Error estimates for general non-linear Cahn–Hilliard equations on evolving surfaces

Cedric Aaron Beschle and Balázs Kovács

Abstract In this paper, we consider the Cahn–Hilliard equation on evolving surfaces with prescribed velocity and a general non-linear potential. High-order evolving surface finite elements are used to discretise the weak equation system in space, and a modified matrix–vector formulation for the semi-discrete problem is derived. The anti-symmetric structure of the weak equation system is preserved by the matrix–vector formulation and it is utilised to prove optimal-order and uniform-in-time error estimates. An extension of the convergence results is given for general non-linear Cahn–Hilliard equations on evolving surfaces. The paper is concluded by a variety of numerical experiments.

Keywords Cahn–Hilliard equation · evolving surfaces · evolving surface finite elements · error estimates · stability · energy estimates · general non-linear problems

Mathematics Subject Classification (2000) 35R01

1 Introduction

This paper studies the Cahn–Hilliard equation on evolving surfaces with prescribed surface velocity. The non-linear potential is only assumed to satisfy locally Lipschitz-type assumptions. The Cahn–Hilliard equation is formulated as a system of second order equations, exhibiting an anti-symmetric structure. The semi-discretisation of the system by higher order evolving surface finite
elements, cf. [10, 18], preserves this anti-symmetric structure, which is utilised to prove a new convergence result. Optimal-order uniform-in-time error estimates in the $L^2$ and $H^1$ norms for both solution variables are proven.

Cahn and Hilliard first described an equation modelling phase separation processes in [5]. Which since found many applications, e.g. in foam modelling, dendritic flow, phase separation models, and image processing.

The Cahn–Hilliard equation on a stationary surface with boundary was first investigated by Du, Ju and Tian in [9]. They study a full discretisation of the Cahn–Hilliard equation with homogeneous Dirichlet boundary conditions, and prove optimal-order error estimates in the $L^2$ norm for $u$, using linear finite elements. Elliott and Ranner were the first to consider the Cahn–Hilliard equation on a closed evolving surface with a prescribed velocity in [15]. They proved optimal-order uniform-in-time error estimates in the $L^2$ and $H^1$ norms in the first variable and optimal-order $L^2$-in-time error estimates in the $L^2$ and $H^1$ norms in the second variable, using a discretisation by linear evolving surface finite elements. Using a new stability proof, the results of this paper improve the error estimates for the second variable from optimal-order $L^2$-in-time to optimal-order uniform-in-time estimates. Furthermore, the proof is built, such that it is extendable to general non-linear Cahn–Hilliard equations in a straightforward way.

The main result of this paper, optimal-order uniform-in-time error estimates, is proven by studying the stability and consistency of the semi-discretisation.

In the stability analysis, the difference between the Ritz map of the exact solution and the numerical solution is estimated in terms of defects and their time derivatives. To account for initial errors in the second variable, a modification of the second-order system is required. The stability proof uses energy estimates, performed in the matrix–vector formulation, and utilises the anti-symmetric structure of the error equations, testing the error equations with the errors and also with their time derivatives. The stability analysis was first developed for Willmore flow in [20]. A uniform-in-time $L^\infty$ bound for the numerical solution is key to estimate the non-linear term. It is obtained from the time-uniform $H^1$ norm error bounds using an inverse estimate and exists for a small time due to a continuous initial function. The stability proof is independent of geometric errors.

We also give details to an extension of the stability proof to rather general non-linear Cahn–Hilliard equations, which include proliferation terms [27, equation (3.1)], where an additional non-linearity appears in both variables in both equations of the system. The generality of the stability proof can also be seen through the related results in [20] and [17]. A further advantage of this stability proof, is that it is expected to be generalisable to a full discrete method based on linearly implicit backward difference methods, on which we intend to report in a subsequent work.

In the consistency analysis the $L^2$ norms of the defects and their time derivatives are estimated. The bounds use geometric error estimates, including interpolation and Ritz map error estimates, bounds on the discrete surface
velocity, and geometric approximation errors for high-order evolving surface finite elements, see [18].

The paper is structured as follows. In Section 2, based on the works [10] and [15], the weak formulation for the Cahn–Hilliard equation on evolving surfaces is derived as a system of equations. In Section 3 the evolving surface finite element method is used to discretise this system of equations in space. The obtained semi-discrete problem is written as a matrix–vector formulation. In Section 4 the novel error estimates proved in this work are stated and discussed in comparison to the existing results by Elliott and Ranner [15]. Section 5 contains the stability part of the proof. Section 6 treats the consistency part of the proof. In Section 7 the two parts are combined to prove the main result. In Section 8 the result is extended to general non-linear Cahn–Hilliard equations. In Section 9 a full discretisation to the problem is given, cf. [2,1]. In Section 10 the theoretical results are complemented by numerical experiments.

2 Cahn–Hilliard equation on evolving surfaces

In the following we consider a smoothly evolving closed hypersurface \( \Gamma(t) \subset \mathbb{R}^{d+1} \), with \( d \leq 3 \), for \( 0 \leq t \leq T \). The initial surface \( \Gamma(0) = \Gamma^0 \) is given (and at least \( C^2 \)), and it evolves with the given and sufficiently smooth velocity \( v \). The surface \( \Gamma(t) \) is given as the image of a smooth mapping \( X : \Gamma^0 \times [0,T] \to \mathbb{R}^{d+1} \), by \( \Gamma(t) = \{ X(p,t) \mid p \in \Gamma^0 \} \). The embedding \( X \) and the velocity \( v \) satisfy the ordinary differential equation (ODE):

\[
\partial_t X(p,t) = v(X(p,t),t), \quad p \in \Gamma^0, \quad 0 \leq t \leq T. \tag{2.1}
\]

Let \( \nu \) denote the unit outward normal vector to \( \Gamma(t) \). Then the surface (or tangential) gradient on \( \Gamma(t) \), of a function \( u : \Gamma(t) \to \mathbb{R} \), is denoted by \( \nabla \! |_{\Gamma(t)} u \), and is given by \( \nabla \! |_{\Gamma(t)} u = \nabla \bar{u} - (\nabla \bar{u} \cdot \nu) \nu \) (the surface gradient is independent of the extension \( \bar{u} \)), while the Laplace–Beltrami operator on \( \Gamma(t) \) is given by \( \Delta \! |_{\Gamma(t)} u = \nabla \! |_{\Gamma(t)} \cdot \nabla \! |_{\Gamma(t)} u \). Moreover, \( \partial^\bullet u \) denotes the material derivative of \( u \), i.e. \( \partial^\bullet u(\cdot,t) = d/dt(u(X(\cdot,t),t)) = \partial_t u(\cdot,t) + v \cdot \nabla u(\cdot,t) \). The space-time manifold will be denoted by \( \mathcal{G}_T = \bigcup_{t \in [0,T]} \Gamma(t) \times \{t\} \). For more details on these notions we refer to [10,11].

In this paper we consider the Cahn–Hilliard equation on evolving surfaces.

It is the fourth-order equation for a scalar function \( u : \mathcal{G}_T \to \mathbb{R} \) given by

\[
\partial^\bullet u + u(\nabla \! |_{\Gamma(t)} v) = \Delta \! |_{\Gamma(t)} \left( \varepsilon \Delta \! |_{\Gamma(t)} u + \varepsilon^{-1} W'(u) \right) \quad \text{on} \ \Gamma(t), \tag{2.2}
\]

with continuous (and sufficiently regular) initial conditions \( u(\cdot,0) = u^0 \) on the initial surface \( \Gamma^0 \). The parameter \( \varepsilon > 0 \) describes the length of the transition regions, we are however not interested in taking the limit \( \varepsilon \to 0 \). The non-linear scalar function \( W : \mathbb{R} \to \mathbb{R} \) is a chemical potential and here it is only assumed to have locally Lipschitz continuous derivatives \( W' \) and \( W'' \). A typical example is a double well potential, i.e. \( W(u) = (u^2 - 1)^2 \). The solution \( u \in [-1,1] \)
models the concentration of surfactant fluids, with $u = \pm 1$ indicating the pure occurrences of each, cf. [5].

The Cahn–Hilliard equation on a stationary surface $\Gamma$ can be derived as the $H^{-1}(\Gamma)$-gradient flow of the Ginzburg–Landau energy

$$\int_\Gamma \left( \frac{\varepsilon}{2} |\nabla_{\Gamma} u|^2 + \varepsilon^{-1} W(u) \right),$$

(2.3)

cf. [15, Remark 2.1]. In [15] it is stated, that to obtain a gradient flow on an evolving surface, a model for the surface velocity $v$ is needed, leading to a coupled system for $u$ and $v$. In the evolving surface case, $w = -\varepsilon \Delta_{\Gamma(t)} u + \varepsilon^{-1} W'(u)$ is the variation of the evolving surface Ginzburg–Landau energy, see [29].

2.1 Second order system and weak formulation

By introducing the auxiliary function $w : \mathcal{G}_T \to \mathbb{R}$ we rewrite the Cahn–Hilliard equation (2.2) into a system of second order partial differential equations: For $u, w : \mathcal{G}_T \to \mathbb{R}$

$$\partial^* u - \Delta_{\Gamma(t)} w = -u(\nabla_{\Gamma(t)} \cdot v) \quad \text{on } \Gamma(t),$$

(2.4a)

$$w + \varepsilon \Delta_{\Gamma(t)} u = \varepsilon^{-1} W'(u) \quad \text{on } \Gamma(t),$$

(2.4b)

with initial data $u(0, \cdot) = u_0$ on $\Gamma^0$.

On the evolving surface $\Gamma(t)$ we recall the definition of standard Sobolev spaces $L^2(\Gamma(t))$, and $H^1(\Gamma(t))$ and its higher order variants, endowed with their usual norms, see [10,12].

The weak formulation of the Cahn–Hilliard system (2.4) reads: Find $u(\cdot, t) \in H^1(\Gamma(t))$ with a time-continuous material derivative $\partial^* u(\cdot, t) \in L^2(\Gamma(t))$ and $w(\cdot, t) \in H^1(\Gamma(t))$ such that for all test functions $\varphi^u(\cdot, t) \in H^1(\Gamma(t))$ and $\varphi^w(\cdot, t) \in H^1(\Gamma(t))$

$$\int_{\Gamma(t)} \partial^* u \varphi^w + \int_{\Gamma(t)} \nabla_{\Gamma(t)} w \cdot \nabla_{\Gamma(t)} \varphi^u = -\int_{\Gamma(t)} u \varphi^u(\nabla_{\Gamma(t)} \cdot v),$$

(2.5a)

$$\int_{\Gamma(t)} w \varphi^w - \varepsilon \int_{\Gamma(t)} \nabla_{\Gamma(t)} u \cdot \nabla_{\Gamma(t)} \varphi^w = \int_{\Gamma(t)} \varepsilon^{-1} W'(u) \varphi^w,$$

(2.5b)

with initial data $u(0, \cdot) = u_0$ on $\Gamma^0$.

It is important to note here that the anti-symmetric structure (up to the parameter $\varepsilon$) of the above systems ((2.4) and (2.5)) will serve as a key property which will be heavily used in the stability analysis.

Using the Leibniz formula [10], an equivalent weak form reads as: Find $u(\cdot, t) \in H^1(\Gamma(t))$ with a time-continuous material derivative $\partial^* u(\cdot, t) \in L^2(\Gamma(t))$
and \( w(\cdot, t) \in H^1(\Gamma(t)) \) such that for all test functions \( \varphi^u(\cdot, t) \in H^1(\Gamma(t)) \), with \( \partial^* \varphi^w(\cdot, t) = 0 \), and \( \varphi^w(\cdot, t) \in H^1(\Gamma(t)) \)

\[
\frac{d}{dt} \left( \int_{\Gamma(t)} w \varphi^u \right) + \int_{\Gamma(t)} \nabla_{\Gamma(t)} w \cdot \nabla_{\Gamma(t)} \varphi^u = 0, \quad \text{(2.6a)}
\]

\[
\int_{\Gamma(t)} w \varphi^w - \varepsilon \int_{\Gamma(t)} \nabla_{\Gamma(t)} u \cdot \nabla_{\Gamma(t)} \varphi^w = \int_{\Gamma(t)} \varepsilon^{-1} W'(u) \varphi^w. \quad \text{(2.6b)}
\]

For existence, uniqueness, and regularity results we refer to [15, Section 4], and recap below.

Assuming that the initial condition \( u_0 \in H^2(\Gamma_0) \) has bounded energy (2.3) then, there exists a unique weak solution pair \( (u, w) \) to the Cahn–Hilliard equation (2.5), whose energy stays bounded by the energy of the initial data. Furthermore, \( u \) and \( w \) have \( H^2(\Gamma(t)) \) regularity, and satisfy the estimate, for \( 0 \leq t \leq T \),

\[
\varepsilon \sup_{s \in (0, t)} \| u \|_{H^2(\Gamma(s))}^2 + \int_0^t \| w \|_{H^2(\Gamma(s))}^2 \, ds \leq C \| u_0 \|_{H^2(\Gamma_0)}.
\]

### 2.2 Abstract formulation

We will use the time-dependent bilinear forms, cf. [12,11], for any \( u, \varphi \in H^1(\Gamma(t)) \):

\[
m(t; u, \varphi) = \int_{\Gamma(t)} u \varphi, \quad a(t; u, \varphi) = \int_{\Gamma(t)} \nabla_{\Gamma(t)} u \cdot \nabla_{\Gamma(t)} \varphi, \quad g(t; v; u, \varphi) = \int_{\Gamma(t)} u \varphi (\nabla_{\Gamma(t)} \cdot v),
\]

We further define \( a^*(t; \cdot, \cdot) = a(t; \cdot, \cdot) + m(t; \cdot, \cdot) \). All bilinear forms are symmetric in \( u \) and \( \varphi \), \( m \) and \( a^* \) are positive definite, while \( a \) is positive semi-definite. Whenever it is possible, without confusion, we will omit the omnipresent time-dependence of the bilinear forms and write \( m(\cdot, \cdot) \) instead of \( m(t; \cdot, \cdot) \).

We note here that the bilinear forms directly generate the (semi-)norms:

\[
\| u \|_{L^2(\Gamma(t))}^2 = m(u, u), \quad \| \nabla_{\Gamma(t)} u \|_{L^2(\Gamma(t))}^2 = a(u, u), \quad \| u \|_{H^1(\Gamma(t))}^2 = a^*(u, u).
\]

The weak formulation (2.5) can then be rewritten, using the bilinear forms from above, as

\[
m(\partial^* u, \varphi) + a(w, \varphi) = -g(v; u, \varphi), \quad m(w, \varphi) - \varepsilon a(u, \varphi) = \varepsilon^{-1} m(W'(u), \varphi).
\]
and (2.6) is rewritten as
\[
\frac{d}{dt}m(u, \varphi) + a(w, \varphi) = 0, \\
m(w, \varphi) - \varepsilon a(u, \varphi) = \varepsilon^{-1} m(W'(u), \varphi).
\]

The transport formula for the above bilinear forms, [12, Remark 3.3], will be used later on, and reads
\[
\frac{d}{dt} m(u, \varphi) = m(\partial^\bullet u, \varphi) + m(u, \partial^\bullet \varphi) + g(v; u, \varphi). \tag{2.10}
\]

3 Semi-discretisation of Cahn–Hilliard equation on evolving surfaces

For the numerical solution of the above examples we consider a high-order evolving surface finite element method. In the following, from [10,11,7,18], we will briefly recall the construction of the discrete evolving surface, the high-order evolving surface finite element space, the lift operation, and the discrete bilinear forms, etc., which will be used to discretise the Cahn–Hilliard equation of Section 2.

3.1 Evolving surface finite elements

The smooth initial surface $\Gamma(0)$ is approximated by a k-order interpolating discrete surface denoted by $\Gamma_h(0) := \Gamma_h^k(0)$, with vertices $p_j$, $j = 1, \ldots, N$, and is given by the (high-order) triangulation, with maximal mesh width $h$. More details and the properties of such a discrete initial surface can be found in Section 2 of [7].

The triangulation of the surface $\Gamma(t)$, denoted by $\Gamma_h(t) := \Gamma_h^k(t)$, is obtained by integrating the ODE (2.1) (with the known velocity $v$) from time 0 to $t$ for all the nodes $p_j$ of the initial triangulation. The nodes $x_j(t)$ are on the exact surface $\Gamma(t)$ for all times. The discrete surface $\Gamma_h(t)$ remains to be an interpolation of $\Gamma(t)$ for all times. We always assume that the evolving (high-order) triangles are forming an admissible triangulation of the surface $\Gamma(t)$, which includes quasi-uniformity, and that the discrete surface is not a global double covering, cf. Section 5.1 of [10]. The discrete tangential gradient on the discrete surface $\Gamma_h(t)$ is given by:
\[
\nabla_{\Gamma_h(t)} \varphi = \nabla \tilde{\varphi}_h - (\nabla \tilde{\varphi}_h \cdot \nu_h) \nu_h,
\]
understood in an elementwise sense, with $\nu_h$ denoting the normal to $\Gamma_h(t)$.

The high-order evolving surface finite element space $S_h(t) \not\subseteq H^1(\Gamma(t))$ on $\Gamma_h(t)$ is spanned by continuous, piecewise linear nodal basis functions on $\Gamma_h(t)$ satisfying for each node $(x_j)^N_{j=1}$
\[
\phi_k(x_j, t) = \delta_{kj}, \quad \text{for} \quad j, k = 1, \ldots, N.
\]
Then the finite element space is given as

\[ S_h(t) = \text{span}\{\phi_1, \ldots, \phi_N\}. \]

The discrete velocity \( V_h \) of the surface \( \Gamma_h(t) \) is the evolving surface finite element interpolation of the surface velocity \( v \) of \( \Gamma(t) \), i.e.

\[ V_h(\cdot, t) = \sum_{j=1}^{N} v(x_j(t), t) \phi_j(\cdot, t). \quad (3.1) \]

The discrete material derivative is then given by

\[ \partial_t^\bullet \varphi_h = \partial_t \varphi_h + V_h \cdot \nabla \varphi_h, \quad \text{for all} \quad \varphi_h \in S_h(t). \quad (3.2) \]

The key transport property of basis functions derived in Proposition 5.4 in [10], is

\[ \partial_t^\bullet \phi_j = 0, \quad \text{for} \quad j = 1, 2, \ldots, N. \quad (3.3) \]

3.2 Lift

Following [10, 7], we define the lift operator \( \cdot^\ell \) to compare functions on \( \Gamma_h(t) \) with functions on \( \Gamma(t) \). For functions \( \varphi_h : \Gamma_h(t) \to \mathbb{R} \), we define the lift as

\[ \varphi_h^\ell : \Gamma(t) \to \mathbb{R} \quad \text{with} \quad \varphi_h^\ell(y) = \varphi_h(x), \quad \forall x \in \Gamma_h(t), \quad (3.4) \]

where \( y = y(x, t) \in \Gamma(t) \) is the unique point on \( \Gamma(t) \) with \( x - y \) orthogonal to the tangent space \( T_y \Gamma(t) \). The inverse lift \( \varphi^{-\ell} : \Gamma_h(t) \to \mathbb{R} \) denotes a function whose lift is \( \varphi : \Gamma(t) \to \mathbb{R} \). Finally, the lifted finite element space is denoted by \( S_h^\ell(t) \), and is given as

\[ S_h^\ell(t) = \{ \varphi_h^\ell | \varphi_h \in S_h(t) \}. \]

3.3 Discrete bilinear forms

The time-dependent discrete bilinear forms on \( S_h(t) \), i.e. the discrete counterparts of \( m, a \) and \( g \), are given, for \( u_h, \varphi_h \in S_h(t) \), by

\[
\begin{align*}
m_h(t; u_h, \varphi_h) &= \int_{\Gamma_h(t)} u_h \varphi_h, \\
a_h(t; u_h, \varphi_h) &= \int_{\Gamma_h(t)} \nabla \Gamma_h(t) u_h \cdot \nabla \Gamma_h(t) \varphi_h \\
g_h(t; V_h; u_h, \varphi_h) &= \int_{\Gamma_h(t)} u_h \varphi_h (\nabla \Gamma_h(t) \cdot V_h),
\end{align*}
\]

(3.5)

As in the continuous case we let \( a_h^\star(t; \cdot, \cdot) = a_h(t; \cdot, \cdot) + m_h(t; \cdot, \cdot) \). The discrete bilinear forms, clearly inherit the properties of their continuous counterparts, such as the transport formula (2.10), see, e.g. [12, 18].
As in the continuous case, the discrete bilinear forms directly generate the discrete (semi-)norms, for \( u_h \in S_h(t) \),

\[
\|u_h\|_{L^2(\Gamma(t))}^2 = m_h(u_h, u_h), \\
\|\nabla R_h(t)u_h\|_{L^2(\Gamma(t))}^2 = a_h(u_h, u_h), \\
\|u_h\|_{H^1(\Gamma(t))}^2 = a_h(u_h, u_h).
\]

According to [10,7], the discrete norms and their continuous counterparts are \( h \)-uniformly equivalent, for \( \varphi_h \in S_h(t) \),

\[
c\|\varphi_h\|_{L^2(\Gamma(t))} \leq \|\varphi_h\|_{L^2(\Gamma(t))} \leq C\|\varphi_h\|_{L^2(\Gamma(t))}, \\
c\|\nabla R_h(t)\varphi_h\|_{L^2(\Gamma(t))} \leq \|\nabla R_h(t)\varphi_h\|_{L^2(\Gamma(t))} \leq C\|\nabla R_h(t)\varphi_h\|_{L^2(\Gamma(t))}.
\]

### 3.4 Semi-discrete problem

The semi-discrete problem corresponding to the Cahn–Hilliard equation (2.5) reads: Find a solution \( u_h(\cdot, t) \in S_h(t) \) with time-continuous discrete material derivative \( \partial_t^* u_h(\cdot, t) \in S_h(t) \) and \( w_h(\cdot, t) \in S_h(t) \) such that for all test functions \( \varphi_h \in S_h(t) \) and \( \varphi_h^w \in S_h(t) \)

\[
m_h(\partial_t^* u_h, \varphi_h) + a_h(w_h, \varphi_h^w) = -g_h(V_h; u_h, \varphi_h^w), \tag{3.7a}
\]

\[
m_h(w_h, \varphi_h^w) - \varepsilon a_h(u_h, \varphi_h^w) = \varepsilon^{-1}m_h(W'(u_h), \varphi_h^w), \tag{3.7b}
\]

with given initial data \( u_h(\cdot, 0) = u_h^0 \) on \( \Gamma_h^0 \), while \( w_h(\cdot, 0) = w_h^0 \) is obtained by solving the elliptic problem (3.7b) (with a known \( u_h^0 \)) once.

Equivalently, the semi-discrete problem corresponding to the weak form (2.6), using the discrete version of the transport formula (2.10) for (3.7a), reads: Find a solution \( u_h(\cdot, t) \in S_h(t) \) with time-continuous discrete material derivative \( \partial_t^* u_h(\cdot, t) \in S_h(t) \) and \( w_h(\cdot, t) \in S_h(t) \) such that for all test functions \( \varphi_h \in S_h(t) \) with \( \partial_t^* \varphi_h = 0 \) and \( \varphi_h^w \in S_h(t) \)

\[
\frac{d}{dt}m_h(u_h, \varphi_h^w) + a_h(w_h, \varphi_h^w) = 0, \tag{3.8a}
\]

\[
m_h(w_h, \varphi_h^w) - \varepsilon a_h(u_h, \varphi_h^w) = \varepsilon^{-1}m_h(W'(u_h), \varphi_h^w), \tag{3.8b}
\]

again, with given initial data \( u_h(\cdot, 0) = u_h^0 \) on \( \Gamma_h^0 \).

### 3.5 Matrix–vector formulation

We collect the nodal values of \( u_h(\cdot, t) = \sum_{j=1}^N u_j(t)\phi_j(\cdot, t) \in S_h(t) \) and \( w_h(\cdot, t) = \sum_{j=1}^N w_j(t)\phi_j(\cdot, t) \in S_h(t) \), the solution pair of the semi-discrete problem (3.7), into the vectors \( \mathbf{u}(t) = (u_1(t), \ldots, u_N(t)) \in \mathbb{R}^N \) and \( \mathbf{w}(t) = (w_1(t), \ldots, w_N(t)) \in \mathbb{R}^N \).
We define the time-dependent matrices, the mass and stiffness matrix, corresponding to the bilinear forms $m_h$ and $a_h$, respectively, and the semi-linear term involving $W'$:

\[
\begin{align*}
M(t)_{kj} &= m_h(\phi_j(\cdot, t), \phi_k(\cdot, t)), \\
A(t)_{kj} &= a_h(\phi_j(\cdot, t), \phi_k(\cdot, t)), \\
W(u(t))_k &= m_h(W'(u_h(\cdot, t)), \phi_k(\cdot, t)),
\end{align*}
\]

We further define the matrix corresponding to the bilinear form $a^*_h$:

\[K(t) = M(t) + A(t).\]

We also note that, via the transport property (3.3), the time derivative of the mass matrix is given by

\[
\dot{M}(t)_{kj} = g_h(V_h(\cdot, t); \phi_j(\cdot, t), \phi_k(\cdot, t)).
\]

The discrete material derivative of a surface finite element function $u_h(t) \in \mathcal{S}_h(t)$, with nodal values $u(t)$, again by using the transport property (3.3) of the basis functions and the product rule, is given by

\[
\partial^*_h u_h(\cdot, t) = \partial^*_h \left( \sum_{j=1}^N u_j(t) \phi_j(\cdot, t) \right) = \sum_{j=1}^N \dot{u}_j(t) \phi_j(\cdot, t).
\]

Thus, the nodal values of $\partial^*_h u_h$ are given by the vector $\dot{u}(t)$.

The finite element semi-discretisation of the Cahn–Hilliard equation (3.7) then reads:

\[
\begin{align*}
M(t)\dot{u}(t) + A(t)w(t) &= -\dot{M}(t)u(t), \\
M(t)w(t) - \varepsilon A(t)u(t) &= \varepsilon^{-1}W(u(t)).
\end{align*}
\]

The anti-symmetric structure of (3.10), which is shared with (2.4) and (3.7), can be best seen in the rewritten form:

\[
\begin{bmatrix}
M(t) \frac{d}{dt} A(t) \\
-\varepsilon A(t) M(t)
\end{bmatrix}
\begin{bmatrix}
u(t) \\
w(t)
\end{bmatrix}
= \begin{bmatrix}
-\dot{M}(t)u(t) \\
\varepsilon^{-1}W(u(t))
\end{bmatrix}
\]

In order to exploit this favourable structure, the stability analysis will use the matrix–vector system (3.10).

For computations, it is however more advantageous to use the equivalent matrix–vector formulation

\[
\begin{align*}
\frac{d}{dt} \left( M(t)u(t) \right) + A(t)w(t) &= 0, \\
M(t)w(t) - \varepsilon A(t)u(t) &= \varepsilon^{-1}W(u(t)),
\end{align*}
\]

where the surface velocity $V_h$ does not appear directly, as compared to the term with $M(t)$ in (3.10).
3.6 A modified problem

The initial value $u(0)$ can be chosen suitably, on the other hand the initial value $w(0)$ is obtained from the second equation of the system (3.10) (or equivalently (3.11)). Our error analysis will require that the errors in both initial values to be $O(h^{k+1})$ in the $H^1(Γ_h)$ norm. For $u$ this can be achieved using the Ritz map (in which case the initial error in $u$ vanishes), however, such an error estimate is still not feasible for $w$.

To obtain optimal-order error estimates we modify the equation (3.10b) (and equivalently (3.11b) as well) using a time-independent correction term. Let $\bar{w}(0) ∈ \mathbb{R}^N$ denote the solution obtained from (3.10b) at time $t = 0$, and let $w^*(0) ∈ \mathbb{R}^N$ contain the nodal value of the Ritz map of $w(0)$, and set

$$\vartheta = M(0)(w^*(0) - \bar{w}(0)) ∈ \mathbb{R}^N.$$ (3.12)

The second equation is then modified, such that the system (3.10) reads:

$$M(t)\dot{u}(t) + A(t)w(t) = -\dot{M}(t)u(t),$$ (3.13a)

$$M(t)w(t) - εA(t)u(t) = ε^{-1}W(u(t)) + \vartheta.$$ (3.13b)

Similarly, the equivalent system (3.11) is modified to:

$$\frac{d}{dt} \left( M(t)u(t) \right) + A(t)w(t) = 0,$$ (3.14a)

$$M(t)w(t) - εA(t)u(t) = ε^{-1}W(u(t)) + \vartheta.$$ (3.14b)

The semi-discrete problems (3.7) and (3.8) are modified accordingly:

The initial value $w(0)$ is obtained by solving the elliptic problem (3.13b) at $t = 0$, which, via (3.12) and (3.10b), yields

$$M(0)w(0) = εA(0)u(0) + ε^{-1}W(u(0)) + \vartheta$$
$$= M(0)w(0) + \vartheta$$
$$= M(0)w^*(0).$$ (3.15)

4 Error estimates

4.1 Error estimates by Elliott and Ranner [15]

In [15] Elliott and Ranner proved the first error estimates for the evolving surface finite element discretisation of the Cahn–Hilliard equation on evolving surfaces. In particular, they proved optimal-order error estimates in the $L^\infty(L^2)$ and $L^\infty(H^1)$ norms for $u$, and in the $L^2(L^2)$ and $L^2(H^1)$ norms for $w$, using linear surface finite elements.
Theorem 4.1 (Elliott and Ranner [15, Theorem 5.1]) Let $u$ and $w$ be the weak solutions of the Cahn–Hilliard equation on an evolving surface (2.4), and assume that they are sufficiently regular, e.g. satisfy (4.1).

Then there exists an $h_0 > 0$ such that for all $h \leq h_0$ the errors between the solutions $u$ and $w$ and the evolving surface finite element solutions $u_h$ and $w_h$ of degree 1 of (3.7), satisfy the optimal-order error estimates in both variables:

$$
\left( \varepsilon \sup_{t \in (0,T)} \| u(\cdot, t) - u_h^t(\cdot, t) \|^2_{L^2(\Gamma(t))} + \int_0^T \| w(\cdot, t) - w_h^t(\cdot, t) \|^2_{L^2(\Gamma(t))} dt \right)^{1/2} \leq C h^2,
$$

and

$$
\left( \varepsilon \sup_{t \in (0,T)} \| \nabla \Gamma(t)(w(\cdot, t) - w_h^t(\cdot, t)) \|^2_{L^2(\Gamma(t))} \right)^{1/2} \leq C h,
$$

where the constant $C > 0$ is independent of $h$, but depends on $\varepsilon$, on the bounds of the Sobolev norms of the solution $u$, $w$, and $\partial^* u$, on the surface evolution, and on the length $T$ of the time interval.

Sufficient regularity conditions on $u$ and $w$ required by Theorem 4.1 ([15, Theorem 5.1]) are:

- $u \in L^\infty(0,T; H^2(\Gamma(t)))$, with $\partial^* u \in L^2(0,T; H^2(\Gamma(t)))$,
- $w \in L^2(0,T; H^2(\Gamma(t)))$,
- and for the surface velocity:
  - $v \in L^\infty(0,T; W^{2,\infty}(\Gamma(t)))$.

The constant $C$ is explicitly given in the cited theorem.

The analysis presented in [15] is not restricted to linear evolving surface finite elements. By a straightforward combination of the proofs used for [15, Theorem 5.1] with the relevant high-order approximation results shown in [18], the above theorem stays valid for degree $k$ evolving surface finite elements with $h^{k+1}$ and $h^k$ in place of $h^2$ and $h$. For a detailed proof we refer to [3, Section 4.1].

4.2 Uniform-in-time error estimates

We next state a new convergence result for the evolving surface finite element semi-discretisation of order $k \geq 1$.

Theorem 4.2 Let $u$ and $w$ be the weak solutions of the Cahn–Hilliard equation on an evolving surface (2.4), and assume that they satisfy the regularity conditions (4.1).
Then there exists an \( h_0 > 0 \) such that for all \( h \leq h_0 \) the errors between the solutions \( u \) and \( w \) and the evolving surface finite element solutions \( u_h \) and \( w_h \) of degree \( k \), with nodal vectors solving (3.14), satisfy the optimal-order uniform-in-time error estimates in both variables, for \( 0 \leq t \leq T \),
\[
\begin{align*}
\|u_h^t(\cdot,t) - u(\cdot,t)\|_{L^2(\Gamma(t))} + h\|u_h^t(\cdot,t) - u(\cdot,t)\|_{H^1(\Gamma(t))} & \leq Ch^{k+1}, \\
\|w_h^t(\cdot,t) - w(\cdot,t)\|_{L^2(\Gamma(t))} + h\|w_h^t(\cdot,t) - w(\cdot,t)\|_{H^1(\Gamma(t))} & \leq Ch^{k+1},
\end{align*}
\]
whereas the time derivatives of the error in \( u \) satisfies, for \( 0 \leq t \leq T \),
\[
\left( \int_0^t \|\partial^* u_h^t(\cdot,s) - u(\cdot,s)\|^2_{L^2(\Gamma(s))} ds + h\|\partial^* u_h^t(\cdot,s) - u(\cdot,s)\|^2_{H^1(\Gamma(s))} ds \right)^{1/2} \leq Ch^{k+1},
\]
where the constant \( C > 0 \) is independent of \( h \) and \( t \), but depends on \( \varepsilon \), on the bounds of the Sobolev norms of the solution \( u \) and \( w \), on the surface evolution, and on the length of the time interval \( T \).

Sufficient regularity conditions on \( u \) and \( w \) required by Theorem 4.2 are:
\[
\begin{align*}
&u, \partial^* u, (\partial^*)^{(2)} u \in L^2(0, T; H^{k+1}(\Gamma(t))), \text{ with } u \in L^\infty(0, T; H^{k+1}(\Gamma(t))), \\
w, \partial^* w, w \in L^2(0, T; H^{k+1}(\Gamma(t))), \text{ with } w \in L^\infty(0, T; H^{k+1}(\Gamma(t))), \\
&u \in L^\infty(0, T; W^{2,\infty}(\Gamma(t))), \\
&v, \partial^* v \in L^\infty(0, T; W^{k+1,\infty}(\Gamma(t))).
\end{align*}
\]

(4.2)

The main differences between Theorem 4.1 by Elliott and Ranner [15] and Theorem 4.2 from above are the following: The former result provides a better scaling in \( \varepsilon \) between the \( L^2 \) norm and \( H^1 \) semi-norm than the latter, note, however, that in both theorems the constant \( C \) (unfavourably) depends on \( \varepsilon \). Our result proves uniform-in-time error estimates in the \( H^1 \) norm not only for the error in \( u \) but also for the error in \( w \), which is not available through the analysis of Elliott and Ranner [15]. Theorem 4.2 also provides error estimates in the errors of the material derivatives for \( u \). Naturally, the additional error estimates require more regularity assumptions on \( u \), \( w \) and \( v \).

Theorem 4.2 will be proven by studying the questions of stability and consistency. The consistency of the algorithm is shown by proving high-order estimates for the defects (the error obtained by inserting the Ritz map of the exact solutions into the method), which are obtained by using geometric and approximation error estimates for high-order evolving surface finite elements from [18], which combines techniques of [10,12] and [7].

The main issue in the proof is stability, i.e. a mesh independent, uniform-in-time bound of the errors in terms of the defects. The main idea of the stability proof was originally developed for Willmore flow [20], and it relies
on energy estimates that exploit the anti-symmetric structure of the Cahn–Hilliard equation, see (2.4), (3.7), and (3.13). The basic idea of the stability proof is concisely sketched in Figure 5.1. In order to estimate the non-linear term, a key issue in the stability proof is to ensure that the $L^\infty$ norm of the error in $u$ remains bounded. The uniform-in-time $H^1$ norm error bounds together with an inverse estimate provide a bound in the $L^\infty$ norm. Similarly, it is also possible to show such an $L^\infty$ norm bound for the error in $w$, provided by our uniform-in-time $H^1$ norm bounds in both $u$ and $w$. This enables to generalise our stability proof, and hence also the convergence proof, to quite general non-linear problems, which contain a non-linear expression in both equations in (2.4), possibly involving the gradient of $u$ as well. The generalisations only require very minor changes in the stability proof, as elaborated in Section 8. Such examples include Cahn–Hilliard equation with a proliferation term [27]. More details to these problems and to their stability analysis are given in Section 8, along with Theorem 8.1 which generalises Theorem 4.2 to these more general non-linear Cahn–Hilliard equations.

5 Stability

5.1 Preliminaries

This section is dedicated to the definition of a few concepts, such as the comparison of various quantities on different discrete surfaces and a generalised Ritz map, which will be all used throughout the stability analysis.

The finite element matrices $M(t)$, $A(t)$, and $K(t)$ induce (semi-)norms which correspond to discrete Sobolev (semi-)norms:

\[
\|w\|_M^2(t) = w^T M(t) w = \|w_h\|_{L^2(\Gamma_h(t))}^2,
\]

\[
\|w\|_A^2(t) = w^T A(t) w = \|\nabla\Gamma_h(t) w_h\|_{L^2(\Gamma_h(t))}^2,
\]

\[
\|w\|_K^2(t) = \|w\|^2_M(t) + \|w\|^2_A(t) = \|w_h\|_{H^1(\Gamma_h(t))}^2,
\]

(5.1)

for a vector $w \in \mathbb{R}^N$ corresponding to the finite element function $w_h \in S_h(t)$.

From [21, Lemma 4.6] we recall the following estimates for the time derivatives of the mass and stiffness matrix, and, additionally, we prove that they also hold for the second order time derivatives.

**Lemma 5.1** For all vectors $w, z \in \mathbb{R}^N$ we have

\[
w^T \dot{M}(t) z \leq c \|w\|_M(t) \|z\|_M(t),
\]

(5.2a)

\[
w^T \dot{A}(t) z \leq c \|w\|_A(t) \|z\|_A(t),
\]

(5.2b)

\[
w^T \ddot{M}(t) z \leq c \|w\|_M(t) \|z\|_M(t),
\]

(5.2c)

\[
w^T \ddot{A}(t) z \leq c \|w\|_A(t) \|z\|_A(t),
\]

(5.2d)

where the constant $c > 0$ is independent of $h$, but depends on the surface velocity $v$. 

Proof The first two estimates were shown in Lemma 4.6 of [21].

We prove the estimate (5.2c) for the second derivative of the mass matrix. For vectors \( \mathbf{w}, \mathbf{z} \in \mathbb{R}^N \) corresponding to discrete functions \( w_h, z_h \in \mathcal{S}_h(t) \) (for \( 0 \leq t \leq T \)), we have \( \partial^* \mathbf{w} = \partial^*_h \mathbf{z} = 0 \) by the transport property (3.3). Using twice the discrete version of the Leibniz formula [10, Lemma 2.2], we obtain

\[
\mathbf{w}^T \tilde{\mathbf{M}}(t) \mathbf{z} = \frac{d^2}{dt^2} \int_{\Gamma_h(t)} w_h z_h = \frac{d}{dt} \int_{\Gamma_h(t)} w_h z_h (\nabla \mathbf{v}_h : V_h)
\]

\[
\int_{\Gamma_h(t)} w_h z_h \partial^*_h (\nabla \mathbf{v}_h : V_h) + \int_{\Gamma_h(t)} w_h z_h (\nabla \mathbf{v}_h : V_h)^2.
\]

To estimate the first integral we recall how to interchange surface differential operators with the material derivative [13, Lemma 2.6], for \( u : \Gamma(t) \rightarrow \mathbb{R} \) and \( u : \Gamma(t) \rightarrow \mathbb{R}^3 \), respectively,

\[
\partial^* (\nabla \mathbf{v}) = \nabla (\partial^* \mathbf{v}) - \mathbf{v}^T \nabla \mathbf{v} \cdot \nabla \mathbf{v},
\]

\[
\partial^* (\nabla \mathbf{v} : \mathbf{w}) = \nabla (\partial^* \mathbf{v} : \mathbf{w}) - \mathbf{v}^T \nabla \mathbf{v} \cdot \nabla \mathbf{w},
\]

and analogously for discrete differential operators. Then the second formula from above is used to estimate the integral, together with the bounds on the discrete velocity \( \mathbf{v}_h \). The latter is implied by the sufficient regularity of the velocity \( \mathbf{v} \), and recalling that \( \mathbf{v}_h \) is the interpolation of \( \mathbf{v} \); cf. (3.1), see Lemma 6.2 or [3, Lemma 3.1.6]. We altogether obtain

\[
\int_{\Gamma_h(t)} w_h z_h \partial^*_h (\nabla \mathbf{v}_h : V_h) \leq \|w_h\|_{L^2(\Gamma_h(t))} \|z_h\|_{L^2(\Gamma_h(t))} \|\partial^*_h (\nabla \mathbf{v}_h : V_h)\|_{L^\infty(\Gamma_h(t))} \leq c \|\mathbf{w}\|_{\mathbf{M}(t)} \|\mathbf{z}\|_{\mathbf{M}(t)}.
\]

The second integral is directly bounded by

\[
\int_{\Gamma_h(t)} w_h z_h (\nabla \mathbf{v}_h : V_h)^2 \leq \|w_h\|_{L^2(\Gamma_h(t))} \|z_h\|_{L^2(\Gamma_h(t))} \|\nabla \mathbf{v}_h : V_h\|_{L^\infty(\Gamma_h(t))}^2 \leq c \|\mathbf{w}\|_{\mathbf{M}(t)} \|\mathbf{z}\|_{\mathbf{M}(t)}.
\]

The estimate for the stiffness matrix is shown by the exact same arguments, cf. [3]. \( \square \)

5.2 Error equations and defects

Before turning to the stability analysis, let us define a Ritz map of the exact solution onto the evolving surface finite element space, from [25,18] we recall the definition of a time-dependent Ritz map on evolving surfaces: \( \tilde{R}_h : H^1(\Gamma(t)) \rightarrow S_h^e(t) \), (here we do not include the velocity term of [25]).

Let \( u(\cdot, t) \in H^1(\Gamma(t)) \) for \( 0 \leq t \leq T \). Then the Ritz map is defined through \( \tilde{R}_h(t)u \in \mathcal{S}_h(t) \) which satisfies, for all \( \varphi_h \in \mathcal{S}_h(t) \),

\[
a_h^*(\tilde{R}_h(t)u, \varphi_h) = a^*(u, \varphi_h).
\]
The Ritz map is then defined as the lift of \( \tilde{R}_h(t) \), i.e. \( R_h(t)u = (\tilde{R}_h(t)u)^t \in S_h(t) \). We will often suppress the omnipresent time-dependency of the Ritz map. In [25] it was shown that the above Ritz map is well-defined, optimal-order error estimates for high-order evolving surface FEM were shown in [18], and are recalled in Lemma 6.4.

Let us consider now the (unlifted) Ritz map of the exact solutions \( u \) and \( w \) of (2.4), which are denoted by
\[
\begin{align*}
    u_h^*(\cdot, t) &= \tilde{R}_h(t)u(\cdot, t) \in S_h(t) \quad \text{and} \quad w_h^*(\cdot, t) = \tilde{R}_h(t)w(\cdot, t) \in S_h(t),
\end{align*}
\]
whose nodal values are collected into the vectors
\[
    \mathbf{u}^*(t) \in \mathbb{R}^N \quad \text{and} \quad \mathbf{w}^*(t) \in \mathbb{R}^N.
\]
The nodal vectors of the Ritz maps of the exact solutions satisfy the system (3.10) only up to some defects \( d_u(t) \) and \( d_w(t) \) in \( \mathbb{R}^N \), corresponding to the finite element functions \( d_h^u(\cdot, t) \) and \( d_h^w(\cdot, t) \) in \( S_h(t) \):
\[
\begin{align*}
    M(t)\dot{\mathbf{u}}^*(t) + A(t)\mathbf{w}^*(t) &= -\dot{M}(t)\mathbf{u}^*(t) + M(t)d_u(t), \quad (5.5a) \\
    M(t)\mathbf{w}^*(t) - \varepsilon A(t)\mathbf{u}^*(t) &= \varepsilon^{-1}W(\mathbf{u}^*(t)) + M(t)d_w(t). \quad (5.5b)
\end{align*}
\]
The errors between the nodal values of the semi-discrete solutions and of the Ritz maps of the exact solutions are denoted by \( e_u(t) = u(t) - u^*(t) \) and \( e_w(t) = w(t) - w^*(t) \) in \( \mathbb{R}^N \). By subtracting (5.5) from (3.13) we obtain that the errors \( e_u \) and \( e_w \) (corresponding to the functions \( e_{u_h} \) and \( e_{w_h} \) in \( S_h(t) \)) satisfy the following error equations:
\[
\begin{align*}
    M(t)e_u(t) + A(t)e_w(t) &= -\dot{M}(t)e_u(t) - M(t)d_u(t), \quad (5.6a) \\
    M(t)e_w(t) - \varepsilon A(t)e_u(t) &= \varepsilon^{-1}\left(W(u(t)) - W(u^*(t))\right) \\
    &\quad - M(t)d_w(t) + \vartheta, \quad (5.6b)
\end{align*}
\]
with zero initial values \( e_u(0) = 0 \) and \( e_w(0) = 0 \), by construction. Since the initial values satisfy (5.6b) at \( t = 0 \), we obtain the useful expression
\[
    \vartheta = M(0)d_w(0). \quad (5.7)
\]

5.3 Stability bounds

**Proposition 5.1** Suppose there exists a constant \( c > 0 \) independent of \( h \) and \( t \) such that the defects are bounded for \( \kappa \geq 2 \) by
\[
\begin{align*}
    \|d_u(t)\|_{M(t)} &\leq c h^\kappa, \quad \|\dot{d}_u(t)\|_{M(t)} \leq c h^\kappa, \\
    \|d_w(t)\|_{M(t)} &\leq c h^\kappa, \quad \|\dot{d}_w(t)\|_{M(t)} \leq c h^\kappa, \quad t \in [0, T]. \quad (5.8)
\end{align*}
\]
Further suppose that, for all \( 0 \leq t \leq T \), the Ritz map \( u_h^* = \tilde{R}_h u \) satisfies the bound \( \|u_h^*(\cdot, t)\|_{L^\infty(S_h(t))} \leq M \).
Then there exists $h_0 > 0$ such that the following error bound holds for $h \leq h_0$ and $0 \leq t \leq T$:

$$
\| e_u(t) \|_{K(t)}^2 + \| e_w(t) \|_{K(t)}^2 + \int_0^t \| \dot{e}_u(s) \|_{K(s)}^2 \, ds
\leq C \int_0^t \| d_u(s) \|_{M(s)}^2 + \| \dot{d}_u(s) \|_{M(s)}^2 + \| d_w(s) \|_{M(s)}^2 \, ds + C \| d_u(t) \|_{M(t)}^2 + C t \| d_w(0) \|_{M(0)}^2.
$$

(5.9)

The constant $C > 0$ is independent of $t$ and $h$, but depends on $0 < \epsilon < 1$, and exponentially on the final time $T$.

**Proof** The proof is based on energy estimates, and its basic idea is very similar to that of [20]. Proving uniform-in-time $H^1$ norm error estimates is essential for handling the non-linear term, which is done by deriving an $L^\infty$ norm bound for the errors using an inverse estimate.

In order to achieve a uniform-in-time stability bound, two sets of energy estimates are needed. These energy estimates strongly exploit the anti-symmetric structure of (2.4). (i) In the first, an energy estimate is proved for $e_u$, but comes with a critical term involving $\dot{e}_u$. (ii) The second estimate uses the time derivative of (5.6b), leads to a bound of this critical term and also to a uniform-in-time bound for $e_w$. The combination of these two energy estimates will give the above stability bound. The structure and basic idea of the proof is sketched in Figure 5.1. In order to handle the semi-linear term we first prove the stability bound on a time interval where the $L^\infty$ norm of $e_u$ is small enough, and then show that this time interval can be enlarged up to $T$.

In the following $c$ and $C$ are generic constants that take different values on different occurrences. Whenever it is possible, without confusion, we omit the
argument \( t \) of time-dependent vectors but not of time-dependent matrices. By \( \varepsilon > 0 \) we will denote a small number, used in Young’s inequality, and hence we will often incorporate \( h \) independent multiplicative constants into those, yet unchosen, factors.

We start by stating that there exists a maximal time \( 0 < t^* \leq T \) such that, for all \( t \leq t^* \),

\[
\|e_{uh}(\cdot, t)\|_{L^\infty(I_h(t))} \leq h^{\frac{\varepsilon^2}{2}} \quad \text{for all } \quad 0 \leq t \leq t^*. \tag{5.10}
\]

Since \( e_{uh}(\cdot, 0) = 0 \) and since \( u_h \) and \( u_h^* \) are continuous in time, we directly infer that \( t^* > 0 \).

Thus, by the assumption that the Ritz map of the exact solution satisfies \( \|u_h^*(t)\|_{L^\infty(I_h(t))} \leq M \), we obtain the following bound for the numerical solution:

\[
\|u_h(\cdot, t)\|_{L^\infty(I_h(t))} = \|u_h^*(\cdot, t) - e_{uh}(\cdot, t)\|_{L^\infty(I_h(t))} \leq \|u_h^*(\cdot, t)\|_{L^\infty(I_h(t))} + \|e_{uh}(\cdot, t)\|_{L^\infty(I_h(t))} \leq 2M,
\]

for all \( 0 \leq t \leq t^* \) and for \( h \leq h_0 \) sufficiently small. Thus, for \( f \in C(\mathbb{R}) \)

\[
\|f(u_h(\cdot, t))\|_{L^\infty(I_h(t))} \leq C,
\]

for all \( 0 \leq t \leq t^* \) and \( h \leq h_0 \) sufficiently small.

We first prove the stated stability bound for \( 0 \leq t \leq t^* \), and then show that indeed \( t^* \) coincides with \( T \).

**Energy estimate (i):** We take the first error equation (5.6a) and test it with \( e_u \), while the second one (5.6b) is tested by \( \varepsilon^{-1}e_w \), to obtain

\[
e_{uh}^TM(t)e_u + e_{uh}^TA(t)e_w = -e_{uh}^TM(t)e_u - e_{uh}^TM(t)d_u,
\]

\[
\varepsilon^{-1}e_{uw}^TM(t)e_u - e_{uw}^MA(t)e_u = \varepsilon^{-2}e_{uw}^T(W(u(t)) - W(u^*(t)))
\]

\[
- \varepsilon^{-1}e_{uw}^T\dot{M}(t)d_w + \varepsilon^{-1}e_{uw}^T\dot{\theta}.
\]

By adding the two equations, and by the symmetry of \( A \), we eliminate the mixed term \( e_{uw}^TA(t)e_u \), and obtain

\[
e_{uh}^TM(t)e_u + \varepsilon^{-1}e_{uw}^TM(t)e_u = -e_{uh}^TM(t)e_u
\]

\[
+ \varepsilon^{-2}e_{uw}^T(W(u(t)) - W(u^*(t)))
\]

\[
- e_{uh}^TM(t)d_u - \varepsilon^{-1}e_{uw}^M(t)d_w + \varepsilon^{-1}e_{uw}^T\dot{\theta}.
\]

Using the product rule and symmetry of \( M \) we rewrite the first term as

\[
e_{uh}^TM(t)e_u = \frac{1}{2} \frac{d}{dt} (e_{uh}^TM(t)e_u) - \frac{1}{2} e_{uh}^T\dot{M}(t)e_u,
\]

which altogether yields

\[
1 \frac{d}{dt} \|e_u\|_{M(t)} + \varepsilon^{-1}\|e_w\|^2_{M(t)} = - \frac{1}{2} e_{uh}^T\dot{M}(t)e_u
\]

\[
+ \varepsilon^{-2}e_{uw}^T(W(u(t)) - W(u^*(t)))
\]

\[
- e_{uh}^TM(t)d_u - \varepsilon^{-1}e_{uw}^M(t)d_w + \varepsilon^{-1}e_{uw}^T\dot{\theta}.
\]
Similarly, we test (5.6a) by $e_w$ and (5.6b) by $\hat{e}_u$, now a subtraction leads to cancelling the mixed term $e_w^T M(t) \hat{e}_u$, and again by the product rule and the symmetry of $A$, we obtain

$$
\varepsilon \frac{1}{2} \frac{d}{dt} \|e_u\|^2_{A(t)} + \|e_w\|^2_{A(t)} = -e_w^T \dot{M}(t) e_u + \varepsilon \frac{1}{2} \dot{e}_u^T \dot{A}(t) e_u \\
- \varepsilon^{-1} e_u^T (W(u(t)) - W(u^*(t))) \\
- e_w^T M(t) d_u + e_u^T M(t) d_w - \dot{e}_u^T \vartheta.
$$

Dividing the second equality by $\varepsilon$ and taking their linear combination yields

$$
\frac{1}{2} \frac{d}{dt} \|e_u\|^2_{K(t)} + \varepsilon^{-1} \|e_w\|^2_{K(t)} = -\varepsilon^{-1} e_w^T \dot{M}(t) e_u - \frac{1}{2} e_u^T \dot{M}(t) e_u + \frac{1}{2} e_u^T \dot{A}(t) e_u \\
+ \varepsilon^{-2} e_w^T (W(u(t)) - W(u^*(t))) \\
- \varepsilon^{-2} e_u^T (W(u(t)) - W(u^*(t))) \\
- e_u^T M(t) d_u - \varepsilon^{-1} e_w^T M(t) d_w \\
- \varepsilon^{-1} e_u^T M(t) d_u + \varepsilon^{-1} e_w^T M(t) d_w \\
+ \varepsilon^{-1} \dot{e}_u^T \vartheta - \varepsilon^{-1} \dot{e}_u^T \vartheta.
$$

(5.13)

The terms on the right-hand side are now estimated separately.

The terms involving time derivatives of matrices are estimated using Lemma 5.1, by

$$
egarepsilon^{-1} e_w^T \dot{M}(t) e_u - \frac{1}{2} e_u^T \dot{M}(t) e_u + \frac{1}{2} e_u^T \dot{A}(t) e_u \leq \varepsilon^{-1} \|e_w\|_{M(t)} \|e_u\|_{M(t)} + c \|e_u\|^2_{K(t)}.
$$

(5.14)

For the non-linear terms, using (5.11) and the local-Lipschitz property of $W'$, we obtain

$$
\varepsilon^{-2} e_w^T (W(u(t)) - W(u^*(t))) = \varepsilon^{-2} \int_{\Gamma_h(t)} c_{wh}(\cdot, t) \left( W'(u_h(\cdot, t)) - W'(u^*_h(\cdot, t)) \right) \\
\leq L \varepsilon^{-2} \|c_{wh}(\cdot, t)\|_{L^2(\Gamma_h(t))} \|u_h(\cdot, t) - u^*_h(\cdot, t)\|_{L^2(\bar{\Gamma}_h(t))} \\
= c \varepsilon^{-2} \|e_w\|_{M(t)} \|e_u\|_{M(t)}.
$$

(5.15)

and similarly

$$
\varepsilon^{-2} \dot{e}_u^T (W(u(t)) - W(u^*(t))) \leq c \varepsilon^{-2} \|e_u\|_{M(t)} \|e_u\|_{M(t)}.
$$

(5.16)

The defect terms are estimated by Cauchy–Schwarz inequality, as

$$
- e_u^T M(t) d_u - \varepsilon^{-1} e_w^T M(t) d_w - \varepsilon^{-1} \dot{e}_u^T M(t) d_u + \varepsilon^{-1} \dot{e}_u^T M(t) d_w \\
\leq \|e_u\|_{M(t)} \|d_u\|_{M(t)} + \varepsilon^{-1} \|e_w\|_{M(t)} \|d_w\|_{M(t)} \\
+ \varepsilon^{-1} \|e_w\|_{M(t)} \|d_u\|_{M(t)} + \varepsilon^{-1} \|\dot{e}_u\|_{M(t)} \|d_w\|_{M(t)}.
$$

(5.17)
The terms involving the correction term are bounded similarly as the defect terms. Using equality (5.7) and the norm equivalence in time [14, Lemma 4.1] (to change the time from 0 to $t$), we obtain

$$
\varepsilon^{-1} \hat{e}_w^2 \vartheta - \varepsilon^{-1} \hat{e}_w^2 \vartheta \\
\leq \varepsilon^{-1} \| \hat{e}_w \|_{M(0)} \| d_{w(0)} \|_{M(0)} + \varepsilon^{-1} \| \hat{e}_u \|_{M(0)} \| d_{w(0)} \|_{M(0)} \\
\leq c \varepsilon^{-1} \| \hat{e}_w \|_{M(t)} \| d_{w(0)} \|_{M(t)} + c \varepsilon^{-1} \| \hat{e}_u \|_{M(t)} \| d_{w(0)} \|_{M(t)}.
$$

(5.18)

Altogether, by the combination of the estimates (5.14)–(5.17) with (5.13), by multiple Young’s inequalities (with $\varrho > 0$ chosen later on) and by absorptions to the left-hand side, we obtain

$$
\frac{1}{2} \frac{d}{dt} \| e_u \|_{K(t)}^2 + \varepsilon^{-1} \| e_w \|_{K(t)}^2 \\
\leq \frac{1}{2} \| \hat{e}_u \|_{K(t)}^2 + c \| e_u \|_{K(t)}^2 \\
+ c \| d_u \|_{M(t)}^2 + c \| d_w \|_{M(t)}^2 + c \| d_w(0) \|_{M(0)}^2.
$$

(5.19)

Integrating from 0 to $t \in (0, t^*)$, and using that $e_u(0) = 0$, we obtain the first energy estimate:

$$
\| e_u(t) \|_{K(t)}^2 + \varepsilon^{-1} \int_0^t \| e_w(s) \|_{K(s)}^2 ds \\
\leq \varrho \int_0^t \| \hat{e}_u(s) \|_{K(s)}^2 ds + c \int_0^t \| e_u(s) \|_{K(s)}^2 ds \\
+ c \int_0^t (\| d_u(s) \|_{M(s)}^2 + \| d_w(s) \|_{M(s)}^2) ds \\
+ c t \| d_w(0) \|_{M(0)}^2.
$$

(5.20)

Note the critical term, with $\| \hat{e}_u(s) \|_{K(s)}$, on the right-hand side, which cannot be bounded or absorbed in any direct way.

Energy estimates (ii) To control the critical term on the right-hand side of (5.20) we will now derive an energy estimate, which includes this term on the left-hand side. To this end, we first differentiate the second equation of (5.6) with respect to time (note that the time-independent $\vartheta$ vanishes), and, after rearranging the terms, we obtain the following system:

$$
M(t) \hat{e}_u + A(t) e_w = - \dot{M}(t) e_u - M(t) d_u, \\
M(t) \hat{e}_w - \varepsilon A(t) \hat{e}_u = - \dot{M}(t) e_w + \varepsilon \dot{A}(t) e_u \\
+ \varepsilon^{-1} \frac{d}{dt} (W(u(t)) - W(u^*(t))) \\
- \dot{M}(t) d_w - M(t) \dot{d}_w.
$$

(5.21)

Testing the error equation system (5.21) twice, similarly as before in Part (i), would not lead to a feasible energy estimate, but to a bound which include a new critical term $\hat{e}_u$. The issue is avoided by separating the two estimates for the error equations, (ii.a) and (ii.b), and then taking their weighted combination in (ii.c).
(ii.a) We test (5.21a) by $\dot{\mathbf{e}}_u$ and (5.21b) by $\varepsilon^{-1} \mathbf{e}_w$, adding the two equations together to cancel the mixed term $\dot{\mathbf{e}}_u^T \mathbf{A}(t) \mathbf{e}_w$, and using the product rule as before, we obtain

\[
\|\dot{\mathbf{e}}_u\|_M(t) + \varepsilon^{-1} \frac{1}{2} \left( \frac{d}{dt} \|\mathbf{e}_w\|_M(t) \right) = -\dot{\mathbf{e}}_u^T \dot{\mathbf{M}}(t) \mathbf{e}_u - \varepsilon^{-1} \frac{1}{2} \mathbf{e}_w^T \dot{\mathbf{M}}(t) \mathbf{e}_w + \mathbf{e}_w^T \dot{\mathbf{A}}(t) \mathbf{e}_u
\]

\[
+ \varepsilon^{-2} \mathbf{e}_w^T \left( \mathbf{W}(\mathbf{u}(t)) - \mathbf{W}(\mathbf{u}^*(t)) \right) \mathbf{e}_u
\]

\[
- \dot{\mathbf{e}}_w^T \dot{\mathbf{M}}(t) \mathbf{d}_u - \varepsilon^{-1} \mathbf{e}_w^T \dot{\mathbf{M}}(t) \mathbf{d}_w - \varepsilon^{-1} \mathbf{e}_w^T \mathbf{M}(t) \dot{\mathbf{d}}_w(t).
\]

The right-hand side terms are again estimated separately. The ones in the first line are bounded, using Lemma 5.1, by

\[
\|\dot{\mathbf{e}}_u\|_M(t) \leq c \|\dot{\mathbf{e}}_u\|_M(t) \|\mathbf{e}_u\|_M(t) + c \varepsilon^{-1} \|\mathbf{e}_w\|_M(t) + c \|\mathbf{e}_w\|_L(t) \|\mathbf{e}_u\|_A(t).
\]

The non-linear term occurs differentiated with respect to time. Therefore, with the help of the transport formula (2.10) and inserting $\mp \int_{\Gamma_h(t)} W''(u^*_h(\cdot, t)) \partial_h^* u_h^*(\cdot, t) \phi_j$ after the second equality, we obtain

\[
\varepsilon^{-2} \mathbf{e}_w^T \frac{d}{dt} \left( \mathbf{W}(\mathbf{u}) - \mathbf{W}(\mathbf{u}^*) \right)
\]

\[
= \varepsilon^{-2} \left( \int_{\Gamma_h(t)} W''(u_h(\cdot, t)) \partial_h^* u_h(\cdot, t) e_{w_h}(\cdot, t) - \int_{\Gamma_h(t)} W''(u_h^*(\cdot, t)) \partial_h^* u_h^*(\cdot, t) e_{w_h}(\cdot, t) \right)
\]

\[
+ \int_{\Gamma_h(t)} (\nabla_{\Gamma_h} \cdot \mathbf{V}_h) W'(u_h(\cdot, t)) e_{w_h}(\cdot, t) - \int_{\Gamma_h(t)} (\nabla_{\Gamma_h} \cdot \mathbf{V}_h) W'(u_h^*(\cdot, t)) e_{w_h}(\cdot, t)
\]

\[
= \varepsilon^{-2} \left( \int_{\Gamma_h(t)} W''(u_h(\cdot, t)) \left( \partial_h^* u_h(\cdot, t) - \partial_h^* u_h^*(\cdot, t) \right) e_{w_h}(\cdot, t) \right)
\]

\[
+ \int_{\Gamma_h(t)} \left( W''(u_h(\cdot, t)) - W''(u_h^*(\cdot, t)) \right) \partial_h^* u_h^*(\cdot, t) e_{w_h}(\cdot, t)
\]

\[
+ \int_{\Gamma_h(t)} (\nabla_{\Gamma_h} \cdot \mathbf{V}_h) \left( W'(u_h(\cdot, t)) - W'(u_h^*(\cdot, t)) \right) e_{w_h}(\cdot, t)
\]

\[
\leq c \varepsilon^{-2} \|\mathbf{e}_w\|_M(t) \|\mathbf{e}_u\|_M(t) + \|\dot{\mathbf{e}}_u\|_M(t),
\]

where we used the local-Lipschitz continuity of $W'$ and $W''$, together with (5.12), and the bounds on $V_h$ obtained by interpolation error estimates (for details, see [3, Lemma 3.1.6]).

The defect terms are bounded, similarly as before, by

\[
- \dot{\mathbf{e}}_u^T \dot{\mathbf{M}}(t) \mathbf{d}_u - \varepsilon^{-1} \mathbf{e}_w^T \dot{\mathbf{M}}(t) \mathbf{d}_w - \varepsilon^{-1} \mathbf{e}_w^T \mathbf{M}(t) \dot{\mathbf{d}}_w(t)
\]

\[
\leq c \|\dot{\mathbf{e}}_u\|_M(t) \|\mathbf{d}_u\|_M(t) + c \varepsilon^{-1} \|\mathbf{e}_w\|_M(t) \|\mathbf{d}_w\|_M(t) + c \varepsilon^{-1} \|\mathbf{e}_w\|_M(t) \|\dot{\mathbf{d}}_w\|_M(t).
\]
The non-linear term is bounded, similarly to (5.24), by
\[ \varepsilon \text{ the product rule again, after a division by } \]
\[ \dot{e} \text{ second from the first equation to cancel the mixed term } \]
\[ \rho > \]
Altogether, by plugging in (5.23)–(5.25) into (5.22), then using Young’s inequalities (possibly with a small number \( \rho > 0 \)) and absorptions to the left-hand side, we obtain the first energy estimate of this part:
\[
\frac{1}{2} \| \dot{e}_u \|_{\dot{M}(t)}^2 + \varepsilon^{-1} \frac{1}{2} \frac{d}{dt} \| \dot{e}_w \|_{\dot{M}(t)}^2 \leq c \| e_u \|_{K(t)}^2 + c \varepsilon^{-1} \| e_w \|_{K(t)}^2 + c (\| d_u \|_{\dot{M}(t)}^2 + \| d_w \|_{\dot{M}(t)}^2).
\] (5.26)

(ii.b) We now test (5.21a) by \( \dot{e}_w^T \) and (5.21b) by \( \dot{e}_u^T \), then subtracting the second from the first equation to cancel the mixed term \( \dot{e}_w^T \dot{M}(t) \dot{e}_u \), then using the product rule again, after a division by \( \varepsilon \) we obtain
\[
\| e_u \|_{A(t)}^2 + \varepsilon^{-1} \frac{1}{2} \frac{d}{dt} \| e_w \|_{A(t)}^2 = - \varepsilon^{-1} \dot{e}_w^T \dot{M}(t) e_u - \varepsilon^{-1} \dot{e}_w^T \dot{M}(t) d_u
\]
\[
- \dot{e}_w^T \dot{A}(t) e_u + \varepsilon^{-1} \dot{e}_u^T \dot{M}(t) e_w + \varepsilon^{-1} \frac{1}{2} \dot{e}_w^T \dot{A}(t) e_w
\]
\[
- \varepsilon^{-2} \dot{e}_u^T \frac{d}{dt} (W(u(t)) - W(u^+(t)))
\]
\[
+ \varepsilon^{-1} \dot{e}_u^T \dot{M}(t) d_w + \varepsilon^{-1} \dot{e}_w^T \dot{M}(t) \dot{d}_w.
\] (5.27)

The terms are again estimated separately. Let us first highlight that it is not possible to directly estimate the terms containing \( \dot{e}_w(t) \) in their current form, because there is no term on the left-hand side to absorb them. Therefore, we first rewrite them using the product rule, and estimate them using Lemma 5.1, to obtain
\[
\dot{e}_w^T \dot{M}(t) e_u = \frac{d}{dt} (\dot{e}_w^T \dot{M}(t) e_u) - e_w^T \dot{M}(t) e_u - e_w^T \dot{M}(t) \dot{e}_u
\]
\[
\leq \frac{d}{dt} (\dot{e}_w^T \dot{M}(t) e_u) + c \| e_w \|_{M(t)} (\| e_u \|_{M(t)} + \| \dot{e}_u \|_{M(t)}),
\] (5.28)
\[
\dot{e}_w^T \dot{M}(t) d_u = \frac{d}{dt} (\dot{e}_w^T \dot{M}(t) d_u) - e_w^T \dot{M}(t) d_u - e_w^T \dot{M}(t) \dot{d}_u
\]
\[
\leq \frac{d}{dt} (\dot{e}_w^T \dot{M}(t) d_u) + c \| e_w \|_{M(t)} (\| d_u \|_{M(t)} + \| \dot{d}_u \|_{M(t)}).
\]

The terms with matrix derivatives are bounded, using Lemma 5.1, by
\[
- \dot{e}_u^T \dot{A}(t) e_u + \varepsilon^{-1} \dot{e}_u^T \dot{M}(t) e_w + \varepsilon^{-1} \frac{1}{2} \dot{e}_w^T \dot{A}(t) e_w
\]
\[
\leq c \| \dot{e}_u \|_{A(t)} \| e_u \|_{A(t)} + c \varepsilon^{-1} \| \dot{e}_u \|_{M(t)} \| e_w \|_{M(t)} + c \varepsilon^{-1} \| e_w \|_{A(t)}^2.
\] (5.29)

The non-linear term is bounded, similarly to (5.24), by
\[
\varepsilon^{-2} \dot{e}_u^T \frac{d}{dt} (W(u(t)) - W(u^+(t))) \leq C \varepsilon^{-2} \| \dot{e}_u \|_{M(t)} (\| \dot{e}_u \|_{M(t)} + \| e_u \|_{M(t)}).
\] (5.30)
The defect terms are bounded, similarly as before, by

$$\begin{align*}
- \hat{\epsilon}^T M(t) \dot{d}_u - \epsilon^{-1} \hat{d}_w^T \dot{M}(t) \dot{d}_w - \epsilon^{-1} \hat{d}_w^T M(t) \ddot{d}_w(t) \\
\leq c\|\hat{\epsilon}^2_u\| M(t) \|\dot{d}_u\| M(t) + c\epsilon^{-1}\|\epsilon w\| M(t) \|\dot{d}_w\| M(t) + c\epsilon^{-1}\|\epsilon w\| M(t) \|\ddot{d}_w\| M(t).
\end{align*}$$

(5.31)

Altogether, by plugging in (5.28)–(5.31) into (5.27), then using Young’s inequalities (possibly with a small number $\varrho > 0$) and absorptions to the left-hand side, we obtain the second energy estimate of this part:

$$\begin{align*}
\frac{1}{2} \|\hat{\epsilon}^2_u\| A(t) + \epsilon^{-1} \frac{1}{2} \frac{d}{dt} \|\epsilon w\| A(t) \leq c_0 \|\hat{\epsilon}^2_u\| M(t) + c\|\epsilon u\|^2 K(t) + c\epsilon^{-1}\|\epsilon w\|^2 K(t) \\
+ c\left(\|d_u\|^2 M(t) + \|\dot{d}_u\|^2 M(t) + \|\dot{d}_w\|^2 M(t) + \|\ddot{d}_w\|^2 M(t)\right) \\
- \epsilon^{-1} \frac{d}{dt} \hat{d}_w^T M(t) \epsilon u - \epsilon^{-1} \frac{d}{dt} \hat{d}_w^T M(t) \epsilon u,
\end{align*}$$

(5.32)

with a particular constant $c_0 > 0$ (independent of $h$, but depending on $W''$, viz. on the constant in (5.24)).

(iii) We now take the weighted combination of the energy estimates from (ii.a) and (ii.b): multiplying the estimate (5.26) by $4c_0$ and adding it to the estimate (5.32). Collecting the terms, (and since $\varrho > 0$ is yet to be chosen, further constants are incorporated into it), we obtain

$$\begin{align*}
2c_0 \|\hat{\epsilon}^2_u\| M(t) + \frac{1}{2} \|\hat{\epsilon}^2_u\| A(t) + \epsilon^{-1} 2c_0 \frac{d}{dt} \|\epsilon w\| M(t) + \epsilon^{-1} \frac{1}{2} \frac{d}{dt} \|\epsilon w\| A(t) \\
\leq c_0 \|\hat{\epsilon}^2_u\| M(t) + c\|\epsilon u\|^2 K(t) + c\epsilon^{-1}\|\epsilon w\|^2 K(t) \\
+ c\left(\|d_u\|^2 M(t) + \|\dot{d}_u\|^2 M(t) + \|\dot{d}_w\|^2 M(t) + \|\ddot{d}_w\|^2 M(t)\right) \\
- \epsilon^{-1} \frac{d}{dt} \hat{d}_w^T M(t) \epsilon u - \epsilon^{-1} \frac{d}{dt} \hat{d}_w^T M(t) \epsilon u.
\end{align*}$$

(5.33)

The first term on the right-hand side of (5.33) can directly be absorbed to the left-hand side, which yields

$$\begin{align*}
c_0 \|\hat{\epsilon}^2_u\| M(t) + \frac{1}{2} \|\hat{\epsilon}^2_u\| A(t) + \epsilon^{-1} 2c_0 \frac{d}{dt} \|\epsilon w\| M(t) + \epsilon^{-1} \frac{1}{2} \frac{d}{dt} \|\epsilon w\| A(t) \\
\leq c\|\epsilon u\|^2 K(t) + c\epsilon^{-1}\|\epsilon w\|^2 K(t) \\
+ c\left(\|d_u\|^2 M(t) + \|\dot{d}_u\|^2 M(t) + \|\dot{d}_w\|^2 M(t) + \|\ddot{d}_w\|^2 M(t)\right) \\
- \epsilon^{-1} \frac{d}{dt} \hat{d}_w^T M(t) \epsilon u - \epsilon^{-1} \frac{d}{dt} \hat{d}_w^T M(t) \epsilon u.
\end{align*}$$
Integrating the above inequality from 0 to \( t \leq t^* \), and then dividing by \( \min\{c_0, 1/2\} \), yields

\[
\|\mathbf{e}_w(t)\|^2_{K(t)} + \varepsilon^{-1} \int_0^t \|\dot{\mathbf{e}}_u(s)\|^2_{K(s)} ds \\
\leq c \int_0^t \|\mathbf{e}_u(s)\|^2_{K(s)} ds + c \varepsilon^{-1} \int_0^t \|\mathbf{e}_w(s)\|^2_{K(s)} ds + ct\|\mathbf{d}_w(0)\|^2_{M(0)} \\
+ c \int_0^t (\|\mathbf{u}(s)\|^2_{M(s)} + \|\dot{\mathbf{u}}(s)\|^2_{M(s)} + \|\mathbf{w}(s)\|^2_{M(s)} + \|\dot{\mathbf{w}}(s)\|^2_{M(s)}) ds \\
- c\varepsilon^{-1}\mathbf{e}_w^T(t)\dot{\mathbf{M}}(t)\mathbf{e}_u(t) + c\varepsilon^{-1}\mathbf{e}_w^T(0)\dot{\mathbf{M}}(0)\mathbf{e}_u(0) \\
- c\varepsilon^{-1}\mathbf{e}_w^T(t)\dot{\mathbf{M}}(t)\mathbf{d}_u(t) + c\varepsilon^{-1}\mathbf{e}_w^T(0)\dot{\mathbf{M}}(0)\mathbf{d}_u(0).
\]

We estimate the newly obtained non-integrated terms on the right-hand side using Lemma 5.1, Cauchy–Schwarz and Young’s inequalities, a further absorption, and using that \( \mathbf{e}_u(0) \) and \( \mathbf{e}_w(0) \) are zero, we then obtain

\[
\|\mathbf{e}_w(t)\|^2_{K(t)} + \varepsilon^{-1} \int_0^t \|\dot{\mathbf{e}}_u(s)\|^2_{K(s)} ds \\
\leq c \int_0^t \|\mathbf{e}_u(s)\|^2_{K(s)} ds + c \varepsilon^{-1} \int_0^t \|\mathbf{e}_w(s)\|^2_{K(s)} ds \\
+ c \int_0^t (\|\mathbf{u}(s)\|^2_{M(s)} + \|\dot{\mathbf{u}}(s)\|^2_{M(s)} + \|\mathbf{w}(s)\|^2_{M(s)} + \|\dot{\mathbf{w}}(s)\|^2_{M(s)}) ds \\
+ c_1\|\mathbf{e}_u(t)\|^2_{M(t)} + c\|\mathbf{d}_u(t)\|^2_{M(t)} + ct\|\mathbf{d}_w(0)\|^2_{M(0)},
\]

(5.34)

with a \( c_1 > 0 \).

This energy estimate now contains the (previously) critical term \( \dot{\mathbf{e}}_u \) on the left-hand side.

Combining the energy estimates: We now take again a \( c_1 \)-weighted linear combination (in order to absorb the term \( c_1\|\mathbf{e}_u\|^2_{M(t)} \) of the two energy estimates (5.20) and (5.34), to obtain

\[
\|\mathbf{e}_u(t)\|^2_{K(t)} + \|\mathbf{e}_w(t)\|^2_{K(t)} + \varepsilon^{-1} \int_0^t \|\dot{\mathbf{e}}_u(s)\|^2_{K(s)} ds + \varepsilon^{-1} \int_0^t \|\mathbf{e}_w(s)\|^2_{K(s)} ds \\
\leq \rho \int_0^t \|\dot{\mathbf{e}}_u(s)\|^2_{M(s)} ds \\
+ c \int_0^t \|\mathbf{e}_u(s)\|^2_{K(s)} ds + c \varepsilon^{-1} \int_0^t \|\mathbf{e}_w(s)\|^2_{K(s)} ds \\
+ c \int_0^t (\|\mathbf{u}(s)\|^2_{M(s)} + \|\dot{\mathbf{u}}(s)\|^2_{M(s)} + \|\mathbf{w}(s)\|^2_{M(s)} + \|\dot{\mathbf{w}}(s)\|^2_{M(s)}) ds \\
+ c\|\mathbf{d}_u(t)\|^2_{M(t)} + ct\|\mathbf{d}_w(0)\|^2_{M(0)},
\]

(5.35)
By choosing $\varrho$ small enough, the first term (previously the critical term) on the left-hand side can now be absorbed. This enables us to use Gronwall’s inequality, which then yields the stated stability estimate on $[0, t^\ast]$.

Now, it only remains to show that, in fact, $t^\ast = T$, for $h$ sufficiently small. The proven stability bound (for $0 \leq t \leq t^\ast$) together with the assumed defect bounds (5.8) imply

$$
\|e_u(t)\|_{K(t)}^2 + \|e_w(t)\|_{K(t)}^2 \leq c h^\kappa, \quad \text{with } \kappa \geq 2.
$$

By an inverse estimate, see, e.g. [4, Theorem 4.5.11], we have, for $0 \leq t \leq t^\ast$,

$$
\|e_{uh}(\cdot, t)\|_{L^\infty(\Gamma_h(t))} \leq c h^{-d/2} \|e_{uh}(\cdot, t)\|_{L^2(\Gamma_h(t))} \leq c \|e_u(t)\|_{K(t)} \leq c Ch^{\kappa-d/2} \leq \frac{1}{2} h^{\kappa-d/2},
$$

(5.36)

for sufficiently small $h$. Therefore the bound (5.10) can be extended to $t \geq t^\ast$, which contradicts the maximality of $t^\ast$ unless we already have $t^\ast = T$. We hence proved the stability bound (5.9) over $[0, T]$, and completed the proof.

\[\square\]

6 Consistency

Before we turn to proving consistency of the spatial semi-discretisation, and to the proof of Theorem 4.2, we collect some preparatory results: error estimates of the nodal interpolations on the surface, for the Ritz map, and some results which estimate various geometric errors. Most of these results were shown in [12, 7, 18].

Let us briefly recall our assumptions on the evolving surface and on its discrete counterpart, from Section 2 and 3.1: $\Gamma(t)$ is a closed smooth (at least $C^2$) surface in $\mathbb{R}^{d+1}$ with $d \leq 3$, evolving with the surface velocity $v$ with $v, \partial v \in L^\infty(0, T; W^{k+1, \infty}(\Gamma(t)))$. The discrete surface $\Gamma_h(t)$ is a $k$-order interpolation of $\Gamma(t)$ at each time, and therefore its velocity $V_h$ is the nodal interpolation of $v$.

6.1 Geometric errors

6.1.1 Interpolation error estimates

The following result gives estimates for the error in the interpolation. Our setting follows that of Section 2.5 of [7].

Let us assume that the surface $\Gamma(t)$ is approximated by the interpolation surface $\Gamma_h(t)$ of order $k$. Then for any $u \in H^{k+1}(\Gamma(t))$, there is a unique $k$-order surface finite element interpolation $\tilde{I}_h u \in \mathcal{S}_h(t)$, furthermore we set $(\tilde{I}_h u)^t = I_h u$. 


Lemma 6.1 Let \( u(\cdot, t) \in H^{k+1}(\Gamma(t)) \) for all \( 0 \leq t \leq T \). The surface interpolation operator of \( I_h \) of order \( k \) satisfies the following error estimates, for \( u = u(\cdot, t) \) and for \( 0 \leq t \leq T \),
\[
\|u - I_h u\|_{L^2(\Gamma(t))} + h\|\nabla_F (u - I_h u)\|_{L^2(\Gamma(t))} \leq c h^{k+1} \|u\|_{H^{k+1}(\Gamma(t))},
\]
\[
\|u - I_h u\|_{L^\infty(\Gamma(t))} + h\|\nabla_F (u - I_h u)\|_{L^\infty(\Gamma(t))} \leq c h^{k+1} \|u\|_{W^{k+1, \infty}(\Gamma(t))},
\]
with a constant \( c > 0 \) independent of \( h \) and \( t \), but depending on \( \mathcal{G}_T \).

6.1.2 Discrete surface velocities

This section gives a definition of a discrete velocity on the exact surface \( \Gamma(t) \) associated to \( V_h \), and explores approximation results for the discrete velocities.

The following result, recalled from [3, Lemma 3.1.6], shows boundedness of the discrete velocity \( V_h \), using the fact that it is the interpolation of \( v \). The proof is based on the interpolation error estimate Lemma 6.1 and the interchange formulas (5.3).

Lemma 6.2 Assume that \( v \) and \( \partial^* v \) are in \( W^{k+1, \infty}(\Gamma(t)) \). Then, for \( h \leq h_0 \) sufficiently small, the following bounds hold:
\[
\|V_h\|_{W^{1, \infty}(\Gamma_h(t))} \leq c \|v\|_{W^{k+1, \infty}(\Gamma(t))},
\]
\[
\|\partial^* V_h\|_{W^{1, \infty}(\Gamma_h(t))} \leq c \|\partial^* v\|_{W^{k+1, \infty}(\Gamma(t))},
\]
\[
\|\partial^*_h (\nabla_{\Gamma_h(t)} \cdot V_h)\|_{L^\infty(\Gamma_h(t))} \leq c (\|\partial^* v\|_{W^{k+1, \infty}(\Gamma(t))} + \|v\|_{W^{k+1, \infty}(\Gamma(t))}^2),
\]
where the constant \( c > 0 \) is independent of \( h \) and \( t \), but depends on \( \mathcal{G}_T \).

We give here an explicit formula for \( v_h \) the discrete surface velocity of \( \Gamma(t) \) associated to \( V_h \): for \( x(t) \in \Gamma_h(t) \) with \( y(t) = x^h(t) \),
\[
v_h(y(t), t) = \partial_t p(x(t), t) + V_h(x(t), t) \cdot \nabla y(x(t), t),
\]
with \( y(x(t), t) \in \Gamma(t) \) denoting the unique solution to the lift, i.e. to \( x(t) = y(x(t), t) + d(x(t), t)\nu(y(x(t), t), t) \) for \( x(t) \in \Gamma_h(t) \). For more details see [12, Definition 4.3].

Apart from the original material derivative \( \partial^* \) on \( \Gamma(t) \), a discrete material derivative associated to the velocity \( v_h \) is also defined, see [12, equation (4.9)], (element-wise) by
\[
\partial^*_h \varphi = \partial_t \varphi + v_h \cdot \nabla \varphi.
\]

That is we have the following three different material derivatives:

for \( \varphi : \Gamma(t) \to \mathbb{R} \) : \( \partial^* \varphi = \partial_t \varphi + v \cdot \nabla \varphi \),
\[
\partial^*_h \varphi = \partial_t \varphi + v_h \cdot \nabla \varphi,
\]
for \( \varphi_h : \Gamma_h(t) \to \mathbb{R} \) : \( \partial^*_h \varphi = \partial_t \varphi_h + V_h \cdot \nabla \varphi_h \).

We note here that it will be always clear from the context whether the discrete material derivative \( \partial^*_h \) is meant on \( \Gamma(t) \) associated to \( v_h \), or on \( \Gamma_h(t) \) associated to \( V_h \).
From [18, Lemma 5.4] we recall high-order error bounds between the velocity $v_h$ of the lifted material points and the surface velocity $v$ (for the case $k = 1, l = 0$ we refer to [12]).

**Lemma 6.3** The difference between the continuous velocity $v$ and the discrete velocity $v_h$ on $\Gamma(t)$ is estimated by

$$\| (\partial^i u_h)(v - v_h)||_{L^\infty(\Gamma(t))} + h\| \nabla_{\Gamma(t)} (\partial^i u_h)(v - v_h)||_{L^\infty(\Gamma(t))} \leq c_l h^{k+1} ,$$

for $l \geq 0$, with a constant $c_l > 0$ independent of $h$ and $t$, but depending on the surface velocity $v$.

Since we need to establish a bound for the discrete material derivatives of both defects $d_u$ and $d_w$, we recall some transport formulas from [12, Lemma 4.2]:

\[
\frac{d}{dt} m(u, \varphi) = m(\partial^i u, \varphi) + m(u, \partial^i \varphi) + g(v; u, \varphi), \quad (6.3a)
\]

\[
\frac{d}{dt} m_h(u_h, \varphi_h) = m_h(\partial^i u_h, \varphi_h) + m(u_h, \partial^i \varphi_h) + g_h(V_h; u_h, \varphi_h). \quad (6.3b)
\]

These formulas will help us to derive equations for $\partial^i d_u$ and $\partial^i d_w$, which can be readily estimated, and will also be often used in the proofs in Section 6.1.4.

### 6.1.3 Error estimates for the generalised Ritz map

From [18, Theorem 6.3 and 6.4] we recall that the generalised Ritz map (5.4) satisfies the following optimal high-order error estimates.

**Lemma 6.4** Let $u : G_T \to \mathbb{R}$ such that $u(\cdot, t)$ and $(\partial^j u(\cdot, t))$ are in $H^{k+1}(\Gamma(t))$ for all $0 \leq t \leq T$ and $j = 1, \ldots, l$. Then the error in the generalised Ritz map (5.4) satisfies the bounds, for $0 \leq t \leq T$ and for $h \leq h_0$ with sufficiently small $h_0$,

\[
\| u - R_h(t) u \|_{L^2(\Gamma(t))} + h\| u - R_h(t) u \|_{H^1(\Gamma(t))} \leq c h^{k+1} \| u \|_{H^{k+1}(\Gamma(t))},
\]

\[
\| \partial_j u_h(v - R_h(t) u) \|_{L^2(\Gamma(t))} + h\| \partial_j u_h(v - R_h(t) u) \|_{H^1(\Gamma(t))} \leq c h^{k+1} \sum_{j=0}^l \| \partial^j u_h \|_{H^{k+1}(\Gamma(t))}, \quad \text{for all } l \geq 1,
\]

where the constant $c > 0$ is independent of $h$ and $t$, but depends on $G_T$.

### 6.1.4 Geometric approximation errors

The time dependent bilinear forms $m, g$ and their discrete counterparts $m_h, g_h$, from (2.7) and (3.5), respectively, satisfy the following high-order geometric approximation estimates, see [18, Lemma 5.6].
Lemma 6.5 Let \( z_h, \varphi_h \in S_h(t) \) with lifts \( z^e_h, \varphi^e_h \in S^e_h(t) \). Then, for all \( h \leq h_0 \) with \( h_0 \) sufficiently small, the following estimates hold

\[
|m(z^e_h, \varphi^e_h) - m_h(z_h, \varphi_h)| \leq ch^{k+1}\|z^e_h\|_{L^2(\Gamma(t))}\|\varphi^e_h\|_{L^2(\Gamma(t))},
\]
\[
|g(v_h; z^e_h, \varphi^e_h) - g_h(V_h; z_h, \varphi_h)| \leq ch^{k+1}\|z^e_h\|_{L^2(\Gamma(t))}\|\varphi^e_h\|_{L^2(\Gamma(t))},
\]

where the constant \( c > 0 \) is independent of \( h \) and \( t \), but depends on \( \mathcal{G}_T \).

Similar results hold for the errors in the bilinear form \( a \), cf. [18, Lemma 5.6], but these are not used herein. The above results furthermore imply

\[
|m(W'(z^e_h), \varphi^e_h) - m_h(W'(z_h), \varphi_h)| \leq ch^{k+1}\|W'(z^e_h)\|_{L^2(\Gamma(t))}\|\varphi^e_h\|_{L^2(\Gamma(t))},
\]
\[
|m(\partial^*_{z_h}W'(z^e_h), \varphi^e_h) - m_h(\partial^*_{z_h}W'(z_h), \varphi_h)| \leq ch^{k+1}\|\partial^*_{z_h}W'(z^e_h)\|_{L^2(\Gamma(t))}\|\varphi^e_h\|_{L^2(\Gamma(t))}.
\]

(6.4)

Below we present and prove a new geometric approximation estimate which relates time derivatives of \( g \) and \( g_h \).

Lemma 6.6 Let \( z_h, \varphi_h \in S_h(t) \) with \( \partial^*_{z_h}z_h, \partial^*_{\varphi_h}\varphi_h \in S_h(t) \), with their corresponding lifts in \( S^e_h(t) \). Then, for all \( h \leq h_0 \) with \( h_0 \) sufficiently small, the following holds

\[
|m((\nabla_{\Gamma(t)} \cdot v_h)^2 z^e_h, \varphi^e_h) + m(\partial^*_{v_h}((\nabla_{\Gamma(t)} \cdot v_h) z^e_h, \varphi^e_h)
- m_h((\nabla_{\Gamma_h(t)} \cdot V_h)^2 z_h, \varphi_h) - m_h(\partial^*_{V_h}((\nabla_{\Gamma_h(t)} \cdot V_h) z_h, \varphi_h))| \leq ch^{k+1}\left(\|z^e_h\|_{L^2(\Gamma(t))}\|\varphi^e_h\|_{L^2(\Gamma(t))} + \|z_h\|_{L^2(\Gamma(t))}\|\varphi_h\|_{L^2(\Gamma(t))}\right)
+ \|\partial^*_{z_h}z^e_h\|_{L^2(\Gamma(t))}\|\varphi^e_h\|_{L^2(\Gamma(t))}
\]

for all \( h \leq h_0 \) with \( h_0 \) sufficiently small, where the constant \( c > 0 \) is independent of \( h \) and \( t \), but depends on the surface velocity \( v \).

Proof Although, this lemma was first proved in [3, Lemma 3.1.8], due to its importance we present it here in full detail.

Let \( \mu_h \) denote the quotient of the measures on \( \Gamma(t) \) and \( \Gamma_h(t) \), we then start by differentiating the integral transformation

\[
m(z^e_h, \varphi^e_h) = m_h(z_h, \varphi_h \mu_h),
\]

with respect to time using the transport formulae (6.3), to obtain

\[
\frac{d}{dt}m(z^e_h, \varphi^e_h) = m(\partial^*_{z^e_h}z^e_h, \varphi^e_h) + m(z^e_h, \partial^*_{\varphi^e_h}\varphi^e_h) + g(v_h; z^e_h, \varphi^e_h)
= \frac{d}{dt}m_h(z_h, \varphi_h \mu_h) = m_h(\partial^*_{z_h}z_h, \varphi_h \mu_h) + m_h(z_h, (\partial^*_{\varphi_h}\varphi_h) \mu_h)
+ g_h(V_h; z_h, \varphi_h \mu_h) + m_h(z_h, (\partial^*_{\varphi_h}\varphi_h) \mu_h).
\]
Using \( \partial^*_h(z^t_h) = (\partial^*_h z_h)^t \) we obtain

\[
g(v_h; z^t_h, \varphi^t_h) - g_h(V_h; z_h, \varphi_h \mu_h) = m_h(\partial^*_h z_h, \varphi_h \mu_h) - m((\partial^*_h z_h)^t, \varphi^t_h) + m_h(z_h, (\partial^*_h \varphi_h) \mu_h) - m(z^t_h, (\partial^*_h \varphi_h)^t) + m_h(z_h, (\partial^*_h \varphi_h) \mu_h) = m_h(z_h, (\partial^*_h \varphi_h) \mu_h).
\]

(6.5)

In particular, for \( \partial^*_h z_h \) in the role of \( z_h \), and with the use of the geometric estimate for the surface measure \( \| \partial^*_h \mu_h \|_{L^\infty} \leq c h^{k+1} \), from [18, Lemma 5.2], we obtain the estimate

\[
g(v_h; \partial^*_h z_h, \varphi^t_h) - g_h(V_h; \partial^*_h z_h, \varphi_h \mu_h) = m_h(\partial^*_h z_h, (\partial^*_h \mu_h) \varphi_h) \leq c h^{k+1} \| \partial^*_h z_h \|_{L^2(I^*_h(t))} \| \varphi_h \|_{L^2(I^*_h(t))},
\]

and, with \( \partial^*_h \varphi_h \) in the role of \( \varphi_h \),

\[
g(v_h; z^t_h, \partial^*_h \varphi^t_h) - g_h(V_h; z_h, (\partial^*_h \varphi_h) \mu_h) = m_h(z_h, \partial^*_h \mu_h \partial^*_h \varphi_h) \leq c h^{k+1} \| \partial^*_h z_h \|_{L^2(I^*_h(t))} \| \partial^*_h \varphi_h \|_{L^2(I^*_h(t))}.
\]

Differentiating equation (6.5) with respect to time, using (6.3), yields

\[
\frac{d}{dt} m_h((\nabla \cdot v_h) z^t_h, \varphi^t_h) = \frac{d}{dt} m_h((\nabla \cdot v_h) z_h, \varphi_h \mu_h) = \frac{d}{dt} m_h(z_h, (\partial^*_h \mu_h) \varphi_h),
\]

computing the derivatives on the left-hand side then leads to

\[
m((\nabla \cdot v_h)^2 z^t_h, \varphi^t_h) - m((\nabla \cdot v_h)^2 z_h, \varphi_h \mu_h) + m(\partial^*_h (\nabla \cdot v_h)^2 z^t_h, \varphi^t_h) - m(\partial^*_h (\nabla \cdot v_h)^2 z_h, \varphi_h \mu_h) = g_h(V_h; (\partial^*_h \varphi_h) \mu_h) - g(v_h; \partial^*_h \varphi_h \mu_h) + \frac{d}{dt} m_h(z_h, (\partial^*_h \mu_h) \varphi_h) + g_h(V_h; z_h, \partial^*_h \varphi_h \mu_h).
\]

The pairs in the first two lines on the right-hand side are already estimated above, while the last term is estimated by the geometric estimate \( \| \partial^*_h \mu_h \|_{L^\infty} \leq c h^{k+1} \). To estimate the remaining derivative term, we first compute the time derivative by (6.3b) and then estimate each term to obtain

\[
\frac{d}{dt} m_h(z_h, (\partial^*_h \mu_h) \varphi_h) = m_h(\partial^*_h z_h, (\partial^*_h \mu_h) \varphi_h) + m_h(z_h, (\partial^*_h \partial^*_h \mu_h) \varphi_h) + m_h(z_h, (\partial^*_h \mu_h) \partial^*_h \varphi_h) + g_h(V_h; z_h, (\partial^*_h \mu_h) \varphi_h) \leq c h^{k+1} \left( \| \partial^*_h z_h \|_{L^2(I^*_h(t))} \| \varphi_h \|_{L^2(I^*_h(t))} + \| \partial^*_h z_h \|_{L^2(I^*_h(t))} \| \partial^*_h \varphi_h \|_{L^2(I^*_h(t))} + \| \partial^*_h \varphi_h \|_{L^2(I^*_h(t))} \right).
\]
using the geometric error estimate \( \| \partial_h^2 \mu_h \|_{L^\infty} \leq c h^{k+1} \), cf. [18, Lemma 5.2].

Altogether by triangle inequalities and by combining the above estimates, we obtain

\[
\begin{align*}
&|m((\nabla v \cdot v_h)^2 z_h^\ell, \varphi_h^\ell) - m_h((\nabla v_h \cdot V_h)^2 z_h, \varphi_h)| \\
&\quad + m(\partial_h^2((\nabla v \cdot v_h)^2 z_h, \varphi_h)) - m_h(\partial_h^2((\nabla v_h \cdot V_h)^2 z_h, \varphi_h)) \\
&\leq |m((\nabla v \cdot v_h)^2 z_h^\ell, \varphi_h^\ell) - m_h((\nabla v_h \cdot V_h)^2 z_h^\ell, \varphi_h^\ell)| \\
&\quad + m(\partial_h^2((\nabla v \cdot v_h)^2 z_h^\ell, \varphi_h^\ell)) - m_h(\partial_h^2((\nabla v_h \cdot V_h)^2 z_h^\ell, \varphi_h^\ell)) \\
&\quad + |m_h((\nabla v_h \cdot V_h)^2 z_h, \varphi_h(\mu_h - 1))| \\
&\quad + |m_h(\partial_h^2((\nabla v_h \cdot V_h)^2 z_h, \varphi_h(\mu_h - 1))| \\
&\leq c h^{k+1} \left( |z_h|_{L^2(\Gamma_h(t))} \| \varphi_h \|_{L^2(\Gamma_h(t))} + |z_h|_{L^2(\Gamma_h(t))} \| \partial_h \varphi_h \|_{L^2(\Gamma_h(t))} \\
&\quad + |z_h|_{L^2(\Gamma_h(t))} \| \partial_h \varphi_h \|_{L^2(\Gamma_h(t))} \right) \\
&\quad + c(\| \mu_h - 1 \|_{L^\infty(\Gamma_h(t))} |z_h|_{L^2(\Gamma_h(t))} \| \varphi_h \|_{L^2(\Gamma_h(t))} \\
&\leq c h^{k+1} \left( |z_h|_{L^2(\Gamma_h(t))} \| \varphi_h \|_{L^2(\Gamma_h(t))} + |z_h|_{L^2(\Gamma_h(t))} \| \partial_h \varphi_h \|_{L^2(\Gamma_h(t))} \\
&\quad + |z_h|_{L^2(\Gamma_h(t))} \| \partial_h \varphi_h \|_{L^2(\Gamma_h(t))} \right),
\end{align*}
\]

where we have used the bounds on the discrete velocity from Lemma 6.2, and the geometric estimate \( \| \mu_h - 1 \|_{L^\infty} \leq c h^{k+1} \), from [18, Lemma 5.2]. \( \square \)

### 6.2 Defect bounds

In this section we prove bounds for the defects and for their time derivatives, i.e. we prove that condition (5.8) of Proposition 5.1 is indeed satisfied.

**Proposition 6.1** Let \( u, w \) solve the Cahn–Hilliard equation on evolving surfaces (2.4). Furthermore, let \( u, w \) and the continuos surface velocity \( v \) be sufficiently smooth, satisfying (4.2) is adequate. Then, for all \( h \leq h_0 \) sufficiently small, and for all \( t \in [0,T] \), the defects are bounded as

\[
\begin{align*}
\| d_u(t) \|_{M(t)} &= \| d_u(t) \|_{L^2(\Gamma_h(t))} \leq c h^{k+1}, \\
\| \partial_h d_u(t) \|_{M(t)} &= \| \partial_h d_u(t) \|_{L^2(\Gamma_h(t))} \leq c h^{k+1}, \\
\| d_w(t) \|_{M(t)} &= \| d_w(t) \|_{L^2(\Gamma_h(t))} \leq c h^{k+1}, \\
\| \partial_h d_w(t) \|_{M(t)} &= \| \partial_h d_w(t) \|_{L^2(\Gamma_h(t))} \leq c h^{k+1},
\end{align*}
\]

where the constant \( c > 0 \) is independent of \( h \) and \( t \), but depends on the bounds on Sobolev norms of \( u, w \) and the surface velocity \( v \).
Proof The Ritz map (5.4) of the exact solutions \( u \) and \( w \) satisfies the discrete problem only up to some defects, \( d_u(\cdot, t) \in S_h(t) \) and \( d_w(\cdot, t) \in S_h(t) \), defined in (5.5). Rewriting these equations using the bilinear form notation from (2.7), we thus have, for an arbitrary \( \varphi_h \in S_h(t) \),

\[
m_h(d_u, \varphi_h) = m_h(\partial^*_h R_h u, \varphi_h) + a_h(R_h w, \varphi_h) + g_h(V_h; R_h u, \varphi_h),
\]

\[
m_h(d_w, \varphi_h) = \varepsilon a_h(R_h u, \varphi_h) + \varepsilon^{-1} m_h(W'(R_h u), \varphi_h) - m_h(R_h w, \varphi_h). \tag{6.7}
\]

Upon subtracting the corresponding equations for the exact solution (2.6) with \( \varphi = \varphi_h^\ell \) and applying the transport formula (6.3a) (with \( \partial^*_h \varphi_h^\ell = 0 \)), from the equations in (6.7), and then adding and subtracting some terms in order to apply the definition of the Ritz map \( R_h \) (5.4), we obtain the following two equations satisfied by the defects \( d_u \) and \( d_w \):

\[
m_h(d_u, \varphi_h) = \left( m_h(\partial^*_h R_h u, \varphi_h) - m(\partial^*_h u, \varphi_h^\ell) \right)
- \left( m_h(R_h w, \varphi_h) - m(w, \varphi_h^\ell) \right)
+ \left( g_h(V_h; R_h u, \varphi_h) - g(v_h; u, \varphi_h^\ell) \right)
= I_u + II_u + III_u, \tag{6.8a}
\]

\[
m_h(d_w, \varphi_h) = -\varepsilon \left( m_h(R_h u, \varphi_h) - m(u, \varphi_h^\ell) \right)
+ \varepsilon^{-1} \left( m_h(W'(R_h u), \varphi_h) - m(W'(u), \varphi_h^\ell) \right)
- \left( m_h(R_h w, \varphi_h) - m(w, \varphi_h^\ell) \right)
= I_w + II_w + III_w. \tag{6.8b}
\]

We now estimate the defects and their material derivatives in the \( L^2(\Gamma(t)) \) norm by bounding each pair on the right-hand sides of the above equations separately, using the geometric estimates from the previous subsection, and using similar techniques as in [12,18]. Since throughout the proofs most norms are on \( \Gamma(t) \), we will omit these below, and write \( L^2, H^{k+1} \) instead of \( L^2(\Gamma(t)), H^{k+1}(\Gamma(t)) \) etc.

**Bound for \( d_u \):** For the pair in the first line, we add and subtract terms to obtain

\[
I_u = \left( m_h(\partial^*_h R_h u, \varphi_h) - m(\partial^*_h R_h u, \varphi_h^\ell) \right) + m(\partial^*_h (R_h u - u), \varphi_h^\ell)
\leq c h^{k+1} \| \partial^*_h R_h u \|_{L^2} \| \varphi_h^\ell \|_{L^2} + c h^{k+1} \left( \| u \|_{H^{k+1}} + \| \partial^* u \|_{H^{k+1}} \right) \| \varphi_h^\ell \|_{L^2}, \tag{6.9}
\]

where we have used Lemma 6.5 together with the fact that \( \partial^*_h (\eta^\ell) = (\partial^*_h \eta)^\ell \), the Ritz map error bound Lemma 6.4. The Ritz map error estimate is again used to show the bound \( \| \partial^*_h R_h u \|_{L^2} \leq c (\| u \|_{H^{k+1}} + \| \partial^* u \|_{H^{k+1}}) \).
By the same techniques, we prove the following bound for $II_u$:

$$II_u \leq \left( m_h(\tilde{R}_h w, \varphi_h) - m(R_h w, \varphi_h^e) \right) - m(R_h w - w, \varphi_h^e)$$

$$\leq ch^{k+1} ||w||_{H^{k+1}} ||\varphi_h^e||_{L^2}$$  \hspace{1cm} (6.10)

The third term $III_u$ is estimated using similar arguments as before, by Lemma 6.5, Lemma 6.4, and Lemma 6.3, as follows

$$III_u = g_h(V_h; \tilde{R}_h u, \varphi_h) - g(v_h; u, \varphi_h^e)$$

$$= \left( g_h(V_h; \tilde{R}_h u, \varphi_h) - g(v_h; R_h u, \varphi_h^e) \right) + g(v_h; R_h u - u, \varphi_h^e)$$  \hspace{1cm} (6.11)

$$\leq ch^{k+1} ||u||_{H^{k+1}} ||\varphi_h^e||_{L^2}.$$  

The estimates (6.9)–(6.11) together, using the norm equivalence (3.6), and the definition of the $L^2(I_h(t))$ norm yields

$$||d_u||_{L^2(I_h(t))} = \sup_{0 \neq \varphi_h \in S_h} \frac{m_h(d_u, \varphi_h)}{||\varphi_h||_{L^2(I_h(t))}} \leq ch^{k+1} \left( ||u||_{H^{k+1}} + ||\partial^* u||_{H^{k+1}} + ||w||_{H^{k+1}} \right).$$  \hspace{1cm} (6.12)

**Bound for $\partial^*_h d_u$:** We start by differentiating the defect equation for $d_u$ (6.8a) with respect to time. Using that $\partial^*_h \varphi_h = \partial^*_h (\varphi_h^e) = 0$, we obtain

$$m_h(\partial^*_h d_u, \varphi_h) = -g_h(V_h; d_u, \varphi_h) + \frac{d}{dt} \left( I_u + II_u + III_u \right).$$

The first term is immediately bounded, using Lemma 6.2, the Cauchy–Schwarz inequality and (6.12), by

$$g_h(V_h; d_u, \varphi_h) \leq ch^{k+1} \left( ||u||_{H^{k+1}} + ||\partial^* u||_{H^{k+1}} + ||w||_{H^{k+1}} \right) ||\varphi_h^e||_{L^2}.$$  \hspace{1cm} (6.13)

The terms differentiated in time are estimated separately, using analogous techniques as before.

For the first term, by the transport formulas (6.3a) and (6.3b), we obtain

$$\frac{d}{dt} I_u = \left( m_h((\partial^*_h)^{(2)} \tilde{R}_h u, \varphi_h) - m((\partial^*_h)^{(2)} u, \varphi_h^e) \right)$$

$$+ \left( g_h(V_h; \partial^*_h \tilde{R}_h u, \varphi_h) - g(v_h; \partial^*_h u, \varphi_h^e) \right) \leq ch^{k+1} \sum_{j=0}^2 ||(\partial^*)^{(j)} u||_{H^{k+1}} ||\varphi_h^e||_{L^2} + ch^{k+1} \sum_{j=0}^1 ||(\partial^*)^{(j)} w||_{H^{k+1}} ||\varphi_h^e||_{L^2},$$  \hspace{1cm} (6.14)

where for the inequality we used the arguments used to show (6.9) and (6.11).
By the same arguments, for the second term we obtain the bound
\[
\frac{d}{dt} I_{II} = - \left( m_h(\partial_h R_h w, \varphi_h) - m(\partial_h w, \varphi_h^l) \right) \\
- \left( g_h(V_h; R_h w, \varphi_h) - g(v_h; w, \varphi_h^l) \right) \\
\leq c h^{k+1} \left( \|u\|_{H^{k+1}} + \|\partial^* u\|_{H^{k+1}} \right) \|\varphi_h^l\|_{L^2}. 
\] (6.15)

By the time differentiation of the third term, using the transport formulas (6.3a) and (6.3b), we obtain
\[
\frac{d}{dt} III_u = \frac{d}{dt} \left( g_h(V_h; R_h u, \varphi_h) - g(v_h; u, \varphi_h^l) \right) \\
= m_h(\partial_h^* (\nabla \Gamma_h \cdot V_h) R_h u, \varphi_h) + m_h(\nabla \Gamma_h \cdot V_h)^2 R_h u, \varphi_h \\
- m(\partial_h^* (\nabla R \cdot v_h) u, \varphi_h^l) - m((\nabla R \cdot v_h)^2 u, \varphi_h^l) \\
+ \left( g_h(V_h, \partial_h^* R_h u, \varphi_h) - g(v_h, \partial_h^* u, \varphi_h^l) \right) =: II^1_I + II^2_I. 
\]

The pair in the third line is estimated by previous arguments just as before, by
\[
II^2_I \leq c h^{k+1} \left( \|u\|_{H^{k+1}} + \|\partial^* u\|_{H^{k+1}} \right) \|\varphi_h^l\|_{L^2}. 
\] (6.16)

The remaining pair in the rectangular brackets is estimated by similar ideas as above, adding and subtracting intermediate terms, using the geometric approximation estimate from Lemma 6.6, Ritz map error estimates Lemma 6.4, and Lemma 6.3 (for \( l = 0 \) and \( 1 \)) and Lemma 6.2:
\[
II^1_I = \left( m_h(\partial_h^* (\nabla \Gamma_h \cdot V_h) R_h u, \varphi_h) + m_h(\nabla \Gamma_h \cdot V_h)^2 R_h u, \varphi_h \right) \\
- m(\partial_h^* (\nabla R \cdot v_h) R_h u, \varphi_h^l) - m((\nabla R \cdot v_h)^2 R_h u, \varphi_h^l) \\
+ m(\partial_h^* (\nabla R \cdot v_h)(R_h u - u), \varphi_h^l) + m((\nabla R \cdot v_h)^2(R_h u - u), \varphi_h^l) \\
\leq c h^{k+1} \left( \|R_h u\|_{L^2} + \|\partial^* R_h u\|_{L^2} \right) \|\varphi_h^l\|_{L^2} + c h^{k+1} \|u\|_{H^{k+1}} \|\varphi_h^l\|_{L^2} \\
\leq c h^{k+1} \left( \|u\|_{H^{k+1}} + \|\partial^* u\|_{H^{k+1}} \right) \|\varphi_h^l\|_{L^2}. 
\] (6.17)

The combination of the estimates (6.13)–(6.17), using the norm equivalence (3.6), yields
\[
\|\partial^* d_u\|_{L^2(\Gamma_h(t))} \leq c h^{k+1} \left( \sum_{j=0}^{2} \|\partial^* (j) u\|_{H^{k+1}} + \sum_{j=0}^{1} \|\partial^* (j) w\|_{H^{k+1}} \right). 
\] (6.18)

**Bound for \( d_w \):**

The \( L^2 \) norm of the defect \( d_w \) (6.8b) is estimated by the same techniques by which the bound (6.10) was shown. Due to the (locally Lipschitz continuous)
non-linear term $W^{'}$ we need an $L^{\infty}$ bound on the Ritz map, which we obtain by

$$
\|R_{h}u\|_{L^{\infty}} \leq \|R_{h}u - I_{h}u\|_{L^{\infty}} + \|I_{h}u\|_{L^{\infty}}
$$

$$
\leq ch^{-d/2}\|R_{h}u - I_{h}u\|_{L^{2}} + \|I_{h}u\|_{L^{\infty}}
$$

$$
\leq ch^{-d/2}(\|R_{h}u - u\|_{L^{2}} + \|u - I_{h}u\|_{L^{2}}) + \|I_{h}u - u\|_{L^{\infty}} + \|u\|_{L^{\infty}}
$$

$$
\leq ch^{k+1-d/2}\|u\|_{H^{k+1}} + (ch^{2} + 1)\|u\|_{W^{2,\infty}},
$$

(6.19)

using an inverse estimate [4, Theorem 4.5.11], interpolation error bounds Lemma 6.1, and for the last term the (sub-optimal) interpolation error estimate of [7, Proposition 2.7] (with $p = \infty$).

By similar techniques as before, and using (6.4) together with (6.19) the pairs for $d_{w}$ are estimated analogously. The bounds for $I_{w}$ and $III_{w}$ are straightforward, while $II_{w}$ is bounded, using the local Lipschitz continuity of $W^{'}$ and (6.19), by

$$
\varepsilon II_{w} = m_{h}(W^{'}(\tilde{R}_{h}u), \varphi_{h}) - m(W^{'}(R_{h}u), \varphi_{h}^{k})
$$

$$
+ m(W^{'}(R_{h}u - W^{'}(u), \varphi_{h}^{k})
$$

$$
\leq ch^{k+1}\|W^{'}(R_{h}u)||L^{2}\|\varphi_{h}^{k}\|_{L^{2}} + c\|R_{h}u - u\|_{L^{2}}\|\varphi_{h}^{k}\|_{L^{2}}
$$

$$
\leq ch^{k+1}(\|W^{'}(R_{h}u)||L^{2} + \|u\|_{H^{k+1}})\|\varphi_{h}^{k}\|_{L^{2}}
$$

$$
\leq ch^{k+1}(\|W^{'}(R_{h}u) - W^{'}(u)||L^{2} + \|W^{'}(u)||L^{2} + \|u\|_{H^{k+1}})\|\varphi_{h}^{k}\|_{L^{2}}
$$

$$
\leq ch^{k+1}(c\|u\|_{H^{k+1}} + \|W^{'}(u)||L^{2} + \|u\|_{H^{k+1}})\|\varphi_{h}^{k}\|_{L^{2}}.
$$

(6.20)

We altogether obtain the estimate

$$
\|d_{w}\|_{L^{2}(\Gamma_{h}(t))} \leq ch^{k+1}\left(\|u\|_{H^{k+1}} + \|w\|_{H^{k+1}} + \|u\|_{W^{2,\infty}}\right).
$$

(6.21)

**Bound for $\partial_{h}^{\ast}d_{w}$:** Just as for $\partial_{h}^{\ast}d_{w}$, we differentiate the expression (6.8b) with respect to time. Using again $\partial_{h}^{\ast}\varphi_{h} = \partial^{\ast}(\varphi_{h}^{k}) = 0$, we obtain

$$
m_{h}(\partial_{h}^{\ast}d_{w}, \varphi_{h}) = -g_{h}(V_{h}; d_{w}, \varphi_{h}) + \frac{d}{dt}(I_{w} + II_{w} + III_{w}).
$$

The first term is estimated using (6.21), while the remaining terms are bounded similarly to (6.15) and (6.20) (using (6.19)).

Altogether, we obtain

$$
\|\partial_{h}^{\ast}d_{w}\|_{L^{2}(\Gamma_{h}(t))} \leq ch^{k+1}\left(\sum_{j=0}^{1}\left(\|\partial^{(j)}u\|_{H^{k+1}} + \|\partial^{(j)}w\|_{H^{k+1}}\right) + \|u\|_{W^{2,\infty}}\right).
$$

(6.22)
7 Proof of Theorem 4.2

Proof (Proof of Theorem 4.2) We combine the stability bound of Proposition 5.1, and the consistency estimates of Proposition 6.1.

The errors are split as follows

\[ u - u_h^\ell = u - R_h u + (u_h^* - u_h)^\ell, \]
\[ w - w_h^\ell = w - R_h w + (w_h^* - w_h)^\ell, \]
\[ \partial^\bullet(u - u_h^\ell) = \partial^\bullet(u - R_h u) + (\partial^\bullet(u_h^* - u_h))^\ell, \]

upon recalling that \( u_h^* = \tilde{R}_h u \) and \( w_h^* = \tilde{R}_h w \).

The first terms in each error are directly and similarly bounded by error estimates for the Ritz map Lemma 6.4 – uniformly in time – by

\[ \|u - R_h u\|_{L^2(\Gamma(t))} + h\|u - R_h u\|_{H^1(\Gamma(t))} \leq ch^{k+1}\|u\|_{H^{k+1}(\Gamma(t))}. \]

The second terms are the errors \( e_{u_h}, e_{w_h} \) and \( \partial_h^\bullet e_{u_h} \), therefore bounded by the combination of the stability estimate (5.9) and the consistency estimates (6.6). In Proposition 5.1 the \( L^\infty \) norm assumption on \( u_h^* = \tilde{R}_h u \) was proved in (6.19). Altogether, we obtain

\[ \|e_{u_h}\|_{H^1(\Gamma_h(t))}^2 + \|e_{w_h}\|_{H^1(\Gamma_h(t))}^2 + \int_0^t \|\partial_h^\bullet e_{u_h}\|_{H^1(\Gamma_h(s))}^2 ds \leq ch^{2(k+1)}. \]

By combining the above estimates we obtain the stated error estimates. \( \square \)

8 Extensions to non-linear problems

In this section we will expand on Cahn–Hilliard equations with a more general non-linearity, such as the ones briefly discussed after Theorem 4.2, which includes Cahn–Hilliard equations with a proliferation term [27, equation (3.1)], and which is just slightly less general than the fourth-order non-linear parabolic PDE in [8].

We will now allow the non-linearities to depend on the gradients of \( u \) and \( w \), and consider the general non-linear Cahn–Hilliard equation:

\[ \partial^\bullet u - \Delta_{\Gamma(t)} u = g(u, \nabla_{\Gamma(t)} u) - u(\nabla_{\Gamma(t)} \cdot v), \quad \text{on } \Gamma(t), \]
\[ w + \Delta_{\Gamma(t)} w = f(u, \nabla_{\Gamma(t)} u), \]

on \( \Gamma(t) \) (8.1)

The corresponding modified evolving surface finite element semi-discretisation of the above problem can be written in the matrix–vector formulation:

\[ M(t)\dot{u}(t) + A(t)w(t) = g(u(t)) - \dot{M}(t)u(t), \]
\[ M(t)w(t) - A(t)u(t) = f(u(t)) + \vartheta, \]

where \( M(t) \) is the mass matrix and \( A(t) \) is the stiffness matrix.
with the same correction term \( \vartheta \) as before, see (3.12). The non-linear terms 
\( g(u(t)) \) and \( f(u(t)) \) are then defined by

\[
\begin{align*}
  g(u(t))|_k &= m_h(g(u_h, \nabla \Gamma_h(t) u_h), \phi_k(\cdot, t)), \\
  f(u(t))|_k &= m_h(f(u_h, \nabla \Gamma_h(t) u_h), \phi_k(\cdot, t)), \\
  k &= 1, \ldots, N.
\end{align*}
\]

The stability analysis of this semi-linear problem only requires the following modifications in the proof of Proposition 5.1:

- Since the non-linear terms now include the gradient of \( u \) as well, we need to establish a uniform-in-time bound of the \( W^{1,\infty}(\Gamma_h) \) norm of the errors in \( u \). This is achieved analogously to (5.10), though the definition of \( t^* \) should now include a \( W^{1,\infty}(\Gamma_h) \) norm bound for the errors \( e_u \), but otherwise the exact same as (5.10). It is key to note that the arguments of (5.36) remain valid, in particular the first inequality holds between \( \| \cdot \|_{W^{1,\infty}(\Gamma_h)} \) and \( \| \cdot \|_{H^{1}(\Gamma_h)} \).

- This \( W^{1,\infty}(\Gamma_h) \) norm bound allows to estimate the non-linear terms \( f \) and \( g \), similarly as it was done for \( W \) in (5.15)–(5.16). For instance, the estimates for the non-linear term \( f \) then take the form:

\[
z^T (f(u) - f(u^*)) \leq c \| z \|_{M(t)} (\| e_u \|_{M(t)} + \| e_u \|_{K(t)}),
\]

and similarly for \( g \), while in Part (ii) the time differentiated terms with \( f \) are bounded as

\[
z^T \frac{d}{dt} (f(u) - f(u^*)) \leq c \| z \|_{M(t)} (\| e_u \|_{K(t)} + \| \dot{e}_u \|_{K(t)}).
\]

For these estimates we refer to the general estimates in Part (v) of the proof of Proposition 7.1 in [19].

In Part (ii), terms tested with \( \dot{e}_w \) are first rewritten using the arguments of (5.28), and then estimated via the inequalities above, while the time derivatives of the errors \( e_u \) need to be absorbed during the combination of (ii.a) and (ii.b). Note that the \( K \) norms of the errors already appear on the right-hand side of (5.35), and hence cause no further issues during the Gronwall argument.

- Note, however, that it is not possible to include \( w \) in the non-linear term \( f \). In Part (ii), such a term would lead to an estimate

\[
\dot{e}_w(t)^T \frac{d}{dt} \left( f(u(t), w(t)) - f(u^*(t), w^*(t)) \right),
\]

which cannot be bounded without a new critical term \( \| \dot{e}_w \|_{K(t)} \), even with the argument of (5.28). On the other hand, with a few further tricks it would be possible to consider non-linearities \( g(u, w) \) which are linear in \( w \), see, e.g. [20], but we do not elaborate on such situations here.

Altogether, the following general stability estimate holds.
Proposition 8.1 Let the conditions of Proposition 5.1 be satisfied. Assume further that the functions \( g \) and \( f \) and their derivatives are locally Lipschitz continuous. Then there exists \( h_0 > 0 \) such that the errors \( e_u \) and \( e_w \), corresponding to the non-linear Cahn–Hilliard equation (8.1), satisfy the following bounds, for \( h \leq h_0 \) and \( 0 \leq t \leq T \),

\[
\|e_u(t)\|_K(t) + \|e_w(t)\|_K(t) + \int_0^t \|\dot{e}_u(s)\|_K(s)ds \\
\leq C \int_0^t \|d_u(s)\|_M(s) + \|\dot{d}_u(s)\|_{M(s)} + \|d_w(s)\|_{M(s)} + \|\dot{d}_w(s)\|_{M(s)}ds \\
+ C\|d_u(t)\|_M(t) + Ct\|d_u(0)\|_{M(0)}.
\]

The constant \( C > 0 \) is independent of \( t \) and \( h \), but depends on \( 0 < \epsilon < 1 \), and exponentially on the final time \( T \).

Therefore, subject to a straightforward revision of the consistency analysis, the above general stability result Proposition 8.1 implies optimal-order error estimates for the general non-linear Cahn–Hilliard equation, analogous to Theorem 4.2.

Theorem 8.1 Let \( u \) and \( w \) be the sufficiently regular weak solutions of the above general non-linear Cahn–Hilliard equation on an evolving surface (8.1).

Then there exists an \( h_0 > 0 \) such that for all \( h \leq h_0 \) the error between the solutions \( u \) and \( w \) and the degree \( k \) evolving surface finite element solutions \( u_h \) and \( w_h \) of the modified system satisfy the optimal-order uniform-in-time error estimates in both variables, for \( 0 \leq t \leq T \),

\[
\|u_h(t) - u(t)\|_{L^2(\Gamma(t))} + h\|u_h(t) - u(t)\|_{H^1(\Gamma(t))} \leq C h^{k+1},
\]

\[
\|w_h(t) - w(t)\|_{L^2(\Gamma(t))} + h\|w_h(t) - w(t)\|_{H^1(\Gamma(t))} \leq C h^{k+1}.
\]

9 Full discretisation via linearly implicit backward difference formulae

We recall the matrix–vector formulation from (3.14):

\[
\frac{d}{dt} \left(M(t)u(t)\right) + A(t)w(t) = 0, \\
M(t)w(t) - \epsilon A(t)u(t) = \epsilon^{-1}W(u(t)).
\]

As a time discretisation, we consider the linearly implicit \( s \)-step backward differentiation formulae (BDF). For a step size \( \tau > 0 \), and with \( t_n = n\tau \leq T \), the discretised time derivative is determined by

\[
\dot{u}^n = \frac{1}{\tau} \sum_{j=0}^s \delta_j u^{n-j}, \quad n \geq s,
\]

(9.2)
while the non-linear term uses an extrapolated value, and reads as:

\[ W(\tilde{u}^n) := W\left(\sum_{j=0}^{s-1} \gamma_j u_{n-1-j}\right), \quad n \geq s. \]

We determine the approximations to the variables \( u^n \) to \( u(t_n) \) and \( w^n \) to \( w(t_n) \) by the fully discrete system of linear equations, for \( n \geq s \),

\[
\begin{pmatrix}
\delta_0 M(t_n) \tau A(t_n) & u_n \\
-\varepsilon A(t_n) M(t_n) & w_n
\end{pmatrix} = \begin{pmatrix}
-\sum_{j=1}^{s} \delta_j M(t_{n-j}) u_{n-j} \\
\varepsilon^{-1} W(\sum_{j=0}^{s-1} \gamma_j u_{n-1-j})
\end{pmatrix},
\]

which is used for the upcoming numerical experiments. The starting values \( u^i \) and \( w^i \) \((i = 0, \ldots, s - 1)\) are assumed to be given. They can be precomputed using either a lower order method with smaller step sizes, or an implicit Runge–Kutta method.

The method is determined by its coefficients, given by \( \delta(\zeta) = \sum_{j=0}^{s} \delta_j \zeta^j = \frac{1}{s} \left(1 - \zeta^s\right) \) and \( \gamma(\zeta) = \sum_{j=0}^{s-1} \gamma_j \zeta^j = (1 - (1 - \zeta^s) / \zeta). \) The classical BDF method is known to be zero-stable for \( s \leq 6 \) and to have order \( s \); see [16, Chapter V]. This order is retained by the linearly implicit variant using the above coefficients \( \gamma_j \); cf. [2,1].

The anti-symmetric structure of the system is preserved, and can be observed in (9.3). Since the idea of energy estimates, using the \( G \)-stability theory of Dahlquist [6] and the multiplier technique of Nevanlinna & Odeh [28], can be transferred to linearly implicit BDF full discretisations (up to order 5), we strongly expect that Proposition 5.1 translates to the fully discrete case, and so does the convergence result Theorem 4.2; and also the generalisations in Section 8. This is strengthened by the successful application of these techniques to the analogous linearly implicit backward difference methods applied to evolving surface PDEs: [26,24,23] showing optimal-order error bounds for various problems on evolving surfaces. The method was also analysed for various geometric surface flows, for \( H^1 \)-regularised surface flows [22], and for mean curvature flow [19], both proving optimal-order error bounds for full discretisations.

10 Numerical experiments

In this section we report on numerical experiments, using (9.3), for the nonlinear Cahn–Hilliard equation on an evolving surface:

- We perform a convergence test for the non-linear Cahn–Hilliard equation with the linear evolving surface FEM and BDF methods of various order. We would like to note here that [15] only presents errors and EOCs for a linear problem (using the linearly implicit Euler method).
- We perform the same experiment as Elliott and Ranner in [15, Section 6.2], i.e. we report on the evolution of the Ginzburg–Landau energy along the surface evolution for the non-linear Cahn–Hilliard equation using the first and second order BDF methods.
In the numerical experiments we use the non-linear Cahn–Hilliard equation on an evolving surface (2.4) with the double-well potential, hence the non-linear term is $W'(u) = u^3 - u$. Formulated as a system the problem reads:

$$
\begin{align*}
\partial_t u - \Delta_{\Gamma(t)} u &= -u(\nabla_{\Gamma(t)} \cdot v) + f & \text{on } \Gamma(t), \\
\epsilon \Delta_{\Gamma(t)} w &= \lambda W'(u) & \text{on } \Gamma(t),
\end{align*}
$$
(10.1)

with an extra inhomogeneity $f(\cdot, t) : \Gamma(t) \to \mathbb{R}$, chosen such that the exact solution is known to be $u(x, t) = e^{-6t}x_1x_2$, while $w$ is also explicitly known through the second equation of (10.1). The surface $\Gamma(t)$ evolves time-periodically from a sphere into an ellipsoid and back. In particular the surface is given as the zero level set of a distance function:

$$
\Gamma(t) = \{ x \in \mathbb{R}^3 | d(x, t) := a(t)^{-1}x_1^2 + x_2^2 + x_3^2 - 1 = 0 \},
$$
(10.2)

with $a(t) = 1 + 0.25 \sin(2\pi t)$. The initial surface $\Gamma(0) = \Gamma^0$ is the unit sphere. The surface evolution is computed using the ODE for the positions (2.1), with $v = -V\nu$, $V = (\partial_t d) / |\nabla d|$ and $\nu = \nabla d / |\nabla d|$.

For the numerical experiments the ODE was solved numerically by the classical 4th order Runge–Kutta method with the smallest time step size present in the experiment.

Various numerical experiments have been carried out using the same evolving surface, in particular also for the Cahn–Hilliard equation by Elliott and Ranner [15], and for other problems as well, see, for instance [10,26].

The initial value $u_0$ is the interpolation of the exact initial value $u_0$. For high-order BDF methods the required additional starting values $u_i$ (for $i = 1, \ldots, q - 1$) are taken as the interpolation of the exact values, if they exist, as well or are otherwise computed using a cascade of steps performed by the preceding lower order method.

10.1 Convergence experiments

In these experiments we have used the parameter $\epsilon = 0.5$. The final time is $T = 1$, the time discretisations use a sequence of time step sizes $\tau = 0.2 \cdot 2^{-i}$ for $i = 1, \ldots, 7$, and a sequence of initial meshes with (roughly quadrupling) degrees of freedom as reported in the figures.

In Figures 10.1–10.4 we report on the $L^\infty(L^2)$ norm errors (left) and $L^\infty(H^1)$ norm errors (right) between the numerical and exact solution for both variables $u$ and $w$, i.e. the plots show the errors

$$
\|u - u_h\|_{L^\infty(L^2)} + \|w - w_h\|_{L^\infty(L^2)} \quad \text{and} \quad \|u - u_h\|_{L^\infty(H^1)} + \|w - w_h\|_{L^\infty(H^1)},
$$

where the norms are understood as

$$
\|u - u_h\|_{L^\infty(L^2)} = \max_{0 \leq n \tau \leq T} \|u(\cdot, n\tau) - (u_h^n)\|^2_{L^2(\Gamma(n\tau))}.
$$
For the first order BDF method, Figure 10.1 shows logarithmic plots of the errors against the mesh width $h$, the lines marked with different symbols correspond to different time step sizes. We also report on temporal convergence in Figure 10.2, where the roles are reversed, the errors are plotted against the time step size $\tau$, and the lines with different markers correspond to different mesh refinements.

In Figure 10.1 we can observe two regions: a region where the spatial discretisation error dominates, matching to the order of convergence of our theoretical results of Theorem 4.2 (note the reference lines), and a region, with small mesh widths, where the temporal discretisation error dominates (the error curves flatten out). For the $H^1$ norm we observe better convergence rates as predicted. For Figure 10.2, the same description applies, but with reversed roles. Although, we do not study convergence of full discretisations, the classical order of the BDF methods can be observed. We note here, that flat error curves, which were completely dominated by a discretisation error, were not plotted.

Figure 10.3 and 10.4 report on the same plots, but for the third order BDF method. Again, both the spatial and temporal convergence, as shown by the figures, are in agreement with the theoretical convergence results of Theorem 4.2 and with the classical orders of the BDF methods (note the reference lines).

The plots for time convergence, Figures 10.2 and 10.4, are supporting our claim that Theorem 4.2 can be extended for full discretisations with linearly implicit BDF methods, which is left to a subsequent work.

![Fig. 10.1] Spatial convergence of the BDF1 / linear ESFEM discretisation for the non-linear Cahn–Hilliard equation on an evolving ellipsoid
Fig. 10.2 Temporal convergence of the BDF1 / linear ESFEM discretisation for the non-linear Cahn–Hilliard equation on an evolving ellipsoid

Fig. 10.3 Spatial convergence of the BDF3 / linear ESFEM discretisation for the non-linear Cahn–Hilliard equation on an evolving ellipsoid

Fig. 10.4 Temporal convergence of the BDF3 / linear ESFEM discretisation for the non-linear Cahn–Hilliard equation on an evolving ellipsoid
10.2 The Ginzburg–Landau energy

The numerical experiments in [15, Section 6.2] reporting on the Ginzburg–Landau energy were repeated here for high-order BDF methods.

We again consider the non-linear Cahn–Hilliard equation (10.1), with $\varepsilon = 0.1$ and with $f = 0$ on the same evolving surface $\Gamma(t)$ as before, but with $a(t) = 1 + 0.25 \sin(10\pi t)$, and with initial value

$$u_0(x) = 0.1 \cos(2\pi x_1) \cos(2\pi x_2) \cos(2\pi x_3).$$

This setting is the same as in [15, Section 6.2].

In Figure 10.5 and 10.6 we report on the time evolution of the Ginzburg–Landau energy (until $T = 0.2$ and $T = 1$) of the BDF2 / linear ESFEM discretisation. In both plots we have used the time step size $\tau = 10^{-4}$ (the same as [15, Section 6.2]), and eight different mesh refinement levels (higher numbering denotes finer meshes). The meshes are not nested refinements of a single coarse grid, the coarsest has 54 while the finest has 10146 nodes.

As it was pointed out by Elliott and Ranner [15] “the energy does not decrease monotonically along solutions”, see Figure 10.5, and as they predicted the solutions converge to a time-periodic solution, the periodicity in their energies can be nicely observed in Figure 10.6.

![Ginzburg-Landau energy](image)

**Fig. 10.5** The Ginzburg–Landau energy over $[0, 0.2]$ for BDF2 / linear ESFEM discretisation with $\tau = 10^{-4}$ and over several spatial refinements.

**Acknowledgement**

We thank Christian Lubich for helpful discussions, in particular on initial values.

The work of Cedric Aaron Beschle is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 251654672 – TRR 161. The work of Balázs Kovács is supported by Deutsche Forschungsgemeinschaft – Project-ID 258734477 – SFB 1173.
Fig. 10.6 The Ginzburg–Landau energy over $[0,1]$ for BDF2 / linear ESFEM discretisation with $\tau = 10^{-4}$ and over several spatial refinements.

References

1. G. Akrivis, B. Li, and C. Lubich. Combining maximal regularity and energy estimates for time discretizations of quasilinear parabolic equations. Math. Comp., 86(306):1527–1552, 2017.
2. G. Akrivis and C. Lubich. Fully implicit, linearly implicit and implicit-explicit backward difference formulae for quasi-linear parabolic equations. Numerische Mathematik, 131(4):713–735, 2015.
3. C. Beschle. Error estimates for the Cahn–Hilliard equation on evolving surfaces. University of Tübingen, Master thesis. 2019.
4. S. C. Brenner and R. Scott. The mathematical theory of finite element methods, volume 15. Springer, Berlin, 2008.
5. J. Cahn and J. Hilliard. Free energy of a nonuniform system. i. interfacial free energy. J. Chem. Phys., 28(2):258–267, 1958.
6. G. Dahlquist. G–stability is equivalent to A–stability. BIT, 18(4):384–401, 1978.
7. A. Demlow. Higher-order finite element methods and pointwise error estimates for elliptic problems on surfaces. SIAM Journal on Numerical Analysis, 47(2):805–827, 2009.
8. T. Dlotko. Fourth order semilinear parabolic equations. Tsukuba J. Math., 16(2):389–405, 1992.
9. Q. Du, L. Ju, and L. Tian. Finite element approximation of the Cahn–Hilliard equation on surfaces. Comput. Methods Appl. Mech. Engrg., 200(29-32):2458–2470, 2011.
10. G. Dziuk and C. M. Elliott. Finite elements on evolving surfaces. IMA Journal of Numerical Analysis, 27(2):262–292, 2007.
11. G. Dziuk and C. M. Elliott. Finite element methods for surface PDEs. Acta Numer., 22:289–396, 2013.
12. G. Dziuk and C. M. Elliott. $L^2$–estimates for the evolving surface finite element method. Mathematics of Computation, 82(281):1–24, 2013.
13. G. Dziuk, D. Kröner, and T. Müller. Scalar conservation laws on moving hypersurfaces. Interfaces and Free Boundaries, 15(2):203–236, 2013.
14. G. Dziuk, C. Lubich, and D. Mansour. Runge–Kutta time discretization of parabolic differential equations on evolving surfaces. IMA Journal of Numerical Analysis, 32(2):394–416, 2012.
15. C. M. Elliott and T. Ranner. Evolving surface finite element method for the Cahn–Hilliard equation. Numer. Math., 129(3):483–534, 2015.
16. E. Hairer and G. Wanner. Solving Ordinary Differential Equations II.: Stiff and differential–algebraic problems. Springer, Berlin, Second edition, 1996.
17. P. Harder and B. Kovács. Error estimates for the Cahn–Hilliard equation with dynamic boundary conditions. 2020. arXiv:2005.03349.
18. B. Kovács. High-order evolving surface finite element method for parabolic problems on evolving surfaces. *IMA Journal of Numerical Analysis*, 38(1):430–459, 2018.
19. B. Kovács, B. Li, and C. Lubich. A convergent evolving finite element algorithm for mean curvature flow of closed surfaces. *Numerische Mathematik*, 143(4):797–853, 2019.
20. B. Kovács, B. Li, and C. Lubich. A convergent evolving finite element algorithm for willmore flow. 2020. (in preparation).
21. B. Kovács, B. Li, C. Lubich, and C. Power Guerra. Convergence of finite elements on an evolving surface driven by diffusion on the surface. *Numerische Mathematik*, 137(3):643–689, 2017.
22. B. Kovács and C. Lubich. Linearly implicit full discretization of surface evolution. *Numer. Math.*, 140(1):121–152, 2018.
23. B. Kovács and C. Power Guerra. Error analysis for full discretizations of quasilinear parabolic problems on evolving surfaces. *Numerical Methods for Partial Differential Equations*, 32(4):1200–1231, 2016.
24. B. Kovács and C. Power Guerra. Higher order time discretizations with ALE finite elements for parabolic problems on evolving surfaces. *IMA J. Numer. Anal.*, 38(1):460–494, 2018.
25. C. Lubich and D. Mansour. Variational discretization of wave equations on evolving surfaces. *Mathematics of Computation*, 84(292):513–542, 2015.
26. C. Lubich, D. Mansour, and C. Venkataraman. Backward difference time discretization of parabolic differential equations on evolving surfaces. *IMA J. Numer. Anal.*, 33(4):1365–1385, 2013.
27. A. Miranville. The Cahn–Hilliard equation and some of its variants. *AIMS Mathematics*, 2(3):479, 2017.
28. O. Nevanlinna and F. Odeh. Multiplier techniques for linear multistep methods. *Numer. Funct. Anal. Optim.*, 3:377–423, 1981.
29. D. O’Connor and B. Stinner. The Cahn–Hilliard equation on an evolving surface. arXiv:1607.05627, 2016.