Journal of Computational Physics 230 (2011) 1686–1706

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Finite element approximation of nematic liquid crystal flows using a saddle-point structure

Santiago Badia [a,*], Francisco Guillén-González [b], Juan Vicente Gutiérrez-Santacreu [c]

[a] International Center for Numerical Methods in Engineering (CIMNE), Universitat Politècnica de Catalunya, Jordi Girona 1-3, Edifici C1, 08034 Barcelona, Spain
[b] Dpto. de Ecuaciones Diferenciales y Análisis Numérico, Universidad Sevilla. Avda. Reina Mercedes, s/n. 41012 Sevilla, Spain
[c] Dpto. de Matemática Aplicada I, Universidad Sevilla. Avda. Reina Mercedes, s/n. 41012 Sevilla, Spain

ARTICLE INFO

Article history:
Received 9 May 2010
Received in revised form 16 November 2010
Accepted 19 November 2010
Available online 27 November 2010

Keywords:
Nematic liquid crystals
Finite element methods
Saddle-point problems
Ericksen–Leslie problem
Ginzburg–Landau problem

ABSTRACT

In this work, we propose finite element schemes for the numerical approximation of nematic liquid crystal flows, based on a saddle-point formulation of the director vector subproblem. It introduces a Lagrange multiplier that allows to enforce the sphere condition. In this setting, we can consider the limit problem (without penalty) and the penalized problem (using a Ginzburg–Landau penalty function) in a unified way. Further, the resulting schemes have a stable behavior with respect to the value of the penalty parameter, a key difference with respect to the existing schemes. Two different methods have been considered for the time integration. First, we have considered an implicit algorithm that is unconditionally stable and energy preserving. The linearization of the problem at every time step value can be performed using a quasi-Newton method that allows to decouple fluid velocity and director vector computations for every tangent problem. Then, we have designed a linear semi-implicit algorithm (i.e. it does not involve nonlinear iterations) and proved that it is unconditionally stable, verifying a discrete energy inequality. Finally, some numerical simulations are provided.

1. Introduction

Liquid crystals are materials that exhibit phases between those of a liquid and those of a crystal. They are made of macromolecules of similar size, usually represented as rods. The nematic phase is considered the simplest liquid crystal phase where the elongated molecules tend to be locally parallel to some preferential direction. However the molecular centers of gravity are allowed to flow freely as in an isotropic fluid, i.e. without a positional order. This uniaxial orientational order is typically modelled by a unit vector called the director vector $\mathbf{d}$. The first phenomenological theory describing statical configurations of a nematic liquid crystal was proposed by Oseen [27] and Frank [15]. They suggest that the director field corresponds to a minimum of the so-called Oseen–Frank free-energy functional, which in the most elementary form is the Dirichlet energy

$$E(\mathbf{d}) = K \int_{\Omega} |\nabla \mathbf{d}|^2 \, d\mathbf{x}, \tag{1}$$

subject to the sphere condition; K is an elastic constant.

It is known that orientational orders affect all the macroscopic properties of the fluid velocity, introducing an anisotropic stress tensor in the linear momentum equations. The continuum theory of nematic liquid crystals was formulated by Ericksen [11,12] and Leslie [21,20], containing the Oseen–Frank elastic energy.

* Corresponding author.
E-mail addresses: sbadia@cimne.upc.edu (S. Badia), guillen@us.es (F. Guillén-González), juanvi@us.es (J.V. Gutiérrez-Santacreu).

0021-9991/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2010.11.033

Our interest is to construct numerical approximations for the motion of a nematic liquid crystal governed by the simplification of the Ericksen–Leslie equations proposed by Lin in [22]. This problem has numerically been treated using the Ginzburg–Landau penalty problem in order to enforce the sphere constraint. In Section 2 we formulate the problem in the saddle-point framework. Such a formulation allows us to enforce the sphere condition with or without penalty in a single setting. One benefit of this approach is that an energy estimate is obtained for both cases. In Section 3 we present a semi-discrete scheme based on low-order finite elements for approximating all the unknowns. This scheme is unconditionally stable and its solution satisfies a discrete energy estimate. In Section 4, three time-stepping schemes are considered. The first two schemes are nonlinear, with a backward-differencing and mid-point discretization, respectively. Different linearizations for these schemes are studied in Section 5. The third scheme is linear, implicit with respect to the linear term and semi-implicit with respect to the nonlinear term. The three schemes are again unconditionally stable. In Section 6, we test our numerical algorithms with a smooth initial condition and a initial condition with two defect points. Finally, we compare the numerical approximations for a test with analytical solution.

2. Problem statement

A micro–macroscopic continuum theory has been developed for the modeling of nematic liquid crystal flows (see [10]), that characterizes this physical phenomenon in terms of the (microscopic) molecular orientation and the (macroscopic) velocity–pressure variables. The simplified Ericksen–Leslie system consists of a set of partial differential equations that reads as follows: find $\mathbf{d}$, $\mathbf{u}$, and $\tilde{p}$ such that

$$\begin{aligned}
&\partial_t \mathbf{d} + (\mathbf{u}\cdot\nabla)\mathbf{d} - \gamma\Delta\mathbf{d} - \gamma|\nabla\mathbf{d}|^2\mathbf{d} = \mathbf{0},\\
&|\mathbf{d}| = 1,\\
&\partial_t \mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} - \nu\Delta\mathbf{u} + \nabla\tilde{p} + \lambda\nabla\cdot((\nabla\mathbf{d})^t\nabla\mathbf{d}) = \mathbf{g},\\
&\nabla\cdot\mathbf{u} = 0,
\end{aligned} \tag{2}$$

in $(\mathbf{x},t)\in\Omega\times(0,T)$, where $\Omega\subset\mathbb{R}^3$ is the spatial bounded domain filled by the liquid crystal, and $(0,T)$ the time interval. The physical constants are the fluid viscosity $\nu>0$, an elasticity constant $\lambda>0$ and a relaxation time $\gamma>0$. The unknown $\mathbf{d}(\mathbf{x},t)\in\mathbb{R}^3$ is the director vector that determines the orientation of the molecules, $\mathbf{u}(\mathbf{x},t)\in\mathbb{R}^3$ is the velocity of the flow and $\tilde{p}(\mathbf{x},t)\in\mathbb{R}$ is the pressure. The data $\mathbf{g}(\mathbf{x},t)\in\mathbb{R}^3$ is a force term. The gradient operator is defined as $\nabla\mathbf{x}=\partial_j\mathbf{x}_i$ and $(\nabla\mathbf{x})^t$ denotes its transpose. In the following, we will consider that the boundary conditions $\partial_{\mathbf{n}}\mathbf{d}=0$ and $\mathbf{u}=0$ are satisfied a.e. on the boundary $\partial\Omega$ if we do not specify the contrary; $\partial_{\mathbf{n}}\mathbf{d}=\nabla\mathbf{d}\cdot\mathbf{n}$ is the normal derivative where $\mathbf{n}$ is the outward normal vector to the boundary. Initial boundary conditions $\mathbf{d}(\mathbf{x},0)=\mathbf{d}^0$ (with $|\mathbf{d}^0|=1$ a.e. in Ω), and $\mathbf{u}(\mathbf{x},0)=\mathbf{u}^0$ (with $\nabla\cdot\mathbf{u}^0=0$ in Ω and $\mathbf{u}^0=0$ on $\partial\Omega$).

In fact, system (2) is a simplification of the classical Ericksen–Leslie theory of liquid crystals obtained after assuming that some physical elastic constants are equal (see [10]). In general, this assumption is not true, but the mathematical nature of the system does not change, and the complications related to its numerical approach are still present in the simplified problem. For this reason, system (3) has been subject of many mathematical analyses (see [17,25,26,4]).

The saddle-point formulation for (2) consists in finding $\mathbf{d}$, $\mathbf{u}$, q, and $\tilde{p}$ such that

$$\begin{aligned}
&\partial_t \mathbf{d} + (\mathbf{u}\cdot\nabla)\mathbf{d} + \gamma(-\Delta\mathbf{d} + q\,\mathbf{d}) = \mathbf{0},\\
&\mathbf{d}\cdot\mathbf{d} = 1,\\
&\partial_t \mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} - \nu\Delta\mathbf{u} + \lambda\nabla\cdot((\nabla\mathbf{d})^t\nabla\mathbf{d}) + \nabla\tilde{p} = \mathbf{g},\\
&\nabla\cdot\mathbf{u} = 0,
\end{aligned} \tag{3}$$

where $q(\mathbf{x},t)\in\mathbb{R}$ is the Lagrange multiplier used to enforce the restriction $|\mathbf{d}|^2=1$ a.e. in $\Omega\times(0,T)$ ($|\cdot|$ denotes the Euclidean norm). It is interesting to observe that problem (3) has only quadratic nonlinear terms whereas problem (2) involves cubic nonlinear terms.

In top of the open questions related to the Navier–Stokes equations, the nonconvex constraint over $\mathbf{d}$ makes the theoretical analysis of the previous problem very difficult to approach. So, a penalized version is usually considered, in which the constraint $|\mathbf{d}|^2=1$ is weakly enforced by adding the Ginzburg–Landau (GL) function $\mathbf{f}_\epsilon(\mathbf{d})$, for $0<\epsilon\ll 1$, where $\mathbf{f}_\epsilon(\mathbf{d}) := \frac{1}{\epsilon^2}(\mathbf{d}\cdot\mathbf{d}-1)\mathbf{d}$. We will also make use of the potential function $F_\epsilon(\mathbf{d}) := \frac{1}{4\epsilon^2}(\mathbf{d}\cdot\mathbf{d}-1)^2$; note that $\nabla_{\mathbf{d}}F_\epsilon(\mathbf{d})=\mathbf{f}_\epsilon(\mathbf{d})$. The GL penalized problem then reads

$$\begin{aligned}
&\partial_t \mathbf{d} + (\mathbf{u}\cdot\nabla)\mathbf{d} + \gamma(-\Delta\mathbf{d} + \mathbf{f}_\epsilon(\mathbf{d})) = \mathbf{0},\\
&\partial_t \mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} - \nu\Delta\mathbf{u} + \lambda\nabla\cdot((\nabla\mathbf{d})^t\nabla\mathbf{d}) + \nabla\tilde{p} = \mathbf{g},\\
&\nabla\cdot\mathbf{u} = 0,
\end{aligned} \tag{4}$$

supplemented with the respective initial and boundary conditions. In fact, it is straightforward to note that the GL penalized problem (4) can also be casted in a saddle-point form as follows:

$$\begin{aligned}
&\partial_t \mathbf{d} + (\mathbf{u}\cdot\nabla)\mathbf{d} + \gamma(-\Delta\mathbf{d} + q\,\mathbf{d}) = \mathbf{0},\\
&\mathbf{d}\cdot\mathbf{d} - \epsilon^2 q = 1,\\
&\partial_t \mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} - \nu\Delta\mathbf{u} + \lambda\nabla\cdot((\nabla\mathbf{d})^t\nabla\mathbf{d}) + \nabla\tilde{p} = \mathbf{g},\\
&\nabla\cdot\mathbf{u} = 0,
\end{aligned} \tag{5}$$

where $q = \frac{1}{\epsilon^2}(\mathbf{d}\cdot\mathbf{d} - 1)$, hence $q\mathbf{d} = \mathbf{f}_\epsilon(\mathbf{d})$. The clear advantage of (5) with respect to (4) is the fact that it formally allows $\epsilon = 0$, i.e. it includes the limit and penalized problem in a unified formulation. This unified approach also permits to make connections between existing methods that seemed essentially different. Further, the saddle-point approach gives a clue about how to deal with the mathematical analysis of the limit problem. The stability of the multiplier q in the limit case can only be attained via an inf–sup conditions à *la* Babuska–Brezzi, with a formidable complication: the inf–sup condition is nonlinear.

For the subsequent numerical analysis, we will consider a reformulation of the coupling term in the fluid momentum equation. After some manipulation, the coupling term can be re-written as:

$$\nabla\cdot((\nabla\mathbf{d})^t\nabla\mathbf{d}) = \partial_j(\partial_j\mathbf{d}_k\partial_i\mathbf{d}_k) = \partial_j^2\mathbf{d}_k\partial_i\mathbf{d}_k + \frac{1}{2}\partial_i\left(|\partial_j\mathbf{d}_k|^2\right) = (\nabla\mathbf{d})^t\Delta\mathbf{d} + \frac{1}{2}\nabla(|\nabla\mathbf{d}|^2).$$

Using the second equation in (5), we have:

$$q(\nabla\mathbf{d})^t\mathbf{d} = \frac{1}{2}q\nabla(|\mathbf{d}|^2) = \frac{1}{2}q\nabla(\epsilon^2 q) = \frac{\epsilon^2}{4}\nabla q^2.$$

Using the **d**-system in (5), we get:

$$\nabla\cdot((\nabla\mathbf{d})^t\nabla\mathbf{d}) = (\nabla\mathbf{d})^t(\Delta\mathbf{d} - q\,\mathbf{d}) + \frac{1}{2}\nabla\left(|\nabla\mathbf{d}|^2 + \frac{\epsilon^2}{2}q^2\right) = \frac{1}{\gamma}(\nabla\mathbf{d})^t(\partial_t\mathbf{d} + (\mathbf{u}\cdot\nabla)\mathbf{d}) + \frac{1}{2}\nabla\left(|\nabla\mathbf{d}|^2 + \frac{\epsilon^2}{2}q^2\right).$$

Note that for (4) one obtains $q(\nabla\mathbf{d})^t\mathbf{d} = q\frac{1}{2}\nabla(|\mathbf{d}|^2) = \mathbf{0}$. Then, we have the above equality for $\epsilon = 0$. We can absorb the second term in the pressure gradient by using the modified pressure $p = \tilde{p} + \frac{\lambda}{2}|\nabla\mathbf{d}|^2 + \frac{\lambda\epsilon^2}{4}q^2$, leading to the following version of the fluid momentum equation:

$$\partial_t\mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} - \nu\Delta\mathbf{u} + \frac{\lambda}{\gamma}(\nabla\mathbf{d})^t(\partial_t\mathbf{d} + (\mathbf{u}\cdot\nabla)\mathbf{d}) + \nabla p = \mathbf{g}. \tag{6}$$

In fact, these manipulations are not new for the GL penalized problem (see e.g. [24,17,4]). This reformulation will allow us to obtain finite element approximations with an energy estimate that mimics the one at the continuous level.

Since the aim of this work is to consider a Faedo–Galerkin approximation of system (4) or (5) based on finite element spaces, we consider the problem in a weak sense as follows: find $(\mathbf{d}(t), q(t), \mathbf{u}(t), p(t)) \in W^{1,3}(\Omega)\cap L^\infty(\Omega)\times H_\epsilon^{-1}(\Omega)\times H_0^1(\Omega)\times L_0^2(\Omega)$ such that

$$(\partial_t\mathbf{d}, \bar{\mathbf{d}}) + \gamma(\nabla\mathbf{d}, \nabla\bar{\mathbf{d}}) + c(\mathbf{u}, \mathbf{d}, \bar{\mathbf{d}}) + \gamma b_{\mathrm{d}}(q, \mathbf{d}, \bar{\mathbf{d}}) = \mathbf{0}, \tag{7a}$$

$$b_{\mathrm{d}}(\bar{q}, \mathbf{d}, \mathbf{d}) - \epsilon^2(q, \bar{q}) = \langle 1, \bar{q}\rangle, \tag{7b}$$

$$(\partial_t\mathbf{u}, \bar{\mathbf{u}}) + \nu(\nabla\mathbf{u}, \nabla\bar{\mathbf{u}}) + c(\mathbf{u}, \mathbf{u}, \bar{\mathbf{u}}) + \frac{\lambda}{\gamma}c(\bar{\mathbf{u}}, \mathbf{d}, \partial_t\mathbf{d} + (\mathbf{u}\cdot\nabla)\mathbf{d}) - b_{\mathrm{u}}(p, \bar{\mathbf{u}}) = \langle\mathbf{g}, \bar{\mathbf{u}}\rangle, \tag{7c}$$

$$b_{\mathrm{u}}(\bar{p}, \mathbf{u}) = 0 \tag{7d}$$

hold for any $(\bar{\mathbf{d}}, \bar{q}, \bar{\mathbf{u}}, \bar{p}) \in H^1(\Omega)\times H_\epsilon^{-1}(\Omega)\times H_0^1(\Omega)\times L_0^2(\Omega)$ a.e. in $t\in(0,T)$. The regularity for **d** is the one needed in order to have all the terms to be integrated in $L^1(\Omega)$; it can be checked by using Sobolev imbedding theorems and the regularity results in Theorem 2.1. The forms associated to this problem are defined as:

$$b_{\mathrm{d}}(q, \mathbf{d}, \bar{\mathbf{d}}) = (q, \mathbf{d}\cdot\bar{\mathbf{d}}), \quad b_{\mathrm{u}}(p, \bar{\mathbf{u}}) = (p, \nabla\cdot\bar{\mathbf{u}}), \quad c(\mathbf{u}, \mathbf{v}, \bar{\mathbf{u}}) = \langle(\mathbf{u}\cdot\nabla)\mathbf{v}, \bar{\mathbf{u}}\rangle.$$

Note that the elastic tensor effect in the **u**-system is denoted by the same form $c(\cdot,\cdot,\cdot)$. Hereafter, $(\cdot,\cdot)$ denotes the inner product in $L^2(\Omega)$, with $\|\cdot\|$ the associated norm, and $\langle\cdot,\cdot\rangle$ the duality product between $(H^1(\Omega))'$ and $H^1(\Omega)$. Since the Lagrange multiplier q will lose regularity from $\epsilon > 0$ to $\epsilon = 0$, we consider the Banach space $H_\epsilon^{-1}(\Omega)$; for any $\epsilon > 0$ it is the Hilbert space $L^2(\Omega)$ but endowed with the norm $\epsilon\|\cdot\| + \|\cdot\|_{-1}$, whereas for $\epsilon = 0$ it is the dual space of $H^1(\Omega)$.

Theorem 2.1. *The continuous problem (7) with $\epsilon = 0$ satisfies the following energy equality:*

$$\begin{aligned}
&\|\mathbf{u}(t)\|^2 + \lambda\|\nabla\mathbf{d}(t)\|^2 + 2\int_0^t\left(\nu\|\nabla\mathbf{u}(s)\|^2 + \frac{\lambda}{\gamma}\|\partial_t\mathbf{d}(s) + (\mathbf{u}(s)\cdot\nabla)\mathbf{d}(s)\|^2\right)\mathrm{d}s\\
&\quad = \|\mathbf{u}^0\|^2 + \lambda\|\nabla\mathbf{d}^0\|^2 + 2\int_0^t\langle\mathbf{g}(s), \mathbf{u}(s)\rangle\mathrm{d}s,
\end{aligned} \tag{8}$$

that holds for any $t\in[0,T]$. On the other hand, for $\epsilon > 0$, the system satisfies

$$\|\mathbf{u}(t)\|^2+\lambda\|\nabla\mathbf{d}(t)\|^2+2\int_0^t\left(\nu\|\nabla\mathbf{u}(s)\|^2+\frac{\lambda}{\gamma}\|\partial_t\mathbf{d}(s)+(\mathbf{u}(s)\cdot\nabla)\mathbf{d}(s)\|^2\right)\mathrm{d}s+\lambda\frac{\epsilon^2}{2}\|q(t)\|^2$$
$$=\|\mathbf{u}^0\|^2+\lambda\|\nabla\mathbf{d}^0\|^2+\lambda\frac{\epsilon^2}{2}\|q^0\|^2+2\int_0^t\langle\mathbf{g}(s),\mathbf{u}(s)\rangle\mathrm{d}s, \tag{9}$$

for any $t\in[0,T]$, where $q^0=(|\mathbf{d}^0|^2-1)/\epsilon^2$.

Proof. Defining $\mathbf{w}(\mathbf{d},\mathbf{u}):=\partial_t\mathbf{d}+(\mathbf{u}\cdot\nabla)\mathbf{d}$ the following identity holds:

$$\frac{\lambda}{\gamma}c(\mathbf{u},\mathbf{d},\mathbf{w})=\frac{\lambda}{\gamma}\|\mathbf{w}\|^2-\frac{\lambda}{\gamma}(\partial_t\mathbf{d},\mathbf{w}). \tag{10}$$

On the other hand, we can re-write Eq. (7a) as:

$$(\mathbf{w},\bar{\mathbf{d}})+\gamma(\nabla\mathbf{d},\nabla\bar{\mathbf{d}})+\gamma b_{\mathrm{d}}(q,\mathbf{d},\bar{\mathbf{d}})=0.$$

Taking $\bar{\mathbf{d}}=\partial_t\mathbf{d}$, we easily get:

$$-(\mathbf{w},\partial_t\mathbf{d})=\frac{\gamma}{2}\partial_t\|\nabla\mathbf{d}\|^2+\gamma\, b_{\mathrm{d}}(q,\mathbf{d},\partial_t\mathbf{d}). \tag{11}$$

The time derivative of Eq. (7b) leads to:

$$2\,b_d(\bar{q},\partial_t\mathbf{d},\mathbf{d})-\epsilon^2(\partial_t q,\bar{q})=0,$$

which for $\bar{q}=q$ allows to write the last term of (11) as follows:

$$b_{\mathrm{d}}(q,\mathbf{d},\partial_t\mathbf{d})=\frac{\epsilon^2}{2}(q,\partial_t q)=\frac{\epsilon^2}{4}\partial_t\|q\|^2.$$

Hence, we finally have that (10) is expressed as:

$$\frac{\lambda}{\gamma}c(\mathbf{u},\mathbf{d},\mathbf{w})=\frac{\lambda}{\gamma}\|\mathbf{w}\|^2+\frac{\lambda}{2}\partial_t\|\nabla\mathbf{d}\|^2+\frac{\lambda\epsilon^2}{4}\partial_t\|q\|^2.$$

The desired energy equality is obtained testing (7c) and (7d) against $(\mathbf{u},p)$, using the previous equality and the skew-symmetry property $c(\mathbf{u},\bar{\mathbf{u}},\bar{\mathbf{u}})=0$ for any $\bar{\mathbf{u}}\in H_0^1(\Omega)$. We get:

$$\partial_t\left\{\|\mathbf{u}\|^2+\lambda\|\nabla\mathbf{d}\|^2+\frac{\lambda\epsilon^2}{2}\|q\|^2\right\}+2\left(\nu\|\nabla\mathbf{u}\|^2+\frac{\lambda}{\gamma}\|\mathbf{w}\|^2\right)=2\langle\mathbf{g},\mathbf{u}\rangle,$$

a.e. in $(0,T)$. For $\epsilon=0$, the energy equality (8) is obtained after integrating the previous equation in the time interval $(0,t)$. The energy equality for the penalized problem (9) is proved by noting that Eq. (7b) holds at $t=0$ with $q^0=(|\mathbf{d}^0|^2-1)/\epsilon^2$ for $\epsilon>0$. □

Pressure stability relies on the well-known inf–sup condition:

$$\inf_{\bar{p}\in L_0^2(\Omega)}\sup_{\bar{\mathbf{u}}\in H^1(\Omega)}\frac{b_{\mathrm{u}}(\bar{p},\nabla\cdot\bar{\mathbf{u}})}{\|\bar{p}\|\|\bar{\mathbf{u}}\|_1}\geqslant\beta_{\mathrm{u}}>0,$$

which is known to be true due to the surjectivity of the divergence operator from $H_0^1(\Omega)$ onto $L_0^2(\Omega)$; $L_0^2(\Omega)$ is the space of $L^2(\Omega)$ functions with zero mean value. We refer to [31,18] for some regularity results for the pressure in the transient Navier–Stokes system. The control over the Lagrange multiplier q is not so well understood. The GL penalty version introduces $L^2(\Omega)$ control over q that is lost when the penalty $\epsilon\searrow 0$. So, this stability does not apply for the singular limit $\epsilon=0$, and the well-posedness of the original problem can only rely on an inf–sup condition. The inf–sup condition

$$\inf_{\bar{q}\in H^{-1}(\Omega)}\sup_{\bar{\mathbf{d}}\in H_0^1(\Omega)}\frac{b_{\mathrm{d}}(\bar{q},\mathbf{d},\bar{\mathbf{d}})}{\|\bar{q}\|_{-1}\|\bar{\mathbf{d}}\|_1}\geqslant\beta_{\mathrm{d}}(\mathbf{d})>0, \tag{12}$$

has been proved very recently in [19] under some regularity assumptions over $\mathbf{d}$ in the frame of the steady harmonic maps problem, which is (7a) and (7b) without the time derivative and convective terms. This regularity is much stronger than the one that the energy estimate (8) provides for $\mathbf{d}$.

With regard to the long-term behavior for a zero forcing term ($\mathbf{g}=0$), we can see that for $t\to\infty$ the energy dissipation in (8) goes to zero:

$$\nu\|\nabla\mathbf{u}(t)\|^2+\frac{\lambda}{\gamma}\|\partial_t\mathbf{d}(t)+(\mathbf{u}(t)\cdot\nabla)\mathbf{d}(t)\|^2\to 0,$$

hence $\mathbf{u}$ goes to the trivial stationary point $\mathbf{u}=0$ and $\|\partial_t\mathbf{d}\|\to 0$. The $\mathbf{d}$ component of the solution exhibits non-trivial stationary states. Such a stationary states are minima of the Oseen–Frank free-energy function (1). It means that there exist steady

solutions with $\|\nabla \mathbf{d}\| > 0$. This fact is shared by both the limit and penalized case. For the penalized case, there is an additional term in the energy, which is the penalty energy $\frac{\lambda \epsilon^2}{2}\|q\|^2$.

3. Spatial discretization

Let $\mathcal{T}_h$ be a partition of Ω into a set of finite elements $\{K\}$. For every element K, we denote by h_K its diameter, and set the characteristic mesh size as $h = \max_{K\in\mathcal{T}_h} h_K$. The space of polynomials of degree less or equal to $k > 0$ in a finite element K is denoted by $\mathcal{P}^k(K)$. The space of continuous piecewise polynomials is defined as:

$$\mathcal{P}_h^k = \left\{ v_h \in \mathcal{C}^0(\Omega) \text{ such that } v_h|_K \in \mathcal{P}^k(K)\ \forall K \in \mathcal{T}_h \right\}. \tag{13}$$

These approximations are usually called H^1-conforming approximations, because of the inter-element continuity. The space $\mathcal{P}_h^k$ is spanned by the set of nodal functions $\{\pi_h^a\}_{a\in\mathcal{N}_h}$, where $\mathcal{N}_h$ is the set of nodes in the mesh.

Therefore, any function $\varphi_h \in \mathcal{P}_h^k$ can be uniquely determined in terms of its nodal values $\{\varphi^a\}_{a\in\mathcal{N}_h}$ as $\sum_{a\in\mathcal{N}_h}\varphi^a\pi_h^a$ (see [5,13]). The nodal interpolation of a continuous function $\varphi \in \mathcal{C}^0(\overline{\Omega})$ is denoted by $\pi_h(\varphi) = \sum_{a\in\mathcal{N}_h}\varphi(\mathbf{x}^a)\pi_h^a$.

Let us consider a conforming finite element discretization of problem (7). The finite element space D_h for the director vector $\mathbf{d}_h$ is chosen to be $(\mathcal{P}_h^1)^d$. We also consider the space Q_h for the Lagrange multiplier q_h to be $\mathcal{P}_h^1$. The constraint form reads as:

$$b_{\mathrm{d}}(q_h, \mathbf{d}_h, \bar{\mathbf{d}}_h) = (q_h, \mathbf{d}_h \cdot \bar{\mathbf{d}}_h). \tag{14}$$

In this case, the constraint is satisfied in a discrete sense, as the incompressibility condition for the fluid problem. In the frame of harmonic maps, Hu et al. have considered the following modification of the constraint form in [19]:

$$b_{\mathrm{d}}(q_h, \mathbf{d}_h, \bar{\mathbf{d}}_h)_h = (q_h, \pi_h(\mathbf{d}_h \cdot \bar{\mathbf{d}}_h)). \tag{15}$$

Since $\pi_h(\cdot) \in Q_h$ with the previous choice of finite element spaces, the constraint equation amounts to saying that $|\mathbf{d}^a| = 1$ for any $a \in \mathcal{N}_h$. Furthermore, the finite element pair $D_h \times Q_h$ has been recently proved to satisfy the corresponding discrete version of the inf–sup condition (12):

$$\inf_{q_h\in Q_h}\sup_{\bar{\mathbf{d}}_h\in D_h}\frac{b_{\mathrm{d}}(q_h, \mathbf{d}_h, \bar{\mathbf{d}}_h)_h}{\|q_h\|_{-1}\|\bar{\mathbf{d}}_h\|_1} \geqslant \beta_{h,\mathrm{d}}(\mathbf{d}_h) > 0, \tag{16}$$

in [19]. It has allowed to prove the well-posedness of the tangent problem that arises from the full Newton linearization of the steady-state harmonic maps problem, in the vicinity of a local minimum under strong regularity assumptions. In this proof, the fact that the projection $\pi_h(\cdot)$ has been used is necessary, so it does not apply to $b_{\mathrm{d}}(q_h, \mathbf{d}_h, \bar{\mathbf{d}}_h)$. In any case, numerical experimentation says that this choice is also stable for the problems considered in Section 6.

Remark 3.1. Let us recall that, for Dirichlet boundary conditions over $\mathbf{d}_h$, the discrete inf–sup condition (16) is violated when there are elements $K \in \mathcal{T}_h$ with all the nodes constrained; we can easily check that there is no control of the Lagrange multiplier at a boundary node that is connected to boundary nodes only. This type of meshes usually leads to problems and should be avoided. In any case, an alternative way to circumvent this problem is to consider homogeneous Dirichlet boundary conditions over q too, as proposed in [19]. Since $|\mathbf{d}^a| = 1$ for appropriate boundary data, $q^a = 0$ is an appropriate condition.

In [19], the saddle-point version of the harmonic maps problem is used together with a Newton linearization, and the corresponding inf–sup condition for the tangent problem is proved. As an alternative, it could be used for the minimization step of the linearized problem in the popular Alouges' method proposed in [1] (see also [3,2] and Section 5.3). In the frame of liquid crystals, we consider two different solvers, that are extensions of these two approaches to the problem at hand.

Let us point out that the $\mathbf{d}_h - q_h$ block matrices in the corresponding linear system are diagonal matrices when using closed (nodal) integration, and so computationally more efficient; closed (nodal) integration of the constraint terms is the way to get a lumped mass matrix for Lagrangian elements. The closed integration of the constraint trilinear form consists of:

$$b_{\mathrm{d}}(q_h, \mathbf{d}_h, \bar{\mathbf{d}}_h)_{hs} = \sum_{a\in\mathcal{N}_h} q^a\mathbf{d}^a \cdot \bar{\mathbf{d}}^a \int_\Omega \pi_h^a \,\mathrm{d}\mathbf{x} = (q_h, \mathbf{d}_h \cdot \bar{\mathbf{d}}_h)_s, \tag{17}$$

where the sub-label s indicates lumped sub-integration of the term. The nodal enforcement of the restriction is even more explicit in this case. Here and in the sequel, we will also use the sub-index s for inner-products $(f,g)_s = \int_\Omega \pi_h(fg)\,\mathrm{d}\mathbf{x}$ and L^2-discrete norm $\|f\|_s = (f,f)_s^{1/2}$ involving finite element functions to denote that closed integration is used.

Remark 3.2. The nodal enforcement of the constraint could also be understood as a collocation method for the constraint equation. In this case, the discrete version of the Lagrange multiplier space consists of $Q_h = \{\delta(\mathbf{x}_a), a \in \mathcal{N}\}$, where $\delta(\mathbf{x}_a) : \mathcal{C}^0(\Omega) \to \mathbb{R}$ is defined by $\langle\delta(\mathbf{x}_a), v\rangle = v(\mathbf{x}_a)$ for $v \in \mathcal{C}^0(\Omega)$. This approach to the problem is not so powerful, because it can only be used for the limit case, since the penalty term is ill-posed.

For the Navier–Stokes sub-problem we consider the standard MINI element, in which the pressure finite element space P_h is taken as $\mathcal{P}^1_h$, and the velocity space V_h is $(\mathcal{P}^1_h)^d \oplus (\mathcal{B}_h)^d$, where

$$\mathcal{B}_h = \{ v_b \text{ such that } v_b|_K \in \mathcal{P}^3(K),\ v_b|_{\partial K} = 0,\ v_b|_K \geqslant 0,\ \forall K \in \mathcal{T}_h \}$$

is the space of bubbles (cubic in dimension 2) at every element (see e.g. [7]). This velocity–pressure finite element pair is known to satisfy the discrete inf–sup condition

$$\inf_{\bar{p}_h \in Q_h} \sup_{\bar{\mathbf{u}}_h \in V_h} \frac{b_{\mathrm{u}}(\bar{p}_h, \bar{\mathbf{u}}_h)}{\|\bar{p}_h\| \|\bar{\mathbf{u}}_h\|_1} \geqslant \beta_{\mathrm{u},h} \geqslant 0$$

for $\beta_{\mathrm{u},h}$ uniform with respect to h. Onwards, we consider the skew-symmetric form (for Dirichlet boundary conditions) of the convective term in the $\mathbf{u}_h$ problem:

$$\tilde{c}(\mathbf{u}_h, \mathbf{v}_h, \bar{\mathbf{u}}_h) = ((\mathbf{u}_h \cdot \nabla)\mathbf{v}_h, \bar{\mathbf{u}}_h) + \frac{1}{2}((\nabla \cdot \mathbf{u}_h)\mathbf{v}_h, \bar{\mathbf{u}}_h).$$

The finite element approximation of system (7) reads as: find $(\mathbf{d}_h(t), q_h(t), \mathbf{u}_h(t), p_h(t)) \in D_h \times Q_h \times V_h \times P_h$ such that

$$(\partial_t \mathbf{d}_h, \bar{\mathbf{d}}_h) + \gamma(\nabla \mathbf{d}_h, \nabla \bar{\mathbf{d}}_h) + c(\mathbf{u}_h, \mathbf{d}_h, \bar{\mathbf{d}}_h) + \gamma b_{\mathrm{d}}(q_h, \mathbf{d}_h, \bar{\mathbf{d}}_h)_{\delta\rho} = 0, \tag{18a}$$

$$b_{\mathrm{d}}(\bar{q}_h, \mathbf{d}_h, \mathbf{d}_h)_{\delta\rho} - \epsilon^2 (q_h, \bar{q}_h)_\rho = (1, \bar{q}_h)_\rho, \tag{18b}$$

$$(\partial_t \mathbf{u}_h, \bar{\mathbf{u}}_h) + \nu(\nabla \mathbf{u}_h, \nabla \bar{\mathbf{u}}_h) + \tilde{c}(\mathbf{u}_h, \mathbf{u}_h, \bar{\mathbf{u}}_h) + \frac{\lambda}{\gamma} c(\bar{\mathbf{u}}_h, \mathbf{d}_h, \partial_t \mathbf{d}_h + (\mathbf{u}_h \cdot \nabla)\mathbf{d}_h) - b_{\mathrm{u}}(p_h, \bar{\mathbf{u}}_h) = \langle \mathbf{g} \bar{\mathbf{u}}_h \rangle, \tag{18c}$$

$$b_{\mathrm{u}}(\bar{p}_h, \mathbf{u}_h) = 0, \tag{18d}$$

at almost every $t \in (0, T]$, for any $(\bar{\mathbf{d}}_h, \bar{q}_h, \bar{\mathbf{u}}_h, \bar{p}_h) \in D_h \times Q_h \times V_h \times P_h$. The sub-label δ takes the values h when using $\pi_h(\cdot)$ in b_{d}; ρ takes the value s when sub-integration is used.

Remark 3.3. Many existing liquid crystal finite element approximations involve an auxiliary variable (see e.g. [17,4]). We can re-formulate the problem, by introducing an auxiliary variable $\mathbf{w}_h$ and its corresponding finite element space W_h; Eq. (18a) is replaced by:

$$(\mathbf{w}_h, \bar{\mathbf{d}}_h) + \gamma(\nabla \mathbf{d}_h, \nabla \bar{\mathbf{d}}_h) + \gamma\, b_{\mathrm{d}}(q_h, \mathbf{d}_h, \bar{\mathbf{d}}_h)_{\delta\rho} = 0,$$
$$(\partial_t \mathbf{d}_h, \bar{\mathbf{w}}_h) + c(\mathbf{u}_h, \mathbf{d}_h, \bar{\mathbf{w}}_h) - (\mathbf{w}_h, \bar{\mathbf{w}}_h) = 0$$

for any $(\bar{\mathbf{d}}_h, \bar{\mathbf{w}}_h) \in D_h \times W_h$. Then, the elastic stress term $c(\bar{\mathbf{u}}_h, \mathbf{d}_h, \partial_t \mathbf{d}_h + (\mathbf{u}_h \cdot \nabla)\mathbf{d}_h)$ is replaced by:

$$\int_\Omega (\nabla \mathbf{d}_h)^t \mathbf{w}_h \cdot \bar{\mathbf{u}}_h \mathrm{d}\mathbf{x}.$$

This approach would in principle introduce extra unknowns to the problem, which would preferably be avoided. In most existing liquid crystal algorithms, always for the penalized GL problem, the space for W_h is taken equal to D_h, i.e. $(\mathcal{P}^1_h)^d$ (see [17,4]). The method is proved to be stable and convergent but, since $(\mathbf{u}_h \cdot \nabla)\mathbf{d}_h \notin D_h$, the problem for $\mathbf{w}_h$ is global and $\mathbf{w}_h$ cannot be locally eliminated. However, taking $W_h \equiv (V_h \cdot \nabla)D_h$ and $D_h \subset W_h$, we simply have $\mathbf{w}_h = \partial_t \mathbf{d}_h + (\mathbf{u}_h \cdot \nabla)\mathbf{d}_h$. Implicitly, this is the approach that has been used above and the one used in [23] for the approximation of the penalized GL method.

The saddle-point approach has a clear advantage with respect to the previous finite element approximations of liquid crystal flows. In the frame proposed herein, we can approximate numerically both the limit problem and GL penalized versions using the same numerical approximation. Existing algorithms [23,25,26,4,17] could not take ϵ arbitrarily small, since the condition number of the matrix blows up with ϵ^{-2}. As a rule of thumb, $\epsilon = 0.05$ was the limit value used in numerical experiments. Since the stability of the saddle-point structure as $\epsilon \searrow 0$ is kept by virtue of the discrete inf–sup condition (16), the linear system is non-singular even for the limit problem. In the numerical experiments section we analyze all these properties, identifying an interesting asymptotic behavior as $\epsilon \searrow 0$ that has not been observed yet. In any case, it is interesting to relate our approach with the existing GL penalized techniques, i.e. for $\epsilon > 0$. From (18b) with the constraint as in (14), the penalty function takes the value

$$q_h = \frac{1}{\epsilon^2} P_{Q_h}(\mathbf{d}_h \cdot \mathbf{d}_h - 1).$$

So, $q_h \mathbf{d}_h$ acts as $\mathbf{f}^h_\epsilon(\mathbf{d}_h) = \frac{1}{\epsilon^2} P_{Q_h}(\mathbf{d}_h \cdot \mathbf{d}_h - 1)\mathbf{d}_h$ in the penalized finite element formulations. Taking a Lagrange multiplier space such that $D_h \cdot D_h \subset Q_h$, the penalized saddle-point problem can be written in the frame of the GL approximation by taking $\mathbf{f}_\epsilon(\mathbf{d}_h) := \frac{1}{\epsilon^2}(\mathbf{d}_h \cdot \mathbf{d}_h - 1)\mathbf{d}_h$. We can easily check that it coincides with the choice of $\mathbf{f}_\epsilon(\mathbf{d}_h)$ in [17,25,26]. For first order finite element approximations of the director field, we can easily prove that $\mathbf{f}_\epsilon(\mathbf{d}_h) = 0$ is only possible for $\mathbf{d}_h$ a constant function. So, these schemes exhibit a locking phenomenon as $\epsilon \searrow 0$. This choice is not appropriate for the limit problem.

When using the constraint form as in (15) with Q_h as $\mathcal{P}_h^1$, we have:

$$q_h = \frac{1}{\epsilon^2} P_{Q_h}(\pi_h(\mathbf{d}_h \cdot \mathbf{d}_h - 1)) = \frac{1}{\epsilon^2}(\pi_h(\mathbf{d}_h \cdot \mathbf{d}_h - 1)),$$

since $P_{Q_h}(\pi_h(\cdot)) = \pi_h(\cdot)$. It is also interesting to note that, when considering b_{d} as in (17), the evaluation of q_h is local at every node of the mesh, with the expression:

$$q_h^a = \frac{1}{\epsilon^2}(\mathbf{d}_h^a \cdot \mathbf{d}_h^a - 1) \quad \forall a \in \mathcal{N}_h.$$

Using this term in $b_{\mathrm{d}}(q_h, \mathbf{d}_h, \bar{\mathbf{d}}_h)_{hs}$, we obtain:

$$b_{\mathrm{d}}(q_h, \mathbf{d}_h, \bar{\mathbf{d}}_h)_{hs} = \sum_{a \in \mathcal{N}_h} \frac{1}{\epsilon^2}(\mathbf{d}_h^a \cdot \mathbf{d}_h^a - 1)\mathbf{d}_h^a \cdot \bar{\mathbf{d}}_h^a \int_\Omega \pi_h^a \mathrm{d}\mathbf{x}.$$

This penalization term is in fact the one used in [4] for the GL term. The method proposed by Walkington and Liu in [25,26] considered $\mathcal{C}^1$ Hermite polynomial approximations for which the lumping technique cannot be used.

4. Time discretization

Let us consider a uniform partition of the time interval $[0,T]$ into N elements (t^n, t^{n+1}) for $n = 0,\ldots,N-1$, where $t^n := n\,k$. The element size is denoted by $k := \frac{T}{N}$. The mid-point value is written as $f^{n+\frac{1}{2}} := \frac{f^{n+1}+f^n}{2}$. We also denote $\frac{f^{n+1}-f^n}{k}$ as $\delta_t f^{n+1}$. Since the forcing term $\mathbf{g}(t)$ does not have pointwise sense in time, we define $\mathbf{g}^{n+1} := \frac{1}{k}\int_{t^n}^{t^{n+1}} \mathbf{g}(s)\,\mathrm{d}s$ and $\mathbf{g}^{n+\frac{1}{2}} := \mathbf{g}^{n+1}$.

We will design both implicit and semi-implicit schemes that satisfy a discrete version of the energy equality (8) for $\epsilon = 0$ or (9) for $\epsilon > 0$. Both the implicit and semi-implicit scheme are unconditionally stable. Further, the semi-implicit scheme is linear. As far as we know, this is the first linear scheme that exhibits unconditional stability.

4.1. Implicit algorithm

The most straightforward approximation of the problem at hand consists of a Backward-Euler first order time integration.

In this case, given $\mathbf{d}_h^n \in D_h$ and $\mathbf{u}_h^n \in V_h$, the problem at the time step t^{n+1} reads as: find $(\mathbf{d}_h^{n+1}, q_h^{n+1}, \mathbf{u}_h^{n+1}, p_h^{n+1}) \in D_h \times Q_h \times V_h \times P_h$ such that

$$\left(\delta_t \mathbf{d}_h^{n+1}, \bar{\mathbf{d}}_h\right) + \gamma\left(\nabla \mathbf{d}_h^{n+1}, \nabla \bar{\mathbf{d}}_h\right) + c\left(\mathbf{u}_h^{n+1}, \mathbf{d}_h^{n+1}, \bar{\mathbf{d}}_h\right) + \gamma b_{\mathrm{d}}\left(q_h^{n+1}, \mathbf{d}_h^{n+1}, \bar{\mathbf{d}}_h\right)_{\delta\rho} = 0, \tag{19a}$$

$$b_{\mathrm{d}}\left(\bar{q}_h, \mathbf{d}_h^{n+1}, \mathbf{d}_h^{n+1}\right)_{\delta\rho} - \epsilon^2\left(q_h^{n+1}, \bar{q}_h\right)_\rho = (1, \bar{q}_h)_\rho, \tag{19b}$$

$$\begin{aligned}&\left(\delta_t \mathbf{u}_h^{n+1}, \bar{\mathbf{u}}_h\right) + \nu\left(\nabla \mathbf{u}_h^{n+1}, \nabla \bar{\mathbf{u}}_h\right) + \tilde{c}\left(\mathbf{u}_h^{n+1}, \mathbf{u}_h^{n+1}, \bar{\mathbf{u}}_h\right) + \frac{\lambda}{\gamma} c\left(\bar{\mathbf{u}}_h, \mathbf{d}_h^{n+1}, \delta_t \mathbf{d}_h^{n+1} + \left(\mathbf{u}_h^{n+1} \cdot \nabla\right)\mathbf{d}_h^{n+1}\right)\\ &\quad - b_{\mathrm{u}}\left(p_h^{n+1}, \bar{\mathbf{u}}_h\right) = \left\langle \mathbf{g}^{n+1}, \bar{\mathbf{u}}_h \right\rangle,\end{aligned} \tag{19c}$$

$$b_{\mathrm{u}}\left(\bar{p}_h, \mathbf{u}_h^{n+1}\right) = 0 \tag{19d}$$

for any $(\bar{\mathbf{d}}_h^{n+1}, \bar{q}_h^{n+1}, \bar{\mathbf{u}}_h^{n+1}, \bar{p}_h^{n+1}) \in D_h \times Q_h \times V_h \times P_h$. However, this first order approximation is only conditionally stable. The proof of an energy equality for this fully discrete system follows the line of the one for the continuum problem. In order to prove $\mathbf{d}_h$-stability, we have to test the equation for the director field (19a) against $\delta_t \mathbf{d}_h^{n+1}$, appearing the term

$$b_{\mathrm{d}}\left(q_h^{n+1}, \mathbf{d}_h^{n+1}, \delta_t \mathbf{d}_h^{n+1}\right)_{\delta\rho} = \frac{1}{2k}\left(q_h^{n+1}, \left|\mathbf{d}_h^{n+1}\right|^2 - \left|\mathbf{d}_h^n\right|^2 + \left|\mathbf{d}_h^{n+1} - \mathbf{d}_h^n\right|^2\right)_\rho,$$

where $\frac{1}{2k}\left(q_h^{n+1}, \left|\mathbf{d}_h^{n+1} - \mathbf{d}_h^n\right|^2\right)_\rho$ cannot be controlled. So, a straightforward first order approximation of the problem at hand is not appropriate. In any case, in the numerical experiments we have performed, this instability has not been activated. One way to circumvent that problem is to replace Eq. (19b) by the discrete time derivative $b_d\left(\bar{q}_h, \mathbf{d}_h^{n+1}, \partial_t \mathbf{d}_h^{n+1}\right)_{\delta\rho} = \frac{\epsilon^2}{2}\left(\delta_t q_h^{n+1}, \bar{q}_h\right)_\rho$; see Section 4.2 for more details and a semi-implicit version of (19).

Alternatively, in order to get an unconditionally stable algorithm, we have considered a Crank–Nicolson time integration scheme. Analogously, given $\mathbf{d}_h^n \in D_h$ and $\mathbf{u}_h^n \in V_h$, the problem at the time step t^{n+1}reads as: find $(\mathbf{d}_h^{n+1}, q_h^{n+1}, \mathbf{u}_h^{n+1}, p_h^{n+1}) \in D_h \times Q_h \times V_h \times P_h$ such that

$$\left(\delta_t \mathbf{d}_h^{n+1}, \bar{\mathbf{d}}_h\right) + \gamma\left(\nabla \mathbf{d}_h^{n+\frac{1}{2}}, \nabla \bar{\mathbf{d}}_h\right) + c\left(\mathbf{u}_h^{n+\frac{1}{2}}, \mathbf{d}_h^{n+\frac{1}{2}}, \bar{\mathbf{d}}_h\right) + \gamma b_{\mathrm{d}}\left(q_h^{n+\frac{1}{2}}, \mathbf{d}_h^{n+\frac{1}{2}}, \bar{\mathbf{d}}_h\right)_{\delta\rho} = 0, \tag{20a}$$

$$b_{\mathrm{d}}\left(\bar{q}_h, \mathbf{d}_h^{n+1}, \mathbf{d}_h^{n+1}\right)_{\delta\rho} - \epsilon^2\left(q_h^{n+1}, \bar{q}_h\right)_\rho = \langle 1, \bar{q}_h\rangle_\rho, \tag{20b}$$

$$\begin{aligned}&\left(\delta_t \mathbf{u}_h^{n+1}, \bar{\mathbf{u}}_h\right) + \nu\left(\nabla \mathbf{u}_h^{n+\frac{1}{2}}, \nabla \bar{\mathbf{u}}_h\right) + \tilde{c}\left(\mathbf{u}_h^{n+\frac{1}{2}}, \mathbf{u}_h^{n+\frac{1}{2}}, \bar{\mathbf{u}}_h\right) - b_{\mathrm{u}}\left(p_h^{n+\frac{1}{2}}, \bar{\mathbf{u}}_h\right)\\ &\quad + \frac{\lambda}{\gamma} c\left(\bar{\mathbf{u}}_h, \mathbf{d}_h^{n+\frac{1}{2}}, \delta_t \mathbf{d}_h^{n+1} + \left(\mathbf{u}_h^{n+\frac{1}{2}} \cdot \nabla\right)\mathbf{d}^{n+\frac{1}{2}}\right) = \left\langle \mathbf{g}^{n+\frac{1}{2}}, \bar{\mathbf{u}}_h\right\rangle,\end{aligned} \tag{20c}$$

$$b_{\mathrm{u}}\left(\bar{p}_h, \mathbf{u}_h^{n+\frac{1}{2}}\right) = 0 \tag{20d}$$

for any $(\bar{\mathbf{d}}_h, \bar{q}_h, \bar{\mathbf{u}}_h, \bar{p}_h) \in D_h \times Q_h \times V_h \times P_h$. We can check that the restriction over $\mathbf{d}_h$ (20b) has been enforced at the time nodes, instead of mid-points. The reason is that, in its present form, the method is unconditionally stable. This system satisfies the semi-discrete version in time of the energy equality in Theorem 2.1, and so it is energy preserving for the limit case with $\epsilon = 0$. It is interesting to note the effect of the initial condition $\mathbf{d}_h^0$ in the equality.

As we will infer from the next theorem, we must take care of the choice for $\mathbf{d}_h^0$. An initial condition that does not satisfy the discrete constraint (20b) introduces an initial energy that blows up as $\epsilon \searrow 0$. The use of this kind of initial approximation is ill-posed for the limit problem.[1] On the other hand, this initial condition has been used as a test problem in some numerical articles based on GL penalty problem approximations (see [23,25,26,4,17]). We refer to the numerical experimentation section for more details.

In the limit case, for a $\mathbf{d}^0$ with $|\mathbf{d}^0| = 1$ a.e. in Ω, a typical finite element projection, e.g. Scott–Zhang or Clement projector, will fail to satisfy the discrete constraint (20b). As an alternative when the constraint form (14) is used, we can consider the following projector for $\mathbf{d}^0$: given $\mathbf{d}^0$, find $\mathbf{d}_h^0 \in D_h$ and $q_h^0 \in Q_h$ such that

$$\gamma\left(\nabla \mathbf{d}_h^0, \nabla \bar{\mathbf{d}}_h\right) + \gamma b_{\mathrm{d}}\left(q_h^0, \mathbf{d}_h^0, \bar{\mathbf{d}}_h\right) = \gamma(\nabla \pi_h(\mathbf{d}^0), \nabla \bar{\mathbf{d}}_h), \quad \forall \bar{\mathbf{d}}_h \in D_h, \tag{21a}$$

$$b_{\mathrm{d}}\left(\bar{q}_h, \mathbf{d}_h^0, \mathbf{d}_h^0\right) = (1, \bar{q}_h), \quad \forall \bar{q}_h \in Q_h. \tag{21b}$$

Assuming that $\mathbf{d}^0 \in H^2(\Omega) \cap W^{1,\infty}(\Omega)$, it has been proved in [19, Theorem 5.3] that this harmonic maps problem admits a unique solution $\mathbf{d}_h^0$ such that $\left\|\mathbf{d}_h^0 - \mathbf{d}^0\right\|_1 \leqslant Ch$.

The use of a constraint-preserving Riesz projector is basic, in order to get an admissible initial condition $\mathbf{d}_h^0$ for the limit problem. In case of using the nodal enforcement, i.e. the constraint form (15) or (17), a more straightforward approach consists on projecting $\mathbf{d}^0$ with a typical finite element projector $P_h\mathbf{d}^0$ and *a posteriori* normalize the value at each node $\left(\mathbf{d}_h^0\right)^a = (P_h\mathbf{d}^0)^a/|(P_h\mathbf{d}^0)^a|$.

In the next theorem, we prove that the scheme (20) satisfies a discrete counterpart of the energy equalities given in Theorem 2.1. As a result, the method is unconditionally stable.

Theorem 4.1. *The discrete solution of system (20) for $\epsilon = 0$ satisfies the following energy equality:*

$$\begin{aligned}&\left\|\mathbf{u}_h^n\right\|^2 + \lambda\left\|\nabla \mathbf{d}_h^n\right\|^2 + 2\nu k\sum_{m=0}^{n-1}\left\|\nabla \mathbf{u}_h^{m+\frac{1}{2}}\right\|^2 + 2\frac{\lambda}{\gamma}k\sum_{m=0}^{n-1}\left\|\delta_t \mathbf{d}_h^{m+1} + \left(\mathbf{u}_h^{m+\frac{1}{2}} \cdot \nabla\right)\mathbf{d}_h^{m+\frac{1}{2}}\right\|^2\\ &\quad 2k\sum_{m=0}^{n-1}\left\langle \mathbf{g}^{m+\frac{1}{2}}, \mathbf{u}_h^{m+\frac{1}{2}}\right\rangle + \left\|\mathbf{u}_h^0\right\|^2 + \lambda\left\|\nabla \mathbf{d}_h^0\right\|^2\end{aligned} \tag{22}$$

for any $n \in \{0, 1, \ldots, N\}$, where the initial condition $\mathbf{d}_h^0$ satisfies the discrete constraint (20b). On the other hand, the penalized version of (20), for $\epsilon > 0$, satisfies:

$$\begin{aligned}&\left\|\mathbf{u}_h^n\right\|^2 + \lambda\left\|\nabla \mathbf{d}_h^n\right\|^2 + \frac{\lambda\epsilon^2}{2}\left\|q_h^n\right\|_\rho^2 + 2k\sum_{m=0}^{n-1}\nu\left\|\nabla \mathbf{u}_h^{m+\frac{1}{2}}\right\|^2 + 2k\sum_{m=0}^{n-1}\frac{\lambda}{\gamma}\left\|\delta_t \mathbf{d}_h^{m+1} + \left(\mathbf{u}_h^{m+\frac{1}{2}} \cdot \nabla\right)\mathbf{d}_h^{m+\frac{1}{2}}\right\|^2\\ &\quad = 2k\sum_{m=0}^{n-1}\left\langle \mathbf{g}^{m+\frac{1}{2}}, \mathbf{u}_h^{m+\frac{1}{2}}\right\rangle + \left\|\mathbf{u}_h^0\right\|^2 + \lambda\left\|\nabla \mathbf{d}_h^0\right\|^2 + \frac{\lambda\epsilon^2}{2}\left\|q_h^0\right\|_\rho^2,\end{aligned} \tag{23}$$

where q_h^0 is defined below in (26). In fact, $q_h^0 = \frac{1}{\epsilon^2}P_{Q_h}\left(\left|\mathbf{d}_h^0\right|^2 - 1\right)$ for the constraint (14) and $q_h^0 = \frac{1}{\epsilon^2}\pi_h\left(\left|\mathbf{d}_h^0\right|^2 - 1\right)$ for the constraint form (15). In case of using (17), $(q_h^0)^a = \frac{1}{\epsilon^2}\left(\left|\left(\mathbf{d}_h^0\right)^a\right|^2 - 1\right)$ and the L^2 norm for q_h is replaced by the lumped one.

Proof. The proof of this result follows that of Theorem 2.1. Let us start re-writing the $\mathbf{d}_h$ constraint Eq. (20b) in an incremental form. Using the fact that $\delta_t(f^{n+1})^2 = 2f^{n+\frac{1}{2}}\delta_t f^{n+1}$, we have that:

[1] Let us point out that a similar situation occurs for the Stokes problem, even though a constraint-preserving (discrete solenoidal) initial velocity is only needed for the obtention of enhanced control over the time derivative of the velocity and subsequently, the pressure (see [8]).

$$b_{\mathrm{d}}\left(\bar{q}_h, \mathbf{d}_h^{n+\frac{1}{2}}, \delta_t \mathbf{d}_h^{n+1}\right)_{\delta\rho} - \frac{\epsilon^2}{2}\left(\delta_t q_h^{n+1}, \bar{q}_h\right)_\rho = 0. \tag{24}$$

Then, taking $\bar{q}_h = q_h^{n+\frac{1}{2}}$ and using that $\delta_t \|q_h^{n+1}\|_\rho^2 = 2\left(q_h^{n+\frac{1}{2}}, \delta_t q_h^{n+1}\right)_\rho$, we have that

$$b_{\mathrm{d}}\left(q_h^{n+\frac{1}{2}}, \mathbf{d}_h^{n+\frac{1}{2}}, \delta_t \mathbf{d}_h^{n+1}\right)_{\delta\rho} = \frac{\epsilon^2}{2}\left(q_h^{n+\frac{1}{2}}, \delta_t q_h^{n+1}\right)_\rho = \frac{\epsilon^2}{4k}\left(\|q_h^{n+1}\|_\rho^2 - \|q_h^n\|_\rho^2\right). \tag{25}$$

At this point, since we have used the restriction (20b), then (25) is only true for $n > 0$, because the restriction (20b) does not generally holds for $\mathbf{d}_h^0$. In the limit case ($\epsilon = 0$) see (21b). For $\epsilon > 0$, but we can define $q_h^0 \in Q_h$ such that

$$b_{\mathrm{d}}\left(\bar{q}_h, \mathbf{d}_h^0, \mathbf{d}_h^0\right)_{\delta\rho} - \epsilon^2\left(q_h^0, \bar{q}_h\right)_\rho = (1, \bar{q}_h)_\rho \quad \forall \bar{q}_h \in Q_h. \tag{26}$$

Using (25) and (26), we finally get (25) also for $n = 0$.

The rest of terms can be treated as at the continuous level. With regard to the time derivatives, we use the fact that $\left(\delta_t f^{n+1}, f^{n+\frac{1}{2}}\right) = \frac{1}{2k}\left(\|f^{n+1}\|^2 - \|f^n\|^2\right)$. The skew-symmetric version of $\tilde{c}$ is required, in order to get this result. Integrating in time, i.e. $k\sum_{m=0}^{N-1}(\cdot)$ we get the energy equality. □

4.2. Semi-implicit algorithm

The implicit algorithms are nonlinear, and so, a linearization technique and subsequent nonlinear iterations have to be performed (see Section 5). Now, we consider a semi-implicit algorithm, which is implicit in the sense that a linear system has to be solved at every iteration, but explicit in terms of nonlinearity; at every time step, the problem to be solved is linear. In the sequel, we propose a new semi-implicit scheme and prove its unconditional stability, showing the good design of the algorithm.

In order to motivate the method, we recall the incremental form (24) of the $\mathbf{d}_h$ constraint (20b). But, in the following algorithm, we will consider the incremental form of the constraint equation with the linearization $\mathbf{d}_h^{n+\frac{1}{2}} \approx \mathbf{d}_h^n$:

$$b_{\mathrm{d}}(\bar{q}_h, \mathbf{d}_h^n, \delta_t \mathbf{d}_h^{n+1}) - \frac{\epsilon^2}{2}\left(\delta_t q_h^{n+1}, \bar{q}_h\right) = 0.$$

Given $\mathbf{d}_h^n \in D_h$ and $\mathbf{u}_h^n \in V_h$, the problem at the time step t^{n+1} reads as: find the finite element functions $\left(\mathbf{d}_h^{n+1}, q_h^{n+1}, \mathbf{u}_h^{n+1}, p_h^{n+1}\right) \in D_h \times Q_h \times V_h \times P_h$ such that

$$\left(\delta_t \mathbf{d}_h^{n+1}, \bar{\mathbf{d}}_h\right) + \gamma\left(\nabla \mathbf{d}_h^{n+1}, \nabla \bar{\mathbf{d}}_h\right) + c_{\mathrm{d}}\left(\mathbf{u}_h^{n+1}, \mathbf{d}_h^n, \bar{\mathbf{d}}_h\right) + \gamma b_{\mathrm{d}}\left(q_h^{n+1}, \mathbf{d}_h^n, \bar{\mathbf{d}}_h\right)_{\delta\rho} = 0, \tag{27a}$$

$$b_{\mathrm{d}}\left(\bar{q}_h, \mathbf{d}_h^n, \delta_t \mathbf{d}_h^{n+1}\right)_{\delta\rho} - \frac{\epsilon^2}{2}\left(\delta_t q_h^{n+1}, \bar{q}_h\right)_\rho = 0, \tag{27b}$$

$$\left(\delta_t \mathbf{u}_h^{n+1}, \bar{\mathbf{u}}_h\right) + \nu\left(\nabla \mathbf{u}_h^{n+1}, \nabla \bar{\mathbf{u}}_h\right) + \tilde{c}\left(\mathbf{u}_h^n, \mathbf{u}_h^{n+1}, \bar{\mathbf{u}}_h\right) + \frac{\lambda}{\gamma} c\left(\bar{\mathbf{u}}_h, \mathbf{d}_h^n, \delta_t \mathbf{d}_h^{n+1} + (\mathbf{u}_h^{n+1} \cdot \nabla)\mathbf{d}^n\right) - b_{\mathrm{u}}\left(p_h^{n+1}, \bar{\mathbf{u}}_h\right) = \langle \mathbf{g}^{n+1}, \bar{\mathbf{u}}_h \rangle, \tag{27c}$$

$$b_{\mathrm{u}}\left(\bar{p}_h, \mathbf{u}_h^{n+1}\right) = 0. \tag{27d}$$

for any $(\bar{\mathbf{d}}_h, \bar{q}_h, \bar{\mathbf{u}}_h, \bar{p}_h) \in D_h \times Q_h \times V_h \times P_h$.

In the next theorem, we prove that in fact, this semi-implicit algorithm is unconditionally stable and satisfies an energy equality.

Theorem 4.2. *System (27) with $\epsilon = 0$ satisfies the following energy equality:*

$$\begin{aligned}
&\left\|\mathbf{u}_h^n\right\|^2 + \lambda\left\|\nabla \mathbf{d}_h^n\right\|^2 + k^2 \sum_{m=0}^{n-1}\left(\left\|\delta_t \mathbf{u}_h^{m+1}\right\|^2 + \lambda\left\|\delta_t \nabla \mathbf{d}_h^{m+1}\right\|^2\right) + 2k\sum_{m=0}^{n-1}\left(\nu\left\|\nabla \mathbf{u}_h^{m+1}\right\|^2 + \frac{\lambda}{\gamma}\left\|\delta_t \mathbf{d}_h^{m+1} + (\mathbf{u}_h^{m+1} \cdot \nabla)\mathbf{d}_h^m\right\|^2\right) \\
&\quad = 2k\sum_{m=0}^{n-1}\langle \mathbf{g}^{m+1}, \mathbf{u}_h^{m+1}\rangle + \left\|\mathbf{u}_h^0\right\|^2 + \lambda\left\|\nabla \mathbf{d}_h^0\right\|^2,
\end{aligned}$$

for any $n \in \{0, 1, \ldots, N\}$. For $\epsilon > 0$, system (27) satisfies:

$$\begin{aligned}
&\left\|\mathbf{u}_h^n\right\|^2 + \lambda\left\|\nabla \mathbf{d}_h^n\right\|^2 + \frac{\lambda\epsilon^2}{2}\left\|q_h^n\right\|_\rho^2 + k^2 \sum_{m=0}^{n-1}\left(\left\|\delta_t \mathbf{u}_h^{m+1}\right\|^2 + \lambda\left\|\delta_t \nabla \mathbf{d}_h^{m+1}\right\|^2 + \frac{\lambda\epsilon^2}{2}\left\|\delta_t q_h^{m+1}\right\|_\rho^2\right) \\
&\quad + 2k\sum_{m=0}^{N-1}\left(\nu\left\|\nabla \mathbf{u}_h^{m+1}\right\|^2 + \frac{\lambda}{\gamma}\left\|\delta_t \mathbf{d}_h^{m+1} + (\mathbf{u}_h^{m+1} \cdot \nabla)\mathbf{d}_h^{m+1}\right\|^2\right) \\
&\quad = 2k\sum_{m=0}^{n-1}\langle \mathbf{g}^{m+1}, \mathbf{u}_h^{m+1}\rangle + \left\|\mathbf{u}_h^0\right\|^2 + \lambda\left\|\nabla \mathbf{d}_h^0\right\|^2 + \frac{\lambda\epsilon^2}{2}\left\|q_h^0\right\|_\rho^2
\end{aligned}$$

for any $n \in \{0, 1, \ldots, N\}$, where q_h^0 is defined as in Theorem 4.1.

Proof. In order to prove the theorem, we need to show some relations. Now, let us define $\mathbf{w}_h^{n+1} := \delta_t \mathbf{d}_h^{n+1} + (\mathbf{u}_h^{n+1}\cdot\nabla)\mathbf{d}_h^n$. We have:

$$c\left(\mathbf{u}_h^{n+1}, \mathbf{d}_h^n, \delta_t \mathbf{d}_h^{n+1} + (\mathbf{u}_h^{n+1}\cdot\nabla)\mathbf{d}_h^n\right) = \int_\Omega (\mathbf{u}_h^{n+1}\cdot\nabla)\mathbf{d}_h^n \cdot \left(\delta_t \mathbf{d}_h^{n+1} + (\mathbf{u}_h^{n+1}\cdot\nabla)\mathbf{d}_h^n\right) = \int_\Omega (\mathbf{u}_h^{n+1}\cdot\nabla)\mathbf{d}_h^n \cdot \mathbf{w}_h^{n+1}$$
$$= \left\|\mathbf{w}_h^{n+1}\right\|^2 - \left(\delta_t \mathbf{d}_h^{n+1}, \mathbf{w}_h^{n+1}\right). \tag{28}$$

By definition of $\mathbf{w}_h^{n+1}$, Eq. (27a) can be written as:

$$\left(\mathbf{w}_h^{n+1}, \bar{\mathbf{d}}_h\right) + \gamma\left(\nabla \mathbf{d}_h^{n+1}, \nabla \bar{\mathbf{d}}_h\right) + \gamma\, b_{\mathrm{d}}\left(q_h^{n+1}, \mathbf{d}_h^n, \bar{\mathbf{d}}_h\right)_{\delta\rho} = 0. \tag{29}$$

So, testing this equation against $\delta_t \mathbf{d}_h^{n+1}$, we easily get

$$-\left(\delta_t \mathbf{d}_h^{n+1}, \mathbf{w}_h^{n+1}\right) = \frac{\gamma}{2}\delta_t \left\|\nabla \mathbf{d}_h^{n+1}\right\|^2 + \frac{\gamma k}{2}\left\|\delta_t \nabla \mathbf{d}_h^{n+1}\right\|^2 + \gamma\, b_{\mathrm{d}}\left(q_h^{n+1}, \mathbf{d}_h^n, \delta_t \mathbf{d}_h^{n+1}\right)$$
$$= \frac{\gamma}{2}\delta_t \left\|\nabla \mathbf{d}_h^{n+1}\right\|^2 + \frac{\gamma k}{2}\left\|\delta_t \nabla \mathbf{d}_h^{n+1}\right\|^2 + \gamma\frac{\epsilon^2}{4}\delta_t \left\|q_h^{n+1}\right\|_\rho^2 + \gamma\frac{\epsilon^2 k}{4}\left\|\delta_t q_h^{n+1}\right\|_\rho^2, \tag{30}$$

where we have invoked the constraint Eq. (27b) and used the fact that $\left(\delta_t f^{n+1}, f^{n+1}\right) = \frac{1}{2}\delta_t \|f^{n+1}\|^2 + \frac{k}{2}\left\|\delta_t f^{n+1}\right\|^2$. Accordingly (28)–(30), we have

$$\frac{\lambda}{\gamma} c\left(\mathbf{u}_h^{n+1}, \mathbf{d}_h^n, \delta_t \mathbf{d}_h^{n+1} + (\mathbf{u}_h^{n+1}\cdot\nabla)\mathbf{d}_h^n\right) = \frac{\lambda}{\gamma}\left\|\mathbf{w}_h^{n+1}\right\|^2 + \frac{\lambda}{2}\delta_t \left\|\nabla \mathbf{d}_h^{n+1}\right\|^2 + \frac{\lambda k}{2}\left\|\delta_t \nabla \mathbf{d}_h^{n+1}\right\|^2 + \frac{\lambda\epsilon^2}{4}\delta_t \left\|q_h^{n+1}\right\|_\rho^2 + \frac{\lambda\epsilon^2 k}{4}\left\|\delta_t q_h^{n+1}\right\|_\rho^2, \tag{31}$$

Taking $(\bar{\mathbf{u}}_h, \bar{p}_h) = (\mathbf{u}_h^{n+1}, p_h^{n+1})$ in (27c) and (27d), using (31) and applying $k\sum_{m=0}^{N-1}(\cdot)$, we prove the theorem. □

As far as we know, this is the first finite element approximation of the liquid crystal problem (7), both for the penalized and limit case, that is unconditionally stable and linear. The penalized method in [4] was unconditionally stable but nonlinear, whereas the method in [17] was linear but conditionally stable. Furthermore, both methods introduced an extra vectorial unknown to the problem, with the corresponding increasement of computational cost. The method proposed herein is more efficient because it does not introduce new unknowns and does not require nonlinear iterations. Furthermore, the method is unconditionally stable. Compared to the implicit algorithm introduced above, this method solves one linear system per time step, without the need to perform nonlinear iterations.

5. Nonlinear solvers

In order to use the implicit algorithms previously introduced, a linearization must be performed. For the subsequent exposition, let us write system (20) (or equivalently (21)) at the time step t^{n+1} in a compact manner as follows:

$$\left\langle \mathcal{L}_{\mathrm{d}}\left(\mathbf{d}_h^{n+1}, q_h^{n+1}, \mathbf{u}_h^{n+1}\right), \left(\bar{\mathbf{d}}_h, \bar{q}_h\right)\right\rangle = 0, \quad \left\langle \mathcal{L}_{\mathrm{u}}\left(\mathbf{u}_h^{n+1}, p_h^{n+1}, \mathbf{d}_h^{n+1}\right), \left(\bar{\mathbf{u}}_h, \bar{p}_h\right)\right\rangle = 0. \tag{32}$$

5.1. Exact Newton scheme

It is clear that both operators are nonlinear. At this point, we can linearize the problem using an exact Newton linearization. Given a previous iterate $\left(\mathbf{d}_h^{n+1,k}, q_h^{n+1,k}, \mathbf{u}_h^{n+1,k}, p_h^{n+1,k}\right)$, the new iterate

$$\left(\mathbf{d}_h^{n+1,k+1}, q_h^{n+1,k+1}, \mathbf{u}_h^{n+1,k+1}, p_h^{n+1,k+1}\right) = \left(\mathbf{d}_h^{n+1,k}, q_h^{n+1,k}, \mathbf{u}_h^{n+1,k}, p_h^{n+1,k}\right) + \left(\delta\mathbf{d}_h^{k+1}, \delta q_h^{k+1}, \delta\mathbf{u}_h^{k+1}, \delta p_h^{k+1}\right)$$

is obtained after solving the linear system:

$$\left\langle \frac{\mathrm{d}\mathcal{L}_{\mathrm{d}}\left(\mathbf{d}_h^{n+1,k}, q_h^{n+1,k}, \mathbf{u}_h^{n+1,k}\right)}{\mathrm{d}(\mathbf{d}_h, q_h, \mathbf{u}_h)}\cdot\left(\delta\mathbf{d}_h^{k+1}, \delta q_h^{k+1}, \delta\mathbf{u}_h^{k+1}\right), \left(\bar{\mathbf{d}}_h, \bar{q}_h\right)\right\rangle = -\left\langle \mathcal{L}_{\mathrm{d}}\left(\mathbf{d}_h^{n+1,k}, q_h^{n+1,k}, \mathbf{u}_h^{n+1,k}\right), \left(\bar{\mathbf{d}}_h, \bar{q}_h\right)\right\rangle,$$
$$\left\langle \frac{\mathrm{d}\mathcal{L}_{\mathrm{u}}\left(\mathbf{u}_h^{n+1,k}, p_h^{n+1,k}, \mathbf{d}_h^{n+1,k}\right)}{\mathrm{d}(\mathbf{u}_h, p_h, \mathbf{d}_h)}\cdot\left(\delta\mathbf{u}_h^{k+1}, \delta p_h^{k+1}, \delta\mathbf{d}_h^{k+1}\right), \left(\bar{\mathbf{u}}_h, \bar{p}_h\right)\right\rangle = -\left\langle \mathcal{L}_{\mathrm{u}}\left(\mathbf{u}_h^{n+1,k}, p_h^{n+1,k}, \mathbf{d}_h^{n+1,k}\right), \left(\bar{\mathbf{u}}_h, \bar{p}_h\right)\right\rangle, \tag{33}$$

where $\frac{\mathrm{d}\mathcal{F}(\mathbf{x}^*)}{\mathrm{d}\mathbf{x}}\cdot\delta\mathbf{x} \in Y'$ denotes the weak Gâteaux derivative of the functional $\mathcal{F}: X \to Y$ at $\mathbf{x}^*$ with respect to $\mathbf{x}$ in the direction $\delta\mathbf{x} \in X$, for X, Y Banach spaces. Problem (33) is the tangent problem of (32) around $\left(\mathbf{d}_h^{n+1,k}, q_h^{n+1,k}, \mathbf{u}_h^{n+1,k}, p_h^{n+1,k}\right)$. Every

nonlinear iteration k of the exact Newton linearization requires to solve a linear system coupling all the unknowns $\left(\delta\mathbf{d}_h^{k+1}, \delta q_h^{k+1}, \delta\mathbf{u}_h^{k+1}, \delta p_h^{k+1}\right)$ in the problem, with the corresponding computational cost; in dimension 3, it involves eight degrees of freedom per node.

5.2. Quasi-Newton scheme

For numerical purposes, it is convenient to decouple the different variables in the problem, in order to reduce the CPU time and memory usage. In particular, since we want to decouple sub-problems (20a), (20b) and (20c), (20d), we consider a quasi-Newton method in which the tangent matrix decouples problems. Given the iterate $\left(\mathbf{d}_h^{n+1,k}, q_h^{n+1,k}, \mathbf{u}_h^{n+1,k}, p_h^{n+1,k}\right)$, the new iterate is obtained after solving the two linear sub-problems:

$$\left\langle \frac{\mathrm{d}\mathcal{L}_\mathrm{d}\left(\left(\mathbf{d}_h^{n+1,k}, q_h^{n+1,k}\right), \mathbf{u}_h^{n+1,k}\right)}{\mathrm{d}(\mathbf{d}_h, q_h)} \cdot \left(\delta\mathbf{d}_h^{k+1}, \delta q_h^{k+1}\right), (\bar{\mathbf{d}}_h, \bar{q}_h) \right\rangle = -\left\langle \mathcal{L}_\mathrm{d}\left(\left(\mathbf{d}_h^{n+1,k}, q_h^{n+1,k}\right), \mathbf{u}_h^{n+1,k}\right), (\bar{\mathbf{d}}_h, \bar{q}_h) \right\rangle,$$

$$\left\langle \frac{\mathrm{d}\mathcal{L}_\mathrm{u}\left(\left(\mathbf{u}_h^{n+1,k}, p_h^{n+1,k}\right), \mathbf{d}_h^{n+1,k+1}\right)}{\mathrm{d}(\mathbf{u}_h, p_h)} \cdot \left(\delta\mathbf{u}_h^{k+1}, \delta p_h^{k+1}\right), (\bar{\mathbf{u}}_h, \bar{p}_h) \right\rangle = -\left\langle \mathcal{L}_\mathrm{u}\left(\left(\mathbf{u}_h^{n+1,k}, p_h^{n+1,k}\right), \mathbf{d}_h^{n+1,k+1}\right), (\bar{\mathbf{u}}_h, \bar{p}_h) \right\rangle.$$

Doing that, we have reduced the number of terms in the tangent problem, since we have neglected cross derivatives, i.e. the variation of $\mathcal{L}_\mathrm{d}$ with respect to $\mathbf{u}_h$ and vice versa. A nice property of this approach is the fact that it allows *modularity*. Separate codes can be used to solve the two sub-problems, only needing to pass the unknowns from one code to the other at every iteration. So, even though this method involves nonlinear iterations, it has been linearized in such a way that the computation of $(\mathbf{u}_h, p_h)$ is segregated from the one for $(\mathbf{d}_h, q_h)$, notably reducing the solver CPU time per iteration with respect to the semi-implicit method.

5.3. Nonlinear block Gauss–Seidel scheme with Alouges' method

The linearization of Eqs. (20a) and (20b), that are equivalent to the harmonic maps problem plus a convection term, can be linearized by extending the strategy proposed by Alouges in [1]. We refer to [1,3] for a mathematical analysis of the Alouges' method in the frame of steady harmonic maps. The idea is to consider a Picard linearization of (20a) and (20b) together with a normalization of the director field approximation. Obviously, this approach only has sense when the constraint over $\mathbf{d}_h$ is exactly enforced on the nodes, using one of the alternatives proposed above, and no penalty is introduced, i.e. $\epsilon = 0$. Let us consider the previous iterate $\left(\mathbf{d}_h^{n+1,k}, q_h^{n+1,k}, \mathbf{u}_h^{n+1,k}, p_h^{n+1,k}\right)$. First, we compute $\left(\tilde{\mathbf{d}}_h^{n+1,k+1}, q_h^{n+1,k+1}\right)$, solution of the linear system

$$\left(\delta_t\tilde{\mathbf{d}}_h^{n+1,k+1}, \bar{\mathbf{d}}_h\right) + \gamma\left(\nabla\tilde{\mathbf{d}}_h^{n+1,k+1}, \nabla\bar{\mathbf{d}}_h\right) + c\left(\mathbf{u}_h^{n+1,k}, \tilde{\mathbf{d}}_h^{n+1,k+1}, \bar{\mathbf{d}}_h\right) + \gamma b_\mathrm{d}\left(q_h^{n+1,k+1}, \mathbf{d}_h^{n+1,k}, \bar{\mathbf{d}}_h\right)_{h\rho} = 0,$$

$$b_\mathrm{d}\left(\bar{q}_h, \mathbf{d}_h^{n+1,k}, \tilde{\mathbf{d}}_h^{n+1,k+1}\right)_{h\rho} = \langle 1, \bar{q}_h\rangle_\rho.$$

Then, we compute $\mathbf{d}_h^{n+1,k+1}$, as the normalization of $\tilde{\mathbf{d}}_h^{n+1,k+1}$ on the nodes. So, at every node, we compute

$$(\mathbf{d}^{n+1,k+1})^a = \frac{(\tilde{\mathbf{d}}^{n+1,k+1})^a}{|(\tilde{\mathbf{d}}^{n+1,k+1})^a|}.$$

Obviously, $\mathbf{d}_h^{n+1,k+1}$ satisfies the nonlinear constraint

$$b_\mathrm{d}\left(\bar{q}_h, \mathbf{d}_h^{n+1,k+1}, \tilde{\mathbf{d}}_h^{n+1,k+1}\right)_{h\rho} = \langle 1, \bar{q}_h\rangle_\rho.$$

The third step of the algorithm consists of solving (20c) and (20d) with the known value of $\mathbf{d}_h^{n+1,k+1}$, e.g. using a Picard linearization: we compute $\left(\mathbf{u}_h^{n+1,k+1}, p_h^{n+1,k+1}\right)$ solution of

$$\left(\delta_t\mathbf{u}_h^{n+1,k+1}, \bar{\mathbf{u}}_h\right) + \nu\left(\nabla\mathbf{u}_h^{n+1,k+1}, \nabla\bar{\mathbf{u}}_h\right) + \tilde{c}\left(\mathbf{u}_h^{n+1,k}, \mathbf{u}_h^{n+1,k+1}, \bar{\mathbf{u}}_h\right)_h + b_\mathrm{u}\left(p_h^{n+1,k+1}, \bar{\mathbf{u}}_h\right)$$
$$+ \frac{\lambda}{\gamma} c\left(\bar{\mathbf{u}}_h, \mathbf{d}_h^{n+1,k+1}, \partial_t\mathbf{d}_h^{n+1,k+1} + \left(\mathbf{u}_h^{n+1,k+1}\cdot\nabla\right)\mathbf{d}_h^{n+1,k+1}\right) = \langle \mathbf{g}^{n+1}, \bar{\mathbf{u}}_h\rangle, b_\mathrm{u}\left(\bar{p}_h, \mathbf{u}_h^{n+1,k+1}\right) = 0.$$

So, the final procedure involves a linearized harmonic map-like system, with a convection term, a normalization of the director field, and the linearized Navier–Stokes equations, e.g. using Picard, evaluating the coupling elastic term with the director field of the second step. The final problem has a computational cost per iteration similar to the quasi-Newton algorithm above, since the two sub-problems have been decoupled.

6. Numerical experimentation

In this section, we perform some numerical experiments, in order to check the behavior of the methods proposed above. We will distinguish between three different numerical methods:

- The implicit method with nodally exact enforcement of the constraint and Crank–Nicolson time integration. Thus, the method consists of system (20) with $\rho = s$ and $\delta = h$, that is to say, using a closed integration rule for the constraint equation and the bilinear form b_d in (15), i.e. Eq. (17). The problem is linearized using the quasi-Newton scheme in Section 5.2, that decouples $\mathbf{d}_h$ and $\mathbf{u}_h$ computation at the linear solver level. We will denote this method as *nodal implicit method*.
- The implicit method with $\mathcal{P}_h^1$ as Lagrange multiplier space and Crank–Nicholson time integration. In this case, the method consists of system (20) with the expression of b_d in (14). Again, we use the quasi-Newton scheme for linearization. We will denote this method as $P1$ implicit method.
- The semi-implicit method (27) with a closed integration rule for the constraint equation and the bilinear form b_d in (15). We will denote this method as *semi-implicit method*.

For all the methods, we will consider both exact and penalized formulations. In all cases, we have used quadrature rules that integrate exactly all the terms in the linear system.

One of the outputs of the simulations are the time behavior of the different energies interacting in the system. Let us define the elastic and kynetic energies respectively as:

$$\mathcal{E}_d(t) = \|\nabla \mathbf{d}_h(t)\|^2, \quad \mathcal{E}_u(t) = \|\mathbf{u}_h(t)\|^2.$$

The validation of the code has been carried out by using the method of manufactured solutions and checking that the implicit methods under consideration are energy preserving, i.e. the equality (22) is satisfied up to convergence tolerance. So, for the implicit methods with $\epsilon = 0$ and zero forcing terms, the energy equality has been checked:

$$\lambda \mathcal{E}_d(t^n) + \mathcal{E}_u(t^n) + \sum_{m=0}^{n-1} \mathcal{E}_{\text{dis}}\left(t^{m+\frac{1}{2}}\right) = \mathcal{E}_0$$

where

$$\mathcal{E}_{\text{dis}}(t) = 2\nu k \|\nabla \mathbf{u}_h(t)\|^2 + 2\frac{\lambda}{\gamma} k \|\delta_t \mathbf{d}_h(t) + \mathbf{u}_h(t) \cdot \nabla \mathbf{d}_h(t)\|^2$$

denotes the energy dissipation and $\mathcal{E}_0$ the initial energy (see (22)). We have analyzed three different test problems. Two of them have been presented in previous numerical works about the approximation of liquid crystals and serve for comparison with pre-existing techniques. The third example is a problem with known analytical solution that clearly serves to assess the algorithms.

6.1. Example 1: a smooth harmonic map

The first test under consideration has been previously solved in [4]. We consider the problem (3) for $\epsilon = 0$ in the square domain $\Omega = (-1,1)^2$ and $\mathbf{g} = \mathbf{0}$. The initial conditions are:

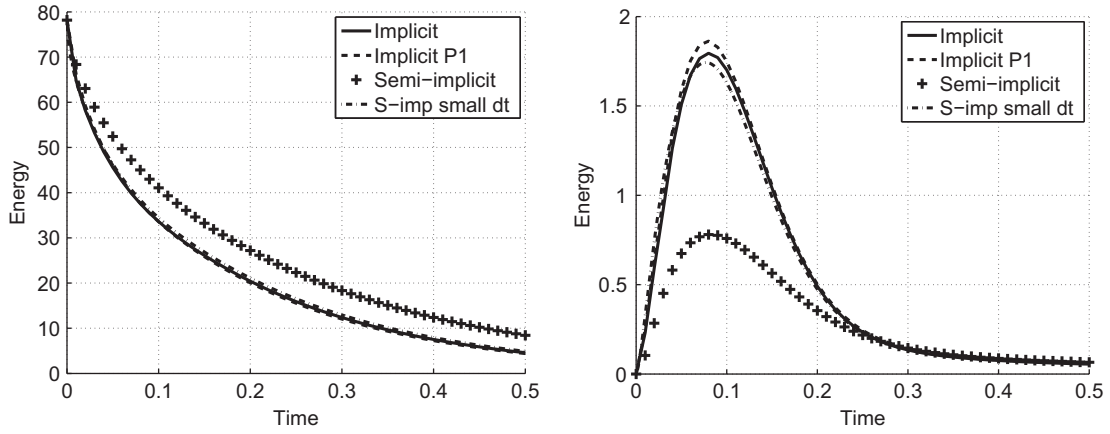
$$\mathbf{u}_0 = \mathbf{0}, \quad \mathbf{d}_0 = (\sin(a), \cos(a))^t, \quad a = 2.0\pi(\cos(x) - \sin(y)),$$

whereas homogeneous Dirichlet and Neumann boundary conditions are enforced over $\mathbf{u}$ and $\mathbf{d}$ respectively. So, the only energy introduced to the system is via the initial energy $\mathcal{E}_0$. The physical parameters are $\lambda = \gamma = 1.0$ and $\nu = 0.1$, and the numerical parameters are chosen as $h = 2^{-5}$, $k = 0.01$, and $\epsilon = 0$, unless otherwise stated.

In Fig. 1 we show the elastic and kynetic energies $\mathcal{E}_d(t)$ and $\mathcal{E}_u(t)$ with respect to time. The initial conditions introduce energy to the system via the initial elastic energy $\lambda\|\nabla \mathbf{d}_0\|^2$, that is related to (23) (see [4, Fig. 5.1]). It is clear out of the results that most of the initial elastic energy is (unsurprisingly) transferred to $\mathcal{E}_d$, whereas the kynetic energy for the velocity $\mathcal{E}_u$ at its peak ($t = 0.1$ s) is only about a 5% of $\mathcal{E}_d$. The steady-state that is reached for the initial condition stated above has no energy as $t \to \infty$, since the steady-state $\mathbf{d}_h$ is constant in space (see [4, Fig. 5.1]).

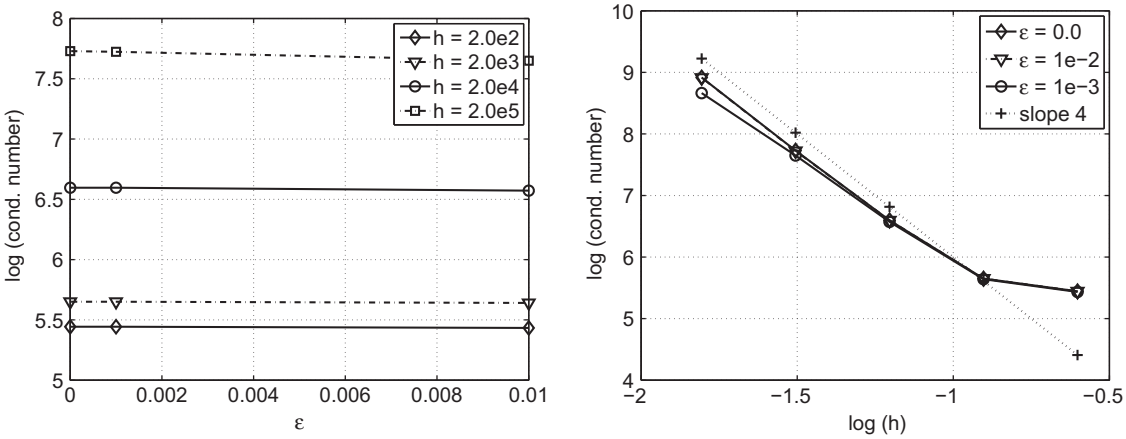
We have plot the energy results for the three methods under consideration, whereas for the semi-implicit method, we have also included the results for $k = 10^{-3}$. Out of these plots, we can conclude that the nodal constraint and the $P1$ Lagrange multiplier lead to very similar results. As expected, the semi-implicit method shows some lag in the dynamics to the steady state, but reducing by 10 the time step size, the results are almost identical to those with an implicit method. With regard to the number of nonlinear iterations, the implicit method required an average of 5.36 iterations per time step in the time interval $(0,1)$, using as convergence criterion $\frac{\|x^{k+1}-x^k\|}{\|x^k\|} < \text{tol}$ for all the unknowns of the problem and $\text{tol} = 10^{-6}$.

Now, we consider $\lambda = 0.0$, reducing the problem to the transient harmonic maps system. We do this since we want to evaluate the dependence of the condition number and the constant in the discrete inf–sup condition (16), for the $\mathbf{d}$ problem; these features are well-known for the Navier–Stokes block. We compute the condition number of the system matrix for a given time step value, a penalty parameter $\epsilon = 0$, 10^{-2}, 10^{-3} and $h = 2^{-i}$ with $i = 2, 3, 4, 5$; the results are collected in



(a) $\mathcal{E}_d(t)$ (b) $\mathcal{E}_u(t)$

Fig. 1. Example 1: $\mathcal{E}_d(t)$ and $\mathcal{E}_u(t)$ plots for the nodal implicit, $P1$ implicit and explicit methods, for $\epsilon = 0$ and $k = 10^{-2}$. The explicit method is also used with $k = 10^{-3}$.



(a) Cond. number vs. ϵ (b) Cond. number vs. h

Fig. 2. Example 1: Condition number of the **d** block-matrix vs. (ϵ, h).

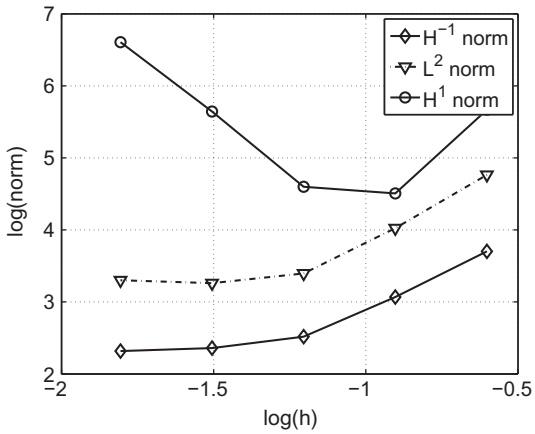


Fig. 3. Example 1: q_h stability vs. h for the L^2, H^1 and discrete H^{-1} norm.

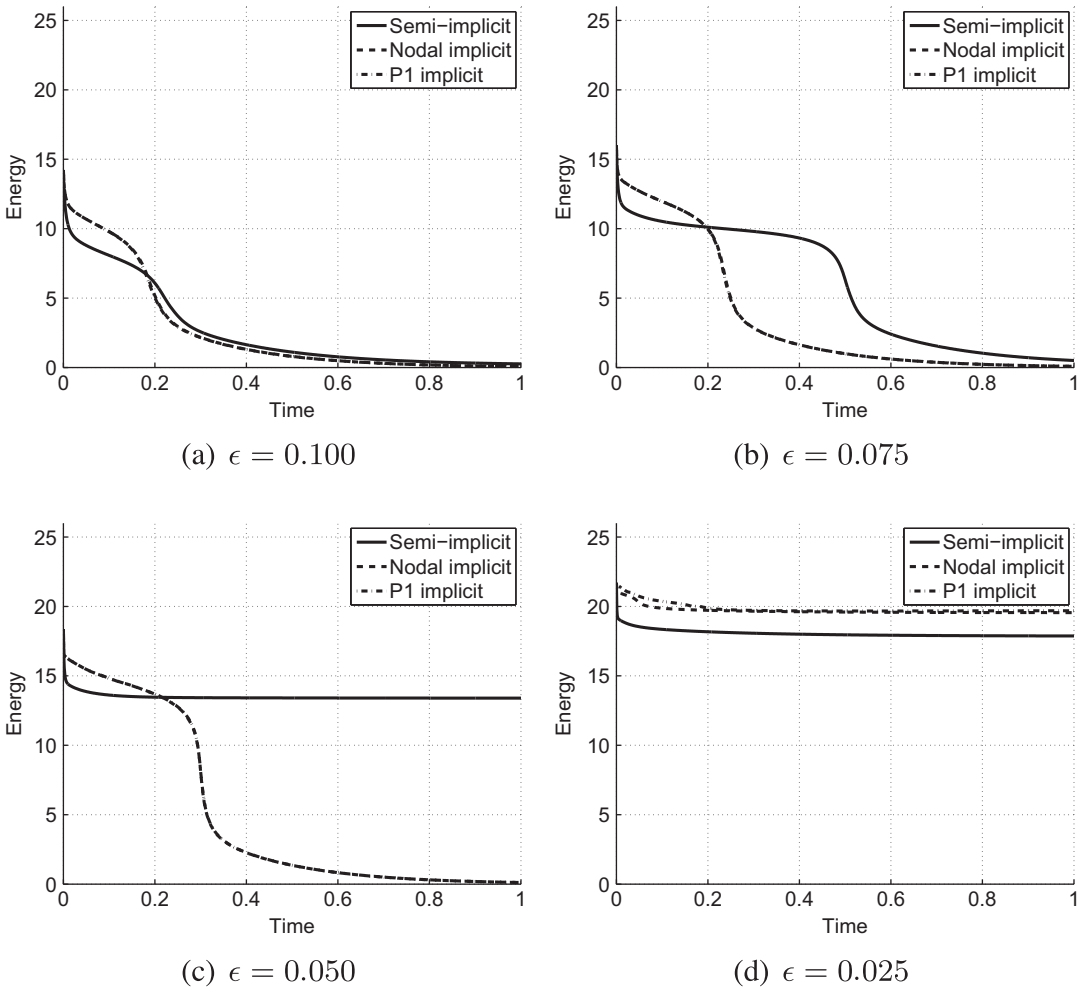


Fig. 4. Example 2: Elastic energy $\mathcal{E}_d(t)$ plots for the nodal implicit, $P1$ implicit and semi-implicit methods. The results are shown for different values of ϵ.

Fig. 2. In Fig. 2(a) we see that the condition number of the matrix is not sensitive to ϵ, which means that the stability of the Lagrange multiplier comes from the discrete inf–sup condition. Let us remark that for the GL formulations in the literature (see e.g. [4,23,17,25,26]) the condition number blows up for $\epsilon \searrow 0$. We also plot in Fig. 2(b) the condition number in terms of the mesh size h for different values of ϵ. We see that the condition number is almost $\mathcal{O}(h^{-4})$. Finally, we want to analyze the constant in the inf–sup condition (16). In order to do this, we evaluate the L^2 norm, the H^1 norm and a discrete H^{-1} norm. Given $q_h \in Q_h$, let us find the discrete Riestz projection $g_h \in Q_h$ such that

$$(g_h, f_h)_{H^1} = \langle q_h, f_h \rangle_{H^{-1}\times H^1}.$$

We define the discrete H^{-1} norm as follows:

$$\|q_h\|_{H_d^{-1}} := \sup_{f_h \in Q_h} \frac{\langle q_h, f_h \rangle_{H^{-1}\times H^1}}{\|f_h\|_1} = \sup_{f_h \in L_h} \frac{(g_h, f_h)_{H^1}}{\|f_h\|_1} = \|g_h\|_1.$$

In Fig. 3 we plot the norm of the Lagrange multiplier at a given time step size, for the uniform partition introduced above. We can easily see that the Lagrange multiplier q_h is stable in both the H^{-1} and L^2 norms. As expected, the H^1 norm blows up with h. This test indicates that for smooth enough solutions, $q_h \in L^2(\Omega)$.

6.2. Example 2: annihilation and stable defects

The second test we consider can be found in any article about GL-based numerical approximations of liquid crystals. It consists on the annihilation of regularized initial singularities. The problem is solved in the square domain $\Omega = (-1,1)^2$ and $\mathbf{g} = \mathbf{0}$. The initial conditions are:

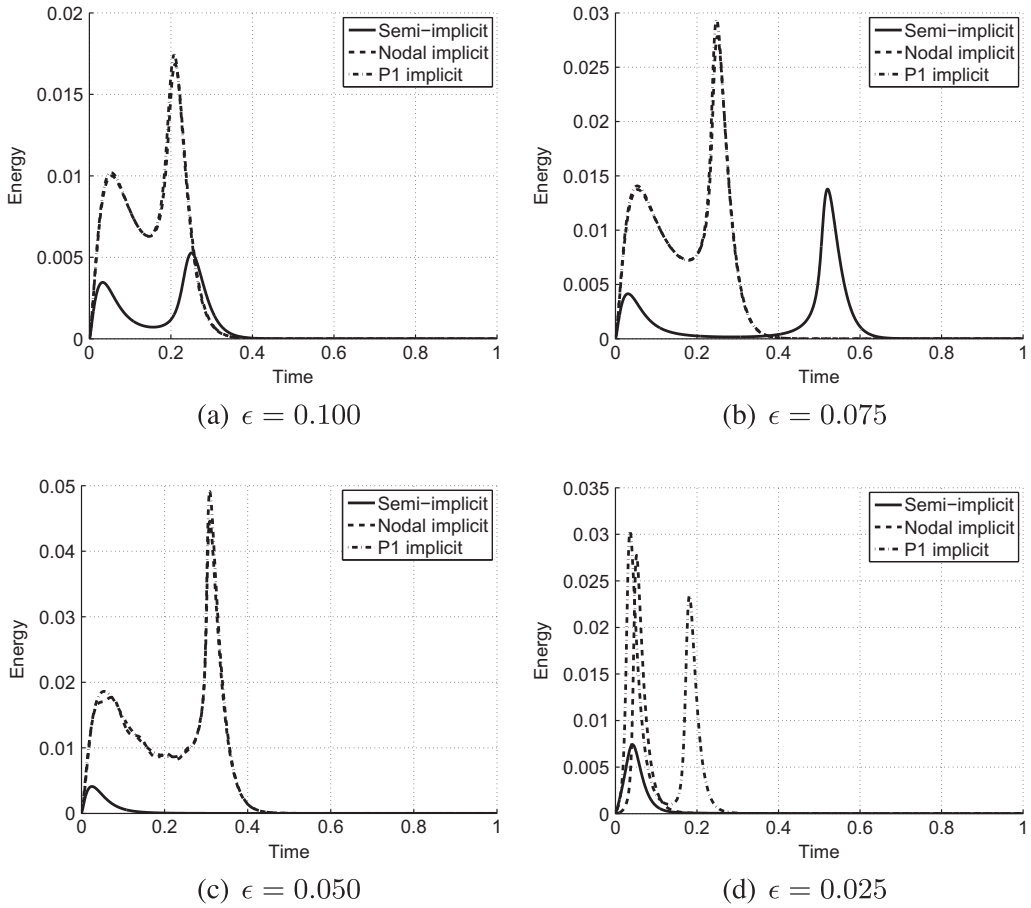


Fig. 5. Example 2: Kynetic energy $\mathcal{E}_u(t)$ plots for the nodal implicit, $P1$ implicit and semi-implicit methods. The results are shown for different values of ϵ.

$$\mathbf{u}_0 = \mathbf{0}, \quad \mathbf{d}_0 = \tilde{\mathbf{d}}/\sqrt{|\tilde{\mathbf{d}}|^2 + \epsilon^2}, \quad \text{for } \tilde{\mathbf{d}} = (x^2 + y^2 - 0.25, y)^t.$$

The physical parameters are $\lambda = \gamma = \nu = 1.0$, and the numerical parameters are chosen as $h = 2^{-5}$, $k = 0.001$, unless otherwise stated. We have considered $\epsilon = 0, 2.5 \cdot 10^{-3}, 5 \cdot 10^{-3}, 7.5 \cdot 10^{-3}, 10^{-2}$.

As commented in [4, Section 5.2] the solution critically depends on the initial data and so, on ϵ. It was also pointed out by Lin and Liu in [23, Example 2.2] that the evolution of the singularities was very sensitive to the algorithm and the mesh selected in the computation. Furthermore, in both references the authors point out that the local energy at the singularity goes to infinity as $\epsilon \searrow 0$. In this article, we put particular attention on the solution with respect to ϵ. The initial condition introduces initial energy to the system in two different ways. On one side, via the initial elastic energy, and on the other side, via the q^0 term. However, the second term reads as:

$$\frac{\lambda\epsilon^2}{2}\|q^0\|^2 = \frac{\lambda}{2}\|\mathbf{d}^0 \cdot \mathbf{d}^0 - 1\|^2 = \frac{\lambda}{2}\int_\Omega \left(\frac{\epsilon^2}{|\tilde{\mathbf{d}}|^2 + \epsilon^2}\right)^2, \tag{34}$$

which tends to zero as $\epsilon \searrow 0$. At the discrete level the energy also depends on the projection that is used for obtaining $\mathbf{d}_h^0$ from $\mathbf{d}^0$. We use a nodal projection for the sake of comparison with previous works. However, this projector is ill-posed as $(\epsilon, h) \searrow 0$ in the singularity, but can be solved by giving an average value for the node over the singularity, if there is one. In the limit case ($\epsilon = 0$), the initial energy is only coming from the initial elastic energy.

In Fig. 4 we show the elastic energy $\mathcal{E}_d(t)$ for the three formulations and different values of ϵ. The implicit saddle-point Ginzburg–Landau formulation proposed herein leads to very similar results to those obtained for classical GL formulation (see [17,4,23]). On the other hand, it is clear that the annihilation time goes to ∞ as $\epsilon \searrow 0$. For $\epsilon = 0.025$ the implicit methods converge to a stable solution with non-zero elastic energy and zero velocity. In fact, as commented above, this long-term behavior is in agreement with the structure of the dynamical system at hand. The steady-state solution for $\mathbf{d}_h$ is a solution of the discrete harmonic maps problem. These numerical results are in concordance with the mathematical analysis of

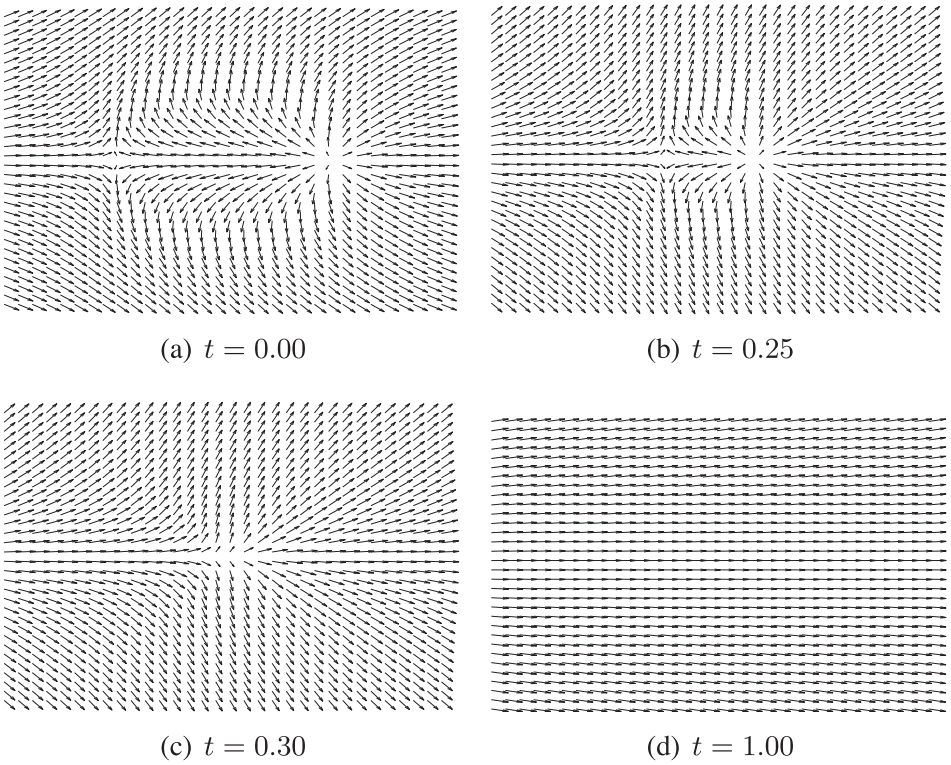


Fig. 6. Example 2: Director vector **d** fields at different time step values for the nodal implicit method and $\epsilon = 5 \cdot 10^{-3}$.

Brezis, Coron and Lieb in [6], which proves that this singular solution minimizes the energy of a stationary pair of point defects. A physical justification of the existence of stable *combed hedgehog* point defects can be found in [28] using different approaches, e.g. Peach–Köhler type arguments (see also [30,29,14,9]). The average number of nonlinear iterations in the time interval (0,1) is 3.04 for the convergence criterium defined above with tol = 10^{-6}.

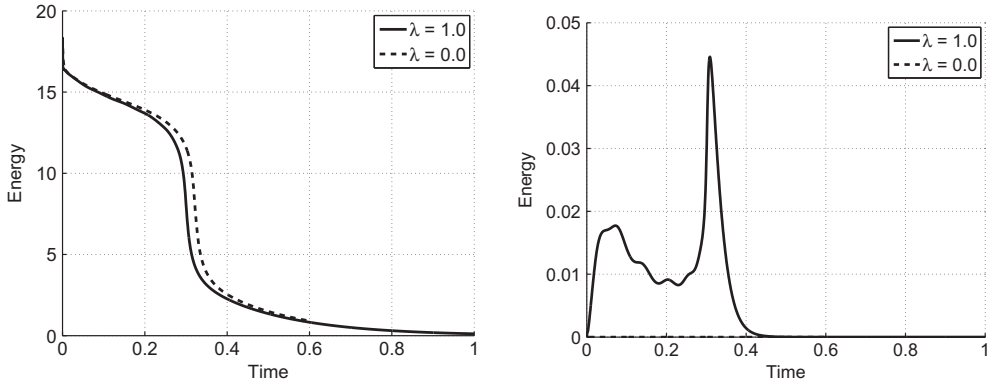
The results obtained with nodal and $P1$ implicit methods are almost identical. We show the director vector fields at different time values for $\epsilon = 5{\cdot}10^{-3}$ in Fig. 6. We see how the singularities approach and annihilate at some point over $t = 0.30$ s, which coincides with an abrupt decreasement of elastic energy (see Fig. 4(c)).

In Fig. 4 we also show the results for the explicit method. It is clear, out of these results, that the effective penalty of the semi-implicit formulation is smaller than the one for the implicit method. For ϵ large enough or small enough, the results are fairly similar. However, in the transition between them, the explicit method exhibits a lag with respect to the implicit solution, or can even tend to a different stationary point. Summarizing, the solution is critically ϵ-dependent. For $\epsilon = 0.0$, the results are almost identical to those for $\epsilon = 0.025$. As far as we know, this behavior for $\epsilon \searrow 0$ has not been previously analyzed because previous GL formulations were ill-posed in this assymptotic regime, and a minimum value of $\epsilon = 0.05$ was common practice. Herein, this problem has been solved by using a saddle-point GL approach.

We also plot the kynetic energy in all cases (see Fig. 5). These results make clear that this test problem is almost a harmonic maps problem, since the energy transferred to **u** is very small. The results for the implicit methods are very similar, but for $\epsilon = 2.5 \cdot 10^{-3}$ the implicit $P1$ method exhibits two picks of kynetic energy whereas the nodal version only one.

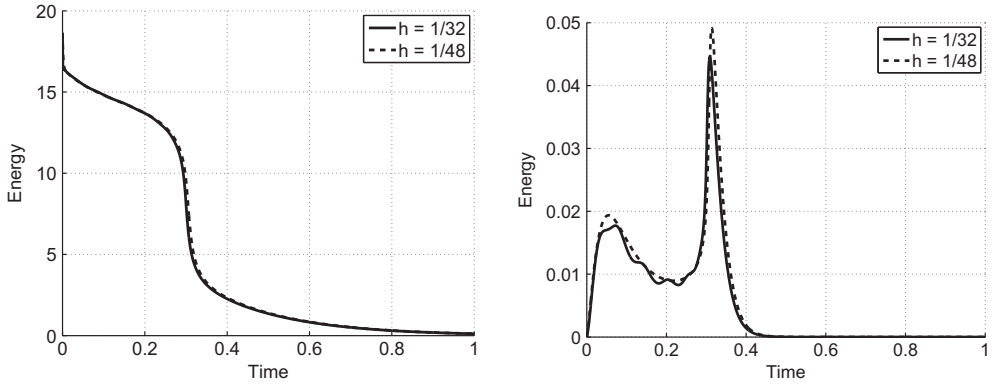
In Fig. 7(a) and (b) we have considered the full liquid crystal problem with $\lambda = 1.0$ and the transient harmonic maps problem, i.e. $\lambda = 0.0$. The elastic energy is almost identical in both cases, and so, the **u** influence almost neglectable. Obviously, the kynetic energy is zero for $\lambda = 0.0$ in Fig. 7(b). With regard to the mesh size, we have compared both elastic and kynetic energy for $h = 2^{-5}, 2^{-6}$, showing that the solution is well converged for the mesh being used. Analogously, we show the results for $k = 10^{-3}, 10^{-2}$, that allow to say that the time step size is also acceptable.

Since the initial energy introduced to the system is not enough for the annihilation of singularities for the limit problem, and the energy transferred to the velocity is erratic, we have modified the initial velocity as $\mathbf{u}^0 = \omega(-y,x)$, with $\omega = 5, 50$. We show some $\mathbf{d}_h$ vector fields at different values for $\omega = 50$ in Fig. 9. A comparison of the elastic energy for these two different initial velocities can be found in Fig. 8. The case with $\omega = 50$ shows richer dynamics, as expected. Since there is no forcing term, the initial velocity is dissipated, and the elastic energy starts to show a linear and slow decay. The initial velocity with $\omega = 5$ seems to be not enough, leading to results similar to those with $\omega = 0$.



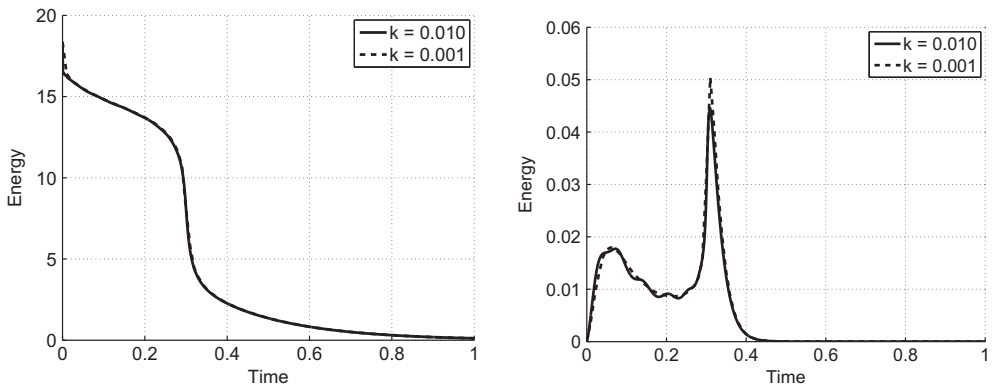
(a) $\mathcal{E}_d(t)$ vs. λ

(b) $\mathcal{E}_u(t)$ vs. λ



(c) $\mathcal{E}_d(t)$ vs. h

(d) $\mathcal{E}_u(t)$ vs. h



(e) $\mathcal{E}_d(t)$ vs. k

(f) $\mathcal{E}_u(t)$ vs. k

Fig. 7. Example 2: $\mathcal{E}_d(t)$ and $\mathcal{E}_u(t)$ plots for the nodal implicit method with $\epsilon = 5 \cdot 10^{-3}$ and different values of (λ, h, k).

6.3. Example 3: magical spiral

We end the numerical experiments section with the *magical spiral* problem. A nice presentation of the problem and the obtention of the analytical solution can be found in [10, pag. 158]. It consists of two concentric cylinders with the following anchoring boundary conditions: the molecules are normal to the inner cylinder and tangential to the external cylinder. The numerical simulations have been performed for $\lambda = \gamma = 1.0$, $\epsilon = 0$ and $k = 0.01$. The initial velocity is zero and the initial director field we have considered is plotted in Fig. 10(a).

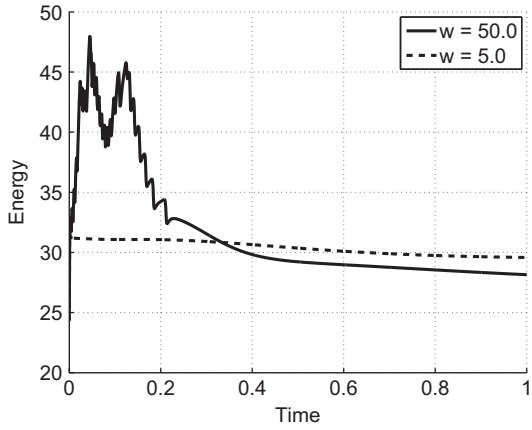


Fig. 8. Example 2: $\mathcal{E}_s(t)$ plot for the nodal implicit method, $\epsilon = 0.0$ and $\omega = 5, 50$.

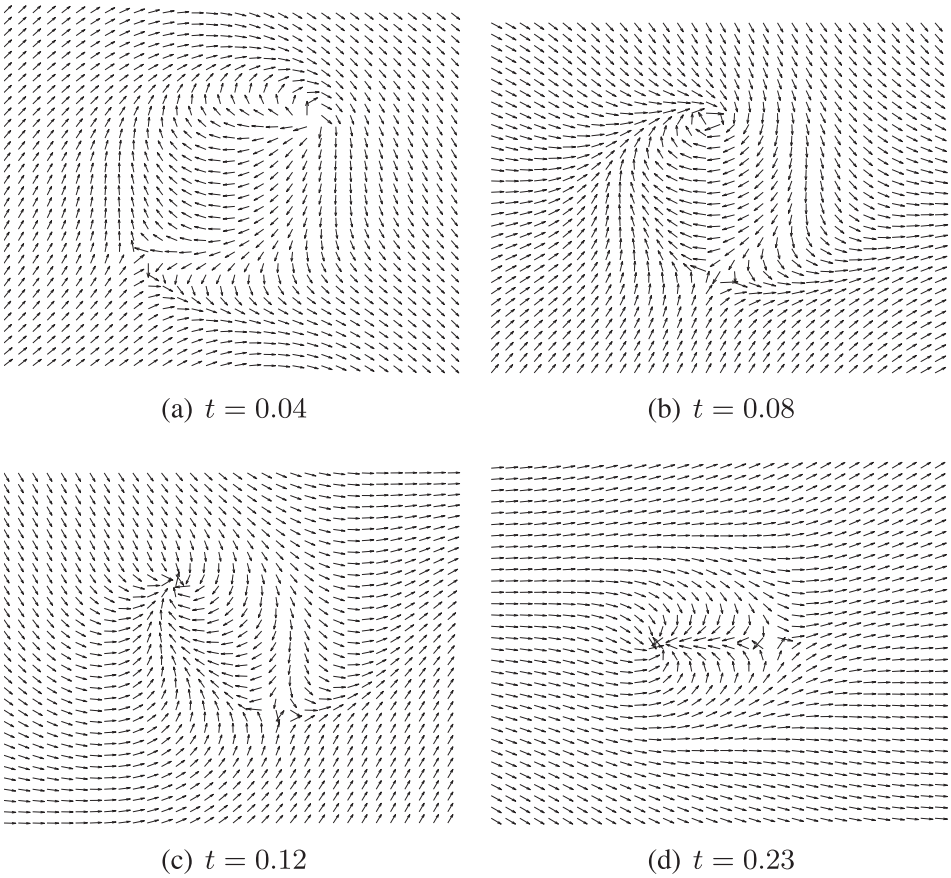
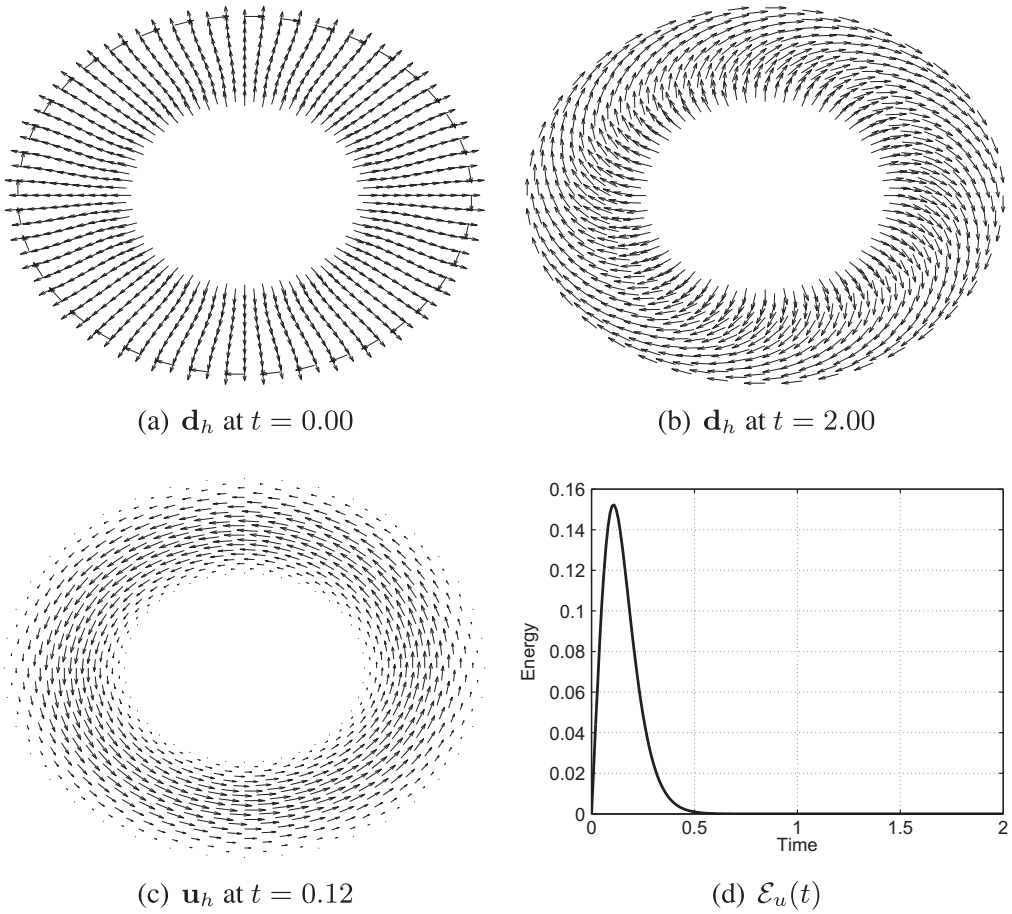


Fig. 9. Example 2: Director vector **d** fields at different time step values for the nodal implicit method, $\epsilon = 0$ and $\omega = 50$.

In Fig. 10(b) we show the magical spiral that is obtained in the steady-state limit. With regard to the velocity, we plot the vector field for the time step $t = 0.12$ s (the one with maximum $\mathcal{E}_u$) in Fig. 10(c), and the kynetic energy in Fig. 10(d). These results have been obtained with the nodal implicit method.

The nice feature of this test problem is the fact that it can be experimentally observed and analytically solved. It can be proved that the angle ψ between **d** and the radial direction (see [10, Fig. 3.23]) is equal to:



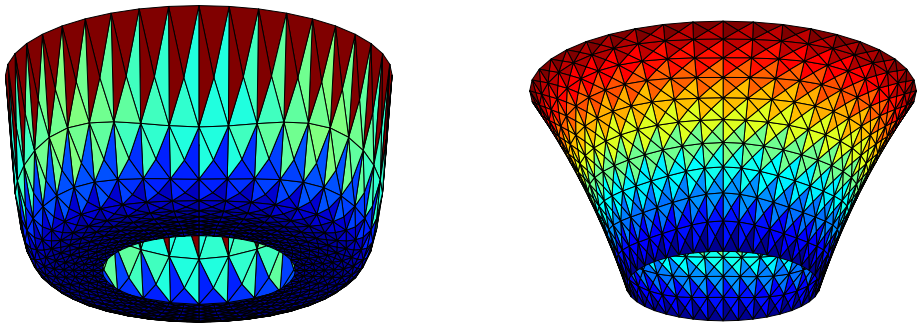
(a) $\mathbf{d}_h$ at $t = 0.00$ (b) $\mathbf{d}_h$ at $t = 2.00$

(c) $\mathbf{u}_h$ at $t = 0.12$ (d) $\mathcal{E}_u(t)$

Fig. 10. Example 3: Director vector **d** fields at different time step values, velocity vector field and $\mathcal{E}_u(t)$ plot for the nodal implicit method with $\epsilon = 0$.

$$\psi = \frac{\pi}{2}\frac{\ln(r/r_0)}{\ln(r_1/r_0)},$$

where r is the radial coordinate and r_0 and r_1 the inner and outer cylinder radius. In this particular case, we have considered $r_0 = 1$ m and $r_1 = 2$ m. We plot the values of ψ for the first time step value and the steady-state case in Fig. 11. In Fig. 12 we plot the error between the numerical approximation for ψ and the exact solution. We easily see that the steady-state solution converges to the exact solution; the error is reduced to the order of $\mathcal{O}(10^{-2})$, which is the numerical error associated to the mesh.



(a) $t = 0.01$ (b) $t = 2.00$

Fig. 11. Example 3: Plots of ψ on Ω at different time step values, obtained with the nodal implicit method and $\epsilon = 0$.

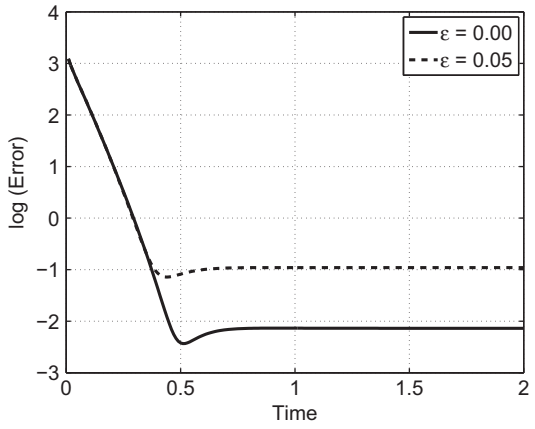


Fig. 12. Example 3: L^2-error between exact and computed values of ψ vs. time for $\epsilon = 0,\ 5 \cdot 10^{-2}$.

In order to evaluate the effect of the penalty in the accuracy of the solution, we have solved the same problem with $\epsilon = 5 \cdot 10^{-2}$. The error plot can also be found in Fig. 12. We can easily see that, even for a fairly coarse mesh, the penalty error is more than one order of magnitude larger than the discretization error. This result illustrates very well the dramatic impact of the penalty term in the accuracy of the liquid crystal simulations. So, it seems much more reasonable to stick to stable saddle-point formulations that do not require any penalization.

7. Conclusions

In this work, we have considered a different approach to the numerical approximation of nematic liquid crystals. The problem is posed in a saddle-point form. We have also considered a penalty version of the problem, analogous to the Ginzburg–Landau penalization, but using the saddle-point structure. The gain doing this is the fact that the resulting numerical schemes are stable for any choice of the penalty $\epsilon \geqslant 0$ when using Lagrangian finite elements, based on the theoretical results in [19] for the steady harmonic maps problem. This is a gain with respect to the previous penalized formulations, since the condition number of the resulting linear system does not blow up as $\epsilon \searrow 0$.[2] Another important consequence is the fact that the limit and penalized problems can be treated in a unified way. On the other hand, the saddle-point formulation involves quadratic nonlinear terms whereas the Ginzburg–Landau formulation includes cubic nonlinearities.

The time integration of the problem has also been considered. We have distinguished between implicit methods and semi-implicit methods in which nonlinear iterations are not performed. As far as we know, the semi-implicit method proposed herein is the first one that is unconditionally stable, which has been possible to obtain by using the saddle-point structure of the director vector problem. Furthermore, it is not needed to include any auxiliary variable. On the other hand, as in the previous approaches, all the unknowns are coupled at the linear system, which makes its solution expensive.

On the contrary, the implicit method includes nonlinear iterations. However, we have designed a quasi-Newton linearization of the problem that allows to decouple the velocity and director vector sub-problems at the linear system level, clearly reducing the CPU cost of the solver per iteration. Another nice feature of the method proposed is the fact that it is energy preserving, satisfying the same energy equality as the continuous problem. As the semi-implicit method, it is unconditionally stable and does not require the introduction of any auxiliary variable.

We have performed a set of test problems comparing different approaches. It is interesting to note that we have obtained interesting results in the asymptotic regime $\epsilon \searrow 0$ for typical test problems including defects. Furthermore, we have checked the accuracy of the method for a problem with analytical solution and assessed the serious effect of penalization over it.

Let us finish comparing the methods considered in this work with the ones in the literature. Since the methods and the interesting features to be used for comparison are many, we have included all in Table 1. From these results, we could easily extract some recommendations. The methods proposed herein are the only ones that unify the limit and penalized problem and their condition number is independent of ϵ. The limit and penalized schemes in [4] are different, and so, not treated in a unified way. The limit problem in [4] is the only one in the existing literature that approximates the original problem, but the method ends up being only conditionally stable. The sense in which the sphere constraint is enforced is much weaker than the one in this article; the restriction does not appear explicitly in the problem.

With regard to CPU cost, the methods proposed herein end up having eight degrees of freedom per node, only beated by the methods in [23,24] (for the penalized problem) and [4] (for the limit problem); the method in [25] does not introduce any auxiliary unknown but it requires $\mathcal{C}^1$ finite element approximations, dramatically increasing the CPU cost. Furthermore, the methods proposed herein are both unconditionally stable, as the ones in [24,4] (for the penalized problem). Another

[2] Furthermore, penalized formulations require to tune ϵ with respect to the mesh and time step sizes in order to obtain accurate results; the authors in [4] claim that this is a drawback of the penalized formulation, since this tuning turns out to be subtle.

Table 1
Comparison of methods. ISP refers to the implicit saddle point method in this work and SSP refers to the semi-implicit one. The rest of the methods are denoted by the reference in which they were proposed. In particular, we have denoted the penalized scheme in [4] with [4, P] whereas the one for the limit problem as [4, L]. With regard to auxiliary unknown, we have considered q as an extra unknown, since it does not appear in previous works.

Method	ISP	SSP	[17]	[23]	[24]	[25]	[26]	[4, P]	[4, L]	[16]
Aux. unknowns (#)	1	1	3	0	0	0	9	3	0	3
Semi-implicit	×	✓	✓	✓	×	×	×	×	×	✓
Uncond. stab. $\forall(\epsilon,h,k)$	✓	✓	×	×	✓	×	×	✓	✓	×
ϵ-indep. cond. number	✓	✓	×	×	×	×	×	×	✓	×
Energy preserving	✓	×	×	×	✓	×	×	×	×	×
Quadratic nonlinear	✓	✓	×	×	×	×	×	×	×	×
$\mathcal{C}^0$ approx. for **d**	✓	✓	✓	✓	✓	×	✓	✓	✓	✓

important benefit of the saddle-point structure is the quadratic nonlinearity of the resulting system, in comparison to the cubic nonlinearity of all the previous GL schemes and the scheme for the limit problem in [4]; this fact simplifies the linearization of the problem and makes its convergence easier.

For small values of ϵ, the use of saddle-point methods should be favoured, since they are unconditionally stable and the condition number of the resulting linear system for the rest of unconditionally stable methods becomes too large to be solved. But this is not the only improvement of these saddle-point methods, as we can see in the table and commented above. For large enough values of the penalty, the scheme in [24] seems to be the most appealing among the ones in the literature. However, it is unclear the physical interest of the results obtained for large values of the penalty, since, out of our numerical experiments, they are inaccurate.

References

[1] F. Alouges, A new algorithm for computing liquid crystal stable configurations: the harmonic mapping case, SIAM J. Numer. Anal. 34 (5) (1997) 1708–1726.
[2] J.W. Barrett, S. Bartels, X. Feng, A. Prohl, A convergent and constraint-preserving finite element method for the p-harmonic flow into spheres, SIAM J. Numer. Anal. 45 (3) (2007) 905–927.
[3] S. Bartels, Stability and convergence of finite-element approximation schemes for harmonic maps, SIAM J. Numer. Anal. 43 (1) (2005) 220–238.
[4] R. Becker, X. Feng, A. Prohl, Finite element approximations of the Ericksen–Leslie model for nematic liquid crystal flow, SIAM J. Numer. Anal. 46 (4) (2008) 1704–1731.
[5] S.C. Brenner, L.R. Scott, The Mathematical Theory of Finite Element Methods, Springer-Verlag, 1994.
[6] H. Brezis, J.M. Coron, E.H. Lieb, Harmonic maps with defects, Commun. Math. Phys. 107 (4) (1986) 649–705.
[7] F. Brezzi, M. Fortin, Mixed and Hybrid Finite Element Methods, Springer-Verlag, 1991.
[8] E. Burman, M.A. Fernández, Galerkin finite element methods with symmetric pressure stabilization for the transient Stokes equations: stability and convergence analysis, SIAM J. Numer. Anal. 47 (1) (2008) 409–439.
[9] C. Chiccoli, O.D. Lavrentovich, P. Pasini, C. Zannoni, Monte Carlo simulations of stable point defects in hybrid nematic films, Phys. Rev. Lett. 79 (22) (1997) 4401–4404.
[10] P.G. de Gennes, J. Prost, The Physics of Liquid Crystals, Oxford University Press, 1995.
[11] J. Ericksen, Continuum theory of nematic liquid crystals, Res. Mech. 21 (1987) 381–392.
[12] J.L. Ericksen, Conservation laws for liquid crystals, Trans. Soc. Rheol. 5 (1961) 22–34.
[13] A. Ern, J.L. Guermond, Theory and Practice of Finite Elements, Springer-Verlag, 2004.
[14] J.J. Feng, C. Zhou, Orientational defects near colloidal particles in a nematic liquid crystal, J. Colloid Interface Sci. 269 (2004) 72–78.
[15] F.C. Frank, On the theory of liquid crystals, Discuss. Faraday Soc. 25 (1958) 19–28.
[16] V. Girault, F. Guillén-González. Mixed formulation, approximation and decoupling algorithm for a penalized nematic liquid crystals model. Mathematics of Computation, (in press).
[17] F. Guillén-González, J.V. Gutiérrez-Santacreu. A linear mixed finite element scheme for a nematic Ericksen–Leslie liquid crystal model. submitted for publication.
[18] J.G. Heywood, R. Rannacher, Finite element approximation of the nonstationary Navier–Stokes problem. I: Regularity of solutions and second-order error estimates for spatial discretization, SIAM J. Numer. Anal. 19 (1982) 275–311.
[19] Q. Hu, X.-C. Tai, R. Winther, A saddle-point approach to the computation of harmonic maps, SIAM J. Numer. Anal. 47 (2) (2009) 1500–1523.
[20] F. Leslie, Some constitutive equations for liquid crystals, Arch. Ration. Mech. Anal. 28 (1968) 265–283.
[21] F.M. Leslie, Theory of flow phenomena in liquid crystals, Adv. Liq. Cryst. 4 (1979) 1–81.
[22] F.H. Lin, Nonlinear theory of defects in nematic liquid crystals: phase transition and flow phenomena, Commun. Pure Appl. Math. 42 (1989) 789–814.
[23] P. Lin, C. Liu, Simulations of singularity dynamics in liquid crystal flows: a C^0 finite element approach, J. Comput. Phys. 215 (2006) 348–362.
[24] P. Lin, C. Liu, H. Zhang, An energy law preserving C0 finite element scheme for simulating the kinematic effects in liquid crystal dynamics, J. Comput. Phys. 227 (2) (2007) 1411–1427.
[25] C. Liu, N.J. Walkington, Approximation of liquid crystal flows, SIAM J. Numer. Anal. 37 (3) (2000) 725–741.
[26] C. Liu, N.J. Walkington, Mixed methods for the approximation of liquid crystal flows, Math. Model. Numer. Anal. 36 (2) (2002) 205–222.
[27] C. Oseen, Theory of liquid crystals, Trans. Faraday Soc. 29 (1933) 883–899.
[28] L.M. Pismen, B.Y. Rubinstein, Motion of interacting point defects in nematics, Phys. Rev. Lett. 69 (1) (1992).
[29] R. Repnik, L. Mathelitsch, M. Svetec, S. Kralj, Physics of defects in nematic liquid defects, Eur. J. Phys. 24 (2003) 481–492.
[30] M. Svetec, S. Kralj, Z. Bradač, S. Žumer, Annihilation of nematic point defects: Pre-collision and post-collision evolution, The Eur. Phys. J. 20 (E) (2006) 71–79.
[31] R. Temam, Navier–Stokes Equations, North-Holland, 1984.