A METHOD OF DECOMPOSING AN EMPIRICAL SERIES INTO ITS CYCLICAL AND PROGRESSIVE COMPONENTS

By Ragnar Frisch

In the last decade's intensive study of all sorts of social and economic time series, it has become clear, it seems to me, that the usual time series technique is not quite adequate for the purpose which the social investigator is pursuing. The technique which is now most in vogue does not seem powerful enough to deal with the more complicated situations which arise when the time series studied represents an *interference phenomenon* between several components: short cycles, long cycles, different orders of trends, etc., and when, furthermore, the cyclical or progressive characteristics of these various components are changing.

Of course, in any time series there are always certain intrinsic features (the relative importance of the erratic element, the degree of complexity in the nature of the underlying components, etc.) which put an absolute, so to speak "natural," limit to the amount of significant information which it is possible to obtain from the series. But although no omnipotent technique can be constructed, yet the technique is not a matter of small importance. If the method of analysis is inadequate we may be forced to give in long before the natural limit of significant information is reached. And I believe that in this respect there is still room for considerable progress over the orthodox time series methods.

The present paper summarizes very briefly some points of an attempt I have engaged upon to push the technique in this field a little step forward. The present statement will give nothing more than a general discussion of the nature of the problem and some hints at the character of the tools I am using. The whole subject will be discussed in more detail, and a series of numerical applications given in a monograph shortly to be published. A preliminary statement of my approach to the problem was mimeographed in April, 1927, and through the courtesy of Professor Wesley C. Mitchell and the Rockefeller Institution circulated to a list of economists and statisticians. Subsequently the theory was considerably generalized and a condensed statement of it was published in the *Skandinavisk Aktuarietidskrift,* 1928. The principles developed in this paper have formed the basis of an extensive numerical work on actual and on "manufactured" series which has been going on in my seminar at Yale the last semester. During this work considerable improvements in detail and in the practical adaptation have, of course, been made, but there has been found no ground for a modification of the basic principles involved.

In the analysis of time series we may, roughly speaking, consider the following four groups of problems:

(1) The decomposition of a given time series. We want to find out on more or less empirical grounds *what is actually present in the series at hand*, that is to say, what sort of components the series contains.

(2) The comparison between different series. We want to compare a certain component in one series with the corresponding component in another series, or more generally, we want to compare a set of components in one series with certain components in other series.

(3) The explanation problem. When we have found that a given series contains certain components, we ask the further question: *How did these things come into the series?* In a sense, this is the crucial question of time series analysis. It is only the answer to this question that can give the ultimate significance test for the observed components. But answering such a question means working out a whole rational explanation of the phenomenon at hand. This is not a question of time series technique any more, but a question of the whole content of the *theory* of the particular phenomenon at hand. Even before such a theory is worked out it may be fruitful to try to give some answer to the simpler question: what is actually present in the series? Time and again we have seen in the natural sciences, as well as in the social sciences, that such a more or less empirical attack on the problem has suggested a new and fruitful starting point for the theoretical research. This simpler, technical question is the decomposition problem, listed above as problem No. 1.

(4) Finally, we have the problem of forecasting, which, if it shall have anything to do with science at all, must be based on a thorough understanding of all the foregoing three questions.

It is only the first of these four problems, namely, the decomposition problem, that shall be considered in the present paper. So far I have purposely avoided any attempt to enter into the field of forecasting. I believe that no systematic and reliable forecasting will be possible until we have obtained more knowledge about the real nature of all the various sorts of cycles and trends whose composite effect is shown in our time series.

In my approach to the decomposition problem there are in particular two things I have had in view. First, I have wanted to develop a method that is *more flexible* than the usual methods of curve fitting. To take an example: a compound interest curve, a polynomial or some other specific formula fitted as a long time trend to an economic time series, will, as a rule, give a good fit for a while and then it will shoot entirely out of the data. In such a case it is customary to speak of a "break in the trend." To me, most of these cases are rather examples of a breakdown in the trend *method* than a real break in the trend of the data. It is true that there do occur cases where a real break in trend takes place, but these cases are extremely rare. In most cases the so-called break in trend is only an apparent thing due to the artificial rigidity of the method. What we need in order to take care of situations of this kind is a method which has the same sort of flexibility features as a moving average, but which is more refined, and furthermore constructed so as to deal with the complicated situations that arise when our time series represents the composite effect of several components.

Second, comparing our desiderata with what we obtain by the periodogram analysis, I have wanted to develop a method which is such that it will actually give components *that we can see.* In the periodogram analysis each cycle is represented by a sort of index number, namely, the magnitude of the Fourier coefficient of the corresponding trial period. If this index is high, we take it as indicating that this sort of cycle is strongly present, and if the index is small, we consider it as indicating a lack of evidence of the presence of this sort of cycle. Of course, we may also by the classical harmonic analysis procedure determine the phase of the cycle in question and from a trigonometric table plot a sine curve with the specified properties. But this is not tracing the component in the sense which I have in mind. I am thinking of a procedure which would make it possible to trace a given component *in its actual historical course* so that we can compare a given historical swing in the component in question with the next swing in the same component. In many sorts of data, and particularly in economic data, it is quite obvious that the cyclical character of a given component is not constant. We may, for instance, have a component that is changing with respect to the length of the period. In the work done in my seminar at Yale we found, for instance, a characteristic lengthening in the zero distance of the "40 month"-cycle in the beginning of the war period. This is very reasonable, I think, on a priori grounds: the unsettlement occurring with the beginning of the War, moratoria, etc., made the business men adopt a policy of "wait and see" which must necessarily have lengthened the zero distance in question. In other cases we may have a change in the phase of the cycle or a change in the amplitude, and so on. And these things are exactly some of the most interesting and fundamental things to study. But, it is impossible to modify the periodogram analysis into a truly *moving* method that would permit such a continuous historical observation of each component, because the periodogram gives significant information only when the range for which it is constructed covers a great number of the cycles in question.[1]

A method by which it shall be possible to trace the historical course of each component must necessarily to a large extent be based on the local properties of the given curve instead of on its total properties. The nature of the cycles and the trends and other components will have to be determined in each point of time primarily by taking account of the properties of the curve in the vicinity of this point, that is to say, by taking account of the slope, convexity, etc. Of course, in an actual case it is impossible to carry this principle to the extreme. We must make a compromise and seek our information not only in the strictly infinitesimal vicinity of the point considered, but also seek it at some short distance from this point. We have to adopt a practical interpretation of the notion of "locality." For instance, we have to revert to finite differences instead of true differentials. And if the erratic element is heavy, we have to perform a mechanical smoothing, or in some other equivalent way modify the operation performed on the series so as to make it extend over a certain length of time. In this respect the operations developed below are perfectly general. According to the nature of the problem at hand they may be squeezed in on a short interval of time or extended to a longer interval. They may, for instance, be extended over a very long interval and constructed in such a way as to give the same sort of information that the periodogram analysis can give.

The principal tools used in my approach to the decomposition problem are *linear* operations. By a linear operation in this connection I simply mean a moving total with constant (positive or negative) weights. If $u(t)$ is a function of time which is known in a set of equidistant points, and if $\Omega_1, \Omega_2 \ldots \Omega_m$ are constant weights independent of time, the linear operation Ω is defined by

$$\Omega u(t) = \sum_{i=1}^{m} \Omega_i . u\left(t + \left(i - \frac{m+1}{2}\right)\delta\right)$$

where δ is the distance between consecutive observations. In the general case no assumption needs to be made about the symmetry of the weights. As a rule, however, it will be found convenient to consider symmetric weights, that is, weights such that $\Omega_i = \Omega_{m-i+1}$.

The use of such linear operations in the study of time series is, of

[1] I am speaking here of the classical form of periodogram analysis. Professor Wiener of the Massachusetts Institute of Technology has told me that he has recently developed a generalized periodogram analysis (*Acta Mathematica*, 1930) which will overcome these difficulties. I have not yet had an opportunity to study Wiener's method carefully, but I have some doubt as to the possibility of actually tracing the historical course of each component by constructing a periodogram. I shall revert to this question in my forthcoming monograph on the decomposition of empirical series.

course, in itself not new. I believe, however, that the systematic way in which I try to utilize them is somewhat novel. I try, for instance, to connect these linear operations with the very definition of the notion of a "component" in a given time series. If no assumption whatsoever is made regarding the nature of the components, then the decomposition problem has no sense. In this case we may postulate the existence of n perfectly arbitrary "components" and make this fit in with the given series, simply by determining an $(n+1)$th component equal to the deviation of the given series from the cumulated effect of the n postulated components. The assumptions by which I give a meaning to the notion of "component" are built on the idea that each component shall represent something oscillating around, or departing for good, from a point of equilibrium. By formulating this assumption in terms of the approximate effect which a linear operation will have on such a component, certain rules for the use of these linear operations in order to *determine* the components are developed.

One of the problems studied in this connection is the general amplification problem. I study certain types of operations Ω that will knock out progressive components *with a small convexity*, and other operations that will knock out small wave-like components, and at the same time amplify cyclical components with periods falling within a certain more or less definitely defined range. In many cases, this process is already quite sufficient to isolate a given sort of cyclical component and exhibit its historical course. In other cases there may be two cycles that are too similar with respect to wave length to make it possible to isolate them one at a time by such a simple amplification. In this case the problem may be attacked by certain propositions obtained from the study of the effect on the composite graph produced by "twin cycles." Or the problem may be attacked by using a certain sort of linear operation which we may call the *key operations*. These are certain simple linear operations that are iterated a number of times and from the results of which is formed a certain algebraic equation, the key equation, whose roots will give information about the length of the periods in the case of cycles or of the yearly progressivity in case of non-cyclical trends. From the knowledge of these characteristics the conclusion as to the other characteristics of the components is easy. This analysis may be carried through either on a local basis, determining the characteristics of the components separately in each point of time and thus obtaining a moving determination of the components, or it may be carried through on the basis of the total properties of the curve. The key equation procedure may also be extended to the simultaneous determination of three or more components.

The first to utilize the idea of a key equation seem to be Fr. Kühnen and H. Bruns.[1] The key equation processes of these earlier writers, however, were not built on the notion of a general linear operation, but on a particular form of such operations, namely, either successive differences or successive point selections. Therefore their theory could not be based on a systematic manipulation of the key operations and of the amplification operations so as to have the equation work under the most favorable circumstances, which is the essence of my approach. This, I believe, explains why their method never had the success of being adopted in practical work to any large extent.

One aspect of the problem to which I attach great importance is what might be called the *Slutsky effect.* This is the fact that linear operations applied to a random variable may produce fluctuations of a more or less cyclical character. I study the laws of such spurious cycle corrections and show in particular that it is possible to construct operations Ω which, when applied to a random variable, will produce nearly rigorous sine curves, the period and amplitude (but not the phase) of which can be predicted almost exactly when the operation Ω is known. I call this the Slutsky effect because I believe the Russian statistician, Eugen Slutsky, was the first to take up this effect in a systematic study. My results go somewhat further than Slutsky's. In particular, I derive coefficients by which the amplitude of observed cycles are compared with the amplitude of spurious cycles created through the Slutsky effect. The knowledge of the laws of spurious cycle creation thus obtained may be combined with the key equation approach in a manner which goes a long way toward eliminating the spurious cycles. The procedure simply consists in, so to speak, setting aside one root of the key equation to take up the spurious effect. This has proved rather effective, particularly if the key equation is formed on a local basis.

If the cyclical components in the series studied are not too similar with respect to wave length, certain phases of the procedure here indicated may be worked out graphically, utilizing the inflection points of the given curve as it appears after a graphical smoothing. In this way, a person trained in graphical smoothing may perform a very rapid rough analysis of what the given curve contains. I have received information telling that this graphical process in the form explained in my first paper has been applied also in other quarters with satisfactory results.

[1] Fr. Kühnen, *Astronomische Nachrichten*, 1909; H. Bruns, *ibid.*, 1911. Later the idea has been used by J. I. Craig, *Monthly Notices of the Royal Astronomical Society*, 1916, and G. Y. Yule, *Philosophical Transactions of the Royal Society of London*, 1927.