

HIGHER-ORDER LINEARLY IMPLICIT FULL DISCRETIZATION OF THE LANDAU–LIFSHITZ–GILBERT EQUATION

GEORGIOS AKRIVIS, MICHAEL FEISCHL, BALÁZS KOVÁCS, AND CHRISTIAN LUBICH

ABSTRACT. For the Landau–Lifshitz–Gilbert (LLG) equation of micromagnetics we study linearly implicit backward difference formula (BDF) time discretizations up to order 5 combined with higher-order non-conforming finite element space discretizations, which are based on the weak formulation due to Alouges but use approximate tangent spaces that are defined by L^2 -averaged instead of nodal orthogonality constraints. We prove stability and optimal-order error bounds in the situation of a sufficiently regular solution. For the BDF methods of orders 3 to 5, this requires that the damping parameter in the LLG equations be above a positive threshold; this condition is not needed for the A-stable methods of orders 1 and 2, for which furthermore a discrete energy inequality irrespective of solution regularity is proved.

1. INTRODUCTION

1.1. Scope. In this paper we study the convergence of higher-order time and space discretizations of the Landau–Lifshitz–Gilbert (LLG) equation, which is the basic model for phenomena in micromagnetism, such as in recording media [26, 36].

The main novelty of the paper lies in the construction and analysis of what is apparently the first numerical method for the LLG equation that is second-order convergent in both space and time to sufficiently regular solutions and that satisfies, as an important robustness property irrespective of regularity, a discrete energy inequality analogous to that of the continuous problem.

We study discretization in time by linearly implicit backward difference formulae (BDF) up to order 5 and discretization in space by finite elements of arbitrary polynomial degree. For the BDF methods up to order 2 we prove optimal-order error bounds in the situation of a sufficiently regular solution and a discrete energy inequality irrespective of solution regularity under very weak regularity assumptions on the data. For the BDF methods of orders 3 to 5, we prove optimal-order error bounds in the situation of a sufficiently regular solution under the additional condition that the damping parameter in the LLG equation be above a method-dependent positive threshold. However, no discrete energy inequality irrespective of solution regularity is obtained for the BDF methods of orders 3 to 5.

The discretization in space is done by a higher-order non-conforming finite element method based on the approach of Alouges [4, 5], which uses a projection to

Received by the editor March 13, 2019, and, in revised form, March 14, 2020.

2010 *Mathematics Subject Classification.* Primary 65M12, 65M15; Secondary 65L06.

Key words and phrases. BDF methods, non-conforming finite element method, Landau–Lifshitz–Gilbert equation, energy technique, stability.

an approximate tangent space to the normality constraint. Contrary to the point-wise orthogonality constraints at the nodes, which define the approximate tangent space in those papers and yield only first-order convergence also for finite elements with higher-degree polynomials, we here enforce orthogonality averaged over the finite element basis functions. With these modified approximate tangent spaces we prove H^1 -convergence of optimal order in space and time under the assumption of a sufficiently regular solution.

Key issues in the error analysis are the properties of the orthogonal projection onto the approximate tangent space, the higher-order consistency error analysis, and the proof of stable error propagation, which is based on non-standard energy estimates and uses both L^2 and maximum norm finite element analysis.

1.2. The Landau–Lifshitz–Gilbert equation. The standard phenomenological model for micromagnetism is provided by the Landau–Lifshitz (LL) equation

$$(1.1) \quad \partial_t \mathbf{m} = -\mathbf{m} \times \mathbf{H}_{\text{eff}} - \alpha \mathbf{m} \times (\mathbf{m} \times \mathbf{H}_{\text{eff}})$$

where the unknown magnetization field $\mathbf{m} = \mathbf{m}(x, t)$ takes values on the unit sphere \mathbb{S}^2 , $\alpha > 0$ is a dimensionless damping parameter, and the effective magnetic field \mathbf{H}_{eff} depends on the unknown \mathbf{m} . The Landau–Lifshitz equation (1.1) can be equivalently written in the Landau–Lifshitz–Gilbert form

$$(1.2) \quad \alpha \partial_t \mathbf{m} + \mathbf{m} \times \partial_t \mathbf{m} = (1 + \alpha^2) [\mathbf{H}_{\text{eff}} - (\mathbf{m} \cdot \mathbf{H}_{\text{eff}}) \mathbf{m}].$$

Indeed, in view of the vector identity $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c}) \mathbf{b} - (\mathbf{a} \cdot \mathbf{b}) \mathbf{c}$, for $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$, we have $-\mathbf{m} \times (\mathbf{m} \times \mathbf{H}_{\text{eff}}) = \mathbf{H}_{\text{eff}} - (\mathbf{m} \cdot \mathbf{H}_{\text{eff}}) \mathbf{m}$, and taking the vector product of (1.1) with \mathbf{m} and adding α times (1.1) then yields (1.2).

Since $\mathbf{m} \times \mathbf{a}$ is orthogonal to \mathbf{m} , for any $\mathbf{a} \in \mathbb{R}^3$, it is obvious from (1.1) that $\partial_t \mathbf{m}$ is orthogonal to \mathbf{m} : $\mathbf{m} \cdot \partial_t \mathbf{m} = 0$; we infer that the Euclidean norm satisfies $|\mathbf{m}(x, t)| = 1$ for all x and for all t , provided this is satisfied for the initial data.

The term in square brackets on the right-hand side in (1.2) can be rewritten as $\mathbf{P}(\mathbf{m}) \mathbf{H}_{\text{eff}}$, where (with \mathbf{I} the 3×3 unit matrix)

$$\mathbf{P}(\mathbf{m}) = \mathbf{I} - \mathbf{m} \mathbf{m}^T$$

is the orthogonal projection onto the tangent plane to the unit sphere \mathbb{S}^2 at \mathbf{m} .

In this paper we consider the situation

$$(1.3) \quad \mathbf{H}_{\text{eff}} = \frac{1}{1 + \alpha^2} (\Delta \mathbf{m} + \mathbf{H}),$$

where $\mathbf{H} = \mathbf{H}(x, t)$ is a given external magnetic field. The factor $1/(1 + \alpha^2)$ is chosen for convenience of presentation, but is inessential for the theory; it can be replaced by any positive constant factor.

With this choice of \mathbf{H}_{eff} , we arrive at the Landau–Lifshitz–Gilbert (LLG) equation in the form

$$(1.4) \quad \alpha \partial_t \mathbf{m} + \mathbf{m} \times \partial_t \mathbf{m} = \mathbf{P}(\mathbf{m}) (\Delta \mathbf{m} + \mathbf{H}).$$

We consider this equation as an initial-boundary value problem on a bounded domain $\Omega \subset \mathbb{R}^3$ and a time interval $0 \leq t \leq \bar{t}$, with homogeneous Neumann boundary conditions and initial data \mathbf{m}_0 taking values on the unit sphere, i.e., the Euclidean norm $|\mathbf{m}_0(x)|$ equals 1 for all $x \in \Omega$.

We consider the following weak formulation, first proposed by Alouges [4, 5]: Find the solution $\mathbf{m} : \Omega \times [0, \bar{t}] \rightarrow \mathbb{S}^2$ with $\mathbf{m}(\cdot, 0) = \mathbf{m}_0$ by determining, at

$\mathbf{m}(t) \in H^1(\Omega)^3$, the time derivative $\partial_t \mathbf{m}$ (omitting here and in the following the argument t) as that function in the tangent space

$$T(\mathbf{m}) := \{\boldsymbol{\varphi} \in L^2(\Omega)^3 : \mathbf{m} \cdot \boldsymbol{\varphi} = 0 \text{ a.e.}\} = \{\boldsymbol{\varphi} \in L^2(\Omega)^3 : \mathbf{P}(\mathbf{m})\boldsymbol{\varphi} = \boldsymbol{\varphi}\}$$

that satisfies, for all $\boldsymbol{\varphi} \in T(\mathbf{m}) \cap H^1(\Omega)^3$,

$$(1.5) \quad \alpha(\partial_t \mathbf{m}, \boldsymbol{\varphi}) + (\mathbf{m} \times \partial_t \mathbf{m}, \boldsymbol{\varphi}) + (\nabla \mathbf{m}, \nabla \boldsymbol{\varphi}) = (\mathbf{H}, \boldsymbol{\varphi}),$$

where the brackets (\cdot, \cdot) denote the L^2 inner product over the domain Ω . The numerical methods studied in this paper are based on this weak formulation.

1.3. Previous work. There is a rich literature on numerical methods for Landau–Lifshitz(–Gilbert) equations; for the numerical literature up to 2007 see the review by Cimrak [17].

Alouges & Jaisson [4, 5] propose linear finite element discretizations in space and linearly implicit backward Euler in time for the LLG equation in the weak formulation (1.5) and prove convergence *without* rates towards nonsmooth weak solutions, using a discrete energy inequality and compactness arguments. Convergence of this type was previously shown by Bartels & Prohl [11] for fully implicit methods that are based on a different formulation of the Landau–Lifshitz equation (1.1). In [6], convergence without rates towards weak solutions is shown for a method that is (formally) of “almost” order 2 in time, based on the midpoint rule, for the LLG equation with an effective magnetic field of a more general type than (1.3).

In a complementary line of research, convergence *with* rates has been studied under sufficiently strong regularity assumptions, which can, however, not be guaranteed over a given time interval, since solutions of the LLG equation may develop singularities. A first-order error bound for a linearly implicit time discretization of the Landau–Lifshitz equation (1.1) was proved by Cimrak [16]. Optimal-order error bounds for linearly implicit time discretizations based on the backward Euler and Crank–Nicolson methods combined with finite element full discretizations for a different version of the Landau–Lifshitz equation (1.1) were obtained under sufficient regularity assumptions by Gao [23] and An [7], respectively. In contrast to [4, 5, 6, 11], these methods do not satisfy an energy inequality irrespective of the solution regularity.

Numerical discretizations for the coupled system of the LLG equation (1.5) with the eddy current approximation of the Maxwell equations are studied by Feischl & Tran [21], with first-order error bounds in space and time under sufficient regularity assumptions. This also yields the first result of first-order convergence of the method of Alouges & Jaisson [4, 5].

There are several methods for the LLG equations that are of formal order 2 in time (though only of order 1 in space), e.g., [35, 31, 19], but none of them comes with an error analysis. Fully implicit BDF time discretizations for LLG equations have been used successfully in the computational physics literature [37], though without giving any error analysis.

To the authors’ knowledge, the second-order linearly implicit method proposed and studied here is thus the first numerical method for the LLG (or LL) equation that has rigorous a priori error estimates of order 2 in both space and time under high regularity assumptions and that satisfies a discrete energy inequality irrespective of regularity.

We conclude this brief survey of the literature with a remark: The existing convergence results either give convergence of a subsequence without rates to a weak solution (without imposing strong regularity assumptions), or they show convergence with rates towards sufficiently regular solutions (as we do here). Both approaches yield insight into the numerical methods and have their merits, and they complement each other. Clearly, neither approach is fully satisfactory, because convergence without rates of some subsequence is nothing to observe in actual computations, and on the other hand high regularity is at best provable for close to constant initial conditions [22] or over short time intervals. We regard the situation as analogous to the development of numerical methods and their analysis in other fields such as nonlinear hyperbolic conservation laws: second-order methods are highly popular in that field, even though they can only be shown to converge with very low order (1/2 or less or only without rates) for available regularity properties; see, e.g., [32, Chapter 3]. Nevertheless, second-order methods are favored over first-order methods in many applications, especially if they enjoy some qualitative properties that give them robustness in non-regular situations. A similar situation occurs with the LLG equation, where the most important qualitative property appears to be the energy inequality.

1.4. Outline. In Section 2 we describe the numerical methods studied in this paper. They use time discretization by linearly implicit BDF methods of orders up to 5 and space discretization by finite elements of arbitrary polynomial degree in a numerical scheme that is based on the weak formulation (1.5), with an approximate tangent space that enforces the orthogonality constraint approximately in an L^2 -projected sense.

In Section 3 we state our main results:

- For the full discretization of (1.5) by linearly implicit BDF methods of orders 1 and 2 and finite element methods of arbitrary polynomial degree we give optimal-order error bounds in the H^1 norm, under very mild mesh conditions, in the case of sufficiently regular solutions (Theorem 3.1). For these methods we also show a discrete energy inequality that requires only very weak regularity assumptions on the data (Proposition 3.1). This discrete energy inequality is of the same type as the one used in [5, 11] for proving convergence without rates to a weak solution.
- For the linearly implicit BDF methods of orders 3 to 5 and finite element methods with polynomial degree at least 2, we have optimal-order error bounds in the H^1 norm only if the damping parameter α is larger than some positive threshold, which depends on the order of the BDF method (Theorem 3.2). Moreover, a stronger (but still mild) CFL condition $\tau \leq ch$ is required. A discrete energy inequality under very weak regularity conditions is not available for the BDF methods of orders 3 to 5, in contrast to the A-stable BDF methods of orders 1 and 2.

In Section 4 we prove a perturbation result for the continuous problem by energy techniques, as a preparation for the proofs of our error bounds for the discretization.

In Section 5 we study properties of the L^2 -orthogonal projection onto the discrete tangent space, which are needed to ensure consistency of the full order and stability of the space discretization with the higher-order discrete tangent space.

In Section 6 we study consistency properties of the methods and present the error equation.

In Sections 7 and 8 we prove Theorems 3.1 and 3.2, respectively. The higher-order convergence proofs are separated into consistency (Section 6) and stability

estimates. The stability proofs use the technique of energy estimates, in an unusual version where the error equation is tested with a projection of the discrete time derivative of the error onto the discrete tangent space. These proofs are different for the A-stable BDF methods of orders 1 and 2 and for the BDF methods of orders 3 to 5. For the control of nonlinearities, the stability proofs also require pointwise error bounds, which are obtained with the help of finite element inverse inequalities from the H^1 error bounds of previous time steps.

In Section 9 we illustrate our results by numerical experiments.

In an Appendix we collect basic results on energy techniques for BDF methods that are needed for our stability proofs.

2. DISCRETIZATION OF THE LLG EQUATION

We now describe the time and space discretization that is proposed and studied in this paper.

2.1. Time discretization by linearly implicit BDF methods. We shall discretize the LLG equation (1.5) in time by the linearly implicit k -step BDF methods, $1 \leq k \leq 5$, described by the polynomials δ and γ ,

$$\delta(\zeta) = \sum_{\ell=1}^k \frac{1}{\ell} (1-\zeta)^\ell = \sum_{j=0}^k \delta_j \zeta^j, \quad \gamma(\zeta) = \frac{1}{\zeta} [1 - (1-\zeta)^k] = \sum_{i=0}^{k-1} \gamma_i \zeta^i.$$

We let $t_n = n\tau$, $n = 0, \dots, N$, be a uniform partition of the interval $[0, \bar{t}]$ with time step $\tau = \bar{t}/N$. For the k -step method we require k starting values \mathbf{m}^i for $i = 0, \dots, k-1$. For $n \geq k$, we determine the approximation \mathbf{m}^n to $\mathbf{m}(t_n)$ as follows. We first extrapolate the known values $\mathbf{m}^{n-k}, \dots, \mathbf{m}^{n-1}$ to a preliminary normalized approximation $\widehat{\mathbf{m}}^n$ at t_n ,

$$(2.1) \quad \widehat{\mathbf{m}}^n := \sum_{j=0}^{k-1} \gamma_j \mathbf{m}^{n-j-1} \bigg/ \left| \sum_{j=0}^{k-1} \gamma_j \mathbf{m}^{n-j-1} \right|.$$

To avoid potentially undefined quantities, we define $\widehat{\mathbf{m}}^n$ to be an arbitrary fixed unit vector if the denominator in the above formula is zero.

The derivative approximation $\dot{\mathbf{m}}^n$ and the solution approximation \mathbf{m}^n are related by the backward difference formula

$$(2.2) \quad \dot{\mathbf{m}}^n = \frac{1}{\tau} \sum_{j=0}^k \delta_j \mathbf{m}^{n-j}, \quad \text{i.e.,} \quad \mathbf{m}^n = \left(- \sum_{j=1}^k \delta_j \mathbf{m}^{n-j} + \tau \dot{\mathbf{m}}^n \right) / \delta_0.$$

We determine \mathbf{m}^n by requiring that for all $\varphi \in T(\widehat{\mathbf{m}}^n) \cap H^1(\Omega)^3$,

$$(2.3) \quad \alpha(\dot{\mathbf{m}}^n, \varphi) + (\widehat{\mathbf{m}}^n \times \dot{\mathbf{m}}^n, \varphi) + (\nabla \mathbf{m}^n, \nabla \varphi) = (\mathbf{H}(t_n), \varphi) \\ \dot{\mathbf{m}}^n \in T(\widehat{\mathbf{m}}^n), \quad \text{i.e.,} \quad \widehat{\mathbf{m}}^n \cdot \dot{\mathbf{m}}^n = 0.$$

Here we note that on inserting the formula in (2.2) for \mathbf{m}^n in the third term of (2.3), we obtain a linear constrained elliptic equation for $\dot{\mathbf{m}}^n \in T(\widehat{\mathbf{m}}^n) \cap H^1(\Omega)^3$ of the form

$$\alpha(\dot{\mathbf{m}}^n, \varphi) + (\widehat{\mathbf{m}}^n \times \dot{\mathbf{m}}^n, \varphi) + \frac{\tau}{\delta_0} (\nabla \dot{\mathbf{m}}^n, \nabla \varphi) = (\mathbf{f}^n, \varphi) \quad \forall \varphi \in T(\widehat{\mathbf{m}}^n) \cap H^1(\Omega)^3,$$

where \mathbf{f}^n consists of known terms. The bilinear form on the left-hand side is $H^1(\Omega)^3$ -coercive on $T(\widehat{\mathbf{m}}^n) \cap H^1(\Omega)^3$, and hence the above linear equation has a

unique solution $\hat{\mathbf{m}}^n \in T(\widehat{\mathbf{m}}^n) \cap H^1(\Omega)^3$ by the Lax–Milgram lemma. Once this elliptic equation is solved for $\hat{\mathbf{m}}^n$, we obtain the approximation $\mathbf{m}^n \in H^1(\Omega)^3$ to $\mathbf{m}(t_n)$ from the second formula in (2.2).

2.2. Full discretization by BDF and higher-order finite elements. For a family of regular and quasi-uniform finite element triangulations of Ω with maximum meshwidth $h > 0$ we form the Lagrange finite element spaces $V_h \subset H^1(\Omega)$ with piecewise polynomials of degree $r \geq 1$. We denote the L^2 -orthogonal projections onto the finite element space by $\Pi_h : L^2(\Omega) \rightarrow V_h$ and $\mathbf{\Pi}_h = \mathbf{I} \otimes \Pi_h : L^2(\Omega)^3 \rightarrow V_h^3$. With a function $\mathbf{m} \in H^1(\Omega)^3$ that vanishes nowhere on Ω , we associate the discrete tangent space

$$(2.4) \quad \begin{aligned} T_h(\mathbf{m}) &= \{\boldsymbol{\varphi}_h \in V_h^3 : (\mathbf{m} \cdot \boldsymbol{\varphi}_h, v_h) = 0 \quad \forall v_h \in V_h\} \\ &= \{\boldsymbol{\varphi}_h \in V_h^3 : \mathbf{\Pi}_h(\mathbf{m} \cdot \boldsymbol{\varphi}_h) = 0\}. \end{aligned}$$

This space is different from the discrete tangent space used in [4, 5], where the orthogonality constraint $\mathbf{m} \cdot \boldsymbol{\varphi}_h = 0$ is required to hold pointwise at the finite element nodes. Here, the constraint is enforced weakly on the finite element space, as is done in various saddle point problems for partial differential equations, for example for the divergence-free constraint in the Stokes problem [14, 25]. In contrast to that example, here the bilinear form associated with the linear constraint, i.e., $b(\mathbf{m}; \boldsymbol{\varphi}_h, v_h) = (\mathbf{m} \cdot \boldsymbol{\varphi}_h, v_h)$, depends on the state \mathbf{m} . This dependence substantially affects both the implementation and the error analysis.

Following the general approach of [4, 5] with this modified discrete tangent space, we discretize (1.5) in space by determining the time derivative $\partial_t \mathbf{m}_h(t) \in T_h(\mathbf{m}_h(t))$ such that (omitting the argument t)

$$(2.5) \quad \begin{aligned} \alpha(\partial_t \mathbf{m}_h, \boldsymbol{\varphi}_h) + (\mathbf{m}_h \times \partial_t \mathbf{m}_h, \boldsymbol{\varphi}_h) + (\nabla \mathbf{m}_h, \nabla \boldsymbol{\varphi}_h) \\ = (\mathbf{H}, \boldsymbol{\varphi}_h) \quad \forall \boldsymbol{\varphi}_h \in T_h(\mathbf{m}_h), \end{aligned}$$

where the brackets (\cdot, \cdot) denote again the L^2 inner product over the domain Ω .

The full discretization with the linearly implicit BDF method is then readily obtained from (2.3): determine $\hat{\mathbf{m}}_h^n \in T_h(\widehat{\mathbf{m}}_h^n)$ such that

$$(2.6) \quad \alpha(\hat{\mathbf{m}}_h^n, \boldsymbol{\varphi}_h) + (\widehat{\mathbf{m}}_h^n \times \dot{\mathbf{m}}_h^n, \boldsymbol{\varphi}_h) + (\nabla \mathbf{m}_h^n, \nabla \boldsymbol{\varphi}_h) = (\mathbf{H}(t_n), \boldsymbol{\varphi}_h) \quad \forall \boldsymbol{\varphi}_h \in T_h(\widehat{\mathbf{m}}_h^n),$$

where $\widehat{\mathbf{m}}_h^n$ and $\dot{\mathbf{m}}_h^n$ are related to \mathbf{m}_h^{n-j} for $j = 0, \dots, k$ in the same way as in (2.1) and (2.2) above with \mathbf{m}_h^{n-j} in place of \mathbf{m}^{n-j} , viz.,

$$(2.7) \quad \dot{\mathbf{m}}_h^n = \frac{1}{\tau} \sum_{j=0}^k \delta_j \mathbf{m}_h^{n-j}, \quad \widehat{\mathbf{m}}_h^n = \sum_{j=0}^{k-1} \gamma_j \mathbf{m}_h^{n-j-1} \left/ \left| \sum_{j=0}^{k-1} \gamma_j \mathbf{m}_h^{n-j-1} \right| \right.$$

To avoid potentially undefined quantities, we define $\widehat{\mathbf{m}}_h^n$ to be an arbitrary fixed unit vector if the denominator in the above formula is zero. (We will, however, show that this does not occur in the situation of sufficient regularity.)

To implement the discrete tangent space $T_h(\widehat{\mathbf{m}}_h^n)$, there are at least two options: using the constraints $\mathbf{\Pi}_h(\mathbf{m} \cdot \boldsymbol{\varphi}_h) = 0$ or constructing a local basis of $T_h(\mathbf{m})$.

(a) *Constraints:* Let ϕ_i for $i = 1, \dots, N_h := \dim V_h$ denote the nodal basis of V_h and denote the basis functions of V_h^3 by $\phi_{\mathbf{i}} = \mathbf{e}_k \otimes \phi_i$ for $\mathbf{i} = (i, k)$, where \mathbf{e}_k for $k = 1, 2, 3$ are the standard unit vectors of \mathbb{R}^3 . We denote by M and A the usual mass and stiffness matrices, respectively, with entries $m_{ij} = (\phi_i, \phi_j)_{L^2(\Omega)}$ and $a_{ij} = (\nabla \phi_i, \nabla \phi_j)_{L^2(\Omega)^3}$. We further introduce the sparse skew-symmetric matrix

$S^n = (s_{i,j}^n) \in \mathbb{R}^{3N_h \times 3N_h}$ with entries $s_{i,j}^n = (\widehat{\mathbf{m}}_h^n \times \phi_i, \phi_j)_{L^2(\Omega)^3}$ and the sparse constraint matrix $C^n = (c_{i,j}^n) \in \mathbb{R}^{3N_h \times N_h}$ by $c_{i,j}^n = (\widehat{\mathbf{m}}_h^n \cdot \phi_i, \phi_j)_{L^2(\Omega)}$. Finally, we denote the matrix of the unconstrained time-discrete problem as

$$K^n = \alpha \mathbf{I} \otimes M + \frac{\tau}{\delta_0} \mathbf{I} \otimes A + S^n.$$

Let $\dot{m}^n \in \mathbb{R}^{3N_h}$ denote the nodal vector of $\dot{\mathbf{m}}_h^n \in T_h(\widehat{\mathbf{m}}_h^n)$. In this setting, (2.6) yields a system of linear equations of saddle point type

$$\begin{aligned} K^n \dot{m}^n + (C^n)^T \lambda^n &= f^n, \\ C^n \dot{m}^n &= 0, \end{aligned}$$

where $\lambda^n \in \mathbb{R}^{N_h}$ is the unknown vector of Lagrange multipliers and $f^n \in \mathbb{R}^{3N_h}$ is a known right-hand side.

(b) *Local basis*: It is possible to compute a local basis of $T_h(\mathbf{m})$ by solving small local problems. To see that, let $\omega \subset \Omega$ denote a collection of elements of the mesh and let $\bar{\omega} \supset \omega$ denote the same set plus the layer of elements touching ω (the patch of ω). A sufficient (and necessary) condition for $\varphi_h \in V_h^3$ with $\text{supp}(\varphi_h) \subseteq \omega$ to belong to $T_h(\mathbf{m})$ is

$$(2.8) \quad (\mathbf{m} \cdot \varphi_h, \psi_h) = 0 \quad \text{for all } \psi_h \in V_h \text{ with } \text{supp}(\psi_h) \subseteq \bar{\omega}.$$

If we denote by $\#\omega$ the number of generalized hat functions of V_h supported in ω , the space of functions in V_h^3 with support in ω is $3\#\omega$ -dimensional. On the other hand, the space of test functions in (2.8) is $\#\bar{\omega}$ -dimensional. We may choose ω sufficiently large (depending only on shape regularity) such that $3\#\omega > \#\bar{\omega}$ and hence (2.8) has at least one solution which is then a local basis function of $T_h(\mathbf{m})$. Choosing different ω to cover Ω yields a full basis of $T_h(\mathbf{m})$.

Let us denote the so obtained basis of $T_h(\widehat{\mathbf{m}}_h^n)$ by (ψ_ℓ^n) , given via $\psi_\ell^n = \sum_i \phi_i b_{i\ell}^n$, and the sparse basis matrix by $B^n = (b_{i\ell}^n)$. Then, the nodal vector $\dot{m}^n = B^n x^n$ is obtained by solving the linear system

$$(B^n)^T K^n B^n x^n = (B^n)^T f^n.$$

An advantage of this approach is that the dimension is roughly halved compared to the formulation with constraints. However, the efficiency of one approach versus the other depends heavily on the numerical linear algebra used. Such comparisons are outside the scope of this paper.

Remark 2.1. The algorithm described above does not enforce the norm constraint $|\mathbf{m}| = 1$ at the nodes. The user might add a normalization step in the definition of \mathbf{m}^n in (2.2). However, here we do not consider this normalized variant of the method, whose convergence properties are not obvious to derive.

Remark 2.2. Differently to [4], we do not use the pointwise discrete tangent space

$$\begin{aligned} T_h^{\text{pw}}(\mathbf{m}) &= \{\varphi_h \in V_h^3 : \mathbf{m} \cdot \varphi_h = 0 \text{ at every node}\} \\ &= \{\varphi_h \in V_h^3 : I_h(\mathbf{m} \cdot \varphi_h) = 0\} = \mathbf{I}_h \mathbf{P}(\mathbf{m}) V_h^3, \end{aligned}$$

where $I_h : C(\bar{\Omega}) \rightarrow V_h$ denotes finite element interpolation and $\mathbf{I}_h = \mathbf{I} \otimes I_h : C(\bar{\Omega})^3 \rightarrow V_h^3$. It is already reported in [4, Section 4] that an improvement of the order with higher-degree finite elements could not be observed in numerical experiments when using the pointwise tangent spaces in the discretization (2.5). Our analysis shows a lack of consistency of optimal order in the discretization with

$T_h^{\text{PW}}(\mathbf{m})$, which originates from the fact that $\mathbf{I}_h \mathbf{P}(\mathbf{m})$ is not self-adjoint. The order reduction can, however, be cured by adding a correction term: in the n th time step, determine $\dot{\mathbf{m}}_h^n \in T_h^{\text{PW}}(\widehat{\mathbf{m}}_h^n)$ such that for all $\varphi_h \in T_h^{\text{PW}}(\widehat{\mathbf{m}}_h^n)$,

$$(2.9) \quad \begin{aligned} & \alpha(\dot{\mathbf{m}}_h^n, \varphi_h) + (\widehat{\mathbf{m}}_h^n \times \dot{\mathbf{m}}_h^n, \varphi_h) + (\nabla \mathbf{m}_h^n, \nabla \varphi_h) \\ & - (\nabla \widehat{\mathbf{m}}_h^n, \nabla (\mathbf{I} - \mathbf{P}(\widehat{\mathbf{m}}_h^n)) \varphi_h) = (\mathbf{P}(\widehat{\mathbf{m}}_h^n) \mathbf{H}(t_n), \varphi_h), \end{aligned}$$

with notation $\widehat{\mathbf{m}}_h^n$ and $\dot{\mathbf{m}}_h^n$ as in (2.7). With the techniques of the present paper, it can be shown that like (2.6), also this discretization converges with optimal order in the H^1 norm under sufficient regularity conditions. Since this paper is already rather long, we do not include the proof of this result. In contrast to (2.6) for the first- and second-order BDF methods, the method (2.9) does not admit an h - and τ -independent bound of the energy that is irrespective of the smoothness of the solution.

3. MAIN RESULTS

3.1. Error bound and energy inequality for BDF of orders 1 and 2. For the full discretization with first- and second-order BDF methods and finite elements of arbitrary polynomial degree $r \geq 1$ we will prove the following optimal-order error bound in Sections 5 to 7.

Theorem 3.1 (Error bound for orders $k = 1, 2$). *Consider the full discretization (2.6) of the LLG equation (1.4) by the linearly implicit k -step BDF time discretization for $k \leq 2$ and finite elements of polynomial degree $r \geq 1$ from a family of regular and quasi-uniform triangulations of Ω . Suppose that the solution \mathbf{m} of the LLG equation is sufficiently regular. Then, there exist $\bar{\tau} > 0$ and $\bar{h} > 0$ such that for numerical solutions obtained with step sizes $\tau \leq \bar{\tau}$ and meshwidths $h \leq \bar{h}$, which are restricted by the very mild CFL-type condition*

$$\tau^k \leq \bar{c} h^{1/2}$$

with a sufficiently small constant \bar{c} (independent of h and τ), the errors are bounded by

$$(3.1) \quad \|\mathbf{m}_h^n - \mathbf{m}(t_n)\|_{H^1(\Omega)^3} \leq C(\tau^k + h^r) \quad \text{for } t_n = n\tau \leq \bar{t},$$

where C is independent of h, τ and n (but depends on α and exponentially on \bar{t}), provided that the errors of the starting values also satisfy such a bound.

The precise regularity requirements are as follows:

$$(3.2) \quad \begin{aligned} & \mathbf{m} \in C^{k+1}([0, \bar{t}], L^\infty(\Omega)^3) \cap C^1([0, \bar{t}], W^{r+1, \infty}(\Omega)^3), \\ & \Delta \mathbf{m} + \mathbf{H} \in C([0, \bar{t}], W^{r+1, \infty}(\Omega)^3). \end{aligned}$$

Remark 3.1 (*Discrepancy from normality*). Since $\mathbf{m}(x, t_n)$ are unit vectors, an immediate consequence of the error estimate (3.1) is that

$$(3.3) \quad \|1 - |\mathbf{m}_h^n|\|_{L^2(\Omega)} \leq C(\tau^k + h^r) \quad \text{for } t_n = n\tau \leq \bar{t},$$

with a constant C independent of n, τ and h . The proof of Theorem 3.1 also shows that the denominator in the definition of the normalized extrapolated value $\widehat{\mathbf{m}}_h^n$ satisfies

$$(3.4) \quad \left\| 1 - \left| \sum_{j=0}^{k-1} \gamma_j \mathbf{m}_h^{n-j-1} \right| \right\|_{L^\infty(\Omega)} \leq C h^{-1/2} (\tau^k + h^r) \leq \frac{1}{2} \quad \text{for } t_n = n\tau \leq \bar{t},$$

which in particular ensures that $\widehat{\mathbf{m}}_h^n$ is unambiguously defined. We note, however, that (3.3) and (3.4) are not guaranteed to hold without sufficient regularity of the solution.

Testing with $\boldsymbol{\varphi} = \partial_t \mathbf{m} \in T(\mathbf{m})$ in (1.5), we obtain (only formally, if $\partial_t \mathbf{m}$ is not in $H^1(\Omega)^3$)

$$\alpha(\partial_t \mathbf{m}, \partial_t \mathbf{m}) + (\nabla \mathbf{m}, \partial_t \nabla \mathbf{m}) = (\mathbf{H}, \partial_t \mathbf{m}),$$

which, by integration in time and the Cauchy–Schwarz and Young inequalities, implies the energy inequality

$$\|\nabla \mathbf{m}(t)\|_{L^2}^2 + \frac{1}{2}\alpha \int_0^t \|\partial_t \mathbf{m}(s)\|_{L^2}^2 ds \leq \|\nabla \mathbf{m}(0)\|_{L^2}^2 + \frac{1}{2\alpha} \int_0^t \|\mathbf{H}(s)\|_{L^2}^2 ds.$$

Similarly, we test with $\boldsymbol{\varphi}_h = \dot{\mathbf{m}}_h^n \in T_h(\widehat{\mathbf{m}}_h^n)$ in (2.6). Then we can prove the following discrete energy inequality, which holds under very weak regularity assumptions on the data.

Proposition 3.1 (Energy inequality for orders $k = 1, 2$). *Consider the full discretization (2.6) of the LLG equation (1.4) by the linearly implicit k -step BDF time discretization for $k \leq 2$ and finite elements of polynomial degree $r \geq 1$. Then, the numerical solution satisfies the following discrete energy inequality: for $n \geq k$ with $n\tau \leq \bar{t}$,*

$$\gamma_k^- \|\nabla \mathbf{m}_h^n\|_{L^2}^2 + \frac{1}{2}\alpha\tau \sum_{j=k}^n \|\dot{\mathbf{m}}_h^j\|_{L^2}^2 \leq \gamma_k^+ \sum_{i=0}^{k-1} \|\nabla \mathbf{m}_h^i\|_{L^2}^2 + \frac{\tau}{2\alpha} \sum_{j=k}^n \|\mathbf{H}(t_j)\|_{L^2}^2,$$

where $\gamma_1^\pm = 1$ and $\gamma_2^\pm = (3 \pm 2\sqrt{2})/4$.

This energy inequality is an important robustness indicator of the numerical method. In [5, 11], such energy inequalities are used to prove convergence without rates for a subsequence $\tau_\ell \rightarrow 0$ and $h_\ell \rightarrow 0$ to a weak solution of the LLG equation for the numerical schemes considered there. Those methods have $\gamma^\pm = 1$, which is inessential in the proofs, and the numerical approximation from those methods is of unit norm at the finite element nodes, which appears essential in the proofs of [5, 11].

As the proof of Proposition 3.1 is short, we give it here.

Proof. The proof relies on the A-stability of the first- and second-order BDF methods via Dahlquist’s G-stability theory as expressed in Lemma A.1 of the Appendix, used with $\delta(\zeta) = \sum_{\ell=1}^k (1 - \zeta)^\ell / \ell$ and $\mu(\zeta) = 1$. The positive definite symmetric matrices $G = (g_{ij})_{i,j=1}^k$ are known to be $G = 1$ for $k = 1$ and (see [27, p. 309])

$$G = \frac{1}{4} \begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix} \quad \text{for } k = 2,$$

which has the eigenvalues $\gamma^\pm = (3 \pm 2\sqrt{2})/4$.

We test with $\boldsymbol{\varphi}_h = \dot{\mathbf{m}}_h^n \in T_h(\widehat{\mathbf{m}}_h^n)$ in (2.6) and note $(\widehat{\mathbf{m}}_h^n \times \dot{\mathbf{m}}_h^n, \dot{\mathbf{m}}_h^n) = 0$, so that

$$\alpha \|\dot{\mathbf{m}}_h^n\|_{L^2}^2 + (\nabla \mathbf{m}_h^n, \nabla \dot{\mathbf{m}}_h^n) = (\mathbf{H}(t_n), \dot{\mathbf{m}}_h^n).$$

The right-hand side is bounded by

$$(\mathbf{H}(t_n), \dot{\mathbf{m}}_h^n) \leq \frac{\alpha}{2} \|\dot{\mathbf{m}}_h^n\|_{L^2}^2 + \frac{1}{2\alpha} \|\mathbf{H}(t_n)\|_{L^2}^2.$$

Recalling the definition of $\dot{\mathbf{m}}_h^n$, we have by Lemma A.1

$$(\nabla \mathbf{m}_h^n, \nabla \dot{\mathbf{m}}_h^n) \geq \frac{1}{\tau} \sum_{i,j=1}^k g_{ij} (\nabla \mathbf{m}_h^{n-i+1}, \nabla \mathbf{m}_h^{n-j+1}) - \frac{1}{\tau} \sum_{i,j=1}^k g_{ij} (\nabla \mathbf{m}_h^{n-i}, \nabla \mathbf{m}_h^{n-j}).$$

We fix \bar{n} with $k \leq \bar{n} \leq \bar{t}/\tau$ and sum from $n = k$ to \bar{n} to obtain

$$\begin{aligned} & \sum_{i,j=1}^k g_{ij} (\nabla \mathbf{m}_h^{\bar{n}-i+1}, \nabla \mathbf{m}_h^{\bar{n}-j+1}) + \frac{1}{2} \alpha \tau \sum_{n=k}^{\bar{n}} \|\dot{\mathbf{m}}_h^n\|_{L^2}^2 \\ & \leq \sum_{i,j=1}^k g_{ij} (\nabla \mathbf{m}_h^{k-i}, \nabla \mathbf{m}_h^{k-j}) + \frac{\tau}{2\alpha} \sum_{n=k}^{\bar{n}} \|\mathbf{H}(t_n)\|_{L^2}^2. \end{aligned}$$

Noting that

$$\begin{aligned} \gamma^- \|\nabla \mathbf{m}_h^{\bar{n}}\|_{L^2}^2 & \leq \sum_{i,j=1}^k g_{ij} (\nabla \mathbf{m}_h^{\bar{n}-i+1}, \nabla \mathbf{m}_h^{\bar{n}-j+1}), \\ \sum_{i,j=1}^k g_{ij} (\nabla \mathbf{m}_h^{k-i}, \nabla \mathbf{m}_h^{k-j}) & \leq \gamma^+ \sum_{i=0}^{k-1} \|\nabla \mathbf{m}_h^i\|_{L^2}^2, \end{aligned}$$

we obtain the stated result. \square

3.2. Error bound for BDF of orders 3 to 5. For the BDF methods of orders 3 to 5 we prove the following result in Section 8. Here we require a stronger, but still moderate stepsize restriction in terms of the meshwidth. More importantly, we must impose a positive lower bound on the damping parameter α of (1.1).

Theorem 3.2 (Error bound for orders $k = 3, 4, 5$). *Consider the full discretization (2.6) of the LLG equation (1.4) by the linearly implicit k -step BDF time discretization for $3 \leq k \leq 5$ and finite elements of polynomial degree $r \geq 2$ from a family of regular and quasi-uniform triangulations of Ω . Suppose that the solution \mathbf{m} of the LLG equation has the regularity (3.2), and that the damping parameter α satisfies*

$$(3.5) \quad \alpha > \alpha_k \quad \text{with} \\ \alpha_k = 0.0913, 0.4041, 4.4348, \quad \text{for } k = 3, 4, 5, \text{ respectively.}$$

Then, for an arbitrary constant $\bar{C} > 0$, there exist $\bar{\tau} > 0$ and $\bar{h} > 0$ such that for numerical solutions obtained with step sizes $\tau \leq \bar{\tau}$ and meshwidths $h \leq \bar{h}$ that are restricted by

$$(3.6) \quad \tau \leq \bar{C}h,$$

the errors are bounded by

$$\|\mathbf{m}_h^n - \mathbf{m}(t_n)\|_{H^1(\Omega)^3} \leq C(\tau^k + h^r) \quad \text{for } t_n = n\tau \leq \bar{t},$$

where C is independent of h, τ and n (but depends on α and exponentially on $\bar{C}\bar{t}$), provided that the errors of the starting values also satisfy such a bound.

Theorem 3.2 limits the use of the BDF methods of orders higher than 2 (and more severely for orders higher than 3) to applications with a large damping parameter α , such as cases described in [24, 39]. We remark, however, that in many situations α is of magnitude 10^{-2} or even smaller [10]. A very small damping parameter α affects not only the methods considered here. To our knowledge, the error analysis

of any numerical method proposed in the literature breaks down as $\alpha \rightarrow 0$, as does the energy inequality.

It is not surprising that a positive lower bound on α arises for the methods of orders $k \geq 3$, since they are not A-stable and a lower bound on α is required also for the simplified linear problem $(\alpha + i)\partial_t u = \Delta u$, which arises from (1.4) by freezing \mathbf{m} in the term $\mathbf{m} \times \partial_t \mathbf{m}$ and diagonalizing this skew-symmetric linear operator (with eigenvalues $\pm i$ and 0) and by omitting the projection $\mathbf{P}(\mathbf{m})$ on the right-hand side of (1.4).

The proof of Theorem 3.2 uses a variant of the Nevanlinna–Odeh multiplier technique [34], which is described in the Appendix for the convenience of the reader. While for sufficiently large α we have an optimal-order error bound in the case of a smooth solution, there is apparently no discrete energy inequality under weak regularity assumptions similar to Proposition 3.1 for the BDF methods of orders 3 to 5.

As in Remark 3.1, the error bounds also allow us to bound the discrepancy from normality.

4. A CONTINUOUS PERTURBATION RESULT

In this section we present a perturbation result for the continuous problem, because we will later transfer the arguments of its proof to the discretizations to prove stability and convergence of the numerical methods.

Let $\mathbf{m}(t)$ be a solution of (1.4) for $0 \leq t \leq \bar{t}$, and let $\mathbf{m}_*(t)$, also of unit length, solve the same equation up to a defect $\mathbf{d}(t)$ for $0 \leq t \leq \bar{t}$:

$$(4.1) \quad \begin{aligned} \alpha \partial_t \mathbf{m}_* + \mathbf{m}_* \times \partial_t \mathbf{m}_* &= \mathbf{P}(\mathbf{m}_*)(\Delta \mathbf{m}_* + \mathbf{H}) + \mathbf{d} \\ &= \mathbf{P}(\mathbf{m})(\Delta \mathbf{m}_* + \mathbf{H}) + \mathbf{r}, \end{aligned}$$

with

$$\mathbf{r} = -(\mathbf{P}(\mathbf{m}) - \mathbf{P}(\mathbf{m}_*))(\Delta \mathbf{m}_* + \mathbf{H}) + \mathbf{d}.$$

Then, \mathbf{m}_* also solves the perturbed weak formulation

$$\alpha(\partial_t \mathbf{m}_*, \varphi) + (\mathbf{m}_* \times \partial_t \mathbf{m}_*, \varphi) + (\nabla \mathbf{m}_*, \nabla \varphi) = (\mathbf{r} + \mathbf{H}, \varphi) \quad \forall \varphi \in T(\mathbf{m}) \cap H^1(\Omega)^3,$$

and the error $\mathbf{e} = \mathbf{m} - \mathbf{m}_*$ satisfies the error equation

$$(4.2) \quad \begin{aligned} \alpha(\partial_t \mathbf{e}, \varphi) + (\mathbf{e} \times \partial_t \mathbf{m}_*, \varphi) + (\mathbf{m} \times \partial_t \mathbf{e}, \varphi) + (\nabla \mathbf{e}, \nabla \varphi) &= -(\mathbf{r}, \varphi) \\ &\forall \varphi \in T(\mathbf{m}) \cap H^1(\Omega)^3. \end{aligned}$$

Before we turn to the perturbation result, we need Lipschitz-type bounds for the orthogonal projection $\mathbf{P}(\mathbf{m}) = \mathbf{I} - \mathbf{m}\mathbf{m}^T$ applied to sufficiently regular functions.

Lemma 4.1. *The projection $\mathbf{P}(\cdot)$ satisfies the following estimates, for functions $\mathbf{m}, \mathbf{m}_*, \mathbf{v} : \Omega \rightarrow \mathbb{R}^3$, where \mathbf{m} and \mathbf{m}_* take values on the unit sphere and $\mathbf{m}_* \in W^{1,\infty}(\Omega)^3$:*

$$\begin{aligned} \|(\mathbf{P}(\mathbf{m}) - \mathbf{P}(\mathbf{m}_*))\mathbf{v}\|_{L^2(\Omega)^3} &\leq 2 \|\mathbf{v}\|_{L^\infty(\Omega)^3} \|\mathbf{m} - \mathbf{m}_*\|_{L^2(\Omega)^3}, \\ \|\nabla((\mathbf{P}(\mathbf{m}) - \mathbf{P}(\mathbf{m}_*))\mathbf{v})\|_{L^2(\Omega)^{3 \times 3}} &\leq 2 \|\mathbf{m}_*\|_{W^{1,\infty}(\Omega)^3} \|\mathbf{v}\|_{W^{1,\infty}(\Omega)^3} \|\mathbf{m} - \mathbf{m}_*\|_{L^2(\Omega)^3} \\ &\quad + 6 \|\mathbf{v}\|_{L^\infty(\Omega)^3} \|\nabla(\mathbf{m} - \mathbf{m}_*)\|_{L^2(\Omega)^{3 \times 3}}. \end{aligned}$$

Proof. Setting $\mathbf{e} = \mathbf{m} - \mathbf{m}_*$, we start by rewriting

$$(\mathbf{P}(\mathbf{m}) - \mathbf{P}(\mathbf{m}_*))\mathbf{v} = -(\mathbf{m}\mathbf{m}^T - \mathbf{m}_*\mathbf{m}_*^T)\mathbf{v} = -(\mathbf{m}\mathbf{e}^T + \mathbf{e}\mathbf{m}_*^T)\mathbf{v}.$$

The first inequality then follows immediately by taking the L^2 norm of both sides of the above equality, using the fact that \mathbf{m} and \mathbf{m}_* are of unit length. The second inequality is proved similarly, using the product rule

$$\begin{aligned} \partial_i(\mathbf{P}(\mathbf{m}) - \mathbf{P}(\mathbf{m}_*))\mathbf{v} &= -\partial_i(\mathbf{e}\mathbf{e}^T + \mathbf{m}_*\mathbf{e}^T + \mathbf{e}\mathbf{m}_*^T)\mathbf{v} \\ &= -(\partial_i\mathbf{e}\mathbf{e}^T + \mathbf{e}\partial_i\mathbf{e}^T + \partial_i\mathbf{m}_*\mathbf{e}^T + \mathbf{m}_*\partial_i\mathbf{e}^T + \partial_i\mathbf{e}\mathbf{m}_*^T + \mathbf{e}\partial_i\mathbf{m}_*^T)\mathbf{v} \\ &\quad + (\mathbf{m}\mathbf{e}^T + \mathbf{e}\mathbf{m}_*^T)\partial_i\mathbf{v}, \end{aligned}$$

the L^∞ bound of $\partial_i\mathbf{m}_*$, and the fact that $\|\mathbf{e}\|_{L^\infty} \leq \|\mathbf{m}\|_{L^\infty} + \|\mathbf{m}_*\|_{L^\infty} \leq 2$. \square

We have the following perturbation result.

Lemma 4.2. *Let $\mathbf{m}(t)$ and $\mathbf{m}_*(t)$ be solutions of unit length of (1.5) and (4.1), respectively, and suppose that, for $0 \leq t \leq \bar{t}$, we have*

$$(4.3) \quad \begin{aligned} \|\mathbf{m}_*(t)\|_{W^{1,\infty}(\Omega)^3} + \|\partial_t\mathbf{m}_*(t)\|_{W^{1,\infty}(\Omega)^3} &\leq R \\ \text{and } \|\Delta\mathbf{m}_*(t) + \mathbf{H}(t)\|_{L^\infty(\Omega)^3} &\leq K. \end{aligned}$$

Then, the error $\mathbf{e}(t) = \mathbf{m}(t) - \mathbf{m}_*(t)$ satisfies, for $0 \leq t \leq \bar{t}$,

$$(4.4) \quad \|\mathbf{e}(t)\|_{H^1(\Omega)^3}^2 \leq C \left(\|\mathbf{e}(0)\|_{H^1(\Omega)^3}^2 + \int_0^t \|\mathbf{d}(s)\|_{L^2(\Omega)^3}^2 ds \right),$$

where the constant C depends only on α, R, K , and \bar{t} .

Proof. Let us first assume that $\partial_t\mathbf{m}(t) \in H^1(\Omega)^3$ for all t . Following [21], we test in the error equation (4.2) with $\varphi = \mathbf{P}(\mathbf{m})\partial_t\mathbf{e} \in T(\mathbf{m})$. By the following argument, this test function is then indeed in $H^1(\Omega)^3$ and can be viewed as a perturbation of $\partial_t\mathbf{e}$:

$$\begin{aligned} \varphi &= \mathbf{P}(\mathbf{m})\partial_t\mathbf{e} = \mathbf{P}(\mathbf{m})\partial_t\mathbf{m} - \mathbf{P}(\mathbf{m})\partial_t\mathbf{m}_* \\ &= \mathbf{P}(\mathbf{m})\partial_t\mathbf{m} - \mathbf{P}(\mathbf{m}_*)\partial_t\mathbf{m}_* - (\mathbf{P}(\mathbf{m}) - \mathbf{P}(\mathbf{m}_*))\partial_t\mathbf{m}_* \\ &= \partial_t\mathbf{m} - \partial_t\mathbf{m}_* - (\mathbf{P}(\mathbf{m}) - \mathbf{P}(\mathbf{m}_*))\partial_t\mathbf{m}_*, \end{aligned}$$

and so we have

$$(4.5) \quad \varphi = \mathbf{P}(\bar{\mathbf{m}})\partial_t\mathbf{e} = \partial_t\mathbf{e} + \mathbf{q} \quad \text{with} \quad \mathbf{q} = -(\mathbf{P}(\mathbf{m}) - \mathbf{P}(\mathbf{m}_*))\partial_t\mathbf{m}_*.$$

By Lemma 4.1 and using (4.3) we have

$$(4.6) \quad \|\mathbf{q}\|_{L^2} \leq 2R\|\mathbf{e}\|_{L^2} \quad \text{and} \quad \|\nabla\mathbf{q}\|_{L^2} \leq CR\|\mathbf{e}\|_{H^1}.$$

Testing the error equation (4.2) with $\varphi = \partial_t\mathbf{e} + \mathbf{q}$, we obtain

$$\begin{aligned} \alpha(\partial_t\mathbf{e}, \partial_t\mathbf{e} + \mathbf{q}) + (\mathbf{e} \times \partial_t\mathbf{m}_*, \partial_t\mathbf{e} + \mathbf{q}) + (\mathbf{m} \times \partial_t\mathbf{e}, \partial_t\mathbf{e} + \mathbf{q}) \\ + (\nabla\mathbf{e}, \nabla(\partial_t\mathbf{e} + \mathbf{q})) = -(\mathbf{r}, \partial_t\mathbf{e} + \mathbf{q}), \end{aligned}$$

where, by (4.1) and Lemma 4.1 with (4.3), \mathbf{r} is bounded as

$$(4.7) \quad \begin{aligned} \|\mathbf{r}\|_{L^2} &\leq \|(\mathbf{P}(\mathbf{m}) - \mathbf{P}(\mathbf{m}_*))(\Delta\mathbf{m}_* + \mathbf{H})\|_{L^2} + \|\mathbf{d}\|_{L^2} \\ &\leq 2K\|\mathbf{e}\|_{L^2} + \|\mathbf{d}\|_{L^2}. \end{aligned}$$

By collecting terms, and using the fact that $(\mathbf{m} \times \partial_t\mathbf{e}, \partial_t\mathbf{e})$ vanishes, we altogether obtain

$$\begin{aligned} \alpha\|\partial_t\mathbf{e}\|_{L^2}^2 + \frac{1}{2}\frac{d}{dt}\|\nabla\mathbf{e}\|_{L^2}^2 &= -\alpha(\partial_t\mathbf{e}, \mathbf{q}) - (\mathbf{e} \times \partial_t\mathbf{m}_*, \partial_t\mathbf{e} + \mathbf{q}) - (\mathbf{m} \times \partial_t\mathbf{e}, \mathbf{q}) \\ &\quad - (\nabla\mathbf{e}, \nabla\mathbf{q}) - (\mathbf{r}, \partial_t\mathbf{e} + \mathbf{q}). \end{aligned}$$

For the right-hand side, the Cauchy–Schwarz inequality and $\|\mathbf{m}\|_{L^\infty} = 1$ yield

$$\begin{aligned} \alpha \|\partial_t \mathbf{e}\|_{L^2}^2 + \frac{1}{2} \frac{d}{dt} \|\nabla \mathbf{e}\|_{L^2}^2 &\leq \alpha \|\partial_t \mathbf{e}\|_{L^2} \|\mathbf{q}\|_{L^2} + R \|\mathbf{e}\|_{L^2} (\|\partial_t \mathbf{e}\|_{L^2} + \|\mathbf{q}\|_{L^2}) \\ &\quad + \|\partial_t \mathbf{e}\|_{L^2} \|\mathbf{q}\|_{L^2} + \|\nabla \mathbf{e}\|_{L^2} \|\nabla \mathbf{q}\|_{L^2} + \|\mathbf{r}\|_{L^2} (\|\partial_t \mathbf{e}\|_{L^2} + \|\mathbf{q}\|_{L^2}). \end{aligned}$$

Young’s inequality and absorptions, together with the bounds in (4.6) and (4.7), yield

$$\alpha \frac{1}{2} \|\partial_t \mathbf{e}\|_{L^2}^2 + \frac{1}{2} \frac{d}{dt} \|\nabla \mathbf{e}\|_{L^2}^2 \leq c \|\mathbf{e}\|_{H^1}^2 + c \|\mathbf{d}\|_{L^2}^2.$$

Here, we note that

$$\frac{1}{2} \frac{d}{dt} \|\mathbf{e}\|_{L^2}^2 = (\partial_t \mathbf{e}, \mathbf{e}) \leq \frac{1}{2} \|\partial_t \mathbf{e}\|_{L^2}^2 + \frac{1}{2} \|\mathbf{e}\|_{L^2}^2, \quad \text{so that} \quad \|\partial_t \mathbf{e}\|_{L^2}^2 \geq \frac{d}{dt} \|\mathbf{e}\|_{L^2}^2 - \|\mathbf{e}\|_{L^2}^2.$$

Combining these inequalities and integrating in time, we obtain

$$\|\mathbf{e}(t)\|_{H^1}^2 \leq c \|\mathbf{e}(0)\|_{H^1}^2 + c \int_0^t \|\mathbf{e}(s)\|_{H^1}^2 ds + c \int_0^t \|\mathbf{d}(s)\|_{L^2}^2 ds.$$

By Gronwall’s inequality, we then obtain the stated error bound.

Finally, if $\partial_t \mathbf{m}(t)$ is not in $H^1(\Omega)^3$ for some t , then a regularization and density argument, which we do not present here, yields the result, since the error bound does not depend on the H^1 norm of $\partial_t \mathbf{m}$. \square

5. ORTHOGONAL PROJECTION ONTO THE DISCRETE TANGENT SPACE

For consistency and stability of the full discretization, we need to study properties of the $L^2(\Omega)$ -orthogonal projection onto the discrete tangent space $T_h(\mathbf{m})$, which we denote by

$$\mathbf{P}_h(\mathbf{m}): V_h^3 \rightarrow T_h(\mathbf{m}).$$

We do not have an explicit expression for this projection, but the properties stated in Lemmas 5.1 to 5.3 will be used for proving consistency and stability. We recall that we consider a quasi-uniform, shape-regular family \mathcal{T}_h of triangulations with Lagrange finite elements of polynomial degree r .

The first lemma states that the projection $\mathbf{P}_h(\mathbf{m})$ approximates the orthogonal projection $\mathbf{P}(\mathbf{m}) = \mathbf{I} - \mathbf{m}\mathbf{m}^T$ onto the tangent space $T(\mathbf{m})$ with optimal order. It will be used in the consistency error analysis of Section 6.

Lemma 5.1. *For $\mathbf{m} \in W^{r+1,\infty}(\Omega)^3$ with $|\mathbf{m}| = 1$ almost everywhere we have*

$$\begin{aligned} \|(\mathbf{P}_h(\mathbf{m}) - \mathbf{P}(\mathbf{m}))\mathbf{v}\|_{L^2(\Omega)^3} &\leq Ch^{r+1} \|\mathbf{v}\|_{H^{r+1}(\Omega)^3}, \\ \|(\mathbf{P}_h(\mathbf{m}) - \mathbf{P}(\mathbf{m}))\mathbf{v}\|_{H^1(\Omega)^3} &\leq Ch^r \|\mathbf{v}\|_{H^{r+1}(\Omega)^3}, \end{aligned}$$

for all $\mathbf{v} \in H^{r+1}(\Omega)^3$, where C depends on a bound of $\|\mathbf{m}\|_{W^{r+1,\infty}(\Omega)^3}$.

The second lemma states that the projection $\mathbf{P}_h(\mathbf{m})$ has Lipschitz bounds of the same type as those of the orthogonal projection $\mathbf{P}(\mathbf{m})$ given in Lemma 4.1. It will be used in the stability analysis of Sections 7 and 8.

Lemma 5.2. *Let $\mathbf{m} \in W^{1,\infty}(\Omega)^3$ and $\widetilde{\mathbf{m}} \in H^1(\Omega)^3$ with $|\mathbf{m}| = |\widetilde{\mathbf{m}}| = 1$ almost everywhere and $\|\mathbf{m}\|_{W^{1,\infty}} \leq R$. There exist $C_R > 0$ and $h_R > 0$ such that for $h \leq h_R$, for all $\mathbf{v}_h \in V_h^3$,*

$$(i) \quad \|(\mathbf{P}_h(\mathbf{m}) - \mathbf{P}_h(\widetilde{\mathbf{m}}))\mathbf{v}_h\|_{L^2(\Omega)^3} \leq C_R \|\mathbf{m} - \widetilde{\mathbf{m}}\|_{L^p(\Omega)^3} \|\mathbf{v}_h\|_{L^q(\Omega)^3},$$

for $(p, q) \in \{(2, \infty), (\infty, 2)\}$, and

$$(ii) \quad \begin{aligned} \|(\mathbf{P}_h(\mathbf{m}) - \mathbf{P}_h(\widetilde{\mathbf{m}}))\mathbf{v}_h\|_{H^1(\Omega)^3} &\leq C_R \|\mathbf{m} - \widetilde{\mathbf{m}}\|_{H^1(\Omega)^3} \|\mathbf{v}_h\|_{L^\infty(\Omega)^3} \\ &\quad + C_R \|\mathbf{m} - \widetilde{\mathbf{m}}\|_{L^2(\Omega)^3} \|\mathbf{v}_h\|_{W^{1,\infty}(\Omega)^3}. \end{aligned}$$

The next lemma shows the $W^{s,p}$ -stability of the projection. It is actually used for $p = 2$ in the proof of Lemmas 5.1 and 5.2 and will be used for $p = 2$ in Section 6 and for $p = \infty$ in Sections 7 and 8.

Lemma 5.3. *There exists a constant depending only on $p \in [1, \infty]$ and the shape regularity of the mesh such that for all $\mathbf{m} \in W^{1,\infty}(\Omega)^3$ with $|\mathbf{m}| = 1$ almost everywhere,*

$$\|\mathbf{P}_h(\mathbf{m})\mathbf{v}_h\|_{W^{s,p}(\Omega)^3} \leq C \|\mathbf{m}\|_{W^{1,\infty}(\Omega)^3}^2 \|\mathbf{v}_h\|_{W^{s,p}(\Omega)^3}$$

for all $\mathbf{v}_h \in V_h^3$ and $s \in \{-1, 0, 1\}$.

These three lemmas will be proved in the course of this section, in which we formulate also three more lemmas that are of independent interest but will not be used in the following sections.

In the following, we use the dual norms

$$\|v\|_{W^{-1,q}} := \sup_{w \in W^{1,p}} \frac{(v, w)}{\|w\|_{W^{1,p}}} \quad \text{for } 1/p + 1/q = 1.$$

The space $W^{-1,1}(\Omega)$ is not the dual space of $W^{1,\infty}(\Omega)$ but rather defined as the closure of $L^2(\Omega)$ with respect to the norm $\|\cdot\|_{W^{-1,1}}$. We also recall that $\Pi_h : W^{s,p}(\Omega) \rightarrow W^{s,p}(\Omega)$ is uniformly bounded for $s \in \{0, 1\}$ and $p \in [1, \infty]$ (see, e.g., [20] for proofs in a much more general setting). By duality, we also obtain uniform boundedness for $s = -1$ and $p \in [1, \infty]$. A useful consequence is that for $v_h \in V_h$,

$$\begin{aligned} \|v_h\|_{W^{-1,q}} &= \sup_{w \in W^{1,p}} \frac{(v_h, \Pi_h w)}{\|w\|_{W^{1,p}}} \\ &\leq \sup_{w \in W^{1,p}} \frac{(v_h, \Pi_h w)}{\|\Pi_h w\|_{W^{1,p}}} \sup_{w \in W^{1,p}} \frac{\|\Pi_h w\|_{W^{1,p}}}{\|w\|_{W^{1,p}}} \lesssim \sup_{w_h \in V_h} \frac{(v_h, w_h)}{\|w_h\|_{W^{1,p}}}. \end{aligned}$$

Lemma 5.4. *There holds $\|v\|_{W^{s,p}(\Omega)} \simeq \sup_{w \in W^{-s,q}(\Omega)} \frac{(v, w)}{\|w\|_{W^{-s,q}(\Omega)}}$ with $1/p + 1/q = 1$ for $p \in [1, \infty]$ and $s \in \{-1, 0, 1\}$.*

Proof. The interesting case is $(s, p) = (1, \infty)$ since all other cases follow by duality. For $v \in W^{1,\infty}(\Omega)$, there exists a sequence of functions $\mathbf{q}_n \in C_0^\infty(\Omega)^3$ with $\|\mathbf{q}_n\|_{L^1} = 1$ such that

$$\|\nabla v\|_{L^\infty} = \lim_{n \rightarrow \infty} (\nabla v, \mathbf{q}_n) = \lim_{n \rightarrow \infty} -(v, \operatorname{div} \mathbf{q}_n) \leq \sup_{\mathbf{q} \in W^{1,1}} \frac{(v, \operatorname{div} \mathbf{q})}{\|\mathbf{q}\|_{L^1}}.$$

Moreover, there holds

$$\|\operatorname{div} \mathbf{q}\|_{W^{-1,1}} \leq \sup_{w \in W^{1,\infty}} \frac{(\mathbf{q}, \nabla w)}{\|\nabla w\|_{L^\infty}} \leq \|\mathbf{q}\|_{L^1}.$$

Combining the last two estimates shows

$$\|\nabla v\|_{L^\infty} \leq \sup_{w \in W^{-1,1}} \frac{(v, w)}{\|w\|_{W^{-1,1}}}.$$

Since

$$\|v\|_{L^\infty} = \sup_{w \in L^1} \frac{(v, w)}{\|w\|_{L^1}} \leq \sup_{w \in W^{-1,1}} \frac{(v, w)}{\|w\|_{W^{-1,1}}},$$

we conclude the proof. \square

Let the discrete normal space $N_h(\mathbf{m}) := V_h^3 \ominus T_h(\mathbf{m})$ be given as the L^2 -orthogonal complement of $T_h(\mathbf{m})$ in V_h^3 . We note that

$$(5.1) \quad N_h(\mathbf{m}) = \{\mathbf{\Pi}_h(\mathbf{m}\psi_h) : \psi_h \in V_h\}$$

by the definition of $T_h(\mathbf{m})$. The functions in the discrete normal space are bounded from below as follows.

Lemma 5.5. *For every $R > 0$, there exist $h_R > 0$ and $c > 0$ such that for all $\mathbf{m} \in W^{1,\infty}(\Omega)^3$ with $|\mathbf{m}| = 1$ almost everywhere and $\|\mathbf{m}\|_{W^{1,\infty}(\Omega)} \leq R$ and for all $h \leq h_R$,*

$$\|\mathbf{\Pi}_h(\mathbf{m}\psi_h)\|_{W^{s,p}(\Omega)^3} \geq c \|\psi_h\|_{W^{s,p}(\Omega)}$$

for all $\psi_h \in V_h$ and $(s, p) \in \{-1, 0, 1\} \times [1, \infty]$.

Proof. (a) We first prove the result for $s \in \{-1, 0\}$. Let $\mathbf{I}_h : C(\Omega) \rightarrow V_h^3$ denote the nodal interpolation operator and define $\mathbf{m}_h := \mathbf{I}_h \mathbf{m} \in V_h^3$.

There holds

$$\|\mathbf{\Pi}_h(\mathbf{m}_h\psi_h)\|_{L^p} \geq \|\mathbf{m}_h\psi_h\|_{L^p} - \|(\mathbf{I} - \mathbf{\Pi}_h)(\mathbf{m}_h\psi_h)\|_{L^p}.$$

Moreover, stability of $\mathbf{\Pi}_h$ in $L^p(\Omega)^3$, for $1 \leq p \leq \infty$, see [20], implies the estimate

$$\|(\mathbf{I} - \mathbf{\Pi}_h)(\mathbf{m}_h\psi_h)\|_{L^p} \leq (1 + C) \inf_{\mathbf{v}_h \in V_h^3} \|\mathbf{m}_h\psi_h - \mathbf{v}_h\|_{L^p}.$$

In turn, this implies

$$\begin{aligned} \|(\mathbf{I} - \mathbf{\Pi}_h)(\mathbf{m}_h\psi_h)\|_{L^p} &\lesssim \|(\mathbf{I} - \mathbf{I}_h)(\mathbf{m}_h\psi_h)\|_{L^p} \\ &= \left(\sum_{T \in \mathcal{T}_h} \|(\mathbf{I} - \mathbf{I}_h)(\mathbf{m}_h\psi_h)\|_{L^p(T)^3}^p \right)^{1/p}. \end{aligned}$$

For each element, the approximation properties of \mathbf{I}_h show

$$\begin{aligned} \|(\mathbf{I} - \mathbf{I}_h)(\mathbf{m}_h\psi_h)\|_{L^p(T)^3} &\lesssim h^{r+1} \|\nabla^{r+1}(\mathbf{m}_h\psi_h)\|_{L^p(T)^3} \\ &\leq h^{r+1} \sum_{i+j=r+1} \|\nabla^{\min\{i,r\}} \mathbf{m}_h\|_{L^\infty(T)^3} \|\nabla^{\min\{j,r\}} \psi_h\|_{L^p(T)^3}. \end{aligned}$$

Thus, multiple inverse estimates yield

$$\|(\mathbf{I} - \mathbf{I}_h)(\mathbf{m}_h\psi_h)\|_{L^p(T)^3} \lesssim h \|\mathbf{m}_h\|_{W^{1,\infty}} \|\psi_h\|_{L^p(T)^3}.$$

Moreover, we have

$$\|\mathbf{m}_h\psi_h\|_{L^p} \geq \|\mathbf{m}\psi_h\|_{L^p} - \|(\mathbf{m} - \mathbf{m}_h)\psi_h\|_{L^p} \geq \frac{1}{2} \|\psi_h\|_{L^p}$$

provided that $\|\mathbf{m} - \mathbf{m}_h\|_{L^\infty} \leq \frac{1}{2}$, which in view of

$$\|\mathbf{m} - \mathbf{m}_h\|_{L^\infty} = \|(\mathbf{I} - \mathbf{I}_h)\mathbf{m}\|_{L^\infty} \lesssim h \|\nabla \mathbf{m}\|_{L^\infty}$$

is satisfied for $h \leq h_R$ with a sufficiently small $h_R > 0$ that depends only on R . Altogether, this shows

$$\|\mathbf{\Pi}_h(\mathbf{m}_h\psi_h)\|_{L^p} \gtrsim \|\psi_h\|_{L^p}$$

for $h \leq h_R$. Similarly we estimate

$$\|\mathbf{\Pi}_h((\mathbf{m} - \mathbf{m}_h)\psi_h)\|_{L^p} \lesssim \|\mathbf{m} - \mathbf{m}_h\|_{L^\infty} \|\psi_h\|_{L^p} \lesssim h \|\nabla \mathbf{m}\|_{L^\infty} \|\psi_h\|_{L^p}.$$

Altogether, we obtain

$$\|\mathbf{\Pi}_h(\mathbf{m}\psi_h)\|_{L^p} \gtrsim \|\mathbf{\Pi}_h(\mathbf{m}_h\psi_h)\|_{L^p} - \|\mathbf{\Pi}_h((\mathbf{m}_h - \mathbf{m})\psi_h)\|_{L^p} \gtrsim \|\psi_h\|_{L^p}$$

for $h \leq h_R$. This concludes the proof for $s = 0$. Finally, for $s = -1$ we note that by using the result for $s = 0$ and an inverse inequality,

$$\begin{aligned} \|(\mathbf{I} - \mathbf{\Pi}_h)(\mathbf{m}\psi_h)\|_{W^{-1,p}} &\lesssim h\|\psi_h\|_{L^p} \\ &\lesssim h\|\mathbf{\Pi}_h(\mathbf{m}\psi_h)\|_{L^p} \lesssim \|\mathbf{\Pi}_h(\mathbf{m}\psi_h)\|_{W^{-1,p}}. \end{aligned}$$

Since $\|\mathbf{m}\psi_h\|_{W^{-1,p}} \gtrsim \|\mathbf{m}\|_{W^{1,\infty}}^{-1} \|\psi_h\|_{W^{-1,p}}$, this concludes the proof for $s \in \{-1, 0\}$.

(b) It remains to prove the result for $s = 1$. Note that the result follows from duality if we show

$$(5.2) \quad \|\mathbf{I}_h(\mathbf{m} \cdot \mathbf{w}_h)\|_{W^{-1,q}} \gtrsim \|\mathbf{w}_h\|_{W^{-1,q}}$$

for all $\mathbf{w}_h \in N_h(\mathbf{m})$. To see this, note that (5.2) implies

$$\begin{aligned} \|\mathbf{\Pi}_h(\mathbf{m}\psi_h)\|_{W^{1,p}} &\geq \sup_{\mathbf{w}_h \in N_h(\mathbf{m})} \frac{(\psi_h, \mathbf{I}_h(\mathbf{m} \cdot \mathbf{w}_h))}{\|\mathbf{w}_h\|_{W^{-1,q}}} \\ &\gtrsim \sup_{\mathbf{w}_h \in N_h(\mathbf{m})} \frac{(\psi_h, \mathbf{I}_h(\mathbf{m} \cdot \mathbf{w}_h))}{\|\mathbf{I}_h(\mathbf{m} \cdot \mathbf{w}_h)\|_{W^{-1,q}}} = \sup_{\omega_h \in V_h} \frac{(\psi_h, \omega_h)}{\|\omega_h\|_{W^{-1,q}}} \simeq \|\psi_h\|_{W^{1,p}}, \end{aligned}$$

where we used in the second to last equality that part (a) for $s = 0$ already shows that $\dim(N_h(\mathbf{m})) = \dim V_h$ and since (5.2) implies that the map $N_h(\mathbf{m}) \rightarrow V_h$, $\mathbf{w}_h \mapsto \mathbf{I}_h(\mathbf{m} \cdot \mathbf{w}_h)$ is injective, it is already bijective. It remains to prove (5.2). To that end, we first show for $\mathbf{w}_h = \mathbf{\Pi}_h(\mathbf{m}\omega_h) \in N_h(\mathbf{m})$ for some $\omega_h \in V_h$, using the reverse triangle inequality, that

$$\begin{aligned} \|\mathbf{m} \cdot \mathbf{w}_h\|_{W^{-1,q}} &\geq \|\omega_h\|_{W^{-1,q}} - \|\mathbf{m} \cdot (\mathbf{I} - \mathbf{\Pi}_h)(\mathbf{m}\omega_h)\|_{W^{-1,q}} \\ &\gtrsim \|\mathbf{m}\|_{W^{1,\infty}}^{-1} \|\omega_h\|_{W^{-1,q}} - \|\mathbf{m} \cdot (\mathbf{I} - \mathbf{\Pi}_h)(\mathbf{m}\omega_h)\|_{W^{-1,q}}. \end{aligned}$$

With $\mathbf{m}_h := \mathbf{I}_h(\mathbf{m}) \in V_h^3$, the last term satisfies

$$\begin{aligned} \|\mathbf{m} \cdot (\mathbf{I} - \mathbf{\Pi}_h)(\mathbf{m}\omega_h)\|_{W^{-1,q}} &\lesssim h\|\mathbf{m}\|_{W^{1,\infty}} \|(\mathbf{I} - \mathbf{\Pi}_h)(\mathbf{m}\omega_h)\|_{L^q} \\ &\lesssim h\|\mathbf{m}\|_{W^{1,\infty}} (\|\mathbf{m} - \mathbf{m}_h\|_{L^\infty} \|\omega_h\|_{L^q} + h\|\mathbf{m}_h\|_{W^{1,\infty}} \|\omega_h\|_{L^q}), \end{aligned}$$

where we used the same arguments as in the proof of part (a) to get the estimate $\|(\mathbf{I} - \mathbf{\Pi}_h)(\mathbf{m}_h\omega_h)\|_{L^q} \lesssim h\|\mathbf{m}_h\|_{W^{1,\infty}} \|\omega_h\|_{L^q}$. The fact $\|\mathbf{m}_h\|_{W^{1,\infty}} \lesssim \|\mathbf{m}\|_{W^{1,\infty}}$, the approximation property $\|\mathbf{m} - \mathbf{m}_h\|_{L^\infty} \lesssim h\|\mathbf{m}\|_{W^{1,\infty}}$, and an inverse inequality conclude

$$(5.3) \quad \|\mathbf{m} \cdot \mathbf{w}_h\|_{W^{-1,q}} \gtrsim \|\mathbf{w}_h\|_{W^{-1,q}}$$

with (hidden) constants depending only on $\|\mathbf{m}\|_{W^{1,\infty}}$ and shape regularity of the mesh.

To prove (5.2), it remains to bound the left-hand side above by $\|\mathbf{I}_h(\mathbf{m} \cdot \mathbf{w}_h)\|_{W^{-1,q}}$. To that end, we note

$$\begin{aligned} \|(I - \mathbf{I}_h)(\mathbf{m} \cdot \mathbf{w}_h)\|_{W^{-1,q}} &\lesssim h\|\mathbf{w}_h\|_{L^q} = h \sup_{\mathbf{v} \in L^p} \frac{(\mathbf{w}_h, \mathbf{v})}{\|\mathbf{v}\|_{L^p}} \\ &\lesssim h \sup_{\mathbf{v} \in N_h(\mathbf{m})} \frac{(\mathbf{w}_h, \mathbf{v})}{\|\mathbf{v}\|_{L^p}} = h \sup_{\mathbf{v} \in V_h} \frac{(\mathbf{I}_h(\mathbf{m} \cdot \mathbf{w}_h), \mathbf{v})}{\|\mathbf{\Pi}_h(\mathbf{m}\mathbf{v})\|_{L^p}} \lesssim h\|\mathbf{I}_h(\mathbf{m} \cdot \mathbf{w}_h)\|_{L^q}, \end{aligned}$$

where we used part (a) for $s = 0$ for the last inequality. An inverse inequality and the combination with (5.3) imply (5.2) for $h > 0$ sufficiently small in terms of $\|\mathbf{m}\|_{W^{1,\infty}}^{-1}$. This concludes the proof. \square

Lemma 5.6. *Define the matrix $M \in \mathbb{R}^{N_h \times N_h}$, where N_h denotes the dimension of V_h , by $M_{ij} := h^{-3}(\mathbf{\Pi}_h(\mathbf{m}\phi_j), \mathbf{\Pi}_h(\mathbf{m}\phi_i))$. Under the assumptions of Lemma 5.5, there exists $C > 0$ such that for $h \leq h_R$,*

$$\|M\|_p + \|M^{-1}\|_p \leq C \quad \text{for } 1 \leq p \leq \infty,$$

where C depends only on the shape regularity.

Proof. Lemma 5.5 shows for $x \in \mathbb{R}^{N_h}$

$$(5.4) \quad Mx \cdot x = h^{-3} \|\mathbf{\Pi}_h(\mathbf{m} \sum_{i=1}^{N_h} x_i \phi_i)\|_{L^2}^2 \gtrsim h^{-3} \left\| \sum_{i=1}^{N_h} x_i \phi_i \right\|_{L^2}^2 \simeq |x|^2,$$

where $|\cdot|$ denotes the Euclidean norm on \mathbb{R}^{N_h} . Let $d(i, j) := \text{dist}(z_i, z_j)h^{-3}$ denote the metric which (approximately) measures the number of elements between the supports of ϕ_i and ϕ_j , corresponding to the nodes z_i and z_j , and let $B_d(z)$ denote the corresponding ball. In the following, we use a localization property of the L^2 -projection, i.e., there exist $a, b > 0$ such that for all $\ell \in \mathbb{N}$,

$$(5.5) \quad \|\mathbf{\Pi}_h(\mathbf{m}\phi_i)\|_{L^2(\Omega \setminus B_\ell(z_i))} \leq ae^{-b\ell} \|\mathbf{m}\phi_i\|_{L^2}.$$

The proof of this bound is essentially contained in the proof of [9, Lemma 3.1]. Since we use the very same arguments below, we briefly recall the strategy: First, one observes that the mass matrix $\widetilde{M} \in \mathbb{R}^{N_h \times N_h}$ with entries $\widetilde{M}_{ij} := h^{-3}(\phi_j, \phi_i)$ is banded in the sense that $d(i, j) \gtrsim 1$ implies $\widetilde{M}_{ij} = 0$, and it satisfies $\widetilde{M}x \cdot x \gtrsim |x|^2$. As shown below, this implies that the inverse matrix \widetilde{M}^{-1} satisfies $|(\widetilde{M}^{-1})_{ij}| \lesssim e^{-bd(i,j)}$ for some $b > 0$ independent of $h > 0$. Note that each entry of the vector field $\mathbf{\Pi}_h(\mathbf{m}\phi_i) \in V_h^3$ can be represented by $\sum_{j=1}^{N_h} x_{k,j} \phi_j$, $k = 1, 2, 3$, and is computed by solving $\widetilde{M}x_k = g_k \in \mathbb{R}^{N_h}$ with $\mathbf{m} = (m_1, m_2, m_3)^T$ and $g_{k,j} := (m_k \phi_i, \phi_j)$. Hence, the exponential decay of \widetilde{M}^{-1} directly implies (5.5).

From the decay property (5.5), we immediately obtain

$$|M_{ij}| \leq \widetilde{a}e^{-\widetilde{b}d(i,j)}$$

for all $1 \leq i, j \leq N_h$ and some $\widetilde{a}, \widetilde{b} > 0$. This already proves $\|M\|_p \leq C$. We follow the arguments from [28] to show that also M^{-1} decays exponentially. To that end, note that (5.4) implies the existence of $c > 0$ such that $\|I - cM\|_2 =: q < 1$ and hence

$$(5.6) \quad M^{-1} = c(I - (I - cM))^{-1} = c \sum_{k=0}^{\infty} (I - cM)^k.$$

Clearly, $I - cM$ inherits the decay properties from M and therefore

$$\begin{aligned} |((I - cM)^{k+1})_{ij}| &\leq \widetilde{a}^{k+1} \sum_{r_1, \dots, r_k=1}^{N_h} e^{-\widetilde{b}(d(i, r_1) + \dots + d(r_k, j))} \\ &\leq \widetilde{a}^{k+1} \left(\max_{s=1, \dots, N_h} \sum_{r=1}^{N_h} e^{-\widetilde{b}d(s, r)/2} \right)^k e^{-\widetilde{b}d(i, j)/2}. \end{aligned}$$

The value of $\max_{s=1,\dots,N_h} \sum_{r=1}^{N_h} e^{-\tilde{b}d(s,r)/2}$ depends only on the shape regularity of the triangulation and on \tilde{b} , but is independent of h (it just depends on the number of elements contained in an annulus of thickness $\approx h$). This implies the existence of $\tilde{c} \geq 1$ such that

$$|((I - cM)^{k+1})_{ij}| \leq \min\{q^{k+1}, \tilde{c}^{k+1} e^{-\tilde{b}d(i,j)/2}\}.$$

Thus, for $\tilde{c}^{k+1} \leq e^{\tilde{b}d(i,j)/4}$, we have $|((I - cM)^{k+1})_{ij}| \leq e^{-\tilde{b}d(i,j)/4}$, whereas for $\tilde{c}^{k+1} > e^{\tilde{b}d(i,j)/4}$, we have $|((I - cM)^{k+1})_{ij}| \leq q^{k+1} < q^{\tilde{b}d(i,j)/(4 \log(\tilde{c}))}$. Altogether, we find some $\tilde{b} > 0$ (we reuse the symbol), independent of h such that

$$|((I - cM)^{k+1})_{ij}| \leq q^{(k+1)/2} |((I - cM)^{k+1})_{ij}|^{1/2} \lesssim q^{(k+1)/2} e^{-\tilde{b}d(i,j)}.$$

Plugging this into (5.6), we obtain

$$|(M^{-1})_{ij}| \lesssim \sum_{k=0}^{\infty} q^{(k+1)/2} e^{-\tilde{b}d(i,j)} \lesssim e^{-\tilde{b}d(i,j)}.$$

This yields the stated result. \square

We are now in a position to prove Lemma 5.3.

Proof of Lemma 5.3. (a) We first consider the case $s = 0$. In view of (5.1), we write $(\mathbf{I} - \mathbf{P}_h(\mathbf{m}))\mathbf{v}_h \in N_h(\mathbf{m})$ as

$$(\mathbf{I} - \mathbf{P}_h(\mathbf{m}))\mathbf{v}_h = h^{-3/2} \sum_{i=1}^{N_h} x_i \mathbf{\Pi}_h(\mathbf{m}\phi_i)$$

for some coefficient vector $x \in \mathbb{R}^{N_h}$ and let $b_i := h^{-3/2}(\mathbf{v}_h, \mathbf{m}\phi_i)$ for $i = 1, \dots, N_h$. Then, there holds $Mx = b$ with the matrix M from Lemma 5.6. This lemma and the L^p -stability of the L^2 -orthogonal projection $\mathbf{\Pi}_h$ [20] imply that for $p \in [1, \infty]$,

$$\begin{aligned} \|(\mathbf{I} - \mathbf{P}_h(\mathbf{m}))\mathbf{v}_h\|_{L^p} &= \|\mathbf{\Pi}_h h^{-3/2} \sum_{i=1}^{N_h} x_i \mathbf{m}\phi_i\|_{L^p} \lesssim \|h^{-3/2} \sum_{i=1}^{N_h} x_i \mathbf{m}\phi_i\|_{L^p} \\ &\lesssim h^{-3/2} \left(\sum_{i=1}^{N_h} h^3 |x_i|^p \right)^{1/p} = h^{3/p-3/2} |x|_p = h^{3/p-3/2} |M^{-1}b|_p \lesssim h^{3/p-3/2} |b|_p. \end{aligned}$$

With $|b_i| \leq h^{-3/2} \|\mathbf{v}_h\|_{L^p(\text{supp}(\phi_i))} h^{3(1-1/p)} = \|\mathbf{v}_h\|_{L^p(\text{supp}(\phi_i))} h^{3/2-3/p}$, this shows

$$\|\mathbf{P}_h(\mathbf{m})\mathbf{v}_h\|_{L^p} \lesssim \|\mathbf{v}_h\|_{L^p}.$$

(b) We now turn to the cases $s = \pm 1$. Define the operator

$$\tilde{\mathbf{P}}_h^\perp(\mathbf{m})\mathbf{v}_h := \mathbf{\Pi}_h(\mathbf{m}\mathbf{\Pi}_h(\mathbf{m} \cdot \mathbf{v}_h))$$

and note that $\tilde{\mathbf{P}}_h^\perp(\mathbf{m})\mathbf{v}_h \in N_h(\mathbf{m})$ as well as $\ker \tilde{\mathbf{P}}_h^\perp(\mathbf{m}) = T_h(\mathbf{m})$ (due to Lemma 5.5). However, $\tilde{\mathbf{P}}_h^\perp(\mathbf{m})$ is no projection. We observe for $\mathbf{v}_h = \mathbf{\Pi}_h(\mathbf{m}\psi_h) \in N_h(\mathbf{m})$ that

$$\begin{aligned} \|(\mathbf{I} - \tilde{\mathbf{P}}_h^\perp(\mathbf{m}))\mathbf{v}_h\|_{W^{-1,p}} &= \|\mathbf{\Pi}_h \mathbf{m}\psi_h - \mathbf{\Pi}_h(\mathbf{m}\mathbf{\Pi}_h(\mathbf{m} \cdot \mathbf{\Pi}_h(\mathbf{m}\psi_h)))\|_{W^{-1,p}} \\ &\lesssim \|\mathbf{m}\|_{W^{1,\infty}} \|\psi_h - \mathbf{m} \cdot \mathbf{\Pi}_h(\mathbf{m}\psi_h)\|_{W^{-1,p}} \\ &= \|\mathbf{m}\|_{W^{1,\infty}}^2 \|(\mathbf{I} - \mathbf{\Pi}_h)(\mathbf{m}\psi_h)\|_{W^{-1,p}} \\ &\lesssim \|\mathbf{m}\|_{W^{1,\infty}}^2 h \|\psi_h\|_{L^p}. \end{aligned}$$

With Lemma 5.5 we conclude

$$\|(\mathbf{I} - \tilde{\mathbf{P}}_h^\perp(\mathbf{m}))\mathbf{v}_h\|_{W^{-1,p}} \lesssim \|\mathbf{m}\|_{W^{1,\infty}}^2 h \|\mathbf{v}_h\|_{L^p}.$$

Since $\tilde{\mathbf{P}}_h^\perp(\mathbf{m})\mathbf{P}_h(\mathbf{m}) = 0$ by definition of $T_h(\mathbf{m})$, we obtain with part (a) and an inverse inequality that for all $\mathbf{v}_h \in V_h^3$,

$$\begin{aligned} \|(\mathbf{I} - \mathbf{P}_h(\mathbf{m}) - \tilde{\mathbf{P}}_h^\perp(\mathbf{m}))\mathbf{v}_h\|_{W^{-1,p}} &= \|(\mathbf{I} - \tilde{\mathbf{P}}_h^\perp(\mathbf{m}))(\mathbf{I} - \mathbf{P}_h(\mathbf{m}))\mathbf{v}_h\|_{W^{-1,p}} \\ &\lesssim \|\mathbf{m}\|_{W^{1,\infty}}^2 h \|(\mathbf{I} - \mathbf{P}_h(\mathbf{m}))\mathbf{v}_h\|_{L^p} \\ &\lesssim \|\mathbf{m}\|_{W^{1,\infty}}^2 h \|\mathbf{v}_h\|_{L^p} \\ &\lesssim \|\mathbf{m}\|_{W^{1,\infty}}^2 \|\mathbf{v}_h\|_{W^{-1,p}}. \end{aligned}$$

The $W^{-1,p}(\Omega)$ -stability of Π_h implies $\|\tilde{\mathbf{P}}_h^\perp(\mathbf{m})\mathbf{v}_h\|_{W^{-1,p}} \lesssim \|\mathbf{m}\|_{W^{1,\infty}}^2 \|\mathbf{v}_h\|_{W^{-1,p}}$ and the triangle inequality concludes the proof for $s = -1$. The case $s = 1$ follows by duality. \square

Proof of Lemma 5.2. (a) ($s = 0$) The projection $\mathbf{v}_h := \mathbf{P}_h(\mathbf{m})\mathbf{v}$ is given by the equation

$$(\mathbf{v}_h, \boldsymbol{\varphi}_h) = (\mathbf{v}, \boldsymbol{\varphi}_h) \quad \forall \boldsymbol{\varphi}_h \in T_h(\mathbf{m}),$$

which in view of the definition of $T_h(\mathbf{m})$ is equivalent to the solution of the saddle point problem (with the Lagrange multiplier $\lambda_h \in V_h$)

$$\begin{aligned} (\mathbf{v}_h, \mathbf{w}_h) + (\mathbf{m} \cdot \mathbf{w}_h, \lambda_h) &= (\mathbf{v}, \mathbf{w}_h) \quad \forall \mathbf{w}_h \in V_h^3, \\ (\mathbf{m} \cdot \mathbf{v}_h, \mu_h) &= 0 \quad \forall \mu_h \in V_h. \end{aligned}$$

By the first equation, we also obtain the identity $\mathbf{\Pi}_h(\mathbf{m}\lambda_h) = (\mathbf{I} - \mathbf{P}_h(\mathbf{m}))\mathbf{v}_h$, which will be used below. Furthermore, $\tilde{\mathbf{v}}_h := \mathbf{P}_h(\tilde{\mathbf{m}})\mathbf{v}$ is given by the same system with $\tilde{\mathbf{m}}$ in place of \mathbf{m} , yielding a corresponding Lagrange multiplier $\tilde{\lambda}_h$. Hence, the differences $\mathbf{e}_h := \mathbf{v}_h - \tilde{\mathbf{v}}_h$ and $\delta_h := \lambda_h - \tilde{\lambda}_h$ satisfy

$$\begin{aligned} (\mathbf{e}_h, \mathbf{w}_h) + (\mathbf{m} \cdot \mathbf{w}_h, \delta_h) &= -(\mathbf{w}_h, (\mathbf{m} - \tilde{\mathbf{m}})\tilde{\lambda}_h) \quad \forall \mathbf{w}_h \in V_h^3, \\ (\mathbf{m} \cdot \mathbf{e}_h, \mu_h) &= -((\mathbf{m} - \tilde{\mathbf{m}}) \cdot \tilde{\mathbf{v}}_h, \mu_h) \quad \forall \mu_h \in V_h. \end{aligned}$$

The classical results on saddle-point problems (see [13, Proposition 2.1]) require two inf-sup conditions to be satisfied. First,

$$\inf_{q_h \in V_h} \sup_{\mathbf{v}_h \in V_h^3} \frac{(\mathbf{m} \cdot \mathbf{v}_h, q_h)}{\|\mathbf{v}_h\|_{H^s} \|q_h\|_{H^{-s}}} > 0$$

holds uniformly in h due to Lemma 5.5. Second,

$$\inf_{\mathbf{w}_h \in T_h(\mathbf{m})} \sup_{\mathbf{v}_h \in T_h(\mathbf{m})} \frac{(\mathbf{v}_h, \mathbf{w}_h)}{\|\mathbf{v}_h\|_{H^s} \|\mathbf{w}_h\|_{H^{-s}}} > 0$$

holds uniformly in h due to the stability estimates from Lemma 5.3 (noting that $\mathbf{v}_h = \mathbf{P}_h(\mathbf{m})\mathbf{v}_h$ and $\mathbf{w}_h = \mathbf{P}_h(\mathbf{m})\mathbf{w}_h$ for $\mathbf{v}_h, \mathbf{w}_h \in T_h(\mathbf{m})$). For the above saddle-point problems, these bounds for $s = 0$ give us an L^2 bound for $\mathbf{e}_h = \mathbf{P}_h(\mathbf{m})\mathbf{v} - \mathbf{P}_h(\tilde{\mathbf{m}})\mathbf{v}$: From [13] we obtain

$$\|\tilde{\mathbf{v}}_h\|_{L^2} + \|\tilde{\lambda}_h\|_{L^2} \lesssim \|\mathbf{v}\|_{L^2}$$

and

$$\|\mathbf{e}_h\|_{L^2} + \|\delta_h\|_{L^2} \lesssim \|(\mathbf{m} - \tilde{\mathbf{m}})\tilde{\lambda}_h\|_{L^2} + \|(\mathbf{m} - \tilde{\mathbf{m}}) \cdot \tilde{\mathbf{v}}_h\|_{L^2}.$$

With the stability from Lemma 5.3 and Lemma 5.5, we also obtain

$$\|\tilde{\mathbf{v}}_h\|_{L^\infty} + \|\tilde{\lambda}_h\|_{L^\infty} \lesssim \|\mathbf{P}_h(\tilde{\mathbf{m}})\mathbf{v}\|_{L^\infty} + \|(\mathbf{I} - \mathbf{P}_h(\tilde{\mathbf{m}}))\mathbf{v}\|_{L^\infty} \lesssim \|\mathbf{v}\|_{L^\infty}.$$

Altogether, this implies

$$\|\mathbf{e}_h\|_{L^2} + \|\delta_h\|_{L^2} \lesssim \|\mathbf{m} - \tilde{\mathbf{m}}\|_{L^p} \|\mathbf{v}\|_{L^q}$$

for $(p, q) \in \{(2, \infty), (\infty, 2)\}$.

(b) ($s = 1$) For the $H^1(\Omega)$ -estimate, we introduce the Riesz mapping J_h between $V_h \subset H^1(\Omega)$ and its dual $V_h \subset H^1(\Omega)'$, i.e., the isometry defined by

$$(v_h, J_h \psi_h)_{H^1} = \langle v_h, \psi_h \rangle \quad \forall v_h \in V_h, \psi_h \in V_h.$$

By $\mathbf{J}_h := \mathbf{I} \otimes J_h$ we denote the corresponding vector-valued mapping on V_h^3 . We consider the bilinear form on $V_h^3 \times V_h^3$ defined by

$$a_h(\mathbf{v}_h, \mathbf{w}_h) = \langle \mathbf{v}_h, \mathbf{J}_h^{-1} \mathbf{w}_h \rangle, \quad \mathbf{v}_h, \mathbf{w}_h \in V_h^3,$$

and reformulate the saddle-point problem for $(\mathbf{v}_h, \lambda_h) \in V_h^3 \times V_h \subset H^1(\Omega)^3 \times H^1(\Omega)'$ as

$$\begin{aligned} a_h(\mathbf{v}_h, \mathbf{w}_h) + \langle \mathbf{m} \cdot \mathbf{J}_h^{-1} \mathbf{w}_h, \lambda_h \rangle &= a(\mathbf{v}, \mathbf{w}_h) \quad \forall \mathbf{w}_h \in V_h^3, \\ \langle \mathbf{m} \cdot \mathbf{v}_h, J_h^{-1} \mu_h \rangle &= 0 \quad \forall \mu_h \in V_h. \end{aligned}$$

As in the case $s = 0$ (algebraically it is the same system), we have $\mathbf{v}_h = \mathbf{P}_h(\mathbf{m})\mathbf{v}$ and $\mathbf{\Pi}_h(\mathbf{m}\lambda_h) = (\mathbf{I} - \mathbf{P}_h(\mathbf{m}))\mathbf{v}$. The system for $\mathbf{e}_h = \mathbf{v}_h - \tilde{\mathbf{v}}_h$ and $\delta_h = \lambda_h - \tilde{\lambda}_h$ reads

$$\begin{aligned} a_h(\mathbf{e}_h, \mathbf{w}_h) + \langle \mathbf{m} \cdot \mathbf{J}_h^{-1} \mathbf{w}_h, \delta_h \rangle &= -\langle (\mathbf{m} - \tilde{\mathbf{m}}) \cdot \mathbf{J}_h^{-1} \mathbf{w}_h, \tilde{\lambda}_h \rangle \quad \forall \mathbf{w}_h \in V_h^3, \\ \langle \mathbf{m} \cdot \mathbf{e}_h, J_h^{-1} \mu_h \rangle &= -\langle (\mathbf{m} - \tilde{\mathbf{m}}) \cdot \tilde{\mathbf{v}}_h, J_h^{-1} \mu_h \rangle \quad \forall \mu_h \in V_h. \end{aligned}$$

The above inf-sup bounds for $s = 1$ and $s = -1$ are precisely the inf-sup conditions that need to be satisfied for these generalized saddle-point problems (see [15, Theorem 2.1]), whose right-hand sides are bounded by

$$|a_h(\mathbf{v}, \mathbf{w}_h)| \leq \|\mathbf{v}\|_{H^1} \|\mathbf{J}_h^{-1} \mathbf{w}_h\|_{H^{-1}} \simeq \|\mathbf{v}\|_{H^1} \|\mathbf{w}_h\|_{H^1}$$

and

$$\begin{aligned} |\langle (\mathbf{m} - \tilde{\mathbf{m}}) \cdot \mathbf{J}_h^{-1} \mathbf{w}_h, \tilde{\lambda}_h \rangle| &\lesssim \|(\mathbf{m} - \tilde{\mathbf{m}}) \tilde{\lambda}_h\|_{H^1} \|\mathbf{w}_h\|_{H^1}, \\ |\langle (\mathbf{m} - \tilde{\mathbf{m}}) \cdot \tilde{\mathbf{v}}_h, J_h^{-1} \mu_h \rangle| &\leq \|(\mathbf{m} - \tilde{\mathbf{m}}) \cdot \tilde{\mathbf{v}}_h\|_{H^1} \|\mu_h\|_{H^1}. \end{aligned}$$

As in the case $s = 0$, we obtain from Lemma 5.3 and Lemma 5.5 that

$$\begin{aligned} \|\tilde{\mathbf{v}}_h\|_{W^{1,\infty}} + \|\tilde{\lambda}_h\|_{W^{1,\infty}} &\lesssim \|\mathbf{P}_h(\tilde{\mathbf{m}})\mathbf{v}\|_{W^{1,\infty}} + \|(\mathbf{I} - \mathbf{P}_h(\tilde{\mathbf{m}}))\mathbf{v}\|_{W^{1,\infty}} \\ &\lesssim \|\mathbf{v}\|_{W^{1,\infty}}. \end{aligned}$$

Hence, we obtain from [15, Theorem 2.1], for $(p, q) \in \{(2, \infty), (\infty, 2)\}$,

$$\begin{aligned} \|\mathbf{e}_h\|_{H^1} &\lesssim \|(\mathbf{m} - \tilde{\mathbf{m}}) \tilde{\lambda}_h\|_{H^1} + \|(\mathbf{m} - \tilde{\mathbf{m}}) \cdot \tilde{\mathbf{v}}_h\|_{H^1} \\ &\lesssim \sum_{s'=0}^1 \left(\|\mathbf{m} - \tilde{\mathbf{m}}\|_{H^1} \|\tilde{\lambda}_h\|_{W^{1-s',q}} + \|\mathbf{m} - \tilde{\mathbf{m}}\|_{W^{s',p}} \|\tilde{\mathbf{v}}_h\|_{W^{1-s',q}} \right) \\ &\lesssim \sum_{s'=0}^1 \|\mathbf{m} - \tilde{\mathbf{m}}\|_{W^{s',p}} \|\mathbf{v}\|_{W^{1-s',q}}. \end{aligned}$$

This implies the $H^1(\Omega)^3$ estimate and hence concludes the proof. \square

Proof of Lemma 5.1. Since $\mathbf{P}_h(\mathbf{m})\mathbf{v}$ is the Galerkin approximation of the saddle point problem for $\mathbf{P}(\mathbf{m})\mathbf{v}$ (as in the previous proof), the Céa lemma for saddle-point problems (see [13, Theorem 2.1]) shows in L^2

$$\begin{aligned} & \|(\mathbf{P}_h(\mathbf{m}) - \mathbf{P}(\mathbf{m}))\mathbf{v}\|_{L^2} \\ & \lesssim \inf_{(\mathbf{w}_h, \mu_h) \in V_h^3 \times V_h} \left(\|\mathbf{P}(\mathbf{m})\mathbf{v} - \mathbf{w}_h\|_{L^2} + \|\mathbf{m} \cdot \mathbf{v} - \mu_h\|_{L^2} \right) \\ & \lesssim h^{r+1} \|\mathbf{m}\|_{W^{r+1, \infty}} \|\mathbf{v}\|_{H^{r+1}} \end{aligned}$$

and similarly in H^1 , using [15, Theorem 2.1],

$$\begin{aligned} & \|(\mathbf{P}_h(\mathbf{m}) - \mathbf{P}(\mathbf{m}))\mathbf{v}\|_{H^1} \\ & \lesssim \inf_{(\mathbf{w}_h, \mu_h) \in V_h^3 \times V_h} \left(\|\mathbf{P}(\mathbf{m})\mathbf{v} - \mathbf{w}_h\|_{H^1} + \|\mathbf{m} \cdot \mathbf{v} - \mu_h\|_{H^1} \right) \\ & \lesssim h^r \|\mathbf{m}\|_{W^{r+1, \infty}} \|\mathbf{v}\|_{H^{r+1}}. \end{aligned}$$

This concludes the proof. \square

6. CONSISTENCY ERROR AND ERROR EQUATION

To study the consistency errors, we find it instructive to separate the issues of consistency for the time and space discretizations. Therefore, we first show defect estimates for the semidiscretization in time, and then turn to the full discretization.

6.1. Consistency error of the semi-discretization in time. The order of both the fully implicit k -step BDF method, described by the coefficients $\delta_0, \dots, \delta_k$ and 1, and the explicit k -step BDF method, that is the method described by the coefficients $\delta_0, \dots, \delta_k$ and $\gamma_0, \dots, \gamma_{k-1}$, is k , i.e.,

$$(6.1) \quad \sum_{i=0}^k (k-i)^\ell \delta_i = \ell k^{\ell-1} = \ell \sum_{i=0}^{k-1} (k-i-1)^{\ell-1} \gamma_i, \quad \ell = 0, 1, \dots, k.$$

We first rewrite the linearly implicit k -step BDF method (2.3) in strong form,

$$(6.2) \quad \alpha \dot{\mathbf{m}}^n + \widehat{\mathbf{m}}^n \times \dot{\mathbf{m}}^n = \mathbf{P}(\widehat{\mathbf{m}}^n)(\Delta \mathbf{m}^n + \mathbf{H}(t_n)),$$

with Neumann boundary conditions.

The consistency error \mathbf{d}^n of the linearly implicit k -step BDF method (6.2) for the solution \mathbf{m} is the defect by which the exact solution misses satisfying (6.2), and is given by

$$(6.3) \quad \mathbf{d}^n = \alpha \dot{\mathbf{m}}_*^n + \widehat{\mathbf{m}}_*^n \times \dot{\mathbf{m}}_*^n - \mathbf{P}(\widehat{\mathbf{m}}_*^n)(\Delta \mathbf{m}_*^n + \mathbf{H}(t_n))$$

for $n = k, \dots, N$, where we use the notation $\mathbf{m}_*^n = \mathbf{m}(t_n)$ and

$$(6.4) \quad \begin{aligned} \widehat{\mathbf{m}}_*^n &= \sum_{j=0}^{k-1} \gamma_j \mathbf{m}_*^{n-j-1} \Big/ \left| \sum_{j=0}^{k-1} \gamma_j \mathbf{m}_*^{n-j-1} \right|, \\ \dot{\mathbf{m}}_*^n &= \mathbf{P}(\widehat{\mathbf{m}}_*^n) \frac{1}{\tau} \sum_{j=0}^k \delta_j \mathbf{m}_*^{n-j} \in T(\widehat{\mathbf{m}}_*^n). \end{aligned}$$

Note that the definition of $\dot{\mathbf{m}}_*^n$ contains the projection $\mathbf{P}(\widehat{\mathbf{m}}_*^n)$, while $\dot{\mathbf{m}}^n$ was defined without a projection (see the first formula in (2.2)), since $\dot{\mathbf{m}}^n = \mathbf{P}(\widehat{\mathbf{m}}^n) \dot{\mathbf{m}}^n$ is automatically satisfied due to the constraint in (2.3).

The consistency error is bounded as follows.

Lemma 6.1. *If the solution of the LLG equation (1.4) has the regularity $\mathbf{m} \in C^{k+1}([0, \bar{t}], L^2(\Omega)^3) \cap C^1([0, \bar{t}], L^\infty(\Omega)^3)$ and $\Delta \mathbf{m} + \mathbf{H} \in C([0, \bar{t}], L^\infty(\Omega)^3)$, then the consistency error (6.3) is bounded by*

$$\|\mathbf{d}^n\|_{L^2(\Omega)^3} \leq C\tau^k$$

for $n = k, \dots, N$.

Proof. We begin by rewriting the equation for the defect as

$$(6.5) \quad \begin{aligned} \mathbf{d}^n &= \alpha \dot{\mathbf{m}}_*^n + \widehat{\mathbf{m}}_*^n \times \dot{\mathbf{m}}_*^n - \mathbf{P}(\mathbf{m}_*^n)(\Delta \mathbf{m}_*^n + \mathbf{H}(t_n)) \\ &\quad - (\mathbf{P}(\widehat{\mathbf{m}}_*^n) - \mathbf{P}(\mathbf{m}_*^n))(\Delta \mathbf{m}_*^n + \mathbf{H}(t_n)). \end{aligned}$$

In view of (1.4), we have

$$\mathbf{P}(\mathbf{m}_*^n)(\Delta \mathbf{m}_*^n + \mathbf{H}(t_n)) = \alpha \partial_t \mathbf{m}(t_n) + \mathbf{m}_*^n \times \partial_t \mathbf{m}(t_n),$$

and can rewrite (6.5) as

$$\begin{aligned} \mathbf{d}^n &= \alpha (\dot{\mathbf{m}}_*^n - \partial_t \mathbf{m}(t_n)) + (\widehat{\mathbf{m}}_*^n \times \dot{\mathbf{m}}_*^n - \mathbf{m}_*^n \times \partial_t \mathbf{m}(t_n)) \\ &\quad - (\mathbf{P}(\widehat{\mathbf{m}}_*^n) - \mathbf{P}(\mathbf{m}_*^n))(\Delta \mathbf{m}_*^n + \mathbf{H}(t_n)), \end{aligned}$$

i.e.,

$$\begin{aligned} \mathbf{d}^n &= \alpha (\dot{\mathbf{m}}_*^n - \partial_t \mathbf{m}(t_n)) + (\widehat{\mathbf{m}}_*^n - \mathbf{m}_*^n) \times \dot{\mathbf{m}}_*^n + \mathbf{m}_*^n \times (\dot{\mathbf{m}}_*^n - \partial_t \mathbf{m}(t_n)) \\ &\quad - (\mathbf{P}(\widehat{\mathbf{m}}_*^n) - \mathbf{P}(\mathbf{m}_*^n))(\Delta \mathbf{m}_*^n + \mathbf{H}(t_n)). \end{aligned}$$

Therefore,

$$(6.6) \quad \mathbf{d}^n = \alpha \dot{\mathbf{d}}^n + \widehat{\mathbf{d}}^n \times \dot{\mathbf{m}}_*^n + \mathbf{m}_*^n \times \dot{\mathbf{d}}^n - (\mathbf{P}(\widehat{\mathbf{m}}_*^n) - \mathbf{P}(\mathbf{m}_*^n))(\Delta \mathbf{m}_*^n + \mathbf{H}(t_n)),$$

with

$$(6.7) \quad \dot{\mathbf{d}}^n := \dot{\mathbf{m}}_*^n - \partial_t \mathbf{m}(t_n), \quad \widehat{\mathbf{d}}^n := \widehat{\mathbf{m}}_*^n - \mathbf{m}_*^n.$$

Now, in view of the first estimate in Lemma 4.1, we have

$$\|(\mathbf{P}(\widehat{\mathbf{m}}_*^n) - \mathbf{P}(\mathbf{m}_*^n))(\Delta \mathbf{m}_*^n + \mathbf{H}(t_n))\|_{L^2} \leq C \|\widehat{\mathbf{m}}_*^n - \mathbf{m}_*^n\|_{L^2},$$

i.e.,

$$(6.8) \quad \|(\mathbf{P}(\widehat{\mathbf{m}}_*^n) - \mathbf{P}(\mathbf{m}_*^n))(\Delta \mathbf{m}_*^n + \mathbf{H}(t_n))\|_{L^2} \leq C \|\widehat{\mathbf{d}}^n\|_{L^2}.$$

Therefore, it suffices to estimate $\dot{\mathbf{d}}^n$ and $\widehat{\mathbf{d}}^n$.

To estimate $\widehat{\mathbf{d}}^n$, we shall proceed in two steps. First we shall estimate the extrapolation error

$$(6.9) \quad \sum_{j=0}^{k-1} \gamma_j \mathbf{m}_*^{n-j-1} - \mathbf{m}_*^n$$

and then $\widehat{\mathbf{d}}^n$.

By Taylor expanding about t_{n-k} , the leading terms of order up to $k-1$ cancel, due to the second equality in (6.1), and we obtain

$$(6.10) \quad \sum_{i=0}^{k-1} \gamma_i \mathbf{m}_*^{n-i-1} - \mathbf{m}_*^n = \frac{1}{(k-1)!} \left[\sum_{j=0}^{k-1} \gamma_j \int_{t_{n-k}}^{t_{n-j-1}} (t_{n-j-1} - s)^{k-1} \mathbf{m}^{(k)}(s) ds - \int_{t_{n-k}}^{t_n} (t_n - s)^{k-1} \mathbf{m}^{(k)}(s) ds \right],$$

with $\mathbf{m}^{(\ell)} := \frac{\partial^\ell \mathbf{m}}{\partial t^\ell}$, whence

$$(6.11) \quad \left\| \sum_{i=0}^{k-1} \gamma_i \mathbf{m}_*^{n-i-1} - \mathbf{m}_*^n \right\|_{L^2} \leq C\tau^k.$$

Now, for a normalized vector \mathbf{a} and a non-zero vector \mathbf{b} , we have

$$\mathbf{a} - \frac{\mathbf{b}}{|\mathbf{b}|} = (\mathbf{a} - \mathbf{b}) + \frac{1}{|\mathbf{b}|} (|\mathbf{b}| - |\mathbf{a}|) \mathbf{b},$$

whence

$$|\mathbf{a} - \frac{\mathbf{b}}{|\mathbf{b}|}| \leq 2|\mathbf{a} - \mathbf{b}|.$$

Therefore, (6.11) yields

$$(6.12) \quad \|\widehat{\mathbf{d}}^n\|_{L^2} \leq C\tau^k.$$

To bound $\dot{\mathbf{d}}^n$, we use the fact that $\mathbf{P}(\mathbf{m}(t_n))\partial_t \mathbf{m}(t_n) = \partial_t \mathbf{m}(t_n) \in T(\mathbf{m}(t_n))$, so that we have

$$\begin{aligned} \dot{\mathbf{d}}^n &= \mathbf{P}(\widehat{\mathbf{m}}_*^n) \frac{1}{\tau} \sum_{j=0}^k \delta_j \mathbf{m}(t_{n-j}) - \partial_t \mathbf{m}(t_n) \\ &= \mathbf{P}(\widehat{\mathbf{m}}_*^n) \left(\frac{1}{\tau} \sum_{j=0}^k \delta_j \mathbf{m}(t_{n-j}) - \partial_t \mathbf{m}(t_n) \right) + (\mathbf{P}(\widehat{\mathbf{m}}_*^n) - \mathbf{P}(\mathbf{m}(t_n))) \partial_t \mathbf{m}(t_n). \end{aligned}$$

By Lemma 4.1 and (6.12), we have for the last term

$$\|(\mathbf{P}(\widehat{\mathbf{m}}_*^n) - \mathbf{P}(\mathbf{m}(t_n))) \partial_t \mathbf{m}(t_n)\|_{L^2} \leq C\tau^k.$$

By Taylor expanding the first term about t_{n-k} , we see that, due to the order conditions of the implicit BDF method, i.e., the first equality in (6.1), the leading terms of order up to $k-1$ cancel, and we obtain

$$(6.13) \quad \begin{aligned} \frac{1}{\tau} \sum_{j=0}^k \delta_j \mathbf{m}(t_{n-j}) - \partial_t \mathbf{m}(t_n) &= \frac{1}{k!} \left[\frac{1}{\tau} \sum_{j=0}^k \delta_j \int_{t_{n-k}}^{t_{n-j}} (t_{n-j} - s)^k \mathbf{m}^{(k+1)}(s) ds \right. \\ &\quad \left. - k \int_{t_{n-k}}^{t_n} (t_n - s)^{k-1} \mathbf{m}^{(k+1)}(s) ds \right], \end{aligned}$$

whence

$$(6.14) \quad \|\dot{\mathbf{d}}^n\|_{L^2} \leq C\tau^k,$$

provided the solution \mathbf{m} is sufficiently regular. Now, (6.6), (6.8), (6.14), and (6.12) yield

$$(6.15) \quad \|\mathbf{d}^n\|_{L^2} \leq C\tau^k.$$

This is the desired consistency estimate, which is valid for BDF methods of arbitrary order k . \square

6.2. Consistency error of the full discretization. We define the Ritz projection $R_h: H^1(\Omega) \rightarrow V_h$ corresponding to the Poisson–Neumann problem via

$$(\nabla R_h \varphi, \nabla \psi) + (R_h \varphi, 1)(\psi, 1) = (\nabla \varphi, \nabla \psi) + (\varphi, 1)(\psi, 1)$$

for all $\psi \in V_h$, and we denote $\mathbf{R}_h = \mathbf{I} \otimes R_h: H^1(\Omega)^3 \rightarrow V_h^3$. We denote again the L^2 -orthogonal projections onto the finite element space by $\Pi_h: L^2(\Omega) \rightarrow V_h$ and $\mathbf{\Pi}_h = \mathbf{I} \otimes \Pi_h: L^2(\Omega)^3 \rightarrow V_h^3$. As in the previous section, we write $\mathbf{P}_h(\mathbf{m})$ for the L^2 -orthogonal projection onto the discrete tangent space at \mathbf{m} . We insert the following quantities, which are related to the exact solution,

$$(6.16) \quad \begin{aligned} \mathbf{m}_{*,h}^n &= \mathbf{R}_h \mathbf{m}(t_n), \\ \widehat{\mathbf{m}}_{*,h}^n &= \sum_{j=0}^{k-1} \gamma_j \mathbf{m}_{*,h}^{n-j-1} \Big/ \left| \sum_{j=0}^{k-1} \gamma_j \mathbf{m}_{*,h}^{n-j-1} \right|, \\ \dot{\mathbf{m}}_{*,h}^n &= \mathbf{P}_h(\widehat{\mathbf{m}}_{*,h}^n) \frac{1}{\tau} \sum_{j=0}^k \delta_j \mathbf{m}_{*,h}^{n-j} \in T_h(\widehat{\mathbf{m}}_{*,h}^n), \end{aligned}$$

into the linearly implicit k -step BDF method (2.6) and obtain a defect $\mathbf{d}_h^n \in T_h(\widehat{\mathbf{m}}_{*,h}^n)$ from

$$(6.17) \quad \begin{aligned} \alpha(\dot{\mathbf{m}}_{*,h}^n, \varphi_h) + (\widehat{\mathbf{m}}_{*,h}^n \times \dot{\mathbf{m}}_{*,h}^n, \varphi_h) &= -(\nabla \mathbf{m}_{*,h}^n, \nabla \varphi_h) + (\mathbf{H}(t_n), \varphi_h) \\ &\quad + (\mathbf{d}_h^n, \varphi_h) \end{aligned}$$

for all $\varphi_h \in T_h(\widehat{\mathbf{m}}_{*,h}^n)$. By definition, there holds $(R_h \varphi, 1) = (\varphi, 1)$ (this can be seen by testing with $\psi = 1$) and hence

$$(\nabla \mathbf{m}_{*,h}^n, \nabla \varphi) = (\nabla \mathbf{m}(t_n), \nabla \varphi) = -(\Delta \mathbf{m}(t_n), \varphi).$$

Thus, we obtain the consistency error for the full discretization by

$$(6.18) \quad \mathbf{d}_h^n = \mathbf{P}_h(\widehat{\mathbf{m}}_{*,h}^n) \mathbf{D}_h^n \quad \text{with} \quad \mathbf{D}_h^n = \alpha \dot{\mathbf{m}}_{*,h}^n + \widehat{\mathbf{m}}_{*,h}^n \times \dot{\mathbf{m}}_{*,h}^n - \Delta \mathbf{m}(t_n) - \mathbf{H}(t_n)$$

for $n = k, \dots, N$. The consistency error is bounded as follows.

Lemma 6.2. *If the solution of the LLG equation (1.4) has the regularity*

$$\begin{aligned} \mathbf{m} &\in C^{k+1}([0, \bar{t}], L^2(\Omega)^3) \cap C^1([0, \bar{t}], W^{r+1, \infty}(\Omega)^3) \quad \text{and} \\ \Delta \mathbf{m} + \mathbf{H} &\in C([0, \bar{t}], W^{r+1, \infty}(\Omega)^3), \end{aligned}$$

then the consistency error (6.18) is bounded by

$$\|\mathbf{d}_h^n\|_{L^2(\Omega)^3} \leq C(\tau^k + h^r)$$

for n with $k\tau \leq n\tau \leq \bar{t}$.

Proof. We begin by defining

$$\mathbf{D}^n := \alpha \partial_t \mathbf{m}(t_n) + \mathbf{m}(t_n) \times \partial_t \mathbf{m}(t_n) - \Delta \mathbf{m}(t_n) - \mathbf{H}(t_n)$$

and note that $\mathbf{P}(\mathbf{m}_*^n) \mathbf{D}^n = 0$. Here we denote again $\mathbf{m}_*^n = \mathbf{m}(t_n)$ and in the following we use also the notations $\dot{\mathbf{m}}_*^n$ and $\widehat{\mathbf{m}}_*^n$ as defined in (6.4). With this, we rewrite the equation for the defect as

$$\begin{aligned} \mathbf{d}_h^n &= \mathbf{P}_h(\widehat{\mathbf{m}}_{*,h}^n) \mathbf{D}_h^n - \mathbf{P}(\mathbf{m}_*^n) \mathbf{D}^n \\ &= \mathbf{P}_h(\widehat{\mathbf{m}}_{*,h}^n) (\mathbf{D}_h^n - \mathbf{D}^n) + (\mathbf{P}_h(\widehat{\mathbf{m}}_{*,h}^n) - \mathbf{P}_h(\widehat{\mathbf{m}}_*^n)) \mathbf{D}^n \\ &\quad + (\mathbf{P}_h(\widehat{\mathbf{m}}_*^n) - \mathbf{P}(\widehat{\mathbf{m}}_*^n)) \mathbf{D}^n + (\mathbf{P}(\widehat{\mathbf{m}}_*^n) - \mathbf{P}(\mathbf{m}_*^n)) \mathbf{D}^n \end{aligned}$$

$$\equiv I + II + III + IV.$$

For the term IV we have by Lemma 4.1

$$\|IV\|_{L^2} \leq 2\|\widehat{\mathbf{m}}_*^n - \mathbf{m}_*^n\|_{L^2} \|\mathbf{D}^n\|_{L^\infty},$$

where the last term $\widehat{\mathbf{m}}_*^n - \mathbf{m}_*^n$ has been bounded in the L^2 norm by $C\tau^k$ in the proof of Lemma 6.1.

The term III is estimated using the first bound from Lemma 5.1, under our regularity assumptions, as

$$\|III\|_{L^2} \leq Ch^r.$$

For the bound on II we use Lemma 5.2 (i) (with $p = 2$ and $q = \infty$), to obtain

$$\|II\|_{L^2} \leq C_R \|\widehat{\mathbf{m}}_{*,h}^n - \widehat{\mathbf{m}}_*^n\|_{L^2} \|\mathbf{D}^n\|_{L^\infty},$$

where, using (7.11), we obtain

$$\|\widehat{\mathbf{m}}_{*,h}^n - \widehat{\mathbf{m}}_*^n\|_{L^2} \leq \frac{2\|\sum_{i=1}^k \gamma_i (\mathbf{R}_h - \mathbf{I}) \mathbf{m}_*^{n-i}\|_{L^2}}{\min|\sum_{i=1}^k \gamma_i \mathbf{m}_*^{n-i}|} \leq Ch^r.$$

The denominator is bounded from below by $1 - C\tau^k$, because $|\mathbf{m}_*^n| = 1$ and $|\sum_{i=1}^k \gamma_i \mathbf{m}_*^{n-i} - \mathbf{m}_*^n| \leq C\tau^k$. For the first term we have

$$\begin{aligned} \|I\|_{L^2} &\leq \|\mathbf{D}^n - \mathbf{D}_h^n\|_{L^2} \\ &\leq \alpha \|\partial_t \mathbf{m}(t_n) - \dot{\mathbf{m}}_{*,h}^n\|_{L^2} + \|\mathbf{m}(t_n) \times \partial_t \mathbf{m}(t_n) - \widehat{\mathbf{m}}_{*,h}^n \times \dot{\mathbf{m}}_{*,h}^n\|_{L^2}. \end{aligned}$$

The terms $\|\partial_t \mathbf{m}(t_n) - \dot{\mathbf{m}}_{*,h}^n\|_{L^2}$ and $\|\mathbf{m}_*^n \times \partial_t \mathbf{m}(t_n) - \widehat{\mathbf{m}}_*^n \times \dot{\mathbf{m}}_{*,h}^n\|_{L^2}$ can be handled as in the proof of Lemma 6.1. Standard error estimates for the Ritz projection R_h (we do not exploit the Aubin–Nitsche duality here) imply

$$\|(\mathbf{I} - \mathbf{R}_h) \dot{\mathbf{m}}_*^n\|_{L^2} \leq ch^r \|\dot{\mathbf{m}}_*^n\|_{H^{r+1}}.$$

Together this yields, under the stated regularity assumption,

$$\|I\|_{L^2} \leq C(\tau^k + h^r),$$

and the result follows. \square

6.3. Error equation. We recall, from (2.6), the fully discrete problem with the linearly implicit BDF method: find $\dot{\mathbf{m}}_h^n \in T_h(\widehat{\mathbf{m}}_h^n)$ such that for all $\varphi_h \in T_h(\widehat{\mathbf{m}}_h^n)$,

$$(6.19) \quad \alpha(\dot{\mathbf{m}}_h^n, \varphi_h) + (\widehat{\mathbf{m}}_h^n \times \dot{\mathbf{m}}_h^n, \varphi_h) + (\nabla \mathbf{m}_h^n, \nabla \varphi_h) = (\mathbf{H}(t_n), \varphi_h).$$

Then, similarly as we have done in Section 4, we first rewrite (6.17): for all $\varphi_h \in T_h(\widehat{\mathbf{m}}_h^n)$,

$$(6.20) \quad \alpha(\dot{\mathbf{m}}_{*,h}^n, \varphi_h) + (\widehat{\mathbf{m}}_{*,h}^n \times \dot{\mathbf{m}}_{*,h}^n, \varphi_h) + (\nabla \mathbf{m}_{*,h}^n, \nabla \varphi_h) = (\mathbf{r}_h^n + \mathbf{H}(t_n), \varphi_h)$$

with

$$(6.21) \quad \mathbf{r}_h^n = -(\mathbf{P}_h(\widehat{\mathbf{m}}_h^n) - \mathbf{P}_h(\widehat{\mathbf{m}}_{*,h}^n))(\Delta \mathbf{m}_*(t_n) + \mathbf{H}(t_n)) + \mathbf{d}_h^n.$$

The error $\mathbf{e}_h^n = \mathbf{m}_h^n - \mathbf{m}_{*,h}^n$ satisfies the error equation that is obtained by subtracting (6.20) from (6.19). We use the notations

$$(6.22) \quad \widehat{\mathbf{e}}_h^n = \widehat{\mathbf{m}}_h^n - \widehat{\mathbf{m}}_{*,h}^n,$$

$$(6.23) \quad \dot{\mathbf{e}}_h^n = \dot{\mathbf{m}}_h^n - \dot{\mathbf{m}}_{*,h}^n = \frac{1}{\tau} \sum_{j=0}^k \delta_j \mathbf{e}_h^{n-j} + \mathbf{s}_h^n,$$

$$\text{with } \mathbf{s}_h^n = (\mathbf{I} - \mathbf{P}_h(\widehat{\mathbf{m}}_{\star,h}^n)) \frac{1}{\tau} \sum_{j=0}^k \delta_j \mathbf{m}_{\star,h}^{n-j}.$$

We have the following bound for \mathbf{s}_h^n .

Lemma 6.3. *Under the regularity assumptions of Lemma 6.2, we have*

$$(6.24) \quad \|\mathbf{s}_h^n\|_{H^1(\Omega)^3} \leq C(\tau^k + h^r).$$

Proof. We use Lemmas 5.1 and 5.3, and the bounds in the proof of Lemma 6.2. We start by subtracting $(\mathbf{I} - \mathbf{P}(\widehat{\mathbf{m}}_{\star,h}^n))\partial_t \mathbf{m}_{\star}^n = 0$, and obtain (with $\partial^\tau \mathbf{m}_{\star,h}^n := \frac{1}{\tau} \sum_{j=0}^k \delta_j \mathbf{m}_{\star,h}^{n-j}$)

$$\begin{aligned} \mathbf{s}_h^n &= (\mathbf{I} - \mathbf{P}_h(\widehat{\mathbf{m}}_{\star,h}^n))\partial^\tau \mathbf{m}_{\star,h}^n - (\mathbf{I} - \mathbf{P}(\widehat{\mathbf{m}}_{\star,h}^n))\partial_t \mathbf{m}_{\star}^n \\ &= (\partial^\tau \mathbf{m}_{\star,h}^n - \partial_t \mathbf{m}_{\star}^n) - (\mathbf{P}_h(\widehat{\mathbf{m}}_{\star,h}^n)\partial^\tau \mathbf{m}_{\star,h}^n - \mathbf{P}(\widehat{\mathbf{m}}_{\star,h}^n)\partial_t \mathbf{m}_{\star}^n). \end{aligned}$$

The first term above is bounded as $O(\tau^k + h^r)$ via the techniques of the consistency proofs, Lemma 6.1 and 6.2. For the second term we have

$$\begin{aligned} &\mathbf{P}_h(\widehat{\mathbf{m}}_{\star,h}^n)\partial^\tau \mathbf{m}_{\star,h}^n - \mathbf{P}(\widehat{\mathbf{m}}_{\star,h}^n)\partial_t \mathbf{m}_{\star}^n \\ &= \mathbf{P}_h(\widehat{\mathbf{m}}_{\star,h}^n)(\partial^\tau \mathbf{m}_{\star,h}^n - \partial_t \mathbf{m}_{\star}^n) + (\mathbf{P}_h(\widehat{\mathbf{m}}_{\star,h}^n) - \mathbf{P}(\widehat{\mathbf{m}}_{\star,h}^n))\partial_t \mathbf{m}_{\star}^n, \end{aligned}$$

where the first term is bounded as $O(\tau^k + h^r)$, using Lemma 5.3 and the previous estimate, while the second term is bounded as $O(h^r)$ by the H^1 estimate from Lemma 5.1. Altogether, we obtain the stated H^1 bound for \mathbf{s}_h^n . \square

We then have the error equation

$$(6.25) \quad \alpha(\dot{\mathbf{e}}_h^n, \boldsymbol{\varphi}_h) + (\widehat{\mathbf{e}}_h^n \times \dot{\mathbf{m}}_{\star,h}^n, \boldsymbol{\varphi}_h) + (\widehat{\mathbf{m}}_h^n \times \dot{\mathbf{e}}_h^n, \boldsymbol{\varphi}_h) + (\nabla \mathbf{e}_h^n, \nabla \boldsymbol{\varphi}_h) = -(\mathbf{r}_h^n, \boldsymbol{\varphi}_h),$$

for all $\boldsymbol{\varphi}_h \in T_h(\widehat{\mathbf{m}}_h^n)$, which is to be taken together with (6.21)–(6.23).

7. STABILITY OF THE FULL DISCRETIZATION FOR BDF OF ORDERS 1 AND 2

For the A-stable BDF methods (those of orders 1 and 2) we obtain the following stability estimate, which is analogous to the continuous perturbation result in Lemma 4.2.

Lemma 7.1 (Stability for orders $k = 1, 2$). *Consider the linearly implicit k -step BDF discretization (2.6) for $k \leq 2$ with finite elements of polynomial degree $r \geq 1$. Let \mathbf{m}_h^n and $\mathbf{m}_{\star,h}^n = \mathbf{R}_h \mathbf{m}(t_n)$ satisfy equations (2.6) and (6.17), respectively, and suppose that the exact solution $\mathbf{m}(t)$ is bounded by (4.3) and $\|\mathbf{H}(t)\|_{L^\infty} \leq M$ for $0 \leq t \leq \bar{t}$. Then, for sufficiently small $h \leq \bar{h}$ and $\tau \leq \bar{\tau}$, the error $\mathbf{e}_h^n = \mathbf{m}_h^n - \mathbf{m}_{\star,h}^n$ satisfies the following bound, for $k\tau \leq n\tau \leq \bar{t}$,*

$$(7.1) \quad \|\mathbf{e}_h^n\|_{H^1(\Omega)^3}^2 \leq C \left(\sum_{i=0}^{k-1} \|\mathbf{e}_h^i\|_{H^1(\Omega)^3}^2 + \tau \sum_{j=k}^n \|\mathbf{d}_h^j\|_{L^2(\Omega)^3}^2 + \tau \sum_{j=k}^n \|\mathbf{s}_h^j\|_{H^1(\Omega)^3}^2 \right),$$

where the constant C is independent of h, τ and n , but depends on α, R, K, M , and \bar{t} . This estimate holds under the smallness condition that the right-hand side is bounded by $\hat{c}h$ with a sufficiently small constant \hat{c} (note that the right-hand side is of size $O((\tau^k + h^r)^2)$ in the case of a sufficiently regular solution).

Combining Lemmas 7.1, 6.2 and 6.3 yields the *proof of Theorem 3.1*: These lemmas imply the estimate

$$\|\mathbf{e}_h^n\|_{H^1(\Omega)^3} \leq \tilde{C}(\tau^k + h^r)$$

in the case of a sufficiently regular solution. Since then $\|\mathbf{R}_h \mathbf{m}(t_n) - \mathbf{m}(t_n)\|_{H^1(\Omega)^3} \leq Ch^r$ and because of $\mathbf{m}_h^n - \mathbf{m}(t_n) = \mathbf{e}_h^n + (\mathbf{R}_h \mathbf{m}(t_n) - \mathbf{m}(t_n))$, this implies the error bound (3.1).

The smallness condition imposed in Lemma 7.1 is satisfied under the very mild CFL condition, for a sufficiently small $\bar{c} > 0$ (independent of h, τ and n),

$$\tau^k \leq \bar{c}h^{1/2}.$$

Taken together, this proves Theorem 3.1.

Proof. (a) Preparations. The proof of this lemma transfers the arguments of the proof of Lemma 4.2 to the fully discrete situation, using energy estimates obtained by testing with (essentially) the discrete time derivative of the error, as presented in the Appendix, which is based on Dahlquist's G-stability theory.

However, testing the error equation (6.25) directly with $\dot{\mathbf{e}}_h^n$ is not possible, since $\dot{\mathbf{e}}_h^n$ is not in the tangent space $T_h(\widehat{\mathbf{m}}_h^n)$. Therefore, as in the proof of Lemma 4.2, we again start by showing that the test function $\varphi_h = \mathbf{P}_h(\widehat{\mathbf{m}}_h^n)\dot{\mathbf{e}}_h^n \in T_h(\widehat{\mathbf{m}}_h^n) \cap H^1(\Omega)^3$ is a perturbation of $\dot{\mathbf{e}}_h^n$ itself:

$$\begin{aligned} \varphi_h &= \mathbf{P}_h(\widehat{\mathbf{m}}_h^n)\dot{\mathbf{e}}_h^n = \mathbf{P}_h(\widehat{\mathbf{m}}_h^n)\dot{\mathbf{m}}_h^n - \mathbf{P}_h(\widehat{\mathbf{m}}_h^n)\dot{\mathbf{m}}_{\star,h}^n \\ &= \mathbf{P}_h(\widehat{\mathbf{m}}_h^n)\dot{\mathbf{m}}_h^n - \mathbf{P}_h(\widehat{\mathbf{m}}_{\star,h}^n)\dot{\mathbf{m}}_{\star,h}^n + (\mathbf{P}_h(\widehat{\mathbf{m}}_{\star,h}^n) - \mathbf{P}_h(\widehat{\mathbf{m}}_h^n))\dot{\mathbf{m}}_{\star,h}^n. \end{aligned}$$

Here we note that $\mathbf{P}_h(\widehat{\mathbf{m}}_h^n)\dot{\mathbf{m}}_h^n = \dot{\mathbf{m}}_h^n \in T_h(\widehat{\mathbf{m}}_h^n)$ by construction of the method (2.6), and $\mathbf{P}_h(\widehat{\mathbf{m}}_{\star,h}^n)\dot{\mathbf{m}}_{\star,h}^n = \dot{\mathbf{m}}_{\star,h}^n \in T_h(\widehat{\mathbf{m}}_{\star,h}^n)$ by the definition of $\dot{\mathbf{m}}_{\star,h}^n$ in (6.4). So we have

$$\varphi_h = \dot{\mathbf{m}}_h^n - \dot{\mathbf{m}}_{\star,h}^n - (\mathbf{P}_h(\widehat{\mathbf{m}}_h^n) - \mathbf{P}_h(\widehat{\mathbf{m}}_{\star,h}^n))\dot{\mathbf{m}}_{\star,h}^n,$$

and hence

$$(7.2) \quad \varphi_h = \mathbf{P}_h(\widehat{\mathbf{m}}_h^n)\dot{\mathbf{e}}_h^n = \dot{\mathbf{e}}_h^n + \mathbf{q}_h^n \quad \text{with} \quad \mathbf{q}_h^n = -(\mathbf{P}_h(\widehat{\mathbf{m}}_h^n) - \mathbf{P}_h(\widehat{\mathbf{m}}_{\star,h}^n))\dot{\mathbf{m}}_{\star,h}^n.$$

The proof now transfers the proof of the continuous perturbation result Lemma 4.2 to the discrete situation with some notable differences, which are emphasized here:

(i) Instead of using the continuous quantities it uses their spatially discrete counterparts, in particular the discrete projections $\mathbf{P}_h(\widehat{\mathbf{m}}_h^n)$ and $\mathbf{P}_h(\widehat{\mathbf{m}}_{\star,h}^n)$, defined and studied in Section 5. In view of the definition (2.1) and (6.16) of $\widehat{\mathbf{m}}_h^n$ and $\widehat{\mathbf{m}}_{\star,h}^n$, respectively, this requires that $\sum_{j=0}^{k-1} \gamma_j \mathbf{m}_h^{n-j-1}(x)$ and $\sum_{j=0}^{k-1} \gamma_j \mathbf{m}_{\star,h}^{n-j-1}(x)$ are bounded away from zero uniformly for all $x \in \Omega$.

(ii) Instead of Lemma 4.1 we use Lemma 5.2 (with $\widehat{\mathbf{m}}_h^n$ and $\widehat{\mathbf{m}}_{\star,h}^n$ in the role of $\widetilde{\mathbf{m}}$ and \mathbf{m} , respectively) to bound the quantity \mathbf{q}_h^n . This requires that $\widehat{\mathbf{m}}_{\star,h}^n$ and $\dot{\mathbf{m}}_{\star,h}^n$ are bounded in $W^{1,\infty}$ independently of h .

Ad (i): In order to show that $|\sum_{j=0}^{k-1} \gamma_j \mathbf{m}_h^{n-j-1}(x)|$ stays close to 1 for all $x \in \Omega$, we need to establish an L^∞ bound for the errors $\mathbf{e}_h^{n-j-1} = \mathbf{m}_h^{n-j-1} - \mathbf{m}_{\star,h}^{n-j-1}$.

We use an induction argument and assume that for some time step number \bar{n} with $\bar{n}\tau \leq \bar{t}$ we have

$$(7.3) \quad \|\mathbf{e}_h^n\|_{L^\infty} \leq \rho \quad \text{for } 0 \leq n < \bar{n},$$

where we choose ρ sufficiently small independent of h and τ . (In this proof it suffices to choose $\rho \leq 1/(4C_\gamma)$, where $C_\gamma = \sum_{j=0}^{k-1} |\gamma_j| = 2^k - 1$.)

Note that the smallness condition of the lemma implies that (7.3) is satisfied for $\bar{n} = k$, because for the L^∞ errors of the starting values we have by an inverse inequality, for $i = 0, \dots, k-1$,

$$\|\mathbf{e}_h^i\|_{L^\infty} \leq Ch^{-1/2} \|\mathbf{e}_h^i\|_{H^1} \leq Ch^{-1/2} (\hat{c}h)^{1/2} = C\hat{c}^{1/2} \leq \rho,$$

provided that \hat{c} is sufficiently small (independent of τ and h), as is assumed.

We will show in part (b) of the proof that with the induction hypothesis (7.3) we obtain also $\|\mathbf{e}_h^{\bar{n}}\|_{L^\infty} \leq \rho$ so that finally we obtain (7.3) for all \bar{n} with $\bar{n}\tau \leq \bar{t}$.

Using reverse and ordinary triangle inequalities, the error bound of [12, Corollary 8.1.12] (noting that $\mathbf{m}(t) \in W^{2,\infty}(\Omega)$ under our assumptions) and the L^∞ boundedness of $\partial_t \mathbf{m}$, and the bound (7.3), we estimate

$$\begin{aligned} (7.4) \quad & \left\| \left| \sum_{j=0}^{k-1} \gamma_j \mathbf{m}_h^{n-j-1} \right| - 1 \right\|_{L^\infty} = \left\| \left| \sum_{j=0}^{k-1} \gamma_j \mathbf{m}_h^{n-j-1} \right| - |\mathbf{m}_\star^n| \right\|_{L^\infty} \\ & \leq \left\| \sum_{j=0}^{k-1} \gamma_j \mathbf{m}_h^{n-j-1} - \mathbf{m}_\star^n \right\|_{L^\infty} \leq \left\| \sum_{j=0}^{k-1} \gamma_j \mathbf{e}_h^{n-j-1} \right\|_{L^\infty} \\ & \quad + \left\| \sum_{j=0}^{k-1} \gamma_j (\mathbf{R}_h \mathbf{m}_\star^{n-j-1} - \mathbf{m}_\star^{n-j-1}) \right\|_{L^\infty} + \left\| \sum_{j=0}^{k-1} \gamma_j (\mathbf{m}_\star^{n-j-1} - \mathbf{m}_\star^n) \right\|_{L^\infty} \\ & \leq \left\| \sum_{j=0}^{k-1} \gamma_j \mathbf{e}_h^{n-j-1} \right\|_{L^\infty} + Ch + C\tau \leq \sum_{j=0}^{k-1} |\gamma_j| \cdot \rho + Ch + C\tau \leq \frac{1}{2}, \end{aligned}$$

provided that h and τ are sufficiently small. The same argument also yields that $\left\| \left| \sum_{j=0}^{k-1} \gamma_j \mathbf{m}_{\star,h}^{n-j-1} \right| - 1 \right\|_{L^\infty} \leq \frac{1}{2}$, and so we have

$$(7.5) \quad \frac{1}{2} \leq \left| \sum_{j=0}^{k-1} \gamma_j \mathbf{m}_h^{n-j-1}(x) \right| \leq \frac{3}{2} \quad \text{and} \quad \frac{1}{2} \leq \left| \sum_{j=0}^{k-1} \gamma_j \mathbf{m}_{\star,h}^{n-j-1}(x) \right| \leq \frac{3}{2}$$

for all $x \in \Omega$. In particular, it follows that $\widehat{\mathbf{m}}_h^n$ and $\widehat{\mathbf{m}}_{\star,h}^n$ are unambiguously defined.

Ad (ii): The required $W^{1,\infty}$ bound for $\mathbf{m}_{\star,h}^n = \mathbf{R}_h \mathbf{m}(t_n)$ follows from the $W^{1,\infty}$ -stability of the Ritz projection: by [12, Theorem 8.1.11] and by the assumed $W^{1,\infty}$ bound (4.3) for $\mathbf{m}(t)$,

$$(7.6) \quad \|\mathbf{m}_{\star,h}^n\|_{W^{1,\infty}} \leq C \|\mathbf{m}(t_n)\|_{W^{1,\infty}} \leq CR.$$

The bounds (7.5) and (7.6) for $n \leq \bar{n}$ imply that also

$$(7.7) \quad \|\widehat{\mathbf{m}}_{\star,h}^n\|_{W^{1,\infty}} \leq CR$$

for $n \leq \bar{n}$ (with a different constant C). Using this bound in Lemma 5.3 and the assumed $W^{1,\infty}$ bound (4.3) for $\partial_t \mathbf{m}(t)$, we obtain with $\delta(\zeta)/(1-\zeta) = \sum_{\ell=1}^k (1-\zeta)^{\ell-1}/\ell =: \sum_{j=0}^{k-1} \mu_j \zeta^j$ that

$$\|\dot{\mathbf{m}}_{\star,h}^n\|_{W^{1,\infty}} = \|\mathbf{P}_h(\widehat{\mathbf{m}}_{\star,h}^n)\|_{W^{1,\infty}} \frac{1}{\tau} \sum_{j=0}^k \delta_j \mathbf{m}_\star^{n-j}\|_{W^{1,\infty}}$$

$$\begin{aligned}
&= \|\mathbf{P}_h(\widehat{\mathbf{m}}_{*,h}^n) \sum_{j=0}^{k-1} \mu_j \frac{1}{\tau} (\mathbf{m}_*^{n-j} - \mathbf{m}_*^{n-j-1})\|_{W^{1,\infty}} \\
&= \|\mathbf{P}_h(\widehat{\mathbf{m}}_{*,h}^n) \sum_{j=0}^{k-1} \mu_j \frac{1}{\tau} \int_{t_{n-j-1}}^{t_{n-j}} \partial_t \mathbf{m}(t) dt\|_{W^{1,\infty}} \\
&\leq C_R \left\| \sum_{j=0}^{k-1} \mu_j \frac{1}{\tau} \int_{t_{n-j-1}}^{t_{n-j}} \partial_t \mathbf{m}(t) dt \right\|_{W^{1,\infty}} \\
&\leq C_R \sum_{j=0}^{k-1} |\mu_j| R.
\end{aligned}$$

We can now establish a bound for \mathbf{q}_h^n as defined in (7.2), using Lemma 5.2 together with the above $W^{1,\infty}$ bounds for $\widehat{\mathbf{m}}_{*,h}^n$ and $\dot{\mathbf{m}}_{*,h}^n$ to obtain

$$(7.8) \quad \|\mathbf{q}_h^n\|_{L^2} \leq c \|\widehat{\mathbf{e}}_h^n\|_{L^2} \quad \text{and} \quad \|\nabla \mathbf{q}_h^n\|_{L^2} \leq c \|\widehat{\mathbf{e}}_h^n\|_{H^1}.$$

With the $W^{1,\infty}$ bound of $\widehat{\mathbf{m}}_{*,h}^n$ we also obtain a bound of \mathbf{r}_h^n defined in (6.21). Using Lemma 5.2 (i) and recalling the L^∞ bound of $\Delta \mathbf{m} + \mathbf{H}$ of (4.3), we find that \mathbf{r}_h^n is bounded by

$$(7.9) \quad \begin{aligned} \|\mathbf{r}_h^n\|_{L^2} &\leq \|(\mathbf{P}_h(\widehat{\mathbf{m}}_h^n) - \mathbf{P}_h(\widehat{\mathbf{m}}_{*,h}^n))(\Delta \mathbf{m}_*^n + \mathbf{H}^n)\|_{L^2} + \|\mathbf{d}_h^n\|_{L^2} \\ &\leq c \|\widehat{\mathbf{e}}_h^n\|_{L^2} + \|\mathbf{d}_h^n\|_{L^2}. \end{aligned}$$

(b) *Energy estimates.* For $n \leq \bar{n}$ with \bar{n} of (7.3), we test the error equation (6.25) with $\varphi_h = \dot{\mathbf{e}}_h^n + \mathbf{q}_h^n$ and obtain

$$\begin{aligned}
\alpha(\dot{\mathbf{e}}_h^n, \dot{\mathbf{e}}_h^n + \mathbf{q}_h^n) + (\widehat{\mathbf{e}}_h^n \times \dot{\mathbf{m}}_{*,h}^n, \dot{\mathbf{e}}_h^n + \mathbf{q}_h^n) + (\widehat{\mathbf{m}}_h^n \times \dot{\mathbf{e}}_h^n, \dot{\mathbf{e}}_h^n + \mathbf{q}_h^n) \\
+ (\nabla \mathbf{e}_h^n, \nabla(\dot{\mathbf{e}}_h^n + \mathbf{q}_h^n)) = -(\mathbf{r}_h^n, \dot{\mathbf{e}}_h^n + \mathbf{q}_h^n).
\end{aligned}$$

By collecting the terms, and using the fact that $(\widehat{\mathbf{m}}_h^n \times \dot{\mathbf{e}}_h^n, \dot{\mathbf{e}}_h^n) = 0$, we altogether obtain

$$\begin{aligned}
\alpha \|\dot{\mathbf{e}}_h^n\|_{L^2}^2 + (\nabla \mathbf{e}_h^n, \nabla \dot{\mathbf{e}}_h^n) &= -\alpha(\dot{\mathbf{e}}_h^n, \mathbf{q}_h^n) - (\widehat{\mathbf{e}}_h^n \times \dot{\mathbf{m}}_{*,h}^n, \dot{\mathbf{e}}_h^n + \mathbf{q}_h^n) \\
&\quad - (\widehat{\mathbf{m}}_h^n \times \dot{\mathbf{e}}_h^n, \mathbf{q}_h^n) - (\nabla \mathbf{e}_h^n, \nabla \mathbf{q}_h^n) - (\mathbf{r}_h^n, \dot{\mathbf{e}}_h^n + \mathbf{q}_h^n).
\end{aligned}$$

We now estimate the term $(\nabla \mathbf{e}_h^n, \nabla \dot{\mathbf{e}}_h^n)$ on the left-hand side from below using Dahlquist's Lemma A.1, so that the ensuing relation (A.2) yields

$$(\nabla \mathbf{e}_h^n, \nabla \dot{\mathbf{e}}_h^n) \geq \frac{1}{\tau} \left(\|\nabla \mathbf{E}_h^n\|_G^2 - \|\nabla \mathbf{E}_h^{n-1}\|_G^2 \right) + (\nabla \mathbf{e}_h^n, \nabla \mathbf{s}_h^n),$$

where $\mathbf{E}_h^n = (\mathbf{e}_h^{n-k+1}, \dots, \mathbf{e}_h^n)$ and the G -weighted semi-norm is given by

$$\|\nabla \mathbf{E}_h^n\|_G^2 = \sum_{i,j=1}^k g_{ij} (\nabla \mathbf{e}_h^{n-k+i}, \nabla \mathbf{e}_h^{n-k+j}).$$

This semi-norm satisfies the relation

$$(7.10) \quad \gamma^- \sum_{j=1}^k \|\nabla \mathbf{e}_h^{n-k+j}\|_{L^2}^2 \leq \|\nabla \mathbf{E}_h^n\|_G^2 \leq \gamma^+ \sum_{j=1}^k \|\nabla \mathbf{e}_h^{n-k+j}\|_{L^2}^2,$$

where γ^- and γ^+ are the smallest and largest eigenvalues of the positive definite symmetric matrix $G = (g_{ij})$ from Lemma A.1.

The remaining terms are estimated using the Cauchy–Schwarz inequality and $\|\widehat{\mathbf{m}}_h^n\|_{L^\infty} = 1$; we altogether obtain

$$\begin{aligned} \alpha \|\dot{e}_h^n\|_{L^2}^2 + \frac{1}{\tau} \left(\|\nabla \mathbf{E}_h^n\|_G^2 - \|\nabla \mathbf{E}_h^{n-1}\|_G^2 \right) &\leq \alpha \|\dot{e}_h^n\|_{L^2} \|\mathbf{q}_h^n\|_{L^2} \\ &+ \|\widehat{e}_h^n\|_{L^2} (\|\dot{e}_h^n\|_{L^2} + \|\mathbf{q}_h^n\|_{L^2}) + \|\dot{e}_h^n\|_{L^2} \|\mathbf{q}_h^n\|_{L^2} \\ &+ \|\nabla e_h^n\|_{L^2} (\|\nabla \mathbf{q}^n\|_{L^2} + \|\nabla \mathbf{s}_h^n\|_{L^2}) + \|\mathbf{r}_h^n\|_{L^2} (\|\dot{e}_h^n\|_{L^2} + \|\mathbf{q}_h^n\|_{L^2}). \end{aligned}$$

We now show an L^2 error bound for \widehat{e}_h^n in terms of $(e_h^{n-j-1})_{j=0}^{k-1}$. Using the fact that for $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3 \setminus \{0\}$,

$$(7.11) \quad \left| \frac{\mathbf{a}}{|\mathbf{a}|} - \frac{\mathbf{b}}{|\mathbf{b}|} \right| = \left| \frac{(|\mathbf{b}| - |\mathbf{a}|)\mathbf{a} + |\mathbf{a}|(\mathbf{a} - \mathbf{b})}{|\mathbf{a}| |\mathbf{b}|} \right| \leq 2 \frac{|\mathbf{a} - \mathbf{b}|}{|\mathbf{b}|},$$

and the lower bounds in (7.5) for both $|\sum_{j=0}^{k-1} \gamma_j \mathbf{m}_h^{n-j-1}|$ and $|\sum_{j=0}^{k-1} \gamma_j \mathbf{m}_{\star,h}^{n-j-1}|$, we can estimate

$$(7.12) \quad \|\widehat{e}_h^n\|_{L^2} = \left\| \frac{\sum_{j=0}^{k-1} \gamma_j \mathbf{m}_h^{n-j-1}}{\sum_{j=0}^{k-1} \gamma_j \mathbf{m}_h^{n-j-1}} - \frac{\sum_{j=0}^{k-1} \gamma_j \mathbf{m}_{\star,h}^{n-j-1}}{\sum_{j=0}^{k-1} \gamma_j \mathbf{m}_{\star,h}^{n-j-1}} \right\|_{L^2} \leq C \sum_{j=0}^{k-1} \|e_h^{n-j-1}\|_{L^2}^2.$$

To show a similar bound for $\|\nabla \widehat{e}_h^n\|_{L^2}$ we need the following two observations: First, the $W^{1,\infty}$ bounds for $\mathbf{m}_{\star,h}^{n-j-1}$ from (7.6) imply $W^{1,\infty}$ boundedness for $\widehat{\mathbf{m}}_{\star,h}^n$ by

$$\left| \partial_j \left(\frac{\mathbf{b}}{|\mathbf{b}|} \right) \right| \leq \left| \frac{\partial_j \mathbf{b}}{|\mathbf{b}|} \right| + \left| \frac{\mathbf{b}(\partial_j \mathbf{b}, \mathbf{b})}{|\mathbf{b}|^3} \right|.$$

Second, similarly we have

$$\begin{aligned} \left| \partial_j \left(\frac{\mathbf{a}}{|\mathbf{a}|} - \frac{\mathbf{b}}{|\mathbf{b}|} \right) \right| &\leq \left| \frac{\partial_j \mathbf{a}}{|\mathbf{a}|} - \frac{\partial_j \mathbf{b}}{|\mathbf{b}|} \right| + \left| \frac{\mathbf{a}(\partial_j \mathbf{a}, \mathbf{a})|\mathbf{b}|^3 - \mathbf{b}(\partial_j \mathbf{b}, \mathbf{b})|\mathbf{a}|^3}{|\mathbf{a}|^3 |\mathbf{b}|^3} \right| \\ &\leq \left| \frac{\partial_j \mathbf{a}}{|\mathbf{a}|} - \frac{\partial_j \mathbf{b}}{|\mathbf{b}|} \right| + \frac{|\mathbf{a}|^3 - |\mathbf{b}|^3}{|\mathbf{a}|^3 |\mathbf{b}|} |\partial_j \mathbf{b}| + \frac{|\mathbf{a}(\partial_j \mathbf{a}, \mathbf{a}) - \mathbf{b}(\partial_j \mathbf{b}, \mathbf{b})|}{|\mathbf{b}|^3} \\ &\leq \left| \frac{\partial_j \mathbf{a}}{|\mathbf{a}|} - \frac{\partial_j \mathbf{b}}{|\mathbf{b}|} \right| + \frac{|\mathbf{a} - \mathbf{b}|(|\mathbf{b}|^2 + |\mathbf{b}||\mathbf{a}| + |\mathbf{a}|^2)}{|\mathbf{a}|^3 |\mathbf{b}|} |\partial_j \mathbf{b}| \\ &\quad + \frac{|\mathbf{a}|^2 |\partial_j \mathbf{a} - \partial_j \mathbf{b}|}{|\mathbf{b}|^3} + \frac{|\mathbf{a}| |\partial_j \mathbf{b}| |\mathbf{a} - \mathbf{b}|}{|\mathbf{b}|^3} + \frac{|\mathbf{a} - \mathbf{b}| |\partial_j \mathbf{b}|}{|\mathbf{b}|^2}. \end{aligned}$$

Combining these two observations, again with \mathbf{m}_h and $\mathbf{m}_{\star,h}$ in the role of \mathbf{a} and \mathbf{b} , respectively, and the upper and lower bounds from (7.5) altogether yield

$$(7.13) \quad \|\nabla \widehat{e}_h^n\|_{L^2}^2 \leq C \sum_{j=0}^{k-1} \|e_h^{n-j-1}\|_{H^1}^2.$$

We estimate further using Young's inequality and absorptions into the term $\|\dot{e}_h^n\|_{L^2}^2$, together with the bounds in (7.8) and (7.9), to obtain

$$\alpha \frac{1}{2} \|\dot{e}_h^n\|_{L^2}^2 + \frac{1}{\tau} \left(\|\nabla \mathbf{E}_h^n\|_G^2 - \|\nabla \mathbf{E}_h^{n-1}\|_G^2 \right) \leq c \sum_{j=0}^k \|e_h^{n-j}\|_{H^1}^2 + c \|\mathbf{d}_h^n\|_{L^2}^2 + c \|\nabla \mathbf{s}_h^n\|_{L^2}^2.$$

Multiplying both sides by τ , summing up from k to $n \leq \bar{n}$, and using an absorption yield

$$\begin{aligned} & \alpha \frac{1}{2} \tau \sum_{j=k}^n \|\dot{\mathbf{e}}_h^j\|_{L^2}^2 + \|\nabla \mathbf{E}_h^n\|_G^2 \\ & \leq \|\nabla \mathbf{E}_h^{k-1}\|_G^2 + c\tau \sum_{j=k}^n \|\mathbf{e}_h^j\|_{H^1}^2 + c\tau \sum_{j=k}^n (\|\mathbf{d}_h^j\|_{L^2}^2 + \|\mathbf{s}_h^j\|_{H^1}^2) + c\tau \sum_{i=0}^{k-1} \|\mathbf{e}_h^i\|_{H^1}^2. \end{aligned}$$

We then arrive, using (7.10), at

$$(7.14) \quad \begin{aligned} \alpha \frac{1}{2} \tau \sum_{j=k}^n \|\dot{\mathbf{e}}_h^j\|_{L^2}^2 + \|\nabla \mathbf{e}_h^n\|_{L^2}^2 & \leq c\tau \sum_{j=k}^n \|\mathbf{e}_h^j\|_{H^1}^2 + c\tau \sum_{j=k}^n (\|\mathbf{d}_h^j\|_{L^2}^2 + \|\mathbf{s}_h^j\|_{H^1}^2) \\ & \quad + c \sum_{i=0}^{k-1} \|\mathbf{e}_h^i\|_{H^1}^2, \end{aligned}$$

with c depending on α .

Similarly as in the time continuous case in the proof of Lemma 4.2, we connect $\|\mathbf{e}_h^n\|_{L^2}^2$ and $\tau \sum_{j=k}^n \|\dot{\mathbf{e}}_h^j\|_{L^2}^2$. We rewrite the identity

$$\frac{1}{\tau} \sum_{j=0}^k \delta_j \mathbf{e}_h^{n-j} = \dot{\mathbf{e}}_h^n - \mathbf{s}_h^n, \quad n \geq k,$$

as

$$\frac{1}{\tau} \sum_{j=k}^n \delta_{n-j} \mathbf{e}_h^j = \dot{\mathbf{e}}_h^n - \mathbf{s}_h^n - \mathbf{g}_h^n, \quad n \geq k,$$

with $\delta_\ell = 0$ for $\ell > k$ and where

$$\mathbf{g}_h^n := \frac{1}{\tau} \sum_{i=0}^{k-1} \delta_{n-i} \mathbf{e}_h^i$$

depends only on the starting errors and satisfies $\mathbf{g}_h^n = 0$ for $n \geq 2k$. With the inverse power series of $\delta(\zeta)$,

$$\kappa(\zeta) = \sum_{n=0}^{\infty} \kappa_n \zeta^n := \frac{1}{\delta(\zeta)},$$

we then have, for $n \geq k$,

$$\mathbf{e}_h^n = \tau \sum_{j=k}^n \kappa_{n-j} (\dot{\mathbf{e}}_h^j - \mathbf{s}_h^j - \mathbf{g}_h^j).$$

By the zero-stability of the BDF method of order $k \leq 6$, the coefficients κ_n are uniformly bounded: $|\kappa_n| \leq c$ for all $n \geq 0$. Therefore we obtain via the Cauchy-Schwarz inequality

$$\begin{aligned} \|\mathbf{e}_h^n\|_{L^2}^2 & \leq 2\tau^2 \left\| \sum_{j=k}^n \kappa_{n-j} (\dot{\mathbf{e}}_h^j - \mathbf{s}_h^j) \right\|_{L^2}^2 + 2\tau^2 \left\| \sum_{j=k}^{2k-1} \kappa_{n-j} \mathbf{g}_h^j \right\|_{L^2}^2 \\ & \leq (2n\tau)\tau c^2 \sum_{j=k}^n \|\dot{\mathbf{e}}_h^j - \mathbf{s}_h^j\|_{L^2}^2 + 2\tau^2 c^2 k \sum_{j=k}^{2k-1} \|\mathbf{g}_h^j\|_{L^2}^2 \end{aligned}$$

$$\leq C\tau \sum_{j=k}^n \|\dot{\mathbf{e}}_h^j\|_{L^2}^2 + C\tau \sum_{j=k}^n \|\mathbf{s}_h^j\|_{L^2}^2 + C \sum_{i=0}^k \|\mathbf{e}_h^i\|_{L^2}^2.$$

Inserting this bound into (7.14) then yields

$$\alpha \|\mathbf{e}_h^n\|_{L^2}^2 + \|\nabla \mathbf{e}_h^n\|_{L^2}^2 \leq c\tau \sum_{j=k}^n \|\mathbf{e}_h^j\|_{H^1}^2 + c\tau \sum_{j=k}^n (\|\mathbf{d}_h^j\|_{L^2}^2 + \|\mathbf{s}_h^j\|_{H^1}^2) + c \sum_{i=0}^{k-1} \|\mathbf{e}_h^i\|_{H^1}^2,$$

and a discrete Gronwall inequality implies the stated stability result for $n \leq \bar{n}$. It then follows from this stability bound, the smallness condition of the lemma and the inverse estimate from H^1 to L^∞ that (7.3) is satisfied also for $\bar{n} + 1$. This completes the induction step for (7.3) and proves the stated error bound. \square

8. STABILITY OF THE FULL DISCRETIZATION FOR BDF OF ORDERS 3 TO 5

Stability for full discretizations using the BDF methods of orders 3 to 5 can be shown under additional conditions on the damping parameter α and the stepsize τ .

Lemma 8.1 (Stability for orders $k = 3, 4, 5$). *Consider the linearly implicit k -step BDF discretization (2.6) for $3 \leq k \leq 5$ with finite elements of polynomial degree $r \geq 2$. Let \mathbf{m}_h^n and $\mathbf{m}_{\star,h}^n$ satisfy (2.6) and (6.17), respectively, and suppose that the regularity assumptions of Lemma 7.1 hold. Furthermore, assume that the damping parameter α satisfies*

$$(8.1) \quad \alpha > \alpha_k := \frac{\eta_k}{1 - \eta_k}$$

with the multiplier η_k of Lemma A.2, and that τ and h satisfy the mild CFL-type condition, for some $\bar{c} > 0$,

$$(8.2) \quad \tau \leq \bar{c}h.$$

Then, for sufficiently small $h \leq \bar{h}$ and $\tau \leq \bar{\tau}$, the error $\mathbf{e}_h^n = \mathbf{m}_h^n - \mathbf{m}_{\star,h}^n$ satisfies the following bound, for $k\tau \leq n\tau \leq \bar{t}$,

$$(8.3) \quad \|\mathbf{e}_h^n\|_{H^1(\Omega)^3}^2 \leq C \left(\sum_{i=0}^{k-1} \|\mathbf{e}_h^i\|_{H^1(\Omega)^3}^2 + \tau \sum_{j=k}^n \|\mathbf{d}_h^j\|_{L^2(\Omega)^3}^2 + \tau \sum_{j=k}^n \|\mathbf{s}_h^j\|_{H^1(\Omega)^3}^2 \right),$$

where the constant C is independent of τ, h and n , but depends on α, R, K, M , and exponentially on $\bar{c}\bar{t}$. The bound holds under the smallness condition that the right-hand side is bounded by $\hat{c}h^3$ with a constant \hat{c} (note that the right-hand side is of size $O((\tau^k + h^r)^2)$ in the case of a sufficiently regular solution and sufficiently accurate starting values).

Together with the defect bounds of Section 6, this stability lemma proves Theorem 3.2. We remark that the thresholds $\alpha_k > 0$ defined here are the same as those appearing in Theorem 3.2.

Proof. The proof of this lemma combines the arguments of the proof of Lemma 7.1 with a nonstandard variant of the multiplier technique of Nevanlinna and Odeh, as outlined in the Appendix. Since the size of the parameter α determines which BDF methods satisfy the stability estimate, the dependence on α will be carefully traced all along the proof.

(a) *Preparations.* As in the previous proof, we make again the induction hypothesis (7.3) for some \bar{n} with $\bar{n}\tau \leq \bar{t}$, but this time with $\rho = c_0 h$ for some positive constant c_0 :

$$(8.4) \quad \|\mathbf{e}_h^n\|_{L^\infty} \leq c_0 h, \quad n < \bar{n}.$$

By an inverse inequality, this implies that $\|\mathbf{e}_h^n\|_{W^{1,\infty}}$ has an h - and τ -independent bound, and hence also $\|\mathbf{m}_h^n\|_{W^{1,\infty}}$ for $n < \bar{n}$. Together with (7.5), this implies

$$(8.5) \quad \|\widehat{\mathbf{m}}_h^n\|_{W^{1,\infty}} \leq C$$

and further

$$(8.6) \quad \|\widehat{\mathbf{e}}_h^n\|_{L^\infty} \leq Ch.$$

As in the Appendix, we aim to subtract η_k times the error equation for time step $n-1$ from the error equation for time step n , and then to test with $\varphi_h = \mathbf{P}_h(\widehat{\mathbf{m}}_h^n) \dot{\mathbf{e}}_h^n \in T_h(\widehat{\mathbf{m}}_h^n)$ (similarly as in the proof of Lemma 7.1). However, this is not possible directly due to the different test spaces at different time steps:

$$(8.7a) \quad \begin{aligned} & \alpha(\dot{\mathbf{e}}_h^n, \varphi_h) + (\widehat{\mathbf{e}}_h^n \times \dot{\mathbf{m}}_{*,h}^n, \varphi_h) \\ & + (\widehat{\mathbf{m}}_h^n \times \dot{\mathbf{e}}_h^n, \varphi_h) + (\nabla \mathbf{e}_h^n, \nabla \varphi_h) = -(\mathbf{r}_h^n, \varphi_h), \end{aligned}$$

for all $\varphi_h \in T_h(\widehat{\mathbf{m}}_h^n)$, and

$$(8.7b) \quad \begin{aligned} & \alpha(\dot{\mathbf{e}}_h^{n-1}, \psi_h) + (\widehat{\mathbf{e}}_h^{n-1} \times \dot{\mathbf{m}}_{*,h}^{n-1}, \psi_h) \\ & + (\widehat{\mathbf{m}}_h^{n-1} \times \dot{\mathbf{e}}_h^{n-1}, \psi_h) + (\nabla \mathbf{e}_h^{n-1}, \nabla \psi_h) = -(\mathbf{r}_h^{n-1}, \psi_h), \end{aligned}$$

for all $\psi_h \in T_h(\widehat{\mathbf{m}}_h^{n-1})$.

As in (7.2), we have

$$(8.8) \quad \varphi_h = \mathbf{P}_h(\widehat{\mathbf{m}}_h^n) \dot{\mathbf{e}}_h^n = \dot{\mathbf{e}}_h^n + \mathbf{q}_h^n, \quad \text{with } \mathbf{q}_h^n = -(\mathbf{P}_h(\widehat{\mathbf{m}}_h^n) - \mathbf{P}_h(\widehat{\mathbf{m}}_{*,h}^n)) \dot{\mathbf{m}}_{*,h}^n,$$

where \mathbf{q}_h^n is bounded by (7.8).

In turn, the test function $\psi_h = \mathbf{P}_h(\widehat{\mathbf{m}}_h^{n-1}) \dot{\mathbf{e}}_h^n \in T_h(\widehat{\mathbf{m}}_h^{n-1})$ is a perturbation of $\varphi_h = \dot{\mathbf{e}}_h^n + \mathbf{q}_h^n$, since using (8.8) we obtain

$$\begin{aligned} \psi_h &= \mathbf{P}_h(\widehat{\mathbf{m}}_h^{n-1}) \dot{\mathbf{e}}_h^n \\ &= \mathbf{P}_h(\widehat{\mathbf{m}}_h^n) \dot{\mathbf{e}}_h^n - (\mathbf{P}_h(\widehat{\mathbf{m}}_h^n) - \mathbf{P}_h(\widehat{\mathbf{m}}_h^{n-1})) \dot{\mathbf{e}}_h^n \\ &= \dot{\mathbf{e}}_h^n + \mathbf{q}_h^n + \mathbf{p}_h^n \quad \text{with } \mathbf{p}_h^n = -(\mathbf{P}_h(\widehat{\mathbf{m}}_h^n) - \mathbf{P}_h(\widehat{\mathbf{m}}_h^{n-1})) \dot{\mathbf{e}}_h^n. \end{aligned}$$

The perturbation \mathbf{p}_h^n is estimated using the second bound in Lemma 5.2 (i) with $p = \infty$, $q = 2$, and noting (8.5). We obtain

$$\begin{aligned} \|\mathbf{p}_h^n\|_{L^2} &\leq \|(\mathbf{P}_h(\widehat{\mathbf{m}}_h^n) - \mathbf{P}_h(\widehat{\mathbf{m}}_h^{n-1})) \dot{\mathbf{e}}_h^n\|_{L^2} \\ &\leq c \|\dot{\mathbf{e}}_h^n\|_{L^2} \|\widehat{\mathbf{m}}_h^n - \widehat{\mathbf{m}}_h^{n-1}\|_{L^\infty} \\ &\leq c \|\dot{\mathbf{e}}_h^n\|_{L^2} \left(\|\widehat{\mathbf{e}}_h^n\|_{L^\infty} + \|\widehat{\mathbf{m}}_{*,h}^n - \widehat{\mathbf{m}}_{*,h}^{n-1}\|_{L^\infty} + \|\widehat{\mathbf{e}}_h^{n-1}\|_{L^\infty} \right) \\ &\leq c \|\dot{\mathbf{e}}_h^n\|_{L^2} \left(\|\widehat{\mathbf{e}}_h^n\|_{L^\infty} + \sum_{j=0}^{k-1} |\gamma_j| \int_{t_{n-j-2}}^{t_{n-j-1}} \|\mathbf{R}_h \partial_t \mathbf{m}(t)\|_{L^\infty} dt + \|\widehat{\mathbf{e}}_h^{n-1}\|_{L^\infty} \right). \end{aligned}$$

We have $\|\mathbf{R}_h \partial_t \mathbf{m}(t)\|_{L^\infty} \leq c \|\partial_t \mathbf{m}(t)\|_{W^{1,\infty}}$ by [12, Theorem 8.1.11]. In view of (8.6) we obtain, for $\tau \leq \bar{C}h$,

$$(8.9) \quad \|\mathbf{p}_h^n\|_{L^2} \leq Ch \|\dot{\mathbf{e}}_h^n\|_{L^2},$$

and by an inverse estimate,

$$(8.10) \quad \|\nabla \mathbf{p}_h^n\|_{L^2} \leq C \|\dot{\mathbf{e}}_h^n\|_{L^2}.$$

We also recall the bound (7.9) for $\|\mathbf{r}_h^n\|_{L^2}$.

(b) *Energy estimates.* By subtracting $\eta_k(8.7b)$ from (8.7a) with the above choice of test functions, we obtain

$$(8.11) \quad \begin{aligned} & \alpha(\dot{\mathbf{e}}_h^n - \eta_k \dot{\mathbf{e}}_h^{n-1}, \dot{\mathbf{e}}_h^n + \mathbf{q}_h^n) + (\widehat{\mathbf{e}}_h^n \times \dot{\mathbf{m}}_{*,h}^n - \eta_k \widehat{\mathbf{e}}_h^{n-1} \times \dot{\mathbf{m}}_{*,h}^{n-1}, \dot{\mathbf{e}}_h^n + \mathbf{q}_h^n) \\ & + (\widehat{\mathbf{m}}_h^n \times \dot{\mathbf{e}}_h^n - \eta_k \widehat{\mathbf{m}}_h^{n-1} \times \dot{\mathbf{e}}_h^{n-1}, \dot{\mathbf{e}}_h^n + \mathbf{q}_h^n) + (\nabla \mathbf{e}_h^n - \eta_k \nabla \mathbf{e}_h^{n-1}, \nabla(\dot{\mathbf{e}}_h^n + \mathbf{q}_h^n)) \\ & - \eta_k [\alpha(\dot{\mathbf{e}}_h^{n-1}, \mathbf{p}_h^n) + (\widehat{\mathbf{e}}_h^{n-1} \times \dot{\mathbf{m}}_{*,h}^{n-1}, \mathbf{p}_h^n) \\ & \quad + (\widehat{\mathbf{m}}_h^{n-1} \times \dot{\mathbf{e}}_h^{n-1}, \mathbf{p}_h^n) + (\nabla \mathbf{e}_h^{n-1}, \nabla \mathbf{p}_h^n)] \\ & = -(\mathbf{r}_h^n - \eta_k \mathbf{r}_h^{n-1}, \dot{\mathbf{e}}_h^n + \mathbf{q}_h^n) - \eta_k (\mathbf{r}_h^{n-1}, \mathbf{p}_h^n). \end{aligned}$$

We estimate the terms of the error equation (8.11) separately and track carefully the dependence on η_k and α .

The term $\alpha(\dot{\mathbf{e}}_h^n - \eta_k \dot{\mathbf{e}}_h^{n-1}, \dot{\mathbf{e}}_h^n)$ is bounded from below, using Young's inequality and absorptions, by

$$\alpha(\dot{\mathbf{e}}_h^n - \eta_k \dot{\mathbf{e}}_h^{n-1}, \dot{\mathbf{e}}_h^n) \geq \alpha(1 - \frac{1}{2}\eta_k) \|\dot{\mathbf{e}}_h^n\|_{L^2}^2 - \frac{\alpha}{2}\eta_k \|\dot{\mathbf{e}}_h^{n-1}\|_{L^2}^2,$$

while the term $(\nabla \mathbf{e}_h^n - \eta_k \nabla \mathbf{e}_h^{n-1}, \nabla \dot{\mathbf{e}}_h^n)$ is bounded from below, via the relation (A.2) and (6.23), by

$$(\nabla \mathbf{e}_h^n - \eta_k \nabla \mathbf{e}_h^{n-1}, \nabla \dot{\mathbf{e}}_h^n) \geq \frac{1}{\tau} \left(\|\nabla \mathbf{E}_h^n\|_G^2 - \|\nabla \mathbf{E}_h^{n-1}\|_G^2 \right) + (\nabla \mathbf{e}_h^n - \eta_k \nabla \mathbf{e}_h^{n-1}, \nabla \mathbf{s}_h^n),$$

with $\mathbf{E}_h^n = (\mathbf{e}_h^{n-k+1}, \dots, \mathbf{e}_h^n)$, and where the G -weighted semi-norm is generated by the matrix $G = (g_{ij})$ from Lemma A.1 for the rational function $\delta(\zeta)/(1 - \eta_k \zeta)$.

The remaining terms outside the rectangular bracket are estimated using the Cauchy-Schwarz and Young inequalities (the latter often with a sufficiently small but fixed h - and τ -independent weighting factor $\mu > 0$) and $\|\widehat{\mathbf{m}}_h^n\|_{L^\infty} = 1$ and orthogonality. We obtain, with varying constants c (which depend on α and are inversely proportional to μ)

$$\begin{aligned} & \alpha(\dot{\mathbf{e}}_h^n - \eta_k \dot{\mathbf{e}}_h^{n-1}, \mathbf{q}_h^n) + (\widehat{\mathbf{e}}_h^n \times \dot{\mathbf{m}}_{*,h}^n - \eta_k \widehat{\mathbf{e}}_h^{n-1} \times \dot{\mathbf{m}}_{*,h}^{n-1}, \dot{\mathbf{e}}_h^n + \mathbf{q}_h^n) \\ & + (\widehat{\mathbf{m}}_h^n \times \dot{\mathbf{e}}_h^n - \eta_k \widehat{\mathbf{m}}_h^{n-1} \times \dot{\mathbf{e}}_h^{n-1}, \dot{\mathbf{e}}_h^n + \mathbf{q}_h^n) + (\nabla \mathbf{e}_h^n - \eta_k \nabla \mathbf{e}_h^{n-1}, \nabla \mathbf{q}_h^n) \\ & \leq (\alpha\mu + \mu + \frac{1}{2}\eta_k) \|\dot{\mathbf{e}}_h^n\|_{L^2}^2 + (\alpha\mu\eta_k + \frac{1}{2}\eta_k) \|\dot{\mathbf{e}}_h^{n-1}\|_{L^2}^2 \\ & \quad + c(\|\mathbf{q}_h^n\|_{L^2}^2 + \|\widehat{\mathbf{e}}_h^n\|_{L^2}^2 + \|\widehat{\mathbf{e}}_h^{n-1}\|_{L^2}^2) + \frac{1}{2}(\|\nabla \mathbf{e}_h^n\|_{L^2}^2 + \eta_k^2 \|\nabla \mathbf{e}_h^{n-1}\|_{L^2}^2 + \|\nabla \mathbf{q}_h^n\|_{L^2}^2) \\ & \leq (\alpha\mu + \mu + \frac{1}{2}\eta_k) \|\dot{\mathbf{e}}_h^n\|_{L^2}^2 + (\alpha\mu\eta_k + \frac{1}{2}\eta_k) \|\dot{\mathbf{e}}_h^{n-1}\|_{L^2}^2 + c \sum_{j=0}^k \|\mathbf{e}_h^{n-j-1}\|_{H^1}^2, \end{aligned}$$

where in the last inequality we used (7.12) and (7.13) to estimate $\widehat{\mathbf{e}}_h^n$.

The terms inside the rectangular bracket are bounded similarly, using (8.9) and (8.10) and the condition $\tau \leq \widetilde{C}h$, by

$$\begin{aligned} & \alpha(\dot{\mathbf{e}}_h^{n-1}, \mathbf{p}_h^n) + (\widehat{\mathbf{e}}_h^{n-1} \times \dot{\mathbf{m}}_{*,h}^{n-1}, \mathbf{p}_h^n) + (\widehat{\mathbf{m}}_h^{n-1} \times \dot{\mathbf{e}}_h^{n-1}, \mathbf{p}_h^n) + (\nabla \mathbf{e}_h^{n-1}, \nabla \mathbf{p}_h^n) \\ & \leq \mu \|\dot{\mathbf{e}}_h^{n-1}\|_{L^2}^2 + ch \|\dot{\mathbf{e}}_h^{n-1}\|_{L^2}^2 + c(\|\widehat{\mathbf{e}}_h^{n-1}\|_{L^2}^2 + \|\nabla \mathbf{e}_h^{n-1}\|_{L^2}^2) \\ & \leq \mu \|\dot{\mathbf{e}}_h^{n-1}\|_{L^2}^2 + c \sum_{j=0}^k \|\mathbf{e}_h^{n-j-1}\|_{H^1}^2. \end{aligned}$$

Here μ is an arbitrarily small positive constant (independent of τ and h), and c depends on the choice of μ .

In view of (7.9), the terms with the defects \mathbf{r}_h^n are bounded by

$$\begin{aligned} & -(\mathbf{r}_h^n - \eta_k \mathbf{r}_h^{n-1}, \dot{\mathbf{e}}_h^n + \mathbf{q}_h^n) - \eta_k (\mathbf{r}_h^{n-1}, \mathbf{p}_h^n) \\ & \leq \mu \|\dot{\mathbf{e}}_h^n\|_{L^2}^2 + c(\|\mathbf{r}_h^n\|_{L^2}^2 + \|\mathbf{r}_h^{n-1}\|_{L^2}^2 + \|\mathbf{q}_h^n\|_{L^2}^2) \\ & \leq \mu \|\dot{\mathbf{e}}_h^n\|_{L^2}^2 + c \sum_{j=0}^k \|\mathbf{e}_h^{n-j-1}\|_{L^2}^2 + c \sum_{j=0}^1 \|\mathbf{d}_h^{n-j}\|_{L^2}^2. \end{aligned}$$

Combination of these inequalities yields

$$\begin{aligned} & \left(\alpha(1 - \frac{1}{2}\eta_k) - \frac{1}{2}\eta_k - \mu\right) \|\dot{\mathbf{e}}_h^n\|_{L^2}^2 - \left(\frac{\alpha}{2}\eta_k + \frac{1}{2}\eta_k + \mu\alpha\eta_k\right) \|\mathbf{e}_h^{n-1}\|_{L^2}^2 \\ & + \frac{1}{\tau} \left(\|\nabla \mathbf{E}_h^n\|_G^2 - \|\nabla \mathbf{E}_h^{n-1}\|_G^2\right) \\ & \leq c \sum_{j=0}^k \|\mathbf{e}_h^{n-j-1}\|_{H^1}^2 + c \sum_{j=0}^1 \|\mathbf{d}_h^{n-j}\|_{L^2}^2 + c \|\nabla \mathbf{s}_h^n\|_{L^2}^2. \end{aligned}$$

Under condition (8.1) we have

$$\omega := \alpha(1 - \eta_k) - \eta_k > 0.$$

Multiplying both sides by τ and summing up from $k+1$ to n with $n \leq \bar{n}$ yields, for sufficiently small μ ,

$$\begin{aligned} & \frac{1}{2}\omega\tau \sum_{j=k+1}^n \|\dot{\mathbf{e}}_h^j\|_{L^2}^2 + \|\nabla \mathbf{E}_h^n\|_G^2 \\ & \leq c\tau \|\dot{\mathbf{e}}_h^k\|_{L^2}^2 + \|\nabla \mathbf{E}_h^k\|_G^2 + c\tau \sum_{j=0}^{n-1} \|\mathbf{e}_h^j\|_{H^1}^2 + c\tau \sum_{j=k}^n \|\mathbf{d}_h^j\|_{L^2}^2 + c\tau \sum_{j=k}^n \|\nabla \mathbf{s}_h^j\|_{L^2}^2. \end{aligned}$$

Moreover, testing the error equation (8.7a) for $n = k$ with $\dot{\mathbf{e}}_h^k + \mathbf{q}_h^k$, we obtain by more straightforward estimates

$$\tau \|\dot{\mathbf{e}}_h^k\|_{L^2}^2 + \|\mathbf{e}_h^k\|_{H^1}^2 \leq c\tau \left(\sum_{i=0}^{k-1} \|\mathbf{e}_h^i\|_{H^1}^2 + \|\mathbf{d}_h^k\|_{L^2}^2 + \|\nabla \mathbf{s}_h^k\|_{L^2}^2 \right),$$

which we insert in the above error bound. The proof is then completed using exactly the same arguments as in the last part of the proof of Lemma 7.1, by establishing an estimate between $\|\mathbf{e}_h^n\|_{L^2}^2$ and $\tau \sum_{j=k}^n \|\dot{\mathbf{e}}_h^j\|_{L^2}^2$ and using a discrete Gronwall inequality, and completing the induction step for (8.4). \square

9. NUMERICAL EXPERIMENTS

To obtain significant numerical results, we prescribe the exact solution \mathbf{m} on given three-dimensional domains $\Omega := [0, 1] \times [0, 1] \times [0, L]$ with $L \in \{1/100, 1/4\}$. The discretizations of these domains will consist of a few layers of elements in z -direction (one layer for $L = 1/100$ and ten layers for $L = 1/4$) and a later specified number of elements in x and y directions. This mimics the common case of thin film alloys as for example in the standard problems of the Micromagnetic Modeling Activity Group at NIST Center for Theoretical and Computational Materials Science (ctcms.nist.gov). Moreover, this mesh structure helps to keep the computational

requirements reasonable and allow us to compute the experiments on a desktop PC. We are aware that these experiments are only of preliminary nature and are just supposed to confirm the theoretical results. A more thorough investigation of the numerical properties of the developed method is needed. This will require us to incorporate preconditioning, parallelization of the computations, as well as lower order energy contributions in the effective field (1.3) to be able to compare to benchmark results from computational physics. This, however, is beyond the scope of this paper, and will be the topic of a subsequent work.

We consider the time interval $[0, \bar{t}]$ with $\bar{t} = 0.2$ and define two different exact solutions. Since within our computational budget either the time discretization error or the space discretization error dominates, we construct the solutions such that the first one is harder to approximate in space, while the second one is harder to approximate in time. Both solutions are constant in z -direction as is often observed in thin-film applications.

9.1. Implementation. The numerical experiments were conducted using the finite element package FEniCS (www.fenicsproject.org) on a desktop computer. As already discussed in Section 2.2, there are several ways to implement the tangent space restriction. We decided to solve a saddle point problem (variant (a) in Section 2.2) for simplicity of implementation. For preconditioning, we used the black-box AMG preconditioner that comes with FEniCS. Although this might not be the optimal solution, it keeps the number of necessary iterative solver steps within reasonable bounds. Assuming perfect preconditioning, the cost per time-step is then proportional to the number of mesh-elements. We observed this behavior approximately, although further research beyond the scope of this work is required to give a definite conclusion.

9.2. Exact solutions. We choose the damping parameter $\alpha = 0.2$ and define $g(t) := (\bar{t} + 0.1)/(\bar{t} + 0.1 - t)$ as well as $d(x) := (x_1 - 1/2)^2 + (x_2 - 1/2)^2$, which is the squared distance of the projection of x to $[0, 1] \times [0, 1]$ and the point $(1/2, 1/2)$. For some constant $C = 400$ (a choice made to have pronounced effects), define

$$(9.1) \quad \mathbf{m}(x, t) := \begin{cases} \begin{pmatrix} Ce^{-\frac{g(t)}{1/4-d(x)}}(x_1 - 1/2) \\ Ce^{-\frac{g(t)}{1/4-d(x)}}(x_2 - 1/2) \\ \sqrt{1 - C^2 e^{-2\frac{g(t)}{1/4-d(x)}} d(x)} \end{pmatrix} & \text{if } d(x) \leq \frac{1}{4}, \\ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} & \text{else.} \end{cases}$$

It is easy to check that $|\mathbf{m}(x, t)| = 1$ for all $(x, t) \in \Omega \times [0, \bar{t}]$. Moreover, $\partial_n \mathbf{m}(x, t) = 0$ for all $x \in \partial\Omega$. We may calculate the time derivative of \mathbf{m} in a straightforward fashion, i.e., $\partial_t \mathbf{m}(x, t) = 0$ for $d(x) > 1/4$ and

$$\partial_t \mathbf{m}(x, t) = \begin{cases} \begin{pmatrix} \frac{-g'(t)}{1/4-d(x)} Ce^{-\frac{g(t)}{1/4-d(x)}}(x_1 - 1/2) \\ \frac{-g'(t)}{1/4-d(x)} Ce^{-\frac{g(t)}{1/4-d(x)}}(x_2 - 1/2) \\ \frac{g'(t)}{1/4-d(x)} C^2 e^{-2\frac{g(t)}{1/4-d(x)}} \frac{d(x)}{\mathbf{m}_3(x, t)} \end{pmatrix} & \text{if } d(x) \leq \frac{1}{4}. \end{cases}$$

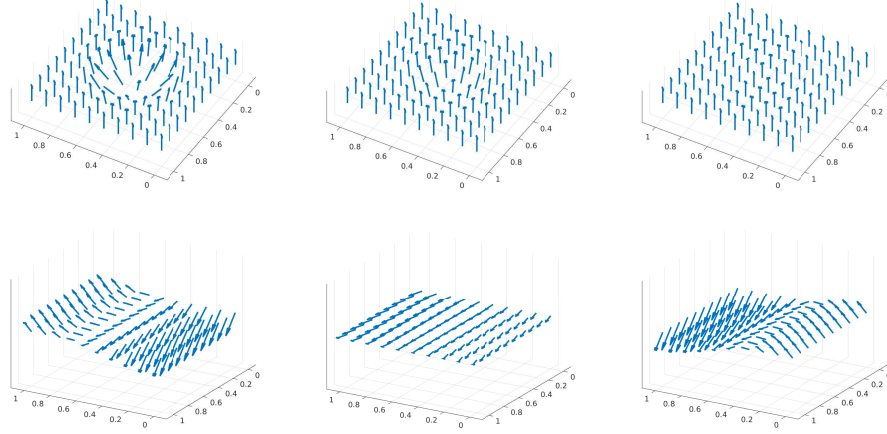


FIGURE 9.1. The first row shows the exact solution $\mathbf{m}(x, t)$ from (9.1) for $x \in [0, 1] \times [0, 1] \times \{0\}$ and $t \in \{0, 0.05, \bar{t}\}$ (from left to right), whereas the second row shows the exact solution $\widetilde{\mathbf{m}}(x, t)$ from (9.2) for $x \in [0, 1] \times [0, 1] \times \{0\}$ and $t \in \{0, 0.2/6, 0.2/3\}$ (from left to right). While the problems are three-dimensional, the solutions are constant in z -direction and we only show one slice of the solution.

Here, \mathbf{m}_3 denotes the third component of \mathbf{m} as defined above.

The second exact solution is defined via

$$(9.2) \quad \widetilde{\mathbf{m}}(x, t) := \begin{pmatrix} -(x_1^3 - 3x_1^2/2 + 1/4) \sin(3\pi t/\bar{t}) \\ \sqrt{1 - (x_1^3 - 3x_1^2/2 + 1/4)^2} \\ -(x_1^3 - 3x_1^2/2 + 1/4) \cos(3\pi t/\bar{t}) \end{pmatrix}.$$

Due to the polynomial nature in the first and the third component, and the well-behaved square-root, the space approximation error does not dominate the time approximation.

9.3. The experiments. We now may compute the corresponding forcings \mathbf{H} resp. $\widetilde{\mathbf{H}}$ to obtain the prescribed solutions by inserting into (1.4), i.e.,

$$\mathbf{H} = \alpha \partial_t \mathbf{m} + \mathbf{m} \times \partial_t \mathbf{m} - \Delta \mathbf{m}.$$

(Note that we may disregard the projection $\mathbf{P}(\mathbf{m})$ from (1.4) since we solve in the tangent space anyway.) We compute \mathbf{H} numerically by first interpolating \mathbf{m} and $\partial_t \mathbf{m}$ and then computing the derivatives. This introduces an additional error which is not accounted for in the theoretical analysis. However, the examples below confirm the expected convergence rates and hence we conclude that this additional perturbation is negligible. Figure 9.1 shows slices of the exact solution at different time steps. Figure 9.2 shows the convergence with respect to the time step size τ , while Figure 9.3 shows convergence with respect to the spatial mesh size h . All experiments confirm the expected rates for smooth solutions.

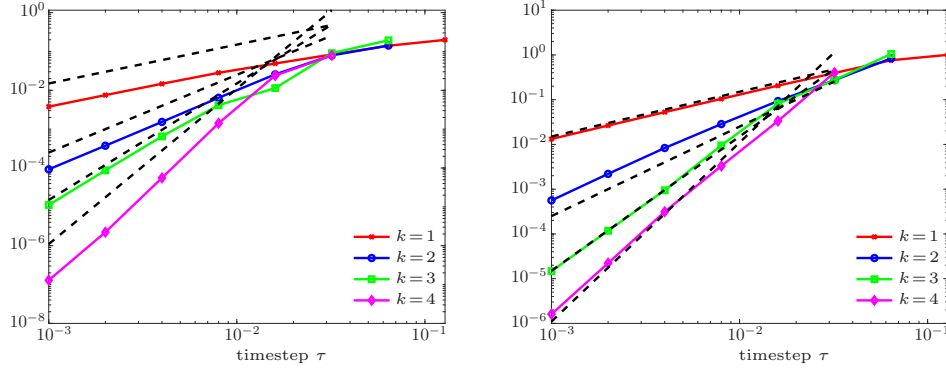


FIGURE 9.2. The plots show the error between computed solutions and exact solution $\widetilde{\mathbf{m}}$ for a given time stepsize with a spatial polynomial degree of $r = 2$ and a spatial mesh size $1/40$ which results in $\approx 6 \cdot 10^4$ degrees of freedom per time step in the left plot. In the right plot we use a thicker domain $D = [0, 1] \times [0, 1] \times [0, 1/4]$ with 10 elements in z -direction. This results in $\approx 4 \cdot 10^5$ degrees of freedom per timestep. We use the k -step methods of order $k \in \{1, 2, 3, 4\}$ and observe the expected rates $\mathcal{O}(\tau^k)$ indicated by the dashed lines. The coarse levels of the higher order methods are missing because the k th step is already beyond the final time \bar{t} .

Finally, we consider an example with nonsmooth initial data and constant right-hand side. The initial data are given by

$$(9.3) \quad \mathbf{m}_0(x) := \begin{pmatrix} x_1 - 1/2 \\ x_2 - 1/2 \\ \sqrt{1 - d(x)} \end{pmatrix} \text{ if } d(x) \leq \frac{1}{4} \text{ and } \mathbf{m}_0(x) := \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \text{ else.}$$

With the constant forcing field $\mathbf{H} := (0, 1, 1)^T$ we compute a numerical approximation to the unknown exact solution. Note that we do not expect any smoothness of the solution (even the initial data is not smooth). Figure 9.4 nevertheless shows a physically consistent decay of the energy $\|\nabla \mathbf{m}(t)\|_{L^2(\Omega)^3}$ over time as well as a good agreement between different orders of approximation. Moreover, the computed approximation shows little deviation from unit length as would be expected for smooth solutions.

APPENDIX: ENERGY ESTIMATES FOR BACKWARD DIFFERENCE FORMULAE

The stability proofs of this paper rely on energy estimates, that is, on the use of positive definite bilinear forms to bound the error e in terms of the defect d . This is, of course, a basic technique for studying the time-continuous problem and also for backward Euler and Crank–Nicolson time discretizations (see, e.g., Thomée [38]), but energy estimates still appear to be not well known for backward difference formula (BDF) time discretizations of order up to 5, which are widely used for solving stiff ordinary differential equations. To illustrate the basic mechanism, we here just consider the prototypical linear parabolic evolution equation in its weak

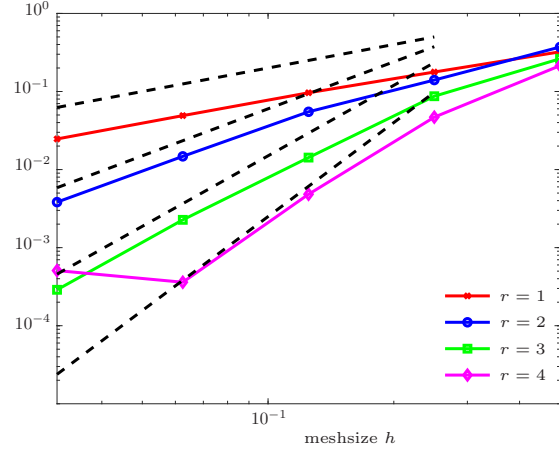


FIGURE 9.3. The plot shows convergence in meshsize h with respect to the exact solution \mathbf{m} from (9.1) on the domain $D = [0, 1] \times [0, 1] \times [0, 1/100]$ with one layer of elements in z -direction. We used the second order BDF method with $\tau = 10^{-3}$ and spatial polynomial degrees $r \in \{1, 2, 3, 4\}$. The mesh sizes range from $1/2$ to $1/32$. We observe the expected rates $\mathcal{O}(h^r)$ indicated by the dashed lines. The finest mesh-size for $r = 4$ does not reach the expected error level. This is due to the fact that the time-discretization errors start to dominate in that region.

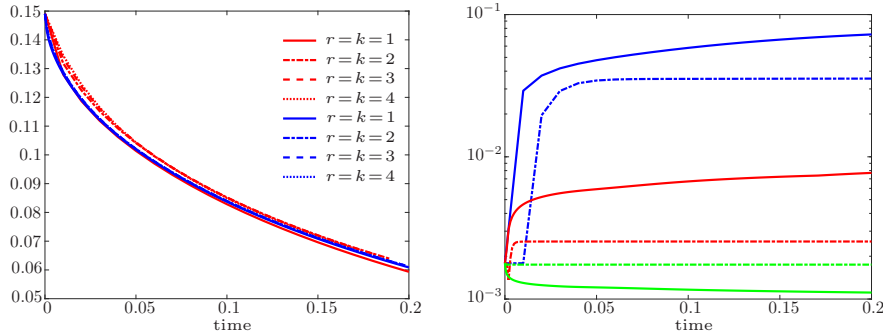


FIGURE 9.4. Left plot: Decay of energies $\|\nabla \mathbf{m}(t)\|_{L^2(\Omega)^3}$ for the approximations to the unknown solution with \mathbf{m}_0 and \mathbf{H} given in (9.3) and one line after (9.3). We plot four approximations of the k -step method with polynomial degree r for $r = k \in \{1, 2, 3, 4\}$. The spatial mesh-size is $1/40$ and the size of the timesteps is 10^{-3} (blue) and 10^{-2} (red). Right plot: Deviation from unit length $\|1 - |\mathbf{m}(t)|^2\|_{L^\infty(\Omega)}$ plotted over time for step sizes $\tau = 10^{-2}$ (blue), $\tau = 10^{-3}$ (red), and $\tau = 10^{-4}$ (green). The solid lines indicate $k = 1$, whereas the dashed lines indicate $k = 2$. The spatial mesh-size is $1/40$ with $r = 1$.

formulation, given by two positive definite symmetric bilinear forms (\cdot, \cdot) and $a(\cdot, \cdot)$ on Hilbert spaces H and V with induced norms $|\cdot|$ and $\|\cdot\|$, respectively, and with V densely and continuously embedded in H . The problem then is to find $u(t) \in V$ such that

$$(A.1) \quad (\partial_t u, v) + a(u, v) = (f, v) \quad \forall v \in V,$$

with initial condition $u(0) = u_0$. If u^* is a function that satisfies the equation up to a defect d , that is,

$$(\partial_t u^*, v) + a(u^*, v) = (f, v) + (d, v) \quad \forall v \in V,$$

then the error $e = u - u^*$ satisfies, in this linear case, an equation of the same form,

$$(\partial_t e, v) + a(e, v) = (d, v) \quad \forall v \in V,$$

with initial value $e_0 = u_0 - u_0^*$. Testing with $v = e$ yields

$$\frac{1}{2} \frac{d}{dt} |e|^2 + \|e\|^2 = (d, e).$$

Estimating the right-hand side by $(d, e) \leq \|d\|_* \|e\| \leq \frac{1}{2} \|d\|_*^2 + \frac{1}{2} \|e\|^2$, with the dual norm $\|\cdot\|_*$, and integrating from time 0 to t results in the error bound

$$|e(t)|^2 \leq |e(0)|^2 + \int_0^t \|d(s)\|_*^2 ds.$$

On the other hand, testing with $v = \partial_t e$ yields

$$|\partial_t e|^2 + \frac{1}{2} \frac{d}{dt} \|e\|^2 = (d, \partial_t e),$$

which leads similarly to the error bound

$$\|e(t)\|^2 \leq \|e(0)\|^2 + \int_0^t |d(s)|^2 ds.$$

This procedure is all-familiar, but it is not obvious how to extend it to time discretizations beyond the backward Euler and Crank–Nicolson methods. The use of energy estimates for BDF methods relies on the following remarkable results.

Lemma A.1. (Dahlquist [18]; see also [8] and [27, Section V.6]) *Let $\delta(\zeta) = \delta_k \zeta^k + \dots + \delta_0$ and $\mu(\zeta) = \mu_k \zeta^k + \dots + \mu_0$ be polynomials of degree at most k (and at least one of them of degree k) that have no common divisor. Let (\cdot, \cdot) be an inner product with associated norm $|\cdot|$. If*

$$\operatorname{Re} \frac{\delta(\zeta)}{\mu(\zeta)} > 0 \quad \text{for } |\zeta| < 1,$$

then there exists a positive definite symmetric matrix $G = (g_{ij}) \in \mathbb{R}^{k \times k}$ such that for v_0, \dots, v_k in the real inner product space,

$$\left(\sum_{i=0}^k \delta_i v_{k-i}, \sum_{j=0}^k \mu_j v_{k-j} \right) \geq \sum_{i,j=1}^k g_{ij} (v_i, v_j) - \sum_{i,j=1}^k g_{ij} (v_{i-1}, v_{j-1}).$$

In combination with the preceding result for the multiplier $\mu(\zeta) = 1 - \eta_k \zeta$, the following property of BDF methods up to order 5 becomes important.

Lemma A.2. (Nevanlinna & Odeh [34]) *For $k \leq 5$, there exists $0 \leq \eta_k < 1$ such that for $\delta(\zeta) = \sum_{\ell=1}^k \frac{1}{\ell}(1-\zeta)^\ell$,*

$$\operatorname{Re} \frac{\delta(\zeta)}{1 - \eta_k \zeta} > 0 \quad \text{for } |\zeta| < 1.$$

The smallest possible values of η_k are

$$\eta_1 = \eta_2 = 0, \quad \eta_3 = 0.0836, \quad \eta_4 = 0.2878, \quad \eta_5 = 0.8160.$$

Precise expressions for the optimal multipliers for the BDF methods of orders 3, 4 and 5 are given by Akrivis & Katsoprinakis [1].

An immediate consequence of Lemma A.2 and Lemma A.1 is the relation

$$(A.2) \quad \left(\sum_{i=0}^k \delta_i v_{k-i}, v_k - \eta_k v_{k-1} \right) \geq \sum_{i,j=1}^k g_{ij}(v_i, v_j) - \sum_{i,j=1}^k g_{ij}(v_{i-1}, v_{j-1})$$

with a positive definite symmetric matrix $G = (g_{ij}) \in \mathbb{R}^{k \times k}$; it is this inequality that plays a crucial role in our energy estimates, and the same inequality for the inner product $a(\cdot, \cdot)$.

The error equation for the BDF time discretization of the linear parabolic problem (A.1) reads

$$(\dot{e}^n, v) + a(e^n, v) = (d^n, v) \quad \forall v \in V, \quad \text{where} \quad \dot{e}^n = \frac{1}{\tau} \sum_{j=0}^k \delta_j e^{n-j},$$

with starting errors e^0, \dots, e^{k-1} . When we test with $v = e^n - \eta_k e^{n-1}$, the first term can be estimated from below by (A.2), the second term is bounded from below by $(1 - \frac{1}{2}\eta_k)\|e^n\|^2 - \frac{1}{2}\eta_k\|e^{n-1}\|^2$, and the right-hand term is estimated from above by the Cauchy–Schwarz inequality. Summing up from k to n then yields the error bound

$$(A.3) \quad |e^n|^2 + \tau \sum_{j=k}^n \|e^j\|^2 \leq C_k \left(\sum_{i=0}^{k-1} (|e^i|^2 + \tau \|e^i\|^2) + \tau \sum_{j=k}^n \|d^j\|_*^2 \right),$$

where C_k depends only on the order k of the method. This kind of estimate for the BDF error has recently been used for a variety of linear and nonlinear parabolic problems [33, 3, 2, 30].

On the other hand, when we first subtract η_k times the error equation for $n-1$ from the error equation with n and then test with \dot{e}^n , we obtain

$$(\dot{e}^n - \eta_k \dot{e}^{n-1}, \dot{e}^n) + a(e^n - \eta_k e^{n-1}, \dot{e}^n) = (d^n - \eta_k d^{n-1}, \dot{e}^n).$$

Here, the second term is bounded from below by (A.2) with the $a(\cdot, \cdot)$ inner product, the first term is bounded from below by $(1 - \frac{1}{2}\eta_k)|\dot{e}^n|^2 - \frac{1}{2}\eta_k|\dot{e}^{n-1}|^2$, and the right-hand term is estimated from above by the Cauchy–Schwarz inequality. Summing up from $k+1$ to n then yields the error bound

$$(A.4) \quad \|e^n\|^2 + \tau \sum_{j=k+1}^n |\dot{e}^j|^2 \leq c_k \left(\sum_{i=0}^k \|e^i\|^2 + \eta_k \tau |\dot{e}^k|^2 + \tau \sum_{j=k}^n |d^j|^2 \right).$$

Moreover, testing the error equation for $n = k$ with \dot{e}^k and once again using the Cauchy–Schwarz and Young inequalities yields

$$\|e^k\|^2 + \tau|\dot{e}^k|^2 \leq \tilde{c}_k \left(\sum_{i=0}^{k-1} \|e^i\|^2 + \tau|d^k|^2 \right).$$

When we insert this bound in (A.4), we finally arrive at the bound, valid for $n \geq k$,

$$(A.5) \quad \|e^n\|^2 + \tau \sum_{j=k}^n |\dot{e}^j|^2 \leq C_k \left(\sum_{i=0}^{k-1} \|e^i\|^2 + \tau \sum_{j=k}^n |d^j|^2 \right).$$

It is this type of estimate that we use in the present paper for the nonlinear problem considered here. This technique has previously been used in [29].

Acknowledgment. The work of Michael Feischl, Balázs Kovács and Christian Lubich is supported by Deutsche Forschungsgemeinschaft – Project-ID 258734477 – SFB 1173.

REFERENCES

- [1] G. Akrivis and E. Katsoprinakis, *Backward difference formulae: new multipliers and stability properties for parabolic equations*, Math. Comp. **85** (2016) 2195–2216.
- [2] G. Akrivis, B. Li, and C. Lubich, *Combining maximal regularity and energy estimates for time discretizations of quasilinear parabolic equations*, Math. Comp. **86** (2017) 1527–1552.
- [3] G. Akrivis and C. Lubich, *Fully implicit, linearly implicit and implicit–explicit backward difference formulae for quasi-linear parabolic equations*, Numer. Math. **131** (2015) 713–735.
- [4] F. Alouges, *A new finite element scheme for Landau–Lifshitz equations*, Discrete Contin. Dyn. Syst. Ser. S. **1** (2008) 187–196.
- [5] F. Alouges and P. Jaisson, *Convergence of a finite element discretization for the Landau–Lifshitz equations in micromagnetism*, Math. Models Methods Appl. Sci. **16** (2006) 299–316.
- [6] F. Alouges, E. Kritsikis, J. Steiner, and J.-C. Toussaint, *A convergent and precise finite element scheme for Landau–Lifshitz–Gilbert equation*, Numer. Math. **128** (2014) 407–430.
- [7] R. An, *Optimal error estimates of linearized Crank–Nicolson Galerkin method for Landau–Lifshitz equation*, J. Sci. Comput. **69** (2016) 1–27.
- [8] C. Baiocchi and M. Crouzeix, *On the equivalence of A-stability and G-stability*, Appl. Numer. Math. **5** (1989) 19–22.
- [9] R. E. Bank and H. Yserentant, *On the H^1 -stability of the L^2 -projection onto finite element spaces*, Numer. Math. **126** (2014) 361–381.
- [10] E. Barati, M. Cinal, D. M. Edwards, and A. Umerski, *Gilbert damping in magnetic layered systems*, Phys. Rev. B **90** (2014) 014420.
- [11] S. Bartels and A. Prohl, *Convergence of an implicit finite element method for the Landau–Lifshitz–Gilbert equation*, SIAM J. Numer. Anal. **44** (2006) 1405–1419.
- [12] S. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Texts in Applied Mathematics, Springer, New York, 2008.
- [13] F. Brezzi, *On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers*, Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge **8** (1974) 129–151.
- [14] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.
- [15] P. Ciarlet Jr., J. Huang, and J. Zou, *Some observations on generalized saddle-point problems*, SIAM J. Matrix Anal. Appl. **25** (2003) 224–236.
- [16] I. Cimrák, *Error estimates for a semi-implicit numerical scheme solving the Landau–Lifshitz equation with an exchange field*, IMA J. Numer. Anal. **25** (2005) 611–634.
- [17] I. Cimrák, *A survey on the numerics and computations for the Landau–Lifshitz equation of micromagnetism*, Arch. Comput. Methods Eng. **15** (2008) 277–309.
- [18] G. Dahlquist, *G-stability is equivalent to A-stability*, BIT **18** (1978) 384–401.

- [19] G. Di Fratta, C. M. Pfeiler, D. Praetorius, M. Ruggeri, and B. Stiftner, *Linear second-order IMEX-type integrator for the (eddy current) Landau–Lifshitz–Gilbert equation*, IMA J. Numer. Anal. **40** (2020) 2802–2838.
- [20] J. Douglas Jr., T. Dupont, and L. Wahlbin, *The stability in L^q of the L^2 -projection into finite element function spaces*, Numer. Math. **23** (1974/75) 193–197.
- [21] M. Feischl and T. Tran, *The eddy current–LLG equations: FEM–BEM coupling and a priori error estimates*, SIAM J. Numer. Anal. **55** (2017) 1786–1819.
- [22] M. Feischl and T. Tran, *Existence of regular solutions of the Landau–Lifshitz–Gilbert equation in 3D with natural boundary conditions*, SIAM J. Math. Anal. **49** (2017) 4470–4490.
- [23] H. Gao, *Optimal error estimates of a linearized backward Euler FEM for the Landau–Lifshitz equation*, SIAM J. Numer. Anal. **52** (2014) 2574–2593.
- [24] I. Garate and A. H. MacDonald, *Influence of a transport current on magnetic anisotropy in gyrotropic ferromagnets*, Phys. Rev. B **80** (2009) 134403.
- [25] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier–Stokes Equations: Theory and Algorithms*, Springer-Verlag, Berlin, 1986.
- [26] B. Guo and S. Ding, *Landau–Lifshitz Equations*, World Scientific, Hackensack, NJ, 2008.
- [27] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II: Stiff and Differential–Algebraic Problems*, 2nd revised ed., Springer-Verlag, Berlin Heidelberg, Springer Series in Computational Mathematics v. 14, 2002.
- [28] S. Jaffard, *Propriétés des matrices “bien localisées” près de leur diagonale et quelques applications*, Ann. Inst. H. Poincaré Non Linéaire **7** (1990) 461–476.
- [29] B. Kovács, B. Li, and C. Lubich, *A convergent evolving finite element algorithm for mean curvature flow of closed surfaces*, Numer. Math. **143** (2019) 797–853.
- [30] B. Kovács and C. Lubich, *Numerical analysis of parabolic problems with dynamic boundary conditions*, IMA J. Numer. Anal. **37** (2017) 1–39.
- [31] E. Kritsikis, A. Vaysset, L. D. Buda-Prejbeanu, F. Alouges, and J.-C. Toussaint, *Beyond first-order finite element schemes in micromagnetics*, J. Comput. Phys. **256** (2014) 357–366.
- [32] D. Kröner, *Numerical Schemes for Conservation Laws*, John Wiley & Sons, Ltd., Chichester; B. G. Teubner, Stuttgart, 1997.
- [33] C. Lubich, D. Mansour, and C. Venkataraman, *Backward difference time discretization of parabolic differential equations on evolving surfaces*, IMA J. Numer. Anal. **33** (2013) 1365–1385.
- [34] O. Nevanlinna and F. Odeh, *Multiplier techniques for linear multistep methods*, Numer. Funct. Anal. Optim. **3** (1981) 377–423.
- [35] D. Praetorius, M. Ruggeri, and B. Stiftner, *Convergence of an implicit-explicit midpoint scheme for computational micromagnetics*, Comput. Math. Appl. **75** (2018) 1719–1738.
- [36] A. Prohl, *Computational Micromagnetism*, Teubner, Stuttgart, 2001.
- [37] D. Suess, V. Tsiantos, T. Schrefl, J. Fidler, W. Scholz, H. Forster, R. Dittrich, and J. J. Miles, *Time resolved micromagnetics using a preconditioned time integration method*, J. Magn. Magn. Mater. **248** (2002) 298–311.
- [38] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, 2nd ed., Springer-Verlag, Berlin, 2006.
- [39] D. Thonig and J. Henk, *Gilbert damping tensor within the breathing Fermi surface model: anisotropy and non-locality*, New J. Phys. **16** (2014) 013032.

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING, UNIVERSITY OF IOANNINA, 451 10 IOANNINA, GREECE, AND INSTITUTE OF APPLIED AND COMPUTATIONAL MATHEMATICS, FORTH, 700 13 HERAKLION, CRETE, GREECE

Email address: `akrivis@cse.uoi.gr`

INSTITUTE FOR ANALYSIS AND SCIENTIFIC COMPUTING (E 101), TECHNICAL UNIVERSITY WIEN, WIEDNER HAUPTSTRASSE 8-10, 1040 VIENNA, AUSTRIA

Email address: `michael.feischl@tuwien.ac.at`

MATHEMATISCHES INSTITUT, UNIVERSITÄT TÜBINGEN, AUF DER MORGENSTELLE, D-72076 TÜBINGEN, GERMANY

Email address: `kovacs@na.uni-tuebingen.de`

MATHEMATISCHES INSTITUT, UNIVERSITÄT TÜBINGEN, AUF DER MORGENSTELLE, D-72076 TÜBINGEN, GERMANY

Email address: `lubich@na.uni-tuebingen.de`