

Maximum time step for high order BDF methods applied to gradient flows

Morgan Pierre

▶ To cite this version:

Morgan Pierre. Maximum time step for high order BDF methods applied to gradient flows. 2021. hal-03438159

HAL Id: hal-03438159 https://hal.science/hal-03438159

Preprint submitted on 21 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MAXIMUM TIME STEP FOR HIGH ORDER BDF METHODS APPLIED TO GRADIENT FLOWS

MORGAN PIERRE

Dedicated to Hélène

ABSTRACT. For a backward differentiation formula (BDF) applied to the gradient flow of a semiconvex function, quadratic stability implies gradient stability. Namely, it is possible to build a Lyapunov functional for the discrete-in-time dynamical system, with a restriction on the time step. The maximum time step which can be derived from quadratic stability has previously been obtained for the BDF1, BDF2 and BDF3 schemes. Here, we compute it for the BDF4 and BDF5 methods. We also prove that the BDF6 scheme is not quadratically stable. Our results are based on the tools developed by Dahlquist and other authors to show the equivalence of A-stability and G-stability.

Keywords: gradient flow, BDF method, semiconvex function, Lasalle's invariance principle, Lojasiewicz inequality, Allen-Cahn equation.

1. INTRODUCTION

Backward differentiation formulae (BDF) were first introduced by Curtiss and Hirschfelder for the numerical resolution of stiff ordinary differential equations [11]. They have been extensively studied (see [15, 16, 27] and references therein). In recent years, BDF methods have been proved to be very powerful for the time discretization of semilinear parabolic problems in various situations. In particular, the Nevanlinna-Odeh multipliers which date back to 1981 [23] have been used to obtain stability and errors estimates on finite time intervals in the works [2, 21].

BDF methods also have very nice properties regarding the asymptotic behaviour of solutions. This is well illustrated by the Allen-Cahn equation, which reads

$$\frac{\partial u}{\partial t} - \Delta u + f'(u) = 0, \quad x \in \Omega, \ t > 0, \tag{1.1}$$

where Ω is a bounded subset of \mathbb{R}^d with smooth boundary (d = 1, 2 or 3). The unknown function u is the order parameter and the function f' is the derivative of a double-well potential. A standard choice is the quartic potential

$$f(s) = \frac{1}{4\varepsilon^2} (s^2 - 1)^2 \quad (s \in \mathbb{R}),$$
(1.2)

where $\varepsilon > 0$ represents the typical thickness of the interface between two phases. The Allen-Cahn equation is usually endowed with homogeneous Neumann boundary conditions. It is the $L^2(\Omega)$ -gradient flow of the energy

$$\mathcal{E}(u) = \int_{\Omega} \frac{1}{2} |\nabla u|^2 + f(u) dx.$$
(1.3)

This can be used to prove that every solution to (1.1) with the quartic potential (1.2) converges to a steady state as t goes to infinity. This result is based on Lasalle's invariance principle and on a Lojasiewicz-Simon inequality, which involves the analycity of f [26]. We refer the reader to the review [17] for more details.

In [13], Elliott and Stuart proposed several time and space discretizations of (1.1) which preserve the gradient flow structure. In particular, they proved that for k = 1, 2 and 3, the k-step BDF method (BDFk) has this property, with a restriction on the time step. They managed to build a Lyapunov functional which is a modified version of (1.3) for the discrete-in-time dynamical system. As a consequence, every solution to the time discrete problem converges to a steady state: this result is based on arguments similar to the time continuous problem [4, 5, 7, 18, 22].

In [27], Stuart and Humphries generalized the result in [13] to the case of semiconvex functions. They also proposed to investigate the gradient stability of the k-step BDF method for k = 4, 5 and 6. In [7], the author and his collaborators generalized the construction in [27] and proved that the BDF4 and BDF5 are gradient stable, with a restriction on the time step. For this purpose, they introduced a notion of *quadratic stability* involving quadratic forms. Quadratic stability implies gradient stability and it is a definition easier to work with, because of its algebraic nature. Numerical simulations suggested that the BDF6 method is not quadratically stable. For $k \geq 7$, k-step BDF methods are not zero stable [15] and so they cannot be gradient stable [7, Remark 2.8].

In [25], the author computed the maximum time step which can be derived from quadratic stability for the BDF3 scheme applied to a gradient flow. The purpose was also to try and have a better understanding of the BDF6 scheme, but the method was incomplete to achieve this goal. In this paper, we manage to compute the maximum time step for the BDF4 and BDF5 schemes, and we prove that the BDF6 scheme is *not* quadratically stable. The idea is to apply to our situation the powerful tools developed by Dahlquist [12], Baiocchi and Crouzeix [6] (see also [16, Section V.6]) to show the equivalence of A-stability and G-stability.

The situation for quadratic stability is therefore very analogous to what happens with the Nevanlinna-Odeh multipliers, which are valid for BDF schemes up to order 5 only. However, for $k \ge 2$, there is a possible gap between quadratic stability and gradient stability [25, Section 3.5]. Therefore, there is still a slight chance to prove that the BDF6 is gradient stable. In a related manner, stability results on finite time intervals have recently been obtained in the work [3] for the BDF6 scheme applied to semilinear parabolic equations.

The manuscript is organized as follows. In Section 2, we compute a maximum value related to quadratic stability for the BDF4, BDF5 and BDF6 schemes, by using tools developped for Dahlquist's equivalence theorem. In Section 3, we show how quadratic stability implies gradient stability with a maximum time step for this approach. We also derive some consequences for the asymptotic behaviour of discretized gradient flows, including the non-autonomous case. Our analysis is made in the Euclidean space \mathbb{R}^N , but we stress that the tools developped here also apply to the infinite dimensional case, for PDEs such as the Allen-Cahn equation [7] or the Cahn-Hilliard equation [5, 8].

2. Optimal constants for quadratic stability

We consider the k-step BDF scheme applied to the gradient flow

$$y'(t) = -\nabla F(y(t)), \quad t \ge 0$$

in \mathbb{R}^N , where $F : \mathbb{R}^N \to \mathbb{R}$ is of class C^1 . The k-step BDF method with fixed time step $\Delta t > 0$ reads: let y_0, \ldots, y_{k-1} in \mathbb{R}^N and for $n = 0, 1, \ldots$, let y_{n+k} solve

$$\frac{1}{\Delta t} \sum_{j=1}^{k} \frac{1}{j} \partial^j y_{n+k} = -\nabla F(y_{n+k}), \qquad (2.1)$$

where $\partial y_{n+k} = y_{n+k} - y_{n+k-1}$ is the backward difference operator and by induction, $\partial^j y_{n+k} = \partial^{j-1}(\partial y_{n+k})$ for $j \ge 2$.

If F is a semiconvex function, in order to find a Lyapunov function for the discrete-in-time dynamical system (2.1), we take the inner product of (2.1) with ∂y_{n+k} (see (3.15)) and we are led to consider the quantity

$$\Gamma_k = \left\langle \sum_{j=1}^k \frac{1}{j} \partial^j y_{n+k}, \partial y_{n+k} \right\rangle, \qquad (2.2)$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in \mathbb{R}^N . By induction, it is easily seen that for $j \geq 1$, $\partial^j y_{n+k}$ is a linear combination of $\partial y_{n+1}, \ldots, \partial y_{n+k}$. Thus, it is convenient to see Γ_k as a quadratic form depending on the variables $\partial y_{n+1}, \ldots, \partial y_{n+k}$ in \mathbb{R}^N , namely

$$\Gamma_k = \left\langle \sum_{j=1}^k \frac{1}{j} \partial^{j-1} \partial y_{n+k}, \partial y_{n+k} \right\rangle.$$

We introduce the polynomial of degree k-1

$$\delta_k(z) = \sum_{j=1}^k \frac{1}{j} (1-z)^{j-1} = \sum_{i=0}^{k-1} \delta_i z^i, \qquad (2.3)$$

so that

$$\Gamma_k = \left\langle \sum_{i=0}^{k-1} \delta_i \partial y_{n+k-i}, \partial y_{n+k} \right\rangle \tag{2.4}$$

The polynomial δ_k is analogous to the "defining polynomial" used for the analysis of multistep schemes (see, e.g., [15, 27]), but we use here the variables $\partial y_{n+1}, \ldots, \partial y_{n+k}$ instead of y_n, \ldots, y_{n+k} . The fundamental idea is to apply Dahlquist's theorem, which is stated below, to the polynomials $\delta(z) = \delta_k(z) - \beta$ and $\mu(z) = 1$, where $\beta \in \mathbb{R}$ is optimal.

Lemma 2.1 ([6, 12] and Section V.6 in [16]). Let $\delta(z)$ and $\mu(z)$ be real polynomials of degree at most k - 1 (at least one of them of exact degree k - 1) that have no common divisor. If

$$\operatorname{Re}\frac{\delta(z)}{\mu(z)} > 0 \quad for \quad |z| < 1, \tag{2.5}$$

then there exist a symmetric positive definite matrix $G = (g_{ij}) \in \mathbb{R}^{k-1} \times \mathbb{R}^{k-1}$ and real numbers a_1, \ldots, a_k such that for all $v_1, \ldots, v_k \in \mathbb{R}^N$,

$$\left\langle \sum_{i=0}^{k-1} \delta_i v_{k-i}, \sum_{j=0}^{k-1} \mu_j v_{k-j} \right\rangle = \sum_{i,j=1}^{k-1} g_{ij} \langle v_{i+1}, v_{j+1} \rangle - \sum_{i,j=1}^{k-1} g_{ij} \langle v_i, v_j \rangle + \left\| \sum_{i=1}^k a_i v_i \right\|^2.$$
(2.6)

Hereafter, $\|\cdot\|$ denotes the Euclidean norm in \mathbb{R}^N . We have

$$\begin{split} \delta_2(z) &= \frac{3}{2} - \frac{1}{2}z, \\ \delta_3(z) &= \frac{11}{6} - \frac{7}{6}z + \frac{1}{3}z^2, \\ \delta_4(z) &= \frac{25}{12} - \frac{23}{12}z + \frac{13}{12}z^2 - \frac{1}{4}z^3, \\ \delta_5(z) &= \frac{137}{60} - \frac{163}{60}z + \frac{137}{60}z^2 - \frac{21}{20}z^3 + \frac{1}{5}z^4, \\ \delta_6(z) &= \frac{49}{20} - \frac{71}{20}z + \frac{79}{20}z^2 - \frac{163}{60}z^3 + \frac{31}{30}z^4 - \frac{1}{6}z^5. \end{split}$$

For $2 \le k \le 6$ we consider the region

$$D_k = \{\delta_k(z) : |z| < 1\}$$
(2.7)

which is drawn in pink in Figures 1-3 and we define

$$\beta_k = \inf \{ \operatorname{Re} \delta_k(z) : |z| < 1 \}.$$

By examing the curve $\theta \mapsto \delta_k(e^{i\theta})$ for $\theta \in [0, 2\pi]$ we find that

$$\beta_k = \min \left\{ \operatorname{Re} \delta_k(z) : |z| = 1 \right\}.$$

and we obtain

$$\beta_2 = 1, \ \beta_3 = \frac{95}{96}, \ \beta_4 = 0.8141 \dots \beta_5 = 0.1855 \dots \text{ and } \beta_6 = -1.37 \dots$$
 (2.8)

The exact value of β_4 is

$$\beta_4 = \frac{664}{729} - \frac{43\sqrt{43}}{2916}$$

and the exact value of β_5 is rather lengthy (see (2.10)). It is easy to check that $\beta_6 < 0$. For instance, we have

$$\beta_6 < \operatorname{Re}\left(\delta_6(e^{-2i\pi/3})\right) = -\frac{9}{10} < 0.$$
(2.9)

Example 2.2. Let us compute β_5 (the other cases are similar). We consider the curve $\{\delta_5(e^{i\theta}) : \theta \in [0, 2\pi]\}$ in \mathbb{C} (cf. Figure 2, right). We have

$$x(\theta) = \operatorname{Re}\left(\delta_k(e^{i\theta})\right) = p_5(\cos(\theta)),$$

where

$$p_5(X) = \frac{8}{5}X^4 - \frac{21}{5}X^3 + \frac{89}{30}X^2 + \frac{13}{30}X + \frac{1}{5}X^3 + \frac{1}$$

Thus, $\frac{d}{d\theta}x(\theta) = -\sin(\theta)p_5'(\cos\theta)$ with

$$p_5'(X) = \frac{32}{5}X^3 - \frac{63}{5}X^2 + \frac{89}{15}X + \frac{13}{30}.$$

The critical points of $x(\theta)$ are obtained for $\sin \theta = 0$ or $p'_5(\cos \theta) = 0$. The only real root of p'_5 is

$$X_1 = \frac{21}{32} - \frac{1121}{96\sqrt[3]{49041 - 16\sqrt{3891895}}} - \frac{1}{96}\sqrt[3]{49041 - 16\sqrt{3891895}}.$$

Among these critical points, the minimal value of $x(\theta)$ is obtained for $\cos \theta = X_1$ and so

$$\beta_5 = p_5(X_1) = 0.1855459753\dots$$
(2.10)



FIGURE 1. Region (2.7) for BDF2 (left) and BDF3 (right).

The analysis above shows:



FIGURE 2. Region (2.7) for BDF4 (left) and BDF5 (right).



FIGURE 3. Region (2.7) for BDF6.

Lemma 2.3. Let $2 \le k \le 6$. Then for every $\beta \le \beta_k$, we have

$$\operatorname{Re}\delta_k(z) - \beta > 0 \quad for \quad |z| < 1, \tag{2.11}$$

where β_k given by (2.8). Moreover, if $\beta > \beta_k$, then property (2.11) fails.

From these considerations, we deduce:

Theorem 2.4. Let $2 \leq k \leq 6$ and let δ_k be defined by (2.3). For each $\beta < \beta_k$, there exist a symmetric positive definite matrix $Q_k = (q_{ij}) \in \mathbb{R}^{k-1} \times \mathbb{R}^{k-1}$ and a symmetric positive definite matrix $R_k = (r_{ij}) \in \mathbb{R}^k \times \mathbb{R}^k$ such that for all $v_1, \ldots, v_k \in \mathbb{R}^N$,

$$\left\langle \sum_{i=0}^{k-1} \delta_i v_{k-i}, v_k \right\rangle = \beta \|v_k\|^2 + \sum_{i,j=1}^{k-1} q_{ij} \langle v_{i+1}, v_{j+1} \rangle - \sum_{i,j=1}^{k-1} q_{ij} \langle v_i, v_j \rangle + \sum_{i,j=1}^k r_{ij} \langle v_i, v_j \rangle.$$
(2.12)

The value β_k is the supremum of the numbers for which this property holds.

Proof. By Lemma 2.3, we may apply Lemma 2.1 to the polynomials $\delta(z) = \delta_k(z) - \beta_k$ and $\mu(z) = 1$. Thus, there exist a positive definite matrix $G = (g_{ij}) \in \mathbb{R}^{k-1} \times \mathbb{R}^{k-1}$ and real numbers a_1, \ldots, a_k such that

$$\left\langle \sum_{i=0}^{k-1} \delta_i v_{k-i}, v_k \right\rangle = \beta_k \|v_k\|^2 + \sum_{i,j=1}^{k-1} g_{ij} \langle v_{i+1}, v_{j+1} \rangle - \sum_{i,j=1}^{k-1} g_{ij} \langle v_i, v_j \rangle + \left\| \sum_{i=1}^k a_i v_i \right\|^2.$$
(2.13)

Let $\beta < \beta_k$. We consider for $\varepsilon > 0$ the symmetric matrix

$$G_{\varepsilon} = (g_{ij}^{\varepsilon}) \in \mathbb{R}^{k-1} \times \mathbb{R}^{k-1}$$

defined by

$$g_{ij}^{\varepsilon} = \begin{cases} g_{ij} & \text{if } i \neq j, \\ g_{ii} + i\varepsilon & \text{if } i = j. \end{cases}$$

By expressing G in terms of G_{ε} , equation (2.13) becomes

$$\left\langle \sum_{i=0}^{k-1} \delta_i v_{k-i}, v_k \right\rangle = \beta \|v_k\|^2 + \sum_{i,j=1}^{k-1} g_{ij}^{\varepsilon} \langle v_{i+1}, v_{j+1} \rangle - \sum_{i,j=1}^{k-1} g_{ij}^{\varepsilon} \langle v_i, v_j \rangle + (\beta_k - \beta - (k-1)\varepsilon) \|v_k\|^2 + \varepsilon \sum_{i=1}^{k-1} \|v_i\|^2 + \left\| \sum_{i=1}^k a_i v_i \right\|^2 (2.14)$$

By continuity, for $\varepsilon > 0$ small enough, the matrix G_{ε} is positive definite and

$$\beta_k - \beta - (k-1)\varepsilon > 0. \tag{2.15}$$

This shows that formula (2.12) holds with $Q_k = G_{\varepsilon}$ and R_k defined by the second line of (2.14).

Conversely, we assume now that (2.12) holds for a number $\beta \in \mathbb{R}$, a symmetric positive definite matrix Q_k and a positive definite matrix R_k . We may assume that N = 1 since $\mathbb{R} \simeq \mathbb{R} \times \{0\}^{N-1} \subset \mathbb{R}^N$. Then (2.12) shows that for all $v_1, \ldots, v_k \in \mathbb{R}$, we have

$$\sum_{i=0}^{k-1} \delta_i v_{k-i} v_k = \beta |v_k|^2 + V_2^T Q_k V_2 - V_1^T Q_k V_1 + V^T R_k V,$$

where

$$V = (v_1, \dots, v_k)^T$$
, $V_1 = (v_1, \dots, v_{k-1})^T$ and $V_2 = (v_2, \dots, v_k)^T$.

If we consider R_k and Q_k as hermitian matrices, this implies that for all $Z = (z_1, \ldots, z_k)^T \in \mathbb{C}^k$, we have

$$\operatorname{Re}\left(\sum_{i=0}^{k-1} \delta_i z_{k-i} \bar{z}_k\right) = \beta |z_k|^2 + Z_2^T Q_k \bar{Z}_2 - Z_1^T Q_k \bar{Z}_1 + Z^T R_k \bar{Z},$$

where

$$Z_1 = (z_1, \dots, z_{k-1})^T$$
 and $Z_2 = (z_2, \dots, z_k)^T$

Let $\zeta \in \mathbb{C}$ such that $|\zeta| > 1$. On choosing $Z = (\zeta^1, \zeta^2, \dots, \zeta^k)$, we obtain

$$|\zeta|^{2k} \operatorname{Re}\left(\sum_{i=0}^{k-1} \delta_i \zeta^{-i}\right) = \beta |\zeta|^{2k} + (|\zeta|^2 - 1) Z_1^T Q_k \bar{Z}_1 + Z^T R_k \bar{Z}.$$

Since $|\zeta| > 1$ and Q_k , R_k are positive definite, this shows that

$$\operatorname{Re}\left(\sum_{i=0}^{k-1} \delta_i \zeta^{-i}\right) - \beta > 0$$

By Lemma 2.3, we necessarily have $\beta \leq \beta_k$ and the proof is complete.

Remark 2.5. Since $\beta_6 < 0$, Theorem 2.4 shows that the BDF6 scheme is not quadratically stable in the sense defined in [7]. The cases k > 6 and k < 6 have been considered in [7]. Therefore, the k-step BDF method is quadratically stable if and only $1 \le k \le 5$.

Remark 2.6. For a given $\beta < \beta_k$, the couple (Q_k, R_k) in Theorem 2.4 is generally not unique. For instance, if k = 2 and $\beta = 1/2 < \beta_2 = 1$, we have

$$\frac{3}{2}\|v_2\|^2 - \frac{1}{2}\langle v_2, v_1 \rangle = \beta \|v_2\|^2 + \frac{1}{4}\|v_2\|^2 - \frac{1}{4}\|v_1\|^2 + \frac{1}{4}\|v_1 - v_2\|^2 + \frac{1}{2}\|v_2\|^2, \quad (2.16)$$

in which case $Q_2 = (1/4)$, or

$$\frac{3}{2}\|v_2\|^2 - \frac{1}{2}\langle v_2, v_1 \rangle = \beta \|v_2\|^2 + \frac{1}{2}\|v_2\|^2 - \frac{1}{2}\|v_1\|^2 + \frac{1}{2}\|v_1 - \frac{1}{2}v_2\|^2 + \frac{3}{8}\|v_2\|^2, \quad (2.17)$$

in which case $Q_2 = (1/2)$.

Remark 2.7. There is a constructive proof of Lemma 2.1, as explained in [16, Section V.6]. In turn, this gives a constructive proof of Theorem 2.4. Formula (2.13) can be obtained as follows. We first introduce the polynomials

$$\rho_k(z) = z^{k-1} \left(\delta_k(1/z) - \beta_k \right) \text{ and } \sigma_k(z) = z^{k-1}.$$

Then, we consider the function

$$E_k(z) = \frac{1}{2} \bigg(\rho_k(z)\sigma_k(1/z) + \rho_k(1/z)\sigma_k(s) \bigg).$$

It can be factorized as

$$E_k(z) = a_k(z)a_k(1/z)$$
 with $a_k(z) = \sum_{i=0}^{k-1} a_{i+1}z^i$. (2.18)

The coefficients a_1, \ldots, a_k in (2.13) are given by the coefficients of the polynomial $a_k(z)$. Next, we consider the polynomial

$$P_k(\omega,\zeta) = \frac{1}{2} \left(\rho_k(\omega)\sigma_k(\zeta) + \rho_k(\zeta)\sigma_k(\omega) \right) - a_k(\omega)a_k(\zeta).$$

By (2.18), P_k vanishes when $\omega \zeta - 1 = 0$. It can therefore be written as

$$P_k(\omega,\zeta) = \left(\omega\zeta - 1\right)\sum_{i,j=1}^{k-1} g_{ij}\omega^{i-1}\zeta^{j-1}.$$

This gives the matrix $G_k = (g_{ij})$ in (2.13). Once formula (2.13) is known, it is easy to find $\varepsilon > 0$ small enough such that G_{ε} is positive definite and (2.15) holds (cf. examples below). We note that in many applications, formula (2.13) is sufficient but in some cases, it is best to have a positive definite matrix R_k (see, e.g., Theorem 3.7).

Example 2.8. If k = 2, we have $\rho_2(z) = z/2 - 1/2$ and $\sigma_2(z) = z$. Thus,

$$E_2(z) = \frac{1}{4} (2 - z - 1/z) = a_2(z)a_2(1/z)$$
 with $a_2(z) = z/2 - 1/2$.

The polynomial P_2 is

$$P_2(\omega,\zeta) = \frac{1}{4}(\omega\zeta - 1),$$

so $G_2 = (1/4)$. Thus, formula (2.13) reads

$$\left\langle \frac{3}{2}v_2 - \frac{1}{2}v_1, v_2 \right\rangle = \beta_2 \|v_2\|^2 + \frac{1}{4}\|v_2\|^2 - \frac{1}{4}\|v_1\|^2 + \frac{1}{4}\|v_2 - v_1\|^2.$$

If we choose $\beta = 1/2$ and $\varepsilon = 1/4$ in (2.14), we recover formula (2.17).

Example 2.9. For k = 3, we have

$$\rho_3(z) = \frac{81}{96}z^2 - \frac{7}{6}z + \frac{1}{3} \text{ and } \sigma_3(z) = z^2.$$

The algorithm described in Remark 2.7 gives

$$a_3(z) = \frac{1}{\sqrt{6}}(z^2 - \frac{7}{4}z + 1)$$

and

$$P_3(\omega,\zeta) = \left(\omega\zeta - 1\right) \left(\frac{1}{6} + \frac{65}{96}\omega\zeta - \frac{7}{24}(\omega+\zeta)\right),\,$$

and so

$$G_3 = \frac{1}{6} \begin{pmatrix} 1 & -7/4 \\ -7/4 & 65/16 \end{pmatrix}.$$

The Cholesky factorization of G_3 is $G_3 = L_3 L_3^T$ with

$$L_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 & 0\\ -7/4 & 1 \end{pmatrix}$$

We recover the same formula as in [25, Theorem 1], where β_3 was obtained by maximizing the constant β in (2.12).

Remark 2.10. In Theorem 2.4, it is possible to replace \mathbb{R}^N by a Hilbert space \mathcal{H} and $\langle \cdot, \cdot \rangle$ by the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ in \mathcal{H} . The Euclidean norm is replaced by the Hilbertian norm in \mathcal{H} , $\|\cdot\|_{\mathcal{H}}$. For the Allen-Cahn equation, $\mathcal{H} = L^2(\Omega)$ [7] and for the Cahn-Hilliard equation, $\mathcal{H} = H^{-1}$ [8].

3. Gradient stability of BDF schemes

3.1. The continuous problem. We consider the gradient flow

$$y'(t) = -\nabla F(y(t)), \quad t \ge 0,$$
 (3.1)

where $y : [0, +\infty) \to \mathbb{R}^N$ and $F \in C^1(\mathbb{R}^N, \mathbb{R})$. Hereafter, we assume that F satisfies the following two conditions,

$$\langle \nabla F(v) - \nabla F(w), v - w \rangle \ge -c_F \|v - w\|^2, \quad \forall v, w \in \mathbb{R}^N,$$
 (3.2)

for some constant $c_F \geq 0$ and

$$\lim_{\|v\|\to+\infty} F(v) = +\infty.$$
(3.3)

Condition (3.3) is a coercivity condition. Condition (3.2) is known as a onesided Lipschitz condition for ∇F . Equivalently, we can say that $v \mapsto \nabla F + c_F v$ is a maximal monotone operator on \mathbb{R}^N or that the function

$$v \mapsto F(v) + \frac{c_F}{2} \|v\|^2$$

is convex on \mathbb{R}^N . A function F which satisfies (3.2) is called a *semiconvex* function.

If F is semiconvex, then there exists a minimum number $c_F^* \ge 0$ for which property (3.2) holds and we denote by c_F this optimal constant. In particular, F is convex if and only if $c_F = 0$.

By the Cauchy-Peano theorem and by monotonicity (see, e.g., [9]), for every $y_0 \in \mathbb{R}^N$, there exists a unique $y \in C^1([0, +\infty), \mathbb{R}^N)$ of (3.1) such that $y(0) = y_0$. On taking the inner product of (3.1) by -y'(t), we obtain that

$$\frac{d}{dt}F(y(t)) = -\|y'(t)\|^2, \qquad (3.4)$$

for all $t \ge 0$. In particular, F(y(t)) is nonincreasing, and the coercivity condition (3.3) implies that y is bounded. By Lasalle's invariance principle [17], the ω -limit set of y(0), defined by

$$\omega(y(0)) = \left\{ y^{\star} \in \mathbb{R}^{N} : \exists t_{n} \to +\infty, \ y(t_{n}) \to +\infty \right\},\$$

is a nonempty compact and connected subset of \mathbb{R}^N , which is included in the set \mathcal{S} of critical points of F, that is

$$\mathcal{S} = \left\{ y \in \mathbb{R}^N : \nabla F(y) = 0 \right\}.$$
(3.5)

If the critical points of F are isolated, this implies that y(t) tends to a critical point y^* of F in \mathbb{R}^N as t tends to $+\infty$. Otherwise, convergence to equilibrium may fail. We refer to [1, 17, 24] for such counterexamples where the ω -limit set is a cycle. Lojasiewicz [19, 20] proved that if F is real analytic, then y(t)converges to a critical point of F as t tends to $+\infty$.

In [10, Section 5.1], Chill and Jendoubi proved that a similar convergence result holds for the non-autonomous gradient flow

$$y'(t) = -\nabla F(y(t)) + g(t), \quad t \ge 0,$$
(3.6)

where $g: [0, +\infty) \to \mathbb{R}$ is a given function which tends to 0 at infinity in an appropriate sense. In this case, the system (3.6) is asymptotically autonomous.

3.2. **BDFk scheme applied to a gradient flow.** Let $\Delta t > 0$ denote the time step. The general k-step backward differentiation formula (BDF) for (3.1) reads: let y_0, \ldots, y_{k-1} in \mathbb{R}^N and for $n = 0, 1, \ldots$, let y_{n+k} solve

$$\sum_{j=1}^{k} \frac{1}{j} \partial^j y_{n+k} = -\Delta t \nabla F(y_{n+k}).$$
(3.7)

The one-step BDF method is the backward Euler scheme:

$$y_{n+1} - y_n = -\Delta t \nabla F(y_{n+1}), \quad n \ge 0.$$
 (3.8)

The two-step BDF method reads

$$\frac{3}{2}y_{n+2} - 2y_{n+1} + \frac{1}{2}y_n = -\Delta t \nabla F(y_{n+2}), \quad n \ge 0.$$
(3.9)

If $F \in C^{k+2}(\mathbb{R}^N, \mathbb{R})$ $(k \leq 6)$ and if the initial conditions are well chosen, the error between the solution y of (3.1) and its approximation given by the BDF scheme (3.7) is of order $O(\Delta t^k)$ on finite time intervals [27, Theorem 3.5.7].

We will consider a non-autonomous version of (3.7), namely

$$\sum_{j=1}^{k} \frac{1}{j} \partial^j y_{n+k} = -\Delta t \nabla F(y_{n+k}) + \Delta t g_{n+k}, \quad n \ge 0,$$
(3.10)

where $(g_n)_{n \in \mathbb{N}}$ is a given sequence in \mathbb{R}^N which belongs to $l^2(\mathbb{N})^N$. In particular, $g_n \to 0$ and the scheme (3.10) is asymptotically autonomous. It is a discrete version of (3.6).

For $1 \le k \le 6$, we define

$$\alpha_k = \sum_{j=1}^k 1/j > 0. \tag{3.11}$$

At every step $n \ge 0$, the vector y_{n+k} in (3.10) is computed from y_n , y_{n+1} , ..., y_{n+k-1} by solving a nonlinear equation. The assumptions (3.2)-(3.3) on F imply that the scheme is solvable for all Δt and uniquely solvable if Δt is small enough [7].

Proposition 3.1. Let y_0, \ldots, y_{k-1} be given in \mathbb{R}^N . For each $\Delta t > 0$, there exists a least one sequence $(y_n)_{n \in \mathbb{N}}$ in \mathbb{R}^N with initial values y_0, \ldots, y_{k-1} which complies with (3.10). If $c_F \Delta t < \alpha_k$, this sequence is unique.

If $c_F \Delta t \ge \alpha_k$, uniqueness may fail [25, Example 1].

3.3. Gradient stability of BDFk schemes. Let us see how quadratic stability implies gradient stability. We assume that $2 \le k \le 5$ since the BDF6 method is not quadratically stable. The case k = 1 is considered in [14].

By Theorem 2.4, for each $\beta \in (0, \beta_k)$, there exist a symmetric positive definite matrix $Q_k^{\beta} = (q_{ij}) \in \mathbb{R}^{k-1} \times \mathbb{R}^{k-1}$ and a symmetric positive definite matrix $R_k^{\beta} = (r_{ij}^{\beta}) \in \mathbb{R}^k \times \mathbb{R}^k$ such that for all $v_1, \ldots, v_k \in \mathbb{R}^N$, formula (2.12) holds. We denote by q_k^{β} and r_k^{β} the positive definitive quadratic forms on $(\mathbb{R}^N)^{k-1}$ and $(\mathbb{R}^N)^k$ associated to Q_k^{β} and R_k^{β} , respectively. Namely,

$$q_k^{\beta}(v_1, \dots, v_{k-1}) = \sum_{i,j=1}^{k-1} q_{ij} \langle v_i, v_j \rangle$$
 and $r_k^{\beta}(v_1, \dots, v_k) = \sum_{i,j=1}^k r_{ij} \langle v_i, v_j \rangle.$

For a sequence $(y_n)_{n \in \mathbb{N}}$ in \mathbb{R}^N , we write

$$\hat{y}_{n+k} = (y_{n+k}, \partial y_{n+k}, \dots, \partial y_{n+2})$$

and

$$\hat{F}_{\beta}(\hat{y}_{n+k}) = F(y_{n+k}) + \frac{1}{\Delta t} q_k^{\beta}(\partial y_{n+2}, \dots, \partial y_{n+k}).$$

In other words, the function \hat{F}_{β} is defined for every $(v_1, v_2, \ldots, v_k) \in (\mathbb{R}^N)^k$ by

$$\hat{F}_{\beta}(v_1, \dots, v_k) = F(v_1) + \frac{1}{\Delta t} q_k^{\beta}(v_k, \dots, v_2).$$
(3.12)

The gradient stability of the BDFk scheme (3.10) reads as follows.

Theorem 3.2. Let $2 \le k \le 5$. Assume that $c_F \Delta t < 2\beta_k$, let $\beta \in (c_F \Delta t/2, \beta_k)$ and set $\varepsilon = \beta - c_F \Delta t/2 > 0$. If $(y_n)_{n \in \mathbb{N}}$ is a sequence in \mathbb{R}^N which complies with the non-autonomous BDFk scheme (3.10), then for all $n \ge 0$ we have

$$\hat{F}_{\beta}(\hat{y}_{n+k}) + \frac{1}{\Delta t} r_k^{\beta}(\partial y_{n+1}, \dots, \partial y_{n+k}) \le \hat{F}_{\beta}(\hat{y}_{n+k-1}) + \frac{\Delta t}{4\varepsilon} \|g_{n+k}\|^2.$$
(3.13)

Proof. Using assumption (3.2), we easily obtain that

$$F(v) - F(w) \ge \langle \nabla F(w), v - w \rangle - \frac{c_F}{2} \|v - w\|^2, \quad \forall v, w \in \mathbb{R}^N.$$
(3.14)

We take the inner product of (3.10) with ∂y_{n+k} . We find

$$\Gamma_k = \Delta t \langle \nabla F(y_{n+k}), y_{n+k-1} - y_{n+k} \rangle + \Delta t \langle g_{n+k}, \partial y_{n+k} \rangle, \qquad (3.15)$$

where Γ_k is defined by (2.2). Using (3.14), the Cauchy-Schwarz inequality and Young's inequality, we obtain

$$\Gamma_k \leq \Delta t \left(F(y_{n+k-1}) - F(y_{n+k}) + \frac{c_F}{2} \|\partial y_{n+k}\|^2 \right) + \varepsilon \|\partial y_{n+k}\|^2 + \frac{\Delta t^2}{4\varepsilon} \|g_{n+k}\|^2.$$

Now, we use (2.4) and formula (2.12). This yields

$$(\beta - \frac{c_F \Delta t}{2} - \varepsilon) \|\partial y_{n+k}\|^2 + q_k^\beta (\partial y_{n+2}, \dots, \partial y_{n+k}) - q_k^\beta (\partial y_{n+1}, \dots, \partial y_{n+k-1}) + r_k^\beta (\partial y_{n+1}, \dots, \partial y_{n+k}) \le \Delta t \left(F(y_{n+k-1}) - F(y_{n+k}) \right) + \frac{\Delta t^2}{4\varepsilon} \|g_{n+k}\|^2,$$

which is equivalent to (3.13).

If $(y_n)_{n \in \mathbb{N}}$ is a sequence in \mathbb{R}^N , we define its ω -limit set by

$$\omega((y_n)_{n\in\mathbb{N}}) = \left\{ y^* \in \mathbb{R}^N : \exists n_p \to +\infty \text{ such that } y_{n_p} \to y^* \right\}.$$

We recall the following standard result (see, e.g., [25, Lemma 2]):

Lemma 3.3. If $(y_n)_{n \in \mathbb{N}}$ is a bounded sequence in \mathbb{R}^N such that $y_{n+1} - y_n \to 0$, then $\omega((y_n)_{n \in \mathbb{N}})$ is a nonempty compact and connected subset of \mathbb{R}^N .

Corollary 3.4. Let $2 \leq k \leq 5$ and assume that $c_F \Delta t < 2\beta_k$. If $(y_n)_{n \in \mathbb{N}}$ is a sequence in \mathbb{R}^N which complies with (3.10), then $\partial y_n \to 0$ and $\omega((y_n)_{n \in \mathbb{N}})$ is a nonempty compact and connected subset of \mathbb{R}^N included in \mathcal{S} (cf. (3.5)).

Proof. By summing relation (3.13) from n = 0 to p, we obtain

$$\hat{F}_{\beta}(\hat{y}_{p+k}) + \frac{1}{\Delta t} \sum_{n=0}^{p} r_{k}^{\beta}(\partial y_{n+1}, \dots, \partial y_{n+k}) \le \hat{F}_{\beta}(\hat{y}_{k-1}) + \frac{\Delta t}{4\varepsilon} \sum_{n=0}^{p} \|g_{n+k}\|^{2}.$$

Since q_k^{β} and r_k^{β} are positive definite, this shows that $(F(y_{n+k}))_{n\in\mathbb{N}}$ is bounded from above. By coercivity of F (cf. (3.3)), $(y_n)_{n\in\mathbb{N}}$ is bounded. The function F is also bounded from below, so by letting p tend to $+\infty$, we obtain

$$\frac{1}{\Delta t}\sum_{n=0}^{\infty}r_k^{\beta}(\partial y_{n+1},\ldots,\partial y_{n+k}) \leq -\inf_{\mathbb{R}}F + \hat{F}_{\beta}(\hat{y}_{k-1}) + \frac{\Delta t}{4\varepsilon}\sum_{n=0}^{\infty}\|g_{n+k}\|^2 < +\infty.$$

In particular, $r_k^{\beta}(\partial y_{n+1}, \ldots, \partial y_{n+k})$ tends to 0 and since r_k^{β} is positive definite, we have $\partial y_{n+1} \to 0$ as well. We may therefore apply Lemma 3.3. If $y^* \in \omega((y_n)_{n \in \mathbb{N}})$, we choose $n_p \to +\infty$ such that $y_{n_p+k} \to y^*$. By letting $n = n_p$ tend to infinity in (3.10), we obtain that $\nabla F(y^*) = 0$. This concludes the proof. \Box

Remark 3.5 (Uniqueness vs. stability condition). For k = 4 or 5, we have $2\beta_k < 2 < \alpha_k$. Thus, for k = 4 or 5, a gradient stable sequence which complies with (3.10) is uniquely defined by its initial values. In contrast, for k = 1, 2 or 3, we have $\alpha_k < 2\beta_k$, so if $\Delta t \in [\alpha_k/c_F, 2\beta_k/c_F)$, there may be several gradient stable sequences with the same initial values [25, Section 3.5]. These remarks are summarized in Table 1.

k	1	2	3	4	5	6
α_k	1	3/2	11/6	25/12	137/60	49/20
$2\beta_k$	2	2	95/48	1.62	0.371	-2.74

TABLE 1. Uniqueness and stability numbers for BDFk methods

We will use the following result, which is a direct consequence of [4, Theorem 2.4] and Łojasiewicz's inequality for analytic functions [19].

Theorem 3.6 ([4, 19]). Let $\Phi : \mathbb{R}^M \to \mathbb{R}$ be real analytic. Consider a bounded sequence $(x_n)_{n \in \mathbb{N}}$ in \mathbb{R}^M which satisfies the following conditions:

H1: There exists a positive constant c_1 such that for each $n \in \mathbb{N}$,

$$\Phi(x_n) - \Phi(x_{n+1}) \ge c_1 ||x_{n+1} - x_n||^2;$$

H2: There exists a positive constant c_2 such that for each $n \in \mathbb{N}$,

$$\|\nabla \Phi(x_{n+1})\| \le c_2 \|x_{n+1} - x_n\|_{\mathcal{F}}$$

Then the whole sequence $(x_n)_{n\in\mathbb{N}}$ converges in \mathbb{R}^M .

Theorem 3.7. Let $2 \le k \le 5$ and assume that $c_F \Delta t < 2\beta_k$. If $F : \mathbb{R}^N \to \mathbb{R}$ is real analytic and if $(y_n)_{n \in \mathbb{N}}$ is a sequence in \mathbb{R}^N which complies with (3.7), then the whole sequence $(y_n)_{n \in \mathbb{N}}$ converges to a critical point of F.

Proof. We apply Theorem 3.6 to the sequence $x_n = (y_n, \ldots, y_{n+k-1})$ with the function $\Phi(x_n) = \hat{F}_{\beta}(\hat{y}_{n+k-1})$ where \hat{F}_{β} is defined by (3.12). Namely, if $x = (a_1, \ldots, a_k) \in (\mathbb{R}^N)^k$, we have

$$\Phi(a_1,\ldots,a_k) = F(a_k) + \frac{1}{\Delta t} q_k^\beta(\partial a_k,\ldots,\partial a_2).$$

Since F is real analytic and q_k^{β} is a polynomial, it is clear that Φ is real analytic. The relation (3.13) shows that assumption **H1** is satisfied (we consider here the autonomous case $(g_n)_{n \in \mathbb{N}} = 0$). Indeed, since r_k^{β} is positive definite on $(\mathbb{R}^N)^k$, it is equivalent to the Euclidean norm on $(\mathbb{R}^N)^k$. The proof of Corollary 3.4 shows that the sequence $(\hat{y}_{n+k})_{n \in \mathbb{N}}$ is bounded, and so $(x_n)_{n \in \mathbb{N}}$ is bounded. Since q_k^{β} is quadratic, for each $1 \leq i \leq k$, $\partial_{a_i} q_k^{\beta} (\partial a_k, \ldots, \partial a_2)$ is a linear function of $(\partial a_2, \ldots, \partial a_k)$. Furthermore, the scheme (3.7) shows that $\nabla F(y_{n+k})$ is a linear combination of

$$(\partial y_{n+1},\ldots,\partial y_{n+k}) = \partial x_{n+1}.$$

This shows that assumption H2 is satisfied and we are in position to apply Theorem 3.6. The proof is complete.

Example 3.8. Theorem 3.7 can be applied to a standard finite element discretization of the Allen-Cahn equation (1.1) with the quartic potential (1.2) and Neumann boundary conditions. In this case, the constant c_F associated to the discrete version of the energy (1.3) is equal to $1/\varepsilon^2$ since [4, Proposition 5.1]

$$f''(s) = (3s^2 - 1)/\varepsilon^2 \ge -1/\varepsilon^2.$$

Thus, the condition on the time step reads $\Delta t < 2\beta_k \varepsilon^2$. Other examples are the Swift-Hohenberg equation [14] or the Cahn-Hilliard equation [5, 8].

The convergence result of Theorem 3.7 can be generalized in several ways. For instance, it is possible to consider the non-autonomous case (3.10). The sequence $(g_n)_{n \in \mathbb{N}}$ is assumed to satisfy a condition which implies that $(g_n)_{n \in \mathbb{N}}$ belongs to $l^1(\mathbb{N})^N$ [14, Remark 3]. The BDF1 and BDF2 schemes have been considered in [14] (see also [4, Scheme (4.4)]). In Theorem 3.7, it is also possible to require less regularity on F: the function F need not even be differentiable [25].

Acknowledgements

The author is thankful to Michel Pierre for suggesting him to investigate in detail the quadratic stability of the BDF6 scheme.

References

- P.-A. Absil, R. Mahony, and B. Andrews. Convergence of the iterates of descent methods for analytic cost functions. SIAM J. Optim., 16(2):531–547, 2005.
- G. Akrivis. Stability of implicit-explicit backward difference formulas for nonlinear parabolic equations. SIAM J. Numer. Anal., 53(1):464–484, 2015.
- [3] G. Akrivis, M. Chen, F. Yu, and Z. Zhou. The energy technique for the six-step BDF method. SIAM J. Numer. Anal., 59(5):2449–2472, 2021.
- [4] N. E. Alaa and M. Pierre. Convergence to equilibrium for discretized gradient-like systems with analytic features. IMA J. Numer. Anal., 33(4):1291–1321, 2013.
- [5] P. F. Antonietti, B. Merlet, M. Pierre, and M. Verani. Convergence to equilibrium for a second-order time semi-discretization of the Cahn-Hilliard equation. *AIMS Mathematics*, 1(3):178–194, 2016.
- [6] C. Baiocchi and M. Crouzeix. On the equivalence of A-stability and G-stability. Appl. Numer. Math., 5(1-2):19–22, 1989.
- [7] A. Bouchriti, M. Pierre, and N. E. Alaa. Gradient stability of high-order BDF methods and some applications. J. Difference Equ. Appl., 26(1):74–103, 2020.
- [8] M. Brachet, P. Parnaudeau, and M. Pierre. Convergence to equilibrium for time and space discretizations of the Cahn-Hilliard equation. in preparation.
- H. Brezis. Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert. North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, 1973.
- [10] R. Chill and M. A. Jendoubi. Convergence to steady states in asymptotically autonomous semilinear evolution equations. *Nonlinear Anal.*, 53(7-8):1017–1039, 2003.
- [11] C. F. Curtiss and J. O. Hirschfelder. Integration of stiff equations. Proc. Nat. Acad. Sci. U.S.A., 38:235–243, 1952.
- [12] G. Dahlquist. G-stability is equivalent to A-stability. BIT, 18(4):384–401, 1978.
- [13] C. M. Elliott and A. M. Stuart. The global dynamics of discrete semilinear parabolic equations. SIAM J. Numer. Anal., 30(6):1622–1663, 1993.
- [14] M. Grasselli and M. Pierre. Convergence to equilibrium of solutions of the backward Euler scheme for asymptotically autonomous second-order gradient-like systems. *Commun. Pure Appl. Anal.*, 11(6):2393–2416, 2012.
- [15] E. Hairer, S. P. Nørsett, and G. Wanner. Solving ordinary differential equations. I, volume 8 of Springer Series in Computational Mathematics. Springer-Verlag, Berlin, second edition, 1993.
- [16] E. Hairer and G. Wanner. Solving ordinary differential equations. II, volume 14 of Springer Series in Computational Mathematics. Springer-Verlag, Berlin, 2010. Stiff and differential-algebraic problems, Second revised edition, paperback.
- [17] A. Haraux and M. A. Jendoubi. The convergence problem for dissipative autonomous systems. Classical methods and recent advances. BCAM SpringerBriefs. SpringerBriefs in Mathematics. Springer, Cham; BCAM Basque Center for Applied Mathematics, Bilbao, 2015.
- [18] T. Horsin and M. A. Jendoubi. On the convergence to equilibria of a sequence defined by an implicit scheme. Discrete Contin. Dyn. Syst. Ser. S, 14(8):3017–3025, 2021.
- [19] S. Lojasiewicz. Ensembles semi-analytiques. I.H.E.S. Notes, 1965. available at https: //perso.univ-rennes1.fr/michel.coste/Lojasiewicz.pdf.

- [20] S. Lojasiewicz. Sur les trajectoires du gradient d'une fonction analytique. In Geometry seminars, 1982–1983 (Bologna, 1982/1983), pages 115–117. Univ. Stud. Bologna, Bologna, 1984.
- [21] C. Lubich, D. Mansour, and C. Venkataraman. Backward difference time discretization of parabolic differential equations on evolving surfaces. *IMA J. Numer. Anal.*, 33(4):1365–1385, 2013.
- [22] B. Merlet and M. Pierre. Convergence to equilibrium for the backward Euler scheme and applications. *Commun. Pure Appl. Anal.*, 9(3):685–702, 2010.
- [23] O. Nevanlinna and F. Odeh. Multiplier techniques for linear multistep methods. Numer. Funct. Anal. Optim., 3(4):377–423, 1981.
- [24] J. Palis, Jr. and W. de Melo. Geometric theory of dynamical systems. Springer-Verlag, New York-Berlin, 1982. An introduction, Translated from the Portuguese by A. K. Manning.
- [25] M. Pierre. Maximum time step for the BDF3 scheme applied to gradient flows. Calcolo, 58(1):Paper No. 3, 17, 2021.
- [26] L. Simon. Asymptotics for a class of nonlinear evolution equations, with applications to geometric problems. Ann. of Math. (2), 118(3):525–571, 1983.
- [27] A. M. Stuart and A. R. Humphries. Dynamical systems and numerical analysis, volume 2 of Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, Cambridge, 1996.

Morgan Pierre, Laboratoire de Mathématiques et Applications, Université de Poitiers, CNRS, F-86073 Poitiers, France.

Email address: morgan.pierre@math.univ-poitiers.fr