Incorporation of Statistical Information for Polynomial and Transcendental Interpolation

FREDERIC H. HERMAN, STUDENT MEMBER, IEEE, AND THADDEUS J. KOBYLARZ, MEMBER, IEEE

Abstract—One of the principal considerations in the implementation of automated testing systems is the "cost" of accumulating test data. Because testing is controlled and recorded by a digital computer, only discrete data points are gathered. Also, an elapsed waiting time is necessary in order to establish the test conditions prior to performing a measurement. Due to these factors, the number of measurements has a major influence on the cost of testing devices automatically. It now follows that automated testing of continuous systems will entail interpolation. This paper treats the situation wherein statistical information is substituted for some of the deterministic data in interpolation and extrapolation procedures. The principle advantage with this approach is that a reduction in the number of measurements required to characterize a device through interpolation and extrapolation is achieved. A consequence of utilizing the statistical data is that the device function is said to lie within statistical bounds. In addition to introducing the proposed approach, this paper describes practical considerations in obtaining the required statistical information. A hypothetical example is given in order to illustrate the application of the proposed techniques.

I. Introduction

A PROBLEM of growing concern is the problem of testing complex devices by automated testing systems. A significant percentage of these devices must be tested in a continuous nonlinear instantaneous mode. That is, the device is characterized by such an input–output response function. In addition to the inputs that the device is intended to respond to, environmental variables such as temperature and humidity may cause significant functional variations, which must be known for design.

Practical considerations usually require that measurements be discrete. That is, at some specific input x_0 the response of the device $f(x_0)$ is measured. Often this limitation of discrete test points results from the fact that automatic testing systems driven by digital computers generate the tests and evaluate the measurements. It is natural to describe the input–output functions by means of interpolation functions using the discrete test points as data.

Exact-fit interpolation procedures are well suited to automated testing systems. They are easily mechanized and the theory of their application is well established in the literature. Practical application of this theory makes use of *a priori* knowledge of the suitability of a weighted sum of standard functions to describe the actual device function. For example, it may be known that a certain order polynomial describes the function within a given interval. Usually an error bound can be estimated from information about some order derivative of the function [1].

When test measurements may be readily obtained, one may simply gather sufficient data to characterize a device via interpolation. However, when such information is costly to obtain and many measurements are required, it is advisable to consider alternative information that may be available [2]. The costs are usually associated with the time requirements or the difficulties in simulating inputs for measurements or in the effects such tests have on the device. For example, accurate testing requires that input transients be avoided. In thermal and mechanical perturbations, the transients may be of significant duration and the delay for steady state contributes significantly to testing time cost.

This paper considers those situations in which the number of measurements required to characterize a particular device is to be minimized. The approach considered is to replace some of the deterministic measurements normally required for interpolation with statistical information. The "statistical data" are in the form of tolerance limits [3] obtained from the known distribution of the device function $f(x)$ at the value of x corresponding to the data point. Since deterministic data values are replaced with statistical ranges of values, the utilization of such information results in two bounding functions, between which the device function is said to lie. This statistical bounding of the device function corresponds to some confidence level which represents an uncertainty over and above the usual error bound of conventional interpolation. It is noted that the utilization of such statistical data has been suggested for nonconventional interpolation and extrapolation elsewhere [4]. This nonconventional scheme uses a differential extrapolation analysis to obtain divergent bounds for a function near a deterministic data point, and it does not assume a particular order of approximation.

The assumptions and background required for this paper are presented in the next section. Following this, the basic approach used in this paper is formalized for the simple case of a single deterministic data value replaced by statistical data. This approach is then generalized to allow more than one point to be replaced. Certain problems are seen to arise from the need to combine several sets of tolerance limits. These problems and the practical aspects of earlier assumptions are discussed. A hypothetical example is given in order to illustrate the application of the proposed techniques. Several implementation strategies are suggested.

II. Assumptions and Background

The functions to be evaluated correspond to input–output or response functions of physical systems. This section will

Manuscript received April 30, 1973; revised July 20, 1973. This paper was accepted for presentation at the 1973 Electrical and Electronic Measurement and Test Instrument Conference (EEMTIC), Ottawa, Ont., Canada, May 15–17, however, it was not formally presented.

The authors are with the Stevens Institute of Technology, Hoboken, N.J.

describe several assumptions with respect to these response functions and the systems they represent. Also, since interpolation and extrapolation techniques will be the basis for the function evaluation, several conventional methods will be indicated. The methods discussed will be those which are well suited to automated testing procedures and also the statistical approach to be developed. The preliminaries of this section will be useful in the development of subsequent sections.

The functions considered are assumed to be instantaneous, continuous, and allowably nonlinear. The independent variable is referred to as the input x and the functional value $f(x)$ is referred to as the output. The function $f(x)$ is to be also viewed as a random variable (RV) f_i, where the index i denotes that this RV corresponds to the unknown functional value at x_i. That is for any value of x in [a, b], the probability density function (pdf) of $f(x)$ is assumed known.[1] In this case x is a deterministic parameter of the pdf $g(x, u)$, and one can write

$$p\{f(x) \leqslant z\} = \int_{-\infty}^{z} g(x, u)\, du. \tag{1}$$

Two types of exact-fit interpolation formulas will be considered. 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. 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 [1]. 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 one moves away from a data point. 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 paper does not consider the conventional techniques related to obtaining error bounds for polynomial interpolation. It is assumed that the reader is familiar with the considerable amount of literature available in this area. Nor is the related problem of determining an adequate order for the interpolation discussed here, except to say that such information must result from prior experience with the functions to be approximated.

[1] It will be seen later that the pdf need not be known precisely for practical application of subsequent developments. Nor does one need to know $g(x, u)$ at more than a finite number of points in $[a, b]$. The relaxation of the required knowledge of the distribution of $f(x)$ will be discussed in a later section.

The polynomial interpolation formula discussed in this paper is the Lagrange form [1]

$$y(x) = \sum_{i=1}^{N} L_i(x) \cdot f(x_i) \tag{2}$$

where

$$L_i(x) = \prod_{\substack{j=1\\ j\neq i}}^{N} (x - x_j) / \prod_{\substack{j=1\\ j\neq i}}^{N} (x_i - x_j). \tag{3}$$

This form requires N values of $f(x)$ at $x_i, i = 1, \cdots N$, and produces an $N - 1$ order polynomial. This form uses no derivative data and is easily mechanized, allowing for dynamic selection of the x_i during the testing. Other formulas can be used, but this particular form is useful in discussing the proposed approach.

The second type of interpolation formula considered is a transcendental formulation. It is useful when $f(x)$ corresponds to a band-limited function [5], in the sense that it contains no frequencies greater than a cutoff frequency, ω_c. Based on the well-known sampling theorem [5], this interpolation uses a finite number of equispaced samples. Several authors [6], [7] 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 $0 < r < 1$. The cardinal interpolation formula is then

$$y(x) = \sum_{m=K-N}^{K+N} f(m/2\omega_c) \cdot \operatorname{sinc}(2\omega_c[x - m/2\omega_c]), \quad .0 < N < \infty \tag{4}$$

where K is an integer chosen such that

$$2\omega_c x - \tfrac{1}{2} \leqslant K(x) \leqslant 2\omega_c x + \tfrac{1}{2} \tag{5}$$

and where

$$\operatorname{sinc} z \equiv \sin(\pi z)/\pi z. \tag{6}$$

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.

Other interpolation formulas, which utilize the assumption of a band-limited function are possible. For example, a similar formula to (4) is possible if both values of $f(x)$ and its derivative are used [6]. It is interesting to note that an error bound can be determined for Lagrange interpolation when $f(x)$ is band limited [8], providing the data are nearly equispaced. Also, the practical determination of the band-limited properties of a function must result from prior experience with the function.

The two types of interpolation formulas considered allow two useful observations to be made. Equations (2) and (4) are summations of data multiplied by interpolation weighting functions.[2] 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

[2] Yen [9] distinguishes these from conventional polynomial coefficients by the term "composing functions."

subsequent developments. It will be useful to represent both (4) and (6) in what follows by the general form

$$y(x) = \sum_{i=1}^{N+1} W_i(x) \cdot f(x_i) \tag{7}$$

where the required substitutions and constraints are determined by the particular form of interpolation formula represented by (7).

III. Single Tolerance Interval

In this section, it is considered how probabilistic information of the form previously described can be used in exact-fit interpolation formulas such as (4) and (6). Suppose that N data points are available for the Nth-order polynomial interpolation of a function that is to be approximated. In place of the $(N+1)$th data point normally required $f(x_{N+1})$ construct tolerance limits [3], [10] using the pdf of $f(x_{N+1})$ [see (1)]. Define a tolerance interval (TI), $[C_1, C_2]$, by the probability

$$p\{f(x_{N+1}) \in [C_1, C_2]\} = 2\alpha, \qquad 0 \leqslant \alpha \leqslant \tfrac{1}{2}. \tag{8}$$

Since (8) does not uniquely determine C_1 and C_2, it is necessary to choose an additional constraint. Usually the TI is defined symmetrically about the distribution mean.

Once the TI of $f(x_{N=1})$ has been obtained, the approach is to use this range of values for the normally required $f(x_{N+1})$ value. By considering $f(x_{N+1})$ now as a RV f_{N+1}, it is appropriate to rewrite (7) as a RV equation. That is

$$y(x) = A(x) + B(x) f_{N+1} \tag{9}$$

where x is now a parameter of the deterministic constants $A(x)$ and $B(x)$. It is clear that if (8) holds, then the RV $y(x)$ is bounded according to[3]

$$p\{A(x) + B(x) \cdot C_1(x) \leqslant y(x) \leqslant A(x) + B(x) \cdot C_2(x)\} = 2\alpha. \tag{10}$$

This result defines two curves to be constructed, between which the function $f(x)$ is said to lie with a probability of 2α. These bounds do not include the conventional interpolation error which effectively widens the bounding. Fig. 1 illustrates a typical interpolation in which a single deterministic data value has been replaced by a TI. It is noted that the determination of whether to use C_1 or C_2 for minimization or maximization of $y(x)$ at a specific x is greatly simplified for exact-fit interpolation formulas. The previous observation that the sign of the weighting functions of (7) can change only at a data point indicates that the sign of $W_{N+1}(x)$ need be determined only once for each interval of x between consecutive data points.

The result of (10) indicates that it is possible to reduce the number of measurements on a particular system provided the density function is available at certain parameter values. In practice, one may precompute the values of C_1 and C_2 at specific values of x, or may determine an interpolation formula for C_1 and C_2 in terms of x at the specific confidence levels of interest. The latter approach is more suited for an adaptive procedure in which the probabilistic data points may be selected dynamically. Precomputing these functions allows for faster processing and for a compact storage of the statistical information.

[3](It is assumed that $B(x)$ is positive here, but for negative values the limits of (10) are interchanged. Also, C_1 and C_2 of (8) are generally functions of x as indicated in (10).)

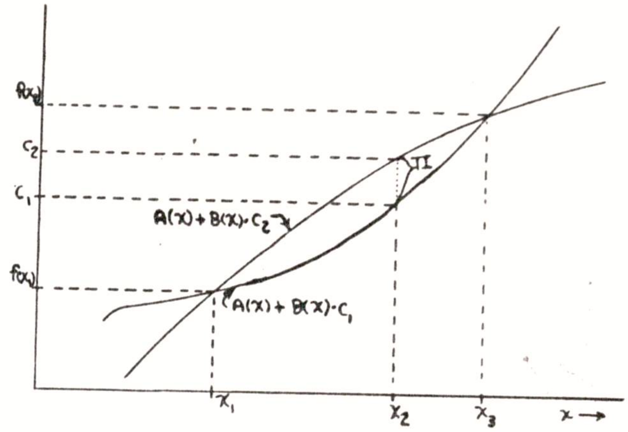

Fig. 1. Polynomial interpolation in which one data point has been replaced by tolerance interval.

IV. Incorporation of Several Tolerance Intervals

The approach introduced in the last section will now be extended for the situation in which several deterministic data points are replaced by the TI. By differentiating (7) with respect to $f_i(x)$, one observes that

$$\frac{\partial y(x)}{\partial f_i(x)} = W_i(x). \tag{11}$$

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 (7) are deterministic data, and the remaining $N+1-k$ values are the RV f_i, represented by the TI. That is,

$$y(x) = \sum_{i=1}^{k} W_i(x) \cdot f(x_i) + \sum_{i=k+1}^{N+1} W_i(x) \cdot f_i$$

and

$$p\{C_{i,1} \leqslant f(x_i) \leqslant C_{i,2}\} = 2\alpha, \qquad i = k+1, \cdots, N+1.$$

The constants $C_{i,1}$ and $C_{i,2}$ represent the lower and the upper tolerance limits for f_i, respectively. Then the "composite tolerance interval" (composite TI) is defined by the expressions

$$D_1 \equiv \sum_{i=1}^{k} W_i(x) \cdot f(x_i) + \sum_{i=k+1}^{N+1} [\min (W_i(x) \cdot C_{i,1}, W_i(x) \cdot C_{i,2})] \tag{12a}$$

and

$$D_2 \equiv \sum_{i=1}^{k} W_i(x) \cdot f(x) + \sum_{i=k+1}^{N+1} [\max (W_i(x) \cdot C_{i,1}, W_i(x) \cdot C_{i,2})] \tag{12b}$$

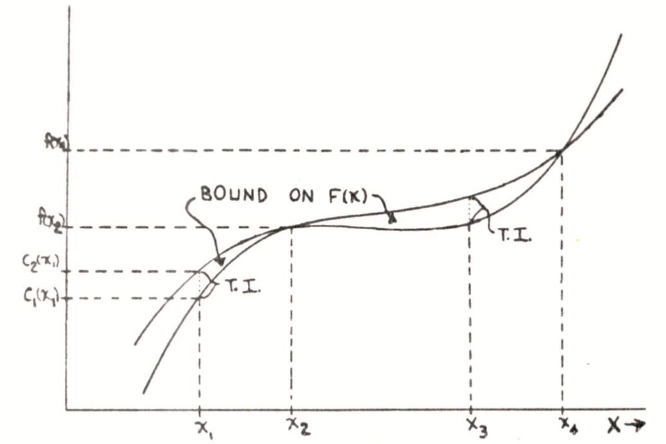

Fig. 2. Polynomial interpolation in which several data points have been replaced by tolerance intervals.

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 $[D_1, D_2]$.

Inspection of (12) reveals that one may consider each of the TI substitutions independently in obtaining a composite TI bound on $f(x)$. Fig. 2 shows a situation in which two TI's have been used in a polynomial interpolation.

While a straightforward procedure has just been indicated for the determination of statistical bounds on $f(x)$, it remains to be shown that the composite TI 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 (7) 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. It will be further assumed that the TI are defined symmetrically about the distribution means, as suggested previously. Under these assumptions, it is possible to speculate on the confidence level associated with the composite TI. It is shown elsewhere [10] that for a large class of pdf's, the confidence level corresponding to the composite TI is at least 2α. This class includes the Laplace, the Gaussian-normal, and the Cauchy pdf's. For other pdf's, which are monomodal and are symmetric or near symmetric, a nearly analytic argument [10] may be used to show that the composite tolerance level is at least 2α. The development of the argument and its application is lengthy and, therefore, omitted here.

The requirement that the pdf of $f(x)$ be known explicitly over the interpolation interval of x is usually difficult to satisfy in practical situations. However, the determination of a likely candidate for the distribution of $f(x)$ and estimates of its parameters are often possible. In order to obtain statistical information, one must be prepared to make an initial concerted measurement effort on a large number of similar randomly selected devices. 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. Since one often uses predetermined data points, the parameters of the device pdf need be determined only at those points where the TI are to be substituted for deterministic data. A further simplification results when one chooses the confidence level ahead of time. In this case, only the TI need be estimated using standard-lot sampling techniques [3].

A final consideration of this section is that of the statistical independence of the RV f_i, which are represented by their respective TI's. 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. In this case, the two values are highly correlated. If the distance between these two points is increased, then it may become reasonable to make the assumption of statistical independence, although the existence of some correlation will usually result in pessimistic statistical bounds.

Sometimes it is possible to remove a form of correlation. By way of illustration, consider the class of devices which have input-output functions of the form

$$f(x) = a_0 + \sum_{k=1}^{N} a_k x^k$$

and suppose that the constants a_k are very nearly the same for all the devices that are being represented by $f(x)$, except for a_0. For this constant, assume that there is great variation from one device to another. This wide variation may represent different output "bias" levels for electronic amplifiers for example. One may easily see that for any particular device, the values of $f(x)$ are highly correlated by the value of a_0. To remove this correlation, a_0 is initially determined for each component tested and then subtracted from function value to give the new function

$$\bar{f}(x) = f(x) - a_0$$

The distribution of $\bar{f}(x)$ is gathered, and then interpolation is performed on this "scaled" function. Thus the correlation due to a_0 is avoided. Other types of normalization can be used for other types of functional correlation.

IV. A Hypothetical Example

Consider a temperature-to-voltage transducer as the device to be tested. The device consists of a small quartz crystal and a monolithic circuit. The circuit consists of an oscillator of which the frequency is controlled by the crystal, and a network which limits the oscillator output and then converts it to a dc voltage so that the output is approximately proportional to the frequency. The crystal has a temperature coefficient, which results in a change in the resonant frequency proportional to the temperature.

Both the crystal and the circuit are separately "calibrated" by laser trimming. The device is then constructed as a hybrid device. 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. It has been found from experience with a large number of these devices, that a second-order polynomial interpolation over intervals of 10°

results in negligible truncation error. The device function must be known over a 50° range. One hundred percent inspection of the devices is desired. Using conventional interpolation, each device must be sampled at 5° intervals, requiring measurement at a total of 11 temperatures. 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.

Because of the large number of devices produced, it is desirable to reduce testing time. From data gathered from conventional testing of several thousand devices, TI at each of the 11 test temperatures has been determined corresponding to a confidence level of 0.99 (99 percentile). The total temperature range is to be broken up into subintervals of 10° as before. However, the midpoint measurement of each subinterval is to be replaced by the TI corresponding to this temperature. By using this statistical data, five of the 11 temperature simulations are eliminated. The bounding of the device function corresponds to the single TI case of Section III and Fig. 1.

VI. Implementation Strategies

Application of the preceding developments to practical problems requires consideration of additional questions. These include the following: where to locate data points, which points and how many should be statistically represented, and when to make additional measurements. These questions may be partially dealt with in terms of conventional interpolation theory. This section considers strategies beyond the conventional techniques which may be studied elsewhere [1].

It is assumed that there exists an *a priori* criteria for deciding the acceptability of a device function. 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. In the example given in the last section, this might correspond to a limit on the relative deviation from linearity.

In deciding where to locate data points, one may be constrained by equispaced requirements of most cardinal interpolation formulatons. 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 alluded to previously. One would choose these data points to 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. 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.

Depending on the allowable flexibility of the testing system, various adaptive strategies can be implemented. 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, the statistical data point could be moved to the overlap region in order to determine whether use of the TI corresponding to this point would result in the overlap being avoided.

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.

References

[1] R. W. Hamming, *Introduction to Applied Numerical Analysis*. New York: McGraw-Hill, 1971.

[2] F. H. Herman, T. J. Kobylarz, A. J. Graf, and P. C. Calella, "Monitoring IC process parameters with statistically enhanced test pattern data," to be published.

[3] A. H. Bowker and G. J. Lieberman, *Engineering Statistics*. Englewood Cliffs, N.J.: Prentice-Hall, 1964.

[4] F. H. Herman and T. J. Kobylarz, "Statistically based interpolation and extrapolation in automated testing systems," presented at the 6th Asilomar Conf. Circuits and Systems. Nov. 1972.

[5] A. Papoulis, *The Fourier Integral and Its Applications*. New York: McGraw-Hill, 1962.

[6] H. D. Helms and J. B. Thomas, "Truncation error of sampling theorem expensions," *Proc. IRE*, vol. 50, pp. 179–184. Feb. 1962.

[7] J. L. Brown, "Bounds for truncation error in sampling expansions of band limited signals," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 440–444, July 1969.

[8] R. Radzyner and P. T. Bason, "An error bound for Lagrange interpolation of low-pass functions," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 669–671, Sept. 1972.

[9] J. L. Yen, "On non-uniform sampling of bandwidth-limited signals," *IRE Trans. Circuit Theory*, vol. CT-3, pp. 251–257, Dec. 1956.

[10] F. H. Herman and T. J. Kobylarz, "Bounds on linear combinations of independent random variables," to be published.