Galerkin approach to approximate solutions of some nonlinear oscillator equations

D. Anderson
Department of Radio and Space Science, Chalmers University of Technology, SE 412 96 Göteborg, Sweden

M. Desaix
School of Engineering, University College of Borås, SE 501 90 Borås, Sweden

M. Lisak and J. Rasch
Department of Radio and Space Science, Chalmers University of Technology, SE 412 96 Göteborg, Sweden

(Received 11 February 2010; accepted 22 April 2010)

An analysis based on the Galerkin method is given of some nonlinear oscillator equations that have been analyzed by several other methods, including harmonic balance and direct variational methods. The present analysis is shown to provide simple yet accurate approximate solutions of these nonlinear equations and illustrates the usefulness and the power of the Galerkin method. © 2010 American Association of Physics Teachers. [DOI: 10.1119/1.3429974]

I. INTRODUCTION

Nonlinear oscillator equations play an important role in physics and engineering and have attracted much interest. Explicit analytical solutions of nonlinear oscillator equations are few, and either numerical solutions or approximate analytical techniques are frequently used. \(^1\) Many solution methods for nonlinear oscillator equations are based on different perturbation techniques and are algebraically cumbersome.

Another powerful and flexible technique for finding approximate solutions to many different equations in physics is variational methods such as Ritz optimization, \(^2\) which often produces analytically simple and accurate results. The idea of the direct variational approach is to reformulate the original equation as a variational problem and then to minimize the corresponding variational functional within a set of trial functions. This idea can be made more specific as follows: Consider the equation \(L[y(t)]=0\), with appropriate boundary conditions on the interval \([a,b]\). Assume that this equation can be rewritten as the variational problem \(\delta L[y]=0\), where the functional \(L[y]\) is determined by the Lagrangian \(L[y(t)]\). That is,

\[
\delta L[y] = \int_a^b \frac{\delta L}{\delta y} \frac{dy}{dt} dt = 0, \quad \text{with} \quad \frac{\delta L}{\delta y} = L[y(t)].
\]  

\(\delta L / \delta y\) denotes the variational derivative, that is, the comitant Euler–Lagrange variational equation is equivalent to \(L[y(t)]=0\). Instead of allowing a general variation \(\delta y\), the Ritz optimization procedure seeks an approximate solution within a restricted set of functions spanned by a function with a specified time dependence but with a degree of flexibility in form obtained by allowing the ansatz function to depend on certain parameters. That is, \(y(t) = y_T(t; a_1, a_2, \ldots, a_n)\), where \(y_T\) satisfies the boundary conditions for all parameters \(a_j\) with \(j=1,2,\ldots,n\). If this ansatz is inserted into the variational functional, the variational problem reduces to an ordinary optimization problem in the \(n\) variables \(a_n\) for the function \(J[y_T]=J[a_1, a_2, \ldots, a_n]\). The corresponding optimization equations \(\delta J / \delta a_j = 0\) determine the optimal choice of the parameters and the corresponding approximation of the solution.

A related but more general approximation method is the method of weighted residuals, which in its simplest form is called the Galerkin method. \(^3\) The first part of this method is the same as in the Ritz optimization procedure. An approximate solution is sought in the form of an ansatz function of specified time dependence but with flexibility allowed by including a number of parameters, that is, \(y(t) = y_T(t; a_1, a_2, \ldots, a_n)\), where \(y_T\) satisfies the boundary conditions for all parameters \(a_j\). Although this ansatz function does not usually satisfy the equation \(L[y]=0\) and instead gives rise to a residual \(R[y_T(t)] = L[y_T] \neq 0\), this residual can be made to vanish in a weighted averaged sense by multiplying it by weight functions \(w_j(t)\) and integrating over the interval. That is,

\[
\int_a^b L[y_T(t; a_1, a_2, \ldots, a_n)] w_j dt = 0, \quad (j = 1, 2, \ldots, n),
\]  

where Eq. (2) provides \(n\) relations for determining the \(n\) unknown parameters \(a_j\).

If the original equation allows a variational reformulation and the weight functions are taken as \(w_j = \frac{\partial L}{\partial y_j} / \partial a_j\), the corresponding variational and Galerkin equations coincide. \(^3\) However, the Galerkin approach is more flexible than the variational approach because it is applicable to a broader range of problems (for example, where a variational reformulation of the original equation is not possible) although at the loss of the optimization feature.

The purpose of this paper is to give a pedagogical demonstration of the application of the Galerkin method to some well known nonlinear oscillator equations and to illustrate the power and flexibility of this method.

II. THE NONLINEAR PENDULUM

The most well known nonlinear oscillator equation is the simple pendulum described by

\[
y'' + \omega_0^2 \sin y = 0, \quad y(0) = A, \quad y(0) = 0,
\]  

where \(y\) corresponds to the angle of oscillation and \(\omega_0\) is a constant determined by the length of the pendulum \(L\) and the
acceleration of gravity \( g \) according to \( \omega_0^2 = g/L \). In most investigations Eq. (3) is simplified by considering only small oscillations around the stationary solution \( y = 0 \), which implies that the equation can be linearized and reduced to the linear equation \( \ddot{y} + \omega_0^2 y = 0 \) with the simple solution: \( y(t) = A \cos(\omega_0 t) \). This solution has the linear property that the oscillation frequency is independent of the oscillation amplitude, a property that does not hold for higher oscillation amplitudes where the oscillation frequency depends on amplitude (a typically nonlinear feature).

In spite of its seeming simplicity, the problem of finding solutions for general initial amplitudes, \( A \), has occupied many eminent mathematical scientists, with the first one being Bernoulli in 1749.\(^5\) The problem is still often studied as an example of the application of different solution techniques.\(^6-10\)

There is an exact analytical solution of Eq. (3), although in the not very explicit form of an incomplete elliptic function. The most important information about the pendulum is the relation between the oscillation time, \( T \), and the amplitude, \( A \), which is given by\(^8\)

\[
\frac{T}{T_0} = \frac{2K(\sin(A/2))}{\pi},
\]

where \( K(k) \) denotes the complete elliptic integral of the first kind and \( T_0 = 2\pi / \omega_0 \) is the period in the limit of small amplitudes. Many ingenious methods have been suggested for finding approximate solutions of Eq. (3) and for finding simple analytical approximations for the frequency-amplitude relation, given exactly by Eq. (4). An early example is the approximation proposed by Bernoulli given by \( T/T_0 = 1 + A^2/16 \). For some recent papers on different approximation procedures to the nonlinear pendulum, see Refs. 5--10.

As an illustration of the power of the Galerkin method, it is interesting to reconsider this nonlinear oscillator equation. A simple trial function is \( y_T = A \cos(\omega t) \), and the problem is to determine the relation between the amplitude of the oscillations, \( A \), and the nonlinear frequency \( \omega = 2\pi / T \). A natural choice for the weight function is \( w(t) = \cos(\omega t) \). The condition of vanishing residual (integrated over a quarter period) implies that

\[
\int_0^{\pi/4} (\ddot{y} + \omega_0^2 \sin y) w(t) dt = 0.
\]

If we use the trial function for \( y(t) \) and the proposed weight function, the oscillation time is found to be

\[
\frac{T}{T_0} = \sqrt{\frac{A}{2J_1(A)}},
\]

where \( J_1(x) \) denotes the Bessel function of order one, and we have used the fact that \( \int_0^{\pi/2} \sin(A \cos x) \cos x dx = \frac{\pi}{2} J_1(A) \).

Equation (6) reduces to \( T/T_0 \rightarrow 1 \) as \( A \rightarrow 0 \), as it should. For a string pendulum the maximum allowable value is \( A = \pi/2 \) and for this amplitude the oscillation time is \( T/T_0 = 1.178 \), compared to the result \( T/T_0 \approx 1.180 \) from Eq. (4) (an error of only 0.2%). A rod pendulum would allow initial amplitudes up to \( A = \pi \) where the exact oscillation time would become infinite. The approximation given by Eq. (6) has qualitatively the same behavior, but reaches infinity only at a higher value of \( A \), corresponding to the first zero of \( J_1(A) \), that is, \( A \approx 3.83 \).

The expression for the oscillation time given by Eq. (6) was also obtained in Ref. 12 using the method of harmonic balance. This solution procedure first expands the nonlinearity \( \sin y \) in powers of \( y \) and then the function \( y(t) \) in a Fourier series containing odd multiples of \( \cos(\omega t) \). After identifying powers of \( \cos(\omega t) \), an expansion of \( T/T_0 \) in terms of \( A \) was found and then \textit{a posteriori} recognized as the expansion involving the first order Bessel function. In the Galerkin approach, the nonlinear relation linking the period and the amplitude of the oscillations is obtained with only a few calculations and using the explicit expression for the integral in Eq. (7).

In a variational approach to the nonlinear pendulum, the nonlinear oscillator equation is rewritten as the variational problem

\[
\delta \int_0^{\omega a = \pi/2} L(y, \dot{y}) dt = 0,
\]

where

\[
L = \frac{1}{2} y^2 + \omega_0^2 \cos y.
\]

In the direct variational approach, the functional is minimized within a smaller set of functions. A natural choice of trial function in this case is again \( y_T = A \cos(\omega t) \), where the frequency \( \omega \) can be considered as fixed and the optimal choice of the corresponding amplitude, \( A \), is to be found. We substitute this trial function into the variational integral and obtain the integrated or “reduced” Lagrangian

\[
\langle L \rangle = \int_0^{\omega a = \pi/2} L(y_T, \dot{y}_T) dt = \frac{\omega \pi A^2}{8} + \frac{\omega_0^2 \pi}{\omega^2} J_0(A),
\]

where \( J_0(x) \) denotes the zero order Bessel function. Variation with respect to the single parameter \( A \) yields [using \( dJ_0(A)/dA = -J_1(A) \)]

\[
\frac{\partial \langle L \rangle}{\partial A} = \frac{\omega \pi A}{4} - \frac{\omega_0^2 \pi}{\omega^2} J_1(A) = 0,
\]

which implies

\[
\left( \frac{T}{T_0} \right)^2 = \frac{A}{2J_1(A)}.
\]

Equation (12) is the same result as before, as it should be because the weight function in the Galerkin analysis is \( w(t) = \cos(\omega t) = \partial y_T / \partial A \).

III. OSCILLATORS WITH POWER LAW NONLINEARITIES

Recently, a series of papers\(^{13-19}\) has compared different techniques for finding approximate analytical solutions of nonlinear oscillator equations of the form

\[
\ddot{y} + f(y) = 0, \quad y(0) = A, \quad \dot{y}(0) = 0,
\]

where \( f(y) \) is an odd nonlinear function. The nonlinear pendulum equation is a special case of this form. Particular at-
tension has been given to nonlinear power law functions of the form \( f(y) = \text{sign}(y)|y|^p \), for example, \( p=1/3, 2, 4/3 \) and \( p=-1, 2, 4/3 \). Equations of the form of Eq. (13) appear in a broad range of applications. In addition to the applications in mechanics and plasma physics mentioned in Refs. 1, 12, and 16–19, equivalent equations also appear in diffusion and heat conduction problems involving concentration dependent diffusion constants and temperature dependent conductivities, respectively (see the Appendix). When analyzing equations of the form in Eq. (13), it is often convenient to reformulate them as an eigenvalue problem by suitable normalization to dimensionless variables (see Ref. 1 and the Appendix). A suitable normalization in the present case is to rescale the amplitude \( y/A \rightarrow y \) and time \( t/(T/4) \rightarrow t \), where \( T \) denotes the period. This choice of a new time variable implies that the first quarter period of the oscillations extends between \( t=0 \) and \( t=1 \). Equation (13) can then be rewritten as

\[
\ddot{y} + \lambda \text{sign}(y)|y|^p = 0, \tag{14}
\]

where \( \lambda = A^{p-1}/(T/4)^2 \). The boundary conditions are \( y(0)=1, \dot{y}(0)=0, \) and \( y(1)=0 \). This reformulation has the advantage that it directly determines the nonlinear scaling of the oscillation time with amplitude once the eigenvalue has been determined \( (T/4)^2 = A^{p-1}/2 \).

Of special interest is the case \( p=-1 \) for which Eq. (13) becomes

\[
\ddot{y} + \lambda y = 0, \tag{15}
\]

where \( \lambda = [T/(4A)]^2 \), which implies that \( T = 4\sqrt{\lambda}A \). In spite of the singular nonlinearity, Eq. (15) still has periodic solutions. As for the nonlinear pendulum, it can also be integrated to give the implicit solution

\[
t = \int_y^A \frac{dy}{\sqrt{2 \ln(A/y)}} = A\sqrt{\pi/2} \text{erf}(\sqrt{\ln(A/y)}), \tag{16}
\]

where \( \text{erf}(x) \) denotes the error function. This solution is not very explicit and several detailed investigations of varying complexity have been made of Eq. (13) to obtain good approximate solutions. These investigations have been based on various approximation methods, for example, the method of harmonic balance, and the direct variational method of Ritz optimization. Particular emphasis has been given to the determination of the nonlinear relation between the oscillator frequency and the amplitude. We reconsider Eq. (15) as a further illustration of the power and usefulness of the method of weighted residuals in the context of the Galerkin method. For the simplest application of the method of weighted residuals, it is convenient to rewrite Eq. (15) in the form used by several authors,

\[
\mathcal{L} = y\ddot{y} + \lambda = 0. \tag{17}
\]

In view of the previous investigations, a simple trial function satisfying the normalized boundary conditions is \( y_T = \cos(\pi t/2) \). For simplicity, we choose the same weight function \( w(t) = \cos(\pi t/2) \). We use this ansatz in Eq. (17) and obtain

\[
\lambda = \frac{-\int_0^1 y_T(t)\ddot{y}_T(t)w(t)dt}{\int_0^1 w^2(t)dt} = \frac{\pi^2}{2} \frac{\int_0^1 \cos^2 t dt}{\int_0^1 \cos^2 t dt} = \frac{\pi^2}{6}, \tag{18}
\]

which implies the relation \( T/A = (\sqrt{8}/3)\pi \approx 5.13 \), in comparison to the exact relation obtained from the implicit solution given by Eq. (16) of \( T/A = \sqrt{8}\pi \approx 5.01 \); the error is only 2.3%.

It is instructive to analyze this problem using the variational approach. The Lagrangian corresponding to Eq. (13) for \( f(y) = 1/y \) is

\[
\mathcal{L} = \frac{1}{2}y^2 - \ln|y|. \tag{19}
\]

As in the previous example, the frequency \( \omega_0 \) can be considered as given and the corresponding amplitude, \( A \), is to be determined. A suitable trial function is \( y_T = A \cos(\omega t) \). If we substitute this function into the variational integral and integrate over a quarter period, we obtain the reduced Lagrangian

\[
\langle \mathcal{L} \rangle = \frac{\pi \omega A^2}{8} - \frac{\pi \ln A}{2\omega} + \frac{\pi}{2} \ln 2. \tag{20}
\]

Variation with respect to \( A \) yields the relation \( \omega A = \sqrt{2} \) or \( T/A = \pi \sqrt{2} = 4.44 \), which is an error of \( \approx 11.4\% \). We see that in this case the Galerkin approach turns out to give a more accurate result than the variational approach despite the fact that the same trial function was used in both cases. The explanation is the fortuitous reformulation of the nonlinear oscillator equation from the original form given by Eq. (15) into that of Eq. (17), which is not in the form of the variational equation corresponding to the Lagrangian (although it is equivalent).

The cosine function may seem a natural choice of trial function in view of its importance for linear oscillators and of previous investigations that have all been based on this function. However, it is not obvious that it is the best choice of trial function in cases involving strongly nonlinear oscillations. An even simpler trial function satisfying the boundary conditions is the parabola \( y_T(t) = 1 - t^2 \). If we use this function and the weight function \( w(t) = y_T(t) \), the Galerkin approach gives \( \lambda = 8/5 \), which implies \( T/A \approx 5.06 \), an error of only 0.94%.

**IV. MORE FLEXIBLE TRIAL FUNCTIONS**

The Galerkin approach is easily generalized to more flexible trial functions, albeit typically at the expense of more complicated calculations. In view of the analysis based on the harmonic balance approach, we reconsider Eq. (17) using the trial function \( y_T = a_1 \cos(\pi t/2) + a_2 \cos(3\pi t/2) \) and weight functions \( w_1 = \cos(\pi t/2) \) and \( w_2 = \cos(3\pi t/2) \), where \( a_1 + a_2 = 1 \) to satisfy the initial condition. This relation together with the two Galerkin equations constitute three equations for determining \( a_1, a_2, \) and \( \lambda \). The subsequent calculations are straightforward but tedious, and the result is \( a_1 = 1.0676, a_2 = -0.0676, \) and \( \lambda \approx 1.6896 \), which implies \( T/A \approx 5.1994 \), that is, an error of 3.7%. This error is larger than that corresponding to the approximation using only the first term! This result emphasizes that the Galerkin approach does not necessarily improve with increasing number of terms. The Galerkin result can be compared to the result of the harmonic balance approach, which to second order gives \( a_1 = 10/9 \approx 1.11, a_2 = -1/9 \approx -0.11, \) and the corresponding
error for the T/A relation of 1.6%. We note that a variational analysis based on this trial function is not convenient because it leads to complicated integrals that can only be evaluated numerically.

To avoid excessive calculations, it is often advantageous to use trial functions of compact form rather than formal expansions. An example of a more flexible choice of trial function is \( y_\alpha(t) = \cos^\alpha(\pi t/2) \), where \( \alpha \) is a parameter to be determined. This choice allows for a more general form than the \( \cos(\pi t/2) \) variation used previously. Because we now have an extra parameter to determine, two weight functions must be chosen. The choice of weight functions in this case is not obvious, but as in the previous application we choose one of the weight functions as \( w_1(t) = y_\alpha(t) = \cos^\alpha(\pi t/2) \). This weight function is very small around \( t=1 \) and consequently suppresses any residual errors in this region. For this reason (and for simplicity), we choose as the second weight function the more "neutral" function, \( w_2(t) = 1 \). The two equations for the eigenvalue \( \lambda \) obtained using these weight functions are

\[
\lambda = -\frac{\int_0^1 y_\alpha(t) \dot{y}_\alpha(t) w_1(t) dt}{\int_0^1 w_1(t) dt} = \frac{\pi^2 \alpha}{6} \frac{G(\alpha/2 + 1)}{G((3\alpha - 1)/2)} \frac{G((3\alpha - 1)/2)}{G(3\alpha/2)}
\]

and

\[
\lambda = -\frac{\int_0^1 y_\alpha(t) \dot{y}_\alpha(t) w_2(t) dt}{\int_0^1 w_2(t) dt} = \frac{\pi^2 \alpha}{8\sqrt{\pi}} \frac{G(\alpha - 1/2)}{G(\alpha)}
\]

where \( G(x) \) denotes the Gamma function. These equations imply the transcendental equation for \( \alpha \),

\[
F(\alpha) = \frac{4\sqrt{\pi}}{3} \frac{G(\alpha)}{G((\alpha + 1)/2)} \frac{G((3\alpha - 1)/2)}{G(3\alpha/2)} = 1.
\]

After finding the value of \( \alpha \), the eigenvalue is given by Eq. (21b). Equation (22) is not possible to solve analytically but can be solved numerically with the result that \( \alpha \approx 0.75 \), which gives the eigenvalue \( \lambda \approx 1.53 \) and the relation \( T/A \approx 4.95 \), an error of only 1.2%.

A corresponding analysis can also be made using the trial function \( y_\alpha(t, \alpha) = (1 - t^\alpha)^\alpha \) and the weight functions \( w_1(t) = y(t, \alpha) \) and \( w_2(t) = 1 \). The result is qualitatively similar,

\[
\lambda = 4\alpha \frac{G(\alpha + 3/2)G(3\alpha - 1)}{G(\alpha)G(3\alpha + 1)} = \frac{\pi^2 \alpha}{4\sqrt{\pi}} \frac{G(\alpha)}{G(\alpha + 1)}
\]

The solution of this system of equations yields \( \alpha \approx 0.82, \lambda \approx 1.56 \), and \( T/A \approx 4.99 \), which constitutes an error of only 0.4%. A comparison between the different approximate solutions and the exact solution given by Eq. (16) is shown in Fig. 1.

A drawback of the latter two trial functions is that they give rise to transcendental equations for the parameter \( \alpha \) that can only be solved numerically. However, they give simple and explicit solutions once \( \alpha \) has been determined. This property can be an advantage if the solution is to be used for subsequent calculations.

V. DISCUSSION

The Galerkin method is a convenient and flexible approach for obtaining approximate solutions of many important equations. When applied to nonlinear oscillator problems, it provides simple, compact, and accurate approximate solutions. The Galerkin method is less systematic in its iteration procedure than, for example, the harmonic balance method, which has been applied to different nonlinear oscillator equations. However, as illustrated by our examples, the accuracy obtained is often surprisingly good and sufficient for most practical purposes.

APPENDIX: THE NONLINEAR DIFFUSION EQUATION

To illustrate the physical relevance of Eq. (13) for power law nonlinearities, we consider the well known one-dimensional diffusion equation given by

\[
\frac{\partial n}{\partial t} = \frac{\partial}{\partial x} \left( D \frac{\partial n}{\partial x} \right),
\]

where \( n \) denotes the diffusing quantity and \( D \) the diffusion constant. In many problems the diffusion coefficient, \( D \), depends nonlinearly on \( n \) and can be approximated by the power law form \( D = D_0 n^{\alpha n} \), where \( n_0 \) is a suitable normalization density, \( D_0 \) is the diffusion coefficient corresponding to this density, and \( \alpha \) is a parameter. The corresponding diffusion equation becomes

\[
\frac{\partial n}{\partial t} = \frac{\partial}{\partial x} \left( D_0 \left( \frac{n}{n_0} \right)^{\alpha n} \frac{\partial n}{\partial x} \right).
\]

For the boundary conditions \( n(t, L) = 0, n(0, 0) = N_0 \) and, using separation of variables, \( n(t, x)/n_0 = T(t)X(x) \). Eq. (A1) can be separated into two independent equations. The equation for the time dependent part can easily be solved to yield \( T(t) = (1 + \alpha t/\tau)^{-1/\alpha} \), where \(-1/\tau\) denotes the separation constant. [Without loss of generality we have assumed \( T(0) = 1 \).]
The equation for $X(x)$ is more complicated, and it is convenient to introduce the function $R(x) = X^{\alpha+1}$, which makes it possible to write the equation for $R$ in the form

$$\frac{d^2 R}{dx^2} + \lambda R^p = 0,$$  \hspace{1cm} (A3)

where $\lambda = L^2 R^{-1}(0)/\tau p D_0$, $R(0)=1$, and $R(1)=0$. The coordinate $x$ has been normalized according to $x/L \rightarrow x$, the amplitude $R$ to $R/R(0) \rightarrow R$, and $p = 1/(1 + \alpha)$. Then the equation (considered on the interval $[0,1]$ where $X$ can be assumed positive) is equivalent to Eq. (14).

4. L. Komzsik, Applied Calculus of Variations for Engineers (CRC, Boca Raton, FL, 2009).

---

**Superb Calculation of the Pressure in a Fluid**

Bob Panoff, Shodor Foundation

**Tune: Supercalifragilisticexpialidocious**

Bernoulli knew he had a rule he used for wings in air
For fluid incompressible he’d never have a scare.
The density of energy’s the same at every spot
A caveat is cavitation in which case it’s not!

Oh, Superb calculation of the pressure in a fluid
Is simple so that anyone with any sense can do it.
We all deserve a force conserved among the objects paired.
Just add to pressure rho gee aitch then add half rho vee squared

A water tower tower’s o’er a town so water goes
Through every pipe, and when you turn the faucet on it flows.
The pressure head is now instead a steady stream, you see,
The pipe’s diameter determines stream velocity.

The sum at every point’s a constant, check it if you care
Each term can change within a range for water or for air.
The key’s to keep the units straight and don’t have any gap
Or else your fluid starts to leak and then you’ll just get Oh... .