

ANNOUNCEMENT ON “SHARP ERROR ESTIMATE OF BDF2 SCHEME WITH VARIABLE TIME STEPS FOR LINEAR REACTION-DIFFUSION EQUATIONS”

ZHANG Ji-wei ¹, ZHAO Cheng-chao ²

(1. School of Mathematics and Statistics, Hubei Key Laboratory of Computational Science,
Wuhan University, Wuhan 430072, China)

(2. Beijing Computational Science Research Center, Beijing 100193, China)

Abstract: In this note we announce the sharp error estimate of BDF2 scheme for linear diffusion reaction problem with variable time steps. Our analysis shows that the optimal second-order convergence does not require the high-order methods or the very small time steps $\tau_1 = \mathcal{O}(\tau^2)$ for the first level solution u^1 . This is, the first-order consistence of the first level solution u^1 like BDF1 (i.e. Euler scheme) as a starting point does not cause the loss of global temporal accuracy, and the ratios are updated to $r_k \leq 4.8645$.

Keywords: BDF2; DOC; DCC; variable time-steps; sharp error estimate

In this note, we revisit the two-step backward differentiation formula (BDF2) with variable time-steps for solving the following reaction-diffusion equation:

$$\begin{aligned} \partial_t u &= \Delta u + \kappa u + f(x, t), & x \in \Omega, t \in (0, T], \\ u(x, 0) &= u_0(x), & x \in \bar{\Omega}, \\ u(x, t) &= 0, & x \in \partial\Omega, t \in [0, T], \end{aligned} \quad (1)$$

where the reaction coefficient $\kappa \in \mathbb{R}$, and Ω is a bounded domain.

Set the generally nonuniform time levels $0 = t_0 < t_1 < t_2 < \cdots < t_N = T$ with the k th time-step size $\tau_k := t_k - t_{k-1}$, the maximum step size $\tau := \max_{1 \leq k \leq N} \tau_k$, and the adjacent time-step ratio

$$r_k = \frac{\tau_k}{\tau_{k-1}}, \quad 2 \leq k \leq N.$$

The BDF2 scheme with variable time-step is given as

$$\mathcal{D}_2 u^n = \Delta u^n + \kappa u^n + f^n, \quad \text{for } 1 \leq n \leq N, \quad (2)$$

* Received date: 2020-10-12

Accepted date: 2020-11-23

Foundation item: Supported by Natural Science Foundation of Hubei Province(2019CFA007); Supported by NSFC (11771035).

Biography: Zhang Jiwei (1979–), male, professor, major in the computational and applied mathematics.

Corresponding author: Zhao Chengchao.

where the BDF2 formula can be unified to the following discrete convolution form

$$\mathcal{D}_2 u^n := \sum_{k=1}^n b_{n-k}^{(n)} \nabla_\tau u^k, \quad n \geq 1, \quad (3)$$

which is simulated by the discrete fractional derivatives [1–4]. Here the discrete convolution kernels $b_{n-k}^{(n)}$ are given as $b_0^{(1)} := 1/\tau_1$ and for $n > 1$,

$$b_0^{(n)} = \frac{1 + 2r_n}{\tau_n(1 + r_n)}, \quad b_1^{(n)} = -\frac{r_n^2}{\tau_n(1 + r_n)} \quad \text{and} \quad b_j^{(n)} = 0 \quad \text{for } 2 \leq j \leq n-1. \quad (4)$$

For the problem (1) with $\kappa = 0$, twenty years ago Becker [5] (one also refers to Thomée’s classical book [6, Lemma 10.6]) presented that if the adjacent step ratios r_k satisfy $0 < r_k \leq (2 + \sqrt{13})/3 \approx 1.868$, the discrete solution can be bounded by

$$\|u^n\| \leq C \exp(C\Gamma_n) \left(\|u_0\| + \sum_{j=1}^n \tau_j \|f^j\| \right) \quad \text{for } n \geq 1,$$

where C is a positive constant and $\Gamma_n := \sum_{k=2}^{n-2} \max\{0, r_k - r_{k+2}\}$. As pointed out in [6] and [7], the magnitudes of Γ_n can be zero, bounded [6, pp. 175] and unbounded [7, Remark 4.1] by selecting certain step-ratio sequence and vanishing step sizes. After that, Emmrich [8] improves the Becker’s constrained condition to $0 \leq r_k \leq 1.91$, but still remains the undesirable factor $\exp(C\Gamma_n)$ in the L_2 -norm stability. Chen *et al.* circumvent the factor $\exp(C\Gamma_n)$ in Becker’s estimate with a bounded factor $\exp(Ct_n)$ with $0 \leq r_k \leq 1.53$, but lack the estimate in the ideal case of $\Gamma_n = 0$. Recently, by using the technique of the discrete orthogonal convolution (DOC) kernels, a nice and interesting work [9] obtains the convergence

$$\begin{aligned} & \|u(t_n) - u^n\| \\ & \leq 2 \exp(4\kappa t_{n-1}) \left(\|u_0 - u^0\| + 2t_n \int_0^{t_1} \|\partial_{tt} u\| dt + 3t_n \max_{1 \leq k \leq n} \tau_k \int_{t_{k-1}}^{t_k} \|\partial_{ttt} u\| dt \right) \end{aligned} \quad (5)$$

with $0 \leq r_k \leq 3.561$. Here the DOC kernels are defined by

$$\sum_{j=k}^n \theta_{n-j}^{(n)} b_{j-k}^{(j)} = \delta_{nk}, \quad \text{for all } 1 \leq k \leq n, \quad (6)$$

where δ_{nk} represents the Kronecker delta symbol with $\delta_{nk} = 1$ if $n = k$ and $\delta_{nk} = 0$ if $n \neq k$. One can see that the right-hand-side second term is the first-order convergence when t_n is large. If the second-order convergence is obtained, it suffers from a restriction condition $|\mathfrak{R}_p| \leq N_0 \ll N$ with the index set defined by

$$\mathfrak{R}_p = \left\{ k \mid 1 + \sqrt{2} \leq r_k \leq (3 + \sqrt{17})/2 \right\}. \quad (7)$$

In this note, by introducing the novel conception of the discrete complementary convolution (DCC) kernels, we achieve the sharp second-order convergence for BDF2 scheme and update the adjacent time-step restriction condition to

$$\mathbf{A1}: 0 < r_k \leq r_{\max} = \frac{1}{6} \left(\sqrt[3]{1196 - 12\sqrt{177}} + \sqrt[3]{1196 + 12\sqrt{177}} \right) + \frac{4}{3} \approx 4.8645, \text{ for } 2 \leq k \leq N,$$

which does not suffer from the restriction condition $|\mathfrak{R}_p| \leq N_0 \ll N$. Here $r_{\max}$ is the positive real solution of $x^3 = (2x + 1)^2$.

To obtain stability analysis of the discrete scheme (2), it is natural to find a DCC kernel $p_{n-j}^{(n)}$ such that

$$\sum_{j=1}^n p_{n-j}^{(n)} \mathcal{D}_2 u^j = \sum_{j=1}^n p_{n-j}^{(n)} \sum_{l=1}^j b_{j-l}^{(j)} \nabla_\tau u^l = \sum_{l=1}^n \nabla_\tau u^l \sum_{j=l}^n p_{n-j}^{(n)} b_{j-l}^{(j)} = u^n - u^0, \quad \forall n \geq 1. \quad (8)$$

One can see that if the identity (8) holds for all $n \geq 1$, it only requires

$$\sum_{j=k}^n p_{n-j}^{(n)} b_{j-k}^{(j)} \equiv 1, \quad \forall 1 \leq k \leq n, 1 \leq n \leq N. \quad (9)$$

The first main contribution in our paper is establishing the positive semi-definiteness of BDF2 convolution kernels, which produces the constrain condition **A1** on the adjacent time-step ratios.

Lemma 0.1 Assume the time step ratios r_k satisfy **A1**. For any real sequence $\{w_k\}_{k=1}^n$, it holds that

$$2w_k \sum_{j=1}^k b_{k-j}^{(k)} w_j \geq \frac{r_{k+1}}{(1+r_{k+1})} \frac{w_k^2}{\epsilon_* \tau_k} - \frac{r_k}{(1+r_k)} \frac{w_{k-1}^2}{\epsilon_* \tau_{k-1}}, \quad k \geq 2, \quad (10)$$

$$2 \sum_{k=1}^n w_k \sum_{j=1}^k b_{k-j}^{(k)} w_j \geq 0, \quad \text{for } n \geq 1, \quad (11)$$

where the positive constant $\epsilon_* = \frac{1}{6} \sqrt[3]{12} (\sqrt[3]{\sqrt{177} + 9} - \sqrt[3]{\sqrt{177} - 9}) \approx 0.4534$ is taken.

The positive semi-definiteness of $\theta_{n-k}^{(n)}$ is derived by the positive semi-definiteness of $b_{n-k}^{(n)}$.

Lemma 0.2 If the BDF2 kernels $b_{n-k}^{(n)}$ defined in (4) are positive semi-definite, then the DOC kernels $\theta_{n-k}^{(n)}$ defined in (6) are also positive semi-definite. This is, it holds for any real sequence $\{\omega_j\}_{j=1}^n$ that

$$\sum_{k=1}^n \omega_k \sum_{j=1}^k \theta_{k-j}^{(k)} \omega_j \geq 0, \quad \forall n \geq 1.$$

A immediate product of the semi-positive definiteness of the BDF2 kernels is the following energy stability for BDF2 scheme (2) (one also refers to [9]).

Theorem 0.1 Assume the condition **A1** holds and $\kappa \leq 0$, then the discrete solution u^n to the BDF2 scheme (2) with variable time steps satisfies

$$\partial_\tau E^k \leq 2 \langle f^k, \partial_\tau u^k \rangle, \quad \forall k \geq 1. \quad (12)$$

Furthermore, the energy has the following estimate:

$$\sqrt{E^n} \leq \sqrt{E^0} + 4C_\Omega \left(\sum_{k=1}^n \|\nabla_\tau f^k\| + \|f^0\| \right), \quad \forall n \geq 1. \quad (13)$$

Here the (modified) discrete energy E^k is defined by

$$E^k := \frac{r_{k+1}\tau_k}{\epsilon^*(1+r_{k+1})} \|\partial_\tau u^k\|^2 + |u^k|_1^2 - \kappa \|u^k\|^2, \quad \text{for } \kappa \leq 0 \text{ and } k \geq 1, \quad (14)$$

where the initial energy $E^0 := |u^0|_1^2 - \kappa \|u^0\|^2$, $\partial_\tau u^k = \nabla_\tau u^k / \tau_k$ and $\epsilon_* = \frac{1}{6} \sqrt[3]{12} (\sqrt[3]{\sqrt{177} + 9} - \sqrt[3]{\sqrt{177} - 9}) \approx 0.4534$ and $\|\cdot\|$ represent the inner product and norm in the space $L^2(\Omega)$.

The second main contribution in our paper is that we achieve the sharp error estimate. To obtain the sharp error estimate, we first establish the stability and consistency results for the BDF2 scheme (2). Before that, several useful properties of DOC and DCC kernels are presented.

Lemma 0.3 [9] The DOC kernels $\theta_{n-j}^{(n)}$ have the following properties:

$$\sum_{j=1}^n \theta_{n-j}^{(n)} = \tau_n, \quad \text{for } n \geq 1, \quad (15)$$

$$\sum_{k=1}^n \sum_{j=1}^k \theta_{k-j}^{(k)} = t_n, \quad \text{for } n \geq 1. \quad (16)$$

Lemma 0.4 The DOC kernel $\theta_{n-j}^{(n)}$ can be explicitly represented by

$$\theta_{n-k}^{(n)} = \begin{cases} \frac{\tau_n(1+r_k)}{1+2r_k} \prod_{i=k+1}^n \frac{r_i}{1+2r_i}, & \text{for } 2 \leq k \leq n. \\ \tau_n \prod_{i=k+1}^n \frac{r_i}{1+2r_i}, & \text{for } k = 1. \end{cases} \quad (17)$$

We remark that the results in Lemma 0.4 is very important for the following convergence analysis and the bound of DCC kernels. Here we emphasize on the connection of DCC and DOC kernels, and present the properties of DCC kernels by propositions as follows.

Proposition 0.1 Let the DCC kernels $p_{n-j}^{(n)}$ and the DOC kernels $\theta_{n-j}^{(n)}$ be defined by (9) and (6) respectively. Then, the DCC kernel and DOC kernel hold

$$p_{n-j}^{(n)} = \sum_{l=j}^n \theta_{l-j}^{(l)}, \quad \forall 1 \leq j \leq n, \quad (18)$$

$$\theta_{n-j}^{(n)} = p_{n-j}^{(n)} - p_{n-1-j}^{(n-1)}, \quad \forall 1 \leq j \leq n, \quad (19)$$

where $p_{-1}^{(n)} := 0, \forall n \geq 0$ is defined.

Proposition 0.2 Let τ be the maximum time-step size and the time-step ratios satisfy $0 < r_k \leq r_*$, where r_* is any given positive constant. The DCC kernels $p_{n-k}^{(n)}$ defined in (9)

satisfy

$$p_{n-j}^{(n)} = \sum_{k=j}^n \frac{\tau_k(1+r_k)}{1+2r_k} \prod_{i=j+1}^k \frac{r_i}{1+2r_i}, \quad 2 \leq j \leq n, \quad (20)$$

$$p_{n-1}^{(n)} = \sum_{k=1}^n \tau_k \prod_{i=2}^k \frac{r_i}{1+2r_i}, \quad (21)$$

$$\sum_{j=1}^n p_{n-j}^{(n)} = t_n, \quad (22)$$

$$p_{n-j}^{(n)} \leq \sum_{k=j}^n \tau_k \left(\frac{r_*}{1+2r_*} \right)^{k-j} \leq \sum_{k=j}^n \frac{\tau_k}{2^{k-j}} \leq 2\tau, \quad (23)$$

where $\prod_{i=j+1}^k = 1$ for $j \geq k$ is defined.

To obtain the stability of the BDF2 scheme (2), we introduce a discrete Grönwall inequality for the following L^2 -norm estimate.

Lemma 0.5 Assume $\lambda > 0$ and the sequences $\{v_j\}_{j=1}^N$ and $\{\eta_j\}_{j=0}^N$ are nonnegative. If

$$v_n \leq \lambda \sum_{j=1}^{n-1} \tau_j v_j + \sum_{j=0}^n \eta_j, \quad \text{for } 1 \leq n \leq N,$$

then it holds

$$v_n \leq \exp(\lambda t_{n-1}) \sum_{j=0}^n \eta_j, \quad \text{for } 1 \leq n \leq N.$$

We now present the stability result of the BDF2 scheme (2).

Theorem 0.2 If the BDF2 kernels $b_{n-j}^{(n)}$ defined in (4) are positive semi-definite (or condition **A1** holds), the discrete solution u^n of the BDF2 scheme (2) is unconditionally stable in the L^2 -norm. If $\kappa > 0$ and the maximum time-step size $\tau \leq 1/(4\kappa)$, it holds

$$\|u^n\| \leq 2 \exp(4\kappa t_{n-1}) \left(\|u^0\| + 2 \sum_{j=1}^n p_{n-j}^{(n)} \|f^j\| \right), \quad \text{for } 1 \leq n \leq N. \quad (24)$$

If $\kappa \leq 0$, it holds

$$\|u^n\| \leq \|u^0\| + 2 \sum_{j=1}^n p_{n-j}^{(n)} \|f^j\|, \quad \text{for } 1 \leq n \leq N. \quad (25)$$

Lemma 0.6 Set

$$\begin{aligned} G^l &:= -\frac{1}{2} \int_{t_{l-1}}^{t_l} (t - t_{l-1})^2 \partial_{ttt} u \, dt, \quad 1 \leq l \leq N, \\ R^j &:= -\frac{1}{2} b_1^{(j)} \tau_{j-1} \int_{t_{j-1}}^{t_j} (2(t - t_{j-1}) + \tau_{j-1}) \partial_{ttt} u \, dt, \quad 2 \leq j \leq N, \\ R^1 &:= \frac{1}{2\tau_1} \int_0^{t_1} t^2 \partial_{ttt} u \, dt - \frac{1}{\tau_1} \int_0^{t_1} s u_{tt}(s) \, ds. \end{aligned} \quad (26)$$

The truncation error $\eta^j := \mathcal{D}_2 u(t_j) - \partial_t u(t_j)$ ($1 \leq j \leq N$) can be expressed the following form

$$\eta^j = \sum_{l=1}^j b_{j-l}^{(j)} G^l + R^j, \quad 1 \leq j \leq N. \quad (27)$$

Theorem 0.3 Assume the conditions in Theorem 0.2 hold, and the truncated error can be expressed by

$$f^n := \sum_{k=1}^n b_{n-k}^{(n)} G^k + R^n, \quad (28)$$

where G^k and R^n are given in (26). If $\kappa > 0$ and the maximum time-step size $\tau \leq 1/(4\kappa)$, it holds

$$\|u^n\| \leq 2 \exp(4\kappa t_{n-1}) \left(\|u^0\| + 2 \sum_{k=1}^n \|G^k\| + 2 \sum_{k=1}^n p_{n-k}^{(n)} \|R^k\| \right), \quad \text{for } 1 \leq n \leq N. \quad (29)$$

If $\kappa \leq 0$, it holds

$$\|u^n\| \leq \|u^0\| + 2 \sum_{k=1}^n \|G^k\| + 2 \sum_{k=1}^n p_{n-k}^{(n)} \|R^k\|, \quad \text{for } 1 \leq n \leq N. \quad (30)$$

Finally, applying the Lemmas 0.6, Proposition 0.2 and Theorem 0.3, we achieve the sharp error estimate.

Theorem 0.4 Let $u(t, x)$ be the solution to problem (1). If the BDF2 kernels $b_{n-k}^{(n)}$ defined in (4) are positive semi-definite (or the condition **A1** holds), then the solution u^n to BDF2 scheme (2) is convergent in the L^2 -norm. If $\kappa > 0$ and the maximum time-step size $\tau < 1/(4\kappa)$, it holds

$$\begin{aligned} \|u(t_n) - u^n\| \leq & 2 \exp(4\kappa t_{n-1}) \left(\|u(0) - u^0\| + 2(\tau_1 + \tau) \int_0^{t_1} \|\partial_{tt} u\| dt + \sum_{k=1}^n \tau_k^2 \int_{t_{k-1}}^{t_k} \|\partial_{ttt} u\| dt \right. \\ & \left. + 2t_n \max_{1 \leq k \leq n} \tau_k \int_{t_{k-1}}^{t_k} \|\partial_{ttt} u\| dt \right), \quad \text{for } 1 \leq n \leq N. \end{aligned} \quad (31)$$

If $\kappa \leq 0$, it holds

$$\begin{aligned} \|u(t_n) - u^n\| \leq & \|u(0) - u^0\| + 2(\tau_1 + \tau) \int_0^{t_1} \|\partial_{tt} u\| dt + \sum_{k=1}^n \tau_k^2 \int_{t_{k-1}}^{t_k} \|\partial_{ttt} u\| dt \\ & + 2t_n \max_{1 \leq k \leq n} \tau_k \int_{t_{k-1}}^{t_k} \|\partial_{ttt} u\| dt, \quad \text{for } 1 \leq n \leq N. \end{aligned} \quad (32)$$

More details can be found in [10].

References

- [1] Li D, Liao H, Wang J, Sun W, Zhang J. Analysis of L1-Galerkin FEMs for time-fractional nonlinear parabolic problems[J]. *Commun. Comput. Phys.*, 2018, 24(1): 86–103.
- [2] Li D, Wang J, Zhang J. Unconditionally convergent L1-Galerkin FEMs for nonlinear time-fractional Schrödinger equations[J]. *SIAM J. Sci. Comp.*, 2017, 39: A3067–A3088.
- [3] Liao H, Li D, Zhang J. Sharp error estimate of the nonuniform L1 formula for reaction-subdiffusion equations[J]. *SIAM J. Numer. Anal.*, 2018, 56(2): 1112–1133.
- [4] Liao H, McLean W, Zhang J. A discrete Grönwall inequality with application to numerical schemes for fractional reaction-subdiffusion problems[J]. *SIAM J. Numer. Anal.*, 2019, 57(1): 218–237.
- [5] Becker J. A second order backward difference method with variable steps for a parabolic problem[J]. *BIT*, 1998, 38(4): 644–662.
- [6] Thomée V. Galerkin finite element methods for parabolic problems, second edition[M]. Springer-Verlag, 2006.
- [7] Chen W, Wang X, Yan Y, Zhang Z. A second order BDF numerical scheme with variable steps for the Cahn-Hilliard equation[J]. *SIAM J. Numer. Anal.*, 2019, 57(1): 495–525.
- [8] Emmrich E. Stability and error of the variable two-step BDF for semilinear parabolic problems[J]. *J. Appl. Math. Comput.*, 2005, 19: 33–55.
- [9] Liao H, Zhang Z M. Analysis of adaptive BDF2 scheme for diffusion equations[J]. *Math. Comput.*, 2020.
- [10] Zhang J, Zhao C. Sharp error estimate of BDF2 scheme with variable time steps for linear reaction-diffusion equations, submitted.