Fully Discrete Orthogonal Collocation Method of Sobolev Equations

Ning Ma, Wenliang Bian, Xiaofei Lu

College of Science, China University of Petroleum, Beijing, China
Email: bjmaning@163.com

Abstract

In this paper, the fully discrete orthogonal collocation method for Sobolev equations is considered, and the equivalence for discrete Galerkin method is proved. An optimal order error estimate is obtained.

Keywords

Sobolev Equations, Orthogonal Collocation Method, Error Estimate

1. Introduction

Sobolev equations are a class of mathematical physics equations, which are widely used in engineering field. Many numerical methods have been proposed, such as the characteristic difference method [1], the $H^1$-Galerkin Finite Element Method [2], the mixed finite element [3] and so on. The collocation method now is widely used in many fields including engineering technology and computational mathematics. Many applications have been proved effectively, e.g. the heat conduction equation [4], stochastic PDEs [5] and reaction diffusion equation [6]. The collocation method has high convergence order and does not need to calculate numerical integration so that the calculation is simple. So now we consider the application of fully discrete collocation method for Sobolev equations. We consider the linear Sobolev equations as follows:

\[
\begin{align*}
  u_t &= \nabla \left( a \nabla u_t + b \nabla u \right) + f(x,y,t), (x,y) \in \Omega, t \in (0,T], \\
  u|_{\partial \Omega} &= 0, (x,y) \in \partial \Omega, t \in [0,T], \\
  u|_{t=0} &= u_0(x,y), (x,y) \in \Omega.
\end{align*}
\]

(1)

In the equations, $u_t$ is the time derivative of $u$, and $\nabla u$ is the gradient of $u$. $\Omega = [0,1] \times [0,1]$, $\partial \Omega$ is the border of $\Omega$. $a = a(x,y,t)$ and $b = b(x,y,t)$ are known bounded differentiable functions.
2. Fully-Discrete Collocation Method

First, time is divided into $n$ equal parts. Let $\Delta t = \frac{T}{n}$ be the time step. Then we introduce the following notations:

$$t_n = n\Delta t, u^n = u(t_n), \partial_t u^n = \frac{u^n - u^{n-1}}{\Delta t}, \nabla u^n = \nabla u(t_n), \partial_t \nabla u^n = \frac{\nabla u^n - \nabla u^{n-1}}{\Delta t}.$$

Then we discrete the spatial region $\Omega$ into grids by points $(x_i, y_j), i = 0, 1, 2, \ldots, M, j = 0, 1, 2, \ldots, N$ and let $x_i, y_j$ are satisfied $0 = x_0 < x_1 < \cdots < x_M = 1, 0 = y_0 < y_1 < \cdots < y_N = 1$. Let $[7]$ $\Omega_{ij} = [x_{ij-1}, x_{ij}] \times [y_{ij-1}, y_{ij}], h_{ij} = x_i - x_{i-1}, h_{ij} = y_j - y_{j-1}, h = \max \{h_{ij}, h_{ij}\}$, $H_3 = \{ v = v(x, y) \in C^3(\Omega) | v|_{\Omega_{ij}} \text{ is a Bi-cubic Hermite polynomial} \}$, $H_3^0 = \{ v = v(x, y) \in H_3 | v|_{\partial \Omega} = 0 \}$.

The four Gauss points $(x_d, y_d); k,l = 1,2$ in $\Omega_{ij}$ are collocation points as follows: $x_d = x_i + h_2 \theta_1, y_d = y_j + h_2 \theta_2, k,l = 1,2$, where $\theta_1 = \frac{3 - \sqrt{3}}{6}, \theta_2 = \frac{3 + \sqrt{3}}{6}$. Then the intermediate variable $q = a\nabla u + b\nabla u$ is introduced so that the orthogonal collocation scheme as follows can be established. Seeking $(U, Q): [0, T] \to H_3 \times H_3$, such that

$$\begin{align*}
  \partial_t U^n - \nabla Q^n - f^n &\big|_{(x_d, y_d)} = 0, \\
  Q^n - (a^n \partial_t \nabla U^n + b^n \nabla U^n) &\big|_{(x_d, y_d)} = 0, \\
  [U]|_{\partial \Omega} &\big|_{t} = \begin{cases} U \big|_{\partial \Omega} & t \in [0, T], \\
  U \big|_{t=0} & u_0(x, y). \end{cases}
\end{align*}$$

(2)

Now we set the following notations $[4]$:

$$\begin{align*}
  \langle u, v \rangle_{ij} &= \sum_{i=1}^{M} \sum_{j=1}^{N} h_{ij}^2 \sum_{k,l=1}^{2} uv(x_{ij}, y_{ij}), \\
  \langle u, v \rangle_{i} &= \sum_{i=1}^{M} h_{i}^2 \sum_{k=1}^{2} uv(x_{i}, y_{i}), \\
  \langle u, v \rangle_{j} &= \sum_{j=1}^{N} h_{j}^2 \sum_{k=1}^{2} uv(x_{i}, y_{j}), \\
  \| u \|^2 &= \langle u, u \rangle, \| u \|^2 = \langle u, u \rangle.
\end{align*}$$

(3)

Next, we are going to prove existence and uniqueness of collocation solution and obtain the error estimate.

3. Discrete Galerkin Method

Consider the following discrete Galerkin scheme

$$\begin{align*}
  \langle \partial_t U^n, z_1 \rangle + \langle \nabla Q^n, \nabla z_1 \rangle - \langle f^n, z_1 \rangle &\big|_{t} = 0, \quad z_1 \in H_3^0, \\
  \langle Q^n, z_2 \rangle - (a^n \partial_t \nabla U^n + b^n \nabla U^n, z_2) &\big|_{t} = 0, \quad z_2 \in H_3^0.
\end{align*}$$

(4)

Theorem 3.1: The solutions of (4) and (2) are equivalent, existent and unique.
Proof: From the Equation (3), it is clear that the solution of (2) must be the solution of (4).

Let \( \{ \xi_i : l = 1, 2, \cdots, 4MN \} = \{(x_{il}, y_{il}), i = 1, \cdots, M, j = 1, \cdots, N, k, l = 1, 2\} \), \( \{Z_j\}_{4MN} \) be a group base of \( H^1_0 \). Thereupon \( \forall u^n(x, y) \in H^1_0 \) can be expressed as \( U^n(x, y) = \sum_{l=1}^{4MN} \beta^n_l Z_l(x, y) \). So (2) and (4) can be written in the form as follows

\[
F \beta^n + G \beta^{n-1} = R, C \beta^n + D \beta^{n-1} = S,
\]

\[
F = (F_{ij})_{4MN \times 4MN}, F_{ij} = Z_j(\xi_i) - a^n \Delta Z_j(\xi_i) - b^n \Delta t \Delta Z_j(\xi_i),
\]

\[
C = (C_{ij})_{4MN \times 4MN}, C_{ij} = \left( Z_j(\xi_i), Z_i \right) - a^n \left( \Delta Z_j(\xi_i), Z_i \right) - b^n \Delta t \left( \Delta Z_j(\xi_i), Z_i \right),
\]

where \( G, D \) are both matrices of \( 4MN \times 4MN \) and \( R, S \) are both vectors of \( 4MN \). Obviously the solution of equation \( F \tau = 0 \) must be satisfied the equation \( C \tau = 0 \), when \( \tau \) is a vectors of \( 4MN \). So \( F \) is nonsingular when \( C \) is nonsingular. Then the solutions of (2) and (4) are unique. To get the existence and uniqueness, we just need to prove \( A(\xi_i)_{4MN \times 4MN} \) where \( A_{ij} = \left( Z_j(\xi_i) - a \Delta Z_j(\xi_i), Z_i \right) \) is nonsingular when \( \Delta t \) is sufficiently small. And the nonsingularity of \( A \) has been proved [8] in. Thus the theorem is proved.

Next we will need to analyse the error estimate of (4).

4. Error Estimate

Define interpolation operators \( (P_1, P_2) \) which satisfied the following conditions

\[
W = P_U v = W - U, \eta = u - W, V = P_q w = V - Q, \xi = q - V,
\]

\[
\left\{ \begin{aligned}
\nabla (q^n - V^n), z \end{aligned} \right\} = 0, \forall z \in H^1_3,
\]

\[
\left\{ a^n \nabla (u^n - W^n) + b^n \nabla (u^n - W^n), \nabla z \right\} = 0, \forall z \in H^1_3,
\]

i.e., \( u - U = v + \eta, q - Q = w + \xi \). Now we can get the error equations

\[
\left\{ \begin{aligned}
(\eta_i, z_i) + \left( \partial_i v^n, z_i \right) + \left( r^n, z_i \right) + \left( \xi^n + w^n, \nabla z_i \right) = 0, & z_i \in H^1_3, \\
(\xi^n + w^n, z_i) - \left( a^n (\nabla r^n + \partial_i \nabla v^n), z_i \right) - \left( b^n \nabla v^n, z_i \right) = 0, & z_i \in H^1_3.
\end{aligned} \right.
\]

where \( r^n = W^n - \partial_i W^n, \nabla r^n = \nabla W^n - \partial_i \nabla W^n \). Then there is the theorem as follows.

**Theorem 4.1:** If \( u(x, y) \) is the accurate solution of (1), \( U(x, y) \) is the solution of the orthogonal collocation method, and \( u(x, y) \) satisfies the condition [4] [7] \( u \in L^\infty(0, T; H^4(\Omega)) \cap L^2(0, T; H^6(\Omega)) \), \( u_i \in L^\infty(0, T; H^6(\Omega)) \), then there is the error estimate as follows

\[
\left\| u^n - U^n \right\| \leq O(h^4 + \Delta t), \left\| q^n - Q^n \right\| \leq O(h^4 + \Delta t).
\]

Proof: First, it is clearly for \( r^n, \nabla r^n \) that

\[
\left\| r^n \right\| \leq C \Delta t \int_{t_{n-1}}^{t_n} \left\| W^n \right\| ds, \left\| \nabla r^n \right\| \leq C \Delta t \int_{t_{n-1}}^{t_n} \left\| \nabla W^n \right\| ds.
\]

Then let \( z_1 = v^n, z_2 = \nabla v^n \) in (5), the equations
\[
\begin{align*}
\left\langle \eta^e, v^e \right\rangle + & \left\langle \partial_t v^e, v^e \right\rangle + \left\langle r^e, v^e \right\rangle + \left\langle \xi^e + w^e, \nabla v^e \right\rangle = 0, \\
\left\langle \xi^e + w^e, \nabla v^e \right\rangle - & \left\langle a^e \left( \nabla r^e + \partial_t \nabla v^e \right), \nabla v^e \right\rangle - \left\langle b^e \nabla v^e, \nabla v^e \right\rangle = 0
\end{align*}
\]

can be got. It is easily calculated to see that
\[
a^e \left\langle \partial_t \nabla v^e, \nabla v^e \right\rangle + \left\langle \partial_t v^e, v^e \right\rangle = -a^e \left\langle \nabla r^e, \nabla v^e \right\rangle - b^e \left\langle \nabla v^e, \nabla v^e \right\rangle - \left\langle r^e, v^e \right\rangle - \left\langle \eta^e, v^e \right\rangle.
\]

Then through the Cauchy inequality, \( \varepsilon \)-inequality and \( \left\langle \partial_t v^e, v^e \right\rangle \geq \frac{1}{2} \partial_t \| v^e \|^2 \), and the functions \( a \) and \( b \) are bounded, it leads to the inequality
\[
\frac{1}{2} \partial_t \| v^e \|^2 + \frac{1}{2} \partial_t \| v^e \|^2 \\
\leq K \left\| \nabla r^e \right\|^2 + \varepsilon \left\| \nabla v^e \right\|^2 + \left\| \nabla v^e \right\|^2 + K \left\| v^e \right\|^2 \\
+ \varepsilon \left\| v^e \right\|^2 + K \left\| \eta^e \right\|^2 + \varepsilon \left\| v^e \right\|^2.
\]

The coefficients \( K, C \) both have nothing to do with \( h, \Delta t \) in the upper equation and following proof. Add the inequality (6) and make summation to the series sum from \( n = 1 \) to \( n \) and multiply \( \Delta t \). Then
\[
\left\| \nabla v^e \right\|^2 + \left\| v^e \right\|^2 \\
\leq K \Delta t \sum_{i=0}^{n-1} \left( \left\| \eta^e \right\|^2 + \left\| \nabla v^i \right\|^2 + \left\| v^i \right\|^2 \right) + K \Delta t^2 \int_0^t \left( \left\| \nabla W_n \right\|^2 + \left\| W_n \right\|^2 \right) ds
\]
is obtained. So it follows from discrete Gronwall lemma that
\[
\left\| \nabla v^e \right\|^2 + \left\| v^e \right\|^2 \leq K \Delta t \sum_{i=0}^{n-1} \left( \left\| \eta^e \right\|^2 \right) + K \Delta t^2 \int_0^t \left( \left\| \nabla W_n \right\|^2 + \left\| W_n \right\|^2 \right) ds \quad (7)
\]
if \( \Delta t \) is small enough.

Second, let \( z_1 = \partial_t v^e, z_2 = \partial_t \nabla v^e \) in (5), the equations
\[
\left\{ \begin{align*}
\left\langle \eta^e, \partial_t v^e \right\rangle + & \left\langle \partial_t v^e, \partial_t v^e \right\rangle + \left\langle r^e, \partial_t v^e \right\rangle + \left\langle \xi^e + w^e, \partial_t \nabla v^e \right\rangle = 0, \\
\left\langle \xi^e + w^e, \partial_t \nabla v^e \right\rangle - & \left\langle a^e \left( \nabla r^e + \partial_t \nabla v^e \right), \partial_t \nabla v^e \right\rangle - \left\langle b^e \nabla v^e, \partial_t \nabla v^e \right\rangle = 0,
\end{align*} \right.
\]
can be got. It is easy to get
\[
\left\langle a^e \left( \partial_t \nabla v^e, \partial_t \nabla v^e \right), \partial_t v^e, v^e \right\rangle \\
= -a^e \left\langle \nabla r^e, \partial_t \nabla v^e \right\rangle - b^e \left\langle \nabla v^e, \partial_t \nabla v^e \right\rangle - \left\langle r^e, \partial_t v^e \right\rangle - \left\langle \eta^e, \partial_t v^e \right\rangle.
\]

Then through Cauchy inequality and \( \varepsilon \)-inequality, (6) and (7) it leads to the inequality
\[
\left\| \partial_t \nabla v^e \right\|^2 + \left\| \partial_t v^e \right\|^2 \\
\leq K \left\| \eta^e \right\|^2 + K \Delta t \sum_{i=0}^{n-1} \left\| \eta^e \right\|^2 + K \Delta t^2 \int_0^t \left( \left\| \nabla W_n \right\|^2 + \left\| W_n \right\|^2 \right) ds,
\]
if \( \Delta t \) is sufficiently small.

At last, let \( z_2 = w^e \) in the second equation of (5), it can be expressed as
\[
\left\langle \xi^e + w^e, w^e \right\rangle - \left\langle a^e \left( \nabla r^e + \partial_t \nabla v^e \right), w^e \right\rangle - \left\langle b^e \nabla v^e, w^e \right\rangle = 0. \quad (7) \text{and (8) implies that}
\]

DOI: 10.4236/jamp.2017.512192
\[
\|w^h\|^2 \leq K \|\varepsilon^h\|^2 + K \|\eta^h\|^2 + K \Delta t^2 \sum_{i=1}^n \left( \int_0^1 \left( \|\nabla W_{a_i}\|^2 + \|W_{a_i}\|^2 \right) ds \right)
\]

(9)

The results
\[
\|\eta\| \leq Ch^4 \left( \sum_{i,j} \|q^{(ij)}\|^2 \right)^{\frac{1}{2}}, \quad \|\eta\| \leq Ch^4 \left( \sum_{i,j} \|q^{(ij)}\|^2 \right)^{\frac{1}{2}},
\]

\[
\|\xi\| \leq Ch^4 \left( \sum_{i,j} \|q^{(ij)}\|^2 \right)^{\frac{1}{2}}, \quad \|\xi\| \leq Ch^4 \left( \sum_{i,j} \|q^{(ij)}\|^2 \right)^{\frac{1}{2}},
\]

(10)

can be obtained from lemma 1.6 in [4], where \(u\) is sufficiently smooth (\(C\) is a positive constant). Moreover (3) in [7] implies that \(\forall f \in H, \|f\| \leq C \|f\|\) is valid. So it follows from (7), (9) and (10) that
\[
\|w^h - U^h\|^2 \leq K \|\varepsilon^h\|^2 + K \|\eta^h\|^2 + K \Delta t^2 \int_0^1 \left( \|\nabla W_{a_i}\|^2 + \|W_{a_i}\|^2 \right) ds \leq K_1 h^8 + K_2 \Delta t^2,
\]
\[
\|q^h - Q^h\|^2 \leq K \|\varepsilon^h\|^2 + K \|\eta^h\|^2 + K \Delta t^2 \int_0^1 \left( \|\nabla W_{a_i}\|^2 + \|W_{a_i}\|^2 \right) ds \leq K_1 h^8 + K_2 \Delta t^2,
\]

where \(K_1\) and \(K_2\) are constants which have nothing to do with \(h\) and \(\eta\). Thus the theorem is proved.

Acknowledgements

Sincere thanks to the Basic Subjects Fund and Science Foundations of China University of Petroleum (Beijing) (NO. 2462015YQ0604, NO. 2462015QZDX02).

References