The Method of Weighted Residuals—A Review

B. A. FINLAYSON and L. E. SCRIVEN

UNIVERSITY OF MINNESOTA,
MINNEAPOLIS, MINNESOTA

Abstract

The method of weighted residuals unifies many approximate methods of solution of differential equations that are being used currently. This review presents the basic method in its historical context and shows some of the many possible modifications that have been used throughout the past fifty years. The relationship between the Galerkin method, which is one version of the method of weighted residuals, and variational methods is outlined. Also included is an extensive listing of published applications of the method of weighted residuals.

Introduction

The method of weighted residuals is an engineer's tool for finding approximate solutions to the equations of change of distributed systems. Experience and intuition can be distilled into a reasonable and sometimes quite accurate first guess, from which it is possible to proceed to successively improved approximations. The analytical form of the approximate solution is often more useful than solutions generated by numerical integration, and the approximate solution usually requires less computation time to generate. The method is applicable to nonlinear and non-self-adjoint problems—one of its most attractive features.

The method of weighted residuals (MWR) includes many approximation methods that are being used currently. It provides a vantage point from which it is easy to see the unity of these methods as well as the relationships between them. This review, after outlining application of the basic method to initial-value, boundary-value, and eigenvalue problems, surveys the history of major contributions to the subject and discusses some of the many modifications of the basic method. The review concludes with a listing of applications of weighted residual methods to problems arising in applied mechanics and related fields. Four practical aspects of MWR in need of further research are identified.

A. Basic Method

The best available treatments of MWR have been those by Crandall [1], who coined the name method of weighted residuals, Ames [2], and Collatz [3], who calls these methods error-distribution principles. The following outline parallels their treatments, in places contrasting them and elaborating on them.

Given a system of differential or integro-differential equations of change and constitutive relations, boundary conditions representing the interactions between the system and its surroundings, and initial conditions representing some base state of interest, the general approach is to assume a trial solution whose functional dependence on position is chosen, but which includes undetermined functions of time. The latter are found by requiring that the trial solution satisfy the differential equation in some specified approximate sense.

Initial Value Problem

Consider the differential equation for \( u(x, t) \):

\[
N(u) - \frac{\partial u}{\partial t} = 0 \quad x \text{ in } V, \ t > 0
\]

where \( N(\cdot) \) denotes a general differential operator involving spatial derivatives of \( u \), \( V \) is a three-dimensional domain with boundary \( S \), and \( t \) represents time. Suppose the initial and boundary conditions are

\[
\begin{align*}
  u(x, 0) &= u_0(x), \quad x \text{ in } V \\
  u(x, t) &= f_d(x, t), \quad x \text{ on } S
\end{align*}
\]

Assume a trial solution of the form

\[
u^*(x, t) = u_d(x, t) + \sum_{i=1}^{N} c_i(t) u_i(x, t)
\]

where the approximating functions, \( u_i \), are prescribed and satisfy the boundary conditions

\[
u_x = f_x, \quad u_i = 0, \quad x \text{ on } S
\]

Then \( u^* \) satisfies the boundary conditions for all functions \( c_i(t) \). It is not necessary that the trial solution
be linear in the $c_i$, but such a choice is usually made for simplicity; no systematic study of alternatives has been reported, so far as the authors know. The differential equation residual and initial residual,

$$R(u^*) = N(u^*) - \frac{\partial u^*}{\partial t}$$

$$R_d(u^*) = u_d(x) - u_g(x, 0) - \sum_{i=1}^{N} c_i(0) u_i(x, 0)$$

are measures of the extent to which the function $u^*$ satisfies the differential equation and initial conditions, respectively. As the number $N$ of approximating functions $u_i$ is increased in successive approximations, one hopes the residuals will become smaller; the exact solution is obtained when both residuals are identically zero. As an approximation to this ideal, the weighted integrals of the residuals are set equal to zero:

$$\langle w_j; R(u^*) \rangle = 0$$

$$\langle w_j; R_d(u^*) \rangle = 0$$

where

$$\langle w_i, v \rangle = \int_{V} w_i v dV$$

represents a spatial average or inner product and $w_i$ is a prescribed weighting function. If $u^*$ is the exact solution, Equations (7) are satisfied regardless of the choice of weighting functions.

The weighting functions can be chosen in several different ways, and each choice corresponds to a different criterion in MWR. Once the choice is made, Equations (7) become a set of $N$ first-order ordinary differential equations in the $N$ unknowns $c_i(t)$. For the linear problem

$$\frac{\partial u_i}{\partial t} = L(u)$$

with approximating functions $u_i$ and $u_g$ that do not themselves depend on time, Equations (7) become simply

$$\sum_{i=1}^{N} \frac{dc_i}{dt} \langle w_j; u_i \rangle = \sum_{i=1}^{N} c_i \langle w_j; L(u_i) \rangle + \langle w_j; L(u_g) \rangle$$

or, in matrix notation

$$\dot{c} = \mathbf{A} \dot{c} + \mathbf{b}$$

The solution to these equations is substituted into Equation (3) to give the approximate solution to the problem. Successive approximations are obtained by increasing $N$ and solving Equation (10) anew. The convergence of successive approximations gives a clue, but not necessarily a definitive one, to the reasonableness of the approximation.

### Boundary Value and Eigenvalue Problems

The method is equally applicable to steady-state and eigenvalue problems. For steady-state problems, the $c_i$ are constants rather than functions of time; for linear problems they are determined as solutions to

$$\overline{B} \overline{c} = \overline{b}.$$  \hspace{1cm} (12)

For nonlinear boundary-value problems it may be useful to assume trial solutions of a more general form than Equation (3), viz.:

$$u^*(x) = \phi(1; c_1, u_1(x))$$

For the linear eigenvalue problem

$$L(u) - \lambda u = 0$$

the approximate solution is determined by

$$\sum_{i=1}^{N} c_i \left( \langle w_j; L(u_i) \rangle - \lambda \langle w_j; u_i \rangle \right) = 0$$

and this set of equations has a non-trivial solution only if

$$\det(A_{ij} - \lambda B_{ij}) = 0.$$ \hspace{1cm} (16)

The values of $\lambda$ for which this is true are the approximations to the first $N$ eigenvalues $\lambda_k$.

### Weighting Functions

The choice of the weighting functions, $w_i$ in (7), corresponds to various criteria in MWR: the historical relationship of the criteria is portrayed in Table I.

In the collocation method, due to Frazer, Jones, and Skan [4], the weighting functions are the Dirac delta functions

$$w_j = \delta(x_j - x);$$

### Table I

<table>
<thead>
<tr>
<th>Date</th>
<th>Investigator</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1921</td>
<td>Pohlhausen [18]</td>
<td>Integral method</td>
</tr>
<tr>
<td>1938</td>
<td>Poritsky [31]</td>
<td>Method of reduction to ordinary differential equations</td>
</tr>
<tr>
<td>1940</td>
<td>Repman [55]</td>
<td>Convergence of Galerkin's method</td>
</tr>
<tr>
<td>1941</td>
<td>Bickley [12]</td>
<td>Collocation, Galerkin, least squares for initial-value problems</td>
</tr>
<tr>
<td>1942</td>
<td>Keldysh [57]</td>
<td>Convergence of Galerkin's method, steady-state</td>
</tr>
<tr>
<td>1949</td>
<td>Faedo [59]</td>
<td>Convergence of Galerkin's method, unsteady-state</td>
</tr>
<tr>
<td>1953</td>
<td>Green [60]</td>
<td>Unification as method of weighted residuals</td>
</tr>
<tr>
<td>1956</td>
<td>Crandall [1]</td>
<td></td>
</tr>
</tbody>
</table>
the differential equation is then satisfied exactly at the N collocation points, x_j. As N is increased, the residual vanishes at more and more points and presumably approaches zero throughout V.

If the weighting functions are

\[ w_j = \begin{cases} 1 & \text{x in } V_j \\ 0 & \text{x not in } V_j \end{cases} \]

then the differential equation is satisfied on the average in each of the N subdomains, V_j; this is the subdomain method [5, 6]. If the V_j are disjoint (which they need not be), the size of one or more subdomains decreases as N is increased, with the result that the differential equation is satisfied on the average in smaller and smaller regions, and presumably the residual approaches zero everywhere. It was Biezeno's presentation [7] of the subdomain method at the First International Congress of Applied Mechanics which prompted Courant's remark [8] that led Crandall to choose the name, "method of weighted residuals." The authors' translation of Courant's remark reads:

"Mr. Courant (Göttingen) indicated afterward that the method advanced by Mr. Biezeno can be viewed from the standpoint of the calculus of variations in the following manner. If a differential equation, as it arises for example in a variational problem, must be satisfied, then we can express it so that the left side of the differential equation, multiplied with an arbitrary function and then integrated, must give us the value zero (vanishing of the first variation). Instead of taking an arbitrary function, we can also take infinitely many determined functions, if these only form a so-called complete function system for the region in question. The piecewise constant functions advanced by Mr. Biezeno are indeed just an especially simple special case of such a complete function system."

The least-squares method, which seems to have been first presented for this type of application by Picone in 1928 [9], uses the weighting functions \( \partial (u^*) / \partial c_j \). The corresponding interpretation is that the mean square residual

\[ I = \int_V [R(u^*)]^2 \, dV \]

(19)

is minimized with respect to the constants c_j.

In the Galerkin method [10], developed in 1915 as the first criterion of what is now known as the method of weighted residuals, the weighting functions w_j are just the approximating functions of u_j. The approximating functions are often members of a complete system of functions, although this property, required for mathematical purposes, is sometimes ignored in practice. The Galerkin method then can be interpreted as making the residual orthogonal to members of the complete set.

A fundamental property (sometimes the definition) of a complete system of functions is that a piecewise continuous function can be orthogonal to each and every member only if the function is identically zero. In the approximation scheme outlined above, the residual is usually continuous (depending upon the differential operator and the choice of approximating functions), and hence the residual can vanish only if it is orthogonal to each member of a complete system of functions. Of course in practice the residual is made orthogonal to no more than a modest, finite number of the members of a complete set. In the original Galerkin method, developed in the study of elastic equilibrium and stability of rods and plates, Galerkin used trial solutions with unknown constant coefficients. Now many similar techniques are often referred to as the Galerkin or generalized Galerkin method: (i) the one given above in which \( c_j = c_j(t) \) for time-dependent problems [12, 1]; (ii) one in which trial solutions are of the more general form \( u^* = f(x, t; c_j) \) with weighting functions \( \partial / \partial c_j \) [13]; and (iii) one in which weighting functions are of the form \( K(u_j) \), rather than \( u_j \), where \( K \) is a specified differential operator [14, 15].

The method of moments is similar to the Galerkin method except that the residual is made orthogonal to members of a system of functions which need not be the same as the approximating functions. Both methods are combined under the single name of orthogonalization methods by Collatz [3]. Yamada [16] and Kravchuk [17] applied the method of moments to ordinary differential equations by using the weighting functions \( \{x^n\} \) regardless of the choice of approximating functions. For the first approximation, the weighting function is unity, and the method of moments in this case is equivalent to the subdomain method and is usually called the integral method, or von Kármán-Pohlhausen method [18, 19]. For the integral method, reviewed in detail by Goodman [20], the differential equation is satisfied on the average over the domain of interest.

**Boundary Methods**

In the foregoing it is presumed that the trial solution satisfies the boundary conditions but not the differential equation. The converse situation can also be treated: the differential equation is satisfied but the boundary conditions are not. Trial solutions of this sort lead to boundary methods, as they are called by Collatz [3]; the procedures are analogous to those above, but with the spatial average, Equation (8), replaced by an average over the boundary.

**Mixed Methods**

The intermediate situation can also be handled: in so-called mixed methods the trial solution satisfies neither the differential equations nor the boundary conditions. In Schuleshko's treatment of mixed methods [21], the differential-equation residual is made orthogonal to one set of weighting functions, using (8) as the inner product, while the boundary residual is simultaneously made orthogonal to another set of weighting functions, using an appropriate surface integral as the inner product. If N weighting functions are used, this leads to
2N conditions, yet in general only N conditions can be satisfied by the N independent c_i. For this procedure to work, some of the conditions must be discarded, as was noticed by Snyder, Spriggs and Stewart [22] in their discussion of the Galerkin method.

On the other hand, Bolotin [23], Mikhlin [11], and Finlayson [24] have pointed out that for the Galerkin method the dilemma can be resolved by adding the differential-equation residuals to the boundary residuals. The combination is made in such a way that the differential-equation residual, when integrated by parts, cancels identical terms of the boundary residual. The situation is analogous to the treatment of natural boundary conditions in the calculus of variations, and indeed only boundary conditions analogous to natural boundary conditions can be handled in this way. Such a combination of equation and boundary residuals represents a generalization beyond the treatment given by Crandall [1] (page 235), who states that MWR cannot be applied unless the trial solution satisfies all boundary conditions. However, Crandall [1] (page 321) does combine the residuals for eigenvalue problems in which an eigenvalue appears in both the differential equation and boundary conditions. The combination of residuals for more general problems is important in establishing the equivalence between the Galerkin method and several so-called variational methods [25, 24] (see below).

B. Refinements on the Basic Method

Other modifications are possible; Duncan [26] makes the approximating functions satisfy derived (or secondary) boundary conditions which are determined by requiring that the differential equation be satisfied on the boundary. Derived boundary conditions are also used in boundary-layer theory in the von Kármán-Pohlhausen method; other compatibility conditions—such as continuity of the velocity and certain of its derivatives at the edge of the boundary layer—are employed as well. Recently it has been shown [27, 28] that additional compatibility conditions are required to assure good results when the integral method is applied to magnetohydrodynamic boundary-layer problems. In those cases the additional conditions are found by differentiating the differential equation in the direction normal to the surface; all trial solutions must then satisfy this equation at both the solid surface and the edge of the boundary layer. A variation of the collocation method is given by Collatz [3], who differentiates an ordinary differential equation and applies the collocation method to the residual of the resulting equation, too.

Kantorovich and Krylov [29] outline a method for two-dimensional problems in which the residual is required to be zero along a line in the domain (such as x = y). The method of reduction to ordinary differential equations, as developed by Kantorovich [30] and independently by Poritsky [31], reduces a partial differential equation to a system of ordinary differential equations. This is the procedure described above for initial-value problems but it can be applied equally well to boundary-value or eigenvalue problems. The spatial averages (Equation (8)) are taken over all the independent variables except one, and the approximate solution is found by solving a set of ordinary differential equations involving this remaining independent variable. While this semi-direct method was originally proposed in the context of variational principles, Kantorovich [32] in 1942 showed its equivalence to the Galerkin method. Even earlier Bickley [12] had applied the Galerkin method to unsteady-state problems in a manner equivalent to the method of reduction to ordinary differential equations. In general, MWR can be used to reduce the number of independent variables in any partial differential equation. The resulting system of equations is simpler (it may be algebraic or ordinary differential equations or even a set of partial differential equations), but its solution remains only an approximate solution to the original problem.

In the collocation method a critical problem is the choice of collocation points. For ordinary differential equations Wright [33] has shown that the residual is minimized if the collocation points are given by the roots of the Chebyshev polynomials.

Naturally the method of weighted residuals can be combined with other methods. Collatz [3] presents a combination of the iteration method and MWR. Yang [34, 35, 36] uses the approximate solution generated by the integral method as the first step in the following procedure for time-dependent problems: the result of the integral method is substituted into those terms involving time-derivatives and equation thereby obtained is solved as a steady-state, nonhomogeneous, partial differential equation. An advantageous coupling of MWR and numerical finite difference methods has been employed by Kaplan [37], Kaplan and Bewick [38], and Kaplan, Marlowe, and Bewick [39] to reduce the computer time necessary to solve certain nuclear reactor problems; the number of independent variables was reduced from four to three or two by using MWR. Other modifications and hybrid schemes are possible and will undoubtedly be proposed as needs arise.

Choice of Approximating Functions

The choice of approximating functions can be crucial in applying MWR. How to arrive at a good, if not the best, selection is an outstanding problem. Certainly any symmetry properties of the system should be exploited but there seems to be no way available at present to do this systematically for all problems. In problems of conventional types it is usually convenient to have the approximating functions satisfy the boundary conditions, and Kantorovich and Krylov [29] show how to construct complete sets of functions which vanish on a boundary of complicated shape. Snyder and Stewart [40] combine this scheme and symmetry arguments to find approximating functions for the velocity vector field in fluid flowing through regularly packed beds of spheres.

Derived boundary conditions can also be used to place restrictions on the approximating functions admitted, and improvement sometimes results [26, 148]. Usually, however, several sets of approximating functions are admissible and it is not possible to choose one as the "best." Heywood and Mofatt [41] even suggest as a qualitative criterion that the approximate
solution be relatively insensitive to different but reasonable choices of the approximating functions.

Methods have been devised for constructing approximating functions especially for eigenvalue problems involving high-order ordinary differential equations of the sort that arise in the theory of convective instability [42, 43, 44]. The approximating functions are just eigenfunctions of one or another lower-order, simpler yet related eigenvalue problem on the same domain. Polynomials are popular approximating functions; they have even been used in cylindrical and spherical domains [45] where proper regard must be taken of possible singularities. Falk [46] uses Hermite polynomials, which are orthogonal on a semi-infinite domain. Other authors [39, 47] emphasize that numerical difficulties (for large \( N \)) can be avoided in the Galerkin method if the approximating functions are orthonormalized.

Selecting approximating functions remains somewhat dependent on the user's intuition and experience, and this is often regarded as a major disadvantage of MWR. Clearly, the question of methods for arriving at optimal choices of approximating functions warrants thorough investigation. Leads may exist in the local solutions and regional expansions used in perturbation methods [47a].

Comparison of Different Criteria

Comparisons of different criteria as applied to the same problem exist only for relatively simple, linear, initial-value problems [12, 25, 48] and boundary-value problems [4, 13, 1]. In the literature on eigenvalue problems the Galerkin method predominates, although there are a few comparisons with the collocation and least-squares methods [1, 4]. The results of these comparisons may be summarized by Crandall's remark [1] (page 375): "The variation between results obtained by applying different criteria to the same trial family...is much less significant than the variations that can result from the choice of different trial families." However, there may be a great difference in the work necessary to obtain the approximate solution when using different criteria. Crandall's experience evidently is based entirely on linear problems. The only comparison for nonlinear problems appears to be the unpublished thesis by Collings [49], as referenced by Ames [2]. Ames comes to the conclusion that the Galerkin method is superior, but cautions that this stand is based on limited experience and may not hold in general [2].

For linear, ordinary differential equations Frazer, Jones and Skan [4] argue that the collocation, least squares, and Galerkin methods are equivalent in the limit as \( N \to \infty \). Other similarities exist between the methods [1]; e.g., when the approximating functions are chosen to be the eigenfunctions of the linear operator, i.e., \( L(u_i) = \lambda_i u_i \), then the least-squares and Galerkin method coincide.

For self-adjoint (hence necessarily linear) eigenvalue problems, the eigenvalues are real, and Crandall [1] emphasizes that the Galerkin method leads to symmetric matrices in Equation (15)—and hence real-valued approximations—whereas the other methods may give complex eigenvalues as approximations to the exact real eigenvalues. The least-squares method is particularly unsuited for linear eigenvalue problems because it turns the linear problem into a nonlinear one [1].

The least-squares procedure for eigenvalue problems as outlined by Becker [50] differs somewhat from that of Crandall [1]; Becker does not have such a difficulty in the first approximation. Whereas Crandall uses the weighting function \( \partial L/\partial c_i \) for the first approximation, Becker uses \( \partial L/\partial \lambda \), where \( \lambda \) is the eigenvalue. Consider the linear eigenvalue problem

\[
L(u) + \lambda u = 0
\]

with \( u = 0 \) on the boundary. For the first approximation with a trial solution \( u^* = c_i u_i \), the residual is

\[
R(u_i) = c_i (L(u_i) + \lambda_i u_i)
\]

The mean square error is then

\[
I = c_i^2 \int V [L(u_i) + \lambda_i u_i]^2 dV
\]

Crandall apparently would determine \( \lambda \) from

\[
\partial L/\partial c_i = 0 = 2c_i \int V [L(u_i) + \lambda_i u_i]^2 dV
\]

which is a quadratic in \( \lambda \) and may lead to complex values of \( \lambda \). Becker, on the other hand, would determine \( \lambda \) from

\[
\partial L/\partial \lambda = 0 = 2c_i \int V [L(u_i) + \lambda_i u_i] u_i dV
\]

which is linear in \( \lambda \) and gives real values as long as the equation and \( u_i \) are real. Becker's procedure appears to be simpler for the first approximation.

For higher approximations both procedures lead to nonlinear equations for this linear problem. Crandall would substitute as weighting functions \( \partial L/\partial c_i, i = 1, 2, \ldots, N \), and Becker would substitute \( \partial L/\partial \lambda, \partial L/\partial c_i, i = 2, 3, \ldots, N \). The latter is thus using the eigenvalue \( \lambda \) as one of the parameters and is also exploiting the fact that the mean square error can be minimized as a function of \( \lambda, \gamma = c_i/c_1 \), rather than \( c_j \) since

\[
l(c_1, c_2, \ldots, c_N, \lambda) = c_1^2 I(1, \lambda_1, \lambda_2, \ldots, \lambda_N, \lambda)
\]

For initial-value problems, the least-squares method must be applied carefully and has certain disadvantages. The method is applicable if the time dependence of the approximate solution is specified—in other words, semi-direct methods cannot be used in the method of least squares. Consider the problem \( \partial u/\partial t = L(u) \) and assume a trial solution of the form \( u^* = u_x + \sum a \rho \eta (x, t) \). Then the functional \( I \), representing the mean square residual, can be minimized:

\[
I = \int_0^T \int V \left( \frac{\partial u}{\partial t} - L(u) \right)^2 dV dt
\]

Of course the solution depends on the value of \( T \). If the upper limit of integration is infinite, the solution may no longer have this ambiguity. This was the ap-

Oftentimes, however, the time dependence of the solution is difficult to guess and the trial solution must involve underdetermined functions of time, \( u^*_i = u_0 + \sum_{i=1}^{N} c_i(t) u_i(x) \). The mean square residual is

\[
l = \int \left( \frac{\partial u}{\partial t} - L(u) \right)^2 dV \tag{27}
\]

Now, however, \( l \) depends on time and involves time derivatives. Consequently it cannot in general be made a minimum for all time by any set of functions \( c_i(t) \); this was shown by Citron [51] and Finlayson and Scriven [25]. Consequently, if a semi-direct method is used to solve this type of problem the term least squares is a misnomer because the mean square residual is not being minimized.

The least-squares method is discussed at length in a monograph by Becker [50]. Listing criteria which he maintains a good variational method must satisfy, he concludes that the least-squares method is the best general criterion of MWR. Becker's list includes the following points: (i) errors should be minimized in some sense; (ii) the functional should be positive definite; (iii) the procedure should be capable of treating initial-value problems, as well as others. These seem somewhat slanted toward the least-squares method; indeed, items (i) and (ii) cannot be realized for all problems except in the least-squares method. Yet no one has shown that a solution is necessarily best because its mean square residual is smallest; such a definition or proof will certainly depend on the particular application. In addition, the least-squares method can be used to treat initial-value problems in only a limited way, as shown above. Furthermore, an important point should be added to the list of desiderata—the method should be simple to apply. As already shown, this criterion immediately eliminates the least-squares method for linear eigenvalue problems because it turns a linear problem into a more difficult nonlinear. Becker realized that his conclusion may not always be valid [50] (page 61): "While the least-squares method seems to be the most useful general approach, in some applications (in which some specific criteria may be added to our 'general list') other methods may be preferable." Becker illustrates the advantages of the method of least-squares by solving a set of nonlinear, time-dependent partial differential equations which model the fuel depletion in a nuclear reactor; he finds results that compare well with the more lengthy numerical solutions.

In this discussion of the various criteria of MWR, the Galerkin method has been distinguished from the method of moments by means of the weighting functions used in the two. In the Galerkin method, the weighting functions must be the same set of functions which are used for the trial solution, whereas in the method of moments the weighting functions can be some other set of functions. This distinction is not always made [52] and is probably unimportant in practice, although the two methods have different histories and may have different convergence properties. There are inconsistencies of terminology in the literature; for example, Kawaguit [53] used the method of moments rather than the Galerkin method as he claimed, for the weighting functions differ from the approximating functions in his work. Another example of confusing nomenclature is the name method of integral relations, which refers to a generalization of the subdomain method; it is adequately reviewed by Belotserkovskii and Chushkin [54].

### C. Convergence Theorems

#### Galerkin Method

After introduction of the Galerkin method in 1915 some twenty-five years elapsed before the convergence of the method was studied. Even today much remains to be done; only a few theorems have been proved, and these pertain exclusively to linear problems. Repman [55] was the first to prove convergence of solutions obtained by the Galerkin method though only for a certain Fredholm-type integral equation. Petrov [56] then studied the convergence of the Galerkin method for eigenvalue problems of fourth-order ordinary differential equations—in particular, the Orr-Sommerfeld equation of hydrodynamic stability theory. Keldysh [57] treated general ordinary differential equations and also second-order elliptic partial differential equations. Miklin [11] later simplified Keldysh's proofs. The equations are of the form

\[
L(u) = - \sum_{i,j=1}^{m} A_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{m} B_i \frac{\partial u}{\partial x_i} + Cu = f \tag{28}
\]

Both Keldysh and Miklin prove that the first derivatives of the Galerkin approximate solution converge in the mean to the first derivatives of the exact solution. Whenever the Rayleigh-Ritz and Galerkin methods coincide (see below), the convergence proofs for the Rayleigh-Ritz method imply convergence of the Galerkin method, too. Thus the Galerkin-convergence proofs given by Kantorovich and Krylov [29] apply only to specific problems with a minimum or maximum principle, whereas the convergence proofs mentioned here are applicable to problems whether or not they have a corresponding variational principle. It has been claimed [40] that completeness of the set of approximating functions is sufficient to assure convergence, but the proofs given by Miklin and others show clearly that this is not enough.

Recently convergence proofs have become available for certain eigenvalue problems associated with hydrodynamic stability investigations [58, 15]. Results applicable to unsteady-state problems are less extensive. Faedo [59] applied the Galerkin method to a hyperbolic differential equation and inspired the important work of Green [60], who proved the uniform convergence of the Galerkin method when ap-
plied to the following equation
\[ \frac{\partial^2 u}{\partial x^2} - \frac{\partial u}{\partial t} - g(x, t) u = f(x, t) \] (29)

The Galerkin method has been used to prove the existence of weak solutions to (i) the Navier-Stokes equations with time dependence [61], and (ii) equations representing the unsteady-state transport equation with a known velocity field [62]:

\[ \sum_{i=1}^{m} \frac{\partial}{\partial x_i} \left( A_{ij} \frac{\partial u}{\partial x_j} \right) - \sum_{i=1}^{m} B_i \frac{\partial u}{\partial x_i} + C u = \frac{\partial u}{\partial t} = F(x, t) \] (30)

Recently the Galerkin method has been applied to the Taylor problem with time-dependent disturbances [63], and a method has been developed to generate improvable, pointwise upper and lower bounds—and hence error bounds—for the solution to Equation (30) [129a].

Other Methods

Convergence proofs are rarely available for the other criteria of MWR. The notable exception is the least-squares method, which is well-treated (for boundary problems) in the notable text by Mikhlin [11]. Mikhlin proves conditions which insure that the method of least squares gives a sequence of approximate solutions which converge in the mean to the exact solution. Furthermore, the mean-square-error of the approximate solution can be determined. He also points out that the least-squares method converges more slowly than the Ritz method (when the latter can be applied) but may give uniform convergence rather than convergence in the mean. Some results for the collocation method are given by Kader [64], while the method of moments is treated by Kravchuk [65].

Nonlinear Problems

Very little is known about the convergence of MWR for nonlinear problems without a corresponding variational principle. Krasnosel’ski [66] presents theorems—mostly without proofs—for the Galerkin method applied to nonlinear integral equations. Glansdorff [67] mentions a forthcoming proof of the convergence of the local potential method, which is identical in application to the Galerkin method; he treats the steady-state heat conduction equation with temperature-dependent thermal conductivity. Of course, as Ames [2] has pointed out, convergence proofs are not as useful as error bounds. Even a computer does not make it possible to calculate infinitely many terms and when truncating the series one always wonders how good the resulting approximate solution is. Comparison of successive approximations is an aid in such a case, but even an approximate solution that seems to be converging may not be converging to the exact solution. The available convergence theorems and error bounds are so scarce that engineer and applied scientist must usually extrapolate from previously tested results for other problems to new situations when applying approximate methods.

D. Comparison to Other Methods

Separation of Variables

The Galerkin method is related to a wide variety of other approximate methods as well as to some exact methods of analysis. In particular, it can be shown [29, 22, 24] that if a problem yields to the method of separation of variables and if the Galerkin method is applied in a certain way*, then the two solutions are the same, provided the Galerkin method is carried through to completion. Of course in numerical calculations, after obtaining an exact solution in the form of an infinite series, one calculates only a finite number of terms as a matter of practical necessity.

Variational Methods

There is also a close relationship between the Galerkin method and the Ritz or Rayleigh-Ritz method when the latter can be applied [14, 69, 70, 1, 21, 3, 29, 71, 72]. In particular, if the same trial functions are used, the resulting calculations are identical. Contrary to a currently prevalent opinion, this equivalence still persists when the trial functions do not satisfy the natural boundary conditions [25, 11, 25, 24], which they need not do in the Rayleigh-Ritz method. The boundary residual is either added or subtracted to the differential equation residual, and the calculations are again equivalent to the Ritz or Rayleigh-Ritz method. The choice of adding or subtracting is dictated either by mathematical convenience—part of the differential-equation residual can be integrated by parts to cancel part of the boundary residual—or by the physics—the differential equation and boundary conditions both come from macroscopic balances taken over the volume and surface, respectively; these macroscopic balances can be combined in only one way, and the residuals are combined in exactly the same way. A very important difference between the Rayleigh-Ritz method and the Galerkin method is that in the former some functional—possibly representing an eigenvalue—is being minimized or maximized. Consequently the approximate values of the functional represent either upper or lower bounds. In the Galerkin method this information is missing; exactly the same values would be obtained, but one would not know that these were upper or lower bounds. However, when the variational integral is of no significance, the Galerkin method, because of its generality, may be preferred. The variational and Galerkin methods are compared schematically in Figure 1.

Most variational principles are merely stationary principles, rather than minimum or maximum principles.

*The approximating functions in the Galerkin method must be the eigenfunctions found by the separation of variables and the Galerkin method must be applied to the initial conditions as well as to the differential equation. Such a result means simply that if the exact solution is contained in the trial function, the Galerkin method will find it.

Though there is basically but a single method, it is convenient to follow the custom (scarcely universal) of distinguishing between the "Rayleigh-Ritz method" when it is applied to minimum or maximum principles and the "Ritz method" when it is applied to merely stationary principles.

741
In such cases, the Ritz method is again equivalent to the Galerkin method. The calculations are identical; the results are identical; but in the variational method one knows that the variational integral is being made stationary, i.e., insensitive to changes in the trial solution. If the variational integral has physical significance and is the quantity of interest, then the variational methods have an advantage over the Galerkin method even though the answers are the same.

**Adjoint Variational Methods**

Variational principles exist for linear problems only if they are self-adjoint. For non-self-adjoint linear problems, variational principles can be formulated for the original equations and their adjoints, and again MWR is related to the corresponding variational methods. The impetus for using the adjoint operator in variational formulations seems to stem from Morse and Feshbach [73], who gave a variational principle for the unsteady-state heat conduction equation; Roussopoulos [74] also gave a variational principle for any linear non-self-adjoint problem. Schmit [75] and Washizu [76] have applied such a principle to the unsteady-state heat conduction equation, while Selengut [77, 78] developed the idea for nuclear reactor problems. Many other examples exist, such as those of Nichols and Bankoff [79] for convective diffusion of heat, Finlayson [24] for convective diffusion of a multicomponent mixture of chemical species; Lewins [80], Slattery [81], Flumerfelt and Slattery [82] for extensions to nonlinear problems; and many authors [83–94, 38] for nuclear reactor and associated problems. In applications to linear non-self-adjoint problems, the method of weighted residuals yields the same results as any of these variational principles as long as the weighting functions for the original equations are taken as the approximating functions for the adjoint, and vice versa [75, 24]. The question then arises as to whether this variational method, which can be regarded as an application of one form of MWR, is preferable to Galerkin's method, which does not require the complication of an auxiliary adjoint system. There is some evidence that the adjoint variational principle leads to slightly better results [95–97], and Clark and Hansen [98] imply that the use of adjoint weighting functions might speed convergence. Kaplan and Bewick [38] claim that the variational method is the best strategy in that it gives better answers more often. However, they go on to say:

"Of course, practical considerations may introduce yet another meaning of the word "best"; namely, "most economical." In this sense we find that the Galerkin method (which uses the trial functions also as weighting functions) is, in most instances, preferable to the variational method, since it gives results which are almost as good but does not require separate calculations of the weighting functions."

In essence, the adjoint variational method trades increased complexity for possibly better results: there still is no clear-cut answer to the question of whether the Galerkin method or the adjoint-variational method is best.

For certain initial-value problems there may be no difference between the variational method and Galerkin's method if the semi-direct approach is used in both. Whenever the corresponding steady-state problem is self-adjoint, it is reasonable to expand the unsteady-state solution and its adjoint in terms of the same functions of position with unknown functions of time as coefficients:

\[
\begin{align*}
\mathbf{u} &= \mathbf{u}_S + \sum_{i=1}^{N} c_i(t) \mathbf{u}_i(\mathbf{x}) \\
\mathbf{u}^* &= \mathbf{u}_S + \sum_{i=1}^{N} c_i^*(t) \mathbf{u}_i(\mathbf{x})
\end{align*}
\]  

(31)  

(32)

Because the approximating functions are the same for both \( \mathbf{u} \) and \( \mathbf{u}^* \), the weighting functions in the adjoint method are the same as those in the Galerkin method. Consequently, the solutions are identical, whether derived by the adjoint method or the more direct Galerkin method.

The adjoint system is also useful for eigenvalue problems. Roberts [99] presents the general theory, and examples can be found in the works of Chandrasekhar [43] and DiPrima [100] as well as others. While introduction of the adjoint system does increase the complexity of the problem—particularly the boundary conditions—some advantage is gained over the straightforward application of MWR because the eigenvalue is made stationary, and hence insensitive to changes in the trial function. This advantage does not usually apply to boundary and initial-value problems since the variational integral is seldom of interest in those cases.
Method of Least Squares

Mikhlin [11] points out that for boundary-value problems the least-squares method for
\[ Lu = f \quad \text{in } V \]  \hspace{0.5cm} (33)
\[ B_i u = 0 \quad \text{on } S \]  \hspace{0.5cm} (34)
is equivalent to applying the calculus of variations to the equation
\[ L^* (Lu - f) = 0 \]  \hspace{0.5cm} (35)
Consider the following minimum principle: minimize the functional
\[ I = \int_V (Lu - f)^2 \, dV + \int_S (B_i u)^2 \, dS \]  \hspace{0.5cm} (36)
among all functions \( u \) having the appropriate continuity and differentiability requirements. The natural boundary conditions corresponding to this variational principle are of the form [30]
\[ N_i (Lu - f) = 0 \]  \hspace{0.5cm} (37)
where \( N_i \) are differential operators. The Euler equation is just equation (35) and the equivalence with equation (33) rests with the premise that the equation
\[ L^* u = 0 \]  \hspace{0.5cm} (38)
\[ B_i u = 0 \]  \hspace{0.5cm} (39)
has only the trivial solution. Note that the natural boundary conditions (37) are similar to the compatibility conditions mentioned above in connection with the integral method for magnetohydrodynamic boundary-layer problems.

Method of the Local Potential

A procedure based on the so-called local potential of Prigogine and Glansdorff [101-104] has been proposed as a variational method for determining approximate solutions to boundary-value [105-108], eigenvalue [109], and more recently initial-value problems [108]. Rosen [110-113] used the same type of computational scheme earlier. The actual applications of these methods have been shown to be equivalent to the Galerkin method [114, 24]. Moreover, it has been demonstrated that the variational integral is not stationary in the local potential method and that no minimum principle exists in applications [114, 24]. Consequently, the advantages usually associated with variational principles are missing from the local potential method, which can be regarded as a disguised application of the Galerkin method. See Ref. 129b for a more detailed critique.

Lagrangian Thermodynamics

The so-called variational methods due to Biot [115-121] and others [122-128, 51] are also equivalent to the Galerkin method [25, 24]. In these Lagrangian thermodynamic methods there is no variational integral which is being made stationary [25, 79]; their sole significance appears to be as means for generating a computational scheme. That scheme is, however, identical to the Galerkin method, which is more straightforward and applicable to a broader range of situations. There is no reason that the Galerkin method should not be preferred, so far as the authors know. See Ref. 129b for a more detailed critique.

E. Applications

The general features of MWR in its numerous versions and various refinements have been presented, and its relationships to certain other approximation methods have been sketched. Which of all these methods are superior, and over just what ranges of circumstances the superiority exists, are matters that can be settled finally only on the basis of representative applications. More systematic comparative studies and evaluations are needed than have been reported to date. Until they are forthcoming the investigator of a new problem can expect little more help than he can get out of seeing how others have handled more or less similar problems. References 130-187 have been selected as much to illustrate pitfalls, shortcomings, and failures as to cite the attractive features and successes of different versions of MWR. The preponderance of recent papers accurately reflects the upsurge of applications of these methods in one field after another; the emphasis on problems of flow and transport is conditioned by interests of the authors. The popularity of the integral methods which originated in boundary-layer studies [18, 20, 34-36, 136-153, 160-163, 166-169] can be discounted in part as a tradition perpetuated by formal instruction beginning with elementary texts in fluid mechanics and heat transfer.

Beyond any guidance he can get from past experience the problem-solver can look for reassurances in comparisons of different forms of trial solutions and of successive approximations in any one form: the appearances of convergence with more numerous adjustable parameters and of insensitivity to form of approximating functions do lend confidence to results. So do close matches with established information on special cases and limiting cases. It is also true that MWR, like variational methods, may yield better estimates of properties of the solution at large, such as an integral or eigenvalue, than of the solution itself. The main advantage and disadvantage of MWR are contained in the same feature, namely, that the results depend on more or less arbitrary decisions by the user. Intuition, experience, any available information all can be rapidly exploited but the reliability of the results is frequently hazy. Hopefully this review sheds light on the basic issues and will be useful to those interested in applying weighted residual methods and related techniques.

F. Areas for Further Research

Of the unsolved problems concerning MWR the following are most important in the opinion of the authors:

(1) Choice of criterion in MWR. Systematic comparative studies using representative (nonlinear) problems are needed. The least-squares procedure for nonlinear problems particularly warrants attention. The Galerkin method and adjoint variational method for linear problems need to be compared.
(2) Development of rational methods for selecting approximating functions, for example by systematically applying symmetry properties of the problem. A related question is whether linear combinations of functions are more effective than nonlinear forms, especially for nonlinear problems.

(3) Definition of mathematical or engineering criteria for identifying optimal approaches in (1) and (2). Optimization theory should be brought to bear on the decisions that have to be made by users of weighted-residual and other approximation methods.

(4) Derivation of error bounds for approximate solutions by MWR. And though they are not as directly useful, convergence proofs are needed, especially for nonlinear problems.

Acknowledgment

This review stems from research supported in part by the National Science Foundation through a Graduate Fellowship received by B. A. Finlayson and in part by the Air Force Office of Scientific Research through AFOSR Grant Number 2196-63.

References

17. Kravchuk, M. F., Application of the method of moments to the solution of linear differential and integral equations (in Ukrainian), Kiev. Soobsch. Akad. Nauk UkSSR 1, 168(1952); referred to by P. Shuleshko in Ref. 21 and L. V. Kantorovich and V. I. Krylov in Ref. 29. Another citation is given by S. G. Mikhlin in Ref. 11.
87. Dougherty, D. E. and Shen, C. N., The space-time
104. Prigogine, I., Evolution criteria, variational properties, and fluctuations, pp. 3-16 in Ref. 68.

**Additional References to Applications**

**Nonlinear Problems**

Galerkin method: *Refs.* 13, 27, 53, 106-108, 115, 121 and


Integral method: *Refs.* 20, 34-36, and


143 Kapur, J. N. and Gupta, R. C., Two-dimensional flow of power-law fluids in the inlet length of a straight channel, I and II, ZAMP 43, 135–141 (1963), 44, 277–284 (1964); AMR 18(1965), Rev. 7500.


146 Rajeshwari, G. K. and Rathna, S. L., Flow of a particular class of non-Newtonian visco-elastic and viscoelastic fluids near a stagnation point, ZAMP 13, 45–57 (1962); AMR 17(1964), Rev. 872.


149 Siegel, R. and Sparrow, E. M., Simultaneous development of velocity and temperature distributions in a flat duct with uniform wall heating, AIChE J. 5, 75–79 (1959); AMR 13(1960), Rev. 405.


Subdomain method: Ref. 54 and


Method of moments: Refs. 16, 129 and


**Linear Initial-value Problems**


Integral method: Refs. 34, 48, 147 and


Least-squares method: Refs. 12, 50, 51.

Method of moments: Refs. 12, 48, 129.

Collocation: Ref. 12.

**Linear Boundary-value Problems**

Galerkin method: Refs. 4, 37, 40, 112, 120, 124, 127, and


Integral method: Refs. 147, 148 and


Method of moments:


Collocation method: Ref. 4 and


Linear Eigenvalue Problems

Galarkin method: Refs. 10, 23, 43, 44, 56, 58, 100, 130, and


183 Duncan, W. J., Applications of the Galerkin method to the torsion and flexure of cylinders and prisms, Phil. Mag. (7) 25, 634–649 (1938).


Least-squares method: Refs. 4, 50, 64.

Collocation method: Ref. 64.

748