required to evaluate the undetermined coefficients U_j. As might be expected, the accuracy of this scheme depends heavily on the location of the collocation points. One approach to selecting these locations, called orthogonal collocation, has been used successfully in nonlinear problems arising in chemical engineering (see, e.g., Finlayson, 1972).

2.2.2 Application of the Method of Weighted Residuals

The concepts of MWR may appear somewhat abstract in the compact notation presented in the preceding section. The following examples, however, should serve to clarify the basic concepts and to demonstrate that, in practice, the method is relatively straightforward. As an example, let us consider the ODE describing the rate of cooling of an object with a temperature T in an environment at temperature T_e. Newton's law of cooling states that the time rate of cooling is proportional to the difference between the object temperature and the environmental temperature

$$\frac{dT}{dt} + k(T - T_e) = 0, \tag{2.2.9}$$

where T = object temperature
T_e = temperature of the environment
t = time
k = a proportionality constant.

Let us calculate the temperature change given the trial functions of Figure 2.5 defined over the interval $0 \leq t \leq 1$. The initial condition is $T = 1$ when $t = 0$. We specify the remaining two parameters in (2.2.9) as $k = 2$ and $T_e = \frac{1}{2}$.

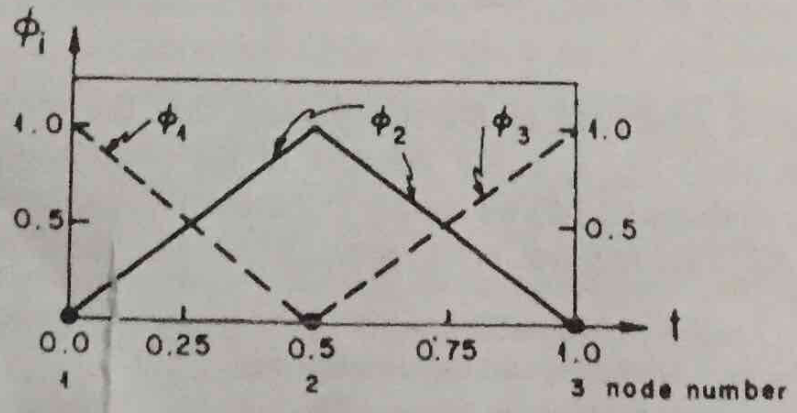

Figure 2.5. Basis functions used to calculate temperature distribution over the interval $0 < t \leq 1$.

The functional form of the bases illustrated in Figure 2.5 is

$$\phi_i = \begin{cases} \dfrac{t - t_{i-1}}{t_i - t_{i-1}}, & t_{i-1} \le t \le t_i \\ \dfrac{t_{i+1} - t}{t_{i+1} - t_i}, & t_i \le t \le t_{i+1}, \end{cases} \tag{2.2.10}$$

where the subscripts refer to nodal values in the finite difference sense used earlier. Note that when $T_1 = 1$, this choice of bases is consistent with the requirement that the trial function, $\hat{T}$ say, fulfills the condition $T(0) = 1$ [T_1 is the undetermined coefficient at node 1; see (2.2.1)].

In the following development, the general MWR problem is formulated in steps 1 and 2. The choice of a specific MWR scheme takes place in step 3, when a specific form of the weighting function is selected.

Step 1

Define the trial function

$$T \approx \hat{T} = \sum_{j=1}^{3} T_j \phi_j(t). \tag{2.2.11}$$

It should be noted here that, because of the definition of $\phi_j(t)$, T_j is the value of $\hat{T}$ at the nodal point j.

Step 2

Formulate the integral equations arising out of the MWR:

$$\int_t R(t) w_i(t)\, dt = 0, \qquad i = 1, 2, 3, \tag{2.2.12}$$

where

$$R(t) \equiv \frac{d\hat{T}}{dt} + k(\hat{T} - T_e).$$

Substitution of (2.2.11) into (2.2.12) yields

$$\int_t \left\{ \sum_{j=1}^{3} T_j \left(\frac{d\phi_j}{dt} + k\phi_j \right) - kT_e \right\} w_i(t)\, dt = 0, \qquad i = 1, 2, 3. \tag{2.2.13}$$

Thus we have three equations, one written for each of the three weighting functions w_i, and three undetermined parameters T_j.

Step 3

Select weighting functions:

1. *Galerkin.* The appropriate weighting function in the Galerkin method is the basis function $\phi_i(t)$. Introduction of this basis function for w_i in (2.2.13) yields

$$(2.2.14) \qquad \sum_{j=1}^{3} T_j \int_0^1 \left\{ \frac{d\phi_j}{dt} + k\phi_j \right\} \phi_i \, dt = \int_0^1 kT_e \phi_i \, dt, \qquad i = 1, 2, 3.$$

Expansion of (2.2.14) yields the following matrix equation:

(2.2.15)

$$\begin{bmatrix} \int_0^{1/2} \left(\frac{d\phi_1}{dt}\phi_1 + k\phi_1\phi_1 \right) dt & \int_0^{1/2} \left(\frac{d\phi_2}{dt}\phi_1 + k\phi_2\phi_1 \right) dt & 0 \\ \int_0^{1/2} \left(\frac{d\phi_1}{dt}\phi_2 + k\phi_1\phi_2 \right) dt & \int_0^{1} \left(\frac{d\phi_2}{dt}\phi_2 + k\phi_2\phi_2 \right) dt & \int_{1/2}^{1} \left(\frac{d\phi_3}{dt}\phi_2 + k\phi_3\phi_2 \right) dt \\ 0 & \int_{1/2}^{1} \left(\frac{d\phi_2}{dt}\phi_3 + k\phi_2\phi_3 \right) dt & \int_{1/2}^{1} \left(\frac{d\phi_3}{dt}\phi_3 + k\phi_3\phi_3 \right) dt \end{bmatrix} \cdot \begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix} = \begin{bmatrix} \int_0^{1/2} kT_e\phi_1 \, dt \\ \int_0^{1} kT_e\phi_2 \, dt \\ \int_{1/2}^{1} kT_e\phi_3 \, dt \end{bmatrix}.$$

Evaluation of the integrals appearing in (2.2.15) gives

$$(2.2.16) \qquad \frac{1}{2} \begin{bmatrix} -1 + \frac{k}{3} & 1 + \frac{k}{6} & 0 \\ -1 + \frac{k}{6} & \frac{2k}{3} & 1 + \frac{k}{6} \\ 0 & -1 + \frac{k}{6} & 1 + \frac{k}{3} \end{bmatrix} \begin{bmatrix} 1 \\ T_2 \\ T_3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \frac{kT_e}{2} \\ kT_e \\ \frac{kT_e}{2} \end{bmatrix},$$

where we have used the initial condition to evaluate T_1. Equation (2.2.16) may be solved using the specified values for k and T_e. Only two of the three equations can be satisfied exactly. The resulting solution depends upon the two equations selected. From the second and third equations, we obtain

$$[T_1, \quad T_2, \quad T_3] = [1, \quad 0.678, \quad 0.571],$$

from the first and third equations

$$[T_1,\ T_2,\ T_3]=[1,\ 0.625,\ 0.550],$$

and from the first and second equations

$$[T_1,\ T_2,\ T_3]=[1,\ 0.625,\ 0.625].$$

The analytical solution for this problem is

$$T(t)=T(0)e^{-kt}+T_e(1-e^{-kt}),$$

which yields

$$[T_1,\ T_2,\ T_3]=[1,\ 0.684,\ 0.568].$$

It is apparent that the most accurate solution is obtained using equations two and three. This corresponds to the case when the equation written at the boundary node is not used explicitly. In practice, this choice is routinely made by simply eliminating from the matrix equation rows and columns associated with specified, or known coefficients (e.g., row 1 would be eliminated since T_1 is specified).

2. *Subdomain*. The appropriate weighting function for the subdomain method is given by

$$w_i=\begin{cases}1, & \dfrac{t_i+t_{i-1}}{2}\leqslant t\leqslant\dfrac{t_i+t_{i+1}}{2}\\ 0, & \text{otherwise.}\end{cases} \tag{2.2.17}$$

Substitution of (2.2.17) into (2.2.13) results in the following integral equations:

$$\sum_{j=1}^{3}T_j\int_0^{0.25}\left\{\frac{d\phi_j}{dt}+k\phi_j\right\}dt=\int_0^{0.25}kT_e\,dt, \tag{2.2.18a}$$

$$\sum_{j=1}^{3}T_j\int_{0.25}^{0.75}\left\{\frac{d\phi_j}{dt}+k\phi_j\right\}dt=\int_{0.25}^{0.75}kT_e\,dt, \tag{2.2.18b}$$

$$\sum_{j=1}^{3}T_j\int_{0.75}^{1.0}\left\{\frac{d\phi_j}{dt}+k\phi_j\right\}dt=\int_{0.75}^{1.0}kT_e\,dt. \tag{2.2.18c}$$

Equation (2.2.18) may be written in matrix form as

(2.2.19)

$$\begin{bmatrix} \int_0^{0.25}\left\{\frac{d\phi_1}{dt}+k\phi_1\right\}dt & \int_0^{0.25}\left\{\frac{d\phi_2}{dt}+k\phi_2\right\}dt & 0 \\ \int_{0.25}^{0.75}\left\{\frac{d\phi_1}{dt}+k\phi_1\right\}dt & \int_{0.25}^{0.75}\left\{\frac{d\phi_2}{dt}+k\phi_2\right\}dt & \int_{0.25}^{0.75}\left\{\frac{d\phi_3}{dt}+k\phi_3\right\}dt \\ 0 & \int_{0.75}^{1.0}\left\{\frac{d\phi_2}{dt}+k\phi_2\right\}dt & \int_{0.75}^{1.0}\left\{\frac{d\phi_3}{dt}+k\phi_3\right\}dt \end{bmatrix} \cdot \begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix} = \begin{bmatrix} \int_0^{0.25} kT_e\,dt \\ \int_{0.25}^{0.75} kT_e\,dt \\ \int_{0.75}^{1.0} kT_e\,dt \end{bmatrix}.$$

Evaluation of the integrals appearing in (2.2.19) provides the final form of the matrix equation:

(2.2.20)
$$\frac{1}{2}\begin{bmatrix} -1+\frac{3k}{8} & 1+\frac{k}{8} & 0 \\ -1+\frac{k}{8} & \frac{3k}{4} & 1+\frac{k}{8} \\ 0 & -1+\frac{k}{8} & 1+\frac{3k}{8} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ T_2 \\ T_3 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \frac{kT_e}{2} \\ kT_e \\ \frac{kT_e}{2} \end{bmatrix}.$$

The accuracy of the solution obtained from (2.2.20) once again depends on the choice of equations. From the second and third equations, we obtain

$$[T_1,\ T_2,\ T_3]=[1,\ 0.684,\ 0.579],$$

from the first and third equations

$$[T_1,\ T_2,\ T_3]=[1,\ 0.600,\ 0.543],$$

and from the first and second

$$[T_1,\ T_2,\ T_3]=[1,\ 0.600,\ 0.680].$$

It is apparent that once again, as in the case of the Galerkin formulation, less accurate solutions are obtained when the equation written for the Dirichlet

node is included in the analysis. The solution generated using the best Galerkin and subdomain formulations appear to be of approximately the same accuracy.

3. *Collocation.* The collocation method is obtained by substituting the Dirac delta function for the weighting function of (2.2.13). This substitution yields the following expression:

$$(2.2.21)\qquad \sum_{j=1}^{3} T_j \int_t \left\{ \frac{d\phi_j}{dt} + k(\phi_j - T_e) \right\} \delta(t - t_i)\, dt = 0, \qquad i = 1,2,3.$$

We are now confronted with the problem of where to choose the collocation points or, alternatively, where to locate the t_i points. Because we have three undetermined coefficients, three collocation points are required. If, however, one point is chosen at $t=0$, the initial condition eliminates one unknown. The remaining two points could be located anywhere within the intervals

$$0 < t < 0.5 \qquad \text{and} \qquad 0.5 < t < 1.0.$$

A point located at $t=0.5$ would present additional complications because the trial function is continuous at that point but its derivative is discontinuous there, so that a value chosen is rather arbitrary. This is equivalent to the choice of the natural knots of a spline function as the collocation points.*

A more accurate collocation scheme can be developed using Gaussian integration points as collocation points (DeBoor and Swartz, 1973). Because we are using piecewise linear trial functions in this example, there is one Gaussian point per element and these points are located at $t=0.25$ and $t=0.75$. Equations (2.2.21) can be written in matrix form as

$$(2.2.22)\qquad \begin{bmatrix} \dfrac{d\phi_1}{dt} + k\phi_1\Big|_{0.0} & 0 & 0 \\ \dfrac{d\phi_1}{dt} + k\phi_1\Big|_{0.25} & \dfrac{d\phi_2}{dt} + k\phi_2\Big|_{0.25} & 0 \\ 0 & \dfrac{d\phi_2}{dt} + k\phi_2\Big|_{0.75} & \dfrac{d\phi_3}{dt} + k\phi_3\Big|_{0.75} \end{bmatrix} \cdot \begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix} = \begin{bmatrix} kT_e|_{0\,0} \\ kT_e|_{0.25} \\ kT_e|_{0.75} \end{bmatrix}.$$

*By a *spline*, p, we refer here to a function on the region D which, in the vicinity of a point within D, can be represented by a polynomial of a specified degree, m say. This function and its first $m-1$ derivatives must be continuous and the mth derivatives are square integrable. The *knots* of the spline p are the real numbers $t_0, t_1, \ldots, t_n$ at which we constrain the behavior of p.

An important characteristic of the collocation scheme is illustrated in this expression. Because the collocation point is not located at a node, the algebraic equation is dependent only on information associated with its host element. Thus the second equation in (2.2.22) does not link T_3 to T_1 as was the case in the methods above.

Evaluation of the coefficient matrix and incorporation of the initial condition yields

$$\text{(2.2.23)} \qquad \begin{bmatrix} 1 & 0 & 0 \\ -2+\frac{k}{2} & 2+\frac{k}{2} & 0 \\ 0 & -2+\frac{k}{2} & 2+\frac{k}{2} \end{bmatrix} \cdot \begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix} = \begin{bmatrix} 1 \\ kT_e \\ kT_e \end{bmatrix}.$$

Equation (2.2.23) provides the following solution for the coefficients:

$$[T_1,\ T_2,\ T_3] = [1,\ 0.667,\ 0.555].$$

Although it is apparent that the collocation scheme performed relatively poorly in this simple example, the method is nevertheless attractive from several points of view. The coefficient matrix is easily generated, even for very complex operators. The unusual matrix structure (not all collocation matrices are triangular and amenable to immediate forward substitution) suggests the possibility of computationally efficient equation solvers, and the straightforward and relatively simplistic nature of the underlying theory suggests that the approach should be readily understood by the practitioner, who may have a minimal background in numerical methods. It should also be pointed out that collocation is generally used either with basis functions of higher-order continuity (e.g., spline functions in C^m,* where $m>0$) or with trial functions exhibiting global support (e.g., a single higher-degree polynomial is used over the whole domain of interest).

It is particularly interesting to examine the relationship between collocation and Galerkin's method. To do this, let us digress for a moment and reconsider the matrix equation that arises out of (2.2.15), the Galerkin formulation, when a one-point Gaussian quadrature is used to evaluate the integrals appearing in the coefficient matrix. Although the integrals in this case are easily evaluated, in general it is advantageous to operate in a dimensionless ξ coordinate system where $-1 \leqslant \xi \leqslant 1$. Gaussian quadrature information is readily available in standard texts for this coordinate system. To employ this approach, a simple

*C^m denotes a class of functions that have continuous derivatives up to the mth order.