Euler-Lagrange Equations for Nonlinearly Elastic Rods with Self-Contact

Friedemann Schuricht*

Mathematisches Institut
Universität zu Köln
Weyertal 86–90
D-50931 Köln, Germany

Heiko von der Mosel**

Mathematisches Institut
Universität Bonn
Beringstraße 1
D-53115 Bonn, Germany

September 6, 2007

Abstract

We derive the Euler-Lagrange equations for nonlinear elastic rods with self-contact. The excluded volume constraint is formulated in terms of an upper bound on the global curvature of the centreline. This condition is shown to guarantee the global injectivity of the deformation mapping of the elastic rod. Topological constraints such as a prescribed knot and link class to model knotting and supercoiling phenomena as observed, e.g., in DNA-molecules, are included using the notion of isotopy and Gaussian linking number. The global curvature as a nonsmooth side condition requires the use of Clarke's generalized gradients to obtain the explicit structure of the contact forces, which appear naturally as Lagrange multipliers in the Euler-Lagrange equations. Transversality conditions are discussed and higher regularity for the strains, moments, the centreline and the directors is shown.

Mathematics Subject Classification: 49K15, 53A04, 57M25, 74B20, 74G40, 74G65, 74K10, 74M15 92C40

*fs@math.uni-koeln.de
**heiko@math.uni-bonn.de

1 Introduction

In nature we observe that bodies can touch but not penetrate each other, since interpenetration of matter is impossible. In particular, deformable bodies can exhibit self-contact as, e.g., if we step on a beer can or if an electrical cord forms knots and wraps around itself. It turns out that the mathematical treatment of this simple physical phenomenon is surprisingly difficult.

In the bio-sciences there is rapidly growing interest in a variety of problems which display the effect of self-contact as an inherent feature. For instance, the supercoiling of DNA (i.e., when the double helix wraps around itself) and knotting phenomena cause self-touching of the molecule. This mechanism controls certain biochemical processes in the cell and is of special interest in structural molecular biology (see [5], [6], [7], [18], [31], [25]). There are also molecular configurations resembling the centreline of *ideal knots* which may be described as maximally tightened knotted ropes touching themselves "everywhere" (see [13]). On a completely different length scale multicellular bacterial macrofibres of *Bacillus subtilis* form a highly twisted helical structure exhibiting self-contact, which seems to be an advantageous configuration of self-organization in the cell population (see [16],[28]). Macroscopic examples are knotted metal wires with isolated contact points or with several regions of line contact. Interestingly certain helical shapes observed in nature coincide with optimal configurations of closely packed strings, which also serve as model for the structure of folded polymeric chains (see [15],[27]).

The previous examples have the common feature that they can be modeled as long slender elastic tubes or rods deforming in space where the constraint prohibiting interpenetration cannot be neglected. In particular, special side conditions and topological constraints enforce the tubular surface to touch itself. Based on the Cosserat theory, describing deformations of nonlinearly elastic rods in space that can undergo flexure, extension, shear, and torsion, the existence of energy minimizing configurations for that class of problems is shown in [11],[30] and [22]. In the present paper we derive the Euler-Lagrange equation and further regularity results as necessary conditions for energy minimizing rods without interpenetration and subjected to topological constraints where we restrict our investigation to inextensible unshearable rods. Starting with solutions whose existence is proved in particular in [11], no additional hypothetical assumptions on the regularity of the rod or on the position and direction of the contact forces are made. It should be emphasized that such a rigorous derivation of variational equations in nonlinear elasticity taking into account self-contact has never been done before. Furthermore, notice that most investigations on contact problems in the literature are based on comparably simple mechanical models enjoying nice convexity properties and thus being accessible by variational inequalities. These in turn, however, do not contain any explicit term describing the contact reaction. In our more general situation, on the other hand, one cannot hope for such convexity properties but, by employing nonsmooth tools more subtle than convex analysis, we are able to derive the explicit contact term as a Lagrange multiplier which provides additional structural information about the contact reactions. Moreover this allows us to obtain further regularity properties of the minimizing configuration. In particular, we rigorously obtain from our analysis that contact forces are directed normally to the lateral surface of the rod — a physically natural fact which is usually invoked as hypotheses into the theory.

The underlying mathematical structure for the description of deformed configurations of an elastic rod is that of a *framed curve.* Here a base curve, interpreted as centreline of a tube of uniform radius, is associated with an orthonormal frame at each point, reflecting the orientation of the cross section attached to that point. The main difficulty in posing an appropriate variational problem modeling the previous examples is to find a mathematically precise and analytically tractable formulation of the condition that the tube not pass through itself, which is often referred to as the *excluded volume constraint.* On the one hand, the method used in [22] delivers very general existence results, but it seems to be unsuitable for the derivation of the Euler-Lagrange equation. The method used in [11], on the other hand, provides a geometrically exact condition for self-avoidance and corresponding existence results for the smaller class of unshearable rods, but, as we will see in this paper, the Euler-Lagrange equation can be derived rigorously, i.e., without hypothetical regularity assumptions for the energy minimizer. Here the excluded volume constraint, which expresses global injectivity for the mapping assigning the deformed position to each material point of the rod, is mathematically transferred to the centreline as a bound on its *global curvature.* This is a nonlocal quantity whose inverse, the *global radius of curvature,* was introduced by Gonzalez and Maddocks [10] in the context of ideal knots. Since this notion is not restricted to smooth curves (as is the case, e.g., for the classical normal injectivity radius), global curvature turns out to be appropriate for the direct methods in the calculus of variations. Let us mention that the use of repulsive potentials along the centreline of the rod to model self-avoidance (as an alternative to our geometrically exact excluded volume constraint) leads to non-trivial analytical and computational difficulties (see [9], [17], [29]). For an appropriate description of self-contact problems for rods, we take into account also topological restrictions for the framed curve as a given knot class for the centreline and a prescribed link between the centreline and a curve on the lateral boundary of the rod.

The mathematical challenge for deriving the Euler-Lagrange equations in the present context lies in the fact that global curvature furnishes a nonsmooth nonconvex side condition of the variational problem. Thus, standard arguments leading to variational inequalities are not applicable (see [14]). Furthermore it would be desirable to obtain explicit structural information about the contact forces, which remains hidden when using variational inequalities. It turns out that, similar to the treatment of contact between nonlinearly elastic bodies and rigid obstacles (see [19], [20], [21]), Clarke's calculus of generalized gradients of locally Lipschitz continuous functionals is the key to succeed (see [4]). It provides a general Lagrange multiplier rule applicable to our situation, and suitable tools to evaluate the structure of the Lagrange multiplier corresponding to the nonsmooth excluded volume constraint. The resulting Euler-Lagrange equation stated in Theorem 4.1 contains an explicit contact term and corresponds to the mechanical equilibrium condition of the rod theory, at least if certain transversality conditions detecting the physically relevant cases are satisfied. This way we recover in particular apparently obvious mechanical properties of frictionless contact forces in a mathematically rigorous way.

In contrast to most treatments for contact problems, our approach allows us to conclude higher regularity properties for energy minimizing states of (unshearable inextensible) rods exhibiting self-contact without hypothetical smoothness assumptions, but merely based on the smoothness of the data. If the density of the elastic energy is strictly convex and sufficiently smooth, then the moments, the first derivatives of the frame field, and the second derivatives of the centreline of the rod in equilibrium have to be Lipschitz continuous, cf. Corollaries 4.3 and 4.5. This, in particular, excludes concentrated contact moments and answers a long standing open question in the engineering community. Our regularity results including the explicit structural information about contact forces may turn out quite useful for numerical computations, where a thorough understanding of contact sets and contact forces seems crucial, see [6].

In Section 2 the reader is introduced to the Cosserat theory of nonlinearly elastic rods to an extent necessary for the purposes of this paper. In particular, the theory is specialized to materials where shear and extension can be neglected and to rods where all cross sections are circular with the same radius. But, in generalization to usual treatments, we have to consider forces as vector-valued measures in order not to invoke a priori structural restrictions for contact forces that, e.g., can indeed have concentrations.

Section 3 is devoted to geometric and topological constraints to be invoked in our rod problems. First we describe the excluded volume constraint in terms of the global curvature, which guarantees global injectivity of the deformation mapping, see Lemma 3.1. For that we review the definition of global curvature and its basic properties. Then the formulation of topological constraints such as a given knot class for the centrecurve and a given link class for a framed curve is introduced using the notion of isotopy and the Gaussian linking number, where we employ an analytic formula for the latter avoiding topological degree theory. We extend this concept to the case where the frame field is not closed as a curve in $SO(3)$. This way we are able to distinguish the infinitely many equilibrium states having the same boundary conditions but differ in knotting and linking (number of rotations of the frame around the centreline).

In Section 4 we state a general variational problem for nonlinearly elastic rods subjected to the geometric excluded volume constraint, to topological restrictions, and to boundary conditions. Then we formulate the Euler-Lagrange equation for that problem, a number of structural properties for contact forces which may occur in the case of self-touching, and further regularity results for the moments and the shape of the rod. In particular we consider the case of a quadratic elastic energy which is important for various applications.

Section 5 contains all the proofs. In Section 5.1 we prove Theorem 4.1 and Corollary 4.2 in several steps. First we show that the topological properties (knot class and link type) of the minimizing solution are stable under small perturbations in an appropriate space of variations. Furthermore we remove some redundancies in the side conditions. This way we obtain a reduced variational problem without topological constraints, a solution of which is given by the solution of the original problem, see Lemma 5.4. We then claim that a nonsmooth Lagrange multiplier rule is applicable to the reduced variational problem. In order to prove this claim we have to compute the derivative of the energy (Lemma 5.6), the derivatives of the functionals occurring in the boundary conditions (Lemma 5.7), and the generalized gradient of a functional involving the global curvature (Lemma 5.10). The Euler-Lagrange equation then follows. Analyzing the properties of the contact forces and certain transversality conditions we finish the proof of Theorem 4.1. The remaining regularity assertions in Corollaries 4.3–4.6 are verified in Section 5.2.

In Appendix A we provide the quite technical computation of the derivative of the mapping assigning the frame vectors to certain shape variables of the rod. Analytically this means to determine the derivative of a solution of an ordinary differential equation with respect to a parameter in a Banach space. A short summary of the relevant facts regarding Clarke's calculus of generalized gradients can be found in Appendix B. Here we present a variant of a nonsmooth chain rule adapted to our application.

Notation. We use $\boldsymbol{x} \cdot \boldsymbol{y}$ to denote the standard Euclidean inner product of $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^3$, and $|\cdot|$ to denote the (intrinsic) distance between two points in $\mathbb{R}^3$ or in some parameter set $J \subset \mathbb{R}$ depending on the context. To denote the enclosed (smaller) angle between two non-zero vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^3$ we use $\sphericalangle(\boldsymbol{x}, \boldsymbol{y}) \in [0, \pi]$. The distance between a point $\boldsymbol{x} \in \mathbb{R}^3$ and a subset $\Sigma \subset \mathbb{R}^3$ will be denoted by $\operatorname{dist}(\boldsymbol{x}, \Sigma)$ and the diameter of Σ will be denoted by $\operatorname{diam}(\Sigma)$. For any $\delta > 0$ we define open neighbourhoods of $\boldsymbol{x}$ and Σ by

$$B_\delta(\boldsymbol{x}) = \{\boldsymbol{y} \in \mathbb{R}^3 \mid |\boldsymbol{y} - \boldsymbol{x}| < \delta\}, \qquad B_\delta(\Sigma) = \{\boldsymbol{y} \in \mathbb{R}^3 \mid \operatorname{dist}(\boldsymbol{y}, \Sigma) < \delta\}.$$

The interior of a set Σ will be denoted by $\operatorname{int} \Sigma$. For Sobolev spaces of functions on the interval $[0, L]$, whose weak derivatives up to order m are p-integrable we use the standard notation $W^{m,p}([0, L])$, and the class of functions of bounded variation is denoted by $BV([0, L])$. For general Banach spaces X with dual space X^* we denote the duality pairing on $X^* \times X$ by $\langle \cdot, \cdot \rangle_{X^* \times X}$.

2 Rod theory

In this section we provide a brief introduction to the special Cosserat theory which describes the behaviour of nonlinearly elastic rods that can undergo large deformations in space by suffering flexure, torsion, extension, and shear. General nonlinear constitutive relations appropriate for a large class of applications can be taken into account. Though mathematically one-dimensional, this theory allows a mechanically natural and geometrically exact three-dimensional interpretation of deformed configurations which is of particular importance for problems where contact occurs. In this paper we restrict our attention to rods where shear and extension can be neglected. This special case can be obtained from the general theory by a simple material constraint. For a more comprehensive presentation we refer to Antman [3, Ch. VIII].

Kinematics. We assume that the position $\boldsymbol{p}$ of the deformed material points of a slender cylindrical elastic body can be described in the form

$$\boldsymbol{p}(s,\xi^1,\xi^2) = \boldsymbol{r}(s) + \xi^1\boldsymbol{d}_1(s) + \xi^2\boldsymbol{d}_2(s) \;\text{ for }\; (s,\xi^1,\xi^2) \in \Omega, \tag{1}$$

where the parameter set Ω is given by

$$\Omega := \{\, (s,\xi^1,\xi^2) \in \mathbb{R}^3 \mid s \in [0,L], \quad (\xi^1)^2 + (\xi^2)^2 \leq \theta^2 \}. \tag{2}$$

Here, $\boldsymbol{r} : [0,L] \to \mathbb{R}^3$ describes the deformed configuration of the centreline of the body. $\boldsymbol{d}_1(s), \boldsymbol{d}_2(s)$ are orthogonal unit vectors describing the orientation of the deformed cross section at the point $\boldsymbol{r}(s) \in [0,L]$. We interpret s as length parameter and ξ^1, ξ^2 as thickness parameters of the rod. With

$$\boldsymbol{d}_3 := \boldsymbol{d}_1 \wedge \boldsymbol{d}_2$$

we get a right-handed orthonormal basis $\{\boldsymbol{d}_1, \boldsymbol{d}_2, \boldsymbol{d}_3\}$ at each $s \in [0,L]$, whose vectors are called *directors*, and which can be identified with an orthogonal matrix $\boldsymbol{D} = (\boldsymbol{d}_1|\boldsymbol{d}_2|\boldsymbol{d}_3) \in SO(3)$ (the right-hand side denotes the matrix with columns $\boldsymbol{d}_1, \boldsymbol{d}_2, \boldsymbol{d}_3$). A deformed configuration of the rod is thus determined by functions $\boldsymbol{r} : [0,L] \to \mathbb{R}^3$ and $\boldsymbol{D} : [0,L] \to SO(3)$, where it is reasonable to consider $\boldsymbol{r} \in W^{1,q}([0,L],\mathbb{R}^3)$ and $\boldsymbol{D} \in W^{1,p}([0,L],\mathbb{R}^{3\times3})$, $p,q \geq 1$.

In the special case of an inextensible unshearable rod we assume that s is the arc length of the deformed centrecurve $\boldsymbol{r}(\cdot)$ and that the deformed cross sections are orthogonal to the base curve, i.e.,

$$\boldsymbol{r}'(s) = \boldsymbol{d}_3(s) \quad \text{for all } \; s \in [0,L], \tag{3}$$

(note that $|d_3(s)| = 1$). Thus, by $\boldsymbol{d}_3 \in W^{1,p}([0,L],\mathbb{R}^3)$, we even have that $\boldsymbol{r} \in W^{2,p}([0,L],\mathbb{R}^3)$. (Observe that $\boldsymbol{d}_3(\cdot)$ is continuous and admits derivatives a.e. on $[0,L]$.) Specifying [11, Lemma 6] to that case we see, that each such configuration uniquely corresponds to shape and placement variables

$$w = (u, \boldsymbol{r}_0, \boldsymbol{D}_0) \ \text{ with } \ u \equiv (u^1, u^2, u^3),$$

in the space

$$X_0^p := L^p([0,L], \mathbb{R}^3) \times \mathbb{R}^3 \times SO(3),$$

such that

$$\begin{aligned} & \boldsymbol{d}_k'(s) = \Big[\sum_{i=1}^{3} u^i(s)\boldsymbol{d}_i(s)\Big] \wedge \boldsymbol{d}_k(s) \ \text{ for } \quad \text{a.e. } s \in [0,L], \quad k = 1,2,3, \\ & \boldsymbol{D}(0) = \boldsymbol{D}_0, \\ & \boldsymbol{r}(s) = \boldsymbol{r}_0 + \int_0^s \boldsymbol{d}_3(\tau)\, d\tau. \end{aligned} \tag{4}$$

The function u is called the *strain* and fixes the shape of the rod while $(\boldsymbol{r}_0, \boldsymbol{D}_0)$ determine its spatial placement. We use the notation $\boldsymbol{p}[w], \boldsymbol{r}[w]$, etc. to indicate that the values are calculated for $w = (u, \boldsymbol{r}_0, \boldsymbol{D}_0) \in X_0^p$. Notice that X_0^p is a subset of the Banach space

$$X^p := L^p([0,L], \mathbb{R}^3) \times \mathbb{R}^3 \times \mathbb{R}^{3\times 3}.$$

By $w^\circ := (u^\circ, \boldsymbol{r}_0, \boldsymbol{D}_0)$ we identify the relaxed (stress-free) *reference configuration*. Note that $\boldsymbol{r}[w^\circ]$ need not be a straight line.

We demand that the map $\boldsymbol{p}$ preserve orientation in the sense that

$$\det\left[\frac{\partial \boldsymbol{p}(s,\xi^1,\xi^2)}{\partial(s,\xi^1,\xi^2)}\right] > 0 \quad \text{for a.e.} \quad (s,\xi^1,\xi^2) \in \Omega, \tag{5}$$

which, due to the special form of $\boldsymbol{p}$, is equivalent to

$$\frac{1}{\theta} \geq \sqrt{(u^1)^2 + (u^2)^2} = |\,\boldsymbol{r}''| \quad \text{a.e. on } [0,L] \tag{6}$$

(cf. [11]). Here, $|\boldsymbol{r}''|$ is the *local curvature* of the base curve $\boldsymbol{r}(.)$, since $\boldsymbol{r}$ is parameterized by arc length. It can be shown that inequality (6) ensures *local* injectivity of $\boldsymbol{p}(.)$ on $\operatorname{int}\Omega$, (argue as in the proof of [22, Prop. 3.1]). Note, on the other hand, that *global* injectivity of $\boldsymbol{p}(.)$ on $\operatorname{int}\Omega$, which prevents interpenetration of the elastic body, is an important and natural requirement in continuum mechanics. While this condition is neglected in many treatments in elasticity, its consideration is a major objective of our investigation here.

In this paper we are particularly interested in configurations where the rod is closed to a ring, i.e., we assume that

$$\boldsymbol{r}(0) = \boldsymbol{r}(L), \quad \boldsymbol{d}_3(0) = \boldsymbol{d}_3(L), \tag{7}$$

and call it a *closed configuration*. Notice that the centreline $\boldsymbol{r}$ is closed in the C^1-sense, i.e., the curve and its tangent closes up at the end points. For the rod this can be rephrased by saying that the cross sections at the end points coincide, but the directors $\boldsymbol{d}_1(0)$, $\boldsymbol{d}_2(0)$, may be different from $\boldsymbol{d}_1(L)$ and $\boldsymbol{d}_2(L)$, respectively.

Forces and equilibrium conditions. In contact problems as considered in the present work, contact forces may occur, which are possibly concentrated, e.g., at some isolated point. Thus we need a more general approach for the treatment of forces than usual (for a more detailed discussion see Schuricht [19]). In particular, we can not assume integrable force densities in general. We identify subbodies of the rod with corresponding subsets of Ω. In particular, we set

$$\Omega_{\mathcal{J}} := \{(s, \xi^1, \xi^2) \in \Omega : s \in \mathcal{J}\} \text{ for } \mathcal{J} \subset [0, L], \text{ and } \Omega_s := \Omega_{[s,L]}. \tag{8}$$

For a given configuration, the material of Ω_s exerts a *resultant force* $\boldsymbol{n}(s)$ and a *resultant couple* $\boldsymbol{m}(s)$ across section s on the material of $\Omega_{[0,s)}$. This definition does not make sense for $s = 0$, but it is convenient to set

$$\boldsymbol{n}(0) := 0 \text{ and } \boldsymbol{m}(0) := 0. \tag{9}$$

We assume that all forces other than $\boldsymbol{n}$ acting on the body can be described by a finite vector-valued Borel measure

$$\Omega' \mapsto \mathfrak{f}(\Omega'), \tag{10}$$

assigning the resultant force to subbodies which correspond to Borel sets $\Omega' \subset \Omega$. We call $\mathfrak{f}$ the *external force*. It generates the *induced couple of* $\mathfrak{f}$ given by

$$\mathfrak{l}_{\mathfrak{f}}(\Omega') := \int_{\Omega'} [\xi^1 \boldsymbol{d}_1(s) + \xi^2 \boldsymbol{d}_2(s)] \wedge d\mathfrak{f}(s, \xi^1, \xi^2). \tag{11}$$

Analogously we assume that all couples different from $\boldsymbol{m}$ and $\mathfrak{l}_\mathfrak{f}$ can be given by a finite vector-valued Borel measure

$$\Omega' \to \mathfrak{l}(\Omega'), \tag{12}$$

which we call the *external couple.*

A configuration of the rod is in equilibrium if the resultant force and the resultant torque about the origin vanish for each part of the rod. In terms of the distribution functions

$$\boldsymbol{f}(s) := \int_{\Omega_s} d\mathfrak{f}(\sigma, \xi^1, \xi^2), \quad \boldsymbol{l}(s) := \int_{\Omega_s} d\mathfrak{l}(\sigma, \xi^1, \xi^2), \tag{13}$$

$$\boldsymbol{l}_f(s) := \int_{\Omega_s} d\mathfrak{l}_\mathfrak{f}(\sigma, \xi^1, \xi^2) = \int_{\Omega_s} [\xi^1 \boldsymbol{d}_1(\sigma) + \xi^2 \boldsymbol{d}_2(\sigma)] \wedge d\mathfrak{f}(\sigma, \xi^1, \xi^2), \tag{14}$$

these requirements are equivalent to the *equilibrium conditions* in integral form

$$\boldsymbol{n}(s) - \boldsymbol{f}(s) \;=\; 0 \;\text{ for }\; s \in [0, L], \tag{15}$$

$$\boldsymbol{m}(s) - \int_s^L \boldsymbol{r}'(\sigma) \wedge \boldsymbol{n}(\sigma)\, d\sigma - \boldsymbol{l}_f(s) - \boldsymbol{l}(s) \;=\; 0 \;\text{ for }\; s \in [0, L]. \tag{16}$$

Notice that the resultant force and the resultant couple of all external actions for the whole body must vanish by (9). For sufficiently smooth external forces and moments we obtain the classical form of the equilibrium conditions by differentiating (15), (16).

Constitutive Relations. We assume that the material of the rod is *elastic*, which means that there is a *constitutive function* $\hat{\boldsymbol{m}}$, such that $\boldsymbol{m}$ is determined by the strain through

$$\boldsymbol{m}(s) = \hat{\boldsymbol{m}}(u(s), s), \tag{17}$$

where $\hat{\boldsymbol{m}}$ is usually assumed to be continuously differentiable in u. Note that (17) can provide the correct values of $\boldsymbol{m}$ only a.e. on $[0, L]$, if the strains are discontinuous as, e.g., in the case where concentrated forces or couples are present (cf. [19]). Let us mention that the resultant force $\boldsymbol{n}$ cannot be determined by a constitutive function in the unshearable inextensible case - it rather enters the theory as a Lagrange multiplier.

The material is called *hyperelastic* if there is a *stored energy density* $W : \mathbb{R}^3 \times [0, L] \to \mathbb{R} \cup \{+\infty\}$, such that

$$\hat{\boldsymbol{m}}(u, s) = \sum_{i=1}^{3} W_{u^i}(u, s)\boldsymbol{d}_i(s). \tag{18}$$

The *total stored energy* of the rod is given by

$$E_{\mathrm{s}}(u, v) = \int_0^L W(u(s), s)\, ds. \tag{19}$$

For our analysis we assume that

(W1) $W(., s)$ is continuously differentiable on $\mathbb{R}^3$ for a.e. $s \in [0, L]$,

(W2) $W(u, .)$ is Lebesgue-measurable on $[0, L]$ for all $u \in \mathbb{R}^3$.

The *Strong Ellipticity Condition* in nonlinear elasticity enforces $W(., s)$ to be convex, i.e., the matrix

$$\frac{\partial \hat{\boldsymbol{m}}}{\partial u}$$

has to be positive definite. It is reasonable to require that the energy density W approaches ∞ under complete compression of the material, i.e.,

$$W(u, v, s) \to \infty \text{ as } \frac{1}{\theta} - \sqrt{(u^1)^2 + (u^2)^2} \to 0. \tag{20}$$

Due to severe analytical difficulties for regularity investigations connected with such a degeneracy we will neglect this condition in our treatment and focus on energy densities that are sufficiently smooth.

In the following we assume that there are no prescribed external couples $\mathfrak{l}$ and, for simplicity, that the given external force depends only on the coordinates (s, ξ^1, ξ^2), but not on the configuration $\boldsymbol{p}[w]$, and denote this special force by $\mathfrak{f}_e$ in contrast to general external forces $\mathfrak{f}$ introduced in (10).

Then $\mathfrak{f}_e$ is conservative and has the *potential energy*

$$\begin{aligned} E_{\mathrm{p}}(w) &:= E_{\mathrm{p}}(\boldsymbol{p}[w]) := -\int_\Omega \boldsymbol{p}[w](s, \xi^1, \xi^2) \cdot d\mathfrak{f}_e(s, \xi^1, \xi^2) \\ &= -\int_\Omega [\boldsymbol{r}[w](s) + \xi^1 \boldsymbol{d}_1[w](s) + \xi^2 \boldsymbol{d}_2[w](s)] \cdot d\mathfrak{f}_e(s, \xi^1, \xi^2). \end{aligned} \tag{21}$$

This way we can cover external forces such as weight or prescribed terminal loads. However, in our investigation later, we also consider *self-contact forces* which do depend on the configuration $\boldsymbol{p}[w]$. But such forces do not enter our analysis through the potential energy, but occur naturally as Lagrange multipliers of some constrained variational problem. In particular, we are going to derive the Euler-Lagrange equation for energy minimizing configurations subjected to an analytical condition preventing interpenetration and to topological constraints described in the next section.

3 Constraints

Global injectivity. In [11] an analytical condition ensuring global injectivity of the mapping $\boldsymbol{p}$ is introduced by means of the global radius of curvature – a nonlocal geometric quantity for curves that goes back to Gonzalez, Maddocks [10], and which is further analyzed in [24]. Note that elements $(\boldsymbol{r}, \boldsymbol{D})$ determining a configuration of a rod are referred to as framed curves in the geometric context of [11] and [24]. We are going to recall the definition of the global radius of curvature, present the related notion of global curvature and its important properties that our analysis later is based on.

Recall that throughout our developments we exclusively deal with centrelines $\boldsymbol{r} : [0, L] \to \mathbb{R}^3$ parameterized by arc length, and with closed configurations, see (7). Therefore it will often be useful to identify the interval $[0, L]$ with the circle $S_L \cong \mathbb{R}/(L \cdot \mathbb{Z})$. The curve $\boldsymbol{r}$ is said to be *simple* if $\boldsymbol{r} : S_L \to \mathbb{R}^3$ is injective. Otherwise there exist $s, t \in S_L$ $(s \neq t)$ for which $\boldsymbol{r}(s) = \boldsymbol{r}(t)$. Any such pair will be called a *double point* of $\boldsymbol{r}$.

For a closed curve $\boldsymbol{r} : S_L \to \mathbb{R}^3$ the *global radius of curvature* $\rho_G[\boldsymbol{r}](s)$ *at* $s \in S_L$ is defined as

$$\rho_G[\boldsymbol{r}](s) := \inf_{\substack{\sigma,\tau \in S_L \setminus \{s\} \\ \sigma \neq \tau}} R(\boldsymbol{r}(s), \boldsymbol{r}(\sigma), \boldsymbol{r}(\tau)), \tag{22}$$

where $R(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \geq 0$ is the radius of the *smallest* circle containing the points $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z} \in \mathbb{R}^3$. For collinear but pairwise distinct points $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ we set $R(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ to be infinite. When $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ are non-collinear (and thus distinct) there is a unique circle passing through them and

$$R(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = \frac{|\boldsymbol{x} - \boldsymbol{y}|}{|2 \sin[\sphericalangle(\boldsymbol{x} - \boldsymbol{z}, \boldsymbol{y} - \boldsymbol{z})]|}\,. \tag{23}$$

If two points coincide, however, say $\boldsymbol{x} = \boldsymbol{z}$ or $\boldsymbol{y} = \boldsymbol{z}$, then there are many circles through the three points and we take $R(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ to be the smallest possible radius namely the distance $|\boldsymbol{x}-\boldsymbol{y}|/2$. We should point out that with this choice the function $R(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ fails to be continuous at points, where at least two of the arguments $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$, coincide. Notice nevertheless that, by definition, $R(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ is symmetric in its arguments.

The *global radius of curvature of* $\boldsymbol{r}$ is defined as

$$\mathcal{R}[\boldsymbol{r}] := \inf_{s\in S_L} \rho_G[\boldsymbol{r}](s). \tag{24}$$

If $\mathcal{R}[\boldsymbol{r}] > 0$, then $\boldsymbol{r}$ is simple, and $\boldsymbol{r} \in C^{1,1} \cong W^{2,\infty}$, see [11, Lemma 2],[1] i.e., $\boldsymbol{r}$ has a Lipschitz continuous tangent field $\boldsymbol{r}'$. Furthermore,

$$\|\boldsymbol{r}''\|_{L^\infty} \leq \frac{1}{\mathcal{R}[\boldsymbol{r}]}\,. \tag{25}$$

Moreover, $\mathcal{R}[\boldsymbol{r}]$ equals the radius of the largest open ball placed tangent to $\boldsymbol{r}(S_L)$ at any point $\boldsymbol{r}(s)$, that can be rotated around the tangent vector $\boldsymbol{r}'(s)$ without intersecting the curve $\boldsymbol{r}(S_L)$, [11, Lemma 3]. This geometric property gives an intuitive idea why deformed rods with a centreline $\boldsymbol{r}$ satisfying $\mathcal{R}[\boldsymbol{r}] \geq \theta$, might have no self-intersections. The following result confirms that unshearable inextensible rods with such a positive lower bound on the global radius of curvature are indeed globally injective. A more general version for unshearable extensible rods covering the following Lemma can be found in [11, Lemma 7].

Lemma 3.1. *Consider a closed configuration* $(\boldsymbol{r}[w], \boldsymbol{D}[w]) \in W^{2,p} \times W^{1,p}$, $p \geq 1$, *for* $w \in X_0^p$, *and suppose that* $\mathcal{R}[\boldsymbol{r}[w]] > 0$. *Then* $\boldsymbol{p}[w]|_{\mathrm{int}\,(\Omega)} : \mathrm{int}\,(\Omega) \to \mathbb{R}^3$ *is globally injective iff* $\mathcal{R}[\boldsymbol{r}[w]] \geq \theta > 0$.

For our further analysis it is necessary to work with the notion of global curvature investigated in detail in [24], which has better regularity properties than the global radius of curvature. The *global curvature of* $\boldsymbol{r}$ *at* $s \in S_L$ is defined as

$$\kappa_G[\boldsymbol{r}](s) := \sup_{\substack{\sigma,\tau\in S_L\setminus\{s\}\\ \sigma\neq\tau}} \frac{1}{R(\boldsymbol{r}(s), \boldsymbol{r}(\sigma), \boldsymbol{r}(\tau))}\,. \tag{26}$$

[1] Even more is true: $\mathcal{R}[\boldsymbol{r}] > 0$ if and only if $\boldsymbol{r} \in C^{1,1}$ and simple, see [24].

Notice that $\kappa_G[\boldsymbol{r}](.)$ can take values in $(0,\infty]$. In analogy to $\mathcal{R}[\boldsymbol{r}]$ we define the *global curvature of* $\boldsymbol{r}$ by

$$\mathcal{K}[\boldsymbol{r}] := \sup_{s\in S_L} \kappa_G[\boldsymbol{r}](s). \tag{27}$$

It is immediate consequence of the definitions that

$$\kappa_G[\boldsymbol{r}](s) = \frac{1}{\rho_G[\boldsymbol{r}](s)}\,, \quad \text{for all} \quad s \in S_L\,, \tag{28}$$

$$\mathcal{K}[\boldsymbol{r}] = \frac{1}{\mathcal{R}[\boldsymbol{r}]}\,. \tag{29}$$

In light of (25) together with (29) we say for curves $\boldsymbol{r}$ with $\mathcal{R}[\boldsymbol{r}] > 0$, that the *global curvature* $\mathcal{K}[\boldsymbol{r}]$ *is locally not attained* iff

$$\|\boldsymbol{r}''\|_{L^\infty} < \mathcal{K}[\boldsymbol{r}]. \tag{30}$$

For curves $\boldsymbol{r}$ with $\mathcal{R}[\boldsymbol{r}] > 0$ we have an alternative analytically more tractable characterization of $\mathcal{K}[\boldsymbol{r}]$. For that let $\boldsymbol{x},\boldsymbol{y},\boldsymbol{t} \in \mathbb{R}^3$ be such that the vectors $\boldsymbol{x}-\boldsymbol{y}$ and $\boldsymbol{t}$ are linearly independent. By P we denote the plane spanned by $\boldsymbol{x}-\boldsymbol{y}$ and $\boldsymbol{t}$. Then there is a unique circle contained in P through $\boldsymbol{x}$ and $\boldsymbol{y}$ and tangent to $\boldsymbol{t}$ in the point $\boldsymbol{y}$. We denote the radius of that circle by $r(\boldsymbol{x},\boldsymbol{y},\boldsymbol{t})$ and set $r(\boldsymbol{x},\boldsymbol{y},\boldsymbol{t}) := \infty$, if $\boldsymbol{x}-\boldsymbol{y}$ and $\boldsymbol{t}$ are collinear. Elementary geometric arguments show that r may be computed as

$$r(\boldsymbol{x},\boldsymbol{y},\boldsymbol{t}) = \frac{|\boldsymbol{x}-\boldsymbol{y}|}{2\left|\frac{\boldsymbol{x}-\boldsymbol{y}}{|\boldsymbol{x}-\boldsymbol{y}|} \wedge \frac{\boldsymbol{t}}{|\boldsymbol{t}|}\right|}\,, \tag{31}$$

which shows that $r(\boldsymbol{x},\boldsymbol{y},\boldsymbol{t})$ is continuous on the set of triples $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{t})$ with the property that $\boldsymbol{x}-\boldsymbol{y}$ and $\boldsymbol{t}$ are linearly independent. But it fails to be continuous at points, where, e.g., two of the arguments coincide. Recall that curves $\boldsymbol{r}$ with $\mathcal{R}[\boldsymbol{r}] > 0$, are of class $C^{1,1}$. Hence, for every pair $(s,\sigma) \in S_L \times S_L$, we can look at the radius $r(\boldsymbol{r}(s),\boldsymbol{r}(\sigma),\boldsymbol{r}'(\sigma))$, and in [24] it is shown that the global curvature $\mathcal{K}[\boldsymbol{r}]$ is characterized by

$$\mathcal{K}[\boldsymbol{r}] = \sup_{\substack{s,\sigma\in S_L\\ s\neq\sigma}} \frac{1}{r(\boldsymbol{r}(s),\boldsymbol{r}(\sigma),\boldsymbol{r}'(\sigma))}\,, \quad \text{if} \quad \mathcal{R}[\boldsymbol{r}] > 0. \tag{32}$$

The following set $A[\boldsymbol{r}]$, where the supremum in (32) is attained, will be of particular interest when deriving the structure of the contact term in the Euler-Lagrange equation in the next section, since it identifies the cross sections touching each other if $\mathcal{K}[\boldsymbol{r}] = \theta^{-1}$.

$$A[\boldsymbol{r}] := \{(s,\sigma) \in [0,L] \times [0,L],\ \sigma \le s : \mathcal{K}[\boldsymbol{r}] = \frac{1}{r(\boldsymbol{r}(s), \boldsymbol{r}(\sigma), \boldsymbol{r}'(\sigma))}\}. \tag{33}$$

By (32) and the definition of $r(.,.,.)$, one has $s \neq \sigma$ for all $(s,\sigma) \in A[\boldsymbol{r}]$, if $\mathcal{R}[\boldsymbol{r}] > 0$. The condition $\sigma \le s$ in the previous definition (which is not part of the corresponding definition in [24]) ensures that each pair of touching cross sections is counted only once. On the other hand, all pairs of touching cross sections are contained in the set $A[\boldsymbol{r}]$. To see this we recall from [24] that for closed curves $\boldsymbol{r}$ with $\mathcal{R}[\boldsymbol{r}] > 0$ and satisfying (30) the identitites

$$|\boldsymbol{r}(s) - \boldsymbol{r}(\sigma)| = 2\mathcal{R}[\boldsymbol{r}] \quad \text{and} \tag{34}$$

$$\boldsymbol{r}'(s) \cdot (\boldsymbol{r}(s) - \boldsymbol{r}(\sigma)) = \boldsymbol{r}'(\sigma) \cdot (\boldsymbol{r}(s) - \boldsymbol{r}(\sigma)) = 0 \tag{35}$$

hold for all $(s,\sigma) \in A[\boldsymbol{r}]$. Consequently, by (31),

$$r(\boldsymbol{r}(s), \boldsymbol{r}(\sigma), \boldsymbol{r}'(\sigma)) = r(\boldsymbol{r}(\sigma), \boldsymbol{r}(s), \boldsymbol{r}'(s)) = 2\mathcal{R}[\boldsymbol{r}]$$

for $(s,\sigma) \in A[\boldsymbol{r}]$. Hence, if $\mathcal{K}[\boldsymbol{r}] = \theta^{-1}$, $\theta > 0$, all pairs of cross sections touching each other are indeed detected by $A[\boldsymbol{r}]$, which we call the set of *contact parameters.*

For our variational approach in Section 4 we need the following continuity result for global curvature proved in more generality in [24].

Lemma 3.2. *Let $\mathcal{L} \subset C^{1,1}([0,L], \mathbb{R}^3)$ be the set of curves $\boldsymbol{r}$ of fixed length $L(\boldsymbol{r}) = L > 0$ and parameterized by arc length. Then $\mathcal{K}[.]$ (and hence $\mathcal{R}[.]$) is continuous on $\mathcal{L}$.*

Topological constraints. We are interested in elastic rods that form a knot of a prescribed type, which can be described by the closed centreline lying in a given knot class. To make this precise we introduce the topological concept of isotopy.

Two continuous closed curves $\boldsymbol{K}_1, \boldsymbol{K}_2 \subset \mathbb{R}^3$ are *isotopic*, denoted as $\boldsymbol{K}_1 \simeq \boldsymbol{K}_2$, if there are open neighbourhoods N_1 of $\boldsymbol{K}_1$, N_2 of $\boldsymbol{K}_2$, and a continuous mapping $\Phi : N_1 \times [0,1] \to \mathbb{R}^3$ such that $\Phi(N_1,\tau)$ is homeomorphic to N_1 for all $\tau \in [0,1]$, $\Phi(\boldsymbol{x},0) = \boldsymbol{x}$ for all $\boldsymbol{x} \in N_1$, $\Phi(N_1,1) = N_2$, and $\Phi(\boldsymbol{K}_1,1) = \boldsymbol{K}_2$.

For simplicity we will frequently write $\boldsymbol{r}_1 \simeq \boldsymbol{r}_2$ instead of $\boldsymbol{r}_1(S_{L_1}) \simeq \boldsymbol{r}_2(S_{L_2})$ for two closed isotopic curves $\boldsymbol{r}_1 : S_{L_1} \to \mathbb{R}^3$ and $\boldsymbol{r}_2 : S_{L_2} \to \mathbb{R}^3$. Roughly speaking, two curves are in the same isotopy class if one can be continuously deformed onto the other.

In [24] the following Lemma concerning C^0-perturbations of knotted curves with a bounded global curvature is shown.

Lemma 3.3. *Let $\boldsymbol{r}$ be a rectifiable closed continuous curve satisfying*

$$\mathcal{K}[\boldsymbol{r}] \leq C_0 \tag{36}$$

for some fixed constant $C_0 < \infty$. Then there exists $\varepsilon = \varepsilon(\boldsymbol{r}, C_0) > 0$, such that for all rectifiable closed continuous curves $\tilde{\boldsymbol{r}}$ with $\mathcal{K}[\tilde{\boldsymbol{r}}] \leq C_0$ and

$$\|\boldsymbol{r} - \tilde{\boldsymbol{r}}\|_{C^0} \leq \varepsilon, \tag{37}$$

one has $\boldsymbol{r} \simeq \tilde{\boldsymbol{r}}$.

The statement of the lemma is no longer true if one removes the assumptions on the global curvature, small knotted regions might pull tight in the uniform topology.

A pair $(\boldsymbol{r}, \boldsymbol{D})$ of a curve $\boldsymbol{r}$ and an associated frame field $\boldsymbol{D}$ is said to be a *framed curve*. Here we consider framed curves with $\boldsymbol{r}(0) = \boldsymbol{r}(L)$ and satisfying (3), which we call closed framed curves. If we prescribe the knot type for the curve $\boldsymbol{r}$ and boundary conditions as, e.g., $\boldsymbol{D}(0) = \boldsymbol{D}(L)$, there are still infinitely many topologically distinct components in the space of closed framed curves. Indeed, every full rotation of the pair $\boldsymbol{d}_1(L), \boldsymbol{d}_2(L)$ within the cross section respects the boundary conditions, but changes the linking number between the centreline and the curve $\boldsymbol{r}(.) + (\theta/2)\boldsymbol{d}_1(.)$, which is a topological invariant. Since such a change of topological type is accompanied by an (often drastic) change of the equilibrium configuration for an elastic rod, we need to prescribe the linking number in order to identify particular solutions, see also the discussion in [2]. The approach in [11] using the concept of homotopies in $SO(3)$ distinguishes only two topologically different classes, since the fundamental group of $SO(3)$ is $\mathbb{Z}_2$.

One way to determine the link between two disjoint closed (but not necessarily simple) curves is to compute the *Gaussian linking number*, which is usually defined in terms of the topological degree, see, e.g., [26, p. 402]. For a pair of absolutely continuous disjoint curves, however, there is an analytically more convenient formula, which we adopt as definition for the linking number. For closed curves $\boldsymbol{r}_1, \boldsymbol{r}_2 \in W^{1,1}([0,L],\mathbb{R}^3)$ with $\boldsymbol{r}_1([0,L]) \cap \boldsymbol{r}_2([0,L]) = \emptyset$, the *linking number* $l(\boldsymbol{r}_1, \boldsymbol{r}_2)$ is given by

$$l(\boldsymbol{r}_1, \boldsymbol{r}_2) = \frac{1}{4\pi} \int_0^L \int_0^L \frac{\boldsymbol{r}_1(s) - \boldsymbol{r}_2(t)}{|\boldsymbol{r}_1(s) - \boldsymbol{r}_2(t)|^3} \cdot [\boldsymbol{r}_1'(s) \wedge \boldsymbol{r}_2'(t)]\, ds\, dt. \tag{38}$$

One can show that $l(\boldsymbol{r}_1, \boldsymbol{r}_2)$ is integer-valued and stable with respect to smooth perturbations preserving the non-intersection property.

For a closed framed curve respecting (3) we want to consider the linking number of the curves $\boldsymbol{r}(.)$ and $\boldsymbol{r}(.) + (\theta/2)\boldsymbol{d}_1(.)$. The problem here is that the second curve might not be closed and that the two curves might intersect each other. The first problem can be solved by closing the curve $\boldsymbol{r}(.) + (\theta/2)\boldsymbol{d}_1(.)$ up in a unique way, namely by

$$\boldsymbol{\beta}^D(s) := \begin{cases} \boldsymbol{r}(s) + \frac{\theta}{2}\boldsymbol{d}_1(s) \quad \text{for } \ s \in [0,L], \\ \boldsymbol{r}(L) + \frac{\theta}{2}[\cos(\phi_D(s-L))\boldsymbol{d}_1(L) + \sin(\phi_D(s-L))\boldsymbol{d}_2(L)] \\ \qquad\qquad\qquad \text{for } \ s \in [L, L+1], \end{cases} \tag{39}$$

where $\phi_D \in [0, 2\pi)$ is the angle between $\boldsymbol{d}_1(0)$ and $\boldsymbol{d}_1(L)$, such that $\phi_D - \pi$ has the same sign as $(\boldsymbol{d}_1(0) \wedge \boldsymbol{d}_1(L)) \cdot \boldsymbol{d}_3(0)$. For technical reasons we identify $\boldsymbol{r}$ with its trivial extension onto $[0, L+1]$ according to

$$\boldsymbol{r}(s) := \boldsymbol{r}(L) \ \text{ for } \ s \in [L, L+1]. \tag{40}$$

Notice that $\boldsymbol{r}, \boldsymbol{\beta}^D \in W^{1,q}([0,L+1],\mathbb{R}^3)$, $1 \le q \le \infty$, if $\boldsymbol{r} \in W^{1,q}([0,L],\mathbb{R}^3)$, and that $\boldsymbol{r}$ and $\boldsymbol{\beta}^D$ are closed. Demanding the global curvature bound $\mathcal{K}[\boldsymbol{r}] \le \theta^{-1}$ we ensure that

$$\boldsymbol{r}([0, L+1]) \cap \boldsymbol{\beta}^D([0, L+1]) = \emptyset \tag{41}$$

by Lemma 3.1 and (29). Thus the linking number of a closed framed curve $(\boldsymbol{r}, \boldsymbol{D})$ satisfying (3), $\mathcal{K}[\boldsymbol{r}] \le \theta^{-1}$, $\boldsymbol{r} \in W^{1,1}([0,L],\mathbb{R}^3)$ and $\boldsymbol{D} \in W^{1,1}([0,L],\mathbb{R}^{3\times 3})$, is well-defined by

$$l(\boldsymbol{r}, \boldsymbol{D}) := l(\boldsymbol{r}, \boldsymbol{\beta}^D). \tag{42}$$

The following perturbation result for $l(\boldsymbol{r}, \boldsymbol{D})$ is shown in [24, Theorem 4.6].

Lemma 3.4. *Let* $(\boldsymbol{r}, \boldsymbol{D}) \in W^{1,p}([0,L], \mathbb{R}^3) \times W^{1,p}([0,L], \mathbb{R}^{3\times 3})$, $p > 1$, *be a closed framed curve satisfying* (3) *and* $\mathcal{K}[\boldsymbol{r}] \le \theta^{-1}$. *Then there is* $\varepsilon > 0$, *such that* $l(\tilde{\boldsymbol{r}}, \tilde{\boldsymbol{D}})$ *is well-defined and*

$$l(\boldsymbol{r}, \boldsymbol{D}) = l(\tilde{\boldsymbol{r}}, \tilde{\boldsymbol{D}}) \tag{43}$$

for all closed framed curves $(\tilde{\boldsymbol{r}}, \tilde{\boldsymbol{D}}) \in W^{1,p}([0,L], \mathbb{R}^3) \times W^{1,p}([0,L], \mathbb{R}^{3\times 3})$ *satisfying*

$$\|\boldsymbol{r} - \tilde{\boldsymbol{r}}\|_{W^{1,p}} \le \varepsilon, \quad \|\boldsymbol{D} - \tilde{\boldsymbol{D}}\|_{W^{1,p}} \le \varepsilon. \tag{44}$$

4 Variational problem, Euler-Lagrange equations and regularity

The variational problem. In this section we state a general variational problem where we seek energy minimizing closed configuration of elastic rods that are globally injective and belong to prescribed knot and link classes. Then we formulate the corresponding Euler-Lagrange equations satisfied by configurations with minimal energy. Finally we provide regularity results.

For elastic rods determined by elements $w = (u, \boldsymbol{r}_0, \boldsymbol{D}_0) \in X_0^p$ we consider stored energy functionals E_{s} of the form (19) where the stored energy density satisfies (W1),(W2), see Section 2. In addition we consider potential energies E_{p} as given in (21). Let $\boldsymbol{D}_0 = (\boldsymbol{d}_{01}|\boldsymbol{d}_{02}|\boldsymbol{d}_{03})$ and $\boldsymbol{D}_1 = (\boldsymbol{d}_{11}|\boldsymbol{d}_{12}|\boldsymbol{d}_{13})$ be given matrices in $SO(3)$ with equal third column vectors $\boldsymbol{d}_{03} = \boldsymbol{d}_{13}$. Furthermore, let $\theta > 0$ be a given positive constant, $\boldsymbol{r}_0 \in \mathbb{R}^3$ be a given vector, $\boldsymbol{K}_0$ a simple closed curve in $\mathbb{R}^3$, as a representative for a prescribed knot class, and $l_0 \in \mathbb{Z}$ representing a given link class.

Then we look at the minimization problem

$$E(w) := E_{\mathrm{s}}(w) + E_{\mathrm{p}}(w) \;\to\; \text{Min!}\,, \quad w \in X_0^p \tag{45}$$

under the constraints

$$\begin{aligned}
\boldsymbol{r}[w](L) &= \boldsymbol{r}_0, && (46)\\
\boldsymbol{D}[w](L) &= \boldsymbol{D}_1, && (47)\\
\mathcal{R}[w] &\ge \theta, && (48)\\
\boldsymbol{r}[w] &\simeq \boldsymbol{K}_0, && (49)\\
l[w] &= l_0. && (50)
\end{aligned}$$

Here and from now on we use the short notation $\mathcal{R}[w]$, $\mathcal{K}[w]$, $A[w]$, $l[w]$ for $\mathcal{R}[\boldsymbol{r}[w]]$, $\mathcal{K}[\boldsymbol{r}[w]]$, $A[\boldsymbol{r}[w]]$ and $l[\boldsymbol{r}[w]]$. Note that (49) is well-defined because of the constraint (48).

Geometrically, the boundary conditions (46) and (47) lead to closed configurations $(\boldsymbol{r}, \boldsymbol{D})$ with a prescribed angle between $\boldsymbol{d}_1[w](0)$ and $\boldsymbol{d}_1[w](L)$, and (48) guarantees that deformations are globally injective by Lemma 3.1. For the derivation of the Euler-Lagrange equations later on we will have to reformulate the variational problem (45)–(50) with a minimum number of equations, see Section 5.

The existence of solutions for the variational problem (45)–(50) was proven in [11, Sec. 4.2.1] under a natural coercivity condition on W but based on the more restrictive notion of link classes in terms of homotopies in $SO(3)$. By [24, Lemma 4.5] these results can be extended to link classes as considered here, see [30].

Euler-Lagrange equations. The basic issues we shall address here are the derivation of the Euler-Lagrange equations for solutions of the variational problem described above and the presentation of regularity results for the minimizing configurations. We impose the standard growth condition on W_u that

$$\text{(W3)} \qquad |W_u(u,s)| \le c|u|^p + g(s) \ \text{ for } \quad \text{a.e. } s \in [0,L],$$

where $c \ge 0$ is a constant and $g \in L^1([0,L])$. This condition excludes energy densities with the property (20), but it circumvents severe analytical difficulties caused by energies satisfying (20), and it is still an open problem in nonlinear elasticity how to handle these energies for regularity considerations. The existence theory, however, covers these more general energies (cf. [11],[22],[30]).

Theorem 4.1. *Suppose W is a stored energy density satisfying* (W1)–(W3). *Let $w = (u, \boldsymbol{r}_0, \boldsymbol{D}_0) \in X_0^p$ be a solution of the variational problem* (45)–(50), *such that the global curvature $\mathcal{K}[w]$ is locally not attained. Then there exist Lagrange multipliers $\lambda_E \ge 0$, $\boldsymbol{f}_0 \in \mathbb{R}^3, \boldsymbol{m}_0 \in \mathbb{R}^3$ and a Radon measure μ on $[0,L] \times [0,L]$ supported in $A[w]$ (cf.* (33)*), not all zero, such that the following*

Euler-Lagrange equations hold:

$$
\begin{aligned}
0 &= \lambda_E \Big[\hat{\boldsymbol{m}}(u(s),s) - \int_{\Omega_s} \big[\xi^1 \boldsymbol{d}_1[w](t) + \xi^2 \boldsymbol{d}_2[w](t)\big] \wedge d\mathfrak{f}_e(t,\xi^1,\xi^2)\Big] \\
&\quad - \lambda_E \int_s^L \boldsymbol{d}_3[w](t) \wedge \int_{\Omega_t} d\mathfrak{f}_e(\sigma,\xi^1,\xi^2)\,dt \qquad (51)\\
&\quad + \boldsymbol{m}_0 + \int_s^L \boldsymbol{d}_3[w](t) \wedge (\boldsymbol{f}_0 - \boldsymbol{f}_c(t))\,dt \quad \text{for a.e. } s \in [0,L],
\end{aligned}
$$

$$
0 = \lambda_E \int_\Omega d\mathfrak{f}_e(t,\xi^1,\xi^2), \qquad (52)
$$

$$
\begin{aligned}
0 &= \int_0^L \boldsymbol{d}_3[w](t) \wedge \Big[\boldsymbol{f}_0 - \boldsymbol{f}_c(t) - \lambda_E \int_{\Omega_t} d\mathfrak{f}_e(s,\xi^1,\xi^2)\Big]\,dt, \\
&\quad -\lambda_E \int_\Omega [\xi^1 \boldsymbol{d}_1[w](t) + \xi^2 \boldsymbol{d}_2[w](t)] \wedge d\mathfrak{f}_e(t,\xi^1,\xi^2), \qquad (53)
\end{aligned}
$$

where for $\tau \in [0,L]$,

$$
\boldsymbol{f}_c(\tau) := \int_{\mathcal{Q}_\tau} \frac{\boldsymbol{r}[w](s) - \boldsymbol{r}[w](\sigma)}{|\boldsymbol{r}[w](s) - \boldsymbol{r}[w](\sigma)|}\,d\mu(s,\sigma), \qquad (54)
$$

$$
\mathcal{Q}_\tau := \{(s,\sigma) \in [0,L]\times[0,L] : \sigma \le \tau \le s\} \ \text{for} \ \tau \in [0,L]. \qquad (55)
$$

(We have used identity (18) *in* (51).*)*

Moreover, if $\mathcal{R}[w] > \theta$ *in condition* (48)*, then* μ *is the zero measure. In addition, we can choose* $\lambda_E = 1$*, if one of the following* transversality *conditions is satisfied:*

(a) $\boldsymbol{p}[w]$ *admits an* isolated active contact pair*, i.e., there is a point* $(s,\sigma) \in \operatorname{supp}\mu$*, and some* $\varepsilon > 0$*, such that*

$$
\Big[(B_\varepsilon(s) \times [0,L]) \cup ([0,L] \times B_\varepsilon(\sigma))\Big] \cap \operatorname{supp}\mu = (s,\sigma). \qquad (56)
$$

(b) $\boldsymbol{p}[w]$ *has a* curved contact free arc*, i.e., there is an open nonempty interval* $J \subset S_L$ *with* $\boldsymbol{d}_3[w] \not\equiv$ *const. on* J*, and such that*

$$
r(\boldsymbol{r}[w](s), \boldsymbol{r}[w](\sigma), \boldsymbol{r}'[w](\sigma)) > \theta \ \text{for all} \ s \in J, \sigma \in [0,L], s \neq \sigma. \qquad (57)
$$

(c) *There is* $s \in [0,L]$, *such that* $\boldsymbol{r}''[w](s)$ *exists, and such that*

$$(58)\quad \boldsymbol{r}''[w](s) \notin \overline{\mathrm{conv}}(\{\rho(\boldsymbol{r}[w](s) - \boldsymbol{r}[w](\sigma)) : \rho > 0, (s,\sigma) \in \operatorname{supp}\mu\}).$$

Remarks. 1. Using the notation introduced in (13),(14), we recover from (51) the integral from of the equilibrium conditions including the contact term involving f_c, if $\lambda_E = 1$. Moreover, (52) and (53) for $\lambda_E = 1$ express the fact that the resultant force of all external actions must vanish for the whole rod, whereas the resultant couple of all the external actions for the whole rod balances the couple induced by the contact action.

2. Observe that the transversality conditions (a) and (c) are only relevant in the case of contact. If $\mathcal{R}[w] > \theta$, then we can omit the assumption that $\mathcal{K}[w]$ is locally not attained, and we always have $\lambda_E = 1$. Condition (58) in (c) excludes certain "clamped" or rigid configurations where one cannot expect transversality, e.g., as in tightly knotted curves with multiple contact points everywhere. Coleman et al. constructed initially straight, homogeneous inextensible rods furnishing strict local minima for certain quadratic elastic energies with points and lines of self-contact, see [5]. It is unclear, however, whether global minimizers obtained by our existence result may exhibit self-contact everywhere along the curve. Even if this happened to be the case, it appears to be very unlikely apart from very specific cases that all contact points violate (58) in condition (c). In view of this we believe that our transversality conditions (a),(b),(c) cover the generic situation for minimizing configurations. Notice that if multiple contact points are excluded, i.e., if

$$\sharp\{\sigma \in S_L \backslash \{s\} : (s,\sigma) \in \operatorname{supp}\mu\} \le 1 \ \text{ for all } \ s \in S_L,$$

then (c) says that one has to find only one active contact pair $(s,\sigma) \in \operatorname{supp}\mu$, such that $\boldsymbol{r}[w](s) - \boldsymbol{r}[w](\sigma)$ is not parallel to $\boldsymbol{r}''[w](s)$, when the latter exists.

3. Note that $A[w]$ does not contain a certain neighbourhood of the diagonal in $[0,L] \times [0,L]$, since $\mathcal{K}[w]$ is locally not attained. Thus cross sections touching each other cannot be arbitrarily close to each other in arc length. We can even find a constant $\eta = \eta(\boldsymbol{r}) > 0$ such that $|s - \sigma| \ge \eta$ for all $(s,\sigma) \in A[w]$ (cf. Lemma 5.8 in Section 5.1).

4. The measure μ is defined on $[0,L]^2$ and supported on $A[w]$ which is merely a subset of the triangle $\{(s,\sigma) \in [0,L]^2 : \sigma < s\}$. This ensures in particular that each pair of touching cross sections occurs only once in $A[w]$. According to (35) proved in Lemma 5.8 in Section 5.1 the vector $\boldsymbol{r}[w](s) - \boldsymbol{r}[w](\sigma)$ is perpendicular to the tangent vectors $\boldsymbol{r}'[w](s)$ and $\boldsymbol{r}'[w](\sigma)$ for all

$(s,\sigma) \in A[w]$. This together with part (iv) of the following corollary can be interpreted mechanically that in the case when $\mathcal{R}[\boldsymbol{r}] = \theta$ the contact forces are always perpendicular to the curve $\boldsymbol{r}$.

For notational convenience we set

$$F(s,\sigma) := \frac{\boldsymbol{r}[w](s) - \boldsymbol{r}[w](\sigma)}{|\boldsymbol{r}[w](s) - \boldsymbol{r}[w](\sigma)|} \quad \text{for} \quad (s,\sigma) \in [0,L]^2. \tag{59}$$

Corollary 4.2. *Let $\boldsymbol{f}_c$ be as in Theorem 4.1.Then*

(i) $\boldsymbol{f}_c \in BV([0,L],\mathbb{R}^3)$ *and thus, it is bounded.*

(ii) *The right and the left limit of $\boldsymbol{f}_c$, denoted by $\boldsymbol{f}_c(\tau\pm)$, exists for each $\tau \in S_L$, and*

$$\begin{aligned} [\boldsymbol{f}_c](\tau) &:= \boldsymbol{f}_c(\tau+) - \boldsymbol{f}_c(\tau-) \\ &= -\int_{\{\tau\}\times[0,L]} F(s,\sigma)\,d\mu(s,\sigma) + \int_{[0,L]\times\{\tau\}} F(s,\sigma)\,d\mu(s,\sigma). \end{aligned} \tag{60}$$

(iii) *For a.e. $\tau \in [0,L]$ there is a nonnegative Radon measure μ_τ on $[0,L]$ such that*

$$\boldsymbol{f}_c'(\tau) = -\int_0^L F(\tau,\sigma)\,d\mu_\tau(\sigma).$$

(iv)

$$[\boldsymbol{f}_c](\tau) \cdot \boldsymbol{r}'[w](\tau) = 0 \quad \textit{for all} \quad \tau \in [0,L], \tag{61}$$

$$\boldsymbol{f}_c'(\tau) \cdot \boldsymbol{r}'[w](\tau) = 0 \quad \textit{for a.e.}\ \tau \in [0,L]. \tag{62}$$

(v) *The tangential component $\tau \mapsto \boldsymbol{f}_c(\tau) \cdot \boldsymbol{r}'[w](\tau)$ is of class $W^{1,\infty}([0,L])$.*

From the Euler-Lagrange equations we can derive further regularity results for $\boldsymbol{r}[w]$, $\boldsymbol{D}[w]$, and $\boldsymbol{m}(s) = \hat{\boldsymbol{m}}(u(s),s)$.

Corollary 4.3. *If all the hypotheses of Theorem 4.1 including one of the transversality conditions* (a), (b) *or* (c) *hold, then $\boldsymbol{m} \in BV([0,L],\mathbb{R}^3)$. If $\mathfrak{f}_e$ has an integrable density $\boldsymbol{\phi}_e$, i.e., if*

$$d\mathfrak{f}_e(t,\xi^1,\xi^2) = \boldsymbol{\phi}_e(t,\xi^1,\xi^2)\,dtd\xi^1 d\xi^2 \tag{63}$$

with $\boldsymbol{\phi}_e \in L^1(\Omega, \mathbb{R}^3)$*, then* $\boldsymbol{m} \in W^{1,1}([0,L], \mathbb{R}^3)$ *with*

$$
\begin{aligned}
\boldsymbol{m}'(s) = &-\boldsymbol{d}_1[w](s) \wedge \int_D \xi^1 \boldsymbol{\phi}_e(s, \xi^1, \xi^2)\, d\xi^1 d\xi^2 \\
&- \boldsymbol{d}_2[w](s) \wedge \int_D \xi^2 \boldsymbol{\phi}_e(s, \xi^1, \xi^2)\, d\xi^1 d\xi^2 \\
&+ \boldsymbol{r}'[w](s) \wedge \left[\boldsymbol{f}_0 - \boldsymbol{f}_c(s) - \int_s^L \int_D \boldsymbol{\phi}_e(t, \xi^1, \xi^2)\, d\xi^1 d\xi^2\, dt \right]
\end{aligned}
\tag{64}
$$

for a.e. $s \in S_L$*, where* $D := B_\theta(0) \subset \mathbb{R}^2$*. In particular, if* $\boldsymbol{\phi}_e$ *is bounded on* Ω*, , then* $\boldsymbol{m} \in W^{1,\infty}([0,L], \mathbb{R}^3)$*. If* $\mathfrak{f}_e = 0$*, then* $\boldsymbol{m}' \in BV([0,L], \mathbb{R}^3)$ *in addition.*

Let us point out that (64) is the *classical differential form of the equilibrium equation.* Furthermore we note that $W^{1,1}([0,L]) \subset C^0([0,L])$.

Under additional assumptions on the stored energy density W we can derive higher regularity for the strain u, the centrecurve $\boldsymbol{r}[w]$ and the corresponding frame field $\boldsymbol{D}[w]$. Instead of (W1)-(W3) we consider W satisfying (W3) and

(W4) $W(.,.)$ is of class $C^2(\mathbb{R}^3 \times [0,L])$ with $W_{uu}(u,s)$ positive definite for all $u \in \mathbb{R}^3$ and $s \in [0,L]$.

Note that (W4) implies (W1) and (W2). For the following result it actually suffices to assume a C^2-dependence of W with respect to $u \in \mathbb{R}^3$ and only a C^1-dependence with respect to $s \in [0,L]$.

Corollary 4.4. *Let all the hypotheses of Theorem 4.1 including one of the transversality conditions* (a), (b), *or* (c), *be satisfied and, in addition, assume that* W *satisfies* (W3)–(W4). *Then* $u \in BV([0,L], \mathbb{R}^3)$*,* $\boldsymbol{D}[w] \in W^{1,\infty}([0,L], \mathbb{R}^{3\times 3})$*,* $\boldsymbol{D}'[w] \in BV([0,L], \mathbb{R}^{3\times 3})$*,* $\boldsymbol{r}[w] \in W^{2,\infty}([0,L], \mathbb{R}^3)$*, and* $\boldsymbol{r}''[w] \in BV([0,L], \mathbb{R}^3)$.

To deduce higher regularity we assume for simplicity that there are no external forces, (instead of assuming higher regularity of $\mathfrak{f}_e$).

Corollary 4.5. *In addition to the hypotheses in Corollary 4.4 assume that* $\mathfrak{f}_e = 0$*. Then* $u \in W^{1,\infty}([0,L], \mathbb{R}^3)$*,* $\boldsymbol{D}[w] \in W^{2,\infty}([0,L], \mathbb{R}^{3\times 3})$ *and* $\boldsymbol{r}[w] \in W^{3,\infty}([0,L], \mathbb{R}^3)$.

Note that this last result implies that the curvature $|\boldsymbol{r}''[w]|$ is Lipschitz continuous.

If there is a contact free arc, i.e. an interval $J \subset S_L$, such that one has (57), then the standard boot strap arguments for problems without contact yield higher regularity for $u, \boldsymbol{D}, \boldsymbol{r}$ and $\boldsymbol{m}$ on J, as long as $W(.,.)$ and $\mathfrak{f}_e$ are sufficiently smooth.

The special case where W is a quadratic function in u plays an important role in various applications:

$$W(u,s) := \frac{1}{2}C(s)(u(s) - u^{\mathrm{o}}(s)) \cdot (u(s) - u^{\mathrm{o}}(s)), \tag{65}$$

where $C : [0, L] \to \mathbb{R}^{3\times3}$ is a Lebesgue measurable function such that $C(\sigma)$ is symmetric with $\lambda^C_{\min}(\sigma) \geq c > 0$ for a.e. $\sigma \in [0, L]$, where $\lambda^C_{\min}(\sigma)$ denotes the smallest eigenvalue of $C(\sigma)$. The function $u^{\mathrm{o}}(\sigma)$ is the stress-free reference strain as a prescribed material parameter. In this special situation we have more detailed regularity information for u and thus also for $\boldsymbol{r}[w]$ and $\boldsymbol{D}[w]$. For simplicity we assume again that there are no external forces present.

Corollary 4.6. *Let $\mathfrak{f}_e = 0$ and assume that one of the transversality conditions* (a), (b) *or* (c) *holds.*

(i) *If $u^{\mathrm{o}} \in L^r([0, L], \mathbb{R}^3)$ and $C \in L^{2r}([0, L], \mathbb{R}^{3\times3})$ with $p \leq r \leq \infty$, then $u \in L^r([0, L], \mathbb{R}^3)$. Moreover, $\boldsymbol{D}[w] \in W^{1,r}([0, L], \mathbb{R}^{3\times3})$ and $\boldsymbol{r}[w] \in W^{2,r}([0, L], \mathbb{R}^3)$.*

(ii) *If $u^{\mathrm{o}} \in W^{1,\infty}([0, L], \mathbb{R}^3)$ and $C \in W^{1,\infty}([0, L], \mathbb{R}^{3\times3})$, then $u \in W^{1,\infty}([0, L], \mathbb{R}^3)$. Moreover, $\boldsymbol{D}[w] \in W^{2,\infty}([0, L], \mathbb{R}^{3\times3})$ and $\boldsymbol{r}[w] \in W^{3,\infty}([0, L], \mathbb{R}^3)$.*

As before, by virtue of boot strap arguments, one gets higher regularity of $u, \boldsymbol{D}[w], \boldsymbol{r}[w]$ and $\boldsymbol{m}$ on parts of the rod without contact, if C, u^{o} and $\mathfrak{f}_e$ are smooth enough.

5 Proofs

5.1 Proof of Theorem 4.1 and and Corollary 4.2

We are going to proceed in several steps, and we always assume that $w = (u, \boldsymbol{r}_0, \boldsymbol{D}_0) \in X^p_0$ is a solution of the variational problem (45)–(50) such that the global curvature $\mathcal{K}[w]$ is locally not attained.

Modified variational problem. First we provide a method to represent small variations of $\boldsymbol{D}_0$ on the manifold $SO(3)$ by variations in a linear space. Notice that small perturbations of $\boldsymbol{D}_0$ have the form $\boldsymbol{D}_0 \overset{\triangle}{\boldsymbol{D}}$, where $\overset{\triangle}{\boldsymbol{D}} \in SO(3)$ is close to the identity. Such matrices $\overset{\triangle}{\boldsymbol{D}}$ can be represented in a unique way by means of the rotation vector $\overset{\triangle}{\boldsymbol{\alpha}} = \overset{\triangle}{\boldsymbol{\alpha}} (\overset{\triangle}{\boldsymbol{D}}) \in \mathbb{R}^3$, where the direction of $\overset{\triangle}{\boldsymbol{\alpha}}$ describes the rotation axis of $\overset{\triangle}{\boldsymbol{D}}$, and the length $| \overset{\triangle}{\boldsymbol{\alpha}} |$ equals the positively oriented rotation angle in $[0, \pi)$. In a neighbourhood of the identity in $SO(3)$, the mapping $\overset{\triangle}{\boldsymbol{D}} \mapsto \overset{\triangle}{\boldsymbol{\alpha}} (\overset{\triangle}{\boldsymbol{D}})$ is continuous and has a continuous inverse mapping $\boldsymbol{U}$ in a neighbourhood of the origin in $\mathbb{R}^3$. In particular, we have $\overset{\triangle}{\boldsymbol{\alpha}} (Id) = 0 \in \mathbb{R}^3$ and $\boldsymbol{U}(0) = Id \in SO(3)$. Thus small perturbations of $\boldsymbol{D}_0 \in SO(3)$ have the form $\boldsymbol{D}_0 \boldsymbol{U}(\overset{\triangle}{\boldsymbol{\alpha}})$ with $\overset{\triangle}{\boldsymbol{\alpha}} \in \mathbb{R}^3$, $\overset{\triangle}{\boldsymbol{\alpha}}$ close to $0 \in \mathbb{R}^3$, and we can identify each slightly perturbed configuration

$$(u + \overset{\triangle}{u}, \boldsymbol{r}_0 + \overset{\triangle}{\boldsymbol{r}}_0, \boldsymbol{D}_0 \overset{\triangle}{\boldsymbol{D}}) \in X_0^p$$

with an element

$$\overset{\triangle}{w} = (\overset{\triangle}{u}, \overset{\triangle}{\boldsymbol{r}}_0, \overset{\triangle}{\boldsymbol{\alpha}}) \in L^p([0, L], \mathbb{R}^3) \times \mathbb{R}^3 \times \mathbb{R}^3$$

by $\overset{\triangle}{\boldsymbol{D}} = \boldsymbol{U}(\overset{\triangle}{\boldsymbol{\alpha}})$.

Since certain arguments in our proof below only work as long as we consider perturbed configurations where the global curvature is locally not attained either, we have to restrict our analysis to variations of the form

$$\overset{\triangle}{w} := (\overset{\triangle}{u}, \overset{\triangle}{\boldsymbol{r}}_0, \overset{\triangle}{\boldsymbol{\alpha}}) \in L^\infty([0, L], \mathbb{R}^3) \times \mathbb{R}^3 \times \mathbb{R}^3 =: Y \tag{66}$$

instead of taking $\overset{\triangle}{u} \in L^p([0, L], \mathbb{R}^3)$. With the norm

$$\| \overset{\triangle}{w} \|_Y := \| \overset{\triangle}{u} \|_{L^\infty} + | \overset{\triangle}{\boldsymbol{r}}_0 | + | \overset{\triangle}{\boldsymbol{\alpha}} | \quad \text{for} \quad \overset{\triangle}{w} = (\overset{\triangle}{u}, \overset{\triangle}{\boldsymbol{r}}_0, \overset{\triangle}{\boldsymbol{\alpha}}) \in Y \tag{67}$$

$(Y, \|.\|_Y)$ is a Banach space, whereas the original set X_0^p is not a linear space. For notational convenience we introduce the modified energy function

$$\breve{E}(\overset{\triangle}{w}) := E((u + \overset{\triangle}{u}, \boldsymbol{r}_0 + \overset{\triangle}{\boldsymbol{r}}_0, \boldsymbol{D}_0 \boldsymbol{U}(\overset{\triangle}{\boldsymbol{\alpha}}))) \quad \text{for} \quad \overset{\triangle}{w} \in B_\delta(0) \subset Y, \tag{68}$$

where $B_\delta(0)$ is a small neighbourhood of $0 \in Y$ with $\delta > 0$ not fixed but sufficiently small. Analogously, we define $\breve{E}_{\mathrm{s}}(\overset{\triangle}{w})$, $\breve{E}_{\mathrm{p}}(\overset{\triangle}{w})$, $\breve{\boldsymbol{r}}[\overset{\triangle}{w}]$, $\breve{\boldsymbol{D}}[\overset{\triangle}{w}]$, $\breve{\mathcal{R}}[\overset{\triangle}{w}]$, etc. Note that $\breve{\boldsymbol{r}}[0] = \boldsymbol{r}[w]$, $\breve{\boldsymbol{D}}[0] = \boldsymbol{D}[w]$, etc.

Now we consider the *modified variational problem*

$$\breve{E}(\overset{\vartriangle}{w}) \longrightarrow \quad \text{Min!}, \quad \overset{\vartriangle}{w} \in Y, \tag{69}$$

subjected to

$$\begin{aligned}
\breve{\boldsymbol{r}}[\overset{\vartriangle}{w}](L) &= \boldsymbol{r}_0 + \overset{\vartriangle}{\boldsymbol{r}}_0, && (70)\\
\breve{\boldsymbol{D}}[\overset{\vartriangle}{w}](L) &= \boldsymbol{D}_1 \boldsymbol{U}(\overset{\vartriangle}{\boldsymbol{\alpha}}), && (71)\\
\breve{\mathcal{R}}[\overset{\vartriangle}{w}] &\geq \theta, && (72)\\
\breve{\boldsymbol{r}}[\overset{\vartriangle}{w}] &\simeq \boldsymbol{K}_0, && (73)\\
\breve{l}[\overset{\vartriangle}{w}] &= l_0. && (74)
\end{aligned}$$

Notice as before that the linking number $\breve{l}[\overset{\vartriangle}{w}]$ is well-defined by (42) for $\overset{\vartriangle}{w} \in Y$, by (72) and Lemma 3.1. Since $L^\infty([0,L],\mathbb{R}^3) \hookrightarrow L^p([0,L],\mathbb{R}^3)$

$$\overset{\vartriangle}{w} = 0 \quad \text{is a local minimizer of (69)–(74).} \tag{75}$$

Reduction of the modified problem. It turns out that some of the constraints of the modified variational problem are redundant which would imply difficulties in obtaining $\lambda_E = 1$ as we claim in the second part of Theorem 4.1. Furthermore we will replace Condition (72) by an equivalent condition with a functional having better differentiability properties than $\breve{\mathcal{R}}[.]$.

First we state the following simple regularity and convergence results for the solutions of the system (4).

Lemma 5.1. (i) *Let l be a nonnegative integer and $1 \leq r \leq \infty$. If $u \in W^{l,r}(I,\mathbb{R}^3)$ then $\boldsymbol{D} \in W^{l+1,r}(I,\mathbb{R}^{3\times3})$ and $\boldsymbol{r} \in W^{l+2,r}(I,\mathbb{R}^{3\times3})$. If $u \in C^{l,\alpha}(I,\mathbb{R}^3)$ for some $\alpha \in [0,1]$, then $\boldsymbol{D} \in C^{l+1,\alpha}(\bar{I},\mathbb{R}^{3\times3})$ and $\boldsymbol{r} \in C^{l+2,\alpha}(\bar{I},\mathbb{R}^{3\times3})$*

(ii) *Let $1 < p < \infty$. If $w_n \rightharpoonup w$ in X^p, where $\{w_n\} \subset X_0^p$, then $w \in X_0^p$ and*

$$\boldsymbol{D}_n \to \boldsymbol{D} \ \text{ in } \ C^0(\bar{I},\mathbb{R}^{3\times3}), \qquad \boldsymbol{r}_n \to \boldsymbol{r} \ \text{ in } \ C^0(\bar{I},\mathbb{R}^3), \tag{76}$$

$$\boldsymbol{D}_n \rightharpoonup \boldsymbol{D} \ \text{ in } \ W^{1,p}(I,\mathbb{R}^{3\times3}), \qquad \boldsymbol{r}_n \rightharpoonup \boldsymbol{r} \ \text{ in } \ W^{2,p}(I,\mathbb{R}^3), \tag{77}$$

where $\boldsymbol{r}_n := \boldsymbol{r}[w_n]$, $\boldsymbol{r} := \boldsymbol{r}[w]$, $\boldsymbol{D}_n := \boldsymbol{D}[w_n]$, $\boldsymbol{D} := \boldsymbol{D}[w]$.

(iii) *Let* $1 < p \leq \infty$. *If* $w_n \to w$ *in* X^p, *where* $\{w_n\} \subset X_0^p$, *then*

$$\boldsymbol{d}_{k,n} \to \boldsymbol{d}_k \ \text{ in } \ W^{1,p}(I,\mathbb{R}^3),\ k=1,2,3, \ \text{ and } \ \boldsymbol{r}_n \to \boldsymbol{r} \ \text{ in } \ W^{2,p}(I,\mathbb{R}^3). \tag{78}$$

Proof. (i) Starting with $l=0$, i.e., $u \in L^r(I,\mathbb{R}^3)$, the right-hand side of the first equation in (4) is in $L^r(I,\mathbb{R}^3)$, hence $\boldsymbol{d}_k'$, $k=1,2,3$, as well, since on the right-hand side, $\boldsymbol{d}_k \in W^{1,p}(I,\mathbb{R}^3) \hookrightarrow C^0(\bar{I},\mathbb{R}^3)$. Thus $\boldsymbol{D} \in W^{1,r}(I,\mathbb{R}^{3\times 3})$. For $l \geq 1$ use bootstrap arguments inductively. The other results follow easily from the last equation in (4).

Part (ii) was essentially proven already in [11, Lemma 8]. The stronger convergence for $\{\boldsymbol{r}_n\}$ follows from the last equation in (4).

(iii) Let $k=1,(k=2,3$ can be treated in the same way). Using the orthonormality of the $\boldsymbol{d}_i$ we can rewrite the equation for $\boldsymbol{d}_1$ in (4) as

$$\boldsymbol{d}_1'(s) = u^3(s)\boldsymbol{d}_2(s) - u^2(s)\boldsymbol{d}_3(s) \ \text{ for a.e. } s \in I. \tag{79}$$

Subtracting (79) from the corresponding equation for $\boldsymbol{d}_{1,n}$ we obtain

$$\begin{aligned} \boldsymbol{d}_{1,n}'(s) - \boldsymbol{d}_1'(s) &= (u_n^3 - u^3)\boldsymbol{d}_{2,n}(s) + u^3(\boldsymbol{d}_{2,n}(s) - \boldsymbol{d}_2(s)) \\ &\quad - (u_n^2 - u^2)\boldsymbol{d}_{3,n}(s) - u^2(\boldsymbol{d}_{3,n}(s) - \boldsymbol{d}_3(s)), \end{aligned} \tag{80}$$

for a.e. $s \in I$. Taking the L^p-norm, we get

$$\begin{aligned} \|\boldsymbol{d}_{1,n}' - \boldsymbol{d}_1'\|_{L^p} &\leq \sum_{i=2}^3 \|u_n^i - u^i\|_{L^p}\|D_n\|_{C^0} + \sum_{i=2}^3 \|\boldsymbol{d}_{i,n} - \boldsymbol{d}_i\|_{C^0}\|u\|_{L^p} \\ &\to 0 \ \text{ as } \ n \to \infty, \end{aligned}$$

where one uses (76) on the right-hand side, which holds even for $p=\infty$, since strong convergence in $L^\infty(I,\mathbb{R}^3)$ implies weak convergence in $L^{\tilde{p}}(I,\mathbb{R}^3)$ for all $\tilde{p} \in [1,\infty)$. Thus $\boldsymbol{d}_{1,n}' \to \boldsymbol{d}_1'$ in $L^p([0,L],\mathbb{R}^3)$ and $\boldsymbol{d}_{1,n} \to \boldsymbol{d}_1$ in $L^p([0,L],\mathbb{R}^3)$ by (76), which implies the first statement in (78). For the second statement in (78) we argue in the same way and use the last equation in (4) in addition. □

For the minimizing configuration we deduce the following regularity properties:

Lemma 5.2. *Let* $w = (u, \boldsymbol{r}_0, D_0) \in X_0^p$ *be a solution of the variational problem* (45)–(50). *Then*

(i) $u^1, u^2 \in L^\infty([0,L])$,

(ii) $\boldsymbol{d}_3[w], \breve{\boldsymbol{d}}_3[\overset{\triangle}{w}] \in W^{1,\infty}([0,L],\mathbb{R}^3)$, *and* $\boldsymbol{r}[w], \breve{\boldsymbol{r}}[\overset{\triangle}{w}] \in W^{2,\infty}([0,L],\mathbb{R}^3)$ *for any* $\overset{\triangle}{w} \in B_\delta(0) \subset Y$.

Proof of Lemma 5.2. Since $\mathcal{R}[w] \geq \theta > 0$, (6),(25),(48) imply

$$\sqrt{(u^1(s))^2 + (u^2(s))^2} \leq \|\boldsymbol{r}''\|_{L^\infty} \leq \mathcal{R}[w]^{-1} \leq \theta^{-1} < \infty \text{ for a.e. } s \in S_L,$$

i.e., $u^1, u^2 \in L^\infty([0,L])$, which shows part (i).

By the differential system (4) we have

$$\begin{aligned} \boldsymbol{d}_3'[w] &= u^2 \boldsymbol{d}_1[w] - u^1 \boldsymbol{d}_2[w], \\ \boldsymbol{r}'[w] &= \boldsymbol{d}_3[w]. \end{aligned}$$

Arguing as in the proof of Lemma 5.1 we obtain (ii) for $\boldsymbol{d}_3[w], \boldsymbol{r}[w]$. If we replace $w = (u, \boldsymbol{r}_0, \boldsymbol{D}_0)$ with $(u + \overset{\triangle}{u}, \boldsymbol{r}_0 + \overset{\triangle}{\boldsymbol{r}}_0, \boldsymbol{D}_0 \boldsymbol{U}(\overset{\triangle}{\boldsymbol{\alpha}}))$ and solve the perturbed differential system

$$\begin{aligned} \breve{\boldsymbol{d}}_k{}'[\overset{\triangle}{w}](s) &= \Big[\sum_{i=1}^{3} (u^i + \overset{\triangle}{u}{}^i)(s) \breve{\boldsymbol{d}}_i[\overset{\triangle}{w}](s) \Big] \wedge \breve{\boldsymbol{d}}_k[\overset{\triangle}{w}](s), \\ \breve{\boldsymbol{r}}'[\overset{\triangle}{w}](s) &= \breve{\boldsymbol{d}}_3[\overset{\triangle}{w}](s), \\ \breve{\boldsymbol{r}}[\overset{\triangle}{w}](0) &= \boldsymbol{r}_0 + \overset{\triangle}{\boldsymbol{r}}_0, \quad \breve{\boldsymbol{D}}[\overset{\triangle}{w}](0) = \boldsymbol{D}_0 \boldsymbol{U}(\overset{\triangle}{\boldsymbol{\alpha}}), \end{aligned} \tag{81}$$

for a.e. $s \in [0,L]$, $k = 1,2,3$, then we get the remaining statement in part (ii) in the same way. $\square$

As a consequence of Lemma 5.2 we observe that small variations of w of the kind described above do not violate the topological constraints.

Lemma 5.3. *Let* $w = (u, \boldsymbol{r}_0, D_0) \in X_0^p$ *be a solution of the variational problem* (45)–(50). *Then*

(i) $\breve{\boldsymbol{r}}[\overset{\triangle}{w}] \simeq \boldsymbol{r}[w]$ *for all* $\| \overset{\triangle}{w} \|_Y$ *sufficiently small.*

(ii) *For all* $\| \overset{\triangle}{w} \|_Y$ *sufficiently small satisfying*

$$\breve{\boldsymbol{r}}[\overset{\triangle}{w}](L) = \breve{\boldsymbol{r}}[\overset{\triangle}{w}](0) \quad \text{and} \quad \breve{\boldsymbol{D}}[\overset{\triangle}{w}](L) = \boldsymbol{D}_1 \boldsymbol{U}(\overset{\triangle}{\boldsymbol{\alpha}}), \tag{82}$$

one has $\breve{l}[\overset{\triangle}{w}] = l[w] = l_0$.

Proof. (i) (78) of Lemma 5.1 implies that

$$\|\boldsymbol{r}[w] - \check{\boldsymbol{r}}[\overset{\triangle}{w}]\|_{W^{2,\infty}} + \|\boldsymbol{D}[w] - \check{\boldsymbol{D}}[\overset{\triangle}{w}]\|_{W^{1,p}} \to 0 \text{ as } \| \overset{\triangle}{w} \|_Y \to 0, \tag{83}$$

and notice that the convergence in $W^{2,\infty}$ is equivalent to convergence in $C^{1,1}$. Hence for $\| \overset{\triangle}{w} \|_Y$ sufficiently small we have

$$\check{\mathcal{K}}[\overset{\triangle}{w}] \leq 2\theta^{-1}, \tag{84}$$

by the continuity of $\mathcal{K}[.]$ with respect to the convergence in (83), see Lemma 3.2. Now apply Lemma 3.3 for $\boldsymbol{r} = \boldsymbol{r}[w]$ and $C_0 = 2\theta^{-1}$ with $\| \overset{\triangle}{w} \|_Y$ so small that $\|\boldsymbol{r}[w] - \check{\boldsymbol{r}}[\overset{\triangle}{w}]\|_{C^0} \leq \varepsilon$, where $\varepsilon = \varepsilon(\boldsymbol{r}, 2\theta^{-1})$ is as in Lemma 3.3.

(ii) If we extend the curves $\boldsymbol{r}[w]$, $\check{\boldsymbol{r}}[\overset{\triangle}{w}]$, and $\boldsymbol{r}[w] + (\theta/2)\boldsymbol{d}_1[w]$, $\check{\boldsymbol{r}}[\overset{\triangle}{w}] + (\theta/w)\check{\boldsymbol{d}}_1[\overset{\triangle}{w}]$ according to (40) and (39), respectively, we readily infer from (82) that all these curves have the interval $[0, L+1]$ as their common domain. Now apply Lemma 3.4 for $\| \overset{\triangle}{w} \|_Y$ sufficiently small to conclude the proof. $\square$

Lemma 5.3 implies that the topological constraints are stable with respect to small variations in Y. Thus they can be removed without affecting the fact that $\overset{\triangle}{w}= 0$ is a local minimizer of the modified variational problem.

In order to replace (72) by an equivalent condition we introduce the functions

$$P[\overset{\triangle}{w}](s,\sigma) := (\check{\boldsymbol{r}}[\overset{\triangle}{w}](s), \check{\boldsymbol{r}}[\overset{\triangle}{w}](\sigma), \check{\boldsymbol{r}}'[\overset{\triangle}{w}](\sigma)), \tag{85}$$

$$H(\boldsymbol{x},\boldsymbol{y},\boldsymbol{t}) := \frac{4|(\boldsymbol{x}-\boldsymbol{y}) \wedge \boldsymbol{t}|^2}{|\boldsymbol{x}-\boldsymbol{y}|^4}, \text{ for } \boldsymbol{x},\boldsymbol{y},\boldsymbol{t} \in \mathbb{R}^3, \boldsymbol{t} \neq 0, \tag{86}$$

and note that according to (31),(32) we may write

$$\check{\mathcal{K}}[\overset{\triangle}{w}]^2 = \sup_{\substack{s,\sigma \in S_L \\ s \neq \sigma}} H(P[\overset{\triangle}{w}](s,\sigma)). \tag{87}$$

By (29) we can replace (72) with

$$g(\overset{\triangle}{w}) := \check{\mathcal{K}}[\overset{\triangle}{w}]^2 - \theta^{-2} \leq 0. \tag{88}$$

To remove the redundancies in the boundary conditions we are going to replace the nine scalar conditions (71) by just three scalar equations, see

(92)–(94) below. (Note that an element of $SO(3)$ has merely three degrees of freedom.)

This way we get the *reduced variational problem*

$$\breve{E}(\overset{\triangle}{w}) \;\to\; \text{Min!}\,, \quad \overset{\triangle}{w}\in Y, \tag{89}$$

subjected to

$$\begin{aligned}
g(\overset{\triangle}{w}) &\;\leq\; 0, && (90)\\
\boldsymbol{g}_0(\overset{\triangle}{w}) := \breve{\boldsymbol{r}}[\overset{\triangle}{w}](L) - (\boldsymbol{r}_0 + \overset{\triangle}{\boldsymbol{r}}_0) &\;=\; 0, && (91)\\
g_1(\overset{\triangle}{w}) := \breve{\boldsymbol{d}}_1[\overset{\triangle}{w}](L)\cdot(\boldsymbol{D}_1\boldsymbol{U}(\overset{\triangle}{\boldsymbol{\alpha}}))_2 &\;=\; 0, && (92)\\
g_2(\overset{\triangle}{w}) := \breve{\boldsymbol{d}}_3[\overset{\triangle}{w}](L)\cdot(\boldsymbol{D}_0\boldsymbol{U}(\overset{\triangle}{\boldsymbol{\alpha}}))_1 &\;=\; 0, && (93)\\
g_3(\overset{\triangle}{w}) := \breve{\boldsymbol{d}}_3[\overset{\triangle}{w}](L)\cdot(\boldsymbol{D}_0\boldsymbol{U}(\overset{\triangle}{\boldsymbol{\alpha}}))_2 &\;=\; 0, && (94)
\end{aligned}$$

where, for $\boldsymbol{M} \in \mathbb{R}^{3\times 3}$, we denoted the k-th column vector by $(\boldsymbol{M})_k$, $k = 1, 2, 3$.

Lemma 5.4. *The reduced variational problem* (89)–(94) *has a local minimizer at* $\overset{\triangle}{w}= 0$.

Proof. In Lemma 5.3 it was shown that small variations do not violate the topological constraints, hence (73) and (74) hold for all $\| \overset{\triangle}{w} \|_Y$ sufficiently small. Conditions (92)–(94) determine the frame $\breve{\boldsymbol{D}}[\overset{\triangle}{w}](L)$ to be equal to $\boldsymbol{D}_1(\boldsymbol{U}(\overset{\triangle}{\boldsymbol{\alpha}}))$. Indeed, $\boldsymbol{d}_{13} = \boldsymbol{d}_{03}$ by assumption on $\boldsymbol{D}_1$. Thus (93),(94) force $\breve{\boldsymbol{d}}_3[\overset{\triangle}{w}](L)$ to be parallel to $(\boldsymbol{D}_1\boldsymbol{U}(\overset{\triangle}{\boldsymbol{\alpha}}))_3$, and by continuity (see Lemma 5.1) we get $\breve{\boldsymbol{d}}_3[\overset{\triangle}{w}](L) = (\boldsymbol{D}_1\boldsymbol{U}(\overset{\triangle}{\boldsymbol{\alpha}}))_3$ for $\| \overset{\triangle}{w} \|_Y$ small. Now (92) implies that $\breve{\boldsymbol{d}}_1[\overset{\triangle}{w}](L)$ is perpendicular to $(\boldsymbol{D}_1\boldsymbol{U}(\overset{\triangle}{\boldsymbol{\alpha}}))_2$, and $\breve{\boldsymbol{d}}_1[\overset{\triangle}{w}](L)$ is automatically perpendicular to $(\boldsymbol{D}_1\boldsymbol{U}(\overset{\triangle}{\boldsymbol{\alpha}}))_3 = \breve{\boldsymbol{d}}_3[\overset{\triangle}{w}](L)$. Again by continuity, we get $\breve{\boldsymbol{d}}_1[\overset{\triangle}{w}](L) = (\boldsymbol{D}_1\boldsymbol{U}(\overset{\triangle}{\boldsymbol{\alpha}}))_1$ for $\| \overset{\triangle}{w} \|_Y$ small. Since $\breve{\boldsymbol{D}}[\overset{\triangle}{w}]$, $\boldsymbol{D}_1\boldsymbol{U}(\overset{\triangle}{\boldsymbol{\alpha}}) \in SO(3)$, we still obtain $\breve{\boldsymbol{d}}_2[\overset{\triangle}{w}](L) = (\boldsymbol{D}_1\boldsymbol{U}(\overset{\triangle}{\boldsymbol{\alpha}}))_2$ for $\| \overset{\triangle}{w} \|_Y$ small. Thus $\breve{\boldsymbol{D}}[\overset{\triangle}{w}](L) = \boldsymbol{D}_1\boldsymbol{U}(\overset{\triangle}{\boldsymbol{\alpha}})$, i.e., (92)–(94) imply (71). (70) and (72) are obviously equivalent to (91) and (90), respectively. Since $\overset{\triangle}{w}= 0$ is a local minimizer of (69)–(74), it is also a local minimizer of (89)–(94). □

We will now derive the Euler-Lagrange equations for the reduced variational problem, instead of (45)–(50) or (69)–(74). For that purpose we have to compute a number of derivatives.

Differentiability of the base curve and the directors. In order to analyze the dependence of the energy functions $\breve{E}_{\mathrm{s}}$, $\breve{E}_{\mathrm{p}}$, and the side conditions on perturbations $\overset{\triangle}{w} \in Y$ we need to understand, how the solutions of the perturbed differential system (81) depend on $\overset{\triangle}{w}$. According to [23, Thm 2.1] the solutions of (81) depend *continuously differentiable* on the perturbations $\overset{\triangle}{u} \in L^\infty([0,L],\mathbb{R}^3)$, $\overset{\triangle}{\boldsymbol{r}} \in \mathbb{R}^3$, and $\overset{\triangle}{D} \in SO(3)$. Since the mapping $\overset{\triangle}{\boldsymbol{\alpha}} \mapsto \boldsymbol{U}(\overset{\triangle}{\boldsymbol{\alpha}})$ is smooth in a small neighbourhood of $0 \in \mathbb{R}^3$, we obtain

Lemma 5.5. *Let w be a solution of* (45)–(50)*. Then the mapping*

$$(\overset{\triangle}{w}, s) \mapsto (\breve{\boldsymbol{r}}[\overset{\triangle}{w}](s), \breve{\boldsymbol{D}}[\overset{\triangle}{w}\,](s))$$

from $B_\delta(0) \times [0,L]$ into $\mathbb{R}^3 \times \mathbb{R}^{3\times 3}$ is continuously differentiable for some sufficiently small $\delta > 0$ (depending on w), i.e.,

$$(\breve{\boldsymbol{r}}[.](.), \breve{\boldsymbol{D}}[.](.)) \in C^1(B_\delta(0) \times [0,L], \mathbb{R}^3 \times \mathbb{R}^{3\times 3}). \tag{95}$$

Note. Since we study continuity and differentiability near the origin in Y, it is sufficient to take a bounded neighbourhood of the origin in $L^\infty([0,L],\mathbb{R}^3)$ as parameter set Λ in [23, Theorem 2.1], which corresponds to perturbations $\overset{\triangle}{u}$. This way [23, Theorem 2.1] implies the desired regularity (95), but only for small intervals instead of for $[0,L]$. Since the system (81) is always uniquely solvable on $[0,L]$ and, by uniform boundedness of the solution, even on $[-\varepsilon, L+\varepsilon]$ for any given $\varepsilon > 0$ (cf. [11, Lemma 6]), we obtain (95) with $[0,L]$ by a covering argument using the compactness of $[0,L]$.

Since $\boldsymbol{r}$ and $\boldsymbol{d}_k$ enter explicitly into the potential energy $\breve{E}_{\mathrm{p}}(.)$ and the side conditions (90)–(94), we need to calculate the Fréchet derivative of the mappings $\overset{\triangle}{w} \mapsto \breve{\boldsymbol{r}}[\overset{\triangle}{w}](s)$ and $\overset{\triangle}{w} \mapsto \breve{\boldsymbol{d}}_k[\overset{\triangle}{w}]$, $k = 1,2,3$, at the origin $0 \in Y$, which we denote by $\partial_w \breve{\boldsymbol{r}}[0](s)$, $\partial_w \breve{\boldsymbol{d}}_k[0](s)$, respectively. Lemma A.1 in the appendix shows that

$$\partial_w \breve{\boldsymbol{d}}_k[0](s)\, \overset{\triangle}{w} = \boldsymbol{z}(s) \wedge \boldsymbol{d}_k[w](s), \quad k = 1,2,3, \tag{96}$$

and thus

$$\partial_w \breve{\boldsymbol{r}}[0](s)\, \overset{\triangle}{w} = \overset{\triangle}{\boldsymbol{r}}_0 + \int_0^s \boldsymbol{z}(\tau) \wedge \boldsymbol{d}_3[w](\tau)\, d\tau, \tag{97}$$

for all $s \in [0,L]$, $\overset{\triangle}{w}= (\overset{\triangle}{u}, \overset{\triangle}{\boldsymbol{r}}_0, \overset{\triangle}{\boldsymbol{\alpha}}) \in Y$. Here, $\boldsymbol{z} = \boldsymbol{z}[\overset{\triangle}{u}]$ is a special characterization of elements $\overset{\triangle}{u} \in L^\infty([0,L],\mathbb{R}^3)$ by the uniquely assigned function

$$\boldsymbol{z}(s) = \boldsymbol{z}(0) + \int_0^s \sum_{i=1}^3 \overset{\triangle}{u}{}^i(\tau)\boldsymbol{d}_i[w](\tau)\,d\tau \tag{98}$$

with

$$\boldsymbol{z}(0) \wedge \boldsymbol{d}_k[w](0) = (\boldsymbol{D}_0\boldsymbol{U}'(0)\,\overset{\triangle}{\boldsymbol{\alpha}})_k, \quad k = 1,2,3, \tag{99}$$

where $\boldsymbol{U}'$ denotes the derivative of $\boldsymbol{U}$ with respect to $\boldsymbol{\alpha}$ at $0 \in \mathbb{R}^3$, see our remark at the end of Appendix A. Note that $\boldsymbol{z} \in W^{1,\infty}([0,L],\mathbb{R}^3)$. In particular,

$$\boldsymbol{z}(0) = 0 \quad \text{ for } \quad \overset{\triangle}{w}= (\overset{\triangle}{u}, \overset{\triangle}{r}_0, 0) \in Y. \tag{100}$$

Differentiability of the energy E.

Lemma 5.6. *Let w be a solution of* (45)–(50). *Then the energy functions*

$$\breve{E}_{\mathrm{s}}, \breve{E}_{\mathrm{p}} : B_\delta(0) \subset Y \longrightarrow \mathbb{R}$$

are continuously differentiable for some sufficiently small $\delta > 0$ (depending on w), and we have

$$\begin{aligned} \breve{E}_{\mathrm{s}}'(0)\,\overset{\triangle}{w} &= \int_0^L W_u(u(t),t)\cdot \overset{\triangle}{u}(t)\,dt \\ &= \int_0^L \boldsymbol{z}'(t)\cdot \sum_{i=1}^3 W_{u^i}(u(t),t)\boldsymbol{d}_i[w](t)\,dt \end{aligned} \tag{101}$$

and

$$\begin{aligned} \breve{E}_{\mathrm{p}}'(0)\,\overset{\triangle}{w} = &-\overset{\triangle}{\boldsymbol{r}}_0 \cdot \int_\Omega d\mathfrak{f}_e(t,\xi^1,\xi^2) \\ &- \int_0^L \boldsymbol{z}'(t)\cdot \int_t^L \boldsymbol{d}_3[w](\tau) \wedge \int_{\Omega_\tau} d\mathfrak{f}_e(s,\xi^1,\xi^2)\,d\tau dt \\ &- \int_0^L \boldsymbol{z}'(t)\cdot \int_{\Omega_t} \left[\xi^1\boldsymbol{d}_1[w](\tau) + \xi^2\boldsymbol{d}_2[w](\tau)\right] \wedge d\mathfrak{f}_e(\tau,\xi^1,\xi^2)dt \\ &- \boldsymbol{z}(0)\cdot \int_\Omega \left[\xi^1\boldsymbol{d}_1[w](t) + \xi^2\boldsymbol{d}_2[w](t)\right] \wedge d\mathfrak{f}_e(t,\xi^1,\xi^2) \\ &- \boldsymbol{z}(0)\cdot \int_0^L \boldsymbol{d}_3[w](t) \wedge \int_{\Omega_t} d\mathfrak{f}_e(\sigma,\xi^1,\xi^2)\,dt, \end{aligned} \tag{102}$$

for all $\overset{\triangle}{w}= (\overset{\triangle}{u}, \overset{\triangle}{\boldsymbol{r}}_0, \overset{\triangle}{\boldsymbol{\alpha}}) \in Y$*, where* $\boldsymbol{z} \in W^{1,\infty}([0, L], \mathbb{R}^3)$ *is given by (98),(99).*

Proof. Recall that

$$\breve{E}_{\mathrm{s}}(\overset{\triangle}{w}) = \int_0^L W(u(s)+ \overset{\triangle}{u}(s), s)\, ds, \tag{103}$$

$$\breve{E}_{\mathrm{p}}(\overset{\triangle}{w}) = -\int_\Omega (\breve{\boldsymbol{r}}[\overset{\triangle}{w}](s) + \xi^1\breve{\boldsymbol{d}}_1[\overset{\triangle}{w}](s) + \xi^2\breve{\boldsymbol{d}}_2[\overset{\triangle}{w}](s)) \cdot d\mathfrak{f}_e(s, \xi^1, \xi^2). \tag{104}$$

Conditions (W1)–(W3) on W imply that $\breve{E}_{\mathrm{s}}(.)$ is Fréchet-differentiable, and we obtain (101) by standard arguments and (98). (Notice that the integral on the right-hand side exists by (W3) and the fact that $u \in L^p([0, L], \mathbb{R}^3)$.) We can differentiate in (104) with respect to $\overset{\triangle}{w}$ under the integral sign, because the integrand as well as its Fréchet derivative have integrable majorants. Using (96)–(99) we obtain for $\overset{\triangle}{w}= (\overset{\triangle}{u}, \overset{\triangle}{\boldsymbol{r}}_0, \overset{\triangle}{\boldsymbol{\alpha}}) \in Y$

$$\begin{aligned}\breve{E}'_{\mathrm{p}}(0)\, \overset{\triangle}{w} \;=\; & -\int_\Omega \left[\overset{\triangle}{\boldsymbol{r}}_0 + \int_0^s \boldsymbol{z}(t) \wedge \boldsymbol{d}_3[w](t)\, dt\right] \cdot d\mathfrak{f}_e(s, \xi^1, \xi^2)\\ & -\int_\Omega \left[\xi^1\boldsymbol{z}(s) \wedge \boldsymbol{d}_1[w](s) + \xi^2\boldsymbol{z}(s) \wedge \boldsymbol{d}_2[w](s)\right] \cdot \mathfrak{f}_e(s, \xi^1, \xi^2).\end{aligned} \tag{105}$$

Applying Fubini's Theorem and integrating by parts we calculate

$$\begin{aligned}&\int_\Omega \int_0^s \boldsymbol{z}(t) \wedge \boldsymbol{d}_3[w](t)\, dt \cdot d\mathfrak{f}_e(s, \xi^1, \xi^2)\\ &= \int_0^L \left[\boldsymbol{z}(t) \wedge \boldsymbol{d}_3[w](t)\right] \cdot \int_{\Omega_t} d\mathfrak{f}_e(s, \xi^1, \xi^2)\, dt\\ &= \int_0^L \boldsymbol{z}(t) \cdot \left[\boldsymbol{d}_3[w](t) \wedge \int_{\Omega_t} d\mathfrak{f}_e(s, \xi^1, \xi^2)\right] dt\\ &= -\int_0^L \boldsymbol{z}'(t) \cdot \int_0^t \boldsymbol{d}_3[w](\tau) \wedge \int_{\Omega_\tau} d\mathfrak{f}_e(s, \xi^1, \xi^2)\, d\tau\, dt\\ &\quad + \left[\boldsymbol{z}(t) \cdot \int_0^t \boldsymbol{d}_3[w](\tau) \wedge \int_{\Omega_\tau} d\mathfrak{f}_e(s, \xi^1, \xi^2) d\tau\right]_{t=0}^{t=L}\end{aligned}$$

$$
\begin{aligned}
&= -\int_0^L \boldsymbol{z}'(t)\cdot \int_0^t \boldsymbol{d}_3[w](\tau)\wedge \int_{\Omega_\tau} d\mathfrak{f}_e(s,\xi^1,\xi^2)\,d\tau\,dt \\
&\quad +\boldsymbol{z}(L)\cdot \int_0^L \boldsymbol{d}_3[w](\tau)\wedge \int_{\Omega_\tau} d\mathfrak{f}_e(s,\xi^1,\xi^2)d\tau \\
&= -\int_0^L \boldsymbol{z}'(t)\cdot \int_0^t \boldsymbol{d}_3[w](\tau)\wedge \int_{\Omega_\tau} d\mathfrak{f}_e(s,\xi^1,\xi^2)\,d\tau\,dt \\
&\quad +\left[\boldsymbol{z}(0)+\int_0^L \boldsymbol{z}'(t)\right]\cdot \int_0^L \boldsymbol{d}_3[w](\tau)\wedge \int_{\Omega_\tau} d\mathfrak{f}_e(s,\xi^1,\xi^2)d\tau\,dt \\
&= \int_0^L \boldsymbol{z}'(t)\cdot \int_t^L \boldsymbol{d}_3[w](\tau)\wedge \int_{\Omega_\tau} d\mathfrak{f}_e(s,\xi^1,\xi^2)\,d\tau\,dt \qquad (106)\\
&\quad +\boldsymbol{z}(0)\cdot \int_0^L \boldsymbol{d}_3[w](\tau)\wedge \int_{\Omega_\tau} d\mathfrak{f}_e(s,\xi^1,\xi^2)d\tau.
\end{aligned}
$$

Similarly we obtain for $i=1,2$,

$$
\begin{aligned}
&\int_\Omega \xi^i(\boldsymbol{z}(t)\wedge \boldsymbol{d}_i[w](t))\cdot d\mathfrak{f}_e(t,\xi^1,\xi^2) \\
&\qquad = \int_\Omega \xi^i\left(\left[\boldsymbol{z}(0)+\int_0^t \boldsymbol{z}'(\tau)\,d\tau\right]\wedge \boldsymbol{d}_i[w](t)\right)\cdot d\mathfrak{f}_e(t,\xi^1,\xi^2) \\
&\qquad = \boldsymbol{z}(0)\cdot \int_\Omega \xi^1 \boldsymbol{d}_i[w](t)\wedge d\mathfrak{f}_e(t,\xi^1,\xi^2) \qquad (107)\\
&\qquad\quad + \int_0^L \boldsymbol{z}'(t)\cdot \int_{\Omega_t} \xi^i \boldsymbol{d}_i[w](s)\wedge d\mathfrak{f}_e(s,\xi^1,\xi^2)\,dt.
\end{aligned}
$$

(105)–(107) verify (102) and conclude the proof. □

Differentiability of $\boldsymbol{g}_0, \boldsymbol{g}_1, \boldsymbol{g}_2, \boldsymbol{g}_3$.

Lemma 5.7. *For some sufficiently small $\delta > 0$ (depending on the minimizer w) the functions $\boldsymbol{g}_0, g_i, i = 1,2,3,$ given in* (91)–(94) *are continuously* *differentiable on* $B_\delta(0) \subset Y$ *with*

$$\begin{aligned} \boldsymbol{g}_0'(0) \overset{\triangle}{w} &= \boldsymbol{z}(0) \wedge \int_0^L \boldsymbol{d}_3[w](t)\, dt \\ &\quad + \int_0^L \boldsymbol{z}'(t) \wedge \int_t^L \boldsymbol{d}_3[w](\tau)\, d\tau\, dt, \end{aligned} \tag{108}$$

$$g_1'(0) \overset{\triangle}{w} = \int_0^L \boldsymbol{z}'(t) \cdot \boldsymbol{d}_{03}\, dt \tag{109}$$

$$g_2'(0) \overset{\triangle}{w} = \int_0^L \boldsymbol{z}'(t) \cdot \boldsymbol{d}_{02}\, dt \tag{110}$$

$$g_3'(0) \overset{\triangle}{w} = -\int_0^L \boldsymbol{z}'(t) \cdot \boldsymbol{d}_{01}\, dt \tag{111}$$

where $\boldsymbol{z} \in W^{1,\infty}([0,L],\mathbb{R}^3)$ *is given by* (98),(99).

Proof. We use (97) to differentiate $\boldsymbol{g}_0(.)$ in (91) and obtain

$$\begin{aligned} \boldsymbol{g}_0'(0) \overset{\triangle}{w} &= \int_0^L \boldsymbol{z}(t) \wedge \boldsymbol{d}_3[w](t)\, dt \\ &= -\int_0^L \boldsymbol{z}'(t) \wedge \int_0^t \boldsymbol{d}_3[w](\tau)\, d\tau\, dt \\ &\quad + \left[\boldsymbol{z}(t) \wedge \int_0^t \boldsymbol{d}_3[w](\tau)\, d\tau\right]_{t=0}^{t=L} \\ &= -\int_0^L \boldsymbol{z}'(t) \wedge \int_0^t \boldsymbol{d}_3[w](\tau)\, d\tau\, dt + \boldsymbol{z}(L) \wedge \int_0^L \boldsymbol{d}_3[w](\tau)\, d\tau \\ &= -\int_0^L \boldsymbol{z}'(t) \wedge \int_0^t \boldsymbol{d}_3[w](\tau)\, d\tau\, dt \\ &\quad + \left[\boldsymbol{z}(0) + \int_0^L \boldsymbol{z}'(t)\, dt\right] \wedge \int_0^L \boldsymbol{d}_3[w](\tau)\, d\tau \\ &= \int_0^L \boldsymbol{z}'(t) \wedge \int_t^L \boldsymbol{d}_3[w](\tau)\, d\tau\, dt + \boldsymbol{z}(0) \wedge \int_0^L \boldsymbol{d}_3[w](t)\, dt, \end{aligned} \tag{112}$$

thus proving (108). Differentiating (92) we get

$$
\begin{aligned}
g_1'(0)\overset{\triangle}{w} &= (\boldsymbol{z}(L)\wedge \boldsymbol{d}_1[w](L))\cdot(\boldsymbol{D}_1\boldsymbol{U}(0))_2 + \breve{\boldsymbol{d}}_1[0](L)\cdot(\boldsymbol{D}_1\boldsymbol{U}'(0)\overset{\triangle}{\boldsymbol{\alpha}})_2 \\
&= \left(\left[\boldsymbol{z}(0)+\int_0^L \boldsymbol{z}'(t)\,dt\right]\wedge \boldsymbol{d}_{11}\right)\cdot \boldsymbol{d}_{12} \\
&\quad + \boldsymbol{d}_{11}\cdot(\boldsymbol{D}_1\boldsymbol{D}_0^{-1}\boldsymbol{D}_0\boldsymbol{U}'(0)\overset{\triangle}{\boldsymbol{\alpha}})_2. \qquad (113)
\end{aligned}
$$

To evaluate the last term we use (99) and notice that the matrix $\boldsymbol{D}_1\boldsymbol{D}_0^{-1}$ is orthogonal, hence

$$
\begin{aligned}
\boldsymbol{d}_{11}\cdot(\boldsymbol{D}_1\boldsymbol{D}_0^{-1}\boldsymbol{D}_0\boldsymbol{U}'(0)\overset{\triangle}{\boldsymbol{\alpha}})_2 &= \boldsymbol{d}_{11}\cdot(\boldsymbol{D}_1\boldsymbol{D}_0^{-1}(\boldsymbol{D}_0\boldsymbol{U}'(0)\overset{\triangle}{\boldsymbol{\alpha}})_2) \\
&= \boldsymbol{d}_{11}\cdot(\boldsymbol{D}_1\boldsymbol{D}_0^{-1}(\boldsymbol{z}(0)\wedge \boldsymbol{d}_{02})) \\
&= ((\boldsymbol{D}_1\boldsymbol{D}_0^{-1})^{-1}\boldsymbol{d}_{11})\cdot(\boldsymbol{z}(0)\wedge \boldsymbol{d}_{02}) \\
&= (\boldsymbol{D}_0\boldsymbol{D}_1^{-1}\boldsymbol{d}_{11})\cdot(\boldsymbol{z}(0)\wedge \boldsymbol{d}_{02}) \\
&= \boldsymbol{d}_{01}\cdot(\boldsymbol{z}(0)\wedge \boldsymbol{d}_{02}) \\
&= \boldsymbol{z}(0)\cdot(\boldsymbol{d}_{02}\wedge \boldsymbol{d}_{01}) \\
&= -\boldsymbol{z}(0)\cdot \boldsymbol{d}_{03}. \qquad (114)
\end{aligned}
$$

Inserting this into (113) leads to the desired formula (109) by $\boldsymbol{d}_{03} = \boldsymbol{d}_{13}$. Similar but simpler is the computation for $g_2'(0)$:

$$
\begin{aligned}
g_2'(0)\overset{\triangle}{w} &= (\boldsymbol{z}(L)\wedge \boldsymbol{d}_3[w](L))\cdot(\boldsymbol{D}_0\boldsymbol{U}(0))_1 + \breve{\boldsymbol{d}}_3[0](L)\cdot(\boldsymbol{D}_0\boldsymbol{U}'(0)\overset{\triangle}{\boldsymbol{\alpha}})_1 \\
&= \left(\left[\boldsymbol{z}(0)+\int_0^L \boldsymbol{z}'(t)\,dt\right]\wedge \boldsymbol{d}_{13}\right)\cdot \boldsymbol{d}_{01} + \boldsymbol{d}_{03}\cdot(\boldsymbol{D}_0\boldsymbol{U}'(0)\overset{\triangle}{\boldsymbol{\alpha}})_1 \\
&= \left(\left[\boldsymbol{z}(0)+\int_0^L \boldsymbol{z}'(t)\,dt\right]\wedge \boldsymbol{d}_{03}\right)\cdot \boldsymbol{d}_{01} + \boldsymbol{d}_{03}\cdot(\boldsymbol{z}(0)\wedge \boldsymbol{d}_{01}) \\
&= \left(\int_0^L \boldsymbol{z}'(t)\,dt\wedge \boldsymbol{d}_{03}\right)\cdot \boldsymbol{d}_{01} = \int_0^L \boldsymbol{z}'(t)\cdot \boldsymbol{d}_{02}. \qquad (115)
\end{aligned}
$$

This shows (110) and (111) is proved in the same way. □

Differentiability of g. We intend to compute the generalized gradient $\partial g(0)$ by the methods presented in Appendix B. In order to guarantee that the function g is accessible to these methods it has to be shown that g is Lipschitz continuous in a neighbourhood of $\overset{\triangle}{w}= 0$, and for that the functions

$H(.,.,.)$ and $P[.](.,.)$ have to meet certain differentiability properties. This is the first and only instance where we actually need that the global curvature $\mathcal{K}[w]$ of the minimizer is locally not attained. For curves $\boldsymbol{r}$ satisfying (30) the global curvature $\mathcal{K}[\boldsymbol{r}]$ can be characterized by a maximum over pairs of parameters in a well-defined compact subset of $[0,L]\times[0,L]$ away from the diagonal, for the proof see [24] and our remark concerning the set $A[\boldsymbol{r}]$ in Section 3.

Lemma 5.8. *Let $\boldsymbol{r}$ be a curve with $\mathcal{R}[\boldsymbol{r}]>0$, such that $\mathcal{K}[\boldsymbol{r}]$ is locally not attained and set*

$$\eta(\boldsymbol{r}) := \frac{1-\mathcal{R}[\boldsymbol{r}]\cdot\|\boldsymbol{r}''\|_{L^\infty}}{\|\boldsymbol{r}''\|_{L^\infty}}, \tag{116}$$

$$\mathcal{Q}=\mathcal{Q}[\boldsymbol{r}] := \{(s,\sigma)\in[0,L]\times[0,L] : L-\eta(\boldsymbol{r})\geq s-\sigma\geq\eta(\boldsymbol{r})\}. \tag{117}$$

Then

(i) $0<\eta(\boldsymbol{r})<L/(2\pi)$,

(ii) $A[\boldsymbol{r}]\cap\mathcal{Q}\neq\emptyset$, *i.e.,*

$$\mathcal{K}[\boldsymbol{r}] = \max_{(s,\sigma)\in\mathcal{Q}} \frac{1}{r(\boldsymbol{r}(s),\boldsymbol{r}(\sigma),\boldsymbol{r}'(\sigma))}, \tag{118}$$

(iii) $\mathcal{K}[\boldsymbol{r}] > (r(\boldsymbol{r}(s),\boldsymbol{r}(\sigma),\boldsymbol{r}'(\sigma)))^{-1}$ *for all* $(s,\sigma)\in[0,L]^2$ *such that* $(s,\sigma)\notin\mathcal{Q}$ *and* $(\sigma,s)\notin\mathcal{Q}$.

The key observation of Lemma 5.8 is that in this case the global curvature is characterized by a maximum over a fixed set. It is important to notice that this characterization is stable with respect to small variations in Y:

Lemma 5.9. *Let w be a minimizing configuration for* (45)–(50), *such that $\mathcal{K}[w]$ is locally not attained. Then there are constants $\delta>0$ and $\tilde{\eta}\in(0,L/2\pi)$ (both depending on the minimizer w) such that*

$$g(\overset{\triangle}{w}) = \max_{(s,\sigma)\in\tilde{\mathcal{Q}}} H(P[\overset{\triangle}{w}](s,\sigma))-\theta^{-2} \quad \textit{for all} \quad \overset{\triangle}{w}\in B_\delta(0)\subset Y, \tag{119}$$

where

$$\tilde{\mathcal{Q}} := \{(s,\sigma)\in[0,L]\times[0,L] : L-\tilde{\eta}\geq s-\sigma\geq\tilde{\eta}\}. \tag{120}$$

In particular, $A[\check{\boldsymbol{r}}[\overset{\triangle}{w}]]\subset\tilde{\mathcal{Q}}$ for all $\overset{\triangle}{w}\in B_\delta(0)$.

Proof. By (78) of Lemma 5.1 and by Lemma 3.2 we have that $\breve{\mathcal{K}}[.]$ and hence also $\breve{\mathcal{R}}[.]$ according to (29), are continuous on Y, i.e.,

$$\breve{\mathcal{K}}[\overset{\triangle}{w}] \to \breve{\mathcal{K}}[0] = \mathcal{K}[w] \text{ and } \breve{\mathcal{R}}[\overset{\triangle}{w}] \to \breve{\mathcal{R}}[0] = \mathcal{R}[w] \text{ as } \| \overset{\triangle}{w} \|_Y \to 0. \tag{121}$$

By virtue of (30), which holds for the minimizing configuration $\boldsymbol{r}[w]$, and by (48), (29) we obtain

$$\|\breve{\boldsymbol{r}}''[\overset{\triangle}{w}]\|_{L^\infty} < \breve{\mathcal{K}}[\overset{\triangle}{w}] \le 2\theta^{-1} \text{ for } \| \overset{\triangle}{w} \|_Y \quad \text{sufficiently small.} \tag{122}$$

Consequently, Lemma 5.8 is applicable to $\breve{\boldsymbol{r}}[\overset{\triangle}{w}]$ for $\| \overset{\triangle}{w} \|_Y$ sufficiently small and, by (118),

$$\breve{\mathcal{K}}[\overset{\triangle}{w}] = \max_{(s,\sigma)\in\mathcal{Q}[\breve{\boldsymbol{r}}[\overset{\triangle}{w}]]} \frac{1}{r(\breve{\boldsymbol{r}}[\overset{\triangle}{w}](s), \breve{\boldsymbol{r}}[\overset{\triangle}{w}](\sigma), \breve{\boldsymbol{r}}'[\overset{\triangle}{w}](\sigma))}.$$

From (116) we see that $\overset{\triangle}{w} \mapsto \eta(\breve{\boldsymbol{r}}[\overset{\triangle}{w}])$ is continuous near the origin in Y. Thus we can assume that

$$\frac{L}{2\pi} > \eta(\breve{\boldsymbol{r}}[\overset{\triangle}{w}]) \ge \frac{1}{2}\eta(\breve{\boldsymbol{r}}[0]) = \frac{1}{2}\eta(\boldsymbol{r}[w]) =: \tilde{\eta}$$

for all $\overset{\triangle}{w} \in B_\delta(0) \subset Y,\ \delta > 0$ sufficiently small. Lemma 5.8 (iii) implies that

$$\breve{\mathcal{K}}[\overset{\triangle}{w}] = \max_{(s,\sigma)\in\tilde{\mathcal{Q}}} \frac{1}{r(\breve{\boldsymbol{r}}[\overset{\triangle}{w}](s), \breve{\boldsymbol{r}}[\overset{\triangle}{w}](\sigma), \breve{\boldsymbol{r}}'[\overset{\triangle}{w}](\sigma))}.$$

with $\tilde{Q}$ defined in (120). By (31), (85), (86) and (88) we finally obtain (119). By the definition of $\tilde{\eta}$ and by Lemma 5.8 we see that $A[\breve{\boldsymbol{r}}[\overset{\triangle}{w}]] \subset \tilde{\mathcal{Q}}$ for all $\overset{\triangle}{w} \in B_\delta(0)$. □

Due to the characterization (119) of g we can apply the nonsmooth chain rule proved in Proposition B.2 of Appendix B to analyze the structure of the generalized gradients $\partial g(0)$. This leads to

Lemma 5.10. *Let w be a minimizer of* (45)–(50), *such that $\mathcal{K}[w]$ is locally not attained. Then the function g as defined in* (88) *is Lipschitz continuous on $B_\delta(0) \subset Y$ for some small $\delta > 0$ depending on w. Furthermore, for any*

$g^* \in \partial g(0)$ *there is a Radon measure* μ^* *on* $[0, L] \times [0, L]$ *with nonempty support on* $A[w]$*, see* (33)*, such that*

$$\begin{aligned} \langle g^*, \overset{\triangle}{w} \rangle_{Y^* \times Y} &= -\int_0^L \boldsymbol{z}'(t) \cdot \int_t^L \boldsymbol{d}_3(\tau) \wedge \boldsymbol{f}_c^*(\tau)\, d\tau\, dt \\ &\quad - \boldsymbol{z}(0) \cdot \int_0^L \boldsymbol{d}_3(t) \wedge \boldsymbol{f}_c^*(t)\, dt, \end{aligned} \tag{123}$$

where

$$\boldsymbol{f}_c^*(\tau) := \int_{\mathcal{Q}_\tau} \frac{\boldsymbol{r}[w](s) - \boldsymbol{r}[w](\sigma)}{|\boldsymbol{r}[w](s) - \boldsymbol{r}[w](\sigma)|}\, d\mu^*(s, \sigma), \tag{124}$$

$$\mathcal{Q}_\tau := \{(s, \sigma) \in [0, L] \times [0, L] : \sigma \le \tau \le s\} \;\; \textit{for} \;\; \tau \in [0, L]. \tag{125}$$

Proof. We consider the representation (119) of g. To verify the assumptions (a)–(c) of Proposition B.2 we observe that the set $T := \tilde{\mathcal{Q}} \subset \mathbb{R}^2$ is compact. We set $X := Y, U := B_\delta(0) \subset Y$ for some sufficiently small $\delta > 0$. Furthermore, define $p(.,.) := P[.](.)$, $G := H$, and

$$N := B_R(\boldsymbol{r}_0) \times B_R(\boldsymbol{r}_0) \times B_{\bar{\delta}}(S^2) \backslash \{(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}) \in [\mathbb{R}^3]^3 : \boldsymbol{x} = \boldsymbol{y}\} \tag{126}$$

for $\bar{\delta} > 0$ sufficiently small, where $B_R(\boldsymbol{r}_0) \subset \mathbb{R}^3$ with

$$R \ge 2 \operatorname{diam} \boldsymbol{r}[w]. \tag{127}$$

According to Lemma 5.5 hypothesis (b) of Proposition B.2 holds true. (Note that $\breve{\boldsymbol{r}}'[\overset{\triangle}{w}] = \breve{\boldsymbol{d}}_3[\overset{\triangle}{w}]$, to which Lemma 5.5 applies.) By (127) the set N is an open neighbourhood of the set $P[B_\delta(0)](\tilde{\mathcal{Q}})$, since $\breve{\boldsymbol{r}}[\overset{\triangle}{w}]$ is uniformly close to $\boldsymbol{r}[w]$ by Lemma 5.1 for small $\| \overset{\triangle}{w} \|_Y$ and $|\breve{\boldsymbol{r}}'[\overset{\triangle}{w}]| = 1$ on $[0, L]$ by (81). Furthermore,

$$\breve{\boldsymbol{r}}[\overset{\triangle}{w}](s) \neq \breve{\boldsymbol{r}}[\overset{\triangle}{w}](\sigma) \;\text{ for all }\; (s, \sigma) \in \tilde{\mathcal{Q}},$$

because the diagonal is excluded in $\tilde{\mathcal{Q}}$ and $\breve{\boldsymbol{r}}[\overset{\triangle}{w}]$ is simple for $\| \overset{\triangle}{w} \|_Y$ sufficiently small, according to $\breve{\mathcal{R}}[\overset{\triangle}{w}] > 0$, see (48), (121), and [11, Lemma 1].

The function $H = H(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t})$ as defined in (86) is continuously differentiable on N with differential

$$H'(\boldsymbol{x},\boldsymbol{y},\boldsymbol{t})\cdot(\overset{\triangle}{\boldsymbol{x}},\overset{\triangle}{\boldsymbol{y}},\overset{\triangle}{\boldsymbol{t}}) = \frac{8}{|\boldsymbol{x}-\boldsymbol{y}|^4}\Big\{(\overset{\triangle}{\boldsymbol{y}}-\overset{\triangle}{\boldsymbol{x}})\big[((\boldsymbol{x}-\boldsymbol{y})\cdot\boldsymbol{t})\boldsymbol{t}-|\boldsymbol{t}|^2(\boldsymbol{x}-\boldsymbol{y}) + \frac{2|(\boldsymbol{x}-\boldsymbol{y})\wedge\boldsymbol{t}|^2}{|\boldsymbol{x}-\boldsymbol{y}|^2}(\boldsymbol{x}-\boldsymbol{y})\big] + \overset{\triangle}{\boldsymbol{t}}\cdot\big\{|\boldsymbol{x}-\boldsymbol{y}|^2\boldsymbol{t}-((\boldsymbol{x}-\boldsymbol{y})\cdot\boldsymbol{t})(\boldsymbol{x}-\boldsymbol{y})\big\}\Big\} \quad \text{for} \quad \overset{\triangle}{\boldsymbol{x}},\overset{\triangle}{\boldsymbol{y}},\overset{\triangle}{\boldsymbol{t}}\in\mathbb{R}^3. \tag{128}$$

H' is bounded on N and thus satisfies (180). Hence we have verified assumptions (a)–(c) and can apply Proposition B.2, i.e., g is Lipschitz continuous near $\overset{\triangle}{w}=0$ and for any $g^*\in\partial g(0)$ there is a probability Radon measure $\bar{\mu}$ on $\tilde{\mathcal{Q}}$ supported on $A[w]\subset\tilde{\mathcal{Q}}$, such that

$$\langle g^*,\overset{\triangle}{w}\rangle_{Y^*\times Y} = \int_{\tilde{\mathcal{Q}}} H'(P[0](s,\sigma))\cdot P_w[0](s,\sigma)\,\overset{\triangle}{w}\,d\bar{\mu}(s,\sigma) \tag{129}$$

for all $\overset{\triangle}{w}\in Y$. Since we have to consider the integrand only on the support of $\bar{\mu}$, we need to evaluate (128) merely for $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{t})=(\boldsymbol{r}[w](s),\boldsymbol{r}[w](\sigma),\boldsymbol{r}'[w](\sigma))$ with $(s,\sigma)\in A[w]$.

The global curvature $\mathcal{K}[w]$ of the minimizer w is locally not attained, so we can use (34) and (35) to obtain for $(s,\sigma)\in A[w]$

$$H'(P[0](s,\sigma))\cdot(\overset{\triangle}{\boldsymbol{x}},\overset{\triangle}{\boldsymbol{y}},\overset{\triangle}{\boldsymbol{t}}) = \frac{8}{(2\mathcal{R}[w])^3}\Big\{(\overset{\triangle}{\boldsymbol{y}}-\overset{\triangle}{\boldsymbol{x}})\cdot\frac{\boldsymbol{r}[w](s)-\boldsymbol{r}[w](\sigma)}{|\boldsymbol{r}[w](s)-\boldsymbol{r}[w](\sigma)|}+\overset{\triangle}{\boldsymbol{t}}\cdot\boldsymbol{r}'[w](\sigma)\mathcal{R}[w]\Big\}. \tag{130}$$

In (129) we have $(\overset{\triangle}{\boldsymbol{x}},\overset{\triangle}{\boldsymbol{y}},\overset{\triangle}{\boldsymbol{t}})=P_w[0](s,\sigma)\,\overset{\triangle}{w}$ for $(s,\sigma)\in A[w]$ and $\overset{\triangle}{w}=(\overset{\triangle}{u},\overset{\triangle}{\boldsymbol{r}}_0,\overset{\triangle}{\boldsymbol{\alpha}})\in Y$, which, by (85), (96) and (97), can be computed as

$$P_w[0](s,\sigma)\,\overset{\triangle}{w}= \Big(\overset{\triangle}{\boldsymbol{r}}_0+\int_0^s \boldsymbol{z}(t)\wedge\boldsymbol{d}_3[w](t)\,dt\ ,\ \overset{\triangle}{\boldsymbol{r}}_0+\int_0^\sigma \boldsymbol{z}(t)\wedge\boldsymbol{d}_3[w](t)\,dt, \boldsymbol{z}(\sigma)\wedge\boldsymbol{d}_3[w](\sigma)\Big). \tag{131}$$

This leads to

$$\langle g^*,\overset{\triangle}{w}\rangle_{Y^*\times Y} = -\frac{1}{\mathcal{R}[w]^3}\int_{\tilde{\mathcal{Q}}}\frac{\boldsymbol{r}[w](s)-\boldsymbol{r}[w](\sigma)}{|\boldsymbol{r}[w](s)-\boldsymbol{r}[w](\sigma)|}\cdot\int_\sigma^s \boldsymbol{z}(t)\wedge\boldsymbol{d}_3[w](t)\,dt d\bar{\mu}(s,\sigma) \tag{132}$$

for $g^* \in \partial g(0)$, $\overset{\triangle}{w} \in Y$. Let us extend the measure $\bar{\mu}$ from $\tilde{\mathcal{Q}}$ to the triangle

$$\bar{\mathcal{Q}} := \{(s,\sigma) \in [0,L] \times [0,L] : s \geq \sigma\} \supset \tilde{\mathcal{Q}}$$

by zero, which we denote by $\bar{\mu}$ again. Then we can replace $\tilde{\mathcal{Q}}$ with $\bar{\mathcal{Q}}$ in (132). By Fubini's Theorem and the special structure of the set $\bar{\mathcal{Q}}$ we can transform the integral on the right-hand side in (132) further, where we also use the notation given in (124), (125) and $\mu^* := \mathcal{R}[w]^{-3}\bar{\mu}$:

$$\begin{aligned}
\langle g^*, \overset{\triangle}{w} \rangle_{Y^* \times Y} &= -\frac{1}{\mathcal{R}[w]^3} \int_0^L \boldsymbol{z}(t) \wedge \boldsymbol{d}_3[w](t) \cdot \int_{\mathcal{Q}_t} \frac{\boldsymbol{r}[w](s) - \boldsymbol{r}[w](\sigma)}{|\boldsymbol{r}[w](s) - \boldsymbol{r}[w](\sigma)|}\, d\bar{\mu}(s,\sigma)\, dt \\
&= -\int_0^L \boldsymbol{z}(t) \cdot (\boldsymbol{d}_3[w](t) \wedge \boldsymbol{f}_c^*(t))\, dt \\
&= \int_0^L \boldsymbol{z}'(t) \cdot \int_0^t \boldsymbol{d}_3[w](\tau) \wedge \boldsymbol{f}_c^*(\tau)\, d\tau\, dt \\
&\qquad - \left[\boldsymbol{z}(t) \cdot \int_0^t \boldsymbol{d}_3[w](\tau) \wedge \boldsymbol{f}_c^*(\tau)\, d\tau \right]_{t=0}^{t=L} \\
&= \int_0^L \boldsymbol{z}'(t) \cdot \int_0^t \boldsymbol{d}_3[w](\tau) \wedge \boldsymbol{f}_c^*(\tau)\, d\tau\, dt \\
&\qquad - \boldsymbol{z}(L) \cdot \int_0^L \boldsymbol{d}_3[w](\tau) \wedge \boldsymbol{f}_c^*(\tau)\, d\tau \\
&= -\int_0^L \boldsymbol{z}'(t) \cdot \int_t^L \boldsymbol{d}_3[w](\tau) \wedge \boldsymbol{f}_c^*(\tau)\, d\tau\, dt \\
&\qquad - \boldsymbol{z}(0) \cdot \int_0^L \boldsymbol{d}_3[w](t) \wedge \boldsymbol{f}_c^*(t)\, dt. \qquad (133)
\end{aligned}$$

This verifies (123). □

Lagrange multiplier rule. By Lemma 5.4 we know that $\overset{\triangle}{w} = 0$ is a local minimizer for the reduced variational problem (89)–(94). We are in the position to apply the Lagrange multiplier rule, Proposition B.1 (iii), to this variational problem, since the energy functions $\breve{E}_{\mathrm{s}}$, $\breve{E}_{\mathrm{p}}$ and the constraints $g, \boldsymbol{g}_0, g_i, i = 1,2,3$, are Lipschitz continuous near $0 \in Y$ according to Lemmas 5.6, 5.7 and 5.10. Hence there exist multipliers $\lambda_E, \lambda \geq 0$, $\boldsymbol{\lambda}_0 \in \mathbb{R}^3$,

$\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$, not all zero, such that by (178)

$$0 \in \lambda_E(\breve{E}_\mathrm{s}'(0) + \breve{E}_\mathrm{p}'(0)) + \lambda \partial g(0) + \boldsymbol{\lambda}_0 \cdot \boldsymbol{g}_0'(0) + \sum_{i=1}^{3} \lambda_i g_i'(0) \tag{134}$$

with

$$\lambda g(0) = 0. \tag{135}$$

In other words, there exists $g^* \in \partial g(0) \subset Y^*$, such that
(136)

$$0 = \left\{ \lambda_E(\breve{E}_\mathrm{s}'(0) + \breve{E}_\mathrm{p}'(0)) + \boldsymbol{\lambda}_0 \cdot \boldsymbol{g}_0'(0) + \sum_{i=1}^{3} \lambda_i g_i'(0) \right\} \overset{\triangle}{w} + \lambda \langle g^*, \overset{\triangle}{w} \rangle_{Y^* \times Y}$$

for all $\overset{\triangle}{w} \in Y$.

Choosing $\overset{\triangle}{w} = (\overset{\triangle}{u}, 0, 0) \in Y$, we have $\boldsymbol{z}(0) = 0$ by (100) and $\overset{\triangle}{\boldsymbol{r}}_0 = 0$.

Inserting the expressions (101), (102),(108)–(111) and (123) into (136) and using (18) we thus arrive at

$$\begin{aligned} 0 \;=\;& \lambda_E \Bigg[\int_0^L \boldsymbol{z}'(t) \cdot \hat{\boldsymbol{m}}(u(t), t)\, dt \\ & - \int_0^L \boldsymbol{z}'(t) \cdot \int_t^L \boldsymbol{d}_3[w](\tau) \wedge \int_{\Omega_\tau} d\mathfrak{f}_e(s, \xi^1, \xi^2)\, d\tau\, dt \\ & - \int_0^L \boldsymbol{z}'(t) \cdot \int_{\Omega_t} [\xi^1 \boldsymbol{d}_1[w](\tau) + \xi^2 \boldsymbol{d}_2[w](\tau)] \wedge d\mathfrak{f}_e(\tau, \xi^1, \xi^2)\, dt \Bigg] \\ & + \int_0^L \boldsymbol{z}'(t) \cdot \left(\int_t^L \boldsymbol{d}_3[w](\tau) \wedge \boldsymbol{\lambda}_0\, d\tau \right) dt \\ & + \lambda_1 \int_0^L \boldsymbol{z}'(t) \cdot \boldsymbol{d}_{03} + \lambda_2 \int_0^L \boldsymbol{z}'(t) \cdot \boldsymbol{d}_{02} - \lambda_3 \int_0^L \boldsymbol{z}'(t) \cdot \boldsymbol{d}_{01} \\ & - \lambda \int_0^L \boldsymbol{z}'(t) \cdot \int_t^L \boldsymbol{d}_3[w](\tau) \wedge \boldsymbol{f}_c^*(\tau)\, d\tau\, dt \end{aligned} \tag{137}$$

for all $\overset{\triangle}{u} \in L^\infty([0, L], \mathbb{R}^3)$. Recall that $\overset{\triangle}{u}$ uniquely determines $\boldsymbol{z}'$ by (98), and notice that $\boldsymbol{z}'$ can be any function in $L^\infty([0, L], \mathbb{R}^3)$ by a suitable choice of $\overset{\triangle}{u} \in L^\infty([0, L], \mathbb{R}^3)$. Thus the Fundamental Lemma in the calculus of variations implies the Euler-Lagrange equation (51) by means of the notation $\boldsymbol{f}_0 := \boldsymbol{\lambda}_0$,

$\boldsymbol{m}_0 := \lambda_1 \boldsymbol{d}_{03} + \lambda_2 \boldsymbol{d}_{02} - \lambda_3 \boldsymbol{d}_{01}$, and $\mu := \lambda\mu^*$. If $\mathcal{R}[w] > \theta$ in (48), i.e., if $g(0) < 0$ in (90), then by (135), $\lambda = 0$, hence $\mu = 0$. Notice that μ^* has nonempty support in $A[w]$, but λ can vanish even if $\mathcal{R}[w] = \theta$.

Now we take variations $\overset{\triangle}{w} = (0, \overset{\triangle}{\boldsymbol{r}}_0, 0) \in Y$ in (136). Thus $\boldsymbol{z}' = 0$ a.e. on $[0, L]$ and $\boldsymbol{z}(0) = 0$, and we obtain by (101), (102), (108)–(111), (123) that

$$0 = \lambda_E \int_\Omega d\mathfrak{f}_e(t, \xi^1, \xi^2),$$

which is (52). Finally we consider variations $\overset{\triangle}{w} = (0, 0, \overset{\triangle}{\boldsymbol{\alpha}}) \in Y$ in (136).Notice that for any vector $\boldsymbol{x} \in \mathbb{R}^3$, there exists $\overset{\triangle}{\boldsymbol{\alpha}} \in \mathbb{R}^3$, such that $\boldsymbol{x} = \boldsymbol{z}(0)$, where $\boldsymbol{z}(0)$ is given by (99), since $\boldsymbol{d}_k[w](0), k = 1, 2, 3$, furnish an orthonormal basis of $\mathbb{R}^3$, see the remark at the end of Appendix A. Thus

$$\begin{aligned} 0 &= \int_0^L \boldsymbol{d}_3[w](t) \wedge \left[\boldsymbol{f}_0 - \boldsymbol{f}_c(t) - \lambda_E \int_{\Omega_t} d\mathfrak{f}_e(s, \xi^1, \xi^2)\right] dt, \\ &\quad -\lambda_E \int_\Omega [\xi^1 \boldsymbol{d}_1[w](t) + \xi^2 \boldsymbol{d}_2[w](t)] \wedge d\mathfrak{f}_e(t, \xi^1, \xi^2), \end{aligned}$$

which is (53).

Before finishing the proof of Theorem 4.1 we first prove Corollary 4.2.

Proof of Corollary 4.2. Set

$$\begin{aligned} R_\tau &:= \{(s, \sigma) \in [0, L]^2 : s \geq \tau\}, \\ S_\tau &:= \{(s, \sigma) \in [0, L]^2 : \sigma > \tau\}. \end{aligned}$$

Let $\pi_1(s, \sigma) := s$ and $\pi_2(s, \sigma) := \sigma$ be projection operators on $[0, L]^2$, and for Borel sets $A \subset [0, L]$ we define the push-forwards

$$\mu^1(A) := \mu(\pi_1^{-1}(A)), \qquad \mu^2(A) := \mu(\pi_2^{-1}(A)),$$

which are Radon measures on $[0, L]$ (cf. [1, p. 32]). By [1, Theorem 2.28] there exist Radon measures μ^1_s, μ^2_σ on $[0, L]$, $s, \sigma \in [0, L]$, such that $s \mapsto \mu^1_s(A)$ is μ^1-measurable and $\sigma \mapsto \mu^2_\sigma(A)$ is μ^2-measurable for all Borel sets $A \subset [0, L]$, and such that for all $\tau \in [0, L]$

$$\begin{aligned} \int_{R_\tau} F(s, \sigma)\, d\mu(s, \sigma) &= \int_\tau^L \int_0^L F(s, \sigma)\, d\mu^1_s(\sigma) d\mu^1(s), \\ \int_{[0,L]^2 \setminus S_\tau} F(s, \sigma)\, d\mu(s, \sigma) &= \int_0^\tau \int_0^L F(s, \sigma)\, d\mu^2_\sigma(s) d\mu^2(\sigma). \end{aligned}$$

Hence

$$
\begin{aligned}
\boldsymbol{f}_c(\tau) &= \int_{R_\tau \setminus S_\tau} F(s,\sigma)\, d\mu(s,\sigma) \\
&= \int_\tau^L \int_0^L F(s,\sigma)\, d\mu_s^1(\sigma) d\mu^1(s) - \int_{[0,L]^2} F(s,\sigma)\, d\mu(s,\sigma) \\
&\quad + \int_0^\tau \int_0^L F(s,\sigma) d\mu_\sigma^2(s) d\mu^2(\sigma),
\end{aligned} \tag{138}
$$

where we used the fact that $\operatorname{supp}\mu \subset A[w] \subset \{(s,\sigma) : \sigma < s\}$.

Since $s \mapsto \int_0^L F(s,\sigma) d\mu_s^1(\sigma)$ is μ^1-measurable and $\sigma \mapsto \int_0^L F(s,\sigma) d\mu_\sigma^2(\sigma)$ is μ^2-measurable (cf. [1, Theorem 2.28]), the function $\boldsymbol{f}_c$ belongs to the space $BV([0,L],\mathbb{R}^3)$, and such functions are bounded. From (138) we readily obtain (ii), and by taking the inner product of (60) with $\boldsymbol{r}'[w](\tau)$ the equation (61) follows.

By the Lebesgue Decomposition Theorem (cf. [8, p. 42]) there are nonnegative functions $\alpha^1, \alpha^2 \in L^1([0,L])$, representing the absolutely continuous part of μ^1, μ^2, such that differentiation of (138) implies for a.e. $\tau \in [0,L]$

$$
\boldsymbol{f}_c'(\tau) = -\alpha^1(\tau) \int_0^L F(\tau,\sigma)\, d\mu_\tau^1(\sigma) + \alpha^2(\tau) \int_0^L F(s,\tau)\, d\mu_\tau^2(s).
$$

By $F(s,\tau) = -F(\tau,s)$ and with the nonnegative measure

$$
\mu_\tau := \alpha^1(\tau)\mu_\tau^1 + \alpha^2(\tau)\mu_\tau^2, \quad \tau \in [0,L],
$$

we arrive at

$$
\boldsymbol{f}_c'(\tau) = -\int_0^L F(\tau,\sigma)\, d\mu_\tau(\sigma) \quad \text{for a.e. } \tau \in [0,L]. \tag{139}
$$

Taking the inner product of (139) with $\boldsymbol{r}'[w](\tau)$ and by (35) we obtain (62).

For the proof of (v) we have to show that the mapping $\tau \mapsto \boldsymbol{f}_c(\tau)\cdot\boldsymbol{r}'[w](\tau)$ is Lipschitz continuous on $[0,L]$. For $t,\tau \in [0,L]$ we have

$$
\begin{aligned}
&|\boldsymbol{f}_c(t)\cdot \boldsymbol{d}_3[w](t) - \boldsymbol{f}_c(\tau)\cdot \boldsymbol{d}_3[w](\tau)| \le \\
&\quad |\boldsymbol{d}_3[w](t) - \boldsymbol{d}_3[w](\tau)||\boldsymbol{f}_c(t)| + |(\boldsymbol{f}_c(t) - \boldsymbol{f}_c(\tau))\cdot \boldsymbol{d}_3[w](\tau)|.
\end{aligned} \tag{140}
$$

By Lemma 5.1, $\boldsymbol{d}_3[w] \in W^{1,\infty}([0,L],\mathbb{R}^3)$, i.e., it is Lipschitz continuous, and $\boldsymbol{f}_c$ is bounded according to assertion (i). Thus it remains to be shown that the second term on the right-hand side is Lipschitz continuous. For $t > \tau$ we can estimate, using $|F(s,\sigma)| = 1$,

$$
\begin{aligned}
&|(\boldsymbol{f}_c(t) - \boldsymbol{f}_c(\tau)) \cdot \boldsymbol{d}_3[w](\tau)| \\
&= \Big| \int_{\mathcal{Q}_t - \mathcal{Q}_\tau} F(s,\sigma) \cdot \boldsymbol{d}_3[w](\tau)\, d\mu(s,\sigma) - \int_{\mathcal{Q}_\tau - \mathcal{Q}_t} F(s,\sigma) \cdot \boldsymbol{d}_3[w](\tau)\, d\mu(s,\sigma) \Big| \\
&\leq \int_{\mathcal{Q}_t - \mathcal{Q}_\tau} |\boldsymbol{d}_3[w](\tau) - \boldsymbol{d}_3[w](\sigma)|\, d\mu(s,\sigma) + \int_{\mathcal{Q}_t - \mathcal{Q}_\tau} |F(s,\sigma) \cdot \boldsymbol{d}_3[w](\sigma)|\, d\mu(s,\sigma) \\
&+ \int_{\mathcal{Q}_\tau - \mathcal{Q}_t} |\boldsymbol{d}_3[w](\tau) - \boldsymbol{d}_3[w](s)|\, d\mu(s,\sigma) + \int_{\mathcal{Q}_\tau - \mathcal{Q}_t} |F(s,\sigma) \cdot \boldsymbol{d}_3[w](s)|\, d\mu(s,\sigma).
\end{aligned}
$$

The second and fourth term on the right-hand side vanish by (35) and each of the two other terms is bounded from above by $l\mu([0,L]^2)|t-\tau|$, where l is the Lipschitz constant for $\boldsymbol{d}_3[w]$. This together with (140) verifies (v). □

Transversality. We finish the proof of Theorem 4.1. We will show that $\lambda_E = 0$ in (51)–(53) leads to a contradiction as long as one of the transversality conditions (a), (b) or (c) holds true. Thus $\lambda_E > 0$ in these cases and, by normalization, $\lambda_E = 1$.

If $\lambda_E = 0$, then (51) leads to

$$
\boldsymbol{m}_0 + \int_s^L \boldsymbol{d}_3[w](t) \wedge (\boldsymbol{f}_0 - \boldsymbol{f}_c(t))\, dt = 0 \quad \text{for a.e. } s \in [0,L]. \tag{141}
$$

Differentiating (141) leads to

$$
\boldsymbol{d}_3[w](s) \wedge (\boldsymbol{f}_0 - \boldsymbol{f}_c(s)) = 0 \ \text{ for a.e. } s \in [0,L], \tag{142}
$$

and by (141)

$$
\boldsymbol{m}_0 = 0. \tag{143}
$$

We infer from (142) that

$$
\boldsymbol{f}_c(s) = b(s)\boldsymbol{d}_3[w](s) + \boldsymbol{f}_0 \ \text{ for a.e. } s \in [0,L], \tag{144}
$$

where $b \in BV([0,L])$, since $\boldsymbol{f}_c \in BV([0,L],\mathbb{R}^3)$. The only possible type of discontinuity of $\boldsymbol{f}_c$ (and hence of b) could be a jump-discontinuity. Assume that say $[\boldsymbol{f}_c](s_0) \neq 0$ for some $s_0 \in S_L$ (recall the notation in (60)). The identity (144) implies

$$\boldsymbol{f}_c(s_0+) = b(s_0+)\boldsymbol{d}_3[w](s_0), \tag{145}$$

$$\boldsymbol{f}_c(s_0-) = b(s_0-)\boldsymbol{d}_3[w](s_0). \tag{146}$$

Subtracting (146) from (145) leads to

$$[\boldsymbol{f}_c](s_0) = [b](s_0)\boldsymbol{d}_3[w](s_0),$$

contradicting (61) of Corollary 4.2, hence $\boldsymbol{f}_c$ and b must be continuous, and the identity (144) holds everywhere on $[0, L]$. Moreover, by Corollary 4.2, part (v), we know that $f = \boldsymbol{f}_c \cdot \boldsymbol{d}_3[w]$ is of class $W^{1,\infty}([0, L])$, so is b by (144). Consequently also $\boldsymbol{f}_c \in W^{1,\infty}([0, L], \mathbb{R}^3)$, and we can take derivatives in (144) to get

$$\boldsymbol{f}_c'(s) = b'(s)\boldsymbol{d}_3[w](s) + b(s)\boldsymbol{d}_3'[w](s) \text{ for a.e. } s \in S_L. \tag{147}$$

From (62) in Corollary 4.2 and by $\boldsymbol{d}_3[w](s) \cdot \boldsymbol{d}_3'[w](s) = 0$ we infer that

$$b(s) \equiv b_0 = const. \text{ on } [0, L]. \tag{148}$$

Thus

$$\boldsymbol{f}_c'(s) = b_0\boldsymbol{d}_3'[w](s) \text{ a.e. on } S_L. \tag{149}$$

Now we are in the position to investigate the different transversality conditions (a)–(c) stated in Theorem 4.1. Let us assume that we are in situation (a). Applying (60) and using the fact that there is an isolated pair $(s, \sigma) \in \operatorname{supp} \mu$ in the sense of (56), we find a constant $\beta \neq 0$, such that

$$[\boldsymbol{f}_c](s) = \beta \frac{\boldsymbol{r}[w](s) - \boldsymbol{r}[w](\sigma)}{|\boldsymbol{r}[w](s) - \boldsymbol{r}[w](\sigma)|}, \tag{150}$$

contradicting the continuity of $\boldsymbol{f}_c$, that we just proved. In other words, $\lambda_E \neq 0$ in this case.

In case (b) we notice that the assumption (57) implies that

$$\Big[(J \times [0, L]) \cup ([0, L] \times J)\Big] \cap \operatorname{supp} \mu = \emptyset.$$

For any $t_1, t_2 \in J$ with $t_1 < t_2$, we thus obtain by (54) (using the notation introduced in (59)) that

$$\begin{aligned}\boldsymbol{f}_c(t_1) - \boldsymbol{f}_c(t_2) &= \int_{\mathcal{Q}_{t_1}} F(s,\sigma)\,d\mu(s,\sigma) - \int_{\mathcal{Q}_{t_2}} F(s,\sigma)\,d\mu(s,\sigma) \\ &= \int_{\mathcal{Q}_{t_1}-\mathcal{Q}_{t_2}} F(s,\sigma)\,d\mu(s,\sigma) - \int_{\mathcal{Q}_{t_2}-\mathcal{Q}_{t_1}} F(s,\sigma)\,d\mu(s,\sigma) = 0.\end{aligned}$$

Hence $\boldsymbol{f}_c$ is constant on J, i.e.,

$$\boldsymbol{f}_c' \equiv 0 \quad \text{on} \quad J. \tag{151}$$

From (149) we get

$$0 = b_0 \boldsymbol{d}_3'[w] \quad \text{a.e.} \quad \text{on} \quad J. \tag{152}$$

The case $b_0 \neq 0$ contradicts our assumption that $\boldsymbol{d}_3[w](s)$ is not constant on J. Thus $b_0 = 0$ and, by (144), $\boldsymbol{f}_c(s) = \boldsymbol{f}_0$ on $[0, L]$. According to Lemma 5.11 below this implies that μ is the zero measure and hence $\boldsymbol{f}_0 = 0$. Thus,by $\boldsymbol{\lambda}_0 = \boldsymbol{f}_0$, all Lagrange multipliers vanish which is impossible. Consequently, $b_0 = 0$ leads to a contradiction, too, and thus $\lambda_E \neq 0$ in case (b).

In case (c) we infer from (149)

$$\boldsymbol{f}_c'(\tau) = b_0 \boldsymbol{r}''[w](\tau) \text{ for a.e. } \tau \in [0, L],$$

because $\boldsymbol{f}_c \in W^{1,\infty}([0, L], \mathbb{R}^3)$. Now use part (iii) of Corollary 4.2 to conclude a contradiction at the parameter s, where

$$\boldsymbol{r}''[w](s) \notin \overline{\mathrm{conv}}(\{\rho(\boldsymbol{r}[w](s) - \boldsymbol{r}[w](\sigma)) : \rho > 0, (s,\sigma) \in \operatorname{supp}\mu\}).$$

□

Lemma 5.11. *If $\boldsymbol{f}_c \equiv$ constant on $[0, L]$, then $\mu = 0$.*

Proof. Since $\boldsymbol{f}_c$ is constant we infer from (54) (where F as defined in (59)) for every $\tau \in [0, L]$

$$\int_{\mathcal{Q}_{\tau+\varepsilon}} F(s,\sigma)\,d\mu(s,\sigma) = \int_{\mathcal{Q}_{\tau-\varepsilon}} F(s,\sigma)\,d\mu(s,\sigma) \text{ for all } \varepsilon > 0,$$

which implies by (55) for all $\tau \in [0, L]$, $\varepsilon > 0$,

$$\int_{\mathcal{Q}^1_{\tau,\epsilon}} F(s,\sigma)\, d\mu(s,\sigma) = \int_{\mathcal{Q}^2_{\tau,\epsilon}} F(s,\sigma)\, d\mu(s,\sigma), \tag{153}$$

where we have set

$$\mathcal{Q}^1_{\tau,\epsilon} := [\tau+\epsilon, L] \times (\tau-\epsilon, \tau+\epsilon] \quad \text{and} \quad \mathcal{Q}^2_{\tau,\epsilon} := [\tau-\epsilon, \tau+\epsilon) \times [0, \tau-\epsilon].$$

Assuming that $\mu \neq 0$ we find a point $(s_0, \sigma_0) \in \operatorname{supp}\mu$, hence $(s_0,\sigma_0) \in \mathcal{Q}^2_{s_0,\epsilon}$ for all sufficiently small $\varepsilon > 0$ by definition of the set $A[\boldsymbol{r}]$ in (33) containing $\operatorname{supp}\mu$. Thus, by continuity of F, there exists a small radius $r > 0$ such that for all $\varepsilon > 0$

$$\int_{\mathcal{Q}^2_{s_0,\epsilon} \cap B_r((s_0,\sigma_0))} F(s,\sigma)\, d\mu(s,\sigma) \neq 0.$$

This together with (153) for $\tau := s_0$ leads to

$$\int_{\mathcal{Q}^1_{s_0,\epsilon}} F(s,\sigma)\, d\mu(s,\sigma) - \int_{\mathcal{Q}^2_{s_0,\epsilon} \setminus B_r((s_0,\sigma_0))} F(s,\sigma)\, d\mu(s,\sigma) \neq 0.$$

Consequently, for each $\varepsilon > 0$ we either find $(t_1^\varepsilon, t_2^\varepsilon) \in \mathcal{Q}^1_{s_0,\epsilon} \cap \operatorname{supp}\mu$, or $(\sigma_1^\varepsilon, \sigma_2^\varepsilon) \in (\mathcal{Q}^2_{s_0,\epsilon} \setminus B_r((s_0,\sigma_0))) \cap \operatorname{supp}\mu$. Since $\operatorname{supp}\mu$ is closed we can let $\varepsilon \to 0$ to obtain either

$$(t_1, t_2) \in ([s_0, L] \times \{s_0\}) \cap \operatorname{supp}\mu,$$

in which case we denote $s_1 := t_1$, or

$$(\sigma_1, \sigma_2) \in ((\{s_0\} \times [0, s_0]) \setminus B_r((s_0,\sigma_0))) \cap \operatorname{supp}\mu,$$

in which case we set $s_1 := \sigma_2$. (Notice that $s_1 \neq s_0$ in either case since $\operatorname{supp}\mu$ stays away from the diagonal according to our remark after (33), compare also Lemmas 5.8 and 5.9.) In any case we have by (34)

$$|\boldsymbol{r}(s_1) - \boldsymbol{r}(s_0)| = 2\theta. \tag{154}$$

Moreover, we can use (153) for $\tau := s_1$, i.e., we obtain for every $\varepsilon > 0$

$$\int_{\mathcal{Q}^1_{s_1,\epsilon}} F(s,\sigma)\, d\mu(s,\sigma) = \int_{\mathcal{Q}^2_{s_1,\epsilon}} F(s,\sigma)\, d\mu(s,\sigma). \tag{155}$$

We fix $\varepsilon > 0$ and distinguish between two cases.

Case I. If for all $(\tau_1, \tau_2) \in \mathcal{Q}^2_{s_1,\epsilon} \cap \operatorname{supp} \mu$

$$F(\tau_1, \tau_2) \cdot (\boldsymbol{r}(s_1) - \boldsymbol{r}(s_0)) > 0,$$

then

$$\int_{\mathcal{Q}^2_{s_1,\epsilon}} F(s, \sigma) \cdot (\boldsymbol{r}(s_1) - \boldsymbol{r}(s_0))\, d\mu(s, \sigma) > 0,$$

which implies by (155) that also

$$\int_{\mathcal{Q}^1_{s_1,\epsilon}} F(s, \sigma) \cdot (\boldsymbol{r}(s_1) - \boldsymbol{r}(s_0))\, d\mu(s, \sigma) > 0.$$

Hence we find a point $(t_3^\varepsilon, t_4^\varepsilon) \in \mathcal{Q}^1_{s_1,\epsilon} \cap \operatorname{supp} \mu$ such that $F(t_3^\varepsilon, t_4^\varepsilon) \cdot (\boldsymbol{r}(s_1) - \boldsymbol{r}(s_0)) > 0$, i.e.,

$$F(t_4^\varepsilon, t_3^\varepsilon) \cdot (\boldsymbol{r}(s_1) - \boldsymbol{r}(s_0)) < 0. \tag{156}$$

Case II. There is some point $((t_5^\varepsilon, t_6^\varepsilon) \in \mathcal{Q}^2_{s_1,\epsilon} \cap \operatorname{supp} \mu$ with

$$F(t_5^\varepsilon, t_6^\varepsilon) \cdot (\boldsymbol{r}(s_1) - \boldsymbol{r}(s_0)) \le 0. \tag{157}$$

As before, using the fact that $\operatorname{supp} \mu$ is closed we can let $\varepsilon \to 0$ to obtain either

$$(t_3, t_4) \in ([s_1, L] \times \{s_1\}) \cap \operatorname{supp} \mu,$$

in which case we denote $s_2 := t_3$, or

$$(t_5, t_6) \in ((\{s_1\} \times [0, s_1]) \cap \operatorname{supp} \mu,$$

in which case we set $s_2 := t_6$. (Notice again that $s_2 \neq s_1$ in either case since $\operatorname{supp} \mu$ stays away from the diagonal.)

In any case, (156) or (157) and the continuity of F imply

$$F(s_1, s_2) \cdot (\boldsymbol{r}(s_1) - \boldsymbol{r}(s_0) \le 0,$$

which by (154) and elementary geometric arguments leads to

$$|\boldsymbol{r}(s_2) - \boldsymbol{r}(s_0)| \ge \sqrt{(2\theta)^2 + |\boldsymbol{r}(s_1) - \boldsymbol{r}(s_0)|^2} = 2\theta\sqrt{2}. \tag{158}$$

Now we can proceed in the same manner (starting with the identity (155) with s_1 replaced by s_2) to obtain a sequence of points $\{s_i\} \subset [0, L]$ satisfying the analogue of (158), i.e.,

$$|\boldsymbol{r}(s_{i+1}) - \boldsymbol{r}(s_0)| \geq \sqrt{(2\theta)^2 + |\boldsymbol{r}(s_i) - \boldsymbol{r}(s_0)|^2} \geq 2\theta\sqrt{i+1}. \tag{159}$$

Hence we have a divergent sequence of curve points $\boldsymbol{r}(s_i)$ as $i \to \infty$ which is absurd, since $\boldsymbol{r}([0, L])$ is bounded. Consequently, our assumption of a nonempty support for μ was wrong. □

5.2 Further proofs

Proof of Corollary 4.3.

We set $\lambda_E = 1$ in (51) due to transversality. The terms involving the external force $\mathfrak{f}_e$ are of class $BV([0, L], \mathbb{R}^3)$. This implies $\boldsymbol{m} \in BV([0, L], \mathbb{R}^3)$. If, in addition, (63) holds for $\mathfrak{f}_e$, then we may use Fubini's Theorem to write

$$\begin{aligned}\int_{\otimes_s} [\xi^1 \boldsymbol{d}_1[w](t) + \xi^2 \boldsymbol{d}_2[w](t)] \wedge d\mathfrak{f}_e(t, \xi^1, \xi^2) = \\ \int_s^L \boldsymbol{d}_1[w](t) \wedge \int_D \xi^1 \boldsymbol{\phi}_e(t, \xi^1, \xi^2)\, d\bar{\mu}(\xi^1, \xi^2)\, dt \\ + \int_s^L \boldsymbol{d}_2[w](t) \wedge \int_D \xi^2 \boldsymbol{\phi}_e(t, \xi^1, \xi^2)\, d\bar{\mu}(\xi^1, \xi^2)\, dt,\end{aligned} \tag{160}$$

and similarly

$$\int_{\otimes_s} d\mathfrak{f}_e(t, \xi^1, \xi^2) = \int_s^L \boldsymbol{\phi}_e(t, \xi^1, \xi^2)\, d\bar{\mu}(\xi^1, \xi^2)\, dt. \tag{161}$$

The terms on the right-hand side of (160) and (161) are absolutely continuous, which is also the case for all the other terms present in (51), hence $\boldsymbol{m} \in W^{1,1}([0, L], \mathbb{R}^3)$ with (64). If $\boldsymbol{\phi}_e$ is bounded, we get $\boldsymbol{m} \in W^{1,\infty}([0, L], \mathbb{R}^3)$, since also $\boldsymbol{f}_c$ is uniformly bounded on S_L by Lemma 4.2, part (i). □

Proof of Corollary 4.4. Note that, since $\boldsymbol{m} \in BV([0, L], \mathbb{R}^3)$, we get that $W_{u^i}(u(.), .) = \boldsymbol{m} \cdot \boldsymbol{d}_i[w] \in BV([0, L])$ for $i = 1, 2, 3$. By the Implicit Function Theorem we obtain (first locally and, by uniqueness, then globally)

$$u(.) = \hat{u}(W_{u^1}(u(.), .), W_{u^2}(u(.), .), W_{u^3}(u(.), .), .),$$

where $\hat{u}$ is a continuously differentiable vector function in its four entries, hence $u \in BV([0,L],\mathbb{R}^3)$. Now (4) implies that $\boldsymbol{D}'[w] \in BV([0,L],\mathbb{R}^{3\times 3})$. In particular $\boldsymbol{D}'[w] \in L^\infty([0,L],\mathbb{R}^{3\times 3})$, and thus $\boldsymbol{D}[w] \in W^{1,\infty}([0,L],\mathbb{R}^{3\times 3})$. Finally, $\boldsymbol{r}''[w] = \boldsymbol{d}_3'[w] \in BV([0,L],\mathbb{R}^3) \cap L^\infty([0,L],\mathbb{R}^3)$ by (3). $\square$

Proof of Corollary 4.5. According to Corollary 4.3 we know that $\boldsymbol{m} \in W^{1,\infty}([0,L],\mathbb{R}^3)$. This implies in a first step that $W_{u^i}(u(.),.) = \boldsymbol{m}\cdot\boldsymbol{d}_i[w] \in W^{1,p}([0,L])$ for $i = 1,2,3$. By the Implicit Function Theorem we find $u \in W^{1,p}([0,L],\mathbb{R}^3)$, hence $\boldsymbol{D} \in W^{2,p}([0,L],\mathbb{R}^{3\times 3})$ by (4). This in turn gives $W_{u^i}(u(.),.) = \boldsymbol{m}\cdot\boldsymbol{d}_i[w] \in W^{1,\infty}([0,L])$ for $i = 1,2,3$ leading to $u \in W^{1,\infty}([0,L],\mathbb{R}^3)$. Again by (4), $\boldsymbol{D}[w] \in W^{2,\infty}([0,L],\mathbb{R}^{3\times 3})$, hence $\boldsymbol{r}[w] \in W^{3,\infty}([0,L],\mathbb{R}^3)$. $\square$

Proof of Corollary 4.6. By Corollary 4.3 $\boldsymbol{m} \in W^{1,\infty}([0,L],\mathbb{R}^3)$, which implies that

$$\tilde{\boldsymbol{m}}(s) := W_u(u(s),s) = C(s)(u(s) - u^{\mathrm{o}}(s)) \tag{162}$$

is of class $W^{1,p}([0,L],\mathbb{R}^3)$. Since C is uniformly positive definite on $[0,L]$, hence invertible with inverse $C^{-1} \in L^r([0,L],\mathbb{R}^{3\times 3})$ as well, we can invert (162) to obtain

$$u(s) = C^{-1}(s)\tilde{\boldsymbol{m}}(s) + u^{\mathrm{o}}(s), \tag{163}$$

from which we deduce $u \in L^r([0,L],\mathbb{R}^3)$. Then (4) implies $\boldsymbol{D} \in W^{1,r}([0,L],\mathbb{R}^3)$, $\boldsymbol{r}[w] \in W^{2,r}([0,L],\mathbb{R}^3)$. (ii) follows from (163) in the same way. In a first step one obtains $u \in W^{1,p}([0,L],\mathbb{R}^3)$, which leads to $\boldsymbol{D} \in W^{2,p}([0,L],\mathbb{R}^{3\times 3})$ by (4). Now $W_{u^i}(u(.),.) = \boldsymbol{m}\cdot\boldsymbol{d}_i[w] \in W^{1,\infty}([0,L],\mathbb{R}^3)$, $i = 1,2,3$, and (162) again gives $u \in W^{1,\infty}([0,L],\mathbb{R}^3)$. The regularity for $\boldsymbol{D}[w]$ and $\boldsymbol{r}[w]$ follows from (4). $\square$

A Fréchet derivatives of the directors

According to Lemma 5.5, the mappings $(\overset{\triangle}{w},s) \mapsto \breve{\boldsymbol{d}}_k[\overset{\triangle}{w}](s)$, $k = 1,2,3$, are continuously differentiable on $B_\delta(0)\times[0,L]$ and thus

$$\partial_w \breve{\boldsymbol{d}}_k[.](.) \in C^0(B_\delta(0)\times[0,L],\mathcal{L}(Y,\mathbb{R}^3)), \quad k = 1,2,3, \tag{164}$$

where $\mathcal{L}(Y,\mathbb{R}^3)$ denotes the space of continuous linear mappings from Y to $\mathbb{R}^3$. In the following we will give an explicit characterization of this derivative.

Lemma A.1. *For all* $\overset{\triangle}{w}=(\overset{\triangle}{u},\overset{\triangle}{\boldsymbol{r}}_0,\overset{\triangle}{\boldsymbol{\alpha}})\in Y$, $s\in[0,L]$, *one has*

$$\partial_w\breve{\boldsymbol{d}}_k[0](s)\,\overset{\triangle}{w}=\boldsymbol{z}(s)\wedge\boldsymbol{d}_k[w](s),\quad k=1,2,3. \tag{165}$$

Here, $\boldsymbol{z}\in W^{1,\infty}([0,L],\mathbb{R}^3)$ *is the function*

$$\boldsymbol{z}(s)=\boldsymbol{z}(0)+\int_0^s\sum_{i=1}^3\overset{\triangle}{u}{}^i(\tau)\boldsymbol{d}_i[w](\tau)\,d\tau, \tag{166}$$

where $\boldsymbol{z}(0)$ *is uniquely determined by*

$$\boldsymbol{z}(0)\wedge\boldsymbol{d}_k[w](0)=(\boldsymbol{D}_0\boldsymbol{U}'(0)\,\overset{\triangle}{\boldsymbol{\alpha}})_k. \tag{167}$$

In particular, we obtain $\boldsymbol{z}(0)=0$ *for* $\overset{\triangle}{w}=(\overset{\triangle}{u},\boldsymbol{r}_0,0)\in Y$.

Proof. The tangent space of $SO(3)$ at the identity is given by the set of skew symmetric matrices $so(3):=\{\boldsymbol{C}\in\mathbb{R}^{3\times3}\,|\,\boldsymbol{C}^T=-\boldsymbol{C}\}$, see [12, vol.II, Ch.17]. Since we know by Lemma 5.5 that $\breve{\boldsymbol{D}}[.](s):=(\breve{\boldsymbol{d}}_1[.](s)|\breve{\boldsymbol{d}}_2[.](s)|\breve{\boldsymbol{d}}_3[.](s))$ is continuously differentiable on $B_\delta(0)\subset Y$, we may look at the Fréchet derivative $\partial_w\boldsymbol{R}[.](s)$ of the function

$$\boldsymbol{R}[.](s):=\breve{\boldsymbol{D}}[.](s)(\boldsymbol{D}[w](s))^{-1}:B_\delta(0)\subset Y\longrightarrow SO(3)$$

for arbitrary fixed $s\in[0,L]$. Notice that $\boldsymbol{R}[0](s)=Id\in SO(3)$ and that $\partial_w\boldsymbol{R}[0](s)$ is a linear mapping of Y into $so(3)$. Hence

$$\partial_w\boldsymbol{R}[0](s)=\partial_w\breve{\boldsymbol{D}}[0](s)(\boldsymbol{D}[w](s))^{-1}$$

is a linear mapping of Y into $so(3)$. That means, by multiplication with $\boldsymbol{D}[w](s)$ from the right, we obtain for each $\overset{\triangle}{w}\in Y$

$$\partial_w\breve{\boldsymbol{D}}[0](s)\,\overset{\triangle}{w}=\boldsymbol{C}(s)\boldsymbol{D}[w](s),\ \text{ for }\ \boldsymbol{C}(s):=\partial_w\boldsymbol{R}[0](s)\,\overset{\triangle}{w}\in so(3). \tag{168}$$

The skew symmetric matrix $\boldsymbol{C}(s)$ is determined by three coefficients $\boldsymbol{z}(s)=(z_1(s),z_2(s),z_3(s))\in\mathbb{R}^3$, depending on $\overset{\triangle}{w}$, via

$$\boldsymbol{C}(s)=\begin{pmatrix}0&-z_3(s)&z_2(s)\\z_3(s)&0&-z_1(s)\\-z_2(s)&z_1(s)&0\end{pmatrix}\ \text{ for }\ s\in[0,L], \tag{169}$$

which gives rise to rewriting (168) in terms of the columns of $\boldsymbol{D}[w](s)$

$$\partial_w \breve{\boldsymbol{d}}_k[0](s) \overset{\triangle}{w} = \boldsymbol{z}(s) \wedge \boldsymbol{d}_k[w](s). \tag{170}$$

It remains to be examined how $\boldsymbol{z}$ depends on $\overset{\triangle}{w}$. For this reason we differentiate (170) with respect to s, which we may do by [23, Cor.2.2] to get

$$\begin{aligned}
\frac{d}{ds}\partial_w \breve{\boldsymbol{d}}_k[0](s) \overset{\triangle}{w} &= \boldsymbol{z}'(s) \wedge \boldsymbol{d}_k[w](s) + \boldsymbol{z}(s) \wedge \boldsymbol{d}_k'[w](s) \\
&= \boldsymbol{z}'(s) \wedge \boldsymbol{d}_k[w](s) + \boldsymbol{z}(s) \wedge \left(\Big[\sum_{i=1}^{3} u^i(s) \boldsymbol{d}_i[w](s) \Big] \wedge \boldsymbol{d}_k[w](s) \right) \\
&= \boldsymbol{z}'(s) \wedge \boldsymbol{d}_k[w](s) + \left(\boldsymbol{z}(s) \wedge \Big[\sum_{i=1}^{3} u^i(s) \boldsymbol{d}_i[w](s) \Big] \right) \wedge \boldsymbol{d}_k[w](s) \qquad (171) \\
&\quad + \Big[\sum_{i=1}^{3} u^i(s) \boldsymbol{d}_i[w](s) \Big] \wedge \Big(\boldsymbol{z}(s) \wedge \boldsymbol{d}_k[w](s) \Big).
\end{aligned}$$

On the other hand, [23, Corollary 2.2] tells us that

$$\frac{d}{ds}\partial_w \breve{\boldsymbol{d}}_k[0](s) \overset{\triangle}{w} = \partial_w \breve{\boldsymbol{d}}_k{}'[0](s) \overset{\triangle}{w} . \tag{172}$$

Using the notation

$$\boldsymbol{h}[\overset{\triangle}{w}](s) := \sum_{i=1}^{3} (u^i(s) + \overset{\triangle}{u}{}^i(s)) \breve{\boldsymbol{d}}_i[\overset{\triangle}{w}](s),$$

(81) and (170) imply that

$$\begin{aligned}
\partial_w \breve{\boldsymbol{d}}_k{}'[0](s) \overset{\triangle}{w} &\underset{(4)}{=} \partial_w \boldsymbol{h}[0](s) \overset{\triangle}{w} \wedge \boldsymbol{d}_k[w](s) + \boldsymbol{h}[0](s) \wedge \partial_w \breve{\boldsymbol{d}}_k[0](s) \overset{\triangle}{w} \\
&\underset{(170)}{=} \partial_w \boldsymbol{h}[0](s) \overset{\triangle}{w} \wedge \boldsymbol{d}_k[w](s) \qquad (173) \\
&\quad + \Big[\sum_{i=1}^{3} u^i(s) \boldsymbol{d}_i[w](s) \Big] \wedge \Big(\boldsymbol{z}(s) \wedge \boldsymbol{d}_k[w](s) \Big).
\end{aligned}$$

By (171)–(173) we conclude that for $k = 1, 2, 3$,

$$\boldsymbol{z}'(s) \wedge \boldsymbol{d}_k[w](s) = \Big[\partial_w h[0](s) \overset{\triangle}{w} - \boldsymbol{z}(s) \wedge \sum_{i=1}^{3} u^i(s) \boldsymbol{d}_i[w](s) \Big] \wedge \boldsymbol{d}_k[w](s). \tag{174}$$

Now using the product rule and (170) we deduce

$$\begin{aligned}\partial_w \boldsymbol{h}[0](s) \overset{\triangle}{w} &= \sum_{i=1}^{3} \overset{\triangle}{u}{}^{i}(s)\boldsymbol{d}_i[w](s) + \sum_{i=1}^{3} u^i(s)\partial_w \breve{\boldsymbol{d}}_i[0](s) \overset{\triangle}{w} \\ &\underset{(170)}{=} \sum_{i=1}^{3} \overset{\triangle}{u}{}^{i}(s)\boldsymbol{d}_i[w](s) + \boldsymbol{z}(s) \wedge \sum_{i=1}^{3} u^i(s)\boldsymbol{d}_i[w](s).\end{aligned}$$

Inserting this into (174) we arrive at the identity

$$\boldsymbol{z}'(s) = \sum_{i=1}^{3} \overset{\triangle}{u}{}^{i} \boldsymbol{d}_i[w](s),$$

since $\{\boldsymbol{d}_k[w](s)\}_{k=1}^3$ furnishes an orthonormal basis of $\mathbb{R}^3$, which immediately implies (166).

To compute the initial value $\boldsymbol{z}(0)$ in dependence on $\overset{\triangle}{w}$ we first evaluate (165) at $s = 0$

$$\partial_w \breve{\boldsymbol{d}}_k[0](0) \overset{\triangle}{w} = \boldsymbol{z}(0) \wedge \boldsymbol{d}_k[w](0), \quad k = 1, 2, 3. \tag{175}$$

Now we differentiate the identity $\breve{\boldsymbol{d}}_k[w](0) = (\boldsymbol{D}_0 \boldsymbol{U}(\overset{\triangle}{\boldsymbol{\alpha}}))_k$ (cf. (81)), and obtain

$$\partial_w \breve{\boldsymbol{d}}_k[0](0) \overset{\triangle}{w} = (\boldsymbol{D}_0 \boldsymbol{U}'(0) \overset{\triangle}{\boldsymbol{\alpha}})_k, \quad k = 1, 2, 3. \tag{176}$$

(175), (176) imply the initial condition (167) which uniquely determines $\boldsymbol{z}(0)$ by $\overset{\triangle}{w}$, since $\boldsymbol{d}_k[w](0)$, $k = 1, 2, 3$, is an orthonormal basis of $\mathbb{R}^3$. For $\overset{\triangle}{w} = (\overset{\triangle}{u}, \overset{\triangle}{r}_0, 0) \in Y$, i.e., $\overset{\triangle}{\boldsymbol{\alpha}} = 0$ we readily get $\boldsymbol{z}(0) = 0$. □

Remark. If $\boldsymbol{z}$ is expressed by means of the matrix $\boldsymbol{C}$ defined in (169), then (167) can be written as

$$\boldsymbol{C}(0) = \boldsymbol{D}_0 \boldsymbol{U}'(0) \overset{\triangle}{\boldsymbol{\alpha}} \boldsymbol{D}_0^{-1}.$$

Since the derivative $\boldsymbol{U}'(0)$ is invertible, we obtain

$$\boldsymbol{U}'(0)^{-1} \left[\boldsymbol{D}_0^{-1} \boldsymbol{C}(0) \boldsymbol{D}_0\right] = \overset{\triangle}{\boldsymbol{\alpha}}, \tag{177}$$

which implies that for any given vector $\boldsymbol{x} \in \mathbb{R}^3$, we find $\overset{\triangle}{\boldsymbol{\alpha}} \in \mathbb{R}^3$ with (177), such that $\boldsymbol{x} = \boldsymbol{z}(0)$, where $\boldsymbol{z}(0)$ determines $\boldsymbol{C}(0)$ via (169).

Note. For $\boldsymbol{\alpha} \in \mathbb{R}^3$ we have the explicit representation

$$\boldsymbol{U}(\boldsymbol{\alpha}) = \mathrm{Id} \; + \frac{\sin|\boldsymbol{\alpha}|}{|\boldsymbol{\alpha}|}\Lambda(\boldsymbol{\alpha}) + \frac{1-\cos|\boldsymbol{\alpha}|}{|\boldsymbol{\alpha}|}\Lambda(\boldsymbol{\alpha})^2,$$

where

$$\Lambda(\boldsymbol{\alpha}) := \begin{pmatrix} 0 & -\alpha_3 & \alpha_2 \\ \alpha_3 & 0 & \alpha_1 \\ -\alpha_2 & \alpha_1 & 0 \end{pmatrix} \quad \text{for} \quad \boldsymbol{\alpha} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}.$$

Hence $\boldsymbol{U}'(0) : \mathbb{R}^3 \to so(3)$, (where $so(3)$ is the tangent space of $SO(3)$ at the identity consisting of the skew matrices in $\mathbb{R}^{3\times 3}$) is given by

$$\boldsymbol{U}'(0)\,\overset{\triangle}{\boldsymbol{\alpha}} = \Lambda(\overset{\triangle}{\boldsymbol{\alpha}}) \;\text{ for all }\; \overset{\triangle}{\boldsymbol{\alpha}} \in \mathbb{R}^3,$$

which has an obvious inverse $\boldsymbol{U}'(0)^{-1} : so(3) \to \mathbb{R}^3$.

B Clarke's generalized gradients

Here we summarize some basic properties of Clarke's generalized gradients for locally Lipschitz continuous functions, and we derive a special chain rule necessary for our analysis. For a more comprehensive presentation of this nonsmooth calculus we refer the reader to Clarke's monograph [4].

Consider a locally Lipschitz continuous function $f : X \to \mathbb{R}$, where X is a real Banach space. The *generalized directional derivative* $f^0(w; \overset{\triangle}{w})$ of f at $w \in X$ in the direction $\overset{\triangle}{w} \in X$ is defined as

$$f^0(w; \overset{\triangle}{w}) := \limsup_{\substack{v \to w \\ t \searrow 0}} \frac{f(v + t\,\overset{\triangle}{w}) - f(v)}{t}.$$

The mapping $\overset{\triangle}{w} \mapsto f^0(w; \overset{\triangle}{w})$ is positively homogeneous and subadditive, and satisfies $|f^0(w; \overset{\triangle}{w})| \le l_f \|\overset{\triangle}{w}\|_X$, where l_f denotes the local Lipschitz constant of f near $w \in X$.

The *generalized gradient* $\partial f(w)$ of f at w is the subset of X^* given by

$$\partial f(w) := \{ f^* \in X^* : \langle f^*, \overset{\triangle}{w} \rangle_{X^* \times X} \le f^0(w; \overset{\triangle}{w}) \;\text{ for all }\; \overset{\triangle}{w} \in X \}.$$

$\partial f(w)$ is a nonempty, bounded, convex and weak*-compact subset of X^*. For continuously differentiable functions f the generalized gradient $\partial f(w)$ is the singleton $\{f'(w)\}$, whereas for convex functions f the set $\partial f(w)$ is the usual subdifferential of convex analysis. For our purposes we need the following additional properties of the generalized gradients:

Proposition B.1. *Let $f, g, g_i, i = 0, \dots, n$ be Lipschitz continuous near $w \in X$. Then the following holds.*

(i) $\partial(\alpha f)(w) = \alpha \partial f(w)$ *for all* $\alpha \in \mathbb{R}$.

(ii) $\partial \sum_{i=0}^{n} g_i(w) \subset \sum_{i=0}^{n} \partial g_i(w)$.

(iii) *(Lagrange Multiplier Rule) Let w be a local minimizer of f subject to the restrictions $g(v) \leq 0$ and $g_i(v) = 0, i = 0, \dots, n$. Then there exist constants $\lambda_f, \lambda \geq 0$, and $\lambda_i \in \mathbb{R}$, not all zero, such that*

$$0 \in \lambda_f \partial f(w) + \lambda \partial g(w) + \sum_{i=0}^{n} \lambda_i \partial g_i(w), \tag{178}$$

and $\lambda g(w) = 0$.

In our analysis we have to deal with functions of the form

$$g(w) := \max_{t \in T} G(p(w,t)), \quad w \in X. \tag{179}$$

We assume that

(a) T is a metrizable sequentially compact topological space.

(b) The map $p : U \times T \to \mathbb{R}^n$, where $U \subset X$ is open and bounded, satisfies

   (1) $p(.,t)$ is continuously differentiable on U for each $t \in T$.

   (2) $p_w(.,.)$ is continuous on $U \times T$.

   (3) $p(w,.)$ is continuous on T for all $w \in U$.

(c) $G : N \to \mathbb{R}$, where $N \subset \mathbb{R}^n$ is an open neighbourhood of the set $p(U,T) \subset \mathbb{R}^n$, is continuously differentiable, and there is a constant $\Lambda \geq 0$, such that

$$|G'(x)| \leq \Lambda \text{ for all } x \in N. \tag{180}$$

Since T is compact, the function g is well defined, and

$$\mathcal{A}(w) := \{t \in T : g(w) = G(p(w,t))\} \tag{181}$$

is a nonempty closed subset of T.

Proposition B.2. *Suppose that (a)–(c) are satisfied. Then g is locally Lipschitz continuous on U, and for each $g^* \in \partial g(w), w \in U$, there is a probability Radon measure μ on T supported on $\mathcal{A}(w)$, such that*

$$\langle g^*, \overset{\triangle}{w} \rangle_{X^* \times X} = \int_T G'(p(w,t)) \cdot p_w(w,t) \overset{\triangle}{w} \; d\mu(t) \;\; \textit{for all} \;\; \overset{\triangle}{w} \in X. \tag{182}$$

Proof. Fix $w_0 \in U$. Since $p_w(.,.)$ is continuous and T is compact, we can find a neighbourhood $U_0 \subset U$ of w_0 such that $p(.,t)$ is Lipschitz continuous on U_0 for all $t \in T$ with a Lipschitz constant independent of $t \in T$ (compare [19, Lemma 6.9]). $G(.)$ is Lipschitz continuous with Lipschitz constant Λ on N by (180). Thus $G(p(.,t))$ is Lipschitz continuous on U_0 with a uniform constant with respect to $t \in T$. Furthermore, for each $t \in T$, the function $G(p(.,t))$ is continuously differentiable on U_0, and the derivative $G'(p(w,t)) \cdot p_w(w,t)$ is continuous on $U_0 \times T$. The continuous function $G(p(w,.))$ is bounded on the compact set T. Thus we can apply [4, Theorem 2.8.2, Corollary 2] to get the assertion. Note that the continuous derivative of a function agrees with its strict derivative introduced and used in [4, p.30,31]. □

Acknowledgements. Both authors were supported by the Max-Planck Institute for Mathematics in the Sciences in Leipzig and the Sonderforschungsbereich 256 at the University of Bonn. The second author enjoyed the hospitality of the Forschungsinstitut für Mathematik at the ETH Zürich.

References

[1] L. AMBROSIO, N. FUSCO, D. PALLARA, *Functions of bounded variation and free discontinuity problems.* Oxford Math. Monographs, Clarendon Press, Oxford (2000).

[2] J.C. ALEXANDER, S.S. ANTMAN, The ambiguous twist of Love. *Quarterly Appl. Math.* **40**, (1982) 83–92.

[3] S.S. ANTMAN, *Nonlinear Problems of Elasticity.* Springer, New York (1995).

[4] F.H. CLARKE, *Optimization and Nonsmooth Analysis.* John Wiley & Sons, New York (1983).

[5] B.D. COLEMAN, D. SWIGON and I. TOBIAS, Elastic stability of DNA configurations. II. Supercoiled plasmids with self-contact. *Physical Review E* **61**, #1 (2000) 759–770.

[6] J.J. DELROW, J.A. GEBE and J.M. SCHURR, Comparison of hard-cylinder and screened Coulomb interactions in the modeling of supercoiled DNAs. *Biopolymers* **42**, (1997) 455–470.

[7] M. DOI and S.F. EDWARDS, *The Theory of Polymer Dynamics.* Clarendon Press, Oxford (1986).

[8] L.C. EVANS, R.F. GARIEPY, *Measure Theory and Fine Properties of Functions.* Studies in Advanced Mathematics, CRC Press Boca Raton, New York, London, Tokyo (1992).

[9] M.H. FREEDMAN, ZHENG-XU HE, and ZHENGHAN WANG, On the Möbius energy of knots and unknots. *Annals of Math.* **139** no. 1, (1994) 1–50.

[10] O. GONZALEZ and J.H. MADDOCKS, Global curvature, thickness, and the ideal shape of knots. *The Proceedings of the National Academy of Sciences, USA* **96** no. 9, (1999) 4769–4773.

[11] O. GONZALEZ, J.H. MADDOCKS, F. SCHURICHT, H. VON DER MOSEL, Global curvature and self-contact of nonlinearly elastic curves and rods. *Calc. Var.* **14** (1), 29–68 (2002).

[12] G. GROSCHE, V.ZIEGLER, D. ZIEGLER and E. ZEIDLER (eds.), *Teubner-Taschenbuch der Mathematik II.* Teubner Stuttgart Leipzig (1995).

[13] V. KATRITCH, J. BEDNAR, D. MICHOUD, R.G. SCHAREIN, J. DUBOCHET and A. STASIAK, Geometry and physics of knots. *Nature* **384**, (1996) 142–145.

[14] D. KINDERLEHRER, G. STAMPACCHIA *An Introduction to Variational Inequalities and their Applications.* Academic Press, New York (1980)

[15] A. MARITAN, C. MICHELETTI, A. TROVATO, J.R. BANAVAR, Optimal shapes of compact strings. *Nature* **406**, (2000) 287–290.

[16] N.H. MENDELSON, Helical growth of Bacillus subtilis. A new model of cell growth. *Proc. Natl. Acad. Sci. USA* **73** no. 5, (1976) 1740–1744.

[17] J. O'HARA, Energy of knots. in A. Stasiak, V. Katritch, and L.H. Kauffman, eds, *Ser. on Knots and Everything* **19**, World Scientific, Singapore (1988) 288–314.

[18] T. SCHLICK and W.K. OLSON, Supercoiled DNA energetics and dynamics by computer simulation. *J. Mol. Biol.* **223**, (1992) 1089–1119.

[19] F. SCHURICHT, A variational approach to obstacle problems for shearable nonlinearly elastic rods. *Arch. Rational Mech. Anal.* **140**, (1997) 103–159.

[20] F. SCHURICHT, Regularity for shearable nonlinearly elastic rods in obstacle problems. *Arch. Rational Mech. Anal.* **145**, (1998) 23–49.

[21] F. SCHURICHT Variational approach to contact problems in nonlinear elasticity. MPI for Mathematics in the Sciences, Leipzig, Preprint **63**, (2000).

[22] F. SCHURICHT, Global injectivity and topological constraints for spatial nonlinearly elastic rods. MPI for Mathematics in the Sciences, Leipzig, Preprint **1**, (2001).

[23] F. SCHURICHT, H. VON DER MOSEL, Ordinary differential equations with measurable right-hand side and parameters in metric spaces. Preprint **676** (2000) SFB 256 Univ. Bonn.

[24] F. SCHURICHT, H. VON DER MOSEL, Global curvature for rectifiable loops. Preprint **742** (2000) SFB 256 Univ. Bonn, MPI for Mathematics in the Sciences, Leipzig, Preprint **80** (2001).

[25] J. SCHVARTZMAN, J. SOGO and A. STASIAK, Noeuds et ADN: Les noeuds de l'ADN éclairent la réplication. *Pour la Science* **259**, (1999) 25–26.

[26] M. SPIVAK, *Differential Geometry* vol. I. Lecture Notes, Brandeis Univ. (1970).

[27] A. STASIAK, J.H. MADDOCKS, Best packing in proteins and DNA. *Nature* **406**, (2000) 251–253.

[28] M.J. TILBY, Helical cell shape and wall synthesis in a bacterium. *Nature (Lond.)* **266**, (1977) 450–452.

[29] H. VON DER MOSEL, Minimizing the elastic energy of knots. *Asymptotic Anal.* **18**, (1998) 49–65.

[30] H. VON DER MOSEL, *Existence and Regularity for Nonlinear Elastic Self-Contact Problems.* Habilitation thesis, Universität Bonn (2001).

[31] T.P. WESTCOTT, I. TOBIAS and W.K. OLSON, Modelling self-contact forces in the elastic theory of DNA supercoiling. *J. Chem. Phys.* **107**, (1997) 3967–3980.

[32] E. ZEIDLER, *Nonlinear functional analysis and its applications*, vol.IIB: *Nonlinear Monotone Operators.* Springer, New York Berlin Heidelberg (1990).