Warsaw, November'21 2008

Algorithm of the Multifractal Detrended Fluctuation Analysis (MF-DFA) and some remarks concerning our case

Ryszard Kutner

1 Two basic sources of multifractality

Two basic sources of multifractality in time series can be distinguished:

(i) the first one can be caused by a broad probability density function for the values of the time series,

(ii) the long-term correlations between elements of time-series could be the second source.

In principle, the combination of the above two, i.e.,

(iii) the hybrid source,

is also possible.

2 The aim of the MF-DFA

The aim of the MF-DFA [1, 2] is to find the spectrum of singularities $f(\eta)$[1] both for stationary as well as nonstationary time series $\{x_k\}_{k=1}^{N}$. In our particular case, time series should consist of intertransaction time intervals, i.e. $x_k \stackrel{\text{def.}}{=} \Delta t_k,\ k = 1, \ldots, N$. The elements $\Delta x_k \stackrel{\text{def.}}{=} x_k - \langle x \rangle$ are formally considered as single displacements, where mean value $\langle x \rangle \stackrel{\text{def.}}{=} \frac{1}{N}\sum_{k=1}^{N} x_k$. In this note we tested MF-DFA by using time series for WIG (the Warsaw Stock Exchange Index), cf. Fig.1.

The algorithm which makes possible to calculate the spectrum $f(\eta)$ consists of five stages given below.

[1] In quoted bibliography the variable α is used instead of η given here, since in our case α is reserved for the exponent defining the stretched exponential function.

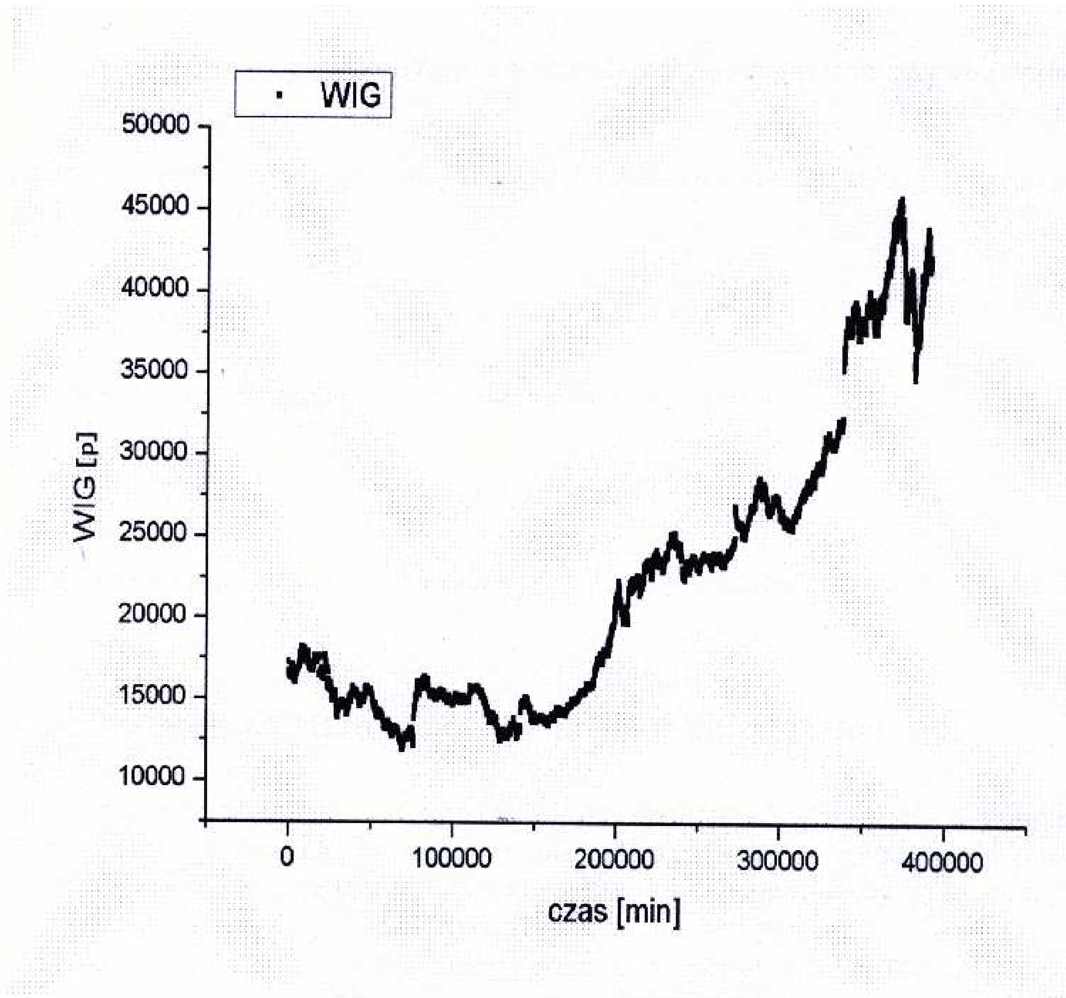

Figure 1: Daily time-series (fixing) of WIG20 vs. time (in minutes) from October 2006 till June 2006 on Warsaw Stock Exchange (this and all remaining figures were taken from [3]).

2.1 The algorithm

2.1.1 Stage 1

In this step we define the sequence of summary displacements (profiles) as follows:

$$Y(j) \stackrel{\text{def.}}{=} \sum_{k=1}^{j} (x_k - \langle x \rangle), \; j = 1, \ldots, N. \tag{1}$$

Note that from this definition one has $Y(N) = 0$.

2.1.2 Stage 2

We divide our support (which consists of numbers $\{k\}_{k=1}^{N}$) into $N_s = INT(N/s)$ nonoverlapping segments of length s. Since in general s and N are incommensurate numbers, a rest part (shorter then s) may remain. In order to incorporate this rest part of the series, the analogous procedure is repeated starting from the opposite end (which begins from the largest number N of the series). As a result of this procedure, $2N_s$ segments are obtained altogether.

2.1.3 Stage 3

The aim of this stage is to detrend profiles $Y(j),\ \ j = 1, \ldots, N$, in each segment of given length s, separately. The least-square fit is made by determination of χ^2-functions (variances) for each segment $\nu = 1, \ldots, N_s$:

$$F^2(\nu, s) = \frac{1}{s}\sum_{j=1}^{s}[Y((\nu-1)s + j) - w_\nu^n(j)]^2, \tag{2}$$

while for segments $\nu = N_s+1, \ldots, 2N_s$, the corresponding variances are defined as follows:

$$F^2(\nu, s) = \frac{1}{s}\sum_{j=1}^{s}[Y((N-\nu-N_s)s + j) - w_\nu^n(j)]^2, \tag{3}$$

where $w_\nu^n(j)$ is the fitting polynomial in the segment ν of the order n fixed for the whole procedure. Hence, we call the Multifractal Detrended Fluctuation Analysis as the MF-DFAn, respectively (i.e. the nth order of the MF-DFA). Thus a comparison of the results for different orders of the MF-DFA allows one to estimatethe the order of the polynomial segment trends in the time series.

2.1.4 Stage 4

In this stage we distinguish two essentially different cases: (i) $q \neq 0$ and (ii) $q \to 0$.

Case (i)
For this case we define the fluctuation function as follows:

$$F_q(s) = \left[\frac{1}{2N_s}\sum_{\nu=1}^{2N_s}[F^2(\nu, s)]^{q/2}\right]^{1/q}. \tag{4}$$

We are interested how this q-dependent fluctuation function depends on the time scale s for diffrerent values of q hence, stages 2 till 4 must be repeated for several time scales s.

Case (ii)
For this case Eq.(4) assumes the form:

$$F_{q\to 0}(s) \to F_0(s) = \exp\left[\frac{1}{4N_s}\sum_{\nu=1}^{2N_s}\ln F^2(\nu, s)\right] \sim s^{h(q=0)}, \tag{5}$$

which can be already numerically processed.

2.1.5 Stage 5

In this stage we determine the scaling behavior of the fluctuation functions by analyzing log-log plots of $F_q(s)$ vs. s for different values of q (cf. Fig.2).

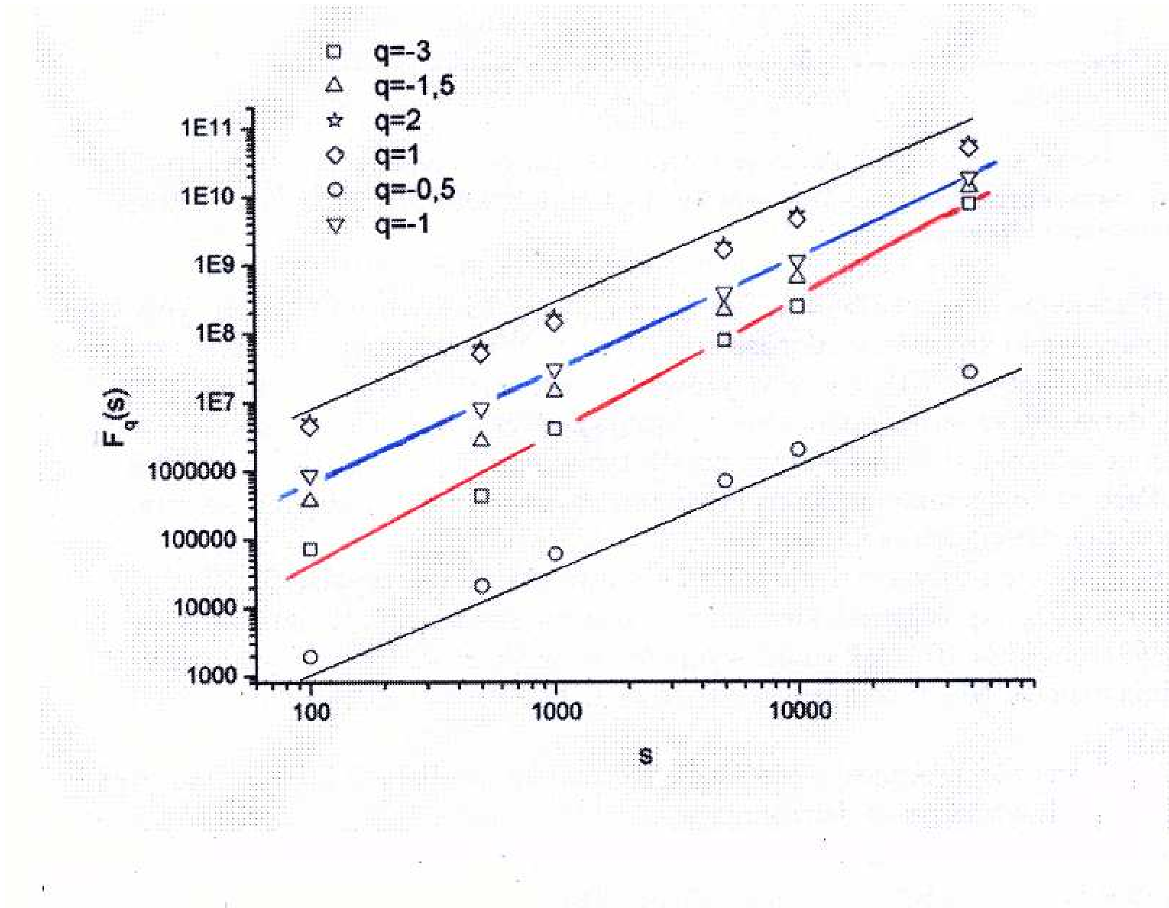

Figure 2: Fluctuation function $F_q(s)$ vs. scale s, for example, for six values of q.

Namely,

$$F_q(s) \sim s^{h(q)}, \tag{6}$$

where $h(q)$ is called generalized Hurst exponent. This exponent defines slopes of straight lines shown in Fig.2, as a function of q. In Fig.3 was presented the behavior of $h(q)$ vs. q.

For completeness, in Fig.4 was shown the dependence of the global scaling exponent

$$\tau(q) \stackrel{\text{def.}}{=} q\,h(q) - 1 \tag{7}$$

versus q while in Fig.5 the emphesized central part of Fig.4 was shown.

Note that relation (7) gives the callibration of the general exponent, since $\tau(q) = -1$, for $q = 0$.

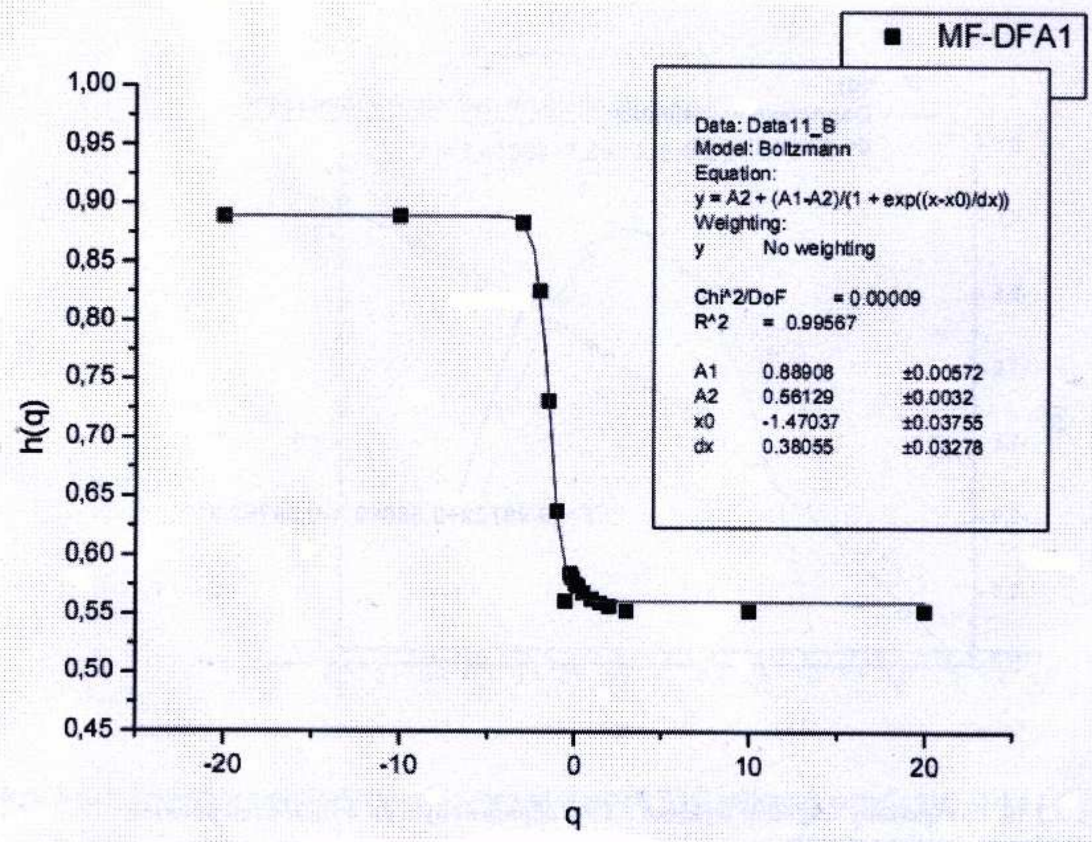

Figure 3: Generalized Hurst exponent $h(q)$ vs. order q.

2.2 The case $h(q)$ close to zero

When absolute value of exponent $\mid h(q) \mid$ is too small, but still greater than zero, then the problem arises (since, it is very inconveniet to process with so small numbers). In such cases a simple modification of the MF-DFA should be made namely, the integration of the cumulative time series $Y(k)$ before the MF-DFA procedure will start. Hence, we replace the single summation in (1) (which defines the profile from original time series $x_{k}{}_{k=1}^{N}$), by a double summation

$$\tilde{Y}(i) = \sum_{k=1}^{i} [Y(k) - \langle Y \rangle]. \tag{8}$$

This integrated profile is considered as a new input one, which is ready for the MF-DFA procedure given above (stages 1 - 5). Right now, we construct generalized fluctuation functions $\tilde{F}_q(s)$ described by scaling law, the analogous to that given in (6):

$$\tilde{F}_q(s) \sim s^{\tilde{h}(q)}, \tag{9}$$

where now exponent

$$\tilde{h}(q) = h(q) + 1 \tag{10}$$

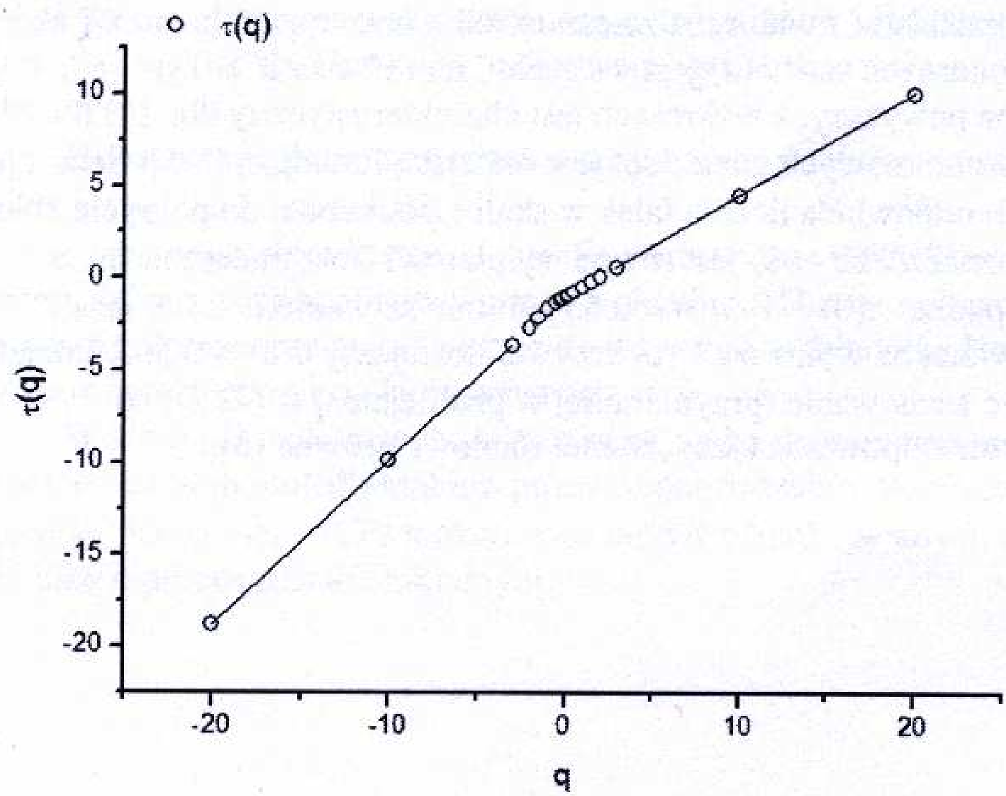

Figure 4: The global scaling exponent $\tau(q)$ vs. order q (small circles); the straight lines (solid curve) were fitted for intermediate and large values of $\mid q \mid$ (≥ 3).

significantly differs from zero, as it should be. Hence, exponent $h(q)$ can be accurately determined even for negative $h(q)$ but distinguishly greater than -1.

2.3 Technical remarks

The technical question arises namely, how large or how small the length s of the single segment should be? Definitely, if s is too small (let say order of few) then the estimation of the variances or χ^2s (cf. Eqs.(2) and (3)) for the single segment is too bad. On the opposite side, if s is too large (let say order of N) then the estimation of the fluctuation function (4) is too bad. Hence, the choice $10^2 \lesssim s \lesssim N/10^2$ seems to be the reasonable one.

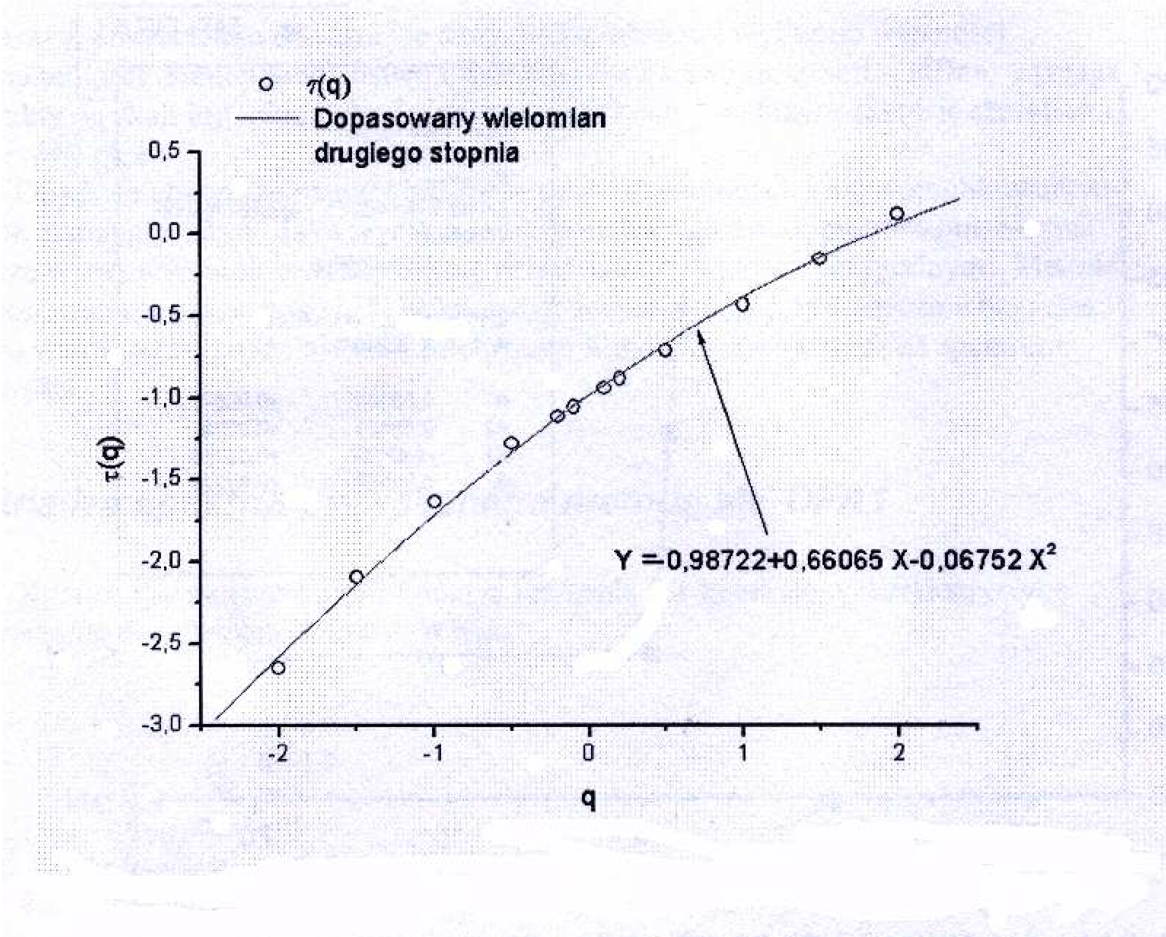

Figure 5: The emphasized central part of the global scaling exponent $\tau(q)$ vs. order q fitted, for example, by the polinomial of the second order (solid curve).

3 Relation to statistical physics

In this section we assume that our time series is a stationary one, which consists of non-negative, normalized elements $x_k,\ k = 1, \ldots, N$. For example, these could be inter-transaction time intervals Δt_k normalized by the full width of time-window $T = \sum_{k=1}^{N} \Delta t_k$, i.e., $x_k = \Delta t_k/T,\ k = 1, \ldots, N$. Next, we replace the variance given by Eq.(2) by the simplified one used in the standard fluctuation analysis (FA):

$$F_{FA}^2(\nu, s) = [Y(\nu s) - Y((\nu - 1)s)]^2. \tag{11}$$

Inserting (11) into Eq.(4) and by using Eq.(6) we obtain

$$F_q^{FA}(s) = \{\frac{1}{2N_s} \sum_{\nu=1}^{2N_s} \left[F_{FA}^2(\nu, s)\right]^{q/2}\}^{1/q} = \{\frac{1}{2N_s} \sum_{\nu=1}^{2N_s} \mid Y(\nu s) - Y((\nu - 1)s) \mid^q\}^{1/q} \sim s^{h(q)}. \tag{12}$$

For simplicity we can assume that s and N are commencurate (i.e., $N_s = N/s$). Hence Eq.(12) transforms into the form

$$\sum_{\nu=1}^{2N_s} | Y(\nu s) - Y((\nu - 1)s) |^q \sim s^{\tau(q)}, \tag{13}$$

where exponent

$$\tau(q) = q\, h(q) - 1. \tag{14}$$

Right now, it is straightforward to find scaling law for the partition function $Z_q(s)$. At first, we can introduce the box probability $p_s(\nu)$:

$$p_s(\nu) = \sum_{k=(\nu-1)s+1}^{\nu s} x_k = Y(\nu s) - Y((\nu - 1)s). \tag{15}$$

Secondly, the partition function can be defined by using this box probability as follows

$$Z_q(s) = \sum_{\nu=1}^{N_s} | p_s(\nu) |^q \sim s^{\tau(q)}. \tag{16}$$

Indeed, this expression relates statistical physics to multifractality in particular case of the stationary tim series. However, the analogous derivation for the nonstationary one is still an open problem.

4 Legendre transformation: spectrum of singularities

Since the Legendre (or contact) transformation can now be well defined, we can introduce the spectrum (distribution) of local dimensions (Hölder exponents or singularities) $f(\eta)$ from the global scaling exponent $\tau(q)$. Namely,

$$\begin{aligned} \eta &= \frac{d\tau(q)}{dq}, \\ f(\eta) &= q\eta - \tau(q) \end{aligned} \tag{17}$$

or equivalently (by using (7))

$$\begin{aligned} \eta &= h(q) + q\,\frac{dh(q)}{dq}, \\ f(\eta) &= q\,[\eta - h(q)] + 1 = q^2\,\frac{dh(q)}{dq} + 1. \end{aligned} \tag{18}$$

Note, to plot $f(\eta)$ versus η the first relation in (18) should be inverted (analytically, numerically or geometrically) to obtain q as a function of η. In other words, we use both relations (18) to determine $f(\eta)$ versus η (cf. Fig.6). Indeed, the derivative $\frac{dh(q)}{dq}$ in both expressions in (18) (which we obtain from empirical data) is mainly responsible for large errors in determination of spectrum f.

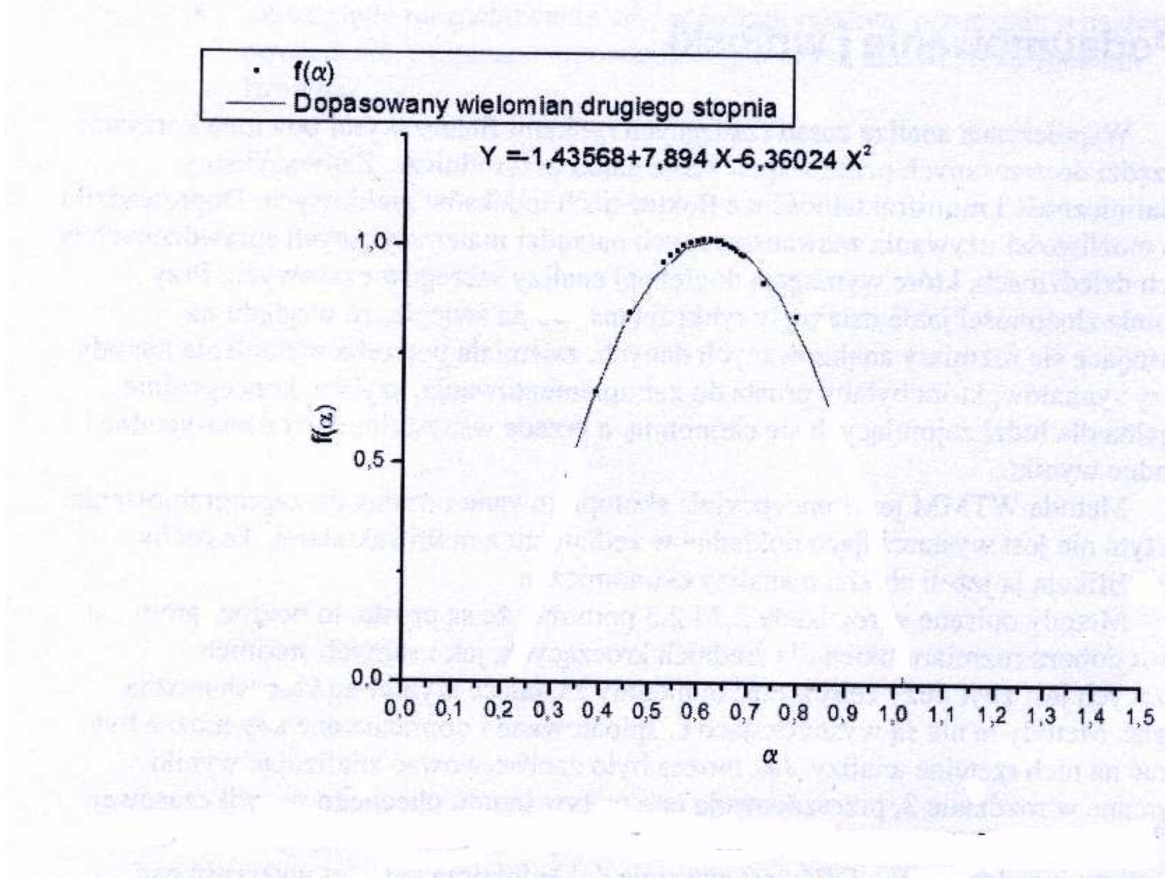

Figure 6: The spectrum of local dimensions $f(\eta)$ vs. η; the empirical curve (dots) is fitted, for example, by the polynomial of the second order.

5 Characteristic examples

5.1 Power-law probability distribution

It is instructive to mention about the particular, extreme case, i.e., drawing of the series $\{x_k\}_{k=1}^{N}$ from the power-law probability distribution function

$$P(x) = \alpha\, x^{-(\alpha+1)},\ \alpha > 0,\ 1 \le x < \infty. \tag{19}$$

It can be estimated that in this case, for $1 \ll s \ll N$, one has

$$F_q(s) \sim \begin{cases} s^{1/q} & \text{for } q > \alpha \\ s^{1/\alpha} & \text{for } q \le \alpha. \end{cases} \tag{20}$$

Hence,

$$h(q) \approx \begin{cases} 1/q & \text{for } q > \alpha \\ 1/\alpha & \text{for } q \le \alpha, \end{cases} \tag{21}$$

and

$$\tau(q) = q\, h(q) - 1 \approx \begin{cases} 0 & \text{for } q > \alpha \\ q/\alpha - 1 & \text{for } q \le \alpha, \end{cases} \tag{22}$$

$$\eta = \frac{d\tau(q)}{dq} \approx \begin{cases} 0 & \text{for } q > \alpha \\ 1/\alpha & \text{for } q \leq \alpha; \end{cases} \tag{23}$$

finally for spectrum of singularities we obtain

$$f(\eta) = q\eta - \tau(q) \approx \begin{cases} 0 & \text{for } \eta = 0 \equiv q > \alpha \\ 1 & \text{for } \eta = 1/\alpha \equiv q \leq \alpha, \end{cases} \tag{24}$$

As it is seen, we obtained a two-point spectrum of singularities. The first point is trivial since, it is independent on α while the location of the second point directly depends on α makeing it possible to determine this exponent from empirics; the later is also discussed in section below.

5.2 Correlations

It is interesting to verify by the MF-DFA whether in our case the elements of time series (i.e. intertransation time intervals) are uncorrelated. For this purpose it is sufficient to compare spectrum of singularities obtained from the original time series with the corresponding one which was constructed by random shuffled the elements of this series. Since we expact that in our case the multifractality wasn't caused by correlations, this shuffling shouldn't change anything (otherwise, we could discover the correlations between elements of our time series).

6 Key open problem

It is an open problem to prove whether the general exponent $\tau(q)$ defined here within the MF-DFA is the same as the corresponding one introduced by our MF-CTRW formalism for $q > 0$. Since the pure theoretical answer seems to be at the moment too ambitious challenge, I suppose that correct fit of the latter to the former one obtained from empirical data, would be a significant step in the proper direction. Unfortunately, our time series concerning futures for WIG20 is to short to give satisfactory empirical dependence $h(q)$ vs. q, necessary to derive spectrum of singularities with sufficiently low dispersion.

References

[1] J.W. Kantelhardt, St. A. Zschiegner, E. Koscielny-Bunde, Sh. Havlin, A. Bunde, H.E. Stanley: *Multifractal detrended fluctuation analysis of nonstationary time series*, Physica A 316 (2002) 87-114.

[2] J.W. Kantelhardt, E. Koscielny-Bunde, H.H.A. Rego, Sh. Havlin, A. Bunde: *Detecting long-range correlations with detrended fluctuation analysis*, Physica A 295 (2001) 441-454.

[3] M. Lech: *Elements of Multifractal Analysis of WIG*, Bachelor Thesis, supervision R. Kutner, Warsaw 2007.