ECONOMIC RESEARCH REPORTS

AN ANALYSIS OF U.S. STOCK PRICE BEHAVIOR USING WAVELETS

BY

James B. Ramsey,
Daniel Usikov,
and
George M. Zaslavsky

RR # 94-06 *February 1994*

C. V. STARR CENTER FOR APPLIED ECONOMICS

NEW YORK UNIVERSITY
FACULTY OF ARTS AND SCIENCE
DEPARTMENT OF ECONOMICS
WASHINGTON SQUARE
NEW YORK, N.Y. 10003

AN ANALYSIS OF U.S. STOCK PRICE BEHAVIOR USING WAVELETS

James B. Ramsey

Dept. of Economics

New York Univ.

Daniel Usikov

Univ. of Maryland

George M. Zaslavsky

Courant Institute of Mathematical Sciences

and Physics Dept.

New York University

The financial and institutional support of the C.V. Starr Center for Applied Economics is gratefully acknowledged.

AN ANALYSIS OF U.S. STOCK PRICE BEHAVIOR USING WAVELETS

ABSTRACT

Using wavelets we re-examine the U.S. stock market price index for any evidence of self-similarity or order that might be revealed at different scales. The wavelet transform localized in time can be used to indicate how the power of the projection of the signal onto the kernel varies with the scale of observation. By comparing how the local power scales vary over time much information about the structure of the data can be obtained. Such evidence is not at all evident from standard analyses of untransformed data, including projections onto a Fourier basis. Wavelets can detect structures in data that are highly localized in time and therefore non-detectable by Fourier transforms.

The main conclusion is that while the data are clearly complex, there seems to be some evidence of non-randomness in the data. There is also some limited evidence of quasi-periodicity in the occurrence of large amplitude shocks to the system.

AN ANALYSIS OF U.S. STOCK PRICE BEHAVIOR USING WAVELETS

Introduction

Economic and financial data make up for their relative scarcity and the lack of control that economists traditionally have had over their construction by the intensity with which they are analyzed. Over the past few decades, there has been a dramatic reversal in opinion concerning the degree of information contained in the data and the beginning of a new approach to the analysis of economic and financial data.

The early post war period until the eighties was characterized by the overriding belief that macro-economic data contained a lot of structural information that could be captured in the estimation of systems of simultaneous equation models with a limited number of external forcing terms driving the economic system. The models were based on concepts of static equilibria that were perturbed to new equilibria by external forces; the role of transients in the adjustment process tended to be ignored. Consequently, the economists' job was to estimate the appropriate parameters relating the steady state of the system to potentially controllable forces to ensure an economic dynamical path that was optimal, at least approximately.

In contrast, financial data, with the exception of interest rates, were regarded as unforecastable; current prices reflect all current information on values so that no gains from trade could be made with publicly available information; this was the famous "efficient markets" hypothesis.

Both ideas were modified during the eighties. While the "efficient markets hypothesis" still prevails in more sophisticated form, there is some disputed, but growing, evidence that the equity stock market is not entirely efficient and that very small, but noticeable, profitable deviations exist at least for a time. The opinion with economic data went the other way. The presence of structure, at least that which can be captured in terms of systems of simultaneous equation models driven by forcing terms, was found to be in large part illusory. This was especially true, when attention was focused on growth rates, or more simply on the variation in first differences.

The "chaos revolution" of the last fifteen years has had a dramatic impact on the modelling and understanding of many scientific disciplines and engineering, as is illustrated in a recent volume of papers on the applications of chaotic phenomena,[1]. Of the many phenomenological implications of nonlinear dynamics, the concepts of self-similarity and fractal dimension have received an enormous amount of attention. The dimension of a wide array of mathematical models and of the realizations of actual experiments have been calculated by a variety of methods. Of these, the Grassberger-Procaccia method has been most popular; see for example, [2-4]. Since these original articles, there have been a number of refinements and extensions of the basic techniques, see for example, [5-7] and in general for an excellent review of current thinking about the discovery of nonlinear systems from time series, see volume 58 of *Physica D*, that is dedicated to the "Interpretation of Time Series from Nonlinear Systems".

Unfortunately, these procedures have not been of much use with economic and financial data; see [8-12]. The overall conclusion with financial data using the Grassberger-Procaccia approach to dimension calculation is that there seems to be little evidence for an attractor in these data; this is especially true with returns in the stock market. The conventional wisdom that stock market returns are an example of geometric Brownian motion is supported by these procedures. However, there are other papers in the financial literature that indicate that at various time scales there are aspects of the data that indicate observable differences with the geometric Brownian motion hypothesis [13-18]. Early use of spectral techniques in the analysis of economic and financial data discovered the "typical spectral shape"; high power at very low frequencies and rapidly declining thereafter, a shape that is consonant with geometric Brownian motion [19,20]. Yet other work that has relied on nonlinear techniques of analysis has discovered some evidences of nonlinearity in the data[21,22].

A major difficulty with most of the previous analysis in this area is that stationarity of the time series has been assumed and the possible presence of highly time localized patterns in the structure of the data has been ignored. Another major difficulty is that dimension calculations using Grassberger-Procaccia procedures, for example, require very large numbers of observations that are simply not available for economic and even for financial data. Consequently, there is a need for alternative ways to examine and evaluate the data that do not depend on the assumption of stationarity, that can detect singularities, and do not require very large numbers of observations. Wavelet analysis can provide a particular answer [23,24].

The research reported in this paper is a first, preliminary step and is both experimental and exploratory in nature. The objective is to examine stock market prices in the form of the Standard and Poor's index of market prices on the New York Stock Exchange for evidence of self-similarity and ordered patterns at some scales by using wavelets. In so far as this first attempt reveals potentially interesting structure in the data, the project will have been successful if it stimulates further research of these data in greater depth.

We wish to emphasize for the reader the following points.

1) While we consider only one series, the stock market index, this data set is well known, reliable, and substantial in quantity, at least by economists' standards.

2) The specific choice of kernel for the wavelet transform is not important, as we argue in the next section. What is important is whether using any kernel some order can be found. Later one can investigate the choice of kernel to find one that is more appropriate to the circumstances at hand.

3) For the impatient reader, the result of the exercise is that without having to specify the economic mechanisms that are in operation in this market, we have been able to detect the presence of a certain amount of pattern in the data as revealed by wavelet analysis.

Self-similarity and Wavelets

The key idea of self-similarity is that the dynamical system has no characteristic scale; the dynamical description of the system is the same at all time and coordinate scales. A corresponding concept is "statistical self-similarity". Statistical self-similarity is the idea that a suitably renormalized sum of random variables has the same distribution as that of the original random variable. In the statistical literature this property is referred to as the infinite divisibility of the probability distribution; the Gaussian distribution is the most common example.

Brownian motion is an example of a self-affine process; self-affine processes differ from self-similar processes only in that for self-affine processes different coordinates scale at different rates. Brownian motion, the supposed model for growth rates in stock prices, is self-affine because if we rescale the time scale by a factor "b" and we rescale the length scale by "$b^{1/2}$" we reproduce the original distribution, Feder[25]. In so far as some scholars claim that stock prices follow a geometric Brownian motion, the case for an examination of the data for evidences of self-similarity or self-affinity has been made.

The wavelet transform provides a useful approach to the discovery of self-similarity, or self-affinity, in data, [23,24]. If {x(t)} denotes a sequence of observations on a signal that is represented as a continuous function of time, the two parameter wavelet transform $T_g(a,u)$ is an integral transform of the signal with respect to a kernel function g(.) that is localized at the point "u" and for which the scale is "a". More precisely, we define:

$$T_g(a,u) = a^{-1/2}\int x(t)g(\frac{x(t)-u}{a})dt; \qquad a>0 \tag{1}$$

The kernel, or analyzing wavelet, g(a,u) satisfies the condition:

$$\int g(y)dy = 0. \tag{2}$$

For a given value in the domain of the signal, u_0, the amplitude of the transform $T_g(a,u_0)$ is maximized when the scale "a" is of the same order as the characteristic scale of the signal x(t) in the neighborhood of u_0. This is an important tool for the visualization of self-similar, or self-affine, properties of fractal and multifractal processes [26]. Further, because the procedure enables one to examine scale variations locally, that is, in the neighborhood of u_0, the usual difficulties associated with nonstationarity of the data are not a problem. Another advantage of the wavelet transform is that the analysis can be applied without knowledge of the underlying mechanism that is generating the data. More importantly, the analysis can be usefully applied if either the data are statistically self-similar or if the sequence of observations is generated by a self-similar function[27]. Holschneider[28] and Argoul[29] have shown that if a function f(x) has the property:

$$|f(x_0+\lambda x) - f(x_0)| \sim \lambda^{\alpha(x_0)} |f(x_0+x) - f(x_0)| \tag{3}$$

then its wavelet transform will scale like:

$$T_g(\lambda a,\ x+\lambda b) \sim \lambda^{\alpha(x_0)} T_g(a,\ x+b) \tag{4}$$

In order to obtain useful estimates of the scaling parameter "a" with reasonable error bars at conventional confidence levels, some averaging of the wavelet transform is needed. Given the presumed power law type of scaling indicated above it makes sense as a first attempt to average the logarithm of the wavelet transform. If x_0 is the point at which the wavelet transform is being calculated, let x_m denote a small interval, then equation(5) defines the averaged wavelet transform.

$$<\log|T_g(a,x_0)|> = \frac{1}{x_m} \int_{x_0-\frac{x_m}{2}}^{x_0+\frac{x_m}{2}} \log|T_g(a,s)|\,ds \tag{5}$$

The same program that was applied to the stock market data and is to be presented below was also applied to the wind tunnel turbulence data discussed in [30]. A review of the turbulence results is instructive as a point of comparison for the results to be discussed in the next section.

The Empirical Results

The raw data consist of an average of daily closing prices of a sample of five hundred stocks that are traded on the New York Stock Exchange. The first observation is January, 3, 1928 and the last is June, 25, 1990. The actual data analyzed were the "growth rates", or "rates of return" that were obtained from the raw data by taking the relative first difference; that is, the variable of interest is:

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{6}$$

where P_t is the observed average daily price for the sampled five hundred stocks. These data are shown in Figures 1 and 2; the former showing the raw price index and the latter the growth rates. Some key dates are highlighted for the interest of the reader.

As explained in the previous section, we may first evaluate the scaling relationships of the data, by examining the color plots of the wavelet transform. Time is recorded on the horizontal axis and the scale of the wavelet transform is recorded on the vertical axis; the scaling range is such that each pixel represents one observation at the top of the graph to each pixel represents 256 observations at the bottom. The length of the horizontal axis is 640 pixels, or points of the discrete time series.

The function g(a,u) used for the transform kernel is:

$$g(1,\,0) = \begin{Bmatrix} 0, & |x| \geq 2 \\ -1/2, & 1 \leq |x| < 2 \\ +1, & |x| < 1 \end{Bmatrix} \tag{7}$$

As stated in the Introduction, a specific choice for the kernel cannot be made because the economists' theoretical knowledge is insufficient for that purpose, other than to rely on the concept of Brownian motion. The kernel defined in equation (7) can be regarded as representing a second difference function. As a first step and in the absence of information specifying a different kernel that in equation (7) has been found by experience to provide a useful general purpose kernel for the purposes at hand.

A subset of the wavelet transforms are presented in Figures 3 to 5 where the zoom level varies from 2^5 to 2^0. Figure 3a shows the central portion of the data, Figure 3b shows the right hand side of Figure 3a scaled up by a factor of 2. The remaining Figures show in turn the next higher factor scaling of the central half of the graph at the previous lower scale level; simply stated, each subsequent graph shows a magnification of the central half of the previous graph by a factor of two.

At each wavelet scaling in each graph; that is, at each point on the vertical axis, the graph shows by color coding the power of the projection of the stock price growth rate signal onto the wavelet transform at the indicated level of scaling. Reading across the graph at a given value for the wavelet scaling, one sees how the power of the projection varies across the time domain. Reading down the graph at a given point in time, one sees how the power varies with the scaling of the wavelet.

This set of six graphs illustrate quite clearly the degree of self-similarity of the data. Compare, for example, the first and last graphs which differ substantially in the time domain involved, the last graph covers the right third of the first graph approximately, and in scaling, where the last graph is an enlargement of the first on the order of 2^5.

Figure 4a provides the clearest example of quasi-periodic structure in that the "peaks" that are visible in that graph occur with a nearly constant period. An alternative way of examining these issues is portrayed in Figure 6. In this Figure the major "peaks" are represented by a Δ. As can be observed there is a considerable degree of symmetry in the Figure over several levels of scaling. This indicates a degree of structure that would not be observed in random data.

The last pair of graphs represent estimates of the scaling relationship that was mentioned in the previous section. Each of the two graphs shows an average plot of the rate of scaling of the power of the wavelet transform expressed as a function of the value of the scale coefficient, "a"; this provides a sequence of power scaling plots for a sequence of subsets of the data.

The graph shown in Figures 7a provides a benchmark by plotting the results for 4000 observations on a set of independently and identically distributed Gaussian random variables. Random data scale given the formulation of the wavelets discussed in the previous section as -0.5, that is, the power of the wavelet in fitting the data declines at approximately the square root of the scaling coefficient. Since with independently and identically distributed data the wavelet transform essentially averages the terms in the sequence over increasingly longer ranges as a increases, the result of square root scaling is not a surprise.

Figure 7b provides the corresponding information for the stock price growth rates at a zoom level of three. The growth rate data do not appear to be much different from the simulated Gaussian data with respect to the scaling relationship. There is some indication, however, that the rate of scaling is highest for the lowest level of zoom and decreases as the zoom and consequently the degree of averaging increases. There are some other subtle differences between the stock market data scaling relationship and that for the simulated Gaussian data. However, the relative paucity of the stock market data cautions against the hasty acceptance of these last differences as phenomenologically significant.

Summary and Conclusions

This has been an initial exploratory investigation into the applicability and usefulness of wavelet analysis in the detection of elements of self-similarity, or self-affinity, and of order in the growth rates of stock price data. Such data have in the past shown remarkable consistency, at least approximately, to the hypothesis of being a random walk of some kind. Previous efforts were impeded by the assumptions of stationarity, the presumed lack of highly localized structures in the time domain, and small numbers of observations. Our procedures do not require stationarity, wavelets were designed to detect highly localized structures in the time domain, and the daily stock market data are very large by economists standards.

Wavelet analysis was used to explore empirically the issue of self-similarity and to look for patterns, possibly only at certain scales, in these data. Although the wavelet transform itself requires substantial amounts of data, the results are nevertheless encouraging for those who believe that economic and financial data do contain more information and potentially observable patterns than has so far been discovered by more traditional means.

There is some evidence for quasi-periodicity in the occurrence of some large amplitude shocks to the system. Self similarity is also evident and this result leads to the conclusion that there is a modest amount of predictability in the data. These results seem to be true notwithstanding the lack of clear evidence of a scaling relationship in the data. In short, the data are something more than the representation of geometric Brownian motion.

References

1] Kim, J. H. and Stringer,J., *Applied Chaos* edt. by Jong Hyun Kim and John Stringer, Publishers: John Wiley & Sons, Inc., New York, 1992.

2] Grassberger, P., and I. Procaccia, "Characterization of Strange Attractors," *Physical Review Letters*, 50 (1983a), pp. 346-349.

3] Grassberger, P., and I. Procaccia, "Estimation of the Kolmogorov Entropy from a Chaotic Signal," *Physical Review A*, 28 (1983b), 2591-2603.

4] Grassberger, P., and I. Procaccia, "Measuring the Strangeness of Strange Attractors," *Physica 9D*, (1983c), 189-208.

5] Provenzale, A., Smith, L.A., Vio, R. and Murante, G., 1992, "Distinguishing between low-dimensional dynamics and randomness in measured time series", *Physica D*, 58, pp.31-49.

6] Theiler, J., Eubank, S., Longin, A., Galdrikian, B., Farmer, J.D., 1992, "Testing for nonlinearity in time series; the method of surrogate data", *Physica D*, 58, pp.77-94.

7] Kostelich, E.J., 1992, "Problems in estimating dynamics from data", *Physica D*, 58, pp.138-152.

8] Hsieh, D., and Lebaron, B., 1988, "Finite Sample Properties of the BDS Statistics," mimeo, Department of Economics, University of Chicago and University of Wisconsin-Madison.

9] LeBaron, B., 1988, "Stock Return Nonlinearities: Some Initial Tests and Findings," Ph.D. Dissertation, Department of Economics, University of Chicago.

10] Scheinkman, J.A. and LeBaron,B., 1987, "Nonlinear Dynamics and

GNP Data," in W. Barnett, J. Geweke, and K. Shell, eds., *Economic Complexity, Chaos, Sunspots, Bubbles, and Nonlinearity*, Cambridge: Cambridge University Press.

11] Ramsey, James B. and Yuan, Hsiao-Jane, 1990, "The statistical properties of dimension calculations using small data sets", *Nonlinearity*, vol.3,pp 155-176.

12] Ramsey, James, B., Sayers, C.L., and Rothman, P. 1990: "The Statistical Properties of Dimension Calculations Using Small Data Sets: Some Economic Applications." *International Economic Review*, Vol. 31, No. 4, pp.991-1020.

13] Fama, Eugene, F., 1990,"Stock Returns, Expected Returns, and Real Activity," *The Journal of Finance* 45 4 1089-1108.

14] Fama, Eugene, F. and Kenneth R. French, 1988, "Permanent and Temporary Components of Stock Prices," *Journal of Political Economy* 96 2, 246-273.

15] Fama, Eugene, F. and Kenneth R. French, 1989,"Business Conditions and Expected Returns on Stocks and Bonds," *Journal of Financial Economics* 25, 23-49.

16] Jegadeesh, Narasimhan, 1990,"Evidence of Predictable Behavior of Security Returns," *The Journal of Finance* 45 3 881-898.

17] Kim, Myung Jig, Charles R. Nelson, 1991,"Mean Reversion in Stock Prices? A Reappraisal of the Empirical Evidence," *Review of Economic Studies* 58, 515-528.

18] Mankiw, N. Gregory, David Romer and Matthew D. Shapiro, 1991, "Stock Market Forecastability and Volatility: A Statistical Appraisal," *Review of Economic Studies* 58 455-477.

19]Granger, C.W.J., in association with M. Hatanaka, 1964, *Spectral Analysis of Economic Time Series*, Princeton University Press: Princeton, NJ.

20] Granger, C. and O. Morgenstern, 1963, "Spectral Analysis of New York Stock Market Prices," *Kyklos* 16, 1-27.

21] Ramsey, J.B., 1990: "Economic and Financial Data as Nonlinear Processes." in *The Stock Market: Bubbles, Volatility, and Chaos*, Edts. Dwyer, Jr., G.P. and Hafer, R.W., Kluwer Academic Publs., Boston.

22] Ramsey, James B. and Cheng-Ping Lian, 1993, "An exploratory analysis of the growth rates of economic and financial data", *Richerche Economiche*,vol. 47, pp.31-64.

23] Mallat, Stephane, 1988, "Review of Multifrequency Channel Decompositions of Images and Wavelet Models," *Robotics Research Technical Report*, New York University Courant Institute Mathematical Sciences,Technical Report No. 412, Robotics Report No. 178, November.

24] Daubechies, Ingrid, 1990, "The Wavelet Transform, Time-Frequency Localization and Signal Analysis," reprinted from *Ieee Transactions on Information Theory*, Vol. 36, No. 5, September.

25] Feder, Jens,1988, *Fractals*, Plenum Press, New York.

26] Arneodo, A., Grasseau, G., & Holschneider, M.,1988, *Phys. Rev.*

Lett., 61, pp. 2281-2284.

27] Scher, H., Shlesinger, M.T., Bendler, J.T., 1991, "Time-Scale Invariance in Transport and Relaxation," in *Physics Today*, pp. 26-34, January.

28] Holschneider,M., 1988, *J. Stat. Phys.*,50, pp.963-993

29] Argoul, F., Arneodo, A., Grasseau, G., Gagne, Y., Hopfinger, E.J. & Frisch, U.,1989, *Nature*, 333, pp. 51-53.

30] Zaslavsky, G.M., Chernikov, A.A., Praskovsky, A.A., Karyakin, M. Yu., and Usikov, D. A., 1992, "Measurements of Local Scaling of Turbulent Velocity Fields at High Reynolds Numbers." in *Topological Aspects of the Dynamics of Fluids and Plasmas*, edited by H. K. Moffat, G. Zaslavsky, P. Comte, and M. Tabor, H.K. Moffatt, G.M. Kluwer Academic Publishers, Boston, p.443.

Figure 1

Standard & Poor's 500 Stock Price Index

No. Observations from Jan. 03 '28.

Figure 2

Standard & Poor's Stock Price Growth Rate

No. Observations from Jan. 04 '28.

Figure 3a

Color Plot of the Power of the Wavelet Transform
of the Stock Price Growth Rate for the
Entire Period of Jan '28 to June '90.
Centered at June '56 at a Zoom Level of 5

Figure 3b

Color Plot of the Power of the Wavelet Transform
of the Stock Price Growth Rate for the Period June '52
to June '90. Centered at Oct. '72 at a Zoom Level of 4

Figure 4a

Color Plot of the Power of the Wavelet Transform
of the Stock Price Growth Rate for the Period July '62
to Dec. '82. Centered at Oct. '72 at a Zoom Level of 3

Figure 4b

Color Plot of the Power of the Wavelet Transform
of the Stock Price Growth Rate for the Period Aug. '67
to Nov. '77. Centered at Oct. '72 at a Zoom Level of 2

Figure 5a

Color Plot of the Power of the Wavelet Transform
of the Stock Price Growth Rate for the Period April '70
to April '75. Centered at Oct. '72 at a Zoom Level of 1

Figure 5b

Color Plot of the Power of the Wavelet Transform
of the Stock Price Growth Rate for the Period July '71
to June '74. Centered at Oct. '72 at a Zoom Level of 0

Figure 6

Schematic Representation of Peaks at Each Zoom Level

At each zoom level k, 2^k data points are represented by each pixel. Triangular markers indicate the position of the larger peaks, or apex structures within each zoom level. Each row is 640 pixels long.

Figure 7a

A Relief Plot for 4,000 Observations on i.i.d. Gaissian (0,1) Deviates at a Zoom Level of 4.

The slope is approximately -0.5. This indicates a $\sqrt{\lambda}$ power decline which represnts an implicit averaging of noise over a wavelet of length λ.

Figure 7b

Relief Plot for 13,400 Observation on Growth Rates for Standard & Poor's 500 Stock Price Index at a Zoom level of 3[a]

[a] The time Period is Jan. '48 to Jan. '68. Half window size in 256.

STANDARD & POOR'S 500 STOCK PRICE INDEX

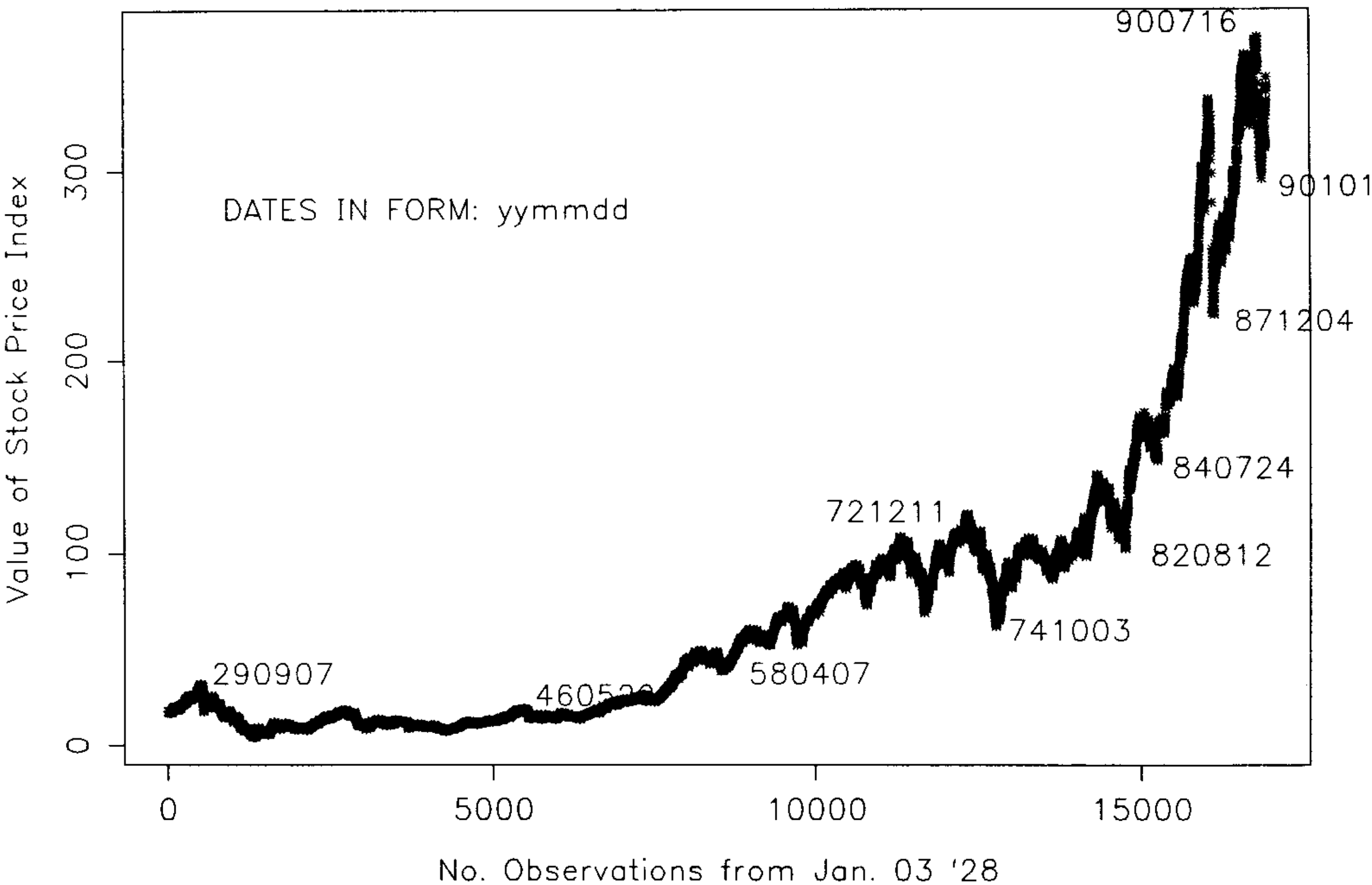


Fig. 1

STANDARD & POOR'S STOCK PRICE GROWTH RATE

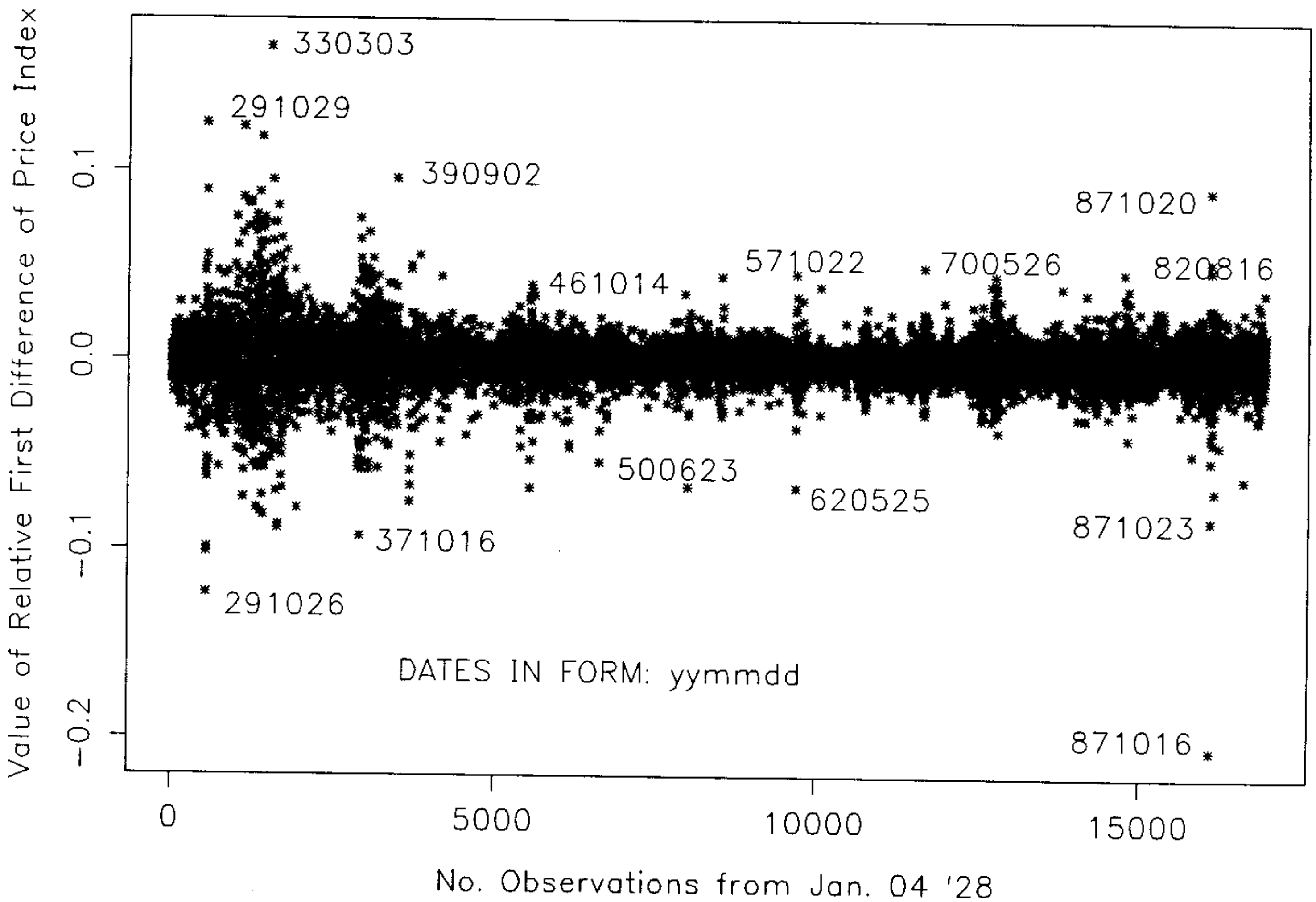


Fig. 2

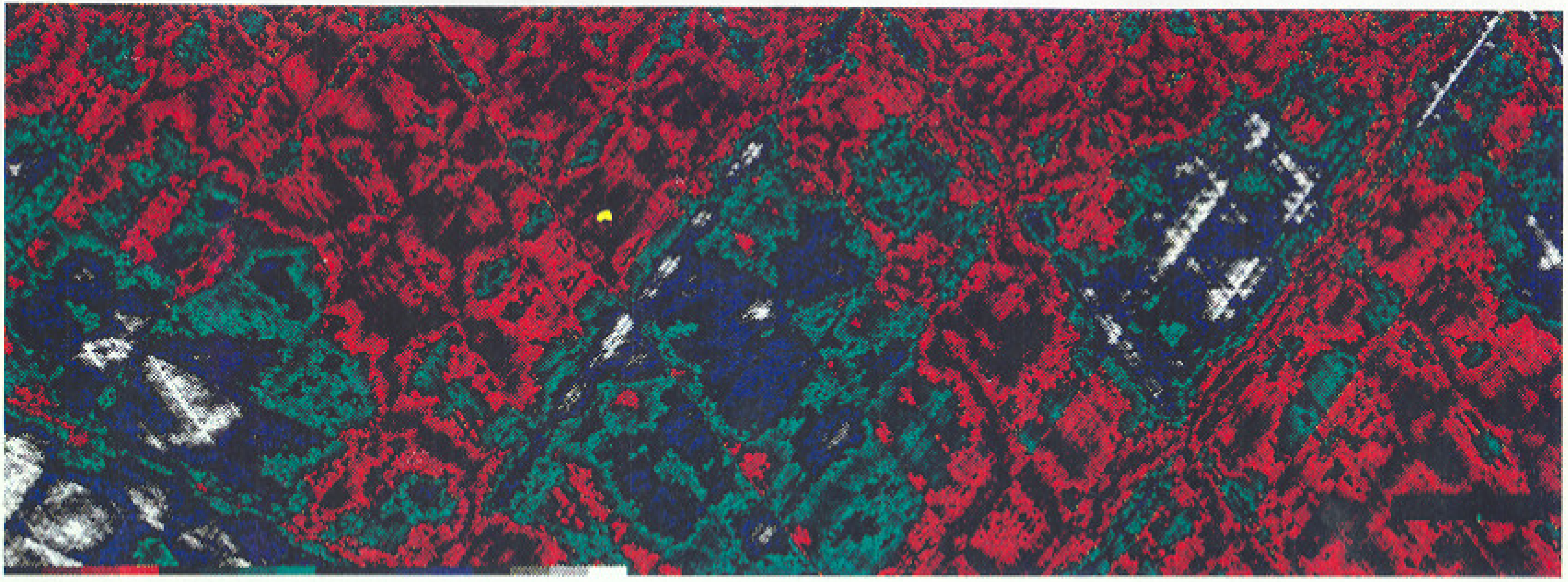


Fig. 3
B

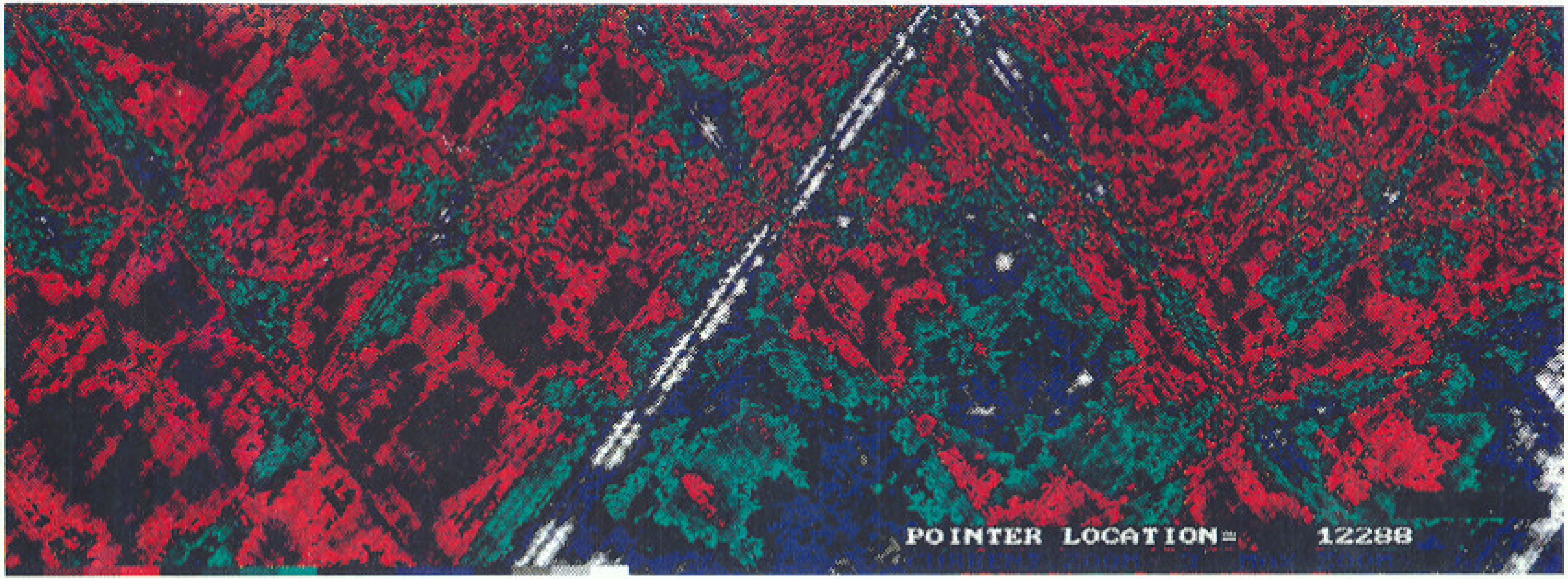


Fig 4
A

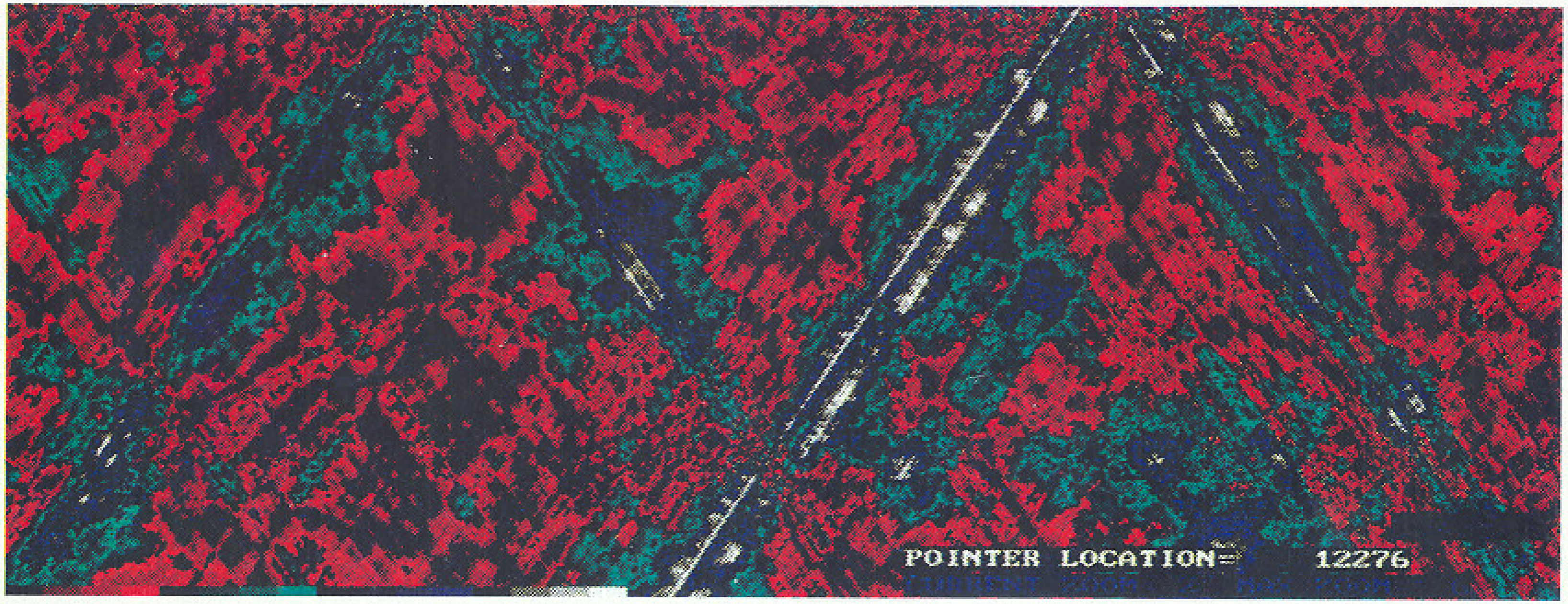
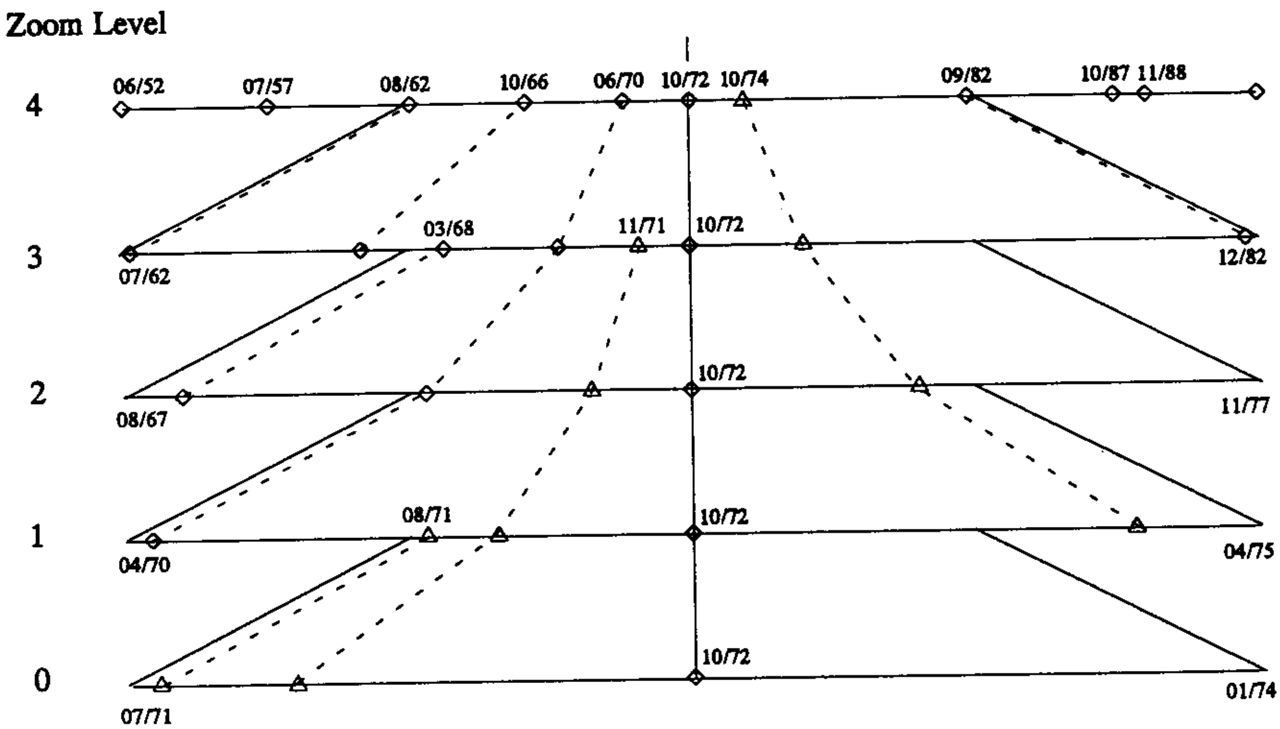


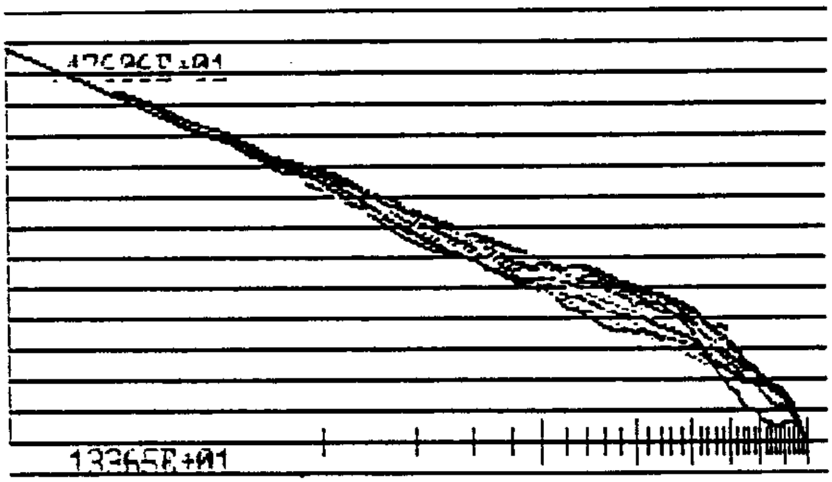
Fig. 4
B

Figure 6

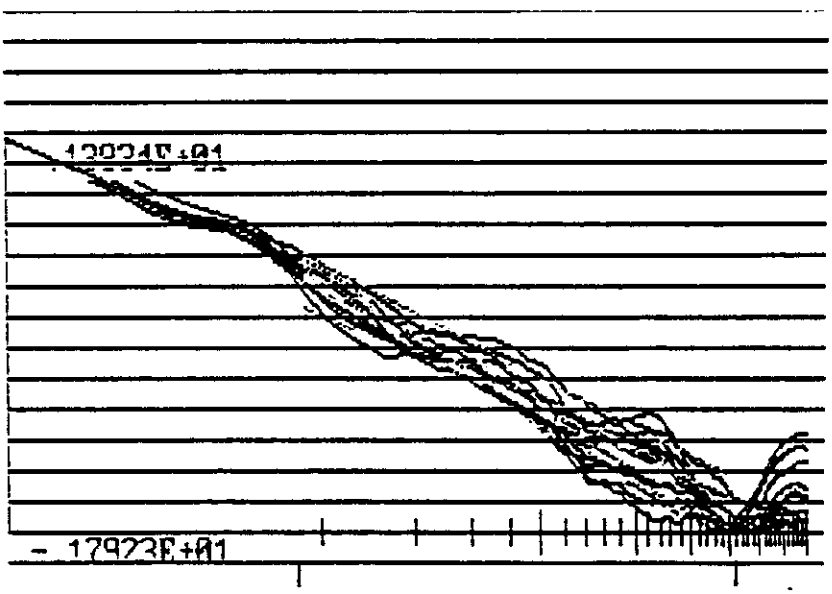
Schematic Representation of Peaks at Each Zoom Level



At each zoom level k, 2^k data points are represented by each pixel. Triangular markers indicate the position of the larger peaks, or apex structures within each zoom level. Each row is 640 pixels long.



a



b

Fig. 7