probable is the fact that by extending the arbitrage procedure to achieve the risk-free profit, due to an increased likelihood of making an error (the number of transition currencies), the risk of failure increases (the spectrum shifts to the left, and from the previous discussion we know that the more anti-persistent the series, the lower the probability of a trend occurring and, consequently, the risk).

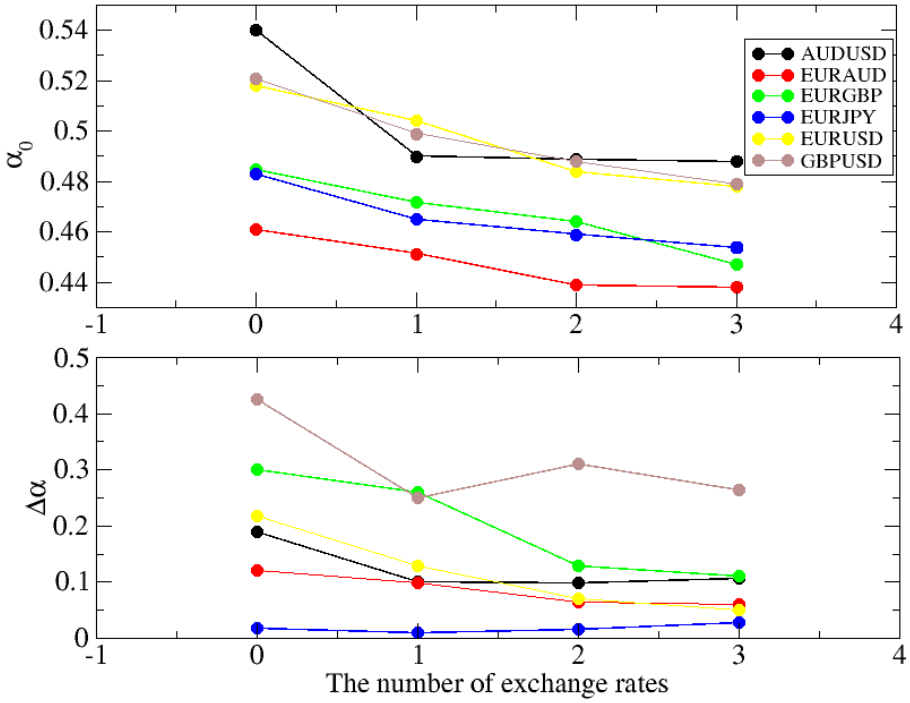

Figure 4.9 The maximum of the multifractal spectra (the upper panel) and the width of the spectrum (the lower panel), as dependent on the number of exchange rates used for determining the base exchange rate.

The situation is slightly different for the width of the spectrum shown in the lower panel in Figure 4.9. For the EURGBP, EURUSD and EURAUD exchanges rates, as the number of transition currencies increases, the spectrum becomes increasingly thinner, and for 4 exchange rates it starts balancing on the borderline characteristic of monofractal processes. In turn, for the EURJPY exchange rate, the spectrum is very narrow and typical of a monofractal, invariably for all recalculated cross-exchange rates. For the AUDUSD and GBPUSD exchange rates, the width decreases for one transition exchange rate, to slightly increase afterwards.

The shift of the spectrum towards the anti-persistent region and the absence of systematic spectrum narrowing depending on the number of transition currencies is surprising, because the authors of study [76] by analyzing, inter alia, symmetric multifractal cascades, have proved that the coincidence of many multifractal processes leads to the destruction of the hierarchical structure in the series, which results in a noise and monofractality. It seems that the difference is associated with the nature of the data examined here, where the exchange rates are interrelated (as shown in chapters 2.3.2 and 4.3.1). So, the sum of series of this type leads to retaining the fractal correlations, which is reflected in the results presented herein.

4.1.2 Analysis of stock market data

Similarly as for the FX market, 29 companies included in the DJIA index were subjected to analysis. From the times series of stock quotes, intervals were removed, in which the price had not changed for 20 minutes and more, to ensure the compactness of the carrier [71]. The parameter q assumed values $q \in [-4,4]$. Figure 4.10 shows functions $F_x(q,s)$ for four companies with the largest (the left-hand side) and the smallest (the right-hand side) spectrum width. It is clearly seen that in both cases the functions are of the power-law type and meet convincingly the scaling relations for the majority of scales. For the calculation of the scaling exponent $h(q)$, the range from $s = 60$ to $s = 20000$ was assumed, where scaling is the best (the range marked with the dashed lines in the figure).

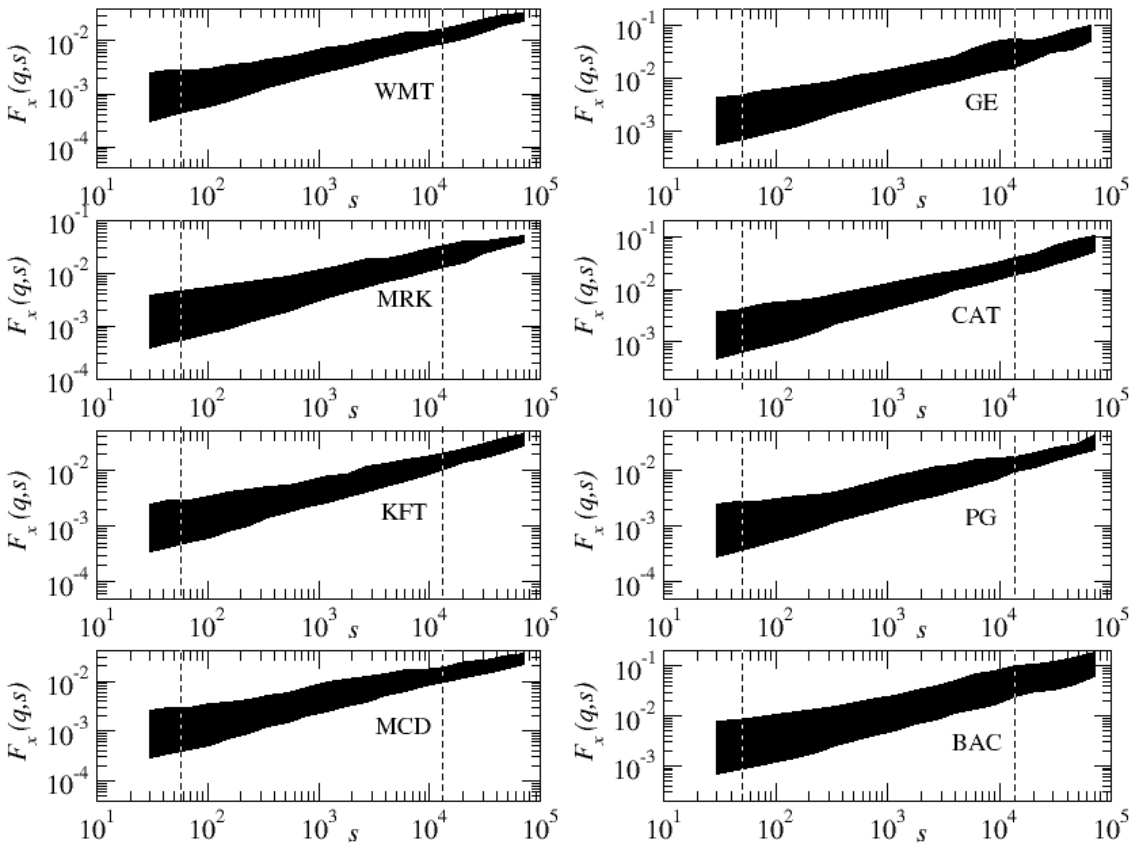

Figure 4.10 The fluctuation function $F_x(q,s)$ for the time series determined for the companies of the largest (the left-hand panel) and the smallest (the right-hand panel) width of the spectrum.

Figure 4.11 presents the multifractal spectra $f(\alpha)$ determined for all companies. They are all asymmetric with the more developed left-hand side being responsible for large

fluctuations. Their maxima lie in the region $\alpha > 0.5$, which is indicative of the persistence of the time series.

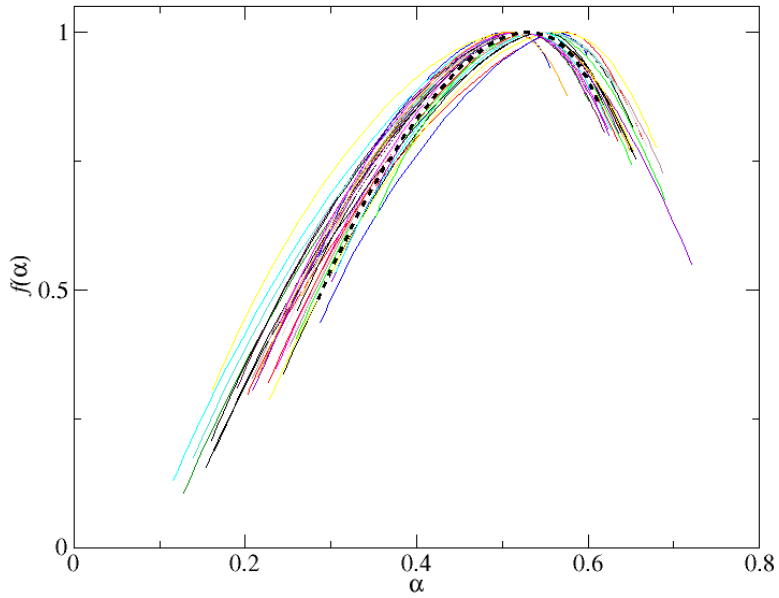

Figure 4.11 Multifractal spectra for 29 DJIA index companies. The spectrum for the index is marked with the dashed line.

The mean spectrum width for the examined set of companies is $\overline{\Delta\alpha} \approx 0.40$,. The companies with the largest spectrum width are WMT (Wall Mart) $\Delta\alpha \approx 0.48$, MRK (Merck & Co) $\Delta\alpha \approx 0.48$, KFT (Kraft Foods) $\Delta\alpha \approx 0.47$ and MCD (McDonald's) $\Delta\alpha \approx 0.45$. In turn, the PG (Procter & Gamble) company has the smallest spectrum width of $\Delta\alpha \approx 0.27$, followed closely by GE (General Electric) with a spectrum width of $\Delta\alpha \approx 0.27$, CAT (Caterpillar) with $\Delta\alpha \approx 0.30$ and BAC (Bank of America Corp.) with $\Delta\alpha \approx 0.31$. In spite of the big differences in spectrum width, all of them are multifractal which indicates a complex structure of the series analyzed.

Interestingly enough, fairly significant differences arise at the sectoral analysis level. The largest spectrum width, being $\overline{\Delta\alpha} \approx 0.44$, is exhibited by financial companies, while industrial goods sector companies have the narrowest spectrum of $\overline{\Delta\alpha} \approx 0.37$. Table 4.1 provides all values by sectors. From this perspective, too, the richness of the fractal structure of the examined data can be noticed. As mentioned, the mean spectrum width value is the greatest in the financial sector. This result could have been expected, because it is financial companies in a broad sense through which the majority of information passes and that are linked with other market participants. Industrial goods sector companies, on the other hand, have the narrowest spectrum,

which indicates a slightly weaker (though still strong) multifractality of the series examined.

Table 4.1 Average multifractal spectrum widths $\overline{\Delta\alpha}$ and the asymmetry factor A_α, given by sectors.

Sector	$\overline{\Delta\alpha}$	A_α
Industrial Goods	0.37	0.5
Consumer Goods	0.38	0.56
Basic Materials	0.40	0.38
Technology	0.40	0.47
Healthcare	0.42	0.59
Services	0.43	0.52
Financial	0.44	0.55

Similarly as for the currency data, for all companies, the time series of mixed data were made and the multifractal were determined for them. In each case, the maximum of the determined spectrum is located in the vicinity of $\alpha \approx 0.5$, and the mean width $\overline{\Delta\alpha}$ is at a level of 0.19. This residual multifractality also results from the length of the analyzed series and is an apparent effect. Moreover, for all companies, the asymmetry factor was calculated from formula (**47**). In the examined set of companies, there is no A_α factor with a negative value, which would indicate a right-hand spectrum. The smallest value of $A_\alpha = 0.08$ is shown by the DD (Disney) company, which is indicative of a nearly symmetric spectrum, while the greatest value of $A_\alpha = 0.77$ is exhibited by BAC (Bank of America) with one of the narrowest spectra.

4.2 The Wavelet Leaders method

In contrast to the above-mentioned multifractal analysis methods based on detrended data analysis, there are also other methods that enable one to capture the richness of fractal structures. The WTMM *(Wavelet Transform Modulus Maxima)* method*)* [64] is a method relying on the wavelet transform, which enables the singularity spectrum of an examined signal to be determined. Its core is the decomposition of the signal into components corresponding to the mother wavelet function. We can think of the wavelet as a wave with the mean value equal to zero, being well located in the time and frequency domains. As demonstrated in study [71], it lends itself well to the analysis of strongly non-stationary time series. However, in

contrast to the detrended fluctuation method, the correctness of the results depends very heavily on the selection of the appropriate wavelet, therefore one needs to be very careful when approaching the interpretation of the results.

In the study *Wavelet Leaders in Multifractal Analysis* [64], the Wavelet Leaders method has been proposed, which enables the determination of the singularity spectrum in a wide scale range and is applicable to processes that contain singularities of the *chirp* type and the *cusp* type[16]. As compared to the WTMM method, this is a more stable method.

In the first step of the method, the signal is decomposed according to this formula:

$$f(x) = \sum_{j,k\in Z} c_{j,k}\psi\left(2^j x - k\right) = \sum_{\lambda\in\Lambda} c_\lambda \psi_\lambda, \tag{48}$$

where ψ is a wavelet and $c_{j,k}$ are wavelet coefficients linked to the segment λ for scale j and position k with the following relationship:

$$\lambda = \lambda_{j,k} = \left[k2^{-j}, (k+1)2^{-j}\right], \tag{49}$$

$$c_{j,k} = 2^j \int_R f(x)\,\psi\left(2^j x - k\right)dx. \tag{50}$$

For every segment λ, the wavelet leader is $d_\lambda = \sup_{\lambda\in\lambda'}|c_{\lambda'}|$, and the distribution function is defined for it using the following relationship:

$$S(q,j) = \frac{1}{2^j}\sum_{\lambda\in\Lambda_j} d_\lambda^q, \tag{51}$$

where Λ_j is a set of segments on scale j. The generalized scaling exponents are given by the formula $\tau(q) = \log_{j\to+\infty}\frac{\log(S(q,j))}{log2^{-j}}$ and are linked to the singularity spectrum with the relationship $d(h) = \inf\{qh - \tau(q)\} + 1$, where h is the Hölder exponent (**34**). Results for the exchange market are shown in the figure below. For the analysis, the package described in study [64] and available on the website http://www.irit.fr/~Herwig.Wendt/software.html was used. For each of the currency pairs examined in the previous sections, the multifractal spectrum was determined for both the rates of return (black colour) and volatility (red colour).

[16] Chirp-type singularities occur in signals, in which the oscillation frequency either increases or decreases with time. In the case of the cusp-type singularity, the tangent at a point cannot be determined.

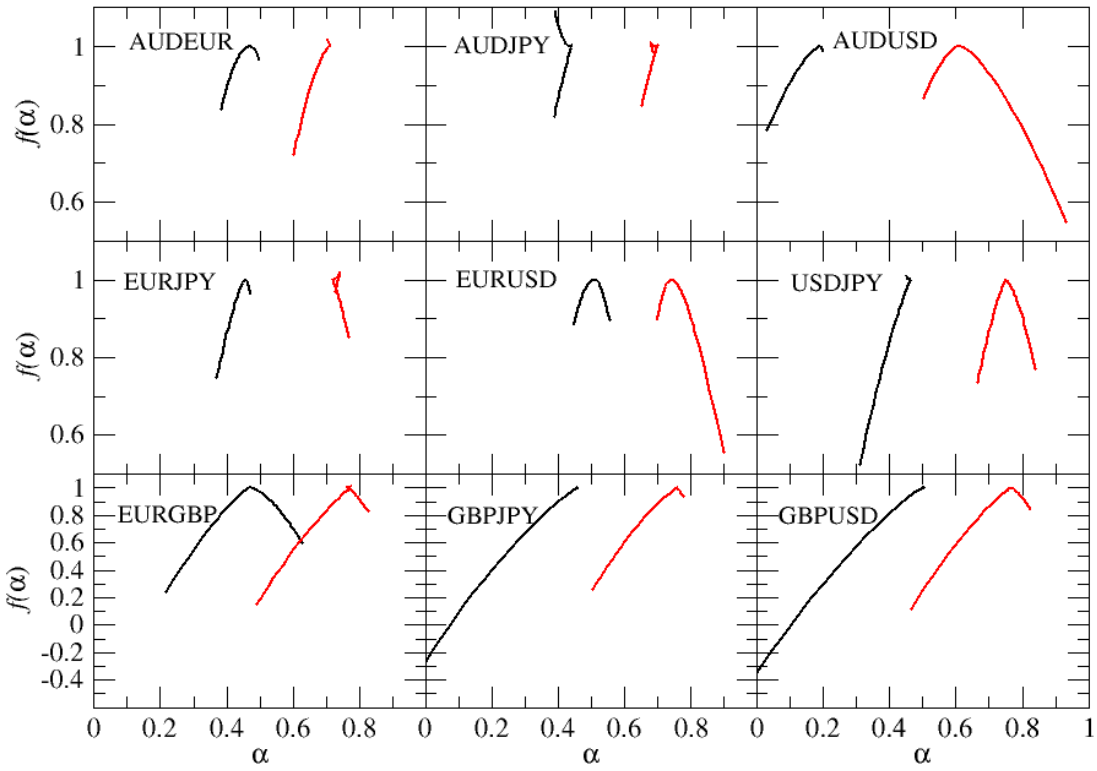

Figure 4.12 The multifractal spectra of all currency pairs determined by the *Wavelet Leaders* method.

The widest spectra determined for the rates of return are exhibited by the currency pairs based on the British pound (the bottom panel in Figure 4.12) with the widths (and the maxima) amounting to, respectively, $\Delta\alpha = 0.56$ ($\alpha_0 = 0.5$) for the GBPUSD pair, $\Delta\alpha = 0.47$ ($\alpha_0 = 0.46$) for the GBPJPY pair, and $\Delta\alpha = 0.42$ ($\alpha_0 = 0.47$) for the EURGBP pair. The narrowest multifractal spectrum of $\Delta\alpha = 0.05$ is shown by the AUDJPY currency pair at the point $\alpha_0 = 0.42$. All of the currency pairs have their maxima in the maximum interval $\alpha_0 \in (0.39; 0.5)$, which indicates little anti-persistence of the series, or the total absence of the linear correlation (GBPUSD and AUDUSD). For the maxima of the determined spectra, the obtained results seems to be consistent with the results obtained using the MF-DFA method. By contrast, for the widths of the spectra, the discrepancies are large. Using formula (**47**) the asymmetry factor has also been determined, which assumes only positive values for all currencies, which means the left-hand asymmetry. The greatest value $A_\alpha = 0.98$ is for the GBPUSD pair, while the least value $A_\alpha = 0.17$, for the EURUSD pair. The values assumed by A_α are greater than those determined with the MF-DFA algorithm, which

may suggest that for FX market data the WL method is more sensitive to large fluctuation.

For the volatility, the British pound-based exchange rates exhibit the widest spectra (spectra maxima); they are very wide, amounting to $\Delta\alpha = 0.36$ ($\alpha_0 = 0.77$) for the GBPUSD pair, $\Delta\alpha = 0.34$ ($\alpha_0 = 0.78$) for the EURGBP pair and $\Delta\alpha = 0.27$ ($\alpha_0 = 0.76$) for the GBPJPY pair. The AUDJPY currency pair has the narrowest spectrum, indicating a practically monofractal behaviour of the cross-correlation, and amounting to $\Delta\alpha = 0.03$ ($\alpha_0 = 0.7$). All of the obtained spectra have similar shapes, compared to the MF-DFA method, but no negative fractal dimensions, so well visible for the pound-based pairs, were demonstrated with the *WL* method. Moreover, for the dollar-based currency pairs (except for GBPUSD, where $A_\alpha = 0.67$), the right-hand spectrum portion, being responsible for low volatility, is very well developed, whereas it is practically invisible for the MF-DFA method. The determined asymmetry factors are: $A_\alpha = -0.49$ for the AUDUSD pair, $A_\alpha = -0.57$ for EURUSD and $A_\alpha = -0.02$ for USDJPY. The remaining exchange rates have left-hand asymmetric spectra.

Similarly as for the FX data, all DJIA index companies were examined using the Wavelet Leaders method (Figure 4.13). The maxima of the spectra $f(\alpha)$ are located in the range $\alpha \in (0.46, 0.59)$, which is largely consistent with the results obtained using the MF-DFA method, where all the results showed a slight persistence of the signal. The average spectrum width $\Delta\alpha$ is 0.5, as compared to $\Delta\alpha = 0.46$ determined by the MF-DFA method. The widest spectrum of $\Delta\alpha = 0.72$ is shown by the AIG company, while the narrowest spectrum of $\Delta\alpha = 0.28$, by XOM.

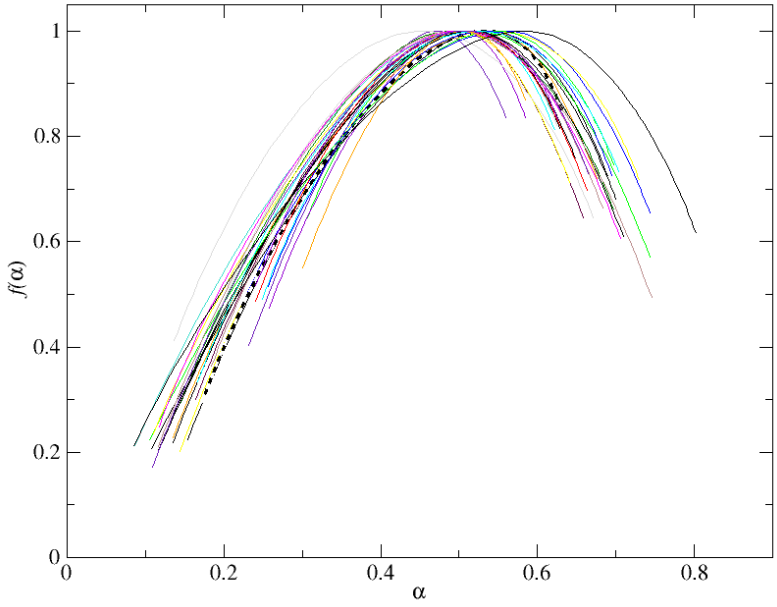

Figure 4.13 Spectra exhibited by DJIA index companies and the index itself (the dashed line), determined by the Wavelet Leaders method.

By comparing the sector analysis given in Table 4.2 with the sector analysis results obtained using the MF-DFA method (Table 4.1), it can be noticed that the Wavelet Leader method indicates a much greater fractal complexity of the examined time series. Only for the Basic Materials sector companies are the obtained results consistent with those obtained by the MF-DFA method.

Moreover, for all companies, the asymmetry factor was calculated from formula (**47**). All of the obtained values of A_α are positive, which means a left-hand spectrum asymmetry. The least value of $A_\alpha = 0.003$ is exhibited by AXP (American Express Company), which implies a nearly symmetric spectrum, while the greatest value of $A_\alpha = 0.58$ is shown by INTC (Intel). In the sector approach, A_α takes on values (Table 4.2) slightly smaller than those determined using the MF-DFA method, which may be due to the fact that for financial data series the WL method better captures the complexity of time series at the smallest fluctuation level.

Table 4.2 The spectrum width $\Delta\alpha$, as dashed down by sectors, determined by the Wavelet Leaders method.

Sector	$\overline{\Delta\alpha}$	A_α
Basic Materials	0.38	0.46

Technology	0.47	0.43
Services	0.48	0.43
Industrial Goods	0.52	0.38
Healthcare	0.56	0.41
Financial	0.56	0.28
Consumer Goods	0.57	0.38

To sum up, the Wavelet Leaders method is a useful tool for examining the fractality of a time series, though, as compared to the MF-DFA method, it is more complex in implementation. The results obtained using this method seem to be largely consistent with the results obtained by the MF-DFA method, which confirms their reliability.

4.3 Multifractal Cross-Correlation Analysis (MFCCA)

Many signals coming from complex systems exhibit relationships of not only the autocorrelation, but also cross-correlation type. To determine the level of such correlation, the Pearson correlation coefficient is commonly used, which identifies only linear relationships. The already mentioned DCCA method, introduced to use by Podobnik and Stanley [65] enables the determination of the coefficient representing power-law cross-correlations between two non-stationary signals. Generalizations on the multifractal case, which have been proposed so far, are MF-DXA introduced by Wei-Xing Zhou [66] and the MF-HXA *(Multifractal Height Cross-Correlation Analysis)* method [83] proposed by Ladislav Kristoufek, which relies on the description of scaling of covariance function of order q. The MF-DXA method, used so far for the analysis of data from areas, such as seismology [84], finances [85; 86] or physics [87], involves the determination of the detrended covariance function of order q between examined time series, although it has one major limitation. In the case of being used for real data, the fluctuation function (formula (**41**)) may assume negative values in this method, which prevents the determination of the scaling exponent for all values of parameter q. Literature provides also proposals for removing this limitation [88], chiefly by introducing a module to the detrended covariance function. However, as will be demonstrated below, this operation does not allow the correct identification of the power-law cross-correlation between two processes.

Study [67] has proposed a new MFCCA (*Multifractal Cross-Correlation Analysis*) method for determining the fractal cross-correlation between two time series, which is free from the limitation of the previous methods. A key component of this algorithm is the correct use of the fluctuation sign, whereby it is possible to determine the covariance function of order q for the complete spectrum. The utility of the model will be shown based on the example of data generated using an appropriate module, as well as real data, as mentioned above. The relationship between the generalized Hurst exponent and parameter λ_q as determined using the MFCCA algorithm, will also be subjected to analysis. The above-mentioned relationship provides information on the strength of cross-correlation in the examined process.

The MFCCA method consists of the following steps, and its flow chart is shown in Figure 4.14. As has been mentioned above, the MFCCA is a generalization of the DCCA method [89] hence the first steps are the same for both of them. We consider two time series, x_i and y_i, where $i = 1,2 \ldots, N$ and N is the length of the series under analysis.

1) For the series, we determine the so-called profile from the following formula:

$$X(j) = \sum_{i=1}^{j} [x_i - \bar{x}], \qquad Y(j) = \sum_{i=1}^{j} [y_i - \bar{y}]. \tag{52}$$

   The mean values in the formula are calculated for the entire series.

2) Each of the signals is divided into $M_s = N/s$ non-overlapping intervals v of a length s. To make sure that all time series points are taken into account in calculations, the entire series is divided twice, once starting from the beginning, and once starting from the end. As a result, $2M_s$ intervals are obtained.

3) For each interval v in the series, the trend is calculated using a polynomial of degree m: $P_{X,v}^{(m)}$ for X and $P_{Y,v}^{(m)}$ for Y. Based on the results obtained in study [63], a polynomial of degree $m = 2$ has been chosen for all calculations.

4) For each segment, the trend is subtracted from the examined series and the detrended cross-covariance is calculated:

$$F_{xy}^2(v,s) = \frac{1}{s}\sum_{k=1}^{s}\left\{\left(X((v-1)s+k) - P_{X,v}^{(m)}(k)\right) * \left(Y((v-1)s+k) - P_{Y,v}^{(m)}(k)\right)\right\}. \tag{53}$$

In contrast to the MFDFA method, $F_{xy}^2(v,s)$ can assume both positive and negative values. Therefore, for the analysis to be correct, the sign of $F_{xy}^2(v,s)$ is indispensable.

5) The next step involves the calculation of the covariance function of degree q, the so-called fluctuation function, averaged over all intervals, using the formula below:

$$F_{xy}^q(s) = \frac{1}{M_s}\sum_{v=1}^{M_s} \text{sign}(F_{xy}^2(v,s))\left|F_{xy}^2(v,s)\right|^{q/2}. \tag{54}$$

In the above formula, $\text{sign}(F_{xy}^2(v,s))$ denotes the sign of function $F_{xy}^2(v,s)$, and parameter q may take on real values and enables series to be analyzed, depending on the signal size. For $q > 0$, the predominant component of the fluctuation function will be large fluctuations, while for $q < 0$, small fluctuations will be amplified, and their contribution to the fluctuation function will be the greatest. What is very important is that when determining the function $F_{xy}^q(s)$, the original sign of amplified (or attenuated) fluctuations will be preserved. This allows the numerical errors of the method due to raising negative values (negative fluctuations) to the power to be avoided. As a result, no complex values occur during calculation and, in addition, no information coming from negative signal values is lost.

For $q = 0$, this formula should be used [62]:

$$F_{xy}^0(s) = \frac{1}{M_s}\sum_{v=1}^{M_s} \text{sign}(F_{xy}^2(v,s))\ln\left|F_{xy}^2(v,s)\right|. \tag{55}$$

The procedure must be repeated for different times scales, s. So obtained fluctuation function $F_{xy}^q(s)$ should scale itself (i.e. be linear on the log-log scale), if a fractal cross-correlation exists between the examined series. If the function fluctuates around zero, or assumes alternately negative and positive values by intervals, then in that case there is no cross-correlation in the signal. In the case, where the function $F_{xy}^q(s)$ is negative for every scale s, then the function $-F_{xy}^q(s)$ is examined.

The power-law character enforces the following relationship:

$$\left[F_{xy}^q(s)\right]^{1/q} = F_{xy}(q,s) \sim s^{\lambda_q}. \tag{56}$$

For $q = 0$, the following relationship is satisfied:

$$\exp(F_{xy}^0(s)) = F_{xy}(0,s) \sim s^{\lambda_0}. \tag{57}$$

λ_q is the exponent characterizing the multifractal cross-correlation. In the event, when λ_q is not dependent on q, then this parameter is equal to λ as obtained using DCCA, and we have a monofractal cross-correlation, otherwise, a multifractal cross-correlation. It is important to note that $\lambda_q = \lambda$ for $q = 2$. The values of scales s_{min} and s_{max} are selected depending on the length of signal N. It is assumed that $s_{max} < N/5$. Figure 4.14 shows a flowchart of the MFCCA algorithm.

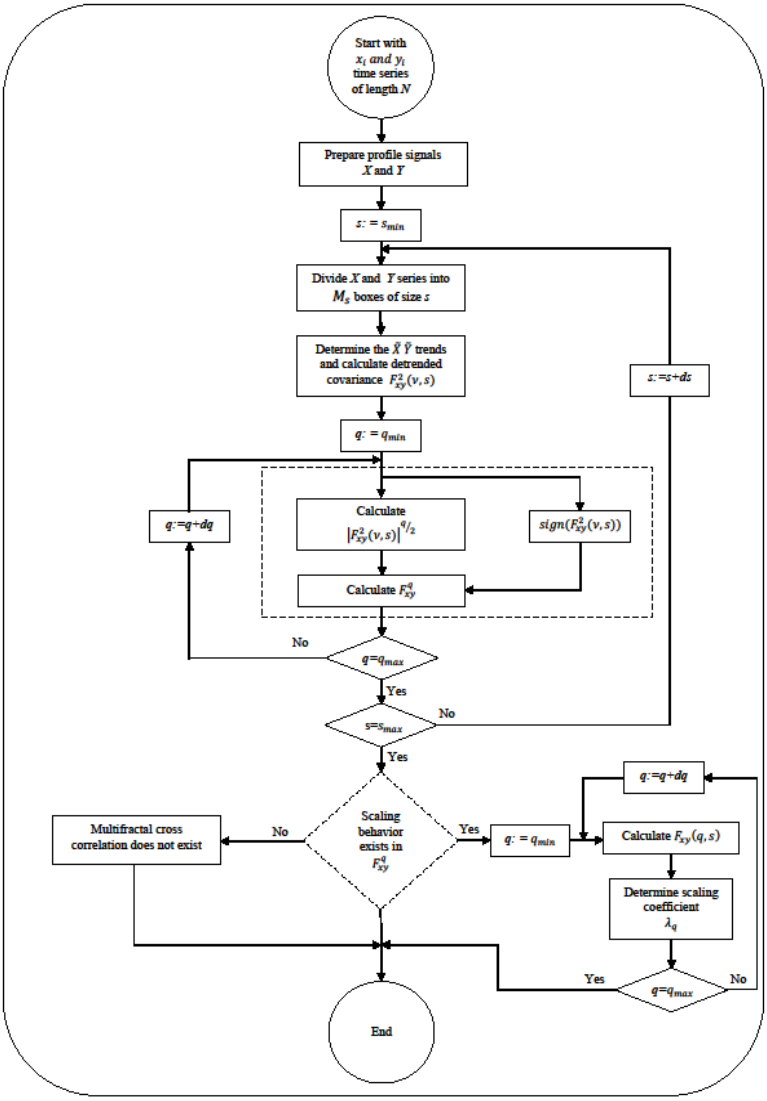

Figure 4.14 A flowchart of the MFCCA algorithm.

To demonstrate the usefulness of the method and its advantage over the MF-DXA method, the following time series were examined:

a) two uncorrelated FARIMA processes [90],

b) two correlated FARIMA processes.

They are an example of monofractal long-range correlated processes. FARIMA processes are non-stationary. To generate a pair of FARIMA processes, the following formulae were used:

$$x_i = \sum_{j=1}^{\infty} a_j(d_x) - x_{i-j} + \epsilon_i, \tag{58}$$

$$y_i = \sum_{j=1}^{\infty} a_j(d_y) - x_{i-j} + \epsilon_i, \tag{59}$$

in which ϵ_i denotes Gaussian noise. In order to correlate two FARIMA processes, the same ϵ_i noise component should be used for generating the both series.

The parameters $d_{x(y)}$ $(-1/2 < d_{x(y)} < 1/2)$ are responsible for the process stationarity and characterize the linear, long-range correlations in the signal. For $d < 0.5$ we have a stationary process, for $0.5 < d < 1$, a non-stationary sub-diffusion process, and for $d > 1$, a super-diffusion process. They are related to the Hurst exponent with the following formula [67]:

$$H = \frac{1}{2} + d_{x(y)}. \tag{60}$$

The weight $a_j\left(d_{x(y)}\right)$ in formulae (**58**) and (**59**) is given by this formula:

$$a_j\left(d_{x(y)}\right) = \frac{\Gamma\left(j - d_{x(y)}\right)}{\left[\Gamma\left(-d_{x(y)}\right)\right]\Gamma(1+j)}, \tag{61}$$

where Γ is the Gamma function. A proposition can be found in the literature [65] which says that the average of individual fractal properties $h_{xy}(q)$ is related to the fractal properties of the two examined series with the following formula:

$$h_{xy}(q) = \frac{h_x(q) + h_y(q)}{2}, \tag{62}$$

where $h_{x(y)}(q)$ is the generalized Hurst exponent, which refers to the fractal properties of the series. For $q = 2$, $h_{x(y)}(q)$ is normal Hurst exponent of a single time series. The relation between $h_{xy}(q)$ and λ_q based on empirical results [67] depends on the time structure of the analyzed series defined by the Hurst exponent. The difference between these two values of $\Delta = \lambda_q - h_{xy}(q)$ defines the degree of similarity between fractal structures (the degree of series correlation). For the same signals $h_{xy}(q) = \lambda_q$, but for signals with different Hurst exponents, the difference is more significant. For FARIMA processes, the relationship:

$$\lambda_q \approx \frac{h_x(q) + h_y(q)}{2} \tag{63}$$

is satisfied, if and only if the time series under analysis have a similar structure and the difference between $h_x(q)$ and $h_y(q)$ is small. This situation is illustrated in the right-hand panel in Figure 4.15, where the relation between λ_q and $h_{xy}(q)$ is shown. In turn, the left-hand panel in Figure 4.15 shows the fluctuation function determined by the MFCCA method for two correlated FARIMA processes with different Hurst exponent values: a) $H_1 = 0.3$, $H_2 = 0.5$ b) $H_1 = 0.3$, $H_2 = 0.8$ c) $H_1 = 0.5$, $H_2 = 0.8$. Each value of q has one corresponding line in the diagram.

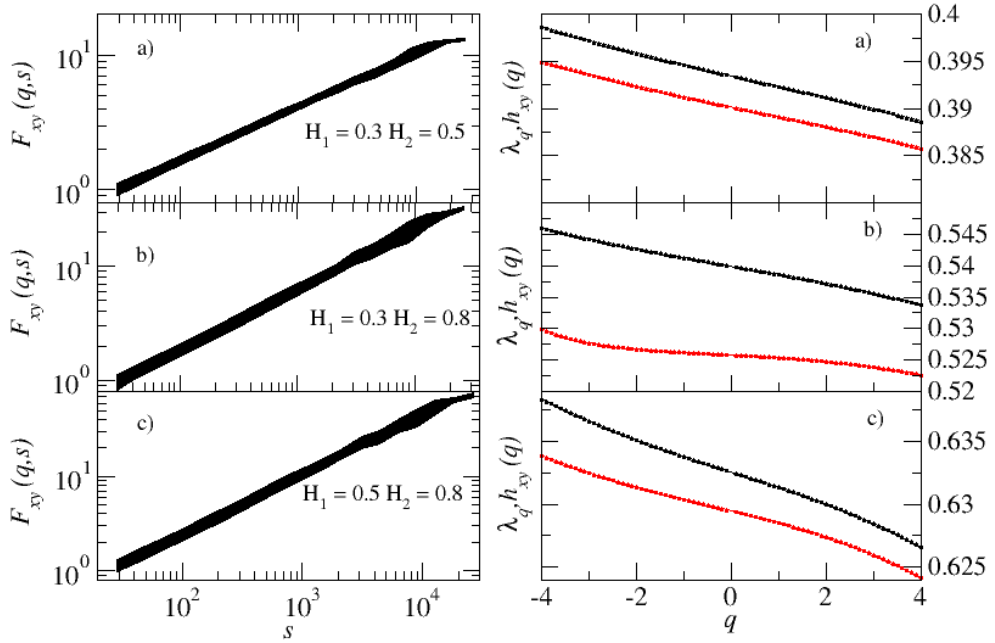

Figure 4.15 The fluctuation function $F_{xy}(q,s)$ determined for three different pairs of processes with a different H exponent (left-hand side) and the determined scaling exponent λ_q and the mean generalized Hurst exponent $h_{xy}(q)$ (right-hand side).

The power-law nature of the function $F_{xy}(q,s)$ is clearly seen within the entire scale range, and the functions lie nearly in parallel to one another, which indicates a uniform character of the cross-correlation occurring between the signals. The difference between the scaling exponent extrema, $\Delta\lambda_q = \max(\lambda_q) - \min(\lambda_q)$, reflecting the complexity of the fractal cross-correlation, is, respectively, 0.0021 (a), 0.003 (b) and 0.0039 (c). Thus, the cross-correlation between the selected FARIMA processes is homogeneous. The parameter $\Delta = \lambda_q - h_{xy}(q)$ assumes the greatest value in case (c), where the difference between the Hurts exponents is the largest.

To better compare the fluctuation function determined for each process separately and the function determined by the cross-correlation method, all the three functions for $q = 2$ are shown in Figure 4.16. For the correlated processes (panel a), cross-correlation fluctuation functions are clearly visible, which scale in a power-law manner from the scale $s = 20$ to the scale $s = 3000$, with the scaling exponent determined for them satisfying approximately relationship (**62**). In the case of uncorrelated processes (panel b), it is not possible to determine the cross-correlation fluctuation function with a single type of scaling.

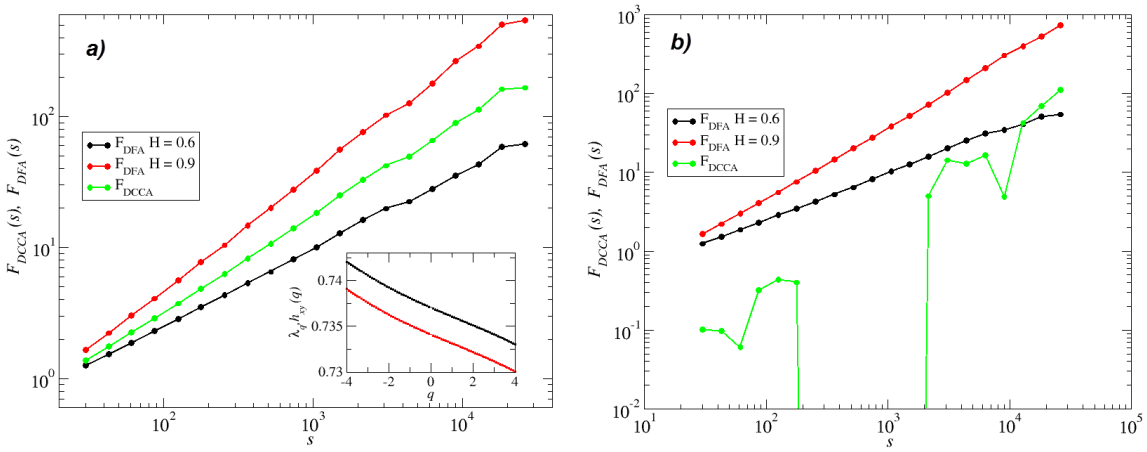

Figure 4.16 The cross-correlation fluctuation function (the green line) determined for two processes with Hurst exponents of H=0.6 (the black line) and H=0.9 (the red line). Panel a) shows correlated processes, while panel b), uncorrelated processes.

Interestingly enough, the MF-DXA algorithm used on the example of uncorrelated data incorrectly identifies non-linear cross-correlations between processes. This situation is illustrated by Figure 4.17. The fluctuation function was determined for each of the moments $q \in [-4,4]$, but due to the fact that for $q < 0$ the function fluctuates heavily, $F_{xy}(q,s)$ is shown in the figure only for positive q values. Scaling is very good from the scale $s = 30$ to $s = 2000$, and the fluctuation functions are parallel to one another, which indicates a monofractal behaviour of the cross-correlation. A confirmation of this fact is the inset in the figure showing λ_q and $h_{xy}(q)$, which vary only slightly with the increase in q. This result is incorrect, as no correlation may exist between two uncorrelated processes.

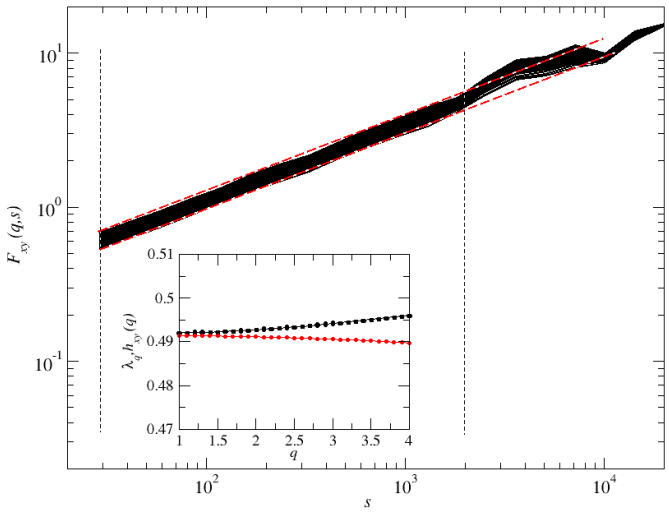

Figure 4.17 The fluctuation function $F_{xy}(q,s)$ for two uncorrelated FARIMA processes, determined with the MF-DXA algorithm. The inset shows the scaling exponent λ_q and $h_{xy}(q)$, as dependent on the parameter q. The black dashed line indicates the scaling interval, and the red dashed line, the power-law fitting to the extreme fluctuation function values.

To determine the degree of correlation differentiation at the level of different signal amplitudes, the following parameter is introduced:

$$d = \Delta\lambda_q - \Delta h_{xy}(q), \tag{64}$$

where $\Delta h_{xy}(q)$ is the mean of the fractal properties determined for the series X and Y. The case of $d = 0$ means that for all values of parameter q there occurs the same correlation between signals for different fluctuation amplitudes. The case of $d \neq 0$ means the change in correlation strength at the level of small and large fluctuations.

4.3.1 The foreign exchange market

For examining the multifractal cross-correlation, the MFCCA algorithm, described in the previous section, was used. The fluctuation functions were determined for parameter $q \in [-4,4]$. For moments $q < 0$, the fluctuation function cannot be determined (which indicates the lack of cross-correlation at the small fluctuation level), therefore $q \in [1,4]$ was assumed for analysis. The analysis covered 5-second returns series of a length of $N = 86400$ for 20 time scales, starting from $s = 30$ through to $s = N/5$. To assure the compactness of the carrier, the intervals of 20 and more zeros were removed from the series. Figure 4.18 shows a sample fluctuation function $F_{xy}(q,s)$, determined for the rates of return (panels a and c) and their volatility (panels b and d) between the exchange rates EURUSD/USDJPY and AUDJPY/EURAUD. Each

line corresponds to one value of q. The insets in the figure contain information on how the scaling exponent λ_q (in black colour) and the mean generalized Hurst exponent $h_{xy}(q)$ (in red colour, formula (**62**)), change with the increase in q. For the returns and for volatility, the fluctuation function $F_{xy}(q,s)$ exhibits a power-law behaviour for scales from $s = 30$ to $s = 3000$ (for the returns and from $s = 30$ to $s = 2000$ (for the volatility). This is indicative of a fractal nature of the cross-correlation of these processes. In all of the examined cases, the value of the scaling exponent λ_q varies slightly with the increase in q, which suggests a multifractal nature of the cross-correlation, both for the returns and for the volatility. Interestingly enough, the mean generalized Hurst exponent $h_{xy}(q)$ for the rates of return varies slightly, whereas for the volatility is practically constant.

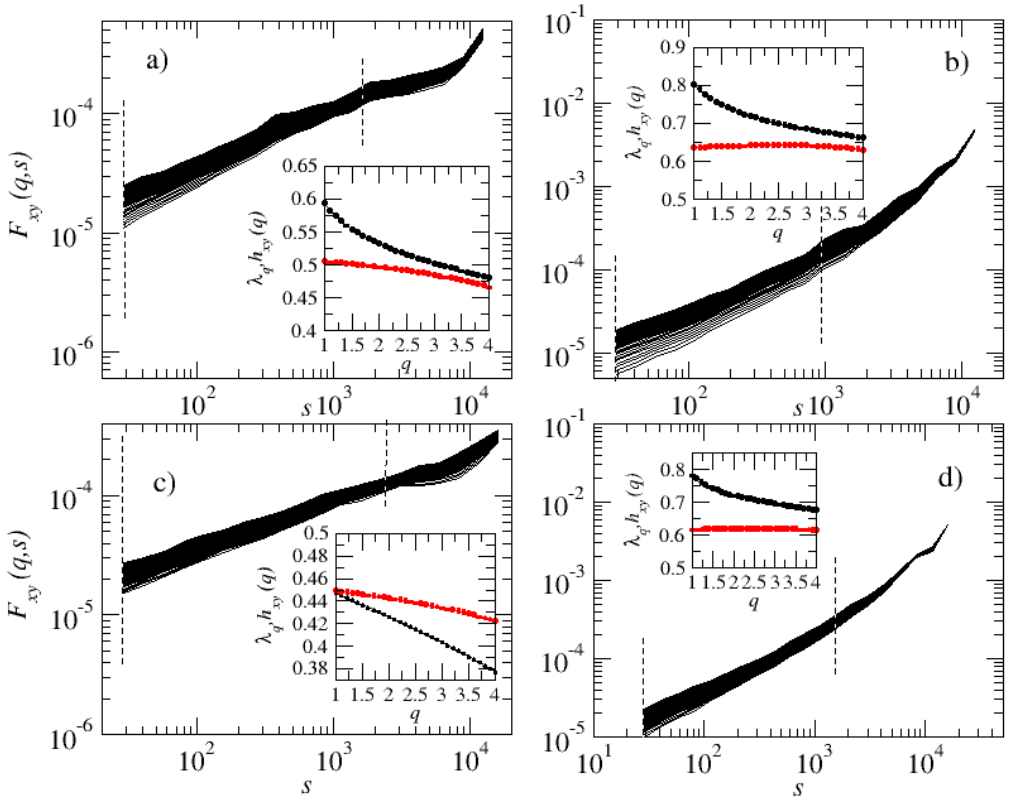

Figure 4.18 (On the left) The fluctuation functions $F_{xy}(q,s)$ determined for two sample rates of return (EURUSD/USDJPY, the top panel, and AUDJPY/EURAUD, the bottom panel). The dashed line indicates the scaling interval. (On the right) The fluctuation function determined for the volatility of these currency pairs. The insets show the scaling exponents λ_q and $h_{xy}(q)$, as dependent on the parameter q.

It should be underlined here that in the case of the rates of return for the currency pairs outside the triangle, AUDJPY/EURGBP and EURUSD/GBPJPY, scaling is much worse

and occurs within a narrow scale interval. In Figure 4.20 , these pairs are indicated by the filling. For the currency pairs EURAUD/GBPJPY, GBPJPY/AUDUSD, AUDJPY/GBPUSD, on the other hand, we can essentially speak of lack of scaling (these pairs are fractally uncorrelated), therefore the estimation of the exponent λ_q was impossible.

The performed analysis shows that all of the currency pairs exhibit cross-correlation both on the returns level, as well as on the volatility level, for currencies forming triangles. Generally, it can be accepted that in the case of volatility for currencies outside the triangle, the λ_q scaling exponent deviation from the mean value of generalized Hurst exponents $h_{xy}(q)$ is greater than for currencies inside the triangle (Figure 4.19).

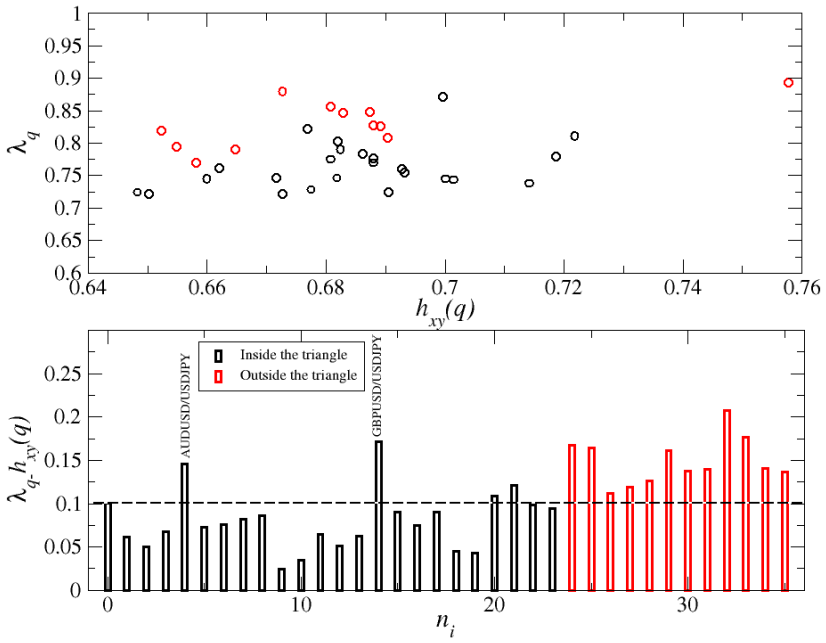

Figure 4.19 The differences between the scaling exponent λ_q and $h_{xy}(q)$ determined for the volatility (q=2).

The situation is slightly different for the returns (Figure 4.20), where it is hard to notice a similar behaviour. There is no feature that would enable one to capture the difference between currencies inside the triangles and those outside the triangle, in both quantitative and qualitative terms.

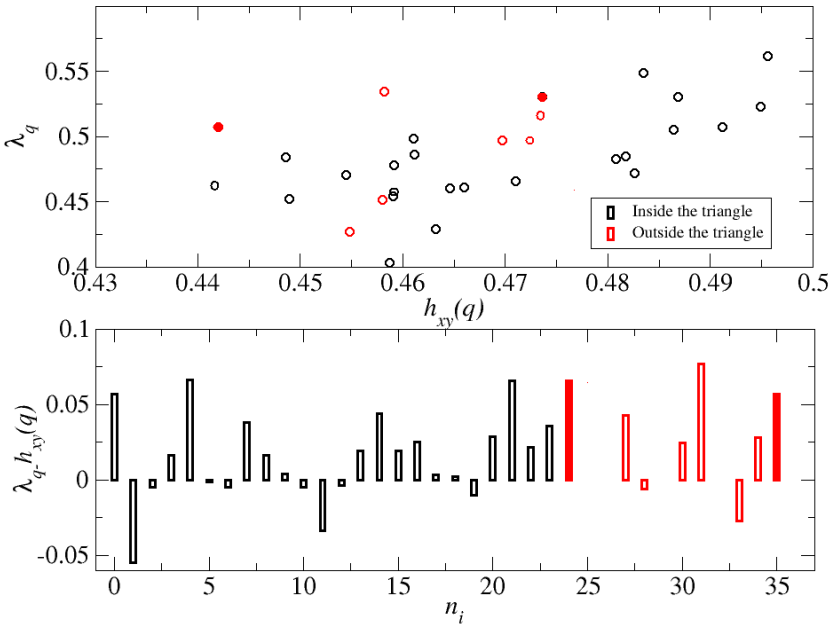

Figure 4.20 The differences between the scaling exponent λ_q and $h_{xy}(q)$ determined for the rates of return and q=2. The filling indicates those pairs, where the fluctuation function scaling is weak.

For all combinations of currencies inside the triangle and outside the triangle, the level of d correlation differentiation was determined (formula (**64**)), as shown in Figure 4.21. Similarly as for the monofractal analysis presented for parameter $q = 2$ (Figures 4.19 and 4.20), so for the multifractal analysis, a difference can be noticed between currencies inside the triangle, as well as outside the triangle, for the volatility (the bottom panel), where d assumes large values, so a change in the strength of correlation takes place at a different fluctuation level. This effect seems to be more pronounced for the currencies outside the triangle. For the rates of return, the pairs of currencies inside triangle cannot be qualitatively distinguished from the pairs of currencies outside the triangle (the top panel).

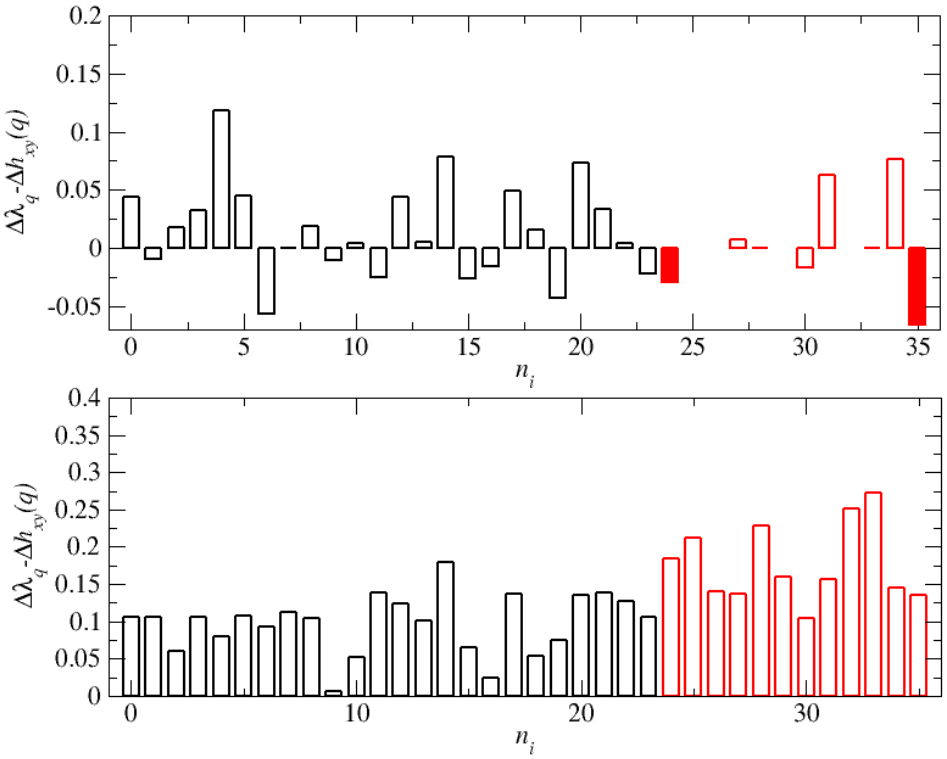

Figure 4.21 The difference between $\Delta\lambda_q$ and $\Delta h_{xy}(q)$ for all currency pairs. The top figure represents the rates of return, while the bottom figure, the volatility. The currencies inside the triangle are marked in black, while the currencies outside the triangle, in red. Those pairs, where the fluctuation function scaling is poorly noticeable is marked by the filling.

4.3.2 The stock market

The quotations of DJIA index companies were also examined for multifractal cross-correlation. Due to the fact that for the General Electric (GE) company the analyzed data were incomplete (data from the period 1.10.2009-03.01.2010 was missing), all time series were truncated to this company's series length that was $t = 172369$ points. All possible combinations of 29 companies were examined in the interval of $-4 \leq q \leq 4$. For all values of parameter $q < 1$, the fluctuation function $F_{xy}(q, s)$ has no well defined scaling for any of the pairs analyzed. It very heavily fluctuates near zero and does not assume the same sign for different values of *s* even by intervals, which is the condition necessary for correlation to occur. This shows that no significant correlations were observed in the signal for small fluctuations, while a reciprocal fractal correlation may only exist at the level of large correlations occurring in the signal. Therefore, the further analysis was limited to the values of $1 \leq q \leq 4$.

An example representative of all combinations of the examined companies, i.e. the fluctuation function determined for the companies CSCO (Cisco Systems, Inc.) and MCD (McDonald's), and the dependence of λ_q and $h_{xy}(q)$ on the parameter q, are represented in Figure 4.22. Two different types of behaviour of the scaling exponent λ_q relative to the averaged generalized Hurst exponent $h_{xy}(q)$ of individual series are indicated there. In the upper panel of Figure 4.22 it can be seen that the values of λ_q are distinctly greater than those of $h_{xy}(q)$, the functions λ_q and $h_{xy}(q)$ are decreasing, and the difference between them is constant in terms of increase in q. This means that the companies are correlated, the complexity of this correlation is not dependent on the fluctuation amplitude and the fractal characteristic of the both processes is different. The fractal strength of the cross-correlation, expressed as $\Delta\lambda_q$, for the above-mentioned companies is 0.13, while the difference in the generalized Hurst exponent (for the extreme values of q is $\Delta h_{xy}(q) = 0.13$.

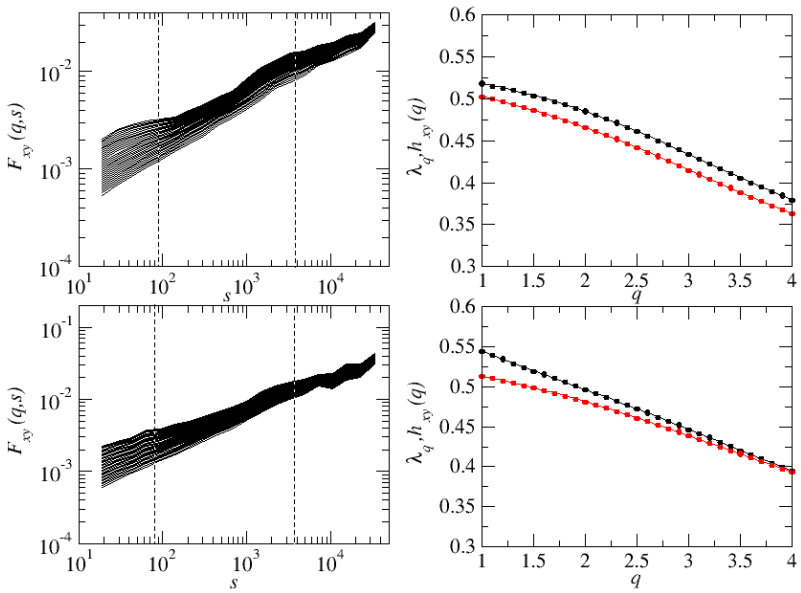

Figure 4.22 The fluctuation functions $F_{xy}(q,s)$ and the magnitudes of λ_q (in black) and $h_{xy}(q)$ (in red) determined for two pairs of companies: CSCO (Cisco Systems, Inc.) and MCD (McDonald's) (the top panel) and DD (Disney) and VZ (Verizon) (the bottom panel). The dashed line indicates the scaling intervals.

However, there is a group of companies, whose fractal properties are different from those described above, since the difference between λ_q and $h_{xy}(q)$ varies with q, and for q close to 4 it is practically zero. This is indicative of the fact that the dynamics of the processes changes with the fluctuation amplitude and its increase, becoming more and more similar to these processes. This can be observed in the bottom panel in Figure 4.22 (the companies DD-Disney and VZ-Verizon), where $\Delta h_{xy}(q) = 0.12$ and $\Delta\lambda_q = 0.15$. Among all the analyzed combinations of companies, the existence of non-linear internal correlations are observed, which are quite significant in view of the fact that $\Delta\lambda_q$ assumes quite considerable values.

Similarly as for the currencies, for all possible combinations among the companies, the parameter d was determined. The histogram in Figure 4.23 shows that little correlation complexity prevails in this set (small and big fluctuations are correlated in a similar manner).

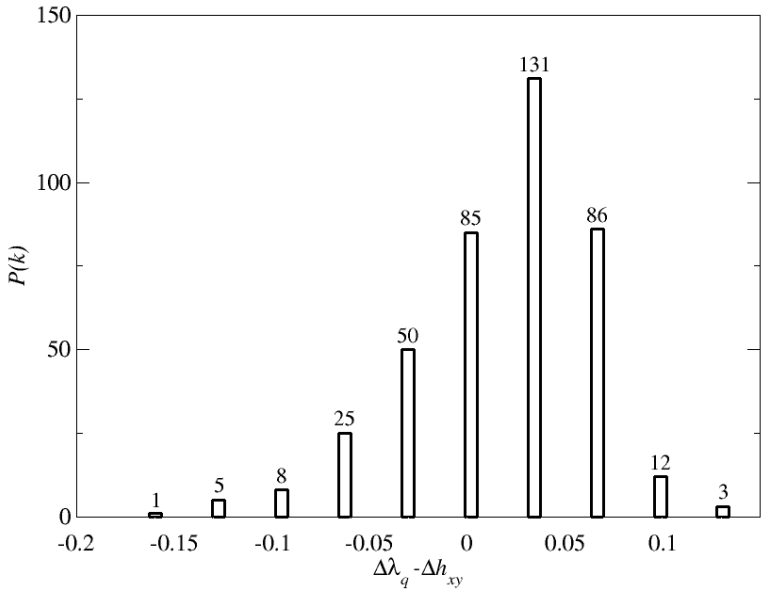

Figure 4.23 The histogram representing the value of d determined for all combinations of 29 DJIA index companies.

The mean value for all set is $d \approx 0.005$, and the values dashed down by sectors are given in Table 4.3. The above shows that the greatest correlation complexity is exhibited by the Services sector companies, while the companies from the Healthcare sector are practically homogeneous in terms of correlation at a varying fluctuation level.

Table 4.3 The value of the d coefficient in sectors break down .

Basic Materials	$d \approx 0.03$
Consumer Goods	$d \approx -0.04$
Financial	$d \approx 0.03$
Healthcare	$d \approx 0.01$
Industrial Goods	$d \approx 0.03$
Services	$d \approx -0.05$
Technology	$d \approx 0.02$

4.4 Determination of the cross-correlation coefficient ρ^q for time series

As has been shown previously, the analysis of cross-correlation between two stable time series is conducted based on the Pearson correlation coefficient, which defines the linear relationship between them. For financial times series, which are not

stable and, as has already been demonstrated, are mutually related in a non-linear manner, such correlation analysis might not be sufficient.

As has been mentioned above, if there exists a cross-correlation between two time series, the following relationship is valid:

$$F_{xy}(q,s) \sim s^{\lambda_q}, \tag{65}$$

where the scaling exponent λ_q determines whether there is a fractal correlation between the series, or not. For the determination of the cross-correlation strength on a specific scale, the coefficient ρ_{DCCA} defined by formula [91] below is proposed:

$$\rho_{DCCA}(s) \approx \frac{F_{xy}^2(s)}{F_x(s)F_y(s)}. \tag{66}$$

Using this coefficient, the cross-correlation level can be determined for non-stationary data on different time scales, and it is not associated with the fractal properties of the signals under analysis. It is the capability to work on non-stationary data that is the essential advantage of $\rho_{DCCA}(s)$ over the Pearson correlation coefficient. The $\rho_{DCCA}(s)$ coefficient is based on the two methods referred to earlier, i.e. DFA and DCCA, and due to the fact that it is defined on a single scale, examines signals can be either fractal or not. Its usefulness has been demonstrated on the example of climatic [92] or stock exchange [91] data.

The $\rho_{DCCA}(s)$ coefficient assumes the following values [93]:

$$\rho_{DCCA}(s) = \begin{cases} 1 - \text{perfect positive correlation between signals} \\ 0 - \text{no correlation between signals} \\ -1 - \text{perfect negative correlation between signals} \end{cases}. \tag{67}$$

It helps not only to determine the strength of correlation between two signals, but also to identify the occurrence of seasonal effects in the signals under analysis. Figure 4.24 shows the $\rho_{DCCA}(s)$ coefficient determined from exchange and financial market data. For the FX market (panel a), a difference is clearly seen for the exchange rates inside the triangle (black colour), where the cross-correlation assumes negative values for all examined scales and is constant for every s, in contrast to the exchange rates outside the triangle, where the cross-correlation takes on values close to zero for all scales. This is a behaviour typical of all the examined currency combinations contained in Table 2.5. The function $\rho_{DCCA}(s)$ behaves differently for volatility (panel c). For the currencies inside the triangle, the $\rho_{DCCA}(s)$ function assumes positive values from the smallest scales and additionally increases with increasing s, attaining 1 with the largest scale. For the currencies outside the triangle, the behaviour is similar, except that for the

smallest scales the values of $\rho_{DCCA}(s)$ are much lower. For sample financial market returns (panel b), a positive cross-correlation can be observed at a constant level for all scales. The coefficient/function $\rho_{DCCA}(s)$ determined among volatility series (panel d) also shows a positive cross-correlation, though it starts fluctuating for higher scales.

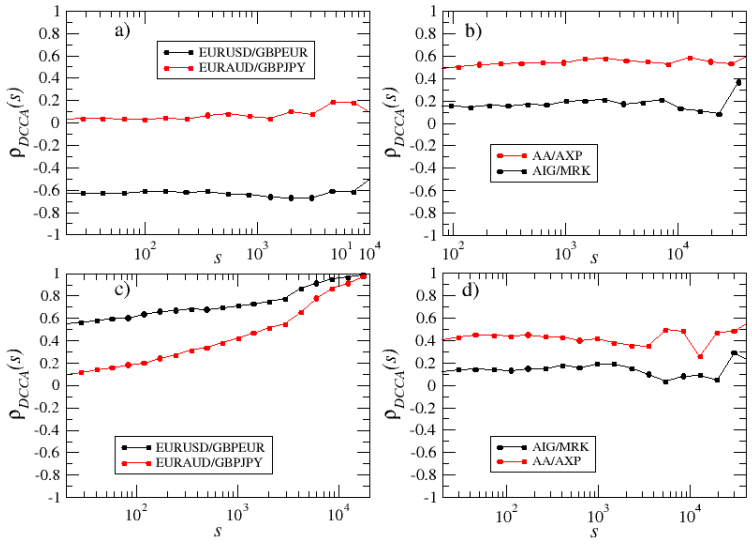

Figure 4.24 The correlation coefficient determined for sample currency pairs (panel a, returns series and c, volatility series) and companies (panel b, return series and d, volatility series).

It seems, however, that so estimated correlation has one limitation. The component functions of the cross-correlation coefficient rely on plain variance and covariance functions, therefore it is not possible to estimate whether the potential correlation comes from varying-amplitude fluctuation to the same extent, or some of the fluctuations have a smaller or greater contribution to this value. To overcome this limitation, study [94] has proposed a q-generalized cross-correlation coefficient which is sensitive to fluctuation amplitude, as shown below:

$$\rho^{q}_{MFCCA}(s) = \frac{F^{q}_{xy}(s)}{F^{\frac{q}{2}}_{x}(s)F^{\frac{q}{2}}_{y}(s)}, \tag{68}$$

where $q \in \mathbb{R}$, F^2_{xy} is the detrended cross-covariance function determined by the MFCCA method, while F^q_x and F^q_y are detrended fluctuation functions of order q, determined by the MF-DFA method for individual time series. It should be noted that the detrended fluctuation function used in the numerator, determined using the MFCCA method, is crucial to the obtaining of reliable results, owing to the fact that is correctly interprets the fluctuation sign. For $q = 2$, formula (**68**) turns into formula (**66**). When using $\rho^{q}_{MFCCA}(s)$, we do not look at the fractality of the series; instead, we look at the

correlations for a specific scale s. For all values of parameter $q > 2$, the increasingly large fluctuations have the greatest contribution to ρ^q_{MFCCA}. Similarly to ρ_{DCCA}, the generalized coefficient assumes values in the interval of $\rho^q_{MFCCA} \in [-1,1]$, for every $q \geq 0$. In the case where $q < 0$, the situation is slightly different. So, in the case of uncorrelated or weakly correlated processes, the module of the numerator in formula (**68**) is much greater than the denominator, as a consequence of which the values of ρ^q_{MFCCA} may go beyond the interval -1,1. To avoid this drawback, the coefficient needs to be redefined in the following manner:

$$\rho^q = \begin{cases} \rho^q_{MFCCA} & \text{for} \quad \left|\rho^q_{MFCCA}\right| \leq 1 \\ (\rho^q_{MFCCA})^{-1} & \text{for} \quad \left|\rho^q_{MFCCA}\right| > 1 \end{cases}. \tag{69}$$

Thus, ρ^q will always stay within the interval $[-1,1]$, and this very coefficient, given by formula (**69**), will be used later on in this study. The figures below show a sample value of the ρ^q coefficient for different values of parameter $q \in \{-4, -2, 2, 4\}$, determined for two correlated 4.25 and uncorrelated 4.26 FARIMA processes. In the case of the correlated processes it is clearly seen that $\rho^q \approx 1$, regardless of the value of q. This is due to the fact that FARIMA processes are stationary, while in the case of non-stationary processes one might expect deviations from 1 with the increase in scale In the case of the uncorrelated processes it is clearly seen that $\rho^q \approx 1$, regardless of the value of q. This is due to the fact that FARIMA processes are stationary, while in the case of non-stationary processes one might expect deviations from 1 with the increase in scale s [94].

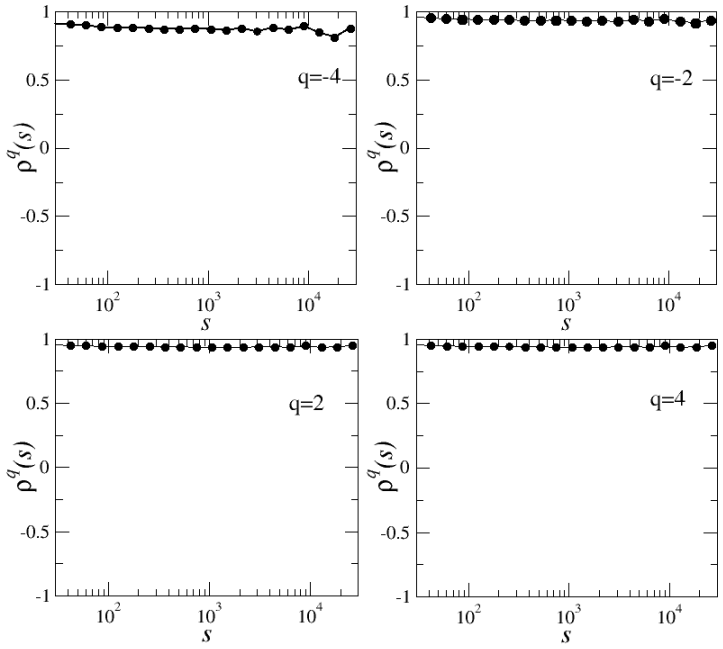

Figure 4.25 The ρ^q coefficient calculated for two correlated FARIMA processes.

For uncorrelated processes (the both time series were subjected to random mixing) the ρ^q coefficient well identifies the absence of correlation between them for $q < 0$; in turn, for $q > 0$, above the scale $s > 3000$, slight deviations from zero are visible in Figure 4.26 . To assess whether these deviations are significant or not, the standard deviation $\overline{\sigma(\rho^q)} = 0.09$ was calculated for 1000 time series as average for all scales, which reaches $\sigma(\rho^q) = 0.25$, for the largest scales. It can be seen, therefore, that even so for the largest scales these deviations are significant, yet they are contained within the limit of error.

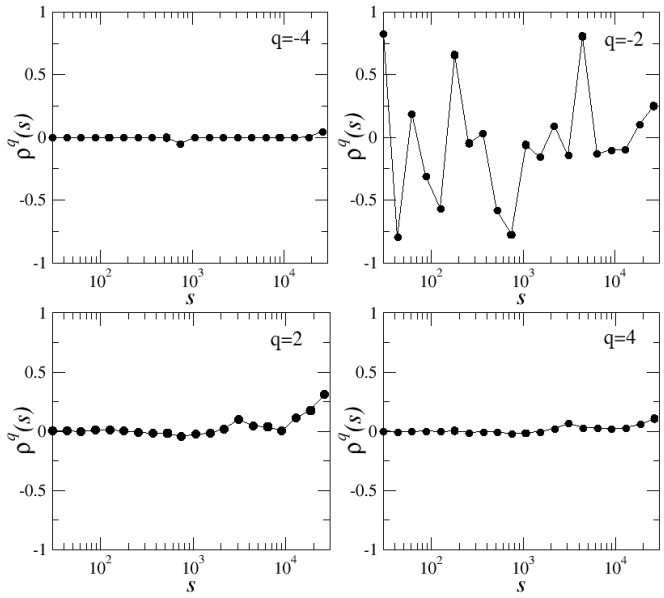

Figure 4.26 The ρ^q coefficient calculated for two uncorrelated FARIMA processes.

It should be emphasized that the detrended fluctuation function used in the numerator, determined using the MFCCA method, is crucial to the obtaining of reliable results, owing to the fact that is correctly interprets the fluctuation sign. Figure 4.27 represents the same two uncorrelated FARIMA processes as in Figure 4.26, except that for the determination of ρ^q, the MF-DXA algorithm was used, in which the function $F_{xy}^2(s)$ (formula (**53**)) is calculated with a module. For $q = -4$, the ρ^q coefficient correctly indicates the absence of correlation for all the examined scales.

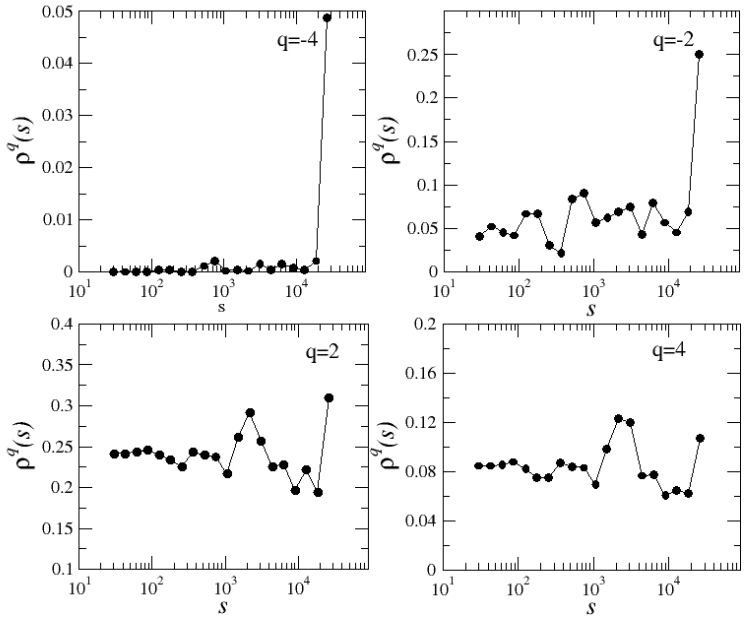

Figure 4.27 The ρ^q coefficient calculated for two uncorrelated FARIMA processes using the MF-DXA algorithm for determining the fluctuation function $F_{xy}^q(s)$.

However, for $q \geq -2$, a clear positive cross-correlation can be noticed throughout the interval, in spite of the fact that such correlation cannot occur between two independent time series. This provides an additional confirmation of the reliability of the MFCCA algorithm, which correctly identifies correlations between time series.

The coefficient ρ^q, introduced in this Chapter and discussed on the example of both artificially generated and real time series, will be used in Chapter 5 for the determination of the network representation of relations between stock exchange listed companies.

5 Network analysis of the financial data – MST trees

This chapter will present and discuss the properties of the network representation of financial markets, understood based on fractal data analysis methods. To the author's knowledge, the method proposed in the present study is the first network representation of this type, where the analysis is made in two dimensions, i.e. with respect both to time scale and to fluctuation amplitude.

A network is best defined as a set of nodes connected with one another by edges [95]. The nodes are the system's components (listed companies) and the edges are interrelations between them (e.g. correlation). So, it can be seen that, from this perspective, the financial market fits perfectly into this pattern. The first study to use network methods for the analysis of cross-correlation on financial markets was that of Mantegna [96] and these methods are currently being successfully employed for the analysis of currency exchange markets [97; 98; 99], commodity markets [100] or stock markets [101; 102; 103]. An advantage of the network formalism is the capability to compare different systems and to isolate the most similar possible features from them. In addition, we get a tool in our hands, which provides information on the system's dynamics and structure in a compact form.

The study of complex networks derives from graph theory and, as an independent research field, has been functioning since, more or less, the mid-20th century. It has found application as a perfect tool for the analysis and description of a wide variety of phenomena and systems [104; 105]. In physics, regular networks (Figure 5.1, panel a) have been in use since long ago, and form a foundation for different theories, e.g. in solid-state physics. A characteristic feature of these systems is a high degree of order, a fixed number of nodes and a small extent of interactions between them; so, this is a deterministic system that is not subject to evolution. Each node in this network has the same multiplicity (degree)[17]. If applied to the description of different dynamic systems, it is little useful. In 1960, two Hungarian mathematicians, Paul Erdös and Alfred Rényi, proposed random graphs for the description of complex networks (model ER) [106]. According to the proposed model, a random network (Figure 5.1, panel b) is constructed using a recurrent procedure. From a set of N unconnected nodes, $N/2$ pairs of nodes are chosen randomly and connected with probability p. In the random graph, the probability of a given node having a multiplicity

[17] Multiplicity (degree) of a node is the number of edges attached to it.

k per $n-1$ possible connections with other nodes is the same as the chance of achieving k successes in $n-1$ trials, which is described by the binomial distribution:

$$P(k) = \binom{n-1}{k} p^k (1-p)^{n-1-k}, \tag{70}$$

which, at the limit $n \gg kz$, where $z = \frac{n(n-1)p}{n} = (n-1)p \approx np$ (mean node degree), turns into the Poisson distribution:

$$P(k) = \frac{z^k e^{-z}}{k!}. \tag{71}$$

This implies that with increasing node multiplicity, the probability of attaching another node decreases, as a consequence of which nodes with a very large number of connections are little probable. Random networks are characterized by a complete lack of regularity of connections between nodes and the absence of their hierarchical ordering. Moreover, phase transitions can be observed in such networks. In the case where the probability p of attaching a node is small, there exist a large number of isolated nodes in the network, and relatively few clusters, compared to the size of the entire network. With the increase in p, these clusters become increasingly large in size and number. In the event when the value of p exceeds the critical value p_c (percolation threshold), transition from the disordered phase (where clusters of sizes much smaller than the network's size only exist) to an ordered phase occurs in the network, whereby a percolation cluster comes into being, which provides the path[18] between arbitrary nodes. This cluster scales itself along with the network size N.

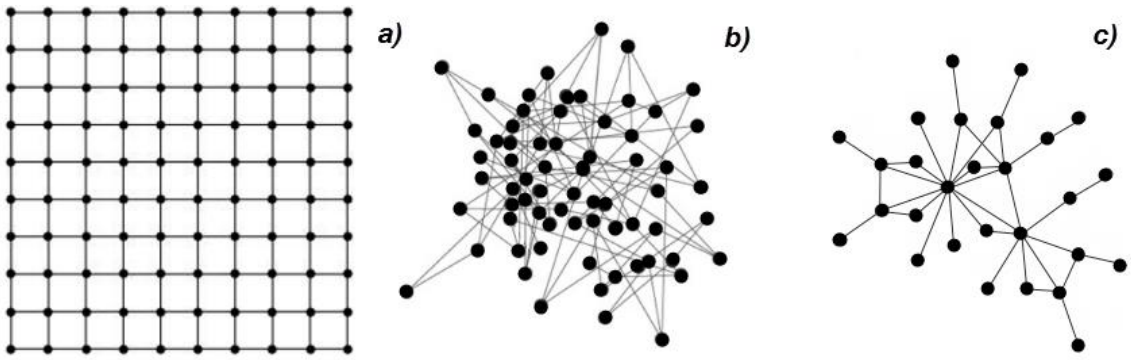

Figure 5.1 Examples of a) a regular, b) a random, and c) a scale-free networks.

Not each of the network nodes is equally important in terms of its function. It may happen that the removal of a node from the network will either go unnoticed, or

[18] The path $l(j,k)$ is a set of edges, which connects the nodes j and k. The length of the path is the number of its edges.

quite the contrary, will lead to an impairment of operation or even disintegration of the network. The number of all edges that a node has is its degree k. The higher its degree, the more important the node is for the entire network because it interacts with a greater number of other nodes. The distribution of node multiplicity $P(k)$ is one of the basic network characteristics, which informs that a randomly chosen network node has k connections with other nodes. Given N nodes and the number of nodes N_k with the defined multiplicity k, this distribution will be expressed by the formula:

$$P(k) = \frac{N_k}{N}. \tag{72}$$

An important feature being common to different real networks is that the node multiplicity distribution is a power-law function $P(k) \sim k^{-\gamma}$ [95], and it is known from the discussion in chapter (**2**) that power-law distributions are scale-free distributions. Scale-free network (Figure 5.1, panel c) was observed, for example, in the network of WWW sites ($\gamma = 2.1$), power networks ($\gamma = 4$) or air connection networks ($\gamma = 2.2$) [107; 108; 109], though, at first glance, these networks (at least the WWW network) could be regarded as random networks. Whether this feature, being characteristic of different types of systems, is accidental or is the effect of some fixed rules was attempted to be solved by two mathematicians, Albert-László Barabási and Réka Albert (model BA). In their study [110] they have shown that there are two mechanisms which generate the power-law distribution of nodes in real networks. These are: the network growth (at each step of network evolution, a new node is added), and the preferential node attachment (one node attaches to another one with a probability proportional to the other node's multiplicity). A characteristic feature of the scale-free network is the resistance to accidental damage to its nodes and a small average path length (formulae (**73**)) [107].

Basic measures to describe a network, in addition to node multiplicity, are the average path length and the node betweenness. For a network with N nodes and the shortest path $l(j,k)$ between the nodes j and k, the average path length L is given by the formula:

$$L = \frac{1}{N(N-1)} \sum_{j \neq k} l(j,k). \tag{73}$$

The greater the average path length L, the more scattered the network is.

The betweenness [111], which, like the multiplicity, describes the significance of a node i in the network, is defined by the following formula:

$$b_i = \sum_{j \neq k} \frac{n_i(j,k)}{n(j,k)}, \tag{74}$$

where $n_i(j,k)$ is the number of the shortest paths between the nodes j and k passing through the node i, and $n(j,k)$ is the total number of the shortest paths. The greater the value of b_i, the more important the node i is to the network. As has been shown [96], at the level of normal linear correlation, the financial market constitutes a set of interconnected companies. Taking companies for network nodes and the coefficient $\rho^q(s)$ (**68**) for edges, weighed networks [19] were made for each value of parameter q.

100 American companies, respectively, the largest and the smallest in terms of capitalization, recorded at intervals of 5 minutes, were used for analysis. Each of the analyzed rate of return series had a length of 40638 points and covered the period from 1.12.1997 to 31.12.1999. The series were also examined for artifacts, i.e. the numerical errors of writing data to a file, which could have disturb the obtained results; so, the series do not have any rates of return with a standard deviation greater than 10, whose occurrence is very little probable.

To avoid any information overflow, the Minimum Spanning Tree (MST) method was used for network structure analysis [112], which enables the structure of a network to be graphically presented in a concise manner and its essential features to be captured analytically [97]. Let there be given a graph G with V nodes, a set of edges $E \in \{\{u,v\}: u,v \in V\}$ and an edge weight C_{xy}. The graph T with V nodes and $D \in E$ edges, having the smallest sum of edge weights ω_e, is the minimum spanning tree. One node V corresponds to one time series. The method relies on the distance measure $d^s_{x,y} = \sqrt{2(1 - C^s_{xy})}$, where C^s_{xy} is the matrix of $N \times N$ coefficients $\rho^q(s)$ (formula (**69**)) determined for scale s (this is the characteristic scale used in the MFCCA and MFDFA methods). To verify whether for the selected scale s the selected distance is well defined for the selected moments q from the interval $q \in [-4,4]$ and meets the conditions of the metric[20] [96], 161700 three-element company combinations was subjected to analysis and checked if the triangle inequality $d^q_{AB} + d^q_{BC} \geq d^q_{AC}$, is satisfied, where d^q_{AB} is the distance between the series a and b. The result of the analysis is shown in Table 5.1 .

[19] In the case, where each edge has a specific numeral value (weight) ω_e, assigned to it, a given network is referred to as the weighed network.

[20] If $x = y$ then $d^s_{x,y} = 0$, $d^s_{x,y} = d^s_{y,x}$, $d^s_{x,y} + d^s_{y,z} \geq d^s_{x,z}$.

Table 5.1 The number of cases, where the triangle inequality for the distance $d_{x,y}^{s}$ is not satisfied. Actual errors are given in parentheses. The letter A denotes the set of the largest companies, while B, that of the smallest companies.

q\scale	Hourly -A-	Daily -A-	Weekly -A-	Hourly -B-	Daily -B-	Weekly -B-
-4	20436 (1501)	0	0	155198	0	0
-3	21505 (2570)	5 (5)	0	155200 (2)	1 (1)	8 (8)
-2	24149 (5214)	3705 (3705)	0	155198	2464 (2464)	1940 (1940)
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0

For $q < 0$, the triangle inequality is not satisfied for the majority of scales. The vast majority of cases is due to the fact that the intervals of zeros longer than the scale s used in the MFCCA and MF-DFA methods have not been removed from the analyzed series. This leads to the situation, where the fluctuation function and, consequently, the coefficient $\rho^{q}(s)$ cannot be determined for small scales. As a result, the triangle inequality is not satisfied numerically. This is particularly visible for smaller companies that are less frequently traded (there are more zeros in the series). There are also cases in the analyzed group, where $\rho^{q}(s)$ is determined correctly, yet the inequality is not satisfied. This means that the distance $d_{x,y}^{s}$ defined for $\rho^{q}(s)$ does not constitute a metric. And since it is not a metric, then the tree built on this distance is not an MST tree. Therefore, moments $q = \{2; 3; 4\}$ were used for determining the MST trees, because only for them is the $d_{x,y}^{s}$ distance unquestionably a metric.

The value of $d_{x,y}^{s}$ measures the distance between two time series X and Y. Thanks to this operation it is possible to exchange the C_{xy} measure, which is proportional to the connection strength, for the $d_{x,y}^{s}$ measure, in which this relationship will be inverse. For fully correlated signals, $d_{x,y}^{s} = 0$, for anti-correlated signals, $\sqrt{2} < d_{x,y}^{s} \leq 2$, while for uncorrelated signals, $d_{x,y}^{s} \cong \sqrt{2}$. Having a given metric, the MST method puts all distances $d_{x,y}^{s}$ between time series (X, Y) in increasing order and combines all nodes with respect to $d_{x,y}^{s}$ in such a manner that each node be attached exactly once. In that case, we obtain a network that is distinguished by the smallest possible sum of edge weights ω_e. In the MST method, each node has a measure of "importance" in the network assigned to it. The more important a node is for the network, the higher is its degree K. In transition from the complete network to the MST

tree, some measures (e.g. the betweenness or the average length of the shortest path) are retained. All trees were determined using Kruskal's algorithm [113] available in Matlab.

The whole analysis was based on the correlation coefficient ρ^q described in section 4.4 and was carried out in two dimensions. One of the dimensions is the time interval between price measurements, therefore individual MST trees are presented in three time scales: an hourly, a daily and a weekly scale. The second dimension is parameter q in the MF-DFA and MFCCA methods, being responsible for the reinforcement of either large or small fluctuations. The aim of the analysis is to answer the question whether and, if so, how the statistical network properties change with the scale and the fluctuation amplitude.

5.1 The 100 largest American companies [21]

For $N = 100$ companies and a given scale s, $N(N-1)/2 = 4950$ coefficients $\rho^q(s)$ were determined for q from the interval $-4 \leq q \leq 4$, based on which a network was constructed, defined in such a manner that each time series x_i was represented in the form of the node i, while the value of coefficient $\rho^q(s)$ represented the edge connecting the nodes $i - j$. A sample function $\rho^q(s)$ determined for the companies GE and CSCO (the hourly scale) is shown in Figure 5.2. It is perfectly visible that for moments $q < 0$, no cross-correlation occurs between the companies for the most of the examined scales s. This is evidenced by two different types of the behaviour of $\rho^q(s)$. For $q = -4,$, the coefficient $\rho^q(s)$ has a value close to 0 practically for all scales, while for $q = -2$ and $q = -3$, it heavily oscillates between negative and positive values, not assuming the same sign for larger intervals of s. For $q > 0$, the correlation is well visible around the value 0.5 ($q = 2$), 0.4 ($q = 3$) and 0.25 ($q = 4$) for all examined scales. A characteristic scale of $s = 1950$ minutes (indicated with the dashed line in the figure) is very well visible for $q > 0$, which represents the full trading week, when the value of $\rho^q(s)$ changes markedly.

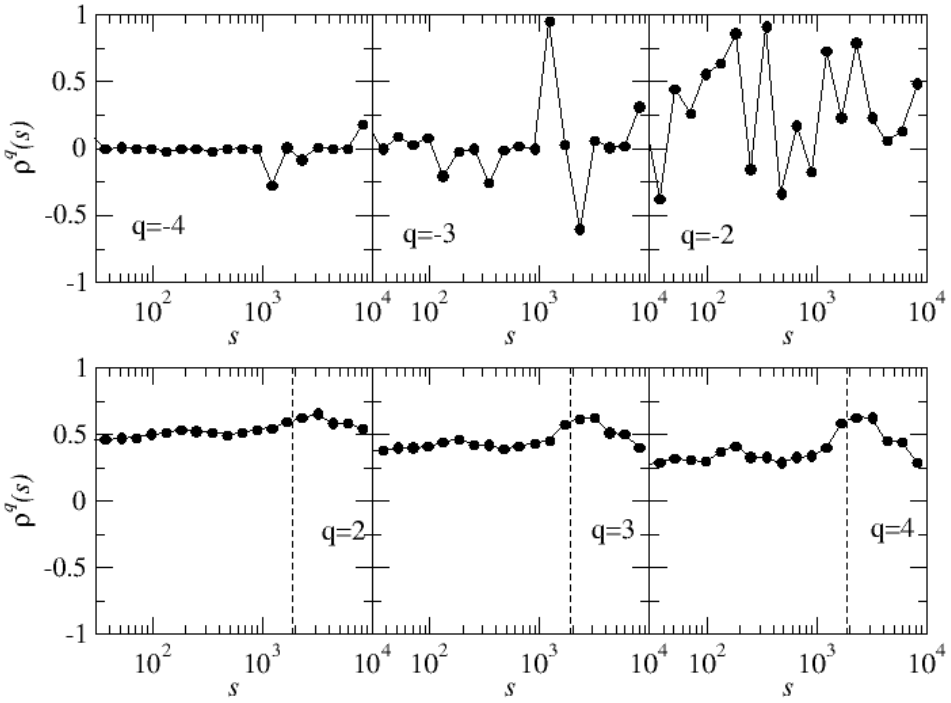

Figure 5.2 The coefficient ρ^q determined for the GE and CSCO company pair (the hourly scale) The dashed line indicates one trading day.

[21] A list of all companies with their division by sectors is given in Table 8.2.

For the MST network analysis performed later on in this study, three characteristic times scales were taken, i.e. the hourly scale ($s = 60$ minutes), the daily scale ($s = 390$ minutes) and the weekly scale ($s = 1950$ minutes). The probability density distribution of $\rho^q(s)$ coefficients, determined for all company combinations and three time scales, is shown in Figure 5.3. The hourly scale is indicated in black colour, the daily scale in red, and the weekly scale in green. For $q = \{-3, -4\}$ and the daily and weekly scales, the distributions are approximately symmetrical, with the distribution maximum being located near the 0 point, which most often means the lack of cross-correlation between companies at the smallest fluctuation level. In the case, where $q = -2$, ρ^q assumes more values from the entire range *(-1;1),* what doesn't mean the cross-correlation between companies as is was stated above. It comes from the fact that ρ^q strongly fluctuates around zero for the entire scale range [94].

For moments $q > 0$, the distributions are asymmetric, their tail is thicker on the right-hand side, and the maximum lies in the range [0.1;0.2]. This is due to the fact that the companies are rarely negatively correlated or uncorrelated with one another, and a slight positive correlation prevails between them on all the examined scales.

For the daily and weekly scales, the shape of the distribution does not change because of the increase in q, and the position of the distribution maximum shifts slightly towards 0, which suggests that reinforcing increasingly large fluctuation amplitudes weakens the correlation between the companies. For the hourly scale, the distribution also shifts towards 0, except that an increase in the number of companies with the lowest cross-correlation occurs in the distribution shape. The effect of weakening of correlation between companies is the strongest for the hourly scale.

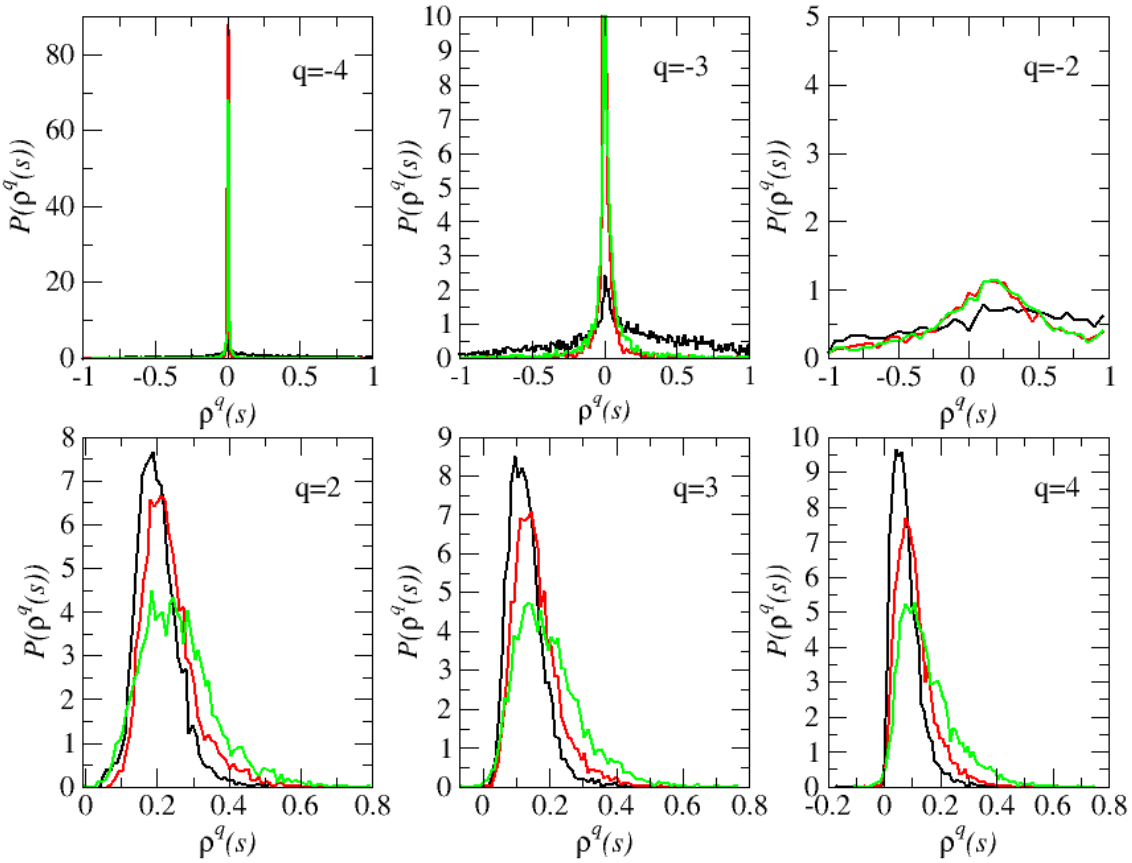

Figure 5.3 The probability density $P(\rho^q(s))$ determined for three time scales: hourly (black colour), daily (red colour) and weekly (green colour).

A sample visualization of the obtained MST trees for the hourly scale and moments $q > 0$ is shown in Figure 5.4 . For this and every other MST tree shown later on in the study, the length of the connection between nodes does not carry any information, in contrast to the thickness and colour of the lines. The thicker and blacker the line, the greater the cross-correlation.

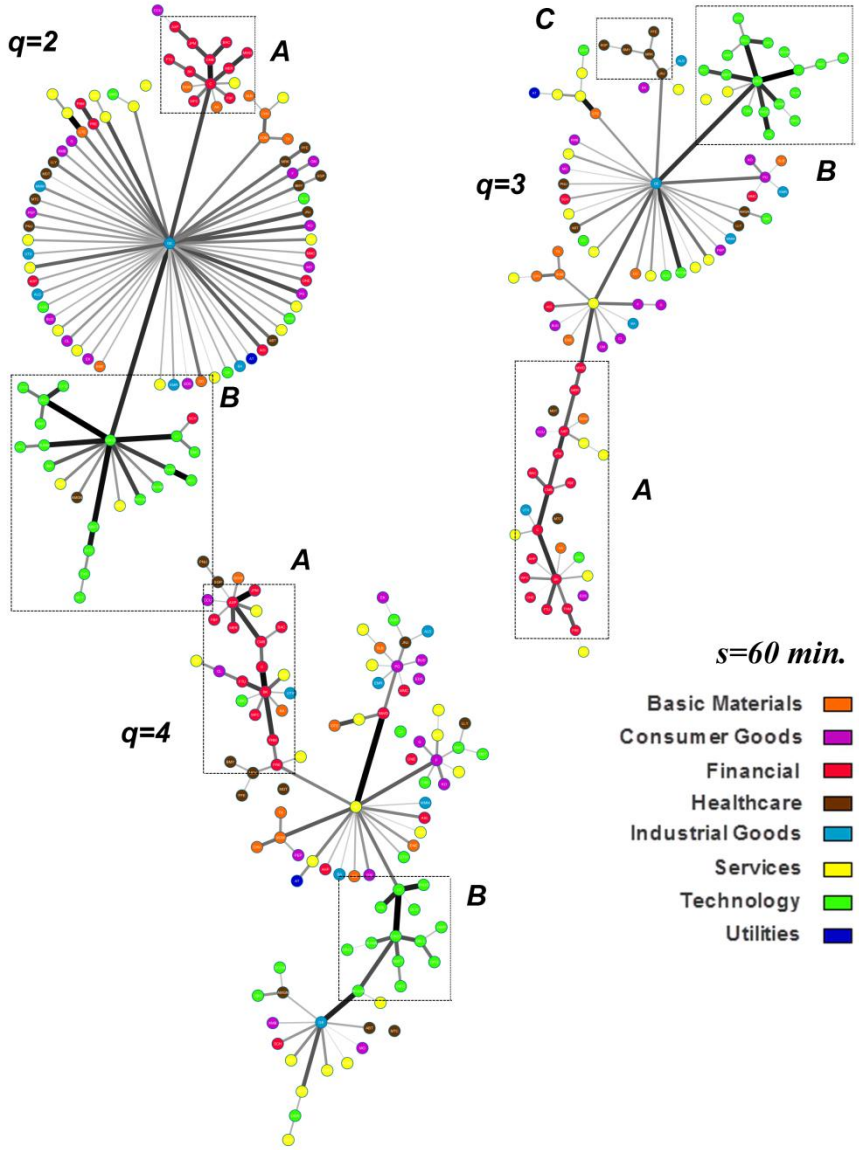

Figure 5.4 An MST tree drawn up for the 100 largest American companies, determined for different values of parameter q and the hourly scale, broken down by sectors. The clusters are indicated with rectangles: A – financial cluster; B – technological cluster; and C – health care cluster.

In all of the analyzed MST trees and for moments $q > 0$, nodes of a high multiplicity k are very clearly visible, and the network itself has the features of a centralized network with visible financial (A - red), technology company (B - green) and health care (C - brown, visible for $q = 3$) clusters. The value of the correlation between these clusters (as well as in the clusters themselves) and the central network node (GE) is one of the greatest, as opposed to the correlation to other network nodes, where this level assumes moderate values, i.e. 0.1–0.4. Interestingly enough, at the plain linear correlation level, networks of this type (i.e. centralized, with clusters) can be observed on the FX market [97; 98].

For $q = 2$, the main node is the GE (General Electric) company with a multiplicity of 51 (super hub), CSCO (Cisco Systems, Inc.) with $k = 12$ (technology cluster) and C (Citigroup Inc.) with $k = 9$ (financial cluster). For $q = 3$, the main network node is also GE with $k = 24$ and HD (Home Depot) and BK (Bank of New York) with $k = 10$. The BK company is directly associated with 7 other financial sector companies, forming a cluster. With a multiplicity of $k = 8$, the CSCO company forms a significantly correlated technology cluster. For $q = 4$, the clusterization is slightly less visible and the network is getting more decentralized. With a multiplicity of $k = 10$, GE is no longer the central network node, whose role is taken over by the HD company with $k = 15$. Still well visible is the technology cluster, with the centre in the CSCO company with $k = 5$, and the financial cluster, this time without the main node, but with a very well visible strong cross-correlation. In addition, clusters with a multiplicity of $k = 8$ are clearly visible in the network, which do not constitute a sectoral connection for BK (The Bank of New York), AXP (American Express Company), F (Ford Motor Co.) and PG (The Procter & Gamble Company).

The presence of GE as the central node (for $q = 2$ and $q = 3$) is not surprising, as this company is a huge conglomerate on the American market, operating either directly or indirectly practically in all market branches, such as energy generation, petroleum production, or machinery and equipment manufacture. It should also be noted that, in the period under examination, the aforementioned companies being the centres of clusters, were the companies of the highest market capitalization among all companies. What is also characteristic of the trees described above is the occurrence of a very large number of companies with a multiplicity of 1 in the vicinity of high-multiplicity nodes.

In terms of topology, the MST trees determined for the daily scale look very much similar to those determined for the hourly scale (Figure 5.5). For all values of parameter q, high-multiplicity nodes are very well visible (GE is again the network's hub), forming three clusters (on the example of $q = 2$): financial (A - red colour, with the centre in the C company with a multiplicity of $k = 7$); technology (B - green colour, with the centre in the CSCO company with a multiplicity of $k = 12$) and health care (C - brown colour, MRK, with a multiplicity of $k = 5$). With the increase in q, the multiplicity of the central nodes of particular clusters slightly decreases.

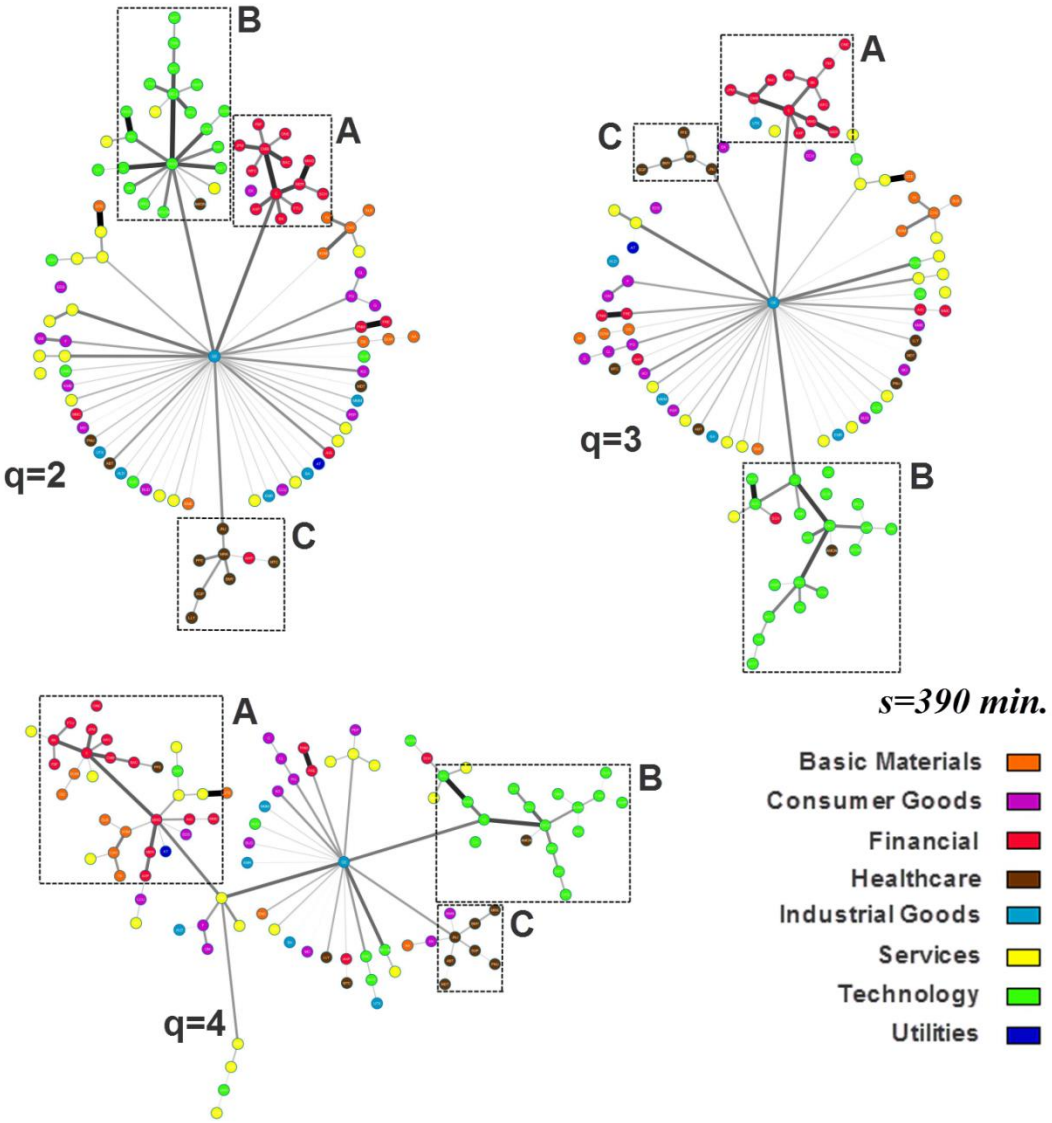

Figure 5.5 An MST tree for the 100 largest American companies, determined for different values of parameter q and the daily scale, broken down by sectors. The clusters are indicated with rectangles: A – financial cluster; B – technological cluster; and C – health care cluster.

The situation is very similar for the tree determined for the weekly scale (Fig. 5.6.). For moments $q = 2$ and $q = 3$, like for the daily scale, three clusters are visible, which are supplemented with a fourth cluster (D - yellow colour) with the central node in the HD company, for moments $q = 3$ and $q = 4$, and a fifth basic materials cluster (E - orange colour) showing up for $q = 3$, with the central node in CHV (Chevron).

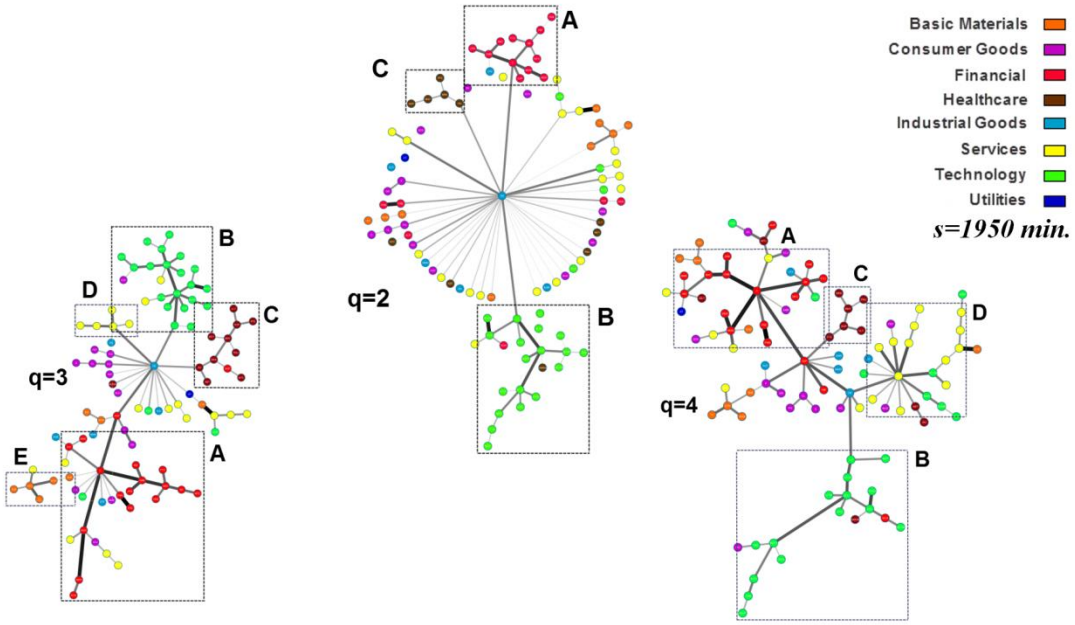

Figure 5.6 An MST tree determined for the 100 largest American companies, for different values of parameter q and the weekly scale, broken down by sectors. The clusters are indicated with dashed line. A – financial cluster; B – technology cluster; C – health care cluster; D – services cluster; and E – basic materials cluster.

It is worth noting that the structure of MST trees is not stable due to the variation in signal amplitudes and the change in the time scale. There are companies (such as GE) which, with the increase in fluctuation amplitude or scale, lose their importance in the network. There are also those, whose multiplicity changes with increasing fluctuation amplitude (MRK is the centre of a cluster for $q = 3$, and is not for $q = 2; 4$). It appears, therefore, that both the fluctuation amplitude and the time scale has the effect on the strength of correlation between companies. In the case of the time scale, a longer time horizon is most often needed for the correlation to form, as confirmed by the weekly scale case, where the clusterization is the greatest and ρ^q assumes the largest values (Figure 5.4). This may also mean that, at the large signal amplitude level, the fluctuations of correlated companies are similar. This is, as it were, a natural behaviour, because the majority of companies head in the same direction as the market does, whereas among themselves, the companies exhibit a more independent dynamic.

Looking at the tree structure on selected scales it can be noticed that, on the quality level, the network does not change significantly, and presumably this is a scale-free network. The scale-free networks of financial data are a well-known phenomenon [101; 114; 115]. The absence of the characteristic scale seems to be confirmed by

Figure 5.7 which shows cumulative distributions of node multiplicities, determined for the hourly, daily and weekly scales, depending on the moment q. For all moments $q > 0$, the distribution is a power-law distribution, which is characteristic of the distribution of multiplicities in scale-free networks. The estimated scaling exponents are: $\alpha = 1.22; 1.45; 1.7$ (for q being, respectively, 2, 3 and 4) for the hourly scale; $\alpha = 1.47; 2.16; 2.4$ for the daily scale; and $\alpha = 1.85; 2.29; 2.09$ for the weekly scale. In study [101] where its authors examined multiplicity distributions for nearly 500 companies listed at the NYSE stock exchange, the scaling exponent was estimated at a level of $\alpha = 2.1$. In turn, study [97] estimated the scaling exponent to be in the range of $\alpha \in (1.37; 1.96)$.

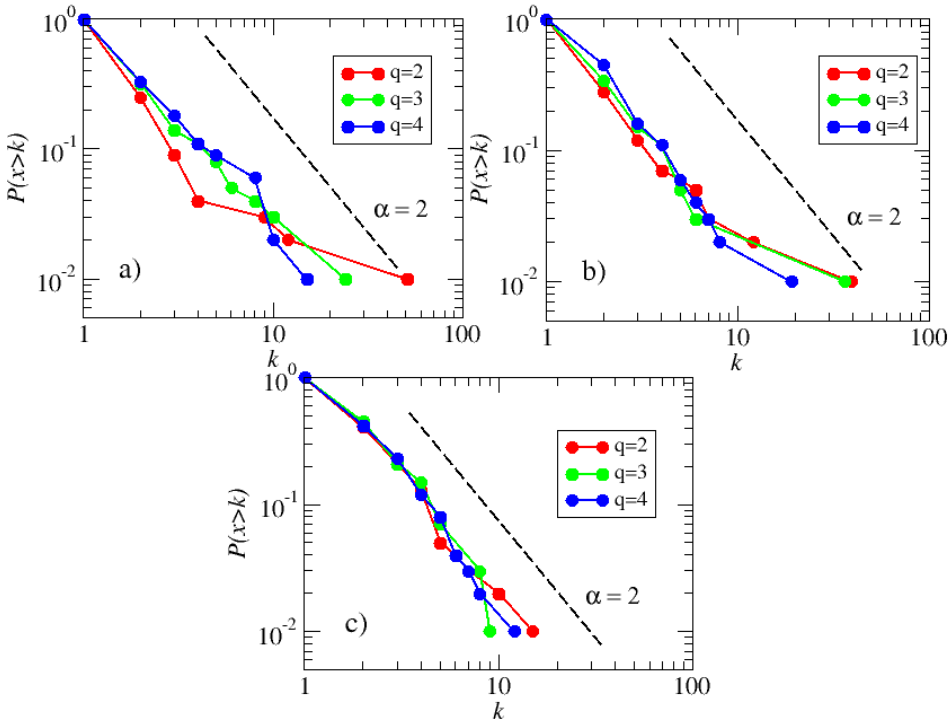

Figure 5.7 Cumulative distribution of node multiplicities for the 100 largest American companies determined for the hourly a), daily b) and weekly scales, respectively, depending on the parameter *q*.

For each tree, the average path length L was also determined (Table 5.2). In view of the fact that the obtained L values are high, the network exhibits the features of a dispersed network. The obtained L values are, for every scale and every moment, greater than the average path length $L_{av} = 2.87$, as determined based on the Pearson correlation coefficient ρ_{xy}, which implies that a network relying on nonlinear correlation at a varying fluctuation level, is more dispersed.

Table 5.2 The average path length L determined for three time scales and different values of parameter q.

Scale | q	2	3	4
Hourly	3.28	5.83	5.97
Daily	3.68	3.4	5.17
Weekly	5.63	6.15	6.16

The betweenness coefficient b_i (Table 5.3) determined for moments $q > 0$, confirms the existence of a large number of nodes significant for network operation efficiency, which lose importance with increasing parameter q. The obtained betweenness values are, on every scale and for every moment, less than the value of the betweenness $b_i = 0.94$ determined for the network relying on the Pearson correlation coefficient ρ_{xy}. This implies that a network built on the coefficient of nonlinear correlations at a large signal amplitude level indicates a smaller number of companies playing a dominant role in the financial market.

Table 5.3 The network betweenness b_i determined for three time scales and different values of parameter q.

Scale | q	2	3	4
Hourly	0.91	0.74	0.76
Daily	0.9	0.9	0.73
Weekly	0.74	0.64	0.62

To sum up, the statistic properties of the network determined for cross-correlations at a large fluctuation level exhibit similarities with respect to scale change. If large fluctuations occur in networks, those networks become heavily centralized, and with the increase in parameter q, the strength of connections between the companies is increasingly high. Very easily visible clusters show up, which group together companies belonging to a single stock market sector. In the case of the technology and financial sectors, they are stable in respect of the change in q, in contrast to the health care cluster which appears only at a higher fluctuation amplitudes (only visible for $q = (3; 4)$) or the basic materials cluster which shows up for the weekly scale only. The examined statistic properties of the MST trees do not change significantly with the change in

scale, either. The scale-free network structure is visible for moment $q > 0$ and for all scales.

It is important to take note of the practical aspect of the performed analysis. As rightly observed in studies [96; 101], MST trees relying on (linear) correlation coefficients can be useful for the optimization of the investment portfolio. In the classic theory of investment portfolio optimization introduced by Markowitz [116], the selection of assets is made based on the expected value and the variance of the portfolio return rate. From among the constructed portfolios, the one is selected, which maximizes the expected value with a given market, or the one which minimizes the risk at the given expected rate of return. By extending the capabilities of such analysis to include nonlinear correlations to be determined additionally on different time scales and for different magnitudes of fluctuations, we get a tool in our hands, which can provide more accurate information on the market under analysis.

5.2 The 100 smallest American companies[22]

In order to verify whether the results obtained in the previous section are typical of large capitalization companies only, a network of 100 American companies of the least capitalization was also subjected to analysis. Like in the previous paragraph, the coefficient $\rho^q(s)$ was determined for three time scales: hourly, daily and weekly, for q parameter values of $-4 \leq q \leq 4$. In Figure 5.8 a sample coefficient $\rho^q(s)$ determined between the companies BZF and EWF, being the largest nodes in the network, is shown for the hourly scale and different values of parameter q. For the moment $q = -4$, lack of correlation is clearly seen. The same is visible for the moments $q = -2$ and $q = -3$, where ρ^q fluctuates around zero what means that there is no correlation on these amplitude level. This effect is characteristic of all of the examined companies. For moments $q > 0$, the correlation is positive for all scales s, although it is much weaker than for the largest companies in the market.

[22] A list of all companies with their membership of respective sectors is given in Table 8.3.

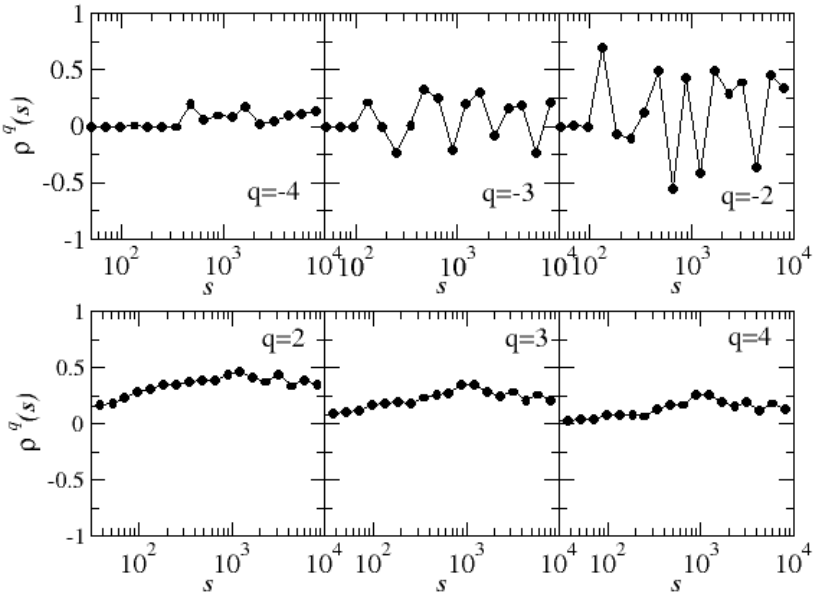

Figure 5.8 The coefficient ρ^q determined between the small-capitalization companies BZF and EWF for the hourly scale.

Figure 5.9 represents the density distribution of coefficients $\rho^q(s)$, determined for all combinations of the smallest companies and three time scales. For the daily scale and weekly scales and for $q < 0$, the shape of the distribution is very similar to that of the one determined for the largest companies. They are approximately symmetrical and for $q \in [-3; -4]$ their maximum is near 0, what confirms the fact that at the lowest fluctuations level, there is no cross-correlation between the companies. For the daily and weekly scales and $q = -2$, there are much fewer cases of mutually uncorrelated companies, and the distribution itself is more flat. But again, it comes from the fact, that ρ^q fluctuates around zero for the entire scale range and it doesn't mean the cross-correlation. For the hourly scale, the distribution ρ^q is stable with respect to change in the negative values of parameter q and, most frequently, there are no cross-correlations between the companies at a low fluctuation level. This behaviour is different from that observed for the largest companies, where the distribution was more flat for the above-mentioned hourly scale and moment $q = -2$.

For positive values of moments q and for the hourly and daily scales, ρ^q coefficient values close to 0 are predominating. For the weekly scale, the distribution is stable with respect to the increase in q, and ρ^q assumes a wider value spectrum, with a preponderance of values indicating a slight positive cross-correlation. So, it appears that, like for the largest companies, at a given fluctuation level, it is the time scale that

plays an important role in the occurrence of correlations. On shorter scales, we can speak of a lack of correlations between the companies, which only become pronounced on the weekly scale. This suggest that a slightly longer time horizon is needed for cross-correlations to occur. For the largest companies, no similar feature could be noticed, because correlations at a varying level of large fluctuation amplitudes were visible for all examined time scales.

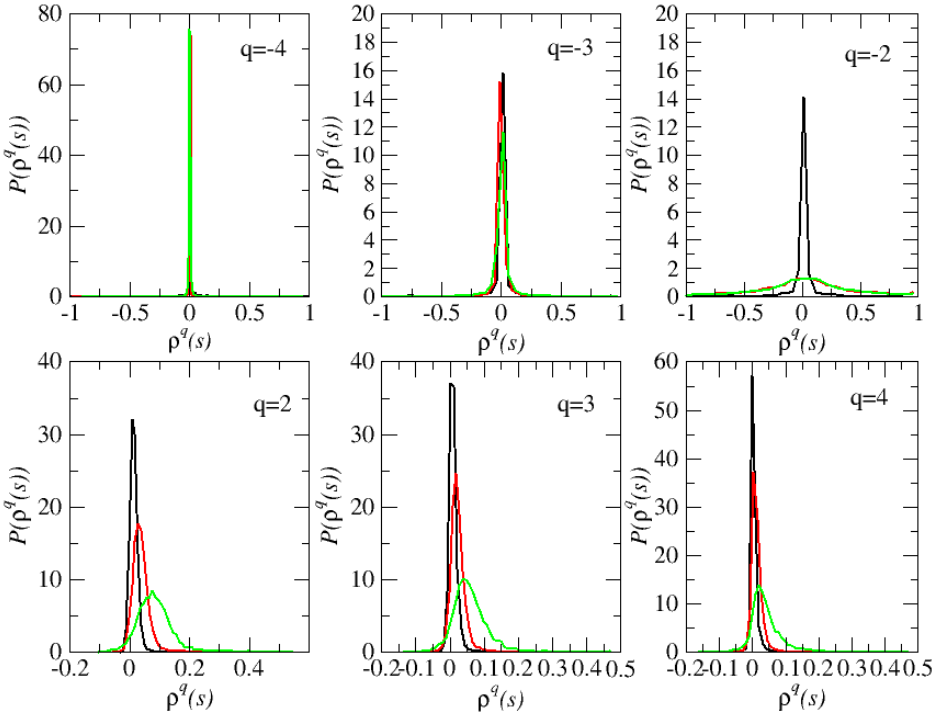

Figure 5.9 The probability density $P(\rho^q(s))$ determined for three time scales: hourly (black colour), daily (red colour) and weekly (green colour).

Like for the largest companies, the MST trees were determined for the smallest companies. The tree determined for the hourly scale and three different values of parameter q is shown and Figure 5.10.

For $q = 2$, there exist central network nodes and these are the EWF (*Earth, Wind & Fire*) and BZF (*Wisdom Tree Brazilian Real Strategy ETF*) investment funds (both of them with a multiplicity of $k = 20$). The value of correlation between them is the highest among all of the examined companies. They form two interconnected financial clusters. The basic materials cluster is also visible in the network, with the centre in the NS (*National Steel Corp Cl-B*) company. The both funds are of the ETF[23] type,

[23] ETF (Exchange-Traded Fund) – an investment fund based on the stock market index or the asset basket. It is most often managed on an algorithmic manner, and its units are quoted on the Stock Exchange similarly to shares. www.investopedia.com access on:12.12.2015.

therefore their investment strategy relies on the tracking of the stock market index. They will be, as it were, a reflection of the market; they turn to their advantage the whole market information, as a consequence of which it could be expected that they will have the greatest multiplicity.

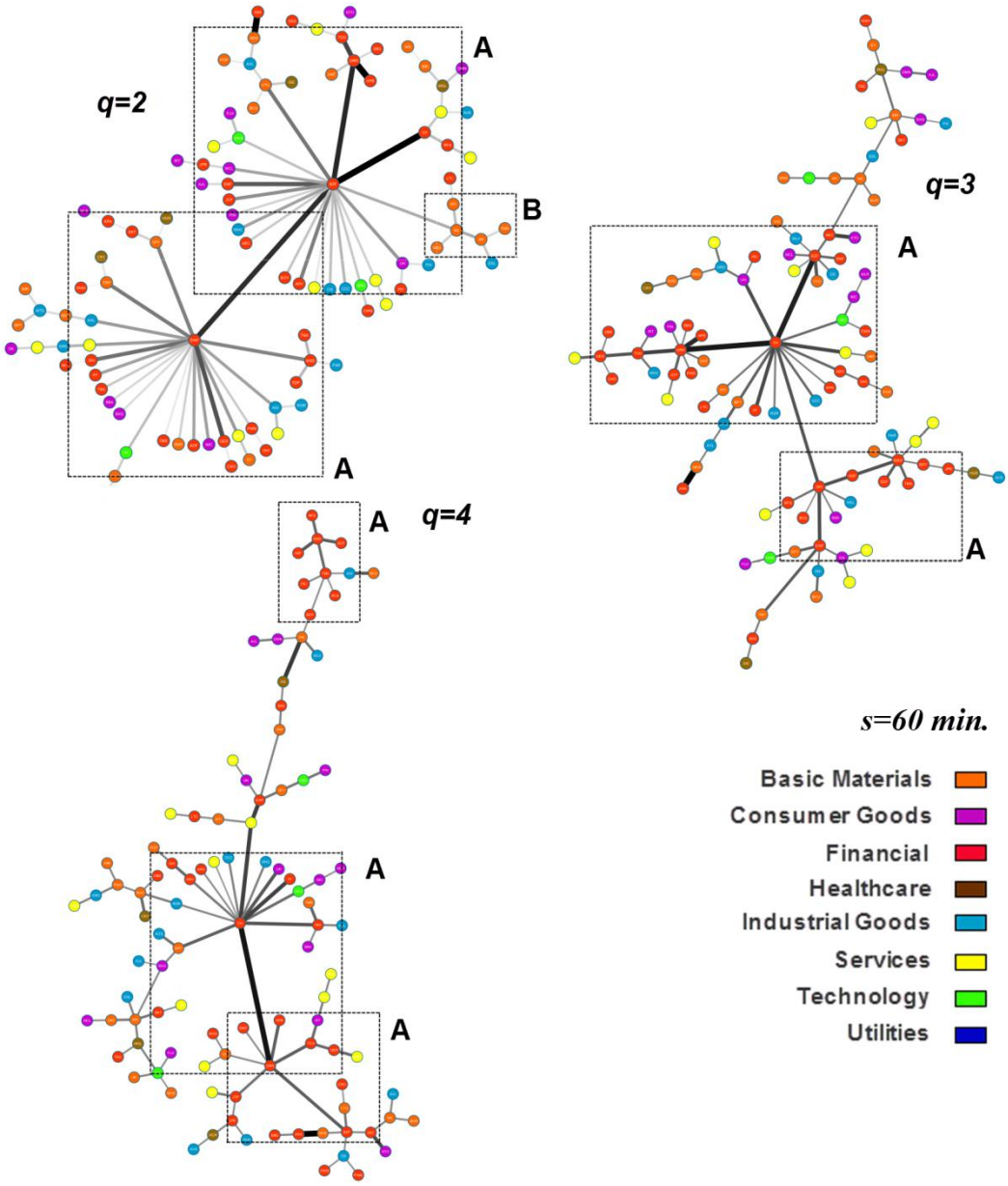

Figure 5.10 An MST tree determined for the 100 smallest American companies, for different q parameter values and the hourly scale. The cluster A constitutes a set of companies from the financial sector (red colour), and the cluster B, from the basic materials sector (orange colour).

For $q = 3$, there are two central network nodes: CH (*Aberdeen Chile Fund Inc.*), with a relatively large multiplicity of $k = 13$ and GHI (*Global High Income Fund Inc.*) with a multiplicity of $k = 7$. These are investment funds, too, and they form interconnected financial clusters. For $q = 4$, the place of the GHI node is taken by

another fund, GRR (*Asia Tigers Fund Inc.),* with a multiplicity of $k = 6$ and, like before, two connected financial clusters are visible. When analyzing the tree structure for the daily scale 5.11 and in the case of $q = 2$, we can clearly see two high-multiplicity nodes: EWF $(k = 34)$ and BZF $(k = 20)$, which are strongly correlated. For $q = 3$, the companies EWF and BZF still have the highest multiplicity ($k = 10$ and $k = 9$, respectively), and the GHI company with a multiplicity of $k = 9$ shows up again. For $q = 4$, the companies BZF and EWF lose their importance in the network, with the predominant nodes being GHI, GRR and TCH (all of them with a multiplicity of $k = 8$). In the cross-section of the entire tree, it is difficult to identify clusters forming a sectoral combination of companies, which is makes is a difference compared to the trees determined for the largest companies, where they were fairly well visible.

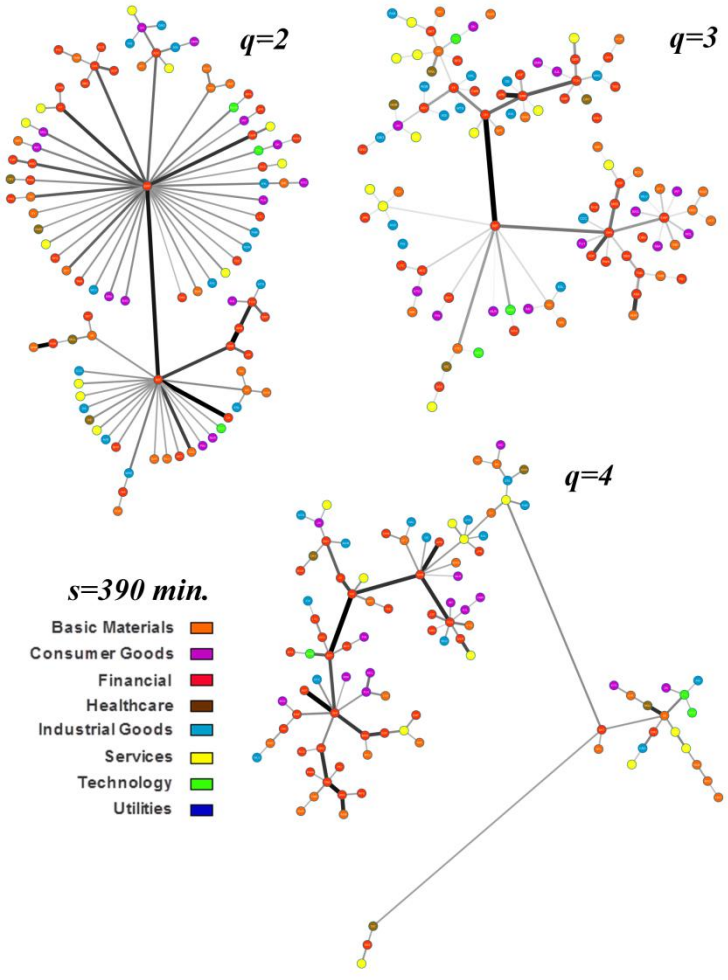

Figure 5.11 An MST tree for the 100 smallest American companies, determined for different *q* parameter values and the daily scale.

The above-described structure of the MST trees determined for the daily scale seems to be similar to that of the trees determined for the weekly scale, as shown in Figure 5.12.

For large fluctuations and for $q = 2$, a single large BXF node with a multiplicity of $k = 27$ occurs in the network, which loses its importance with the increase in parameter q. Like for the daily scale, it is hard to explicitly delineate sectoral clusters, and this is the only difference between to the analysis of the largest companies.

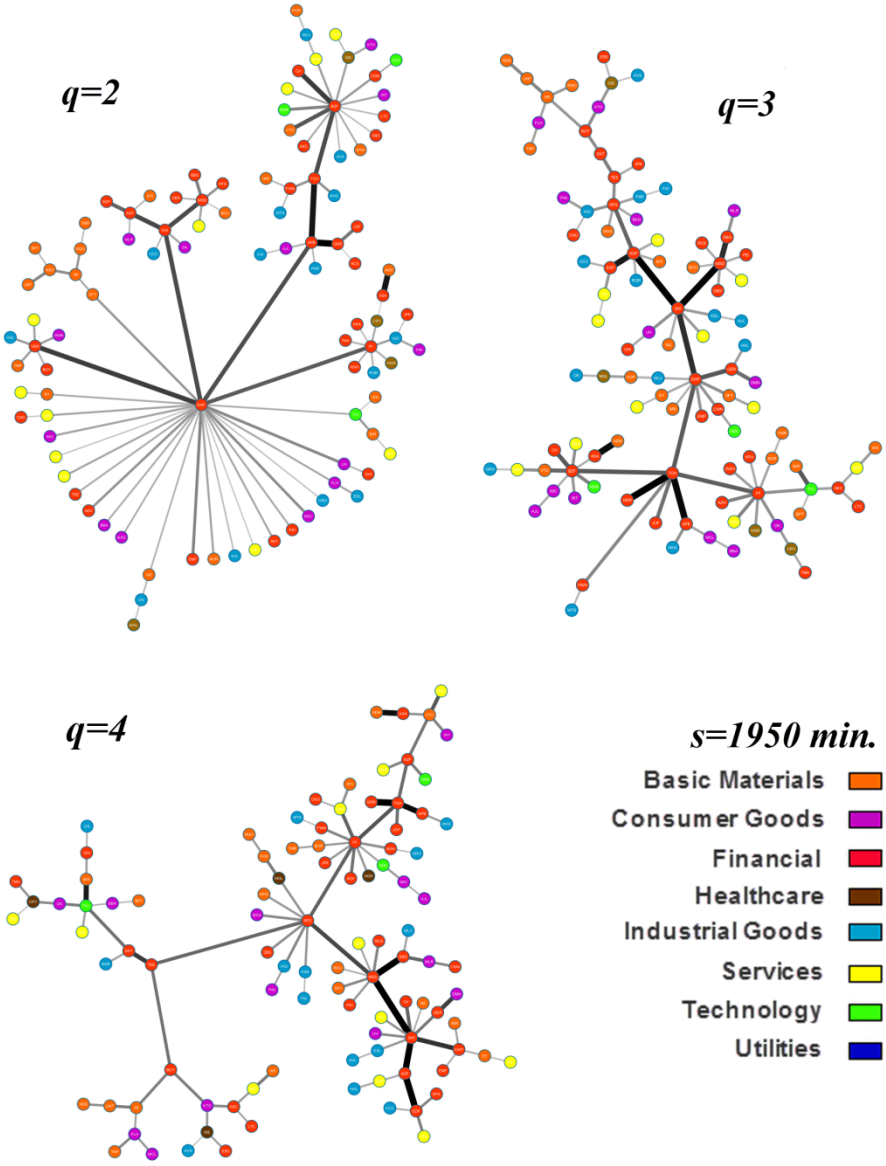

Figure 5.12 An MST tree determined for the 100 smallest American companies, for different q parameter values and the weekly scale.

The node multiplicity distribution represented in Fig. 5.13 seems to indicate the lack of the characteristic scale for moments $q > 0$ for all time scales. The estimated scaling exponents are: $\alpha = 1.95; 2.03; 2.25$ (for q being, respectively, 2, 3 and 4) for the hourly scale; $\alpha = 2.13; 1.64; 1.96$ for the daily scale; while $\alpha = 1.68; 1.55; 1.57$ for the

weekly scale. These are scaling exponents that lie in a value range similar to those obtained for the largest companies.

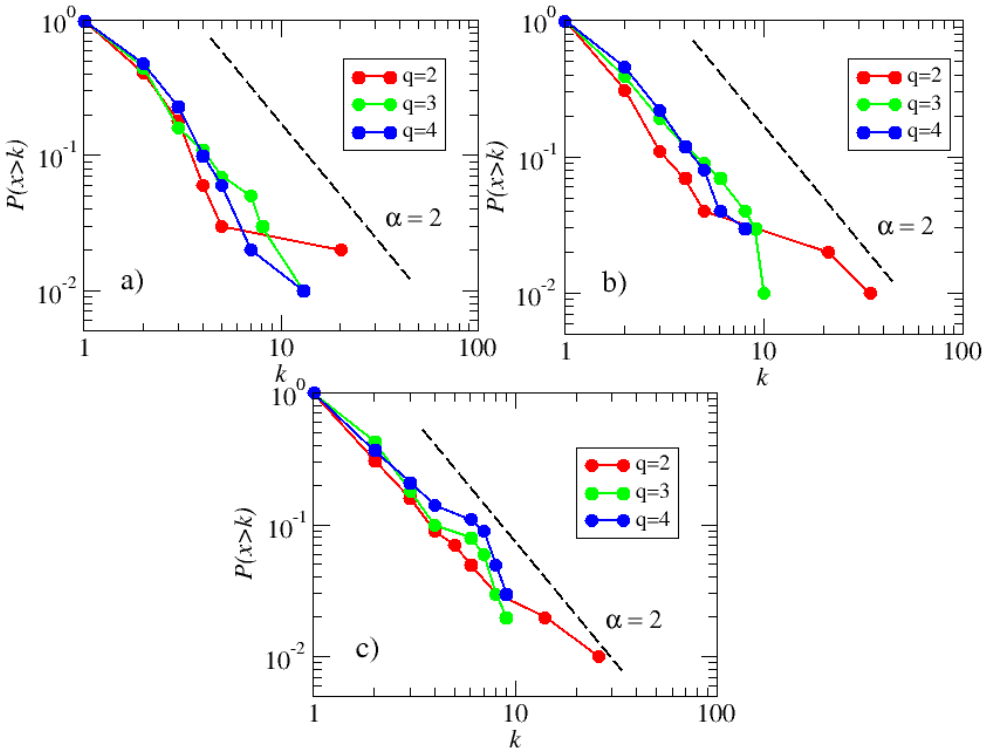

Figure 5.13 Cumulative distribution of node multiplicities for the 100 smallest American companies determined for the hourly a), daily b) and weekly scales, respectively, depending on the parameter *q*.

Like for the companies analyzed in the previous section, the statistic network properties were estimated based on the average path length L and the betweenness b_i. The average path length determined for the network based on the Pearson cross-correlation coefficient ρ_{xy} amounts to $L_{av} = 4.64$, being smaller than any of the determined values of the analyzed trees, except for the moment $q = 2$.

The values and the differences between the average path length values in Table 5.4, as dependent on parameter q, are similar to those determined for the largest companies (Table 5.2).

Table 5.4 The average path length L determined for three time scales and different q parameter values.

Scale | *q*	2	3	4
Hourly	4.25	5.85	6.12
Daily	3.63	5.42	6.99
Weekly	4.47	5.94	5.80

The situation looks similarly at the level of the betweenness (Table 5.5) determined for the network. For the hourly scale, the betweenness values are constant, whereas for the daily and weekly scales, the b_i values decrease with the increase in q, which means that the nodes most important for network operation lose their importance. For all moments $q > 0$ $b_i \in (0.57; 0.84)$, and it is in this range that the $b_i = 0.77$ determined for the ρ_{xy} Pearson correlation coefficient is contained.

Table 5.5 The network betweenness b_i for the 100 smallest companies, determined for three time scales and different q parameter values.

Scale | q	2	3	4
Hourly	0.76	0.77	0.77
Daily	0.84	0.66	0.59
Weekly	0.84	0.60	0.57

Like for the average path length, b_i betweenness values and their change with the increase in the parameter are similar to those determined for the largest companies (Table 5.3).

The analysis made above shows that, at the level of the statistic network properties determined for fluctuation amplitudes of a varying size, there is no significant difference between the largest and the smallest company networks. However, a sectoral clusterization is visible for the largest companies, which is practically absent in the smallest company network.

6 Summary and conclusions

The subject of the research carried out in this study was the analysis of financial data derived from the exchange and capital markets. Main emphasis has been laid on the identification and description of cross-correlations between individual elements of the market. The analyzed data included transaction exchange rates taken from the FX market, time series of triangular relation deviation on return level, the one-minute exchange rates of twenty nine DJIA index companies and the five-minute quotations of the 100 largest and the 100 smallest companies listed on the American market, which were examined in a network representation.

The multifractal analysis was performed using the *MFDFA* and *Wavelet Leaders* methods. The multifractal cross-correlations were identified using a new *MFCCA* method. In addition, the generalized cross-correlation coefficient was used, which, along with the network representation of relations between stock market companies, enables the two-dimensional analysis of relationships between companies with respect to time and fluctuation amplitude.

The analysis was started with the examination of the distributions of returns on different time scales, which made it possible to verify one of the stylized facts, according to which the tails of the cumulative distribution should have power scaling with an exponent of $\alpha = 3$. As has been demonstrated, for financial markets (DAX, SP500) in the period under consideration, deviations from the inverse cubic law can be observed already on time scales of the order of one minute [4] (for greater scales, the distribution tails become increasingly Gaussian in character). The study shows that in the case of the FX market, distributions are not stable, and the scaling exponent is contained in the range of $\in (2.9; 5.05)$ and strongly depends on considered time scale. However, it is hard to indicate a characteristic scale for particular exchange rate, where a break in scaling occurs and the tails bend back towards the Gaussian distribution. The situation is slightly different for the series of deviations from the ideal triangular relation. Starting from the smallest time scale, the distribution tails are very close to the Gaussian distribution. An exception is the EURGBPJPY triangle, where the slopes are in the range of $\alpha \in (3.21; 4.98)$. This result may suggest a much faster information flow on exchange markets compared to asset markets.

In the case of DJIA data, for the smallest time scale (1 minute), the scaling exponents of individual companies assume values approximately close to 2.5, which

means that the deviation from the inverse cubic law is little. The index consisting of those companies, starting from the scale of $t = 10$ min., starts slowly approaching the Gaussian distribution.

The analysis of the linear correlations for returns, made later on in this study, has confirmed the absence of any significant correlations for both financial and currency data. The autocorrelation function, for both sets of data, is characteristic of Brownian processes and after 10 seconds (FX data both for the rates of return and for the series of deviations from the triangular relation) or two minutes (financial data) it attains a noise level. Before this happens, however, it attains negative values which may suggest a negative correlation between successive changes in the series. The observed fact appears to be, however, a computational artefact, caused by too large a number of transactions not changing the asset price. Cross-correlation between currencies shows that generally currencies inside the triangle are cross-correlated negatively (barring one exception), while currencies not being inside the triangle, except for two instances, are slightly positively correlated. A currency pair which is the strongest connected with the others is the EURUSD pair (the most often traded currency pair in the world), while the weakest connected one is the AUDJPY pair. The examination of correlations between the series of deviations from the triangular relation has shown that such series are correlated even up to the scale of 15 seconds in the case, where two common currencies occur between the series. If there are no common currencies, the series are not correlated. For financial data and the majority of all possible combinations between companies, cross-correlation exists. Breaking down companies by sectors, the strongest inter-sector relationships occur in the Industrial Goods sector, and the weakest, in the Services sector.

At the level of the volatility of currency and financial data series, a power-law behaviour of the decay of the autocorrelation and cross-correlations functions and the clear presence of a daily trend in financial data, manifesting itself in distinct peaks occurring every 390 minutes, were observed. In addition, all the examined exchanges rates are cross-correlated for events distant nearly 7 trading hours and for a delay up to 50 seconds, a slightly lower level of cross-correlation can be observed in the currencies outside the triangle, compared to the currencies inside the triangle.

The main focus in the study was put on nonlinear correlations analysis so the next point discussed was autocorrelation multifractal analysis. For illustrative purpose singularity spectra and fluctuation functions were shown for monofractal processes, as

well as for two different multifractal models of the Markov switching processes and multiplicative cascade types. For real exchange rates, the estimated singularity spectra indicate a multifractal nature of the examined processes. They are all asymmetric, with the left-hand part much more developed. This asymmetry may be due to the fact that large fluctuations predominate in the structure of the analyzed signal. In turn, the location of the spectrum maximum suggests a lack of linear correlations or a slight anti-persistence in the examined series, which is consistent with results obtained by traditional methods.

The shapes of the spectra determined for exchange rate volatility series are also (left-sided) asymmetric, wide, and indicative of a great complexity of the analyzed signal. Their maxima are located in an area indicating high persistence of the series. At that level, negative fractal dimensions were observed, which are more and more often appear in the literature. Using the Fourier surrogate method it has been shown that it is nonlinear correlations that are responsible for the fractal nature of the currency time series. The triangular relation deviation series for the analyzed type of data have not exhibited the fractal nature in spite of the fact that those were already previously observed and described in the literature [31]. The explanation of this fact may be the too small length of the series, in which large fluctuations or respective correlations have not had enough time to develop yet.

In addition, so-called *cross rate*s, or exchange rates calculated using transition currencies, were also examined. It has been demonstrated that with the increase in the number of transition currencies the spectrum maximum shifts to the left towards the region of uncorrelated data and, in the majority of instances, it becomes increasingly narrow.

In the case of stock market data, all estimated singularity spectra show perfectly the multifractal nature of the time series. Spectra reveal left sided asymmetry. The location of the spectra maxima indicates a weak persistent nature of these processes. In the sectoral approach, the widest structure variety is exhibited by Financial Sector companies, while the narrowest, by companies from the Industrial Materials sector. It has been demonstrated that in this case, too, it is nonlinear correlations that underlie the fractal nature of the time series.

Both the share market and exchange market data were also examined with the Wavelet Leaders algorithm, and thus obtained results were largely consistent with those obtained by the MFDFA method in terms of the identification of the fractal nature of the

data. The spectra determined by the WL method turned out to be wider, which is a well-known effect, as shown in study [117].

The analysis of nonlinear interrelationships was started with the description of the MFCCA method, which, as has been shown, is able to correctly identify the correlations between time series. On the example of correlated and uncorrelated FARIMA processes generated by the algorithm, the usefulness of the method was demonstrated. Then, currency data were examined for nonlinear cross-correlations. For the rates of return and moments $q < 0$, the absence of relevant fluctuation function scaling has been shown, which means the lack of cross-correlation at the level of low fluctuations. The performed analysis shows also that all of the currency pairs exhibit cross-correlation both on the returns level, as well as on the volatility level, for currencies forming triangles. Generally, it can be accepted that in the case of volatility for currencies outside the triangle, the λ_q scaling exponent deviation from the mean value of generalized Hurst exponents $h_{xy}(q)$ is greater than for currencies inside the triangle. A parameter d has been introduced, being the difference between the span of the scaling exponent λ_q and the span of the average Hurst exponent of the analyzed time series $h_{xy}(q)$, illustrating the complexity of the cross-correlation. The analysis of the volatility has shown that d assumes values greater for currencies outside the triangle than for currencies inside the triangle. For return series, d greater than 0 was also observed, but the only criterion allowing currencies inside the triangles to be distinguished from those outside the triangle was the inability to estimate the scaling exponent for a few cases of currencies outside the triangle.

Similarly for financial markets, no correlations were observed at the low fluctuation level, therefore the analysis was narrowed to moments $q > 1$. Among all combinations of correlations between companies, two different types of scaling exponent behaviour were observed. In one case, the difference between λ_q and $h_{xy}(q)$ is independent of q (the correlation strength is constant at a varying signal amplitude level), while in the other, with the increase in q, this difference decreases practically to zero (change in correlation strength at a varying signal amplitude level). For all combinations of companies, parameter d most often takes on values close to 0, which means the same correlation strength at the level of increasingly high fluctuations. In the sectoral cross-section, change in correlation strength at the high and low fluctuation levels is the most pronounced for Services sector companies.

To assess the strength of fractal cross-correlation between two signals on a particular scale, the generalized coefficient $\rho^q(s)$ was used, which relies on the MFDFA and MFCCA methods. This allowed the numerical estimation of the nonlinear correlation on the selected time scale s, while considering the fluctuation amplitude. The coefficient was used for making a comprehensive analysis of the largest and the smallest American companies listed in the period 1997–1999, while considering three characteristic time scales: hourly, daily and weekly.

The analysis of $\rho^q(s)$ coefficients, determined for the largest companies, shows that, for the hourly, daily and weekly scales, there is no correlation between companies at the level of the lowest fluctuations ($q < 0$). For large fluctuations and moment $q > 0$, in the majority of instances, a small correlation exists between companies, which becomes weaker with the increase in q.

A similar analysis of $\rho^q(s)$ coefficient, made for the smallest companies, has shown that for moments $q < 0$ most often there is no correlation between companies on all time scales. For positive values of moments q and for the daily and weekly scales, also no correlation essentially exists between companies. For the weekly scale, it can be noticed that $\rho^q(s)$ assumes, in the majority of cases, values different from 0, and the distribution itself is stable with respect to increase in q. This is a difference between the behaviour of the largest companies, which implies that for the smallest (and, at the same, most seldom traded) companies, a longer time horizon is needed for nonlinear correlations to occur at the large fluctuation level.

For network analysis, the MST representation was used. To verify whether MST trees can be built based on the $\rho^q(s)$ coefficient, it was checked between each three-element company combination to see if the adopted distance $d_{x,y}^s$ between network nodes meets the metric conditions. From the analysis made on all time scales and moments $-4 \leq q \leq 4$ it was demonstrated that for $q < 0$ the length defined with $\rho^q(s)$ did not meet the metric conditions (the triangle inequality condition for the distance was not satisfied), therefore the analysis was limited to moments $q = \{2; 3; 4\}$, for which the triangular relation was met.

In the case of the largest company network determined based on large fluctuations, we have a scale-free, heavily centralized network with the visible financial, technology and health care sectors for the hourly and daily scales, supplemented with services and basic materials sectors for the weekly scale. The tree structure is not stable

in respect of change in times scale and signal amplitude. For the smallest scales, the GE company is the central network node (for the weekly scale and highest moment, this is the HD company), and other nodes with a multiplicity > 1 appear in the network, which also perform the function of hubs. The average betweenness coefficient b_i determined for the network assumes values closer to 1, which confirms the existence of many nodes in the network, which are important to its operation, but which, however, lose importance with the increase in fluctuation amplitude. The average path lengths L determined for the networks are indicative of a more dispersed nature of the networks on all scales comparing to the one, determined for Pearson correlation coefficient.

The analysis made for the smallest companies is similar. The determined network is a scale-free and centralized network, whose central nodes are investment funds that make investments based on stock exchange indexes. For the hourly scale, the financial and basic materials clusters are visible, but in contrast to the largest company network, there are no sectoral clusters for the daily and weekly scales clearly visible. Admittedly, there are nodes with a large multiplicity, but without the hallmarks of sectoral membership. The betweenness and average path length results take on values similar to those obtained for the largest companies. It appears, therefore, that the obtained structure and properties of the network of companies are independent of their capitalization. The only difference is that strong sectoral relationships can be observed for the largest companies at the big price movement level.

To sum up, the analyses carried out in the study and the obtained results provide a contribution to the knowledge that we have on the complex and rich structure of financial markets in a broad sense, and disclose new facts about the interrelations among their components. The multifractal analysis of the interrelations between market elements shows the existence of such interrelation at the level of large fluctuations. Particularly promising is the use of the q-generalized nonlinear correlation coefficient ρ^q, which, combined with the network representation of interrelations, helps to perceive previously unknown facts, namely the existence of nonlinear correlations between capital market companies at the large fluctuation level, which are additionally unstable in respect of the change (amplification) of the signal amplitude and the time scales examined.

7 Bibliography

1. **Smith A.** *An Inquiry into the Nature and Causes of the Wealth of Nations.* Oxford : Liberty Classics/Oxford Iniversity Press, 1981.
2. **Mantegna R.N., Stanley H.E.** *Ekonofizyka-wprowadzenie.* Warszawa : PWN, 2001. 83-01-13524-7.
3. **Bachelier L.** Theorie de la speculation. *Annales Scientifiques de L'Ecole Normale Superieure.* 1900, Vol. 17, 21.
4. **Drożdż S., Forczek M., Kwapień J., Oświęcimka P., Rak R.** Stock market return distributions: from past to present. *Physica A.* 2007, Vol. 383, 59.
5. **Plerou V., Gopikrishnan P., Nunes Amaral L.A., Meyer M., Stanley H.E.** Scaling of the distribution of price fluctuations of individual companies. *Phys. Rev. E.* 1999, Vol. 60, 6519.
6. **Mandelbrot B.** The Variation of Certain Speculative Prices. *The Journal of Business.* 1963, Vol. 36, 394.
7. **Stanley H.E., Mantegna R.N.** Stochastic Process with Ultraslow Convergence to a Gaussian: The Truncated Lévy Flight. *Phys. Rev. Lett.* 73, 1994, 2946.
8. **Gopikrishnan P., Plerou V., Nunes Amaral L.A.** Scaling of the Distributions of Fluctuations of Financial Market Indices. *Phys. Rev. E.* 1999, Vol. 60, 5305.
9. **Pareto V.** Cours d'Economie Politique. *Journal of Political Economy.* 1896, Vol. 6, 4.
10. **Bacry E., Delour J., Muzy J. F.** Multifractal Random Walk. *Phys. Rev. E.* 64, 2001, 026103.
11. **Muzy J.F., Delour J., Bacry E.** Modelling fluctuations of financial time series: from cascade process to stochastic volatility model. *Eur. Phys. J. B.* 2000, Vol. 17, 537.
12. **Mandelbrot B., Fisher A., Calvet L.** A Multifractal Model of Asset Returns. *Cowles Foundation Discussion Paper, Yale University.* 1997.
13. **Sreenivasan K.R.** Fractals and Multifractals in Fluid Turbulence. *Annu. Rev. Fluid Mech.* 1991, Vol. 23, 539.
14. **Kolmogorow A.N.** The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *Dokl. Akad. Nauk SSSR.* 1941.
15. **Mandelbrot B.** Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier. *Journal of Fluid Mechanics 62.* 1974, Vol. 62, 331.
16. **Frisch U. Sulem P.L., Nelkin M.** A simple dynamical model of intermittent turbulence. *Journal of Fluid Mechanics.* 1978, Tom 87, 719.
17. **Frisch U.** From Global Scaling, a la Kolmogorov, to Local Multifractal Scaling in Fully Developed Turbulence. *Proceedings of The Royal Society A.* 1991, Vol. 434, 89.
18. **Meneveau C., Sreenivasan K.R.** Simple multifractal cascade model for fully developed turbulence. *Phys. Rev. Lett.* 1987, Vol. 59, 1424.
19. **Meneveau C., Sreenivasan K.R**. The multifractal nature of turbulent energy dissipation. *Journal of Fluid Mechanics.* 1991, Vol. 224, 429.
20. **Sreenivasan K.R., Meneveau C.** The fractal facets of turbulence. *Journal of Fluid Mechanics, 173.* 1986, Vol. 173, 357.
21. **Ghashghaie S., Breymann W., Peinke J., Talkner P., Dodge Y.** Turbulent cascades in foreign exchange markets. *Nature.* 1996, Vol. 381, 767.
22. **Arneodo A., Bouchaud J.P., Cont R., Muzy J.F., Sornette D.** Comment on "Turbulent cascades in foreign exchange markets". *Science & Finance, Capital Fund Management, Science & Finance (CFM) working paper archive.* 1996, Vol. 9607120.

23. **Mantegna R.N., Stanley H.E.** Turbulence in Financial Markets. *Nature.* 1996, Vol. 283, 587.
24. **Shayeganfar F.** Levels of complexity in turbulent time series for weakly and high Reynolds number. *Physica A.* 2012, Vol. 391, 3151.
25. **Lux T.** Turbulence in financial markets: the surprising explanatory power of simple cascade models. *Quantitative Finance.* 2001, Vol. 1, 632.
26. **Ecke R.** The Turbulence Problem: An Experimentalist's Perspective. *Los Alamos Science.* 2005, Vol. 29, 124.
27. **Mandelbrot B., Van Ness J.W.** Fractional Brownian motions, fractional noises. *SIAM Review.* 1968, Vol. 10, 422.
28. **Tsallis C.** Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics.* 1988, Vol. 52, 479.
29. **Tsallis C., Levy S. V. F., Souza A. M. C., Maynard R.** Statistical-Mechanical Foundation of the Ubiquity of Lévy Distributions in Nature. *Phys. Rev. Lett.* 1995, Vol. 75, 3589.
30. **Drożdż S., Kwapień J., Oświecimka P., Rak R.** Quantitative features of multifractal subtleties in time series. *EPL (Europhysics Letters).* 2009, Vol. 88, 60003 .
31. **Drożdż S., Kwapień J., Oświęcimka P., Rak R.** The foreign exchange market: return distributions,multifractality, anomalous multifractality and Epps effect. *New Journal of Physics.* 2010, Vol. 12, 105003.
32. **Kwapień J., Drożdż S., Speth J.** Alternation of different fluctuation regimes in the stock market dynamics. *Physica A* . 2003, Vol. 330, 605.
33. **Matia K., Nunes Amaral L.A, Goodwin S.P., Stanley H.E.** Different scaling behaviors of commodity spot and future prices. *Phys. Rev. E.* 202, Vol. 66, 04510.
34. **Drożdż S., Kwapień J., Grümmer F., Ruf F., Speth J.** Are the Contemporary Financial Fluctuations Sooner Converging to Normal? *Acta Phys. Pol. B.* 2003, Vol. 34, 4293.
35. **Mantegna R.N., Stanley H.E.** Scaling Behaviour in the Dynamics of an Economic Index. *Nature.* 1995, Vol. 376, 46.
36. **Liu Y., Gopikrishnan P., Cizeau P., Meyer M., Peng C.-K., Stanley H.E.** Statistical properties of the volatility of price fluctuations. *Phys. Rev. E.* 1999, Vol. 60, 1390.
37. **Mandelbrot B.** *The Fractal Geometry of Nature.* New York : W.H. Freeman and Company, 1983.
38. **Mandelbrot B.** *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk.* New York : Springer, 1997.
39. **Falconer K.** *Fractal Geometry: Mathematical Foundations and Applications.* England : John Wiley & Sons Inc., 2003.
40. **Seuret S., Véhel J.L.** The local Hölder function of a continuous function. *Applied and Computational Harmonic Analysis.* 2002, Vol. 13, 263.
41. **Stauffer D., Aharony A.** *Introduction to Percolation Theory.* US : Taylor-Francis, 1992.
42. **Dommers S., Giardinà C., Hofstad van der R.** Ising Models on Power-Law Random Graphs. *Journal of Statistical Physics.* 2010, Vol. 141, 638.
43. **Amaral L.A.N, Scala A., Barthelemy M., Stanley H.E.** Classes of small-world networks. *PNAS.* 2000, Vol. 97, 11149.
44. **Zipf G.K.** *Selective studies and the principle of relative frequency in language.* Harvard : Harvard University Press, 1932.
45. **Kwapień J., Drożdż S.** Physical approach to complex systems. *Physics Reports.* 2012, Vol. 515, 115.

46. **Gell-Man M.** What is Complexity. *Complexity, John Wiley & Sons, Inc.* 1995, Vol. 1, 16.
47. **Górski A.Z., Drożdż S., Mokrzycka A., Pawlik J.** Accuracy Analysis of the Box-Counting Algorithm. *Acta Physica Polonica A.* 2012, Vol. 121, 28.
48. **Arneodo A., Bacry E., Muzy J.F.** The thermodynamics of fractals revisited with wavelets. *Physica A.* 1995, Vol. 213, 232.
49. **Grassberger P., Procaccia I.** Measuring the strangeness of strange attractors. *Physica D.* 1983, Vol. 9, 189.
50. **Tél T.** Fractals, multifractals, and thermodynamics. *Zeitschrift für Naturforschung A.* 1988, Vol. 43, 1154.
51. **Wawrzaszek A., Macek W.M.** Observation of the multifractal spectrum in solar wind turbulence by Ulysses at high latitudes. *Journal of Geophysical Reserarch-Space Physics.* 2010, Vol. 115, A07104.
52. **Xue C., Shang P., Jing W.** Multifractal detrended cross-correlation of BVP model time series. *Nonlinear Dynamics.* 2012, Vol. 69, 263.
53. **Li Z., Lu X.** Cross-correlations between agricultural commodity futures markets in the US and China. *Physica A.* 2012, Vol. 391, 3930.
54. **Udovichenko V.V., Strizhak P. E.** Multifractal Properties of Copper Sulfide Film Formed in Self-Organizing Chemical System. *Theoretical and Experimental Chemistry.* First, 202, Vol. 38, 259.
55. **Viswanathan G. M., Buldyrev S. V., Havlin S., Stanley H. E.** Quantification of DNA Patchiness Using Long-Range Correlation Measures. *Biophys J.* 1997, Vol. 72, 866.
56. **Oświęcimka P., Kwapień J., Celińska I., Drożdż S., Rak R.** *Computational approach to multifractal music.* s.l. : arXiv:1106.2902.
57. **Ivanova K., Ausloos M.** Application of the detrended fluctuation analysis (DFA) method for describing cloud breaking. *Physica A.* 1999, Vol. 274, 349.
58. **Liu Y., Cizeau P., Meyer M., Peng C.-K.** Correlations in economic time series. *Physica A.* 1997, Vol. 245, 437.
59. **Górski A. Z., Drożdż S., Speth J.** Financial multifractality and its subtleties: an example of DAX. *Physica A: Statistical Mechanics and its Applications.* 2002, Vol. 316, 496.
60. **Czarnecki Ł., Grech D.** Multifractal dynamics of stock markets. *Acta Physica Polonica A.* 2010, Vol. 117, 623.
61. **Peng C.-K., Buldyrev S. V., Havlin S., Simons M., Stanley H. E., Goldberger A. L.** Mosaic organization of DNA nucleotides. *Phys. Rev. E.* 1994, Vol. 49, 1685.
62. **Kantelhardt J.W, Zschiegner S.A, Koscielny-Bunde E., Bunde A., Havlin S., Stanley H.E.** Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A.* 2002, Vol. 316, 87.
63. **Oświęcimka P., Drożdż S., Kwapień J., Górski A.Z.** Effect of detrending on multifractal characteristics. *Acta Physica Polonica, A.* 2013, Vol. 123, 597.
64. **Jaffard S., Lashermes B., Abry P.** *Wavelet Analysis and Applications Applied and Numerical Harmonic Analysis.* Germany : Springer, 2007. 201.
65. **Podobnik B., Stanley H.E.** Detrended Cross-Correlation Analysis: A New Method for Analyzing Two Nonstationary Time Series. *Phys. Rev. Lett.* 2008, Vol. 100, 084102.
66. **Zhou W.-X.** Multifractal detrended cross-correlation analysis for two nonstationary signals. *Phys. Rev. E.* 2008, Vol. 77, 066211.
67. **Oświęcimka P., Drożdż S., Forczek M., Jadach S., Kwapień J.** Detrended Cross-Correlation Analysis Consistently Extended to Multifractality. *Phys. Rev. E.* 2014, Vol. 89, 023305 .

68. **Hurst H.E.** Long-term storage capacity of reservoirs. *Trans. Amer. Soc. Civil Engineers.* 1951, Vol. 116, 770.
69. **Benmehdi S., Makarava N., Benhamidouche N., Holschneider M.** Bayesian estimation of the self-similarity exponent of the Nile River. *Nonlinear Processes in Geophysics.* 2011, Vol. 18, 441.
70. **Peng C.K., Havlin S., Stanley H.E., Goldberger A.L.** Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos.* 1995, Vol. 5, 82.
71. **Oświęcimka P., Kwapień J., Drożdż S.** Wavelet versus detrended fluctuation analysis of multifractal structures. *Phys. Rev. E.* 2006, Vol. 74, 016103.
72. **Kantelhardt J. W., Koscielny-Bunde E., Regoa H.H.A, Havlin S., Bunde A.** Detecting long-range correlations with detrended fluctuation analysis. *Physica A.* 2001, Vol. 295, 441.
73. **Calvet L. E., Fisher A.J.** How to Forecast Long-Run Volatility: Regime Switching and the Estimation of Multifractal Processes. *Journal of Financial Econometrics.* 2004, Vol. 2, 49.
74. **Oświęcimka P., Kwapień J., Drożdż S.** Multifractality in the stock market: increments versus waiting times. *Physica A.* 2005, Vol. 347, 626.
75. **Mandelbrot B.** Multifractal Power Law Distributions: Negative and Critical Dimensions and Other "Anomalies",Explained by a Simple Example. *Journal of Statistical Physics.* 2003, Vol. 110, 739.
76. **Drożdż S., Oświęcimka P.** Detecting and interpreting distortions in hierarchical organization of complex time series. *Phys. Rev. E.* 2015, Vol. 91, 030902.
77. **Mandelbrot B., Evertsz C.J.G., Hayakawa Y.** Exactly self-similar left-sided multifractal measures. *Phys. Rev. A.* 1990, Vol. 42, 4528.
78. **Jaffard S.** The multifractal nature of Levy processes. *Probability Theory and Related Fields.* 1999, Vol. 114, 207.
79. **Zhou W.-X.** The components of empirical multifractality in financial returns. *EPL (Europhysics Letters).* 2009, Vol. 88, 28004.
80. **Kwapień J., Oświęcimka P. , Drożdż S.** Components of multifractality in high-frequency stock returns. *Physica A.* 2005, Vol. 350, 466.
81. **Matia K., Ashkenazy Y., Stanley H.E.** Multifractal properties of price fluctuations of stocks and commodities. *Europhysics Letters.* 2003, Vol. 61, 422.
82. **Grech D., Pamuła G.** On the multifractal effects generated by monofractal signals. *Physica A.* 2013, Vol. 392, 5845.
83. **Kristoufek L.** Multifractal Height Cross-Correlation Analysis: A New Method for Analyzing Long-Range Cross-Correlations. *EPL (Europhysics Letters).* 2011, Vol. 95, 68001.
84. **Shadkhoo S., Jafari G.R.** Multifractal detrended cross-correlation analysis of temporal and spatial seismic data. *Phys. J. B.* 2009, Vol. 72, 679.
85. **Podobnik B., Horvatic D., Petersen A. M., Stanley H.E.** Cross-correlations between volume change and price change. *Proc. Natl. Acad. Sci. USA.* 2009, Vol. 106, 22079.
86. **Zhao X., Shang P., Shi W.** Multifractal cross-correlation spectra analysis on Chinese stock markets. *Physica A.* 2014, Vol. 402, 84.
87. **Dashtian H., Jafari G.R., Koohi Z.L., Masihi M., Sahimi M.** Analysis of Cross Correlations Between Well Logs of Hydrocarbon Reservoirs. *Transport in Porous Media.* 2011, Vol. 90, 445.

88. **He L.-Y., Chen S.-H.** Nonlinear bivariate dependency of price–volume relationships in agricultural commodity futures markets: A perspective from Multifractal Detrended Cross-Correlation Analysis. *Physica A.* 2011, Vol. 390, 297.
89. **Kristoufek L.** Measuring correlations between non-stationary series with DCCA coefficient. *Physica A.* 2014, Vol. 402, 291.
90. **Hosking J.R.M.** Fractional Differencing. *Biometrika.* 1981, Vol. 68, 165.
91. **Zebende G.F.** DCCA cross-correlation coefficient: Quantifying level of cross-correlation. *Physica A.* 2011, Vol. 390, 614.
92. **Vassolera R.T., Zebende G.F.** DCCA cross-correlation coefficient apply in time series of air temperature and air relative humidity. *Physica A.* 2012, Vol. 391, 2438.
93. **Podobnik B., Jiang Z.-Q., Zhou W.-X., Stanley H.E.** Statistical tests for power-law cross-correlated processes. *Phys. Rev. E.* 2011, Vol. 84, 066118.
94. **Kwapień J., Oświęcimka P., Drożdż S.** Detrended fluctuation analysis made flexible to detect range of cross-correlated. *Phys. Rev. E.* 2015, Vol. 92, 052815.
95. **Albert R., Barabasi A.L.** Statistical mechanics of complex networks. *Reviews of Modern Physics.* 2012, Vol. 74, 47.
96. **Mantegna R.N.** Hierarchical structure in financial markets. *Eur. Phys. J. B.* 1999, Vol. 11, 193.
97. **Kwapień J., Gworek S., Drożdż S., Górski A.** Analysis of a network structure of the foreign currency exchange market. *Journal of Economic Interaction and Coordination.* 2009, Vol. 1, 55.
98. **Naylor M.J., Rose L., Moyle B. J.** Topology of foreign exchange markets using hierarchical structure mehods. *Physica A.* 2007, Vol. 382, 199.
99. **Jang W., Lee J., Chang W.** Currency crises and the evolution of foreign exchange market: Evidence from minimum spanning tree. *Physica A.* 2011, Vol. 390, 707.
100. **Sieczka P., Hołyst J.A.** Correlations in commodity markets. *Physica A.* 2009, Vol. 388, 1621.
101. **Onnela J.-P., Chakraborti A., Kaski K., Kertesz J., Kanto A.** Dynamics of market correlations: Taxonomy and portfolio analysis. *Phys. Rev. E.* 2003, Vol. 68, 056110.
102. **Bonanno G., Vandewalle N., Mantegna R. N.** Taxonomy of Stock Market Indices. *Phys. Rev. E.* 2000, Vol. 62, 7615.
103. **Bonanno G., Caldarelli G., Lillo F., Micciché S.,Vandewalle N., Mantegna R. N.** Networks of equities in financial markets. *The European Physical Journal B.* 2004, Vol. 38, 363.
104. **Ravasz E., Barabsi A.L.** Hierarchical organization in complex networks. *Phys. Rev. E.* 2003, Vol. 67, 026112.
105. **Jeong H., Tombor B., Albert R., Oltvai Z. N. Barabási A.-L.** The large-scale organization of metabolic networks. *Nature.* 2000, Vol. 407, 651.
106. **Erdos P., Renyi A.** On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences.* 1960, Vol. 5, 17.
107. **Albert R., Jeong H., Barabási A.L.** Error and attack tolerance in complex networks. *Nature.* 2000, Vol. 406, 378.
108. **Boccaletti S., Latora V., Moreno Y., Chavez M. , Hwang D.U.** Complex networks: Structure and dynamics. *Physics Reports.* 2006, Vol. 424, 175.
109. **Neal Z.** The devil is in the details: Differences in air traffic networks by scale, species, and season. *Social Networks.* 2014, Vol. 38, 63.
110. **Barabasi A.L., Albert R.** Emergence of Scaling in Random Networks. *Science.* 286, 1999, 509.

111. **Barthélemy M.** Betweenness centrality in large complex networks. *The European Physical Journal B.* 2004, Vol. 38, 163.
112. **Wiliński M.J., Sienkiewicz A., Gubiec T., Kutner R., Struzik Z.R.** Structural and topological phase transitions on the German Stock Exchange. *Physica A.* 2013, Vol. 392, 5963.
113. **Kruskal J.B.** On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem. *American Mathematical Society.* 1956, Vol. 7, 48.
114. **Górski A. Z., Drożdż S., Kwapień J.** Scale free effects in world currency exchange network. *Eur. Phys. J. B.* 2008, Vol. 66, 91.
115. **Vandewalle N., Brisbois F., Tordoir X.** Non-random topology of stock markets. *Quantitative Finance.* 2001, Vol. 1, 372.
116. **Markowitz H.M.** Portfolio Selection. *The Journal of Finance.* 1952, Vol. 7, 77.
117. **Serranoa E., Figliola A.** Wavelet Leaders: A new method to estimate the multifractal singularity spectra. *Physica A: Statistical Mechanics and its Applications.* 2009, Vol. 388, 2793.
118. **Cont R.** Empirical properties of asset returns:stylized facts and statistical issues. *Quantitative Finance.* 2001, Vol. 1, 223.
119. **Sornette D.** *Critical Phenomena In Natural Sciences.* New York : Springer Science & Business, 2006.

8 The list of figures and tables

Table 8.1 The list of DJIA stocks with sector assignment.

Symbol	Name	Sector
AA	Alcoa Inc.	Basic Materials
CVX	Chevron Corporation	Basic Materials
DD	E. I. du Pont de Nemours and Company	Basic Materials
XOM	Exxon Mobil Corporation	Basic Materials
KFT	Mondelez International, Inc.	Consumer Goods
KO	The Coca-Cola Company	Consumer Goods
PG	The Procter & Gamble Company	Consumer Goods
AIG	American International Group, Inc.	Financial
AXP	American Express Company	Financial
BAC	Bank of America Corporation	Financial
JPM	JPMorgan Chase & Co.	Financial
TRV	The Travelers Companies, Inc.	Financial
JNJ	Johnson & Johnson	Healthcare

PFE	Pfizer Inc.	Healthcare
MRK	Merck & Co. Inc.	Healthcare
BA	The Boeing Company (BA)	Industrial Goods
CAT	Caterpillar Inc.	Industrial Goods
MMM	3M Company	Industrial Goods
UTX	United Technologies Corporation	Industrial Goods
GE	General Electric Company	Industrial Goods
HD	The Home Depot, Inc.	Services
MCD	McDonald's Corp.	Services
WMT	Wal-Mart Stores Inc.	Services
CSCO	Cisco Systems, Inc.	Technology
IBM	International Business Machines Corporation	Technology
INTC	Intel Corporation	Technology
MSFT	Microsoft Corporation	Technology
T	AT&T, Inc.	Technology
VZ	Verizon Communications Inc.	Technology

Table 8.2 The list of 100 biggest stock in US with sector assignment.

Symbol	**Name**	**Sector**	**Symbol**	**Name**	**Sector**
AA	Alcoa	Basic Materials	PNU	Pharmacia & Upjohn	Healthcare
CHV	Chevron	Basic Materials	ALD	AlliedSignal	Industrial Goods
DD	DuPont De Nemours	Basic Materials	BA	Boeing	Industrial Goods
GTE	Gran Tierra Energy	Basic Materials	GE	General Electric	Industrial Goods
XOM	Exxon	Basic Materials	MMM	3M Corp.	Industrial Goods
TX	Texaco	Basic Materials	UTX	United Technologies	Industrial Goods
ENE	Enron	Basic Materials	EMR	Emerson Electric Co	Industrial Goods
DOW	The Dow Chemical Company	Basic Materials	ARC	Atlantic Richfield	Services
SLB	Schlumberger Limited	Basic Materials	BEL	Bell	Services
BUD	Anheuser-Busch	Consumer Goods	BLS	BellSouth	Services
CCU	Compania Cervecerias Unidas S.A	Consumer Goods	CBS	CBS	Services
CL	Colgate-Palmolive	Consumer Goods	CCL	Carnival Corp.	Services

G	Gillette	Consumer Goods	CMCS	Comcast Corporation	Services
GM	General Motors	Consumer Goods	COX	Cox Communications	Services
KBM	Kimberly-Clark	Consumer Goods	DH	Dayton Hudson	Services
EDS	Electronic Data Systems	Consumer Goods	DIS	Walt Disney	Services
MO	Philip Morris	Consumer Goods	GCI	Gannett Co., Inc	Services
KO	Coca-Cola	Consumer Goods	HD	Home Depot	Services
F	Ford Motor	Consumer Goods	WMT	Wall-Mart Stores	Services
EK	Eastman Kodak	Consumer Goods	MCD	McDonald's Corp	Services
PEP	PepsiCo	Consumer Goods	VIA	Viacom	Services
PG	Procter & Gamble	Consumer Goods	LOW	Lowe's Companies Inc.	Services
AHP	American Home Products	Financial	TWX	Time Warner	Services
AIG	AIG	Financial	T	AT&T	Services
AXP	American Express	Financial	QWST	Qwest Communications Intl.	Services
BAC	Bank America	Financial	SBC	SBC Communications	Services
BK	Bank of New York	Financial	AOL	America Online	Technology
C	Citigroup	Financial	AUD	Automatic Data Processing	Technology
CMB	Chase Manhattan	Financial	CA	Computer Associates	Technology
JPM	J.P. Morgan	Financial	CPQ	Compaq Computer	Technology
MMC	Marsh & McLennan	Financial	CSCO	Cisco Systems	Technology
WFC	Wells Fargo	Financial	DELL	Dell	Technology
MER	Merrill Lynch	Financial	GTW	Gateway 2000	Technology
FBF	Fleet Financial Group	Financial	HWP	Hewlett Packard	Technology
FNM	Fannie Mae	Financial	IBM	IBM	Technology
ONE	Bank One	Financial	INTC	Intel Corp.	Technology
MWD	Morgan Stanley	Financial	LU	Lucent Technologies	Technology
FRE	Freddie Mac	Financial	YHOO	Yahoo!	Technology
FTU	First Union	Financial	WCOM	MCI WorldCom	Technology
SCH	Charles Schwab	Financial	USW	US West	Technology
ABT	Abott Labolatories	Healthcare	UMG	MediaOne Group	Technology
AMGN	Amgen Inc.	Healthcare	TXN	Texas Instruments	Technology
BMY	Bristol-Myers Squibb	Healthcare	EMC	EMC Corporation	Technology
LLY	Eli Lilly and Company	Healthcare	MOT	Motorola	Technology

MDT	Medtronic	Healthcare	MSFT	Microsoft Corp.	Technology
JNJ	Johnson & Johnson	Healthcare	GLW	Corning Inc.	Technology
MTC	Monsanto	Healthcare	SUNW	Sun	Technology
MRK	Merck	Healthcare	QCOM	Qualcomm	Technology
SGP	Schering-Plough	Healthcare	ORCL	Oracle	Technology
PFE	Pfizer	Healthcare	AT	Alltel	Utilities

Table 8.3 The list of 100 smallest stock in US market with sector assignment.

Symbol	Name	Sector	Symbol	Name	Sector
ADF	Acm Managed Dollar Incm Fd In	Financial	JPR	Jp Realty Inc	Financial
ADV	Advest Group Inc The	Financial	KPA	Innkeepers Usa Trust	Financial
AEC	Associated Estates Realty Cor	Financial	KTO	K2 Inc	Consumer Goods
AEM	Agnico-Eagle Mines Ltd	Basic Materials	LTC	Tc Properties Inc	Financial
AGI	Alpine Group Inc. The	Industrial Goods	MAG	Magna Intl Cl-A Sub Vtg Shs	Services
AJL	Mway Japan Ltd Ads 1/2 Com	Consumer Goods	MCL	Moore Corp Ltd	Consumer Goods
APB	Asia Pacific Fund Inc	Financial	MGL	Magellan Health Services Inc	Healthcare
ASA	Asa Ltd	Financial	MLR	Miller Industries In	Consumer Goods
AVL	Aviall Inc	Industrial Goods	MSD	Morgan Stanley Emer Mkts Debt	Financial
AVS	Aviation Sales Company	Industrial Goods	NS	National Steel Corp Cl-B	Basic Materials
BBA	Bombay Co Inc The	Consumer Goods	NSS	Ns Group Inc	Basic Materials
BCU	Borden Chem Plastic L.P.	Basic Materials	NWK	Network Equipment Tech Inc	Technology
BIR	Birmingham Steel Corp	Basic Materials	OHI	Omega Healthcare Investors Inc	Financial
BOY	Boykin Lodging Company	Financial	OMI	Wens Minor Inc Hldg Co	Services
BPT	Bp Prudhoe Bay Royalty Tr	Basic Materials	OS	Oregon Steel Mills Inc	Industrial Goods
BUR	Burlington Res Coal Seam Gas	Basic Materials	PAR	Coastcast Corp	Industrial Goods
BXG	Bluegreen Corp	Consumer Goods	PEI	Penn Real Estate Invst Tst Sbi	Financial

BZF	Brazil Fund Inc	Financial	PIN	Amf Bowling Inc	Consumer Goods
CCC	Calgon Carbon Corp	Industrial Goods	POP	Pope Talbot Inc	Basic Materials
CH	Chile Fund Inc	Financial	PWN	Cash America International Inc	Financial
CNO	Cornerstone Propane Ptners L.P	Financial	RCS	Rcm Strategic Global Govt Fd	Financial
COP	Consolidated Products Inc	Services	RFS	Rfs Hotel Investors Inc	Financial
CRY	Cryolife Inc.	Healthcare	RGO	Ranger Oil Limited	Basic Materials
CWN	Crown American Realty Trust	Financial	RGR	Sturm Ruger Co Inc	Industrial Goods
DBS	Debt Strategies Fund Inc.	Financial	RTI	Rmi Titanium Co New	Basic Materials
DK	Donna Karan International Inc	Consumer Goods	SAH	Sonic Automotive Inc Cl A	Services
DMN	Dimon Corp	Consumer Goods	SBG	Salomon Bro2008Wldwide Govt Tr	Financial
EDF	Merging Mkts Incm Fd Ii Inc	Financial	SFY	Swift Energy Co	Basic Materials
EMF	Templeton Emerging Mkt Fund	Financial	SIE	Sierra Health Services Inc	Healthcare
ESL	Esterline Technologies Corp	Industrial Goods	SKT	Tanger Factory Outlet Centers	Financial
EWF	European Warrant Fund Inc	Financial	SRT	Startek Inc.	Services
EY	Ethyl Corp	Basic Materials	SSS	Sovran Self Storage Inc	Financial
FA	Fairchild Corp Cl-A	Services	TCH	Templeton China World Fund Inc	Financial
FC	Franklin Covey Co.	Services	TEE	National Golf Properties Inc	Financial
FF	First Financial Fund Inc	Financial	TMA	Thornborg Mortgage Asset Corp	Financial
FIX	Comfort Systems Usa Inc	Industrial Goods	TMR	The Meridian Resource Corp	Basic Materials
FLH	Fila Hldgs Spa Adr 5Ord Of Lit	Consumer Goods	TRP	Transcanada Pipelines Limited	Basic Materials
GDC	General Datacomm Ind Inc	Technology	TYL	Tyler Corporation	Technology
GER	Germany Fund Inc	Financial	UAG	United Auto Group Inc Vtg	Services

GHI	Global High Incm Dollar Fd Inc3	Financial	UH	U.S. Home Corp New	Consumer Goods
GRO	Mississippi Chemical Corp	Industrial Goods	UNT	Unit Corporation	Basic Materials
GRR	Asia Tigers Fund Inc The	Financial	USV	U.S.Restaurant Props Inc.	Services
HGR	Hanger Orthopedic Group Inc	Healthcare	VTO	Vitro Sociedad Anonima Ads	Basic Materials
HXL	Hexcel Corp New	Industrial Goods	WHC	Wackenhut Corrections Corp	Industrial Goods
IIR	Iri International Corp	Basic Materials	WHX	Whx Corp Holding Co	Basic Materials
IMC	Intl Multifoods Corp	Consumer Goods	WLV	Wolverine Tube Inc	Industrial Goods
IMR	Imco Recycling Inc	Basic Materials	WS	Weirton Steel Corp	Basic Materials
IRT	Irt Property Co	Consumer Goods	WSO	Watsco Inc	Services
ITX	International Tech Cp New	Services	WTS	Watts Industries Inc Cl A	Industrial Goods
JOF	Japan Otc Equity Fund Inc	Financial	WXH	Winston Hotels Inc	Financial

Table 8.4

All calculations were performed using Matlab environment on Zeus computer in Academic Computer Centre Cyfronet AGH in Kraków.