ECONOMIC RESEARCH REPORTS

THE APPLICATION OF WAVE FORM DICTIONARIES TO STOCK MARKET INDEX DATA

BY

James B. Ramsey
and
Zhifeng Zhang

RR # 94-05 *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

THE APPLICATION OF WAVE FORM DICTIONARIES TO STOCK MARKET INDEX DATA

by

James B. Ramsey
Dept. of Economics
New York Univ.
New York, New York
10003

Dr. Zhifeng Zhang
Dept. of Mathematics
Stanford Univ.
Stanford, Ca.
94305

February 14, 1994

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

THE APPLICATION OF WAVE FORM DICTIONARIES TO STOCK MARKET INDEX DATA

ABSTRACT

A matching pursuit algorithm is used to implement the application of wave form dictionaries to decompose the signal in the stock market (Standard and Poor's 500) index. A wave form dictionary is a class of transforms that generalizes both windowed Fourier transforms and wavelets. Each wave form is parametrized by location, frequency, and scale. Such transforms can analyze signals that have highly localized structures in either time or frequency space as well as broad band structures.

The Standard and Poor's 500 stock market index is found to be highly complex, but not a random walk. There are bursts of high energy that arise suddenly with very localized energy and die out equally quickly. In addition there is evidence of Dirac delta functions representing impulses, or shocks, to the system that seem to cluster more than would be expected under an hypothesis of random variation. It would appear that the energy of the system is largely internally generated, rather than the result of external forcing. Finally, there is apparently some evidence for a quasi-periodic occurrence of oscillations that are well localized in time, but that involve almost all frequencies.

THE APPLICATION OF WAVE FORM DICTIONARIES TO STOCK MARKET INDEX DATA

INTRODUCTION

Two ideas are believed to be known about the behavior of stock market prices. The relative returns, or growth rates, are approximately a random walk, alternatively stated, prices are approximately a geometric Brownian motion, Fama(1990), Fama and French(1988a, b); and the series is nonstationary as indicated by the heteroskedastic behavior of the variances, Bollerslev et al(1990), LeBaron(1992), Poterba et al(1986). Consequently, to analyze such data requires techniques that allow for the presence of nonstationarity, do not make excessive demands on the data, and do not require knowledge of the detailed structure of the generating process which the economist does not have.

Traditional methods, such as correlation analysis, Fama(1990);, ARIMA and ARCH models, Mankiw et al(1991), Bollerslev et al(1990), and spectral methods, Granger and Morgenstern(1963), Daniel and Torous(1990), have corroborated in large part the basic assumptions made above; although there is evidence that the Brownian motion idea is only approximately right, Fama and French(1988), LeBaron(1990), Poterba and Summers(1986). Recently, ideas of "mean reversion" and "special day of the week" and "end of the month " effects have clarified the degree of approximation that is involved, Kim et al(1991), Cutler et al(1989), and Schwert(1989). There is very little evidence for spectral power at almost any frequency, although there is a surprising degree of coherence between indices of monthly production and monthly values of the Standard and Poor's stock market index, see for example, Ramsey and Thomson(1994). Finally, some recent work by Ramsey and Zaslavsky(1994) using wavelet analysis indicates that the stock market data are not random walks, but that the evidence for structure is at best weak; the data are complex for sure.

Recently, Stephen Mallat and colleagues have pioneered the estimation of dictionaries of wave forms, Mallat and Zhang(1993). This approach is a generalization of both fourier analysis and of wavelet analysis, Daubechies(1991). As such the approach is particularly suited to the difficulties faced by an economist trying to understand the behavior of stock market prices. As will be explained in the next section, the benefit from the more general approach is that one can allow for all frequencies up to the Nyquist, including Dirac delta functions, and more importantly, the frequency evaluations can be localized in time and by scale of the windowing function.

This short chapter is in three sections. The first will review the properties of dictionaries of wave forms, establish some notation, and indicate how to estimate the coefficients in the wave form approximation. The second section will report on the results of applying the analysis to the S and P 500 stock market index. The last section provides a summary of the results and explores the implications for future research.

MATCHING PURSUIT AND DICTIONARIES OF WAVEFORMS

Many signals are not stationary. Indeed, some signals, such as speech and perhaps financial data, are highly nonstationary and the nonstationarity often involves intermixing Dirac delta function impulses with discrete shifts in frequencies. Linear expansions in terms of a single basis, whether it is a Fourier, wavelet, or any other basis, are not flexible enough. The basis that might be suitable for one part of the data, may be most unsuitable for another part of the data. A Fourier basis provides a poor representation of signals that are tightly localized in time; a wavelet basis is not well adapted to represent functions that have Fourier representations that are highly localized in frequency space. Each of these examples of an expansion basis for representing signals works well for the types of signals for which they were designed. But if one suspects, either that the signal is highly nonstationary, or that a signal is a mixture of discrete and continuous changes, then more robust, but less specific, tools of analysis are needed. Flexible decompositions are important for representing signals that are characterized by localizations in time and frequency that vary widely over the whole signal. Impulses need to be decomposed over functions that are well concentrated in time; while spectral lines that are well localized in frequency are best represented by waveforms that have a narrow frequency support. Unfortunately, one cannot obtain high resolution in both time and frequency spaces at the same location in time-frequency space. Consequently, a local choice has to be made as to whether time or frequency localization will best represent the signal. This means that a flexible method is required.

The analytical approach used in this paper generalizes both windowed fourier transforms and wavelets. The former approach enables one to estimate frequencies locally and to separate nearby frequencies. The latter approach concentrates on the effects of scaling of the data and enables one to estimate local time effects and to separate nearby impulses. Using wave form dictionaries, we will be able to handle both difficulties in one transformation.

Definition of Time-Frequency Atoms

A general family of time-frequency atoms can be generated by scaling, translating, and modulating a single window function $g(t) \in L^2(R)$. We suppose that the windowing function g(t) is real and centered at 0. We also impose the conditions that $\|g\| = 1$, where $\|.\|$ represents a suitable norm, that the integral of g(t) is non-zero and that $g(0) \neq 0$. For any scale parameter $s > 0$, frequency modulation ξ, and translation u, we define the triplet $\gamma = (s, u, \xi)$ and define the "atom", $g_\gamma(t)$ by:

$$g_\gamma(t) = \frac{1}{\sqrt{s}} g\left(\frac{t-u}{s}\right) e^{i\xi t} \quad . \tag{1}$$

The index γ is an element of the set $\Gamma = R^+ \times R^2$. The factor $\frac{1}{\sqrt{s}}$ normalizes the norm of $g_\gamma(t)$ to 1. The function $g_\gamma(t)$ is centered at the abscissa u and its energy is concentrated in a neighborhood of u, whose size is proportional to s. Its Fourier transform is centered at the frequency $\omega = \xi$ and has an energy concentrated in a neighborhood of ξ, whose size is proportional to 1/s.

The dictionary of time-frequency atoms D is defined by $D = \{g_\gamma(t)\}_{\gamma \in \Gamma}$. The dictionary so defined is a highly redundant set of functions that includes window Fourier frames and wavelet frames, Daubechies(1991) ; redundancy means that the set Γ is not restricted to a basis set. When the signals include time-frequency structures of very different types, that is, when non-stationarity is an important and significant aspect of the signal, one can not choose *a priori* a single frame that is well adapted to perform the expansion for all the constituent structures. Rather, for any given signal, we need to find a sequence of atoms from the dictionary that best match the signal structures in order to obtain a compact decomposition. In the next section we study such adaptive decompositions from redundant dictionaries.

2. Matching Pursuit

Let **H** represent a signal space. We define a dictionary as a family D = $\{g_\gamma(t)\}_{\gamma \in \Gamma}$ of vectors in **H**, such that $\|g_\gamma\| = 1$. We impose the condition that the linear expansion of vectors in D is dense in **H**. In general D is a redundant family of vectors that contains more vector elements than is required for a basis. The smallest complete dictionaries are bases. Consequently, when a dictionary is redundant a signal will not have a unique representation as a sum of vector elements, or atoms. Unlike the case of restricting attention to a basis for the space **H**, we have some degrees of freedom in choosing a signal's particular representation. This freedom allows us to choose a subset of the dictionary that is tailored to the signal in question and which will provide the most compact representation. We choose that subset of the dictionary for which the signal energy is concentrated in as few terms as possible. The chosen atoms highlight the dominant signal features as measured by the energy of the signal captured by the atoms.

Let D be a dictionary of vectors in **H**. An optimal approximation of f ∈ **H** is the expansion:

$$\tilde{f} = \sum_{n=1}^{N} a_n g_{\gamma n}, \tag{2}$$

where N is the number of terms in the expansion for a given degree of approximation. a_n and $g_{\gamma n} \in$ D are chosen in order to minimize

$$\|f - \tilde{f}\|,$$

where $\|.\|$ represents a suitable norm.

Because of the impossibility of computing numerically an optimal solution, we utilize a "greedy algorithm", Mallat and Zhang(1993), that computes a useful sub-optimal approximation. Let $f \in$ H. We want to compute a linear expansion of f over a set of vectors selected from D in such a way as to capture best the signal's inner structure. This is accomplished by successive approximations of f through orthogonal projections of the signal onto elements of D. Let $g_{\gamma 0} \in$ D. The vector f can be decomposed into

$$f = \langle f, g_{\gamma 0}\rangle g_{\gamma 0} + Rf, \tag{3}$$

where Rf is the residual vector after approximating f in the direction of $g_{\gamma 0}$. Clearly $g_{\gamma 0}$ is orthogonal to Rf and is itself normalized to 1, so that:

$$\|f\|^2 = |\langle f, g_{\gamma 0}\rangle|^2 + \|Rf\|^2. \tag{4}$$

To minimize $\|Rf\|$, we must choose $g_{\gamma 0} \in$ D such that $|\langle f, g_{\gamma 0}\rangle|$ is maximum. In some cases, it is only possible to find a vector $g_{\gamma 0}$ that is almost the best in the sense that

$$|\langle f, g_{\gamma 0}\rangle| \geq \alpha \sup_{\gamma \in \Gamma} |\langle f, g_\gamma\rangle|, \tag{5}$$

where α is an optimality factor that satisfies $0 < \alpha \leq 1$.

A matching pursuit is an iterative algorithm that at successive stages decomposes the residue Rf from a prior projection by projecting that residue onto a vector of D, as was done for f. This procedure is repeated for each residue that is obtained from a prior projection. Let $R^0f = f$. Suppose that the n^{th} order residue R^nf, for some $n \geq 0$ has been computed. At the nest stage we choose an element $g_{\gamma n} \in$ D which approximates the residue R^nf:

$$|<R^nf, g_{\gamma n}>| \geq \alpha \sup_{\gamma \in \Gamma} |<R^nf, g_{\gamma}>| \ . \tag{6}$$

The residue R^nf is itself decomposed into:

$$R^nf = <R^nf, g_{\gamma n}> g_{\gamma n} + R^{n+1}f \ , \tag{7}$$

which defines the residue to order n+1. Since $R^{n+1}f$ is orthogonal to $g_{\gamma n}$

$$\|R^nf\| = |<R^nf, g_{\gamma n}>|^2 + \|R^{n+1}f\|^2 \ . \tag{8}$$

If this decomposition is continued up to order m, then f has been decomposed into the concatenated sum

$$f = \sum_{n=0}^{m-1} (R^nf - R^{n+1}f) + R^mf \ . \tag{9}$$

Substituting equation (6) into (8) yields

$$f = \sum_{n=0}^{m-1} <R^nf, g_{\gamma n}> g_{\gamma n} + R^mf \ . \tag{10}$$

Similarly, $\|f\|^2$ is decomposed in terms of the concatenated sum Equation (7) yields an energy conservation equation

$$\|f\|^2 = \sum_{n=0}^{m-1}\left(\|R^n f\|^2 - \|R^{n+1} f\|^2\right) + \|R^m f\|^2 . \tag{11}$$

$$\|f\|^2 = \sum_{n=0}^{m-1} |<R^n f, g_{\gamma n}>|^2 + \|R^m f\|^2 . \tag{12}$$

Although the decomposition is non-linear, we maintain an energy conservation as if it were a linear orthogonal decomposition. A major task is to understand the behavior of the residue $R^m f$ as m increases. By transposing a result proved by Jones(1987) for projection pursuit algorithms, Huber(1985), one can prove that the matching pursuit algorithm as outlined above does converge, even in infinite dimensional spaces, Mallat and Zhang(1993). The following theorem is from Mallat and Zhang(1993).

Theorem 1 *Let $f \in H$. The residue defined by the induction equation (6) satisfies*

$$\lim_{m \to +\infty} \|R^m f\| = 0 . \tag{13}$$

Hence

$$f = \sum_{n=0}^{+\infty} <R^n f, g_{\gamma n}> g_{\gamma n} , \tag{14}$$

and

$$\|f\|^2 = \sum_{n=0}^{+\infty} |<R^n f, g_{\gamma n}>|^2. \tag{15}$$

When H has finite dimension, $\|R^m f\|$ decays exponentially to zero.

3. Implementation of a Matching Pursuit Algorithm

When the dictionary is very redundant, the search for the vectors that best match the signal residues can be limited to a sub-dictionary $D_\alpha = \{g_\gamma\}_{\gamma\in\Gamma_\alpha} \subset$ D. We suppose that Γ_α is a finite index set included in Γ such that for any $f \in H$

$$\sup_{\gamma\in\Gamma_\alpha} |<f, g_\gamma>| \geq \alpha \sup_{\gamma\in\Gamma} |<f, g_\gamma>| . \tag{16}$$

Depending upon the chosen value of α and the degree of dictionary redundancy, the set Γ_α can be made much smaller than Γ.

The matching pursuit is initialized by computing the inner product $(<f, g_\gamma>)_{\gamma\in\Gamma_\alpha}$. The process continues by induction. Suppose that we have already computed $(<R^n f, g_\gamma>)_{\gamma\in\Gamma_\alpha}$, for $n \geq 0$. We search in D_α for an element $g_{\tilde{\gamma}n}$ such that

$$|<R^n f, g_{\tilde{\gamma}_n}>| = \sup_{\gamma\in\Gamma_\alpha} |<R^n f, g_\gamma>| . \tag{17}$$

In order to find a dictionary element that approximates f better than $g_{\tilde{\gamma}n}$, we search using a Newton method for an index γ_n within a neighborhood of $\tilde{\gamma}_n$ in Γ where $|<f, g_\gamma>|$ reaches a local maximum. The optimization produces the following inequality sequence:

$$|<R^n f, g_{\gamma n}>| \geq |<R^n f, g_{\tilde{\gamma}_n}>| \geq \alpha \sup_{\gamma\in\Gamma} |<R^n f, g_\gamma>| . \tag{18}$$

Once the vector $g_{\gamma n}$ is selected, we compute the inner product of the new residue $R^{n+1}f$ with any $g_\gamma \in D_\alpha$, with an updating formula derived from equation (6)

$$<R^{n+1}f,\ g_{\gamma_n}> = <R^n f,\ g_\gamma> - <R^n f,\ g_{\gamma_n}><g_{\gamma_n},\ g_\gamma> \ . \tag{19}$$

Since both $<R^n f, g_\gamma>$ and $<R^n f, g_{\gamma_n}>$ have been stored, this update requires us to compute only the expression $<g_{\gamma_n}, g_\gamma>$. Dictionaries are generally built so that this inner product is recovered with a small number of operations. Mallat and Zhang(1993) describe how to compute efficiently the inner product of two discrete Gabor atoms in order(1) operations.

The number of times we sub-decompose the residues of a given signal f depends upon the desired precision ϵ. The number of iterations is the minimum p such that

$$\|R^p f\| = \|f - \sum_{n=0}^{p-1} <R^n f,\ g_{\gamma_n}> g_{\gamma_n}\| \le \epsilon\, \|f\| \ . \tag{20}$$

The energy conservation equation (11) proves that this last equation is equivalent to

$$\|f\|^2 - \sum_{n=0}^{p-1} |<R^n f,\ g_{\gamma_n}>|^2 \le \epsilon^2 \|f\| \ . \tag{21}$$

Since we do not compute the residue $R^n f$ at each iteration we test the validity of (20) in order to stop the process.

4. Matching Pursuit With Time-Frequency Dictionaries

For dictionaries of time-frequency atoms, a matching pursuit yields an adaptive time-frequency transform. It decomposes any function $f(t) \in L^2(R)$ into a sum of complex time-frequency atoms that best match its residues. This section studies the properties of this particular matching pursuit decomposition. We derive a new type of time-frequency energy distribution by summing the Wigner distribution of each time-frequency atom.

Since a time-frequency atom dictionary is complete, Theorem 1 proves that a matching pursuit decomposes any function $f(t) \in L^2(R)$ into

$$f = \sum_{n=0}^{+\infty} <R^n f,\ g_{\gamma n}> g_{\gamma n}\ , \tag{22}$$

where $\gamma_n = (s_n, u_n, \xi_n)$ and

$$g_{\gamma n}(t) = \frac{1}{\sqrt{s_n}} g\left(\frac{t-u_n}{s_n}\right) e^{i\xi_n t}\ . \tag{23}$$

The term $<R^n f,\ g_{\gamma n}>$ in equation (22) is the same term as "a_n" in equation (2).

These atoms are chosen in sequence to approximate the sequence of residues of f. For signals of size N, we can discretize γ and facilitate the computations in the following manner. We redefine γ by:

$$\gamma = (2^j,\ p2^{j-1},\ k\pi 2^{-j}) \tag{24}$$

where $s = 2^j$, $u = 2^{j-1}p$, $\xi = k\pi 2^{-j}$, and (j, p, k) are integers ; $0 < j < \log_2 N$, $0 \le p < N2^{-(j+1)}$, $0 \le k < 2^{j+1}$. Using these definitions, equation (22) for the nth atom becomes:

$$g_{\gamma_n}(t) = \frac{1}{\sqrt{2^{j_n}}} g\left(\frac{t - 2^{j-1}p_n}{2^{j_n}}\right) e^{ik_n \pi 2^{-j_n} t} \tag{25}$$

where j_n , p_n, and ξ_n are the values taken by the triplet (j, p, ξ) for the nth atom.

From the decomposition of any $f(t)$ within a time-frequency dictionary, we derive a new time-frequency energy distribution, by adding the Wigner distribution for each selected atom. The cross Wigner distribution for two functions $f(t)$ and $h(t)$ is defined by:

$$W[f,\, h](t,\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(t + \frac{\tau}{2}) \bar{h}(t - \frac{\tau}{2}) e^{-i\omega\tau} d\tau \ . \tag{26}$$

The Wigner distribution of $f(t)$ is $Wf(t,\omega) = W[f,f](t,\omega)$. Since the Wigner distribution is quadratic, we derive from the atomic decomposition (21) of $f(t)$ that

$$\begin{aligned} Wf(t,\, \omega) = & \sum_{n=0}^{+\infty} | < R^n f,\, g_{\gamma n} > |^2 W g_{\gamma_n}(t,\, \omega) \\ & + \sum_{n=0}^{+\infty} \sum_{m=0,\, m \neq n}^{+\infty} < R^n f,\, g_{\gamma_n} > \overline{< R^m f,\, g_{\gamma m} >} W[g_{\gamma n},\, g_{\gamma m}](t,\omega) \ . \end{aligned} \tag{27}$$

The double sum corresponds to the cross terms of the Wigner distribution. It regroups the terms that one usually tries to remove in order to obtain a clear picture of the energy distribution of $f(t)$ in the time-frequency plane. We thus only keep the first sum and define

$$Ef(t,\, \omega) = \sum_{n=0}^{+\infty} | < R^n f,\, g_{\gamma n} > |^2 W g_{\gamma_n}(t,\, \omega) \ . \tag{28}$$

A similar decomposition algorithm over time-frequency atoms was derived independently by Qian and Chen [7], in order to define this energy distribution in the time-frequency plane. From the well known dilation and translation properties of the Wigner distribution and the expression (22) of a time-frequency atom, we derive that for $\gamma = (s,\xi,u)$

$$Wg_{\gamma}(t,\omega) = Wg(\frac{t-u}{s}, s(\omega - \xi)) \ , \tag{29}$$

and hence

$$Ef(t,\omega) = \sum_{n=0}^{+\infty} |<R^n f, g_{\gamma_n}>|^2 Wg(\frac{t-u_n}{s_n}, s_n(\omega - \xi_n)) \ . \tag{30}$$

The Wigner distribution also satisfies:

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} Wg(t,\omega)dtd\omega = \|g\|^2 = 1 \ , \tag{31}$$

so that the energy conservation equation (14) implies:

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} Ef(t,\omega)dtd\omega = \|f\|^2 \ . \tag{32}$$

We can interpret $Ef(t,\omega)$ as an energy density of f in the time-frequency plane (t,ω). Unlike the Cohen class distributions, it does not include cross product terms. It also remains positive if $Wg(t,\omega)$ is positive, which is the case when $g(t)$ is Gaussian. On the other hand, the energy density $Ef(t,\omega)$ does not satisfy marginal properties, as opposed to certain Cohen class distributions, Cohen(1989). The importance of these marginal properties for signal processing is however not clear. If $g(t)$ is the Gaussian window

$$g(t) = 2^{1/4} e^{-\pi t^2} , \tag{33}$$

then

$$Wg(t,\omega) = 2e^{-2\pi(t^2+(\frac{\omega}{2\pi})^2)} . \tag{34}$$

The time-frequency atoms $g_\gamma(t)$ are then called Gabor functions. The time-frequency energy distribution $Ef(t,\omega)$ is a sum of Gaussian pulses whose locations and variances along the time and frequency axes depend upon the parameters (s_n, u_n, ξ_n).

THE EMPIRICAL RESULTS

The data that were used are the relative first differences, or growth rates, of an index of the end of day prices on five hundred stocks that are traded on the New York Stock Exchange; this index is known as the Standard and Poor's 500 stock index. There are 16384, or 2^14, observations. The historical period covered is January 3, 1928 to November 18, 1988. A time series plot of the growth rates is shown in Figure 1. The data are characterized by a dense core of rapid oscillations interspersed with episodic bursts of oscillations of very large amplitude in some cases and many bursts of substantial variance relative to the ambient variation.

Given the discussion in the previous section, the behavior represented in Figure 1 indicates that analysis by means of dictionaries of wave forms might well prove to be a useful exploratory tool. The major reason is that prior analysis of these data indicate that Dirac delta functions might play a major role and that while oscillations at specific frequencies over the entire time domain are unlikely, specific frequencies occurring in short bursts, that is, "chirps", might well be observable. The result of running the matching pursuit algorithm on the dictionary of wave forms is illustrated in Figures 2 and 3. In both Figures, "time" runs from left to right along the ordinate, and frequency in cycles per month is represented on the abscissa from zero to one half with zero frequency at the bottom.

Figure 2 shows the Wigner distribution estimated over the 2048 days from November 1934 to September 1941. The high energy line in the middle of the graph represents the 1939 stock market collapse. Figure 3 represents 2048 observations from November 1980 to June 1990 and the high energy line just to the right of the center represents the October 1987 crash.

Both Figures are representative of all the figures that were created. The bulk of the total energy represented in the graphs is contained in either Dirac delta functions, or in very narrow bands of energy along the time axis. There is some evidence of chirps in both graphs that are concentrated about the isolated high energy lines; there are also some very low intensity chirps that are distributed at random.

The major axis for the chirps is predominantly parallel to the frequency axis; that is, there are bursts of groups of frequencies that are well localized in time. This observation is less true in the 1980 to 1990 period. In the later period there is more evidence of frequencies holding for some months at a time, dying out, then recurring, but not at regular intervals.

The localized bursts of high intensity energy have no advance build up in intensity and die out equally abruptly. When eruptions do occur, the energy is distributed over most of the frequency range, although relatively less energy is observed at very low and very high frequencies.

Except for the areas of high energy bursts, the wave form distribution is similar to that of noise. The distribution of relative intensities, as indicated by the variation in shading, for the Standard and Poor's stock index plots indicate that these are not random data, even if they seem to have a very complicated structure. Random data tend to have a much more uniform distribution of intensities, whereas the stock data have periods of relative quiescence interspersed by periods of high intensity oscillations. In addition, the more intense activity areas cluster far more with the stock data than is true for random data.

Figure 4 portrays in descending value the magnitudes of the weight functions "a_n" that are defined in equations (2) and (22). Two sequences of weights are shown, one for the S & P 500 index and one for the same number of observations on a Normal distribution; in this latter case, the number of waveforms that were calculated was three thousand, the same as for the S & P data. Consequently, the two graphs are, except for scale, identical in construction.

A data series with simple structure can be easily represented by a relatively small number of wave forms, so that the shape of the plot of weights is one of rapidly declining levels. In comparing the shape of the weight curves for the Normal noise driven results and for the S & P results, we see that despite the lack of evidence for any structure in the data so far, there is relatively more rapid decline in the values of the weights for the S & P index. The S & P is well approximated by the first thousand wave forms, whereas the noise results require nearly three thousand to achieve the same degree of approximation.

We have pointed out that the majority of the power in the wave forms is in time localized bursts or even in Dirac delta functions. Consequently, an immediate question is whether there is information in the distribution over time of the Dirac delta function wave forms. Figure 5 plots the occurrence of the Dirac delta functions over time. While there is no obvious regularity in the distribution of the delta functions, there is more clustering than would be obtained with noisy data. One interesting sidelight is that neither the October 1929, nor the October 1987, stock market crashes show up in the plot as highly significant delta functions, even though both periods are the centers of very narrow regions in the time domain with high bursts of energy. This indicates that the major stock market crashes and even the lesser ones, while of short duration, are not the output of isolated impulses.

From the discussion in the previous section, the reader will recall that the decomposition of the data by frequency components is best done at the higher scales, that is, for larger values for $s_{n;}$ the frequencies are resolved more precisely than at the lower scales. Compare Figures 6 and 7; the former shows the Wigner distribution of frequencies by position when the octave scales, "j" in equation (24), are restricted to the range of 12 to 14, 14 being the maximum level. The latter figure shows the exact same type of graph for random (Normal) data. The greater degree of regularity in the Standard and Poor's stock market data is readily apparent. What may not be so clear visually is that many of the plotted points for the Standard and Poor's data seem to be quasi-periodic. While there is definitely greater regularity in the stock market data, we are not yet prepared to claim that the observed quasi-periodicity is a genuine phenomenon and not merely an artifact of our processing that we have not yet discovered. However, if these results are at least in part true, the implication is that quasi-periodically there are bursts of energy at almost all frequencies. Such results, coupled with no strong evidence of power at specific frequencies in between the bursts, would produce spectra that would be difficult to distinguish from a noise spectrum by conventional methods.

SUMMARY AND CONCLUSIONS

The results in this analysis have confirmed much of the current thinking about stock market data; they are not random walks, but the structure of the data are very complex. The analysis cited above has been able to clarify these vague notions and to refine some of the aspects of the complexity.

Our first piece of evidence that the data are not random is that the number of structures needed to provide a given level of decomposition is far less than is needed for random data. This result is even stronger when coupled with the recognition that the stock market data are best characterized by periods of relative quiet interspersed by periods of highly time localized intense activity. These periods of intense energy occur without any apparent warning, build in intensity very rapidly, and die out equally rapidly. Nonetheless, these bursts are not the result of isolated impulses that would be represented by Dirac delta functions. There are numerous cases of isolated impulses, but the occurrences of these are separate in general from the intense bursts. The occurrence of the Dirac delta functions represent impulses, or "shocks" to the system from external sources, whereas the energy bursts that are so prominent in the data seem to be internal to the system. The bursts contain high energy at most frequencies, relatively less at very low and very high frequencies.

There is also some tentative evidence that there are occurrences of activity across a range of frequencies at quasi-periodic intervals.

This research has posed more questions than it has answered. The most important finding that will stimulate much interest is that the occurrence of sudden bursts of energy at all frequencies perhaps should not be associated with isolated impulses, or shocks. While a speculative conclusion, it is reasonable to infer from these results that the bursts of high intensive activity are not the result of isolated, unanticipated, external shocks, but more likely the result of the operation of a dynamical system with some form of intermittency in the dynamics.

References

Bollerslev, T., Chou, R. Y., Jayaraman, N., Kroner, K. F., "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence", Journal of Econometrics, 52, (1), 1990: 5-60.

L. Cohen, "Time-frequency distributions: a review", Proceedings of the IEEE, vol. 77, . 7, p. 941-979, July 1989.

Cutler, D., J. M. Poterba and L. H. Summers, "What Moves Stock Prices?," *Journal of Portfolio Management*, Spring 1989, pp. 4-12.

Daniel, K. and W. Torous, "Common Stock Returns and the Business Cycle," UCLA Working Paper, Los Angeles, CA., 1990

I. Daubechies, "Ten Lectures on Wavelets", CBMS-NSF Series in Appl. Math., SIAM, 1991.

G. Davis, S. Mallet and M. Avenaleda, "Chaotic Adaptive Time-Frequency Decompositions", Technical Report, Computer Science, NYU, April 1994.

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

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

Fama, E.F. and K. French, "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, 1988, *22*, 3-25.

Granger, C.W.J. and Oscar Morgenstern, "Spectral Analysis of New York Stock Market Prices," *Kyklos*, 1963, *16*, 1-27.

P.J. Huber, "Projection Pursuit", *The Annals of Statistics*, vol. 15, No. 2, 1985, 435-475.

L.K. Jones, "On a conjecture of Huber concerning the convergence of projection pursuit regression", *The Annals of Statistics*, vol. 15, No. 2, 1987, 880-882.

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

LeBaron, Blake, "Nonlinear Forecasts for the S&P Stock Index," in Martin Casdagli and Stephen Eubank, ed., *Nonlinear Modeling and Forecasting*, Diagnostic Testing for Nonlinearity, Chaos, and General Dependence in Time Series Data, Redwood City, Ca.: Addison-Wesley, 1992, 381-394.

S. Mallet and Z. Zhang, "Matching Pursuit with Time-Frequency Dictionaries", *IEEE Trans. on Signal Processing*, Dec. 1993.

Mankiw, N. Gregory, David Romer and Matthew D. Shapiro, "Stock Market Forecastability and Volatility: A Statistical Appraisal," *Review of Economic Studies*, 1991, *58*, 455-577.

Poterba, James, M. and Lawrence Summers, H., "The Persistence of Volatility and Stock Market Fluctuations," *The American Economic Review*, *76*(5), 1986, 1142-1151.

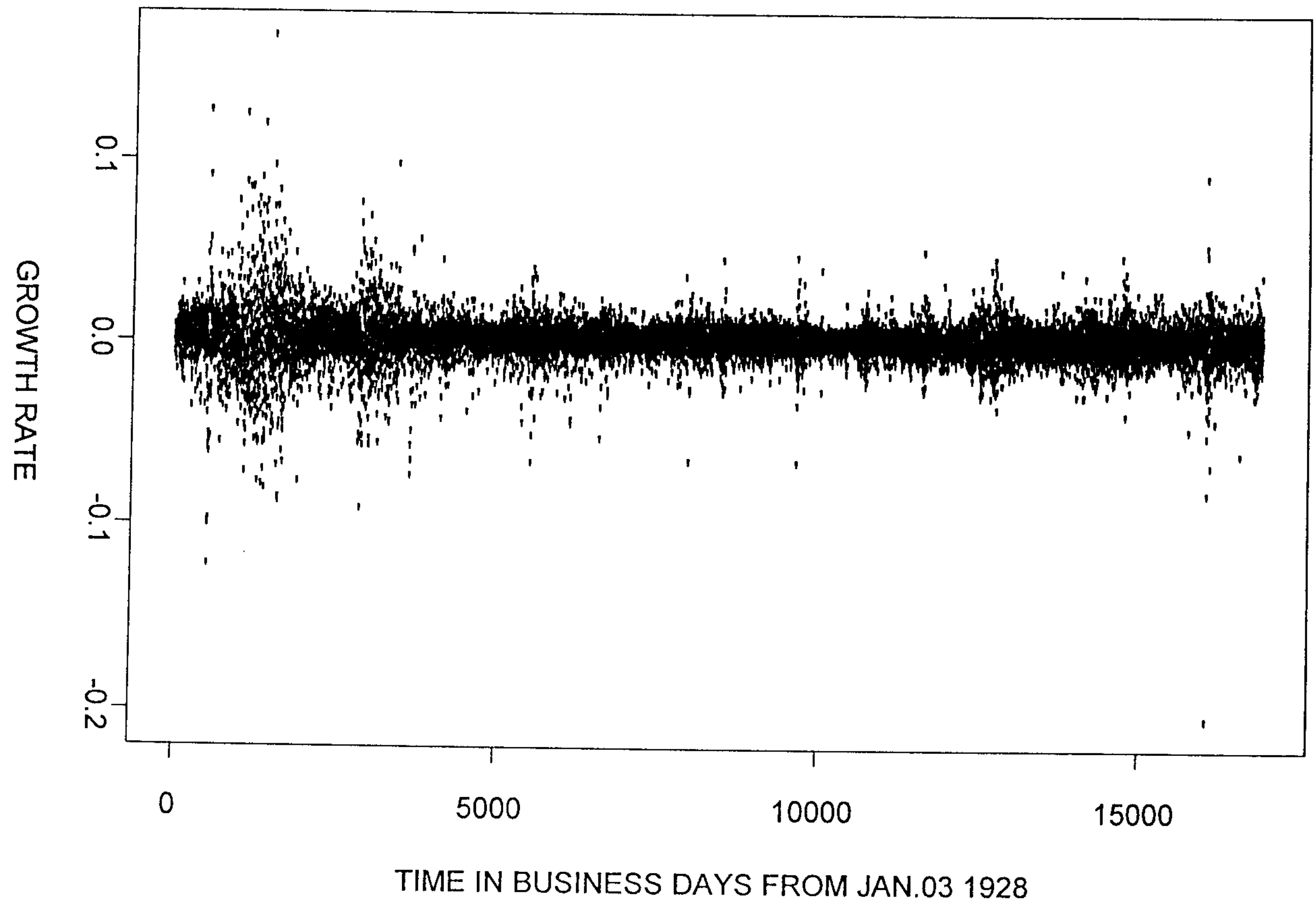
Ramsey, James B. and Thomson, David J., "A Reanalysis of the Spectral Properties of Some Economic Time Series", Dept. of Economics, Working Paper Series, New York Univ. 1994.

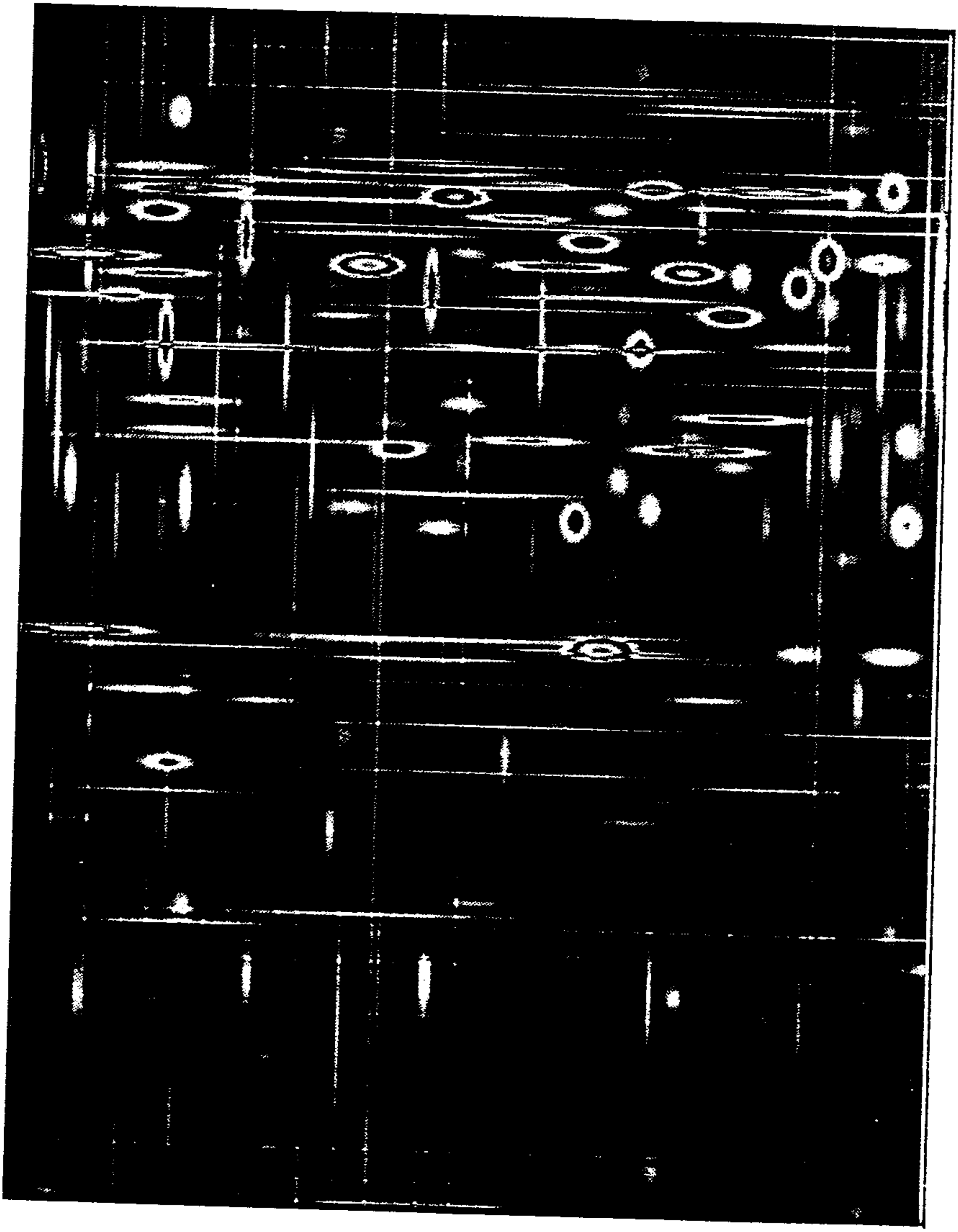
Ramsey, James B. and Zaslavsky, George, "A Wavelet Analysis of U. S. Stock Price Behavior", Dept. of Economics Working Paper Series, New York Univ., New York, February 1994.

S. Qian and D. Chen, "Signal Representation via Adaptive Normalized Gaussian Functions", *IEEE Trans. on Signal Processing*, vol. 36, no. 1, Jan. 1994.

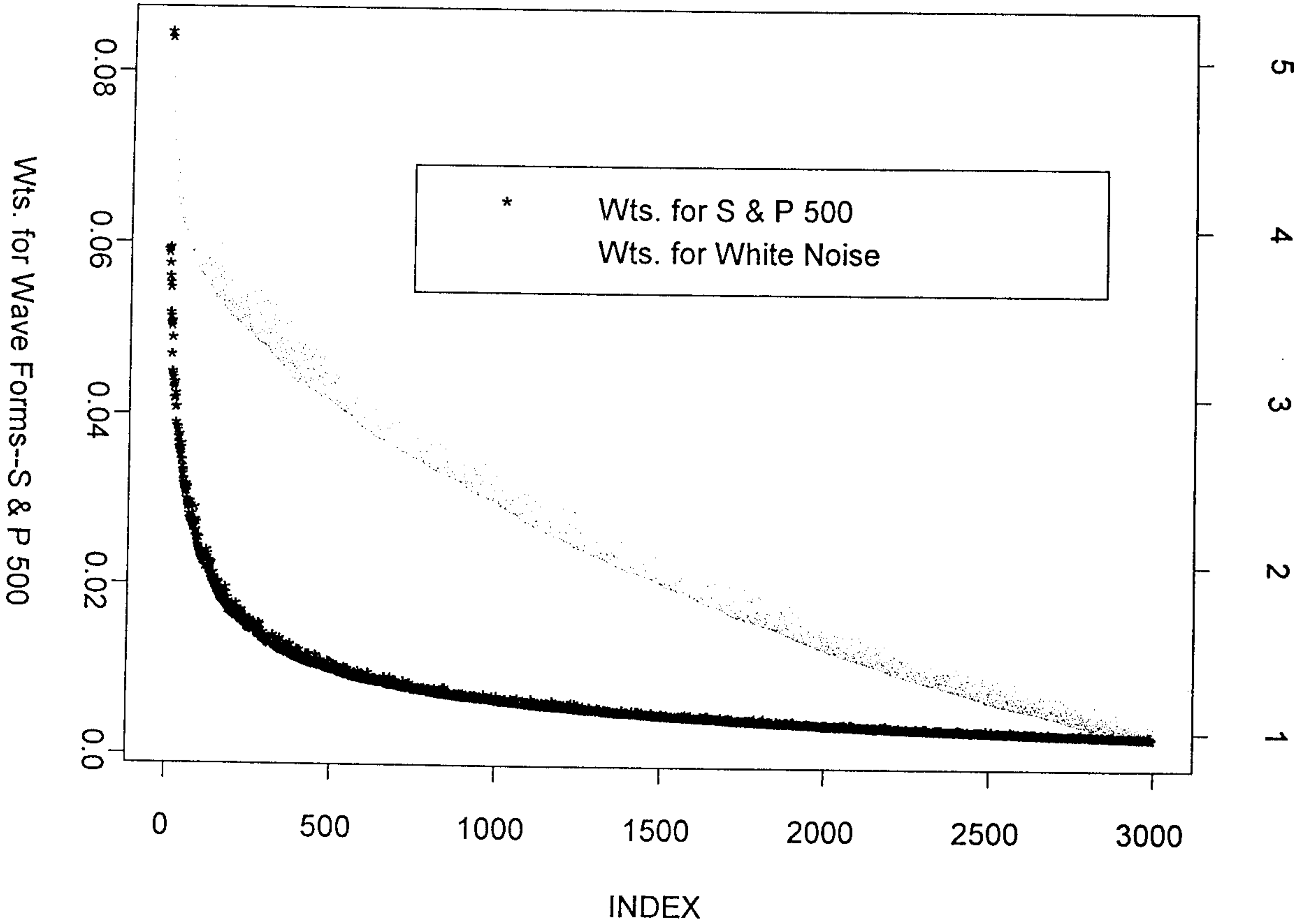
LIST OF FIGURE CAPTIONS

STANDARD & POOR'S 500 INDEX: GROWTH RATES

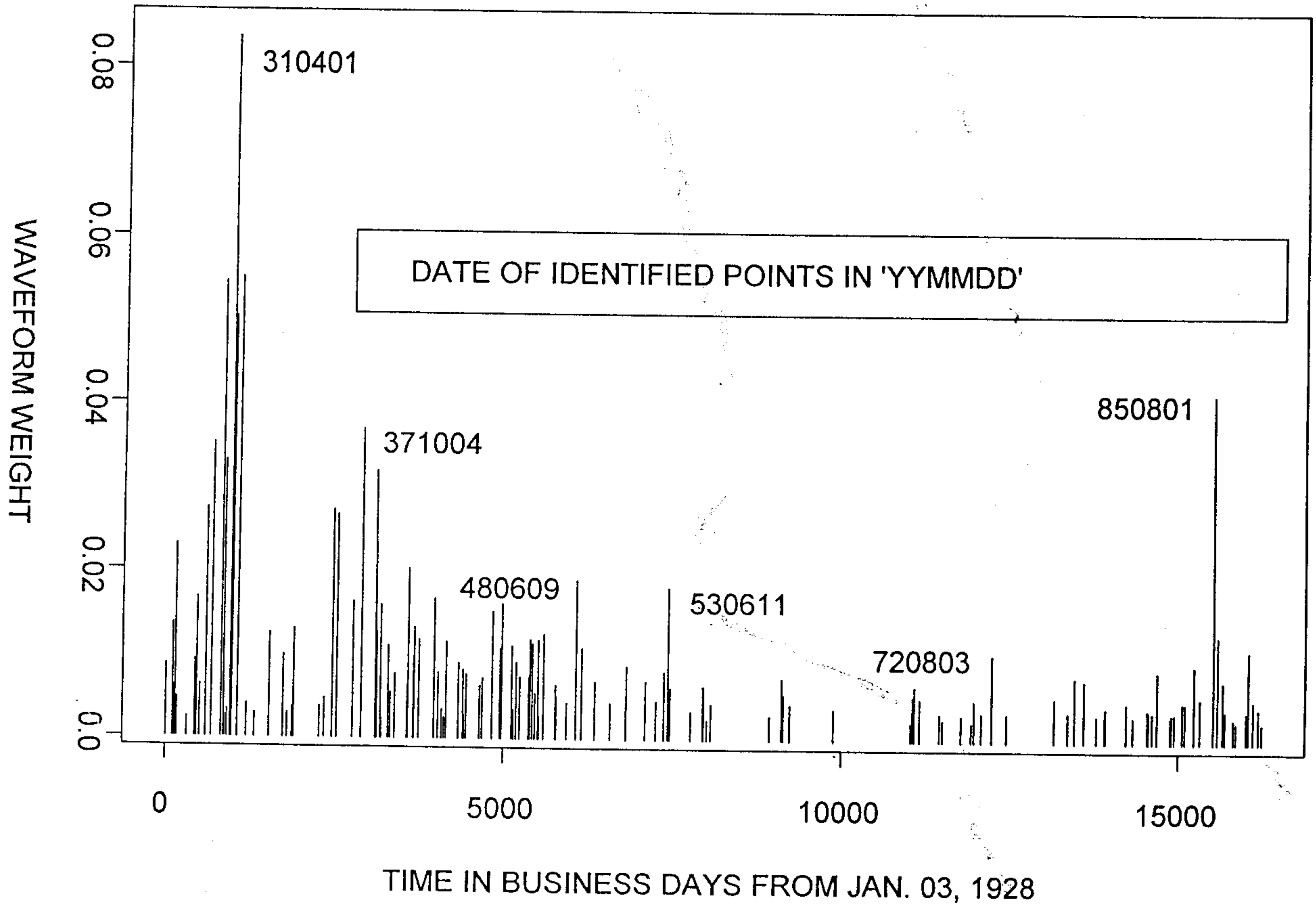




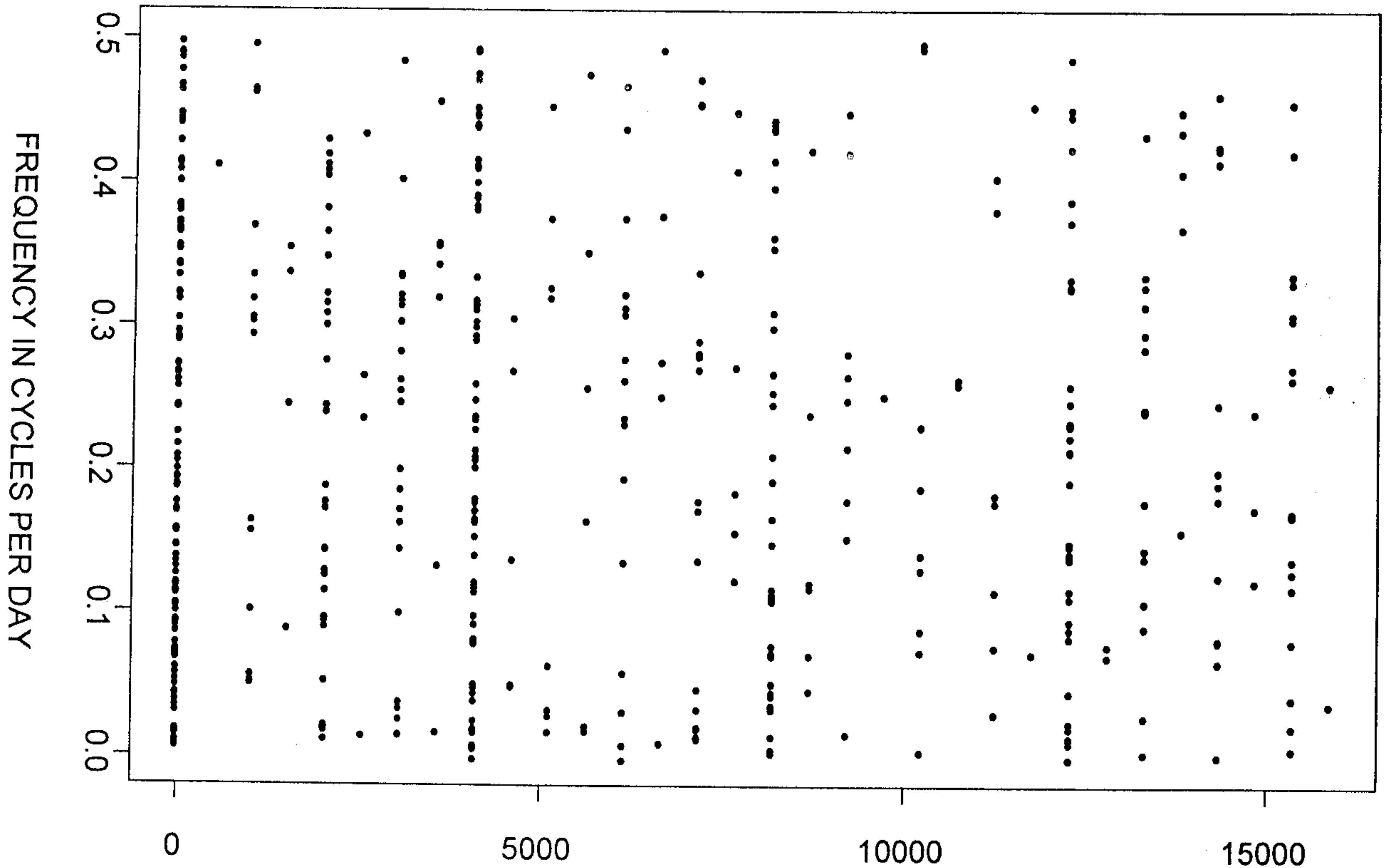
COMPARISON OF WAVE FORM WEIGHTS



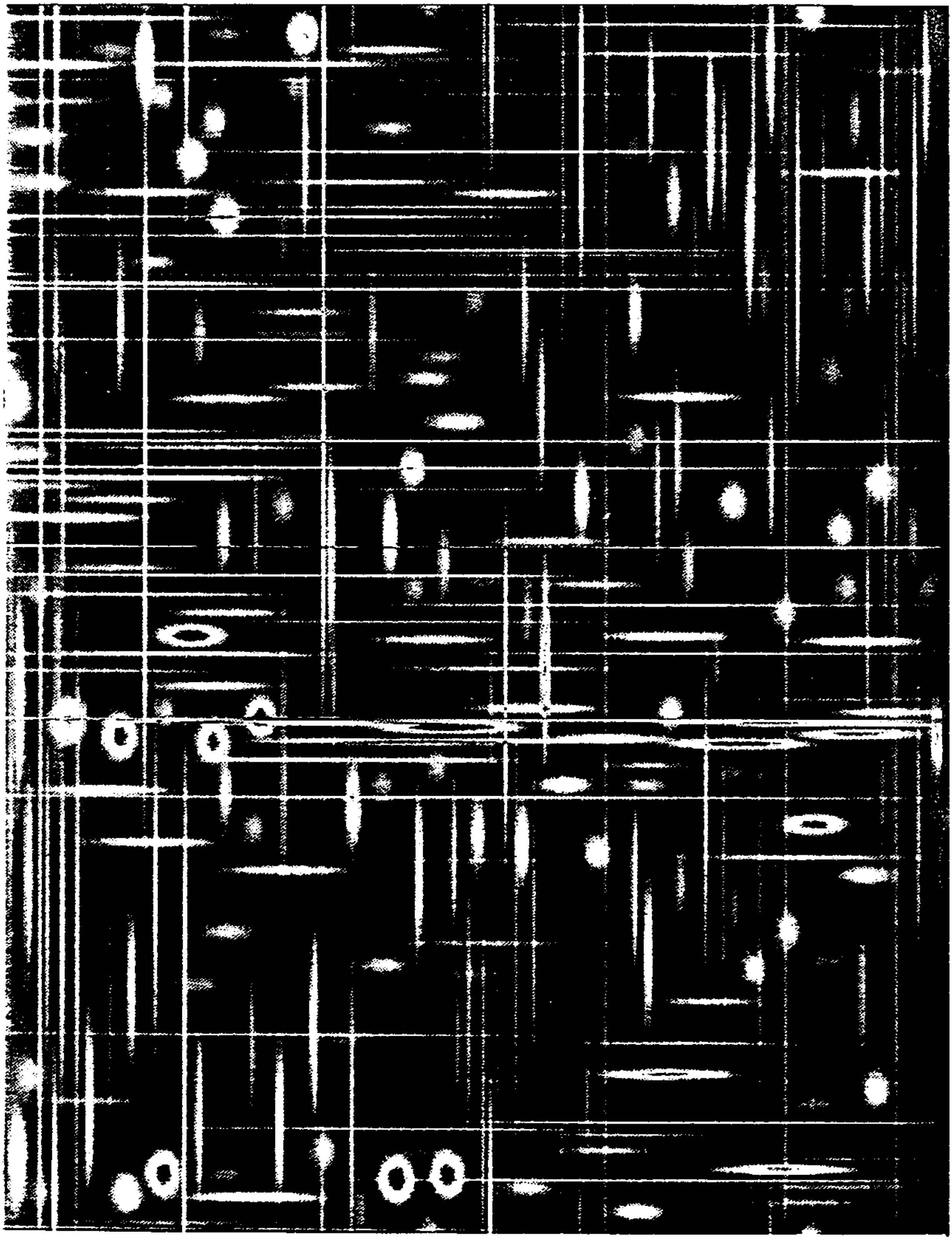
WEIGHTS FOR DIRAC DELTA FNCT. WAVEFORMS



PLOT OF FREQUENCIES BY TIME PERIOD



TIME IN BUSINESS DAYS FROM JAN. 03,1928
Data Plotted only for Scales Greater than 11



PLOT OF FREQUENCIES BY TIME PERIOD