Nodally Exact Ritz Discretizations of 1D Diffusion-Absorption and Helmholtz Equations by Variational FIC and Modified Equation Methods

C. A. Felippa and E. Oñate

Publication CIMNE PI 237, July 2003

Revised May 2004

International Center for Numerical Methods in Engineering

c/ Gran Capitán s/n, Edificio C-1 UPC, 08034 Barcelona, Spain

TABLE OF CONTENTS

Nodally Exact Ritz Discretizations of 1D Diffusion-Absorption and Helmholtz Equations by Variational FIC and Modified Equation Methods

CARLOS A. FELIPPA* AND EUGENIO OÑATE†

** Department of Aerospace Engineering Sciences*
and Center for Aerospace Structures
University of Colorado
Boulder, Colorado 80309-0429, USA

† International Center for Numerical Methods in Engineering (CIMNE)
Edificio C-1, c. Gran Capitán s/n
Universidad Politécnica de Cataluña
08034 Barcelona, Spain

Abstract. This article presents the first application of the Finite Calculus (FIC) in a Ritz-FEM variational framework. FIC provides a steplength parametrization of mesh dimensions, which is used to modify the shape functions. This approach is applied to the FEM discretization of the steady-state, one-dimensional, diffusion-absorption and Helmholtz equations. Parametrized linear shape functions are directly inserted into a FIC functional. The resulting Ritz-FIC equations are symmetric and carry a element-level free parameter coming from the function modification process. Both constant- and variable-coefficient cases are studied. It is shown that the parameter can be used to produce nodally exact solutions for the constant coefficient case. The optimal value is found by matching the finite-order modified differential equation (FOMoDE) of the Ritz-FIC equations with the original field equation. The three ingredients of the method (FIC, Ritz and MoDE) are extendible to multiple dimensions.

Keywords: finite calculus, variational principles, Ritz method, functional modification, stabilization, finite element, diffusion, absorption, Helmholtz, nodally exact solution, modified differential equation.

§1. The Finite Calculus

The Finite Calculus (FIC) has been developed over the past five years [18–30] as a general purpose tool for improving the stability and accuracy of interior discretizations of equations of mathematical physics and engineering. Consider a problem governed by the residual equation

$$\mathbf{r}(\mathbf{u}) = \mathbf{0}, \tag{1}$$

where $\mathbf{u}$ is an array of n primary variables. These in turn are functions of the independent variables $\mathbf{x}$, which may include time. Generally (1) is an ordinary or partial differential equation, to be solved by numerical methods.

Introduce n characteristic lengths h_i collected in array $\mathbf{h}$, where each h_i is paired with the function u_i. These lengths can be viewed as linked, through as yet unspecified means, to mesh or grid dimensions. Using flux balance arguments [21,22] a modified residual is constructed

$$\mathbf{r}(\mathbf{u}) + \mathbf{r}_h(\mathbf{u}, \mathbf{h}) = \mathbf{0}. \tag{2}$$

The simplest form of $\mathbf{r}_h$ is $-\frac{1}{2}\nabla\mathbf{r}\,\mathbf{h}$, where $\nabla\mathbf{r}$ is the gradient matrix of $\mathbf{r}$ with respect to the independent variables. The discretization process, which is usually Galerkin-based FEM, is applied to (2) instead of (1). Consistency with the latter requires that $\mathbf{r}_h \to 0$ as $h_i \to 0$.

But the philosophy of FIC, as emphasized in its name, is that in practice the h_i remain *finite*. The key goal is to pick $\mathbf{r}_h$ and $\mathbf{h}$ so that stability and accuracy characteristics of the solution for *a given mesh* are improved. Further analysis of localized phenomena, such as sharp boundary layers, can be carried out by multiscale devices [6,15,25]. The FIC analysis process is diagramed in Figure 1.

FIC has been primarily used [18–30] for the solution of fluid mechanics equations involving flow, advection, diffusion, ocean waves and chemical reactions. For those applications it competes with stabilization schemes such as SUPG, residual free bubbles and subgrid scale methods [3,4,8,9,15].

In a study of FIC methods for solid mechanics [31] it was found that a variational form formally analogous to the Minimum Potential Energy principle could be obtained by modifying the displacement, strain and stress fields in a manner similar to that done for the residual in the foregoing description, and adjusting their variations.

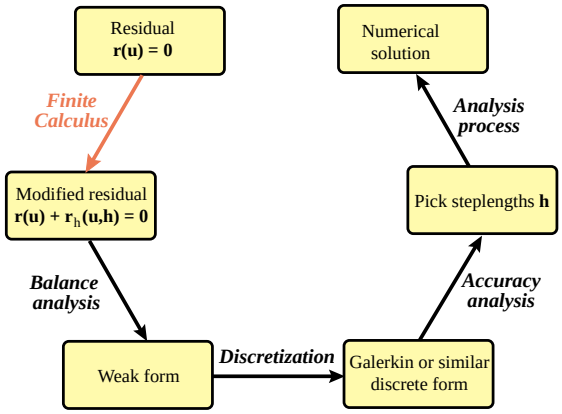

Figure 1. The weak-form-based FIC analysis process.

The approach technically falls into the class of variational principles with noncommutative variations [38], also called modified variational principles in the literature [7]. That finding provides the departure point for the present study.

§2. Modified Variational Forms

Suppose that (1) is derivable from a functional $J[\mathbf{u}]$ in the sense that $\mathbf{r}(\mathbf{u}) = \mathbf{0}$ are the Euler-Lagrange equations of J. The first variation is

$$\delta J[\mathbf{u}] = \delta \mathbf{u}^T \, \mathbf{r}(\mathbf{u}). \tag{3}$$

Define a modified primary variable field:

$$\tilde{\mathbf{u}} \stackrel{\text{def}}{=} \mathbf{u} + \mathbf{u}_h(h), \tag{4}$$

such that $\mathbf{u}_h \to \mathbf{0}$ as $h \to 0$. The choice considered here, suggested by a previous study, is

$$\tilde{\mathbf{u}} = \mathbf{u} - \tfrac{1}{2} h \boldsymbol{\alpha}^T \nabla \mathbf{u}. \tag{5}$$

Here h is an overall characteristic length, array $\boldsymbol{\alpha}$ collects scaling parameters α_i, and the factor $-\frac{1}{2}$ is for convenience in matching to the standard FIC method.

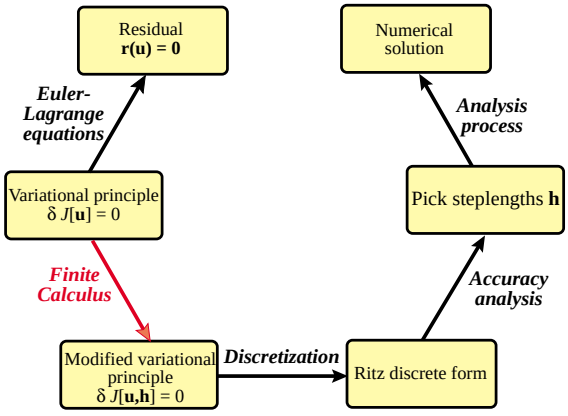

Figure 2. The variational FIC analysis process.

Substituting (5) into J yields the modified functional

$$\tilde{J}_h = J[\tilde{\mathbf{u}}] = J + J_h \tag{6}$$

in which the augmentation term J_h vanishes as $h \to 0$. The Euler-Lagrange equation changes to

$$\delta \tilde{J}_h[\mathbf{u}] = \delta \mathbf{u}^T \left(\mathbf{r}(\mathbf{u}) + \tilde{\mathbf{r}}_h(\mathbf{u}) \right) \tag{7}$$

This has formally the same configuration as (3), and shares with it the property that as $h \to 0$ the Euler-Lagrange equation reduces to (1). But in general starting with $\mathbf{r}_h(\mathbf{u})$ of FIC, namely that in (2), does not reproduce $\tilde{\mathbf{r}}_h$. To avoid confusion we qualify (7) as the *FIC variational residual*. The functional $\tilde{J}_h$ will be called the FIC-modified functional, or FIC functional for brevity. (The superposed tildes are eventually dropped for brevity when there is no danger of confusion.)

The numerical approximation is obtained by working with $\tilde{J}_h$ in the usual way, assuming that h is known. The residual may be used to study stability and accuracy properties of the approximation. The analysis process is diagramed in Figure 2.

§3. The Modified Equation Method

The "Accuracy Analysis" stage of Figure 2 is carried out by the *method of modified equations*. Since this is not a well known technique for differential equations, an outline along with a brief history and two examples is provided in this Section for completeness.

§3.1. Backward Error Analysis

The conventional way to analyze accuracy of a discrete approximation is through *forward error analysis*: the amount by which the discrete solution fails to satisfy the source differential form. To make this measure practical, it is computed using local estimators such as truncation or residual errors (in FEM, through recovery from element patches). This technique furnishes *a posteriori* error indicators, and is well developed in the literature.

Backward error analysis takes the reverse approach to accuracy. Given the computed solution, it asks: which problem has the method actually solved? In other words, we seek an ODE or PDE which, *if exactly solved*, would reproduce the computed solution. This ODE or PDE is called the *modified differential equation*, hereby abbreviated to MoDE. The difference between the modified equation and the original one provides an estimate of the error. An important practical advantage is that the MoDE *can be generated without actually solving the discrete problem*.

This approach is now routinely used for matrix computations after Wilkinson's definitive work in the 1960s [42–44] and has become standard part of numerical linear algebra courses. But it is less known in differential equations. This neglect is unfortunate, since the concept follows common sense. Application problems involve physical parameters such as mass, damping, stiffness, conductivity, diffusivity, etc., which are known approximately. Transient loading actions (e.g, earthquakes, winds, waves) may be subject to high uncertainties. If the modified equation models a "nearby problem" with parameters within the range of experimental uncertainty, it is as good as the original one. This "defect correction" can be used as basis for controlling accuracy *a priori*, before any computations are actually carried out.

§3.2. MoDE Processing

Let $\mathbf{r}_h(\mathbf{u}, \mathbf{h}, \boldsymbol{\alpha}) = \mathbf{0}$ denote a discretization of an ordinary or partial differential equation $\mathbf{r}(\mathbf{u}) = \mathbf{0}$. [As in Sections 1–2, $\mathbf{h}$ collects lengths (in space, time or both) related to mesh or grid dimensions whereas $\boldsymbol{\alpha}$ collects free parameters]. MoDE generation involves three stages:

Step 1: Patch discretization → DDMoDE. The discrete equations at a typical node (a patch in FEM terminology) are rendered continuous in the independent variable(s). This produces a difference-differential form (called delay-differential form when time is the independent variable) of MoDE, called DDMoDE.

Step 2: DDMoDE→ IOMoDE. The difference portion of the DDMoDE is converted to differential form by Taylor series expansion in the mesh dimensions collected in $\mathbf{h}$. This step gives a differential

equation of infinite order, abbreviated to IOMoDE.

Step 3: IOMoDE→FOMoDE. The IOMoDE is reduced to a finite order differential equation, or FOMoDE. This is done by systematic elimination of higher order derivatives. The process typically produces an infinite series in the discretization dimensions. This series can be ocassionally identified and summed in closed form. Technically this is (by far) the most difficult step. It generally requires the use of a computer algebra system (CAS) to be viable.

By comparing the FOMoDE to the original problem one can learn structural aspects of the discretization that go beyond comparison of physical parameter values. For example: preservation of Hamiltonian flow or of conservation laws in the discrete system. These are impossible or difficult to analyze with the conventional truncation error measures.

The procedural steps just outlined are flow-charted in Figure 3. This chart also shows parameter matching step to achieve nodal exactness, which is discussed in more detail in Section 3.4 below.

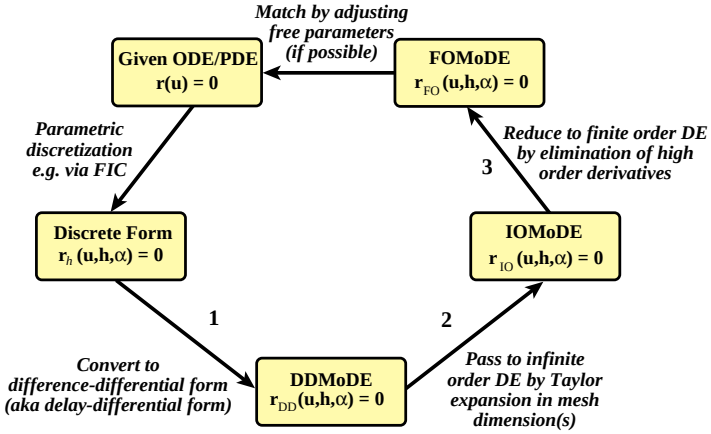

Figure 3. Stages of the modified equation method. Achieving nodal exactness requires "closing the loop." As discussed in §3.4, this may involve additional assumptions.

§3.3. A Brief History of Modified Differential Equations

Modified differential equations as truncated IOMoDE forms originally appeared in conjunction with finite difference discretizations for computational fluid dynamics (CFD). The prescription for constructing them can be found in Richtmyer and Morton's textbook [34, p. 331]. Modified forms were used to interpret numerical dissipation and dispersion in the Lax-Wendroff treatment of shocks, and to derive corrective operators. Similar ideas were used by Hirt [14] and Roache [35]. A drawback of this early work is that there is no guarantee that truncation retains the relevant behavior for finite mesh dimensions, since the discarded portion could be well be dominant in coarse discretizations.

Warming and Hyett [40] were the first to describe the correct procedure for eliminating high order time derivatives of PDE space-time discretizations on the way to the FOMoDE. (Space dimensions were treated by Fourier methods.) They attributed the "modified equation" name to Lomax [17]. The FOMoDE forms were used for studying accuracy and stability of several CFD operators. In this work the original equation typically models flow effects of conduction and convection, **h** includes grid dimensions in space and time, and feedback is used to adjust parameters in terms of improving stability as well as reducing spurious oscillations (e.g. by artificial viscosity or upwinding) and dispersion.

The first MoDE use to study space FEM discretizations for structural mechanics can be found in [39]. However the derivative elimination and force lumping procedures were faulty, which led to incorrect conclusions. This was corrected by Park and Flaggs [32,33], who, being aware of the methods of [40] used modified equations for a systematic study of C^0 beam, plate and shell FEM discretizations.

The method has recently attracted attention from the numerical mathematics community since it provides an effective tool to understand long-time structural behavior of computational dynamic systems, both deterministic and chaotic. Recommended references are [10–13,36]. Web accessible Maple scripts for the IOMoDE→FOMoDE reduction process are presented in [2].

Little of the work to date has used modified equations for optimal selection of free parameters. One exception is [11].

§3.4. Attaining Nodal Exactness

Suppose that the discretization $\mathbf{r}_h(\mathbf{u}, \mathbf{h}, \boldsymbol{\alpha}) = \mathbf{0}$ contains free parameters collected in $\boldsymbol{\alpha}$. As discussed in Sections 1–2, this is always the case for FIC discretizations, whether variationally based or not. Obviously the free parameters will carry over to the three MoDE forms: $\mathbf{r}_{\rm DD}(\mathbf{u}, \mathbf{h}, \boldsymbol{\alpha}) = \mathbf{0}$, $\mathbf{r}_{\rm IO}(\mathbf{u}, \mathbf{h}, \boldsymbol{\alpha}) = \mathbf{0}$ and $\mathbf{r}_{\rm FO}(\mathbf{u}, \mathbf{h}, \boldsymbol{\alpha}) = \mathbf{0}$. Assuming that the last form is available, the question is whether the parameters can be chosen so that

$$\mathbf{r}_{\rm FO}(\mathbf{u}, \mathbf{h}, \boldsymbol{\alpha}_M) \equiv \mathbf{r}(\mathbf{u}), \quad \text{for any } \mathbf{h}. \tag{8}$$

Here subscript M stands for "matching." If this is possible, the discretization $\mathbf{r}_h(\mathbf{u}, \mathbf{h}, \boldsymbol{\alpha}_M)$ becomes *nodally exact*. That is, will give the exact answer at the nodes of any discretization. For FEM discretizations this scheme may be labeled a *nodally exact patch test*, since the MoDE equations are necessarily obtained from an element patch.

The idea is straightforward and attractive but fraught with technical difficulties. In particular:

- Exact matching may be possible only with drastic restrictions on dimensionality, system properties and discretization. For instance: constant coefficients, no source terms, regular meshes. If an exact match is impossible, some "measure of fit" (projection, minimization, etc) has to be chosen.
- Solutions may be imaginary, non-unique, inexistent, or fiendishly hard to compute.
- The FOMoDE may contain "parasitic terms" not present in the governing equation, which cannot be cancelled out by choosing parameters. For example: the source is the Laplace equation $u_{xx} + u_{yy} = 0$ whereas the FOMoDE holds a parameter-free cross-derivative term u_{xy}. The emergence of parasitic terms was in fact observed by Park and Flaggs in their studies of C^0 plate and shell elements [32,33]. Such occurrences can be often traced to consistency defects in the discretization; in that study the presence of such terms flagged element locking.
- Attaining a closed form for the FOMoDE will not be generally possible in more than one dimension, or for variable coefficients. Truncation may be required. In that case the fit can at most be expected to deliver a better solution over a fixed mesh.
- Symbolic manipulations may be prohibitive, even with the help of a computer algebra system.

On the positive side, the approach is completely general, and not linked to any discretization method. The provenance of $\mathbf{r}_h(\mathbf{u}, \mathbf{h}, \boldsymbol{\alpha})$: finite elements, finite differences, boundary elements, etc., is irrelevant. It is not restricted by problem dimensionality, and does not require knowledge of exact solutions.

For FEM discretizations, the first procedure to achieve nodal exactness was Tong's adjoint technique [37]; see also [45, App. 7]. This requires finding exact homogeneous solutions of $\mathbf{r}(\mathbf{u}) = \mathbf{0}$, to be inserted as weight functions in a Petrov-Galerkin discretization. Related schemes are based on localized enrichment by homogeneous and/or particular solutions, for example [5,6]. All of these methods rely on the Galerkin approach rather than the Ritz method used here.

§3.5. Example 1: From IOMoDE to FOMoDE All the Way

As previously noted, Step 3 of the modified equation method is technically challenging. The process will be illustrated here through a specific example. The result will be used later for treating the constant coefficient case of the diffusion-absorption and Helmholtz equations. The given IOMoDE is

the homogeneous, even-derivative, infinite-order ODE:

$$-\frac{\mu}{2a^2}u(x)+\frac{1}{2!}u''(x)+\frac{a^2\chi^2}{4!}u''''(x)+\frac{a^4\chi^4}{6!}u''''''(x)+\ldots=0,\quad a>0,\ \ \mu\neq 0,\ \ 0<\chi\leq 1. \tag{9}$$

Here μ and χ are dimensionless real parameters whereas a, which is a characteristic problem dimension, has dimension of length. Primes denote differentiation with respect to the independent space variable x. Parameter χ goes to zero as the mesh of the source problem is refined. Following a variant of Warming and Hyett's derivative elimination procedure, (9) is differentiated $2(n-1)$ times ($n=1,2,\ldots$) with respect to x while discarding all odd derivatives. The equations are truncated to a maximum derivative order $2n$, and a linear system in the even derivatives u'', u'''', ... is set up. The configuration of the elimination system is illustrated for $n=4$:

$$\begin{bmatrix} 1/2! & a^2\chi^2/4! & a^4\chi^4/6! & a^6\chi^6/8! \\ -\frac{1}{2}\mu a^{-2} & 1/2! & a^2\chi^2/4! & a^4\chi^4/6! \\ 0 & -\frac{1}{2}\mu a^{-2} & 1/2! & a^2\chi^2/4! \\ 0 & 0 & -\frac{1}{2}\mu a^{-2} & 1/2! \end{bmatrix}\begin{bmatrix} u'' \\ u'''' \\ u'''''' \\ u'''''''' \end{bmatrix}=\begin{bmatrix} \frac{1}{2}\mu a^{-2}\,u \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{10}$$

The coefficient matrix of system (10) is Toeplitz and Hessenberg. Solving for u'' yields a truncated FOMoDE, which is then expanded in ascending Taylor series in $\lambda=\mu\chi^2$:

$$u''=\frac{56\mu a^{-2}(360+60\lambda+\lambda^2)\,u}{20160+5040\lambda+252\lambda^2+\lambda^3}=\mu a^{-2}\left(1-\frac{1}{12}\lambda+\frac{1}{90}\lambda^2-\frac{1}{560}\lambda^3+\ldots\right)u. \tag{11}$$

Increasing the order n, the coefficients of the power series in λ are found to be generated by the recursion $c_1=1$, $c_{n+1}=-\frac{1}{2}n^2c_n/[(n+1)(2n+1)]$, $n\geq 1$, which produces the sequence $\{1,-1/12,1/90,-1/560,1/3150,-1/16632,\ldots\}$. The generating function [41] can be found by *Mathematica's* package `RSolve` by entering `<<DiscreteMath‘RSolve‘; g=GeneratingFunction[ a[n+1]==-n*n/(2*(n+1)*(2*n+1))*a[n],a[1]==1,a[n],n,λ]; Print[g]`. To verify the answer do `Print[Series[g,{λ,0,8}]`. The result is

$$\frac{4}{\lambda}\left(\operatorname{arcsinh}\frac{\sqrt{\lambda}}{2}\right)^2=1-\frac{\lambda}{12}+\frac{\lambda^2}{90}-\frac{\lambda^3}{560}+\frac{\lambda^4}{3150}-\frac{\lambda^5}{16632}+\frac{\lambda^6}{84084}-\frac{\lambda^7}{411840}+\ldots \tag{12}$$

This yields the second-order FOMoDE

$$u''=\frac{4}{a^2\chi^2}\left(\operatorname{arcsinh}\frac{\sqrt{\lambda}}{2}\right)^2u=\frac{4}{a^2\chi^2}\left(\operatorname{arcsinh}\frac{\chi\sqrt{\mu}}{2}\right)^2u. \tag{13}$$

To give an example of matching suppose that the original ODE from which (9) comes is $u''=(w/a^2)u$, where w is constant. For nodal exactness, $w=(4/\chi^2)\big(\operatorname{arcsinh}(\frac{1}{2}\chi\sqrt{\mu}\big)^2$. If μ is the free parameter, solving for it gives

$$\mu=\frac{4}{\chi^2}\left(\sinh\frac{\chi\sqrt{w}}{2}\right)^2=\frac{2\big(\cosh(\chi\sqrt{w})-1\big)}{\chi^2}. \tag{14}$$

§3.6. Example 2: From IOMoDE to FOMoDE Part Way

The second example illustrates a case in which the reduction to FOMoDE is incomplete because only part of the series is easily identified as expressible in closed form. The answer may be still useful, however, if the discarded (unprocessed) part becomes negligible for the envisioned applications. The result will be used later for treating the variable coefficient case of the diffusion-absorption and Helmholtz equations. The given IOMoDE is:

$$\begin{aligned} &-\frac{\mu}{2a^2}u(x)+\frac{1}{2!}u''(x)+\frac{a^2\chi^2}{4!}u''''(x)+\frac{a^4\chi^4}{6!}u''''''(x)+\ \ldots \\ &+\frac{\nu}{a}\left(u'(x)+\frac{a^2\chi^2}{3!}u'''(x)+\frac{a^4\chi^4}{5!}u''''(x)+\ldots\right)=0,\quad a>0,\ \ \mu\neq 0,\ \ 0<\chi\leq 1. \end{aligned} \tag{15}$$

Here the even-derivative series is the same as in Example 1, with μ, χ and a retaining the same meaning. The new ingredient is the appearance of an odd-derivative series, which is multiplied by the dimensionless parameter ν. Furthermore μ and ν have the expressions

$$\mu=\frac{\mu_0+\mu_2\,\phi}{\mu_1+\mu_3\,\phi},\qquad \nu=\frac{\mu_4\,\phi+\mu_5\phi^2}{\mu_1+\mu_3\,\phi},\quad \mu_0\neq 0,\ \ \mu_1\neq 0, \tag{16}$$

in which μ_0 through μ_4 are dimensionless functions of the original problem. Parameter ϕ, which has dimensions of length inverse, measures the deviation from a previously solved problem and therefore it can be regarded as a *perturbation variable*. Indeed if $\phi=0$, ν vanishes and we are back in Example 1. Proceeding as before, (15) is repeatedly differentiated with respect to x, retaining derivatives of order up to n. The configuration of the elimination system is illustrated for $n=4$:

$$\begin{bmatrix} \nu\,a^{-1} & 1/2! & \nu a\chi^2/3! & a^2\chi^2/4! \\ -\frac{1}{2}\mu a^{-2} & \nu\,a^{-1} & 1/2! & \nu a\chi^2/3! \\ 0 & -\frac{1}{2}\mu a^{-2} & \nu\,a^{-1} & 1/2! \\ 0 & 0 & -\frac{1}{2}\mu a^{-2} & \nu\,a^{-1} \end{bmatrix} \begin{bmatrix} u' \\ u'' \\ u''' \\ u'''' \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\mu a^{-2}\,u \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{17}$$

Comparing to (10), the main change is that now all derivative orders appear in the left-hand side vector. Solving for u'' yields a truncated FOMoDE, which is then expanded in ascending Taylor series, first in ϕ and then in $\lambda=(\mu_0/\mu_1)\chi^2$:

$$u''=\left[\frac{\mu_0}{\mu_1 a^2}\left(1-\frac{1}{12}\lambda+\frac{1}{90}\lambda^2-\ldots\right)+\frac{\mu_d}{\mu_1 a}\left(1-\frac{1}{6}\lambda+\frac{1}{30}\lambda^2-\ldots\right)\phi+O(\phi^2)\right]u. \tag{18}$$

in which $\mu_d=\mu_1\mu_2-\mu_0\mu_3$. (Actually to get the terms shown one needs to go $n\geq 6$.) The first series has been identified in Example 1 as given by (12), except that now λ is $(\mu_0/\mu_1)\chi^2$. The coefficients of the second one are found to be generated by the recursion $c_1=1$, $c_{n+1}=-n^2c_n/[2n(2n+1)]$, $n\geq 1$. Again using `RSolve` we obtain

$$\frac{2\,\mathrm{arcsinh}\frac{\sqrt{\lambda}}{2}}{\sqrt{\lambda(1+\frac{1}{4}\lambda)}}=1-\frac{\lambda}{6}+\frac{\lambda^2}{30}-\frac{\lambda^3}{140}+\frac{\lambda^4}{630}-\frac{\lambda^6}{2772}+\ldots \tag{19}$$

This yields the second-order FOMoDE

$$u''=\left[\frac{4}{a^2\chi^2}\psi^2+\frac{2(\mu_1\mu_2-\mu_0\mu_3)\psi}{\mu_1 a\sqrt{\lambda(1+\frac{1}{4}\lambda)}}\,\phi+O(\phi^2)\right]u,\quad \text{with}\ \ \psi=\mathrm{arcsinh}\frac{\sqrt{\lambda}}{2}. \tag{20}$$

Note that μ_4 and μ_5 are missing from the terms shown above. In fact μ_4 makes its first appearance in the $O(\phi^3)$ series.

The series in ϕ^2, ϕ^3, etc., are significantly more complicated and no identification was attempted. Completing this FOMoDE to higher powers in ϕ thus remains an open problem.

§4. The Diffusion-Absorption Problem

The remaining sections illustrate the variational FIC discretization and the construction of modified equations for the steady-state, one-dimensional diffusion-absorption equation. For the constant coefficient case of this particular application, the loop in Figure 3 can be successfully closed.

This problem has been recently examined by Oñate, Miquel and Hauke [30] from a FIC-Galerkin standpoint. That study includes advection terms which are not considered here. The governing differential equation that models a one-dimensional, steady state, diffusion-absorption process is

$$\frac{d}{dx}\left(k\frac{du}{dx}\right) - su + Q = 0, \quad \text{in} \quad x \in [x_m, x_p] \tag{21}$$

In this equation u is the state variable, $x \in [x_m, x_p]$ is the problem domain, $k \geq 0$ is the diffusion, $s \geq 0$ is the absorption (also called dissipation or destruction parameter) and Q the source term. Using primes to denote differentiation with respect to x, the foregoing ODE can be abbreviated to

$$\left(k\,u'\right)' - s\,u + Q = 0. \tag{22}$$

With the flux defined as $q = k(du/dx) = k\,u'$, the boundary conditions can be stated as

$$u = \hat{u} \quad \text{on} \quad \Gamma^u, \qquad q = \hat{q}, \quad \text{on} \quad \Gamma^q. \tag{23}$$

where Γ^u and Γ^q are the Dirichlet and Neumann boundaries, respectively. For the one-dimensional problem these consist of four combinations taken at the ends of the problem domain. This problem admits a classical variational formulation. Introduce the functional

$$J[u] = \int_{x_m}^{x_p} \left(\tfrac{1}{2}k(u')^2 + \tfrac{1}{2}su^2 - Qu\right)\,dx. \tag{24}$$

Taking the first variation $\delta J = 0$ over admissible functions $u(x)$ that satisfy the essential BCs yields the differential equation (21) as Euler-Lagrange equation, and the flux constraints in (23) as natural boundary conditions.

§4.1. The Model Problem

Following [30] and assuming $k \neq 0$, a model form of (21) is obtained by introducing the dimensionless coefficient

$$w = \frac{sa^2}{k}, \tag{25}$$

where $a = x_p - x_m$ is the length of the problem domain. This coefficient characterizes the relative importance of absorption over diffusion. The model problem domain is adjusted to extend from $x_m = -\frac{1}{2}a$ to $x_p = \frac{1}{2}a$ for convenience. We will assume zero source: $Q = 0$, and Dirichlet boundary conditions at both ends: $u(-\frac{1}{2}a) = u_m$ and $u(\frac{1}{2}a) = u_p$. We can therefore state the model problem as

$$u'' - \frac{w}{a^2}u = 0 \;\; \text{for } x \in [-\tfrac{1}{2}a, \tfrac{1}{2}a], \qquad u(-\tfrac{1}{2}a) = u_m, \quad u(\tfrac{1}{2}a) = u_p. \tag{26}$$

The associated functional is

$$J[u] = \int_{-a}^{a} \left((u')^2 + \frac{w}{2a^2}u^2\right)\,dx. \tag{27}$$

where variation is taken over continuous $u(x)$ that satisfy the Dirichlet BCs.

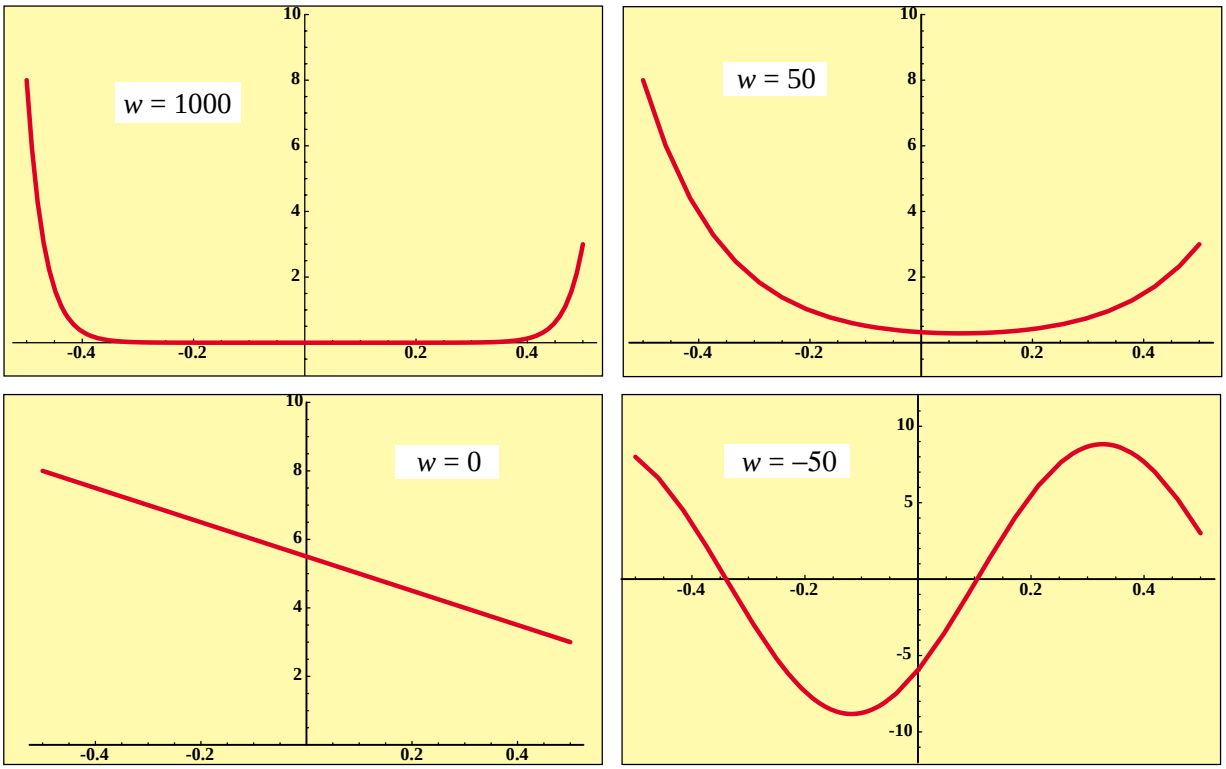

Figure 4. Behavior of the exact solution of the diffusion-absorption model equation for four values of w, with Dirichlet BCs $u(-\frac{1}{2}) = 8$ and $u(\frac{1}{2}) = 3$.

§4.2. Exact Solutions

If $w \neq 0$ the exact solution of the model problem (26) is

$$u(x) = \frac{\sinh\left(\sqrt{w}(\frac{1}{2} - x/a)\right) u_m + \sinh\left(\sqrt{w}(\frac{1}{2} + x/a)\right) u_p}{\sinh(\sqrt{w})}. \tag{28}$$

This form becomes $0/0$ if $w = 0$ and suffers from cancellation errors if $|w|$ is very small, say $|w| < 10^{-6}$. For that case a Taylor series about $w = 0$ gives, to first order in w:

$$\begin{aligned} u(x) \approx & u_m((12a^3 - 24a^2x) + w(-3a^3 + 2a^2x + 12ax^2 - 8x^3))/(24a^3) + \\ & u_p((12a^3 + 24a^2x) + w(-3a^3 - 2a^2x + 12ax^2 + 8x^3))/(24a^3). \end{aligned} \tag{29}$$

The exact solution is displayed in Figure 4 for $u_m = u(-\frac{1}{2}) = 8$, $u_p = u(\frac{1}{2}) = 3$, $w = 1000$, 50, 0 and -50. If $w = 0$ the solution is a straight line. As w grows, exponential-growth boundary layers appear at Dirichlet boundaries. This is illustrated by the upper plots in Figure 4. If $w = 1000$ the solution is very small over most of the problem domain except for two sharp boundary layers near $x = \pm\frac{1}{2}a$.

If $w < 0$ the solution (28) involves complex exponentials and the response is oscillatory, as illustrated in the bottom right of Figure 4. This case is not relevant to diffusion-absorption. On taking $\kappa^2 = -w/a^2$, however, (26) becomes the one-dimensional Helmholtz equation of linear acoustics: $u'' + \kappa^2 u = 0$, used to model, for instance, wave propagation along an elastic bar. This ODE is also applicable to some problems in solid mechanics involving elastic foundations.

§4.3. Conventional Ritz

A standard FEM solution is easily constructed by the Ritz variational formulation. Divide the domain into N^e two-node elements of length $L^e = a/N^e = \chi a$. The end nodes are i and j, with coordinates x_i and x_j, and node values u_i and u_j, respectively. Assume the piecewise linear interpolation

$$u(x) = u_i N_i(x) + u_j N_j(x), \tag{30}$$

where $N_i(x) = (x - x_i)/L^e$, $N_j = (x_j - x)/L^e$ and $L^e = x_j - x_i$ are the well known linear shape functions. Substitution into (27) gives the element stiffness equations

$$\mathbf{S}^e\mathbf{u}^e = \frac{1}{L^e}\begin{bmatrix} 1+\frac{1}{3}\zeta & -1+\frac{1}{6}\zeta \\ -1+\frac{1}{6}\zeta & 1+\frac{1}{3}\zeta \end{bmatrix}\begin{bmatrix} u_i \\ u_j \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \chi = L^e/a, \quad \zeta = w\chi^2 = s(L^e)^2/k. \tag{31}$$

If $w = 0$ this element relation gives, upon assembly, the linear response correctly. However if $w \neq 0$, the use of (31), a scheme that may be labeled "unstabilized Ritz," displays a known defect: if w is large the solution oscillates over coarse meshes. This is illustrated in Figure 8(a) for $N_e = 8$ elements and $w = 1000$. Negative u values are physically incorrect if $w > 0$, which renders the solution useless.

This shortcoming is usually treated by Galerkin stabilization schemes, for example with suitably adjusted weight functions. But this amounts to using a adjoint-equation approach for what is essentially a self-adjoint problem.

§4.4. The FIC Functional

We try to stay within the Ritz framework and piecewise-linear shape functions, but change the functional by the method outlined in Section 2. For this problem, the FIC function modification technique consists of formally replacing

$$\tilde{u}(x) = u(x) - \tfrac{1}{2}hu'(x), \qquad \tilde{u}'(x) = u'(x) - \tfrac{1}{2}hu''(x). \tag{32}$$

These $\tilde{u}(x)$ and $\tilde{u}'(x)$ are inserted into (26). The tildes are then suppressed for brevity. This scheme yields a modified functional $J_h[u]$, where h is the FIC steplength. That h was derived in the original FIC by flux balancing arguments [21]. In the present case h may be simply viewed as a free parameter with dimension of length.

For piecewise linear shape functions $u''(x)$ vanishes over each element, and the second replacement in (32) may be skipped. With this simplification the modified functional is

$$J_h[u] = \int_{-a}^{a} \left((u')^2 + \tfrac{1}{2}\frac{w}{a^2}\left(u - \tfrac{1}{2}hu'(x)\right)^2 \right) dx. \tag{33}$$

The Euler-Lagrange equation given by $\delta J_h[u] = 0$ is

$$(1 + \frac{wh^2}{4a^2})u'' - \frac{w}{a^2}u = 0. \tag{34}$$

From this the FIC variational residual follows as $\delta J_h = \delta u\left[(1 + \frac{1}{4}wh^2/a^2)u'' - (w/a^2)u\right]$.

The expression (34) shows that a nonzero h injects artificial diffusion if $w > 0$. Furthermore, the sign of h makes no difference in the interior of the problem domain. As $h \to 0$ the original ODE (26) is recovered. But the key idea behind FIC is to keep h finite and directly related to mesh size.

§5. The Ritz FIC Equations

The FIC functional (33) is used in conjunction with the piecewise-linear interpolation (30) to construct stabilized Ritz equations for the model diffusion-absorption problem. The steplength $h = h^e$ may in fact change from element to element. For convenience define $h^e = \alpha^e L^e$ where α^e is a dimensionless parameter to be determined. The analysis of the present Section is restricted to equal size elements and the same α for all elements. The latter restriction is removed in Section 7, which studies a variable coefficient variant of the original problem.

With exact element integration (equivalently, a two-point Gauss integration rule), the following Ritz FIC element equations are obtained:

$$\frac{1}{L^e}\begin{bmatrix} 1+(\frac{1}{3}+\frac{1}{2}\alpha+\frac{1}{4}\alpha^2)\zeta & -1+(\frac{1}{6}-\frac{1}{4}\alpha^2)\zeta \\ -1+(\frac{1}{6}-\frac{1}{4}\alpha^2)\zeta & 1+(\frac{1}{3}-\frac{1}{2}\alpha+\frac{1}{4}\alpha^2)\zeta \end{bmatrix}\begin{bmatrix} u_i \\ u_j \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{35}$$

in which $\chi = L^e/a$ and $\zeta = w\chi^2 = s(L^e)^2/k$.

§5.1. Patch Equations

The stiffness equations for a patch of two equal-size elements comprising nodes i, j, k, as pictured in Figure 5, are

$$\frac{1}{L^e}\begin{bmatrix} 1+(\frac{1}{3}+\frac{1}{2}\alpha+\frac{1}{4}\alpha^2)\zeta & -1+(\frac{1}{6}-\frac{1}{4}\alpha^2)\zeta & 0 \\ -1+(\frac{1}{6}-\frac{1}{4}\alpha^2)\zeta & 2+(\frac{2}{3}+\frac{1}{2}\alpha^2)\zeta & -1+(\frac{1}{6}-\frac{1}{4}\alpha^2)\zeta \\ 0 & -1+(\frac{1}{6}-\frac{1}{4}\alpha^2)\zeta & 1+(\frac{1}{3}-\frac{1}{2}\alpha+\frac{1}{4}\alpha^2)\zeta \end{bmatrix}\begin{bmatrix} u_i \\ u_j \\ u_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{36}$$

with $\zeta = w\chi^2$. The remaining issue is to find the value of α to be inserted into the discrete system. Four methods are studied below. The first one is loosely based on the discussion in [30]. The second and third methods rely on the modified differential equation (MoDE). The fourth method is based on knowledge of the exact solution, and is used to verify the third one.

The second method uses a truncated IOMoDE whereas the third one uses FOMoDE to "close the loop" in the flowchart of in Figure 3, and thus achieve nodal exactness. The first three methods are extendible to two and three dimensions. The fourth one relies on the availability of the exact solution, which is not usually available beyond one dimension.

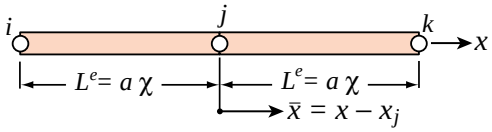

Figure 5. A patch of two equal length Ritz-FIC elements.

Note that all methods actually find α^2 and not α. Which sign of the square root is taken makes no difference for the model problem. This sign does becomes relevant when the discrete equations are obtained by reduced integration, as discussed in Section 5.6.

§5.2. Finding α by Positivity

Consider the patch equations (36). Suppose $u_i > 0$ and $u_k > 0$ are prescribed. Solving for u_j from the second equation gives

$$u_j = \frac{12+(3\alpha^2-2)\zeta}{24+(6\alpha^2+8)\zeta}(u_i+u_k), \quad \zeta = w\chi^2. \tag{37}$$

The denominator is positive for any $\{\alpha, \chi\}$ if $w > 0$. It follows that the condition for $u_j \geq 0$ is $12+(3\alpha^2-2)\zeta \geq 0$, whence $\alpha^2 \leq (2/3) - 4/\zeta = (2/3) - 4/(w\chi^2) = \alpha_P^2$. (Here the P subscript stands for "positivity.") This yields

$$\alpha_P^2 = \frac{2}{3} - \frac{4}{w\chi^2}. \tag{38}$$

This approach gives a useful bound: if $\zeta = w\chi^2 \le 6$, α may be set to zero without impairing positivity. For example, if $w = 600$, a mesh of 10 or more elements can have $\alpha = 0$, since $\chi = 1/10$ and $\zeta = w\chi^2 = 6$. As discussed later, this substitution does not imply an accurate solution.

Inserting this α_P^2 into the element matrix $\mathbf{S}^e$ of (35) cancels out the off-diagonal terms. The assembled $\mathbf{S}$ is therefore *diagonal*. The solution for zero source and Dirichlet conditions at both ends is therefore *zero at all interior nodes*. This mimics well the physical behavior for large and positive w; say $w > 10000$. For positive but smaller w this solution can be way off, but it shows that (38) may be viewed as a *lower bound* on acceptable values of α^2, whereas the choice $\alpha_C^2 = 2/3$ found below is an *upper bound*. The numerical results discussed in Section 6, however, show that these bounds are of little practical value for moderate values of w, and totally useless if $w < 0$.

§5.3. Finding α by Truncated IOMoDE

The next two schemes rely on the modified equation method. To obtain the infinite-order MoDE (IOMoDE) following the techniques of [32,33] proceed as follows. The patch equation for the center node j is $S_{ij}u_i + S_{jj}u_j + S_{jk}u_k = 0$, in which $S_{ij} = S_{jk} = [-1+(\frac{1}{6}-\frac{1}{4}\alpha^2)w\chi^2]/L^e$ and $S_{jj} = [2+(\frac{2}{3}+\frac{1}{2}\alpha^2)w\chi^2]/L^e$. Convert to difference-differential modified equation (DDMoDE) by formally replacing $u_j \to u(x)$, $u_i \to u(x-L^e)$ and $u_k \to u(x+L^e)$ to get $S_{ij}u(x-L^e) + S_{jj}u(x) + S_{jk}u(x+L^e) = 0$. Express $u(x \pm L^e)$ in terms of $u(x)$ by Taylor series

$$\begin{aligned} u(x-L^e) &= u - L^e u' + \tfrac{1}{2}(L^e)^2 u'' - \tfrac{1}{6}(L^e)^3 u''' + \ldots, \\ u(x+L^e) &= u + L^e u' + \tfrac{1}{2}(L^e)^2 u'' + \tfrac{1}{6}(L^e)^3 u''' + \ldots, \end{aligned} \tag{39}$$

Equations (39) are Laplace transformed (replacing each derivative operator $()' \equiv (d/dx)$ by the Laplace transform variable p) and inserted into the DDMoDE. Solve for the Laplace transformed $u(x)$ and backtransform to get the IOMoDE. This can be worked over to the compact form

$$-\frac{6w}{\gamma a^2}u + \frac{1}{2!}u'' + \frac{1}{4!}a^2\chi^2 u'''' + \frac{1}{6!}a^4\chi^4 u'''''' + \ldots = 0, \tag{40}$$

in which $\chi = L^e/a$ is the inverse of the number of elements and $\gamma = 12 + (3\alpha^2 - 2)w\chi^2$. Making $\chi \to 0$ truncates (40) to the second derivative: $u'' = (12w/\gamma a^2)u$. Requiring consistency with the original ODE (26), which is $u'' = (w/a^2)u$, gives $\gamma = 12 + (3\alpha^2 - 2)w\chi^2 = 12$, whence $3\alpha^2 - 2 = 0$ and

$$\alpha_C^2 = \frac{2}{3}. \tag{41}$$

The value (41) will be called the "consistent α" because it is obtained by a ODE consistency argument as $\chi \to 0$. Numerical computations show that use of α_C overestimates the diffusion for all positive w. Consequently it is "safe" in the sense of providing physically correct solutions. But these can be highly inaccurate for small or moderate w. The examples of Section 6 illustrate this point. However this method gives a useful limit: if $w \to +\infty$ on a fixed mesh, $\alpha^2 \to \alpha_C^2 = 2/3$ for consistency.

§5.4. Finding Nodally Exact α via FOMoDE

To get a nodally exact solution using the method outlined in Section 3.4, it is necessary to convert the IOMoDE (40) to finite order (FOMoDE) through elimination of higher order derivatives and series identification. This operation is carried out in Example 1, worked out in Section 3.5. Setting $\mu = 12w/\gamma$, with $\gamma = 12 + (3\alpha^2 - 2)w\chi^2$, into (40) produces (9). The FOMoDE (13), reproduced for

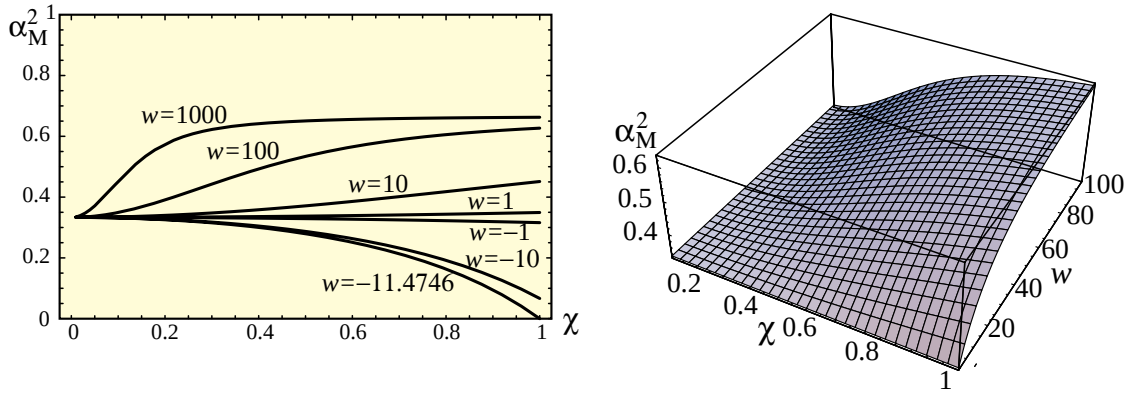

Figure 6. Variation of α_M^2 as a function of $\chi = L^e/a$ and w. Left figure: 2D plots for $\chi \in [0, 1]$ and fixed sample $w \in [1000, -11.4746]$. At the latter value, $\alpha_M^2 = 0$ at $\chi = 1$. Right figure: 3D plot of α_M^2 versus $w \in [0, 100]$ and $\chi = [0, 1]$.

convenience, is $u'' = \left[\left(\text{arcsinh}\frac{1}{2}a\chi\sqrt{\mu}\right)^2/(a^2\chi^2)\right]u$. Matching this to $u'' = (w/a^2)u$ gives the value (14) for μ. Converting to α yields

$$\alpha_M^2 = \frac{2}{3} - \frac{4}{w\chi^2} + \frac{1}{\sinh^2\left(\frac{1}{2}\chi\sqrt{w}\right)}, \tag{42}$$

where subscript M stands for "matching the original ODE." Using $\alpha = \alpha_M$ furnishes a *nodally exact solution* under the following conditions: elements of equal length, constant w, zero source and Dirichlet BCs. Expression (42) is valid for nonzero w, including negative values. Thus it applies to the Helmholtz equation. These conclusions are verified by the numerical experiments reported in Section 6.

The Taylor series of α_M^2 as $\chi = L^e/a \to 0$ (i.e., as the mesh is refined with more elements) is

$$\alpha_M^2 = \frac{1}{3} + \frac{w\chi^2}{60} - \frac{w^2\chi^4}{1512} + \frac{w^3\chi^6}{43200} - \frac{w^4\chi^8}{1330560} + \ldots \tag{43}$$

which shows that $\alpha_M^2 \to 1/3$ if $\chi \to 0$ or $w \to 0$. Figure 6 illustrates the range of variation of α_M^2 as function of w and χ. The latter varies between 0 and 1, attaining 1 only for one element over the domain length a. For positive w, α_M^2 is always positive but less than $\alpha_C^2 = 2/3$. For $w = 0$, $\alpha_M^2 = 1/3$. For negative w, α_M^2 is positive as long as $w\chi^2 > -11.4746$. If $w < -11.4746$, α_M^2 can be negative for coarse meshes. Since $\chi = L^e/a = 1/N^e$, the minimum number of elements N^e to get a positive α_M^2 is $N^e > 1/\sqrt{-11.4746/w}$. The condition is met if there are 2 or more elements per wavelength. If not verified, α_M is imaginary, and complex numbers appear in $\mathbf{S}^e$. If the mesh consists of equal elements, and w is constant, complex numbers occur only in the first and last rows of $\mathbf{S}$, and disappear altogether on applying Dirichlet BCs. Aside from this case the use of complex arithmetic is necessary, although the Ritz equations remain symmetric.

For computer implementation, exponential functions in (42), which may cause numerical accuracy problems for $w > 10^5$, can be avoided by using Padé approximants to α_M^2. The (2,2), (4,4) and (6,6) diagonal approximants computed by *Mathematica* are

$$\begin{aligned}
\alpha_{M22}^2 &= \frac{1260 + 113w\chi^2}{30(126 + 5w\chi^2)}, \qquad \alpha_{M44}^2 = \frac{3270960 + 339948w\chi^2 + 4787w^2\chi^4}{189(51920 + 2800w\chi^2 + 39w^2\chi^4)}, \\
\alpha_{M66}^2 &= \frac{1966225060800 + 218635457040w\chi^2 + 4430449320w^2\chi^4 + 27010573w^3\chi^6}{60(98311253040 + 6016210200w\chi^2 + 115773966w^2\chi^4 + 671585w^3\chi^6)}.
\end{aligned} \tag{44}$$

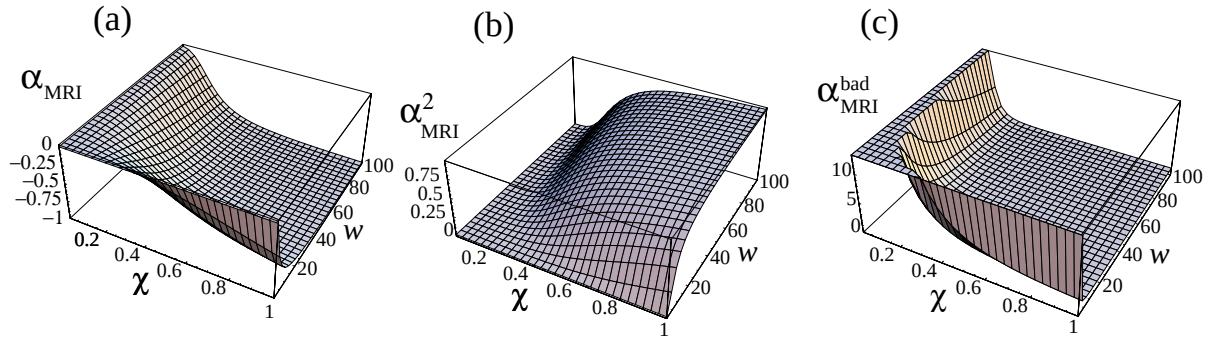

Figure 7. Nodally exact αs for reduced integration as functions of $\chi = L^e/a$ and w: (a) α_{MRI} of (48), (b) α^2_{MRI}, (c) root given by exact solution matching that blows up as $\chi \to 0$.

For $\chi < \frac{1}{4}$ and $w < 10000$ these provide at least 1, 2 and 3 digits of accuracy, respectively. For moderate $w\chi^2$ the higher approximants give 10–12 digits of accuracy. As $w\chi^2 \to \infty$, α^2_{M22}, α^2_{M44}, α^2_{M66} and α^2_{M88} approach the limits 0.75333333333333, 0.64943698277032, 0.6703190462364 and 0.66590125338669, respectively. Since α^2 should not exceed $2/3$ on account of the consistency condition discussed in Section 5.3, a cutoff may have to be implemented for some approximants.

§5.5. Verifying the Nodally Exact α by Exact Solution

A valuable verification of (42) can be obtained directly by equating u_j in (37) to the exact node value for a BVP posed over the 2-element patch with prescribed node values u_i and u_k:

$$u_j^{exact} = \tfrac{1}{2}(u_j + u_k)\,\mathrm{sech}\big(\chi\sqrt{w}\big) = \frac{12 + (3\alpha^2 - 2)w\chi^2}{24 + (6\alpha^2 + 8)w\chi^2}(u_i + u_k). \tag{45}$$

Solving for α^2 and simplifying gives back (42). This method, however, cannot be used if the exact solution is not available, as it happens in two and three dimensional patches, whereas the modified equation method does not rely on such knowledge.

§5.6. Effect of Reduced Integration

The foregoing Ritz-FIC equations have been constructed with exact element integration, which is equivalent to using a two-point Gauss rule. If a one-point Gauss reduced integration rule is used, the element equations become

$$\frac{1}{L^e}\begin{bmatrix} 1 + \frac{1}{4}(1+\alpha)^2\zeta & -1 + \frac{1}{4}(1-\alpha^2)\zeta \\ -1 + \frac{1}{4}(1-\alpha^2)\zeta & 1 + \frac{1}{4}(1+\alpha)^2\zeta \end{bmatrix}\begin{bmatrix} u_i \\ u_j \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \zeta = w\chi^2. \tag{46}$$

The two-element patch equations are

$$\frac{1}{L^e}\begin{bmatrix} 1 + \frac{1}{4}(1+\alpha)^2\zeta & -1 + \frac{1}{4}(1-\alpha^2)\zeta & 0 \\ -1 + \frac{1}{4}(1-\alpha^2)\zeta & 2 + \frac{1}{2}(1+\alpha)^2\zeta & -1 + \frac{1}{4}(1-\alpha^2)\zeta \\ 0 & -1 + \frac{1}{4}(1-\alpha^2)\zeta & 1 + \frac{1}{4}(1+\alpha)^2\zeta \end{bmatrix}\begin{bmatrix} u_i \\ u_j \\ u_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{47}$$

Both α and α^2 now appear in the difference equation, and survive in the three forms of the MoDE. Proceeding as before one obtains the nodally exact α as

$$\alpha_{MRI} = \frac{\sigma - \sqrt{1 + 4(2\sigma - \sigma^2 - 1)/\zeta}}{1 - \sigma}, \quad \zeta = w\chi^2, \quad \sigma = \mathrm{sech}(\chi\sqrt{w}) = \mathrm{sech}(\sqrt{\zeta}). \tag{48}$$

Functions $\alpha_{MRI}(w,\chi)$ and $\alpha^2_{MRI}(w,\chi)$ are displayed in Figure 7(a) and (b), respectively. Note that as $\zeta = w\chi^2 \to 0$, both α_{MRI} and α^2_{MRI} approach zero. In fact, $\alpha_{MRI} = -\frac{1}{6}w\chi^2 - \frac{13}{720}w^2\chi^4 + \ldots$. This is unlike α^2_M in (42), which approaches $\frac{1}{3}$ as per the series (43).

The expression (48) is more complicated than that for α^2_M given in (42). Using the exact solution to find α_{MRI} gives two solutions for α_{MRI}: $\left[\sigma \pm \sqrt{1+4(2\sigma-\sigma^2-1)/\zeta}\right]/(1-\sigma)$, which appear as roots of a quadratic. Taking the minus sign reproduces (48), whereas taking the plus sign gives an α that "blows up" if $w\chi^2 \to 0$, as shown in Figure 7(c). Either value gives nodally exact solutions but the associated coefficient matrices are totally different. In summary, the reduced integration approach for this problem does not offer obvious advantages over exact integration.

§5.7. Source Terms

The MoDE treatment of a smooth source term $q(x)$ in $u'' - (w/a^2)u = q(x)$ can be done by entirely analogous techniques, but there are representation choices. One is based on expanding $q(x)$ in Taylor series: $q(x) = q(x_j) + q'(x_j)(x-x_j)/a + \ldots$ at node j, and inserting into the FIC functional to derive a consistent node force term q_j by the usual methods. Then $q(x_j)$, $q'(x_j)\ldots$ appear in the RHS of the elimination system (10), and the solution series is identified into the FOMoDE. Alternatively $q(x)$ can be expanded in Fourier series [33]. Delta function source terms can be processed directly.

§6. Numerical Results for Constant Coefficients

This section present numerical results obtained with the Ritz-FIC method for the model problem (26). The problem domain is taken to have unit length ($a = 1$) extending from $x_m = -\frac{1}{2}a = -\frac{1}{2}$ through $x_p = \frac{1}{2}a = \frac{1}{2}$. The boundary conditions are of Dirichlet type: $u(-\frac{1}{2}) = 8$ and $u(\frac{1}{2}) = 3$. The domain is divided into 8 elements of equal size; thus $\chi = L^e/a = \frac{1}{8}$. Four values of w: 1000, 50, -50 and -1000, are tested. Results are graphically collected in Figure 8, and discussed below.

§6.1. Results for $w = 1000$

The solution to $w = 1000$ exhibits two sharp boundary layers. Over the propagation region, which extends roughly over the middle six elements of this discretization, $u(x)$ takes small positive values, of order 10^{-3} or less. The problem is discretized using four choices of α: $\alpha = 0$ (conventional Ritz), $\alpha^2_C = 2/3$, $\alpha^2_P = 101/375 = 0.410667$ and $\alpha^2_M = 0.490503$. Numerical results are shown in Figure 8(a), and listed in Table 1 along with the exact solution.

As expected the solution for α^2_M is nodally exact. The results for $\alpha = 0$ oscillate giving unacceptable negative values. The results for α_C and α_P give the correct physical behavior, and bound the boundary layer behavior on both sides. Although the difference of results computed for α_C and α_P with the exact solution are masked in the scale of the plot, the discrepancies at interior points are clear from Table 1.

§6.2. Results for $w = 50$

This case $w = 50$ pertains to moderate absorption. The boundary layers are diffuse and the exact solution resembles a second degree parabola. The problem is again discretized using four α choices: $\alpha = 0$, $\alpha^2_C = 2/3$, $\alpha^2_P = -334/75 = -4.45333$ and $\alpha^2_M = 0.345961$. The numerical results are plotted in Figure 8(b), and listed in Table 2 along with the exact solution.

Again the solution for α^2_M is nodally exact. The solutions for $\alpha = 0$ and $\alpha^2_C = 2/3$ bound the exact solution, maintain positivity and display reasonable accuracy. The results for α_P are way off as can be expected from the rationale for its construction.

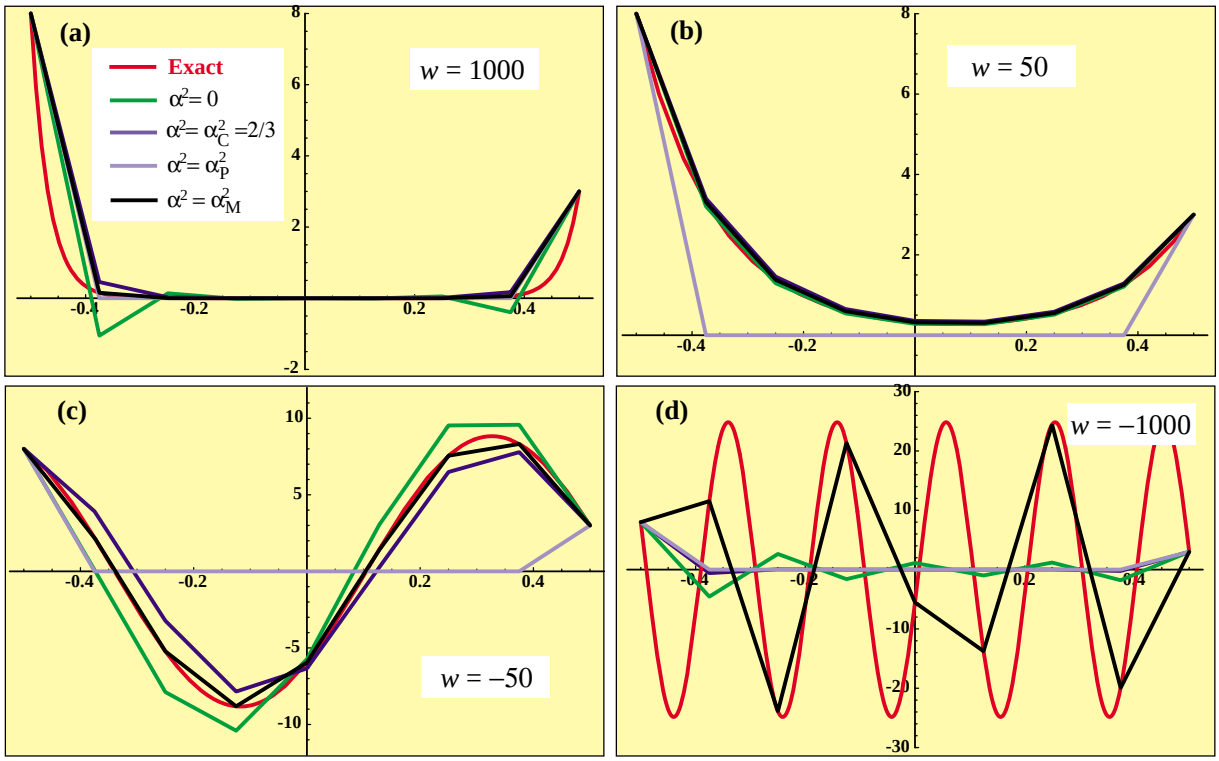

Figure 8. Ritz-FIC results for 8-element discretization of the diffusion-absorption model problem with: (a) $w = 1000$, (b) $w = 50$, (c) $w = -50$, (d) $w = -1000$; Dirichlet BCs $u(-\frac{1}{2}) = 8$ and $u(\frac{1}{2}) = 3$, for several choices of α, compared to the exact solution.

§6.3. Results for $w = -50$

This case $w = -50$ pertains to moderate negative absorption. (As previously noted, this is not physically relevant for diffusion-absorption, but it holds for the Helmholtz equation in acoustic and some elastic foundation problems in solid mechanics.) Negatives values of $u(x)$ are now physically admissible.

The exact solution lacks boundary layers and is oscillatory, going roughly through one wavelength over the problem domain; cf. Figure 8(c). The problem is discretized with four α choices: $\alpha = 0$, $\alpha_C^2 = 2/3$, $\alpha_P^2 = 434/75 = 5.78667$ and $\alpha_M^2 = 0.319898$. Note that α_M^2 is still positive because $w\chi^2 = -50/64 > -11.4746$, and no imaginary numbers appear. The numerical results are plotted in Figure 8(c), and listed in Table 3 along with the exact solution. Again the solution for α_M^2 is nodally exact. The results for $\alpha = 0$ and $\alpha_C^2 = 2/3$ bound the exact solution and follow its shape reasonably well. The solution for α_P is worthless.

§6.4. Results for $w = -1000$

Setting $w = -1000$ produces a rapidly oscillatory exact solution that goes roughly through five wavelengths over the problem domain, as depicted in Figure 8(d). The problem is discretized with $\alpha = 0$, $\alpha_C^2 = 2/3$, $\alpha_P^2 = 0.92667$ and $\alpha_M^2 = -0.261754$. Here α_M^2 is negative because $w\chi^2 = -1000/64 = -11.875 < -11.4746$. Although α_M^2 produces a nodally exact solution, an 8-element discretization over five wavelengths is obviously inadequate to fit the oscillation frequency. To correctly capture the physical behavior more elements should be used.

The effect of injecting more elements is illustrated in Figure 9, which shows results for 8, 16, 32 and

Table 1. Ritz-FIC 8-element solutions, $w = 1000$.

Node	Exact	$\alpha = 0$	$\alpha_C^2 = 0.666667$	$\alpha_P^2 = 0.410667$	$\alpha_M^2 = 0.490503$
1	8	8	8	8	8
2	0.1536	−1.05141	0.455371	0	0.1536
3	0.00294911	0.138198	0.0259205	0	0.00294911
4	0.0000566306	−0.0182785	0.00147722	0	0.0000566306
5	$1.49484\,10^{-6}$	0.00328186	0.000115477	0	$1.49484\,10^{-6}$
6	0.0000212544	−0.007124	0.000558065	0	0.0000212544
7	0.00110592	0.0518597	0.00972042	0	0.00110592
8	0.0575999	−0.394284	0.170764	0	0.0575999
9	3	3	3	3	3

Table 2. Ritz-FIC 8-element solutions, $w = 50$.

Node	Exact	$\alpha = 0$	$\alpha_C^2 = 0.666667$	$\alpha_P^2 = -4.45333$	$\alpha_M^2 = 0.345961$
1	8	8	8	8	8
2	3.3105	3.20689	3.40025	0	3.3105
3	1.38017	1.29421	1.45694	0	1.38017
4	0.60014	0.543987	0.651862	0	0.60014
5	0.320303	0.282379	0.356053	0	0.320303
6	0.307424	0.274406	0.338411	0	0.307424
7	0.55077	0.512905	0.585152	0	0.55077
8	1.25316	1.2121	1.28904	0	1.25316
9	3	3	3	3	3

Table 3. Ritz-FIC 8-element solutions, $w = -50$.

Node	Exact	$\alpha = 0$	$\alpha_C^2 = 0.666667$	$\alpha_P^2 = 5.78667$	$\alpha_M^2 = 0.319898$
1	8	8	8	8	8
2	2.19055	0.089892	3.91683	0	2.19055
3	−5.22172	−7.88235	−3.22636	0	−5.22172
4	−8.81328	−10.406	−7.84896	0	−8.81328
5	−5.95623	−5.73652	−6.33956	0	−5.95623
6	1.25896	2.89827	0.122625	0	1.25896
7	7.55298	9.52964	6.48901	0	7.55298
8	8.32053	9.57371	7.78585	0	8.32053
9	3	3	3	3	3

Table 4. Ritz-FIC 8-element solutions, $w = -1000$.

Node	Exact	$\alpha = 0$	$\alpha_C^2 = 0.666667$	$\alpha_P^2 = 0.922667$	$\alpha_M^2 = -0.261754$
1	8	8	8	8	8
2	11.5432	−4.55513	−0.590353	0	11.5432
3	−23.8971	2.63742	0.0435651	0	−23.8971
4	21.3673	−1.60393	−0.00322138	0	21.3673
5	−5.52955	1.10817	0.000326198	0	−5.52955
6	−13.7521	−0.983943	−0.00122306	0	−13.7521
7	24.4687	1.18959	0.016338	0	24.4687
8	−19.9456	−1.79406	−0.221383	0	−19.9456
9	3	3	3	3	3

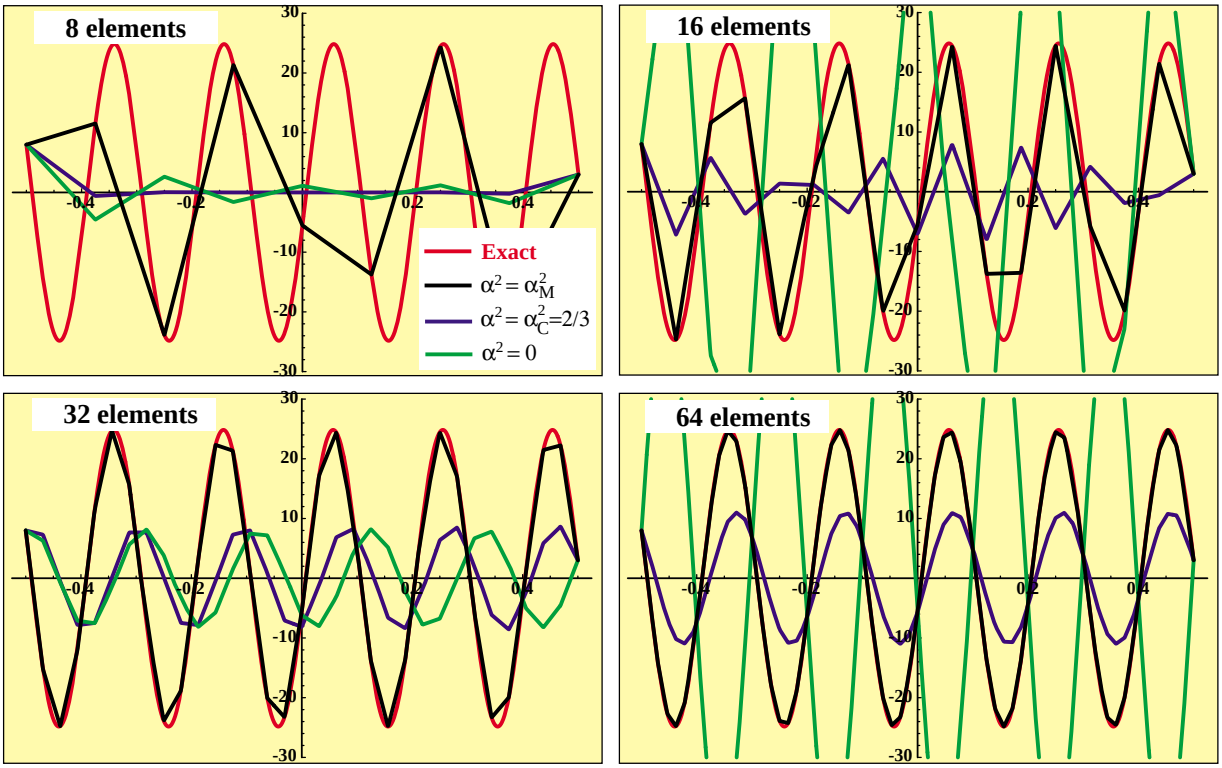

Figure 9. Ritz-FIC convergence study for highly oscillatory case $w = -1000$. Boundary conditions: $u(-\frac{1}{2}) = 8$ and $u(\frac{1}{2}) = 3$. Shown are results for 8, 16, 32 and 64 elements, and three choices of α: α_M, α_C and conventional Ritz $\alpha = 0$. Results for α_P omitted as they are worthless.

64 elements. The latter places about 10 elements per wavelength, which should be adequate as per well known empirical rules for approximating sinusoidal waveforms. The beneficial effect of nodal exactness is evident. The results for other choices of α display erratic behavior, even for fine discretizations.

§7. A Variable Coefficient ODE

This Section studies the application of Ritz FIC to the variable coefficient (VC) equation

$$ku'' - s(x)u + Q = 0 \qquad (49)$$

where $k \geq 0$ is constant but s is a linear function of x. As previously, the computational domain is $x \in [x_m, x_p]$ with Dirichlet boundary conditions $u(x_m) = u_m$ and $u(x_p = u_p$. Of particular interest is the case where s changes sign over the computational domain. In that case $u(x)$ will have boundary-layer exponential decay behavior over the portion where $s > 0$, transitioning to oscillations wherever $s < 0$. See Figure 10.

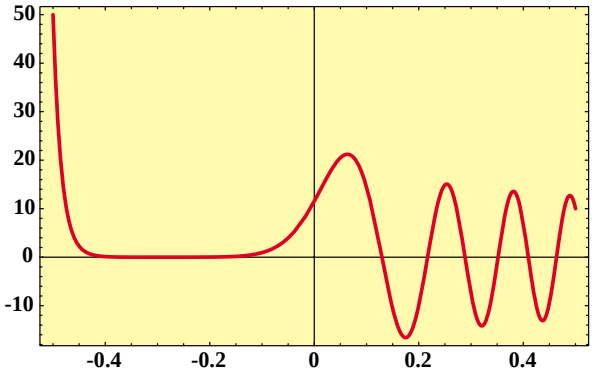

Figure 10. Exact solution of model problem (50) for $a = 1$, $w_m = 50$, $w_p = 10$ and $w = 100 - 8000(x/a)$.

This goal is to provide a benchmark of whether the present methods can handle simultaneouly the diffusion-absorption and Helmholtz type of equations. The restriction to linear variation in x is imposed

to have a closed form solution, in terms of Airy functions, available for comparison. Although these functions are rarely useful in classical mechanics they find applications in optics, quantum mechanics, electromagnetics, and radiative transfer.

§7.1. The VC Model Problem

As before we restrict the computational domain to $\pm\frac{1}{2}a$, define $w(x) = s(x)\,a^2/k$, which is now a linear function in x explicitly given as $w = w_0 + w_1(x/a)$. The model problem is defined as

$$u'' - \frac{w}{a^2}u = 0 \quad \text{with } w = w_0 + w_1\frac{x}{a}, \quad \text{for } x \in [-\tfrac{1}{2}a,\, \tfrac{1}{2}a], \quad u(-\tfrac{1}{2}a) = u_m,\ u(\tfrac{1}{2}a) = u_p. \tag{50}$$

The associated functional is obtained by changing w in (27) to

$$J[u] = \int_{-a}^{a} \left((u')^2 + \frac{w_0 + w_1\frac{x}{a}}{2a^2}u^2 \right) dx. \tag{51}$$

where variation is taken over continuous $u(x)$ that satisfy the Dirichlet BCs.

The solution of (50) can be expressed in closed form in terms of the Airy functions $Ai(x)$ and $Bi(x)$ [1, Sec. 10.4] as follows. Assuming that $w_1 \neq 0$, compute:

$$\begin{aligned} c &= (w_1^2)^{1/3} \quad b_1 = \frac{2w_0 + w_1}{2c}, \quad b_2 = \frac{2w_0 - w_1}{2c}, \quad x_w = \frac{w_0 + w_1(x/a)}{c}, \\ d &= Ai(b_2)Bi(b_1) - Ai(b_1)Bi(b_2), \quad U_p(x) = Ai(b_2)Bi(x_w) - Ai(x_w)Bi(b_2), \\ U_m(x) &= Ai(x_w)Bi(b_1) - Ai(b_1)Bi(x_w), \quad u(x) = \frac{u_m U_m(x) + u_p U_p(x)}{d}. \end{aligned} \tag{52}$$

If $w_1 = 0$ the Airy solution (52) fails. In that case $w = w_0$ is constant and the exact solution in terms of exponentials, given in Section 4.2, should be used. As an example Figure 10 shows the exact solution for a case where $w = 100 - 8000(x/a)$ changes from a sharp boundary layer to wavy behavior.

§7.2. Discretization

To construct a Ritz-FIC finite element discretization, the modified piecewise linear shape functions (32) can be reused with an important modification: α is allowed to vary with x (not only element by element, but also within an element). Inserting the modified shape function into (51) and performing the variation with respect to node displacements gives the Ritz FIC equations for an element of nodes i-j:

$$\begin{bmatrix} S^e_{ii} & S^e_{ij} \\ S^e_{ij} & S^e_{jj} \end{bmatrix} \begin{bmatrix} u_i \\ u_j \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{53}$$

The entries of $\mathbf{S}^e$ are computed by two-point Gauss integration rule, which evaluates the functional (51) exactly for linearly varying $w(x)$ if the value of α at the Gauss nodes is separately chosen. Denoting by w_j, α_j $(j = 1, 2)$ the values of $w(x)$ and $\alpha(x)$, respectively, at Gauss point j,

$$\begin{aligned} S^e_{ii} &= \frac{1}{L^e}\left[1 + \frac{w_1(3 + \sqrt{3} + 3\alpha_1)^2 + w_2(3 - \sqrt{3} + 3\alpha_2)^2}{72}\chi^2\right], \\ S^e_{ij} &= \frac{1}{L^e}\left[-1 + \frac{w_1(2 - 2\sqrt{3}\alpha_1 - 3\alpha_1^2) + w_2(2 + 2\sqrt{3}\alpha_2 - 3\alpha_2^2)}{24}\chi^2\right], \\ S^e_{jj} &= \frac{1}{L^e}\left[1 + \frac{w_1(3 - \sqrt{3} - 3\alpha_1)^2 + w_2(3 + \sqrt{3} - 3\alpha_2)^2}{72}\chi^2\right]. \end{aligned} \tag{54}$$

If $\alpha_1 = \alpha_2 = \alpha$ and $w_1 = w_2 = w$ the entries reduce to the constant coefficient element equations (35).

Over each element w is assumed to vary linearly from w_i at node i to w_j at node j. Gauss point values w_1 and w_2 are computed by interpolation. The selection of α_1 and α_2 is studied next.

§7.3. The Modified Equation

To obtain a modified equation we consider again a patch of two elements as shown. Coefficient w is taken to vary linearly over the patch and is expressed as

$$w(\bar{x}) = w_0 + \phi\bar{x}. \tag{55}$$

See Figure 11. Here $\bar{x} = x - x_j$ is the distance from the center patch node, and $\phi = dw/dx$.

To make further progress it is necessary to make an assumption on the variation of α over the patch although of course only the Gauss point values will appear in the element computations. The assumption is that

$$\alpha(\bar{x}) = \alpha_0 + \beta\phi\bar{x}. \tag{56}$$

where β is a free parameter. From this the Gauss point values are obtained by interpolation from node values.

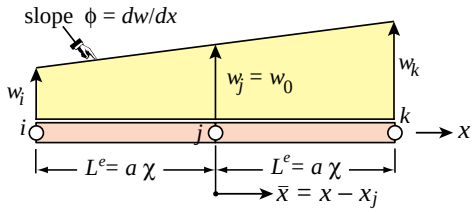

Figure 11. Variation of w over patch.

The computation of the DDMoDE and IOMoDE forms follows the same techniques used to derive (39), but now starting from (53) and (54), and need not be repeated here. The results is the IOMoDE studied in Example 2, which is defined by (15) and (16), and in which

$$\begin{aligned}
&\mu_0 = 12w_0, \quad \mu_1 = 12 - w_0(3\alpha_0^2 - 2)\chi^2, \quad \mu_2 = 6(\alpha_0 + w_0\beta)a\chi, \\
&\mu_3 = -(\alpha_0 + w_0\beta)a\chi^3, \quad \mu_4 = (6\alpha_0^2 - 4 + 12w_0\,\alpha_0\beta)a\chi^2, \quad \mu_5 = -4a^2\beta\chi^3.
\end{aligned} \tag{57}$$

The FOMoDE obtained in Example 2 is (20), which includes terms up to $O(\phi^2)$. In that equation, $\mu_d = \mu_1\mu_2 - \mu_0\mu_3 = 18a\chi(\alpha_0 + w_0\beta)(4 + w_0\alpha_0^2\chi^2)$.

§7.4. Finding α

Matching to the source equation $u'' = (w_0 + \phi\bar{x})u$ requires $4\psi^2 = w_0\chi^2$ and $2\mu_d\psi = \mu_1\sqrt{\lambda(1 + \frac{1}{4}\lambda\chi}$. The first condition says that α_0 must be determined by (42) with w replaced by w_0; that is $\alpha_0^2 = \frac{2}{3} - 4/w_0\chi^2 + 1/\sinh^2\left(\frac{1}{2}\chi\sqrt{w_0}\right)$. The second condition determines the value of β. The latter, however, is not necessary in the actual implementation because giving α_0 at each node determines fully the variation of α over each element.

In the numerical tests reported below a slightly different procedure is followed: α_1 and α_2 were computed directly from interpolated Gauss-point w values:

$$\alpha_1^2 = \frac{2}{3} - \frac{4}{w_1\chi^2} + \frac{1}{\sinh^2\left(\frac{1}{2}\chi\sqrt{w_1}\right)}, \quad \alpha_2^2 = \frac{2}{3} - \frac{4}{w_2\chi^2} + \frac{1}{\sinh^2\left(\frac{1}{2}\chi\sqrt{w_2}\right)} \tag{58}$$

This procedure obviously can handle any variation of $w(x)$ over the computational domain. It covers stepped coefficients in layered problems if nodes are placed at discontinuity points.

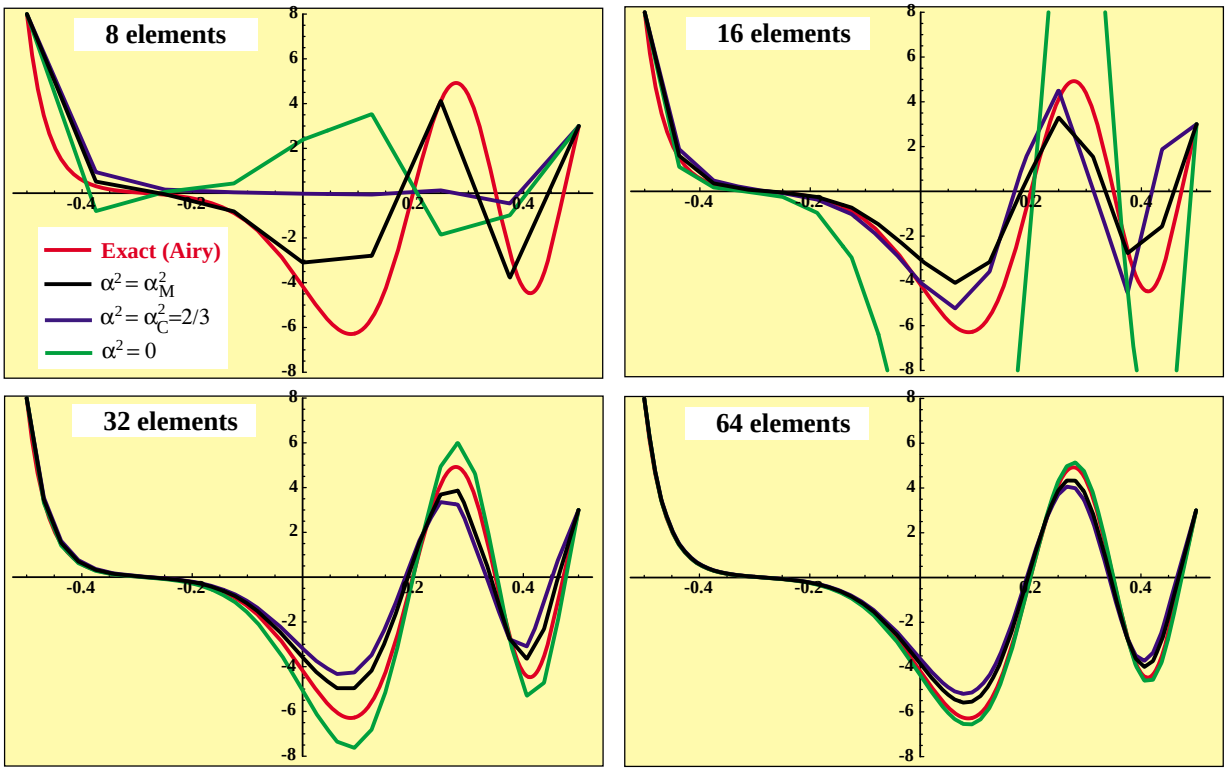

Figure 12. Ritz-FIC convergence study of exponential-to-oscillatory variable-coefficient model equation (50) with $w = -1600x/a$ (w varies from $+800$ to -800). Boundary conditions: $u(-\frac{1}{2}) = 8$ and $u(\frac{1}{2}) = 3$. Results shown for 8, 16, 32 and 64 elements, and three choices of α: α_M, α_C and 0. For the α_M choice, α is evaluated at the 2 element Gauss points via (58) and the positive square roots taken.

§7.5. Numerical Results for Variable Coefficients

A wide range of w variations and number of elements was tested using three choices of α: (i) conventional Ritz with $\alpha = 0$ at all points, (ii) "consistent" $\alpha_C = \sqrt{2/3}$ at all points, and (iii) the matching α_M given by (58) at the element Gauss points. The "positivity" choice $\alpha = \alpha_P$ is not reported as it gave consistently inferior results.

The results shown in the convergence study of Figure 12 are representative. In this case $w = -800x/a$ varies from 800 on the left to -800 at the right. The exponential-like solution over $x < 0$ and the boundary layer near $x = -\frac{1}{2}$ are accurately captured by the three α choices. This behavior was consistently observed in all tests whenever w was positive. On the other hand, convergence difficulties are evident over the transition and the oscillatory regions of the response. The performance of the conventional Ritz $\alpha = 0$ was noticeably erratic within oscillatory regions for coarse meshes.

Unlike the constant-coefficient case, the results obtained with α_M are not nodally exact because the FOMoDE (20) is incomplete. The determination of a complete FOMoDE remains an open problem.

§8. Conclusions

This article has presented a synthesis of three techniques: FIC, variational Ritz and modified differential equations. The major new contributions are:

1. The FIC approach to functional modification. This permits effective stabilization of the diffusion-absorption problem while staying within the ordinary Ritz framework of finite elements. No separate choice of trial and weight functions is necessary.
2. The use of the modified equation (MoDE) approach to find a value of the stabilization parameter that is nodally exact for all values of the absorption-to-diffusion ratio, including negative values that morphs the original ODE into the Helmholtz equation.

An important advantage of a Ritz discretization over its Galerkin counterparts is that symmetric matrices are obtained. This permits reuse of symmetric equation solvers often available in finite element software. Symmetry also simplifies eigenvalue calculations for problems such as linear acoustics in the frequency domain. This advantage, however, holds only as long as FIC parameters remain real, as otherwise complex numbers appear in the Ritz equations. In the application considered here, avoiding complex numbers requires the use of at least two elements per wavelength in the case of the Helmholtz equation.

A potential disadvantage of the Ritz approach, again when compared to Galerkin, is that the FIC steplength h — or equivalently its dimensionless counterpart α — appears quadratically in the element equations, because the shape functions that carry h go into a quadratic functional. (The method of least squares would produce a similar result.) This brings up the problem of choosing signs when solving quadratic equations. In the present application that difficulty is inconsequential.

The main attraction of the modified equation approach is that availability of exact solutions of the source ODE is not required to construct a nodally exact discretization. This feature is important for application of the method in two and three dimensions. For the one-dimensional constant-coefficient problem discussed here, the same nodally exact discretization can be also obtained by patch matching as shown in Section 5.5.

The logical extension of the present combination of methods is the study of two and three dimensional space discretizations by considering regular finite element patches. Since exact solutions for such problems are rarely available, the modified equation method appears to be a promising choice for improving nodal solutions over fixed meshes. The Ritz ingredient, however, may have to be dropped in problems, such as advection, which are not easily formulated in a variational framework.

Acknowledgements

The work of the first author has been partly supported through a fellowship awarded by the Spanish Ministerio of Educación y Cultura to visit CIMNE during May-June 2002, and partly by the National Science Foundation under grant High-Fidelity Simulations for Heterogeneous Civil and Mechanical Systems, CMS-0219422.

References

[1] M. Abramowitz and I. A. Stegun, (Eds.), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th printing, Dover, 1972.

[2] M. O. Ahmed and R. M. Corless, The method of modified equations in Maple, *Electronic Proc. 3rd Int. IMACS Conf. on Applications of Computer Algebra*, Maui, July 1997. PDF accessible at `htpp://math.umn.edu/ACA/1997.html`

[3] F. Brezzi, T. J. R. Hughes, L. D. Marini, A. Russo and E. Süli, A priori error analysis of residual-free bubbles for advection-diffusion problems, *SIAM J. Numer. Anal.*, **36**, 1933–1948, 1999.

[4] F. Brezzi and A. Russo: Stabilization techniques for the finite element method, in *Applied and Industrial Mathematics, Venice-2*, 1998, R. Spigler Ed., Kluwer, 47-58, 2000.

[5] C. Farhat, I. Harari and L. Franca, The discontinuous enrichment method, *Comp. Meths. Appl. Mech. Engrg.*, **190**, 6455-6479, 2001.

[6] C. Farhat, I. Harari, and U. Hetmaniuk, The discontinuous enrichment method for multiscale analysis, *Comp. Meths. Appl. Mech. Engrg.*, **192**, 3195–3209, 2003.

[7] B. M. Finlayson, *The Methods of Weighted Residuals and Variational Principles*, Academic Press, 1972.

[8] L. Franca, C. Farhat, M. Lesoinne and A. Russo, Unusual stabilized finite element methods and residual-free-bubbles, *Int. J. Num. Meth. Fluids*, **27**, 159–168 1998.

[9] L. Franca, C. Farhat, A. P. Macedo and M. Lesoinne. Residual-free bubbles for the Helmholtz equation, IJNME, **40**, 4003–4009, 1997.

[10] D. Griffiths and J. Sanz-Serna. On the scope of the method of modified equations. *SIAM J. Sci. Statist. Comput.*, **7**, 994–1008, 1986.

[11] A. Kolesnikov and A. J. Baker, Efficient implementation of high order methods for the advection-diffusion equation, *Proc. 3^{rd} ASME/JSME Joint Fluids Engrg Conf.*, San Francisco, CA, 1999.

[12] E. Hairer, Backward analysis of numerical integrators and symplectic methods, *Annals Numer. Math.*, **1**, 107–132, 1994.

[13] E. Hairer, C. Lubich and G. Wanner, *Geometric Numerical Integration: Structure Preserving Algorithms for Ordinary Differential Equations*, Springer, Berlin, 2002.

[14] C. W. Hirt, Heuristic stability theory for finite difference equations, *J. Comp. Physics*, **2**, 339–342, 1968.

[15] T. J. R. Hughes, Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods, *Comp. Meths. Appl. Mech. Engrg.*, **127**, 387–401, 1995.

[16] P. E. Kloeden and K. J. Palmer (eds), *Chaotic Numerics*, Amer. Math. Soc., Providence, RI, 1994.

[17] H. Lomax, P. Kutler and F. B. Fuller, The numerical solution of partial differential equations governing convection, AGARDograph 146, 1970.

[18] E. Oñate, J. García and S. R. Idelsohn, Computation of the stabilization parameter for the finite element solution of advective-diffusive problems, *Int. J. Numer. Meth. Fluids*, **25:1**, 1385–1407, 1997.

[19] E. Oñate, J. García and S. R. Idelsohn, Computation of the stabilization parameter for the finite element solution of advective-diffusive problems, *New Advances in Adaptive Computer Methods in Mechanics*, P. Ladeveze, and J. T. Oden, eds., Elsevier, Amsterdam, 1998.

[20] E. Oñate, Derivation of the stabilization equations for advective-diffusive fluid transport and fluid flow problems, *Comp. Meths. Appl. Mech. Engrg.*, **151:1–2**, 233–267, 1998.

[21] E. Oñate and M. Manzan, A general procedure for deriving stabilized space-time finite element methods for advective-diffusive problems, *Int. J. Numer. Meth. Engrg.*, **31**, 203–207, 1999.

[22] E. Oñate and M. Manzan, Stabilization techniques for finite element analysis of convection-diffusion fluid problems, in *Comp. Anal. of Heat Transfer*, G. Comini and B. Sunden (Eds.), WIT Press, 2000.

[23] E. Oñate, A stabilized finite element method for incompressible viscous flows using a finite increment calculus formulation, *Comp. Meths. Appl. Mech. Engrg.*, **182:1–2**, 355–370, 2000.

[24] E. Oñate, J. García, A finite element method for fluid-structure interaction with surface waves using a finite calculus formulation, *Comp. Meths. Appl. Mech. Engrg.*, **191:6–7**, 635–660, 2001.

[25] E. Oñate, Multiscale computational analysis in mechanics using finite calculus: an introduction, *Comp. Meths. Appl. Mech. Engrg.*, **192**, 3043–3059, 2003.

[26] E. Oñate, R. L. Taylor, O. C. Zienkiewicz and J. Rojek, A residual correction method based on finite calculus, *Engrg. Comput.*, **20**, 629–658, 2003.

[27] E. Oñate, R. L. Taylor, O. C. Zienkiewicz and J. Rojek, Finite calculus formulation for analysis of incompressible solids using linear triangles and tetrahedra, *Int. J. Numer. Meth. Engrg.*, **59**, 1473–1500, 2004.

[28] E. Oñate, Possibilities of finite calculus in computational mechanics, *Int. J. Numer. Meth. Engrg.*, **60**, 255-281, 2004.

[29] E. Oñate, J. García and S. R. Idelsohn, Ship Hydrodynamics, in *Encyclopedia of Computational Mechanics*, T. J. R. Hughes, R. de Borst, E. Stein (eds.), to appear 2004.

[30] E. Oñate, J. Miquel and G. Hauke, A stabilized finite element method for the one-dimensional advection diffusion-absorption equation using finite calculus, *Int. J. Numer. Meth. Engrg.*, submitted.

[31] E. Oñate and C. A. Felippa, Variational formulation of the finite calculus equations in solid mechanics, CIMNE Report in preparation.

[32] K. C. Park and D. L. Flaggs, An operational procedure for the symbolic analysis of the finite element method, *Comp. Meths. Appl. Mech. Engrg.*, **46**, 65–81, 1984.

[33] K. C. Park and D. L. Flaggs, A Fourier analysis of spurious modes and element locking in the finite element method, *Comp. Meths. Appl. Mech. Engrg.*, **42**, 37–46, 1984.

[34] R. L. Richtmyer and K. W. Morton, *Difference Methods for Initial Value Problems*, Interscience Pubs (Wiley), New York, 2nd ed., 1967.

[35] P. J. Roache, *Computational Fluid Mechanics*, Hermosa Publishers, Albuquerque, 1970.

[36] A. M. Stuart and A. R. Humphries, *Dynamic Systems and Numerical Analysis*, Cambridge Univ. Press, Cambridge, 1996.

[37] P. Tong, Exact solution of certain problems by the finite element method, *AIAA*, **7**, 179–180, 1969.

[38] B. D. Vujanovic and S. E. Jones, *Variational Methods in Nonconservative Phenomena*, Academic Press, 1989.

[39] J. E. Waltz, R. E. Fulton and N. J. Cyrus, Accuracy and convergence of finite element approximations, *Proc. Second Conf. on Matrix Methods in Structural Mechanics*, WPAFB, Ohio, Sep. 1968, in *AFFDL TR 68-150*, 995–1028, 1968.

[40] R. F. Warming and B. J. Hyett, The modified equation approach to the stability and accuracy analysis of finite difference methods, *J. Comp. Physics*, **14**, 159–179, 1974.

[41] H. S. Wilf, *Generatingfunctionology*, Academic Press, New York, 1991.

[42] J. H. Wilkinson, Error analysis of direct methods of matrix inversion, *J. ACM* **8**, 281–330, 1961.

[43] J. H. Wilkinson, *Rounding Errors in Algebraic Processes*, Prentice-Hall, Englewood Cliffs, N.J., 1963.

[44] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford Univ. Press, Oxford, 1965.

[45] O. C. Zienkiewicz and R. E. Taylor, *The Finite Element Method*, 4th ed., McGraw-Hill, London, Vol. I: 1988.