Furthermore, Eq. (2) can be derived in a very elementary way by adding extra rows B_c to B and extra coordinates q_c to the deflection vector q, so that Pian's Eq. (6) becomes

$$q^* = \begin{bmatrix} q \\ q_c \end{bmatrix} = \begin{bmatrix} B \\ B_c \end{bmatrix} \alpha = B^* \alpha \qquad (6)$$

This can always be done so that B^* is nonsingular.

Then, by inverting Eq. (10) of Ref. 2, the flexibility matrix can be written

$$F = B^* G^{-1} B^{*T} \qquad (7)$$

with F satisfying the equation

$$FQ^* = q^* \qquad (8)$$

Now consider that certain forces Q^* are applied to the element causing deflections q^*, and then the forces at the newly added coordinates are removed, but the deflections q maintained. The deflections q_c, at the released coordinates, will adjust so as to minimize the strain energy, and by Eqs. (3) and (4), Eq. (7) reduces, on inversion, to Eq. (2).

References

[1] Pian, T. H. H., "Derivation of element stiffness matrices," AIAA J. **2**, 576–577 (1964).

[2] Gallagher, R. H., "Techniques for the derivation of element matrices," AIAA J. **1**, 1431–1432 (1963).

N-Segment Least-Squares Approximation

R. Aris* and M. M. Denn†
University of Minnesota, Minneapolis, Minn.

The question of the best piecewise approximation to a given function by linear segments[1–3, 5] and by step functions[4] has been considered recently. It may be worthwhile to call attention to a property of the least-square criterion of fit which allows a rather easy calculation of the break points. Let $f(x)$, $a \le x \le b$, be approximated by N segments of a function of fixed form $\phi(x;\ \mathbf{c})$, where $\mathbf{c} = (c_1, \ldots c_k)$ is a vector of constants to be chosen in each segment. The approximation will be best in the least-square sense if the Nk constants $\mathbf{c}_n$, $n = 1, \ldots N$, and the $(N - 1)$ break points x_n, $n = 1, \ldots N - 1$, where $x_0 = a < x_1 < \ldots x_{N-1} < x_N = b$, are so chosen as to minimize

$$F_N(a, b) = \sum_{n=1}^{N} \int_{x_{n-1}}^{x_n} [f(x) - \phi(x;\ \mathbf{c}_n)]^2 dx \qquad (1)$$

This criterion of fit has the property that, over intervals where $f(x)$ is continuous, the segments are either continuous or, at a point of discontinuity, are equidistant from the curve $f(x)$. For if the break points x_{n-1}, x_n are chosen, the best value of the constants $\mathbf{c}_n$ satisfy the equations

$$\int_{x_{n-1}}^{x_n} [f(x) - \phi(x;\ \mathbf{c})](\nabla_{\mathbf{c}}\phi(x;\ \mathbf{c})] = 0 \qquad (2)$$

where $\nabla_{\mathbf{c}}\phi$ denotes the vector of partial derivatives of ϕ with respect to the components of $\mathbf{c}$. Since x_n occurs once as an upper limit and once as a lower limit, the derivative of the right-hand side of (1) with respect to x_n becomes, after using (2),

$$[f(x_n) - \phi(x_n;\ \mathbf{c}_n)]^2 - [f(x_n) - \phi(x_n;\ \mathbf{c}_{n+1})]^2$$

If x_n is chosen to minimize F_N, we have either

$$\phi(x_n;\ \mathbf{c}_n) = \phi(x_n;\ \mathbf{c}_{n+1})$$

or

$$f(x_n) = \tfrac{1}{2}\{\phi(x_n;\ \mathbf{c}_n) + \phi(x_n;\ \mathbf{c}_{n+1})\}$$

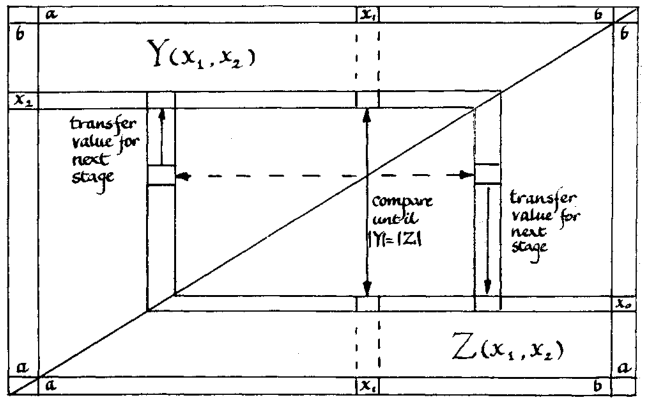

Fig. 1 Tabular method of generating best break points.

The value of this property may be illustrated by a tabular procedure for $N = 2$; it will be evident that a graph could be used in a similar way. Let the best fitting single segment for the interval (y, z), $a < y < z < b$, be calculated. The constants $\mathbf{c}(y, z)$ and the fit $F_1(y, z)$ may be recorded in double-entry table. Let

$$Y(y, z) = f(y) - \phi[y;\ \mathbf{c}(y, z)]$$

$$Z(y, z) = f(z) - \phi[z;\ \mathbf{c}(y, z)]$$

be tabulated as in Fig. 1. In the lower right-hand triangle, the deviation at the upper end of a single segment fit to the interval (x_0, x_1) is entered in a row corresponding to x_0 and column x_1. In the upper left-hand triangle, the entry $Y(y_1, x_2)$ is in a row corresponding to x_2 and column x_1. To find the best break point for two segments, we have only to compare entry of Y in the x_2 row with that of Z in the x_0 row and the same column in order to see at what value of x_1 we have either $Y = Z$ or $Y = -Z$. If more than one possibility exists, the best choice may easily be made by reference to the table of F_1. With a sufficient number of entries, linear interpolation should suffice.

The process can be continued indefinitely by replacing the lower right-hand triangle by a table of $Z_2(x_0, x_2)$, the deviation at the upper end of the best two-segment fit. But this is merely to move the entry $Z(x_1, x_2)$ to the row x_0 as indicated by the arrows. If the same thing is done in the upper right-hand triangle to give $Y_2(x_2, x_4)$, the lower-end deviation of the best two-segment fit over the interval (x_2, x_4), the table may now be used for calculating the best four-segment fit. If $N = 2^p$, only p repetitions of the process are required. Of course, record would be kept of the break points at each stage, and then the best constants could be obtained immediately from the single-segment tables.

References

[1] Aris, R., *Discrete Dynamic Programming* (Blaisdell, New York, 1964).

[2] Bellman, R. E. and Dreyfus, S. E., *Applied Dynamic Programming* (Princeton University Press, Princeton, N. J., 1962).

[3] Bellman, R. E. and Kotkin, B., "On the approximation of curves by line segments using dynamic programming—II," Rand Corp. RM-2978-PR (1962).

[4] Lubowe, A. G., "Optimal functional approximation using dynamic programming," AIAA J. **2**, 376–377 (1964).

[5] Stone, H., "Approximation of curves by line segments," Math. Computation **15**, 40–47(1961).

Received May 1, 1964.

* Professor, Department of Chemical Engineering.
† Graduate Student, Department of Chemical Engineering.