# Variational Dual Solutions of Chern-Simons Theory

Amit Acharya*       Janusz Ginster†       Ambar N. Sengupta‡

## Abstract

A scheme for generating weakly lower semi-continuous action functionals corresponding to the Euler-Lagrange equations of Chern-Simons theory is described. Coercivity is deduced for such a functional in appropriate function spaces to prove the existence of a minimizer, which constitutes a solution to the Euler-Lagrange equations of Chern-Simons theory in a relaxed sense. A geometric analysis is also made, especially for the gauge group $SU(2)$, relating connection forms on the bundle to corresponding forms in the dual scheme.

## 1   Introduction

The goal of this paper is to define a dual variational principle corresponding to the Euler-Lagrange equations of the Chern-Simons functional and to prove existence of minimizers of the dual functional. The functional has the property that its critical points formally define solutions to the Euler-Lagrange equations of Chern-Simons theory [1]. Our approach is an application of a recently developed scheme for generating variational principles corresponding to a given set of equations and side conditions [2, 3, 4, 5, 6, 7, 8, 9]; the equations can represent dissipative or conservative systems. The scheme treats the primal PDE under consideration as constraints and invokes a more-or-less arbitrarily designable strictly convex, auxiliary potential to be optimized. Then, a *convex* dual variational principle for the Lagrange multiplier (dual) fields emerges involving a dual-to-primal (DtP) mapping (an adapted change of variables), with the special property that its Euler-Lagrange equations are exactly the primal PDE system, interpreted as equations governing the dual fields using the DtP mapping. These ideas are closely related to those of Brenier [10, 11] on 'hidden convexity' of nonlinear PDE.

The Chern-Simons functional is not bounded above or below and hence neither minimizers nor maximizers exist (see the discussion in Section 3). This renders the Direct Method of the Calculus of Variations not feasible to prove existence of solutions to the Euler-Lagrange equations. Instead, the proposed method in our approach associates the original problem to the Euler-Lagrange equations of a *dual* functional. We prove that for specific choices of auxiliary potentials this dual functional turns out to be bounded below, lower-semicontinuous and coercive. In particular, the Direct Method of the Calculus of Variations can be invoked to show existence of minimizers of the dual functional. Assuming that the dual functional is regular enough and the DtP mapping is well-defined, these minimizers correspond to extremals of the Chern-Simons theory. This is why we refer to them as *variational dual solutions* of the theory.

An outline of the paper is as follows: In Sec. 2 we present a largely self-contained account of the Chern-Simons functional and the associated variational equations (11), describing flat connections,

---

*Department of Civil & Environmental Engineering, and Center for Nonlinear Analysis, Carnegie Mellon University, Pittsburgh, PA 15213, email: acharyaamit@cmu.edu.

†Weierstrass Institute, Mohrenstrasse 39, 10117 Berlin, Germany, email: ginster@wias-berlin.de.

‡Department of Mathematics, University of Connecticut, Storrs, CT 06269, email: ambarnsg@gmail.com.

which we spell out more explicitly (14) for the case where the gauge group is $SU(2)$; following this, in section 2.3, we work out the geometry of the dual scheme for the variational equations. In Theorem 2.1, we show that the critical points of the dual functional $S_H$ correspond to flat connections with the prescribed boundary values. The explicit expression for the dual functional for a quadratic auxiliary potential is developed in Sec. 2.4 and the consistency of the scheme is shown, i.e., that there exists at least one dual functional whose critical point delivers a solution to the Euler-Lagrange equations of Chern-Simons theory satisfying the prescribed boundary conditions, for every solution of the theory. In Sec. 3 we undertake a variational analysis of dual functionals corresponding to Chern-Simons theory, and prove the existence of variational dual solutions of the same (Theorem 3.1).

## 2 Geometric background

One of the foundational results in geometry is the Gauss-Bonnet theorem, which says that the integral of the curvature over a closed surface is $2\pi$ times the Euler characteristic, a topological invariant, of the surface. A far-reaching generalization was proved by Chern [12] a hundred years later, showing that the integral of the Pfaffian (a polynomial in the matrix entries) of the curvature 2-form over a closed orientable $(2n)$-dimensional manifold equals $(2\pi)^n$ times the Euler characteristic of the manifold. In proving the Gauss-Bonnet theorem and Chern's theorem, it is useful to construct a potential $\Phi$ for the Pfaffian of the curvature. These ideas generalize to the setting of connections on principal bundles: the Pfaffian of the curvature 2-form, corresponding to a connection 1-form $\omega$ on a principal $G$-bundle $\pi : P \to M$, is a closed $(2n)$-form that descends to a closed $(2n)$-form $\mathrm{Pf}(\Omega)$ on $M$. The $(2n)$-form on $P$ has a global potential $\Phi$ that descends locally, but, in general, not globally, to a $(2n-1)$-form on the base manifold $M$ (this is a special case of ([1, Proposition 3.2]). In the case of a 4-dimensional manifold, this potential is the Chern-Simons form, with a suitable normalization constant. If a closed orientable manifold were to be split into two 4-dimensional contractible submanifolds with a common 3-dimensional boundary, then the integral of $\mathrm{Pf}(\Omega)$ over each submanifold equals, by Stokes' theorem, the integral of the local potential over the boundary 3-manifold. Our focus is on the Chern-Simons 3-form for the group $SU(2)$ over an open subset of $\mathbb{R}^3$.

### 2.1 The Chern-Simons form

We work with a smooth 1-form $A$, defined on a compact oriented 3-manifold $M$, possibly with boundary, with values in a Lie algebra $\mathfrak{G}$, which is equipped with a symmetric, bilinear pairing $\langle \cdot, \cdot \rangle$ on $\mathfrak{G} \times \mathfrak{G}$, satisfying the ad-invariance condition (59). We define the *Chern-Simons* 3-form to be:

$$\mathrm{cs}(A) = \left\langle A \wedge dA + \frac{1}{3} A \wedge [A \wedge A] \right\rangle, \tag{1}$$

with $\langle \cdot \rangle$ being as defined in (55). We take the Chern-Simons action to be the integral of this form over $\mathbb{R}^3$ :

$$\mathrm{CS}(A) = \int_M \mathrm{cs}(A). \tag{2}$$

(As remarked earlier, a certain normalization constant multiplier is needed in other contexts.)

An example of interest is when
$$M = \Sigma \times [0,1],$$

2

where $\Sigma$ is a compact, oriented surface. We can consider this as embedded as a submanifold-with-boundary in $\mathbb{R}^3$, by considering the surface $\Sigma$ to be sitting inside $\mathbb{R}^3$. A different example, is given by simply a closed ball in $\mathbb{R}^3$.

### 2.1.1 The case of the gauge group $SU(2)$

Let us specialize to the case of $SU(2)$, with Lie algebra $su(2)$. Notation and basic facts about $SU(2)$ and $su(2)$ are summarized in section A.1. We consider $su(2)$-valued smooth 1-forms $A$ on a 3-manifold, and we work locally with a coordinate chart $(x^1, x^2, x^3)$. (The form $A$ is to be thought of as a connection form on a bundle pulled down by means of a section.) The Chern-Simons 3-form for $A$ is then

$$\mathrm{cs}(A) = -\mathrm{Tr}\left(A \wedge dA + \frac{2}{3} A \wedge A \wedge A\right). \tag{3}$$

We work with a fixed orthonormal basis $\{E_1, E_2, E_3\}$ of $su(2)$, and the standard basis 1-forms $\{dx^1, dx^2, dx^3\}$ for the space of ordinary real-valued 1-forms. Then we can write an $su(2)$-valued 1-form $A$ as

$$A = A_p \otimes dx^p = \sum_{J=1}^{3} A_{Jp} E_J \otimes dx^p, \tag{4}$$

where each

$$A_p = \sum_{J=1}^{3} A_{Jp} E_J$$

is an $su(2)$-valued function on $\Omega$, with each $A_{JP}$ being a real-valued function on $\Omega$. Then the wedge product $A \wedge dA$ is

$$A \wedge dA = \sum_{p,q,r} A_p \partial_q A_r \, dx^p \wedge dx^q \wedge dx^r \tag{5}$$

and

$$A \wedge A \wedge A = \sum_{p,q,r} A_p A_q A_r \, dx^p \wedge dx^q \wedge dx^r. \tag{6}$$

Using skew-symmetry of the wedge product it can be checked that these are both $su(2)$-valued. Then

$$
\begin{aligned}
\mathrm{cs}(A) &= -\mathrm{Tr}\left(A \wedge dA + \frac{2}{3} A \wedge A \wedge A\right) \\
&= -\mathrm{Tr}\left(A_p \partial_q A_r + \frac{2}{3} A_p A_q A_r\right) dx^p \wedge dx^q \wedge dx^r \\
&= -\mathrm{Tr}\left(A_p \partial_q A_r + \frac{1}{3} A_p [A_q, A_r]\right) dx^p \wedge dx^q \wedge dx^r \\
&= -\mathrm{Tr}\left(A_{Jp} \partial_q A_{Kr} (E_J E_K) + \frac{1}{3} A_{Jp} A_{Kq} A_{Lr} E_J [E_K, E_L]\right) dx^p \wedge dx^q \wedge dx^r \\
&= 2\varepsilon_{pqr}\left(A_{Jp} \partial_q A_{Jr} + \frac{2}{3} \varepsilon_{JKL} A_{Jp} A_{Kq} A_{Lr}\right) dx^1 \wedge dx^2 \wedge dx^3,
\end{aligned}
\tag{7}
$$

where we have used the commutation relations (73), the orthonormality (74), and (75).

## 2.2 The Euler-Lagrange equations for the Chern-Simons action

To compute the directional derivative of the CS action, for any smooth 1-form $v$ on $M$, a compact oriented 3-manifold with boundary, we have

$$
\begin{aligned}
\frac{d}{dt}\Big|_{t=0} &\mathrm{CS}(A+tv) \\
&= \int_M \left\langle v \wedge dA + A \wedge dv + \frac{1}{3}\left(v \wedge [A \wedge A] + A \wedge [v \wedge A] + A \wedge [A \wedge v]\right)\right\rangle \\
&= \int_M \left\langle v \wedge dA + dv \wedge A + \frac{1}{3}\left(v \wedge [A \wedge A] + A \wedge [v \wedge A] + A \wedge [A \wedge v]\right)\right\rangle \\
&\qquad \text{(using (56))} \\
&= \int_M \left\langle d(v \wedge A) + 2v \wedge dA + \frac{1}{3}\left(v \wedge [A \wedge A] + v \wedge [A \wedge A] + v \wedge [A \wedge A]\right)\right\rangle \\
&\qquad \text{(using (62))} \\
&= \int_{\partial M} \langle v \wedge A\rangle + 2\int_M \left\langle v \wedge \left(dA + \frac{1}{2}[A \wedge A]\right)\right\rangle,
\end{aligned}
\tag{8}
$$

where $F^A$ is the curvature of $A$:

$$
F^A := dA + \frac{1}{2}[A \wedge A].
\tag{9}
$$

Thus,

$$
d\mathrm{CS}|_A v = \int_{\partial M} \langle v \wedge A\rangle + 2\int_M \langle v \wedge F^A\rangle.
\tag{10}
$$

Taking $v|_{\partial M} = 0$, which means that we fix the values of $A$ on the boundary, the condition that $A$ is a critical point for the Chern-Simons integral is that its curvature $F^A$ be 0:

$$
\text{Euler-Lagrange equations for CS}: \quad F^A = 0.
\tag{11}
$$

Consider now the case $M = \Sigma \times [0,1]$, where $\Sigma$ is a compact, oriented, surface. We denote by $t$ the coordinate along $[0,1]$. We take $A$ to be of the form $A_{\Sigma \times \{t\}}$, with the coefficient of $dt$ being 0; then $A$ being flat means that $A_{\Sigma \times \{t\}}$ is a flat connection over the surface $\Sigma \times \{t\}$ and is the same connection, viewed over $\Sigma$, for all $t$. In other words, $A$ corresponds to a path of flat connections over $\Sigma$.

### 2.2.1 The Euler-Lagrange equations for the $SU(2)$ case

Working now with $A$ being an $su(2)$-valued 1-form on $\Omega \subset \mathbb{R}^3$, we have the variation computation:

$$
\begin{aligned}
\delta\mathrm{CS}(A) &\\
= 2\epsilon_{pqr} &\int_\Omega \Big[(\partial_q A_{Jr})\delta A_{Jp} - (\partial_q A_{Jp})\delta A_{Jr} \\
&\qquad + \frac{2}{3}\epsilon_{JKL}\left\{(\delta A_{Jp})A_{Kq}A_{Lr} + A_{Jp}(\delta A_{Kq})A_{Lr} + A_{Jp}A_{Kq}\delta A_{Lr}\right\}\Big]\,d\mathrm{vol} \\
= 2 &\int_\Omega \Big[\epsilon_{pqr}\partial_q A_{Jr} - \epsilon_{rqp}\partial_q A_{Jr} \\
&\qquad + \frac{2}{3}\left\{\epsilon_{pqr}\epsilon_{JKL}A_{Kq}A_{Lr} + \epsilon_{qpr}\epsilon_{KJL}A_{Kq}A_{Lr} + \epsilon_{rqp}\epsilon_{LKJ}A_{Lr}A_{Kq}\right\}\Big]\delta A_{Jp}\,d\mathrm{vol}
\end{aligned}
$$

4

$$= 2 \int_{\Omega} \left[ \epsilon_{pqr} \left( \partial_q A_{Jr} + \partial_q A_{Jr} \right) + \frac{2}{3} \epsilon_{pqr} \epsilon_{JKL} \left\{ A_{Kq} A_{Lr} + A_{Kq} A_{Lr} + A_{Kq} A_{Lr} \right\} \right] \delta A_{Jp} \, d\text{vol} \qquad (12)$$

$$= 4 \int_{\Omega} \epsilon_{pqr} \left[ \partial_q A_{Jr} + \epsilon_{JKL} A_{Kq} A_{Lr} \right] \delta A_{Jp} \, d\text{vol}.$$

For this variation to be zero for all variations in $A$, we have the condition

$$\epsilon_{pqr} \left[ \partial_q A_{Jr} + \epsilon_{JKL} A_{Kq} A_{Lr} \right] = 0. \qquad (13)$$

For each fixed $p$, this gives

$$\partial_q A_{Jr} - \partial_r A_{Jq} + \epsilon_{JKL} \left\{ A_{Kq} A_{Lr} - A_{Kr} A_{Lq} \right\} = 0. \qquad (14)$$

This is again the flatness condition (11), just written in local coordinates.

## 2.3 Geometric description of the dual scheme for Chern-Simons

As always, we work with an open subset $\Omega$ of $\mathbb{R}^3$, such that $\Omega \cup \partial\Omega$ is a compact, oriented manifold, in the natural structure inherited from $\mathbb{R}^3$. Moreover, we also work with a Lie algebra $\mathfrak{G}$, equipped with an ad-invariant inner-product $\langle \cdot, \cdot \rangle$, and suitably smooth $\mathfrak{G}$-valued 1-forms.

We have seen in (11) that the extrema of the Chern-Simons functional are given by flat connections. In this section we carry out, in geometric language, the *dual method* [2, 5]: this (i) writes a connection $A$ (the primal variable here) as $A^{(H)}(\lambda)$, in terms of a dual variable $\lambda$, and (ii) designs a variational principle involving a functional $S_H(\lambda)$ such that the Euler-Lagrange equations for $S_H(\lambda)$ automatically produce both the flatness condition (11) for $A$ and a given boundary condition for $A$ on $\partial\Omega$. We accomplish this in Theorem 2.1.

### 2.3.1 Lie-algebra-valued forms

At each $p \in \Omega$, we have the space

$$\mathfrak{G} \otimes T_p^* \Omega$$

of linear maps from $T_p \Omega$ with values in $\mathfrak{G}$. The union

$$\mathfrak{G} \otimes (T^* \Omega) := \cup_{p \in \Omega} (\mathfrak{G} \otimes T_p^* \Omega)$$

is a bundle over $\Omega$, with projection map

$$\mathfrak{G} \otimes (T^* \Omega) \to \Omega : A_p \mapsto p, \qquad (15)$$

for $A_p \in \mathfrak{G} \otimes T_p^* \Omega$ and all $p \in \Omega$. A section of the bundle (15) is a 1-form with values in $\mathfrak{G}$.

### 2.3.2 The auxiliary function $H$

We will work with a smooth function $H$ defined on the bundle space $\mathfrak{G} \otimes (T^* \Omega)$, with the restriction of $H$ to the space $\mathfrak{G} \otimes T_p^* \Omega$ denoted by $H_p$. Then for any $A_p \in \mathfrak{G} \otimes (T_p^* \Omega)$, the map

$$dH_p \Big|_{A_p} : \mathfrak{G} \otimes T_p^* \Omega \to \mathbb{R}; \qquad v \mapsto dH_p \Big|_{A_p} v = \frac{d}{dt} \Big|_{t=0} H_p(A_p + tv) \qquad (16)$$

is linear in $v$ and so there exists a unique element

$$\nabla H(A_p) \in \mathfrak{G} \otimes T_p^* \Omega$$

5

such that

$$dH_p\Big|_{A_p} v = \langle \nabla H(A_p), v \rangle \qquad \text{for all } v \in \mathfrak{G} \otimes T_p^*\Omega. \tag{17}$$

Moreover, $\nabla H$, viewed as a $\mathfrak{G}$-valued function on $\mathfrak{G} \otimes (T^*\Omega)$, is smooth. Especially when working in coordinates, it will be simpler if we write $\nabla H(A)$ instead of $\nabla H(A_p)$.

To express the gradient in coordinates, let $E_1, \ldots, E_{\dim \mathfrak{G}}$ be an orthonormal basis of $\mathfrak{G}$; then $\{E_Z \otimes dx^q\}$ is an orthonormal basis of $\mathfrak{G} \otimes T_p^*\Omega$, and

$$\partial_{Zr} H(A) := dH_p\Big|_A (E_Z \otimes dx^q) = \langle \nabla H(A), E_Z \otimes dx^q \rangle. \tag{18}$$

So we can write the gradient as

$$\nabla H(A) = \partial_{Zr} H(A) E_Z \otimes dx^r. \tag{19}$$

An example that we will explore further later is with $H(A_p)$ given by

$$\frac{1}{2}\|A_p - \bar{A}_p\|^2 = \frac{1}{2}\sum_{j=1}^{3} \langle A_p(e_j) - \bar{A}_p(e_j), A_p(e_j) - \bar{A}_p(e_j) \rangle,$$

where $e_1, e_2, e_3$ are orthonormal vectors in $T_p\Omega$, $\langle \cdot, \cdot \rangle$ is the inner-product on $\mathfrak{G}$, and $\bar{A}_p$ is some chosen fixed element of $\mathfrak{G} \otimes T_p^*\Omega$; for this function, the gradient is given by

$$\nabla \frac{1}{2}\|A - \bar{A}\|^2 = \sum_{j=1}^{3} \left( A(e_j) - \bar{A}(e_j) \right) e^j,$$

where $e^1, e^2, e^3$ form the basis of $T_p^*\Omega$ dual to $e_1, e_2, e_3$.

### 2.3.3 The connection space $\mathcal{A}_k$

Let

$$\mathcal{A}_k$$

be the set of all $C^k$ sections, with $k \geq 1$, of the bundle $\mathfrak{G} \otimes (T^*\Omega) \to \Omega$. Thus, $\mathcal{A}_k$ is the space of $\mathfrak{G}$-valued $C^k$-smooth connections over $\Omega$. Although in standard differential geometry, a connection is always $C^\infty$, we can work, for our purposes, with $C^k$ forms, where $k \geq 1$.

### 2.3.4 The pre-dual functional $\hat{S}_H$

We introduce now the *pre-dual* functional

$$\hat{S}_H : \mathcal{A}_k \times \mathcal{A}_k \to \mathbb{R}, \tag{20}$$

given by

$$\hat{S}_H[A, \lambda] = -\int_{\partial\Omega} \langle \lambda \wedge A^{(b)} \rangle + \int_{\Omega} \left( \left\langle d\lambda \wedge A + \frac{1}{2}\lambda \wedge [A \wedge A] \right\rangle + H(A)\, d\text{vol} \right). \tag{21}$$

Here $A^{(b)}$ is a fixed $\mathfrak{G}$-valued $C^k$ 1-form on $\partial\Omega$, which we are not displaying in $\hat{S}_H[A, \lambda]$ for the sake of notational simplicity. (If we were to work at the level of a bundle $P$ over $\Omega$, then a connection would be a $\mathfrak{G}$-valued 1-form on $P$; however, we keep things simple by using a global section of $P$ over $\Omega$, which makes it possible to bring all relevant forms and calculations down to $\Omega$.)

6

**Proposition 2.1.** *The functional $\hat{S}_H$ has directional derivatives everywhere. If $A, \lambda, v$ are $\mathfrak{G}$-valued $C^k$ 1-forms on $\Omega$, with $k \geq 1$, we have*

$$\partial_1 \hat{S}_H(A, \lambda)|v \overset{\text{def}}{=} \frac{d}{dt}\Big|_{t=0} \hat{S}_H[A + tv, \lambda]$$
$$= \int_\Omega \langle *(d^A \lambda) + \nabla H(A), v \rangle \, d\text{vol}, \tag{22}$$

*where $d^A \lambda$ is the covariant derivative:*

$$d^A \lambda = d\lambda + [A \wedge \lambda], \tag{23}$$

*and $*$ is the Hodge dual operation and the gradient $\nabla H(A)$ is the $\mathfrak{G}$-valued 1-form on $\Omega$ for which*

$$H'(A)v = \langle \nabla H(A), v \rangle \tag{24}$$

*holds for all $v \in \mathcal{A}_k$. Moreover,*

$$\partial_2 \hat{S}_H(A, \lambda)|\beta \overset{\text{def}}{=} \frac{d}{ds}\Big|_{s=0} \hat{S}_H(A, \lambda + s\beta)$$
$$= \int_{\partial\Omega} \langle \beta \wedge (A - A^{(b)}) \rangle + \int_\Omega \langle \beta \wedge F^A \rangle \tag{25}$$

*for all $\beta \in \mathcal{A}_k$.*

*Proof.* By straightforward computation, we have

$$\frac{d}{dt}\Big|_{t=0} \hat{S}_H[A + tv, \lambda] = \int_\Omega \left( \left\langle d\lambda \wedge v + \frac{1}{2}\lambda \wedge [v \wedge A] + \frac{1}{2}\lambda \wedge [A \wedge v] \right\rangle + H'(A)v \, d\text{vol} \right). \tag{26}$$

Next, using (56),(61) and (62), we have:

$$\left\langle d\lambda \wedge v + \frac{1}{2}\lambda \wedge [v \wedge A] + \frac{1}{2}\lambda \wedge [A \wedge v] \right\rangle = \langle d\lambda \wedge v + \lambda \wedge [v \wedge A] \rangle$$
$$= \langle d\lambda \wedge v + [A \wedge \lambda] \wedge v \rangle \tag{27}$$
$$= \langle d^A \lambda \wedge v \rangle,$$

We can then write (26) as

$$\partial_1 \hat{S}_H(A, \lambda)|v = \int_\Omega \left( \langle d^A \lambda \wedge v \rangle + H'(A)v \, d\text{vol} \right)$$
$$= \int_\Omega \langle *(d^A \lambda) + \nabla H(A), v \rangle \, d\text{vol}. \tag{28}$$

The directional derivative with respect to $\lambda$ is:

$$\partial_2 \hat{S}_H(A, \lambda)|\beta \overset{\text{def}}{=} \frac{d}{ds}\Big|_{s=0} \hat{S}_H(A, \lambda + s\beta)$$
$$= \int_{\partial\Omega} \langle \beta \wedge (A - A^{(b)}) \rangle + \int_\Omega \langle \beta \wedge F^A \rangle \tag{29}$$

$\square$

7

We equip all spaces of $C^k$ differential forms, with values in $\mathfrak{G}$ or in $\mathbb{R}$, with a family of seminorms given by the suprema of derivatives of the components of the form. This makes each such space of forms a Fréchet space (Example 1.1.5 in [13]).

**Proposition 2.2.** *The function*

$$\hat{S}_H : \mathcal{A}_k \times \mathcal{A}_k \to \mathbb{R} : (A, \lambda) \mapsto \hat{S}_H(A, \lambda) \tag{30}$$

*is directionally $C^1$.*

*Proof.* Combining (28) and (22), we have

$$\hat{S}'_H(A, \lambda)|(v, \beta) = \int_\Omega \langle *(d^A\lambda) + \nabla H(A), v \rangle \, d\mathrm{vol} + \int_{\partial\Omega} \langle \beta \wedge (A - A^{(b)} \rangle + \int_\Omega \langle \beta \wedge F^A \rangle. \tag{31}$$

This is continuous in $(A, \lambda, v, \beta)$. $\qquad\square$

### 2.3.5 Coordinate expressions

We now express the integrand in the first term on the right of (31) in coordinates. Let $E_1, \ldots, E_{\dim\mathfrak{G}}$ be an orthonormal basis of $\mathfrak{G}$, and we use the coordinate system $(x^1, x^2, x^3)$ on $\Omega$. Then we can write $A$ and $\lambda$ as

$$\begin{aligned} A &= \sum_{r=1}^{3} \sum_{Z=1}^{\dim\mathfrak{G}} A_{Zr} \, E_Z \otimes dx^r \\ \lambda &= \sum_{r=1}^{3} \sum_{Z=1}^{\dim\mathfrak{G}} \lambda_{Zr} \, E_Z \otimes dx^r. \end{aligned} \tag{32}$$

Then

$$\begin{aligned} *(d^A\lambda) + \nabla H(A) &= *(\partial_q \lambda_p dx^q \wedge dx^p + [A_q, \lambda_p] dx^q \wedge dx^p) + \partial_{Di} H(A) E_D \otimes dx^i \\ &= \epsilon_{qpi} (\partial_q \lambda_p + [A_q, \lambda_p]) \otimes dx^i + \partial_{Di} H(A) E_D \otimes dx^i \\ &= (\epsilon_{qpi} \partial_q \lambda_{Dp} E_D + \epsilon_{rpi} A_{Cr} \lambda_{Zp} [E_C, E_Z]) \otimes dx^i + \partial_{Di} H(A) E_D \otimes dx^i \\ &= [\epsilon_{qpi} \partial_q \lambda_{Dp} + A_{Cr} \lambda_{Zp} 2\tilde{\epsilon}_{CZD} \epsilon_{rpi} + \partial_{Zk} H(A)] E_D \otimes dx^k \\ &\qquad \text{where } \tilde{\epsilon}_{CZD} = \tfrac{1}{2} \langle [E_C, E_Z], E_D \rangle_{\mathfrak{G}} \\ &= [\epsilon_{qpi} \partial_q \lambda_{Dp} + 2\lambda_{Zp} \tilde{\epsilon}_{CZD} \epsilon_{pir} A_{Cr} + \partial_{Di} H(A)] E_D \otimes dx^i. \end{aligned} \tag{33}$$

By the definition here of $\tilde{\epsilon}_{CZD}$ and by (77), we have

$$\tilde{\epsilon}_{CZD} = \epsilon_{CZD} \qquad \text{for } \mathfrak{G} = su(2). \tag{34}$$

Now let $n$ be the "outward" unit vector field normal to $\partial\Omega$; then the orientation on $\partial\Omega$ is the one that makes an orthonormal frame $(u_1, u_2)$ on $\partial\Omega$ positively oriented if

$$u_1 \wedge u_2 \wedge n = e_1 \wedge e_2 \wedge e_3,$$

where $(e_1, e_2, e_3)$ is the standard basis of $\mathbb{R}^3$. If $\alpha$ is a 2-form on the surface $\partial\Omega$, then, since $(u_1, u_2, n)$ is orthonormal,

$$\alpha = \alpha(u_1, u_2) u^1 \wedge u^2 = (\alpha \wedge n_q dx^q)(u_1, u_2, n) u^1 \wedge u^2 = (\alpha \wedge n_q dx^q)(e_1, e_2, e_3) u^1 \wedge u^2,$$

8

where $(u^1, u^2)$ is the frame of $T^*(\partial\Omega)$ dual to $(u_1, u_2)$. We will denote $u^1 \wedge u^2$ by $da$, viewed effectively as a surface measure. The expression for $\hat{S}_H[A, \lambda]$ in (22), when expressed in terms of coordinates and components, works out to

$$
\begin{aligned}
\hat{S}_H[A, \lambda] \\
= &-\int_{\partial\Omega} \lambda_{Zp} A_{Zr}^{(b)}[(dx^p \wedge dx^r \wedge n_q dx^q)(e_1, e_2, e_3)]\, da + \int_{\Omega} A_{Zr}(\partial_q \lambda_{Yp})\langle E_Y, E_Z\rangle dx^q \wedge dx^p \wedge dx^r \\
&+ \int_{\Omega} \frac{1}{2}\lambda_{Zp} A_{Bq} A_{Cr}\langle E_Z, [E_B, E_C]\rangle dx^p \wedge dx^q \wedge dx^r + \int_{\Omega} H(A)\, dvol \\
= &-\int_{\partial\Omega} \lambda_{Zp} A_{Zr}^{(b)} n_q \epsilon_{prq}\, da + \int_{\Omega} A_{Zr}(\partial_q \lambda_{Zp})\epsilon_{qpr}\, dx + \frac{1}{2}\int_{\Omega} \lambda_{Zp} A_{Bq} A_{Cr} 2\epsilon_{BCZ}\epsilon_{pqr}\, dx + \int_{\Omega} H(A)\, dx \\
= &\int_{\Omega} [A_{Zr}\epsilon_{rqp}\partial_q \lambda_{Zp} + \lambda_{Zp}\epsilon_{pqr}\epsilon_{ZBC}A_{Bq}A_{Cr} + H(A)]\, dx + \int_{\partial\Omega} -\lambda_{Zp}\epsilon_{prq}A_{Zr}^{(b)} n_q\, da
\end{aligned}
\tag{35}
$$

where we used (77) for the second equality, and $dx$ denotes the volume form $dvol = dx^1 \wedge dx^2 \wedge dx^3$.

### 2.3.6   The dual action $S_H(\lambda)$

We assume now that the function $H$ is such that for each $\lambda \in \mathcal{A}_k$ (or a suitably large subset of it) there is an $A^{(H)}(\lambda) \in \mathcal{A}_k$ for which

$$
* \left(d^{A^{(H)}(\lambda)}\lambda\right) + \nabla H((A^{(H)}(\lambda)) = 0,
\tag{36}
$$

and that the *dual to primal map*

$$
\lambda \mapsto A^{(H)}(\lambda)
$$

is directionally $C^1$ (Definition B.1) as a mapping between Fréchet spaces. Then we have the dual-space action

$$
S_H(\lambda) = \hat{S}\big[A^{(H)}(\lambda), \lambda\big].
\tag{37}
$$

The existence of $A^{(H)}(\lambda)$ satisfying (36) needs to be established for each choice of the function $H$. In the following section, we will work through this for a specific choice for $H$ in which case the dual functional can be written out using the coordinate expression (7) for the Chern-Simons form.

**Theorem 2.1.** *With notation and assumptions as above, $S_H$ has a critical point at $\lambda$ if and only if*

$$
\begin{aligned}
F^{A^{(H)}(\lambda)} &= 0 \\
A^{(H)}(\lambda)|_{\partial\Omega} &= A^{(b)}
\end{aligned}
\tag{38}
$$

Note that there is no boundary constraint on the field $\lambda$, and that the boundary value for $A$ given in the second equation in (38) appears automatically as part of the condition for $\lambda$ to be a critical point of $S_H$.

*Proof.* The function $\hat{S}_H$ is directionally $C^1$ and so, given our assumption that $\lambda \mapsto A^{(H)}(\lambda)$ is also $C^1$, it follows that Hamilton's Fréchet space chain rule [13, Theorem 3.3.4] holds (see (81)).

9

Therefore, we can compute the directional derivative of $S_H$ at $\lambda$ along $\beta$:

$$
\begin{aligned}
\frac{d}{dt}\Big|_{t=0} S_H(\lambda + t\beta) &= \frac{d}{dt}\Big|_{t=0} S\left(A^{(H)}(\lambda + t\beta), \lambda + t\beta\right) \\
&= \partial_1 S\left(A^{(H)}(\lambda), \lambda\right) \Big| \left(A^{(H)}\right)'(\lambda)|\beta + \partial_2 S\left(A^{(H)}(\lambda), \lambda\right)|\beta \\
&= 0 + \partial_2 S\left(A^{(H)}(\lambda), \lambda\right)|\beta \\
&= \int_{\partial\Omega} \left\langle \beta \wedge \left(A^{(H)}(\lambda) - A^{(b)}\right)\right\rangle + \int_\Omega \langle \beta \wedge F^{A^{(H)}(\lambda)}\rangle.
\end{aligned}
\tag{39}
$$

Thus, $\lambda$ is a critical point of $S_H$ if and only if the last line in (39) is 0 for all $\beta$, which is equivalent to the equations (38). $\qquad\square$

## 2.4 Quadratic $H$

In this section we develop the explicit form of $S_H[\lambda]$ for the specific choice of a 'shifted' quadratic form for $H$. The language is that of simple vector calculus. We write field arguments of functionals within square brackets.

As seen in (7), the Chern-Simons action integral for a domain $\Omega \subset \mathbb{R}^3$ is given by:

$$
CS[A] = \int_\Omega 2\epsilon_{pqr}\left(A_{Zp}\partial_q A_{Zr} + \frac{2}{3}\epsilon_{ZBC}A_{Zp}A_{Bq}A_{Cr}\right) dx,
$$

where we assume Dirichlet boundary conditions $A(x) = A^{(b)}(x), x \in \partial\Omega$.

Recall from (13) that the Euler Lagrange equations for the Chern-Simons action integral are given by the flatness condition:

$$
\epsilon_{pqr}\partial_q A_{Zr} + \epsilon_{pqr}\epsilon_{ZBC}A_{Bq}A_{Cr} = 0.
\tag{40}
$$

In addition to $A$, we now introduce $\lambda$, also an $su(2)$-valued 1-form; thus, we work with

$$
A, \lambda: \quad \mathbb{R}^3 \supset \Omega \to \mathbb{R}^{3\times3}
$$

and define the Lagrangian $\mathcal{L}$ of the pre-dual functional as the weak form of (40) with $\lambda$ as a test function, along with an additive shifted quadratic potential for $H$; as well, all boundary conditions of the primal problem are imposed as natural boundary condition terms of the functional:

$$
\begin{aligned}
\widehat{S}[A, \lambda] &= \int_\Omega A_{Zr}\,\epsilon_{rqp}\,\partial_q \lambda_{Zp} + \lambda_{Zp}\,\epsilon_{pqr}\,\epsilon_{ZBC}\,A_{Bq}\,A_{Cr} + \frac{1}{2}k\,|A - \bar{A}|^2\,dx + \int_{\partial\Omega} -\lambda_{Zp}\epsilon_{pqr}A_{Zr}^{(b)}n_q\,da \\
&=: \int_\Omega \mathcal{L}(A, \lambda, \operatorname{curl}\lambda)\,dx + \int_{\partial\Omega} -\lambda : (A^{(b)} \times n)\,da.
\end{aligned}
\tag{41}
$$

Here, $n$ is the outward unit normal of $\partial\Omega$, the cross-product is row-wise, $da$ is the volume form on $\partial\Omega$, and the symbol $:$ represents contraction on two indices. (In more detail, $n_p$ is the unit normal to $T_p\partial\Omega$ such that $p + tn_p \notin \Omega$ for all positive $t$, and $da$ is the Riemannian volume form on $\partial\Omega$.)

Now,

$$
\frac{\partial \mathcal{L}}{\partial A_{Di}} = 0 \iff k\left(\delta_{CD}\,\delta_{ir} + \frac{2}{k}\lambda_{Zp}\,\epsilon_{pir}\,\epsilon_{ZDC}\right)(A_{Cr} - \bar{A}_{Cr}) = -\epsilon_{iqp}\,\partial_q\lambda_{Dp} - 2\lambda_{Zp}\,\epsilon_{pir}\,\epsilon_{ZDC}\,\bar{A}_{Cr}.
$$

So that the DtP mapping is defined through

$$k\,\mathbb{K}_{DiCr}\left(A^{(H)}_{Cr} - \bar{A}_{Cr}\right) = -\epsilon_{iqp}\,\partial_q\lambda_{Dp} - 2\lambda_{Zp}\,\epsilon_{pir}\,\epsilon_{ZDC}\,\bar{A}_{Cr};\qquad \mathbb{K}_{DiCr} := \delta_{DC}\,\delta_{ir} + \frac{2}{k}\lambda_{Zp}\,\epsilon_{pir}\,\epsilon_{ZDC}$$

$$(42)$$

The dual functional can be expressed as, *with the notation* $A^{(H)} = \mathsf{A}$,

$$
\begin{aligned}
S[\lambda] &= \widehat{S}[A^{(H)}, \lambda]\\
&= \int_\Omega dx\; \mathsf{A}_{Zr}\,\epsilon_{rqp}\,\partial_q\lambda_{Zp} + \lambda_{Zp}\,\epsilon_{pqr}\,\epsilon_{ZBC}\,\mathsf{A}_{Bq}\,\mathsf{A}_{Cr} + \frac{1}{2}k\,|\mathsf{A} - \bar{A}|^2 - \int_{\partial\Omega}\lambda : (A^{(b)} \times n)\,da\\
&= \int_\Omega dx\; \epsilon_{rqp}\,\partial_q\lambda_{Zp}\left(\mathsf{A}_{Zr} - \bar{A}_{Zr}\right)\\
&\qquad + \bar{A}_{Zr}\,\epsilon_{rqp}\,\partial_q\lambda_{Zp}\\
&\qquad + \lambda_{Zp}\,\epsilon_{pqr}\,\epsilon_{ZBC}\left(\mathsf{A}_{Bq} - \bar{A}_{Bq}\right)\left(\mathsf{A}_{Cr} - \bar{A}_{Cr}\right)\\
&\qquad + \lambda_{Zp}\,\epsilon_{pqr}\,\epsilon_{ZBC}\left(\mathsf{A}_{Bq} - \bar{A}_{Bq}\right)\bar{A}_{Cr}\\
&\qquad + \lambda_{Zp}\,\epsilon_{pqr}\,\epsilon_{ZBC}\,\bar{A}_{Bq}\left(\mathsf{A}_{Cr} - \bar{A}_{Cr}\right)\\
&\qquad - \lambda_{Zp}\,\epsilon_{pqr}\,\epsilon_{ZBC}\,\bar{A}_{Bq}\,\bar{A}_{Cr}\\
&\qquad + \lambda_{Zp}\,\epsilon_{pqr}\,\epsilon_{ZBC}\,\bar{A}_{Bq}\,\bar{A}_{Cr} + \lambda_{Zp}\,\epsilon_{pqr}\,\epsilon_{ZBC}\,\bar{A}_{Bq}\,\bar{A}_{Cr}\\
&\qquad + \frac{1}{2}k\left(\mathsf{A}_{Zr} - \bar{A}_{Zr}\right)\left(\mathsf{A}_{Zr} - \bar{A}_{Zr}\right)\\
&\qquad - \int_{\partial\Omega}\lambda : (A^{(b)} \times n)\,da.
\end{aligned}
$$

$$(43)$$

The linear terms in $(\mathsf{A} - \bar{A})$ are

$$
\begin{aligned}
&\epsilon_{rqp}\,\partial_q\lambda_{Zp}\left(\mathsf{A}_{Zr} - \bar{A}_{Zr}\right) + \lambda_{Zp}\,\epsilon_{pqr}\,\epsilon_{ZBC}\left(\mathsf{A}_{Bq} - \bar{A}_{Bq}\right)\bar{A}_{Cr} + \lambda_{Zp}\,\epsilon_{pqr}\,\epsilon_{ZBC}\,\bar{A}_{Bq}\left(\mathsf{A}_{Cr} - \bar{A}_{Cr}\right)\\
&= (\mathsf{A}_{Di} - \bar{A}_{Di})\,\epsilon_{iqp}\,\partial_q\lambda_{Dp} + \lambda_{Zp}\,\epsilon_{pir}\,\epsilon_{ZDC}\,\bar{A}_{Cr}(\mathsf{A}_{Di} - \bar{A}_{Di}) + \lambda_{Zp}\,\epsilon_{pri}\,\epsilon_{ZCD}\,\bar{A}_{Cr}(\mathsf{A}_{Di} - \bar{A}_{Di})\\
&= (\mathsf{A}_{Di} - \bar{A}_{Di})\,\epsilon_{iqp}\,\partial_q\lambda_{Dp} + 2\lambda_{Zp}\,\epsilon_{pir}\,\epsilon_{ZDC}\,\bar{A}_{Cr}(\mathsf{A}_{Di} - \bar{A}_{Di})\\
&= -(-\epsilon_{iqp}\,\partial_q\lambda_{Dp} - 2\lambda_{Zp}\,\epsilon_{pir}\,\epsilon_{ZDC}\,\bar{A}_{Cr})(\mathsf{A}_{Di} - \bar{A}_{Di})
\end{aligned}
$$

The quadratic terms in $(\mathsf{A} - \bar{A})$ are

$$
\begin{aligned}
&\lambda_{Zp}\,\epsilon_{pqr}\,\epsilon_{ZBC}\left(\mathsf{A}_{Bq} - \bar{A}_{Bq}\right)\left(\mathsf{A}_{Cr} - \bar{A}_{Cr}\right) + \frac{1}{2}k\left(\mathsf{A}_{Zr} - \bar{A}_{Zr}\right)\left(\mathsf{A}_{Zr} - \bar{A}_{Zr}\right)\\
&= \lambda_{Zp}\,\epsilon_{pir}\,\epsilon_{ZDC}\left(\mathsf{A}_{Di} - \bar{A}_{Di}\right)\left(\mathsf{A}_{Cr} - \bar{A}_{Cr}\right) + \frac{1}{2}k\,\delta_{CD}\,\delta_{ir}\left(\mathsf{A}_{Di} - \bar{A}_{Di}\right)\left(\mathsf{A}_{Cr} - \bar{A}_{Cr}\right)\\
&= \frac{1}{2}k\left(\delta_{CD}\,\delta_{ir} + \frac{2}{k}\lambda_{Zp}\,\epsilon_{pir}\,\epsilon_{ZDC}\right)\left(\mathsf{A}_{Di} - \bar{A}_{Di}\right)\left(\mathsf{A}_{Cr} - \bar{A}_{Cr}\right)\\
&= \frac{1}{2}k\,\mathbb{K}_{DiCr}\left(\mathsf{A}_{Di} - \bar{A}_{Di}\right)\left(\mathsf{A}_{Cr} - \bar{A}_{Cr}\right)\\
&= \frac{1}{2}(-\epsilon_{iqp}\,\partial_q\lambda_{Dp} - 2\lambda_{Zp}\,\epsilon_{pir}\,\epsilon_{ZDC}\,\bar{A}_{Cr})\left(\mathsf{A}_{Di} - \bar{A}_{Di}\right);
\end{aligned}
$$

where we used (42) in the last step. So, the linear plus quadratic terms in $(\mathsf{A} - \bar{A})$ are given by

$$
\begin{aligned}
&\left(-1 + \frac{1}{2}\right)\left(-\epsilon_{iqp}\,\partial_q\lambda_{Dp} - 2\lambda_{Zp}\,\epsilon_{pir}\,\epsilon_{ZDC}\,\bar{A}_{Cr}\right)\left(\mathsf{A}_{Di} - \bar{A}_{Di}\right)\\
&= -\frac{1}{2}\left(-\epsilon_{iqp}\,\partial_q\lambda_{Dp} - 2\lambda_{Zp}\,\epsilon_{pir}\,\epsilon_{ZDC}\,\bar{A}_{Cr}\right)\left(\mathsf{A}_{Di} - \bar{A}_{Di}\right),
\end{aligned}
$$

11

and since
$$\left(\mathsf{A}_{Di} - \bar{A}_{Di}\right) = \frac{1}{k} \left(\mathbb{K}^{-1}\right)_{DiCr} \left(-\epsilon_{rgh} \partial_g \lambda_{Ch} - 2\lambda_{Bj} \epsilon_{jrn} \epsilon_{BCI} \bar{A}_{In}\right) \tag{44}$$

from the DtP mapping, the linear plus quadratic terms become

$$-\frac{1}{2}\left(-\epsilon_{iqp} \partial_q \lambda_{Dp} - 2\lambda_{Zp} \epsilon_{pir} \epsilon_{ZDC} \bar{A}_{Cr}\right) \frac{1}{k} \left(\mathbb{K}^{-1}\right)_{DiCr} \left(-\epsilon_{rgh} \partial_g \lambda_{Ch} - 2\lambda_{Bj} \epsilon_{jrn} \epsilon_{BCI} \bar{A}_{In}\right).$$

The 'constant' terms in $(\mathsf{A} - \bar{A})$ are given by

$$(-1 + 1 + 1)\,\lambda_{Zp}\, \epsilon_{pqr}\, \epsilon_{ZBC}\, \bar{A}_{Bq}\, \bar{A}_{Cr} = \lambda_{Zp}\, \epsilon_{pqr}\, \epsilon_{ZBC}\, \bar{A}_{Bq}\, \bar{A}_{Cr}.$$

Thus, the dual Chern-Simons functional for $H = \frac{1}{2}k|A - \bar{A}|^2$ is given by

$$\mathsf{P}_{Di} := -\epsilon_{iqp} \partial_q \lambda_{Dp} - 2\lambda_{Zp} \epsilon_{pir} \epsilon_{ZDC} \bar{A}_{Cr}$$

$$\mathbb{K}_{DiCr} := \delta_{DC} \delta_{ir} + \frac{2}{k}\lambda_{Zp} \epsilon_{pir} \epsilon_{ZDC}$$

$$S[\lambda] = \int_\Omega \frac{1}{2}\left(-\mathsf{P}_{Di} \frac{1}{k} \left(\mathbb{K}^{-1}\right)_{DiCr} \mathsf{P}_{Cr} + 2\lambda_{Zp} \epsilon_{pqr} \epsilon_{ZBC} \bar{A}_{Bq} \bar{A}_{Cr}\right) dx - \int_{\partial\Omega} \lambda : (A^{(b)} \times n)\, da. \tag{45}$$

We note that if $\bar{A}$ is a solution to (40), then $\lambda = 0$ is a critical point of $S$.

## 3 Variational Analysis

In this section, we present a variational analysis for the E-L equation of the C-S functional based on the Direct Method of the Calculus of Variations for an appropriately defined dual functional related to the dual action functional $S_H$, (37).

Let us first note that the C-S-functional itself is not suitable for applying the Direct Method of the Calculus of Variations in order to establish existence of critical points by proving the existence of a minimizer. Indeed, let $\varphi \in C_c^\infty(\Omega; [0, \infty))$ with $\int_\Omega \varphi(x)^3\, dx > 0$ and $A_{Jk}^t(x) = t\varphi(x)\delta_{Jk}$, for $t \in \mathbb{R}$, then it follows by (7) that:

$$\begin{aligned}
CS(A^t) &= 2 \int_\Omega \epsilon_{pqr}(A_{Jp}^t \partial_q A_{Jr}^t + \frac{2}{3}\epsilon_{JKL} A_{Jp}^t A_{Kq}^t A_{Lr}^t)\, dx \\
&= 2 \int_\Omega \frac{2}{3}t^3\varphi(x)^3 \epsilon_{pqr} \epsilon_{JKL} \delta_{Jp}\delta_{Kq}\delta_{Lr}\, dx \\
&= 8t^3 \underbrace{\int_\Omega \varphi(x)^3\, dx}_{>0}.
\end{aligned}$$

By sending $t \to \pm\infty$ it is evident that the $CS$ functional cannot have a minimizer or a maximizer with finite energy in any function space that includes $C_c^\infty(\Omega; \mathbb{R}^{3\times3})$.

In this section, we consider the dual functional

$$\tilde{S}_H[\lambda] = \sup_{A \in X} \int_\Omega A : \operatorname{curl}\lambda + \lambda_{Zp}\epsilon_{pqr}\epsilon_{ZDC}A_{Dq}A_{Cr} - H(A)\, dx - \int_{\partial\Omega} \lambda : (A^{(b)} \times n)\, da, \tag{46}$$

where $n$ is the outer normal to $\partial\Omega$ and the function space $X$ and the domain of $\tilde{S}_H$ will be chosen such that the integral is well-defined. For regular functions $\lambda$ and $A^{(b)}$ the integral over the

boundary $\partial\Omega$ is with respect to the surface measure on $\partial\Omega$. For less regular functions this term has to be understood in the sense of traces as will be explained once the function spaces for $A$ and $\lambda$ will be chosen for specific choices of $H$.

Let us briefly discuss the connection between (46) and the approach introduced in Section 2. Consider

$$T_{Zp}(A) := \epsilon_{pqr}\epsilon_{ZDC}A_{Dq}A_{Cr}$$
$$B(A,\lambda,n) := -\lambda : (A \times n) = -\lambda_{Zp}\epsilon_{prq}A_{Zr}n_q = \lambda_{Zp}\epsilon_{pqr}A_{Zr}n_q = A_{Zr}\epsilon_{rpq}\lambda_{Zp}n_q =: A : (\lambda \times n)$$

Then, if $A \mapsto H(A) \in \mathbb{R}$ is a strictly convex function with positive definite Hessian everywhere,

- $\sup_\lambda \left( \inf_A \int_\Omega A : \operatorname{curl}\lambda + \lambda : T(A) + H(A)\, dx \quad + \quad \int_{\partial\Omega} B(A^{(b)},\lambda,n)\, da \right)$

- $\inf_\lambda \left( \sup_A \int_\Omega -A : \operatorname{curl}\lambda - \lambda : T(A) - H(A)\, dx \quad - \quad \int_{\partial\Omega} B(A^{(b)},\lambda,n)\, da \right)$

- $\inf_\lambda \left( \sup_A \int_\Omega A : \operatorname{curl}\lambda + \lambda : T(A) - H(A)\, dx \quad + \quad \int_{\partial\Omega} B(A^{(b)},\lambda,n)\, da \right)$

- $\sup_\lambda \left( \inf_A \int_\Omega -A : \operatorname{curl}\lambda - \lambda : T(A) + H(A)\, dx \quad - \quad \int_{\partial\Omega} B(A^{(b)},\lambda,n)\, da \right)$

are four[1] viable formulations of our dual scheme whose corresponding minimizers (or maximizers) can be used to define variational dual solutions (see Definition 3.1 below) for the Euler-Lagrange equations of the Chern-Simons action. In this Section, we analyze the third option above. The functional $S[\lambda]$ defined in Sec. 2.4 is related to the first option above with its maximizer defining weak solutions to the C-S Euler-Lagrange system in the presence of appropriate regularity. The viability arises from the fact that the inner $\sup_A$ ($\inf_A$) may be thought of as formally equivalent to evaluating $\mathcal{L}(A(x), \mathcal{D}(x))$ at its critical point w.r.t $A(x)$ for each fixed $\mathcal{D}(x)$ (the latter generated from an appropriate class of $\lambda$ fields); such a critical point is defined by $A^{(H)}(\mathcal{D}(x))$. Hence, if the dual functional $\tilde{S}$ is regular enough in a minimizer, $x \mapsto A^{(H)}(\mathcal{D}(x))$ will satisfy the Euler-Lagrange equation of $\tilde{S}$ which, by the discussion in Section 2.3, is exactly the Euler-Lagrange equation of the Chern-Simons action functional (40).

In light of the discussion above, we define the notion of a *variational dual solution* and a *dual solution*, c.f. [4].

**Definition 3.1.** *Given a Borel-measurable function $H$ we say that $\lambda$ is a variational dual solution to the Chern-Simons equation if it minimizes the dual functional $\tilde{S}_H$. Moreover, we say that $\lambda$ is a dual solution to the Chern-Simons equation if the dual functional $\tilde{S}_H$ is differentiable at $\lambda$, $\lambda$ satisfies the weak form of the corresponding Euler-Lagrange equation for $\tilde{S}_H$ and there exists a smooth dual-to-primal mapping $\lambda \to A^{(H)}(\lambda)$ corresponding to $\tilde{S}_H$.*

First, by the discussion above every dual solution gives rise to a weak solution of the Chern-Simons equation. The notion of variational dual solutions is a weaker concept and not every variational dual solution is necessarily a dual solution or gives rise to a weak solution of the Chern-Simons equation. However, the choice of the function $H$ does significantly influence this relationship, see also the discussion in [4, Section 5].

In this work, we focus on the existence of variational dual solutions for reasonable choices of $H$, i.e., we show the following:

---

[1]We note that the second statement is the negative of the first and the fourth is the negative of the third. The second statement is equal to the third whenever for each $\lambda$ in the set to which it belongs, $-\lambda$ also belongs to the set.

**Theorem 3.1.** *For $H(A) = \ell|A|^\alpha$, $\ell > 0$, $\alpha \geq 2$, there exists a variational dual solution to the Chern-Simons equation.*

**Remark 3.1.** *It can be seen from the proof that the statement of the theorem stays true for all convex functions $H$ such that $|H(A)| \leq \ell\left(|A|^\alpha + 1\right)$, c.f. also Remark 3.2. Note that this includes the case $1 \leq \alpha < 2$. However, in this case the dual functional is degenerate in the sense that $\tilde{S}_H[\lambda] = +\infty$ for all $\lambda \neq 0$. Hence, a variational dual solution exists but is not meaningful.*

In particular, this shows that the problem of finding an *extremal* point for the Chern-Simons functional can (at least in a weak sense) be replaced by finding a *minimizer* of an appropriate dual functional.

By the definition of $\tilde{S}_H$ it can be checked after fixing appropriate function spaces that $\tilde{S}_H$ is convex and lower semi-continuous with respect to the appropriate weak (or weak*) convergence of $\lambda$ and $\operatorname{curl}\lambda$ in these spaces. This will be made more precise below once the function spaces and $H$ will be fixed. Hence, in order to prove Theorem 3.1, i.e., to establish the existence of varitional dual solutions, using the Direct Method of the Calculus of Variations the key step is to show that $\tilde{S}_H$ is (weakly) coercive on an appropriate function space for $\lambda$. This will be done in Proposition 3.2 (for $\alpha > 2$) and Proposition 3.4 (for $\alpha = 2$). For a more detailed discussion we refer to [4, Section 5].

In order to simplify notation we define for a convex function $H : \mathbb{R}^{3\times3} \to \mathbb{R}$, the function $g_H : \mathbb{R}^{3\times3} \times \mathbb{R}^{3\times3} \to \mathbb{R} \cup \{+\infty\}$ as

$$g_H(\lambda, \mu) = \sup_{A \in \mathbb{R}^{3\times3}} A : \mu + \lambda_{Zp}\epsilon_{pqr}\epsilon_{ZBC}A_{Bq}A_{Cr} - H(A).$$

As in [4] it can then be argued that it holds in the setting discussed below that

$$\tilde{S}_H[\lambda] = \int_\Omega g_H(\lambda, \operatorname{curl}\lambda)\,dx - \int_{\partial\Omega} \lambda : (A^{(b)} \times n)\,da.$$

## 3.1 The case $\alpha > 2$

Let $\infty > \alpha > 2$, $\alpha' := \frac{\alpha}{\alpha-1}$ and $\beta := \left(\frac{\alpha}{2}\right)' = \frac{\alpha}{\alpha-2}$. Note that $\beta \geq \alpha'$ since $\alpha/2 \leq \alpha$.

The case of $\alpha = 4$ is worked out in Appendix C with slightly modified constants as a gentle example of the general case dealt with herein.

In the case of $\alpha$-growth of $H$, the natural space for the primal variables $A$ is $X = L^\alpha(\Omega; \mathbb{R}^{3\times3})$ and, accordingly, $(\lambda, \operatorname{curl}\lambda) \in L^\beta(\Omega; \mathbb{R}^{3\times3}) \times L^{\alpha'}(\Omega; \mathbb{R}^{3\times3})$, where $\operatorname{curl}\lambda$ has to be understood in the sense of distributions. Note that these choices guarantee by Hölder's inequality that the integral

$$\int_\Omega A : \operatorname{curl}\lambda + \lambda_{Zp}\epsilon_{pqr}\epsilon_{ZBC}A_{Bq}A_{Cr} - H(A)\,dx$$

is well-defined.

On the other hand, let $\lambda \in L^\beta$ with $\operatorname{curl}\lambda \in L^{\alpha'}$ and a measurable function $A : \Omega \to \mathbb{R}^{3\times3}$ such that $g_H(\lambda(x), \operatorname{curl}\lambda(x)) = A(x) : \operatorname{curl}\lambda(x) + \lambda_{Zp}(x)\epsilon_{pqr}\epsilon_{ZBC}A_{Bq}(x)A_{Cr}(x) - H(A(x))$. Then the first order optimality condition for $A$ reads as

$$0 = (\operatorname{curl}\lambda(x))_{ij} + \lambda_{Zp}(x)\epsilon_{pjr}\epsilon_{ZiC}A_{Cr}(x) + \lambda_{Zp}(x)\epsilon_{pqj}\epsilon_{ZBi}A_{Bq}(x) - (DH(A(x)))_{ij}.$$

Since $|DH(A(x))| = \alpha\ell|A(x)|^{\alpha-1}$ it follows that

$$|A(x)|^{\alpha-1} \leq C'\left(|\operatorname{curl}\lambda(x)| + |\lambda(x)| \cdot |A(x)|\right),$$

which implies that

$$|A(x)|^\alpha \leq \tilde{C}\left(|\operatorname{curl}\lambda(x)|^{\frac{\alpha}{\alpha-1}} + |\lambda(x)|^{\frac{\alpha}{\alpha-2}}\right) = \tilde{C}\left(|\operatorname{curl}\lambda(x)|^{\alpha'} + |\lambda(x)|^{\beta}\right) \in L^1(\Omega),$$

i.e., $A \in L^\alpha(\Omega; \mathbb{R}^{3\times3})$.

Next, note that for functions $\lambda \in L^\beta(\Omega; \mathbb{R}^{3\times3}) \subseteq L^{\alpha'}(\Omega; \mathbb{R}^{3\times3})$ with $\operatorname{curl}\lambda \in L^{\alpha'}(\Omega; \mathbb{R}^{3\times3})$ there exists a linear, continuous trace operator for the term $\lambda \times n$ into the space $W^{-1+\frac{1}{\alpha},\alpha'}(\partial\Omega; \mathbb{R}^{3\times3})$, the dual of the space $W^{1-\frac{1}{\alpha},\alpha}(\partial\Omega; \mathbb{R}^{3\times3})$ which in turn is the trace space of $W^{1,\alpha}(\Omega, \mathbb{R}^{3\times3})$, c.f. see [14, Theorem 18.40] and [15, Theorem 1, page 204] for the case $\alpha = 2$ (the other cases are similar). In particular, assuming that $A^{(b)} \in W^{1-\frac{1}{\alpha},\alpha}(\partial\Omega; \mathbb{R}^{3\times3})$ (i.e., $A^{(b)}$ is the trace of a function in $W^{1,\alpha}(\Omega; \mathbb{R}^{3\times3})$) then the well-defined interpretation of the integral $\int_{\partial\Omega} A^{(b)} : (\lambda \times n)\, da$ is the dual pairing $\langle A^{(b)}, \lambda \times n\rangle_{W^{1-\frac{1}{\alpha},\alpha}(\partial\Omega;\mathbb{R}^{3\times3}), W^{-1+\frac{1}{\alpha},\alpha'}(\partial\Omega;\mathbb{R}^{3\times3})}$. This expression is continuous with respect to the weak convergences $\lambda_k \rightharpoonup \lambda$ in $L^\beta$ and $\operatorname{curl}\lambda_k \rightharpoonup \operatorname{curl}\lambda$ in $L^{\alpha'}$. Moreover, it holds

$$\left|\langle A^{(b)}, \lambda \times n\rangle_{W^{1-\frac{1}{\alpha},\alpha}, W^{-1+\frac{1}{\alpha},\alpha'}}\right| \leq \|A^{(b)}\|_{W^{1-\frac{1}{\alpha},\alpha}} \|\lambda \times n\|_{W^{-1+\frac{1}{\alpha},\alpha'}}$$

$$\leq C'\|A^{(b)}\|_{W^{1-\frac{1}{\alpha},\alpha}}\left(\|\lambda\|_{L^\beta} + \|\operatorname{curl}\lambda\|_{L^{\alpha'}}\right). \tag{47}$$

It can then be shown that in this setting $\tilde{S}_H$ is lower-semicontinuous with respect to $\lambda_k \rightharpoonup \lambda$ in $L^\beta$ and $\operatorname{curl}\lambda_k \rightharpoonup \operatorname{curl}\lambda$ in $L^{\alpha'}$. Hence, the main step to prove existence of a minimizer for $\tilde{S}_H$ using the Direct Method of the Calculus of Variations is to prove coercivity of $\tilde{S}_H$.

Note that in order to show coercivity of $\tilde{S}_H$ for all $H(A) = \ell|A|^\alpha$, $\ell > 0$, it suffices to consider one specific choice for $\ell$. Indeed, for general $\ell'$ one uses for $\tilde{\ell} = \frac{\ell'}{\ell}$ that

$$\sup_A A : \mu + \lambda_{Zp}\epsilon_{pqr}\epsilon_{ZBC}A_{Bq}A_{Cr} - \ell'|A|^\alpha \tag{48}$$

$$= \sup_A \tilde{\ell}\left(A : (\tilde{\ell}^{-1}\mu) + (\tilde{\ell}^{-1}\lambda_{Zp})\epsilon_{pqr}\epsilon_{ZBC}A_{Bq}A_{Cr} - \ell|A|^\alpha\right). \tag{49}$$

Assuming that coercivity holds for $\tilde{S}_H$ for the specific $\ell > 0$, the above yields also coercivity with respect to $\tilde{\ell}^{-1}\lambda$ and $\tilde{\ell}^{-1}\operatorname{curl}\lambda$ for general $\ell'$ which in turn implies coercivity with respect to $\lambda$ and $\operatorname{curl}\lambda$. For technical reasons that become apparent in the proof, we will consider $\ell = \frac{1}{\alpha \cdot 2^{\alpha/2} \cdot 16 \cdot 3^\beta}(4\alpha')^{-\frac{\alpha}{2-\alpha'}}$.

The main step to show the coercivity of $\tilde{S}_H$ is to prove prove lower bounds on the function $g_H$.

**Proposition 3.1.** *Let* $H(A) = \frac{1}{\alpha \cdot 2^{\alpha/2} \cdot 16 \cdot 3^\beta}(4\alpha')^{-\frac{\alpha}{2-\alpha'}}|A|^\alpha$. *Then there exists* $c > 0$ *such that*

$$g_H(\lambda, \mu) \geq c\left(|\mu|^{\alpha'} + |\lambda|^\beta\right).$$

*Proof.* Fix $\lambda, \mu \in \mathbb{R}^{3\times3}$. We distinguish the following cases.

1. Assume that $|\lambda| \leq \frac{1}{4\alpha'}|\mu|^{2-\alpha'}$. Let $A = \mu|\mu|^{\alpha'-2}$ and write

$$K_{Zp} = \epsilon_{pqr}\epsilon_{ZBC}A_{Bq}A_{Cr} = \epsilon_{ZDC}(A_D \times A_C)_p.$$

First, note by Young's inequality that

$$|K|^2 = 4|A_1 \times A_2|^2 + 4|A_2 \times A_3|^2 + 4|A_3 \times A_1|^2$$

$$\leq 4|A_1|^2|A_2|^2 + 4|A_2|^2|A_3|^2 + 4|A_3|^2|A_1|^2 \leq 4|A_1|^4 + 4|A_2|^4 + 4|A_3|^4 \leq 4|A|^4. \tag{50}$$

15

Since, $\alpha(\alpha' - 1) = \frac{\alpha}{\alpha - 1} = \alpha'$ it follows

$$g_H(\lambda, \mu) \geq \mu : A - 2|\lambda||A|^2 - \frac{1}{\alpha}|A|^\alpha = \frac{1}{\alpha'}|\mu|^{\alpha'} - 2|\lambda||\mu|^{2(\alpha'-1)}$$

$$\geq \frac{1}{2\alpha'}|\mu|^{\alpha'} \geq \frac{1}{4\alpha'}\left(|\mu|^{\alpha'} + |\lambda|^\beta\right).$$

For the last estimate we used that $|\lambda|^\beta \leq \frac{1}{(4\alpha')^\beta}|\mu|^{(2-\alpha')\beta} \leq |\mu|^{\alpha'}$ as $(2 - \alpha')\beta = \frac{\alpha-2}{\alpha-1}\beta = \frac{\alpha}{\alpha-1} = \alpha'$.

2. Assume that $\frac{1}{4 \cdot 3^{\beta/2}}|\lambda|^{\frac{1}{2-\alpha'}} \leq |\mu| \leq (4\alpha')^{\frac{1}{2-\alpha'}}|\lambda|^{\frac{1}{2-\alpha'}}$. For $A = (4|\lambda|)^{-1}\mu$ we find using $\frac{1}{2-\alpha'} = \frac{\alpha-1}{\alpha-2} = \frac{\beta}{\alpha'}$, $\alpha\left(\frac{1}{2-\alpha'} - 1\right) = \frac{\alpha}{\alpha-2} = \beta$ and $\frac{2}{2-\alpha'} - 1 = \frac{\alpha'}{2-\alpha'} = \beta$

$$g_H(\lambda, \mu) \geq \mu : A - 2|\lambda||A|^2 - \frac{1}{16 \cdot 3^\beta}\left(4\alpha'\right)^{-\frac{\alpha}{2-\alpha'}}|A|^\alpha$$

$$= \frac{1}{4}|\lambda|^{-1}|\mu|^2 - \frac{1}{8}|\lambda|^{-1}|\mu|^2 - \frac{1}{4^\alpha \cdot 16 \cdot 3^\beta}\left(4\alpha'\right)^{-\frac{\alpha}{2-\alpha'}}|\lambda|^{-\alpha}|\mu|^\alpha$$

$$\geq \frac{1}{8 \cdot 16 \cdot 3^\beta}|\lambda|^{-1+\frac{2}{2-\alpha'}} - \frac{1}{16 \cdot 16 \cdot 3^\beta}\left(4\alpha'\right)^{-\frac{\alpha}{2-\alpha'}}|\lambda|^{-\alpha}|\mu|^\alpha$$

$$\geq \frac{1}{8 \cdot 16 \cdot 3^\beta}|\lambda|^\beta - \frac{1}{16 \cdot 16 \cdot 3^\beta}\left(4\alpha'\right)^{-\frac{\alpha}{2-\alpha'}}\left(4\alpha'\right)^{\frac{\alpha}{2-\alpha'}}|\lambda|^{\alpha(-1+\frac{1}{2-\alpha'})}$$

$$= \frac{1}{8 \cdot 16 \cdot 3^\beta}|\lambda|^\beta - \frac{1}{16 \cdot 16 \cdot 3^\beta}|\lambda|^\beta$$

$$= \frac{1}{16 \cdot 16 \cdot 3^\beta}|\lambda|^\beta \geq \frac{1}{32 \cdot 16 \cdot 3^\beta}\left(4\alpha'\right)^{-\frac{\alpha'}{2-\alpha'}}\left(|\lambda|^\beta + |\mu|^{\alpha'}\right).$$

In the last estimate we used that $|\mu|^{\alpha'} \leq (4\alpha')^{\frac{\alpha'}{2-\alpha'}}|\lambda|^\beta$.

3. Assume that $|\mu| \leq \frac{1}{4 \cdot 3^{\beta/2}}|\lambda|^{\frac{1}{2-\alpha'}}$. Let $V \in \{1, 2, 3\}$ such that it holds for the $V^{th}$ row, $\Lambda_V \in \mathbb{R}^3$, of $\lambda \in \mathbb{R}^{3\times3}$ that $m^2 := \lambda_{Vi}\lambda_{Vi}$ (no sum on $V$) $= |\Lambda_V|^2 \geq \frac{1}{3}|\lambda|^2$, $m > 0$. Choose an orthonormal basis $(E_1, E_2, E_3)$ with $E_V = \frac{1}{m}\Lambda_V$ and $\frac{1}{2}\epsilon_{VYZ}E_Y \times E_Z = E_V$. Now, define $A \in \mathbb{R}^{3\times3}$ row-wise as

$$A_{Yi} = m^{(\beta-1)/2} E_{Yi} \quad \text{if } Y \neq V \qquad \text{and} \qquad A_{Vi} = 0.$$

It follows that $|A| = \sqrt{2}m^{(\beta-1)/2} = \sqrt{2}|\Lambda_V|^{(\beta-1)/2}$ and

$$\epsilon_{pqr}\epsilon_{ZBC}A_{Bq}A_{Cr} = \begin{cases} 2m^{\beta-1} E_{Vp} = 2\lambda_{Vp}|\Lambda_V|^{\beta-2} & \text{if } Z = V, \\ 0 & \text{otherwise.} \end{cases} \tag{51}$$

Then we estimate using $\alpha\frac{\beta-1}{2} = \beta$ and $\frac{\beta-1}{2} + \frac{1}{2-\alpha'} = \frac{1}{\alpha-2} + \frac{\alpha-1}{\alpha-2} = \frac{\alpha}{\alpha-2} = \beta$

$$g_H(\lambda, \mu) \geq -|\mu||A| + 2|\Lambda_V|^\beta - \frac{1}{2^{\alpha/2}}|A|^\alpha \geq -2^{1/2}|\mu||\Lambda_V|^{(\beta-1)/2} + |\Lambda_V|^\beta$$

$$\geq -2|\mu||\lambda|^{(\beta-1)/2} + \frac{1}{3^{\beta/2}}|\lambda|^\beta \geq \frac{1}{2 \cdot 3^{\beta/2}}|\lambda|^\beta \geq \frac{1}{4 \cdot 3^{\beta/2}}\left(|\lambda|^\beta + |\mu|^{\alpha'}\right).$$

$\square$

**Proposition 3.2.** *For $H(A) = \frac{1}{\alpha \cdot 2^{\alpha/2} \cdot 16 \cdot 3^\beta} (4\alpha')^{-\frac{\alpha}{2-\alpha'}} |A|^\alpha$ and $A^{(b)} \in W^{1-\frac{1}{\alpha},\alpha}(\partial\Omega; \mathbb{R}^{3\times 3})$ the dual functional $\tilde{S}_H$ is weakly coercive on the Banach space $\mathcal{A} := \{\lambda \in L^\beta(\Omega; \mathbb{R}^{3\times 3}) : \mathrm{curl}(\lambda) \in L^{\alpha'}(\Omega; \mathbb{R}^{3\times 3})\}$, i.e., for every sequence $(\lambda_k)_k \subseteq \mathcal{A}$ such that $\sup_k \tilde{S}_H[\lambda_k] < \infty$ there exists a (not relabeled) subsequence and $\lambda \in \mathcal{A}$ such that $\lambda_k \rightharpoonup \lambda$ in $L^\beta$ and $\mathrm{curl}(\lambda_k) \rightharpoonup \mathrm{curl}(\lambda)$ in $L^{\alpha'}$.*

*Proof.* By Proposition 3.1 and the trace estimate (47) it follows for a sequence $(\lambda_k)_k \subseteq \mathcal{A}$ satisfying $\sup_k \tilde{S}_H[\lambda_k] \leq C < \infty$ for all $k \in \mathbb{N}$ that

$$c(\|\lambda_k\|_{L^\beta}^\beta + \|\mathrm{curl}\,\lambda_k\|_{L^{\alpha'}}^{\alpha'}) - C'\|A^{(b)}\|_{W^{1-\frac{1}{\alpha},\alpha}} \left(\|\lambda_k\|_{L^\beta} + \|\mathrm{curl}\,\lambda_k\|_{L^{\alpha'}}\right) \leq \tilde{S}_H[\lambda_k] \leq C.$$

Since $\lim_{\rho,\sigma \to +\infty} c(\rho^\beta + \sigma^{\alpha'}) - C'\|A^{(b)}\|_{W^{1-\frac{1}{\alpha},\alpha}} (\rho + \sigma) = +\infty$, it follows that the terms $\|\lambda_k\|_{L^\beta}$ and $\|\mathrm{curl}(\lambda_k)\|_{L^{\alpha'}}$ are uniformly bounded. By the usual compactness arguments this implies the existence of a (not relabeled) subsequence and functions $\lambda \in L^\beta(\Omega; \mathbb{R}^{3\times 3})$ and $\mu \in L^{\alpha'}(\Omega; \mathbb{R}^{3\times 3})$ such that $\lambda_k \rightharpoonup \lambda$ in $L^\beta$ and $\mathrm{curl}(\lambda_k) \rightharpoonup \mu$ in $L^{\alpha'}$. Eventually, it can then be checked that actually $\mathrm{curl}(\lambda) = \mu$. $\square$

Note that by similar arguments one can show directly from Proposition 3.1 and (47) that the dual functional $\tilde{S}_H$ is bounded from below.

**Remark 3.2.** *Note that for $\tilde{H}(A) = \ell|A|^\alpha + \tilde{\ell}|A|^2$, $\ell, \tilde{\ell} > 0$, $\alpha > 2$, the functional $\tilde{S}_{\tilde{H}}$ is also coercive on the space $\{\lambda \in L^\beta(\Omega; \mathbb{R}^{3\times 3}) : \mathrm{curl}(\lambda) \in L^{\alpha'}(\Omega; \mathbb{R}^{3\times 3})\}$. Indeed, since $\ell|A|^\alpha + \tilde{\ell}|A|^2 \leq (\ell+\tilde{\ell})|A|^\alpha + \tilde{\ell} =: H(A) + \tilde{\ell}$ it follows that $\tilde{S}_{\tilde{H}} \geq \tilde{S}_H - \tilde{\ell}|\Omega|$. Hence, the coercivity of $\tilde{S}_{\tilde{H}}$ follows from the shown coercvity of $\tilde{S}_H$. The advantage of using $\tilde{H}$ might be that its Hessian is strictly positive definite which facilitates that the DtP-mapping can at least locally be well-defined in a smooth way.*

## 3.2 Quadratic $H$

In this section we consider the quadratic function $H(A) = \ell|A|^2$. With the same argument as in the case $\alpha > 2$ it suffices to consider $\ell = \frac{1}{2}$.

In this setting, the natural space for the primal functions $A$ is $X = L^2(\Omega; \mathbb{R}^{3\times 3})$. Accordingly, for $(\lambda, \mathrm{curl}\,\lambda) \in L^\infty(\Omega; \mathbb{R}^{3\times 3}) \times L^2(\Omega; \mathbb{R}^{3\times 3})$ the integral expression $\int_\Omega A : \mathrm{curl}\,\lambda + \lambda_{Zp}\epsilon_{pqr}\epsilon_{ZBC}A_{Bq}A_{Cr} - H(A)\,dx$ is well-defined. Similarly to the setting in Section 2.4, there exists a continuous, linear trace operator from $\{\lambda \in L^2(\Omega; \mathbb{R}^{3\times 3}) : \mathrm{curl}\,\lambda \in L^2(\Omega; \mathbb{R}^{3\times 3})\}$ for the term $\lambda \times n$ into the space $W^{-1/2,2}(\partial\Omega; \mathbb{R}^{3\times 3})$ which is the dual of $W^{1/2,2}(\partial\Omega; \mathbb{R}^{3\times 3})$ (the trace space of $W^{1,2}(\Omega; \mathbb{R}^{3\times 3})$). Hence, the boundary term in (46) has a well-defined interpretation as the dual pairing $\langle A^{(b)}, \lambda \times n\rangle_{W^{1/2,2},W^{-1/2,2}}$ if $A^{(b)} \in W^{1/2,2}(\partial\Omega; \mathbb{R}^{3\times})$ and it holds

$$\left|\langle A^{(b)}, \lambda \times n\rangle_{W^{1/2,2},W^{-1/2,2}}\right| \leq C'\|A^{(b)}\|_{W^{1/2,2}}(\|\lambda\|_{L^\infty} + \|\mathrm{curl}\,\lambda\|_{L^2}). \tag{52}$$

Again, in this setting it can be checked that $\tilde{S}_H$ is lower-semicontinuous with respect to $\lambda_k \overset{*}{\rightharpoonup} \lambda$ in $L^\infty(\Omega; \mathbb{R}^{3\times 3})$ and $\mathrm{curl}\,\lambda_k \rightharpoonup \lambda$ in $L^2(\Omega; \mathbb{R}^{3\times 3})$. Hence, in the following we concentrate on showing coercivity of $\tilde{S}_H$ with respect to this convergence. Again, the key will be to show lower bounds for the function $g_H$.

**Proposition 3.3.** *Let $H(A) = \frac{1}{2}|A|^2$. Then there exists $c > 0$ such that*

$$g_H(\lambda, \mu) = +\infty \text{ if } |\lambda| > \frac{3}{2} \text{ and } g_H(\lambda, \mu) \geq c|\mu|^2 \text{ if } |\lambda| \leq \frac{3}{2}.$$

17

*Proof.* First, let $|\lambda| > \frac{3}{2}$ and consider $A \in \mathbb{R}^{3\times3}$ as in case 3 of the proof of Proposition 3.1, i.e. let $\Lambda_V$ be a row of $\lambda$ such that $|\Lambda_V|^2 \geq \frac{1}{3}|\lambda|^2$ and $A \in \mathbb{R}^{3\times3}$ such that $|A| = 2^{1/2}|\Lambda_V|^{1/2}$ and, as in (51) with $\beta = 2$,

$$\epsilon_{pqr}\epsilon_{ZBC}A_{Bq}A_{Cr} = \begin{cases} 2\lambda_{Vp} & \text{if } Z = V, \\ 0 & \text{otherwise.} \end{cases}$$

Then we obtain for $t > 0$ and the matrix $\tilde{A} = tA$

$$g_H(\lambda, \mu) \geq -|\mu||\tilde{A}| + \lambda_{Zp}\epsilon_{pqr}\epsilon_{ZBC}A_{Bq}A_{Cr} - \frac{1}{2}|\tilde{A}|^2$$
$$= -2^{1/2}t|\mu||\Lambda_V|^{1/2} + 2t^2|\Lambda_V|^2 - t^2|\Lambda_V|$$
$$\geq -2^{1/2}t|\mu||\lambda|^{1/2} + \frac{2}{3}t^2|\lambda|^2 - t^2|\lambda| \xrightarrow{t\to\infty} +\infty.$$

Next, let $|\lambda| \leq \frac{3}{2}$. Then we obtain using (50) for $A = \frac{1}{7}\mu$

$$g_H(\lambda, \mu) \geq \mu : A - 2|A|^2|\lambda| - \frac{1}{2}|A|^2 \geq \mu : A - \frac{7}{2}|A|^2 = \frac{1}{14}|\mu|^2.$$

$\square$

Using this lower bound we show the coercivity of $\tilde{S}_H$.

**Proposition 3.4.** *For $H(A) = \frac{1}{2}|A|^2$ and $A^{(b)} \in W^{1/2,2}(\partial\Omega; \mathbb{R}^{3\times3})$ the dual functional $\tilde{S}_H$ is weakly(\*) coercive on the Banach space $\{\lambda \in L^\infty(\Omega; \mathbb{R}^{3\times3}) : \mathrm{curl}(\lambda) \in L^2(\Omega; \mathbb{R}^{3\times3})\}$, i.e. for all sequences $(\lambda_n, \mathrm{curl}(\lambda_n))_n \subseteq L^\infty(\Omega; \mathbb{R}^{3\times3}) \times L^2(\Omega; \mathbb{R}^{3\times3})$ such that $\sup_n \tilde{S}_H[\lambda_n] < \infty$ there exists a (not relabeled) subsequence and $\lambda \in L^\infty(\Omega; \mathbb{R}^{3\times3})$ with $\mathrm{curl}(\lambda) \in L^2(\Omega; \mathbb{R}^{3\times3})$ such that $\lambda_n \xrightarrow{*} \lambda$ in $L^\infty(\Omega; \mathbb{R}^{3\times3})$ and $\mathrm{curl}(\lambda_n) \rightharpoonup \mathrm{curl}(\lambda)$ in $L^2(\Omega; \mathbb{R}^{2\times2})$.*

*Proof.* The result follows similarly to Proposition 3.2 using Proposition 3.3 and (52). $\square$

We close this section by briefly commenting on the case of a slightly modified quadratic function $\tilde{H}$ around a base state $\bar{A}$.

**Remark 3.3.** *Let $\bar{A} \in L^2(\Omega; \mathbb{R}^{3\times3})$ and $\tilde{H}(x, A) = \frac{1}{2}|A - \bar{A}(x)|^2$. Then we can rewrite for every fixed $x \in \Omega$ and $\mu, \lambda \in \mathbb{R}^{3\times3}$*

$$g_{\tilde{H}(x,\cdot)}(\lambda, \mu) := \sup_{A\in\mathbb{R}^{3\times3}} \mu : A + \lambda_{Zp}\epsilon_{pqr}\epsilon_{ZBC}A_{Bq}A_{Cr} - \tilde{H}(x, A)$$
$$= \sup_{A\in\mathbb{R}^{3\times3}} (\mu + \bar{A}(x)) : A + \lambda_{Zp}\epsilon_{pqr}\epsilon_{ZBC}A_{Bq}(x)A_{Cr}(x) - H(A) - \frac{1}{2}|\bar{A}(x)|^2.$$

*Consequently, by the result above, we obtain*

$$g_{\tilde{H}(x,\cdot)}(\lambda, \mu) = +\infty \text{ if } |\lambda| > \frac{3}{2} \text{ and } g_{\tilde{H}(x,\cdot)}(\mu, \lambda) \geq c|\mu + \bar{A}(x)|^2 - \frac{1}{2}|\bar{A}(x)|^2 \text{ if } |\lambda| \leq \frac{3}{2}.$$

*It is straightforward to check that this estimate suffices to obtain coercivity of the dual functional for a modified quadratic function $\tilde{H}$ including a base state $\bar{A} \in L^2(\Omega; \mathbb{R}^{3\times3})$.*

# A   Appendix: Lie-algebra valued forms

We summarize here some notation and essential algebraic facts about differential forms with values in vector spaces and Lie algebras.

Let $A$ be a $k$-form on a manifold $M$, with values in a vector space $V$, and $B$ an $l$-form on $M$ with values in a vector space $W$. Then, in terms of local coordinates $(x^1, \ldots, x^n)$, we define

$$A \wedge B \stackrel{\text{def}}{=} A_{i_1 \ldots i_k} \otimes B_{j_1, \ldots, i_l} dx^{i_1} \wedge \ldots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \ldots \wedge dx^{j_l}, \tag{53}$$

which is a $(k + l)$-form, with values in the vector space $V \otimes W$. (If $V$ and $W$ are equal, both being a vector space of matrices, it is more common to apply the matrix product to $V \otimes V$ and redefine $A \wedge B$ as a matrix-valued form.) This extends to wedge products of more forms in the natural way. If $\Phi : V \to Z$ is a linear mapping, where $Z$ is also a vector space, then, for a $V$-valued $k$-form $A$ as above, we define the $Z$-valued $k$-form $\Phi(A)$ by:

$$\Phi(A) = \Phi(A_{i_1 \ldots i_k}) dx^{i_1} \wedge \ldots \wedge dx^{i_k}. \tag{54}$$

A bilinear map

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R} : (u, v) \mapsto \langle u, v \rangle$$

induces a linear map

$$\langle \, \cdot \, \rangle : V \otimes V \to \mathbb{R} : u_i \otimes v_i \mapsto \langle u_i, v_i \rangle.$$

Now let $A$ be a $V$-valued $k$-form and $B$ a $V$-valued $l$-form; then, as in (53), $A \wedge B$ is a $V \otimes V$-valued $(k + l)$-form, and so, we can follow the recipe in (54) to obtain

$$\langle A \wedge B \rangle_{k+l} \stackrel{\text{def}}{=} \langle A \wedge B \rangle \stackrel{\text{def}}{=} \langle A_{i_1 \ldots i_k}, B_{j_1, \ldots, i_l} \rangle dx^{i_1} \wedge \ldots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \ldots \wedge dx^{j_l}, \tag{55}$$

if $A$ and $B$ are both $V$-valued; note that, here $\langle A \wedge B \rangle$ is an ordinary $(k + l)$-form, not an inner-product of $A$ and $B$. *We will often drop the subscript $k + l$ in $\langle \, \cdot \, \rangle_{k+l}$ when it is clear from the context.* If the bilinear form $\langle \cdot, \cdot \rangle$ is symmetric, then

$$\langle A \wedge B \rangle = (-1)^{kl} \langle B \wedge A \rangle. \tag{56}$$

Now let $\mathfrak{G}$ be a Lie algebra and

$$\mathfrak{G} \otimes \mathfrak{G} \to \mathfrak{G} : v \otimes w \mapsto [v, w] \tag{57}$$

the linear map arising from the Lie bracket on $\mathfrak{G}$. Then for $A$ a $\mathfrak{G}$-valued $k$-form and $B$ a $\mathfrak{G}$-valued $l$-form, we have the $\mathfrak{G}$-valued $(k + l)$-form

$$[A \wedge B] \stackrel{\text{def}}{=} [A_{i_1 \ldots i_k}, B_{j_1, \ldots, i_l}] dx^{i_1} \wedge \ldots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \ldots \wedge dx^{j_l}. \tag{58}$$

Assume that $\langle \cdot, \cdot \rangle$ is a symmetric, bilinar form on $\mathfrak{G}$. We assume, further, that the following *ad-invariance* relations hold:

$$\langle [u, v], w \rangle = -\langle v, [u, w] \rangle, \tag{59}$$

for all $u, v, w \in \mathfrak{G}$. If $\mathfrak{G}$ is a matrix Lie algebra, then the pairing $(u, v) \mapsto \text{Tr}(uv)$ satisfies this condition, as is seen from (76). Symmetry of $\langle \cdot, \cdot \rangle$ and the relation (59) imply the cyclic relations:

$$\langle [u, v], w \rangle = \langle [w, u], v \rangle = \langle [v, w], u \rangle \tag{60}$$

**Lemma A.1.** *Let $A$ be a $k-$form, $B$ an $m$-form, and $C$ and $n$-form, all defined on a manifold, with values in a Lie algebra $\mathfrak{G}$, which is equipped with a symmetric, bilinear pairing $\langle \cdot, \cdot \rangle$ on $\mathfrak{G} \times \mathfrak{G}$, satisfying the ad-invariance condition (59). Then*

$$[A \wedge B] = -(-1)^{km}[B, A] \tag{61}$$

*and*

$$\langle A \wedge [B \wedge C] \rangle_{k+m+n} = (-1)^{(k+m)n} \langle C \wedge [A \wedge B] \rangle_{k+m+n} = (-1)^{(n+m)k} \langle B \wedge [C \wedge A] \rangle_{k+m+n}. \tag{62}$$

*If $\mathfrak{G}$ is a matrix Lie algebra, then, for any $\mathfrak{G}$-valued 1-form $A$, we have*

$$[A \wedge A] = 2\Pi(A \wedge A), \tag{63}$$

*where $\Pi : \mathfrak{G} \otimes \mathfrak{G} \to \mathfrak{G} : u \otimes v \mapsto uv$, the matrix product.*

As a rule, for matrix-valued forms, we drop $\Pi$ and simply write $A \wedge A$ and $A \wedge A \wedge A$. The utility of (62) is that it allows us to move each of the terms $A$, $B$, $C$ to the left.

*Proof.* Equations (61) and (62) are readily verified. Next, for any $\mathfrak{G}$-valued 1-form $A$, we have

$$\begin{aligned}
[A \wedge A] &= [A_i, A_j]dx^i \wedge dx^j \\
&= A_i A_j dx^i \wedge dx^j - A_j A_i dx^i \wedge dx^j \\
&= A_i A_j dx^i \wedge dx^j + A_j A_i dx^j \wedge dx^i \\
&= 2A_i A_j dx^i \wedge dx^j \\
&= 2A \wedge A,
\end{aligned} \tag{64}$$

where we have dropped $\Pi$. $\qquad\square$

The *curvature* form $F^A$, associated to $A$, is defined by

$$F^A = A + \frac{1}{2}[A \wedge A]. \tag{65}$$

When $A$ is matrix-valued, we can then write this as

$$F^A = A + A \wedge A. \tag{66}$$

## A.1 The group $SU(2)$

Some of our results are worked out for the group $SU(2)$. Here we summarize some basic facts and notation about $SU(2)$ and its Lie algebra $su(2)$.

The matrix group $SU(2)$ is, explicitly, given by

$$SU(2) = \left\{ \begin{pmatrix} x_1 + ix_2 & -x_3 + ix_4 \\ x_3 + ix_4 & x_1 - ix_2 \end{pmatrix} : x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1 \right\}. \tag{67}$$

Thus, $SU(2)$ is essentially the unit sphere in $\mathbb{R}^4$, with the identity matrix $I$ corresponding to the point $(1, 0, 0, 0)$. The Lie algebra $su(2)$ is the tangent space $T_I SU(2)$ and is thus

$$su(2) = T_I SU(2) = \left\{ i \begin{bmatrix} y_3 & y_1 - iy_2 \\ y_1 + iy_2 & -y_3 \end{bmatrix} : y_1, y_2, y_3 \in \mathbb{R}^3 \right\}. \tag{68}$$

A basis of $su(2)$ is given by $i\sigma_1, i\sigma_2, i\sigma_3$, where we are using the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{69}$$

Now consider the inner-product on complex matrices given by

$$\langle A, B \rangle = \frac{1}{N} \mathrm{Tr}(A^* B) = \frac{1}{N} \sum_{j,k=1}^{N} A_{jk}^* B_{jk}, \tag{70}$$

the scaling by $1/N$ is to normalize $\langle I, I \rangle$ to have the value 1. The matrices

$$E_1 = i\sigma_1, \ E_2 = i\sigma_2, \ E_3 = i\sigma_3$$

form a basis of $su(2)$. We note also that since each element of $su(2)$ is skew-hermitian, the inner-product (70) reduces on $su(2)$ to

$$\langle A, B \rangle = -\frac{1}{2} \mathrm{Tr}(AB). \tag{71}$$

There are the standard identities:

$$E_1^2 = E_2^2 = E_3^2 = -I$$
$$E_1 E_2 = -E_2 E_1 = -E_3, \ E_3 E_1 = -E_1 E_3 = -E_2, \ E_2 E_3 = -E_3 E_2 = -E_1, \tag{72}$$

which imply the commutation relations:

$$[E_1, E_2] = -2E_3, \ [E_3, E_1] = -2E_2, \ [E_2, E_3] = -2E_1. \tag{73}$$

Orthonormality of the basis $\{E_J\}$ of $su(2)$ arises from the relations:

$$-\mathrm{Tr}(E_J E_K) = 2\delta_{JK}. \tag{74}$$

We have then

$$-\mathrm{Tr}\big(E_J[E_K, E_L]\big) = -\mathrm{Tr}\big(E_L[E_J, E_K]\big) = -\mathrm{Tr}\big(E_K[E_L, E_J]\big) = 4\epsilon_{JKL}. \tag{75}$$

In fact, for any square matrices $X, Y, Z$, all of the same size, we have

$$\mathrm{Tr}(X[Y, Z]) = \mathrm{Tr}(Z[X, Y]) = \mathrm{Tr}(Y[Z, X]). \tag{76}$$

We note that, by definition of the inner-product (71), equation (75) implies:

$$\langle [E_J, E_K], E_L \rangle = 2\epsilon_{JKL}. \tag{77}$$

# B    Appendix: Function space calculus

We summarize some notions and results on Fréchet spaces, using the work of Hamilton [13].

A *Fréchet space* is a real or complex vector space $X$, equipped with a topology arising from a sequence of seminorms $\{\|\cdot\|_n\}_{n\geq 0}$, such that the topology is Hausforff and complete (every Cauchy sequence converges). As an example ([13, Example 1.1.5]), for a vector bundle $E$ over a compact

Riemannian manifold $M$, the vector space $C^\infty(M, V)$ of all $C^\infty$ sections $M \to V$ is a Fréchet space with respect to the topology induced by the family of norms $\|\cdot\|_n$ given by

$$\|f\|_n = \sum_{k=0}^{n} \sup_{p \in M} |D^k f(p)|,$$

where $D^k f$ is the $k$-th covariant derivative of $f$. The larger space $C^k(M, V)$, for finite $k$, is a Banach space with the norm $\|\cdot\|_k$.

In the sequel, $X$, $Y$, and $Z$ are Fréchet spaces.

**Definition B.1.** *For a function $F : U \to Y$, where $U$ is an open subset of $X$, and both $X$ and $Y$ are Fréchet spaces, the* directional derivative $F'(x)h$ or $DF|_x h$ *is defined by*

$$F'(x)h = DF|_x h = \lim_{t \to 0} \frac{1}{t} \left[ F(x + th) - F(x) \right]. \tag{78}$$

*We will say that $F$ is* directionally $C^1$ *on $U$ if $F'(x)h$ exists for all $(x, h) \in U \times X$, and $(x, h) \mapsto F'(x)h$ is continuous.*

According to [13, Theorem 3.2.5], if $F$ is directionally $C^1$ then $h \mapsto F'(x)h$ is linear for all $x \in U$.

By [13, Lemma 3.3.1], a map $F : U \to Y$, where $U$ is a convex open subset of $X$, is directionally $C^1$ if and only if there is a continuous map $L : U \times U \times X \to Y$, with $h \mapsto L(f_0, f_1)h$ being linear for all $f_0, f_1 \in U$ such that

$$F(f_1) - F(f_0) = L(f_0, f_1)(f_1 - f_0). \tag{79}$$

If this holds, then

$$F'(f)h = L(f, f)h \tag{80}$$

for all $f \in U$ and $h \in X$. The proof is obtained by writing $F(f_1) - F(f_0)$ as the integral $\int_0^1 \ldots dt$ of the derivative of the function $[0, 1] \to Y : t \mapsto F'((1 - t)f_0 + t f_1)h$. From this follows Hamilton's Fréchet space *chain rule*[13, Theorem 3.3.4]:

$$D(G \circ F)|_f h = DG|_{F(f)} DF|_f h \tag{81}$$

whenever $F : U \to Y$ and $G : V \to Z$ are composable Fréchet $C^1$ functions, where $U$ is an open subset of $X$ and $V$ is an open subset of $Y$m with $f \in U$ and $h \in X$.

# C   Appendix: Coercivity for Quartic $H$

Consider the quartic function $H(A) = \ell |A|^4$ with $\ell = \left(\frac{3}{8}\right)^6 \frac{1}{4 \cdot 18}$.

In this setting, the natural space for the primal variables $A$ is $X = L^4(\Omega; \mathbb{R}^{3 \times 3})$ and, accordingly, $(\lambda, \operatorname{curl} \lambda) \in L^2(\Omega; \mathbb{R}^{3 \times 3}) \times L^{4/3}(\Omega; \mathbb{R}^{3 \times 3})$, where $\operatorname{curl} \lambda$ has to be understood in the sense of distributions. Note that these choices guarantee by Hölder's inequality that the integral

$$\int_\Omega A : \operatorname{curl} \lambda + \lambda_{Zp} \epsilon_{pqr} \epsilon_{ZBC} A_{Bq} A_{Cr} - H(A) \, dx$$

is well-defined. Next, note that for functions $\lambda \in L^2(\Omega; \mathbb{R}^{3 \times 3}) \subseteq L^{4/3}(\Omega; \mathbb{R}^{3 \times 3})$ with $\operatorname{curl} \lambda \in L^{4/3}(\Omega; \mathbb{R}^{3 \times 3})$ there exists a linear, continuous trace operator for the term $\lambda \times n$ into the space $W^{-3/4,4}(\partial\Omega; \mathbb{R}^{3 \times 3})$, the dual of the space $W^{3/4,4}(\partial\Omega; \mathbb{R}^{3 \times 3})$ which in turn is the trace space of

$W^{1,4}(\Omega,\mathbb{R}^{3\times3})$, c.f. see [14, Theorem 18.40] and [15, Theorem 1, page 204]. In particular, assuming that $A^{(b)} \in W^{3/4,4}(\partial\Omega;\mathbb{R}^{3\times3})$ (i.e., $A^{(b)}$ is the trace of a function in $W^{1,4}(\Omega;\mathbb{R}^{3\times3})$) then the well-defined interpretation of the integral $\int_{\partial\Omega} A^{(b)} : (\lambda \times n)\, da$ is the dual pairing $\langle A^{(b)}, \lambda \times n\rangle_{W^{3/4,4},W^{-3/4,4}}$. This expression is continuous with respect to the weak convergences $\lambda_k \rightharpoonup \lambda$ in $L^4$ and $\operatorname{curl}\lambda_k \rightharpoonup \operatorname{curl}\lambda$ in $L^{4/3}$. Moreover, it holds

$$\left|\langle A^{(b)}, \lambda \times n\rangle_{W^{3/4,4},W^{-3/4,4}}\right| \leq \|A^{(b)}\|_{W^{3/4,4}}\|\lambda \times n\|_{W^{-3/4,4}} \leq C'\|A^{(b)}\|_{W^{3/4,4}}\left(\|\lambda\|_{L^2} + \|\operatorname{curl}\lambda\|_{L^{4/3}}\right).\tag{82}$$

It can then be shown that in this setting $\tilde{S}_H$ is lower-semicontinuous with respect to $\lambda_k \rightharpoonup \lambda$ in $L^4$ and $\operatorname{curl}\lambda_k \to \operatorname{curl}\lambda$ in $L^{4/3}$. Hence, the main step to prove existence of a minimizer for $\tilde{S}_H$ using the Direct Method of the Calculus of Variations is to prove coercivity of $\tilde{S}_H$. In order to show the coercivity of $\tilde{S}_H$ we first prove lower bounds on the function $g_H$.

**Proposition C.1.** *Let* $H(A) = \left(\frac{3}{8}\right)^6 \frac{1}{4\cdot18}|A|^4$. *Then there exists* $c > 0$ *such that*

$$g_H(\lambda,\mu) \geq c\left(|\mu|^{4/3} + |\lambda|^2\right).$$

*Proof.* Fix $\lambda, \mu \in \mathbb{R}^{3\times3}$. We distinguish the following cases.

1. Assume that $|\lambda| \leq \frac{3}{8}|\mu|^{2/3}$. Let $A = \mu|\mu|^{-2/3}$. Then using (50) it follows

$$g_H(\lambda,\mu) \geq \mu : A - 2|\lambda||A|^2 - \frac{1}{8}|A|^4 = \frac{7}{8}|\mu|^{4/3} - 2|\lambda||\mu|^{2/3} \geq \frac{1}{8}|\mu|^{4/3} \geq \frac{1}{16}\left(|\mu|^{4/3} + |\lambda|^2\right).$$

2. Assume that $\frac{1}{12}|\lambda|^{3/2} \leq |\mu| \leq \left(\frac{8}{3}\right)^{3/2}|\lambda|^{3/2}$. For $A = (4|\lambda|)^{-1}\mu$ we find

$$g_H(\lambda,\mu) \geq \mu : A - 2|\lambda||A|^2 - \left(\frac{3}{8}\right)^6\frac{1}{4\cdot18}|A|^4$$

$$= \frac{1}{4}|\lambda|^{-1}|\mu|^2 - \frac{1}{8}|\lambda|^{-1}|\mu|^2 - \left(\frac{3}{8}\right)^6\frac{1}{4\cdot18\cdot256}|\lambda|^{-4}|\mu|^4$$

$$\geq \frac{1}{8\cdot12^2}|\lambda|^2 - \left(\frac{3}{8}\right)^6\frac{1}{4\cdot18\cdot256}\left(\frac{8}{3}\right)^6|\lambda|^2$$

$$\geq \frac{1}{16\cdot12^2}|\lambda|^2 \geq \frac{1}{32\cdot12^2}\left(\frac{3}{8}\right)^2\left(|\lambda|^2 + |\mu|^{4/3}\right).$$

3. Assume that $|\mu| \leq \frac{1}{12}|\lambda|^{3/2}$. Let $V \in \{1,2,3\}$ such that it holds for the $V^{th}$ row, $\Lambda_V \in \mathbb{R}^3$, of $\lambda \in \mathbb{R}^{3\times3}$ that $m^2 := \lambda_{Vi}\lambda_{Vi}$ (no sum on $V$) $= |\Lambda_V|^2 \geq \frac{1}{3}|\lambda|^2$, $m > 0$. Next, let $Y, Z \in \{1,2,3\}$ such that $\{V,Y,Z\} = \{1,2,3\}$. Choose an orthonormal basis $(E_1, E_2, E_3)$ with $E_V = \frac{1}{m}\Lambda_V$ and $\frac{1}{2}\epsilon_{VYZ}E_Y \times E_Z = E_V$. Now, define $A \in \mathbb{R}^{3\times3}$ row-wise as

$$A_{Yi} = \sqrt{m}\,E_{Yi} \quad \text{if } Y \neq V \qquad \text{and} \qquad A_{Vi} = 0.$$

It follows that $|A| = \sqrt{2m}$ and

$$\epsilon_{pqr}\epsilon_{ZBC}A_{Bq}A_{Cr} = \begin{cases} 2m\,E_{Vp} = 2\lambda_{Vp} & \text{if } Z = V, \\ 0 & \text{otherwise.} \end{cases}\tag{83}$$

Then we estimate

$$g_H(\lambda,\mu) \geq -|\mu||A| + 2|\Lambda_V|^2 - \frac{1}{4}|A|^4 \geq -2^{1/2}|\mu||\Lambda_V|^{1/2} + |\Lambda_V|^2$$

$$\geq -2|\mu||\lambda|^{1/2} + \frac{1}{3}|\lambda|^2 \geq \frac{1}{6}|\lambda|^2 \geq \frac{1}{12}\left(|\lambda|^2 + |\mu|^{4/3}\right).$$

23

$\square$

**Proposition C.2.** *For $H(A) = \left(\frac{3}{8}\right)^6 \frac{1}{4\cdot 18}|A|^4$ and $A^{(b)} \in W^{3/4,4}(\partial\Omega;\mathbb{R}^{3\times 3})$ the dual functional $\tilde{S}_H$ is weakly coercive on the Banach space $\mathcal{A} := \{\lambda \in L^2(\Omega;\mathbb{R}^{3\times 3}) : \mathrm{curl}(\lambda) \in L^{4/3}(\Omega;\mathbb{R}^{3\times 3})\}$, i.e., for every sequence $(\lambda_k)_k \subseteq \mathcal{A}$ such that $\sup_k \tilde{S}_H[\lambda_k] < \infty$ there exists a (not relabeled) subsequence and $\lambda \in \mathcal{A}$ such that $\lambda_k \rightharpoonup \lambda$ in $L^2$ and $\mathrm{curl}(\lambda_k) \rightharpoonup \mathrm{curl}(\lambda)$ in $L^{4/3}$.*

*Proof.* By Proposition C.1 and the trace estimate discussed above Proposition C.1 for a sequence $(\lambda_k)_k \subseteq \mathcal{A}$ satisfying $\sup_k \tilde{S}_H[\lambda_k] \leq C < \infty$ it follows for all $k \in \mathbb{N}$ that

$$c(\|\lambda_k\|_{L^2}^2 + \|\mathrm{curl}\,\lambda_k\|_{L^{4/3}}^{4/3}) - C'\|A^{(b)}\|_{W^{3/4,4}}(\|\lambda_k\|_{L^2} + \|\mathrm{curl}\,\lambda_k\|_{L^{4/3}}) \leq \tilde{S}_H[\lambda_k] \leq C.$$

Since $\lim_{\alpha,\beta\to+\infty} c(\alpha^2 + \beta^{4/3}) - C'\|A^{(b)}\|_{W^{3/4,4}}(\alpha + \beta) = +\infty$, it follows that the terms $\|\lambda_k\|_{L^2}$ and $\|\mathrm{curl}(\lambda_k)\|_{L^{4/3}}$ are uniformly bounded. By the usual compactness arguments this implies the existence of a (not relabeled) subsequence and functions $\lambda \in L^2(\Omega;\mathbb{R}^{3\times 3})$ and $\mu \in L^{4/3}(\Omega;\mathbb{R}^{3\times 3})$ such that $\lambda_n \rightharpoonup \lambda$ in $L^2$ and $\mathrm{curl}(\lambda_n) \rightharpoonup \mu$ in $L^{4/3}$. Eventually, it can then be checked that actually $\mathrm{curl}(\lambda) = \mu$. $\square$

Note that by similar arguments one can show directly from Proposition C.1 and (82) that the dual functional $\tilde{S}_H$ is bounded from below.

## Acknowledgments

## References

[1] Shiing Shen Chern and James Simons. "Characteristic forms and geometric invariants". In: *Ann. of Math. (2)* 99 (1974), pp. 48–69. ISSN: 0003-486X. DOI: 10.2307/1971013. URL: https://doi.org/10.2307/1971013.

[2] Amit Acharya. "Variational principle for nonlinear PDE systems via duality". In: *Quarterly of Applied Mathematics* 81 (2023), pp. 127–140.

[3] Amit Acharya. "A dual variational principle for nonlinear dislocation dynamics". In: *Journal of Elasticity* 154.1 (2023), pp. 383–395.

[4] Siddharth Singh, Janusz Ginster, and Amit Acharya. "A Hidden Convexity of Nonlinear Elasticity". In: *published online, Journal of Elasticity* (2024). URL: https://arxiv.org/abs/2401.08538.

[5] Amit Acharya. "A hidden convexity in continuum mechanics, with application to classical, continuous-time, rate-(in) dependent plasticity". In: *Mathematics and Mechanics of Solids* (2024). DOI: https://doi.org/10.1177/10812865241258154.

[6] Amit Acharya and Ambar N. Sengupta. "Action principles for dissipative, non-holonomic Newtonian mechanics". In: *Proceedings of the Royal Society, A* 480 (2024), p. 20240113.

[7] Amit Acharya, Bianca Stroffolini, and Arghir Zarnescu. *Variational dual solutions for incompressible fluids.* 2024. URL: https://arxiv.org/abs/2409.04911.

[8] Uditnarayan Kouskiya and Amit Acharya. "Hidden convexity in the heat, linear transport, and Euler's rigid body equations: A computational approach". In: *Quarterly of Applied Mathematics* 82 (2024), pp. 673–703.

[9] Uditnarayan Kouskiya and Amit Acharya. "Inviscid Burgers as a degenerate elliptic problem". In: *published online, Quarterly of Applied Mathematics* (`https://arxiv.org/abs/2401.08814`, 2024).

[10] Yann Brenier. "Hidden Convexity in some nonlinear PDEs from geometry and physics". In: *Journal of Convex Analysis* 17 (2010), pp. 945–959.

[11] Yann Brenier. "The initial value problem for the Euler equations of incompressible fluids viewed as a concave maximization problem". In: *Communications in Mathematical Physics* 364 (2018), pp. 579–605.

[12] Shiing-shen Chern. "A simple intrinsic proof of the Gauss-Bonnet formula for closed Riemannian manifolds". In: *Ann. of Math. (2)* 45 (1944), pp. 747–752. ISSN: 0003-486X. DOI: `10.2307/1969302`. URL: `https://doi.org/10.2307/1969302`.

[13] Richard S. Hamilton. "The inverse function theorem of Nash and Moser". In: *Bull. Amer. Math. Soc. (N.S.)* 7.1 (1982), pp. 65–222. ISSN: 0273-0979,1088-9485. DOI: `10.1090/S0273-0979-1982-15004-2`. URL: `https://doi.org/10.1090/S0273-0979-1982-15004-2`.

[14] G. Leoni. *A first course in Sobolev spaces.* American Mathematical Soc., 2017.

[15] R. Dautray and J.-L. Lions. *Mathematical analysis and numerical methods for science and technology: volume 3 Spectral Theory and Applications.* Vol. 3. Springer Science & Business Media, 1999.