INTRODUCTION

The estimation of the power spectrum of a stationary time series from partial knowledge of its autocorrelation function has been a classical problem [11] every since Norbert Wiener [5] showed the fourier transform relationship between the power spectrum and the autocorrelation function. Over the years, hundreds of technical papers, books, seminars and courses have been concerned with the particularly practical problem of estimating the power spectrum of a sampled, band-limited, stationary time series, where the information that is known about the time series is the autocorrelation values, $R(t)$, $|t| \leq N$. Almost all of this research has been based on the use of window functions, whose properties can be easily analyzed by use of the fourier transform convolutional theorem. While readily understood, this conventional window function approach produces spectral estimates which are negative and/or spectral estimates which do not agree with their known autocorrelation values. These two affronts to common sense were the main reasons for the development of maximum entropy spectral analysis, which is the subject of this thesis.

The conventional approach to estimating the power spectrum from $R(t)$, $|t| \leq N$, is to assume that $R(t)=0$ for $|t|>N$ and to take the fourier transform of $W(t)R(t)$, $|t| \leq N$, where $W(t)$ is a weighting function. As can be seen, this is conceptually equivalent to taking the fourier transform of $W(t)R(t)$, $(-\infty<t<\infty)$, where $W(t)=0$ for $|t|>N$. In fourier transform theory, it is well known that the fourier transform of a product of two time functions is equal to the convolution of their fourier transforms in the frequency domain. This makes the conventional estimation procedure mathematically elegant since the estimated spectrum is thus equal to the true spectrum convolved by the fourier transform of the weighting function. Much research has been done on the properties of various weighting functions and their corresponding window functions. The selection of a window function involves a compromise between resolution in the final spectrum and the extent to which any one spectral component contaminates the estimates of other nearby components. In addition to these opposing requirements of making the main peak of the window function narrow but minimizing the size of the side lobes of the function, the further condition is sometimes imposed that the resulting spectral estimate be non-negative. While window theory is interesting, it is actually a problem that has been artificially induced into the estimation problem by the assumption that $R(t)=0$ for $|t|>N$ and by the willingness to change perfectly good data by the weighting function. If one were not blinded by the mathematical elegance of the conventional approach, making unfounded assumptions as to the values of unmeasured data and changing the data values that one knows would be totally unacceptable from a common sense and, hopefully, from a scientific point of view. To overcome these problems, it is clear that a completely different philosophical approach to spectral analysis is required.

Given the specific knowledge contained in the limited set of autocorrelation values, plus the general knowledge that power spectra are non-negative, we know that there is normally an infinite number of spectra in agreement with this information. Thus, without additional knowledge, finding the true spectrum, or, alternatively, displaying all possible spectra, is a hopeless task. What is a reasonable goal is to find a single function, $P(f)$, which will be respesentative of the class of all possible spectra.

Outside of requiring P(f) to be a member of the class of possible spectra, i.e., to be in agreement with all the known facts, the basis for determining a most representative member is highly subjective. However, to resolve the problem, some choice must be made. Maximum entropy spectral analysis is based on choosing the spectrum which corresponds to the most random or the most unpredictable time series whose autocorrelation function agrees with the known values. This assumption, which corresponds to the concept of maximum entropy as used in both statistical mechanics and information theory, will be shown to be maximally non-committal with regard to the unknown values of the autocorrelation function.

Besides giving us a spectral estimate which is a non-negative function of frequency and which agrees with the autocorrelation values, maximum entropy spectral analysis has many other desireable and surprising attributes. The first of these is that the resolution of the maximum entropy spectral estimate is greater, and at times much greater, than those obtained from conventional estimates. Since the most random time series has a white or flat spectrum, one would expect the maximum entropy assumption to produce the maximally white spectrum that is consistant with the autocorrelation values. In a sense, this does happen, but very often the "most white" spectrum has to have sharp spectral peaks of much higher resolution than given by the conventional spectral estimates. After all, the conventional estimates do not agree with the data and may not be constrained to be non-negative. Thus, why should they give sharp peaks when the true autocorrelation values require them? In addition, the side lobe problem that was caused by window functions is absent in maximum entropy spectral estimates. Thus, spectral peaks that are close together in frequency can be resolved by the new method.

The second main attribute is that maximum entropy spectra can be calculated from the $R(t)$, $|t| \leq N$, values in little more time than is required for conventional spectral estimates. Thus, maximum entropy spectral analysis is a highly practical theory.

Finally, the third major attribute of the new philosophy is that it has opened the door to the development of many advanced analysis procedures involving multichannel and multidimensional spectra. One test of a good theory has always been the number of new fields of investigation that it produces. A particularly important example is the development of the Burg technique which is not needed or of any real importance to conventional spectral analysis. These developments can be contrasted with the sterility of the conventional methods which have not advanced in the last twenty years.

This thesis is conceptually divided into three major parts. The first and largest is concerned with single channel spectral analysis from autocorrelation information. The single channel theory has been quite thoroughly developed both in theory and in practice. All of the important theorems of the single channel theory are presented and proven, often more than once in different contexts. In addition, in II-C, the philosophy behind the highly important Burg technique is explained and the mathematics presented in a generally useful form. Finally, in section II-D, it is argued that the most important description of the second order statistics of a stationary time series is in terms of reflection coefficients, both from a practical and a theoretical point of view. It is seriously suggested that the old Wiener definition of the power spectrum should be replaced by a new definition involving maximum entropy spectra. In appendix A, an interesting theory of pure line spectra is presented as a limiting case of maximum entropy spectra.

The second major part of the thesis deals with multichannel spectra. Again, all of the important theorems are presented but in less detail than in the single channel case. In general, multichannel spectra do not have the necessary structure to permit the correspondingly rich development as occurs in the single channel case. Because of this, no simple extension of the Burg technique to the multichannel case is possible. A particularly noteworthy proof is the one showing that the multichannel prediction error filter is minimum phase.

The final major portion of the thesis is Chapter IV where the maximum entropy philosophy is considered to be a special case of the general variational principle approach to estimation of functions. A general solution technique is discussed and philosophical comments are presented concerning the consistency and usefulness of measurements.

A quick flip through this thesis will show that it is not a self contained work but requires that the reader be fairly proficient in time series analysis. Undoubtedly the best source for background material for the single channel theory is Jon F. Claerbout's book, Fundamentals of Geophysical Data Processing, [1]. Background material for the multichannel theory can be found in Enders A. Robinson's book, Multichannel Time Series Analysis with Digital Computer Programs, [2].

One may note that the thesis is written entirely in terms of complex time series. The reason for making this choice was that the more general theory is not appreciably harder and does in fact make the theory of reflection coefficients much more complete. Furthermore, there are many practical problems which need the complex theory.