JOURNAL OF NORTHEASTERN UNIVERSITY

Volume 26 Issue 02, 2023 ISSN: 1005-3026 https://dbdxxb.cn/ Original Research Paper

A PRIORI ERROR ANALYSIS OF THE FEM SOLUTION FOR GENERAL LINEAR SECOND ORDER ODES

Wahran Shaker[1], Mohammad Sabawi[2*]

[1, 2] *Mathematics Department, College of Education for Women, Tikrit University, Iraq*

[1]wahran.hameed523@st.tu.edu.iq

Corresponding Author: [2*]mohammad.sabawi@tu.edu.iq

Abstract

In this paper, a priori error analysis has been investigated for continuous (conforming) Galerkin finite element method used for solving a general scalar linear second-order ordinary BVPs. We derived optimal order a priori error bounds in the H_0^1 (energy) and L_2 norms utilising the Ritz Projection and standard a priori error analysis techniques and tools.

Keywords: A priori error analysis, finite element methods, ordinary differential equations.

1. Introduction

The finite element method (FEM) is a broad family of numerical and approximate methods used for solving ordinary differential equations (ODEs) and partial differential equations (PDEs) and also it is used for solving integro-differential equations (IDEs). The FEMs have many excellent numerical features made them popular and widely used in scientific computing. The main advantage of the FEM is its ability for solving a wide variety of problems on different computational domains with different shapes. For example, finite difference methods (FDMs) can solve problems on rectangular and triangular meshes while generally, FEMs can handle geometries of any shapes.

The study of numerical solutions of ODEs has attracted the attention of researchers and important achievements have made. Estep [2] in 1995 investigated the use of FEM for time integration of initial value problems (IVPs) of ODEs and the researcher obtained an asymptotic error estimates. In 2000, Schötzau and Schwab [8] studied and used hp-version DG time-stepping method for solving initial value ODEs and also, they derived new explicit a priori error estimates. In 2001, Estep and Stuart [3] considered and investigated the dynamical behaviour of the DG method for ODEs. Wihler in 2005 in [9] examined the use of p-version continuous Galerkin (CG) time-stepping method for approximating the solution of the nonlinear IVPs and the author obtained explicit a priori error estimates in L_2 - and H^1-norms. Schieweck in 2010[7] presented and examined a discontinuous Galerkin-Petrov (DGP) time discretisation of evolutionary problems in Hilbert Space.

Matthies and Schieweck in 2011 [6] used higher order CG and DG variational methods for obtaining the approximate solution of a nonlinear system of ODEs. Janssen and Wihler in 2014 [5] examined the hp - CG and DG time-stepping methods for nonlinear IVPs and they

Submitted: 26/05/2023 **Accepted**: 27/06/2023

discussed the existence of their discrete solutions. Zhao and Wei in 2014 [10] presented and used a unified DG framework for nonlinear ODEs. Holm and Wihler in 2016 [4] analysed and used CG and DG time-stepping methods of arbitrary order for solving nonlinear IVPs in real Hilbert spaces. In 2016, Quarteroni and coworkers [1] examined and studied a new high order discontinuous Galerkin (DG) FEM for the time integration of Cauchy problem ODEs and they proved and obtained a priori error bounds for the proposed method.

In this paper, we derived a priori error estimates of the FEM solution of generic linear second-order ordinary BVPs using conforming Galerkin finite element method. The main contribution of this paper is deriving optimal order a priori error bounds in the H_0^1 norm. Also, we obtained optimal and suboptimal order a priori error bounds in the L_2 norm for general scalar linear second-order (BVPs) ODEs.

This paper is organised as follows. In section 2 we give the necessary notations and relevant preliminaries of the topic. Section 3 is devoted for the a priori error analysis for the general scalar linear second-order ordinary BVPs. The conclusions are given in section 4.

2. Problem Setting and Notation[1]

We consider a general scalar linear second-order BVP ODE

$$\begin{aligned} -au'' &\quad +bu' + cu = f, \text{ in } I, \\ u(\alpha) &\quad = u(\beta) = 0, \end{aligned} \tag{1}$$

where $I = (\alpha, \beta)$, the solution function $u \in H_0^1(I)$, the source function $f \in L_2(I)$ and the coefficients $a > 0, b, c \geq 0$. For simplicity of notations, we use $\mathcal{H} = H_0^1(I)$, $\|\cdot\|_{\mathcal{H}} = \|\cdot\|_0$ for the energy norm and $\|\cdot\|_{L_2(I)} = \|\cdot\|$ for the L_2 norm.

3. A Priori Error Analysis

The a priori error analysis is very important topic in the study of error and convergence analysis of differential equations using FEMs and other methods. In a priori error analysis we are interested in finding an error estimator of the form

$$\| e \|_V = \|u - u_h\|_V \leq E(u, f; V). \tag{2}$$

Notice that in general the bound in the a priori error analysis depends upon the data of the problem, the source function f, the exact solution u of the problem and the space V. The a priori error bounds in general are not computable since they depend on the exact solution of the problem u which in most cases is unknown. The a priori error analysis is used in the study of convergence of the exact solution of the original problem. It is used in finding the order of convergence of the exact solution and it tells us the required information about how the convergence is fast or how it is slow. In our problem, the a priori error bound depends on the data of the problem, the exact solution u of the original problem (1) and the space V. In this section, we consider deriving a priori error bounds for a generic scalar linear second-order ordinary BVPs. The error is split in the following form

$$\begin{aligned} e = u - u_h &= (u - R_h u) + (R_h u - u_h) \\ &= \rho + \theta, \end{aligned} \tag{3}$$

where $R_h u \in \mathcal{H}_h$ is the Ritz projection of the exact solution $u \in \mathcal{H}$, where $\mathcal{H}_h$ is a finite dimensional subspace of $\mathcal{H}, \rho = u - R_h u$ represents the Ritz projection error which is available in the literature. The idea here is to bound the quantity $\theta = R_h u - u_h \in \mathcal{H}_h$ by a bound depends on ρ (since we do not have a bound for θ), consequently, the whole error e can then be bounded in terms of ρ, i.e.,

$$\| e \|_V = \|u - u_h\|_V = \|(u - R_h u) + (R_h u - u_h)\|_V = \| \rho + \theta \|_V \leq \| \rho \|_V + \| \theta \|_V, \tag{4}$$

where $V = \{\mathcal{H}, L_2\}$. Then, we need to bound θ by a bound in terms of ρ, i.e.,

$$\| \theta \|_V \leq \mathcal{B}(\rho), \tag{5}$$

finally, the whole error is bounded by a bound in terms of ρ

$$\begin{aligned} \| e \|_V = \| \rho + \theta \|_V \leq \| \rho \|_V + \mathcal{B}(\rho) \\ = \hat{\mathcal{B}}(\rho). \end{aligned} \tag{6}$$

3.1. A Priori Error Analysis of Linear Elliptic Problems

In this section, we derive a priori error bounds for a general scalar linear second-order BVP ODEs

Theorem 3.1 ($\mathbf{H_0^1}$ a Priori Error Bounds for Generic Scalar Linear Second-Order BVP ODEs) The finite element approximate solution u_h of the problem (1), satisfies the following a priori H_0^1 error estimate

$$\| e \|_0 = \|u - u_h\|_0 \leq C_5 h \|u''\|. \tag{7}$$

Proof. We start by writing the problem (1) in the weak form: find $u \in \mathcal{H}$ such that

$$a \int_I u'v'dx + b \int_I u'vdx + c \int_I uvdx = \int_I fv, \forall v \in \mathcal{H}. \tag{8}$$

The finite element problem is: find $u_h \in \mathcal{H}_h \subset \mathcal{H}$ such that

$$a\int_I u_h' v' dx + b\int_I u_h' v dx + c\int_I u_h v dx = \int_I f v dx, \forall v \in \mathcal{H}. \quad (9)$$

Subtracting (9) from (8), then testing with $v = v_h \in \mathcal{H}_h$ to have

$$a\int_I (u-u_h)' v_h' dx + b\int_I (u-u_h)' v_h dx + c\int_I (u-u_h) v_h dx = 0, \forall v_h \in \mathcal{H}_h. \quad (10)$$

Inserting (3) in (10) yields

$$a\int_I (\rho+\theta)' v_h' dx + b\int_I (\rho+\theta)' v_h dx + c\int_I (\rho+\theta) v_h dx = 0, \quad (11)$$

which implies that

$$\begin{aligned} a & \int_I \rho' v_h' dx + a\int_I \theta' v_h' dx + b\int_I \rho' v_h dx + b\int_I \theta' v_h dx \\ +c & \int_I \rho v_h dx + c\int_I \theta v_h dx = 0, \end{aligned} \quad (12)$$

which results in

$$\begin{aligned} a & \int_I \theta' v_h' dx + b\int_I \theta' v_h dx + c\int_I \theta v_h dx = -a\int_I \rho' v_h' dx \\ -b & \int_I \rho' v_h dx - c\int_I \rho v_h dx. \end{aligned} \quad (13)$$

Note that

$$\int_I \rho' v_h' dx = \int_I (u - R_h u)' v_h' dx = 0, \forall v_h \in \mathcal{H}_h.$$

Therefore, we obtain

$$\begin{aligned} a\int_I \theta' v_h' dx + b\int_I \theta' v_h dx + c\int_I \theta v_h dx \\ = -b\int_I \rho' v_h dx - c\int_I \rho v_h dx. \end{aligned} \quad (14)$$

Now testing (14) with $v_h = \theta$, we get

$$a\int_I (\theta')^2 dx + b\int_I \theta'\theta dx + c\int_I \theta^2 dx = -b\int_I \rho'\theta dx - c\int_I \rho\theta dx, \tag{15}$$

which is equivalent to

$$a \parallel \theta \parallel_0^2 + b\int_I \theta'\theta dx + c \parallel \theta \parallel^2 = b\int_I -\rho'\theta dx + c\int_I -\rho\theta dx. \tag{16}$$

Using Cauchy-Schwarz inequality on both sides leads to

$$a \parallel \theta \parallel_0^2 + b \parallel \theta \parallel_0 \parallel \theta \parallel + c \parallel \theta \parallel^2 \leq b\|\rho'\| \parallel \theta \parallel + c \parallel \rho \parallel \parallel \theta \parallel. \tag{17}$$

Utilising Young's inequality results in

$$\hat{a} \parallel \theta \parallel_0^2 + c \parallel \theta \parallel^2 \leq b \parallel \rho \parallel_0^2 + c \parallel \rho \parallel^2, \tag{18}$$

where $\hat{a} = 2a + b$. Since $\parallel \theta \parallel_0^2, \parallel \theta \parallel^2 > 0$, then using the inequality

$$\text{If } \gamma + \delta \leq m \text{ then } \gamma \leq m, \text{ where } \gamma, \delta, m \geq 0, \tag{19}$$

when $\gamma = \hat{a} \parallel \theta \parallel_0^2, \delta = c \parallel \theta \parallel^2$ and $m = b \parallel \rho \parallel_0^2 + c \parallel \rho \parallel^2$ in (18), we have

$$\parallel \theta \parallel_0^2 \leq \hat{b} \parallel \rho \parallel_0^2 + \hat{c} \parallel \rho \parallel^2, \tag{20}$$

where $\hat{b} = b/\hat{a}$ and $\hat{c} = c/\hat{a}$. Using Ritz projection error bounds, we have

$$\parallel \rho \parallel = \|u - R_h u\| \leq C_1 h^2 \|u''\|, \tag{21}$$
$$\parallel \rho \parallel_0 = \|(u - R_h u)'\| \leq C_2 h \|u''\|. \tag{22}$$

Substituting (21) and (22) in (20) we get

$$\parallel \theta \parallel_0^2 \leq C_3 h^2 \|u''\|^2, \tag{23}$$

where $C_3 = \hat{c}^2 C_1^2 + \hat{b}^2 C_2^2$. Taking the square root, we arrive at

$$\parallel \theta \parallel_0 \leq C_4 h \|u''\|, \tag{24}$$

where $C_4 = C_3^{1/2}$. Now, using (4), we have

$$\begin{aligned} \parallel e \parallel_0 &= \|u - u_h\|_0 = \parallel \rho + \theta \parallel_0 \leq \parallel \rho \parallel_0 + \parallel \theta \parallel_0 \\ &\leq C_2 h \|u''\| + C_4 h \|u''\| \leq C_5 h \|u''\|, \end{aligned} \tag{25}$$

where $C_5 = C_2 + C_4$. Note that the bound in (25) is the final a priori estimate of the finite element error.

Theorem 3.2 (L_2 a Priori Error Bounds for Generic Scalar Linear Second-Order BVP ODEs) The finite element approximate solution u_h of the problem (1), satisfies the following a priori L_2 error estimates:

1. When $b > 0$ (in the presence of the integral $\int_J e'\theta dx$ and without performing an integration by parts on it) we have the following L_2 suboptimal error bound

$$\| e \| = \|u - u_h\| \leq C_8 h \|u''\|. \tag{26}$$

2. When $b = 0$ (in the absence of the integral $\int_J e'\theta dx$) we have the following L_2 optimal error bound

$$\| e \| = \|u - u_h\| \leq C_9 h^2 \|u''\|. \tag{27}$$

3. When $b > 0$ (in the presence of the integral $\int_J e'\theta dx$ and with performing an integration by parts on it) we have the following L_2 optimal error bound

$$\| e \| = \|u - u_h\| \leq C_{10} h^2 \|u''\|. \tag{28}$$

Proof. 1. Assume that $b > 0$ and by following the same steps in Theorem (3.1) We arrive at (18). Now, reusing the inequality (19) when $\gamma = c \| \theta \|^2, \delta = \hat{a} \| \theta \|_0^2$ and $m = b \| \rho \|_0^2 + c \| \rho \|^2$ in (18), we have

$$\| \theta \|^2 \leq \tilde{b} \| \rho \|_0^2 + \| \rho \|^2, \tag{29}$$

where $\tilde{b} = b/c$. Now, applying Ritz projection error bounds in (21) and (22) in (29) we have

$$\| \theta \|^2 \leq C_6 h^2 \|u''\|^2, \tag{30}$$

where $C_6 = C_1^2 + \tilde{b} C_2^2$. Taking the square root gives us

$$\| \theta \| \leq C_7 h \|u''\|, \tag{31}$$

where $C_7 = C_6^{1/2}$. Finally, inserting (21) and (31) in (4) we have

$$\| e \| \leq \| \rho \| + \| \theta \| \leq C_8 h \|u''\|, \tag{32}$$

where $C_8 = C_1 + C_7$.

2. Notice that the bound in (32) is suboptimal due to the presence of the term $b\int_J e'\theta dx$. Now, to obtain an optimal error bound let $b = 0$ in (17), then we have

$$\| \theta \| \leq \| \rho \| \leq C_1 h^2 \|u''\|. \tag{33}$$

Inserting (21) and (33) in (4) leads to the bound

$$\| e \| \leq \| \rho \| + \| \theta \| \leq 2 \| \rho \| \leq C_9 h^2 \|u''\|, \tag{34}$$

where $C_9 = 2C_1$.

3. Note that we can overcome the suboptimality of the term $b\int_{\mathcal{J}} e'\theta dx$ by using integration by parts

$$b\int_{\mathcal{J}} e'\theta dx = b[e\theta]_{\alpha}^{\beta} - b\int_{\mathcal{J}} e\theta' dx = -b\int_{\mathcal{J}} e\theta' dx. \tag{35}$$

Plugging (35) in (11) and testing it with $v_h = \theta$ and after some mathematical manipulations we find

$$a_0 \| \theta \|_0^2 + c \| \theta \|^2 \leq d \| \rho \|^2, \tag{36}$$

where $a_0 = 2a$ and $d = b + c$. Applying the inequality (19) in (36) when $\gamma = c \| \theta \|^2, \delta = a_0 \| \theta \|_0^2$ and $m = d \| \rho \|^2$ yields

$$\| \theta \| \leq \hat{d} \| \rho \|, \tag{37}$$

where $\hat{d} = (d/c)^{1/2}$. Now, inserting (21) and (38) in (4) we have

$$\| e \| \leq \| \rho \| + \| \theta \| \leq C_{10} h^2 \|u''\|, \tag{38}$$

where $C_{10} = (1 + \hat{d})C_1$.

4. Conclusions

We studied the error analysis of the finite element solution of generic scalar linear second-order ordinary BVPs in 1D. Continuous Galerkin finite element method (CGFEM) with piecewise linear polynomials are used for the space discretisation. Optimal order a priori error bounds are obtained using Ritz projection and standard tools in the H_0^1 (energy) and L_2 norms.

5. References

[1] P. Antonietti, N. Santo, I. Mazzieri, and A. Quarteroni, A high-order discontinuous Galerkin approximation to ordinary differential equations with applications to elastodynamics, MOX, Dipartimento di Matematica Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy), (2016).

[2] D. Estep, A posteriori error bounds and global error control for approximation of ordinary differential equations, SIAM, 32 (1995), pp. 1-48.

[3] D. Estep and A. Stuart, The dynamical behavior of the discontinuous Galerkin method and related difference schemes, AMS, 71 (1995), pp. 1073-1103.

[4] B. Holm and T. Wihler, Continuous and discontinuous Galerkin time stepping methods for nonlinear initial value problems with application to finite time blow-up, arXiv:1407.5520v4, (2016).

[5] B. Janseen and T. Wihler, Existence results for the continuous and discontinuous Galerkin time stepping methods for nonlinear initial value problems, arXiv:1407.5520v1, (2014).

[6] G. Matthies and F. Schieweck, Higher order variational time discretizations for nonlinear systems of ordinary differential equations, Preprint Otto von Guericke Universitat Magdeburg, (2011), p. 130.

[7] F. Schieweck, A stable discontinuous Galerkin-Petrov time discretization of higher order, J. Numer. Math., 18 (2010), pp. 429-456.

[8] D. Schötzau and C. Schwab, Time discretization of parabolic problems by the hp-version of the discontinuous Galerkin finite element method, SIAM J. Numer. Anal., 38 (2000), pp. 837-875.

[9] T. Wihler., An a-priori error analysis of the hp-version of the continuous Galerkin fem for nonlinear initial value problems, J. Sci. Comput., 25 (2005), p. 523? 549.

[10] S. Zhao and G. Wei, A unified discontinuous Galerkin framework for time integration, Math. Method. Appl. Sci., 37 (2014), pp. 1042-1071.