THE INCORPORATION OF PRIOR INFORMATION IN INTERPOLATION AND EXTRAPOLATION ALGORITHMS AS APPLIED TO AUTOMATED TESTING SYSTEMS

by

Frederic Herman

A DISSERTATION

Submitted to the Faculty of the Stevens Institute of Technology in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Frederic Herman
Frederic Herman, Candidate

ADVISORY COMMITTEE

Thaddeus J. Kabylar, Chairman

Gerald Gerstenfeld

Paul M. Chirlian

Howard D. Helm

STEVENS INSTITUTE OF TECHNOLOGY
Castle Point, Hoboken, New Jersey
1975

ACKNOWLEDGMENT

My thanks go to the members of my thesis advisory committee, who have made numerous suggestions that have improved both the clarity of the exposition and the style of writing. I wish to especially thank Dr. T. J. Kobylarz for his patience and insight in advising me in my research which is presented in the thesis.

My special thanks go to my wife, Joan, for helping me through the difficult periods of my doctorate program, and for the long hours of typing the many drafts of the thesis.

TABLE OF CONTENTS

THE INCORPORATION OF PRIOR INFORMATION IN INTERPOLATION AND EXTRAPOLATION ALGORITHMS AS APPLIED TO AUTOMATED TESTING SYSTEMS

by

Frederic Herman

Thaddeus J. Kobylarz
Chairman, Advisory Committee

1975

ABSTRACT

This thesis is concerned with the determination of interpolation and extrapolation procedures for automated testing systems. The principle motive for considering such research is that of reducing the number of measurements required to characterize a particular device function. The approach taken is to use prior test data from similar components of a device type. This information is incorporated into the interpolation or extrapolation procedures either in the form of tolerance intervals (TI's) replacing deterministic data, or else is used to adapt the procedure to the deterministic measurements made on the device being tested.

The type of device functions considered in the thesis are restricted to a continuous, nonlinear, instantaneous category. A further restriction is that certain derivatives

of the functions must also be continuous. Some situations allow the relaxation of the instantaneous requirement.

The interpolation formulas are seen to represent linear combinations of random variables (RV's). A basic problem arises concerning the need to combine TI's of the RV's. A "bounding algorithm" (BA) is presented which avoids the analytic convolution and which forms a statistical bound directly from the individual RV TI's. A property called "linear conformity" is introduced in order to define a measure of how conservative this composite bound is. An investigation using the Fourier transform examines conditions under which linear conformity exists. The analysis assumes that the combined RV's are statistically independent. Several theorems are developed which consider the effect of this assumption on the bounds formed when RV's are dependent.

A family of probability density functions (pdf's) called the "exponential transform family" is described. Several theorems are proved regarding the BA bounds on a linear combination of RV's having these pdf's. A heuristic classification procedure is developed which enables one to establish the linear conformity property for arbitrary pdf transforms under certain conditions.

In a second approach, prior measurements of devices are organized as patterns of information. The patterns are used

to determine the coefficients of an iterative extrapolation formula via a training procedure. The resulting extrapolator is at least as accurate or better than corresponding conventional extrapolators, if the set of training patterns is representative of the functions to be subsequently extrapolated. Convergence of the training procedure is proven for consistent training patterns.

1. INTRODUCTION

With the rapid transition from discrete to monolithic circuits, the problem of testing devices[1] is becoming severely difficult for manufacturers and users of these devices. A significant percentage of these devices must be tested[2] in a continuous, nonlinear, instantaneous mode. While circuit probing has been a dominant testing technique for discrete circuits, it is obviously unfeasible for testing integrated circuits (IC's) because of the extremely small size of circuit elements. One is thus faced with testing a highly complex black box device through terminal excitations and measurements. Excitations may also involve environmental variables such as temperature and humidity. Static testing of a few combinations of inputs cannot be expected to suffice for situations where the functional behavior of the device is critical. It may not be easy to simulate the exact system environment that the device will be subjected to or this environment may vary over a wide range of conditions. Non-electrical inputs such as temperature and other mechanical

---

[1]The word "device" is intended to mean a physical system that is not readily divisible into constituent parts due to the way it is manufactured. Such devices can be very complex systems. Although the devices referred to in this thesis are that of monolithic integrated circuits, it will be seen that mechanical and chemical systems could also have been considered.

[2]The word "test" is intended to mean the determination of a measurable response of the device to an excitation.

stresses may cause significant functional variations which must be known for design.

In order to economically test continuous IC's, automated testing systems have become a necessity. Fortunately, the increased complexity of monolithic devices which has compounded the testing problem has also established the potential for a solution. Large scale integration of digital electronic circuits has allowed data processing equipment to shrink both in size and price. But while the cost of making measurements decreases, the number of measurements per device required to adequately test continuous devices increases.

It is clear that exhaustive discrete measurements of a continuous, nonlinear device's input-output function is not possible since an infinite number of input values exist even for a scalar input. In order to characterize this function, it is necessary to interpolate or extrapolate the measurements that are made. Also, one may be restricted in the range of the input variable that can be generated for the testing of the device. Sometimes this restriction comes about because the device is altered or destroyed by subjecting it to excessive voltages, currents, or temperatures. Frequently it is desirable to be able to test devices in the field using a testing system much simpler than that available in the laboratory. Another important limitation in simulating input values occurs in life and storage testing. Here, one would like to extrapolate in time to predict the

behavior of the particular device based on what has been observed in the past.

This theoretical thesis is concerned with the application of interpolation and extrapolation procedures to automated testing systems. An automated testing system is often required to make decisions as to the reliability of the device in its planned mode of operation, based on sample measurements. The principle problem to be considered is that of reducing the number of measurements required to characterize a particular device through the use of interpolation and extrapolation algorithms. A primary aspect of the approach taken is to use prior test data from similar components of a device type. This information will be incorporated into the interpolation or extrapolation procedures either as statistical data, or will be used to adapt the procedure itself to the deterministic measurements made on the device being tested.

It must be stressed from the beginning that the approaches considered will always require an initial concerted measurement effort. In gathering data to be organized as statistical estimates or to be used in training procedures, for the approaches in the thesis, one must be prepared to measure many components of a device type and measure each component's characteristics much more extensively than would be required just to determine the reliability of these components. The savings of such information gathering

programs can be realized when hundreds of thousands of components of the same device type are manufactured and are subsequently tested using this a priori information. Additional savings accrue when one is also able to simplify the automated testor used by the vendors and purchasers of devices. Simplified field testing can also be important when testing must be performed by personnel under critical military or aerospace situations.

The type of input-output function considered in this thesis will be restricted to a single-valued, continuous, nonlinear, instantaneous function. This device function may be further restricted in that sometimes the derivatives of the function may also be required to be continuous. When this additional restriction is required, a statement to this fact will be made.

The requirement that the device function be instantaneous is of major consequence. By requiring that the output be independent of past inputs, there are many interesting devices that must be excluded. Some of these include magnetic memory devices which are often characterized by hysteresis curves. It is stressed that the reason for excluding such behavior as hysteresis is that such effects add an additional level of complexity to the already difficult study of automated testing. It is expected that the techniques to be developed in this thesis can be extended to dynamic input-output functions, but that such an extension

must be deferred as a future area of research. In some specific applications, one may be able to transform a dynamic response function into an instantaneous function. As an example, one may measure a particular characteristic of a hysteresis loop as a function of temperature. This characteristic is measured once per cycle, and may be an instantaneous function of the temperature.

As a final consideration of the requirement that the device function be instantaneous, it is noted that time may be considered an input. When one considers the long term storage effects on devices, for example, time becomes the principle independent variable in the device function. It will be assumed, however, that the effects of time will only be important over periods many orders of magnitude greater than the time required to make a single measurement on a device.

In describing the particular problem area considered in this thesis and the new results obtained, it is appropriate to indicate the organization of the thesis and the content of each of the chapters. In the following chapter, several segments of the literature will be reviewed which relate to those areas specifically considered within developments of the thesis, and to the general problem of device testing. It is the dual purpose of this review to acquaint the reader with a general background to the problem of automated testing of continuous, instantaneous, nonlinear devices, and also to

indicate the alternative approaches considered by the author. A part of the chapter is concerned with establishing the importance of the principle problem area considered in the thesis. This area concerns the economic limitation on the number of measurements that can be made on a device. Specifically, one wants to minimize the number of measurements made on a given device that is required to characterize this device.

The literature review considers the various existing techniques of characterizing the performance of a device. Each of the areas reviewed are discussed with respect to their relevance to device testing. A conclusion made is that conventional extrapolation and interpolation techniques possess desirable features for automated testing, but that they suffer from an important deficiency. This deficiency concerns the failure of the conventional techniques in utilizing information on prior devices tested. In order to overcome this deficiency, two general approaches are suggested. One approach uses conventional interpolation formulas and prior information organized as statistical distributions. Tolerance limits obtained using the distributions are substituted for data normally obtained by direct measurement of the device under test. The second general approach is to use prior information to obtain extrapolation procedures adapted to the device function to be approximated. A training procedure is suggested which evolves from

considerations of pattern recognition theory. The adaptive extrapolation techniques are particularly useful for the problem of long term reliability prediction for devices. This application area is reviewed in order to determine alternative approaches that exist in the literature.

Comparison between the new results in the thesis and existing methods are explained in more detail throughout the thesis. When a parallel approach exists to a development in a chapter, the differences and similarities will be pointed out to the reader.

The third chapter is devoted to preliminary considerations of the problem of combining prior information with measurements made on a particular device. The material in this chapter provides the reader with a background to the initial approaches considered by the author. The chapter contains three different topics which represent potential starting points for later results. The first topic is concerned with fundamental techniques which combine deterministic device function measurements with "statistical data." Because measurements can represent statistical estimates of physical quantities rather than the quantities themselves, it is necessary to define what is meant by the term statistical data. If a great number of similar devices are measured at the same input conditions, these measurements may be organized as a distribution of a statistical quantity. This quantity is the device function evaluated at the

particular input conditions for which the measurements were obtained. In this thesis, the term "statistical data" will always mean the value or the range of values of the device function obtained from a statistical distribution. The term "deterministic data" will always mean the value obtained by direct measurement made on a particular device.

The results of the initial topic is that interpolation and extrapolation procedures are developed which combine deterministic measurements of a device function with statistical data which is in the form of bounds on the device function first derivative. Rather than determining an approximation to the device function, statistical bounds are obtained between which the device function is said to lie with a particular probability. This approach is not a polynomial approximation and therefore assumes no particular order of approximation. The limitations, assumptions, and possible practical applications are discussed.

This preliminary topic introduces a basic problem. The problem concerns the need to linearly combine tolerance limits of random variables. By assuming that the random variables are independent in the statistical sense, the complexity of the combination problem is reduced. However, there still is a need to simplify the process in which the bounds on the random variables can be combined. A particular combining algorithm is described which was used in the preliminary topic described above. The second topic of the

chapter considers alternative algorithms with respect to their suitability for the statistical interpolation techniques first discussed. Justification for the algorithm used is deferred until the fourth chapter, since an extensive mathematical analysis must be developed.

The last topic of the third chapter concerns a special aspect of device testing. The problem relates to the need to determine the variations of process parameter effects over the area of a monolithic circuit. Due to the extremely small size of the circuit elements and interconnections, it is economically unfeasible to make parametric measurements at points other than those specifically designed for this purpose. With the increasing amount of integration in these monolithic circuits, it becomes increasingly more unlikely that a few test points can satisfactorily represent the process parameter effects over the entire circuit. The approach taken is to combine prior statistical distributions of spatial variations with the few measurements that can be made on these circuits. Conventional Bayesian analysis techniques are employed. The novelty of the approach is in the way that the prior statistics are gathered. A "test wafer" is described which enables a distribution of parameter variations to be obtained. While this last topic is somewhat different from the remaining developments of the thesis, it is included because of its extreme importance in the testing of integrated circuits. It is also noted that the topic

addresses the principle problem studied in this thesis. It enables a reduction in the number of deterministic measurements that must be made on a particular device to satisfactorily characterize it.

The fourth chapter is devoted to obtaining bounds on linear combinations of independent random variables. The need for these bounds has already been indicated in the discussion of the earlier chapter on preliminary results. Further need for this analysis results whenever several deterministic data are replaced with statistical data in the form of tolerance limits. The first part of the chapter formalizes the algorithm which is used to combine the statistical data bounds for individual random variables into composite bounds. The algorithm provides a simple means of obtaining a bounds on the linear sum of random variables. It is assumed that the individual random variables are statistically independent.[1] Following this definition, a detailed mathematical analysis is performed in which the effect of combining several similarly distributed random variables is investigated. Fourier analysis is the vehicle for this analysis. Several theorems are stated and proven which relate to properties of the distribution which corresponds to linear sums of the random variables. The major

[1] In Chapter 3, it is shown that a conservative analysis will usually result, when the random variables are assumed statistically independent when, in fact, they are dependent.

result of the chapter is the determination of several conditions on the distributions of the individual random variables which allow the use of the bounding algorithm. Under these conditions, the probability corresponding to the tolerance limits of the individual random variables also corresponds to the tolerance limits obtained using the bounding algorithm, as a conservative approximation. That is, the probability for the composite bounds is at least that corresponding to the individual random variable tolerance limits. A family of distributions is shown to allow such a combination of tolerance limits. Several heuristic procedures are developed which are useful in considering arbitrary probability distributions with respect to use of the bounding algorithm. The advantages and limitations of the results of this chapter are discussed with respect to practical considerations.

Following this analysis, Chapter 5 contains techniques which allow the incorporation of statistical data in polynomial and transcendental curve fitting. The conventional techniques of exact-fit interpolation and extrapolation possess certain desirable properties. These properties are that the techniques are easily mechanized by simple automated testing systems, and that there exists a great deal of theory related to practical application of these techniques. Conventional exact-fit approximations do not normally utilize any information other than deterministic

data. This chapter treats the situation wherein statistical data is substituted for deterministic data. A consequence of this incorporation of statistical data is the reduction in the required number of measurements on a device that can characterize the device function. The statistical data that is substituted is in the form of tolerance limits and corresponds to a tolerance level or probability. The use of this statistical data results in a statistical bounding of the device function with a corresponding tolerance level. The general approach taken in this chapter is seen to follow from the third chapter concerning preliminary considerations. The assumptions made in Chapter 5 are examined with respect to situations in which they can be expected to be physically valid. Practical applications of the techniques developed in the chapter are also discussed.

In Chapter 6, an alternative approach to combining prior information with deterministic measurements is taken. This approach evolves from a particular class of testing problem, long term reliability prediction. This problem is seen to be an extrapolation problem in which one must use device measurements usually gathered at regular intervals of time, to approximate the device function behavior at some future point in time.

Conventional extrapolation formulas are usually insensitive to the type of function being approximated. While the type of function governs the order of the approximation

used, the coefficients that multiply the data are determined independently of the data values. In this chapter a different approach is taken. The data are considered to represent patterns of information. By using patterns obtained via measurements on many devices, the extrapolation coefficients are determined with a training procedure. This adaptive process results in an extrapolation formula which is more accurate for the patterns of data used to obtain it. If this so-called "training set" is representative of the device under test, the extrapolator is seen to be more accurate than conventional techniques.

The techniques of this chapter are obtained utilizing pattern recognition theory as their vehicle. Other attempts to apply pattern recognition theory to the problem of long term reliability prediction are compared to the results of this chapter.

Chapter 6 also considers the implementation of the adaptive extrapolation techniques. A simulator is described which performs iterative extrapolation of a device function. Given an initial sequence of N consecutive values of the device function, the simulator computes subsequent values of the function corresponding to regular time intervals. A second system is used in conjunction with the simulator. This second system, called a recognizer, determines when the simulated output of the simulator corresponds to a predefined failure mode. Several approaches to the design of

the recognizer are discussed. It will be seen that the design follows from existing pattern recognition techniques.

An important feature of the adaptive extrapolation is that it is immediately extended to multidimensional functions. Conventional extrapolation suffers when more than one dimensional functions are considered because of the rapidlw growing algorithm complexity. The adaptive extrapolator is shown to avoid some of the pitfalls of existing methods.

The concluding chapter of the thesis summarizes recommendations for further research. As a final note of this introductory chapter, it is pointed out that although the primary area of application is that of testing monolithic electronic devices, the results of this thesis apply to other types of systems as well. From time to time, reference will be made to IC's in the development of testing procedures, but it should be clear that mechanical and chemical systems could also have been considered.

2. LITERATURE REVIEW

2.1 Chapter Introduction This thesis is concerned with the theory for applying interpolation and extrapolation procedures to automated testing systems. The reason for considering this problem is to reduce the number of measurements required to characterize a particular device through the use of interpolation and extrapolation algorithms. A primary aspect of the approach taken is to use prior test data from similar components of a device type. Such information is incorporated into the interpolation or extrapolation procedures either as statistical data,[1] or it is used to adapt an algorithm to the deterministic measurements made on the device being tested. Initially, a review of literature directly related to the thesis problem area is made. Included in this review are papers that contain results directly applicable to developments in the thesis. Existing methods for characterizing a device function are reviewed in order to indicate the need for the new approaches taken in later chapters.

The succeeding section reviews several topics that are expected to provide the reader with a background useful in understanding the thesis developments, although these topics

---

[1] In this thesis, the term "statistical data" means the value or range of values of the device function obtained from a statistical distribution.

are not directly relevant to the thesis problem area. That is, these topics do not directly apply to interpolation and extrapolation.

The final section expounds on the conclusion that the reviewed literature fails to cite the use of prior information for interpolation and extrapolation. In order to overcome this deficiency, two general approaches are suggested in the section. The developments of these approaches represent the major work undertaken in the thesis.

2.2 Review of Literature Related to the Thesis Problem Area

In this section, literature is reviewed which directly relates to developments in later chapters. The relationships between these topics and the thesis area of investigation are also explained.

Since one concern of this thesis is with the incorporation of statistical information for automated testing system interpolation and extrapolation procedures, it is useful to begin the section with the consideration of how statistics are usually employed in automated testing. A review of texts on engineering statistics such as Bowker and Lieberman [1964], Guttman and Wilks [1967], and Mood and Graybill [1963] reveals that research in production testing techniques has traditionally been concentrated in the area of sampling statistics. Sampling statistics is concerned with limiting the number of devices or objects

inspected that is required to characterize the entire population from which the sample devices are drawn. In contrast, the problem considered in this thesis is primarily concerned with limiting the number of measurements required to characterize the performance of a particular device. One may conclude from this comparison that the area of sampling statistics is of limited use in the proposed area of investigation. The techniques of sampling statistics are more directly relevant to the means of obtaining statistics required as a priori information for the thesis developments. The only direct application of sampling concepts made in the thesis corresponds to the special problem in the next chapter. This special problem can be thought of as a sampling problem in that measurements are made at the sub-device level.

The developments of this thesis share a common feature in that they combine deterministic measurements on a particular device with statistical information gathered for devices of the same device type. A branch of statistical methods which is concerned with such a combination is called Bayesian analysis (Morgan [1968]). Several formulations, using Bayes' rule, as a basis, can be made. For example, a parameter θ of a probability distribution for observing some physical quantity may be unknown, but one may have some information about the approximate probability

that θ equals θ_1 (i=1,...n). Call this probability $P_\theta(\theta_i)$. Then for a given value of probability distribution parameter θ, one may compute the probability of observing a particular value y_0 of the physical quantity, if all other information about its probability distribution is known. Since this probability is actually conditional with respect to the value of θ_i used, one has the conditional probability

$$P_{y|\theta}(y_0|\theta_i).$$

Then by Bayes' rule,

$$P_{\theta|y}(\theta_i|y_0) = \frac{P_{y|\theta}(y_0|\theta_i)\cdot P_\theta(\theta_i)}{\sum_{j=1}^{n} P_{y|\theta}(y_0|\theta_j)}$$

where $P_{\theta|y}(\theta_i\ y_0)$ is the conditional probability that the parameter value was θ_i given an observation of y equal to y_0. The left-hand side of the above equation represents a posterior probability resulting from combining the information about an observation with the prior probability distribution of θ. Generally Bayesian analysis allows revision of a prior probability distribution. Morgan [1968] develops Bayesian analysis for application to decision processes in which the value of information is incorporated. In a paper by Sage [1969], Bayesian analysis is applied to estimation of parameters of a linear systems model when a measurement error exists. This error is assumed Gaussian.

It is shown that the Bayesian estimate of the parameters have a smaller error variance than that of a least squares estimate. Tsokos [1972] uses Bayesian analysis to determine a posterior distribution of a Weibull failure model in which both the shape and scale factors of the distribution must be estimated. This procedure is used to determine a failure model of a specific device.

Bayesian analysis is thus useful in obtaining improved approximations to a probability distribution associated with some population attribute. It may also be used in a situation analogous to parameter estimations of a device model. Breipohl [1969] presents a tutorial treatment of Bayes rule and how it can be used to develop the well-known Kalman filter which is typically used to recursively estimate the state vectors of a control system. This application of Bayesian analysis is suggestive of a more general area of investigation. One may be concerned with the effect of new observations of an event, such as a system's response, on one's prior information related to the occurrence of this event. Bayesian analysis deals specifically with the effect of events on prior probability distributions (Jaynes [1968]). The effect in this case is the revision of the probability distribution. Bayesian analysis is directly applied to the special problem considered in the next chapter. The other developments of the thesis represent

alternative approaches to Bayesian analysis in their treatment of the more general methods of combining statistical and deterministic information.

Since an important application of the thesis developments is the automated testing of devices to determine if particular devices are acceptable or have failed, it is appropriate to briefly consider fault detection in continuous devices. When the devices are circuits, one may use standard sensitivity analysis[1] and knowledge of the circuit topology, to determine the effect of circuit element variations on the circuit response when a series of "standard" inputs are applied to the circuit. The problem then is to determine a minimum sequence of tests comprised of combinations of the standard inputs which will detect failures of circuit elements. One is reminded that the failures for continuous circuits are generally non-catastrophic. That is, an element failure is a parametric drift beyond some allowable limits. Thus fault detection can be viewed as an inverse mapping of sensitivity analysis. Investigations along these lines have been limited. Seshu and Wasman [1966] used circuit topology and nominal values to obtain tests on linear circuits using frequency domain techniques.

---

[1]See Herskowitz [1968] for examples of sensitivity analysis by computer.

They recognized that the pole and zero locations of a transfer function and the gain factor must be affected by circuit parameters if these parameters are to affect the performance of the system. The circuit topology and nominal values of the circuit elements enabled them to obtain a set of "signatures" of parameter variation. This analysis enabled them to obtain an input-output test sequence in which a frequency response yielded a complete testing of the circuit.

The above technique is only useful if the circuit is linear and if the parameter variations give rise to different changes in the transfer function. The second restriction becomes increasingly more difficult to meet as the size of the system grows. As the corner frequencies become larger in number and closer together, the frequency response must be obtained more and more accurately, requiring more and more measurements to be made. A severe problem with high density integrated circuits is that a large system must be tested with a limited number of test points and that distributed capacitance, inductance, and resistance are not only unavoidable but sometimes are used to the advantage of the design. Thus the above frequency domain analysis suffers from the system complexity, non-discreteness, and often nonlinearity. It might be added that systems are rarely completely linear from the testing

standpoint since one must be concerned with the effect of biases and gain levels (Grossman [1973]). Often this fact is not considered in practical testing (McAleer [1971]). A set of tests are performed at standard conditions and the results are interpreted as being valid over a wide range of environmental conditions that the device is expected to see. This blind use of test results often proves unsatisfactory in high reliability applications.

The literature appears to be lacking in other attempts at fault detection in continuous circuits. The factors which limited the signature approach of Seshu and Wasman [1966] probably make a more general topological approach unattainable. It is noted that interpolation and extrapolation of device functions is an appropriate approach to device characterization primarily when sub-device topological analysis is unfeasible.

Since the thesis is concerned with automated testing system interpolation and extrapolation procedures, an important area of literature review is that of curve fitting techniques. This importance is realized when one considers the usual motivation behind curve-fitting. The data used in the curve-fit is usually assumed to correspond to a deterministic function representing a physical process. The curve-fit results in the determination of an approximating function. From this determined function, one infers

what the physical process, functional values are under conditions different from those corresponding to the curve-fit data. That is, the curve-fit enables one to interpolate between the data points.

With curve-fitting techniques, a standard functional form may be used to approximate a device response function, and the problem is that of determining the constants of the functional form (Hamming [1962]). Normally, one uses data obtained from direct measurement on the device in the curve-fitting scheme. This data provides the information that determines the function constants. One usually does not think of the curve-fitting formulation as being dependent on the data. Such considerations as whether to use an exact-fit or a least squares fit, what order fit to use, and whether a polynomial formulation is appropriate at all, implicitly follow from known or expected properties of the actual device function (Hamming [1962]). Additionally, the presence of noise in the measurements is one of a number of factors that influence the choice of the curve-fit techniques.

When the data used for a curve-fit contains a large random measurement error, often referred to as "noise", it is appropriate to determine a curve that will minimize the effect of the noise while representing the data. Polynomial curve-fitting techniques which minimize the sum of the squares of the deviations of data from their estimate are

referred to as least squares approximations (Ralston [1965]). There are other approaches that deal with random error, but the ultimate choice centers on the question of how well the approximation works in practice. In determining the order of the approximating polynomial, the criteria of minimizing the least squares error is often used. It is noted that this error represents the deviation of the approximating function from the data, but this error does not indicate the performance of the approximation between data points. In choosing the order, Kussmaul [1969] shows that when a measurement error of zero mean and fixed variance exists, it is always better to use a higher rather than a lower order polynomial approximation with respect to the actual order.

While least square approximations are well suited to noisy data, exact-fit approximations have distinct advantages. Each fit polynomial approximations are easy to compute for a number of reasons. Since the coefficients are determined by closed-form equations, the inversion of matrices normally required by least squares is avoided. Hamming [1962] indicates that these matrices are often ill-conditioned in that the determinant of the matrix is often quite close to zero. Use of orthogonal functions may be a means of avoiding this problem, but Hamming points out that, "experience shows that if the orthogonalization is

attempted by means of the Gram-Schmidt process, then the same difficulty arises in a different disguise."

Exact-fit approximations are particularly suited to automatic testing systems when the positioning of the data is predetermined (Hamming [1962]). In this case the coefficients of the approximating polynomial can be precomputed since they are independent of the data values. As Hamming points out, the primary disadvantage of the exact-fit polynomial is that as the order of the fit increases, the polynomials become increasingly more oscillatory between data points. Since the primary purpose of the approximation is to allow interpolation between data points, it is often necessary to avoid high order polynomial curve-fitting. The most simple way is to break up the total interval over which the function must be approximated into subintervals and consider each subinterval separately. This type of approach is referred to as a piece-wise interpolation approximation. A more involved piece-wise formulation makes use of spline functions. Greville [1969] provides a tutorial introduction to these functions, indicating that they result in a "best" approximation according to practice and, in part, theory. This implied optimality turns out to be analogous to least squares approximation optimality, except that the criteria is now that of a mean square error minimization between the spline and the actual function. A

problem arises in the solution of the spline coefficients very much like that discussed for least squares regression. Schumaker [1969] gives a partial compilation of some algorithms for the determination of the unknown coefficients of particular types of splines. He reports that the matrices to be evaluated are "ill-conditioned." In order to circumvent this problem, heuristically formulated algorithms appear to be the ones used in practice. Even then, these techniques require nothing short of a general purpose computer to solve the coefficients.

Although polynomial approximations are easy to use and often work satisfactorily in practice, there are situations in which other types of approximating functions are preferable. Transcendental curve-fitting may perform better when the device function is known to be exponential or bandlimited. Of particular interest are the cardinal series sampling expansions. This type of approximation is often used to reconstruct a bandlimited function which is sampled at regular intervals (Papoulis [1962]). By bandlimited, it is meant that the function contains no frequency above some maximum. The cardinal series expansions are based on the well-known sampling theorem due to Shannon [1948] and they require an infinite number of samples for the theoretical formulation. Since practical considerations restrict a finite number of data, a truncated series is used. Helms

and Thomas [1962] and Brown [1969] obtain an upper bound on the truncation error which results when a finite series is used. Although most applications assume that the data are regularly spaced, one may obtain an expansion with arbitrarily spaced data. Yen [1956] provides such a formulation for interpolating nonregular samples that is essentially a Lagrange formula. Others have sought to show a connection between the cardinal series expansions and the more conventional polynomial interpolation formulas. For example, Radzyner and Basen [1972] determine an error bound for Lagrange interpolation. They required that the interpolated function be bandlimited, that the data be equispaced, and that approximately equal numbers of samples lie on either side of the interpolation point.

Adaptive curve fitting was also suggested by Chang [1968]. Here, a curve is to be approximated by a series of intervals of duration inversely proportionate to the "curliness" of the curve. Unfortunately, this approach requires the evaluation of integrals, among other computational mechanics. The piecing together of these areas may be dealt with using spline theory, but the criterion of performance is usually how good the spline approximates the data and not how well one can expect it to perform for other points. With this inadequacy, along with the difficulty in obtaining the spline adaptively, such an

"adaptive" procedure does not appear to be of practical consequence for device testing.

To the extent that conventional interpolation and extrapolation formulas are used in the determination of bounds in Chapter 5, one may expect existing curve-fitting techniques to be applicable to the thesis developments. The major difference in applying curve-fitting techniques to the methods of Chapter 5 lies in the fact that these methods use statistical information in place of some of the deterministic data normally required by interpolation and extrapolation formulas.

A particular type of extrapolation problem is considered in the thesis. This problem is concerned with the extrapolation of a time dependent device function over large intervals of time. This extrapolation is to be performed using as a data, sequence of observations of the function over an initial period of time. These observations or measurements are made at regular intervals and the extrapolation interval is from the last observation to a later point in time. The meaning of the word "large" here is that the extrapolation interval is at least an order of magnitude larger than the measurement interval.

This extrapolation problem corresponds to an important class of practical problems arising in the automated testing of certain types of devices. These problems arise from

determining long term storage effects on devices. In the case of electronic devices such as monolithic circuits, the additional problem of life testing may be similar to storage testing in that time rather than the number of cycles of operation (oscillations or state changes) affects the mean-time before failure (MTBF). As an indication of the extreme importance of this problem, one may refer to the experimental studies of storage and wearout mechanisms reported by Black and Hall [1972], Elliot [1973], Milek [1969], Schlegel [1969], and Welker [1973].

The extrapolation problem is to predict the long term storage effects for a particular device based on a short time observation of the device characteristics. Since practical considerations require that this observation or testing time be minimized, it is clear that one must minimize the number of measurement intervals required to characterize the device. The extrapolation problem is therefore consistent with the basic problem considered in the thesis.

Since the approximation error of conventional extrapolation formulas are usually proportional to the extrapolation distance raised to a positive power (Hamming [1962]), it is apparent that such techniques are not suited to long term prediction applications. In reviewing literature related to predicting storage and wearout effects, it is

noted that many investigators[1] avoid the prediction problem completely by considering only the failure distributions of the device type. Besides the fact that no information about a particular device is used in determining its future performance, such an approach suffers from its dependence on the definition of failure which may not agree with the definition made by a would-be user of this analysis.

In order to avoid the above shortcomings of conventional extrapolation and the failure distribution approach, it is useful to consider the area of pattern recognition theory. In many practical situations, one may be interested in classifying the predicted device storage effects into two categories. That is, whether the device is expected to be acceptable or unacceptable at a future point in time. Using a pattern recognition approach, measurements made on the device would be organized as a vector called a pattern. Each measurement might, for example, correspond to a component of the pattern. These components are called features (Nilsson [1965]). The most simple pattern recognizer is a system which would generate a binary response to each pattern presented to the recognizer. This response

---

[1]This approach is a special topic of reliability analysis. In addition to the papers already reviewed, three other examples of the approach are Pollock [1967], Sarkar [1971], and Taylor [1973].

indicates which category the presented pattern is classified as by the recognizer.

One of the primary considerations in using pattern recognition theory is the selection of the pattern components. This selection process is called feature extraction. Levine [1969] presents a survey of work in the area of feature extraction. He emphasizes a conclusion by many investigators that the methods of feature extraction are often empirical and use many ad hoc strategies.

The set of all possible patterns in a particular application can be considered to represent a vector space or at least span a vector space. Much of the developments of pattern recognition techniques assume that there exists a hyperplane space which divides the pattern space such that all the patterns belonging to one of the pattern categories lie on one side of the hyperplane, and the remaining patterns lie on the other side (Nilsson [1965] and Sebestyen [1962]). In this case, the patterns are said to be linearly separable. The meaning and ramifications of linear separability are further explored later in the thesis, and one may obtain an introduction to the general theory of pattern classifiers in Nilsson [1965]. It is important to bring up the term of linear separability in that much of the important theory concerning pattern classifiers makes an assumption of this property.

It happens that in many important situations, the patterns are not linearly separable. As a way around this, a transformation can sometimes be obtained which results in a new pattern space which is linearly separable. Sebestyen [1962] presents some of the types of transformations that may be tried. There is no general theory which can be used to generate particular transformations, and like the area of feature extraction, empirical methods and ad hoc strategies must be resorted to. One such strategy reported by Kruskal [1972], attempts to find a linear transformation which will reveal a "cluster structure" in a multi-dimensional pattern space. This technique uses an "index of condensation" which is, "intended to indicate, for a given configuration of points, the extent to which these points are condensed around point centers or low-dimensional structures. This index does not depend on a tentative description of the low-dimensional structures around which the condensation presumably occurs." Unfortunately, author Kruskal admits that his intuitively derived condensation indices fail to perform satisfactorily even for well defined geometrically clustered points. He also warns that, "even if a good index of condensation is discovered, its practical value depends in part on the feasibility of optimizing it over a suitable set of linear transformations."

Work has also been conducted on adaptive procedures which form nonlinear hypersurfaces for the separation of patterns (Kobylarz [1968]). For this situation, the requirement of linear separability is relieved. The paper includes a theorem which interrelates the dimension of the N^{th} order Euclidean space and a minimum "order" adaptive scheme that will guarantee separability for all consistent patterns.

An important feature of pattern recognition theory techniques is that one may obtain the classification mapping by a training procedure under certain conditions. This training procedure, described by Nilsson [1965], requires a set of patterns known as a training set of which it is known beforehand which patterns belong to each of the pattern categories. By repeatedly applying the training patterns to the pattern classifier and making corrections of certain weights of the classifier according to a training algorithm, one may obtain the required mapping under convergent conditions. Nilsson further states that convergence usually means that after a finite number of training cycles, a given percentage of the training patterns will be classified correctly. The condition of convergence usually requires that the training set be linearly separable or at least approximately separable. Nilsson [1965] and Sebestyen [1962] consider many types of training algorithms and proofs of convergence.

Once trained, the classifier is assumed to classify other patterns correctly. Obviously the performance of the classifier for patterns not belonging to the training set depends strongly on how representative the training set was of the entire pattern set. The quality of the training set in characterizing the entire pattern set must be considered in applying pattern classification techniques to practical problems.

Pattern recognition theory has been applied to the problem of predicting long term reliability of electronic devices. Pokrovsky [1972] proposed using device parameters as features in such an approach. Here k parameters are measured at N regular intervals of time. A pattern consists of the kN parameter values. A pattern is considered to belong to either a failure class or a reliable class. The failure class pattern corresponds to devices that eventually fail in some way before a maximum life span T. A failure is considered to occur when one of the parameters reaches some a priori tolerance threshold. This approach suffers from several drawbacks. First, the larger the required N, the greater the dimension of the pattern recognizer. Because of this, one needs an increasing number of training patterns to effectively train the recognizer as N increases. Second, there is no reason to expect that the pattern space is linearly separable nor is there any way to predict the

order of a nonlinear machine that will separate the training set. But even more important, this approach is sensitive to the definition of failure (i. e. the critical thresholds) and to the particular conditions under which the initial testing was carried out. It is the characteristic of engineering design that what is unacceptable to one design is acceptable to another. Except when failure is catastrophic, there is often no clear-cut threshold of unacceptability since the threshold depends on specific design criteria. Moreover, the above approach requires that the environment under which the training samples were obtained by the same for the patterns that the recognizer later classifies. This rules out classification of patterns obtained under changes in the environment as well as being useful for analysis of accelerated testing. For moderate increases in temperature or voltage, one can often assume that the failure mechanisms will be speeded up.[1] Moreover, from one device to the next, the dynamics of the same type of failure mechanism can be expected to vary. Thus any approach which assumes a fixed rate of failure or a fixed threshold of failure is limited in its usefulness.

---

[1]This idea is often the basis of accelerated testing of electronic devices. See Elliot [1973], Bruitz [1969], and Kooi [1968] for discussions of this idea based on device physics.

In section 2.4, it will be explained how the basic concepts of a trainable pattern recognizer can be used in a new approach to the problem of long term extrapolation. This approach will be developed in Chapter 6.

2.3 Review of Several Topics Indirectly Related to Device Testing

Several topics will be briefly reviewed in this section which, although not directly related to the developments in this thesis, are of general interest to the problem of device testing. The relevance of each topic to device testing will also be explained in this section.

The first topic considered is reliability theory. It is natural to consider this topic when studying automated testing since reliability analysis often indicates what kind of information should be determined for a device. However, reliability analysis gives little insight into how this information may be obtained. Reliability theory is often useful in relating the failure rate characteristics and often the functional characteristics of individual devices to the characteristics of a larger system in which the devices operate (Bracchi and Somalvico [1971]) and (Brown and Martz [1971]). In circuit design applications for example, the circuit performance is often expressed in terms of its sensitivity to variations of component values (Herskowitz [1968]). Thus the allowable operating characteristics of a device is often dictated by reliability

considerations of a specific system design. Furthermore, "failure" of a device may be defined by the application at hand. As an example, military specifications (MIL-SPEC) are often set for critical system components according to special operating requirements.

When reliability statistics are known for systems elements, the overall systems reliability can be calculated by mechanized procedures. Cohn and Ott [1971] use a combination of decision-tree analysis and prior probability of element failures to generate an adaptive test procedure. A cost is associated with each test, and the decision-tree is constructed in such a way to minimize the expected cost of determining the system status. Fleming [1971] uses the system topology and prior known failure statistics of system blocks to determine systems reliability. The system must be serial in structure although blocks may be made of parallel redundant sub-blocks. D. B. Brown [1971] develops a similar technique in which topology is expressed as a Boolean function.

More refined statistical analysis of circuits and other systems allows more complex topology. Bracchi and Somalvico [1971] describe three approaches used by others and introduce a fourth. In all four, one is trying to determine the effect of component variability on circuit performance for the purpose of design. Each approach

requires that the component variability be known and that it be in the form of probability density functions. The major difference between the methods is in the way that the required component variability information is combined to form an overall measure of circuit reliability.

In the developments of this thesis, an assumption will be often made that an acceptability region has been determined within which a device function is required to lie, for the device to be acceptable. This function may be an input-output function, or it may be a parameter function dependent on an environmental variable or on time. In any case, it is through techniques such as reliability analysis that one can determine the acceptability constraints on device functions.

A second topic that requires some consideration is that of design of experiments (Wald [1943]). The reason for this concern is that it is possible to draw an analogy between a special type of device testing problem and the theory of design of experiments.

In order to describe the special testing problem, it is useful to consider the following hypothetical situation. A device is known to have a potential mode of long term failure, approximately described by one of several possible mathematical models. These models are functions of a multidimensional input vector $\underline{x}$. The problem is to

determine optimal input vectors that enable one to determine which failure model is to be used to describe the failure mechanism and to determine the unknown parameters for the selected model. Since the models being considered correspond to a mechanism that produces an eventual failure, the practical usefulness of determining the model and its parameters is to predict the long term reliability of the device. In other words, the response of the device for the various optimal input vectors is used to extrapolate present and past device behavior to a future point in time.

With the above testing problem in mind, one may briefly consider several important papers in the area of design of experiments. In what follows, it is assumed that a model equation is known and that a non-negligible experimental (measurement) error exists. This error is assumed to be a random error, normally distributed, with known variance and zero mean.

Wald [1943] considered the problem of testing a linear hypothesis. He considered N independently, normally distributed varieties y_i, $i = 1,...N$ having a common variance σ^2. The expected values of y_i were assumed to be given by

$$E[y_i] = \beta_1 x_{1,i} + \beta_2 x_{2,i} + \beta_p x_{p,i}$$

where the x_{ji}'s were assumed known and could be considered as inputs, while the β_j's were unknown parameters. The β_j's

were population regression coefficients of y on x. The linear hypothesis was that the unknown coefficients β_j satisfy a set of linear equations,

$$g_{i,1}\beta_1+\ldots+g_{i,p}\beta_p = g_i \quad (i=1,\ldots,r;r\leq p).$$

The problem was to test this hypothesis through N observations, and to pick the corresponding p-dimensioned vectors $\underline{x}_i$ (i=1,...N) in such a way as to maximize the sensitivity of the test. In his paper, Wald determined that the optimum choice of the vectors $\underline{x}_i$ occurs when the variance of the least square estimate $\underline{b}$ of $\underline{\beta}$ is a minimum. The actual mechanics for determining the $\underline{x}_i$'s in a domain D are discussed in this paper.

A later paper by Box and Lucas [1959] considered a similar problem, but allowed a nonlinear situation. Here they dealt with a response function

$$f(x_1,x_2,\ldots,x_k;\ \theta_1,\ldots,\theta_p) = f(\underline{x};\underline{\theta})$$

where the x_i's are inputs and the θ_i's are parameters. The authors presented the formulation as a design of an experiment for some physical process, run N times. The problem was to choose the Nk values of input values of x_{ij}, where the subscripts referred to the j^{th} run value of input i.

$$\underset{\sim}{D} = \{x_{ij}\}$$

called the design matrix. The ultimate aim was to obtain a design matrix which enabled high accuracy in the determination of the model parameter, $\underline{\theta}$.

The authors' criteria for the selection of the design matrix was to choose D so that the determinant $|(\underset{\sim}{F}' \underset{\sim}{F})^{-1}|$ is made as large as possible, where the matrix $\underset{\sim}{F} = |f_{rj}|$. The experimental error in measuring the response is assumed to have a zero mean and a constant, although unknown, variance. Inherent in this criteria was that a least squares estimate of $\underline{\theta}$ is used as the method of "fitting" the experimental data.

Since the derivatives, f_{rj} are dependent on the θ's, the practical problem of finding these values may be difficult. If approximate estimates of the θ's are available, one may assume that f is linear in θ_r near $\underline{\theta}_0$. Using this approximation, the authors used a geometrical formulation to obtain a solution for simple problems. They also suggested numerical methods which could be used when the linear approximation was not valid.

In another paper, Box and Hunter [1965] extend the previous problem by considering also a mean of determining the adequacy of the experimental model, represented by the response function $f(\underline{\theta}; \underline{x}, t)$. For each "run" of the experiment, the inputs $\underline{x}$ are chosen and observations of the response are made at specific values of time. For a given run, a least squares estimate of the parameters is obtained, denoted $\underline{\theta}_j$. If N runs are conducted, each with different $\underline{x}$ vectors, a least squares estimate of the

parameters, $\underline{\theta}$, is possible with the complete set of data, in addition to the run estimate. Adequacy is measured as the dependency of $\underline{\theta}_j - \underline{\theta}$ on the levels $\underline{x}_j$. Thus data is used here to evaluate the model rather than the parameters of the model, although both tasks are required in practice.

The techniques just discussed are potentially applicable to the hypothetical testing situation discussed earlier. Each experimental run required by the techniques would correspond to the observation of a particular device over a period of time preceeding a failure. Since one cannot initially predict if the failure mechanism is present in a device, many devices would have to be tested in order to guarantee that among these devices, a sufficient number would exhibit a failure. Using the design of experiments approach, one could therefore determine the failure model and its parameters, and could also determine the model's adequacy.

In this thesis, alternative approaches will be taken to device testing, in connection with failure prediction. One major reason for rejecting the experimental design approach is that one often has no a priori models for the failure mechanism. Even if the mathematical models are available, it is questionable whether one could expect the model parameters to be the same from device to device. Nevertheless, the theory of design of experiments seems

important enough to the author to include in a general literature review. It is expected that there are special situations in practical device testing applications in which the theory may be successfully applied.

The last topic discussed in this section is the concept of entropy (Shannon [1948]). The reason for including this topic in the literature review is that entropy can be a useful tool in studying many statistical processes. If one views the additional information provided by a measurement as the opposite of noise, one can interpret this information change as a reduction in entropy. In this thesis, measurement information for a particular device will be combined with prior statistical information gathered for many devices. Although entropy theory will not be directly utilized in most of the developments,[1] it is expected that this theory may be generally useful to the reader in understanding these developments. A brief review of several applications of entropy theory to generally relevant problems is therefore given.

Since Shannon [1948] showed how the concept of entropy could be applied to the study of communications channels, various other investigators have sought to apply the concept

---

[1]The principle of minimum information (Evans [1969]) (maximum entropy) will be directly used in the proof of a theorem given in the following chapter.

to other statistical processes. One such application was to the area of design of experiments. Box and Hill [1967] showed how the previously discussed techniques of parameter estimation could be extended to the situation wherein several equally likely models were conditions for a mechanistic process. Entropy was used here as a means for discriminating between the models and as a basis for determining additional testing conditions in a sequence of experimental runs. Hill, et al. [1968] and Hill and Hunter [1969] extended this initial technique to a more general problem of combining parameter estimation with model discrimination.

Entropy theory has also been used to study the concept of prior probabilities and to justify the utilization of these probability measures in Bayesian analysis.[1] Jaynes [1968] demonstrated that the concepts of entropy and prior probabilities are consistent, provided they are properly applied. Using an argument of maximum entropy, he proved an equivalence of the two concepts. Jaynes also suggested that the principle of maximum entropy could be applied to such areas as reliability theory.

Evans [1969] considered Jaynes' paper [1968] in relation to reliability theory. Redefining the principle

---

[1] Bayesian analysis will be utilized in a development of the next chapter.

of maximum entropy as a principle of minimum information,[1] he indicated that in any practical problem, unknown probabilities should be allocated so that the resulting information is minimized within the problem constraints. Regulinski [1969] also suggested that reliability theory could be aided by the incorporation of the entropy concept. He drew a parallel between Shannon's communications channel and situations found in reliability theory, and proposed several problems to be solved, although offering no solutions. An example of one success in applying the principle of maximum entropy to reliability is found in a paper by Simkins [1972]. Using a Weibull model for a hazard function, he used entropy as a figure of merit in this reliability analysis application. Bayesian analysis was used as a means of updating the model parameters after each of a sequence of deterministic tests.

In conclusion of this section, the topics of reliability theory, design of experiments, and the concept of entropy have been briefly reviewed. It is expected that these topics will provide the reader with a backgorund that will be useful in the understanding of this thesis developments.

---

[1]The term information is used here in the sense of Shannon [1948].

2.4 Conclusions Made of the Reviewed Literature In this section several important conclusions are made of the reviewed literature. This conclusion indicates a void in existing techniques regarding theory for applying interpolation and extrapolation procedures to the special applications of automated testing systems. It has been seen that exact-fit approximations are often preferable techniques when a device function is to be represented using a number of discrete measurements. These techniques, especially the polynomial approximations, are readily mechanized and often give sufficiently accurate results. However, these conventional techniques fail to utilize all the information that is often available. In practical situations, one measures the functional behavior of large numbers of devices of the same type. Furthermore, it is reasonable to require an initial concerted measurement effort for several lots of devices if the information obtained can be used to economize the testing of devices on a production basis. This economy can take the form of simpler testing systems than the initial laboratory systems and a reduction in the testing time per device for the production or field system. Such an initial intense data gathering effort would allow estimates of the distributions of device functional values.

It appears that the statistical information just discussed should be incorporated into the conventional curve-fitting techniques in order to achieve a reduction in the number of measurements required to characterize a device. The characterization amounts to the interpolation of the device function between data points. Since the requirements of such a characterization is to determine if the device function lies between tolerance limits within which the device is said to function properly, it is clear one may use tolerance limits for the data used to determine the approximation formula. While such an approach may seem quite reasonable, one finds a void in the literature along these lines. The initial ideas of this approach are developed in the next chapter. In a subsequent chapter, the approach is formalized for exact-fit polynomial and transcendental interpolation and extrapolation.

One may observe that the use of tolerance limits for data directly in approximation formulas accomplishes direct incorporation of prior information into conventional techniques. Another new approach developed in this thesis will utilize concepts of pattern recognition theory in order to determine extrapolation formulas that are better suited to the long term extrapolation problem described in the last section. Prior information in the form of a training set of patterns will be used to modify the

conventional techniques of extrapolation. The intention of such an approach is to obtain extrapolation algorithms that are more accurate than conventional formulas, thus allowing a reduction in the required number of measurements to characterize a device. This reduction is seen to allow earlier predictions of future reliability of devices. This approach will additionally require a classification procedure to interpret the extrapolated device functions.

It is noted that the developments in this thesis concerning adaptive extrapolation will help to fill a void existing in the literature. Lucky [1968] considered a problem analogous to adaptive extrapolation in a paper on filter theory. Here the extrapolation was continuous and was used in a data compression system for redundancy removal. In this situation, the function of the extrapolator was that of a predictor. Furthermore, discrete data could also be handled if the system was constructed as a predictive digital filter. The difficulty in applying his results to long term reliability prediction is two-fold. First, the observations that are used to train the extrapolator must be contiguous. This does not allow for observations that are dispersed throughout a large interval of time to be used as training patterns. Nor can training patterns observed for many devices be used in this predictive filter approach. The second difficulty in using

Lucky's techniques is that multivariate extrapolation is not generally attainable.

In conclusion, this literature review has examined topics related to the theory for applying interpolation and extrapolation procedures to automated testing systems. From the reviewed literature, a conclusion is made that there exists several voids in this subject area. This thesis will attempt to partially fill the voids with the expectation of providing improved techniques useful for many automated testing situations.

3. PRELIMINARY CONSIDERATIONS

3.1 Introduction This chapter is devoted to preliminary considerations of the problem of combining prior information with measurements made on a particular device. The material in this chapter is to provide the reader with a background for the initial approaches considered by the author. The chapter contains three different topics which represent potential starting points for later results. The first topic is concerned with fundamental techniques which combine deterministic device function measurements with "statistical data." Because this thesis will use the terms statistical data and deterministic data frequently, it is important to point out the intended meanings of both. If a great number of similar devices are measured at the same input conditions, these measurements may be organized as a distribution of a statistical quantity. This quantity is the device function evaluated at the particular input conditions for which the measurements were obtained. In this thesis, the term "statistical data" will always mean the value or the range of values of the device function obtained from a statistical distribution. The term "deterministic data" will always mean the value obtained by direct measurement made on a particular device. The term "function" will always signify a "single-valued function."

The preliminary topic introduces a basic problem. The problem demonstrates the need to linearly combine tolerance limits of random variables. By assuming that the random variables are independent in the statistical sense, the complexity of the combination problem is reduced. However, there still is a need to simplify the process in which the bounds on the random variables can be combined. A particular combining algorithm will be described which is used in the preliminary topic described above. The second topic of the chapter considers alternative algorithms with respect to their suitability for the statistical interpolation techniques first discussed.

The last topic of the chapter concerns a special aspect of device testing. The problem relates to the determination of process parameter effects over the area of a monolithic circuit. While this last topic is somewhat different from the remaining developments of the thesis, it is included because of its importance in the testing of integrated circuits. It is also noted that the topic addresses the principle problem studied in this thesis. That is, it enables a reduction in the number of deterministic measurements that must be made on a particular device to satisfactorily characterize it.

3.2 Basic Concepts Consider a device having an input-output function f(x) which is continuous, instantaneous, and allowably nonlinear. The function f(x) represents a means

of characterizing the relation between a measurable response (output) and a controllable entity (input) which acts on the device. The input, x, may be an electrical variable such as a voltage or current, for example. This input variable, x, may also be an environmental quantity such as temperature or pressure. The function, f(x), will be referred to as a "device function" or a "response function" in this thesis.

The function f(x) will be assumed continuous throughout this thesis. In this section, the derivative of f(x), f'(x), will also be required to be continuous. Another requirement of f(x) will be that it is an instantaneous function of x. That is, the function value f(x) at any instant in time depends at most on the input value x at the same instant, but not on past or future values of the input.

Let it be further assumed that the probability density function, h(x, z), of f'(x) be known. That is, for any x in the open interval c, one may write (Feller [1968])

$$p\{a \leq f'(x) \leq b\} \equiv \int_a^b h(x,z)dz \qquad (3.1)$$

where x is a parameter of the probability density function (pdf), and z represents the explicit variable of the pdf variation. (One may wonder at the practicality of requiring the pdf at every point in an interval of the parameter x, let alone at a specific point. The acquisition of such prior information will be discussed at the end of this section).

Suppose a particular device is tested by simulating the input x at some point x_o and measuring the response $f(x_o)$. Assuming that the measurement error is negligible, if f(x) were measured for some x differentially close to x_o, the value of f(x) at this second point must be approximately that of $f(x_o)$, due to the continuity assumption of f(x). Intuitively, one suspects that the certainty of knowing the value of f(x) must vary inversely with the distance from x_o, with no other assumed prior knowledge other than the constraints on f(x).

In this section, it will be considered how to combine prior statistical information in the form of the pdf of f'(x) with deterministic measurements in order to obtain a bound on f(x) near a deterministic data point. The process by which this is accomplished here will be referred to as statistically based interpolation and extrapolation (Herman and Kobylarz [1972]).

Let the measurement of $f(x_o)$ be made. Consider the interval $[x_o, -\Delta x\ x_o + \Delta x]$, where the positive constant Δx is chosen sufficiently small so that f(x) is nearly linear in the interval. That is, the Taylor series of f(x) at x_o may be written

$$f(x) = f(x_o) + (x-x_o)\, f'(x_o) + \varepsilon \qquad (3.2)$$

where the truncation error ε is assumed small for the interval $[x_o-\Delta x,\ x_o + \Delta x]$. Presently, the magnitude of Δx that allows a linear approximation will be assumed known. The

consideration of the size of Δx will become unnecessary when the interval is allowed to become differential.

Corresponding to a confidence level α, it is possible to define the tolerance limits (TL's) a and b, for the value $f'(x_o)$, by

$$\alpha = p\ \{a \leq f'(x_o) \leq b\} = \int_a^b h(x_o, z)\,dz \tag{3.3}$$

provided that some further convention be chosen in the definition. In the case of a symmetric or near symmetric pdf, one usually requires that the TL's be symmetrically placed about the distribution mean, μ. That is, the lower TL, a, is defined by

$$\alpha/2 = p\ \{a \leq f'(x_o) \leq \mu(x_o)\} = \int_a^{\mu(x_o)} h(x_o, z)\,dz \tag{3.4a}$$

and the upper TL is defined by

$$\alpha/2 = p\ \{\mu(x_o) \leq f'(x_o) \leq b\} = \int_{\mu(x_o)}^b h(x_o, z)\,dz \tag{3.4b}$$

where

$$\mu(x_o) \equiv \int_{-\infty}^{\infty} z \cdot h(x_o, z)\,dz. \tag{3.5}$$

In what follows, the definition of equation (3.4) will be used to establish the TL's. The two limits may be used to define an interval [a, b] for $f'(x_o)$. This interval will be referred to as a tolerance interval (TI) throughout the thesis.

Suppose the device has been tested at x_o so that the value of $f(x_o)$ is known for this device. Additionally, let the TI be established for $f'(x_o)$ for some value of the confidence level, α, using equation (3.4). Substituting the approximation of equation (3.2) for $x=x_o+\Delta x$ (neglecting ε) into (3.3) yields, after manipulation,

$$\alpha = p\ \{f(x_o)+a\cdot\Delta x \leq f(x_o+\Delta x) \leq f(x_o)+b\cdot\Delta x\}. \tag{3.6}$$

For a negative excursion along the x axis, $(x=x_o-\Delta x)$ assuming the same Δx is appropriate, one has

$$\alpha = p\ \{f(x_o)-b\cdot\Delta x \leq f(x_o-\Delta x) \leq f(x_o)-a\cdot\Delta x\}. \tag{3.7}$$

Figure 3.1 depicts equations (3.6) and (3.7).

The preceding results may be summarized by:

Theorem 3.1 Let $f(x)$ be a continuous, instantaneous, nonlinear function, and let $f(x_o)$ be known. Let the derivative of $f(x)$, $f'(x)$, be continuous also, and let the TI for $f'(x_o)$ be known for a particular confidence level α as defined by (3.4).

If for some truncation error ε, a positive bound $\hat{\varepsilon}$, and Δx exist, such that

$$|f(x)-f(x_o)-(x-x_o)\cdot f'(x_o)| \leq \hat{\varepsilon}$$

for x in $[x_o-\Delta x,\ x_o+\Delta x]$, then the probability that

$$\begin{aligned} &f(x_o)+a\cdot\Delta x-\hat{\varepsilon} \leq f(x) \leq f(x_o)+b\cdot\Delta x+\hat{\varepsilon},\quad x_o \leq x \leq x_o+\Delta x \\ &f(x_o)-b\cdot\Delta x-\hat{\varepsilon} \leq f(x) \leq f(x_o)-a\cdot\Delta x+\hat{\varepsilon},\quad x_o-\Delta x \leq x \leq x_o \end{aligned} \tag{3.8}$$

is greater than or equal to α.

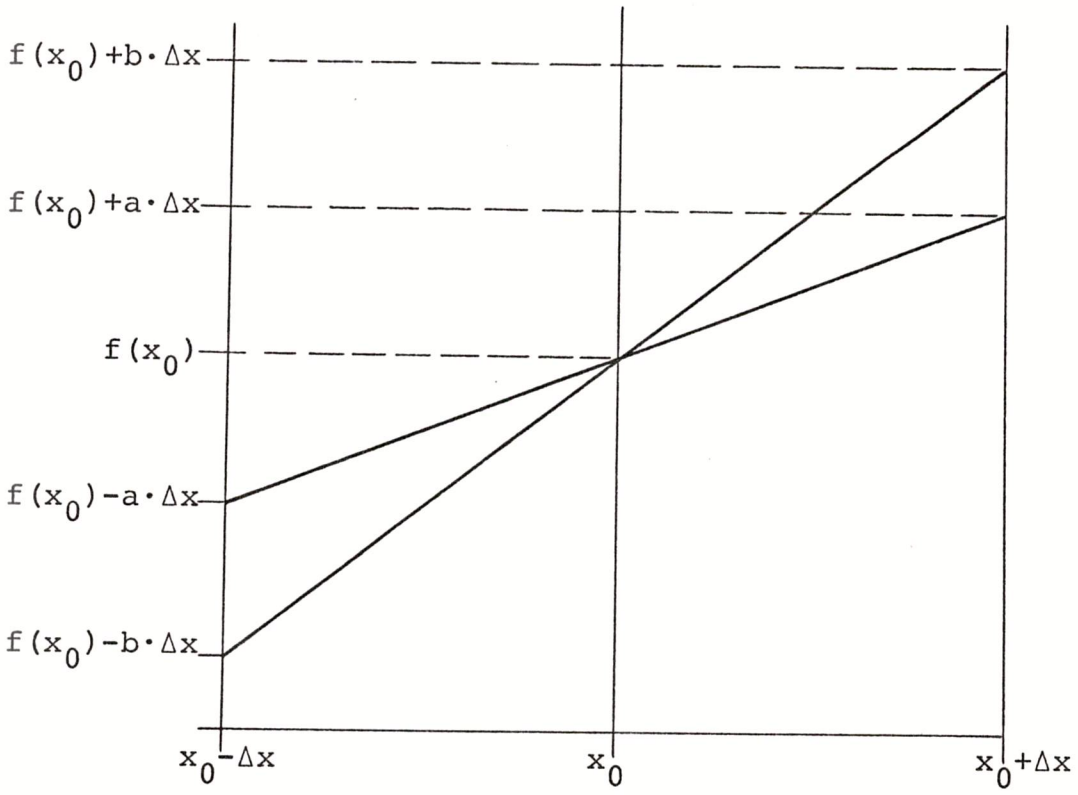

Figure 3.1 Construction of a statistical bound on f(x) near the point x_0. The constants a and b represent the tolerance limits for f'(x) corresponding to a confidence level α, as defined by equation 3.4. The bound on the truncation error is assumed negligible here, but could be added to the bounds obtained by using equation 3.8.

Definition 3.2 The bound on f(x) corresponding to Theorem 3.1, for $\hat{\varepsilon}=0$ is called a "statistical bound" on f(x).

Theorem 3.1 allows the combination of statistical information of the TI for $f'(x_o)$ and the deterministic measurement of $f(x_o)$. When the truncation error of equation (3.2) cannot be neglected, it must be added to the statistical bound as indicated in equation (3.8). Often a bound on this error is known in terms of the distance Δx, and in terms of a maximum bound on one of the derivatives of f(x). This a priori information effectively widens the overall bound on f(x) in the interval $[x_o - \Delta x, x + \Delta x]$ in that the magnitude of the error bound is added to the upper statistical bound and subtracted from the lower statistical bound as in (3.8).

One may extend the application of Theorem 3.1 to the case where the value of f(x) can be measured at more than one x. To illustrate this extension, assume that Theorem 3.1 can be applied at any x in some interval c of x much larger than Δx. That is, for the known positive constants $\hat{\varepsilon}$ and Δx

$$|f(x) - (x-x_i)\cdot f'(x_i) - f(x_i)| < \hat{\varepsilon}$$

for x in $[x_i-\Delta x, x_i+\Delta x]$ and $[x-\Delta x, x+\Delta x]$ in c. Assume for the discussion that follows that the approximation error $\hat{\varepsilon}$ can be neglected. Assume that $f(x_o)$ and $f(x_1)$ are measured, where $x_1 = x_o+2\Delta x$. At x_o and x_1, obtain the statistical bounds on f(x) using Theorem 3.1. Figure (3.2) shows the

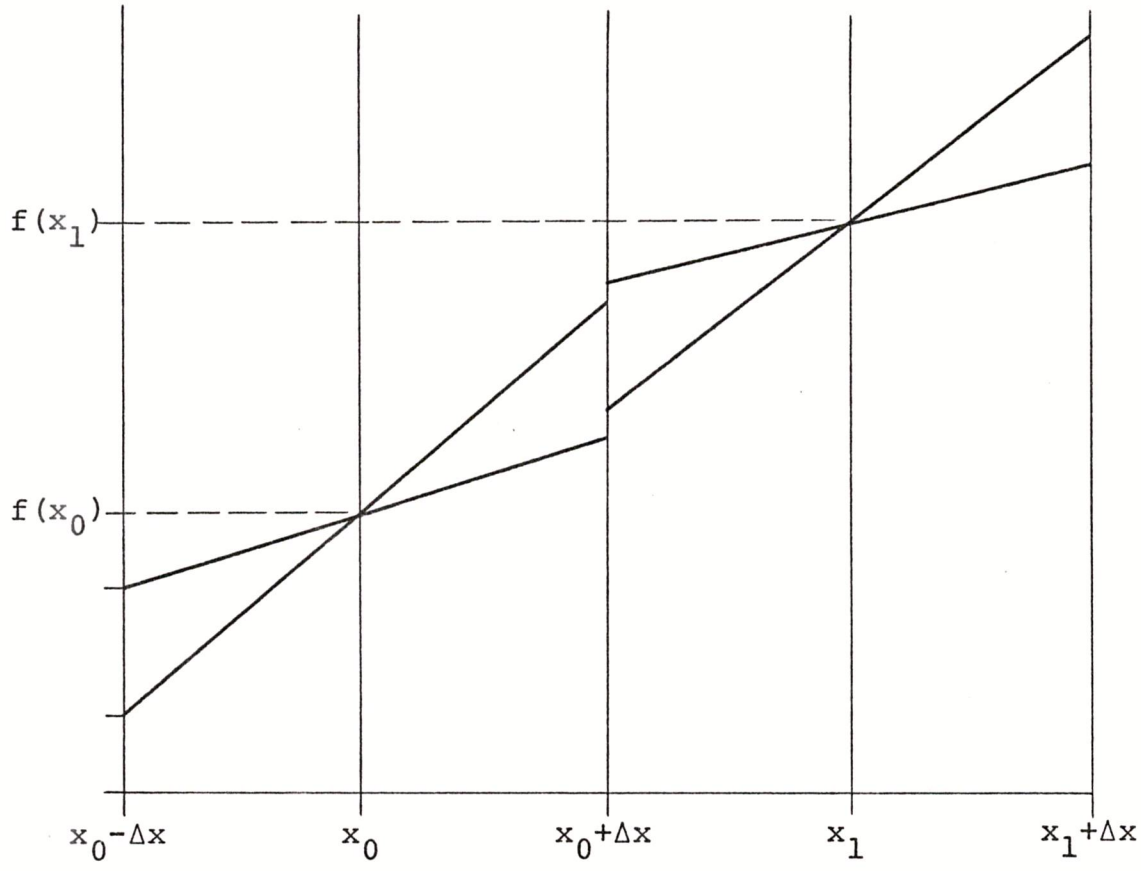

Figure 3.2 Extension of Theorem 3.1 to the case where two deterministic data points are used to form a statistical bound on f(x). Note that $x_1=x_0+2\Delta x$. Also note that the two bounds on $f(x_0+\Delta x)$ do not coincide. The bound on the truncation error is assumed negligible here, but could be added to the bounds obtained by using equation 3.8.

two bounds so obtained. Consider the intersection of these two bounds at $x_o+\Delta x$. In general the two bounds for $f(x_o+\Delta x)$ will not coincide, but will usually have some region of intersection provided α is chosen large enough. Lack of intersection may be interpreted by considering the two extrapolated values of $f(x_o+\Delta x)$ as random variables (RV's). If these two RV's are considered independent, then the probability of the event that $f(x_o+\Delta x)$ is within one of the two bounds, but not the other is $2\alpha(1-\alpha)$. For an $\alpha=.95$, the probability of non-intersection is thus .095. On the other hand, the probability that $f(x_o+\Delta x)$ does not lie within either of the two bounds is $(1-\alpha)^2$ or .0025. The probability that $f(x_o+\Delta x)$ lies within both bounds (intersection) is α^2 or .9025. Finally, the probability that $f(x_o+\Delta x)$ lies within either/or both bounds (union) is $\alpha(2-\alpha)$ or .9975 for $\alpha=.95$.

The results of the previous paragraph and the remarks concerning the various conditions of intersection raise at least two questions. The first addresses the problem of how to combine the two bounds on $f(x)$ resulting from application of Theorem 3.1 in order to obtain a combined bound on $f(x)$ between x_o and x_1. The second question is that of determining a confidence level to be associated with this combined bound.

Although one can invent many strategies to answer the first question, there are two approaches that deserve

particular attention. Each of these two approaches defines a combined bound on $f(x_o+\Delta x)$ in terms of the two TI's of $f(x_o+\Delta x)$ as the intersection of the individual bounds. The second strategy forms the combined bound on $f(x_o+\Delta x)$ as the union. Aside from this difference, both approaches form the combined bound on $f(x)$ between x_o and x_1 as follows. The bound on $f(x)$ between x_o and $x_o+\Delta x$, and between $x_o+\Delta x$ and x_1 is constructed by, in effect connecting straight lines between the combined bound on $f(x_o+\Delta x)$ and the deterministic data points at x_o and x_1. These two strategies are illustrated in Figure 3.3. The combined bound on $f(x)$ corresponding to the intersection approach is shown as dashed lines and the bounds corresponding to the union approach is shown as dotted lines. Discontinuity of bounds have been avoided, since they are contrary to the $f(x)$ continuity assumption.

The union-combined bound is more conservative than the intersection bound in that the former contains the latter bound. In trying to form a basis for choosing one approach over the other, it is possible to offer several heuristic arguments. One suspects that a combination of the individual TI's for $f(x_o+\Delta x)$ should allow an improvement to be made over either of the bounds considered separately. That is, corresponding to a particular confidence level, the bound on $f(x_o+\Delta x)$ using one of the TI's should become narrower when the second TI is additionally obtained. This intuitive reasoning indicates that the intersection

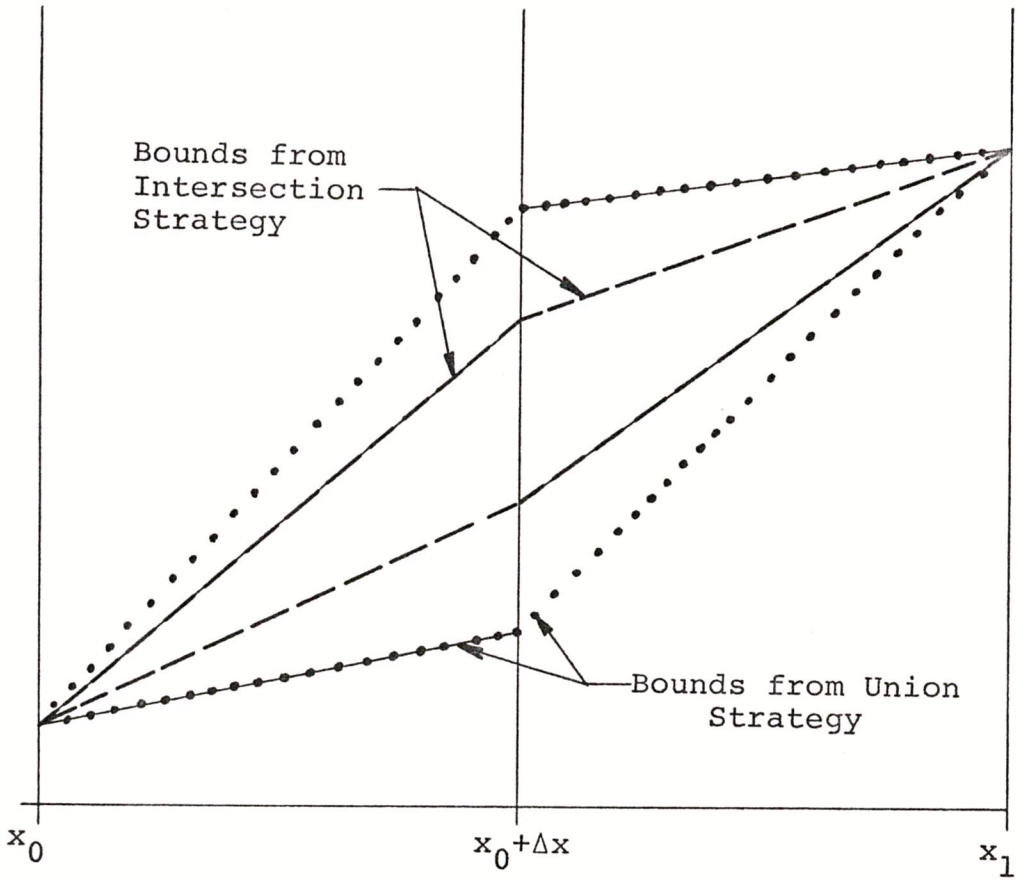

Figure 3.3 Illustration of the construction of combined bounds using two different strategies for combining the individual statistical bounds on f(x) near x_0 and near x_1. The dashed lines represent the bounds obtained using the intersection strategy. The dotted lines represent the bounds obtained using the union strategy. The solid lines represent the individual statistical bounds as constructed in Figure 3.2. Note that the solid lines coincide with either the dotted or the dashed lines within the interval $[x_0, x_1]$.

strategy is easier to justify than the union approach. In order to formalize this rationalization, it is useful to consider an information theory principle. The principle of minimum information is a way of assigning probabilities in the absence of complete specifications for the probabilities. Evans [1969] explains that, "the principle of minimum information states that the unknown probabilities are allocated so that the information (as defined by Shannon [1948]) is a minimum within the specified constraints. If this is done, we will not have added any information to our situation other than that presented by our constraints. Any other allocation of probabilities would give a larger value of information, thus suggesting that we had unwittingly added more information than implied by the specifications of the situation. The assignment of probability made according to this principle may not correspond to the true values, but they are the best we can do without adding more information than we actually had."

In order to apply the principle of minimum information to the selection of a proper strategy (i.e. choose the intersection or the union strategy), it is necessary to put the selection problem into the proper perspective. In either of the two strategies, the ultimate goal is to be able to assign a probability to two possible states. These two states are:

S_1: f(x) lies totally within the combined bounds obtained with the chosen strategy in the interval $[x_o, x_1]$.

S_2: the negation of S_1

The obvious requirements on the probabilities of S_1 and S_2 are:

$$0 \leq p\{S_1\} \leq 1$$
$$p\{S_1\} + p\{S_2\} = 1 \tag{3.9}$$

Since the confidence level of the individual TI's on $f(x_o+\Delta x)$ is preassigned, the assignment of the probabilities $p\{S_i\}$ represents the choice between the two strategies for combining the TI's. Therefore, if the principle of minimum information is to be applied, the strategy resulting in the minimum information for the resultant bounding of f(x) in (x_o, x_1) must be the "best" choice.

The event that $f(x_o+\Delta x)$ lies within the intersection of the two TI's on $f(x_o+\Delta x)$, remembering the assumption of independence of these two TI's, is

$$p_I\{S_1\} = \alpha^2 \tag{3.10a}$$

and from (3.9) and (3.10a)

$$p_I\{S_2\} = 1 - \alpha^2. \tag{3.10b}$$

The event that $f(x_o+\Delta x)$ lies within the union of the two TI's on $f(x_o+\Delta x)$ is

$$p_U\{S_1\} = 2\alpha(1-\alpha)+\alpha^2=\alpha(2-\alpha) \tag{3.11a}$$

and from (3.9) and (3.11a)

$$p_U\{S_2\} = (1-\alpha)^2. \tag{3.11b}$$

The definition of information given by Shannon [1948][1] is

$$I \equiv \sum_i p\{S_i\} \ell n[p\{S_i\}] \tag{3.12}$$

where ℓn denotes the natural logarithm. Denoting the information of the intersection strategy as $I_I(\alpha)$, substituting (3.10) in (3.12) yields

$$I_I(\alpha) = \alpha^2 \ell n[\alpha^2] + (1-\alpha^2) \ell n[1-\alpha^2] \tag{3.13}$$

Similarly, the information of the union strategy $I_U(\alpha)$ is found by substituting (3.11) in (3.12). Thus

$$I_U(\alpha) = \alpha(2-\alpha) \ell n[\alpha(2-\alpha)] + (1-\alpha)^2 \ell n[(1-\alpha)^2] \tag{3.14}$$

It is important to note that

$$I_I(\alpha) = I_U(1-\alpha)$$

which is easily verified by substitution. Therefore,

$$I_I(\alpha) = I_U(\alpha), \text{ for } \alpha = \tfrac{1}{2}.$$

By inspection, one also sees that

$$I_I(0)=I_U(0)=I_I(1)=I_U(1)=0,$$

which is an inherent requirement of the definition of information. Finally, one may verify by numerical computation that

$$I_I(\alpha) > I_U(\alpha), \quad 0<\alpha<\tfrac{1}{2}$$

and

$$I_I(\alpha) < I_U(\alpha), \quad \tfrac{1}{2}<\alpha<1$$

It is possible to summarize the results of the last few paragraphs by a theorem.

---

[1] Information of a state.

Theorem 3.3 Suppose a TI has been obtained for a RV by two independent observers, who possibly have different (prior) information. In combining these two TI's into an overall TI for the RV, it is possible to choose between representing the bound on the RV as either a union or an intersection of the individual TI's, where the choice is to be based on the principle of minimum information.

In particular, if α is the confidence level associated with each of the individual TI's, then the choice of the union (intersection) combination is consistent with the principle of minimum information for α less than (greater than) one-half. The point of indifference corresponds to α equal to one-half.

The application of this theorem to the problem of determining a statistical bound on $f(x)$ in $[x_o, x_1]$ provides a rationale for combining two seemingly contradictory bounds on $f(x)$. The two TI's may differ since they are obtained with different prior information. The information corresponding to the TI emanating from x_o and the information corresponding to the TI emanating from x_1 are statistically dependent, in general. It is this dependency that is missing in the specification of the hypothesized situation. The result of this incomplete knowledge is the uncertainty of how to combine the two TI's for $f(x_o+\Delta x)$. Although Theorem 3.3 offers a philosophical solution to this

uncertainty, one still needs to justify the principle of minimum information as the criterion to be used.

Finally, one may consider that if the information regarding dependency was incorporated into the probabilities of equations (3.10) and (3.11), a marginal increase in the value of I_U and I_I should result. For this to happen, the confidence level of each combined bounds for f(x) between x_o and x_1 should increase.

The foregoing lengthy discussion has enabled the extension of Theorem 3.1 to the situation wherein a statistical bound on f(x) is obtained between two deterministic data points in close proximity. This bound is illustrated in Figure 3.3. Theorem 3.3 has allowed an interpretation to be placed on the way in which the TI's were combined to form an overall bound on f(x). The extension of Theorem 3.1 can be generalized by continuing this bounding process in the positive and negative x directions to include any overall interval of x. The bound between any two adjacent points is obtained as was done above for the two points x_o and x_1. For a sufficiently large confidence level ($\alpha > \frac{1}{2}$), the intersection strategy depicted in Figure 3.3 is used. Should the two TI's combined by this strategy fail to intersect, one may require that the interval of x within which this non-intersection occurs be subdivided. That is, at the midpoint of the interval, f(x) is measured. The statistical bounding of f(x) is now performed between the endpoints of the

subdivided interval and the midpoint. The bounding scheme can be made more adaptive by requiring additional measurements of f(x) in regions where the obtained bound is deemed excessive by some a priori criteria.

It is clear that the above scheme will often require an enormous number of test measurements. When Δx must be kept very small, this simple statistically based interpolation scheme may offer little or no advantage over conventional approaches to characterizing f(x). In order to obtain an improvement over the previous scheme, it is now considered how to use the prior statistical information to a fuller extent. Consider the following construction. Referring to Figure 3.1, assume a TI has been established at $x_o+\Delta x$ using Theorem 3.1. Let $x_1=x_o+\Delta x$. Consider the two points, $(x_1, f(x_o)+b\cdot\Delta x)$ and $(x_1, f(x_o)+a\cdot\Delta x)$. At each of these two points, construct bounds on f(x) between x_1 and $x_1+\Delta x=x_2$ using Theorem 3.1. Since the constants of (3.8) are generally functions of x, the two TI's for $f(x_2)$ are, neglecting $\hat{\varepsilon}$,

$$[f(x_o)+b(x_o)\cdot\Delta x+a(x_1)\cdot\Delta x, f(x_o)+b(x_o)\cdot\Delta x+b(x_1)\cdot\Delta x]$$

and

$$[f(x_o)+a(x_o)\cdot\Delta x+a(x_1)\cdot\Delta x, f(x_o)+a(x_o)\cdot\Delta x+b(x_1)\cdot\Delta x].$$

The above construction is illustrated in Figure 3.4. The two TI's just obtained at x_2 can be combined in a number of ways. The algorithm to be used combines the TI's by taking the minimum limit of the two lower bounds and the maximum of the two upper bounds. That is, $f(x_2)$ is bounded by

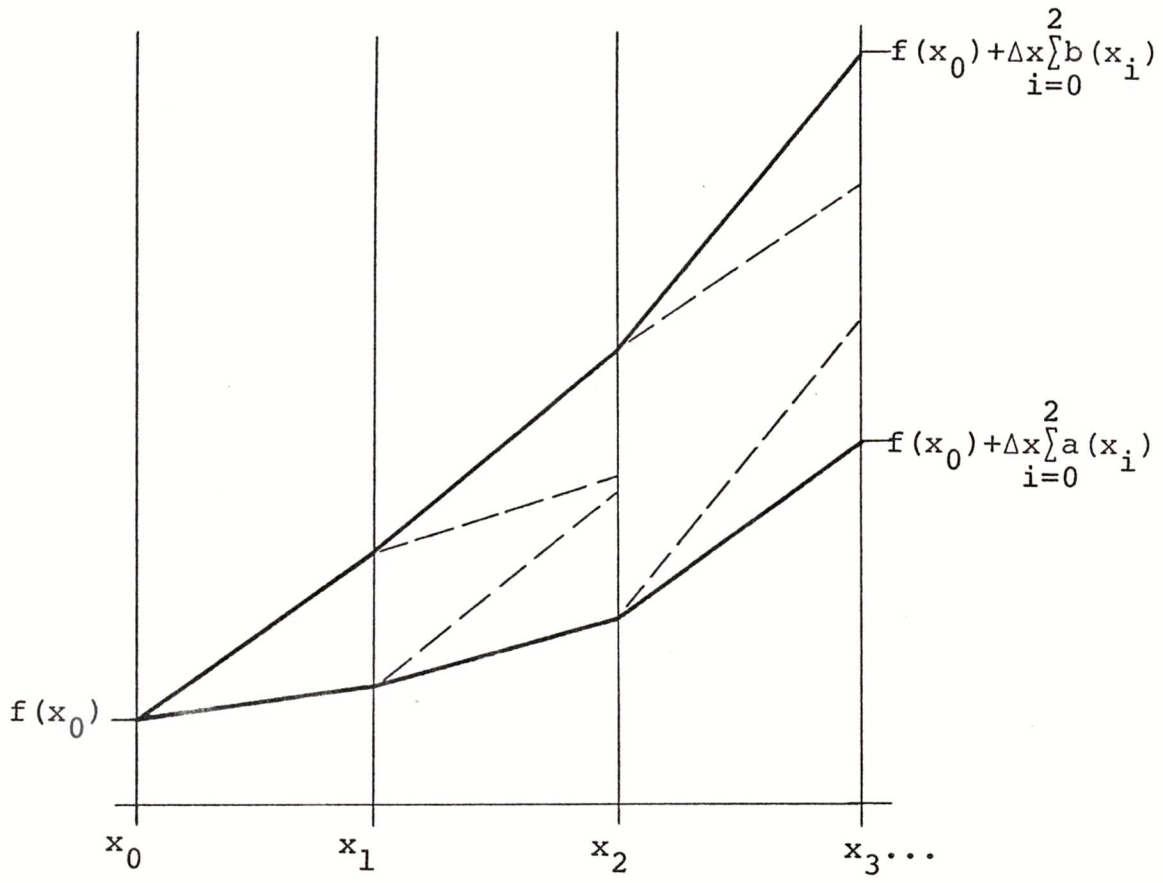

Figure 3.4 Iterative statistical extrapolation. Note that $\Delta x=x_{k+1}-x_k$ for all k. The bounding algorithm is used to combine the tolerance intervals corresponding to each of the $f'(x_i)$. Use of this algorithm results in the divergent bound on $f(x)$ shown as solid lines.

$$f(x_o)+a(x_o)\cdot\Delta x+a(x_1)\cdot\Delta x \leq f(x_2) \leq f(x_o)+b(x_o)\Delta x+b(x_1)\cdot\Delta x. \quad (3.15)$$

Moreover, f(x) is bounded between x_o and x_2 as shown by the solid lines in Figure 3.4. The construction just described can be iterated to produce a bounding of f(x) over many Δx. Equation (3.15) is easily generalized for the case of N iterations. In this case, neglecting the truncation error bound $\hat{\varepsilon}$,

$$f(x_o)+\Delta x\cdot\sum_{i=o}^{N-1} a(x_i) \leq f(x_N) \leq f(x_o)+\Delta x\cdot\sum_{i=o}^{N-1} b(x_i). \quad (3.16)$$

This iterative extrapolation of statistical bounds, referred to hereafter as "iterative extrapolation," is seen to be useful even in the case that Δx must be kept small. The major limitation of the iterative extrapolation is that the bounds are obviously divergent. The practical limit of the total extrapolation interval depends on the spread of the derivative TI's used at each iteration, and acceptable width of the bounds. The acceptability of the obtained bounds is dictated by the intended use of the device to which f(x) corresponds. In practice, this a priori information evolves from reliability analysis of the total system in which the device is to be used. Therefore, when the bounds obtained by iterative extrapolation exceed the acceptable range for f(x), either the device must be rejected, or additional measurements must be made. The extension of iterative extrapolation to the multiple measurement case will be

considered shortly, but first it is necessary to discuss the significance of the bounds obtained by the iterative extrapolation.

When iterative extrapolation bounds are constructed as in Figure 3.4, it is necessary to consider the question of what confidence level can be associated with these composite bounds. The answer to this question is complex and requires the assessment of the nature of the combination of information that occurs in the iterative process. Although the following chapter is devoted to the study of this question, subsequent discussions in this section will require a brief inspection of this topic.

Consider equation (3.16). The values $a(x_i)$ and $b(x_i)$ represent the lower and upper TL's of $f'(x_i)$ respectively. One can see that (3.16) corresponds to

$$f(x_N) = f(x_o) + \Delta x \cdot \sum_{i=o}^{N-1} f'_L(x_i) \qquad (3.17)$$

where $f'_L(x_i)=a(x_i)$ for the lower limit and $f'_L(x_i)=b(x_i)$ for the upper limit of $f(x_N)$. Consider that the values of $f'(x_i)$ and $f(x_N)$ can be viewed as RV's in (3.17). If the $f'(x_i)$'s are assumed independent, then equation (3.17) is a linear combination of independent RV's. In order to assess the probability associated with a TI for $f(x_N)$, one usually must first determine the pdf of $f(x_N)$. It will be seen in the next chapter that this involves an N-fold convolution of the pdf's of $f'(x_i)$. Once the pdf of $f(x_N)$ has been

obtained, the confidence level of any TI can be determined by integrating the pdf between the two limits of the TI.

In practice, it is often advisable to avoid the above procedure, as will be explained shortly. A different approach involves the linear combination of the TI's for the $f'(x_i)$ in (3.17). In order to explain this second approach, it is useful to make the following definition.

Definition 3.4 (Bounding Algorithm) Consider the linear combination of RV's

$$Y = \sum_{i=o}^{N} C_i \cdot X_i \tag{3.18}$$

where the C_i's are constants and the X_i's are RV's with known pdf's. Corresponding to the confidence level α, let the TI's for each of the X_i be determined according to the convention of equation (3.4), such that

$$p\{a_i \leq X_i \leq b_i\} = \alpha, \text{ for all } i.$$

The "composite tolerance interval" (CTI) for Y is given by

$$\left(\sum_{i=o}^{N} \min[C_i \cdot a_i, C_i b_i] \leq Y \leq \sum_{i=o}^{N} \max[C_i a_i, C_i b_i] \right) \tag{3.19}$$

where

$$\min[u,v] \equiv \begin{cases} u, & u \leq v \\ v, & \text{otherwise} \end{cases}$$

and

$$\max[u,v] \equiv \begin{cases} u, & u \geq v \\ v, & \text{otherwise} \end{cases}$$

This process of combination is referred to hereafter as the "bounding algorithm."

It is first noted that equation (3.16) may be written as a special form of (3.19) by letting $C_i=\Delta x$ for all i and noting that $a(x_i)\leq b(x_i)$ by assignment. In this case, Y corresponds to $f(x_N)-f(x_o)$. Notice that the construction of Figure 3.4 is an application of the bounding algorithm.

The CTI produced by the bounding algorithm represents an enormous simplification over the analytic techniques involving convolution and integration, provided a confidence level can be associated with this bound. In the following chapter, the association with a confidence level will be investigated in detail. It will be shown that in many important cases, the CTI confidence level will be at least α. The reader is cautioned that although it appears that the CTI is a "worst case" combination, there are situations in which the confidence level of this bound is less than α. Also, there are a number of other possible ways to combine TI's. Some of these combinations will be considered as the topic of the next section. In order to refer to the case where the CTI confidence level is at least that of the individual TI's combined using the bounding algorithm, it is useful to consider the following definitions.

Definition 3.5 (Linear Conformity) Consider the linear combination of RV's, $x_1,\ldots,x_N$

$$Y = \sum_{i=1}^{N} C_i X_i.$$

Suppose the RV's are statistically independent. If a CTI for Y is obtained using the bounding algorithm and confidence level α corresponding to the individual TI's of the algorithm, then the collection of the pdf's for the X_i's are said to be linearly conformal, if and only if the confidence level corresponding to the CTI is at least α, for all values of the C_i's and with no restriction on the value of α.

When the N pdf's are linearly conformal, and in addition, have the same general distribution form, with perhaps, different distribution parameters, this common distribution is said to be "self-linearly conformal."

It will be possible, in most cases, to refer to the two defined conditions as "linear conformity." The reader will be able to infer the distinction between a collective property of several different general pdf's, and that of a property of a single pdf from the context of discussion.

It is now possible to return to the main development of the iterative extrapolation technique. Referring to equation (3.16), consider the situation in which the iteration interval Δx is allowed to approach 0. If the pdf of f'(x) is known in the interval $(x_o<x<x_r)$, as previously assumed,

then it is clear that the TL's of f'(x) as defined by equation (3.4) can be written as functions of x in that interval; i.e. a(x) and b(x). Consider the upper limit of (3.16),

$$f(x_N) \leq f(x_o) + \sum_{i=o}^{N-1} b(x_o+i\cdot\Delta x)\cdot\Delta x \tag{3.20}$$

in which the substitution $x_i = x_o+i\cdot\Delta x$ has been made. Suppose the function b(x) is continuous over the domain $x_o \leq x \leq x_r$. Let

$$x_o<x_1<\ldots<x_{N-1}<x_N=x_r$$

be defined such that

$$x_{i+1}-x_i=\Delta x=(x_r-x_o)/N \tag{3.21}$$

Let

$$S_N = \sum_{i=o}^{N-1} b(x_i)\cdot\Delta x.$$

Then from the Fundamental Theorem of integral calculus (Thomas [1962]), letting $N \to \infty$ and $\Delta x \to 0$,

$$\lim S_N = \int_{x_o}^{x_r} b(x)dx. \tag{3.22}$$

Realizing that the same analysis is valid for a(x) continuous in $x_o \leq x \leq x_r$, one may write (3.16) as

$$f(x_o) + \int_{x_o}^{x_r} a(x)dx \leq f(x_r) \leq f(x_o) + \int_{x_o}^{x_r} b(x)dx. \tag{3.23}$$

Several observations can be made about (3.23). The first is that by going to a differential iteration interval, the CTI for f(x) can be obtained in terms of two definite integrals and the deterministic value $f(x_o)$. Since the

integrals depend only on the values x_o and x_r and the functions a(x) and b(x), a more convenient form of (3.23) can be obtained. For some arbitrary point x_p within $[x_o, x_r]$, one may write the indefinite integrals

$$A(x) = \int a(x)dx \tag{3.24a}$$

and

$$B(x) = \int b(x)dx \tag{3.24b}$$

Then (3.23) can be written as

$$f(x_o)+A(x_r)-A(x_o) \leq f(x_r) \leq f(x_o)+B(x_r)-B(x_o) \tag{3.25}$$

For a CTI corresponding to points to the left of x_o, one similarly obtains

$$f(x_o)+B(x_1)-B(x_o) \leq f(x_1) \leq f(x_o)+A(x_1)-A(x_o) \quad \text{for } x_1 \leq x_o \tag{3.26}$$

Since one can predetermine the integral functions A(x) and B(x) independently of $f(x_o)$, construction of a CTI for f(x) can be further simplified in practice. A way of determining the continuous statistical bound [A(x),B(x)] will be considered shortly. First, it must be considered how going to a differential iteration interval affects the rate of divergence of the CTI. Referring to Figure 3.4, consider a single iteration interval in which Δx is chosen so small that the TL's for f'(x) in this interval are essentially constant for a fixed α. Suppose the interval is subdivided into two equal intervals of x. In this case, one may see that iteration over these two subintervals will produce a CTI identical to a simple statistical bound obtained as in

Figure 3.1. Therefore, the subdivision did not increase the divergence of the bound on f(x). This result indicates that the divergence is not a function of the number of iterations, but rather a function of the overall length of the interval in which f(x) is bounded.

The results of equations (3.23), (3.25), and (3.26) may be summarized in a theorem.

Theorem 3.6 Corresponding to some confidence level α, let the TL's for the derivative f'(x) be given as functions of x, a(x) and b(x) according to the convention of (3.4), such that $x_l \leq x \leq x_r$. Assume that the pdf of f'(x) varies in $[x_l, x_r]$ only to the extent that its parameters are allowably functions of x. Given the functional value of $f(x_o)$, where x_o is in $[x_l, x_r]$, if the pdf of f'(x) is self-linearly conformal, for any N in (3.20), then f(x) is statistically bounded by the CTI

$$f(x_o) + \int_{x_o}^{x} a(u)\,du \leq f(x) \leq f(x_o) + \int_{x_o}^{x} b(u)\,du, \quad x_o \leq x \leq x_r \tag{3.27a}$$

and

$$f(x_o) + \int_{x_o}^{x} b(u)\,du \leq f(x) \leq f(x_o) + \int_{x_o}^{x} a(u)\,du, \quad x_l \leq x \leq x_o \tag{3.27b}$$

with a corresponding confidence level of at least α. Furthermore, using the integral functions of equation (3.24) the bound of (3.27) is also written as

$$f(x_o)+A(x)-A(x_o)\leq f(x)\leq f(x_o)+B(x)-B(x_o), \quad x_o\leq x\leq x_r \tag{3.28a}$$

$$f(x_o)+B(x)-B(x_o)\leq f(x)\leq f(x_o)+A(x)-A(x_o), \quad x_l\leq x\leq x_o \tag{3.28b}$$

With the exception of two considerations, this theorem has been proven above. One of these considerations corresponds to the implied requirement that the values of f'(x) at any two points in the interval $[x_l,x_r]$ are statistically independent. Obviously, this is impossible since the two values are related by the function f(x) itself. In the next section, it will be shown that the assumption of independence causes no problem in the validity of Theorem 3.6. In fact, for the extreme case in which one derivative value is a deterministic function of the other, the confidence level for the CTI of (3.28) can be equal to α.

The second comment on Theorem 3.6 concerns the implied generalization of Definition 3.5 regarding linear conformal pdf's. That is, it has been required that the linear combination of <u>any</u> N independent RV's leads to a linearly conformal situation.

In the next chapter, this generalization is proven valid for an important class of pdf's. It is also shown heuristically why this generalization can be expected to be valid for other pdf's.

It may occur to the reader that Theorem 3.6 can be extended to provide bounds on f(x) between a pair of deterministic data points. Such an extension can be accomplished in much the same way that Theorem 3.1 was extended.

Consider Figure 3.5, in which deterministic measurements on $f(x)$ have been obtained at x_o and x_1. At each of these two points, a CTI for $f(x)$ in $[x_o,x_1]$ and corresponding to the same confidence level α is obtained using Theorem 3.6. Consider an x_m between x_o and x_1. In general, the two statistical bounds on $f(x)$ at x_m have a non-void intersection, but will not coincide. In determining how to combine these two bounds for $f(x_m)$, one may note that the same situation exists here that was faced in Figure 3.2 at $x_o+\Delta x$. Envoking Theorem 3.3, assuming $\alpha>\frac{1}{2}$, one concludes that the intersection of these two bounds should be used as the overall bound on $f(x_m)$. If this is done at each point in $[x_o,x_1]$, the bound on $f(x)$ in this interval is the intersection of the two regions bounded by the separate CTI's. The combined CTI is thus the cross-hatched region in Figure 3.5. Note that the same comments about non-intersection and excessive bounds that were given for the construction of Figures 3.2 and 3.3 apply to Figure 3.5. That is, the conditions of non-intersection and excessive bounds can be used to indicate the need for measurement of $f(x)$ in the regions having such conditions.

The remainder of this section will deal with practical considerations of the assumptions and requirements of preceding developments. At the heart of these developments has been the need to know the pdf of $f'(x)$, usually as a continuous function of x. Except when each member of a

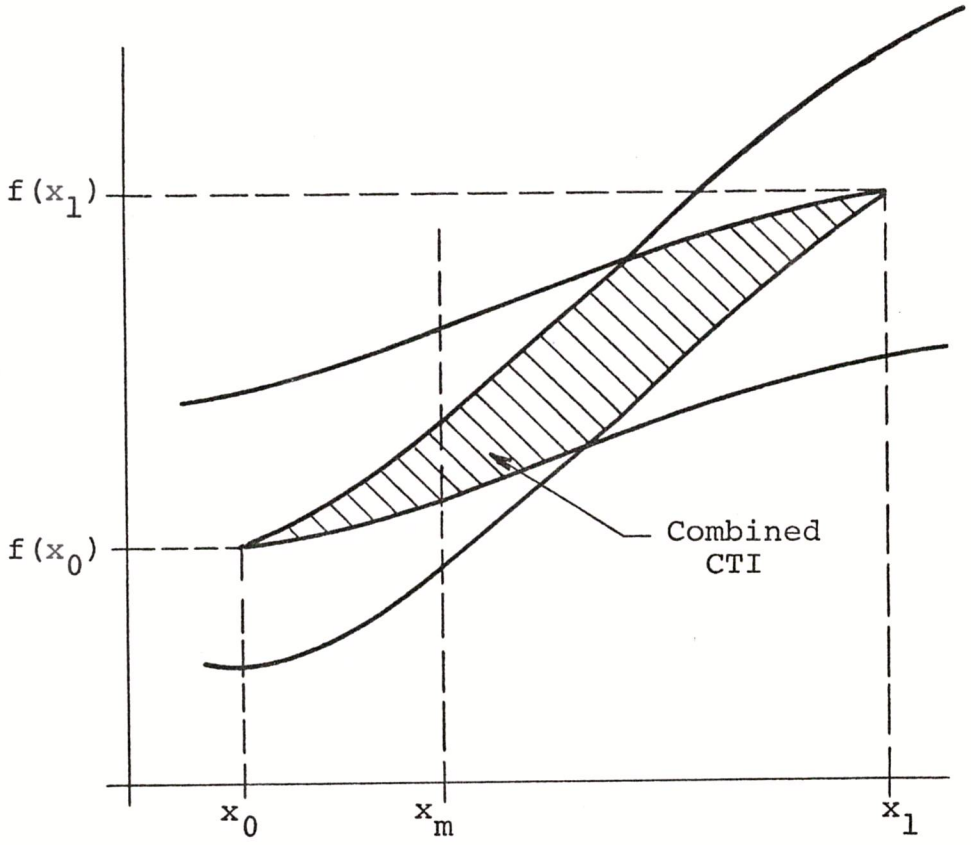

Figure 3.5 Extension of Theorem 3.6 to the case where two deterministic data points are used to form a statistical bound on f(x). The cross-hatched region represents the combined CTI for f(x) between x_0 and x_1. Note that x_m is any point between x_0 and x_1. At x_m the combined CTI is the intersection of the two separate CTI's.

sample space (population) of a RV can be observed, the probability distribution of the RV cannot be measured. However, it is often possible to determine a statistical estimate of the distribution which will suffice for practical purposes.

In order to show reasons for avoiding the determination of a pdf, consider how one obtains an estimate of a pdf for $f'(x_o)$. If the pdf is not assumed to be of a standard form, then estimation involves measuring $f'(x_o)$ for many similar devices[1] and plotting a frequency distribution curve of the measurement values. Alternatively, a cumulative distribution function is often plotted in practice. To conveniently represent the result of this frequency plot, curve-fitting techniques can be employed to obtain a function which approximates the frequency distribution curve. A further approximation results by subsequently using this function as the pdf for $f'(x_o)$.

It is easy to see why the approach to estimating the pdf of $f'(x_o)$ is unsatisfactory for application to the developments of this section. A separate frequency plot is required at each x that the pdf is required to be known. Since the number of values of x that the pdf is used can be enormous and since the above estimation procedure does not allow for any interpolation between points, other estimation

---

[1]It is assumed that the function f(x) corresponds to a device belonging to a very large population of devices of the same device type.

approaches are necessary. Fortunately, one can often find a standard distribution function to represent the statistical data. Estimation of the pdf then involves determining estimates of the distribution parameters. Usually these estimates are in the form of averages. As an example, $f'(x_o)$ may have a nearly normal distribution. If M devices are tested to determine their corresponding measured values of $f'(x_o)$, $y_i'(x_o)$, then the distribution mean $\mu(x_o)$ and variance $\sigma^2(x_o)$ may be estimated by (Bowker and Lieberman [1964])

$$\mu_e(x_o) = \frac{\sum_{i=1}^{M} y_i'(x_o)}{M} \tag{3.29}$$

$$\sigma_e^2(x_o) = \frac{\{\sum_{i=1}^{M} y_i'^2(x_o)\} - \{\sum_{i=1}^{M} y_i'(x_o)\}^2/M}{M-1} \tag{3.30}$$

where the subscript e indicates that the results are estimates. One notes that since the estimates are effectively sums and sums of squares of continuous functions, they must also be continuous (Thomas [1962]). Since these estimates involve averages of functions, one intuitively expects them to vary more slowly than the functions themselves. In order to represent the pdf of $f'(x)$ over an interval of x, the parameter estimates can be determined at regular intervals of x, and then conventional interpolation techniques can be used to provide values of the parameters between these points. One must determine, by inspection of the gathered statistical data, the magnitude of the spacing for

these interpolation points by observing how rapidly the parameter function changes between points.

The "standard function" approach to determining an estimate of the pdf of f'(x) is easier to implement than the more general approach first considered. However, a difficulty arises in determining how close the actual distribution is to the standard form. Moreover, the techniques of this section require only the TI for f'(x) rather than the pdf itself. Since one usually can decide ahead of time what value confidence level to use, it appears that an estimation of the TI's directly from the statistical data is preferred. The TL's are determined at regular intervals of x, and then conventional interpolation of these values is used to represent the TL's between these interpolation points. One must inspect the statistical data in order to determine the appropriate spacing of the interpolation points. This spacing will, of course, influence the amount of preliminary testing of the sampled devices. It is also important to note that for each value of the confidence level α, a separate TI must be determined at each x. Usually one is interested in only one or two values of α.

The determination of the TL's from the statistical data can be accomplished using well established methods of estimation theory. This area of statistics is a topic in most applied statistics texts and will not be presented here in any detail. It is worth indicating two general approaches

to estimating TL's. The first assumes that the pdf is Gaussian-normal. The TL's for this "two-sided" TI is determined using the sample mean, sample variance, number of samples, confidence level α, and another probability β. The second probability corresponds to the assurance that at least $100(\alpha)\%$ of the actual distribution will lie within the estimated TI (Bowker and Lieberman [1964]). The second estimation approach may be used when no assumption about the underlying distribution is made. Assume at some point x_i, the statistic corresponding to the m^{th} largest measured value of $f'(x_p)$ and the statistic corresponding to the m^{th} smallest measured value of $f'(x_p)$ for a sample of M devices are determined. Then the estimate of the TI at this x_p is determined by both of these "order" statistics as TL's. The probability β that this estimate contains at least $100(\alpha)\%$ of the actual distribution for $f'(x_p)$ is given by (Guttman and Wilks [1965])

$$\beta = 1 - e^{-M(1-\alpha)} \{ \sum_{i=o}^{2m-1} [M(1-\alpha)]^i / i! \} \qquad (3.31).$$

provided that β and M are large.

Suppose one wishes to determine an estimate of the TI for $f'(x)$ at some point x_p corresponding to an $\alpha = .98$ for a sample of M=400 devices. Assume that one wants to be at least 95% sure that this estimate will actually contain at least $100(1-\alpha)\%$ of the $f'(x_p)$ distribution. By trial and error one may verify that in this case m should be equal to

2. That is, the second highest and second lowest sample of $f'(x_p)$ can be used as the TL's. There are obviously other ways to use (3.31), but in any case one must use some intuitive judgment in deciding the ultimate values of α and β.

The estimation techniques that have been just indicated should serve only as examples of how to practically acquire the statistical data required by this section's developments. There are other applicable approaches to be found in the general area of estimation and other statistics. It is clear that an initial concerted measurement effort is needed to provide one with the required statistical information. The benefits of such an investment is realized when many thousands of devices must be later tested to determine their device functions.

A second central requirement of the general statistical approach of this section is that the derivative of the device function can be accurately measured. In some situations the derivative may be difficult to measure since one needs to measure the device function at nearly equal values of x. When there is an accuracy problem, it may be impossible to obtain the required statistical information. The reader is advised that it is assumed that accurate measurements can be made, although possibly at a high economic cost. This cost can result from expensive instrumentation, time-consuming calibration and operator intervention, and the amount of time to perform the measurements themselves.

For example, high accuracy analogue measurements require a larger conversion time when analogue-to-digital converters are employed in the measurement. It is assumed that the initial high-cost statistical data gathering effort can be justified by a subsequent savings in the complexity and the increased throughput of production and field testing systems. The initial testing effort might involve hundreds of devices while production testing may involve tens of thousands of devices.

In conclusion of this section, it is seen that several new techniques have been developed that are potentially useful for device testing applications. By providing statistical bounds on the device function f(x), the number of measurements required to characterize a particular device may be reduced, in some cases, from the number required by conventional methods. This reduction is seen to result from the incorporation of statistical information not utilized by conventional techniques.

If one examines the techniques used in practice to characterize a device function, it is possible to conclude that these techniques are often arbitrarily devised. Usually, manufacturers test devices by measuring the device function at a few "typical" points along the device curve. Acceptance of the device corresponds to these functional values lying within "guaranteed" bounds. These bounds often result from the manufacturers' knowledge that a certain

per cent of the devices produced will meet these limits. When the devices are to be used under non-typical conditions or over a variety of conditions, a user must sometimes hand pick devices out of a lot by subjecting the devices to more informative tests. In actually determining the function for a device, one may often sample the device function at close intervals. The closeness of samples is determined by trying different spacing to determine the appropriate spacing that gives sufficient accuracy for the user's purposes. Once this is done, curve-fitting techniques may be applied. Except for the determination of spacing and the selection of an interpolation formula, statistical information is not used.

The general approach developed in this section is therefore a departure from conventional practices. These new techniques not only provide a statistical bound on $f(x)$, but more importantly provide insight into the basic problem of generating testing procedures. An indication is given of what statistical information to gather and how to organize it. When the continuous bounds of Theorem 3.6 are constructed between deterministic measurement points, a piecewise bound on $f(x)$ can be constructed which is independent of any assumption of order of approximation for $f(x)$. Should the statistical bounds extend into a user defined unacceptable region, one may make additional measurements of $f(x)$ in these regions until the bound on $f(x)$ is totally within the

acceptability region or else the statistical bound is totally outside this acceptability region for some interval of x. This latter condition would indicate that the device should be rejected.

There are obvious variations on the applications of the methods developed in this section. These variations primarily concern where to make additional measurements of f(x). It is seen that by working with statistical bounds, there can be some basis for a testing procedure. It is also obvious that the number of measurements required to characterize f(x) will be determined not only by where each measured value lies but also the tightness of the acceptability region. In this sense, the proposed general approach is seen to be optimal and adaptive in that measurements are made only when and where it appears statistically necessary. The suitability of a statistical bounding method depends on the particular application, whether the cost of making measurements on devices is a critical problem, the feasibility of gathering required statistical information, and whether the required computations can be mechanized economically in an automatic testing system.

The techniques developed here will be further developed to apply to the situation in which polynomial and transcendental interpolation is known to be a suitable method of characterizing a device function. This extension will be investigated in Chapter 5. In addition to the statistical

data utilized in the preceding techniques, the prior information regarding order of approximation and bandwidth properties associated with the device function will be incorporated into the process of obtaining statistical bounds on f(x).

3.3 Discussion of the Bounding Algorithm and Some Alternatives

In the last section, a need was seen to arise for a convenient way in which to obtain a TI for a linear combination of independent random variables. In Chapter 5, this need will again arise from the techniques to be developed there. Moreover, in many other situations, one is faced with obtaining a bound for a physical quantity which can be represented by a linear combination of independent, statistical quantities. When the statistical distribution for these independent quantities is known explicitly or when tolerances (TI's) are known for the quantities, the availability of a convenient method of determining a statistical bound on the linear combination from the individual statistics is seen to be of importance.

In this section, conventional approaches to statistically bounding linear combinations of independent RV's will be briefly considered. The reasons why such techniques should be avoided in some situations will be considered. As an alternative to these conventional approaches, the bounding algorithm previously defined in Defintion 3.5 will be

discussed from the standpoint of practicality. The need for this algorithm will be further defended because of the requirements of its application made in the last section and its additional use in Chapter 5. In addition to the bounding algorithm, several alternative algorithms that also use the direct combination of TI will be compared to this first algorithm. The basis of this comparison will be the expected performance of these algorithms in obtaining practical statistical bounds on device functions.

Consider the linear combination of independent RV's

$$Y = \sum_{i=1}^{N} C_i X_i \tag{3.32}$$

where the C_i's are constants and the X_i's are the RV's. In general, to determine a TI for the RV, Y using a conventional approach, one must first determine the pdf of this variable. If the pdf's of the RV's X_i are known, then one must usually perform a weighted N-fold convolution of these pdf's.[1] That is, if $g_i(u)$ is the pdf for X_i, then the pdf of $C_i X_i$ is $g_i(\frac{u}{C_i})/C_i$ and the pdf of Y denoted $g_Y(u)$ is (Papoulis [1962])

---

[1]The exception to this is if each of the X_i's are normally distributed, with mean u_i and variance σ_i^2. In this case, Y is normally distributed with mean $= \sum_{i=1}^{N} C_i u_i$ and variance $= \sum_{i=1}^{N} C_i \cdot \sigma_i^2$. (Bowker & Lieberman [1964])

$$g_Y(u) = \frac{\{g_1(u/C_1)*g_2(u/C_2)*\ldots*g_N(u/C_N)\}}{\prod_{i=1}^{N} |C_i|} \tag{3.33}$$

where '*' denotes the convolution operator. Equation (3.33) normally requires an N-fold integration. The computation of $g_Y(u)$ is more conveniently accomplished by Fourier Transforms, since convolution in the real domain corresponds to multiplication in the frequency domain. If the Fourier Transform of $g_i(u)$, $G_i(\omega)$, exists, that is

$$g_i(u) \leftrightarrow G_i(\omega) \tag{3.34}$$

then one can show [Papoulis [1962])

$$\frac{g_i(\frac{u}{C_i})}{|C_i|} \leftrightarrow G_i(C_i\omega)$$

and

$$g_y(u) \leftrightarrow \prod_{i=1}^{N} G_i(C_i\omega). \tag{3.35}$$

The result of equation (3.35) indicates that one first determines the Fourier Transforms of the N, $G_i(\omega)$, then scales the transforms with the C_i, next multiplies the N scaled transforms together, and finally performs an inverse transform to obtain $g_Y(u)$. Since the operation of transformation involves integration which must be performed numerically if computation is digital, one resorts to the various Fast Fourier Transform (FFT) algorithms to avoid large computational time (Cooley, Lewis & Welch [1967]). Such algorithms are easily mechanized but require both appreciable

storage space and computational time allocations in a minicomputer testing system. Special hardware FFT processors are becoming available so that numerical convolution is not an unreasonable approach in some situations.

There are, however, several reasons for avoiding the above direct approach even when a FFT processor is available. As was mentioned in the last section, it is considerably more difficult to obtain the entire pdf of a device function or its derivative than to obtain an estimate of TL's. Further, since the distribution information must be determined at many values of the input variable, a considerable amount of memory would be required to store the statistical information in the form of pdf's. For example, to store a pdf with 256 point resolution, at 100 different values of the input variable would require 25,700 words of memory (including the values of the input variable), while to store the TL's corresponding to three confidence levels, at 100 input variable values would require 700 words.

A greater disadvantage of the direct convolution approach is seen to exist when one wants to vary one or more of the "constants" C_i. When statistically based polynomial interpolation is considered in Chapter 5, the C_i's will be deterministic functions of the input variable. In order to obtain bounds on f(x) over an interval of this variable, the C_i's will be seen to change continuously over this interval. Where convolution is used to obtain a TI for Y in equation

(3.32), all the transformation operations, transform scaling, multiplications, and inverse transformations would have to be performed as many times as there are points at which the bound of f(x) needs to be known. Thus, direct use of pdf's even with a high speed FFT processor could easily result in prohibitive computational time.

Still another disadvantage of working directly with the pdf's is that of aliasing error (Hamming [1962]). This error occurs when the FFT is used for transforming a function which is not ideally bandlimited. The functions being transformed are pdf's and they rarely have ideal cut-off characteristics. In fact, some distributions such as the Cauchy distribution have non-finite second moments. Thus, some aliasing error will usually occur. The most pronounced effect of this error occurs in the "tails" of $g_Y(u)$. Since one usually picks a confidence level close to unity for practical applications, the TL's obtained by integration of $g_Y(u)$ are very sensitive to small variations of the area under the tails of $g_Y(u)$. Thus, even a small amount of aliasing error can be disasterous in determining the TL's.

It is therefore evident that a method of directly combining the TL's of the individual RV's avoids several important problems that arise in the direct manipulation of pdf's. Nevertheless, one needs to be aware of the general characteristics of the pdf's corresponding to the X_i's since such information is necessary in evaluating the approximate

confidence level of the TI formed by the bounding algorithm. In the next chapter, such evaluation considerations will be studied in detail.

In determining a statistical bound on the device function f(x), one wants to be able to give some assurance or probability that f(x) actually lies within the bound. It seems reasonable then that if a confidence level is to be associated with the bound on f(x), and this confidence level is only approximately determined, then this estimate should be conservative. That is, the estimate of the confidence level is only meaningful if it is a lower limit on the actual confidence level for the bound on f(x). Therefore, a requirement of any potential algorithm for combining TL's is that one may claim a lower bound confidence level for the TI that results from application of the algorithm. In particular, the conditions under which this minimum confidence level can be claimed must be known.

The bounding algorithm (BA) of Definition 3.5 fulfills the above requirement whenever the pdf's of the RV's of equation (3.32) are linearly conformal.[1] In this case, if the TI's of each X_i have a confidence level of α, then the composite TI resulting from the combination of the individual TI's via the BA has a confidence level of at least α.

---

[1] See Definition 3.6

The conditions of linear conformal pdf's are considered in detail in the next chapter.

In addition to the preceding requirement, the BA has other important advantages. This algorithm forms composite TL's by effectively forming a linear combination (in the algebraic sense) of the individual TL's. Therefore, the expressions for the composite TL's bear close resemblance to the RV equation (3.32). The contributions of the individual TL values to the composite TL's are therefore functionally independent. That is, the partial derivative of either of the two composite TL's with respect to any of the individual TL's is not a function of the other individual TL's. As a result of the algebraic linearity of the BA, the interchange of deterministic data and statistical data is especially straightforward. In the last section, the constructions of iterative statistic extrapolation were simplified by the ability to consider the iteration as a recursive process. The linearity also allowed the transition of the iteration to proceed from a discrete to a continuous process via Theorem 3.6. When polynomial and transcendental interpolation is considered in Chapter 5, the algebraic linearity of the BA will result in a relatively simple analysis of the composite TI's.

Another advantage of the BA relates to its ability to handle the situation in which the RV's of equation (3.32) are not statistically independent. In order to analyze the

effect of statistical dependence, consider the N=2 case for (3.32)

$$Y = X_1 + C_2X_2. \tag{3.36}$$

Suppose the two RV's are functionally related by

$$X_2 = R(X_1). \tag{3.37}$$

Corresponding to the confidence level α, suppose the TL's for X_1 are a_1 and b_1, $a_1<b_1$. That is

$$p\{a_1 \leq X_1 \leq b_1\} = \alpha. \tag{3.38}$$

Consider the case where R is monotonic. Then if $a_1 \leq X_1 \leq b_1$, $a_2 \leq X_2 \leq b_2$ where

$$a_2 = \min [R(a_1), R(b_1)]$$

$$b_2 = \max [R(a_1), R(b_2)]$$

and[1]

$$p\{(a_2 \leq R(X_1) \leq b_2) \,|\, (a_1 \leq X_1 \leq b_1)\} = 1. \tag{3.39}$$

For the constant C_2 in equation (3.36), let

$$a_3 = \min [a_2\ C_2,\ b_2\ C_2]$$

$$b_3 = \max [a_2\ C_2,\ b_2\ C_2].$$

Then it follows from (3.39) that

$$p\{(a_3 \leq C_2 \cdot R(X_1) \leq b_3) \,|\, (a_1 \leq X_1 \leq b_1)\} = 1, \tag{3.40}$$

or in simpler notation

$$p\{E_2 | E_1\} = 1$$

where

E_2 denotes the event $a_3 \leq C_2 \cdot R(X_1) \leq b_3$

---

[1] $p\{(R_1) \mid (R_2)\}$ denotes $p\{(R_1)$ given that $(R_2)\}$

and

E_1 denotes the event $a_1 \leq X_1 \leq b_1$.

Then the joint occurrence of E_1 and E_2 may be combined as

$$a_1+a_3 \leq X_1+C_2 \cdot R(X_1) \leq b_1+b_3. \tag{3.41}$$

But

$$p\{E_1 \text{ and } E_2\} = p\{E_2|E_1\} \cdot p\{E_1\} = 1 \cdot \alpha = \alpha. \tag{3.42}$$

Observe that if (3.37) had been substituted directly in (3.36), the expression bounded in (3.41) would be the RV, Y. The lower and upper limits correspond to the TL's obtained by the BA under the assumption of X_1 and X_2 independent. One may further observe that if

$$C_2 \cdot \frac{\partial R(X_1)}{\partial X_1} \geq 0 \tag{3.43}$$

then $a_3 = C_2R(a_1)$ and $b_3 = C_2R(b_1)$. In this case the limits of (3.41) are the same that result from the direct substitution of a_1 and b_1 in

$$Y = X_1 + C_2R(X_1).$$

Therefore, the BA provides "worst case" TL's for X_1 and X_2 monotonically dependent. That is, for equation (3.38) true and R monotonic,

$$p\{a_1+a_3 \leq Y \leq b_1+b_3\} \geq \alpha. \tag{3.44}$$

Further, for the situation of (3.43),

$$p\{a_1+a_3 \leq Y \leq b_1+b_3\} = \alpha. \tag{3.45}$$

It is possible to generalize the preceding discussion into a Theorem.

Theorem 3.7 Given the RV equation

$$Y = X_1 + \sum_{i=2}^{N} C_i \cdot X_i \tag{3.46}$$

in which the C_i's are deterministic constants. Let the TI for X_1 be given for the confidence level α, such that

$$p\{a_1 \leq X_1 \leq b_1\} = \alpha.$$

If the RV's, X_i are monotonic, deterministic functions of the RV, X_1 such that

$$X_i = R_i(X_1)$$

then the constants

$$\begin{aligned} a_i &\equiv \min [C_i \cdot R_i(a_1), C_i \cdot R_i(b_1)] \\ b_i &\equiv \max [C_i \cdot R_i(a_1), C_i \cdot R_i(b_1)] \end{aligned} \tag{3.47}$$

determine the TI's for $C_i X_i$ such that

$$p\{a_i \leq C_i X_i \leq b_i\} = \alpha \quad \text{for } i=2,\ldots,N. \tag{3.48}$$

Furthermore, Y may be bounded by

$$a_Y = \sum_{i=1}^{N} a_i \leq Y \leq \sum_{i=1}^{N} b_i = b_Y \tag{3.49}$$

such that

$$p\{a_Y \leq Y \leq b_Y\} \geq \alpha.$$

Furthermore, if

$$C_i \frac{\partial R(X_1)}{\partial X_1} = \geq 0 \quad \text{for all } i=2,\ldots,N$$

then $p\{a_Y \leq Y \leq b_Y\} = \alpha$.

The bound of (3.49) corresponds to the CTI formed by the BA.

The proof of this theorem is a straightforward extension of the N=2 case by induction. All that one must recognize is that (3.48) monotonically maps the TI for X_1 onto each of the TI's for X_i via R_i.

Consider the effect of a different situation of statistical dependence on the performance of the BA. In this situation, a linear combination of RV's contains some variables that are linear functions of one of the other RV's in the combination.

Theorem 3.8 Suppose the RV's, $X_{1,1}$, $X_{2,1}$,...,$X_{N,1}$ are statistically independent and have pdf's which are linearly conformal.[1] Define another set of RV's as

$$X_{i,j} \equiv D_{i,j} X_{i,1} + L_{i,j}, \quad j=2,\ldots,k_i \tag{3.50}$$

where the k_i's are non-negative integers, and the $D_{i,j}$'s and $L_{i,j}$ are constants. That is, the $X_{i,j}$'s are linear functions of $X_{i,1}$. Then the pdf's of all the $X_{i,j}$ are linearly conformal. That is, given the linear combination

$$Y = \sum_{i=1}^{N} \sum_{j=1}^{k_i} C_{i,j} \cdot X_{i,j} \tag{3.51}$$

where $C_{i,j}$ are arbitrary constants, and given the TI's for each of the $X_{i,j}$ such that

$$p\{a_{i,j} \leq X_{i,j} \leq b_{i,j}\} = \alpha, \quad \begin{matrix} i = 1,\ldots,N \\ j = 2,\ldots k_i \end{matrix} \tag{3.52}$$

---

[1] That is, the collection of pdf's is linearly conformal.

one may apply the BA to (3.51) without the prior knowledge of (3.50) and obtain a CTI for Y. The probability that Y lies in this CTI will be at least α.

The following lemma will be useful in establishing the proof of Theorem 3.8.

Lemma 3.9 If the variable X is bounded by

$$a \leq X \leq b \tag{3.53}$$

then for any constants C_i, i=1,...,N

$$\sum_{i=1}^{N} \min[a \cdot C_i, b \cdot C_i] \leq X \cdot \sum_{i=1}^{N} C_i \leq \sum_{i=1}^{N} \max[a \cdot C_i, b \cdot C_i]. \tag{3.54}$$

Furthermore, if all the C_i's have the same sign, then

$$\sum_{i=1}^{N} \min[a \cdot C_i, b \cdot C_i] = \min[X \cdot \sum_{i=1}^{N} C_i] \quad \text{for } X \varepsilon [a,b] \tag{3.55a}$$

and

$$\sum_{i=1}^{N} \max[a \cdot C_i, b \cdot C_i] = \max[X \cdot \sum_{i=1}^{N} C_i] \quad \text{for } X \varepsilon [a,b] \tag{3.55b}$$

Proof The proof of (3.54) can be divided into two separate parts. Each part may be accomplished by induction.

Part 1. [Left-hand side of (3.54)]

For the N=1 case, consider that from (3.53) the minimum value of $X \cdot C_1$ is

$$\min[a \cdot C_1, b \cdot C_1].$$

Therefore, for N=1, the left side of (3.54) and equation (3.55a) are satisfied.

Suppose

$$\sum_{i=1}^{N-1} \min[a\cdot C_i, b\cdot C_i] \leq X\cdot \sum_{i=1}^{N-1} C_i$$

is true. To each side of this inequality add $X\cdot C_N$. Then

$$\sum_{i=1}^{N-1} \min[a\cdot C_i, b\cdot C_i] + X\cdot C_N \leq X\cdot \sum_{i=1}^{N-1} C_i .$$

But, from the case for N=1

$$\min[a\cdot C_N, b\cdot C_N] \leq X\cdot C_N .$$

Therefore

$$\sum_{i=1}^{N} \min[a\cdot C_i, b\cdot C_i] \leq X\cdot \sum_{i=1}^{N} C_i .$$

Part 2. [Right-hand side of (3.54)].

Replace "min" by "max", and "$\leq$" by "$\geq$", in part 1. Then

$$\sum_{i=1}^{N} \max[a\cdot C_i, b\cdot C_i] \geq X\cdot \sum_{i=1}^{N} C_i .$$

To prove (3.55), let all the C_i be positive. Then

$$\min[a\cdot C_i, b\cdot C_i] = a\cdot C_i, i=1,\ldots N$$

$$\max[a\cdot C_i, b\cdot C_i] = b\cdot C_i, i=1,\ldots N$$

and

$$\min[X\cdot \sum_{i=1}^{N} C_i] = a\cdot \sum_{i=1}^{N} C_i$$

$$\max[X\cdot \sum_{i=1}^{N} C_i] = b\cdot \sum_{i=1}^{N} C_i .$$

Then for positive C_i's, (3.55) is proved. Similarly, letting the C_i's be all negative completes the proof of the lemma.

Proof of Theorem 3.8 Consider the linear combination of RV's, $X_{i,1}$

$$Y = \sum_{i=1}^{N} C_{i,1} \cdot X_{i,1} \tag{3.56}$$

From the assumption that the pdf's of $X_{i,1}$ are linearly conformal, and using the TI's indicated in (3.52) for the j=1 case, it is required that, for any $C_{i,1}$'s,

$$p\{E_1\} \equiv p\{\sum_{i=1}^{N} \tilde{a}_{i,1} \leq \sum_{i=1}^{N} C_{i,1} \cdot X_{i,1} \leq \sum_{i=1}^{N} \tilde{b}_{i,1}\} \geq \alpha, \tag{3.57}$$

where the notation

$$\begin{aligned} \tilde{a}_{i,1} &\equiv \min[a_{i,1} \cdot C_{i,1}, b_{i,1} \cdot C_{i,1}] \\ \tilde{b}_{i,1} &\equiv \max[a_{i,1} \cdot C_{i,1}, b_{i,1} \cdot C_{i,1}] \end{aligned} \tag{3.58}$$

has been adopted. Additionally, define

$$\begin{aligned} \tilde{a}_{i,j} &\equiv \min[a_{i,1} \cdot C_{i,j} \cdot D_{i,j}, b_{i,1} \cdot C_{i,j} \cdot D_{i,j}] \\ \tilde{b}_{i,j} &\equiv \max[a_{i,1} \cdot C_{i,j}, D_{i,j}, b_{i,1} \cdot C_{i,j} \cdot D_{i,j}]. \end{aligned} \tag{3.59}$$

Since the values of the $C_{i,1}$'s are unrestricted, substitute

$$C^*_{i,1} = C_{i,1} + \sum_{j=2}^{k_i} C_{i,j} \cdot D_{i,j} = \sum_{j=1}^{k_i} C_{i,j} \cdot D_{i,j}$$

where $D_{i,1} = 1$ for all i. The event E_1 in (3.57) may then be rewritten as

$$\sum_{i=1}^{N} \tilde{a}^*_{i,1} \leq \sum_{i=1}^{N} X_{i,1} \sum_{j=1}^{k_i} C_{i,j} \cdot D_{i,j} \leq \sum_{i=1}^{N} \tilde{b}^*_{i,1} \tag{3.60}$$

where, from the definition of (3.58), it is required that

$$\tilde{a}^*_{i,1} = \min[a_{i,1} \cdot \sum_{i=1}^{N} \sum_{j=1}^{k_i} C_{i,j} \cdot D_{i,j} \cdot b_{i,1} \cdot \sum_{i=1}^{N} \sum_{j=1}^{k_i} C_{i,j} \cdot D_{i,j}]$$

and

$$\tilde{b}^*_{i,1} = \max[a_{i,1} \cdot \sum_{i=1}^{N} \sum_{j=1}^{k_i} C_{i,j} \cdot D_{i,j}, b_{i,1} \cdot \sum_{i=1}^{N} \sum_{j=1}^{k_i} C_{i,j} \cdot D_{i,j}].$$

Observe that

$$\tilde{a}^*_{i,1} = \min[X_{i,1} \sum_{j=1}^{k_i} C_{i,j} \cdot D_{i,j}]$$

and

$$\tilde{b}^*_{i,1} = \max[X_{i,1} \sum_{j=1}^{k_i} C_{i,j} \cdot D_{i,j}].$$

Now recalling Lemma 3.9, it must be true that

$$\sum_{j=1}^{k_i} \min[a_{i,1} \cdot (C_{i,j} \cdot D_{i,j}), b_{i,1} \cdot (C_{i,j} \cdot D_{i,j})] \leq \tilde{a}^*_{i,1}$$

and that

$$\sum_{j=1}^{k_i} \max[a_{i,1} \cdot (C_{i,j} \cdot D_{i,j}), b_{i,1} \cdot (C_{i,j} \cdot D_{i,j})] \geq \tilde{b}^*_{i,1}.$$

Using the definition of (3.59) in the last result, one obtains

$$\sum_{j=1}^{k_i} \tilde{a}_{i,j} \leq \tilde{a}^*_{i,1}$$

and

$$\tilde{b}^*_{i,1} \leq \sum_{j=1}^{k_i} \tilde{b}_{i,j}.$$

Thus the event E_1 that was rewritten as (3.60) implies that

(E_2)

$$\sum_{i=1}^{N}\sum_{j=1}^{k_i}\tilde{a}_{i,j}\leq\sum_{i=1}^{N}X_{i,1}\sum_{j=1}^{k_i}C_{i,j}\cdot D_{i,j}\leq\sum_{i=1}^{N}\sum_{j=1}^{k_i}\tilde{b}_{i,j}. \tag{3.61}$$

Adding the quantity

$$\sum_{i=1}^{N}\sum_{j=1}^{k_i}L_{i,j}\cdot C_{i,j},L_{i,1}=0 \text{ for all } i$$

to (3.61) and factoring the middle expression yields

$$\sum_{i=1}^{N}\sum_{j=1}^{k_i}(\tilde{a}_{i,j}+L_{i,j}\cdot C_{i,j})\leq\sum_{i=1}^{N}\sum_{j=1}^{k_i}C_{i,j}(X_{i,1}\cdot D_{i,j}+L_{i,j})\leq \sum_{i=1}^{N}\sum_{j=1}^{k_i}(\tilde{b}_{i,j}+L_{i,j}\cdot C_{i,j}). \tag{3.62}$$

But from the definition in (3.50) and from $D_{i,1}=1$ and $L_{i,1}=0$, the event E_2 finally can be written as

$$\sum_{i=1}^{N}\sum_{j=1}^{k_i}(\tilde{a}_{i,j}+L_{i,j}\cdot C_{i,j})\leq\sum_{i=1}^{N}\sum_{j=1}^{k_i}C_{i,j}\cdot X_{i,j}\leq\sum_{i=1}^{N}\sum_{j=1}^{k_i}(\tilde{b}_{i,j}+L_{i,j}\cdot C_{i,j}) \tag{3.63}.$$

Consider the result of (3.63). The middle expression is the linear combination of (3.51). From (3.50), (3.52), and from (3.59)

$$\tilde{a}_{i,j}+L_{i,j}\cdot C_{i,j}\leq C_{i,j}\cdot X_{i,j}\leq\tilde{b}_{i,j}+L_{i,j}\cdot C_{i,j}$$

represents the TI for $C_{i,j}\cdot X_{i,j}$ that is implied by

$$a_{i,1}\leq X_{i,1}\leq b_{i,1}.$$

It therefore follows that (3.63) represents the CTI that results when the BA is used to combine the TI's of the $X_{i,j}$

of the linear combination of (3.51). Note that no restrictions were placed on the values of $C_{i,j}$, $D_{i,j}$, and $L_{i,j}$ other than that which the identity $X_{i,j} = D_{i,j} \cdot X_{i,1} + L_{i,j}$ for j = 1 obviously requires.

Finally, the event E_2 in its final form corresponding to (3.63) is implied by E_1. But since

$$p\{E_1\} \geq \alpha$$

it must also be true that

$$p\{E_2\} \geq \alpha.$$

This concludes the proof of Theorem 3.8.

Theorems 3.7 and 3.8 indicate that the BA will produce a CTI, for a linear combination of RV's, which is in many important cases, either a conservative estimate of the actual TI for the sum, or is equal to the actual TI. In particular, Theorem 3.7 indicates that when all the RV's are monotonically related to each other, the BA provides either a conservative statistical bound or the actual TI with respect to the RV linear sum. Theorem 3.8 indicates that when the statistical relationship between any two RV's is that of linearity dependence, the BA again provides a conservative combination of the individual TI's.

The last two theorems have important significance for applications of the previously developed techniques of this chapter. The values of a device response for different input values are related through the device function. Even

for a complex function having many degrees of freedom, there is some degree of statistical dependence between two different function values of the same function. If these two function values are treated as RV's and subsequently combined in a linear combination, then the dependence of one RV on the other can be important. In determining the TI for the linear combination, one may wish to apply the BA. It is not sufficient to simply establish the linear conformity of the pdf's for the two RV's, since this property is concerned with the situation in which the RV's are independent. But from the last two theorems, it is seen that even for the extreme case of linear dependence between the two RV's, the BA results in either a conservative estimate of the CTI or else in the TI that would result from direct use of the dependency relationship.

Consider the following simple example. Suppose the two RV's X_1 and X_2 are combined as

$$Y = X_1 + X_2$$

and suppose $X_2 = 2X_1$. Then if the α level TI for X_1 corresponds to

$$p\{-A \leq X_1 \leq A\} = \alpha \tag{3.64}$$

it must be true that

$$p\{-2A \leq X_2 \leq 2A\} = \alpha. \tag{3.65}$$

One may conclude by inspection that whether the BA is used to obtain a CTI on Y using (3.64) and (3.65), or $X_2 = 2X_1$ is substituted directly as

$$Y = X_1 + 2X_1 = 3X_1,$$

the TI computed for Y is (-3A, 3A). Suppose X_1 (and therefore X_2) was normally distributed. Then by naively assuming X_1 and X_2 independent, the TI for Y could be directly computed (Bowker & Lieberman [1964]) as

$$|Y| \leq \sqrt{A^2 + 4A^2} = A\sqrt{5} < 3A.$$

This result is an underestimate of the true TI for Y.

Several alternative algorithms that combine TI's directly will now be considered. The first is the sum of squares (SS) algorithm. This algorithm is based on the linear combination of independent, normally distributed RV's. If

$$Y = \sum_{i=1}^{N} C_i \cdot X_i$$

and the X_i's are normally distributed and independent and if the TI's for the X_i's correspond to

$$p\{a_i \leq X_i \leq b_i\} = \alpha$$

then the TI for Y, $[a_y, b_y]$, corresponds to (Bowker & Lieberman [1962])

$$p\{|Y - \tfrac{1}{2}\sum_{i=1}^{N} C_i(a_i+b_i)| \leq \tfrac{1}{2}\sqrt{\sum_{i=1}^{N} C_i^2(b_i-a_i)^2}\} = \alpha.$$

The SS algorithm is important since normal distributions are frequently encountered in practice. However, as was seen in the simple example just given, when two or more of the RV's in a linear combination are dependent, the SS

algorithm may give an underestimate of the actual TI for the RV sum. In view of this drawback, the SS algorithm must be used with caution, and cannot be generally applied.

There are a number of other disadvantages in using the SS algorithm. In discussing these disadvantages, it is useful to consider the general class of algorithms which form TI's corresponding to

$$|Y-\tfrac{1}{2}\sum_{i=1}^{N} C_i(a_i+b_i)| \leq \tfrac{1}{2}\left(\sum_{i=1}^{N} C_i^k (b_i-a_i)^k\right)^{1/k} \text{ for } k>0. \qquad (3.66)$$

It is seen that the SS algorithm corresponds to the k=2 case, while the BA corresponds to the k=1 case. Note that k is not restricted to an integer value. Except for the k=1 case, the algorithms of (3.66) compute TL's for Y which are nonlinear functions of the individual TL's, a_i and b_i. Because of this nonlinearity, these algorithms do not allow the straightforward construction of statistical bounds that is possible with the BA. For example, the ability to obtain continuous bounds on the device function by using a differential extrapolation interval (Theorem 3.6) would be impossible. Equation (3.20) would become

$$f(x_N) \leq f(x_o)+\tfrac{1}{2}\Delta x \cdot \sum_{i=o}^{N-1}[b(x_i)+a(x_i)]+ \frac{\Delta x}{2}\left[\sum_{i=o}^{N-1}\left(b(x_i)-a(x_i)\right)^k\right]^{1/k} \qquad (3.67)$$

Since the bound on $f(x_N)$ is no longer a linear function of the $a(x_i)$ and $b(x_i)$ values, it is not possible to make the

transition between the summation and integration operations as before. In fact, if one considers the case where $C_{min} \leq b(x_i)-a(x_i) \leq C_{max}$ for $x_o \leq x \leq x_r$, then using (3.21), the value of the second expression of the $f(x_N)$ upper bound is between

$$\tfrac{1}{2}(x_r-x_o)\cdot C_{min}\cdot N^{(1/k)-1} \leq \frac{\Delta x}{2}\cdot\left[\sum_{i=o}^{N-1}\left(b(x_i)-a(x_i)\right)^k\right]^{1/k}$$

$$\leq \frac{(x_r-x_o)}{2}\cdot C_{max}\cdot N^{(1/k)-1}.$$

But for $k<1$, $N\to\infty$ implies that the bound on $f(x_N)$ is infinite while for $k>1$, it is implied that

$$f(x_o)+\int_{x_o}^{x_r}\left[\frac{b(x)+a(x)}{2}\right]dx \leq f(x_N) \leq f(x_o)+\int_{x_o}^{x_r}\left[\frac{b(x)+a(x)}{2}\right]dx .$$

But this second result corresponds to a zero width TI which can only correspond to probability of zero. Therefore, it is not possible to obtain a meaningful continuous TI.

It appears then that the class of algorithms given in (3.66) are less suitable than the BA for the techniques developed in the last section. In Chapter 5, it will be seen that the BA is also more suited to the techniques developed there. In Chapter 4, the algorithms of (3.66) will be considered in terms of the confidence level that may be associated with their TI estimates. This treatment will assume statistical independence of the combined RV's. It will be shown that the k=2 case (SS algorithm) may be used as a minimum TI estimate in certain cases. Finally, it is

easily seen that for k>1, the TI's given by the algorithms may be underestimates of the actual TI when the combined RV's are dependent. If k<1, the TI obtained will be more conservative than the BA and will therefore result in overly pessimistic bounds.

There are other possible algorithms that may be used to directly combine TI's. For example, one may select the TI with the maximum width from the combined RV TI's, and then assign an overall TI for the linear sum as some function of this maximum value. Algorithms such as this one result in either overly pessimistic bounds or else bounds for which it is impossible to assign a confidence level.

In conclusion of this section, it has been shown that the direct combination of TI's is a more practical approach in obtaining statistical bounds on linear combinations of RV's. The BA was shown to possess certain advantageous properties that were found lacking in several alternative algorithms. When some of the RV's to be linearly combined are statistically dependent, the BA was proven to yield conservative TI estimates, that in certain limiting cases approached the actual TI of the linear combination.

In the next chapter, the property of self-linear conformity introduced in this chapter will be formally defined. A detailed investigation of this property will be carried out. An important class of pdf's will be shown to possess this property for any size linear sum. A heuristic

analysis will allow other pdf's to be analyzed to determine if they are self-linearly conformal.

In the next section of this chapter, a special problem is considered. This problem is concerned with the combining of statistical data with deterministic device measurements, as before. However, in this case, one is concerned with determining the variations of process parameter effects over the area of a monolithic circuit, rather than determining the functional variations with input changes. This application topic is included since it represents a novel implementation of a conventional technique to combine statistical and deterministic data.

3.4 Special Problem This section will consider a special applications problem which arises in the testing of monolithic circuits. This problem differs from the general function testing problem considered in earlier sections and to be considered in later chapters. Here one is primarily concerned with monitoring the variations of the fabrication process parameter effects over the physical area of the monolithic circuit. These effects of the process parameters are called "effective parameters" in order to distinguish them from the process parameters themselves. Because of the extremely small physical size of the circuit and since the circuit elements cannot be isolated for the purpose of making measurements, the number of points at which a circuit

can have measurements of the effective parameters is limited.

While the problem of this section differs in terms of what is to be determined, this problem is concerned with the same underlying goal of other sections. That is, a device is to be characterized by a limited number of deterministic measurements. As will be explained, there is a cost associated with the number of such measurements used to test a particular device. The approach will be to combine prior information (statistics) with the deterministic measurements made on a particular device in order to obtain a revised estimate of the distribution of the effective parameter variation over the circuit area. This section differs from other sections in that only existing mathematical techniques are required. The implementation of these techniques, however, involves a new approach to statistics gathering. This approach introduces the concept of a "test wafer", described below (Herman, et al [1973]).

In order to define the problem of monitoring process parameters in detail, it is necessary to acquaint the reader with certain aspects of monolithic circuit fabrication. These devices are fabricated on disks called wafers. The wafer contains many devices as shown in an array. The dimensions of a single circuit or "chip" is usually less than a tenth of an inch square, although large-scale-integration (LSI) circuits may be somewhat larger.

Connections to the chips are made by wires bonded to the chip at points of the chip called "lands". These lands must be several times the size of the circuit elements due to practical limitations of present bonding techniques. Because of the relative large size of these contact points, the number and placement of such areas is constrained. Circuit probing at points other than the lands is possible by manual microprobing equipment, but automated probing at points other than lands is not presently feasible. Higher density integration achieved by such technologies as MOS integration aggrevates the probing limitations. Even if probing were not such a severe problem, making most types of parametric measurements at a certain region of the circuit requires that an isolated circuit element located in this region be accessible for the measurement.

The significance of effective parameters for device testing requires some explanation. The many stages of monolithic circuit fabrication involve process variables or parameters. These parameters correspond to the controlled conditions of the process such as temperature, duration of the process, and chemical composition of the diffusion material. Each parameter has some effect on the properties of the monolithic circuit elements. Since all of these elements for each of the devices on a wafer are fabricated in the same processing sequence or "heat", the process parameters for these elements are the same. The differences

between one element and another result from masking techniques in which specific elements are selectively subject to a particular process step while other elements are masked out. While the effect of a process parameter is fairly uniform over circuit elements subject to the same conditions, some variation of this effect is inevitable. While the variation of the effective parameter is negligible over very small circuit areas, it can become quite significant over the area of a LSI device. With the present trend of increasing scale of integration, the physical area of future devices may be many times that of the largest devices currently available.

Recent efforts in improving integrated circuit reliability have led to more critical evaluation of the entire manufacturing process. Schroen, et al [1972] indicates that, "a significant fraction of device failures are due to device parameter changes caused by changes in dopants, impurities, or defects in critical areas superimposed on broad parameter distributions."

If the integrated circuit process technology is to be controlled, there must exist means for monitoring the effect of the process parameters. It is possible to correlate measurable circuit parameters and the parameters of the fabrication process (Schroen, et al [1972] and vanBeek [1972]). For example, diffusion characteristics control transistor properties such as gain and leakage currents. By measuring

the electrical characteristics of a single transistor, it is possible to infer what the process parameters were which produced it. Of course, such an inference corresponds to the wafer region at which the transistor resides.

Techniques for the process control of monolithic circuit quality and reliability at the wafer level have been studied and implemented by many research institutes as well as by circuit manufacturers (Baron and Myers, Miller [1969], Herr [1968 a, b], Anon (Westinghouse Corp.), and Schegel). The control techniques vary from a single test transistor and capacitor on each chip to several such "test patterns" on a wafer. The current approach is to use the parametric measurements made at a few test sites as an indication of the process parameter effects over the entire wafer. That is, if all of these measurements fall within acceptable ranges, the wafer is accepted. Otherwise it is rejected. The acceptable range of each parameter is intended to reflect the range within which the circuit elements will operate satisfactorily and within which the failure mechanisms will be minimized.

Since the test elements require levels for test connection, the number of test sites that can be placed on a chip must be minimized. Due to topology constraints, it is not practical to place the test elements other than at the circuit perimeter. As the size of the circuit becomes large, the use of test transistors to monitor effective parameters

becomes less meaningful. One cannot determine spatial parameter variations using existing practical techniques.

In order to overcome the usual difficulty of accounting for significant parameter variation over the area of a monolithic circuit, a statistical approach will be taken. Prior knowledge of spatial parameter variations will be combined with the information acquired by testing the few test sites that can be economically placed on the production wafer. The concept of a "test wafer" will be described shortly which will enable a practical gathering of the required prior statistics.

The statistical approach to be considered involves determining the statistical distribution of the inferred effective parameter for different points on a wafer. That is, at some point on the wafer, it is assumed that electrical measurements made on a test element enable one to infer what the process parameter value was during the wafer fabrication. This inferred value, referred to as the effective parameter value, is considered to be a RV, X. The statistical distribution for X is defined to correspond to any point on the wafer. It is assumed that such measurements are made "far enough" apart on the wafer such that each effective parameter value can be considered to be an independent sample of the RV, X. Furthermore, X is assumed to

be normally[1] distributed with known variance σ_X^2 and unknown mean μ_X. It is also assumed that μ_X, considered as a RV, has a priori density function which is normal with mean μ_M and variance σ_M^2. The physical interpretation for these statistics is that many wafers are produced under the same process conditions. Corresponding to each of these wafers, X is distributed on the wafer with the density parameters σ_X^2 and μ_X. While the variance σ_X^2 is assumed to be the same for each wafer, the mean is allowably different for each wafer. For many wafers, if the actual mean is somehow determined, one can estimate the mean μ_M and variance σ_X^2 of the wafer mean, μ_X.

At each of N test sites on a particular wafer, the values of X, x_i, i=1,...,N corresponding to these test sites are determined. Using the statistics σ_X^2, σ_M^2, and μ_M along with the N measurements, the distribution for X on the measured wafer is to be estimated using Bayesian analysis (Morgan [1968], Crellin [1972]).

As a first step, the prior distribution for μ_X can be revised to reflect the condition that the N observations (measurements) have been made. Using a result of Breipohl [1969] based on Bayes Theorem, the posterior distribution of μ_M is again normal with respective mean and variance

---

[1]Gaussian-normal density function.

$$\mu_{M,N} = \frac{\sigma_M^2 \sum_{i=1}^{N} x_i + \mu_M \sigma_X^2}{N\,\sigma_M^2 + \sigma_X^2} \tag{3.68}$$

$$\sigma_{M,N}^2 = \frac{\sigma_M^2\,\sigma_X^2}{N\sigma_M^2 + \sigma_X^2} \tag{3.69}$$

where the second subscript denotes that these statistics are conditional on the N observations. One may note that as $N \to \infty$

$$\lim_{N\to\infty} \mu_{M,N} = \lim_{N\to\infty} \sum_{i=1}^{N} x_i/N = \mu_M$$

which is the actual mean for X on the measured wafer. Similarly

$$\lim_{N\to\infty} \sigma_{M,N}^2 = 0$$

as one would expect.

Breipohl's result can be carried a step further. It has been stated that the distribution of X is normal with known variance σ_X^2 and unknown mean μ_X. Therefore, since knowing the value of μ_X determines the distribution of X, it is clear that one can consider this distribution as being conditional on μ_X. That is the normal pdf for X conditioned on $\mu_X = \theta$ is (Bowker and Lieberman [1964])

$$g_{X|\mu_X}(x|\theta) = \frac{1}{(2\pi\,\sigma_X^2)^{\frac{1}{2}}} \exp\,[-(x-\theta)^2/2\,\sigma_X^2] \tag{3.70}$$

where x is the usual pdf variable. That is

$$p\{X \leq a \mid \mu_X = \theta\} = \int_{-\infty}^{a} g_{X|\mu_X}(x|\theta)\, dx.$$

From (3.68) and (3.69), the posterior pdf of μ_X is written

$$g_{\mu_X}(\theta) = \frac{1}{(2\pi\ \sigma^2_{M,N})} \exp\ \{\frac{-(\theta-\mu_{M,N})^2}{2\ \sigma^2_{M,N}}\}. \qquad (3.71)$$

To obtain the absolute density function for X, one integrates the product of (3.70) and (3.71) over the range of θ. Then the absolute density function is

$$g_X(x) = \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_{M,N}\sigma_X} \exp\ \{-\tfrac{1}{2}[\frac{(x-\theta)^2}{\sigma^2_X} + \frac{(\theta-\mu_{M,N})^2}{\sigma^2_{M,N}}]\}d\theta. \qquad (3.72)$$

If the terms in the exponential are expanded, put over a common denominator, then collected as coefficients of θ^2, θ^1, and θ^0, and then finally normalized with respect to the θ^2 coefficient, the exponential expression if (3.72) can be written

$$\exp\{-\tfrac{1}{2}.\ \frac{[\ \theta^2 - 2(\mu_{M,N}\sigma^2_X + y\sigma^2_{M,N})\theta\ +\ (y^2\sigma^2_{M,N} + \sigma^2_X\mu_{M,N})\ ]}{(\sigma^2_X+\sigma^2_{M,N})\qquad\qquad (\sigma^2_X+\sigma^2_{M,N})}}{[\sigma^2_X\sigma^2_{M,N}/(\sigma^2_X+\sigma^2_{M,N})]}\}. \qquad (3.73)$$

The expression of (3.73) becomes upon completing the square (of the quadratic)

$$\exp\{-\tfrac{1}{2}.\ \frac{[\theta-(\frac{\mu_{M,N}\sigma^2_X+y\sigma^2_{M,N}}{\sigma^2_X+\sigma^2_{M,N}})]^2}{[\sigma^2_X\sigma^2_{M,N}/(\sigma^2_X+\sigma^2_{M,N})]}\}.\exp\{-\tfrac{1}{2}.\ \frac{[y-\mu_{M,N}]^2}{[\sigma^2_X+\sigma^2_{M,N}]}\}. \qquad (3.74)$$

The second exponential of (3.74) contains no θ terms and can be taken outside the integral of (3.72). One notices that the first exponential of (3.74) will become a Gaussian-normal pdf if it is divided by

$$[2\pi\sigma_X^2\ \sigma_{M,N}^2/(\sigma_X^2+\sigma_{M,N}^2)]^{\frac{1}{2}} \tag{3.75}$$

But the integral of a Gaussian-normal pdf over the $-\infty<\theta<\infty$ interval is equal to unity. Therefore (3.72) is

$$g_X(y) = \frac{[2\pi\sigma_X^2\sigma_{M,N}^2/(\sigma_X^2+\sigma_{M,N}^2)]^{\frac{1}{2}}}{2\pi\sigma_{M,N}\sigma_X} \cdot \exp[-\tfrac{1}{2}(x-\mu_{M,N})^2/(\sigma_X^2+\sigma_{M,N}^2)]$$

which may be reduced to

$$g_X(y) = \frac{1}{[2\pi(\sigma_X^2+\sigma_{M,N}^2)]^{\frac{1}{2}}} \cdot \exp[-\tfrac{1}{2}(x-\mu_{M,N})^2/(\sigma_X^2+\sigma_{M,N}^2)] \tag{3.76}$$

This result shows that the estimated distribution of X is normal with mean $\mu_{X,N} = \mu_{M,N}$ given by (3.68) and with variance

$$\sigma_{X,N}^2 \equiv \sigma_X^2 + \sigma_{M,N}^2. \tag{3.77}$$

where $\sigma_{M,N}^2$ is given in (3.69). Note that as the number of observations increases, $\sigma_{X,N}^2$ approaches σ_X^2 while the $\mu_{X,N}$ approaches the sample mean of the N observations.

In order to use the result of (3.76), the statistics σ_X^2, σ_M^2, and μ_M are required. It has been stated that σ_X^2 is assumed to be the same for all wafers. This claim can be defended by considering the manufacturing process of the wafers (or some equivalent). Consider that this variance represents the spread of the values of X for points over a

particular wafer. Thus σ_X^2 corresponds to the variation of the effective process parameter over a wafer. It seems reasonable that this variation should correspond to the uniformity of environmental conditions within the chamber in which the wafers are processed. The uniformity may be affected by wafer size, the sequence, and nature of the process steps. Since the process parameters are closely maintained, the uniformity of parameter effects over a wafer should be virtually constant for all wafers subject to the same intended conditions. The variation from wafer to wafer and consequently from heat to heat are more likely to be characterized by the variation of the mean, μ_X, from wafer to wafer. This last variation, while small, is significant in its effect on the monolithic circuits produced. By assuming σ_X^2 constant, one neglects the second order effects of the effective parameter variations.

In view of the limited number of measurements that are possible with the "production wafer" containing the monolithic circuits, it is somewhat questionable how accurate the estimates of σ_X^2, μ_M, and σ_M^2 can be obtained using only such measurement information. It is proposed here that a special wafer called a "test wafer" be used. This wafer is to entirely consist of test patterns or elements that are presently placed in limited quantities on the production wafers. The idea is to place such elements as closely together as possible in order to determine X at as many points

as possible. Using these test wafers, one could achieve a vast improvement in the estimate of σ_X^2 simply from the fact that the number of samples could be increased by several orders of magnitude. Using the average of the estimate of σ_X^2, one could determine a practical estimate of σ_X^2 to be used in (3.76). Also, one would certainly be able to determine if the assumption that σ_X^2 is constant was reasonable. Using the same measurements for the test wafers, the mean of X for each of these wafers could be determined. By testing a reasonable number of the test wafers, one could obtain estimates for the two distribution parameters σ_M^2 and μ_M corresponding to the wafer mean. One could observe whether the assumption, that the wafer mean and the RV X are normally distributed, is a valid approximation. Finally, it is clear that the test wafers are to be subject to the same process parameters as the production wafers. It is also suggested that after the initial data gathering effect in which many test wafers are tested, one should continue to occasionally include test wafers in production runs. In this way, the prior distribution parameters σ_M^2 and μ_M could be continually revised to reflect any general shifts in the fabrication process.

The specific manner in which the pdf of (3.76) is utilized depends on the application. However, several general statements concerning the method of implementing this distribution can be made. One first realizes that the

acceptability ranges for the electrical circuit element parameters depend on the circuit configuration and the desired performance of the monolithic device. The relationship between these acceptability ranges and the circuit performance is studied under topics such as sensitivity analysis and circuit analysis and design (Herskowitz [1968]). As was pointed out earlier, there exists relationships between the circuit element parameters and the effective parameters. Therefore one can determine an acceptability range for the effective parameters for each circuit element.

A further tolerance constraint on effective parameters can come from a known correlation between these parameters and the dynamics of various failure mechanisms. Investigations such as Schroen [1972] have demonstrated that short term parameter drift of test transistors resulting from various step-stress conditions can be indicative of the rates of some "wear-out" mechanisms.

Assuming that corresponding to each regular circuit element, one can determine an acceptability range for each effective parameter[1], one must next consider how to compute a measure of probability that a particular device on a production wafer will operate satisfactorily. For this device, the effective parameters for each circuit element

---

[1]It is possible that the ranges for each effective parameter may be interdependent. However, it will be assumed here that the ranges are independently defined.

must be within the required bounds. Thus for each circuit element, one could integrate the pdf of (3.76), corresponding to each of the effective parameters over the acceptability range defined for the element and the parameter. Assuming that the effective parameters can be considered statistically independent, the product of each such integration represents the probability that the circuit element will operate acceptably. Finally, the probability that the entire device is acceptable can be taken as the product of the probabilities corresponding to the acceptability of each circuit element. In order to simplify the analysis just described, one may determine an overall acceptability range for each effective parameter as the intersection of each of the element defined acceptability ranges. Let the pdf (3.76) for the i^{th} effective parameter be integrated over the overall acceptability range and call the value resulting from this integration p_i. If there are k_i circuit elements that are dependent on the i^{th} effective parameter, the probability P_D that the device is acceptable is

$$P_D = \prod_{i=1}^{N_X} p_i^{k_i}. \tag{3.78}$$

where N_X is the number of effective parameters measured. A problem exists in the determination of the k_i's due to the nature of monolithic devices. Circuit elements for these devices are distributed to some degree so that it is unclear how many circuit elements the device has. Also, there is

some degree of statistical dependence between the effective parameters for adjacent elements. Thus in practice, one may have to resort to a heuristic selection of the constants k_i. These constants may be called "density constants" since they reflect the density of circuit elements in the device.

Having selected the values for the density constants, one is then in a position to make a decision of whether to accept or reject the wafer. Acceptance can be contingent on m of a total n device on the wafer being acceptable. One must select a confidence level α corresponding to this event. Then the probability of at least m of n devices acceptable must satisfy (Feller [1968])

$$\sum_{i=m}^{n} \binom{n}{i} P_D^i (1-P_D)^{n-i} \geq \alpha \tag{3.79}$$

where $\binom{n}{i} = n!/[(i!)(n-i)!]$, and where P_D is given in (3.78).

One may elect to avoid computing the values of p_i required for (3.78) each time a wafer is tested. One observes that the pdf given by (3.76) for the i^{th} effective parameter is a function of the sum of effective parameter values determined by the test site measurements [see (3.68) and (3.69)] once N, σ_X^2, σ_M^2, and μ_M are given. Corresponding to an acceptability range for the effective parameter, one may determine p_i versus the above sum as a table of values. Then one could use the table to determine p_i for each wafer in terms of the sum of effective parameter values determined by the measurements made on the wafer test elements.

Since the statistics σ_M^2, σ_X^2, and, μ_M are determined from test wafer measurements, it is possible to apply this information to testing wafers containing different types of devices, providing the same process sequence applies to the different devices. It is therefore possible to test the feasibility of a device design prior to fabricating it. One must simply determine acceptability ranges for the effective parameters using a sensitivity analysis performed for the device. Afterwards, the density constants k_i are determined for the circuit configuration. A Monte Carlo simulation using the distribution of μ_M and subsequently the distribution function (3.70), may be used to generate the test site effective parameter values. The acceptability of the simulated wafer could then be determined as though measurements had actually been made on the wafer rather than the test wafer. By repeating the simulation many times, one could determine the expected number of wafers that would be accepted. This last result could be used to decide if the device design was economically feasible.

Figure 3.6 shows the general steps that may be followed in performing the above feasibility study and in gathering the prior statistics required for testing. Many variations of this procedure are possible. As mentioned earlier, the specific procedure used will depend on the application.

Figure 3.7 is an example of a testing procedure for production wafers. It has been assumed here that an overall

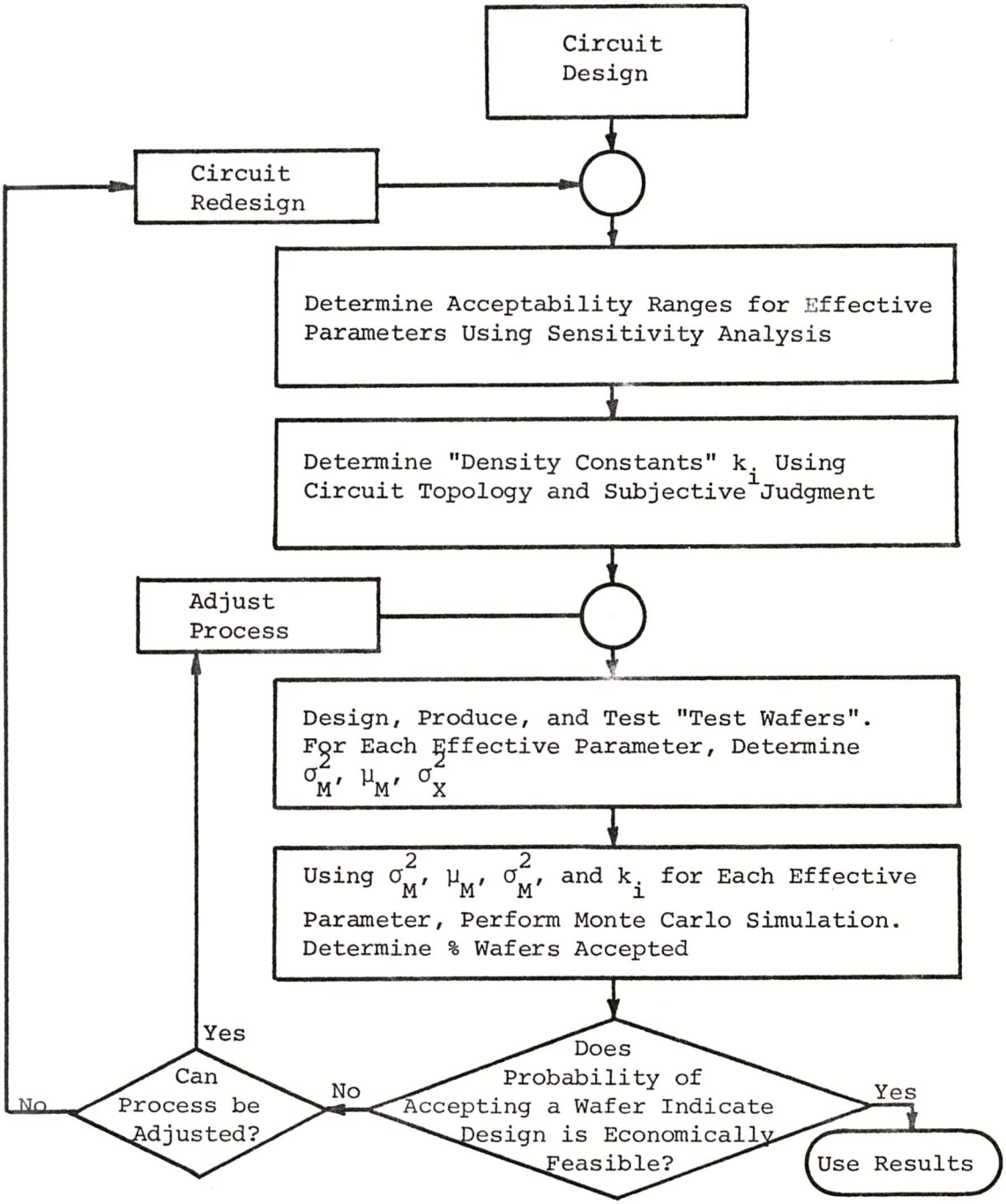

Figure 3.6 General steps for performing a feasibility analysis when test wafers are used to obtain parameter distribution statistics for later production wafer testing.

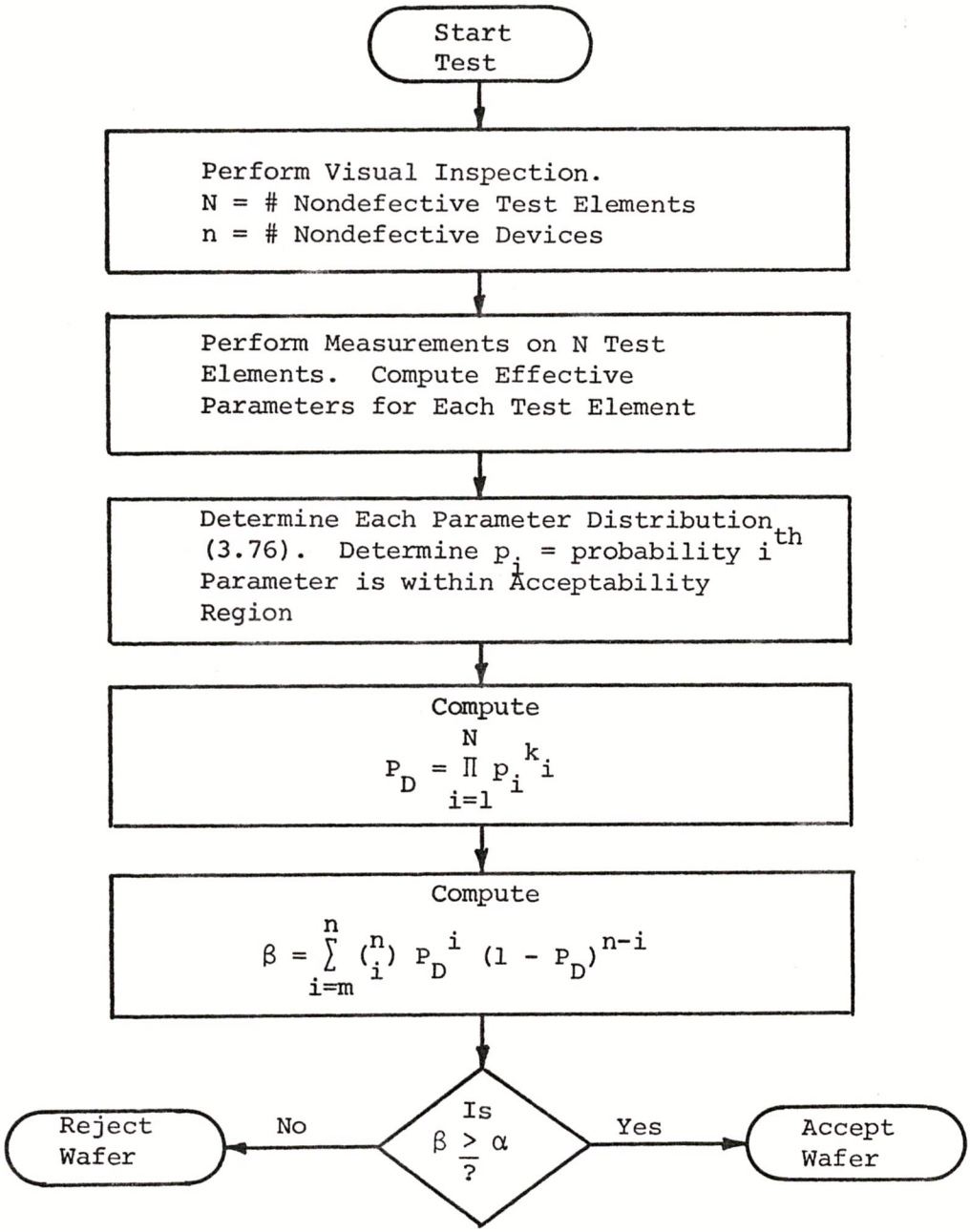

Figure 3.7 Example of production wafer testing using test wafer statistical approach. Note that visual inspection precedes process parameter phase of wafer testing procedure.

acceptability range for each of N_p effective parameters can be determined. It has also been assumed that k_i, σ_M^2, σ_X^2 and μ_M have been determined for each of the effective parameters. In this example, the number of test sites on the production wafer, N, may differ from wafer to wafer. A reason for this is that visual inspection of the wafer is assumed to precede the determination of the test site parameters. Visual inspection is a common practice in which "cosmetic" defects are visually detected. Such defects can be scratches, isolated spots of foreign material, and mask misalignment defects. The practice is to immediately reject all devices having such defects. Since some test sites can contain cosmetic defects, one would similarly avoid measurements involving either test sites having defects or test sites residing in the circuit area of defective devices. Notice that acceptance of a wafer is contingent on the probability that at least m devices are good out of n devices containing no cosmetic defects. Thus, the effect of a cosmetic defect in Figure 3.7 is to reduce the probability of acceptance.

As a final consideration of the test wafer statistical approach, one could consider the case where either the distribution of the effective parameter or the distribution of its mean is non-normal. In this case, one may not be able to obtain closed-form results as was possible with normal distributions. The Bayesian-revised distribution

parameters must then be obtained by numerical integration. The distribution of the wafer effective parameter must similarly be obtained numerically. Such computations must be repeated for each production wafer tested. When such computations are feasible in a practical production testing environment, the test wafer approach is still practical.

In conclusion, this section has introduced a new approach to monitoring process parameter effects for production monolithic device fabrication. By using the new concept of a test wafer, it is possible to gather the required statistics for a Bayesian estimate of the effective parameter distribution on a production wafer. A consequence of the test wafer approach is a rational testing procedure for the acceptance of production wafers.

3.5 Conclusion of Preliminary Considerations This chapter has examined several approaches to combining prior statistics, gathered for a population of similar devices, with deterministic measurements made on a particular device. The general goal has been to characterize the particular device through this combination of information such that more information is available than would be by utilizing only the above deterministic data. This general goal can also be interpreted as a reduction in the number of deterministic measurements required to characterize a device. It has been assumed that the number of such measurements represents the cost of testing the device.

Section 3.2 was concerned with device characterization by means of the input-output device function. The theorems presented in this section contain the essence of the developed preliminary techniques. A bounding algorithm was introduced to enable the simple combination of TI's for the RV's which were required to be combined by the preliminary techniques.

In section 3.3, this algorithm and a number of alternatives were considered with respect to practical application to device function testing. Several theorems were discussed which indicated that the bounding algorithm used in this section is preferably applied in practice, especially when the combined RV's are statistically dependent.

In Chapter 5, the preliminary techniques developed in section 3.2 will be applied to the situation in which a polynomial or transcendental interpolation formula is ordinarily used to characterize a device function.

Section 3.4 presented an application for the use of prior statistics to spatial interpolation of parameter effect variations in monolithic devices. This special problem was included because of its importance and because it represents a conventional approach to the combination of statistical and deterministic data.

In the next chapter, an investigation of the linear conformality property for pdf's will be presented. This property has been seen to be salient to the developments of

this chapter and will also be important to the techniques of Chapter 5.

4. BOUNDS ON LINEAR COMBINATIONS OF INDEPENDENT RANDOM VARIABLES

4.1 Introduction A central concern within this thesis is that of determining bounds on linear combinations of random variables (RV's). It is assumed in most of the thesis developments, that the combined RV's are statistically independent. It was shown in the preceding section, that the assumption of statistical independence often leads to an optimistic bound estimate via analytic methods, when in fact the RV's of a linear combination are dependent. It was also shown that the bounding algorithm (BA) introduced in the last chapter maintains a conservative bound estimate even when the RV's are linearly dependent, provided that this bound is conservative for the linear independent case. The condition of a conservative bound was introduced in the last chapter as the condition of linear conformity.

In this chapter, an investigation is made of the conditions with respect to the probability density functions (pdf's) of the combined RV's which guarantees a conservative bound computed by the BA. The bound formed is an estimate of the tolerance interval (TI) corresponding to a linear combination of independent RV's. The computation of the BA is a direct combination of the TI's for each of the combined RV's. By direct, it is meant that the convolution process associated with conventional methods of linear RV combinations is avoided.

It is first shown that the BA does not always result in a conservative estimate of the TI for a linear combination of RV's. After formal definition of terms, several important properties of the Fourier transform of pdf's are developed. These properties are required for the subsequent chapter developments.

An important family of functions is discussed and several theorems regarding the bounds formed by the BA for members of this family are proven. This family, called the "exponential transform family" contains such standard distributions as the delta, the Cauchy, and the Gaussian distribution. In addition to analyzing the bound formed by the BA, for this family, the chapter considers the bound formed by the sum-of-squares (SS) algorithm, previously defined. It is shown that the SS bound may be used as a lower limit or minimum width TI for a linear combination of independent RV's. Except for a combination of all Gaussian distributed RV's, the SS bound is optimistic. That is, this bound is narrower than the TI that would be formed by conventional means.

A heuristic classification procedure is presented which enables a graphical analysis of an arbitrary pdf transform under certain conditions. These conditions include a requirement of symmetry of the pdf. The heuristic is then developed from a near-analytic argument.

The chapter is concluded with a discussion of the assumptions made during the chapter developments. An example application of the heuristic procedure is given, and the heuristic results are compared to an analytic solution. This example corresponds to the example given early in the chapter.

4.2 Definitions and Assumptions In this section, the bounding algorithm, defined in the last chapter is restated. The condition of linear conformity, (lc) which is an important consideration in applying the BA to a linear combination of independent RV's, is defined. An example is given, which demonstrates that lc cannot be assumed for all situations. In the preceding chapter, the BA was shown to be a preferred method of bounding linear combinations of RV's in several important classes of situations, provided the condition of lc could be established. In studying the condition of lc, several assumptions will be made to simplify this analysis. These assumptions, stated in this section, will be discussed at the end of the chapter in terms of their applicability to practical situations.

Definition 4.1 (Bounding Algorithm) Consider the linear combination of RV's

$$Y = \sum_{i=1}^{N} C_i \cdot X_i \tag{4.1}$$

where the C_i's are constants and the X_i's are RV's with known pdf's. Corresponding to the confidence level α, let the TI's for each of the X_i be determined according to the convention of equation (3.4), such that

$$p\{a_i \leq X_i \leq b_i\} = \alpha, \text{ for all } i.$$

Then the "composite tolerance interval" (CTI) for Y is defined by

$$\left(\sum_{i=1}^{N} \min[C_i \cdot a_i, C_i b_i] \leq Y \leq \sum_{i=0}^{N} \max[C_i a_i, C_i b_i]\right) \tag{4.2}$$

where

$$\min[u,v] \equiv \begin{pmatrix} u & , u \leq v \\ v & , \text{otherwise} \end{pmatrix}$$

and

$$\max[u,v] \equiv \begin{pmatrix} u & , u \geq v \\ v & , \text{otherwise} \end{pmatrix}.$$

This process of combination is referred to as the bounding algorithm (BA).

As mentioned at the beginning of this chapter, one is concerned with the situation wherein the CTI as just defined is a conservative bound. The following definitions describe two conditions of conservative bounds.

Definition 4.2 (Linear Conformity) Consider the linear combination of statistically independent RV's, $X_1 \ldots X_N$,

$$Y = \sum_{i=1}^{N} C_i X_i. \tag{4.3}$$

If a CTI for Y is obtained using the BA and the confidence level for each of the individual TI's is α, then the collection of pdf's for the X_i's are said to be "linearly conformal" (lc), if and only if, the confidence level for the CTI is at least α for all values of the C_i's, and with no restriction on α.

When the N pdf's are lc, and in addition have the same general distribution form, with perhaps different distribution parameters, this common distribution is said to be "self-linearly conformal".

Definition 4.3 (Conditional Linearly Conformity) If for the previous definition, there exists a minimum value of α for which the CTI confidence level is at least that of the individual TI's for all values of the C_i's, then the pdf is said to be conditionally linearly conformal.

The BA may appear to the casual observer as a "worst-case" bound formation. However, there are situations in which the BA results in a TI which has a confidence level less than that corresponding to the individual TI's. Stated differently, the CTI formed by the BA can be an underestimate of the bound on Y corresponding to the confidence

level α associated with the X_i in (4.3). In order to verify the possibility of an underestimate of the CTI, consider the following discrete probability distribution given in Table 4.1.

k	p(X=k)
0	.2
$\pm$1	.1
$\pm$2	.05
$\pm$3	.01
$\pm$4	0
$\pm$5	0
$\pm$6	.02
$\pm$7	.05
$\pm$8	.1
$\pm$9	.05
$\pm$10	.02

Table 4.1

DISTRIBUTION FOR WHICH ALGORITHM IS NOT VALID

Suppose both the RV's X_1 and X_2 have the above distribution. Let

$$Y = X_1 + X_2 \tag{4.4}$$

Using the summation

$$p\{Y=j\} = \sum_{i=-10}^{10} p\{X_i=i\} \cdot p\{X_2=j-i\}$$

and noting that

$$p\{|X_2|>10\} = 0$$

the distribution for Y is obtained, and is given in Table 4.2.

k	p (Y=k)	k	p (Y=k)
0	.0968		
$\pm$1	.0750	$\pm$11	.0110
$\pm$2	.0450	$\pm$12	.0034
$\pm$3	.0184	$\pm$13	.0024
$\pm$4	.0083	$\pm$14	.0065
$\pm$5	.0120	$\pm$15	.0120
$\pm$6	.0291	$\pm$16	.0158
$\pm$7	.0494	$\pm$17	.0120
$\pm$8	.0640	$\pm$18	.0065
$\pm$9	.0494	$\pm$19	.0020
$\pm$10	.0290	$\pm$20	.0004

Table 4.2

CONVOLVED DISTRIBUTION IN WHICH BA NOT VALID

Examination of the two tables reveals that if an α is selected corresponding to

$$|X|\leq k, k = 0, 1, 2, 3,$$

then

$$p\{-2k\leq Y\leq 2k\}<p\{-k\leq (X_1 \wedge X_2)\leq k\}$$

and the algorithm fails to give suitable results. This example will again be considered later in the chapter.

In the remaining sections of this chapter, it will be assumed that all pdf's under consideration are symmetric, and that the RV's of linear combinations are statistically

independent. It will be further assumed that all RV's of a linear combination possess the same pdf with allowably different distribution parameters, and that these pdf's are explicitly known as a priori information. These assumptions will be discussed at the end of the chapter in order to relate their applicability to practical situations.

4.3 Fourier Analysis In this section, several important properties of probability density functions (pdf's) are discussed. These properties pertain to the Fourier transforms of the pdf's. The reason these properties are discussed is that they will be useful in the investigation of linear conformity, to be discussed in the next two sections. Fourier analysis will be the basis for the developments of this chapter.

In establishing TL's for the RV formed from a linear combination of independent RV's, it is apparent that the properties of the convolution integral remain the central concern. This is true since convolution is the process by which the pdf of the linear combination can be obtained. In studying the convolution process it is natural to turn to the Fourier transform. The reason for this is that convolution transforms to a simple multiplication via the Fourier Integral, and certain properties are more easily discerned.

In discussing the convolution and its multiplicative equivalent in the frequency domain, one will find several

properties of pdf's to be useful. It is assumed in what follows with no loss of generality that the mean of the pdf is zero. For the RV, X with non-zero mean, the RV, $X_s = X-\mu$ can be equivalently considered, where μ is the mean of the pdf corresponding to X.

Consider now the moments of the pdf of X, g(u)

$$m_k = \int_{-\infty}^{\infty} u^k \, g(u)\,du. \tag{4.5}$$

From the moment theorem (Papoulis [1962])

$$(-j)^k \, m_k = \left.\frac{d^k G(\omega)}{d\omega^k}\right|_{\omega=0} \tag{4.6}$$

where $j = (-1)^{\frac{1}{2}}$.

For k=0, (4.5) and (4.6) combine as

$$G(0) = \int_{-\infty}^{\infty} g(u)\,du = 1. \tag{4.7}$$

By definition, the right-hand side of (4.7) equals unity.

For k=1, (4.5) and (4.6) become

$$\left.\frac{dG(\omega)}{d\omega}\right|_{\omega=0} = -j \int_{-\infty}^{\infty} u \cdot g(u)\,du \equiv -\mu j. \tag{4.8}$$

Since it has been assumed that g(u) is symmetric about the origin, the mean, μ , is zero. Also since g(u) is real and an even function of u, G(ω) must be real and even (Papoulis [1962]). For an even function, equation (4.8) implies

$$\lim_{\omega\to 0+} \left[\frac{dG(\omega)}{d\omega}\right] + \lim_{\omega\to 0-} \left[\frac{dG(\omega)}{d\omega}\right] = 0.$$

Thus the special case in which the slope of $G(\omega)$ is undefined at the origin, but is finite and of opposite sign to the right and to the left of $\omega = 0$, is also considered.

The last moment of interest is m_2

$$\int_{-\infty}^{\infty} u^2 \, g(u)\,du = - \left. \frac{d^2 G(\omega)}{d\omega^2} \right|_{\omega=0}. \tag{4.9}$$

For a zero mean, the second moment is also the variance so that

$$\sigma^2 = - \left. \frac{d^2 G(\omega)}{d\omega^2} \right|_{\omega=0}. \tag{4.10}$$

Since the variance must be positive, except for degenerate cases (although it may be infinite), and since

$$|G(\omega)| = \left| \int_{-\infty}^{\infty} g(u) e^{-j\omega u} \, d_u \right| \leq \int_{-\infty}^{\infty} g(u) d_u = 1 \tag{4.11}$$

the shape of $G(\omega)$ near the origin is concave downward except for $d[G(\omega)]/d\omega$ discontinuous at the origin. This situation occurs for distributions such as the Cauchy distribution, that is[1]

$$\frac{\beta}{\pi(\beta^2+u^2)} \leftrightarrow e^{-\beta|\omega|} \tag{4.12}$$

$$\beta \geq 0.$$

The result of such distributions is that the second derivative has a negative impulse at the origin. In this case, the function $G(\omega)$ is concave upward. The reason for the

---

[1]The symbol "$\leftrightarrow$" denotes that the expressions separated by $\leftrightarrow$ are Fourier transform pairs.

concern with the shape of the transform of a pdf will become apparent in the next two sections when the linear conformity property is studied.

The last several developments may be summarized as the following theorem.

Theorem 4.4 Given $g(u)$, the pdf of the RV, X. If $g(u)$ is symmetrical about the origin of u, then the Fourier transform $G(\omega)$ of $g(u)$ has the following properties:

1. $G(\omega)\big|_{\omega=0} = 1$

2. $\lim_{\omega \to 0+} \left[\frac{dG(\omega)}{d\omega}\right] + \lim_{\omega \to 0-} \left[\frac{dG(\omega)}{d\omega}\right] = 0.$

3. $\sigma^2 = -\left.\frac{d^2G(\omega)}{d^2\omega}\right|_{\omega=0} \geq 0$

4. $G(\omega) \leq 1 \quad -\infty < \omega < \infty$

There is another development that will be useful in subsequent sections. This development concerns the operation of computing a probability corresponding to a TI taken symmetrically about the mean of a symmetric distribution.

Suppose $g(u)$ is some symmetric pdf and u is defined as the displacement from the mean. That is, the mean is

$$\mu = 0,$$

and the tolerance limits are defined as $(-a, a)$, where $a \geq 0$.

There exists a probability α, corresponding to a, such that

$$p\{|X| \leq a\} = \int_{-a}^{a} g(u)du = \alpha. \tag{4.13}$$

Now if g(u) possesses a Fourier transform, $G(\omega)$, using the inverse transform integral, equation (4.13) may be re-written as (Papoulis [1962])

$$p\{|X| \leq a\} = \int_{-a}^{a} [\frac{1}{2\pi} \int_{-\infty}^{\infty} G(\omega) e^{j\omega u} d\omega] du = \alpha. \tag{4.14}$$

Interchanging the order of integration and integrating with respect to u results in

$$p\{|X| \leq a\} = \frac{1}{\pi} \cdot \int_{-\infty}^{\infty} G(\omega) \frac{\sin \omega a}{\omega} d\omega. \tag{4.15}$$

It was possible to interchange the order of integration since the integrand was square integrable (Papoulis [1962]). That is

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} |G(\omega) e^{j\omega u}|^2 d_\omega \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} |G(\omega)|^2 d_\omega < \infty. \tag{4.16}$$

The right-hand inequality can be justified by

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} |G(\omega)|^2 d\omega = [g(u) * g(u)]\Big|_{u=0} = h(0) < \infty$$

when g(u) contains no impulses at the origin.

It is useful to depart from the main development of this chapter in order to make some observations about the integral of (4.15). Although these observations are not directly used in subsequent developments, they provide some

insight into the probability integral in the "u" domain and the integral of (4.15).

Consider equation (4.15), and in particular the effect of changing a. As this variable increases, the integral must also increase since it represents (4.14). The function $\sin a\omega/\omega$ has the value 'a' at the origin. The envelope of this function is $1/\omega$, except near the origin. As 'a' increases, the period of the sinusoid increases, and $\sin a\omega/\omega$ approaches the envelope curve more quickly. The area under $\sin a\omega/\omega$ for any particular half period

$$[\frac{k\pi}{a}, \frac{(k+1)\pi}{a}]$$

is independent of 'a'. That is, by the substitution $v = a\omega$, one obtains

$$\int_{k\pi/a}^{(k+1)\pi/a} \frac{\sin a\omega}{\omega}\, d\omega = \int_{k\pi}^{(k+1)\pi} \frac{\sin v}{v}\, dv. \qquad (4.17)$$

Also, the k increases, the area under the curve for two successive half periods approach each other and are of opposite sign. Thus, when $\sin a\omega/\omega$ is a weighting function as in equation (4.15), and the function multiplied changes slowly over a period of $\sin a\omega/\omega$, as k increases, the contribution to the total integral for each successive period approaches zero. The interpretation to be made is that as 'a' increases in the integral of (4.15), the behavior near the origin of $G(\omega)$, tends to dominate the entire integral.

The dominant region approaches the origin as 'a' increases. In fact, as $a \to \infty$, (4.15) becomes

$$\lim_{a\to\infty} \frac{1}{\pi} \int_{-\infty}^{\infty} G(\omega) \frac{\sin \omega a}{\omega} \, d\omega$$

$$= \frac{1}{\pi} \int_{-\infty}^{\infty} G(\omega) \lim_{a\to\infty} \left[\frac{\sin \omega a}{\omega}\right] d\omega. \tag{4.18}$$

The limit of the preceding corresponds to the definition of an impulse. That is

$$\lim_{\to\infty} \left(\frac{\sin \omega a}{\omega}\right) = \pi \, \delta \, (\omega). \tag{4.19}$$

Then (4.18) is

$$\frac{1}{\pi} \int_{-\infty}^{\infty} G(\omega) \pi \, \delta(\omega) d\omega = G(0) = 1.$$

This result is expected in view of (4.14) and the definition of a pdf.

The properties developed emphasize the behavior of the pdf transform close to the ω origin. As was just observed, it is this region which dominates the probability corresponding to a TI taken symmetrically about the pdf mean.

Returning to the main development of this section, one may finally consider the convolution and its transform. Consider equation (4.1) for which the pdf of X_i is $g_i(u)$

and the X_i's are statistically independent.[1] Repeating (3.33), the pdf $g_y(u)$ of the RV, Y is written as the N-fold convolution

$$g_y(u) = \left[\frac{g_1(u/C_1) * g_2(u/C_2)(*\ldots*g_N(u/C_N)}{\prod_{i=1}^{N} |C_i|} \right] \tag{4.20}$$

where "*" denotes the convolution operator. That is

$$g_1(u) * g_2(u) = \int_{-\infty}^{\infty} g_1(z) \cdot g_2(u-z)dz. \tag{4.21}$$

The transform of $g_y(u)$ can similarly be written, repeating (3.35), as

$$g_y(u) \leftrightarrow \prod_{i=1}^{N} G_i\ (C_i\omega). \tag{4.22}$$

The approach used in the following sections will be to substitute the right-hand side of (4.22) in (4.15) and replace the constant, 'a', of (4.15) with some combinational form of the tolerance limits defined for the individual RV's X_i. These TI's defined symmetrically about the individual distribution means will correspond to a common confidence level α. The main concern will be to determine whether the composite TI determined by the combinational form corresponds to a confidence level of at least α. In this case, the combinational form will represent a conservative

---

[1]The symbol u is used here as the independent variable representing the value of the RV in the "real" domain.

estimate of the actual symmetrically defined TI corresponding to a confidence level of α.

4.4 Exponential Transform Family In this section a family of functions is considered, referred to as the "exponential transform family."

Definition 4.5 (Exponential Transform Family) A function is a member of the exponential transform family if, and only if, its Fourier transform can be written as

$$F_{\theta,k}(\omega) = \exp(-\theta|\omega^k| \text{ where } \theta \geq 0 \quad k \geq 0. \tag{4.23}$$

A function belonging to this family is distinguished by the two parameters, θ and k.

This family is considered for several reasons. First, the members for k equal to zero, one, and two correspond to the delta function, the Cauchy distribution, and the Normal distribution, respectively. Second, each subset of the family of members with a common value of k, exhibits closure under the operation of convolution in the u^1 domain, and therefore multiplication in the ω domain. This property allows a simplified analysis to be made which is not possible for pdf's in general. Before proceeding, it must

[1]By "u domain" it is meant real domain of the RV's.

be pointed out that the members of this family do not, in general, correspond to pdf's. For example, it is easy to show that for integer k>2, the function of (4.23) cannot correspond to a non-negative function in the u domain. Indeed, the second derivative of the function with respect to ω is zero at the origin. Its inverse transform has a zero second moment from (4.6). This can only be true if the function is negative for some regions of u. Therefore, for k an integer greater than two, the function of (4.23) is not the transform of a pdf. However, since the Fourier transform of many pdf's resemble the members of the exponential transform family in shape, one may speculate on the behavior of such a pdf in terms of the properties of the family member to which the pdf resembles.

The following theorem considers the property of linear conformity with respect to family members which are pdf's.

Theorem 4.6 Given the linear combinations of independent RV's

$$Y = C_1X_1 + C_2X_2 \tag{4.24}$$

where C_1, C_2 are positive constants, let X_1 and X_2 have similar pdf's which are Fourier transformable as

$$\text{pdf } \{X_1\} = g_1(u) \leftrightarrow \exp(-\theta_1|\omega^k|) \quad \theta_1 \geq 0 \tag{4.25}$$

$$\text{pdf } \{X_2\} = g_2(u) \leftrightarrow \exp(-\theta_2|\omega^k|) \quad \theta_2 \geq 0 \tag{4.26}$$

and k be some positive constant identical for the two distributions. The general pdf corresponding to the particular value of k and with general parameter $\theta \geq 0$ is lc with itself, if and only if, $k \geq 1$. In this case the BA applied to (4.24) will result in a conservative CTI or RV, Y. That is, for any α, such that

$$0 \leq \alpha \leq 1$$

if

$$p\{|X_1| \leq a_1\} = \int_{-a_1}^{a_1} g_1(u)\, du = \alpha \tag{4.27}$$

and

$$p\{|X_2| \leq a_2\} = \int_{-a_2}^{a_2} g_2(u)\, du = \alpha \tag{4.28}$$

then

$$p\{|Y| \leq a_1 C_1 + a_2 C_2\} \geq \alpha, \tag{4.29}$$

if, and only if, $k \geq 1$. (Note that $a_1, a_2 \geq 0$.)

Proof Consider first, equations (4.27) and (4.28). From these equations and from equation (4.15), it is required that

$$\frac{1}{\pi}\int_{-\infty}^{\infty} \exp(-\theta_1|\omega^k|)\, \frac{\sin\omega a_1}{\omega}\, d\omega = \frac{1}{\pi}\int_{-\infty}^{\infty} \exp(-\theta_2|\omega^k|) \cdot \frac{\sin\omega a_2}{\omega}\, d\omega. \tag{4.30}$$

Making the substitution

$$a_1 Z = a_2 \omega$$

in the second integral, and noting that

$$\frac{dZ}{d\omega} = \frac{a_2}{a_1} = \frac{Z}{\omega}$$

one sees that (4.30) becomes, after multiplying through by π,

$$\int_{-\infty}^{\infty} \exp(-\theta_1|\omega^k|)\frac{\sin\omega a_1}{\omega}\, d\omega = \int_{-\infty}^{\infty} \exp(-\theta_2|[a_1 Z/a_2]^k|)\frac{\sin Z a_1}{Z}\, dZ \tag{4.31}$$

$$= \int_{-\infty}^{\infty} \exp(-\theta_2[a_1/a_2]^k|Z^k|)\frac{\sin Z a_1}{Z}\, dZ.$$

Since the two latter integrals are required to be equal, it can be easily shown by taking indefinite integrals that the constant coefficients in the exponent must be equal. That is, since a_1 and a_2 are positive

$$\theta_1 = \theta_2\ (a_1/a_2)^k \tag{4.32}$$

or

$$a_2 = (\theta_2/\theta_1)^{1/k}\ a_1. \tag{4.33}$$

The above result establishes the relation between the parameter θ and its effect on the interval constant 'a' for a fixed confidence level, α. Now consider the transform of the RV, Y. From (4.22), one can write the pdf of Y, $g_y u)$ in terms of its transform. That is[1]

$$g_y(u) \leftrightarrow G_1(C_1\omega)\cdot G_2(C_2\omega) = \exp\ (-\theta_1|[C_1\omega]^k|-\theta_2|[C_2\omega]^k|)$$

$$= \exp\ (-[\theta_1 C_1{}^k+\theta_2 C_2{}^k]|\omega^k|). \tag{4.34}$$

Using (4.32), (4.34) becomes

---

[1]For negative C_i's, absolute values should be taken in exponentials.

$$g_y(u) \leftrightarrow \exp(-\theta_1[C_1{}^k + C_2{}^k(a_2/a_1)^k]\cdot|\omega^k|)$$

$$= \exp(-\theta_1\gamma^k|\omega^k|) \tag{4.35}$$

where $\gamma^k = C_1{}^k + (C_2a_2/a_1)^k$.

The probability expression of equation (4.29) may be obtained by using equation (4.15). That is

$$p\{|Y| \leq a_1C_1 + a_2C_2\} = \frac{1}{\pi}\int_{-\infty}^{\infty}\exp(-\theta_1\gamma^k|\omega^k|)\cdot$$

$$sin[(C_1a_1 + C_2a_2)\omega]d\omega. \tag{4.36}$$

Let $\gamma\omega = v$.

Note that $\frac{d\omega}{dv} = \frac{\omega}{v}$.

Equation (4.36) is rewritten as

$$p\{|Y| \leq a_1C_1 + a_2C_2\} = \frac{1}{\pi}\int_{\infty}^{\infty}\exp(-\theta_1|v^k|\cdot$$

$$\frac{sin[(C_1a_1 + C_2a_2)v/\gamma]}{\gamma}\,dv. \tag{4.37}$$

From (4.15) and the properties of a pdf, it is clear that the integral of (4.37) increases monotonically with the constant

$$(C_1a_1 + C_2a_2)/\gamma.$$

Consider that the right-hand side of (4.37) has the form of the integrals in (4.30). Since the proof is concerned with establishing the conditions in which the probability of (4.37) is at least α, one may see from (4.15) and from (4.30), that it is equivalent to determine the conditions such that

$$(C_1a_1 + C_2a_2)/\gamma \geq a_1. \tag{4.38}$$

Using (4.32), (4.33), and (4.35),

$$(C_1a_1+C_2a_2)/\gamma=[C_1a_1+C_2(\theta_2/\theta_1)^{1/k}a_1]/[C_1^k+C_2^k(\theta_1/\theta_2)]^{1/k}$$

$$=a_1[C_1+C_2(\theta_2/\theta_1)^{1/k}]/$$

$$[C_1^k+C_2^k(\theta_2/\theta_1)]^{1/k}. \qquad (4.39)$$

Thus, condition (4.38) can be written as, for $C_1>0$

$$a_1[1+(C_2/C_1)(\theta_2/\theta_1)^{1/k}]/[1+(C_2/C_1)^k(\theta_2/\theta_1)]^{1/k}\geq a_1. \qquad (4.40)$$

The constants in (4.40) are all non-negative, so the inequality can be written, after raising both sides to the k^{th} power, as

$$(1+[C_2/C_1]^k(\theta_2/\theta_1)]^{1/k})^k\geq 1+(C_2/C_1)^k(\theta_2/\theta_1). \qquad (4.41)$$

Let $\zeta^k = (C_2/C_1)^k\ (\theta_2/\theta_1)$.

Thus (4.41) is

$$(1+\zeta)^k \geq 1+\zeta^k. \qquad (4.42)$$

Since ζ is a non-negative constant, it is obvious that from a Taylor series expansion the inequality of (4.42) requires

$$k \geq 1.$$

QED

It should be clear that for the trivial cases $C_1 = 0$ or $C_2 = 0$, the TL's of Y are just a constant magnification of the TL's corresponding to X_1 and X_2. Also, for $a_1 = 0$ or ∞, the inequality again becomes an equality.

Thus it has been shown that the necessary and sufficient conditions for the distribution of the exponential family to be linearly conformal is that the parameter k be greater than or equal to unity.

The above theorem can be extended to a linear combination of "N", RV's. That is

Theorem 4.7 Given the linear combination of N independent RV's, X_i, with zero mean,

$$Y = \sum_{i=1}^{N} C_i X_i \tag{4.43}$$

where C_i are positive constants. If RV's X_i have pdf's which are Fourier transformable as

$$\text{pdf } \{X_i\} = g_i(u) \leftrightarrow \exp(-\theta_i |\omega^k|), \ \theta_i \geq 0 \tag{4.44}$$

and k is a positive constant common to all N distributions, then the general distribution function of (4.44) is self-linearly conformal for any positive integer N if and only if $k \geq 1$. That is, for any α, such that

$$0 \leq \alpha \leq 1$$

if

$$p\{|X_i| \leq a_i\} = \int_{-a_i}^{a_i} g_i(u)\, du = \alpha, \tag{4.45}$$

then

$$p\{|Y| \leq \sum_{i=1}^{N} a_i C_i\} \geq \alpha$$

if and only if $k \geq 1$.

Proof Consider the inductive proof. From the previous theorem, the case N = 2 has been proven. Assume that the case N holds. Consider the N + 1 case. Rewrite (4.43) as

$$Y = \sum_{i=1}^{N} C_i X_i + C_{N+1} X_{N+1} = \hat{X} + C_{N+1} X_{N+1}. \tag{4.46}$$

The pdf of $\hat{X}$ has the same form as that of the individual X_i's. The transform of its pdf can be written as

$$\exp\ (-\hat{\theta}|\omega^k|)$$

and the constant $\hat{a}$ can be defined as

$$\hat{a} \equiv \sum_{i=1}^{N} C_i \cdot a_i.$$

By assumption that the N case holds

$$p\ \{|\hat{X}| \leq \hat{a}\} \geq \alpha, \text{ if } k \geq 1$$

$$< \alpha, \text{ if } k < 1.$$

It is clear that there exists another constant, $\tilde{a}$, such that

$$p\ \{|\hat{X}| \leq \tilde{a}\} = \alpha.$$

Obviously, $\tilde{a} \leq \hat{a}$, if $k \geq 1$ and

$$\tilde{a} > \hat{a}, \text{ if } k < 1.$$

Equation (4.46) can be seen to correspond to the N = 2 case. From the previous theorem, it follows that

$$p\ \{|Y| \leq \tilde{a} + a_{N+1} \cdot C_{N+1}\} \geq \alpha, \ k \geq 1$$

$$< \alpha, \ k < 1.$$

But

$$\tilde{a} + a_{N+1} C_{N+1} \leq \hat{a} + a_{N+1} \cdot C_{N+1} = \sum_{i=1}^{N+1} (a_i C_i), \ k \geq 1$$

or

$$\tilde{a} + a_{N+1}\, C_{N+1} > \hat{a} + a_{N+1} \cdot C_{N+1} = \sum_{i=1}^{N+1} (a_i C_i),\ k < 1$$

and from (4.15) it is obvious that one must also have

$$p\{|Y| \underline{\leq} \hat{a} + a_{N+1} \cdot C_{N+1}\} \underline{\geq} p\{|Y| \underline{\leq} \tilde{a} + a_{N+1} \cdot C_{N+1}\} \underline{\geq} \alpha,\ k \underline{\geq} 1$$

$$p\{|Y| \underline{\leq} \hat{a} + a_{N+1} \cdot C_{N+1}\} < p\{|Y| \underline{\leq} \tilde{a} + a_{N+1} \cdot C_{N+1}\} < \alpha,\ k > 1.$$

Therefore, the N+1 sum of (4.46) represents a lc collection of RV's, if and only if $k \underline{\geq} 1$. This proves the inductive step.

QED

It was assumed that the constants $C_i \underline{\geq} 0$. This was done only as a means of simplifying the notation. Since the pdf's are symmetric, and their transform are also real and symmetric, it should be clear that the same convolved distributions are obtained for both positive and negative C_i. Also, the TL's formed by the bounding algorithm can be written as

$$\sum_{i=1}^{N} a_i \cdot |C_i|$$

if the C_i can be negative also.

When dealing with a linear combination of RV's whose distributions correspond to a particular k-member of the exponential transform family, the above theorems can be used to determine if the pdf's of the linear sum are

linearly conformal. When these RV distributions closely resemble but are not equal to a family member, one may heuristically use the procedure associated with this family member.

It is instructive to consider the shapes of the family member transforms. The curves for various values of k are shown in Figure 4.3. θ is taken as unity except in one curve. For $k < 1$, the curves are steep near the origin and flater away from the origin in comparison to the other family members.

A heuristic classification procedure will be presented in the next section which will compare an arbitrary curve with the k=1 exponential transform family. From this comparison, it will be possible to determine if a distribution which corresponds to the arbitrary curve is linearly conformal with itself in many cases.

Before developing the heuristic procedure of the next section, the TI formed by the sum-of-squares (SS) algorithm will be briefly considered. Using the exponential transform family, an analysis will be made to determine the confidence level that may be associated with the combinational bound on a linear sum of RV's when the SS algorithm is used.

The result of the analysis to be made is stated as the following theorem:

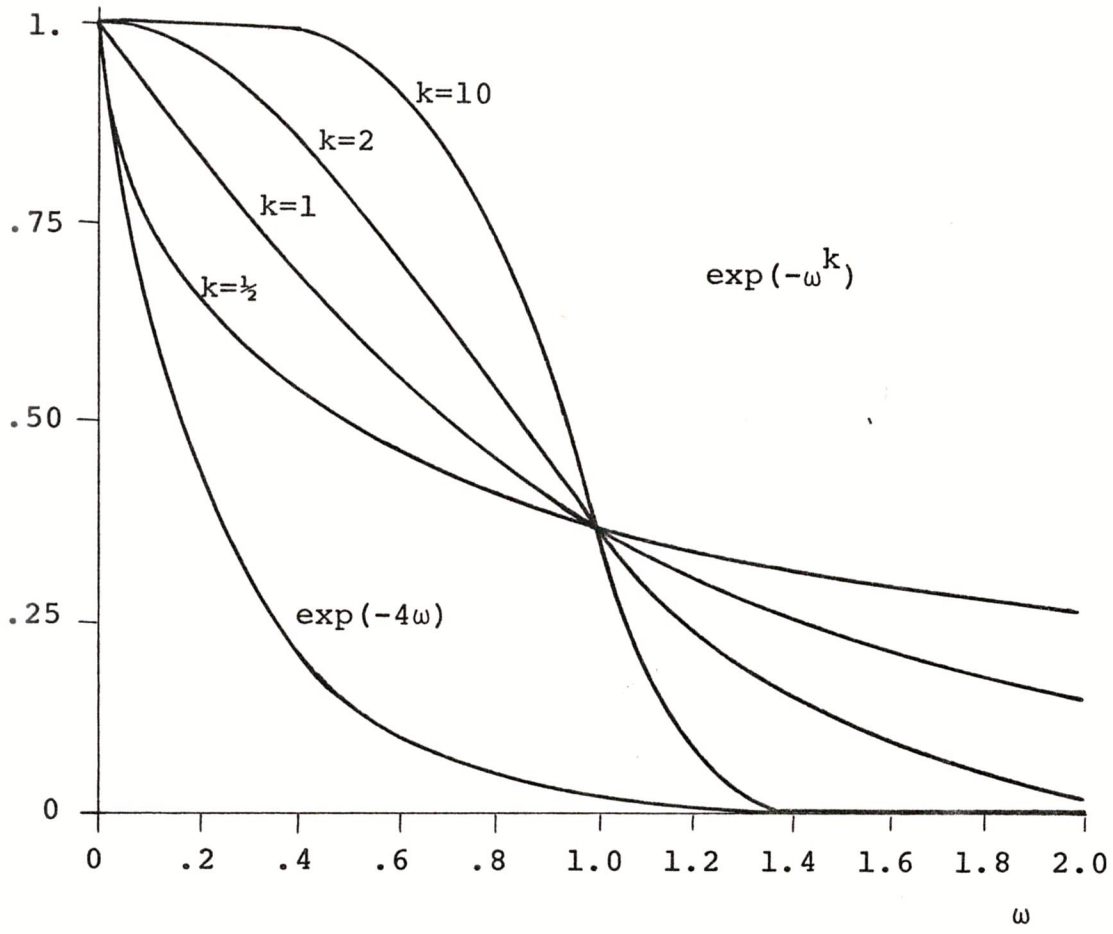

Figure 4.3 Exponential Transform Family. The general form of the family transformed in the ω domain is $\exp(-\theta|\omega^k|)$. Except for the $\exp(-4\omega)$ curve, $\theta=1$, for the curves shown. The negative ω portion of the curves, not shown, is a reflection of plotted portion, reflected on the $\omega=0$ line.

Theorem 4.8 Given the linear combination of N independent RV's, X_i,

$$Y = \sum_{i=1}^{N} C_i X_i \tag{4.47}$$

where C_i are positive constants. Let RV's X_i have pdf's which are Fourier transformable as

$$\text{pdf } \{X_i\} = g_i(u) \leftrightarrow \exp\ (-\theta_i |\omega^k|),\ \theta_i \geq 0 \tag{4.48}$$

where k is a positive constant common to all N distributions. It is tacitly assumed that the TI for each RV is formed symmetrically about a RV zero mean such that each TI corresponds to a confidence level of probability α. That is, for $g_i(u)$ the pdf of RV, X_i, one has

$$p\ \{|X_i| \leq a_i\} = \int_{-a_i}^{a_i} g_i(u)\ du = \alpha. \tag{4.49}$$

From the above, the TI formed using the SS algorithm will correspond to a confidence level of at least α, if and only if $k \geq 2$.

Proof The development follows the analysis of the previous two theorems. Consider the proof of Theorem 4.6. The development through equation (4.35) still applies. The SS algorithm forms the TI

$$|Y| \leq (a_1^2\ C_1^2 + a_2^2 \cdot C_2^2)^{1/2}.$$

Substitute this expression in place of $(a_1C_1+a_2C_2)$ in equations (4.36) through (4.38). Then the condition to be determined is that of finding out under what conditions

$$(C_1^2\, a_1^2 + C_2^2\, a_2^2)^{1/2}/\gamma \geq a_1 \tag{4.50}$$

holds. Using (4.32), (4.33), and (4.35), the left-hand side of (4.50) can be written as

$$(C_1^2\, a_1^2 + C_2^2\, a_2^2)^{1/2}/\gamma = \frac{[C_1^2\, a_1^2 + C_2^2\, (\theta_2/\theta_1)^{2/k}.a_1^{\,2}]^{1/2}}{[C_1^k + C_2^k\, (\theta_2/\theta_1)]^{1/k}}$$

$$= \frac{a_1[1 + (C_2/C_1)^2(\theta_2/\theta_1)^{2/k}]^{1/2}}{[1 + (C_2/C_1)^k(\theta_2/\theta_1)]^{1/k}}.$$

Let $\zeta^k = (C_2/C_1)^k\ (\theta_2/\theta_1)$ and substitute into the above equation. Then the inequality of (4.50) becomes

$$\frac{a_1[1 +\zeta^2]^{1/2}}{[1 +\zeta^k]^{1/k}} \geq a_1.$$

Cancelling the positive a_1, one must determine k such that

$$f(\zeta,\ k) \equiv \frac{[1 + \zeta^2]^{1/2}}{[1 + \zeta^k]^{1/k}} \geq 1 \tag{4.51}$$

is true. Consider the function $f(\zeta,\ k)$ as ζ is varied from zero to infinity (note that ζ is non-negative since θ_1, θ_2, C_1, C_2 are positive). It is instructive to examine the locations of local minimas and maximas. To determine these locations, one takes the partial derivative

$$\frac{df(\zeta, k)}{d\zeta} = \frac{\zeta[1 + \zeta^k]^{1/k - 1} [1 - \zeta^{k-2}]}{[1 + \zeta^2]^{1/2} [1 + \zeta^k]^{2/k}}. \tag{4.52}$$

For $k \neq 2$, the derivative of (4.52) is zero only at $\zeta=0$, 1, and at infinity. Also note that $f(\zeta, k)$ is unity at $\zeta=0$ and infinity. For k=2, the partial derivative is identically zero, as observed from (4.51). These last two results allow a simple analysis of the $f(\zeta, k)$ function in $(0 \leq \zeta \leq \infty)$. It is simple to show that

$$\begin{aligned} f(\zeta, k) &\leq 1 \text{ when } f(1, k) < 1 \\ &= 1 \text{ when } f(1, k) = 1 \\ &\geq 1 \text{ when } f(1, k) > 1. \end{aligned} \tag{4.53}$$

Consider f(1, k). From (4.51), one obtains

$$f(1, k) = \frac{2^{1/2}}{2^{1/k}}$$

from which it is clear that

$$\begin{aligned} f(1, k) &< 1 \quad , \quad k < 2 \\ &= 1 \quad , \quad k = 2 \\ &> 1 \quad , \quad k > 2. \end{aligned} \tag{4.54}$$

Therefore, the expression of (4.51) is satisfied for $k \geq 2$.

To complete the proof, the N case is assumed and the N+1 case then investigated. This analysis parallels the proof of Theorem 4.7.

Suppose the N case holds for the theorem. For the N+1 linear combination, (4.47) can be written as

$$Y = \sum_{i=1}^{N} C_i X_i + C_{N+1} \cdot X_{N+1} = \hat{X} + C_{N+1} \cdot X_{N+1}. \qquad (4.55)$$

The pdf of $\hat{X}$ has the same form as that of the individual X_i's. The transform of its pdf can be written as

$$\exp\ (-\hat{\theta} |\omega^k|)$$

and the constant $\hat{a}$ can be defined as

$$\hat{a}^2 \equiv \sum_{i=1}^{N} C_i^2 \cdot a_i^2.$$

By assumption that the N case holds

$$p\ \{|\hat{X}| \leq \hat{a}\} \geq \alpha, \text{ if } k \geq 2$$

$$< \alpha, \text{ if } k < 2.$$

It is clear that there exists another constant, $\tilde{a}$, such that

$$p\ \{|\hat{X}| \leq \tilde{a}\} = \alpha.$$

Obviously

$$\tilde{a} \leq \hat{a}, \text{ if } k \geq 2$$

and

$$\tilde{a} > \hat{a}, \text{ if } k < 2.$$

Equation (4.55) is seen to correspond to the N = 2 case. Therefore, it follows that

$$p\ \{|Y| \leq \sqrt{\hat{a}^2 + a_{N+1}^2 \cdot C_{N+1}^2}\} \geq \alpha, \ k \geq 2$$

$$< \alpha, \ k < 2.$$

But $\tilde{a}^2 + a_{N+1}^2 C_{N+1}^2 \leq \hat{a}^2 + a_{N+1}^2 \cdot C_{N+1}^2 = \sum_{i=1}^{N+1} a_i \cdot C_i$, $k \geq 2$.

$\tilde{a}^2 + a_{N+1}^2 C_{N+1}^2 > \hat{a}^2 + a_{N+1}^2 \cdot C_{N+1}^2 = \sum_{i=1}^{N+1} a_i \cdot C_i$, $k < 2$.

Therefore the N+1 case is proven. This concludes the inductive step.

QED

Theorem 4.8 suggests an important relationship between the SS algorithm and the BA. From the discussion at the beginning of this section, it is seen that the SS algorithm yields a conservative bound on a linear combination of independent RV's each having a k-exponential transform pdf, only when k=2. For $k > 2$, the exponential transform function does not correspond to a pdf, and for $k < 2$, the TI formed corresponds to a confidence level of probability less than that of the individual RV TI's. The k=2 function corresponds to the Gaussian-normal pdf, for which the SS algorithm provides the TI for the linear sum of RV's. Therefore, at least in the special situation in which exponential transform pdf's are convolved, the SS algorithm provides a minimum "width" TI for the linear sum.

It is also possible to argue from the information theoretic approach that the SS algorithm provides a minimum width TI. One may regard the statement that, for RV, Y

$$p\{|Y| \leq f(\alpha)\} = \alpha, \quad f(\alpha) \geq 0 \tag{4.56}$$

for a known function f (α), as a state having some measure of information. When

$$\int_{-f(\alpha)}^{f(\alpha)} g(u)\, du = \alpha, \tag{4.57}$$

the interval [-f(α), f(α)] is the actual TI for Y corresponding to confidence level α taken symmetrically about the zero mean of pdf g(u). When the functional value f(α) is less than the TL magnitude for Y at probability α, it is clear that (4.56) corresponds to a greater amount of information than that of f(α) equal to the TL magnitude. This fact follows from the range of RV Y being more constrained by a smaller width TI. Certainly

$$\int_{-a_1}^{a_1} g(u)\, du < \int_{-a_2}^{a_2} g(u)\, du \quad \text{for } a_1 < a_2. \tag{4.58}$$

Since whatever bounding algorithm used to bound a linear combination of RV should not add more information than is available, it is clear that the approximate TI formed by the algorithm should be as wide or wider than the actual TI of which it approximates.

Consider the following normalization of scale for a RV, Y having a pdf $g_Y(u)$ with finite variance

$$g_{Y_N}(u) \equiv \sigma_Y\, g_Y\, (\sigma_Y u). \tag{4.59}$$

The variance of the scaled RV, Y_N is by definition

$$\sigma_{Y_N}^2 \equiv \int_{-\infty}^{\infty} u^2 g_{Y_N}(u)\,du = \int_{-\infty}^{\infty} u^2 \sigma_Y g_Y(\sigma_Y \cdot u)\,du$$

$$= \int_{-\infty}^{\infty} \left(\frac{v}{\sigma_Y}\right)^2 \sigma_Y g_Y(v)\,dv/\sigma_Y$$

$$= \int_{-\infty}^{\infty} \left(\frac{1}{\sigma_Y}\right)^2 v^2 g(v)\,dv = \frac{\sigma_Y^2}{\sigma_Y^2} = 1.$$

From the definition of a pdf, it is required that

$$\int_{-\infty}^{\infty} g_{Y_N}(u)\,du = 1$$

$$\int_{-\infty}^{\infty} g_{Y_N}(u)\,du = \int_{-\infty}^{\infty} \sigma_Y g_Y(\sigma_Y u)\,du = \int_{-\infty}^{\infty} \sigma_Y g_Y(v)\,dv/\sigma_Y = 1.$$

It should be clear that the same information of state exists for the scaled RV Y_N as for Y provided that the statement of (4.56) and (4.57) are adjusted by replacing $f(\alpha)$ by $f(\alpha)/\sigma_Y$, and $g_Y(u)$ by $\sigma_Y \cdot g(\sigma_Y u)$.

Shannon [1948] has proven that if $g_Y(u)$ is a one-dimensional distribution, then the Gaussian form of $g_Y(u)$ gives a maximum entropy, subject to the condition that the standard deviation of Y be fixed at σ. Shannon's definition of entropy for a continuous function results in the entropy being a function of the coordinate system scale. Since the normalization of (4.59) provides for a standardization of scale relative to the variance of the unscaled pdf, it is possible to compare different pdf's in terms of

entropy or its negative, information of state (in the sense of (Evans [1969]).

Since a non-Gaussian normalized pdf has less entropy (or more information) than a normalized pdf, the application of an arbitrary monotonic function $f(\alpha)$ with implied probability carries more information for a non-Gaussian normalized RV than a Gaussian-normal RV. Suppose $f(\alpha)$ is the SS algorithm TL magnitude as a function of the confidence level α, and scaled by multiplicative factor $1/\sigma_Y$. When g_{Y_N} is a Gaussian distribution, the state of information corresponding to (4.56) is the negative of the entropy for g_Y. But in this case, the SS algorithm exactly forms the TI for RV Y. If g_{Y_N} is not Gaussian-normal, then (4.56) corresponds to a greater amount of information for $f(\alpha)$ still corresponding to the scaled TL's using the SS algorithm. From (4.58), it follows that the actual TI for a non-Gaussian pdf must be wider than the SS TI. That is, the additional information conveyed by the statement of (4.56) given the pdf of RV Y must be zero if $f(\alpha)$ is the actual TL magnitude as a function of confidence level α.

The preceding discussion indicates that the SS algorithm may be used to determine a lower bound on the TI width for a linear sum of independent RV's, provided that the variance of each RV is non-zero and finite. The BA results in an upper bound for many cases. From the previous

chapter it was shown that the BA may provide the actual TI when the RV's are no longer independent.

As a final topic of this section, the Central Limit Theorem (Mood and Graybill [1963]) is considered. The Central Limit Theorem provides additional insight into the convolution of independent RV's. This theorem states that under certain conditions, the convolutions of N, pdf's tends to the Gaussian-normal distributions as $N \to \infty$, regardless of the individual distributions. The conditions which ensure the validity of the theorem are that the sum of the variance of the pdf's tend to infinity as $N \to \infty$, and that the second moment of each pdf is finite (Mood and Graybill [1963]). While the first condition is usually easy to satisfy, the second may not be. For example, the Cauchy distribution, equation (4.12) does not have finite even moments m_k for k greater than zero.

The Central Limit Theorem bears some importance to the application of the BA. Consider applications for which a TI must be established for a linear sum of many independent RV's of the same distribution type. If it can be established by direct convolution that the algorithm is valid for a small number of convolutions, then it is a good heuristic rule to assume that this validity holds for an integer multiple of this number. This heuristic assumes that the distribution of the individual RV's satisfies the

conditions of the Central Limit Theorem. Its justification rests on the rapidity that many distributions approach the normal distribution under convolution (Mood and Graybill [1963] and Bowker and Lieberman [1964]) and the fact that the normal distribution is lc for any number of convolutions.

The analysis in this section has proven that the BA determines a conservative estimate of the TI for a linear combination of independent RV's, when the pdf's of these RV's are of several important forms. The family of functions corresponding to these forms is called the exponential transform family, and includes the Gaussian and the Cauchy distributions. In practice, one often encounters distributions which do not correspond to any standard functions, although they may approximate these standard forms. Since practical knowledge of a statistical distribution usually corresponds to a collection of discrete data, it is not unlikely that the most convenient representation of the distribution function is graphical. Using numerical techniques, one may obtain an approximation to the Fourier transform of the unknown distribution.

In the next section a heuristic graphical procedure is described which allows one to classify an arbitrary pdf transform curve as being self-linearly conformal under certain situations. These pdf's are restricted to symmetrical, real transforms. The heuristic procedure corresponds

to a graphical comparison of the arbitrary pdf transform to a standard function form. This standard form is the transform of the Cauchy pdf.

4.5 Heuristic Classification of Arbitrary Distributions

For the exponential transform family, it was determined that the family member corresponding to k=1 was the critical function which partitioned the family into two classes. For $k \geq 1$, the members were SLC, while for $k<1$, they were not SLC. It is natural to look for a classification method which allows one to compare a non-family member to the k=1 member in such a way as to ascertain for which ranges of α the linear conformity property applies.

A classification procedure is possible for transforms of arbitrary pdf's which meet certain conditions. This procedure is heuristic due to certain approximations made in its development. These approximations tend to be pessimistic, as seen shortly.

The following notation is used. The standard function $F(\omega)=\exp(-\theta|\omega|)$ is compared to an arbitrary symmetric function $G(\omega)$. The heuristic procedure considers the linear combination

$$Y = X_1 + X_2 \tag{4.60}$$

in which X_1 and X_2 are RV's having the same pdf g(u) which transforms to $G(\omega)$. If the TI for X_1 and X_2 are defined as in (4.13), then the confidence level corresponding to the

TI formed by the BA is, from (4.15) and (4.22)

$$p\{|Y|\leq 2a\} = \frac{1}{\pi}\int_{-\infty}^{\infty} G^2(\omega)\frac{\sin(\omega\cdot 2a)d\omega}{\omega}. \qquad (4.61)$$

The heuristic classification procedure considers the region in the positive ω plane for $\omega \leq \pi/a$ for each value of the tolerance limit a. At each such value, the heuristic is used to determine if the probability of (4.61) is greater than or less than α, (α is the confidence level for each individual TI of X_1 and X_2.) Recall that if this probability is greater than α for any choice of tolerance limit a, and the linear combination of (4.60) can take on non-unity weights, then $G(\omega)$ is SLC. If there is a minimum value of a for which the probability of (4.61) is greater than α, then $G(\omega)$ is conditionally SLC.

It is shown shortly why the unity coefficient combination of X_1 and X_2 may be taken as the critical case for the determination of the SLC property of $G(\omega)$. Since the heuristic must be applied at all values of "a" to classify $G(\omega)$, this procedure undoubtedly appears cumbersome to the reader. However, the graphical method can be applied very easily if $G(\omega)$ is plotted on semi-logarithmic graph paper, as will be demonstrated by an example.

Heuristic Classification Procedure

1. Sketch $G(\omega)$ on semi-log paper (ω, linear scale), for positive ω.

2. The standard function $F(\omega)$ is a straight line, whose slope depends on θ. Rotate this straight line to observe its intersection with $G(\omega)$ beginning at $G(0)$ to $G(\omega_{1/4})$, where $\omega_{1/4}$ is the minimum ω where $G(\omega) = .25$.

3. At each intersection point ω_I, determine whether the probability of (4.61) is greater than the confidence level corresponding to the TI $[-\pi/2\omega_I, \pi/2\omega_I]$ for X_1 and X_2. This determination is made by observing which of three following possible relations exist between $G(\omega)$ and the rotated $F(\omega)$:

   a. Conforms to Figure 4.4a. In Figure 4.4a, $G(\omega)$ is less than $F(\omega)$ for $0<\omega<\omega_I$ and then greater than $F(\omega)$ for $\omega_I<\omega<2\omega_I$. Note that there are no other intersections in $0<\omega<2\omega_I$. According to the heuristic, the relation

   $$p\{|Y| \leq 2a\} < p\{|X_1| \leq a\} = p\{|X_2| \leq a\} \qquad (4.62)$$

   is valid for $a = \pi/2\omega_I$.

   b. Conforms to Figure 4.4b. In Figure 4.4b, the reverse situation is shown. $G(\omega)>F(\omega)$ for $0<\omega<\omega_I$, and $G(\omega)<F(\omega)$ for $\omega_I <\omega<2\omega_I$. According to the heuristic, the relation

   $$p\{|Y| \leq 2a\} > p\{|X_1| \leq a\} = p\{|X_2| \leq a\} \qquad (4.63)$$

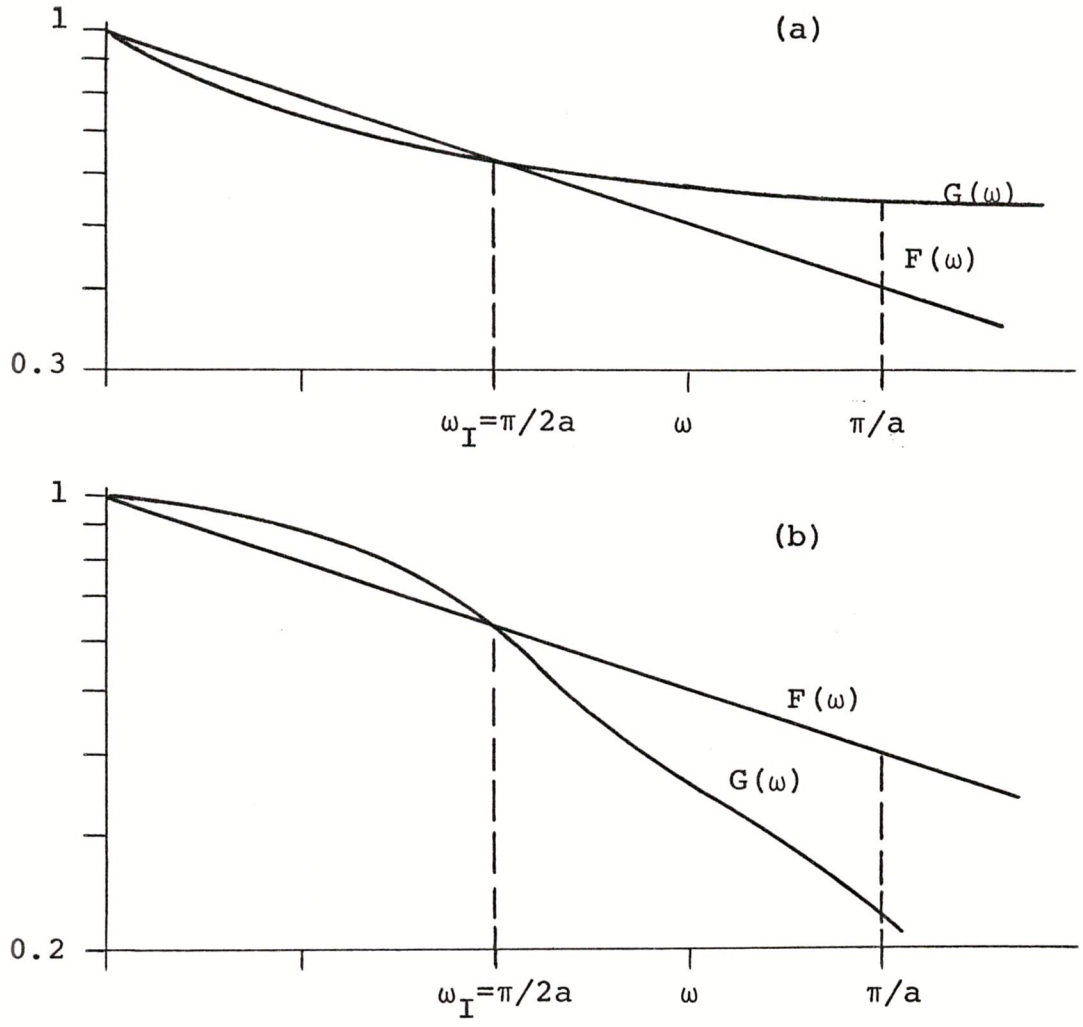

Figure 4.4 Heuristic Classification of an Arbitrary Transform, $G(\omega)$. $F(\omega)$ is a standard function $\exp(-\theta\omega)$ where θ is chosen so that the intersection of the two curves occurs at $\omega_I=\pi/2a$. In (a), $G(\omega)$ is not self-linearly conformal. In (b) $G(\omega)$ is self-linearly conformal. Note that the ordinate scale is logarithmic.

is valid for $a = \pi/2\omega_I$. Note that there are no other intersections between $G(\omega)$ and $F(\omega)$ in $0<\omega<2\omega_I$ than at ω_I.

c. $G(\omega)$ and $F(\omega)$ intersect at other points as well as ω_I in $0<\omega<2\omega_I$. The heuristic is indeterminate for such situations, with respect to the corresponding value of "a". In view of the desire to be conservative in stating whether (4.62) or (4.63) is valid, for a particular tolerance limit "a", the relation (4.62) is assumed for the indeterminate situation just described.

4. Cease the rotation of $F(\omega)$ when the intersection exists at $\omega_{1/4}$. The heuristic is now slightly modified. The determination of whether (4.62) or (4.63) is valid for $a \leq \pi/2\omega_{1/4}$ is made entirely on the determination of the $a=\pi/2\omega_{1/4}$ tolerance limit (that is for $\omega_I = \omega_{1/4}$). One must still have only one intersection of $F(\omega)$ and $G(\omega)$ in the range $0<\omega< \omega_I$ for 1c. If this situation exists, $G(\omega)$ is unconditionally 1c. The same indeterminacy can result as in step 3. That is, if multiple intersections occur in $0<\omega<2\omega_I$, the condition of (4.62) is assumed by default.

Before developing the heuristic procedure, it is instructive to give an example of how the procedure is applied. Consider the transform pair

$$.234\exp(-9|u|)+.12[\exp(-|u-.81|)+\exp(-|u+.81|)]$$

$$\leftrightarrow .4212/(\omega^2+.81)+.48\cos(8\omega)/(\omega^2+1) \qquad (4.64)$$

The pdf given in (4.64) is a continuous approximation to the discrete example of Table 4.1. Figure 4.5 illustrates the heuristic method applied to (4.64). Note that a semi-log plot was used, so that the standard function of the heuristic is a straight line drawn from the point (1,0) through the $G(\omega)$ curve. The intersection of this standard curve indicates that the heuristic is determinate for $\pi/2a \leq .24$. That is, for an intersection at $\pi/2a$, no other intersection occurs for $\omega<\pi/a$. The minimum value of the tolerance limit "a" is therefore $a_{min}=\pi/.48=6.545$. Relation (4.63) is therefore valid for $G(\omega)$ given in (4.64) when $a \geq 6.545$.

To compare the heuristic result with a formal analysis, one must determine when

$$(1/\pi)\int_{-\infty}^{\infty} G(\omega)\frac{\sin\ \omega a}{\omega}d\omega \leq (1/\pi)\int_{-\infty}^{\infty} G^2(\omega)\frac{\sin\ 2\omega a}{\omega}d\omega \qquad (4.65)$$

is true. Using complex integration of the left and right-hand sides of (4.65) for $G(\omega)$ given in (4.64) results, respectively in[1]

---

[1]The derivation of these results is lengthy and omitted here. However, the mechanics are a straightforward application of the Cauchy Integral Theorem (Papoulis [1962]).

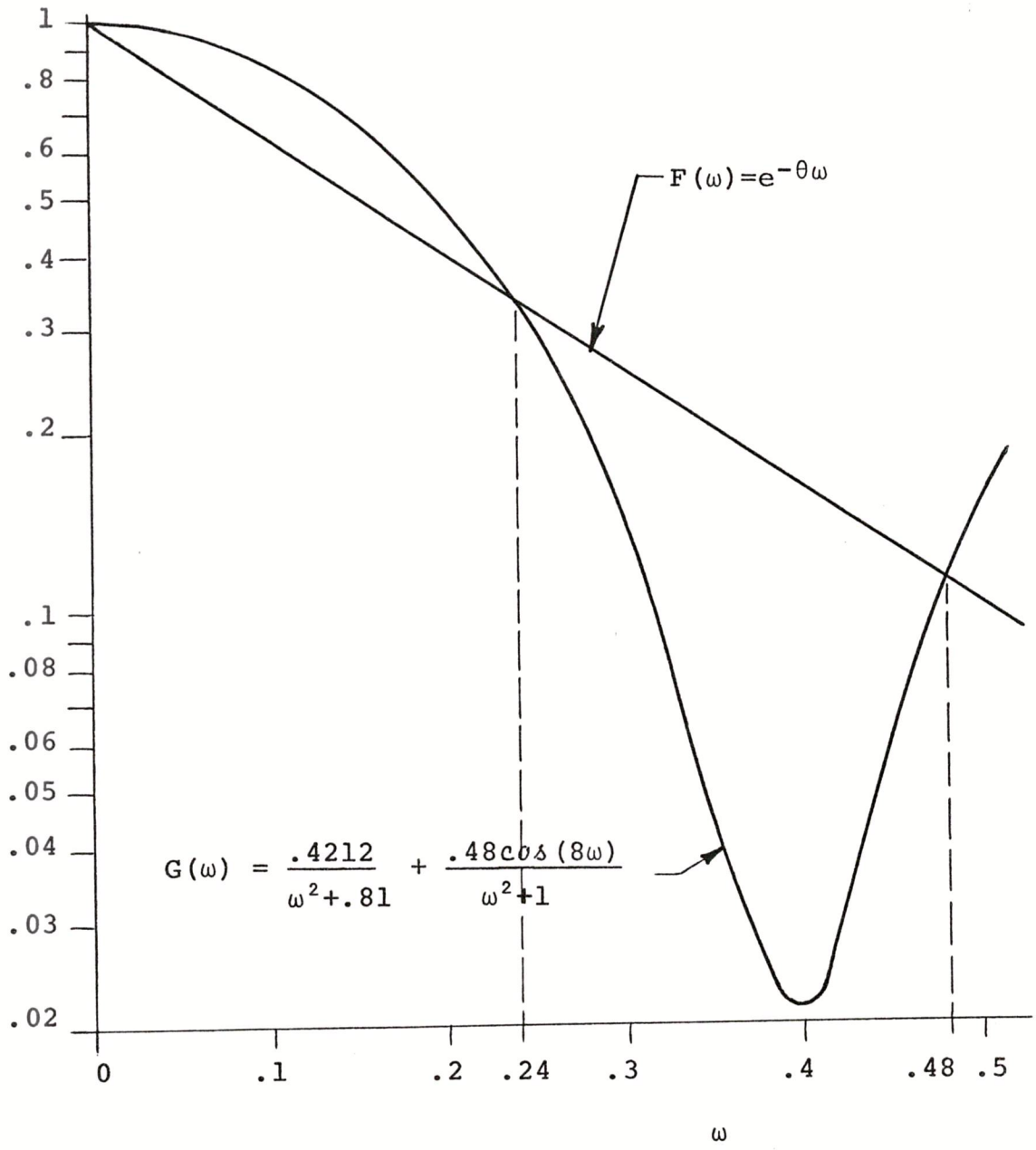

Figure 4.5 Example application of the heuristic classification method to a conditionally, linearly conformal probability distribution. Note that the ordinate scale is logarithmic.

$$\frac{1}{\pi}\int_{-\infty}^{\infty} G(\omega)\ \frac{\sin\omega a}{a} = .52(1-e^{-.9a})+.48\ \sinh(a)\cdot e^{-8} \qquad (4.67)$$

valid when "a" $\leq$ 8

and

$$\frac{1}{\pi}\int_{\infty}^{\infty} G^2(\omega)\ \sin\omega a \simeq .2704(1-e^{-.9a}[1+.9a])+.1152[1-e^{-2a}(a+1)]$$
$$+2.1282\ [\frac{e^{-7.2}\sinh(1.8a)}{.81} - e^{-8}\sinh(2a)] \qquad (4.68)$$

valid when "a" $\leq$ 4.

The value of "a", for which (4.67) equals (4.68), was determined by the bisection-method (Hamming [1962]) on a CDC 6500 computer. It was determined that

$$a_{analytic} = 3.6047.$$

The heuristic method therefore gives a conservative value of the minimum TL's. To see how conservative this bound is, the analytic and heuristic minimum values for "a" were substituted in the right-hand side of (4.67), giving

$$p\ \{|X| \leq a_{min}\} \simeq .575$$

$$p\ \{|X| \leq a_{analytic}\} \simeq .503.$$

The development of the heuristic is given shortly. First, the requirement of equal coefficients for the RV's X_1 and X_2 within equation (4.60) is considered. One may find this requirement for applying the heuristic excessively restrictive. However, for symmetric distributions, it is possible to show that equal coefficients within the linear

combination represents either a minimum or a maximum probability combination with respect to the relative variation of weights.

Consider the linear combination

$$Y = X_1 + C \cdot X_2. \tag{4.69}$$

This combination represents the various relative coefficients for the two RV's, X_1 and X_2, where the coefficient of X_1 is taken as unity without loss of generalization. For $C = 0$, the probability associated with the TI on Y is obviously equal to that for the corresponding TL on X_1. Assuming that X_1 and X_2 have identical pdf's, the BA is applied as usual for some confidence level α on X_1 and X_2. The confidence level of the TI formed by the algorithm is, from equations (4.15), (4.22) and from the assumption that $G(\omega)$ is symmetric,

$$p\{|X_1+CX_2| \leq a+C \cdot a\} = \frac{2}{\pi} \int_0^{\infty} G(\omega) G(C\omega) \frac{\sin(a[1+C]\omega)}{\omega} d\omega \tag{4.70}$$

where $G(\omega)$ is the transform of the pdf for X_1 and X_2 and "a" is the usual TI constant as in equation (4.13)

It is possible to show that the probability of (4.70) has a min-max point at $C = 1$ with respect to C. Substituting the new variable of integration

$$z = [1 + C]\omega$$

equation (4.70) becomes

$$p\{|X_1+CX_2| \leq a+C\cdot a\}$$

$$= \frac{2}{\pi}\int_0^\infty \frac{G(z/[1+C])\cdot G(Cz/[1+C])\sin(az)dz}{z}. \qquad (4.71)$$

Differentiating with respect to C, the right-hand side of (4.71) becomes, denoting the derivative by a prime,

$$\frac{2}{\pi}\int_0^\infty G'(z/[1+C])G(Cz/[1+C])(-z/[1+C]^2)\frac{\sin az\, dz}{z}$$

$$+\frac{2}{\pi}\int_0^\infty G(z/[1+C])G'(Cz/[1+C])(z/[1+C]^2)\frac{\sin az\, dz}{z}.$$

If C = 1 this expression becomes

$$\frac{2}{\pi}\int_0^\infty G\left(\frac{z}{2}\right) G'\left(\frac{z}{2}\right)\left(\frac{z}{4}\right)\frac{\sin az}{z}(-1+1)\,dz = 0$$

which is the result required to be shown. Although this may not be the only min-max point for $G(\omega)$, one can easily show that only one such point exists for the exponential transform family by using the function $\exp(-\theta|\omega^k|)$ in the preceding analysis. For other functions, one may at least intuitively justify using equal coefficients as the worst (or best) case for classifying $G(\omega)$.

The heuristic classification method presented in this section will now be developed. The method is concerned with determining when (see (4.62) and (4.63))

$$(1/\pi)\int_{-\infty}^{\infty} G^2(\omega)\sin(2\omega a)d\omega/\omega \geq (1/\pi)\int_{-\infty}^{\infty} G(\omega)\sin(\omega a)d\omega/\omega \qquad (4.72)$$

is true for the function $G(\omega)$. Several approximations are made in the following analysis. The first approximation is that only the first half-period of the *sinc* function (that is, the function $sin\omega a/\omega$) is considered. This approximation is based on the fact that the *sinc* function falls off as $1/\omega$ and that the magnitude of $G(\omega)$ decreases monotonically for many pdf transforms.

A second approximation is that $sin2\omega a/\omega$ and $sin\omega a/\omega$ are each replaced by a rectangular function having the same area under them as the original functions and with widths equal to the corresponding half-period. It is noted that this second approximation is exact for $G(\omega)=\exp(-\theta|\omega|)$.

Incorporating the two approximations and removing the constants by cancellation, one obtains the approximated representation of (4.72) for symmetric $G(\omega)$, as

$$\int_0^{\pi/2a} G^2(\omega)d\omega \geq (1/2)\int_0^{\pi/a} G(\omega)d\omega. \qquad (4.73)$$

Recall the discussion of equation (4.17) which indicates that as "a" increases, the behavior of $G(\omega)$ and $G^2(\omega)$ near the origin (in equation 4.72) tends to dominate both integrals. Therefore, inequality (4.73) improves as an approximate representation when "a" increases.

The analysis considers the two classes of $G(\omega)$ represented in Figure 4.4. It is to be shown that Figure 4.4a corresponds to the invalidity of (4.73), while Figure 4.4b

preserves the inequality. For each of these two classes, two cases are considered. The first case is when $G(\pi/2a)\geq.25$, and the second is when $G(\pi/2a)\leq.25$.

First consider that $G(\pi/2a)\geq.25$, for both classes shown in Figure 4.4. At some ω_I such that $G(\omega_I)\geq.25$, construct a standard function, $F(\omega)=\exp(-\theta|\omega|)$ such that it intersects $G(\omega)$ at ω_I. Define

$$\Delta(\omega)=G(\omega)-F(\omega). \tag{4.74}$$

The intersection point is to correspond to $\pi/2a$. Substituting (4.74) in (4.73), one obtains

$$\int_0^{\pi/2a}[F(\omega)+\Delta(\omega)]^2 d\omega \overset{?}{\geq} (1/2)\int_0^{\pi/a}[F(\omega)+\Delta(\omega)]d\omega \tag{4.75}$$

where the direction of the inequality is unknown. One can easily show that for the standard (exponential) function $F(\omega)$

$$\int_0^{\pi/2a}F^2(\omega)d\omega = (1/2)\int_0^{\pi/a}F(\omega)d\omega. \tag{4.76}$$

Expanding the left-hand side of (4.75) and subtracting (4.76) from both sides, yields

$$\int_0^{\pi/2a}\Delta(\omega)[\Delta(\omega)+2F(\omega)]d\omega \overset{?}{\geq} (1/2)\int_0^{\pi/2a}\Delta(\omega)d\omega+(1/2)\int_{\pi/2a}^{\pi/a}\Delta(\omega)d\omega$$

or (4.77)

$$\int_0^{\pi/2a}\Delta(\omega)[\Delta(\omega)+2F(\omega)-1/2]d\omega-(1/2)\int_{\pi/2a}^{\pi/a}\Delta(\omega)d\omega \overset{?}{\geq} 0.$$

Since $G(\omega)$ and $F(\omega)$ are at least .25 in $0\leq\omega\leq\pi/2a$, $G(\omega)+F(\omega)-1/2\geq 0$, $0<\omega<\pi/2a$. In Figure 4.4a, $\Delta(\omega)\leq 0$ for $0\leq\omega\leq\pi/2a$, and $\Delta(\omega)\geq 0$ for $\pi/2a\leq\omega\leq\pi/a$. In this situation, both integrals of (4.77) are non-positive. Therefore inequality (4.73) is not satisfied.

Now consider Figure 4.4b and (4.77). Here $\Delta(\omega)\geq 0$ for $0<\omega<\pi/a$, and $\Delta(\omega)\leq 0$ for $\pi/2a\leq\omega\leq\pi/a$. Both integrals of (4.77) are non-negative, and (4.77) is satisfied (and 4.63 is valid).

The case where $G(\pi/2a)\leq .25$ is now considered. The function $F(\omega)$ is now constructed so that it intersects $G(\omega)$ at $G(\omega_I)=.25$ (if $G(\omega)=.25$ at more than one point, $\omega_{1/4}$ assumes the smallest value of those corresponding to $G(\omega)=.25$). Inequality (4.75) is now written, using (4.74), as

$$\int_0^{\omega_{1/4}} [F(\omega)+\Delta(\omega)]^2 d\omega+\int_{\omega_{1/4}}^{\pi/2a} [F(\omega)+\Delta(\omega)]^2 d\omega \overset{?}{\geq} (1/2)\int_0^{\pi/a} [F(\omega)+\Delta(\omega)]d\omega. \tag{4.78}$$

Subtracting (4.76), bringing all terms to the left-hand side, and using (4.74), one obtains

$$\int_0^{\omega_{1/4}} \Delta(\omega)\,[F(\omega)+G(\omega)-(1/2)]\,d\omega+\int_{\omega_{1/4}}^{\pi/a} \Delta(\omega)\,[F(\omega)+G(\omega)-(1/2)\,d\omega - (1/2)\int_{\pi/2a}^{\pi/a} [\Delta(\omega)]d\omega \overset{?}{\geq} 0. \tag{4.79}$$

Since $F(\omega)$ and $G(\omega)$ are at least .25 for $\omega \leq \omega_{1/4}$, and less than .25 for $\omega_{1/4} < \omega \pi/2a$,

$$F(\omega)+G(\omega)-(1/2) \geq 0, \quad \omega \leq \omega_{1/4}$$

$$F(\omega)+G(\omega)-(1/2) \leq 0, \quad \omega_{1/4} \leq \omega \leq \pi/2a.$$

In Figure 4.4a, $\Delta(\omega) \leq 0$ for $\omega \leq \omega_{1/4}$ and $\Delta(\omega) \geq 0$ for $\omega \geq \omega_{1/4}$. The three integrals of (4.79) are non-positive. Therefore, inequality (4.73) is not satisfied.

Now consider Figure 4.4b and (4.79). Here $\Delta(\omega) \geq 0$ for $\omega \leq \omega_{1/4}$ and $\Delta(\omega) \leq 0$ for $\omega \geq \omega_{1/4}$. The three integrals are non-negative satisfying the inequality (which implies 4.63 is valid).

As a final note, one may have realized that the above analysis is conservative. This arises since the integrands of (4.77) and (4.79) were required to be either non-negative or non-positive throughout the range of integration. Such a requirement suggests that the value of "a", determined from the heuristic, which satisfies (4.63), will be a conservative estimate (larger than needed).

As a final consideration of this section, the heuristic procedure is generalized.

Theorem 4.9 Let $G_i(\omega)$ be any of N non-negative transforms for an associated pdf, $g_i(u)$. If each $g_i(u)$ is classified by the heuristic as SLC, then the N-fold convolution of the $g_i(u)$ will be classified by the heuristic procedure as SLC.

Conversely, if each $g_i(u)$ is classified as not SLC, then the N-fold convolution of the $g_i(u)$ will be classified as not SLC.

Proof Let the pdf's $g_i(u)$ be each classified SLC by the heuristic, and let their transforms be denoted $G_i(\omega)$. Then at any ω, say ω_I, there exists standard functions $F_i(\omega)$ such that

$$F_i(\omega_I)=\exp(-\theta_i|\omega_I|)=G_i(\omega_I).$$

By assuming each $g_i(u)$ is classified as SLC, it follows that $G_i(\omega)\geq\exp(-\theta_i|\omega|)$ for $\omega\leq\omega_I$, and $G_i(\omega)<\exp(-\theta_i|\omega|)$ for $\omega\geq\omega_I$. But

$$\prod_{i=1}^{N} F_i(\omega) = \prod_{i=1}^{N} \exp(-\theta_i|\omega_I|)=\exp\left(-\sum_{i=1}^{N} \theta_i|\omega_I|\right) \tag{4.80}$$

and

$$\begin{aligned} &\prod_{i=1}^{N} G_i(\omega)\geq\exp\left(-\sum_{i=1}^{N} \theta_i|\omega|\right), \quad \omega\leq\omega_I \\ &\prod_{i=1}^{N} G_i(\omega)\leq\exp\left(-\sum_{i=1}^{N} \theta_i|\omega|\right), \quad \omega\geq\omega_I. \end{aligned} \tag{4.81}$$

One recognizes (4.80) and (4.81) as the criteria used by the heuristic to classify a pdf with transform $\prod_{i=1}^{N} G_i(\omega)$, where the standard function has parameter (exponent) $\theta_T=\sum_{i=1}^{N} \theta_i$. Since $\prod_{i=1}^{N} G_i(\omega)$ is the transform of the N-fold convolution of the N pdf's $g_i(u)$, it is clear that the heuristic classifies the N-fold convolution as SLC.

The converse can be proven by reversing the sense of the inequalities in the first part of the proof.

QED

Theorem 4.9 indicates that for a summation of N RV's each having SLC pdf's, the pdf of the sum will also be SLC. For pdf's satisfying the Central Limit Theorem (Mood and Graybill [1963]), this result is expected since the pdf of the sum approaches a Normal pdf as N increases. However, Theorem 4.9 is more general by handling pdf's with unbounded variances.

4.6 Concluding Remarks The BA, defined in this chapter, provides a simple means of obtaining a TI for a linear sum of independent RV's in terms of the TI's of the summed RV's. It has been the philosophy of this thesis that when statistical bounds on a quantity are determined through an approximation, that the probability associated with such bounds should be conservative. In the case of a TI estimated by the BA, a conservative estimate means that the confidence level associated with the estimated TI must be less than or equal to the analytically determined confidence level. That is, the pdf's of the linear sum are linearly conformal.

When some of the RV's, combined in a linear sum, are in fact monotonically dependent, and their pdf's are lc, the BA determines a conservative bound estimate, which is

the CTI[1] for the linear sum (from Theorems 3.8 and 3.9). When the RV's are linearly dependent, then the CTI is no longer an approximation, but rather an exact computation of the TI for the linear sum. That is, the actual confidence level for the CTI equals the confidence level for the individual RV's. One concludes that the condition of statistical independence assumed in this chapter results in more conservative analysis than when the RV's of a linear combination are monotonically dependent.

It was shown in the beginning of the chapter, that the BA does not give a conservative estimate of the TI for a linear sum of RV's, having non-lc pdf's. In order to study the conditions in which the BA does determine a conservative TI estimate, several properties of the transforms of pdf's have been examined. It was found that the $(\sin x)/x$ function was of central importance in convolution considerations. The Central Limit Theorem was discussed in relation to the TI of linear combinations of RV's. One important result was that for a linear sum of RV's which satisfy the conditions of the Central Limit Theorem, the distribution of the RV representing the linear sum approached the normal distribution as the number of terms of the sum increased. Since

---

[1]The Composite Tolerance Limit (CTI) was defined earlier in Definition 4.1.

the normal distribution was shown to be SLC, the conjecture made was to treat the pdf of large sums of RV's as lc even if some of the distributions were not SLC. Also inferred was that if a minimum confidence level could be found for which a RV was lc, it would also be lc for larger levels.

The exponential transform family was defined. This family of functions was shown to be partitioned into two classes with respect to the property of linear conformity. By comparing other functions to members of the family, one may deduce whether the compared functions exhibit the property of linear conformity. The exponential transform family contains the delta, Cauchy, and Normal pdf's, and is therefore important in its own right. Finally, a heuristic method was developed which enables the classification of a much larger number of functions by a simple graphical procedure. The heuristic classification procedure assumed that $G(\omega)$ was symmetric and real. Since all pdf's are positive real functions, the real part of any pdf transform must be symmetric (Papoulis [1972]). When the pdf is not symmetric, its transform will contain an imaginary part.

A final word about non-symmetric pdf's may be offered. When a distribution is only slightly non-symmetric, an approach that can be considered is to treat only the real part of the Fourier transform of this pdf. For a large lack of symmetry, the TI definition becomes somewhat arbitrary.

It appears that nothing much can be said in relation to linear conformity in this case. When several RV's having highly skewed pdf's are summed, the resulting distribution becomes less skewed and eventually symmetric. Therefore, in this case one may group the individual RV's as sub-summations, and determine linear conformity by analytic means for each group. Since each grouped RV's will have a more symmetric pdf, one may still be able to apply the techniques of this chapter.

As a final note to this chapter, it is not to be assumed that $G(\omega)$ must be non-negative for the heuristic method to be valid. Indeed, the transform of the rectangular distribution is of the form

$$p_T(u) \leftrightarrow \frac{\sin\omega T}{\omega T} \tag{4.82}$$

where the distribution is centered about the "u" origin and has width 2T. This distribution can be shown to be lc. Application of the heuristic method indicates that there is no indeterminate region in the ω domain, and that the transform given in (4.82) corresponds to the lc case.

In the next chapter, an important application of the techniques developed to this point in the thesis is developed. This application considers the direct utilization of TI's in classical polynomial and transcendental interpolation, and allows the replacement of deterministic data by TI's.

5. INCORPORATION OF STATISTICAL INFORMATION FOR POLYNOMIAL AND TRANSCENDENTAL INTERPOLATION

5.1 Assumptions and Background This chapter is concerned with applying statistical information to the classical problem of approximating continuous functions by interpolation formulas. It is appropriate to begin the discussion with assumptions about the functions to be approximated and assumptions about the statistical distributions associated with the incorporated statistical information. Additionally, this section provides a brief background to the interpolation formulas considered.

The functions to be approximated are assumed to be instantaneous, continuous, and allowably non-linear functions of an independent variable. It is further assumed that the derivatives of the function exist within the interval of interpolation. The maximum order of the derivatives required to exist depends on the proposed classical interpolation formula. That is, if the interpolation formula uses the derivative values of order as great as k, then the first k derivatives of the function are required to exist.

The independent variable is referred to as the input, x, and the functional value, f(x), is referred to as the output. The functional value f(x) at x_i is to be viewed as a random variable (RV), denoted f_i. In particular, it is assumed that a probability density function (pdf) exists for the functional value f_i. That is, the unconditional

probability

$$p\{f_i \leq a\} = \int_{-\infty}^{a} g(x_i, v)\, dv \qquad (5.1)$$

is defined in terms of the pdf $g(x_i, v)$, where x_i is a determinate parameter. It is assumed that the function $g(x_i, v)$ is continuous in x_i and v.

Before continuing with the main discussion, it is useful to indicate to the reader the physical interpretation to be attached to the pdf with regard to f(x). In this thesis, the principle application of the developments is to the automated testing of continuous nonlinear devices. For a particular manufacturer's device type, a randomly selected device of this type may be characterized in part by a continuous input-output function f(x). It is implied by equation (5.1) that prior to any testing of this selected device, it is possible to attach a probability to the event that the actual value of f_i for the device lies within a given range or tolerance interval (TI). Once the particular device is measured at x_i, the value of $f(x_i)$ is determined and is no longer a RV. That is, it is assumed that repeated measurement of the same device at x_i results in the same measured value. However, the value of f(x) at some x different from x_i remains a random quantity. Intuitively, one recognizes that the value at the second point is not entirely statistically independent of the first measured value. It will be assumed in what follows that two different

x_i's will be sufficiently different so that the functional values of f(x) at these two points may be considered independent RV's. No attempt will be made to form a criteria for the sufficiency of separation in what follows. In a later section of the chapter, the problem of non-independence of functional values for a particular selected device will be discussed. It will be shown that the assumption of statistical independence does not invalidate the developments of the chapter under certain conditions of actual dependence.

The pdf $g(x_i, v)$ is assumed to exist for each x_i of interest within the interpolation interval. In practical situations it would be difficult to determine a pdf at a single point x_0, let alone at all points within an interval of parameter x. However, the techniques of this chapter require the knowledge of a TI for a specific confidence level, where the TI is defined according to some a priori convention. For example, the TI limits or tolerance limits (TL's) defined symmetrically about the distribution mean are assumed known at specific points x_i within the range of x that f(x) is to approximated. By using conventional interpolation of the limits, the TL's may be obtained as functions of x. The error resulting from such an approximation will not be considered in detail, but it will be briefly treated in the chapter conclusions. This error will be

assumed negligible. Also considered in the last part of the chapter will be practical approaches to gathering the TL values.

This chapter considers the incorporation of statistical information (in the form of TI's) in conventional interpolation formulas. The class of interpolation formulas considered is exact-fit formulas. The term "exact-fit" is meant to convey the idea that the formula matches the data points exactly, as opposed to a "best-fit" match. These data points normally correspond to the pairs $(x_i, f(x_i))$ representing measured values of $f(x)$ at particular x_i. Two types of exact-fit formulas are considered in this chapter. These are polynomial interpolation and cardinal or transcendental interpolation defined below. The techniques of the chapter are not restricted to these two types, as will be explained shortly.

The first type of formula is that of polynomial interpolation. The various formulas differ in the order of derivatives of $f(x)$ used as data. For example, the first k derivatives of $f(x)$ at each of m points may be used to obtain $km - 1$ order polynomial interpolation. It is easily shown that no matter what formulation is chosen, the same polynomial is produced for any given set of data points (Hamming [1962]). It is, therefore, true that when several

methods, which use the same data are being considered, the basis for choice should be computational efficiency.

In deciding what polynomial formulation to use, practical considerations dictate that high-order derivatives be avoided and that the order of the polynomial be kept small. Increasing measurement error is associated with higher order derivatives. One limits the order of the polynomials used since high-order polynomials are oscillatory, and any error bound that may be obtained can become enormous as the separation from a data point increases. In order to avoid this problem, one may break up the interval of x over which the interpolation is to be performed, and consider each subinterval separately. Thus, only the data points within the subinterval are used, and a lower order polynomial interpolation results.

This chapter does not consider the conventional techniques related to obtaining error bounds for polynomial interpolation. For further information on this topic the reader is referred to Ralston [1965] and Hamming [1962]. The related problem of determining an adequate order for the interpolation is not discussed here, except to say that such information must result from prior experience with the functions to be approximated.

The particular polynomial interpolation formulas will now be presented. The first polynomial interpolation

formula uses only values of f(x) and avoids derivatives. The formulation to be used is the Lagrange form

$$y(x) = \sum_{i=1}^{N} L_i(x) f(x_i) \tag{5.2}$$

where (Hamming [1962])

$$L_i(x) = \prod_{\substack{j=1 \\ j\neq i}}^{N} (x-x_j) \Big/ \prod_{\substack{j=1 \\ j\neq i}}^{N} (x_i-x_j) \tag{5.3}$$

and y(x) is the interpolated value. This form requires N values of f(x) at x_i, i=1,...N, and produces an N-1 order polynomial. As will be seen later in the chapter, the Lagrange form is well suited to automatic computation, especially in adaptive procedures for which the values of x_i are chosen dynamically.

A second polynomial interpolation form uses the values of f(x) and its derivatives f´(x) at each of N points, x_i, i=1,2,...N. The formulation selected is the Hermite form (Hamming [1962]):

$$y(x) = \sum_{i=1}^{N} H_i(x) f(x_i) + \sum_{i=1}^{N} h_i(x) f'(x_i) \tag{5.4}$$

where

$$h_i(x) = (x-x_i) \prod_{\substack{j=1 \\ j\neq i}}^{N} (x-x_j)^2 \Big/ \prod_{\substack{j=1 \\ j\neq i}}^{N} (x_i-x_j)^2 \tag{5.5}$$

$$H_i(x) = [1+a_i(x-x_i)] \prod_{\substack{j=1 \\ j\neq i}}^{N} (x-x_j)^2 \Big/ \prod_{\substack{j=1 \\ j\neq i}}^{N} (x_i-x_j)^2 \tag{5.6}$$

and where

$$a_i = -2 \sum_{\substack{j=1 \\ j \neq i}}^{N} 1/(x_i - x_j).$$

These two formulations will be used in subsequent developments. It should be clear that a mixture of these two methods is possible in which f(x) values are used at some x_i and $f'(x)$ values are used at others. These variations add to the complexity of computations and for this reason such interpolation formulas may be less desirable, in practice. Formulas similar to the Lagrange and Hermite forms may be considered when the required derivatives of function f(x) can be accurately measured. However, higher order derivatives are usually more difficult to measure, as previously pointed out. For this reason, formulas requiring higher order derivatives should usually be avoided.

Another type of exact-fit interpolation formula considered is the transcendental formulation which is a cardinal series summation. This formulation is useful when the function to be approximated corresponds to a band-limited function (Papoulis [1962]), in the sense that the function contains no frequencies greater than a cutoff frequency, ω_c. Based on the well-known sampling theorem (Papoulis [1962]), this interpolation uses a finite number of equispaced samples. Several authors (Helms and Thomas [1962] and

Brown [1969]), consider the error that results from truncating the cardinal expansion of the sampling theorem. Practical limitations require that the interpolated function contain frequency components no greater than $r\omega_c$, where $0<r<1$ and is related to the error bound given below. The cardinal interpolation formula is then

$$y(x) = \sum_{m=K-N}^{K+N} f(m/2\omega_c) \cdot sinc(2\omega_c[x-m/2\omega_c]) \qquad (5.7)$$

for $0<N<\infty$

where K is an integer chosen such that

$$2\omega_c x - \tfrac{1}{2} \leq K(x) \leq 2\omega_c x + \tfrac{1}{2}$$

and where

$$sinc\ z \equiv sin\ (\pi z)/\pi z.$$

The constant K is thus chosen so that an equal number of sample points lie on either side of x so that K is a function of x. The upper bound on the error is written in terms of the maximum of f(x). Let this maximum be written as:

$$M \equiv \max|f(x)| \qquad -\infty<x<\infty$$

Considering N and r in (5.7), the error bound is:

$$|E(x)| \leq 4M/\pi^2 N(1-r).$$

Other interpolation formulas, which utilize the assumption of a band-limited function are possible. For example, a similar formula to (5.7) is possible if both values of f(x) and its derivative are used (Helms and Thomas [1962]).

This formula is somewhat harder to implement for the proposed approach and is not considered here, although it is possible to apply the same techniques as for (5.7).

For the cardinal series interpolation formulas, it is necessary to have a priori knowledge of the bandwidth of f(x). Such information apparently takes the place of the knowledge of the minimum order of a polynomial interpolation. It is interesting to note that an error bound can be determined for Lagrange interpolation when f(x) is band-limited (Radzyner and Bason [1972]), providing the data are nearly equispaced. The practical problem of estimating bandwidth is briefly discussed at the end of the chapter when implementation strategies are discussed.

The two types of exact-fit interpolation formulas considered allow several observations to be made. Equations (5.2) and (5.7) are summations of data multiplied by interpolation weighting functions.[1] These functions are independent of the data values. Also, the weighting functions can be seen to change sign only at data points. These two facts will be useful in the subsequent developments. It will be useful to represent both (5.2) and (5.7) in what follows by the general form

---

[1]Yen [1956] distinguishes these from conventional polynomial coefficients by the term "composing functions".

$$y(x) = \sum_{i=1}^{N} w_i(x) \cdot f(x_i) \qquad (5.8)$$

where the required substitutions and constraints are determined by the particular form of interpolation formula represented by (5.8). In fact, the developments that follow will be valid for other interpolation formulations which have the form of (5.8).

This chapter will be concerned with the direct replacement of one or more data points normally required in equation (5.8) by a tolerance interval for the replaced data value. It is important to note that such an approach does not invalidate the error bounds normally associated with interpolation formulas of the type considered since such error bounds are deterministic in nature. Nor does the proposed incorporation of statistical information result in a new or unconventional interpolation formula.[1] The last two assertions will be discussed further after the proposed techniques of the chapter are developed.

In the next section, the replacement of a single deterministic data value will be considered. This development will parallel an approach introduced in Chapter 3. A generalization resulting in the replacement of several data

---

[1] In the next chapter, a different approach to the incorporation of statistical information for interpolation will result in unconventional interpolation formulas.

points by statistical information will be studied in a subsequent section. This generalization will utilize techniques developed in Chapter 4.

5.2 Single Tolerance Interval Interpolation This section concerns how probabilistic information in the form of TI's can be used to replace deterministic data in an exact-fit interpolation formula. The deterministic data corresponds to measured values of a device input-output response at particular input values. These measured values $f(x_i)$, more compactly f_i, at x_i are distinguished from what will be called "statistical data". The statistical data value for the function $f(x)$ at x_i is a statistical bound for $f(x_i)$ corresponding to a specific confidence level α. This bound is a tolerance interval (Bowker and Lieberman [1964]) determined according to some convenient convention. For example, the convention may be to take this TI about the distribution mean for RV f_i such that the area under the pdf to either side of the mean and within the TI is equal. That is, for TI [a,b], confidence level, α, and pdf mean μ,

$$\int_a^{\mu} g(x_i,v)dv = \int_{\mu}^{b} g(x_i,v)dv = \alpha/2 \qquad (5.9)$$

where $g(x_i,v)$ is the pdf for f_i as in (5.1). Many other possible conventions exist. For example, one may define the TI such that the "tails" of the distribution have equal probability. Again for TI [a,b], this convention requires

that

$$\int_{-\infty}^{a} g(x_i,v)dv = \int_{b}^{\infty} g(x_i,v)dv = (1-\alpha)/2. \qquad (5.10)$$

For a symmetric pdf (5.9) and (5.10) are equivalent.

In the following discussion, no assumption will be made about the particular convention used to define the TI of a random variable. The TI for f_i will be assumed to be given, a priori, and will be represented by the TL's a_i and b_i. That is, the TI is $[a_i,b_i]$ for f_i.

Suppose that N-1 data points are available for the general interpolation formula of (5.8). (Note that if (5.8) corresponds to the cardinal series of (5.7), then f(x) is assumed band-limited and the location of the x_i's in what follows are equispaced. If (5.8) corresponds to equation (5.2) then (5.8) is an $(N-1)^{th}$ order interpolation. For (5.4), the data values are either values of f(x) or its derivative at the x_i's.) In place of, for example, the N^{th} data point $f(x_N)$, substitute the TL's for $f(x_N)$ (or $f'(x_N)$ for Hermite interpolation and a derivative value being replaced). That is,

$$p\{f(x_N) \in [a_N,b_N]\} = \alpha \qquad (5.11)$$

where $0 \leq \alpha \leq 1$

and the TI is further constructed according to a predefined convention.

Once the TI of $f(x_N)$ has been obtained, the approach is to use this range of values for the normally required $f(x_N)$ value. By considering $f(x_N)$ now as a RV f_N, it is appropriate to rewrite (5.8) as a RV equation. That is

$$y(x) = C_1(x) + C_2(x) \cdot f_N \tag{5.12}$$

where x is now a parameter of the deterministic constants $C_1(x)$ and $C_2(x)$. It is clear from Theorem 3.1 that if (5.11) holds, then RV y(x) is bounded according to[1]

$$p\{C_1(x)+C_2(x)\cdot a_N \underline{\leq} y(x) \underline{\leq} C_1(x)+C_2(x)\cdot b_N\} = \alpha. \tag{5.13}$$

This result defines two curves to be constructed, between which the function f(x) is said to lie within a probability of α. These bounds do not include the conventional interpolation error which effectively widens the bounding. Such an error is usually determined as a function of the distance between an interpolation point and a data bound and also a function of an upper bound on a particular order derivative of f(x) within some interval. The incorporation of a deterministic error bound is considered in a later section.

Figure 5.1 illustrates a typical interpolation in which a single deterministic data value has been replaced by a

---

[1]It is assumed that $C_2(x)$ is positive here, but for negative values, the limits of (5.13) are interchanged. Also, the TL's a_N and b_N are generally functions of x, although not indicated in (5.13).

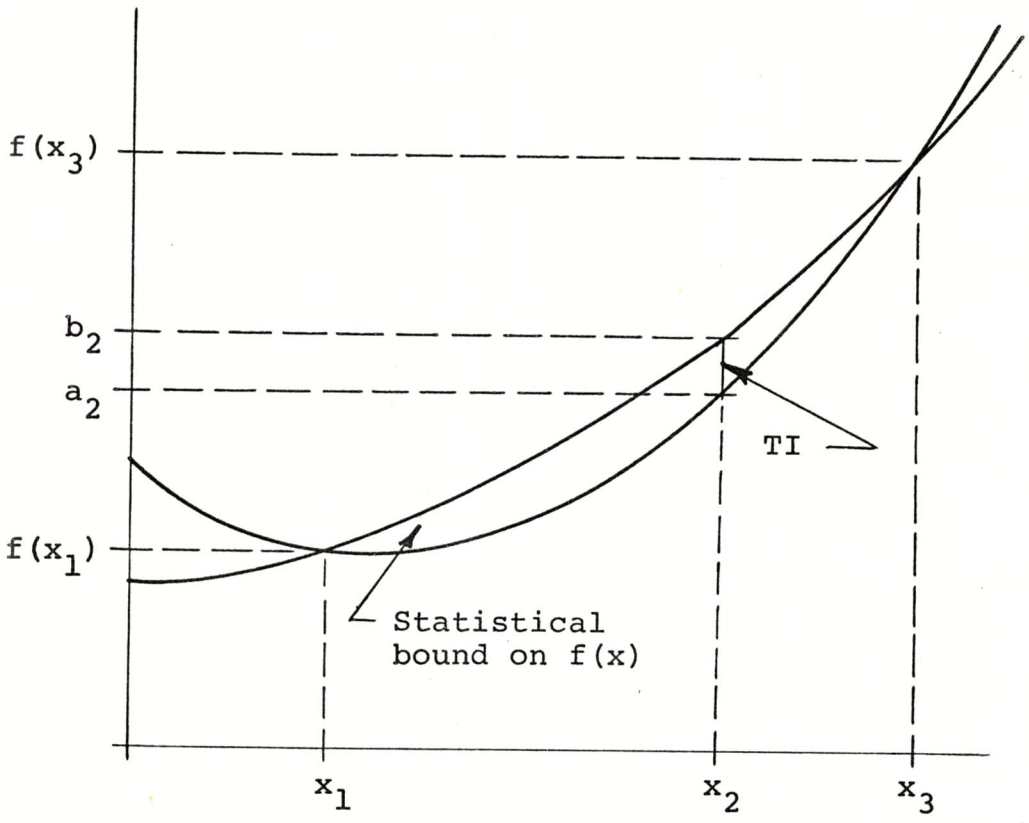

Figure 5.1 Polynomial interpolation in which one data point has been replaced by a tolerance interval (TI). At x_1 and x_3, deterministic data has been used. At x_2, the tolerance interval has been substituted for a deterministic data value.

TI. At points x_1 and x_3, the data are deterministic. At x_2, a TI has replaced the conventionally required data point.

Before considering the case where more than one deterministic data are replaced by statistical data, one must determine the effect of a change in the value of a data point, $f(x_i)$, on the interpolated value of $f(x)$ at some general point x. Two reasons for this consideration can be given. First, it has been implicitly assumed by the statement of (5.13) that the extremes of the TI (that is the TL's) correspond to the maximum width bounds at an interpolation point rather than other values for f_N within the TI. A second reason for studying the relation between the interpolation value and a data value is that of computational simplicity. That is, one must know which value within the TI to substitute in (5.12) to either minimize or maximize $y(x)$ for each x within the overall interpolation interval. The above two questions regarding the formation of a statistical bound are answered by the following two simple theorems.

Theorem 5.1 Let the exact-fit interpolation formula be of the form

$$y(x) = \sum_{k=0}^{K} \sum_{i=1}^{N} w_{ki}(x) \cdot f^{(k)}(x_i), \quad k \geq 0, \quad N \geq 0 \tag{5.14}$$

where $f^{(k)}(x_i)$ is the k^{th} derivative of function $f(x)$ at x_i

and $w_{ki}(x)$ is the interpolation weighting function (composing function, Yen [1956]). (The interpolated value of $f(x)$, $y(x)$, is an approximation and, in general, results in an interpolation error.) Let a tolerance interval for $f^{(k)}(x_i)$ be given as $[a_{ki}, b_{ki}]$, where $a_{ki} \leq b_{ki}$. Then $y(x)$ is maximized at x_0 by setting $f^{(k)}(x_i)$ according to the following

$$f^{(k)}(x_i) = \begin{cases} b_{ki} \; ; \; w_{ki}(x_0) \geq 0 \\ a_{ki} \; ; \; w_{ki}(x_0) \leq 0 \end{cases} \tag{5.15}$$

and minimized at x_0 by setting $f^{(k)}(x_i)$ according to the following

$$f^{(k)}(x_i) = \begin{cases} a_{ki} \; ; \; w_{ki}(x_0) \geq 0 \\ b_{ki} \; ; \; w_{ki}(x_0) \leq 0. \end{cases} \tag{5.16}$$

Proof Theorem 5.1 can be verified by inspection of (5.14). Indeed

$$\frac{\partial y(x_0)}{\partial f^{(k)}(x_i)} = w_{ki}(x_0).$$

Since the weighting functions w_{ki} are independent of $f^{(k)}(x_i)$. That is, $y(x)$ is a linear function of each of the data values. The choice of substituting a_{ki} or b_{ki} for $f^{(k)}(x_i)$ to maximize or minimize $y(x)$ at any point x_0 depends entirely on the sign of the weighting function of $f^{(k)}(x_i)$ at x_0.

The linear relationship between data values and interpolation values in exact-fit interpolation formulas of the polynomial and the equispaced cardinal transcendental formulas results in an obvious advantage over other interpolation methods. Had a least squares interpolation (Hamming [1962]) or a spline function approach (Greville [1969]) been used, determination of the values of $f^{(k)}(x_i)$ in $[a_{ki},b_{ki}]$ which would minimize or maximize $y(x)$ at x_0 would have been considerably more difficult. The increased difficulty of these other methods results since the effective weights of the data are dependent on the data. Because bound formation is required throughout an interpolation interval, it is clear that the chosen interpolation formulas are superior in terms of computational ease.

The next theorem indicates an additional advantage of the Lagrange interpolation formula.

Theorem 5.2 The interpolation weighting functions for Lagrange interpolation change their arithmetic signs only at the data points x_i. In particular, the *sign* of $w_i(x)$ at x_0 for Lagrange interpolation is equal to

$$sign\ [w_i(x_0)] = (-1)^{n_1+n_2}$$

where

n_1 = # data points to the right ($x_j > x_0$) of x_0 excluding x_i,

n_2 = # data points to the right of x_i ($x_j > x_i$).

Proof Theorem 5.2 is verified by inspection of the expressions for either the weighting function in (5.3) or in (5.7). Observe that in (5.3), the weighting function for the data value $f(x_i)$ is

$$L_i(x) = \prod_{\substack{j=1 \\ j\neq i}}^{N} (x-x_j) / \prod_{\substack{j=1 \\ j\neq i}}^{N} (x_i-x_j).$$

For x_j "to the right of" x, the term $(x-x_j)$ is negative. That is $x_j>x$. Similarly x_j to the right of x_i results in (x_i-x_j) negative. The *sign* computation in the theorem therefore follows directly.

This last theorem indicates that the Lagrange interpolation allows the *sign* of a weighting function to be determined just once within each interval between consecutive data points. Furthermore, the *sign* is determined solely from the relative positioning of the data points. Theorem 5.2 thereby allows a somewhat more simple bound formation for the Lagrangian interpolation.

The result of equation (5.13) shows it is possible to reduce the number of measurements required to characterize a device function f(x) provided the TI's, at various points within the interpolation interval, are available. In practice, one may predetermine the TL's at specific values of x or may instead determine an interpolation formula for these limits as a function of x, such that the limits correspond to a particular confidence level of interest.

(Note that these TL interpolation formulas need be determined only once.) The latter approach is more suited for an adaptive procedure in which the probabilistic data points (corresponding to substituted TI's) may be selected dynamically. Adaptive location of data points is briefly discussed later in this chapter.

5.3 Multiple Tolerance Interval Interpolation The approach introduced in the last section will now be extended for the situation in which several deterministic data points are replaced by TI's. It was just shown that the value of a data point used in the exact-fit interpolation formulas (e. g., equation 5.2) has a linear effect on the interpolated value. Furthermore, Theorem 5.1 indicates that when a data value is replaced by a TI, the maximization or minimization of the interpolation function at any x with respect to the value of a particular replaced data value, may be accomplished independently of other data values. Thus, the overall effect of substituting several TI for deterministic data may be considered as a superposition of several independent effects. That is if several data points are represented by specific TI's, the values of each $f(x_k)$ can be selected one by one in the same manner as was done for the single TI case.

Consider the situation in which the first k values of f(x) in equation (5.8) are deterministic data, and the

remaining N-k values are the RV f_i represented by TI. That is,

$$y(x) = \sum_{i=1}^{k} w_i(x) \cdot f(x_i) + \sum_{i=k+1}^{N} w_i(x) f_i \tag{5.17}$$

and

$$p\{a_i \leq f_i \leq b_i\} = \alpha, \quad i=k+1,\ldots,N. \tag{5.18}$$

The constants a_i and b_i represent the lower and upper TL's for f_i, respectively. Then the composite TI (CTI) is defined by the expressions

$$D_1 \equiv \sum_{i=1}^{k} w_i(x) \cdot f(x_i) + \sum_{i=k+1}^{N} [\min(w_i(x) \cdot a_i, w_i(x) \cdot b_i)] \tag{5.19}$$

and

$$D_2 \equiv \sum_{i=1}^{k} w_i(x) \cdot f(x_i) + \sum_{i=k+1}^{N} [\max(w_i(x) \cdot a_i, w_i(x) \cdot b_i)] \tag{5.20}$$

where min and max refer to the minimum and the maximum, respectively, of two values within the parenthesis. Then $y(x)$ is bounded by the TI $[D_1, D_2]$.

Theorem 5.1 indicated that one may consider each of the TI substitutions independently in obtaining a CTI bound on $f(x)$. The CTI, represented by D_1 and D_2, is based on a somewhat questionable assumption that the RV's f_i (and in turn the TI's on these RV's) can be considered in a statistically independent manner. In the next section, it will be argued why such an assumption still leads a conservative statistical bound.

Figure (5.2) shows an example of the replacement of two deterministic data values by TI's. The TI's have been substituted between x_1 and x_3. Note that a bound on f(x) results as shown earlier in Figure 5.1.

While a straightforward procedure has just been indicated for the determination of statistical bounds on f(x), it remains to be shown that the CTI for f(x) can be related to the confidence level of the individual TI's. One observes that the result of substituting several RV's in an interpolation formula such as (5.8) is that a linear combination of RV's is formed. These RV's are not, in general, statistically independent. However, it will be assumed that all the x_i at which deterministic data are replaced by a TI will be far enough apart from each other and from the other deterministic data points so that the RV's f_i at these x_i can be considered independent.

The problem of relating the confidence level of the CTI to the confidence level of the individual TI's is seen to coincide with the problem studied in the previous chapter. The TI for y(x) given in (5.19) and (5.20) is the TI formed by the bounding algorithm (BA) of the last chapter. If each of the TI's of the f_i in (5.17) corresponds to a common α as in (5.18), then the condition that y(x) is bound by D_1 and D_2 with probability of at least α is the condition of linear conformity discussed in the last chapter.

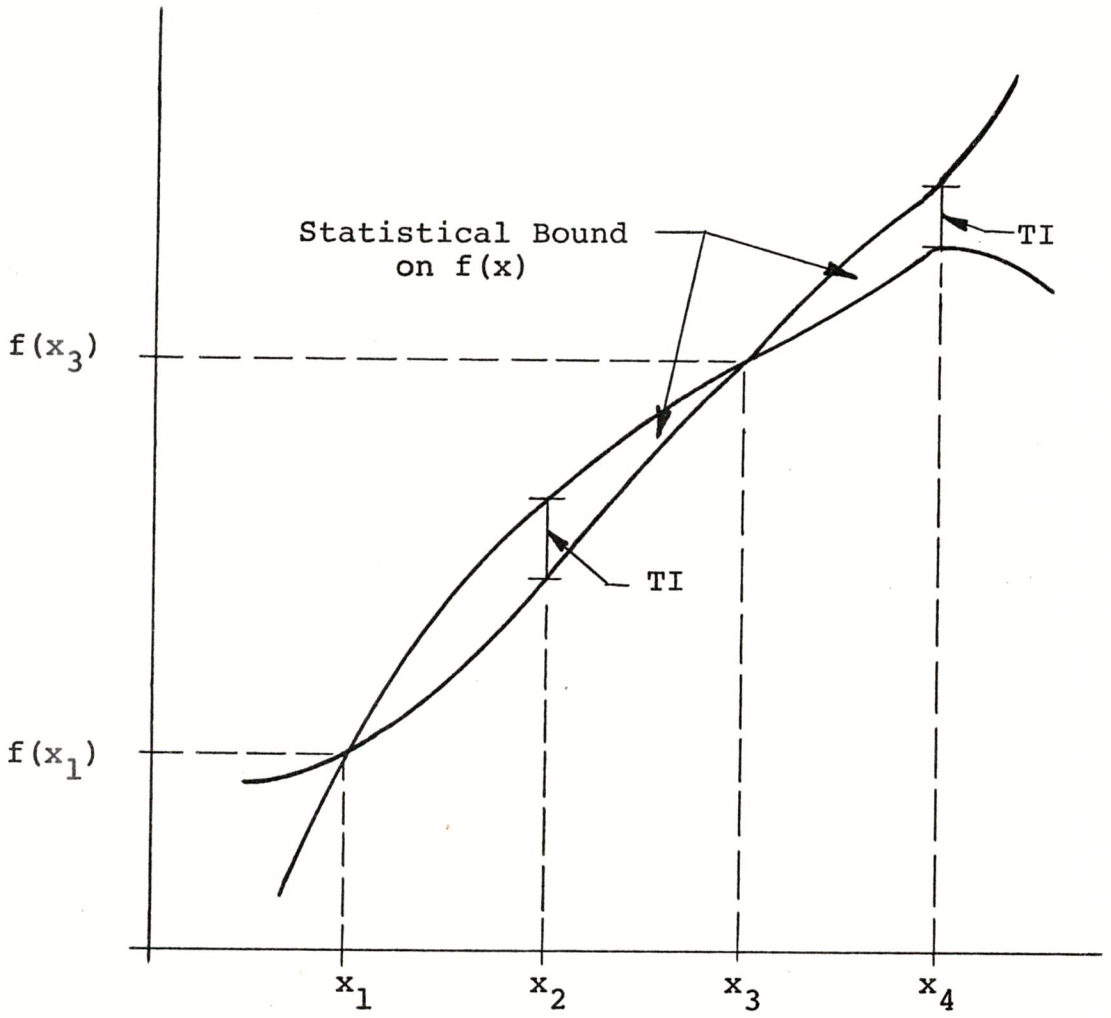

Figure 5.2 Multiple Tolerance Interval Interpolation. At x_1 and x_3, deterministic data has been used. Tolerance intervals have been substituted for deterministic data at x_2 and x_4.

That is, the problem considered here is that of determining if the collection of pdf's, representing the f_i's, are linearly conformal.

While only the TI's for the individual RV's are required to form a bound as in (5.19) and (5.20), the determination of whether the collection, of pdf's corresponding to the individual RV's, is linearly conformal requires one to know the form of these pdf's. In the most general case, one must perform a weighted convolution of the pdf's to determine this condition. However, if the pdf's are of the same form (i. e. Gaussian, rectangular, etc.) the techniques of Chapter 4 may allow a simpler analysis. If the collection of the RV's corresponds to a member of the exponential transform family of Chapter 4, Theorem 4.7 can be used to determine linear conformity. If the pdf's are of the same form, although of an arbitrary functional form, then the heuristic graphical method described in the last chapter is useful, provided that the transform of the pdf form can be sketched or approximately determined. Theorem 4.9 can be used when a subset of pdf's are self lc. As an example, suppose the RV's f_i,i=1,2,3 have pdf $g_1(u)$ and RV's f_i,i=4,5,6 have pdf $g_2(u)$. Let g_1 and g_2 be each classified as SLC with respect to the heuristic method of the last chapter and furthermore let the collection of g_1 and g_2 be (mutually) lc. This last requirement allows one to form three pairs of RV's. Each

pair is made up of one unique member from each of the above two subsets of RV's. Each of these pairs can be represented as a new RV formed as a linear combination of the two RV's forming a pair. Furthermore, the three pdf's of the new RV's are identical and are SLC. This arises from Theorem 4.9 and from the requirement that g_1 and g_2 are SLC. Since g_1 and g_2 are mutually lc, the BA is preserved for each pair. Hence, combining the three pairs (equivalent to combining these individual RV's) having identical SLC pdf's forms a RV which obeys the BA. This last remark, of course, assumes the validity of the heuristic in determining linear conformity of g_1 and g_2.

The determination of linear conformity is important in assessing the confidence level to be associated with the CTI of (5.19) and (5.20). In applying the approach of replacing deterministic data by TI's, one may be unable to directly determine linear conformity of the associated pdf's. Such a situation could result, for example, from a difficulty in determining the pdf forms. The following conjectures have evolved from the author's experience with pdf's and their transforms, but to this time, no attempt has been made for direct verification. The conjectures are:

1. Mono-model, symmetrical pdf's appear to be always lc.
2. pdf's which are self lc appear to be lc as a collection. That is, it appears that a sufficient condition for the pdf

of a collection of RV's in a linear combination to be lc is that each of these pdf's, although differing, are by themselves self lc.

In the following section, the problem of statistical dependence between the f_i of equation (5.17) will be addressed. Practical considerations such as the placement of data points will also be discussed. A hypothetical example will be given in order to illustrate the main techniques of this chapter.

5.4 Statistical Dependence The first topic of this section is the effect of statistical dependence of the RV's, f_i. The reason statistical dependence is important is because two RV's $f(x)$ and $f(x+\Delta x)$ are obviously highly correlated when differentially close in x. However, from the assumption that $f(x)$ is continuous, it is meaningless to consider any randomness between two values of $f(x)$ differentially close in terms of x. If the distance between these two points is increased, then it may become reasonable to make the assumption of decreased statistical correlation. Therefore, if the measured points of $f(x)$ are sufficiently far apart in x, the assumptions of statistical independence can become valid for practical purposes.

In section 3.3, statistical dependence of RV's in a linear combination was studied. Theorems 3.8 and 3.9 are of direct benefit to the present consideration of statistical

dependence between the RV's, f_i. By using the BA to combine the TI's of the f_i, one is assured through Theorem 3.8 that the CTI formed will still be conservative under monotonic dependence between the f_i. Theorem 3.9 considers the situation wherein the set of f_i can be divided into subsets such that within each subset, the f_i are linearly dependent. However f_i from different subsets are statistically independent. In this situation of statistical dependence, Theorem 3.9 assures a conservative CTI. That is, the collection of RV's is lc.

It is concluded that the assumption of statistical independence made in this chapter does not prevent the application of the chapter's techniques to the actual situation of partial dependence between the f_i. While one must guard against a technique which erroneously adds information, one must also avoid a technique which results in a significant loss of information. Such a loss of information can result from naively ignoring existing statistical dependence. In the application of statistical information to interpolation, there are a number of strategies that may be used for avoiding an unnecessary loss of information. These strategies involve intuitive judgment in the placement of data points.

5.5 Implementation Strategies When considering the location of statistical and deterministic data points, it must first be assumed that there exist a priori criteria for deciding

the acceptability of a device function.[1] That is, from the intended use of the device, one must have bounds for the device function which corresponds to proper operation of the device.

When deciding where to locate data points, one may be constrained by equispaced requirements of most cardinal interpolation formulations. The data points may otherwise be positioned in critical segments of the interpolation interval. Such critical segments can exist when the TI greatly overlaps the acceptability limits described above. (The TI's being considered relate to those formed during the initial data gathering phase and should not be confused with the TI's formed by the BA. Section 3.2 considers a practical means for determining the TI's required here.) These critical segment data points should be represented deterministically, while regions in which the TI was small and well within the acceptability limits would be most reasonably represented by statistical data. That is, a deterministic measurement made in a region where the TI extended beyond an acceptability region would represent a greater amount of information about the acceptability of a

---

[1]It must be noted that the conventional interpolation error must be added to the statistical bounds on the device function. In practice, one would incorporate this error in the acceptability bounds by subtracting and adding it to the upper and lower limits, respectively.

device, then would a measurement made where the TI for f(x) was well within the acceptability region. However, the data points must not be placed close together or clustered. Rather, the data should be evenly spaced within the overall interpolation interval, although not necessarily equispaced.

The decision on how many data points to use depends on a priori information about the device function f(x). In the practical problem of device testing, one usually knows the general form of the device characteristic. Such a function is usually an idealized approximation of the actual function exhibited by a device. For example, an amplifier may have a nominally "linear" input-output function. However real amplifiers have some non-linearity, often represented by a measure of distortion. If distortion were the critical factor in acceptability, one might use a polynomial interpolation of order greater than one. However, high polynomial functions are oscillatory. It is preferable to use a piecewise representation of the device function consisting of low order interpolation over subintervals. In any case, the problem of determining the proper order interpolation to use is a classical problem of numerical analysis, for which much theory is available (Hamming [1962] and Ralston [1965]). This problem is merely cited in passing, since the order of interpolation determines the number of data points required.

The substitution of TI's for deterministic data may be accomplished in an adaptive manner. For example, a minimum number of measurements could be made in the initial phase of the testing procedure. If the statistical bounds were found to overlap the acceptability bounds (containing the interpolation error bound), a measurement could be made at the point of maximum overlap, replacing one of the statistical data values. This process could continue until either the overlap condition was removed, a measurement was outside the acceptance region, or all the statistical data were replaced by measurements. Alternatively, a new statistical data point could be selected at the overlap region, in order to determine whether use of the TI corresponding to this point would result in avoiding the overlap.

No precise rules can be given for the exact procedure to use since such techniques are heuristic and strongly depend on the application. The preceding ideas merely indicate the kind of considerations that must be made in practice. One must use as much information as is available in designing a procedure for a particular testing problem.

When the device response function is known to be band-limited, a transcendental interpolation may be preferred to a polynomial formulation. This is especially true for response function that is approximately periodic. As pointed out previously, it is often advantageous to incorporate

the deterministic behavior of the response into the function form so that statistical data points will be nearly independent. As mentioned earlier, one may use Lagrange interpolation rather than the cardinal series formula if the device function is not only bandlimited, but also has a known finite total "energy" (Radzyner and Bason [1972]). In this case one can obtain a bound on the interpolation error in terms of the bandwidth, the energy, and the number of data points used. A disadvantage in this approach is that the constraints associated with transcendental interpolation apply. However, when these constraints are not overly burdensome, and when it is difficult to obtain an error bound for a particular order polynomial interpolation directly, it may be advisable to consider this approach.

A final consideration of implementation is that of gathering the information considered earlier as a priori. Determination of TI estimates and of the form of the f_i pdf's evidently requires a concerted data collection effort. Such data normally is available from conventional testing. That is, if conventional interpolation were being used, then this conventional device testing would require the measurement of the device function at a number of input values. By fixing the input values of such testing, and by collecting the values measured, the required statistics would eventually

evolve. One could construct frequency plots[1] to approximate a pdf and determine the form of the probability distribution. Using estimation formulas such as discussed in Chapter 3, an estimate of the TI of f_i at a particular x_i could be computed.

In order to determine the interpolation formula order, an approximate general functional form for the device is first determined. Several approaches are possible. Depending on the actual situation, a device model equation might be available. Alternatively, one might perform a special testing effort in which a number of devices are tested at intervals of the input variable much smaller than required for normal testing.

Determination of the bandwidth of a device function may be more difficult than gathering other statistical information. One approach to bandwidth measurement could be to approximate the device function for a number of particular components. Using a numerical method such as the Fast Fourier Transform (Cooley et al. [1967]), the transform of these approximated device functions could be computed. By plotting these computed transform functions, one might be able to graphically observe a cut-off frequency.

---

[1]Statistical frequency histograms (Bowker and Lieberman [1964]).

In order to integrate the ideas of this chapter, a simple comprehensive hypothetical example is offered. This example considers a device which is a temperature-to-voltage transducer. At the heart of the transducer is a quartz crystal. The crystal has a temperature coefficient which results in a change in the resonant frequency proportional to the temperature. The crystal and an associated circuit are separately "calibrated" by laser trimming. Although the device is fairly accurate due to the calibration of the two components, some nonlinearity of the device function results from mounting effects of the crystal plus parasitic temperature effects in the circuit. The device function must be known over a 50 degree range. One hundred percent inspection of the devices is desired. Although ten devices can be put in the temperature control chamber at once, it takes several seconds to reach equilibrium at any temperature setting. A minicomputer presently performs the measurements and controls the temperature chamber.

In order to carry through this example, it will be necessary to simulate a large population of devices. Assume that the actual device function for any particular device has the form

$$f(x)=[A+B\ \sin\ (cx+D)]x \tag{5.21}$$

where the independent variable, x, is the temperature in centigrade degrees. The laser calibration results in

f(0)=0 and f(50=5.0 for all devices. Assume that parameters B, C, and D are statistically independent with Gaussian pdf's having the following distribution parameters

$$\mu_B = 0.001 \quad \sigma_B = 0.0005$$

$$\mu_C = 0.1 \quad \sigma_C = 0.05$$

$$\mu_D = 0.0 \quad \sigma_D = 1.0.$$

Then paramater A is effectively adjusted by calibration such that f(50)=5.0 is satisfied. It is noted that (5.21) and the above distribution parameters are not known in practice, but are being assumed so that simulation of sampling can be accomplished.

The simulation required for this example was accomplished in two steps. The first step involved the simulation of N_S samples of the device. Each sample required the generation of parameters B, C, and D for the simulated device. These three parameters were generated as independent, quasi-random samples from pdf's corresponding to each parameter. The method of generating a sample required a random number generator utility subprogram available on a CDC 6500 computer. This random number generator routine generates a number between 0.0 and 1.0 such that the distribution of such numbers is rectangular. In order to provide a simple method of simulating random sampling from a Gaussian distribution, an algorithm based on the Central Limit Theorem was used. According to this algorithm, a sample of n values from the

above random number generator will generate a Gaussian pdf sample corresponding to a distribution with mean and variance (Hamming [1962])

$$\mu=n/2$$

$$\sigma^2=n/12.$$

According to Hamming [1962], "a very good approximation" results for n>12. The value of n used was 48. Therefore, the random sample of 48 random numbers had a mean of 24 and a variance of 4. In order to compute a corresponding sample of a parameter distribution, the simulated sample was corrected using the parameter pdf mean and variance. For example for B, it is required that for the sum of 48 random numbers, s

$$\frac{(s-24)^2}{4} = \frac{(s_B-\mu_B)^2}{\sigma_B^2}.$$

That is, the simulated sample value for parameter B, s_B

$$s_B=\mu_B + \sigma_B\ (s-24)/2. \qquad (5.22)$$

Once the samples of B, C, and D were simulated for a device, and A had been computed according to f(50)=5.0, the function values at x_i=5, 10,...,40, 45 were evaluated from (5.21). At each of these x_i, a TI was computed corresponding to α=.90 from the N_S samples. The method of this TI estimation was that of (3.31) repeated below

$$\beta = 1-\exp[-N_S(1-\alpha)]\cdot \sum_{i=0}^{2M_S-1} [N_S(1-\alpha)]^i/i! \qquad (5.23)$$

where β is the probability that the M_S smallest and the M_S largest sample contains at least $100(\alpha)\%$ of the distribution, for N_S samples. For N_S=200 and β = .80 one has M_S = 9. The results of the simulation is given in Table 5.3. The computations of the simulation were programmed, and run on a CDC 6500 computer.

A plot of the pdf of the simulated RV f(25) is shown in Figure 5.4. This plot was obtained as a histogram of 4000 simulated values. The histogram was computed during the simulation. Note that this plot resembles a Laplace density function in shape.

Figure 5.5 is a plot[1] of three typical device functions resulting from the simulated sampling. Equation (5.21) was used to compute these three functions.

The second step of the example problem concerns the actual implementation. To perform this second simulation, it is first required that acceptability be defined and that the actual interpolation scheme be decided on. Assume that the manufacturer of the devices will sell two grades of the devices. Devices which have a maximum relative error from the ideal curve of f(x)=.1x of 3% will be marketed as having such a guaranteed specification. Devices with greater

---

[1]In order to accentuate the nonlinear feature of the device function curves, the deviation from the ideal response curve was plotted.

	TOLERANCE LIMITS	
x	a(x)	b(x)
5.000000	.4872611	.5152810
10.00000	.9770230	1.030502
15.00000	1.467573	1.546183
20.00000	1.948372	2.063373
25.00000	2.426035	2.576323
30.00000	2.906947	3.085482
35.00000	3.405550	3.585149
40.00000	3.918767	4.079551
45.00000	4.449593	4.547508

Table 5.3 TOLERANCE INTERVALS ESTIMATES FOR EXAMPLE PROBLEM CORRESPONDING TO α=.90.

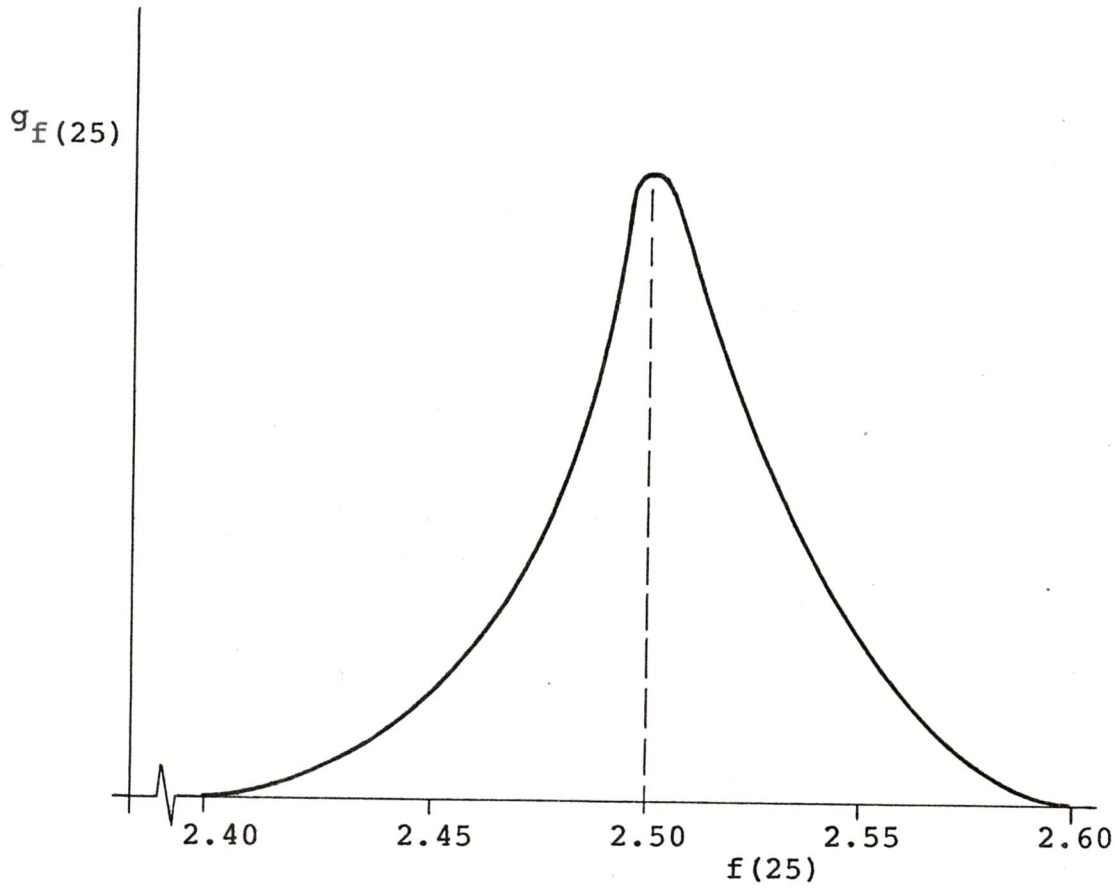

Figure 5.4 Approximation of the probability density function $g_{f(25)}$ for f(25) of example problem. 4000 simulated device samples were used to construct a histogram of f(25). Note that the curve is approximately symmetric and resembles a Laplace density function in its shape.

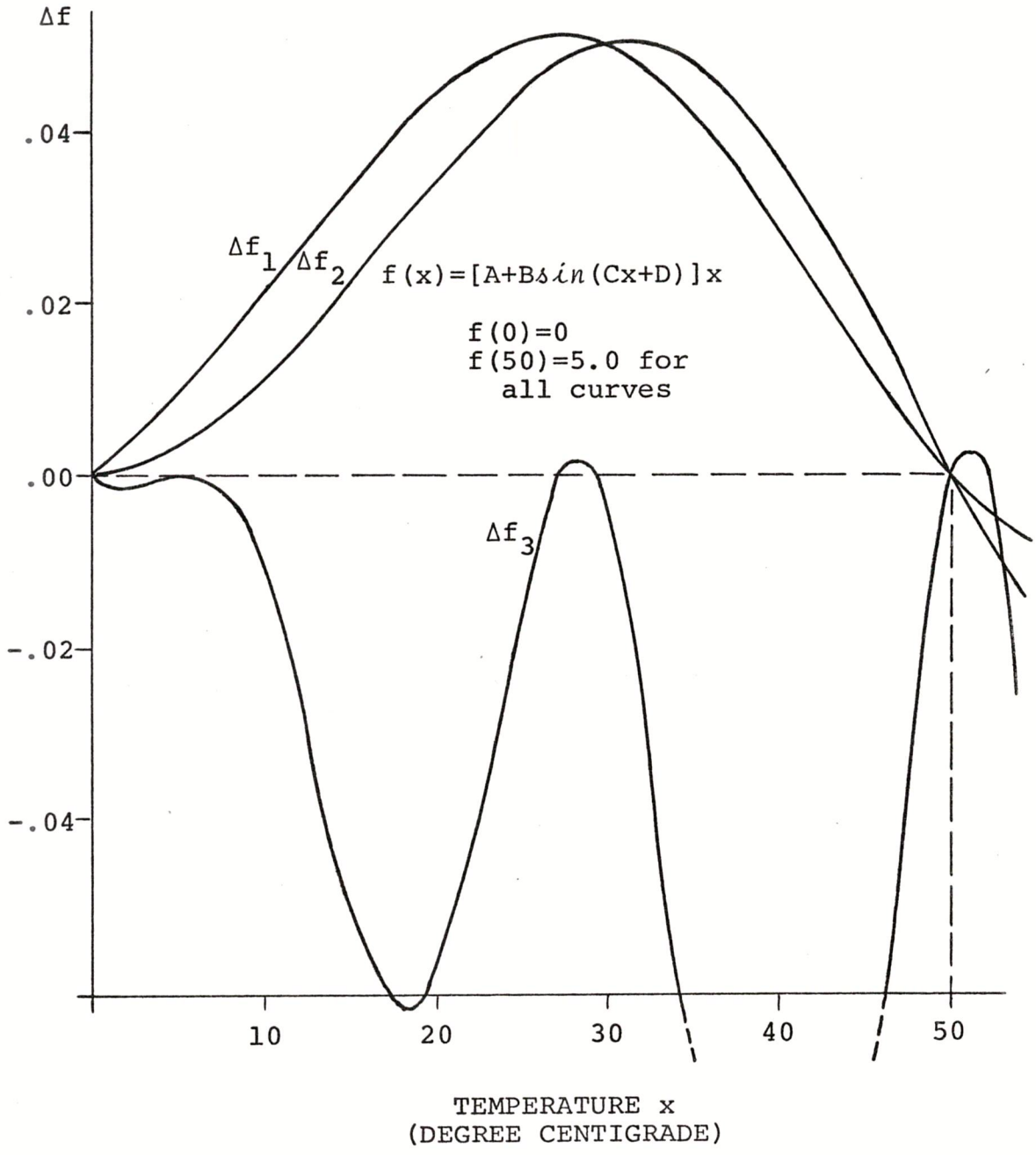

Figure 5.5 Plot of deviation ($\Delta f_i = f_i - .1x$) from the ideal response function ($f(x) = .1x$) by three simulated sample devices. Note that f_3 varies most widely.

nonlinearity will be marketed as having a guaranteed maximum nonlinearity of 5%. To set the acceptability bounds, an estimate of expected interpolation error must be approximated. Assume that a second order polynomial approximation is to be used over 10 degree intervals and that the three data points are to be fixed at the interval ends and at the half-way point. For example, in the interval [0,10], measurements are normally made at x=0,5,10. Assume from practical experience that a linearly growing interpolation error bound E(x) is known to exist

$$E(x) = \pm E_3(28) \cdot \frac{x}{28} = 2.3E\text{-}4 \cdot x \tag{5.24}$$

or as a relative error, E_R

$$E_R = 2.3E\text{-}3. \tag{5.25}$$

Since the acceptability bound given above and the interpolation error are of the same form, it is a simple matter to combine the two. Assuming statistical independence between the two bounds, the decision rule for grade classification is:[1]

$$\text{If } \frac{|y(x)-.1\cdot x|}{.1\cdot x} \leq .0277 \text{ , 3.0\% guaranteed maximum nonlinearity}$$
$$\frac{|y(x)-.1\cdot x|}{.1\cdot x} \geq .0277 \text{ , 5.0\% guaranteed maximum nonlinearity} \tag{5.26}$$

[1]It is assumed that catastrophic failures are detected during calibration, so that only working devices are considered in this example.

The first classification will be denoted as "good", the second as "bad".

If all deterministic data is used to determine the device characteristic, nine additional measurements are required in addition to the calibration data at x=0,50. To demonstrate the advantage of the statistical approach of this chapter simply, the middle data point of each interpolation subinterval will be replaced by the corresponding TI from Table 5.3. Thus, only four deterministic measurements are required (x=10,20,30,40) representing a 55.6% reduction in additional deterministic measurements.

In order to evaluate the performance of the statistical approach, a simulation of production testing was programmed on a CDC 6500 computer. As indicated earlier, lots of ten devices are tested as a group. Devices were simulated as before. As a means of evaluation, both the actual and interpolated device functions were computed at 500 points within (0,50). Measurements were simulated by computing the actual device function value at the appropriate x. A total of 100 runs were simulated (1000 devices). The result for each device was classified in one of five possible categories:

1. Device good and was accepted[1]

---

[1]Accepted means interpolated function met first condition of (5.26).

2. Device good and wasn't accepted

3. Device bad and was accepted

4. Device bad and wasn't accepted

5. Device bad from deterministic data

The results are

Category	# of Devices in Category	% of Total Devices in Category
1	922	92.4
2	25	2.5
3	6	0.6
4	4	0.4
5	43	4.3

One can see by the above results that even for the simple fixed testing procedure used in the example, a significant reduction in the required number of tests was achieved. Assuming that the computations required could be performed within the time interval required to effect a temperature change within the testing fixture (stated earlier as several seconds), throughput would have been doubled by the utilization of the TI's. An adaptive procedure resulting in additional testing productivity could have been achieved. For example, only one of the three data could have been deterministic for each interpolation subinterval. If more than, say, three devices were outside the decision bounds of (5.28) according to the interpolation function,

one of the two statistical data points would be replaced by an additional measurement. As few as two measurements in addition to calibration measurements would be required per device. That is, as few as two additional temperatures would have to be generated for a particular lot.

In conclusion, the incorporation of statistical information in exact-fit interpolation can easily result in a significant reduction of the number of deterministic measurements required to characterize a device. The statistical information required for this method is that of TI's for $f(x_i)$ at a number of x_i. Gathering such statistics may be accomplished in many cases by initially measuring $f(x_i)$ for a large number of devices at the x_i normally required by conventional interpolation.

In the next chapter, statistic information representing previous measurements of devices will again be used to reduce the number of measurements required to characterize a device in terms of a response function. However, the incorporation of this prior information will be used to adapt an extrapolation formula to the previous measurement information. In particular, the extrapolation will be iterative.

6. ADAPTIVE EXTRAPOLATION FOR LONG TERM PREDICTION

6.1 Introduction In this chapter, the extrapolation of a device function is considered. This extrapolation is considered "long term" in that the function behavior must be predicted for values of an independent variable greatly distant from the value of this variable corresponding to data. Usually the independent variable is time and the data are values of the device function at regular intervals of time. That is, the problem is to extrapolate the functional in time. The interval of extrapolation is usually several orders of magnitude greater than the incremental spacing of the data. Other problems of long term extrapolation concern the prediction of a device function response to a physical input variable. In this second general case, the reason for extrapolation is that either one does not wish to subject the device to the extrapolated input value, or else cannot generate this input value conveniently.

Before considering the chapter topic further, one may relate the problem of extrapolation to the problems considered in previous chapters. It is generally true that polynomial extrapolation of the type considered in these chapters does not perform well (Hamming [1962]). That is, the error of the extrapolated value increases rapidly with the length of the extrapolation interval. Transcendental interpolation does not normally allow an extension to extrapolation. Therefore, the previously developed methods are not

usually practical in extrapolating measurements over great intervals. Conversely, the techniques to be discussed will be iterative and are therefore not generally applicable to interpolation. It is concluded that the methods of statistical interpolation, previously discussed, complement the topic studied in this chapter. That is, the two topics are applicable to mutually exclusive problem areas. Additionally, the information required for the long term extrapolation approach to be developed need be less organized than the statistic information requirements of statistical interpolation.

In order to indicate to the reader the motive for the approach taken in this chapter, it is useful to briefly consider the intended area of application for this approach. As indicated in Chapter 2, a problem of growing concern for manufacturers and users of certain devices is that of predicting the life of, or the storage effects on such devices. In the case of many electronic devices such as monolithic circuits, life testing may be similar to storage testing in that time, rather than the number of cycles of operation (oscillations or state changes), has a predominate effect on the mean time before failure (MTBF). Hence, long term storage effects can be of interest to those examining operating lifetimes.

In either of these two testing problems, one is faced with the more general problem area of this thesis. That is,

one wishes to minimize the amount of information required to characterize a device function. However, the need to limit this information differs from the rational of the previous chapters. The present need stems from the fact that the additional information is either unavailable or is prohibitively costly. Additional testing corresponds to longer periods of testing time. Increments of time are often weeks, months, or even years, since the device characteristic changes slowly with time. The period of testing represents a cost.

As previously indicated, situations arise in which the independent variable, used during testing a device, must be limited in range. As an example, electronic detonators can be subject to only a limited current without suffering an irreversible change. Yet, one is primarily interested in predicting the behavior of such devices at destructive operating currents.

It seems reasonable, that in order to reduce the amount of data (information) required to characterize a device performance through extrapolation of the data, one should try to incorporate information resulting from testing other similar devices. Rather than replacing deterministic data with statistical data as in previous chapters, the approach will be to adapt a general extrapolation formula to the statistical data. This approach will be called "adaptive extrapolation" in this chapter. (One should note that the

use of the word "adaptive" as used here refers to the process by which the extrapolator is trained by the statistical information.) The requirements of the proposed approach are that it utilizes the statistical data represented by measurements made on previously tested devices, and it performs at least "as well" as a conventional extrapolation formula when both methods use the same data for the "tested" device. Both these requirements will be stated more precisely later.

In the next section, the basic approach of adaptive extrapolation will be developed. This will be accomplished in part through the application of pattern recognition theory. The statistical information represented by previous device tests will be organized as a set of training patterns. This set of patterns will then be used to "train" the extrapolator. The performance of the trained extrapolator will be compared to that of a conventional extrapolation formula for the case when the training set is generated from a polynomial function. The development will be primarily concerned with extrapolation of a function of a single variable. However, the extrapolation of a multivariant function will be seen to follow simply from the basic development

Following the basic development, the condition and meaning of inconsistent data will be discussed. As will be explained, inconsistent data arises in most practical situations. The training algorithm will be modified in

order to handle such data. A measure of inconsistency will result which will be indicative of the adequacy of the extrapolator.

Following the discussion of inconsistency, several forms of measurement error will be considered. The performance of the adaptive extrapolator with respect to such an error will be analyzed.

Finally, implementation of the adaptive extrapolation technique to long term storage and life testing will be described. Here, the extrapolator will be used to simulate the functioning of a device. By additionally using a pattern recognizer to determine when a condition of failure has been reached, it will be possible to implement a system to perform acceptable testing of devices.

The reader, acquainted with pattern recognition theory, may have observed that the preceding information has avoided the use of clustering for the classification of patterns (Sebestyn [1962]). Clustering is inappropriate for the work since it relates to patterns defined by sets of inequalities. Whereas in general pattern recognition theory, one attempts to pass a hypersurface through a set of points, finding an extrapolation formula represents the determination of a hypersurface on which the set of points lie. This distinction arises from the usual pattern recognition problem being concerned with inequalities, while the extrapolation problem is instead concerned with equalities. Pattern

recognition theory is used in this chapter only to the extent that the mechanics of iterative training procedures can be applied to the problem of finding an optimum extrapolation formula.

6.2 Development of Adaptive Extrapolation Approach In this section, the nature of the extrapolation problem considered in the chapter is defined. Assumptions concerning this problem are made with respect to the functions being extrapolated and the form of extrapolation formula being used. Conventional polynomial extrapolation is reviewed by means of linear algebra analysis. By using linear algebra, one is able to discuss the existence and uniqueness of exact extrapolations of polynomial functions. Following this analysis, an iterative[1] training procedure is given which forms the basis of the adaptive extrapolation approach. This procedure is used to obtain a particular extrapolation formula from groups of measurements made of a device function. Convergence of the training procedure is proven for certain situations.

---

[1]Ralston [1965] used the coordinate "iterative" interpolation (extrapolation) in an entirely different sense than that used in this chapter. His method allows the order of a conventional approximation to be increased and the additional weights to be determined directly from the previous weights. These preceding weights are determined by conventional means.

A fundamental initial assumption, relieved later, is that the function to be extrapolated is an Nth order polynomial

$$f(x) = \sum_{i=0}^{N} c_i \cdot x^i. \tag{6.1}$$

Throughout this chapter, f(x) and the extrapolation of f(x) will be considered at equally spaced intervals of the independent variable x; that is at x_j, such that

$$x_{j+1} - x_j = \Delta x \tag{6.2}$$

for all j. Then for some origin[1] of x, x_0, one can write

$$x_j = x_0 + j \cdot \Delta x. \tag{6.3}$$

The function f(x) is assumed to correspond to the response of a particular device to the input x. The device is assumed to belong to a large collection of "similar" devices. By the word similar, it is meant that other devices have a response f(x) such that the device functions are of some general form.

The extrapolation problem is that of determining the value $f(x_{k+1})$ as a function of the (k+1) previous values of f(x), for a particular device. That is, one must determine the extrapolation function

$$y(x_{k+1}) = H\,[f(x_0),\ f(x_1),\ldots f(x_k)] \tag{6.4}$$

which satisfies some measure of accuracy in its computation

---

[1]The variable x corresponds to time in many typical extrapolation problems.

of $f(x_{k+1})$. The class of extrapolation formulas primarily considered in this chapter will be limited to stationary linear relationships of the data, $f(x_j)$. That is, (6.4) is limited to the functions

$$y(x_{k+1}) = \sum_{j=0}^{k} w_j \cdot f(x_j). \tag{6.5}$$

Consideration of non-linear stationary functions of the data will be postponed until the last section of the chapter. The restriction of linear extrapolators is made in this chapter in order to simplify the chapter developments. However, it will be shown later that extrapolators which combine data as products can be handled by the techniques that are to follow. Another restriction implied by (6.4) and (6.5) is that the extrapolation formula must be a scalar, single variable function. The relaxation of this restriction will be considered in the last two sections of this chapter.

A conventional formulation of extrapolation of the general form given in (6.5) is that of the Lagrange form, introduced in an earlier chapter. For equi-spaced data, one may substitute (6.3) in the Lagrange form given in (5.2) and (5.3). Then

$$y(x_{k+1}) = \sum_{j=0}^{k} L_j(x_{k+1}) \cdot f(x_j) \tag{6.6}$$

where

$$L_j(x_{k+1}) = \frac{\prod_{\substack{s=0 \\ s\neq j}}^{k} [(k+1)\Delta x - s\cdot\Delta x]}{\prod_{\substack{s=0 \\ s\neq j}}^{k} (j\Delta x - s\cdot\Delta x)} \tag{6.7}$$

$$= \frac{\prod_{\substack{s=0 \\ s\neq j}}^{k} (k+1-s)}{\prod_{\substack{s=0 \\ s\neq j}}^{k} (j-s)}$$

$$L_j(x_{k+1}) = \frac{(k+1)!(-1)^{k-j}}{(k+1-j)(j!)(k-j)!} = \binom{k+1}{j}(-1)^{k-j} \tag{6.8}$$

where the right-hand side of (6.8) is a binomial coefficient. Note that the coefficients are independent of the spacing Δx.

The extrapolation formula of (6.6) and (6.7) is a k^{th} order formula and computes $f(x_{k+1})$ exactly if $f(x)$ is a polynomial of order k or less (Ralston [1965]). It is clear that when $f(x)$ is a polynomial of order k or less, (6.6) and (6.7) may be employed as a simple means of computing $f(x_{k+1})$. However, physical functions encountered in practice are often not polynomial. Advantages of using other than conventional polynomial formulas for non-polynomial extrapolation will be discussed later. The adaptive procedure of this chapter used to generate non-conventional extrapolation formulas will be discussed in the remainder of this section. This is done to demonstrate its

performance when f(x) is a polynomial function of degree N, as given in equation (6.1). Since the adaptive procedure is iterative, it is also necessary to discuss convergence of the method. Convergence will result in a solution for the coefficients, w_i in (6.5). Before discussing the adaptive procedure, it is necessary to study the possible solutions of (6.5) which result in exact extrapolation of the polynomial function given in (6.1). This study is required for determining convergence of the adaptive procedure.

Suppose that at k+2 consecutive intervals of x, the values of f(x) are determined. That is, for f(x) a device response function, measurements are made of f(x) at $x=x_0$, $x_1,\ldots,x_k,x_{k+1}$. Let the preceding sequence of k+2 measurements be denoted an experiment. Denote the k+1 sequence of the first k+1 measurements

$$\underline{f}_1 = (f_{1,0}\ f_{1,1}\cdots f_{1,k}) \tag{6.9}$$

where the first subscript corresponds to the experiment number and the second subscript corresponds to the measurement of f(x) at x_j. The vector

$$\underline{f}_1^a = (f_{1,0}\ f_{1,1}\cdots f_{1,k}\ f_{1,k+1}) \tag{6.10}$$

will be called an "augmented" vector.

Let m additional experiments be performed by measuring the same device at intervals starting at different initial x_0's and/or by measuring other device functions with the same order polynomial function f(x), but with allowably different coefficients c_i [see equation (6.1)]. Using the

notation of (6.5) and (6.9), one wishes to obtain the extrapolation of f(x), such that for all r

$$f_{r,k+1} = \sum_{j=0}^{k} w_j \cdot f_{r,j} = \underline{f}_r \cdot \underline{W}. \tag{6.11}$$

Note that the right-hand side of (6.11) is a dot product of two vectors, where the column vector $\underline{W}$ is

$$\underline{W} = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_k \end{bmatrix}. \tag{6.12}$$

Then for the m+1 experiments, one may write the vector equation

$$\begin{bmatrix} f_{1,k+1} \\ f_{2,k+1} \\ \vdots \\ f_{m+1,k+1} \end{bmatrix} = \begin{bmatrix} f_{1,0} f_{1,1} \cdots f_{1,k} \\ f_{2,0} \quad \vdots \\ \vdots \quad \vdots \\ f_{m+1,0} \quad f_{m+1,k} \end{bmatrix} \cdot \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_k \end{bmatrix} \tag{6.13}$$

or in matrix notation (Nering [1967])

$$\underline{F} = \underset{\sim}{f} \cdot \underline{W} \tag{6.14}$$

where $\underline{F}$ is the column vector of k+1 values

$$\underline{F} = \begin{bmatrix} f_{1,k+1} \\ f_{2,k+1} \\ \vdots \\ f_{m+1,k+1} \end{bmatrix}$$

In discussing the possible vectors $\underline{W}$ which satisfy (6.14), several concepts and terms of linear algebra theory will be used. It is assumed that the reader is familiar with the basic theory of vector spaces, dimensionality of vector spaces, the meaning of a vector space "basis", and linear mappings of vector spaces. Texts on linear algebra and matrix theory such as Nering [1967] contain the definitions and theory used in the subsequent discussion.

The vectors obtained from the above mentioned experiments may be envisioned as belonging to a larger set of vectors. It is essential to show that for a particular N corresponding to (6.1) and a particular Δx corresponding to (6.2), the set of all possible vectors $\underline{f}_r$ is a vector space. A second concern related to the set of the experiment vectors is to determine the dimension of this vector space in terms of the integer N of (6.1). It will then be possible to discuss the set of solutions $\underline{W}$ satisfying (6.13). Two theorems will be proven in connection with the above two conjectures about the set of experiment vectors.

Theorem 6.1 Let the general function f(x) be defined by equation (6.1) for a particular non-negative integer N. Let the positive constant Δx be given. Define the (k+1)-tuple vector

$$\underline{f}_r = (f_{r,0}, f_{r,1}, \ldots, f_{r,k}) \tag{6.15}$$

such that

$$f_{r,p} \equiv f_r(x_p) = \sum_{i=0}^{N} c_{i,r} \cdot x_p^{\,i} \tag{6.16}$$

where

$$x_p \equiv x_{0,r} + p \cdot \Delta x. \tag{6.17}$$

Denote the set of all such vectors, arising from allowably different constants $x_{0,r}$ and $c_{i,r}$, i=0,1,...,N, as the set $\{\underline{f}_r\}$.

If the set $\{\underline{f}_r\}$ is defined over the set of real $x_{0,r}$ and the set of real (N+1)-tuples $\{(c_{0,r}, c_{1,r}, \ldots, c_{N,r})\}$, then $\{\underline{f}_r\}$ is a vector space.

Proof To show that the set $\{\underline{f}_r\}$ is a vector space, it is sufficient to show that for two arbitrary vectors $\underline{f}_1$ and $\underline{f}_2$ in the set, and for any two arbitrary constants γ_1, and γ_2, the vector sum

$$\gamma_1 \cdot \underline{f}_1 + \gamma_2 \underline{f}_2$$

is also in the set (Nering [1967]). Using (6.16) and (6.17), one may write

$$\gamma_1 \cdot f_{1,p} = \sum_{i=0}^{N} \gamma_1 c_{i,1} (x_{0,1} + p \cdot \Delta x)^i. \tag{6.18}$$

Observe that one may expand the binomial in (6.18) and collect terms with common powers of $(p\Delta x)$. The highest power of $(p\Delta x)$ will be N. Therefore (6.18) can be written as

$$\gamma_1 \cdot f_{1,p} = \sum_{i=0}^{N} \hat{c}_{i,1} (p\Delta x)^i. \tag{6.19}$$

Equation (6.19) may be viewed as a transformation of coefficients from the triple $(\gamma_1, c_{i,1}, x_{0,1})$ to scalar $\hat{c}_{i,1}$. A similar transformation could be performed for $\gamma_i \cdot f_{2,p}$, so that one can write

$$\gamma_2 \cdot f_{2,p} = \sum_{i=0}^{N} \hat{c}_{i,2} (p\Delta x)^i. \tag{6.20}$$

To complete the proof, one must show that the vector

$$\underline{f}_3 = \gamma_1 \cdot \underline{f}_1 + \gamma_2 \cdot \underline{f}_2 \tag{6.21}$$

belongs to $\{\underline{f}_r\}$. Using (6.19) and (6.20), and the notation of (6.15),

$$f_{3,p} = \sum_{i=0}^{N} (\hat{c}_{i,1} + \hat{c}_{i,2})(p\Delta x)^i. \tag{6.22}$$

One observes that (6.22) requires

$$f_0(x) = \sum_{i=0}^{N} (\hat{c}_{i,1} + \hat{c}_{i,2}) x^i = \sum_{i=0}^{N} c_3 \, x^i.$$

Clearly $\underline{f}_3$ belongs to $\{\underline{f}_r\}$. Therefore, it is proved that $\{f_r\}$ is a vector space.

QED

One can see that the set of augmented vectors, $\{\underline{f}_r^a\}$, is also a vector space. Having determined that $\{\underline{f}_r\}$ is a vector space, one may now consider the relation between the dimension of the vector space, and the integers N of (6.1) and the number of components of the vectors $\underline{f}_r$, k+1. The following theorem will be useful in subsequent discussions of the existence and uniqueness of solutions to (6.14).

Theorem 6.2 Let the vector space $\{\underline{f}_r\}$ be defined as in Theorem 6.1, corresponding to the constants N, Δx, and k. Then the dimension (Nering [1967]) of $\{\underline{f}_r\}$ is

$$\dim \{\underline{f}_r\} = \begin{cases} k+1 & , \ k \leq N \\ N+1 & , \ k \geq N \end{cases} \tag{6.23}$$

Proof Let the vectors $\underline{f}_1, \underline{f}_2, \ldots, \underline{f}_s$ be a basis[1] for $\{\underline{f}_r\}$. That is, the s vectors are linearly independent, where s is the dimension of $\{\underline{f}_r\}$. Consider a matrix with each row being a vector in the basis. For k>N, one may use equation (6.8) to write the $f_{r,N+1}$ component of a vector in the basis as a linear combination of the preceding components, that is, $f_{r,0}, \ldots, f_{r,N}$ for any basis vector. Therefore, the entire (N+1) column is linearly dependent on the first (N+1) columns. Each column to the right of the N+1 column

[1] See Nering [1967] for definitions of "basis", dimension, and linear independence.

is similarly dependent. Therefore, the maximum number of linearly independent columns in the matrix of the basis vectors is N+1.

For the $k \leq N$ case, all (k+1) columns are linearly independent. One may also conclude that the maximum number of linearly independent rows is also N+1 for k>N case, and (k+1) for the $k \leq N$ case (Nering [1967]).

To complete the proof, it must be shown that s, the number of vectors (rows) in the basis matrix, is equal to the column rank ζ. First note that s cannot exceed ζ, from the previous discussion. To show that s cannot be less than ζ, an argument by contradiction is given.

For s less than ζ, an additional row is added to the matrix which is independent of the first s rows. Such a row can be obtained by row echelon operations (Nering [1967]). By the assumption that the first s rows was a basis for $\{\underline{f}_r\}$, the additional row cannot belong to $\{\underline{f}_r\}$. However, corresponding to the first ζ components of $\underline{f}_{s+1}$, the new row, a polynomial of degree $(\zeta-1)$ can be generated. Specifically, this polynomial is formed using Lagrangian interpolation to interpolate the first ζ values of the vector at the corresponding values of x. Therefore, the vector $\underline{f}_{s+1}$ belongs to $\{\underline{f}_r\}$, and a contradiction exists. That is, the first s vectors could not have been a basis. Therefore (6.23) is proved.

QED

The theorem just proved can be used to determine the existence of possible solution to (6.13). Assume that $k \geq N$, where k+1 regularly spaced values of f(x) are to be used to extrapolate polynomial functions of the form given in (6.1). Then from Theorem 6.2, the matrix $\underset{\sim}{f}$ will have rank less than or equal to N+1. Now consider the augmented matrix composed of the augmented vectors $\underline{f}_r^a$ defined in (6.10), and corresponding to the same m+1 experiments. Adding an additional component to each row vector in the previous matrix does not increase the rank of the matrix, since the new component is linearly dependent on the previous components of the row. Since the augmented matrix has the same rank as the matrix $\underset{\sim}{f}$, solutions to (6.13) exist. By the assumption that $k \geq N$, the dimension of $\{\underline{f}_r\}$ is N+1. It is therefore possible to obtain N+1 linearly independent vectors $\underline{f}_r$, representing what may be called "independent experiments". One could, in practice, continue gathering such experiment vectors until N+1 independent vectors were obtained. Then assuming that N+1 independent experiments are represented by equation (6.13), the dimension of the solution space for vector $\underline{W}$ is (k-N) (Nering [1967]). For N=k, the solution space has dimension zero, and therefore a unique non-trivial solution exists. Since for N=k, the Lagrangian extrapolation for equi-spaced data [given in

(6.7) and (6.8)] applies, the solution of (6.13) for this case must be identical to the Lagrangian form. Thus, the following theorem has been proved:

Theorem 6.3 Let the vector space $\{\underline{f}_r^a\}$ be generated from the vector space of Nth order polynomials, where the vector $\underline{f}_r^a$ is the (k+2)-tuple of (6.10). If k=N, the solution of the matrix equation of (6.13) is unique and is the vector $\underline{W}$ having components equal to the Lagrangian form of equation (6.8).

A consequence of f(x) being a polynomial function, is that a different Δx can be used for each $\underline{f}_r$. To prove this claim, one need only observe that

$$f(x_p) = \sum_{i=0}^{N} c_i \cdot (\Delta x \cdot p)^i = \sum_{i=0}^{N} c_i (\Delta x)^i \cdot p^i$$

$$= \sum_{i=0}^{N} \hat{c}_i \cdot p^i. \qquad (6.24)$$

Therefore, any vector $\underline{f}_r$ may be normalized with respect to Δx by a manipulation similar to (6.24). That is, the variable x may be "scaled".

If the vector space $\{\underline{f}_r\}$ is of lower dimension than N+1, no solution to (6.13) will exist, as has been indicated above. The rank of a matrix of augmented vectors will be greater than the same matrix with the k+2 column removed. When no solution exists, the set of vectors in the matrix may be termed "inconsistent". This term arises when one

interprets (6.13) as a set of m+1 equations. By inconsistent, one means that at least one of these equations contradicts the remaining equations. As an example, the two equations

$$\begin{aligned} w_1 + w_2 &= 0 \\ w_1 + w_2 &= 1 \end{aligned} \tag{6.25}$$

are obviously contradictory. That is, the matrix relationship

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

is inconsistent. When one must deal with non-polynomial functions, the experiment vectors will, as a rule, be inconsistent for an arbitrary k+1 component.[1] That is, the Taylor series of many functions is that of an infinite power series. Recall it is assumed in this section that the functions to be extrapolated are polynomials of known order N. The vectors $\underline{f}_r$ are of dimension equal to N+1. One may solve for a solution of (6.13) using matrix inversion under such assumptions. However, an iterative procedure will be used to obtain a solution. This procedure will be extended to the case of inconsistent vectors in the next section, in

[1]The problem of inconsistency does not arise with conventional curve fitting since conventional formulas are derived using the assumption that the function f(x) is a polynomial with the same order as that of the approximating formula.

order to handle the extrapolation of non-polynomial functions for which no solution to (6.13) exists.

In what follows, the vector $\underline{f}_r$ will now be called "training patterns" and the set of such vectors resulting from the m+1 experiments as a "training set". Furthermore, the augmented vectors $\underline{f}_r^a$ will be called "augmented training patterns" and the set of the m+1 such vectors will be called "agumented training set". It will be convenient to normalize all the augmented patterns in the augmented training set with respect to the Euclidian length of the corresponding non-augmented training pattern. That is, each component

$$f_{r,j}, \text{ for } j = 0,\ldots,k+1$$

of pattern (vector) $\underline{f}_r^a$ is divided by $\{\sum_{j=0}^{k} f_{r,j}^2\}^{\frac{1}{2}}$. It is assumed in what follows that such a normalization of each augmented pattern has previously been performed.[1]

The training set is used to obtain an extrapolation formula in a manner analogous to training a linear pattern recognizer. This set of patterns are repeatedly presented to the extrapolator. A particular pattern in the set is used to correct the current solution vector $\underline{W}$ according to an error correction algorithm. One may define a "cycle" as the correction procedure performed for each pattern of the set. The training patterns are considered in a fixed

---

[1]Note that the scaling of a vector $\underline{f}_r$ in (6.13) by a constant does not alter the validity of this equation.

sequence, and this sequence is repeated until the training is determined to be complete. (Completeness will be elaborated on shortly.) It is possible to number the patterns in increasing order, realizing that for the set of training patterns, the index is modulo m+1. Denote the index by I and the approximation to the solution vector $\underline{W}_s$ as $\underline{W}_I$. The initial vector $\underline{W}_1$ is arbitrary. At each iteration of step, a new approximation to $\underline{W}_s$ is computed as

$$\underline{W}_{I+1} = \underline{W}_I + (f_{I,k+1} - \underline{f}_I \cdot \underline{W}_I) \cdot \underline{f}_I. \tag{6.26}$$

Notice that the quantity within the parenthesis is the difference between the actual value $f_{I,k+1}$ at x_{k+1} and the extrapolated value resulting from using the current approximation $\underline{W}_I$ as the weights of the extrapolator of (6.5). Since $\underline{f}_I$ is a unit vector by the method of normalization of the patterns, the correction of the I^{th} estimate of $\underline{W}_s$ is a vector, of length equal to the error within the parenthesis, and of the same direction as the vector $\underline{f}_I$, the non-augmented pattern. One may note that

$$\begin{aligned} \underline{f}_I \cdot \underline{W}_{I+1} &= \underline{f}_I \cdot \underline{W}_I + \underline{f}_I \cdot (f_{I,k+1} - \underline{f}_I \cdot \underline{W}_I) \cdot \underline{f}_I \\ &= \underline{f}_I \cdot \underline{W}_I + (f_{I,k+1} - \underline{f}_I \cdot \underline{W}_I) \cdot (1) \\ &= f_{I,k+1}. \end{aligned}$$

Therefore, the correction of $\underline{W}_I$ results in the new approximation for $\underline{W}_s$ being an exact extrapolator for the pattern $\underline{f}_I$.

Equation (6.26) corresponds nearly to a fractional correction rule of an error correction procedure for

training a pattern recognition system (Nilsson [1965]). (The advantage of putting the extrapolation problem in the context of pattern recognition theory is that known properties of convergence of the pattern recognition training algorithms can be used.) The main variation of (6.26) from a fractional rule correction algorithm can be resolved if $\underline{W}_I$ is augmented, and the augmented patterns corresponding to (6.10) are used as the training patterns of a zero threshold recognition system (Nilsson [1965]). In this case, (6.26) could be written

$$\underline{W}^a_{I+1} = \underline{W}^a_I - (\underline{f}^a_I \cdot \underline{W}^a_I) \cdot \underline{f}^a_I. \qquad (6.27)$$

Remember that the index I is modulo (m+1) for $\underline{f}_I$. Also note that had the patterns not been normalized as previously indicated, (6.26) would require a normalization of $\underline{f}_I$ outside the parenthesis. Since one anticipates many iterations of the set of training vectors before convergence, there is a computational advantage in doing the normalization before the training is initiated.

Before proceeding, one may note that the preceding iteration approach appears to resemble another functional inversion scheme. Specifically, a comparable method is concerned with the inversion of quasi-linear functions by iteration (Stern [1965]). Denote this approach as the QFI approach. This approach has the advantage that the extrapolator can be quasi-linear, while the approach used in this chapter assumes a linear formula. As represented by Stern,

the dimensionality of the quasi-linear function must equal that of the vector $\underline{W}$. In the linear case, the quasi-linear function is the product

$$\underset{\sim}{f} \cdot \underline{W}_I$$

and the reinforcement

$$\underline{W}_{I+1} = W_I + h \cdot (\underline{F} - \underset{\sim}{f} \cdot W_I),$$

where h is a convergence coefficient. Because of the similarity between Stern's reinforcement formula and the one used in this chapter, it may be possible to extend Stern's work to allow an arbitrary number of training patterns. This and investigations of other iterative function inversion schemes is deferred as future work.

A severe deficiency involves the stability of the QFI method. The iteration algorithm requires the determination of a sufficiently small convergence coefficient h in order to guarantee convergence. Direct determination of the actual h appears to be impossible, although a pessimistic value can be found from the pattern matrix. Stern [1965] indicates that in practice, one determines an h by trial and error. One notes that the situation of an inconsistent pattern set and the situation of using too large an h may be indistinguishable in practice.

A primary concern with using the iterative training procedure is that of convergence. The proof of convergence of the proposed training algorithm is found in a paper by Agmon [1954] in which a somewhat more general problem was

considered. The problem was that of solving the system of linear, consistent inequalities

$$\lambda_i(\underline{a}) = \sum_{j=1}^{M} U_{i,j}\, a_j + V_i \geq 0 \quad i=1,\ldots,M. \tag{6.28}$$

The coefficients $U_{i,j}$ and V_j being known numerically, the specific problem was to numerically determine values of a_j that satisfy (6.28). Comparing (6.28) with (6.13), one sees that, with minor changes, the two are essentially the same provided the equality is taken in (6.29).

As mentioned above, the training sequence is actually a repetitious cycling through the training set. A logical definition of convergence would be that, for a particular training pattern, the error at each cycle

$$f_{I,N+1} - \underline{f}_I \cdot W_I$$

would decrease successively from cycle to cycle. Thus, one requires for convergence that

$$|\underline{W}_s - \underline{W}_{I+m+1}| \leq |\underline{W}_s - \underline{W}_I| \tag{6.29}$$

where m+1 is the number of training patterns in the training set. Agmon proves that (in the notation used here)

$$|\underline{W}_s - \underline{W}_{I+m+1}| \leq \theta |\underline{W}_s - \underline{W}_I|$$

where θ depends only on the $f_{j,r}$ and

$$0 < \theta < 1.$$

The following theorem is therefore proved.

Theorem 6.4 Let the m+1 set of training patterns

$$\underline{f}_r = (f_{r,0},\ldots,f_{r,k})$$

be given such that they have been normalized according to

$$\{\sum_{j=0}^{k} f_{r,j}^2\}^{\frac{1}{2}} = 1 \quad \text{for all } r.$$

For an arbitrary non-zero (k+1)-tuple $\underline{W}_1$, define the iteration

$$\underline{W}_{I+1} = \underline{W}_I + (f_{I,k+1} - \underline{f}_I \cdot \underline{W}_I) \cdot \underline{f}_I \tag{6.30}$$

where the index I is modulus m+1 for the $\underline{f}_I$. If there exists a non-trivial solution $\underline{W}_s$ such that

$$\underline{f}_r \cdot \underline{W}_s = f_{r,k+1} \quad r = 1,\ldots,m+1,$$

then the procedure of (6.30) converges absolutely to $\underline{W}_s$ according to

$$|\underline{W}_s - \underline{W}_{I+m+1}| \underline{<} \theta |\underline{W}_s - \underline{W}_I|$$

where

$$0<\theta<1.$$

The constant θ is a rate of convergence. Agmon [1954] gives a method of estimating θ from matrix of the components of the training patterns. Although this evaluation is performed simply for a linearly independent set of consistent training patterns, it does not appear to be adaptable to patterns which are inconsistent or patterns which are dependent, and therefore is not presented here. It is noted that the iterative training procedure cannot converge when the patterns are inconsistent, since in this case, no solution $\underline{W}_s$ exists. Should the rank of the matrix of

non-augmented training vectors equal N+1, where the extrapolated functions are Nth order polynomials, and k=N, the training procedure results in the same solution as (6.8) due to the solution being unique, and due to the guarantee of convergence to a solution. With k>N, many solutions exist so that convergence to a solution other than the one given in (6.8) is probable.

When the training set can be inconsistent, it is important to determine how the training procedure will perform, in order to distinguish between convergence and non-convergence of the iteration. The following theorem indicates that the sequence of weight vectors generated by the training procedure converges to a steady-state oscillation with period equal to that of the pattern sequence.

Theorem 6.5 Consider the fixed sequence of weight vectors $\underline{W}_{I+1}$ [see equation (6.30)] generated by the training algorithm of Theorem 6.4, for the m+1 training patterns of k+1 components. Let the training set be allowably inconsistent. Unless all the patterns form parallel hyperplanes, there exists a unique weight vector $\underline{W}_{s,I}$ corresponding to each training pattern $\underline{f}_I$ (where index I for $\underline{W}_{s,I+1}$ and $\underline{f}_I$, is modulus m+1) such that for I>1,

$$|\underline{W}_{s,I+1} - \underline{W}_{I+1}| = \psi_I |\underline{W}_{s,I} - \underline{W}_I| \qquad (6.31)$$

where

$$\psi_I = \frac{|\underline{f}_I \cdot \underline{f}_{I-1}|}{|\underline{f}_I||\underline{f}_{I-1}|} \leq 1. \tag{6.32}$$

According to Theorem 6.5, the sequence of weight vectors approaches a fixed periodic sequence of weights $\{W_{s,I}\}$ with the same periodicity as that of the training vectors. Note that this sequence depends on the relative ordering of the patterns, but not on the initial arbitrary weight vector $\underline{W}_1$. The reader, familiar with non-linear control theory, may draw an analogy between the stable oscillation of the $\underline{W}_{s,I}$ and the stable oscillation of a limit cycle (Schwarz and Friedland [1965]). One may similarly regard the sequence of weights $\underline{W}_I$ resulting from the training procedure as a trajectory. The terms, "limit cycle" for the steady-state weights $\underline{W}_{s,I}$, and "trajectory" for the sequence of weights $\underline{W}_I$ will be adopted.

Proof of Theorem 6.5 A geometric argument will be used. Consider the weight space. Its dimension is k+1, since the weight vectors are (k+1)-tuples. Corresponding to each pattern, one may construct a pattern hyperplane. This hyperplane is a k-dimensional surface of points $\underline{W}$ which satisfy

$$f_{I,k+1} = \underline{W} \cdot \underline{f}_I \tag{6.32a}$$

for the pattern $\underline{f}_I$. One recognizes that the above expression is the equation of a k-dimensional surface which is

orthogonal to the vector $\underline{f}_I$ (Thomas [1962]). Consider equation (6.30). The correction to weight vector $\underline{W}_I$ is made in the direction $\underline{f}_I$. Moreover, the correction results in the new vector $\underline{W}_{I+1}$, being on the $\underline{f}_I$ hyperplane. (One may verify this by forming the dot product of $\underline{W}_{I+1}$ as determined by (6.30) with $\underline{f}_I$, remembering that the normalization of $\underline{f}_I$ results in $\underline{f}_I \cdot \underline{f}_I = 1$.)

Use the term "limit cycle" to denote the steady-state sequence of repeated weight vectors, $\underline{W}_{s,I}$. The proof consists of three parts. First, a limit cycle is assumed to exist and the sequence of weight vectors $\underline{W}_I$ generated by the training algorithm of Theorem 6.4 is shown to converge absolutely to the limit cycle. The second step of the proof is to show that if a limit cycle exists, it must be unique. Finally, the existence of the limit cycle is demonstrated.

Assume that a limit cycle exists. Consider the I^{th} step of the training algorithm ($I>1$). Except for the first iteration in which the arbitrary starting vector $\underline{W}_1$ is used, the weight vector $\underline{W}_I$ is on the hyperplane corresponding to pattern $\underline{f}_{I-1}$. (Note that the index of $\underline{f}$ is modulo m+1.) The limit cycle passes through the hyperplane at the weight ector $\underline{W}_{s,I}$. At the I^{th} iteration, $\underline{W}_{I+1}$ is computed using (6.30). This new vector is on the $\underline{f}_I$ hyperplane, and the line connecting $\underline{W}_I$ and $\underline{W}_{I+1}$ has the direction of $\underline{f}_I$ [see equation (6.30)]. The limit cycle also approaches the $\underline{f}_I$ hyperplane with the vector direction of $\underline{f}_I$, and intersects

the hyperplane at $\underline{W}_{s,I+1}$. That is, the line connecting $\underline{W}_{s,I}$ and $\underline{W}_{s,I+1}$ is parallel to the line connecting $\underline{W}_I$ and $\underline{W}_{I+1}$. Let the plane containing these two parallel lines be constructed. Refer to Figure 6.1. Construct a line in the plane parallel to the line connecting $\underline{W}_{s,I+1}$ and $\underline{W}_{I+1}$ so that it also intersects $\underline{W}_{s,I}$ (line B). One observes that a triangle is formed by $\underline{W}_{s,I}$, $\underline{W}_I$, and the point A. Since the vector $(\underline{W}_{I+1} - \underline{W}_I)$ must be perpendicular to the $\underline{f}_I$ hyperplane [see (6.32a)], the angle at A is a right angle. Since the patterns $\underline{f}_I$ and $\underline{f}_{I-1}$ are perpendicular to their respective hyperplanes, the angle subtending these vectors is equal to the angle ϕ in the figure. From elementary plane trigonometry

$$\cos \phi = \frac{|\underline{f}_I \cdot \underline{f}_{I-1}|}{|\underline{f}_I||\underline{f}_{I-1}|} \leq 1.$$

But from the figure it is clear that

$$|\underline{W}_{s,I+1} - \underline{W}_{I+1}| = \cos \phi |\underline{W}_{s,I} - \underline{W}_I|.$$

Therefore (6.31) and (6.32) are proved. One concludes that if the limit cycle exists, the weight vector trajectory approaches the limit cycle trajectory absolutely, unless the patterns are all parallel. Should this degenerate case exist, the weight vector trajectory is immediately on a limit cycle after the initial training iteration. That is, the weight trajectory is along a single straight line which intersects all the pattern hyperplanes. In the degenerate case, an infinite number of limit cycles exist. The

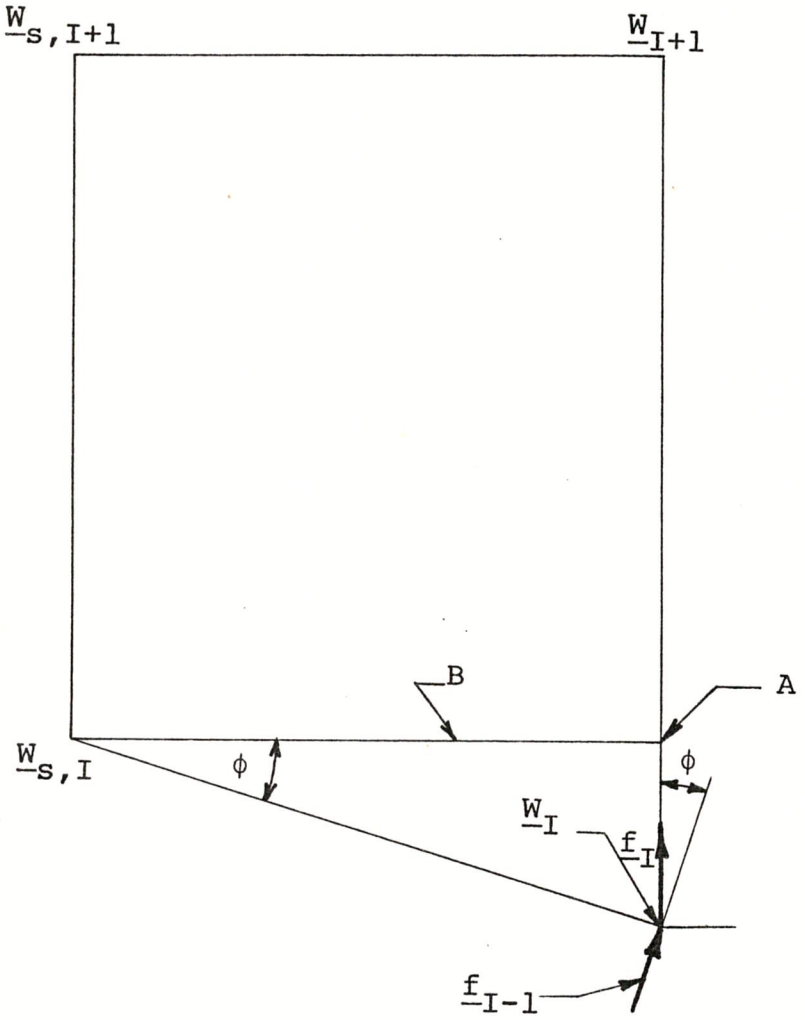

Figure 6.1 Convergence of a weight trajectory to a limit cycle trajectory. Note that the line connecting $\underline{W}_I$ and $\underline{W}_{s,I}$ is in the $\underline{f}_{I-1}$ hyperplane, and the line connecting $\underline{W}_{I+1}$ and $\underline{W}_{s,I+1}$ is in the $\underline{f}_I$ hyperplane. Also note that the directions of $\underline{f}_{I-1}$ and $\underline{f}_I$ are shown along the weight trajectory.

specific limit cycle traversed depends on the choice of the starting vector $\underline{W}_1$.

The second step of the proof is to show that should a limit cycle exist, it is unique except for the degenerate case of all parallel pattern planes. In a construction similar to that of Figure 6.1, one immediately realizes that the distance between two limit cycles decreases absolutely, at each iteration, according to

$$|\underline{W}_{s,I+1} - \underline{W}'_{s,I+1}| = \psi_I \, |\underline{W}_{s,I} - \underline{W}'_{s,I}|$$

where ψ_I is again given by (6.32), and where the prime denotes a second limit cycle. In other words, if a limit cycle exists, in the non-degenerate case, it must be unique.

Finally, the existence of a limit cycle for non-unique solutions will be proven. Consider Figure 6.2, in which two different weight trajectories are shown. From a geometric argument similar to Figure 6.1, it is clear that

$$|W_{I+1} - \underline{W}'_{I+1}| = \psi_I \, |\underline{W}_I - \underline{W}'_I|,$$

where $0 \leq \psi_I < 1$.

That is, any two trajectories must approach each other absolutely (except for the degenerate case). Then it must also be true that

$$|\underline{W}_{I+2(m+1)} - \underline{W}_{I+(m+1)}| < |\underline{W}_{I+(m+1)} - \underline{W}_I|.$$

The distance between the weight vectors for successive cycles of m+1 iterations decreases absolutely. In fact, the rate of convergence may be expressed as

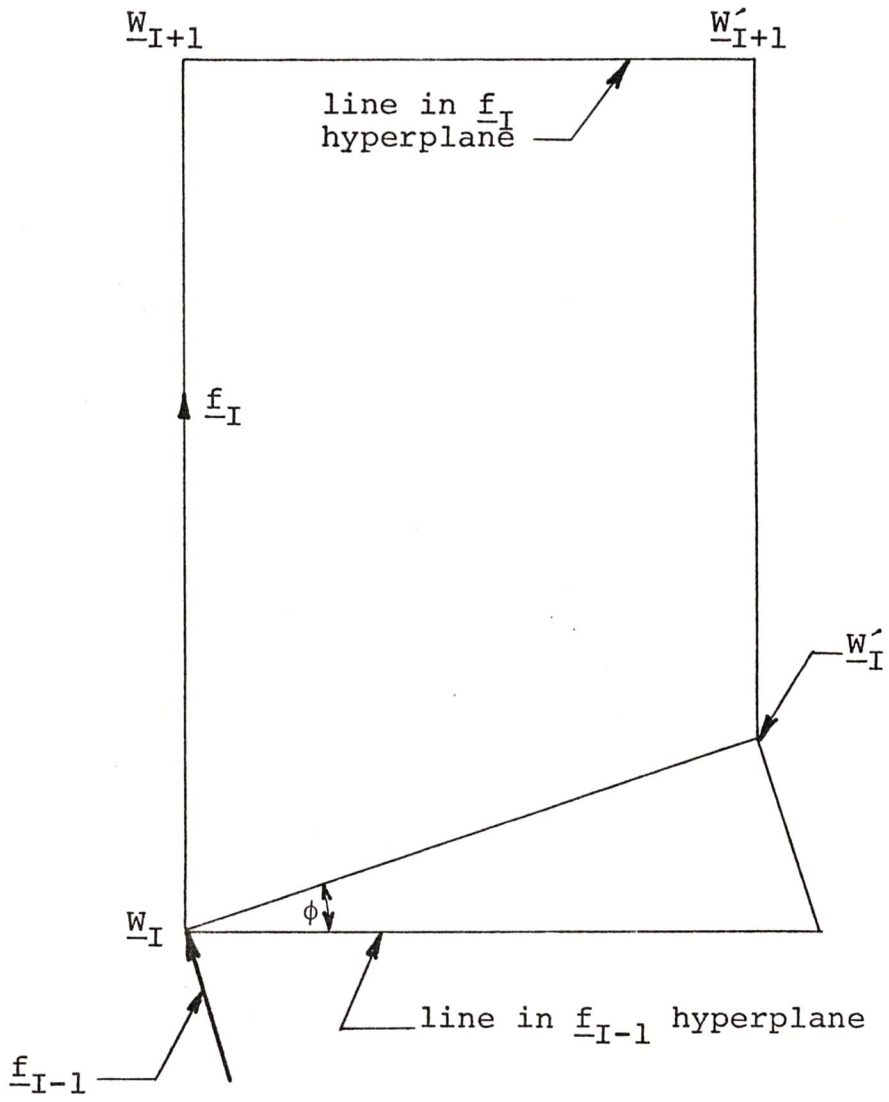

Figure 6.2 Convergence of two weight trajectories to a single trajectory. Note that the directions of $\underline{f}_{I-1}$ and $\underline{f}_I$ are shown along the $\underline{W}$ trajectory.

$$\prod_{I=1}^{m+1} \psi_I = \frac{|\underline{W}_{I+2(m+1)} - \underline{W}_{I+(m+1)}|}{|\underline{W}_{I+(m+1)} - \underline{W}_I|}. \qquad (6.33)$$

One concludes that a limit cycle must exist since eventually the distance $|\underline{W}_{I+(m+1)} - \underline{W}_I|$ approaches zero.

QED

One may observe a special case for (6.31) and (6.32). If $\underline{f}_I \cdot \underline{f}_{I-1} = 0$ for any I, the weight trajectory must immediately reach the limit cycle at the I^{th} iteration of the training procedure. One may also observe that the convergence rate given by (6.33) applies also for the consistent case. Using (6.33), one may ultimately determine the number of complete cycles required to either converge to or reach the limit cycle, for any desired degree of accuracy.

In order to give an indication of the adaptation of a simple extrapolator to a training set, a second order extrapolator (k=1) was trained with the augmented patterns

$$\underline{f}_1^a = (1, 4, 9, 16)/\sqrt{98}$$
$$\underline{f}_2^a = (9, 4, 1, 0)/\sqrt{98}$$
$$\underline{f}_3^a = (-1, 1, 7, 17)/\sqrt{51}$$

where the vectors have been normalized. Figure 6.3 shows the three weights $\underline{W}_{I,j}$ for j = 0, 1, 2 at the beginning of each cycle. The initial approximation $\underline{W}_0 = (0, 0, 0)$ has been used. The conventional solution (1, -3, 3) is shown

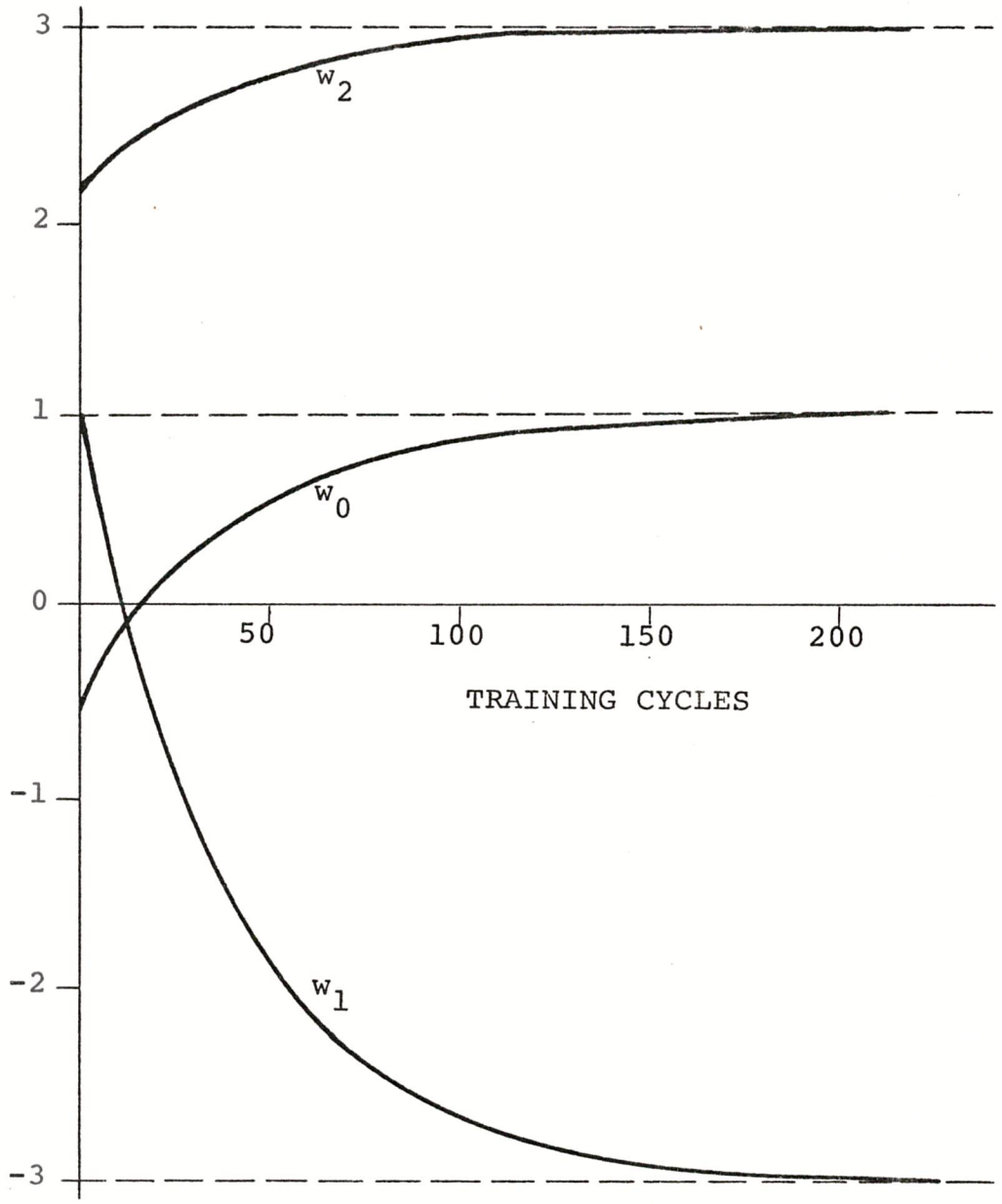

Figure 6.3 Example of training a second order extrapolator with patterns corresponding to second order polynomials. The weight vector (w_0, w_1, w_2) converges to the solution (1, -3, 3) which coincides with Lagrangian extrapolation.

as the asymptotes of the weight curves. For this example, 211 cycles were required for convergence, defined as the condition where the relative change between the vector length of $\underline{W}_I$ at the beginning of two successive cycles was less than .0001.

It is clear that the adaptive or trained extrapolator works as well as conventional extrapolators in the situation where the extrapolated curve is a polynomial of order less than or equal to k for a k^{th} order extrapolation formula, provided that k+1 independent training patterns are used in the training set. While for single variable applications, it is usually easier to use conventional techniques, for multi-variable polynomials, it is harder to determine the extrapolation formulas conventionally, so that the adaptive approach becomes more attractive. In most physical situations, the functions are not polynomials, nor are measurements consistent. Thus, the situation of inconsistency due to various types of error including truncation error is a central issue in both adaptive and conventional extrapolation techniques. In the next section, adaptive extrapolation is extended to inconsistent data, and the occurrence of such data is discussed. It is then shown that for many such occurrences, the adaptive approach is superior to conventional extrapolation.

6.3 Inconsistent Training Patterns In this section, the situation is considered wherein the patterns used to train the extrapolator are inconsistent. Inconsistent patterns are defined and their occurrence explained. The techniques introduced in the last section are modified in order to handle an inconsistent training set. Convergence is proven for the modified training algorithm.

In the previous section, it was explained that the training procedure would be equivalent to solving equation (6.13) for the solution vector $\underline{W}_s$ consisting of k+1 components. The augmented vectors $\underline{f}_r^a$

$$\underline{f}_r^a = (f_{r,0}, f_{r,1}, \cdots f_{r,k}, f_{r,k+1})$$

represent the rows of the matrix $\underline{\underline{f}}$ augmented by the column vector $\underline{f}_{k+1}$ of (6.14). The set of vectors (patterns) $\underline{f}_r^a$ is defined as being inconsistent when the rank of the augmented matrix $\underline{\underline{f}}^a$ is greater than the rank of the matrix $\underline{\underline{f}}$. In the language of pattern recognition theory, inconsistency corresponds approximately to the set of patterns $\{\underline{f}_r^a\}$ not being linearly separable. In pattern recognition problems, one considers a vector space dimension d_p corresponding to patterns to be classified into classes. The usual problem is then to determine hyperplanes that separate the patterns into two regions according to their class (Nilsson [1965]). The distinction between this problem and the extrapolator training problem considered in this

chapter, is that the training patterns lie on the hypersurface, while the patterns used to obtain a classifier in the classification problem lie on either side of the hypersurface to be determined.[1] A clue to the solution of the inconsistent training situation is that the surface on which inconsistent patterns lie is of higher dimension than expected. To obtain a solution without raising the order of the extrapolator, one may consider a tolerance error between a solution and members of the training set such that the distances between training patterns and the solution hyperplane are within a criteria of acceptance. Before developing the techniques for handling an inconsistent training set, it is instructive to review some other solutions to the inconsistency problem. It is also important to indicate how inconsistent patterns for the adaptive extrapolator can arise.

There exist several approaches to inconsistent patterns. One often goes to a non-linear recognition system in which the pattern components (or features as they are usually called) are combined as products in addition to being used individually. Often transformations such as clustering (Sebestyn [1962]) can be determined from considerations of

---

[1]The reason for this distinction is that pattern recognition problems deal with inequalities, while the adaptive extrapolation problem deals with equalities.

the particular problem, increasing the number of components or features may result in consistency.

A number of investigators, including Highleyman [1962], Smith [1968], and Mengert [1970], have approached the inconsistency problem using the performance of a linear pattern classifier as their decision criteria. A decision surface corresponding to the solution vector $\underline{W}_S$ in (6.13), is sought which minimizes the number of wrong classifications. Highleyman [1962] minimizes a loss function corresponding to the sum of the losses resulting from each wrongly classified pattern. By first choosing some continuous loss function which is a function of the distance from each pattern to the current hyperplane, and then using the method of steepest descent, the optimal decision surface can be obtained by iteration. Mengert [1970] uses the same optimality criteria in a search-oriented algorithm which is essentially a gradient method that selectively eliminates some of the patterns. Smith [1968] uses a simplex approach, and the performance criteria is essentially the same as the other two methods.

While one could adopt any one of the above approaches for training the extrapolator, one may instead consider a mechanically simpler approach. In the above techniques, updating $\underline{W}_I$ requires a searching through the training set for an optimal direction of change, in which a computation

is performed for each pattern. Evidently, there may be as much computation in a single iteration of this type of technique as exists for an entire cycle of the consistent training procedure given above. Also, while the "search" techniques are conceptually simple, they are many times more complex than the normal training procedure in terms of their mechanical implementation.

There are several drawbacks in using a simplex approach to train the extrapolator. The simplex formulations can require substantially more storage for the manipulated matrix than is required for storing the patterns alone. Techniques such as partitioning can be applied so that only a small portion of this matrix need be stored in a computer's main memory during execution of the algorithm, but a price is paid in terms of the time required to perform partitioning operations. Since the simplex algorithm requires random access of its matrix, considerable time may be spent on scanning the secondary storage medium where the bulk of the matrix lies, when the pattern set is large. A second problem can exist for ill-conditioned matrices on which the simplex method operates. This problem stems from the "pivot operation" (Nering [1967]) of the simplex method in which an element of the matrix divides a matrix row. When the selected element is very small, large computational

errors can result. Since each step of the algorithm modifies matrix elements, a large computational error from any step can propagate through the succeeding computations.

Before developing the approach to be taken, one may consider how inconsistent patterns arise in the extrapolation problem. It has already been pointed out that no solution may exist to (6.13) if the order of the extrapolation, k, is less than the order of a polynomial function being extrapolated. When the extrapolated function is not a polynomial, there may be no finite order extrapolator which can exactly determine the extrapolated value of the function. In practice, errors due to measurement inaccuracies and roundoff error occur. These errors can easily result in inconsistent patterns even when the extrapolation function, f(x), is ideally a polynomial of degree less than k. One can therefore intuitively expect to be able to obtain a set of independent training patterns of number greater than k+1, in practical situations.

When f(x) is a continuous function, an error bound can be obtained for conventional extrapolation formulas. Assuming measurement error is also bounded, an overall error bound due to extrapolation will exist. It is clear that there may exist a different extrapolation formula resulting in a smaller amount of error. Since inconsistency can be treated as a measure of error, one may consider a training

procedure which produces an inconsistency-tolerant extrapolator.

The approach taken is to incorporate an error threshold within the training algorithm. Assume that training has failed to converge, as indicated by a non-diminishing error

$$E_I = f_{I,k+1} - \underline{f}_I \cdot \underline{W}_I \tag{6.34}$$

for a particular pattern in successive cycles. Using an error threshold, τ, one may now allow the training algorithm to tolerate an error magnitude less than τ. That is, the vector $\underline{W}_I$ is adjusted only by the amount of error over this threshold. Equation (6.30) can now be written as

$$\underline{W}_{I+1} = \begin{cases} \underline{W}_I & ; |E_I| \leq \tau \\ \underline{W}_I + (E_I - \tau \text{sgn}\,[E_I]) \cdot \underline{F}_I & ; \text{otherwise} \end{cases} \tag{6.35}$$

where sgn $[E_I]$ is the sign of the value E_I, and E_I is defined by (6.34).

The value of τ may, in general, depend on $f_{I,k+1}$ or on the components of $\underline{f}_I$. As an example of the many possibilities, an absolute error threshold, a threshold relative to $f_{I,k+1}$, or a threshold relative to $|\underline{f}_I|$ may be used. No attempt is made to determine the "best" measure of error to use since such a decision must be made relative to a specific application. An obvious measure of the "goodness" of an error criteria must ultimately be judged by the error that propagates over many iterations. In practice, one may test several alternatives for a

particular situation at hand. The error criteria illustrated in this chapter is known as minimax (Hamming (1962)).

The threshold considered here is chosen for its simplicity and intuitive appeal. The threshold is defined relative to the pattern length of the current iteration. That is, for the I^{th} iteration, $\underline{W}_I$ is adjusted by pattern $\underline{f}_I$ such that

$$\underline{W}_{I+1} = \begin{cases} \underline{W}_I & ; |E_I| \leq \tau_0 |\underline{f}_I| \\ \underline{W}_I + (E_I - \tau_0 |\underline{f}_I| \text{sign}[E_I]) \cdot \underline{f}_I & ; \text{otherwise} \end{cases} \tag{6.36}$$

Note that

$$\tau \equiv \tau_0 \, |\underline{f}_I| .$$

Since the training patterns have already been normalized by the factor $|\underline{f}_I|$, the threshold is a constant. To explain why the threshold algorithm is a reasonable one, one may consider the following heuristic argument. The error, E_I, may be written

$$\begin{aligned} E_I &= f_{I,k+1} - \underline{f}_I \cdot \underline{W}_I \\ &= \underline{f}_I \cdot \hat{\underline{W}} - \underline{f}_I \cdot \underline{W}_I . \end{aligned} \tag{6.37}$$

The vector $\hat{\underline{W}}$ represents an ideal solution which would exactly compute $f_{I,k+1}$.

Although no $\hat{\underline{W}}$ vector can exist in an inconsistent situation, it is still useful to hypothesize its existence. Equation (6.37) may then be normalized with respect to $|\underline{f}_I|$ as

$$\frac{\underline{E}_I}{|\underline{f}_I|} = \frac{\underline{f}_I}{|\underline{f}_I|} \cdot [\hat{\underline{W}} - \underline{W}_I]$$

where the factor within the brackets represents a vector difference between the ideal solution and $\underline{W}_I$. One may also write

$$\frac{\underline{f}_I}{|\underline{f}_I|} \equiv \underline{U}_I$$

where $\underline{U}_I$ is a unit length vector having the direction of $\underline{f}_I$. The criteria is to minimize the maximum projection of the error between the ideal and the solution vectors, where the projection is in the direction of each vector in the vector space $\{\underline{f}_r\}$. If one can assume that the training set is "representative" of the entire vector space, then by choosing a $\underline{W}_s$ which results in the same projected length $\underline{U}_I \cdot [\hat{\underline{W}} - \underline{W}_s]$ for each vector in the training set, one may speculate that, on the average, the projected length for all vectors in the space is of similar magnitude. One may qualitatively conjecture that by increasing the sampling of training vectors, this set becomes more representative.

To obtain $\underline{W}_s$, the following procedure is used. Starting with $\tau_0 = 0$, perform the training sequence using the correction rule of (6.36). If convergence occurs, $\underline{W}_s$ is obtained. If convergence does not occur, increase τ_0 by a fixed amount $\Delta\tau_0$. Again perform training. If convergence fails to occur, again increase τ_0 by the same increment.

If convergence occurs, the resulting vector $\underline{W}_s$ can be used, or else one may try a value of τ_0 between the last value and the current value, by a trial and error procedure. It is intuitively clear that by increasing τ_0 sufficiently, convergence must eventually occur since the increasing threshold allows the patterns to become more inconsistent.

To properly verify convergence of (6.36) for a sufficiently large τ_0, known properties of convergence of error correction training algorithms are employed. According to Nilsson [1965], an algorithm of the form

$$W_{I+1} = \underline{W}_I + \lambda E_I \underline{f}_I \tag{6.38}$$

will produce convergence provided $0<\lambda\leq 2$. One may see that (6.36) corresponds to a $\lambda\leq 1$. That is λ is unity for $\tau_0=0$, and less than unity for $\tau_0>0$. The error correction algorithms of (6.38) do not depend on the order of sequence for the $\underline{f}_I$ for convergence, although the rate of convergence is affected. Therefore, for the case that $|E_I|\leq\tau_0|\underline{f}_r|$ in (6.36), it is though the vector $\underline{f}_r$ was skipped in the sequence of training patterns.

To complete the discussion of (6.36) one wishes to be sure that having raised τ_0 to a sufficiently large value to obtain a convergence of the training algorithm, there is no other solution $\underline{W}_s'$ corresponding to a τ_0' less than $\tau_0-\Delta\tau_0$ such that

$$|f_{I,k+1} - \underline{f}_I \cdot \underline{W}_s'|\leq\tau_0'|\underline{f}_I|$$

for all vectors $\underline{f}_I$ in the training set. It is remembered that $\Delta\tau_0$ was the increment of τ_0 made at each point where convergence failed. The verification that no such τ_0' exists is possible by the fact that for $\tau_0-\Delta\tau_0$, the training would have converged, thus not requiring a further increment of τ_0 to be made for convergence. The lack of existence of τ_0' does not preclude a smaller increment of τ_0 to be used once convergence has been established for τ_0 but not $\tau_0-\Delta\tau_0$. The following theorem may now be stated.

Theorem 6.6 Given the training set $\{\underline{f}_r^a | r=1,\ldots,m+1\}$ and an arbitrary vector $\underline{W}_1$. Let each of the $\underline{f}_r^a$ be normalized (scaled) such that $|\underline{f}_r|=1$. Define the sequence, $\underline{f}_I$,

$$\underline{f}_1, \underline{f}_2, \ldots, \underline{f}_{m+1}, \underline{f}_{m+2}, \ldots$$

where the subscript of the vectors in this sequence is modulo m+1. The sequence

$$\underline{W}_1, \underline{W}_2, \ldots, \underline{W}_I, \ldots \tag{6.39}$$

defined by (with E_I defined by (6.34)

$$\underline{W}_{I+1} = \begin{cases} \underline{W}_I & |E_I| \leq \tau_0 \\ \underline{W}_I + (E_I - \tau_0 \operatorname{sgn}[E_I]) \cdot \underline{f}_I & \text{otherwise} \end{cases} \tag{6.40}$$

converges to a solution vector $\underline{W}_s$ provided the constant $\tau_0 \geq 0$ is sufficiently large.

Moreover, if the sequence in (6.39) converges for τ_0 but not for $\tau_0-\Delta\tau_0$, $\Delta\tau_0>0$, then there does not exist a $\underline{W}_s'$ such that

$$|f_{r,k+1} - \underline{f}_r \cdot \underline{W}_s'| < \tau_0 - \Delta\tau_0$$

for all $\underline{f}_r$ in the training set. Note that lack of convergence does not depend on the initial vector $\underline{W}_1$ since, by Theorem 6.5, the limit cycle resulting in the non-convergent case is unique. It is also noted that the normalization of the training vectors has led to a simpler error correction statement given as (6.40). That is, $|\underline{f}_r| = 1$ has been substituted in (6.36) to yield (6.40). It is also clear that Theorem 6.6 is true for other threshold algorithms.

Before elaborating on some advantages of an adaptive extrapolator, a simple example is given to illustrate the performance of an extrapolator adapted to a training set compared to the performance of a conventional extrapolator of the same order. A second order extrapolator is used to extrapolate functions of the form

$$f(x) = A\, e^{-Bx} \tag{6.41}$$

To demonstrate the effect of using data from different functions, assume that the constants A and B can have values with the range

$$1 \leq A \leq 2$$

$$.1 \leq B \leq .3$$

The conventional extrapolator is

$$f_3 = 3f_2 - 3f_1 + f_0. \tag{6.42}$$

The increment of x between the $f(x_j)$ is $\Delta x = .2$. To generate a training set, nine sample functions were used, with parameters given in the table below.

Function	A	B
1	1.0	.1
2	1.0	.2
3	1.0	.3
4	1.5	.1
5	1.5	.2
6	1.5	.3
7	2.0	.1
8	2.0	.2
9	2.0	.3

For each of these nine functions, ten patterns were generated with values of $x_0 = 0, .6, 1.2, \ldots, 4.8, 5.4$. The same increment of x was used, that is $\Delta x = .2$, between the data points of the patterns. For example, the pattern corresponding to $x_0 = 1.2$ consists of the four components

$$[f(1.2), f(1.4), f(1.6), f(1.8)].$$

Using these 90 training patterns, a second order extrapolator was obtained using (6.40), after the patterns had been normalized. The training resulting in a final value of τ_0 of .00006, and a solution vector

$$\underline{W}_s = (.99996209, -3.0000357, 2.99996638) \qquad (6.43)$$

The smallest increment of τ_0 was .00001. Using the solution vector of (6.43), the function

$$f(x) = 1.75\, e^{-.2x}$$

was extrapolated with the initial data of f(0), f(.2), and

f(.4). The extrapolated function using $\underline{W}_s$ was computed at increments of x = .2 to x = 20. The conventional extrapolator of (6.42) was also used to compute f(x). These two extrapolators as well as the actual function are shown in Figure 6.4. Notice that the adaptive extrapolator results in an improved computation, even though the weights are only slightly different from the conventional weights, (1, -3, 3).

The above example used patterns from several different functions. In practice, one may have data for the function to be extrapolated for many points along the initial portion of the function curve. Obviously, the more such data points, the greater the number of patterns for the particular curve will be available. It is clear that patterns for the actual curve being extrapolated should be used in preference to patterns for other similar functions. In the next section, the hypothetical example will be used to show how one might displace the patterns for other functions as more data becomes available for the function "under test".

In the remainder of this section, truncation error will be discussed in order to show the superiority of an adaptively determined extrapolator in certain situations. This type of error arises when f(x) is a non-polynomial function, or is a polynomial function of order greater than the order of the extrapolator. When f(x) is non-singular in the interval of x, (x_0, x_{k+1}) the function may be approximated by a k-order polynomial which corresponds to the first k+1 terms of

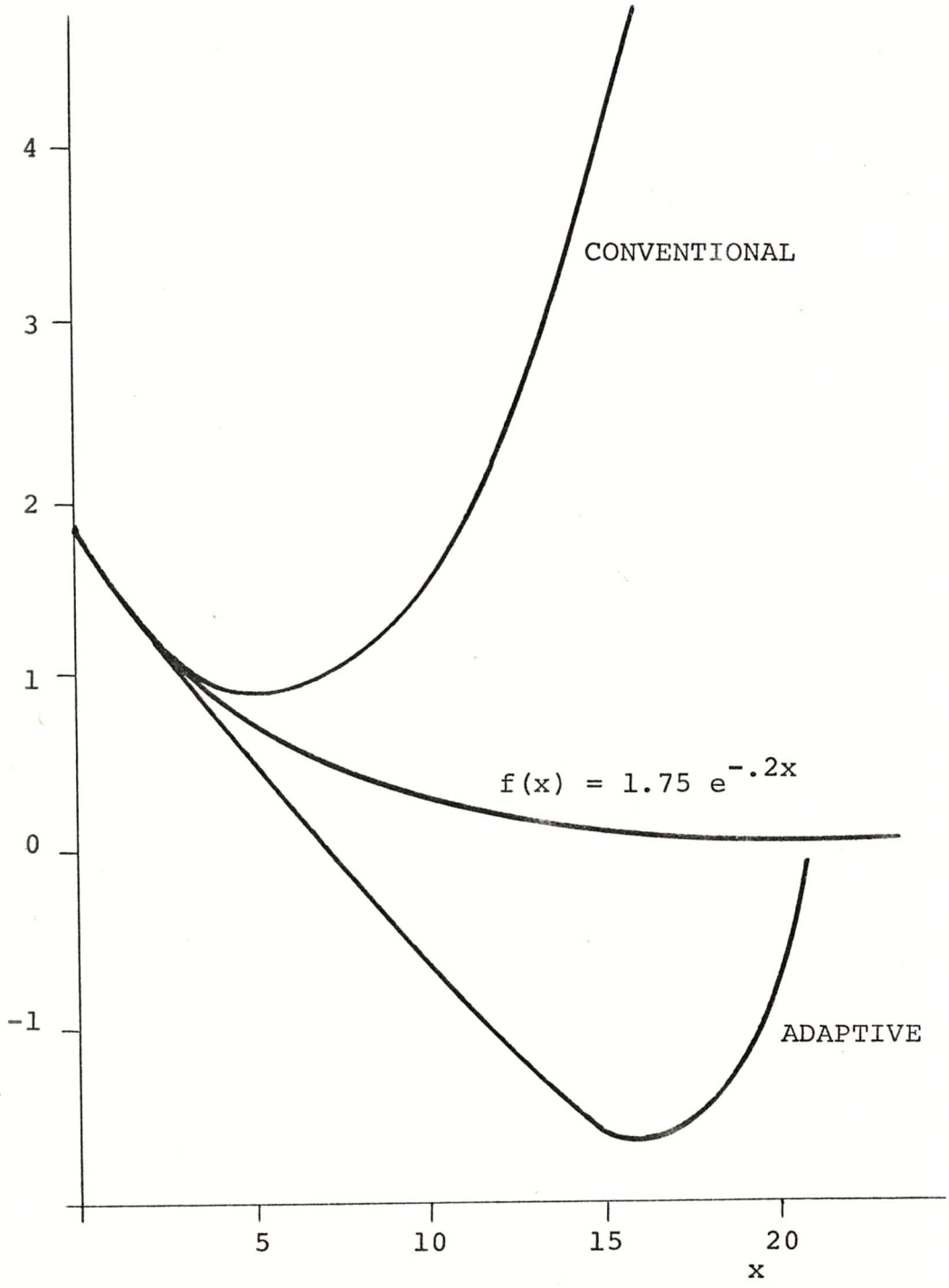

Figure 6.4 Example of a second order adaptive extrapolation. A second order conventional extrapolator is also shown. f(x) is the actual function being extrapolated.

a Taylor series for f(x) expanded about x_0 (Hildebrand [1962]). Conventional extrapolation formulas such as the Lagrange form of (6.7) use such an approximation. Using difference of values f(x), an error similar to the remainder term of the truncated Taylor series can be obtained. This error may be written as (Hamming [1962])

$$e_T = \frac{(x_{k+1}-x_0)\cdot(x_{k+1}-x_1)\ldots(x_{k+1}-x_k)\cdot f^{(k+1)}(x_*)}{(k+1)!} \qquad (6.44)$$

where x_* is some point within (x_0, x_{k+1}). Since

$$x_{j+1} - x_j = \Delta x$$

for all j by the assumption of equi-spaced data, (6.44) may be simplified and written as

$$e_T = (\Delta x)^{k+1} \cdot f^{(k+1)}(x_*). \qquad (6.45)$$

In order to determine the truncation error bound, it is necessary to be able to evaluate the k+1 derivative in the interval and determine the minimum and the maximum. Besides being difficult when f(x) is not known explicitly, this bound is usually pessimistic. Hamming [1962] states that the coefficients of successive differences tend to be proportional to each other in extrapolation formulas generated by the well-known lozenge diagram. This diagram is often used to generate a large number of conventional interpolation formulas. By using this property of the coefficients, it is possible to incorporate the effect of the higher-order differences in extrapolation by combining them with lower order differences. The result is a low order formula

which "contains" the effect of high order differences to some extent. Generally speaking, it is necessary to consider the particular function to be extrapolated in order to determine the appropriate combination of the differences. While normally one must do this analytically, the adaptive approach to extrapolation incorporates the combination of high order differences automatically. That is, the extrapolation weights are adapted to the training patterns such that the maximum inconsistency is minimized. Since the measure of inconsistency is that of the error of extrapolating the training data, the training procedure must result in the extrapolation weights which minimize this error.

Adaptive extrapolation cannot only reduce the truncation error, but can eliminate the error in certain cases. Consider the important case when $f(x) = A \exp [Bx]$. Replacing x by $(x-\Delta x+\Delta x)$, one may write

$$A \exp [Bx] = A \exp [B(x-\Delta x)] \cdot \exp [B\Delta x]. \qquad (6.46)$$

Therefore,

$$f(x) = (\exp[B\Delta x]) \cdot f(x-\Delta x). \qquad (6.47)$$

Notice that the coefficient of $f(x-\Delta x)$ is a constant for a particular selection of A and B. One concludes that a zero order adaptive extrapolator can be obtained which exactly extrapolates a particular f(x). Other important functions exist which allow similar results. For example, the function $f(x) = \sin [Bx+c]$ may be written

$$\sin[Bx+c]=W_2\cdot\sin[B(x-\Delta x)+c]+W_1\cdot\sin[B(x-2\Delta x)+c]$$

$$=W_2\ \sin[Bx+c]\cos[B\Delta x]-W_2\ \cos[Bx+c]\sin[B\Delta x]$$

$$+W_1\ \sin[Bx+c]\cos[2B\Delta x]-W_1\cos[Bx+c]\sin[2B\Delta x].$$

Let $W_2\ \sin[B\Delta x] = -W_1\ \sin[2B\Delta x]$

or

$$W_2 = -W_1\ \frac{\sin[2B\Delta x]}{\sin[B\Delta x]} = -2W_1\ \cos[B\Delta x].$$

Collecting common terms of $\cos[Bx+c]$ yields

$$W_2\ \sin[B\Delta x] = -W_1\ \sin[2B\Delta x]$$

or

$$W_2 = -2W_1\ \cos B\Delta x.$$

Similarly, common terms of $\sin[Bx+c]$ yields

$$W_2\ \cos[B\Delta x] + W_1\ \cos[2B\Delta x] = 1$$

or

$$W_1 = \frac{1-W_2\cos B\Delta x}{\cos^2 B\Delta x-\sin^2 B\Delta x} = \frac{1+2W_1\cos^2 B\Delta x}{\cos^2 B\Delta x-\sin^2 B\Delta x}. \qquad (6.48)$$

Manipulating (6.48), one obtains

$$W_1 = -1$$

and therefore

$$W_2 = 2\cos B\Delta x.$$

Similar results can be obtained for $\cos[Bx+c]$ and other functions. Table 6.5 gives several such recursion formulas. The recursion for the exponential function may be generalized to a function which is a linear combination of several different exponentials. The following theorem states the generalization.

f(x)	Recursion	W_1	W_2
A sin[Bx+c]	$f(x)=W_2 \cdot f(x-\Delta x)+W_1 \cdot f(x-2\Delta x)$	-1	$2cos B\Delta x$
A cos[Bx+c]	$f(x)=W_2 \cdot f(x-\Delta x)+W_1 \cdot f(x-2\Delta x)$	-1	$2cos B\Delta x$
A exp[Bx+c]	$f(x)=W_1 \cdot f(x-\Delta x)$	exp[$-B\Delta x$]	---
A $sinh$[Bx+c]	$f(x)=W_2 \cdot f(x-\Delta x)+W_1 \cdot f(x-2\Delta x)$	-1	$2cosh[B\Delta x]$
A $cosh$[Bx+c]	$f(x)=W_2 \cdot f(x+\Delta x)+W_1 \cdot f(x-2\Delta x)$	-1	$2cosh[B\Delta x]$

Table 6.5 SEVERAL RECURSION FORMULAS FOR EQUI-SPACED EXTRAPOLATION

Theorem 6.7 Let the function f(x) be given as the linear sum of $\hat{n}$ functions $\hat{g}_i(x)$

$$f(x) = \sum_{i=1}^{\hat{n}} A_i \cdot \hat{g}_i(x). \tag{6.49}$$

Corresponding to a fixed Δx, let each $g_i(x)$ have the recursion property

$$\hat{g}_{i,j} = \hat{B}_i \cdot \hat{g}_{i,j-1} \tag{6.50}$$

where $\hat{g}_{i,j}$ denotes $\hat{g}_i(x_j)$ and $x_j - x_{j-1} = \Delta x$, for all j. Let n of the $\hat{n}$ constants $\hat{B}_i$ be different. Then for the function f(x), there exists a recursion of n previous values of f(x) such that, for $f(x_j)$ written f_j

$$f_n = \sum_{j=0}^{n-1} C_j \cdot f_j \tag{6.51}$$

Proof Without loss of generality, the constants A_i in (6.49) may be taken as unity. Indeed one can combine these constants with the original $\hat{g}_i$ functions. The recursion of (6.50) will still be valid. Similarly, if two or more $\hat{g}_i$ functions have the same recursion constant, $\hat{B}_i$, these functions may be summed together to yield a new function which has the common recursion constant $\hat{B}_i$ (see 6.50).

In what follows, the above two simplifications are assumed to have been made. That is,

$$f(x) = \sum_{i=1}^{n} g_i(x) \tag{6.52}$$

where $g_i(x)$ are the resulting functions. Similarly, the recursion formulas are written as before

$$g_{i,j} = B_i \cdot g_{i,j-1} \tag{6.53}$$

where now all the B_i are different.

Using the index notation of (6.53), (6.52) becomes

$$f_j = \sum_{i=1}^{n} g_{i,j}. \tag{6.54}$$

Substitution of (6.54) in (6.51) yields

$$\sum_{i=1}^{n} g_{i,n} = \sum_{j=0}^{n-1} C_j \cdot \sum_{i=1}^{n} g_{i,j}. \tag{6.55}$$

Observe that (6.53) may be iterated backwards to yield

$$g_{i,j} = B_i^j \cdot g_{i,0}. \tag{6.56}$$

Substitution of (6.56) in (6.55) yields

$$\sum_{i=1}^{n} B_i^n g_{i,0} = \sum_{j=0}^{n-1} C_j \cdot \sum_{i=1}^{n} B_i^j \cdot g_{i,0}$$

$$= \sum_{i=1}^{n} g_{i,0} \cdot \sum_{j=0}^{n-1} C_j B_i^j. \tag{6.57}$$

Observe that (6.56) is required to hold for all $g_{i,0}$. This requirement implies the n conditions (resulting from considering common terms of $g_{i,0}$ for each i)

$$B_i^n = \sum_{j=0}^{n-1} C_j \cdot B_i^j. \tag{6.58}$$

These n conditions may be expressed in matrix form as

$$\begin{bmatrix} B_1^n \\ \cdot \\ \cdot \\ \cdot \\ B_n^n \end{bmatrix} = \begin{bmatrix} 1 & B_1 & B_1^2 \ldots B_1^{n-1} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ 1 & B_n & B_n^2 \ldots B_n^{n-1} \end{bmatrix} \begin{bmatrix} C_0 \\ \cdot \\ \cdot \\ \cdot \\ C_{n-1} \end{bmatrix} \qquad (6.59)$$

The determinant of the n x n matrix is recognized as the Vandermonde determinant of order n, which is non-zero for all <u>different</u> indeterminates, B_i (Nering [1967]). Therefore, the matrix is non-singular. One may obtain a unique set of recursion constants, C_i, by matrix operations. Therefore the recursion of (6.51) exists.

<u>QED</u>

From the preceding discussion, it has been shown that adaptive extrapolation often results in better accuracy in that truncation error is reduced or eliminated. This method of determining an extrapolation function suffers from the usual deficiencies of pattern recognition schemes in that it is difficult to predict how well the method will work in a particular application. The success of applying the adaptive approach is greatly dependent on how well the training patterns represent the function to be extrapolated. It has been shown that an adaptively determined extrapolator performs at least as well as a conventional extrapolator in terms of the error criteria incorporated into the

training algorithm. However, the criterion corresponds to a single iteration and does not indicate how the extrapolation error will propagate over many iterations. The adaptive method does give an indication of the adequacy of the order of extrapolation through the error threshold required for convergence. One may increase the number of data values used in extrapolation and may also incorporate products of data values to achieve a non-linear extrapolation. The problem of choosing the order of extrapolation and the number of product terms appears to be resolvable by trial and error only, using the training set and the resulting error threshold.

In the next section, it is shown how the adaptively determined extrapolator may be incorporated in a long term storage prediction application. The extension to a multidimensional extrapolation will also be made.

6.4 Application of Adaptive Extrapolation to Long Term Storage Prediction

In this section, the technique developed in this chapter is applied to the problem of predicting long term storage effects for a hypothetical device. The adaptive extrapolator is incorporated in a more complex system called a simulator. The simulator, in general, contains several extrapolators. Each extrapolator corresponds to one of several normally measurable parameters for the device. Using the parameter values measured at regular intervals of

time, combined in the form of patterns, a training set is used to train each of the extrapolators in the simulator. Once trained, the simulator is provided with an initial pattern. The simulator extrapolates each parameter to provide a revised pattern, now called a "next state" for the device. This next state pattern is fed back to the simulator input and the extrapolation repeated. By iterating the extrapolation and updating of the input pattern, multi-dimensional extrapolation is performed. The iterations simulate future behavior of the device. An additional system, called a recognizer, monitors the simulator output, and determines when a terminal state has been reached. A terminal state is defined as a condition of device failure. Additionally, the overall system may terminate prior to a failure condition when a predetermined number of iterations have been completed. In what follows, the entire system just described is explained in detail. This system is meant to indicate the manner of application. An actual application would require a trial and error procedure to determine a suitable order for each extrapolator as well as the possible conversion to non-linear pattern separation, i. e. the use of products of parameter values as pattern components.

For the purpose of simplifying the presentation, it is supposed that there are three measurable parameters for a device, f, g, and h. The three parameters are assumed to

be dependent. That is, they are hypothesized as being coupled. The independent variable is time, denote "t", and one may write the parameters as

$$f(t) = f[g(t), h(t), t]$$

$$g(t) = g[f(t), h(t), t]$$

$$h(t) = h[f(t), g(t), t].$$

An extrapolator for each parameter is trained using an individual set of augmented training patterns. Each pattern of the set is composed of k+1 values of f(t), g(t), and g(t) at $t_0, t_1, \ldots, t_k$. (It is stressed that although the same number of values for each of the three parameters is used as data for each of the three extrapolators, different numbers for each parameter and for each of the three extrapolators can be used in practice. That is, for three device parameters, a total of nine numbers could be specified.) The augmented patterns for g(t) can be written as

$$\underline{v}^a_{g,i} = (f_0, f_1, \ldots, f_k, g_0, g_1, \ldots g_k, h_0, \ldots, h_k; g_{k+1}).$$

Note that the last component of $\underline{v}^a_{g,i}$ is g_{k+1} the value that the extrapolator should compute when trained. Also observe that only individual parameter values are used. One could also include products of two or more of the existing components ad additional pattern components or "features". The number M+1 of such patterns for each

parameter is assumed to be the same, although a different number could be used for each extrapolator.

Training of each extrapolator is performed separately using the algorithm of the last section. A different threshold will generally result for each extrapolator. This error threshold serves as an indication of whether more features are required. That is, if any excessive error threshold exists for an extrapolator, additional components can be added as indicated above.

The training set for each extrapolator contains patterns obtained by testing similar devices. For example, transistors of a particular type may be subjected to measurements in order to determine such pattern data. A number of these devices may be tested over a long time interval to provide "typical" behavior under the intended storage conditions.

The simulator is shown as a block diagram in Figure 6.6. Since the patterns used for simulation are not augmented and are the same for all three extrapolators, they are denoted as the vector $\underline{V}_I$ where the subscript indicates the iteration of the simulation. The quantity $\underline{V}_1$ is the initial vector and represents the most recent k+1 values of each of the parameters, f(t), g(t), and h(t). Initially, $\underline{V}_1$ is presented to each of the three extrapolators, assumed already trained. The output of each extrapolator is the simulated next value

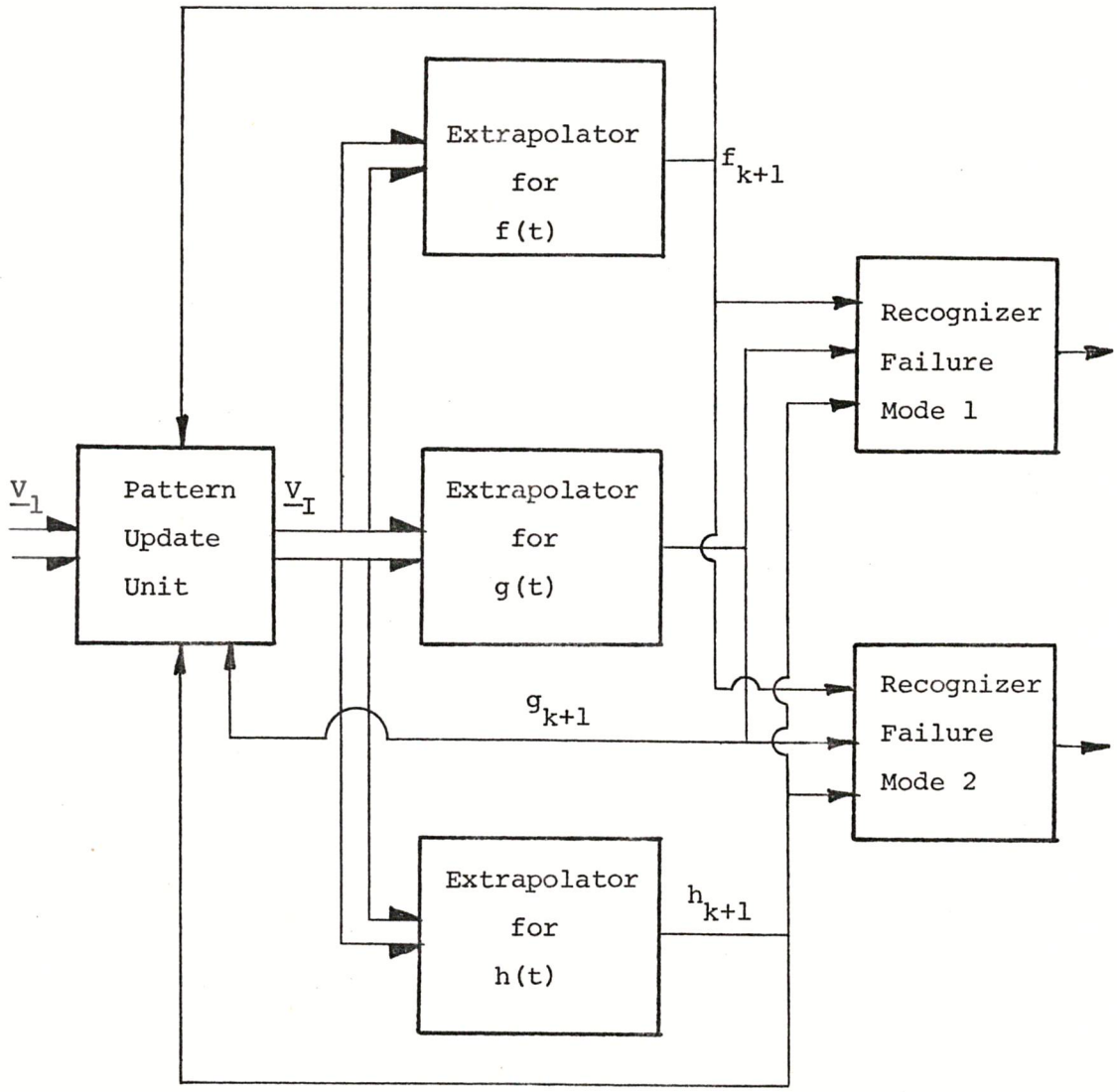

Figure 6.6 Configuration of a three parameter extrapolator. The pattern update unit revises the state vector $\underline{V}_I$ using the "next" values of f_{k+1}, g_{k+1}, and h_{k+1} of the three device parameters. $\underline{V}_1$ is the initial state vector or pattern. Each of the extrapolators are separately trained using a set of training vectors. The failure mode recognizer determines if a particular failure mode has been reached.

of each parameter. These three values are fed back to the block denoted the Pattern Update Unit in Figure 6.6. This block performs the function of replacing f_k, g_k, and h_k by the respective simulated values as determined by each of the extrapolators. Similarly f_{I+1} replaces f_I, g_{I+1} replaces g_I and so forth. The revised pattern, $\underline{V}_2$ is then presented to the extrapolators as before, and the process repeated. This iterative procedure is repeated until either a predetermined amount of simulated time is reached, or when the combination of the three extrapolator outputs represents a failure.

As shown in the figure, the next state parameter values, in addition to being fed back to the simulator input, are presented to two blocks called failure mode recognizers. The function of each recognizer is to determine whether the current parameter combination corresponds to a particular failure mode. While failure modes may sometimes correspond to a single parameter exceeding an a priori acceptability range, a more complex situation usually prevails. In this more complex case, failure modes correspond to an interdependence of the parameters. When such relationships are known, the recognizer simply tests this relationship in performing its decision function. Often the decision function can be implemented as an analogue discriminant function machine (Nilsson [1962]) commonly encountered in pattern recognition theory.

When the recognizer is required to determine the existence of a physical failure mechanism, the relationship between values of parameters and these mechanisms may not be intrinsically apparent. For an electronic component, one may often require determination of the dominant mechanism by physically dissecting the device, or by examining the active material with an electron microscope.[1] Obviously, such an analysis usually amounts to destructive testing. By measuring the parameters just prior to this examination, it is possible to classify the parametric patterns according to the failure mechanism physically determined for the device. By gathering a large number of such patterns and separating them into classes corresponding to dominant failure modes, one has the basis for training a pattern classifier. The well established techniques of pattern recognition theory can be applied so that the recognizer, which examines the output of the simulator, not only recognizes a terminal condition, but also discriminates between different failure mechanisms.

The overall system of the simulator and recognizer form a aystem that can predict the long term storage effects for a particular device. Besides being a useful screening techniques for devices that must be stored for long periods

---

[1]See for example studies by Behera and Speer [1972], Gonzales [1973], and Burity, R. S. et al. [1967].

of time before they are used, it serves as a way of generalizing the results of long term testing to environmental conditions other than that existing when this initial testing was performed. Changing the environmental stress on a device often results in changing the time constants of failure mechanisms, but necessarily not the mechanisms themselves. If such stresses as voltage, temperature, etc. are only moderately changed, one can expect that the rate of failure will change but that the extrapolation formulas will remain approximately valid. This conjecture is defined on the basis that the coefficients of conventional extrapolation are independent of the spacing interval "t", although the error term is sensitive to it. Since the rate of the mechanisms differs to some extent from device to device, even under the same environmental conditions, using patterns from many devices guarantees some generalization.

6.5 Conclusions Adaptive extrapolation has been described with application to the problem of predicting long term reliability of electronic devices. A training procedure was described and shown to be convergent even when the data used to train the extrapolator was inconsistent.

When compared to conventional exact-fit extrapolation, the adaptive method was shown to give comparable results, if the function being extrapolated was a polynomial of order less than or equal to that of the extrapolation.

When this function was not a polynomial, it was shown that adaptive extrapolation could result in a vast improvement in accuracy, providing the data used to train the extrapolator was representative of the function. It was further shown how adaptive extrapolation is easily extended to multivariable functions.

The application to predicting long term reliability of devices was described. Once a simulator was trained, the initial behavior of a particular device could be iteratively extrapolated to large time intervals. If a recognizer was connected to the simulator output, it would be possible to determine a terminal state for the device. This recognizer could be designed from a priori information alone or with additional training patterns using the usual pattern recognition theory techniques.

7. FUTURE WORK

Several areas for future work were indicated in previous chapters which represent possible improvements and extensions to the techniques developed in the thesis. This chapter summarizes these areas of investigation.

A simplification made in discussing the property of linear conformity (lc) and the bounding algorithm (BA) was that random variables (RV's) combined in a linear sum were statistically independent. Theorems 3.7 and 3.8 addressed the more likely situation of statistical dependence of physical quantities. In what may seem an extreme case, monotonic dependence of RV's was shown not to invalidate results obtained for statistically independent RV's. However, investigation of non-monotonic dependence should also be made in order to concretely ascertain that monotonic dependencies represent a "worst-case" situation.

A topic allied with determining the effects of statistical dependence on the linear combinations of RV's is that of determining transformations which will improve prior statistical information by reducing the variance. The consequence of reducing the variance is a simultaneous reduction of the TI for the BA. As an example of such an improvement, consider the class of devices which have input-output functions of the form

$$f(x) = a_0 + \sum_{k=1}^{N} a_k x^k.$$

Assume that all constants a_k, but a_0, are very nearly the same for all the devices that are being represented by f(x). For this constant, assume that there is great variation from one device to another. This wide variation may represent, for instance, different output "bias" levels for electronic amplifiers. Note that $f(0)=a_0$ for a device, and therefore a_0 is directly determined by a single measurement. Since a_0 varies greatly from device to device, its effect on the distribution of $f(x_i)$ for a particular x_i is to greatly increase the variance of this distribution. To remove this undesirable effect of a_0, one may choose the transformation

$$\bar{f}(x) = f(x) - a_0.$$

That is, the value of a_0 is subtracted from the data gathered in forming the statistical distribution for $f(x_i)$. The variance of this distribution is thereby reduced. When using the methods of Chapter 5 for the interpolation of a new device function, the value of a_0 is first determined for this particular function and then added to the tolerance limits representing the distribution information previously gathered. A proposed area of future work is the study of transformations such as the one just presented that would result in improved statistical bounds.

A third area for investigation concerns the heuristic classification procedure developed in Chapter 4. This procedure required several simplifications to be made in order to obtain results. In the development of the heuristic, the standard function $F(\omega)=\exp(-\theta|\omega|)$ emerged as a critical probability density function (pdf) with respect to the property of linear conformity. Application of the heuristic utilized a semi-logarithmic plot since in this case $F(\omega)$ is a straight line. While the Fourier transform was used in the analysis of Chapter 4 due to several desirable properties of the transform, other transformations may be possible which can result in an analytic rather than a heuristic classification procedure for pdf's. In particular, one may be able to utilize the known properties of the standard function $F(\omega)$, including those just mentioned, which will lead to a more powerful analysis procedure.

Several conjectures mentioned in Chapter 5 regarding generalizations of the linear conformity property may be studied as future work. Although these conjectures appear to be valid for the specific pdf's studied by the author, no attempt has been made for direct verification. One of these conjectures is that mono-model, symmetrical pdf's appear always to be lc. Another is that pdf's which are self lc appear to be lc as a collection. It is possible that such conjectures may be more easily verified through

transforms other than the Fourier transform, which was utilized in the thesis.

A final suggested area for future investigation concerns the iterative training algorithms of Chapter 6. These algorithms were used to perform functional inversion and were linear reinforcement formulas. As suggested in the chapter, other iterative function inversion schemes could be investigated which might lead to faster convergence than that of the algorithms presented in the chapter. However, the basic concept of adapting an extrapolator to prior test patterns would be unaffected by the use of different training procedures.

REFERENCES

1. Agmon, S., "The Relaxation Method for Linear Inequalities," Canadian J. Math., vol. 6, pp. 382-392, 1954.
2. Anon., "Manufacturing In-process Control and Measuring Techniques for Integrated Circuits," Westinghouse Elec. Corp., Molecular Electronics Div., Elkridge, MD, Rept. No. IR-8-140 (II).
3. Baron, F. J. and Myers, C. F., "Toward Making LSI Work," Motorola Monitor, vol. 6, No. 3.
4. Behera, S. K. and Speer, D. P., "A Procedure for the Evaluation and Failure Analysis of M. O. S. Memory Circuits Using the Scanning Electron Microscope in Potential Contrast Mode," 10th Annual Proceedings on Reliability Physics, pp. 5-11, 1972.
5. Black, J. R. and Hall, E. L., "Semiconductor Wearout-Its Analysis and Prevention," Motorola Solid-State Technology, vol. 2, No. 1, pp. 2-7, 1972.
6. Bowker, A. H. and Lieberman, G. J., Engineering Statistics, Prentice-Hall, Inc., 1964.
7. Box, G. E. P. and Hill, W. J., "Discrimination Among Mechanistic Models," Technometrics, vol. 9, No. 1, pp. 57-71, February 1967.

8. Box, G. E. P. and Hunter, W. G., "The Experimental Study of Physical Mechanisms," Technometrics, vol. 7, No. 1, pp. 23-42, February 1965.

9. Box, G. E. P. and Lucas, H. L., "Design of Experiments in Non-Linear Situations," Biometrika, vol. 46, pp. 77-90, 1959.

10. Bracchi, G. and Somalvico, M., "The Correlation Method for Computer-Aided Statistical Analysis," IEEE Trans. Rel., vol. R-20, No. 3, pp. 96-101, August 1971.

11. Breipohl, A. M., "Kalman Filtering and Its Application to Reliability," IEEE Trans. Rel., vol. R-18, No. 3, pp. 127-130, August 1969.

12. Brown, D. B., "A Computerized Algorithm for Determining the Reliability of Redundant Configurations," IEEE Trans. Rel., vol. R-20, No. 3, pp. 121-124, August 1971.

13. Brown, D. B. and Martz, H. F., "Simulation Model for the Maintenance of a Deteriorating Component System," IEEE Trans. Rel., vol. R-20, No. 1, pp. 28-32, February 1971.

14. Brown, J. L., "Bounds for Truncation Error in Sampling Expansions of Band-Limited Signals," IEEE Trans. Info. Theory, vol. IT-15, No. 4, pp. 440-444, July 1969.

15. Buritz, R. S., et al., "Investigation of Microcircuit Surface Metallurgy," Report TIC 2749.50/1819, Hughs Aircraft Company, Aerospace Group, January 1969.

16. Chang, S. S. L., "Adaptive Curve Fitting and Suboptimization," IEEE Trans. Auto. Control, vol. AC-13, pp. 719-721, December 1968.

17. Cohn, M. and Ott, G., "Design of Adaptive Procedures for Fault Detection and Isolation," IEEE Trans. Rel., vol. R-20, No. 1, pp. 7-10, February 1971.

18. Cooley, J. W., Lewis, P. A. W., and Welch, P. D., "Application of the Fast Fourier Transform to Computation of Fourier Integrals, Fourier Series, and Convolution Integrals," IEEE Trans. Audio Electroacoust., vol. AW-15, pp. 79-84, June 1967.

19. Crellin, G. L. (editor) "Special Issue on Bayesian Reliability Techniques," IEEE Trans. Rel., vol. R-21, pp. 127-194, August 1972.

20. Elliott, T. W., "Accelerated Testing of Air-to-Air Guided Missiles," Proceedings, 1973 Annual Reliability and Maintainability Symp., pp. 105-108, January 1973.

21. Evans, R. A., "The Principle of Minimum Information," IEEE Trans. Rel., vol, R-18, No. 3, pp. 87-90, August 1969.

22. Feller, W., An Introduction to Probability Theory and Its Applications, vol. 1, John Wiley & Sons, 1968.

23. Fleming, J. L., "Relcomp: A Computer Program for Calculating System Reliability and MTBF," IEEE Trans. Rel., vol. R-20, No. 3, pp. 102-107, August 1971.

24. Gonzales, A. J., "Failure Analysis Applications of the Scanning Electron Microscope," 11th Annual Proceedings on Reliability Physics, pp. 179-184, 1973.

25. Greville, T. N. E., "Introduction to Spline Functions," Theory and Applications of Spline Functions, Academic Press, pp. 1-35, 1969.

26. Grossman, M., "Focus on Fixed Passive Components," Electronic Design, No. 12, pp. 56-67, June 1973.

27. Guttman, I., and Wilks, S. S., Introductory Engineering Statistics, John Wiley & Sons., Inc., 1967.

28. Hamming, R. W., Numerical Methods for Scientists and Engineers, McGraw-Hill Book Co., 1962.

29. Helms, H. D., and Thomas, J. B., "Truncation Error of Sampling-Theorem Expansions," Proc. IRE, vol. 50, pp. 179-184, February 1962.

30. Herman, F. H. and Kobylarz, T. J., "Statistically Based Interpolation and Extrapolation in Automated Testing Systems," Sixth Asilomar Conference on Circuits and Systems, pp. 434-437, 1972.

31. Herman, F. H. et al., "Statistics for Improvements of IC Process Control," Seventh Asilomar Conference on Circuits and Systems, Proceedings published by Western Periodicals Co., North Hollywood, CA, November 1973.

32. Herr, E. A., "Techniques for the Control of Integrated Circuit Quality and Reliability," General Electric Semiconductor Prod. Dept., Syracuse, NY, Rept. No. IR-8-232.

33. Herr, E. A., et al., "The TEG-A Test Element for the Control of Quality and Reliability of Integrated Circuits," General Electric Semiconductor Prod. Dept., Syracuse, NY, Rept. No. R 68EGP1.

34. Herskowitz, G. J., Computer-Aided Integrated Circuit Design, McGraw-Hill Book Co., 1968.

35. Highleyman, W. H., "Linear Decision Functions, with Applications to Pattern Recognition," Proc. IRE, vol. 50, pp. 1501-1514, June 1962.

36. Hill, W. J., et al., "A Joint Design Criterion for the Dual Problem of Model Discrimination and Parameter Estimation," Technometrics, vol. 10, No. 1, pp. 145-160, February 1968.

37. Hill, W. J. and Hunter, W. G., "A Note on Designs for Model Discrimination: Variance Unknown Case," Technometrics, vol. 11, No. 2, pp. 396-400, May 1969.

38. Jaynes, E. T., "Prior Probabilities," IEEE Trans. Sys. Science and Cybernetics, vol. SSC-4, No. 3, pp. 227-241, September 1968.

39. Kobylarz, T. J., "Adaption In Linear and Nonlinear Threshold Models of Neurons," IEEE Trans. Info. Theory, November 1968.

40. Kooi, C. F., "Physical Analysis of Stress Testing for Failure of Electronic Components," IEEE Trans. Rel., vol. R-17, No. 2, pp. 80-84, June 1968.

41. Kruskal, J. B., "Linear Transformations of Multivariate Data to Reveal Clustering," In Shepard, R. N., et al., Multidimensional Scaling, vol. 1, Seminar Press, pp. 179-191, 1972.

42. Kussmaul, K., "Protection Against Assuming the Wrong Degree in Polynomial Regression," Technometrics, vol. 11, No. 4, pp. 677-682, November 1969.

43. Levine, M. D., "Feature Extraction: A Survey," Proc. IEEE, vol. 57, No. 8, pp. 1391-1407, August 1969.

44. Lucky, R. W., "Adaptive Redundancy Removal in Data Transmission," Bell System Tech. J., vol. 47, pp. 549-573, April 1968.

45. McAleer, H. T., "A Look at Automatic Testing," IEEE Spectrum, pp. 63-78, May 1971.

46. Mengert, P. H., "Solution of Linear Inequalities," IEEE Trans. Comp., vol. C-19, No. 2, pp. 124-131, February 1970.

47. Milek, J. T., "Reliability of Hybrid Microelectronic Circuits," Hughes Aircraft Company, Rept. AD 689 757, March 1969.

48. Miller, F., "Test Sites for Microelectronics," 1969 Western Electronic Show, San Francisco, CA, pp. 1-10, August 1969.

49. Mood, A. M. and Graybill, F. A., Introduction to the Theory of Statistics, McGraw-Hill Book Co., 1963.

50. Morgan, B. W., An Introduction to Bayesian Statistical Decision Processes, Prentice-Hall, Inc., pp. 1-25, 1968.

51. Nering, E. D., Linear Algebra and Matrix Theory, John Wiley & Sons, Inc., Chap. 1, 1967.

52. Nilsson, N. J., Learning Machines, McGraw-Hill Book Co., 1965.

53. Papoulis, A., The Fourier Integral and Its Applications, McGraw-Hill Book Co., 1962.

54. Pokrovsky, F. N., "On Reliability Prediction by Pattern Recognition," Proc., 1972 Annual Rel. and Maintainability Symp., pp. 367-375, January 1972.

55. Pollack, S. M., "A Bayesian Reliability Growth Model," Technical Report AD 663 279, W. S. Naval Postgraduate School, Monterey, CA, June 1967.

56. Radzyner, R., and Bason, P. T., "An Error Bound for Lagrange Interpolation of Low-Pass Functions," IEEE Trans. Info. Theory, vol. IT-16, No. 5, pp. 669-671, September 1972.

57. Ralston, A., A First Course in Numerical Analysis, McGraw-Hill Book Co., 1965.

58. Regulinski, T. L., "Elements of Information Theoretic Models: Tutorial Introduction," IEEE Trans. Rel., vol. R-18, No. 3, pp. 82-87, August 1969.

59. Sage, A. P., "Error Analysis, Sensitivity, and Prior Statistics," IEEE Trans. Ed., vol. E-12, No. 2, pp. 117-124, June 1969.

60. Sarkar, T. K., "An Exact Lower Confidence Bound for the Reliability of a Series System Where Each Component has an Exponential Time to Failure Distribution," Technometrics, vol. 13, No. 3, pp. 535-546, August 1971.

61. Sasieni, M., et al., Operations Research - Methods and Problems, John Wiley & Sons, Inc., 1967.

62. Schlegel, E. S., "Study of Failure Modes of Multilevel Large Scale Integrated Circuits," Philco-Ford Corp., Microelectronics Div., Blue Bell, PA, Rept. No. NASA-CR-99100, 1969.

63. Schumaker, L. L., "Some Algorithms for the Computations of Interpolating and Approximating Spline Functions," Theory and Applications of Spline Functions, Academic Press, pp. 87-102, 1969.

64. Schwarz, R. J. and Friedland, B., Linear Systems, McGraw-Hill Book Co., 1965.

65. Sebestyen, G. S., Decision Making Processes in Pattern Recognition, The Macmillian Company, 1962.

66. Seshu, S. and Waxman, R., "Fault Isolation in Conventional Linear Systems - A Feasibility Study," IEEE Trans. Rel., vol. R-15, No. 1, pp. 11-16, May 1966.

67. Shannon, C. E., "A Mathematical Theory of Communication," The Bell System Technical Journal, vol. 27, pp. 379-423 and 623-656, 1948.

68. Shroen, W. H., et al., "Reliability Improvement by Process Control," 10th Annual Proc. Reliability Physics, (IEEE), pp. 42-48, 1972.

69. Simkins, D. J., "Weibull, Bayes, and Entropy," Proc. 1972 Annual Rel. and Maintainability Symp. (IEEE), pp. 42-48, 1972.

70. Smith, F. W., "Pattern Classifier Design by Linear Programming," IEEE Trans. Computers, vol. C-17, No. 4, pp. 367-372, April 1968.

71. Stern, T. E., Theory of Nonlinear Networks and Systems, Addison-Wesley, Reading, MA, 1965.

72. Taylor, D. S., "A Reliability and Comparitive Analysis of Two Standby System Configurations," IEEE Trans. Rel., vol. R-22, pp. 13-19, April 1973.

73. Thomas, G. B., Calculus and Analytic Geometry, Addison-Wesley, Reading, MA, 1962.

74. Tsokos, C. P., "A Bayesian Approach to Reliability: Theory and Simulation," Proc. 1972 Annual Rel. and Maintainability Symp., (IEEE), pp. 78-87, 1972.

75. vanBeek, H. W., "MOS/LSI Circuit Design: Designing-in Reliability," 10th Annual Proc. Rel. Physics, (IEEE), pp. 36-41, 1972.

76. Wald, A., "On the Efficient Design of Statistical Investigations," Ann. Math. Stat., vol. 14, pp. 134-140, 1943.

77. Welker, E. L., "Some Statistical Techniques Useful in System Aging Studies," Proc. 1973 Annual Rel. and Maintainability Symp., (IEEE), pp. 10-16, 1973.

78. Yen, J. L., "On Nonuniform Sampling of Bandwidth-Limited Signals," IRE Trans. Circuit Theory, vol. CT-3, pp. 251-257, December 1956.