ments such as N, DN, E, and EN a large amount in the same direction. Thus B is an interesting impairment, appearing to affect the dimensional position of impairments added before it more than it does the original picture.

Although we have not yet investigated this point experimentally, it appears likely from the above observations that the dimensional character of any impairment combination will be affected by the order in which the impairments are introduced. Impairments which are physically independent in nature appear to interact when introduced into the same television picture.

Conclusions

Multidimensional scaling analysis indicated that observers participating in our experiments used four dimensions in their characterization of the impairments. We interpreted the dimensions as 1) overall picture clarity, 2) a distinction between overlay impairment and object impairment, 3) the amount of purely spatial or stationary overlay patterning, and 4) the amount of spatiotemporal or moving overlay patterning. Our knowledge of whether these dimensions are used more generally by observers when judging impaired television pictures must await further experimentation with a wider selection of impairments than we ourselves employed.

Something of the individual character and interaction of the four basic impairments was revealed by the analysis. The primary effect of echo and noise was to provide an overlay pattern on the picture, stationary in the case of echo and moving or scintillating in the case of noise. The character of the differential quantizing noise was more complex, having components along three dimensions, those of overall clarity, moving overlay patterning, and object-overlay distortion. The character of the band-limitation impairment appeared to be that of a modifier of the character of previous impairments, its effect when introduced singly being mainly along the clarity dimension, with a smaller component along the object-overlay dimension. The interaction between the four impairments was complex but interpretable.

The study points to a possible binary classification of television impairments depending on whether they distort objects in a picture or whether they mask these objects by means of an overlay pattern. A further subdivision of masking impairments into moving and stationary types is suggested.

Certain fundamental factors or dimensions appear to be emerging from the application of multidimensional scaling to acoustical and visual communication systems. Judgments of similarity and preference/quality on both analog and digital impairments introduced singly and multiply in both sensory modalities appear to yield comparable results.

Acknowledgment

The authors would like to record their thanks to Mr. J. A. Allnatt, Dr. J. E. Thompson, Ms. G. Reid, and Mr. T. White of the British Post Office Research Department for very helpful discussions and to Dr. D. J. Goodman and Ms. B. J. McDermott of Bell Laboratories for invaluable advice and assistance with multidimensional scaling programs. The experimental work was conducted with the technical assistance of Mr. M. Brown and Mr. T. Quinlan of the Visual Systems Research Laboratory, University of Essex.

References

[1] N. W. Lewis and J. W. Allnatt, "Subjective quality of television pictures with multiple impairments," *Electron. Lett.*, vol. 1, pp. 187–188, 1965.

[2] J. R. Cavanaugh and A. M. Lessman, "Subjective effects of differential gain and differential phase distortions in NTSC color television pictures," *J. Soc. Motion Pict. Telev. Eng.*, vol. 80, pp. 614–619, 1972.

[3] C. A. Siocos, "Subjective impairment of television pictures—Considerations regarding the summability of impairment units," *J. Soc. Motion Pict. Telev. Eng.*, vol. 81, pp. 917–920, 1972.

[4] J. W. Allnatt, "Opinion-rating model applied to television-transmission-system studies," in *Proc. IEE*, vol. 122, no. 8, pp. 769–774, 1975.

[5] R. N. Shepard, A. K. Romney, and S. B. Nerlove, *Multidimensional Scaling*. New York and London: Seminar Press, 1972.

[6] E. J. Jackson, "Scaling techniques for subjective judgments of image quality," *J. Soc. Motion Pict. Telev. Eng.*, vol. 83, pp. 891–894, 1974.

[7] *CCIR Recommendation 500*, "Method for the subjective assessment of the quality of television pictures," XIIIth Plenary Assembly, Geneva, 1974.

[8] W. W. Cooley and P. R. Lohnes, *Multivariate Data Analysis*. New York: Wiley, 1971.

[9] D. J. Carroll and J. J. Chang, "Analysis of individual differences in multidimensional scaling via an N-way generalization of 'Eckart–Young' decomposition," *Psychometrika*, vol. 35, pp. 283–319, 1970.

[10] D. J. Carroll and M. Wish, "Models and methods for three-way multidimensional scaling," in *Contemporary Developments in Mathematical Psychology, Vol. II*, D. H. Krawtz, R. C. Atkinson, D. R. Luce, and P. Suppes, Eds. San Francisco: Freeman, 1974.

[11] W. S. Torgerson, *Theory and Methods of Scaling*. New York: Wiley, 1972.

[12] B. J. McDermott, "Multidimensional analyses of circuit quality judgments," *J. Acoust. Soc. Amer.*, vol. 45, pp. 774–781, 1969.

[13] B. J. McDermott, C. Scagliola, and D. Goodman, "Perceptual and objective evaluation of speech processed by adaptive differential PCM," *Bell Syst. Tech. J.*, vol. 57, pp. 1597–1618, 1978.

Failure Prediction for an On-Line Maintenance System in a Poisson Shock Environment

K. S. LU AND R. SAEKS, FELLOW, IEEE

***Abstract*—A failure prediction algorithm for application in a periodic on-line maintenance system operating in a Poisson shock environment is described. The system under test is measured at periodic maintenance intervals with the data derived therefrom being used to estimate system lifetime and determine an optimal replacement time. The resultant algorithm is simulated and compared with various fixed replacement schedules.**

I. Introduction

Although considerable effort has been expended during the past decade to develop techniques for fault detection and diagnosis in both analog and digital electronic circuits [10], little attention has been given to the possibility of formulating algorithms for fault prediction. To accurately predict a fault, a device must be tested at periodic maintenance intervals. If the device fails or does not operate correctly, it is replaced immediately. The device may be assumed good if its characteristics are in tolerance. However, if the characteristics are slightly off nominal but the device still operates correctly, one can attempt to predict if the device will fail before the next scheduled maintenance interval. If device failure is predicted, it can be replaced before failure occurs as part of planned preventative maintenance.

Manuscript received April 3, 1978; revised September 18, 1978 and February 1, 1979. This work was supported in part by the Joint Services Electronics Program at Texas Tech University under ONR Contract 76-C-1136.

R. Saeks is with the Department of Electrical Engineering, Texas Tech University, Lubbock, TX 79409.

K. S. Lu is with the Texas Instruments Incorporated, Dallas, TX.

0018-9472/79/0600-0356$00.75

With the advent of the low-cost microprocessor, on-line fault prediction is possible and practical [9]. For this purpose, a curve fitting algorithm for on-line fault prediction was first introduced by Saeks, Liberty, and Tung [11]–[13] in 1975. The disadvantage of this algorithm, however, is that the second-order polynomial model employed is too simple to describe the aging curve of a real-world component. Employing the Poisson-shock model for the wear process introduced by Esary, Marshall, and Proschan [1], [2], [6], another curve fitting fault prediction algorithm which overcomes these disadvantages is discussed in the present paper [9].

In the following section a model for the failure dynamics of a system component parameter is formulated. Here it is assumed that the failure is due to the component being subjected to a sequence of Poisson distributed shocks [3], [7], with the measurable parameter being controlled by an unknown difference equation whose underlying discrete "component time" process is defined by the number of shocks to which the component has been subjected. Since both the failure dynamics (i.e., the difference equation) and the relationship between "component time" and real time are unknown, our failure model is doubly stochastic. The third section of the paper is devoted to the formulation of an algorithm for estimating the component failure dynamics, and its "lifetime" is defined to be the number of shocks required to cause component failure. This is followed by the formulation of an "optimal" replacement theory wherein the optimal real time at which to replace a component is computed in terms of its estimated "lifetime." Finally, the results of a simulation of the algorithm in both an ideal and noisy environment are presented and compared with the simulated performance for several fixed replacement schedules.

II. Failure Dynamics

Let $C(N)$ represent values of a particular component parameter, where the "component time" N denotes the number of shocks the component has received. It is assumed that the drifting parameters can be described by a first-order[1] difference equation of the form:

$$C(N+1) = C(N) - a_0 - a_1 N - a_2 N^2 - \cdots - a_h N^h$$
$$C(0) = 1. \tag{1}$$

Here the coefficients and order of the "forcing polynomial" are assumed to be unknown and must be estimated as part of the fault prediction process. A little algebra together with the standard recursive formula for solving a difference equation will reveal that

$$C(N) = 1 - \sum_{j=0}^{N-1} \sum_{i=0}^{h} a_i j^i. \tag{2}$$

Now, if the tolerance limit for the component parameter is normalized to $C = 0$, we may define the lifetime of the component to be the smallest integer N for which $C(N) \leq 0$. This integer, which we denote by L, then represents the number of shocks necessary to cause the component to fail.

Since the failure model of (1) is dependent on "component time," i.e., the number of shocks the component has received, rather than real time, it remains to define the relationship between "component time" and real time. Following common practice in reliability theory [1], we assume that this relationship is determined by a Poisson process. Indeed, this is the unique point process which has the scaling properties required for such an application [3]. Here the probability of N shocks occurring in the time interval t is

$$P_N(t) = e^{-kt} \frac{(kt)^N}{N!}, \qquad N = 0, 1, 2, \cdots \tag{3}$$

where k is a given constant representing the average number of shocks per unit time. Therefore, (kt) is the average number of shocks in the time interval t.

III. Estimation of Failure Dynamics and Lifetime

In a periodic maintenance system, the performance of a component is measured at each maintenance interval nT. That is to say, $(C_1, C_2, \cdots, C_g)$ is the performance data taken at maintenance times $(T, 2T, \cdots, gT)$. The estimation problem can be stated as, "Given performance data $(C_1, C_2, \cdots, C_g)$, T and k, estimate the unknown constants $(a_0, a_1, \cdots, a_h)$ of the failure dynamics." Since it is assumed that the system is subjected to Poisson shock with constant k, the expected number of shocks in each maintenance interval is kT.[2] As such, if we assume that C_m is the value of the component parameter at $N = mkT$, then upon substituting $C_m = C(mkT)$ in (2), we obtain

$$\sum_{j=0}^{mkT-1} a_0 j^0 + \sum_{j=0}^{mkT-1} a_1 j^1 + \cdots + \sum_{j=0}^{mkT-1} a_h j^h = 1 - C_m$$

where $m = 1, 2, 3, \cdots, g$, or in the matrix form:

$$JA \triangleq \begin{bmatrix} \sum_{j=0}^{kT-1} j^0 & \sum_{j=0}^{kT-1} j^1 & \cdots & \sum_{j=0}^{kT-1} j^h \\ \sum_{j=0}^{2kT-1} j^0 & \sum_{j=0}^{2kT-1} j^1 & \cdots & \sum_{j=0}^{2kT-1} j^h \\ \vdots & \vdots & & \vdots \\ \sum_{j=0}^{gkT-1} j^0 & \sum_{j=0}^{gkT-1} j^1 & \cdots & \sum_{j=0}^{gkT-1} j^h \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_h \end{bmatrix} = \begin{bmatrix} 1 - C_1 \\ 1 - C_2 \\ \vdots \\ 1 - C_g \end{bmatrix} \triangleq Z. \tag{4}$$

Since the number of data points g is typically much greater than the order of the polynomial assumed in the failure model h, it is not expected that (4) admits an exact solution. Rather, we attempt to solve for a coefficient vector A which minimizes the error between JA and Z. In particular, if one adopts a least squares error criterion, the optimal A is given by

$$A^0 = J^{-G} Z \tag{5}$$

where J^{-G} denotes the generalized inverse of J [8]. Indeed, if as is typically the case, J has full column rank, then $J^{-G} = (J^t J)^{-1} J^t$ where t denotes matrix transposition. As such, we take $A^0 = \text{col}\,(a_0^0, a_1^0, \cdots, a_h^0)$ as our estimate of the coefficients of the difference equation characterizing the failure dynamics of our drifting parameter C as per (1).

To estimate the failure dynamics of a drifting parameter, the proper choice of the order h is, in general, quite difficult and depends upon physical considerations and engineering experience. Once h is preselected, however, coefficients to best approximate the failure dynamics can be readily computed via (5). The

[1] The concepts described herein carry over without modification to the case where the failure model is characterized by higher order difference equations. The first-order model, however, suffices to illustrate the theory and is hence used throughout the present paper.

[2] Although not theoretically necessary, we assume that kT is an integer.

accuracy of the resultant estimate, however, is highly dependent on the choice of the order h and on the number of measurements which are taken g. To find a new set of coefficients for a different combination of h and g, the entire calculation procedure is typically repeated from the very beginning, a process which is impractical in the on-line maintenance system. Fortunately, sequential refinement schemes for obtaining new sets of coefficients without repeating the entire calculation can be developed [5], [8]. As such, it is possible to sequentially update one's estimates of the parameters $a_0, a_1, \cdots, a_h$ as additional measurements are taken and/or to increase the order of the model for the failure dynamics without repetitious matrix inversion. Our algorithm for estimation of the failure dynamics underlying the measured data may thus be readily implemented on-line with the computational power presently available in today's microprocessors. The matrix algebraic details of the required sequential refinement schemes are straightforward [5], [8] and readily available in the literature. As such, they will not be repeated here.

In practice, given g measurements $C_1, C_2, \cdots, C_g$ taken at maintenance intervals $T, 2T, 3T, \cdots, gT$, one sequentially estimates the coefficients of the failure dynamics $a_0, a_1, \cdots, a_h$, increasing h until no further error reduction is achieved. The resultant set of coefficients is then used in (2) to determine the component lifetime L. Upon solving the equation, the resultant estimated lifetime is found to be the smallest integer L, such that

$$\sum_{j=0}^{L-1} \sum_{i=0}^{h} a_i j^i \geq 1. \tag{6}$$

Of course, if the measured data is not decaying towards zero, i.e., the component is not failing, this inequality will have no solution, in which case we take L to be infinite [4].

IV. Replacement Theory

Although the algorithm outlined in the preceeding section yields an "optimal" estimate of the number of shocks required to cause failure, the time at which the Lth shock takes place is statistical in nature, and hence, it still remains to determine the optimal (in an appropriate sense) time at which to replace the component. One such criterion is formulated in the following. For this purpose, it is assumed that L has been computed to our satisfaction and we desire to choose a time T_r at which to replace the component as a function of L. Given L and T_r we denote the resultant probability of on-line failure (i.e., failure before T_r) by P_f. $P_r = 1 - P_f$ then denotes the probability that the component is replaced at time T_r before it fails. Similarly, we let $\tilde{T}_f$ denote the expected time to failure for those components which fail on-line, we let $\hat{T}$ denote the expected time to failure for all components, and we let T^* denote the expected time to failure for the components if they were operated to failure without replacement (i.e., $T^* = \hat{T}|_{T_r \to \infty}$). Finally, we let $f_L(t)$ denote the probability density function that the component receives the Lth shock at time t, given that the component fails on-line, whereas $p_i(t)$ represents the density function of the Poisson distribution with parameter (k), and $E_L(t)$ represents the corresponding distribution function; i.e.,

$$p_i(t) = \frac{(kt)^i}{i!} e^{-kt}, \qquad i = 0, 1, 2, \cdots \tag{7}$$

$$E_L(t) = \sum_{i=0}^{L-1} p_i(t). \tag{8}$$

With the aid of some elementary calculus [9], P_f, P_r, $\tilde{T}_f$, and $\hat{T}$, as well as their derivatives with respect to T_r, can be computed analytically. As such, upon defining an appropriate cost measure, an explicit formula for determining an "optimal" T_r (given L) can be derived. We begin with the derivation of the explicit formula for the various quantities involved in our replacement theory.

Since a component will be replaced by our algorithm if and only if it is still operating at time T_r, i.e., if it has not yet received L shocks at time T_r, the probability of replacement is just the probability of receiving less than L shocks by time T_r. We thus have the following.

Property 1:

$$P_r = E_L(T_r).$$

Proof:

$$P_r = \sum_{i=0}^{L-1} \frac{(kT_r)^i}{i!} e^{-kT_r} = \sum_{i=0}^{L-1} P_i(T_r) = E_L(T_r). \tag{9}$$

Property 2:

$$P_f = 1 - E_L(T_r).$$

Property 3:

$$\int_0^{T_r} p_i(t)\, dt = (1/k)\{1 - E_{i+1}(T_r)\}.$$

Proof:

$$\int_0^{T_r} p_i(t)\, dt = \int_0^{T_r} \frac{(kt)^i}{i!} e^{-kt}\, dt = \frac{k^i}{i!} \int_0^{T_r} t^i e^{-kt}\, dt. \tag{10}$$

Using the identity

$$\int x^m e^{ax}\, dx = e^{ax} \sum_{r=0}^{m} (-1)^r \frac{m!\, x^{m-r}}{(m-r)!\, a^{r+1}} \tag{11}$$

this becomes

$$\begin{aligned} \int_0^{T_r} p_i(t)\, dt &= \frac{k^i}{i!} e^{-kt} \sum_{r=0}^{i} (-1)^r \frac{i!\, t^{i-r}}{(i-r)!\, (-k)^{r+1}} \Bigg|_0^{T_r} \\ &= \frac{k^i}{i!} \left\{ e^{-k.0} \frac{i!}{k^{i+1}} - e^{-kT_r} \sum_{r=0}^{i} \frac{i!\, T_r^{i-r}}{(i-r)!\, k^{r+1}} \right\} \\ &= \frac{1}{k} \left\{ 1 - e^{-kT_r} \sum_{r=0}^{i} \frac{(kT_r)^{i-r}}{(i-r)!} \right\} \\ &= \frac{1}{k} \left\{ 1 - e^{-kT_r} \sum_{j=0}^{i} \frac{(kT_r)^j}{j!} \right\} \\ &= \frac{1}{k} \left\{ 1 - \sum_{j=0}^{i} p_j(T_r) \right\} \\ &= \frac{1}{k} \{ 1 - E_{i+1}(T_r) \}. \end{aligned} \tag{12}$$

Property 4:

$$f_L(t) = \frac{p_{L-1}(t)}{1/k(1 - E_L(T_r))}.$$

Proof: To derive this conditional density function we partition the interval $(0, T_r)$ into N segments of length $\Delta = T_r/N$, and we compute the probability that the Lth shock takes place in the ith time interval $((i-1)\Delta, i\Delta]$. Since this can be caused by having $L - 1$ shocks before $(i-1)\Delta$ and at least one shock in the interval $((i-1)\Delta, i\Delta]$, or by having $L - 2$ shocks before $(i-1)\Delta$ and at least two shocks in the interval $((i-1)\Delta, i\Delta]$, etc., the probability of failure in the ith interval is given by

$$\sum_{j=1}^{L} p_{L-j}((i-1)\Delta)[1 - E_j(\Delta)]$$
$$= \sum_{j=1}^{L} p_{L-j}((i-1)\Delta) \sum_{q=j}^{\infty} \frac{(\Delta k)^q}{q!} e^{-\Delta k}$$
$$= \sum_{r=1}^{\infty} \frac{1}{r!}\left[\sum_{s=1}^{r} p_{L-s}((i-1)\Delta)\right](\Delta k)^r e^{-\Delta k}. \quad (13)$$

Taking the probability density function at a point t in the interval $((i-1)\Delta, i\Delta]$ to be limiting value of this quantity divided by Δ as Δ goes to zero [7], it is observed that the terms of (13) containing powers of (Δk) greater than 1 go to zero in the limit. As such, the probability density function for the Lth shock to take place at time t is given by

$$\operatorname*{limit}_{\Delta \to 0} \frac{p_{L-1}((i-1)\Delta)(\Delta k)e^{-k}}{\Delta} = k p_{L-1}((i-1)\Delta). \quad (14)$$

Finally, since we are interested only in the conditional probability density function that the Lth shock will take place at time t, given that the component fails on-line, the quantity of (14) must be normalized, yielding

$$f_L(t) = \frac{k p_{L-1}((i-1)\Delta)}{P_f} = \frac{p_{L-1}(t)}{\frac{1}{k}(1 - E_L(T_r))} \quad (15)$$

as was to be shown.

Property 5:

$$\tilde{T}_f = \frac{L}{k}\frac{1 - E_{L+1}(T_r)}{1 - E_L(T_r)}.$$

Proof: Since T_f is the expected lifetime of the components which fail before replacement,

$$T_f = \int_0^{T_r} t f_L(t)\, dt$$
$$= \int_0^{T_r} \frac{t P_{L-1}(t)}{\frac{1}{k}\{1 - E_L(T_r)\}}\, dt$$
$$= \frac{\int_0^{T_r} t \frac{(kt)^{L-1}}{(L-1)!} e^{-kt}\, dt}{\frac{1}{k}\{1 - E_L(T_r)\}}$$
$$= \frac{\frac{L}{k}\int_0^{T_r} \frac{(kt)^L}{L!} e^{-kt}\, dt}{\frac{1}{k}\{1 - E_L(T_r)\}}$$
$$= \frac{-L\int_0^{T_r} p_L(t)\, dt}{\{1 - E_L(T_r)\}}. \quad (16)$$

From Property 3, (6) thus reduces to the desired equality.

Property 6:

$$\hat{T} = \frac{L}{k}\{1 - E_{L+1}(T_r)\} + T_r E_L(T_r).$$

Proof:

$$\hat{T} = P_f \tilde{T}_f + P_r T_r$$
$$= \{1 - E_L(T_r)\}\frac{L}{k}\frac{1 - E_{L+1}(T_r)}{1 - E_L(T_r)} + T_r E_L(T_r)$$
$$= \left\{\frac{L}{k}\, 1 - E_{L+1}(T_r)\right\} + T_r E_L(T_r). \quad (17)$$

Property 7:

$$T^* = \frac{L}{k}.$$

Property 8:

$$\frac{d(P_f)}{d(kT_r)} = p_{L-1}(T_r) \quad \text{and} \quad \frac{d(P_r)}{d(kT_r)} = -p_{L-1}(T_r).$$

Proof: This result follows simply by differentiating the expressions for P_f and P_r of properties 1 and 2 analytically:

$$P_r = E_L(T_r)$$
$$= \sum_{i=0}^{L-1} \frac{(kT_r)^i}{i!} e^{-kT_r}$$
$$= e^{-kT_r} + \sum_{i=1}^{L-1} \frac{(kT_r)^i}{i!} e^{-kT_r}. \quad (18)$$

Thus

$$\frac{d(P_r)}{d(kT_r)} = -e^{-kT_r} + \sum_{i=1}^{L-1} \frac{i(kT_r)^{i-1}}{i!} - \frac{(kT_r)^i}{i!} e^{-kT_r}$$
$$= \sum_{i=1}^{L-1} \frac{(kT_r)^{i-1}}{(i-1)!} e^{-kT_r} - \sum_{i=0}^{L-1} \frac{(kT_r)^i}{i!} e^{-kT_r}$$
$$= E_{L-1} - E_L$$
$$= -p_{L-1}(T_r). \quad (19)$$

Moreover, since

$$P_f = 1 - P_r, \quad (20)$$
$$\frac{d(P_f)}{d(kT_r)} = \frac{d(1 - P_r)}{d(kT_r)} = p_{L-1}(T_r). \quad (21)$$

Property 9:

$$\frac{d(k\tilde{T}_f)}{d(kT_r)} = L\frac{[1 - E_L(T_r)\}p_L(T_r) - \{1 - E_{L+1}(T_r)\}p_{L-1}(T_r)}{[1 - E_L(T_r)]^2}.$$

Proof: From Property 3,

$$k\tilde{T}_f = L\frac{1 - E_{L+1}(T_r)}{1 - E_L(T_r)}. \quad (22)$$

Thus by direct differentiation

$$\frac{d(k\tilde{T}_f)}{d(kT_r)} = L\frac{\{1 - E_L(T_r)p_L(T_r) - \{1 - E_{L+1}(T_r)\}p_{L-1}(T_r)}{\{1 - E_L(T_r)\}^2}. \quad (23)$$

Property 10:

$$\frac{d(k\hat{T})}{d(kT_r)} = E_L(T_r).$$

Proof: From Property 6,

$$\hat{T} = \frac{L}{k}\{1 - E_{L+1}(T_r)\} + T_r E_L(T_r). \quad (24)$$

Hence

$$k\hat{T} = L\{1 - E_{L+1}(T_r)\} + kT_r E_L(T_r). \tag{25}$$

Thus by direct differentiation

$$\frac{d(k\hat{T})}{d(kT_r)} = L\{p_L(T_r)\} + (kT_r(-p_{L-1}(T_r)) + E_L(T_r)$$

$$= Lp_L(T_r) - kT_r p_{L-1}(T_r) + E_L(T_r). \tag{26}$$

Since

$$Lp_L(T_r) = L\,\frac{(kT_r)^L}{L!}\,e^{-kT_r}$$

$$= (kT_r)\,\frac{(kT_r)^{L-1}}{(L-1)!}\,e^{-kT_r}$$

$$= kT_r p_{L-1}(T_r), \tag{27}$$

this reduces to

$$\frac{d(k\hat{T})}{d(kT_r)} = E_L(T_r) \tag{28}$$

as required.

Given the above statistics for replacement, on-line failure, and expected time to failure of a component with estimated lifetime L and assumed replacement time T_r, we desire to choose T_r (given L) which minimizes some appropriate cost function. Intuitively, this cost function should represent both the cost of on-line failure and the cost of wasted component lifetime due to replacing components before failure [12], [13]. We therefore adopt the cost functional

$$\text{cost} = C_f P_f + C_W(kT^* - k\hat{T}). \tag{29}$$

Here C_f and C_W are appropriate weight factors representing the cost of on-line failure and the cost of component lifetime wastage, respectively. Thus the first term in the cost functional represents the expected cost of a failure (i.e., the probability of an on-line failure times the cost of such a failure), whereas the second term in the cost functional represents the expected cost of wasted component lifetime (i.e., the expected lifetime reduction times the cost per unit time for such a lifetime reduction, with k serving as a normalizing factor).

To minimize the cost functional of (29), one simply substitutes the values for $P_f(T_r)$, T^*, and $\hat{T}(T_r)$ computed in the preceding pages, differentiating the cost with respect to kT_r and setting it equal to zero. This then results in the equality [4]

$$0 = C_f p_{L-1}(T_r) - C_W E_L(T_r) \tag{30}$$

where $d(P_f)/d(kT_r)$ is given by property 9 and $d(k\hat{T})/d(kT_r)$ is given by property 10. Thus the choice of an optimal T_r (given L) is reduced to the solution of a single nonlinear equation in one unknown. The solutions of this equation are plotted in Fig. 1 for a number of values of L and C_f/C_W. Indeed, it can be readily shown that (30) has exactly one solution for $T_r > 0$. Moreover, the function

$$R_L(t) = C_f p_{L-1}(t) - c_W E_L(t) \tag{31}$$

takes on negative values for $0 < t < T_r$ and positive values for $T_r < t$; hence in an on-line maintenance system one need not even solve (30). Rather, one simply evaluates $R_L(t)$ at the time of the next scheduled maintenance. If this results in a negative number, the next scheduled maintenance precedes the optimal replacement time, and hence we should wait at least until the next scheduled maintenance (when we will have more data) to replace the component. On the other hand, if the evaluation of $R_L(t)$ at the next scheduled maintenance time results in a positive value, the optimal replacement time will have passed by the next scheduled maintenance, and hence the component should be replaced at the present maintenance interval.

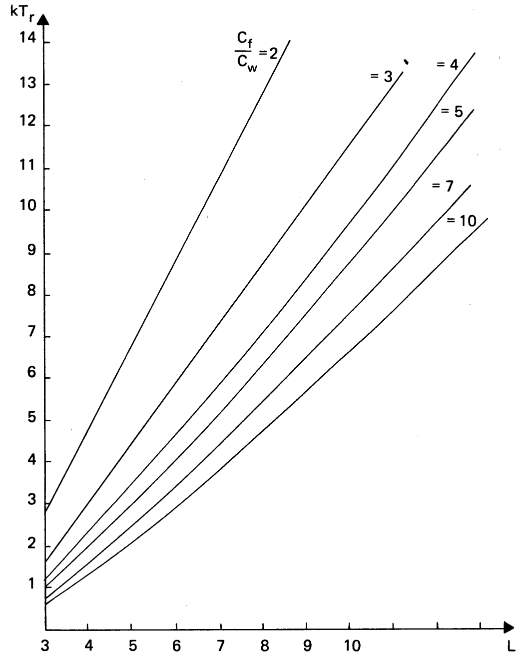

Fig. 1. Replacement time (kT_r) versus Lifetime L with different weight constant.

V. The Algorithm

Summarizing the on-line maintenance algorithm resulting from the above theory takes the following form. At the gth scheduled maintenance interval (at time gT) one measures the component parameter C_g. If C_g is already out of tolerance, the component is simply replaced, and no further analysis is required. If, however, C_g is in tolerance ($C_g > 0$ in our notation), it is used together with the values of the component parameter measured at the previous maintenance intervals to estimate the dynamics of the failure model for the component. Here sequential refinement schemes may be used both to include the effect of C_g on the estimates made at the $(g-1)$st maintenance interval and to increase the order of the polynomial used to represent the component failure dynamics. Once the component failure dynamics have been satisfactorily estimated, one evaluates (31) to estimate whether or not to replace the component. If $R_L((g+1)T) \geq 0$, the component is replaced, whereas if $R_L((g+1)T) < 0$, the component is not replaced, and the system is returned to service until the next scheduled maintenance.

VI. Simulations

A computer simulation of an on-line periodic maintenance system based on the above described algorithm was performed for 600 maintenance intervals with a new component replacing the old component after each replacement decision and/or on-line failure [4]. The system was subjected to computer-generated Poisson shocks with constant $k = 0.1$ shocks per hour and a maintenance interval of $T = 20$ h. The simulation was first run using identical components with $L = 28$ (expected lifetime of 14 maintenance intervals) and then repeated using random components and noisy data measurements.

TABLE I

TOTAL REPLACEMENTS AND FAILURES WITHIN 600 MAINTENANCE INTERVALS FOR DIFFERENT C_f/C_W

C_f/C_w	No. of replacement	No. of failure
50	48	7
75	56	1
100	52	2
150	54	2
200	54	2

TABLE II

TOTAL REPLACEMENTS AND FAILURES WITHIN 600 MAINTENANCE INTERVALS FOR VARIOUS FIXED REPLACEMENT STRATEGIES

Constant replacement time	No. of replacement	No. of failure
every 6 intervals	100	0
every 7 intervals	85	0
every 8 intervals	75	0
every 9 intervals	65	1
every 10 intervals	59	1
every 11 intervals	48	6
every 12 intervals	39	11

TABLE III

OVERALL COST WITH DIFFERENT METHODS AND DIFFERENT C_f/C_W

Cost \ C_f/C_w / Methods	50	75	100	150	200
every 6 intervals	1600	1600	1600	1600	1600
every 7 intervals	1096	1096	1096	1096	1096
every 8 intervals	900	900	900	900	900
every 9 intervals	698	723	748	798	848
every 10 intervals	530	555	580	630	680
every 11 intervals	612	762	912	1212	1512
every 12 intervals	750	1025	1300	1850	2400
the algorithm	690	471	512	668	768

TABLE IV

TOTAL REPLACEMENTS AND FAILURES WITHIN 600 MAINTENANCE INTERVALS FOR VARIOUS FIXED REPLACEMENT STRATEGIES AND THE ALGORITHM AT DIFFERENT NOISE LEVELS

noise level	20%		30%		40%		60%	
method	No. of replace	No. of fail	No. of replace	No. of fail	No. of replace	No. of fail	No. of replace	No. of fail
every 6 intervals	100	0	100	0	100	0	94	6
every 7 intervals	85	0	85	0	84	1	78	8
every 8 intervals	75	0	72	3	71	4	64	12
every 9 intervals	64	2	63	3	60	7	52	17
every 10 intervals	56	4	51	9	45	15	45	18
every 11 intervals	45	10	45	10	45	10	39	20
every 12 intervals	36	15	35	16	36	17	31	23
the algorithm	56	3	55	5	55	5	50	11

TABLE V

OVERALL COST FOR DIFFERENT METHODS AT DIFFERENT NOISE LEVELS

cost \ noise level / method	0%	20%	30%	40%	60%
every 6 intervals	1600	1600	1600	1600	2200
every 7 intervals	1096	1096	1096	1280	2008
every 8 intervals	900	900	1200	1300	2128
every 9 intervals	748	848	948	1376	2432
every 10 intervals	580	880	1380	1980	2364
every 11 intervals	912	1340	1340	1340	2452
every 12 intervals	1300	1728	1828	1984	2612
the algorithm	512	752	980	752	1608

For the case where identical components were employed, Table I gives the total number of replacements and failures resulting from the application of the algorithm over the 600 simulated maintenance intervals with different values of C_f/C_W. For comparison, Table II shows the total number of replacements and failures which would have resulted from a fixed replacement strategy ranging from six to twelve maintenance intervals. Since the cost function is

$$\text{cost} = C_f P_f + C_W(kT^* - k\hat{T}) \tag{32}$$

the overall cost can be expressed as

$$\text{cost} = \frac{C_f}{C_W}\,(\text{number of failures}) + 0.1\,(280^*\,(\text{number of components used}) - 12\,000). \tag{33}$$

The overall costs resulting from the application of our algorithm and the various fixed replacement schedules may be computed from the data in Tables I and II. The resultant costs for different values of C_f/C_W are given in Table III.

Note that since $L = 28$ for each component in this simulation, an optimal replacement strategy of approximately ten maintenance intervals can be computed from (30) without estimating L. As such, it is not surprising that this fixed replacement strategy resulted in lower costs than the algorithm. It should, however, be noted that the algorithm did not have the advantage of an *a priori* knowledge of L, and yet it still outperformed all of the fixed replacement strategies except the optimal strategy (that is, optimal once L is known).

In our second simulation, random noise was added to the data to simulate both the effects of imperfect measurement and the effect of components with random failure characteristics. Various simulations were run as before for 600 maintenance intervals each, with $k = 0.1$ and $T = 20$ and with noise levels ranging between 20 and 60 percent of the tolerance interval. The results of these simulations are given in Tables IV and V. Except for a single case, which we believe to be anomolous, the algorithm outperformed any fixed replacement strategy.

VII. CONCLUSION

In the preceeding pages, we have described a curve fitting algorithm for the prediction of failures in analog devices. The algorithm was tested in a variety of situations and found to be surprisingly effective in predicting failures with relatively little wastage of component lifetime and on-line failure cost.

REFERENCES

[1] J. D. Esary, A. W. Marshall, and F. Proschan, "Shock models and wear processes," *Ann. Probability*, vol. 1, 1973.
[2] J. D. Esary, Ph.D. dissertation, Univ. of California, Berkeley, 1957.
[3] N. I. Johnson, *Discrete Distribution*. New York: Wiley.
[4] K. S. Lu, Ph.D. dissertation, Texas Tech University, Lubbock, TX, 1977.
[5] ——, M.S.E.E. thesis, Texas Tech University, Lubbock, TX, 1973.
[6] A. W. Marshall, "Some comments on the hazard gradient," *J. Stoch. Proc. Applic.*, vol. 3, 1975.
[7] A. Papoulis, *Probability, Random Variable and Stochastic Processes*. New York: McGraw-Hill, 1965.
[8] C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications*. New York: Wiley, 1971.
[9] R. Saeks, "An approach to built-in testing," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-14, pp. 813–818, 1978.
[10] R. Saeks and S. R. Liberty, *Rational Fault Analysis*. New York: Marcel Dekker, 1977.
[11] L. Tung, M.S. thesis, Texas Tech University, Lubbock, TX, 1975.
[12] L. Tung and R. Saeks, "An experiment in fault prediction," in *Proc. 4th Int. Symp. Reliability Electronics*, Budapest, Oct. 1977, pp. 249–257.
[13] L. Tung, S. R. Liberty, and R. Saeks, "Fault prediction towards a mathematical theory," in *Rational Fault Analysis*. New York: Marcel Dekker, 1977, pp. 135–142.

Truncated Sampling Expansions for Band-Limited Random Signals

R. B. KERR

Abstract—**Truncated expansions based upon the sampling theorem but containing only a few terms can be very useful for practical interpolations of band-limited random signals. Such interpolations can be utilized, for example, to improve the accuracy of algorithms used to find the dynamic response of systems to band-limited inputs. The statistical accuracy of such interpolations is investigated, and one very simple application, which yields a Runge–Kutta simulation algorithm of improved accuracy with very little increase in computation, is indicated.**

I. INTRODUCTION

The sampling theorem for band-limited signals, which is so invaluable theoretically to signal analysis and information theory, can also be put to very practical use in the interpolation of sampled signals, if expansions of only a few terms are employed. Besides applications involving straight signal reconstruction (image processing, etc.) truncated sampling expansions can be used to improve the accuracy of digital-simulation algorithms used to determine dynamic-system responses, when the system inputs are sampled at a rate equal to or exceeding the Nyquist rate. Of course, the simulation accuracy can always be improved by increasing the sampling rate. This, however, is not always possible or desirable. Suppose, for example, that we are trying to determine the response statistics of an ocean platform or a moored ship excited by random ocean-wave forces. In this case simulation runs of long duration are required, and if actual wave data are used, they will probably be available only at a fixed sampling interval. Even if the input data are synthetically generated, it is desirable to be able to use a rate not too much in excess of the Nyquist rate. In cases such as this, the use of a few terms of the sampling expansion can often improve simulation accuracy with very little increase in computation.

Manuscript received July 28, 1978; revised November 17, 1978.
The author is with the School of Engineering, Duke University, Durham, NC 27706.

In Section II, we investigate the statistical accuracy of truncated sampling interpolations of uniform band-limited stationary random signals sampled at the Nyquist rate, and in Section III we point out a very simple way in which the accuracy of a Runge–Kutta simulation algorithm can be improved.

II. PRACTICAL CARDINAL FUNCTION INTERPOLATION OF BAND-LIMITED RANDOM SIGNALS

Consider a stationary random process of mean zero and variance σ^2, with a uniform band-limited spectrum. We will assume that samples are taken at an interval $T = 1/2w$ s (w being the single-sided bandwidth) and are therefore mutually uncorrelated. Let us approximate the sample functions $f(t)$ of our random process over each interval $(n-1)T$ to nT by the truncated expansion:

$$f_a(t) = \sum_{i=j+n}^{k+n} f_i \operatorname{sinc}(2wt - i) \tag{1}$$

where

$$\operatorname{sinc} x = \frac{\sin \pi x}{\pi x} \tag{2}$$

and the number of terms used in the truncated sum is $k - j + 1$. We may define the mean-square error function:

$$E(t) = \langle [f(t) - f_a(t)]^2 \rangle \tag{3}$$

($\langle \cdot \rangle$ denotes statistical expectation). $E(t)$ is zero at each sampling point and is a periodic time function since the random process is stationary and the points (samples) used for the interpolation are related to each interval in the same way. Hence, as an error measure, we may define the time average of $E(t)$:

$$\overline{E(t)} = \frac{1}{T}\int_0^T E(t)\, dt. \tag{4}$$

Another useful error measure is the mid-interval mean-square error $E(T/2)$, which will be the maximum mean-square error if "symmetric" data is used for the interpolation—i.e., if the same number of samples on either side of the interval being interpolated are used in the sum (1); in this case $j + k = -1$. (If "present and past" data are used ($k = 0$), then $t = T/2$ will not be the time at which the maximum statistical error occurs.)

To aid in the calculation of $\overline{E(t)}$, we introduce the following notation:

$$\phi_n(t) \triangleq \operatorname{sinc}(2wt - n) = \operatorname{sinc}\left(\frac{t}{T} - n\right) \tag{5}$$

$$\zeta_n \triangleq \frac{1}{T}\int_0^T \phi_n^2(t)\, dt. \tag{6}$$

In order to evaluate $\overline{E(t)}$ for various size expansions, we find that we need to calculate ζ_n. This can be done in terms of the sine-integral functions by a simple change of variable, and we obtain

$$\zeta_n = \frac{1}{\pi}[\operatorname{Si}(n2\pi) - \operatorname{Si}(n-1)2\pi] \tag{7}$$

where

$$\operatorname{Si}(x) = \int_0^x \frac{\sin y}{y}\, dy. \tag{8}$$

It may be noted that the ζ_n simply represent the areas of the "lobes" of $\operatorname{sinc}^2(t/T)$, as indicated in Fig. 1. (Note that $\zeta_n = \zeta_{1-n}$.)