Physica D 31 (1988) 277–283
North-Holland, Amsterdam

APPROACH TO AN IRREGULAR TIME SERIES ON THE BASIS OF THE FRACTAL THEORY

T. HIGUCHI
Geophysics Research Laboratory, University of Tokyo, Tokyo 113, Japan

Received 18 June 1987
Revised manuscript received 9 January 1988
Communicated by M. Mimura

We present a technique to measure the fractal dimension of the set of points $(t, f(t))$ forming the graph of a function f defined on the unit interval. First we apply it to a fractional Brownian function [1] which has a property of self-similarity for all scales, and we can get the stable and precise fractal dimension. This technique is also applied to the observational data of natural phenomena. It does not show self-similarity all over the scale but has a different self-similarity across the characteristic time scale. The present method gives us a stable characteristic time scale as well as the fractal dimension.

1. Introduction

The power spectrum analysis has been conventionally used as a useful and efficient method for analyzing an irregular time series till now. Especially when the power spectrum follows the power law; $P(f) \propto f^{-\alpha}$, the exponent α is considered to be the index for representing the irregularity of a time series. In the actual analysis of the experimental and observational data, the power spectrum obtained by the fast Fourier transform method (FFT) shows noisy fluctuations superposed on the power law spectrum. In order to get a stable power law index, we have to take an ensemble average of the power spectra over a long interval in which the fluctuations are assumed to be statistically stationary. However, since the statistical characteristics of fluctuations often vary for a short time interval, it is not appropriate to take an average of the power spectra over a long interval. Recently studies have been done about quantitative investigation of a time series which shows non-periodic and turbulent behavior, since new ideas for describing an irregular time series have been developed [1–7]. The fractal dimension is introduced as the index for describing the irregularity of a time series in place of the power law index [1].

Burlaga and Klein [7] have presented a method to calculate the length of a curve and have obtained stable values of the fractal dimension of large-scale fluctuations of the interplanetary magnetic field. In this paper we modify their method to calculate the fractal dimension and apply it not only to the simulated data to show that by using our method we can get a stable value of the fractal dimension, but also to the time series of the natural phenomena which show the turbulent behavior.

A fractal curve has the property that each part can be considered a reduced scale image of the whole [1]. The simulated data in this paper has the property of a fractal curve over all time scales. In other words, the power spectrum follows a single power law over all ranges of frequency. In contrast, a Kolmogoroff power spectra is valid only within the inertial range [8]. When a time series changes its structure in time domain across a certain characteristic frequency, it is difficult to determine power law indices and a characteristic time scale from the power spectrum. The technique developed in this paper can easily give us stable indices and time scale corresponding to the characteristic frequency even for a small number of data.

2. Method for calculating the fractal dimension

We now consider a finite set of time series observations taken at a regular interval:

$$X(1),\ X(2),\ X(3), \ldots,\ X(N).$$

From given time series, we first construct a new time series, X_k^m, defined as follows:

$$X_k^m;\ X(m),\ X(m+k),\ X(m+2k), \ldots,\ X\left(m + \left[\frac{N-m}{k}\right] \cdot k\right) \quad (m = 1,2,\ldots,k),$$

where [ ] denotes the Gauss' notation and both k and m are integers. m and k indicate the initial time and the interval time, respectively. For a time interval equal to k, we get k sets of new time series. In the case of $k = 3$ and $N = 100$, three time series obtained by the above process are described as follows:

$$X_3^1;\ X(1),\ X(4),\ X(7), \ldots,\ X(97),\ X(100),$$
$$X_3^2;\ X(2),\ X(5),\ X(8), \ldots,\ X(98),$$
$$X_3^3;\ X(3),\ X(6),\ X(9), \ldots,\ X(99).$$

We define the length of the curve, X_k^m, as follows:

$$L_m(k) = \left\{ \left(\sum_{i=1}^{\left[\frac{N-m}{k}\right]} |X(m+ik) - X(m+(i-1)\cdot k)| \right) \frac{N-1}{\left[\frac{N-m}{k}\right] \cdot k} \right\} \Big/ k.$$

The term, $N - 1/[(N-m)/k] \cdot k$ represents the normalization factor for the curve length of subset time series. We define the length of the curve for the time interval k, $\langle L(k) \rangle$, as the average value over k sets of $L_m(k)$. If $\langle L(k) \rangle \propto k^{-D}$, then the curve is fractal with the dimension D.

To test that our method for determining a fractal dimension is valid, we show numerical application to simulated data. First, we apply our technique to the simulated data $Y(i)$ $(i = 1,2,\ldots,N)$ with the fractal dimension D of 1.5. $Y(i)$ are generated by

$$Y(i) = \sum_{j=1}^{1000+i} Z(j),$$

where $Z(j)$ is a Gaussian noise with the mean zero and a standard deviation of 1. The value of 1000 is arbitrarily chosen to eliminate the effect of sampling $Z(1)$ on $Y(1)$. We use the following values of the interval time k; $k = 1,2,3,4$ and, $k = [2^{(j-1)/4)}]$ $(j = 11,12,13,\ldots)$ for k larger than 4, where [ ] denotes Gauss' notation. The function $\langle L(k) \rangle$ is described as $\langle L(k) \rangle \propto k^{-D}$ for a statistically self-similar curve [1]. Then, if $\langle L(k) \rangle$ is plotted against k on a doubly logarithmic scale, the data should fall on a straight line with a slope $-D$. The logarithm of the length, $\log \langle L(k) \rangle$, for a time series $Y(i)$ with $N = 2^{17}$, is plotted as a function of $\log k$ in fig. 1. The value of the vertical axis is multiplied by an arbitrary factor. The error bar denotes the standard deviation of $\log(L(k))$. In this calculation, the maximum value of k, $k_{\max}$, is 2^{11}. The straight line is fitted according to the least-square procedure. For smooth rectifiable

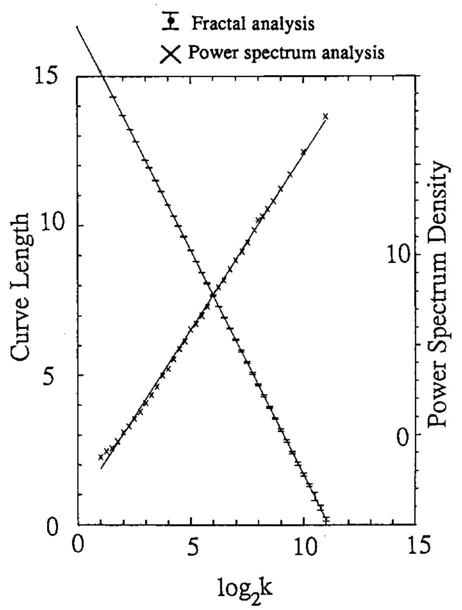

Fig. 1. The logarithm of the curve length $\log\langle L(k)\rangle$ for the time series $Y(i)$ is given as a function $\log k$. The straight line is fitted to the points by the least-square method. The slope represents minus fractal dimension ($D = 1.500 \pm 0.0004$). The error bar in the fractal analysis represents the standard deviation of $\log(L(k))$. The power spectrum obtained by using FFT is denoted by crosses. We take an average of 1024 power spectra with 256 data points by shifting a time window by 128 points. From the relation of the power law index to the fractal dimension, we obtain the fractal dimension of 1.503 ± 0.0067.

Table I
Results of application to simulated data ($k_{max} = 2^{11}$).

D	$N = 2^{15}$	2^{16}	2^{17}
Theoretical value	1.5		
Our method	1.513 ± 0.0007	1.508 ± 0.0011	1.500 ± 0.0004
BK's method	1.528 ± 0.0027	1.522 ± 0.0025	1.511 ± 0.0028
FFT	1.520 ± 0.0120	1.503 ± 0.0097	1.503 ± 0.0067

curves, the value of D is equal to the topological dimension. Since in this case statistically self-similar curves are embedded in a plane, D is a fraction, $1 < D < 2$. In the case of $Y(i)$, a fractal dimension is theoretically equal to 1.5 [1]. Table I shows the results of application of present procedure to time series, $Y(i)$, for $N = 2^{15}$, $N = 2^{16}$ and $N = 2^{17}$. A stable value of the fractal dimension D can be obtained by our method.

3. Comparison with another method

Burlaga and Klein [7] have presented another technique for calculating the length of the curve and estimated the index H characterizing the fractional Brownian function [1]. The index H is related to the fractal dimension D by the relation equation specified as $D = H + 1$. Although their method is superior to the conventional FFT method with regard to the efficiency of calculating the power law index, their method is based on the inappropriate definition of the curve length to obtain the fractal dimension. Before calculating the curve length, they first took an average of $B(t)$ (corresponding to $X(i)$ in this paper) between $t = t_i$ to $t = t_i + \tau$ (τ corresponds to the interval time k). The difference between our technique and Burlaga and Klein's method (BK's method) lies in their operation of averaging the time series prior to calculating the curve length. We consider the case of $k = 3$ in order to clearly show the difference. The curve length defined by BK's method can be rewritten

$$L_{\mathrm{BK}}(3) = \left\{ \left| \frac{X(4) + X(5) + X(6)}{3} - \frac{X(1) + X(2) + X(3)}{3} \right| + \left| \frac{X(7) + X(8) + X(9)}{3} - \frac{X(4) + X(5) + X(6)}{3} \right| + \left| \frac{X(10) + X(11) + X(12)}{3} - \frac{X(7) + X(8) + X(9)}{3} \right| + \cdots \right\} \Big/ 3.$$

The curve length $\langle L(k) \rangle$ defined in this paper is given by

$$\langle L(3) \rangle = \left\{ \frac{|X(4) - X(1)| + |X(5) - X(2)| + |X(6) - X(3)|}{3} + \frac{|X(7) - X(4)| + |X(8) - X(5)| + |X(9) - X(6)|}{3} + \frac{|X(10) - X(7)| + |X(11) - X(8)| + |X(12) - X(9)|}{3} + \cdots \right\} \Big/ 3.$$

As the interval time k becomes large, the curve length defined by BK's method tends to be short, and thus the fractal dimension D obtained by BK's method tends to be larger than the real value. Especially for the case of the time series with the fractal dimension of nearly 2, the length of the curve calculated by BK's method is significantly shorter than the real length and BK's method gives us a wrong fractal dimension. Although the fractal dimension in this study cannot be larger than 2, the fractal dimension of the Gaussian noise calculated by BK's method is around 2.5 and is obviously wrong. Since our technique directly represents the definition of an index H [1], we can estimate not only the index H but also the fractal dimension D more accurately than those given by BK's method.

By using BK's method, we calculate the curve length of the same data set as were previously examined in fig. 1. We specify the curve length obtained by BK's method as $L(k)_{\mathrm{BK}}$. The logarithm of the ratio of the curve length, $L(k)_{\mathrm{BK}} / \langle L(k) \rangle$, as a function of $\log k$ is shown in fig. 2. It is clear that the curve lengths calculated by BK's method are always shorter than those given by our method. Using BK's method, we calculate the curve length of the same data set $Y(i)$, and compare the fractal dimension with that given by our method in table I. The technique presented in this paper gives more precise and stable fractal dimensions than those by BK's method.

The power spectrum analysis has been used to analyze the irregular and non-periodic time series. A curve with a single power law spectrum is self-similar and the index of a power law spectrum α is simply related to the fractal dimension D by the relation $D = (5 - \alpha)/2$ [9]. The data examined in fig. 1 ($N = 2^{17}$) are applied to the power spectrum analysis by FFT. A time window with 256 data points is shifted by 128 points in order to take an ensemble average of the power spectra. Then we obtain 1024 samples of the power spectrum. The average power spectrum denoted by crosses is superposed on fig. 1. The slope of the straight line fitted to $\log k$ versus the logarithm of the power spectrum gives us the power law index. From the relation of the power law index to the fractal dimension, we obtain a fractal dimension of

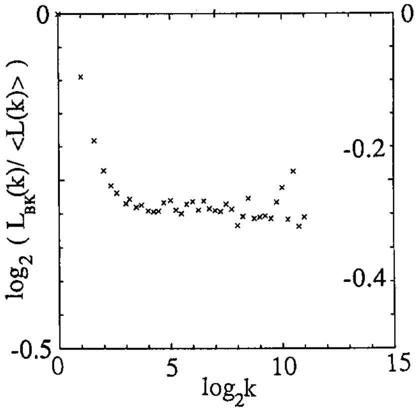

Fig. 2. The quantity $\log(\langle L(k)\rangle/L(k)_{BK})$ as a function of $\log k$ is shown in this figure. It is clear that the curve length calculated by BK's method is always shorter than by our method.

$D = 1.503 \pm 0.0067$. The standard deviation of the fractal dimension obtained by FFT is larger than that obtained by the fractal analysis developed in this paper. Since the FFT method intrinsically requires the ensemble average of the power spectrum to get a stable power spectrum, it takes numerous operations of taking the ensemble average of the power spectrum in order to get as precise and stable values as those given by the present technique. This demerit of the FFT method tends to be significant for a lot of data points. The fractal analysis is an efficient and economical alternative to the spectral analysis for examining the irregular and self-similar time series.

4. Application to natural phenomena

The time series used in the above application is self-similar over all time scales. In cases of natural phenomena, statistical properties of time series often depend on the time scale of measurement. In short, the index D in $\langle L(k)\rangle \propto k^{-D}$ may not be constant over all time scales. It is probable that the value of D changes continuously in some cases and discontinuously at a critical time scale in other cases.

Although the fractal concept is based on the idea that the set of interest is self-similar all over the scale, the fractal dimension can be defined within the range in which the property of self-similarity holds [1]. For example, a *Kolmogoroff* power spectrum is valid within the inertial range [8].

It was shown that the earth's magnetosheath is the turbulent region between the shock surface and earth's magnetopause and that the magnetic field data observed there behaves irregular and their power spectrum follows the power law $P(f) \propto f^{-\alpha}$. Fairfield [10] reported that the power law index α at frequency ranges below and above the proton gyro-frequency are roughly -1 and -3, respectively. Since the exponent α is closely related to the fractal dimension [1, 8], it is considered that the magnetic field has the characteristic time scale around the period of the proton gyration.

We examine the magnetic field data obtained by the satellite ISEE1 [11], by the same procedure as used for simulated data. The fractal analysis is carried out for each of the components of the magnetic field. We demonstrate in fig. 3a the result applied to the component which is in the plane tangent to the shock surface. The number of data used for calculating the fractal dimension is 1200. The interval time for the magnetic field data, τ, is related to k used in application to the simulated data by $\tau = 0.25 \times k$, where 0.25 denotes the sampling time interval. The maximum value of the interval time, τ_{max}, is 64 s. Fig. 3a shows an

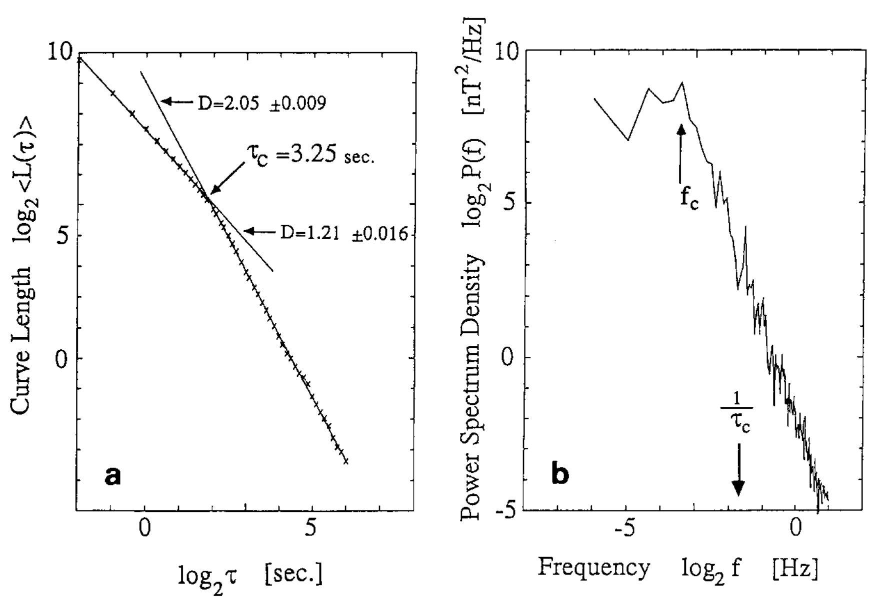

Fig. 3a. The length $\langle L(k) \rangle$ for a magnetic field data on a doubly logarithmic scale as in fig. 1. The curve breaks at the value of $\tau = 3.75$ s. In the range above time scale of larger than $\tau = 3.75$, the curve is linear with a slope equal to -2.05. In the lower range a slope of -1.21 is seen. Fig. 3b shows the power spectrum of the magnetic field data examined in fig. 3a. The power spectrum is calculated by FFT. The critical frequency, f_c, appears at about $\frac{1}{5}$ of $1/\tau_c$.

example of the length of the curve plotted in a doubly logarithmic scale in the same format as in fig. 1. We see that the graph breaks at $\tau = 3.25$. The period of the proton gyration is 0.86 s then. Above the time scale τ_c characterizing the large scale, the curve is linear with $D = 2.05 \pm 0.009$. In the range $\leq \tau_c$ the slope of the curve is equal to 1.21 ± 0.016. The critical time scale τ_c for another data set observed in the magnetosheath is also about several times as long as the proton gyro-period.

Fig. 3b shows the power spectrum of the same magnetic field data as examined in fig. 3a. The power spectrum is calculated by FFT. The power spectrum bends at about 0.0625 Hz corresponding to the period of 16 s, and is nearly flat below this critical frequency, f_c. The power spectrum above f_c shows a power low form, $P(f) \propto f^{-\alpha}$, with $\alpha = 2.78$. The fractal dimension, $D = (5 - \alpha)/2$ for $1 < \alpha < 3$, is about 1.11, and roughly agrees with 1.2 obtained by the present fractal analysis for the time scale below τ_c. When $0 \leq \alpha < 1$, the power spectrum implies that the curve has the fractal dimension $D = 2$ [1]. The flat power spectrum below f_c with $\alpha = 0$ is in accordance with $D = 2$ for $\tau > \tau_c$ obtained by the fractal analysis. It should be noted that the critical time scale τ_c in the fractal analysis corresponds to $\frac{1}{4} - \frac{1}{2}$ wavelength of the critical period, $1/f_c$, in the power spectrum. Thus the agreement between the two methods is satisfactory. In the power spectrum method, however, the low frequency end is intrinsically unstable because of the limited sample length. Without the fractal analysis given in fig. 3a, we could not say from the spectrum result in fig. 3b that the time series is self-similar below f_c. In order to obtain the reliable information covering the low-frequency side and obtain the critical time scale, we have to get many samples of the power spectrum by analyzing the time series for a longer interval. Since the magnetic field observed at the magnetosheath drastically changes its characteristics during a short interval, we cannot assume the stationarity of the magnetic field data for a long time interval. Then it is necessary to quantitatively examine characteristics of the time series for a short time. As compared with the power spectrum analysis, the fractal analysis can give us the stable critical time scale and power indices describing the characteristics of the time series.

5. Concluding remarks

The purpose of this paper is to introduce a method for measuring the fractal dimension and calculating a more precise and stable characteristic time scale than those presented before. We hope that this technique is useful for analyzing non-periodic and irregular time series.

Acknowledgements

I thank Prof. S. Kokubun for his heartful encouragement and evaluation of this paper. I am grateful to Prof. C.T. Russell for providing the time series of the magnetic field data on board ISEE1 satellite.

References

[1] B. Mandelbrot, Fractals: Form, Chance and Dimension (Freeman, San Francisco, 1977).
[2] B. Mandelbrot and J.R. Wallis. Some long-run properties of geophysical records, Water Resource Research 5 (1969) 321.
[3] D. Rulle, Five turbulent problems, Physica D 7 (1983) 40.
[4] P. Grassberger and I. Procaccia, Measuring the strangeness of strange attractors, Physica D 9 (1983) 189.
[5] K.J. Falconer, The Geometry of fractal sets (Cambridge Univ. Press, Cambridge, 1985).
[6] J.D. Farmar, E.T. Otto and J.A. Yorke, The dimension of chaotic attractors, Physica D 7 (1983) 153.
[7] L.F. Burlaga and L.W. Klein, Fractal structure of the interplanetary magnetic field, J. Geophys. Res. 91 (1986) 347.
[8] A.N. Kolmogoroff, C.R. Acad. Sci. USSR 30 (1941) 301.
[9] M.V. Berry, Diffractals, J. Phys. A: Math. Gen. 12 (1979) 781.
[10] D.H. Fairfield, Magnetic fields of the magnetosheath, Rev. Geophys. and Space Phys. 14 (1976) 117.
[11] C.T. Russell and R.C. Elphic, Initial ISEE magnetometer results: magnetopause observations, Space Sci. Rev. 22 (1978) 681.