Nonlinear Dynamics (2006) 45: 183–206
DOI: 10.1007/s11071-005-9002-z

Galerkin Projections and Finite Elements for Fractional Order Derivatives

SATWINDER JIT SINGH and ANINDYA CHATTERJEE*
Mechanical Engineering of Department, Indian Institute of Science, Bangalore 560012, India;
**Author for correspondence (e-mail: anindya100@gmail.com)*

(Received: 15 December 2004; accepted: 22 September 2005)

Abstract. Ordinary differential equations (ODEs) with fractional order derivatives are infinite dimensional systems and nonlocal in time: the history of the state variable is needed to calculate the instantaneous rate of change. This nonlocal nature leads to expensive long-time computations ($\mathcal{O}(t^2)$ computations for solution up to time t). A finite dimensional approximation of the fractional order derivative can alleviate this problem. We present one such approximation using a Galerkin projection. The original infinite dimensional system is replaced with an equivalent infinite dimensional system involving a partial differential equation (PDE). The Galerkin projection reduces the PDE to a finite system of ODEs. These ODEs can be solved cheaply ($\mathcal{O}(t)$ computations). The shape functions used for the Galerkin projection are important, and given attention. The approximation obtained is specific to the fractional order of the derivative; but can be used in any system with a derivative of that order. Calculations with both global shape functions as well as finite elements are presented. The discretization strategy is improved in a few steps until, finally, very good performance is obtained over a user-specifiable frequency range (not including zero). In particular, numerical examples are presented showing good performance for frequencies varying over more than 7 orders of magnitude. For any discretization held fixed, however, errors will be significant at sufficiently low or high frequencies. We discuss why such asymptotics may not significantly impact the engineering utility of the method.

Key words: fractional derivative, Galerkin projection, finite dimensional approximation.

1. Introduction

Ordinary differential equations (ODEs) involving fractional order derivatives are used to model a variety of systems, of which an important engineering application lies in viscoelastic damping [1–4]. Another important application of fractional derivatives lies in control theory (see, e.g., [5–8]). Linear ODEs with fractional order derivatives or differintegrals are studied using Laplace [1, 2] and Fourier transforms [3]. Linear systems with half order derivatives have also been solved using an eigenvector expansion [9].

Nonlinear ODEs involving fractional order derivatives can be solved numerically. But the infinite dimensional nature of these systems leads to high computation cost for simulation of long-time behaviors, as we will discuss later.

In this paper, we present a numerical technique to solve ODEs with fractional order derivatives using a Galerkin projection for reducing the infinite dimensional system to a finite dimensional one. The approximation obtained is specific to the fractional order of the derivative of interest; but it can be used without further change in any system where a derivative of that order appears. This novel method is easy to apply and is computationally more efficient than direct integration-based numerical methods for long-time simulations. Both global shape functions as well as finite elements are used for the Galerkin projections. The discretization strategy is refined in a few steps to provide motivation for the final strategy adopted. For that final strategy, numerical accuracy obtained is excellent. The method is expected to

be most useful for nonlinear ODEs with fractional order derivatives, but can, of course, be used for linear systems as well. The principal advantage offered by our method, while providing demonstrated accuracy, lies in time savings relative to straightforward numerical integration. Other finite dimensional approximations, as in [10, 11], will give similar time savings. However, the conceptual basis of our approximation is different and, in our opinion, simpler.

2. Background

A fractional derivative of order $n + q$ is given, using the Riemann–Louville definition [12], as

$$D_b^{n+q}[x(t)] \equiv \frac{d^{n+q}x(t)}{[d(t-b)]^{n+q}} = \frac{1}{\Gamma(-q)}\frac{d^n}{dt^n}\left[\int_b^t \frac{x(\tau)}{(t-\tau)^{1+q}}d\tau\right], \tag{1}$$

where $-1 < q < 0$ and n is a positive integer. For many practical problems it is assumed that the system starts from rest, so that $x(t) \equiv 0$ for $t \leq 0$. In such cases D_b^{n+q} becomes D_0^{n+q}, and we will henceforth drop the b subscript assuming that the system starts at $b = 0$. Hence the fractional derivative of Equation (1) becomes

$$D^{n+q}[x(t)] = \frac{1}{\Gamma(-q)}\frac{d^n}{dt^n}\left[\int_0^t \frac{x(\tau)}{(t-\tau)^{1+q}}d\tau\right].$$

For many practical problems in dynamics, the order of the fractional derivative lies between 0 and 1, and we will consider fractional values restricted between these limits as well. Accordingly, substituting $n = 1$ in the above, we get

$$D^{1+q}[x(t)] = \frac{1}{\Gamma(-q)}\frac{d}{dt}\left[\int_0^t \frac{x(\tau)}{(t-\tau)^{1+q}}d\tau\right],$$

which in turn can be rewritten as

$$D^{\alpha}[x(t)] = \frac{1}{\Gamma(1-\alpha)}\frac{d}{dt}\left[\int_0^t \frac{x(\tau)}{(t-\tau)^{\alpha}}d\tau\right],$$

where $0 < \alpha < 1$. Two equivalent forms of the above are given as

$$D^{\alpha}[x(t)] = \frac{1}{\Gamma(1-\alpha)}\left[\frac{x(0)}{t^{\alpha}} + \int_0^t \frac{\dot{x}(\tau)}{(t-\tau)^{\alpha}}d\tau\right] = \frac{1}{\Gamma(1-\alpha)}\left[\frac{x(0)}{t^{\alpha}} + \int_0^t \frac{\dot{x}(t-\tau)}{\tau^{\alpha}}d\tau\right]. \tag{2}$$

In each equivalent expression of Equation (2), the first term has a singularity at $t = 0$ but disappears if $x(0) = 0$ (as in, e.g., [13]; we assume the same here), giving

$$D^{\alpha}[x(t)] = \frac{1}{\Gamma(1-\alpha)}\int_0^t \frac{\dot{x}(t-\tau)}{\tau^{\alpha}}d\tau. \tag{3}$$

The above integral involves a singular kernel. We mention in passing that substituting $\bar{\tau} = \tau^{(1-\alpha)}$ gives

$$D^{\alpha}[x(t)] = \frac{1}{\Gamma(2-\alpha)} \int_0^{t^{1-\alpha}} \dot{x}(t - \bar{\tau}^{1/(1-\alpha)})\, d\bar{\tau}. \tag{4}$$

In this way, the singularities can be removed for easier computation; the price paid is that the history of $\dot{x}$, which appears inside the integral, may need to be resampled (using, e.g., splines) before numerical evaluation [14]. However, direct integration of the solution history at each time step is not the subject of this paper; our aim is to avoid such integration.

Equation (4) shows that the fractional order derivative is nonlocal. It requires the history of $\dot{x}(\tau)$ from $\tau = 0$ to $\tau = t$. Any ODE involving such a fractional derivative is therefore infinite dimensional.

While numerically solving an ODE with such a fractional derivative through direct evaluation of the integral of Equation (4), we face the following problem. Assuming time steps of length Δt, the integral evaluated at instant $t = k\,\Delta t$ requires $\mathcal{O}(k)$ arithmetic operations. To reach the instant $t = k\,\Delta t$, therefore, we need

$$\mathcal{O}\left(\sum_{i=1}^{k} i\right) = \mathcal{O}(k^2)$$

operations. In other words, for a simulation time duration t, we need $\mathcal{O}(t^2)$ calculations. For long times, this is prohibitively high. Yet long times may be unavoidable for, e.g., studying the steady state dynamics of lightly damped or chaotic systems, or nonlinear systems under random forcing using Monte Carlo methods (see, e.g., [15]).

An approximation scheme that does not suffer from this $\mathcal{O}(t^2)$ requirement, but can instead compute solutions in $\mathcal{O}(t)$ time, would be useful. Such schemes may involve finite dimensional and local approximations for the fractional order derivative (the scheme we present here does).

Our scheme is comparable with that of [16] which, though not identical, is quite similar in spirit to ours. In this paper, however, we go beyond the treatment of [16] and develop a Galerkin projection as well as finite elements for the dimensional reduction; in this way, we are able to provide a clearer picture of the performance and accuracy of the method. We also mention the work of [17], where there is no finite dimensional approximation *per se*, but the integral required for evaluating the fractional derivative is numerically approximated after subdividing the interval $(0, t)$ into a number of exponentially decreasing contiguous subintervals: the net result is something like $\mathcal{O}(t \ln t)$ time, which is almost as good as $\mathcal{O}(t)$. Unlike [17], our approach *a priori* approximates the fractional derivative operator itself, rather than making approximations while computing the value of the fractional order derivative. Finally, we mention a paper by Schmidt and Gaul [18], which critiques [16] and points out asymptotic aspects of the approximation: in particular, short time and high frequency asymptotic behavior are *not* captured by Yuan and Agrawal's approximation. Our scheme, too, has this apparent flaw. This flaw may be important from a restricted mathematical viewpoint. However, as we discuss at the end of this paper, it is insignificant from at least some engineering perspectives.

3. Transfer Functions and Padé Approximants

A way to obtain finite dimensional approximations might be to directly approximate the transfer function (see, e.g., [19]) of the fractional derivative by a rational function, which can then be converted into an

equivalent set of ODEs. Considering half-order derivatives, for example, the transfer function of interest is $\sqrt{s}$.

A well-known method of obtaining rational approximations is that of Padé approximants, which are rational functions that match the Taylor series expansion of the original function up to a given number of terms. Unfortunately, we have found that Padé approximants for $\sqrt{s}$ give *unstable* transfer functions. Thus, approximating the transfer function directly is nontrivial. Note that considerations of transfer functions can indeed be used to construct very good finite dimensional approximations (see [11] and references therein).

In this paper, we do not attempt to directly approximate the transfer function. Following an indirect route, we convert the given infinite dimensional system involving a fractional order derivative into a different infinite dimensional system involving a partial differential equation (PDE) in which α, the fractional order of the derivative (assumed to lie between 0 and 1), appears as a free parameter. We then approximate solutions of the PDE using a Galerkin projection and, finally finite elements.

4. An Infinite Dimensional System

Consider the PDE (which could also be viewed as an ODE in t with a free parameter ξ)

$$\frac{\partial}{\partial t}u(\xi,t)+\xi^{(1/\alpha)}u(\xi,t)=\delta(t),\quad u(\xi,0^-)\equiv 0, \tag{5}$$

where $\alpha>0$ and $\delta(t)$ is the Dirac delta function. The solution is

$$u(\xi,t)=h(\xi,t)=\exp(-\xi^{1/\alpha}t), \tag{6}$$

where the notation $h(\xi,t)$ is used to denote "impulse response function." On integrating h with respect to ξ between 0 and ∞ we get a function only of t, given by

$$g(t)=\int_0^\infty h(\xi,t)\,d\xi=\frac{\Gamma(1+\alpha)}{t^\alpha}. \tag{7}$$

Abstractly, $g(t)$ is simply the impulse response of a linear, constant coefficient system starting from rest.

Now if we replace the forcing $\delta(t)$ in Equation (5) with some sufficiently well-behaved function $\dot{x}(t)$, then the corresponding response $r(t)$ of the same system, again starting from rest at $t=0$, is (the last two expressions are equivalent)

$$r(t)=\int_0^t g(t-\tau)\dot{x}(\tau)\,d\tau=\Gamma(1+\alpha)\int_0^t\frac{\dot{x}(\tau)}{(t-\tau)^\alpha}\,d\tau=\Gamma(1+\alpha)\int_0^t\frac{\dot{x}(t-\tau)}{\tau^\alpha}\,d\tau.$$

On comparison with Equation (3), we find that

$$r(t)\equiv\Gamma(1+\alpha)\Gamma(1-\alpha)D^\alpha[x(t)],$$

provided $x(t)\equiv 0$ for $t\le 0$, and $0<\alpha<1$.

In this way, we have replaced an α order derivative by the following operations:

1. Solve
$$\frac{\partial}{\partial t}u(\xi,t)+\xi^{(1/\alpha)}u(\xi,t)=\dot{x}(t),\quad u(\xi,0^-)\equiv 0. \tag{8}$$
2. Then integrate to find
$$D^\alpha x(t)=\frac{1}{\Gamma(1-\alpha)\Gamma(1+\alpha)}\int_0^\infty u(\xi,t)\,d\xi. \tag{9}$$

There is no approximation so far. The system chosen above was prompted by Chatterjee [20].

Equation (8) represents an infinite dimensional system, and so we have replaced one infinite dimensional system (fractional derivative) with another. The advantage gained is that we can use a Galerkin projection to reduce Equation (8) to a finite dimensional system of ODEs. In this way, a fractional derivative will be replaced by a finite number of ODEs, and numerical solution up to time t will require $\mathcal{O}(t)$ operations.

5. Galerkin Projection

For the Galerkin projection, we assume that Equation (8) is satisfied by

$$u(\xi,t)\approx\sum_{i=1}^{n}a_i(t)\phi_i(\xi),$$

where n is finite, the ϕ_i are to be chosen by us, and the a_i are to be solved for. The ϕ_i are called shape functions. The choice of shape functions will be discussed in the next section. Now we outline the Galerkin procedure for Equation (8).

Substituting the approximation for $u(\xi,t)$ in Equation (8), we define

$$R(\xi,t)=\sum_{i=1}^{n}\left\{\dot{a}_i(t)\phi_i(\xi)+\xi^{(1/\alpha)}a_i(t)\phi_i(\xi)\right\}-\dot{x}(t),$$

where $R(\xi,t)$ is called the residual. This residual is made orthogonal to the shape functions, yielding n equations:

$$\int_0^\infty R(\xi,t)\phi_m(\xi)\,d\xi=0,\quad m=1,2,\ldots,n. \tag{10}$$

The integrals above need to exist; this will influence the choice of ϕ_i in the next section. Equations (10) constitute n ODEs, which can be written in the form

$$\mathbf{A}\dot{\mathbf{a}}+\mathbf{B}\mathbf{a}=\mathbf{c}\dot{x}(t), \tag{11}$$

where $\mathbf{A}$ and $\mathbf{B}$ are $n\times n$ matrices, $\mathbf{a}$ is an $n\times 1$ vector containing a_is, and $\mathbf{c}$ is an $n\times 1$ vector. The entries of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{c}$ are

$$A_{mi}=\int_0^\infty\phi_m(\xi)\phi_i(\xi)\,d\xi,\quad B_{mi}=\int_0^\infty\xi^{(1/\alpha)}\phi_m(\xi)\phi_i(\xi)\,d\xi,\quad c_m=\int_0^\infty\phi_m(\xi)\,d\xi. \tag{12}$$

In this paper, we assume that the shape functions are always chosen such that the above integrals exist. The choices of shape functions may therefore depend on α. In all numerical examples that follow, we have used shape functions for which the above integrals are finite.

During numerical solution of (say) a second order system including both $\ddot{x}$ as well as $D^{\alpha}[x(t)]$, we will use the quantities x and $\dot{x}$ as parts of the state vector, along with the a_i above. Having access to $\dot{x}$ at each instant, therefore, we can solve Equation (11) numerically to obtain the a_i. Finally

$$D^{\alpha}[x(t)] \approx \frac{1}{\Gamma(1+\alpha)\Gamma(1-\alpha)}\mathbf{c}^{\mathrm{T}}\mathbf{a}, \tag{13}$$

where the T superscript denotes matrix transpose.

Advantages of the approximation proposed here, as we will demonstrate with examples, are that (i) by suitable choice of shape functions we can hope to obtain somewhat uniform performance over a given frequency range, (ii) systematic refinement of the approximation involves straightforward calculations, and (iii) the approximation depends on α, the fractional order of the derivative, but not on the system where this derivative appears.

6. Choice of Shape Functions

The shape functions are chosen keeping in mind that high frequency behavior (on short time scales) corresponds to large values of ξ, and low frequency behavior (on long time scales) to small values of ξ, due to the role ξ plays in $\exp(-\xi^{(1/\alpha)}t)$ (see Equation (6)).

Since the shape functions need to be defined over an infinite domain, we introduce an auxiliary variable z with a bounded domain as

$$z = \frac{\xi^2 + \{\xi/(1+\xi^2)\}}{1+\xi^2}. \tag{14}$$

A plot of $z(\xi)$ is given in Figure 1. The choice of z as a function of ξ is somewhat arbitrary. It is merely one choice that satisfies the following conditions:

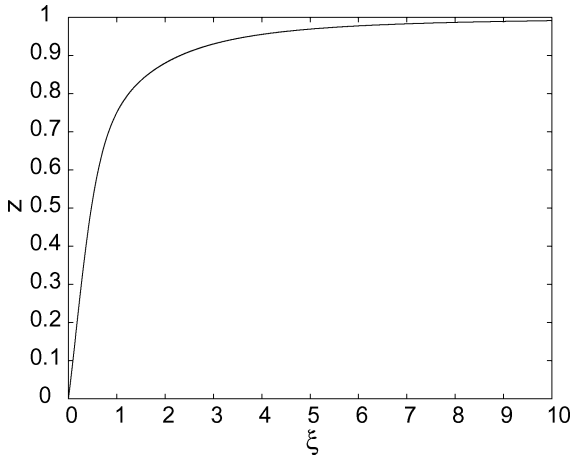

Figure 1. Plot of z versus ξ.

1. z increases monotonically from 0 to 1 as ξ increases from 0 to ∞.
   This condition is needed to inevitably map the infinite interval to the unit interval. Figure 1 verifies this condition.
2. z is approximately linear in ξ for small ξ.
   Figure 1 verifies this condition as well. This condition is useful for capturing reasonable variations for small ξ, which gives good performance over long times (low frequencies) as mentioned above.

For the given choice of $z(\xi)$ and temporarily looking at $\alpha = 1/2$ (we will consider other α values later), we choose the following shape functions:

$$\phi_1(\xi) = 1 - z \quad \text{and} \quad \phi_i(\xi) = \sin\left((i-1)\pi z\right), \quad \text{for } i = 2, 3, \ldots, n.$$

Again, there is arbitrariness in the choice of shape functions. If, for any i,

$$\lim_{z \to 1} \phi_i \neq 0, \tag{15}$$

then not all integrals involved in the approximation remain bounded. Thus, we have assumed that the function to be represented is arbitrary on the unit interval except that it is continuous, and zero at the right endpoint. The above choice of shape functions can, for large enough n, accurately match any such function. This is because any function $f(z)$ which satisfies $f(1) = 0$, upon subtraction of $(1 - z)f(0)$, gives a function that is zero at both $z = 0$ as well as $z = 1$; and such functions can be represented using Fourier sine series.

The shape functions chosen above are to be substituted in Equation (10), which still includes the free parameter $0 < \alpha < 1$, to obtain matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{c}$. Numerical evaluation of the integrals is required.

The matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{c}$ computed for $\alpha = 1/2$ and $n = 7$ are given in the appendix. Calculations with $\alpha = 1/2$ but other values of n were carried out as well, but only graphical results are presented in Section 8 for those cases.

7. Accuracy

The approximation mentioned previously turns out to be reasonably good, but we will eventually do much better (Section 12). For now, we check the above approximation by noting from Equations (11) and (13) that

$$\mathcal{F}(D^\alpha x(t)) = \frac{1}{\Gamma(1+\alpha)\Gamma(1-\alpha)} \mathbf{c}^{\mathrm{T}} \mathcal{F}(\mathbf{a}), \quad \text{when } x(t) \equiv 0 \text{ for } t \leq 0,$$

where $\mathcal{F}(\cdot)$ denotes the Fourier transform.

We choose $\alpha = 1/2$ and $n = 10$ for examination. In Figure 2, in the first and third plot, we compare the magnitude and phase of the frequency response function (FRF) obtained using the above expression with that of $(i\omega)^{1/2}$. The second plot presents relative error of the approximation. The approximation is good for frequencies over about 3 orders of magnitude, and can be refined further. The error is not uniform over the frequency range of interest, but can be made so with suitable choice of shape functions. This will be demonstrated later using finite elements.

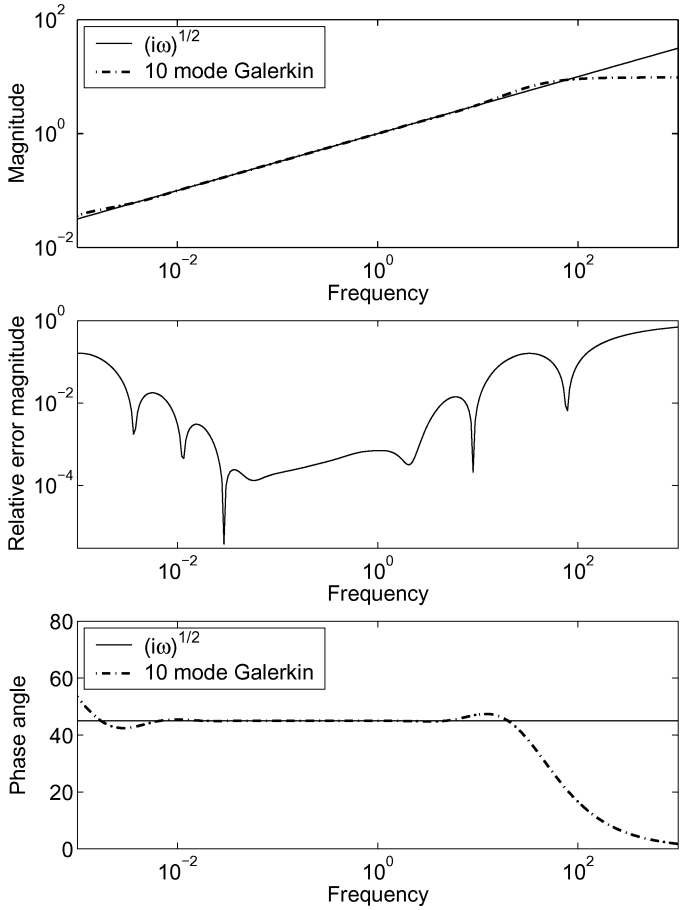

Figure 2. Comparison between magnitude and phase of $(i\omega)^{1/2}$, and approximated FRF using 10 mode Galerkin.

8. Numerical Examples

8.1. A Linear System

Consider the system

$$D^2x(t) + D^{1/2}x(t) + x(t) = \sin(2\pi t), \tag{16}$$

with $x(t) \equiv 0$ for $t \leq 0$ and $\dot{x}(0) = 0$.

Now we present the numerical results obtained using our method along with direct numerical solutions where the fractional derivative was calculated by evaluating the integral of Equation (4) (this numerical integration was also used in [14]).

For completeness, we outline the numerical procedure followed for $n = 7$. Combining Equations (11), (13), and (16), we now write

$$\ddot{x} + \frac{2}{\pi}\mathbf{c}^{\mathrm{T}}\mathbf{a} + x = \sin(2\pi t), \tag{17}$$

$$\mathbf{A}\dot{\mathbf{a}} + \mathbf{B}\mathbf{a} = \mathbf{c}\dot{x}, \tag{18}$$

with initial conditions $x(0) = 0$, $\dot{x}(0) = 0$, and $\mathbf{a}(0) = \mathbf{0}$, where matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{c}$ are evaluated by numerical integration as described earlier, and have been provided for reference in the Appendix.

Results obtained using the Galerkin approximation for $n = 3,\ 5,\ 7,$ and 10 are shown first in Figure 3, where it is seen that the solutions for $n = 3$ and 5 show some small mismatch, but the solutions for $n = 7$ and 10 are indistinguishable to plotting accuracy. Thus, satisfactory convergence is obtained for reasonable n and $\alpha = 1/2$.

In Figure 4 we compare our approximate solutions with direct numerical solutions of the original system (using the technique used by Wahi and Chatterjee [14]). Both $x(t)$ and $\dot{x}(t)$ plots agree very well with the direct numerical solution. As mentioned earlier, however, the present technique requires $\mathcal{O}(t)$ computations for the solution up to time t, while the direct numerical technique requires $\mathcal{O}(t^2)$ computations.

8.2. A Nonlinear System

Consider

$$D^2x(t) + D^{1/2}x(t) - x(t) + x(t)^3 = \sin(2\pi t), \tag{19}$$

with $x(t) \equiv 0$ for $t \leq 0$ and $\dot{x}(0) = 0$.

In Figure 5 we compare our approximation with the direct numerical solution. Again, there is very good agreement.

Having described the basic method and noted its apparently promising accuracy, we now move to finite element approximations. The shape functions of Section 6 will be replaced with ones defined over

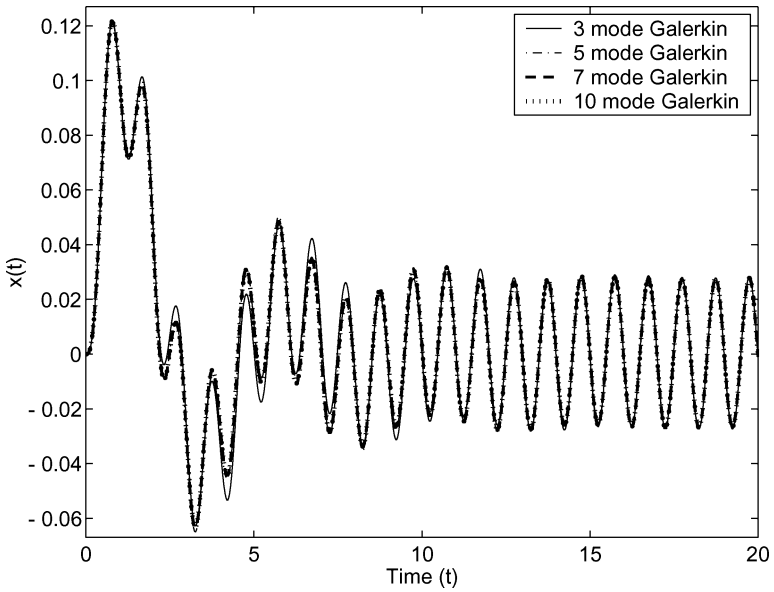

Figure 3. $x(t)$ against time, for $n = 3,\ 5,\ 7,$ and 10.

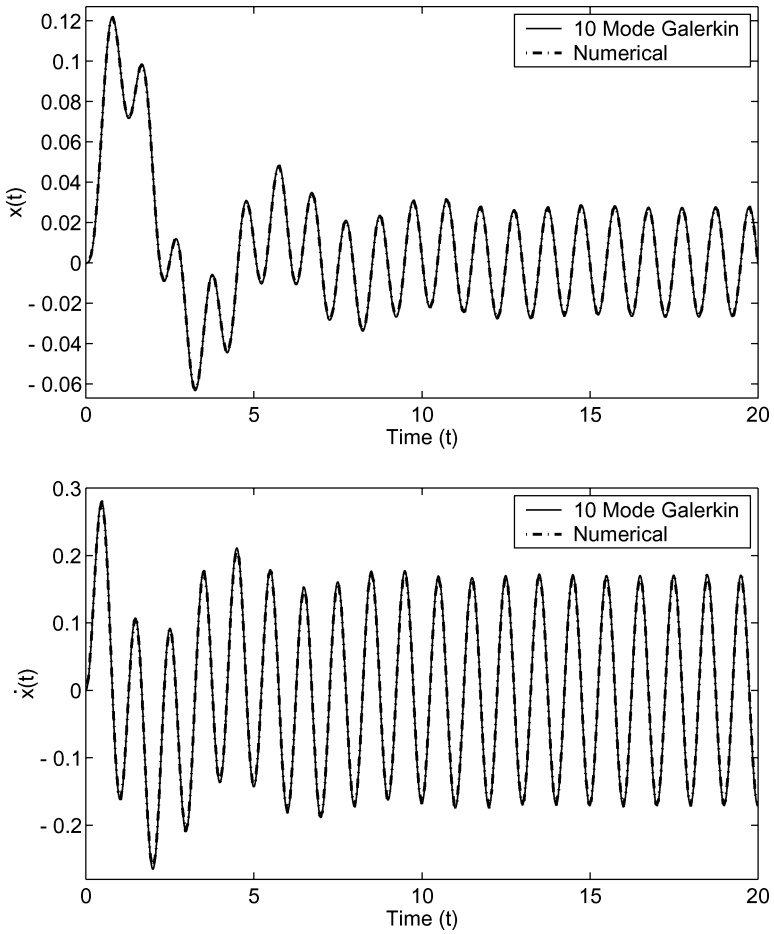

Figure 4. $x(t)$ and $\dot{x}(t)$ against time, for Galerkin ($n = 10$) and direct numerical solution.

finite elements. We will use the two simplest choices of element basis functions: the piecewise constant and the "hat" functions (the shape function for the last element will need to be different, as discussed later). We will also eventually consider values of α other than 1/2.

9. Finite Element Formulation: Uniform Element Size

We first develop an approximation using finite elements of uniform length defined on the unit interval. This will help clarify the need for nonuniform elements.

We define the following auxiliary variable $\eta(\xi)$ (note the change from the z used in Equation (14))

$$\eta(\xi) = \frac{\xi}{1+\xi}, \tag{20}$$

which maps the infinite domain to the unit interval [0, 1]. Comparing with Equation (14), we note that now ξ can be easily solved for in terms of η. A minor price paid is that, for large ξ, η converges to 1

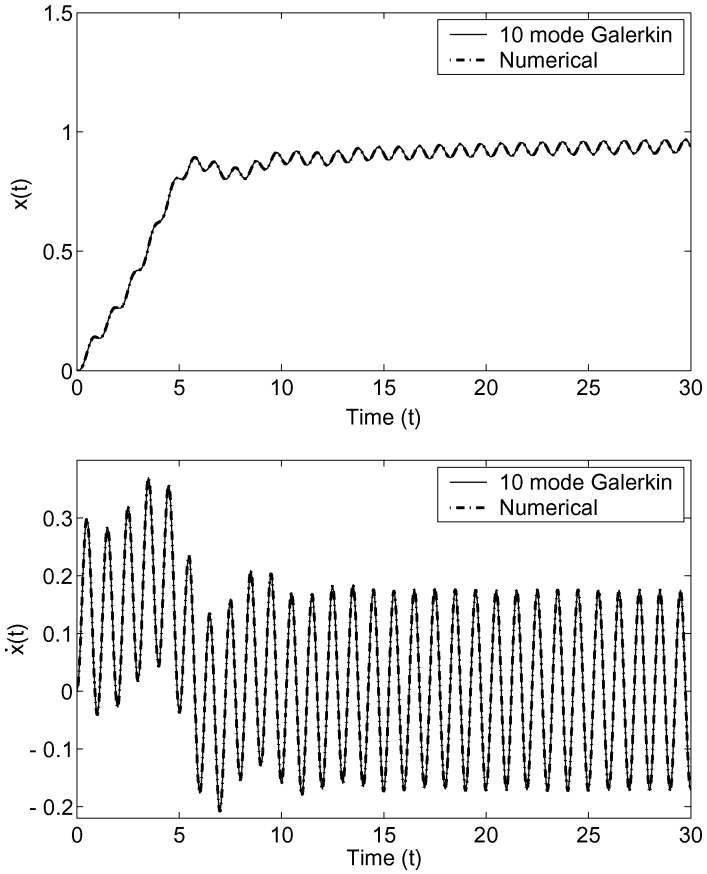

Figure 5. Results for Equation (19).

more slowly than z. This affects the restrictions we must place on the shape function(s) used on last element so as to ensure existence of all integrals involved in the approximation; we will use suitable shape functions for that element later.

Here, again, we temporarily restrict attention to $\alpha = 1/2$.

9.1. Piecewise Constant Case

For the piecewise constant case, the shape functions used are as follows (see Figure 6)

$$\phi_i(\eta) = \begin{cases} 1, & p_{i-1} < \eta \le p_i, \\ 0, & \text{elsewhere,} \end{cases} \qquad \text{for } i = 1, 2, \ldots, n-1,$$

$$\phi_n(\eta) = \begin{cases} \left(\dfrac{1-\eta}{1-p_{n-1}}\right)^2, & p_{n-1} < \eta \le 1, \\ 0, & \text{elsewhere,} \end{cases}$$

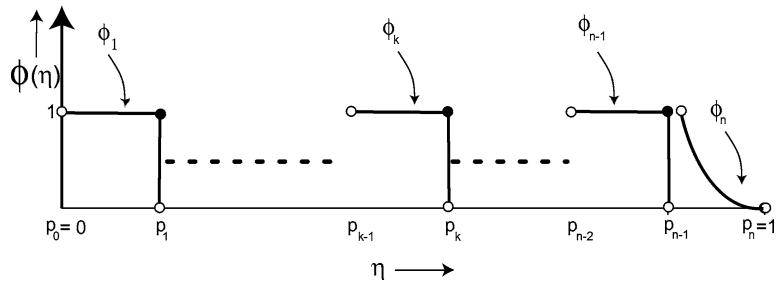

Figure 6. Piecewise constant shape functions. The solid and hollow circles at p_{n-1} are shown nearly coincident for visibility alone; in reality, they coincide exactly. The shape function on the last element is chosen to ensure boundedness of integrals.

where the element nodal point $p_0 = 0$, and $p_i - p_{i-1} = 1/n$ for each i. Notice that the shape function on the nth subinterval is the only one that needs attention to ensure boundedness of integrals; this choice works for $\alpha = 1/2$.

Now a Galerkin projection similar to Section 5 is performed after changing the integration variable to η, giving

$$\int_0^1 \left(\sum_{i=1}^{n} \left\{ \dot{a}_i(t)\phi_i(\eta) + \left(\frac{\eta}{1-\eta} \right)^{(1/\alpha)} a_i(t)\phi_i(\eta) \right\} - \dot{x}(t) \right) \phi_m(\eta) \frac{1}{(1-\eta)^2} d\eta = 0, \tag{21}$$

for $m = 1, 2, \ldots, n$. Equations (21) constitute n ODEs, which can be written in the form of Equations (11). On combining them with the ODE at hand, we get an initial value problem which can be solved numerically in $\mathcal{O}(t)$ time as before.

9.2. Hat Functions

For better accuracy, we can increase the smoothness of the shape functions. Here, we use the "hat" shape functions which are defined as follows (see Figure 7):

$$\phi_1(\eta) = \begin{cases} \dfrac{p_1 - \eta}{p_1}, & 0 \le \eta \le p_1, \\ 0, & \text{elsewhere}, \end{cases}$$

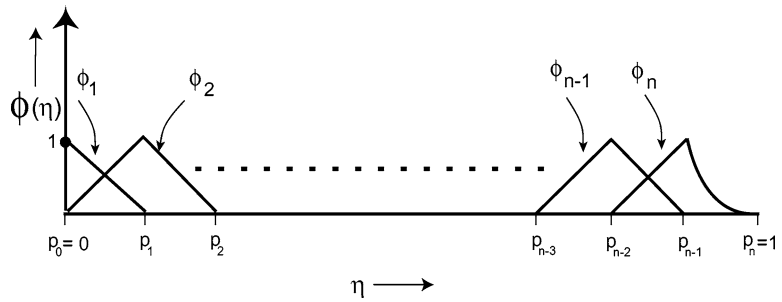

Figure 7. Hat shape functions.

$$\phi_{i+1}(\eta) = \begin{cases} \dfrac{\eta - p_{i-1}}{p_i - p_{i-1}}, & p_{i-1} \le \eta \le p_i, \\ \dfrac{p_{i+1} - \eta}{p_{i+1} - p_i}, & p_i \le \eta \le p_{i+1}, \\ 0, & \text{elsewhere,} \end{cases} \quad \text{for } i = 1, 2, \ldots, n-2,$$

and

$$\phi_n(\eta) = \begin{cases} \dfrac{\eta - p_{n-2}}{p_{n-1} - p_{n-2}}, & p_{n-2} \le \eta \le p_{n-1}, \\ \left(\dfrac{1-\eta}{1-p_{n-1}}\right)^2, & p_{n-1} \le \eta \le 1, \\ 0, & \text{elsewhere,} \end{cases} \tag{22}$$

where again $p_0 = 0$, and $p_i - p_{i-1} = 1/n$ for each i.

Here, we have chosen the last shape function to be identical over the last subinterval to that used with the piecewise constant shape functions (valid for $\alpha = 1/2$). Once again, the Galerkin projection gives n ODEs in the matrix form Equation (11), which on combining with the ODE at hand reduces to an initial value problem which is solved numerically.

We now check accuracy by comparing FRFs, as described in Section 7 (some of these results are also presented in [21]). The FRF for the less accurate piecewise constant case is calculated for 75 elements, whereas only 15 elements are used for the case of hat functions. Results are shown in Figure 8. It is clear that both approximations work well over a significant frequency range. Closer examination of

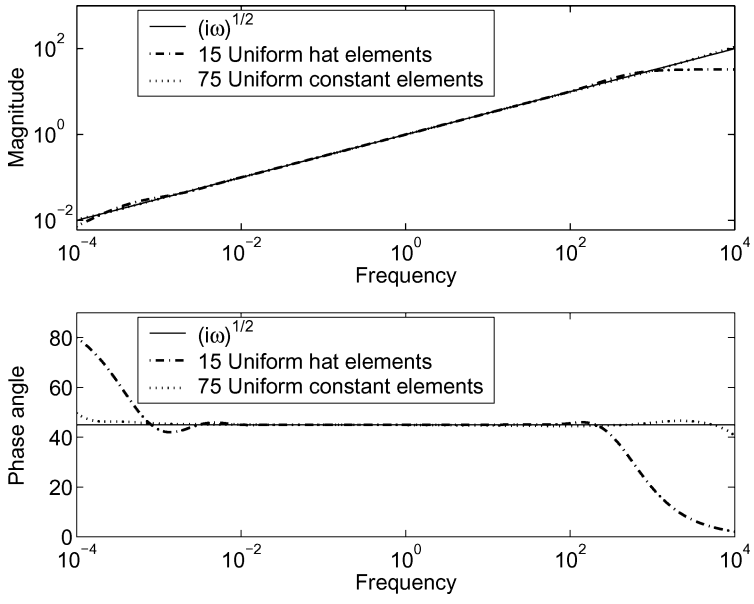

Figure 8. Piecewise constant versus hat function elements. Comparison between magnitude and phase of $(i\omega)^{1/2}$, and approximated FRFs, using 15 hat functions and 75 piecewise constant functions.

relative errors verified (plots not presented here) that the hat functions give superior accuracy within some frequency range, even with far fewer elements.

10. Nonuniform Element Sizes

On using uniform element sizes, the relative error in the FRF of the approximation varies significantly with frequency even over the range of good performance. This nonuniformity in the error can be reduced by using nonuniform element sizes (which is the great strength of finite element approximations).

The larger relative error for very low as well as high frequencies is due to the lack of refinement of elements for small and large values of η. The approximation can be improved by taking smaller elements near 0 and 1 in the η domain. Here, for demonstration, we have used nodal points that are equally spaced on a logarithmic scale in the ξ domain, as follows. We first define

$$\mathbf{y} = \text{logspace}(-\beta_1, \beta_2, n-1),$$

where "logspace" is shorthand for $n-1$ points that are logarithmically equally spaced between $10^{-\beta_1}$ and 10^{β_2}. We then set

$$p_i = \frac{y_i}{1+y_i}, \quad i = 1, 2, \ldots, n-1 \tag{23}$$

to get an $(n-1) \times 1$ array of nonuniformly spaced points in the interval (0, 1). We finally add two more nodes, at 0 and 1, to get an $(n+1) \times 1$ array of nodal locations.

FRFs calculated using the above spacing are shown in Figure 9 for 50 piecewise constant and 15 hat functions. It is seen that the relative error is now more uniform over a user-definable frequency range and the phase angle error is very small over that range as well. In particular, relative error in magnitude is less than 1% for frequency varying by 7 orders of magnitude; and the phase error over this range of frequencies is also quite small. This example demonstrates the potential for using finite element approximations that can be systematically refined as far as desired in any frequency range of interest (not including exactly zero) in a dynamic simulation. In many mechanical systems, the steady state static behavior is in any case determined by stiffness and not damping, and so large relative errors in approximating the low frequency behavior of the damping term may have negligible effects on the computed response in any case. Similarly, for many mechanical systems, the very high frequency behavior is dominated by inertia and not damping, with similar implications for approximation errors at high frequency.

11. Calculations with $\alpha \neq 1/2$

If $\alpha = 2/3$, say, then the shape functions used earlier for $\alpha = 1/2$ continue to give bounded integrals. However, if $\alpha = 1/3$, say, then they do not. For this case, if we suitably change the shape function on the last element (compare Equation (24) below with Equation (22); the exponent of 2 there is changed to 3 below) then all integrals involved are bounded. This points to an issue which we have not really addressed so far: what is the most suitable choice of shape function for the last element, for a given α? What are the consequences of an inappropriate choice? These issues will be pre-empted by Section 12.

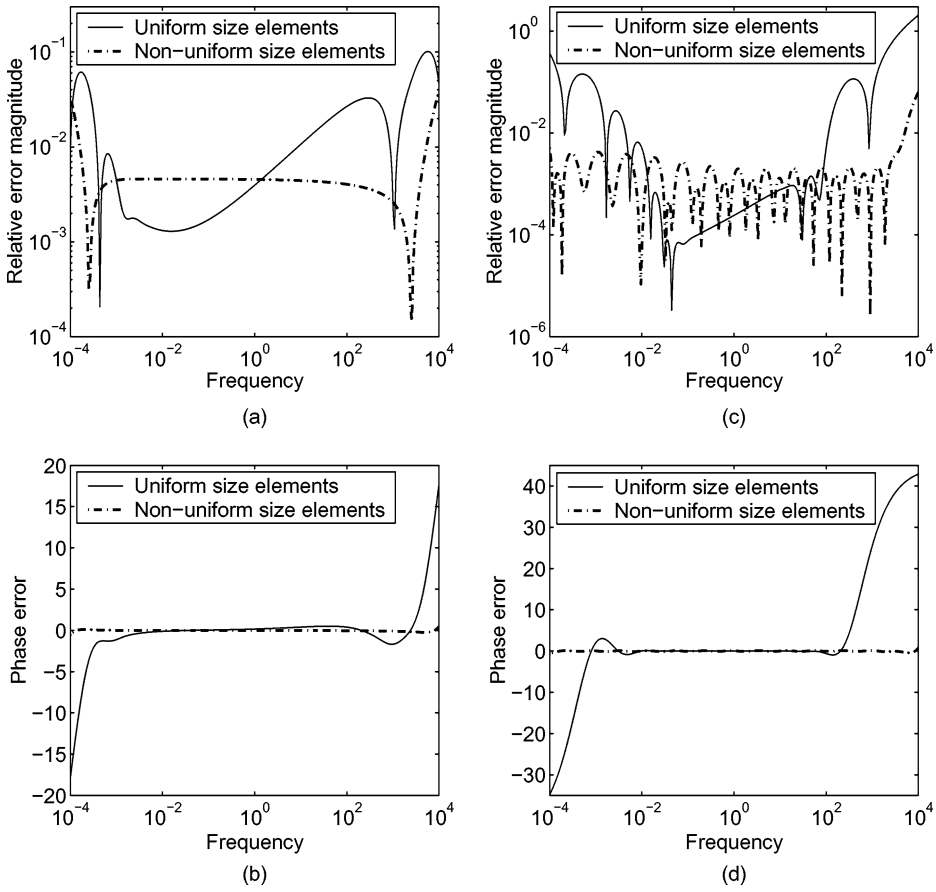

Figure 9. Uniformly versus nonuniformly spaced elements. (a) and (b) Relative error and phase error (compared with $(i\omega)^{1/2}$) in FRFs using 50 elements (piecewise constant) with uniform element sizes and nonuniform elements with nodes at logspace(−2, 2). (c) and (d) Relative error and phase error (compared with $(i\omega)^{1/2}$) in FRFs using 15 elements (hat functions) with uniform element sizes and nonuniform elements with nodes at logspace(−2, 2).

The changed shape function is

$$\phi_n(\eta) = \begin{cases} \dfrac{\eta - p_{n-2}}{p_{n-1} - p_{n-2}}, & p_{n-2} \le \eta \le p_{n-1}, \\ \left(\dfrac{1-\eta}{1-p_{n-1}}\right)^3, & p_{n-1} \le \eta \le 1, \\ 0, & \text{elsewhere.} \end{cases} \tag{24}$$

In Figure 10, we present the comparisons in FRFs for $\alpha = 1/3$ and $\alpha = 2/3$. We used 15 uniform finite elements and the same hat functions as before, except for the modification of Equation (24) in place of (22) when $\alpha = 1/3$. The comparison is good for $\alpha = 1/3$. In contrast, the results obtained for $\alpha = 2/3$ are not good. However, on approximating the 2/3 order derivative using two successive 1/3 order derivatives, we obtain good results.

We mention that we obtained similar results with derivatives of order 0.4 and 0.8; of these, the approximation for order 0.4 was much better than that for order 0.8, but the order 0.8 was satisfactorily

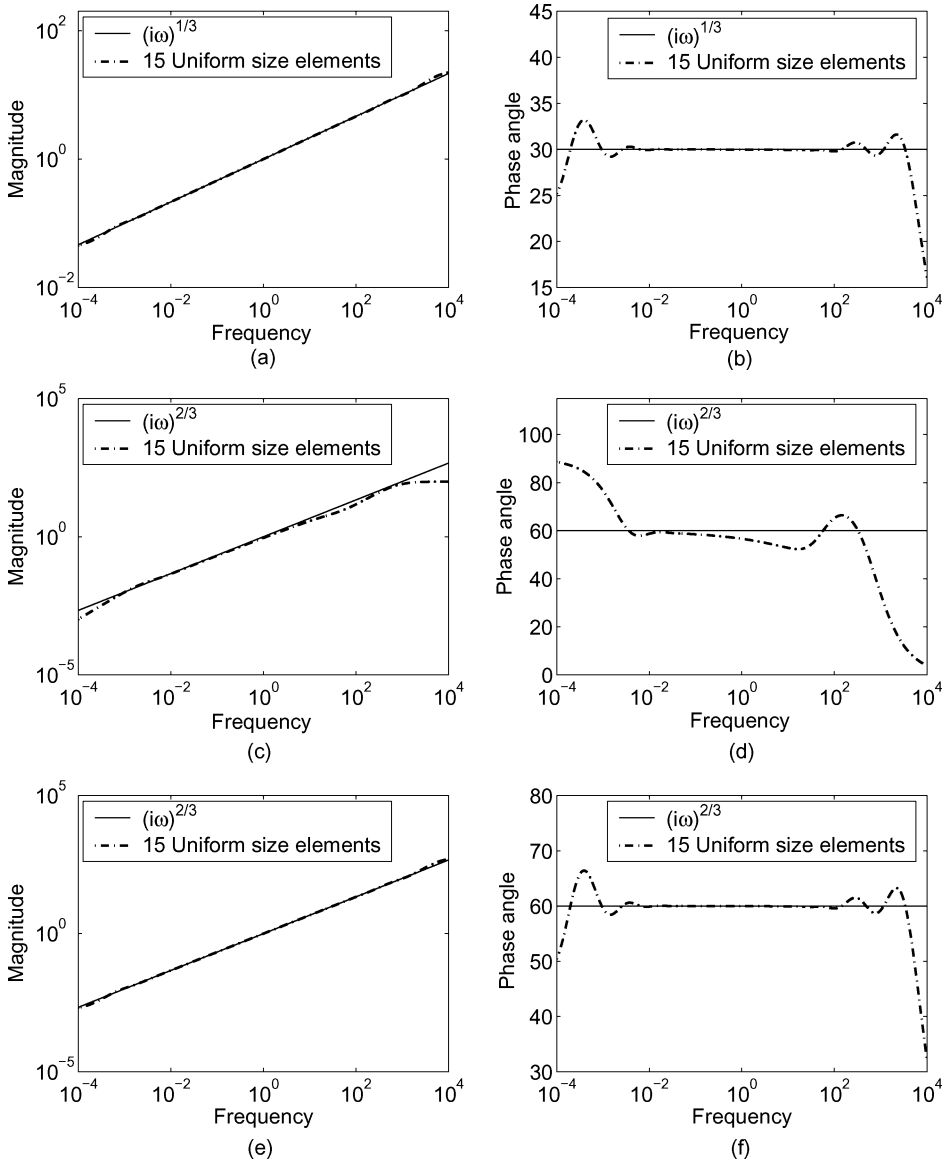

Figure 10. Magnitude and phase angle comparison in FRFs. (a) and (b) 15 uniform hat elements and $\alpha = 1/3$. (c) and (d) 15 uniform hat elements and $\alpha = 2/3$. (e) and (f) 15 uniform hat elements and two successive derivatives of order $1/3$ to achieve $\alpha = 2/3$.

approximated as two successive derivatives of order 0.4. Similar results, again, were obtained for orders 0.45 and 0.9. The approximations obtained for $\alpha = 0.2$ and 0.25 were also good. All these calculations were done using 15 uniform elements and hat functions. These results are not presented here for reasons of space.

On the basis of these studies, it appears that α-order derivatives for relatively higher values of α are not well approximated in the present scheme; but two successive $\alpha/2$-order derivatives, using the same scheme, give better results. The key to this puzzle lies in the α-independence of the mapping from ξ to η in Equation (20).

In what follows, we consider an α-dependent mapping. This gives the best performance so far, and all previous calculations may be seen as motivating Section 12.

12. The Final Scheme: An α-Dependent Mapping

Here, we replace Equation (20) by the following

$$\eta(\xi) = \frac{\xi^{(1/\alpha)}}{1 + \xi^{(1/\alpha)}} \tag{25}$$

which is again a monotonic mapping of $[0, \infty]$ to [0, 1], but now it depends on the order of the fractional derivative α. The advantage gained by using this mapping is that, now, we have more control on the performance of the method for a given frequency range. This is because of the role that ξ and t play in $\exp(-\xi^{(1/\alpha)}t)$. Here, we can consider $T^* \equiv 1/\xi^{*(1/\alpha)}$ for some time T^*. It suggests that frequency

$$F^* \equiv \xi^{*(1/\alpha)} = \frac{\eta^*}{1 - \eta^*}. \tag{26}$$

Thus, any frequency F^* corresponds to an α-independent point η^* on the unit interval. In other words, a given frequency F^* corresponds to a unique point η^* on the unit interval, independent of α.

We use hat function elements to perform the Galerkin projection (similar to Section 9). The only difference in this case is that, the last shape function is also a hat function (compare with Equation (22); the exponent of 2 there is changed to 1 here)

$$\phi_n(\eta) = \begin{cases} \dfrac{\eta - p_{n-2}}{p_{n-1} - p_{n-2}}, & p_{n-2} \le \eta \le p_{n-1}, \\ \dfrac{1 - \eta}{1 - p_{n-1}}, & p_{n-1} \le \eta \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

With the above-mentioned mapping, for any value of α between 0 and 1, all the involved integrals remain bounded for the last hat shape function element as well.

We directly use nonuniform size elements. To this end, we change Equation (23) to

$$p_i = \frac{y_i^{(1/\alpha)}}{1 + y_i^{(1/\alpha)}}, \quad i = 1, 2, \ldots, n - 1. \tag{27}$$

to get an $(n - 1) \times 1$ array of nonuniformly spaced points in the interval (0, 1); add two more nodes at 0 and 1; and get an $(n + 1) \times 1$ array of nodal locations.

We now come to an interesting point regarding the choice of mesh points in the nonuniform finite element discretization. While the map from ξ to η is now α-dependent, the *choice of mesh points* can be made using

$$p_i = \frac{y_i^2}{1 + y_i^2}, \quad i = 1, 2, \ldots, n - 1$$

with no negative consequences (see Equation (26) with $\alpha = 1/2$). The advantage is that the frequency range of interest can be specified easily in this way. We have used this in the numerical examples below.

In Figure 11, we present the comparisons in FRFs for $\alpha = 1/3$, $\alpha = 1/2$, and $\alpha = 2/3$. Fifteen nonuniform finite elements were used. The performance is very good for all cases over a significant frequency range. The percentage error in magnitude and phase angle for $\alpha = 1/3$, $\alpha = 1/2$, and $\alpha = 2/3$ are shown in Figure 12. The errors are below 1% for more than 7 orders of magnitude of frequency. Calculations for other values of α were also done, and similar results were obtained (not presented here). Similarly, we have also verified that taking more elements gives smaller errors over the same frequency range.

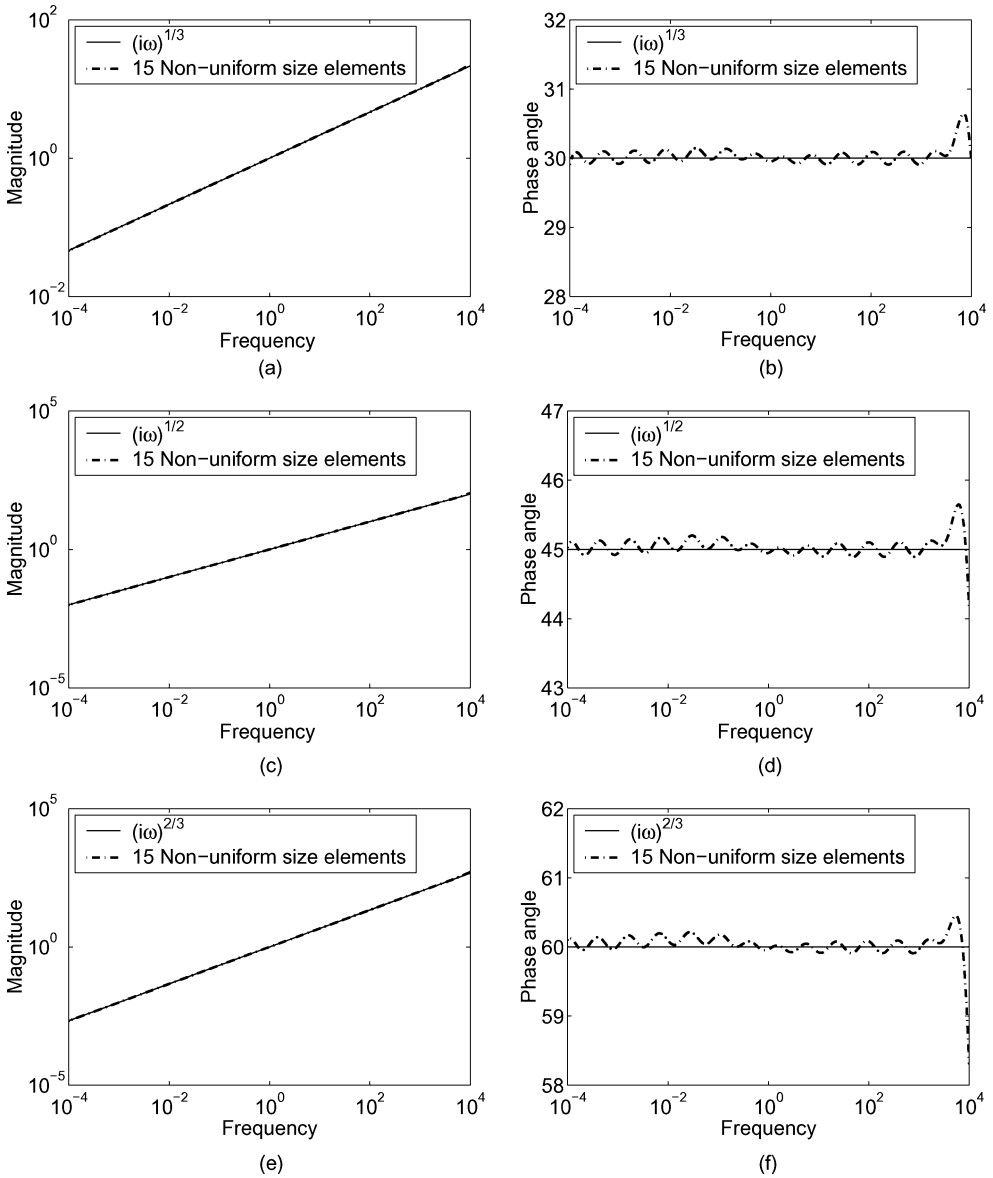

Figure 11. Magnitude and phase angle comparison in FRFs. (a) and (b) 15 nonuniform hat elements and $\alpha = 1/3$. (c) and (d) 15 nonuniform hat elements and $\alpha = 1/2$. (e) and (f) 15 nonuniform hat elements and $\alpha = 2/3$.

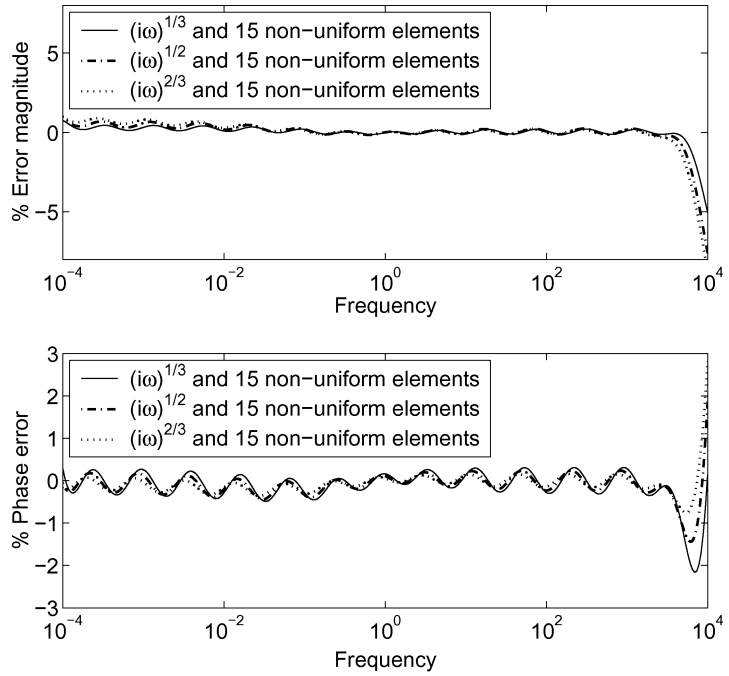

Figure 12. Percent error in magnitude and phase angle for $\alpha = 1/3$, $\alpha = 1/2$, and $\alpha = 2/3$.

For verification by readers who wish to implement this method, the system matrices for seven elements are given in the Appendix.

13. Modeling Issues and Asymptotics

No matter how many elements we take in the FE mesh, the match in the FRF will be good only over some nonzero finite range of frequencies. The very short time or very high frequency asymptotic behavior may *always be wrong*. See, for example, the discussion of [16] in [18]. An anonymous reviewer of this paper, who pointed us to these two papers, also raised similar concerns regarding this feature of our approximation.

This unavoidable feature may, however, have acceptably small implications for engineering practice. We do not suggest that Schmidt and Gaul lack understanding of engineering practice! Rather, their paper is a mathematical one; it raises mathematical concerns and discusses mathematical issues. Here, we present our views on what those issues might or might not imply for actual applications of fractional order derivatives in numerical work with experimentally fitted models for the viscoelastic behavior of physical materials tested in laboratories using imperfect machines with finite frequency ranges and finite measurement precision.

Consider some real material whose experimentally observed damping behavior can be well approximated using fractional order derivatives. We could, of course, also describe this behavior using a large number of (integer order) spring-dashpot combinations (as pointed out by Schmidt and Gaul [18]). The parameters of such integer order spring-dashpot combinations may be difficult to estimate robustly in experiments, however. Why?

As seen earlier, the Galerkin procedure gives very good approximations to fractional order derivatives for many different choices of mesh points. In other words, the same approximately-fractional-order behavior of the real material can be described by many different combinations of integer-order or classical spring-dashpot combinations; these combinations will do an *experimentally indistinguishable* job of capturing the experimental data, which will *always* span only a finite frequency range. In this way, the classical integer-order approach requires identification of many parameters that cannot really be uniquely determined. The parameter estimation problem is not only bigger, but more ill-posed. In contrast, a model involving fractional order derivatives may match the data over the relevant frequency range; and will also involve identification of fewer parameters in a better-posed problem. For this reason, *description* of damping should be done, wherever indicated, using such fractional order derivatives. This makes parameter identification easier for any individual experimenter; but, more importantly, it allows different experimenters in different laboratories to obtain the *same* parameter estimates, without which material behavior cannot be standardized for widespread engineering use.

However, once a suitable model with fractional order derivatives has been identified and standardized, *simulations* using that model can use different approximation techniques; it matters little what the approximation scheme is, provided it is good enough. The only issue for a given calculation is whether the final computed results are *accurate enough*. But what is accuracy?

For the numerical analyst, accuracy means correspondence with the original and exact fractional order derivative behavior. The approximation should be good over all frequencies and time scales that are important in the calculation. If the results are not reliable for some very high frequencies, the analyst notes it, but uses the reliable part of the results anyway. This is the same spirit in which reentrant corners and cracks in elastic bodies are often modeled using ordinary finite element codes: the results are not invalidated simply because even very small finite elements may not exactly capture the singularities in the solution. Rather, a careful analyst keeps a watch on how far from the singularity one must go before the numerical results are reliable.

For the engineer, in addition to the numerical issue, accuracy also means correspondence with the behavior of the original real material we started with. Any difference between exact and approximate mathematical solutions, in behavior regimes where there is no experimental data, are academic curiosities without practical implication in many cases. There is, in the end, no theorem which says that test results for a given piece of rubber can be extrapolated outside the test range; whether we believe it or not is a matter of individual academic taste.

Do our numerics limit the size of the allowable test range? We think not. As seen above, with merely 15 nonuniform hat function elements, we get errors less than 1% over a frequency range spanning more than 7 orders of magnitude which in a real setup might correspond to a frequency range, e.g., from 0.01 Hz to 100 kHz, or from 0.0001 Hz to 1 kHz. Are there mechanical models for anything at all where the range of validity exceeds this? If there are, we can take (say) 20 elements instead of 15. Note, also, that errors less than 1% in estimating material constants are considered small; for example, the Young's modulus of steel is routinely specified as 210 GPa, i.e., to two significant digits only. But, if necessary, the error in our approximation could be made less than 0.1% with more elements.

Finally, if the engineer believes (as discussed in [20]) that the fractional order derivative behavior observed in experiments is actually an artifact of many complex internal dissipation mechanisms, each without memory, then the very-low and very-high (outside the fitting range) frequency behavior of the material may actually *not* match the fitted fractional order behavior. In other words, where the approximation may not match the model, there the model does not match the material anyway.

It is with this viewpoint that we have presented the errors in our numerical approximations above using the frequency domain. While it may well appear in nonlinear differential equations, the fractional

derivative itself is very much linear. The error in the approximation of that derivative can therefore be viewed usefully in the frequency domain. A *post facto* study of the spectrum of the computed solution can show whether the solution was in fact confined within the frequency range where the approximation is good; if not, as is clear from our numerical results, the approximation can be improved by the analyst.

14. Conclusions

A problem with direct numerical simulation of fractionally damped systems is that simulations up to time t require $\mathcal{O}(t^2)$ computations. We have developed a novel Galerkin projection technique for reducing the infinite-dimensional systems associated with fractional order derivatives to finite dimensional systems. The reduction to a finite dimensional system is accomplished by first transforming the infinite dimensional, fractionally damped ODE to another infinite dimensional system consisting of an ODE coupled with a PDE. The Galerkin projection is then performed on the PDE. The approximation obtained is specific to α, the fractional order of the derivative; it can be used for any system where such a derivative appears.

With the present finite dimensional approximation, simulations up to time t of fractionally damped systems can now be done in $\mathcal{O}(t)$ time. It is emphasized that such time savings are only relative to straightforward numerical integration. Other finite dimensional approximations, such as in [10] and [11], will give similar time savings. However, the conceptual basis of our approximation is both different as well as, in our opinion, simpler.

We have presented calculations with global shape functions, and then with finite elements as well. Use of continuous, piecewise differentiable shape functions and finite elements of nonuniform size, along with an α-dependent mapping of the half line to the unit interval, has been shown, for a relatively small number of elements, to give excellent approximations with fairly uniform and small error over a user-specified frequency range. The present approximation is expected to be useful for long-time simulations of nonlinear dynamic systems with fractional order terms.

Appendix A: System Matrices

The matrices described in Section 5 are given here for $n = 7$. We present these for verification by readers implementing the calculation, not direct use. Users are advised to calculate the elements of these matrices to more digits of precision.

$$\mathbf{A} = \begin{bmatrix} 0.3836 & 0.5094 & -0.1495 & 0.3400 & -0.1403 & 0.2400 & -0.1048 \\ 0.5094 & 1.0663 & -0.8553 & 0.7806 & -0.6667 & 0.5728 & -0.4899 \\ -0.1495 & -0.8553 & 1.8469 & -1.5220 & 1.3534 & -1.1565 & 1.0063 \\ 0.3400 & 0.7806 & -1.5220 & 2.4197 & -2.0119 & 1.7869 & -1.5487 \\ -0.1403 & -0.6667 & 1.3534 & -2.0119 & 2.8532 & -2.4040 & 2.1533 \\ 0.2400 & 0.5728 & -1.1565 & 1.7869 & -2.4040 & 3.2196 & -2.7512 \\ -0.1048 & -0.4899 & 1.0063 & -1.5487 & 2.1533 & -2.7512 & 3.5533 \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} 0.3836 & 1.1486 & -2.0856 & 2.8635 & -3.5233 & 4.0857 & -4.5792 \\ 1.1486 & 3.4909 & -6.4706 & 8.9340 & -11.0176 & 12.7999 & -14.3592 \\ -2.0856 & -6.4706 & 12.4249 & -17.4881 & 21.7339 & -25.3767 & 28.5574 \\ 2.8635 & 8.9340 & -17.4881 & 25.2248 & -31.8473 & 37.4914 & -42.4180 \\ -3.5233 & -11.0176 & 21.7339 & -31.8473 & 40.9823 & -48.8886 & 55.7325 \\ 4.0857 & 12.7999 & -25.3767 & 37.4914 & -48.8886 & 59.2233 & -68.2649 \\ -4.5792 & -14.3592 & 28.5574 & -42.4180 & 55.7325 & -68.2649 & 79.6822 \end{bmatrix},$$

and

$$\mathbf{c} = \begin{bmatrix} 1.0708 \\ 2.4137 \\ -3.1911 \\ 4.2328 \\ -4.6777 \\ 5.3009 \\ -5.6135 \end{bmatrix}.$$

The system matrices obtained using α-dependent mapping in Section 12 are as follows for $n = 7$.

$$\mathbf{A} = \begin{bmatrix} 1.3822e{-}1 & 2.3129e{-}2 & 0 & 0 & 0 & 0 & 0 \\ 2.3129e{-}2 & 1.0951e{-}1 & 2.8911e{-}2 & 0 & 0 & 0 & 0 \\ 0 & 2.8911e{-}2 & 1.7846e{-}1 & 5.4390e{-}2 & 0 & 0 & 0 \\ 0 & 0 & 5.4390e{-}2 & 2.7544e{-}1 & 1.0044e{-}1 & 0 & 0 \\ 0 & 0 & 0 & 1.0044e{-}1 & 4.7224e{-}1 & 1.7158e{-}1 & 0 \\ 0 & 0 & 0 & 0 & 1.7158e{-}1 & 9.4760e{-}1 & 3.0324e{-}1 \\ 0 & 0 & 0 & 0 & 0 & 3.0324e{-}1 & 2.0920 \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} 1.3765e{-}4 & 9.2261e{-}5 & 0 & 0 & 0 & 0 & 0 \\ 9.2261e{-}5 & 1.8197e{-}3 & 9.6345e{-}4 & 0 & 0 & 0 & 0 \\ 0 & 9.6345e{-}4 & 1.8634e{-}2 & 1.1022e{-}2 & 0 & 0 & 0 \\ 0 & 0 & 1.1022e{-}2 & 1.5438e{-}1 & 1.1486e{-}1 & 0 & 0 \\ 0 & 0 & 0 & 1.1486e{-}1 & 1.3348 & 1.1040 & 0 \\ 0 & 0 & 0 & 0 & 1.1040 & 1.4873e1 & 1.1838e1 \\ 0 & 0 & 0 & 0 & 0 & 1.1838e1 & 3.0839e2 \end{bmatrix},$$

and

$$\mathbf{c} = \begin{bmatrix} 1.6135e{-}1 \\ 1.6155e{-}1 \\ 2.6176e{-}1 \\ 4.3027e{-}1 \\ 7.4426e{-}1 \\ 1.4224 \\ 3.7947 \end{bmatrix}.$$

Acknowledgments

Comments from an anonymous reviewer led us to develop Section 12, and to write Section 13, which made us think about many practical issues. These have helped to improve the paper.

References

1. Bagley, R. L. and Torvik, P. J., 'Fractional calculus: A different approach to the analysis of viscoelastically damped structures', *AIAA Journal* **21**(5), 1983, 741–748.
2. Bagley, R. L. and Torvik, P. J.,' Fractional calculus in the transient analysis of viscoelastically damped structures', *AIAA Journal* **23**(6), 1985, 918–925.
3. Gaul, L., Klein, P., and Kempfle, S., 'Impulse response function of an oscillator with fractional derivative in damping description', *Mechanics Research Communications* **16**(5), 1989, 297–305.
4. Makris, N., 'Fractional derivative model for viscous damper', *ASCE Journal of Structural Engineering* **117**, 1991, 2708–2724.
5. Chen, Y. and Moore, K. L., 'On D^{α}-type iterative learning control', in *Proceedings of the 40th IEEE Conference on Decision and Control*, Orlando, Florida, USA, 2001, pp. 4451–4456.
6. Debnath, L., 'Recent applications of fractional calculus to science and engineering', *International Journal of Mathematics and Mathematical Sciences* **2003**(54), 2003, 3413–3442.
7. Moreau, X., Ramus-Serment, C., and Oustaloup, A., 'Fractional differentiation in passive vibration control', *Nonlinear Dynamics* **29**(1–4), 2002, 343–362.
8. Podlubny, I., Petras, I., Vinagre, B. M., Chen, Y., O'Leary, P., and Dorcak, L., 'Realization of fractional order controllers', *Acta Montanistica Slovaca* **8**(4), 2003, 233–235.
9. Suarez, L. E. and Shokooh, A., 'An eigenvector expansion method for the solution of motion containing fractional derivatives', *ASME Journal of Applied Mechanics* **64**(3), 1997, 629–635.
10. Chen, Y., Vinagre, B. M., and Podlubny, I., 'Continued fraction expansion approaches to discretizing fractional order derivatives–An expository review', *Nonlinear Dynamics* **38**(1/2), 2004, 155–170.
11. Oustaloup, A., Levron, F., Mathieu, B., and Nanot, F. M., 'Frequency-band complex noninteger differentiator: Characterization and synthesis', *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications* **47**(1), 2000, 25–39.
12. Oldham, K. B., *The Fractional Calculus*, Academic Press, New York, 1974.
13. Koh, C. G. and Kelly, J. M., 'Application of fractional derivative to seismic analysis of base-isolated models', *Earthquake Engineering and Structural Dynamics* **19**, 1990, 229-241.
14. Wahi, P. and Chatterjee, A., 'Averaging oscillations with small fractional damping and delay terms', *Nonlinear Dynamics* **38**(1/2), 2004, 3–22.
15. Shinozuka, M., 'Monte carlo solution of structural dynamics', *Computers & Structures* **2**(5/6), 1972, 855–874.
16. Yuan, L. and Agrawal, O. P., 'Numerical scheme for dynamic systems containing fractional derivatives', *Journal of Vibration and Acoustics* **124**, 2002, 321–324.
17. Ford, N. J. and Simpson,A. C., 'The numerical solution of fractional differential equations: Speed versus accuracy', *Numerical Algorithms* **26**, 2001, 333–346.

18. Schmidt, A. and Gaul, L., 'On a critique of a numerical scheme for the calculation of fractionally damped dynamical systems', *Mechanics Research Communications* **33**, 2006, 99–107.
19. Ogata, K., *System Dynamics*, Prentice Hall, New Jersey, 1998.
20. Chatterjee, A., 'Statistical origins of fractional derivatives in viscoelasticity', *Journal of Sound and Vibration* **284**, 2005, 1239–1245.
21. Singh, S. J. and Chatterjee, A., 'Fractional damping: Statistical origins and galerkin projections', in *Proceedings of the ENOC-2005, the Fifth Euromech Nonlinear Dynamics Conference*, Eindhoven, The Netherlands, August 7–12, 2005.