ELASTICA AS A DYNAMICAL SYSTEM

LARRY BATES, ROBIN CHHABRA AND JĘDRZEJ ŚNIATYCKI

Abstract. The elastica is a curve in $\mathbf{R}^3$ that is stationary under variations of the integral of the square of the curvature. Elastica is viewed as a dynamical system that arises from the second order calculus of variations, and its quantization is discussed.

1. Introduction

Ever since the beginning of the calculus of variations, second order problems such as the classical problem of the elastica have been considered. The peculiar situation that distinguishes most of the interesting examples in second order problems from the more familiar first order theory is that they are parametrization independent, and so the theory of such problems has a somewhat distinctive tone from that of the first order theory. A comprehensive review of this theory, as it was understood up until the 1960s, may be found in the monograph by Grässer [8].[1] By way of contrast, this paper seeks to understand how to systematically exploit the symmetry, conservation laws, and the associated first order canonical formalism, especially as they relate to the integration of the Euler-Lagrange equations of elastica.

A principal motivation for this paper was to understand that portion of the theory of second order variational problems that could reasonably be expected to be useful for elucidating the common behaviour of several geometric functionals on curves. Three examples that motivated our study are the elastica, the shape of a real Möbius band in terms of the geometry of the central geodesic [22], and the curve of least friction. Here we present elastica, and hope to report on the other two in due course.

For reasons not entirely clear to us, the geometric theory of higher order variational problems seems to have developed in a manner largely detached from the needs and concerns of concrete problems. This is a startling contrast to recent developments in geometric mechanics and their understanding of stability, bifurcation, numerical schemes, the incorporation of nonholonomic constraints, etc. The consequences of this are at least two-fold: first, it leads to a palpable sense of dread[2] when faced with trying to look up a formulation of some part of the theory that will cleanly explain how to compute something obvious, and second, a real disconnect between the theoretical insights and the actual computational methods. This disconnect is vividly illustrated in the problem of elastica.

[1]This monograph is especially noteworthy for its comprehensive bibliography.
[2]This may be reduced to mere frustration by those less ignorant than the authors.

Planar elasticae (the equilibrium shape of a linearly elastic thin wire) were considered at least as early as 1694 by James Bernoulli.[3] However, it was not until about 1742 that Daniel Bernoulli convinced Euler to solve the problem by using the isoperimetric method (the old name for the calculus of variations before Lagrange.) From a variational point of view, the elastica is idealized as a curve that minimizes the integral over its length of the square of the curvature (that is, minimize $\int \kappa^2\, ds$), and is thus naturally treated as a second order problem in the calculus of variations. Exhaustive results were then published by Euler in 1744 [6]. Since Euler's results were so comprehensive, it is not surprising that the study of elastica remained somewhat dormant until taken up again by Max Born in his thesis [2]. More recently a striking result was obtained in 1984 by Langer and Singer [13] when they demonstrated the existence of closed elastica that were torus knots. Their proof was noteworthy because they eschewed the usual variational machinery and employed clever *ad hoc* geometric arguments such as an adapted cylindrical coordinate system to aid their integration. In fact, a significant motivation for this paper was to see to what extent their techniques could be understood by a more pedestrian use of the second order calculus of variations that looked more like just 'turning the crank' on the variational machine, and thus had the comfort of familiarity of technique. Since it is not our intent to duplicate their calculations, but gain some group-theoretical insight into the integration procedure, we study a different problem where the arclength is not constrained.

Some features of the elastica problem instantly spring to mind in the modern geometrically oriented reader. The first is that the problem is manifestly invariant by the action of the Euclidean group. The second is that it would be very nice to have a theory that explained how to reduce the symmetry using the concomitant conservation laws that Emmy Noether taught us are in the problem, and then wind up with some form of reduced Euler-Lagrange equations. Assuming we can solve these reduced equations, and hence know the curvature and torsion of the elastic curve, we would expect a good theory to show us methods to determine the shape of our curve that go beyond a mere referral to the fundamental theorem of curves stating that the curvature and torsion of the curve determine it up to a Euclidean motion. Given all this, what we actually find when we look at the published work on elastica (such as [13] or [3]) is that it proceeds somewhat differently. In particular, almost none of the actual computations seem to follow any method that resembled the current theory. There are good reasons for this, and it is not due to ignorance of those geometers but a reflection that the theory at that time was presented in such a way as to simply be unhelpful, and unable to easily identify the geometric meaning of some of their calculations. This is the best explanation we have of the situation at the time and why it was still necessary a decade after [13] appeared for Foltinek (see [7]) to write a paper demonstrating the integration constants for elastica in terms of the conserved Noetherian momenta. Further work on symmetry and integration appeared in the article by Nesterenko and Scarpetta [18]. Another work that gave a detailed study of the conserved momenta was Coronado

[3]See the delightful discussion by Levien in [15] or [14].

[4]. However, Coronado used the spatial coordinate x_1 as the parameter. Due to the parameter invariance, it is a valid procedure in the open dense domain in which $x_1 \neq 0$. This leads to a nonsingular Hamiltonian system to which the standard tools may be applied. In particular, the author analyzes for which values of the Noether invariants the reduced equations of motion can be integrated in the region where the solution exists. A disadvantage of Coronado's approach is that the choice of x_1 as the parameter obscures the geometric structure of the theory.

The plan of this paper is to first discuss the Euler-Lagrange equations for the elastica in arbitrary parametrization and the arclength parametrization. Relations between the conservation laws and the natural equations of the curve (the 'reduced equations', if you will) are derived. Then the conservation laws and symmetry group are systematically employed to integrate the equations. We then compare our approach to that of Langer and Singer in order to have an understanding of the appearance of the conserved quantities in the reduced equations as well as the role of a preferred subgroup of the Euclidean group in the integration. This yields a symmetry group theoretical explanation of the axis of the cylindrical coordinate system so cleverly (but mysteriously) employed by Langer and Singer. This is followed by a discussion of parametrization invariance and the Hamiltonian formalism. Elastica is then studied as a constrained Hamiltonian system which is invariant under the group SE(3) of rigid motions of space as well as the reparametrization group $\mathrm{Diff}_+ \mathbf{R}$ and obtain the corresponding momentum maps. We show that the constraint equations are equivalent to the vanishing of the $\mathrm{Diff}_+ \mathbf{R}$ momentum map $\mathscr{J}$, and the vanishing of the energy function. Reducing the $\mathrm{Diff}_+ \mathbf{R}$ symmetry of $\mathscr{J}^{-1}(0)$, we obtain a Hamiltonian system on the cotangent bundle of the unit sphere bundle over $\mathbf{R}^3$ with a single energy constraint. We solve the equations of motion for this system for every choice of initial data.

The paper concludes with the geometric quantization of elastica, and discusses the quantum representations of the groups SE(3) and $\mathrm{Diff}_+ \mathbf{R}$ as well as the quantum implementation of constraints.

It seems that it is a requirement that all authors on higher order calculus of variations have their own theoretical and notational preferences and foibles, and we are no exception to the rule. However, in order to spare the reader the tedium of wading through all of this before getting to the example of elastica, this material is summarized in the appendix. Thus, the appendix provides the necessary theoretical background of the second order variational calculus, especially as it pertains to symmetry and the Noether theory, as well as serving to fix notational conventions.

Finally, what is not in the paper. We do not explicitly construct SE(3) reduced spaces, nor do we discuss the Poisson bracket formalism. We also do not discuss the meaning of other groups, such as dilations, which, while not a symmetry group in the strict sense, are still of interest in understanding the structure of the solutions of the elastica.

It is our pleasure to thank the referee for a careful reading of a previous version of this paper and making many helpful suggestions, resulting in a much improved presentation.

2. Classical elastica

2.1. The variational equations. The elastica functional is given by

$$A[\sigma] = \int_{t_0}^{t_1} \kappa^2 \, |\dot{x}| \; dt,$$

where

$$\sigma : [t_0, t_1] \to [t_0, t_1] \times \mathbf{R}^3 : t \mapsto (t, x(t))$$

corresponds to a curve $\sigma : t \mapsto x(t)$ in $\mathbf{R}^3$, and κ is the curvature of σ. Since the curvature of the curve depends on its second derivatives, this is naturally a second order variational problem. As the curvature in Cartesian coordinates $x = (x^1, x^2, x^3) \in \mathbf{R}^3$ is

$$\kappa^2 = \frac{|\ddot{x}|^2}{|\dot{x}|^4} - \frac{\langle \dot{x}, \ddot{x} \rangle}{|\dot{x}|^6},$$

the elastica Lagrangian is

$$L(x, \dot{x}, \ddot{x}) = \frac{|\ddot{x}|^2}{|\dot{x}|^3} - \frac{\langle \dot{x}, \ddot{x} \rangle^2}{|\dot{x}|^5}. \tag{1}$$

It is defined on $\{(t, x, \dot{x}, \ddot{x}) \in J^2 \mid \dot{x} \neq 0\}$.

Proposition 2.1. *The elastica Lagrangian* (1) *is invariant under translations and rotations in* $\mathbf{R}^3$ *and is independent of parametrization.*

Proof. The expression (1) for L is independent of x and depends only on Euclidean scalar products of $\dot{x}$ and $\ddot{x}$. Hence, L is invariant under translations and rotations. Moreover, the curvature κ of a curve is independent of its parametrization, and $|\dot{x}| \, dt = ds$ is the element of arclength. Therefore, $L \, dt = \kappa^2 \, ds$ is independent of parametrization. q.e.d.

For elastica, Ostrogradski's momenta are

$$p_{\dot{x}} = 2\frac{\ddot{x}}{|\dot{x}|^3} - 2\frac{\langle \dot{x}, \ddot{x} \rangle \, \dot{x}}{|\dot{x}|^5},$$

and

$$p_x = -\frac{2}{\langle \dot{x}, \dot{x} \rangle^{5/2}} (\langle \dot{x}, \dot{x} \rangle \, \dddot{x} - \langle \dot{x}, \dddot{x} \rangle \, \dot{x}) - \frac{\langle \ddot{x}, \ddot{x} \rangle \, \dot{x}}{\langle \dot{x}, \dot{x} \rangle^{5/2}} + 6\frac{\langle \dot{x}, \ddot{x} \rangle \, \ddot{x}}{\langle \dot{x}, \dot{x} \rangle^{5/2}} - 5\frac{\langle \dot{x}, \ddot{x} \rangle^2 \, \dot{x}}{\langle \dot{x}, \dot{x} \rangle^{7/2}}.$$

The Euler-Lagrange equations

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} + \frac{d^2}{dt^2}\frac{\partial L}{\partial \ddot{x}} = 0$$

can be written in the form

$$\frac{\partial L}{\partial x} - \frac{d}{dt} p_x = 0.$$

Since the Lagrangian is parameter-independent, the Euler-Lagrange equations are necessarily degenerate in that they do not determine the fourth derivative $x^{(4)}$ uniquely. Let

$$(x^{(4)})^{\|} = \frac{\left\langle x^{(4)}, \dot{x}\right\rangle}{|\dot{x}|^2}\dot{x} \quad \text{and} \quad (x^{(4)})^{\perp} = x^{(4)} - (x^{(4)})^{\|}$$

denote the components of $x^{(4)}$ that are parallel and perpendicular to $\dot{x}$, respectively. The Euler-Lagrange equations written in terms of this decomposition are

$$(x^{(4)})^{\perp} = 6\,|\dot{x}|^2 \langle \dot{x}, \ddot{x}\rangle \dddot{x} + 4\langle \dot{x}, \dddot{x}\rangle \ddot{x} + \frac{5}{2}|\ddot{x}|^2 \ddot{x} - 10\langle \dot{x}, \ddot{x}\rangle \langle \dot{x}, \dddot{x}\rangle - \frac{5}{2}\frac{|\ddot{x}|^2 \langle \dot{x}, \ddot{x}\rangle}{|\dot{x}|^2}\dot{x} + \frac{35}{4}\frac{\langle \dot{x}, \ddot{x}\rangle^3}{|\dot{x}|^6}\dot{x} \tag{2}$$

They determine $(x^{(4)})^{\perp}$, but leave the component $(x^{(4)})^{\|}$ undetermined.

On the other hand, the parametrization invariance of the problem allows us to select parametrization by arclength. In the following, assume that s is the arclength parameter of the curve, and $' = d/ds$ is differentiation with respect to arclength. Therefore

$$\left|x'\right|^2 = \langle x', x'\rangle = 1, \tag{3}$$

and, by differentiation

$$\langle x', x''\rangle = 0, \tag{4}$$

$$\langle x', x'''\rangle + \langle x'', x''\rangle = 0, \tag{5}$$

$$\left\langle x', x^{(4)}\right\rangle + 3\langle x'', x'''\rangle = 0, \tag{6}$$

as well (here $x^{(4)}$ is the fourth derivative with respect to the arclength parameterization.) Substitution into (2) and (6) yields

$$(x^{(4)})^{\perp} = -\frac{3}{2}\left|x''\right|^2 x'' \quad \text{and} \quad (x^{(4)})^{\|} = -3\langle x'', x'''\rangle x'. \tag{7}$$

These equations determine the elastica completely. In other words, the choice of a parametrization determines an equation for the component $(x^{(4)})^{\|}$.

Remark 2.2. This yields an equation of the form

$$x^{(4)} = f(x, x', x'', x''')$$

to which theorems in differential equations apply that guarantee the local existence and uniqueness of solutions.

2.2. **The Frenet frame.** The elastica equations (7) are conveniently studied in the moving frame (T, N, B), where $T = x'$ is the unit tangent vector, N the normal vector and B the binormal vector of the curve $s \mapsto x(s)$. The Frenet equations are

$$T' = \kappa N, \tag{8}$$

$$N' = -\kappa T + \tau B, \tag{9}$$

$$B' = -\tau N, \tag{10}$$

with $\kappa = |x''|$ the curvature and τ the torsion of the curve. In order to relate the torsion τ to the derivative variables, observe that

$$x''' = \kappa' N + \kappa N' = \kappa' N - \kappa^2 T + \kappa\tau B, \tag{11}$$

which implies that, if $\kappa \neq 0$,

$$\tau = \kappa^{-1} \langle B, x''' \rangle = \kappa^{-1} \langle T \times N, x''' \rangle = \kappa^{-2} \langle x' \times x'', x''' \rangle . \tag{12}$$

Differentiating (11) and the Frenet equations imply

$$x^{(4)} = -3\kappa\kappa' T + (\kappa'' - \kappa^3 - \kappa\tau^2)N + (2\kappa'\tau + \kappa\tau')B. \tag{13}$$

This, together with the perpendicular equation (7) implies that

$$2\kappa'\tau + \kappa\tau' = 0, \tag{14}$$

$$2\kappa'' + \kappa^3 - 2\kappa\tau^2 = 0. \tag{15}$$

The parallel equation (7) does not lead to any new condition because $\kappa = |x''|$ implies that $\langle x'', x''' \rangle = \kappa\kappa'$. Equation (14) can be immediately integrated to yield

$$\kappa^2\tau = c, \qquad c \text{ a constant.} \tag{16}$$

If $\kappa \neq 0$, substituting $\tau = \frac{c}{\kappa^2}$ into equation (15) and integrating gives

$$(\kappa')^2 + \frac{1}{4}\kappa^4 + \frac{c^2}{\kappa^2} = \text{ constant.} \tag{17}$$

Integration of equation (17) determines completely the functions $\kappa(s)$ and $\tau(s)$ in terms of the initial data $\kappa(s_0)$, $\dot{\kappa}(s_0)$ and $\tau(s_0)$. Thus, in order to find the solution $t \mapsto x(t)$, it suffices to integrate Frenet's equations assuming that the curvature κ and the torsion τ are known functions of s. This can be achieved using the conservation laws for elastica.

2.3. **Conserved momenta.** Since the Lagrange form for elastica is invariant under translations, it follows that the linear momentum $p = \frac{\partial L}{\partial \dot{x}} - \frac{d}{dt}\left(\frac{\partial L}{\partial \ddot{x}}\right)$ is conserved. In the arclength parametrization

$$-2(x''' - \langle x', x''' \rangle x') - \langle x'', x'' \rangle x' = p = \text{ constant.} \tag{18}$$

This implies

$$p = -2x''' - 3 \langle x'', x'' \rangle x'. \tag{19}$$

Similarly, rotational invariance of the Lagrange form implies that the angular momentum

$$\mathscr{J}X_{ij} = x^i p_{x^j} - x^j p_{x^i} + x'^i p_{x'^j} - x'^j p_{x'^i}$$

is conserved (see example A.19.) Setting $l = (l_1, l_2, l_3)$, where

$$l_i = \epsilon_{ijk} \mathscr{J}X_{jk},$$

gives

$$l = x \times p_x + x' \times p_{x'}. \tag{20}$$

The expression (1) for the elastica Lagrangian in arbitrary parametrization yields, in the arclength parametrization

$$p_{x'} = 2x'' - x',$$

which implies that

$$l = x \times p + 2x' \times x''. \tag{21}$$

Proposition 2.3. *The conserved momenta in the moving frame (T, N, B) are*

$$p = -\kappa^2 T - 2\kappa' N - 2\kappa\tau B, \tag{22}$$
$$l = x \times p + 2\kappa B. \tag{23}$$

Proof. Equation (11) implies that

$$p = -2(\kappa' N - \kappa^2 T + \kappa\tau B) - 3\kappa^2 T = -\kappa^2 T - 2\kappa' N - 2\kappa\tau B.$$

while (8) and (10) yield

$$x' \times x'' = T \times (\kappa N) = \kappa(T \times N) = \kappa B. \tag{24}$$

q.e.d.

Proposition 2.4. *Scalar equations for the curvature and torsion (16) and (17) can be rewritten in the form*

$$\kappa^2\tau = -\frac{1}{4}\langle l, p\rangle, \tag{25}$$
$$(\kappa')^2 + \frac{1}{4}\kappa^4 + \frac{\langle l, p\rangle^2}{16\kappa^2} = \frac{1}{4}|p|^2. \tag{26}$$

Proof. Equations (22) and (23) imply

$$\langle l, p\rangle = 2\kappa\langle B, p\rangle = 2\kappa\left\langle B, -\kappa^2 T - 2\dot{\kappa} N - 2\kappa\tau B\right\rangle = -4\kappa^2\tau.$$

Equation (22) implies

$$\begin{aligned} |p|^2 &= \kappa^4 + 4(\kappa')^2 + 4\kappa^2\tau^2 = 4\left((\kappa')^2 + \frac{1}{4}\kappa^4 + \frac{\kappa^4\tau^2}{\kappa^2}\right) \\ &= 4\left((\kappa')^2 + \frac{1}{4}\kappa^4 + \frac{\langle l, p\rangle^2}{16\kappa^2}\right). \end{aligned}$$

q.e.d.

Suppose that κ and τ are known as functions of s (the differential equations imply that they may be expressed as elliptic functions.) It remains to show how the integration of the Frenet equations is aided by the conservation laws (22) and (23)

Theorem 2.5. *The velocity of the elastica in the direction of the conserved momentum p is*

$$\langle x', p\rangle = -\kappa^2. \tag{27}$$

Hence

$$\langle x, p\rangle = \langle x(s_0), p\rangle - \int_{s_0}^{s} \kappa^2 ds. \tag{28}$$

It remains to determine the component of the motion perpendicular to p. If x' is not parallel to p, then the vectors $x' \times p$ and $(x' \times p) \times p$ span the plane of directions perpendicular to p. Their lengths are

$$\left|x' \times p\right| = \left|-2\kappa' B + 2\kappa\tau N\right| = \sqrt{4(\kappa')^2 + 4\kappa^2\tau^2} = \sqrt{|p|^2 - \kappa^4},$$

and

$$\left|(x' \times p) \times p\right| = \sqrt{|p|^4 - |p|^2 \kappa^4} = |p| \sqrt{|p|^2 - \kappa^4}.$$

Let D and E denote the unit vectors in the direction of $x' \times p$ and $(x' \times p) \times p$, respectively. Equations (22) and (23) give

$$D = \frac{x' \times p}{|x' \times p|} = \left(|p|^2 - \kappa^4\right)^{-1/2} (-2\kappa' B + 2\kappa\tau N). \tag{29}$$

Similarly,

$$E = \frac{(x' \times p) \times p}{|(x' \times p) \times p|} = \frac{-(4(\kappa')^2 + 4\kappa^2\tau^2)T + 2\kappa^2\kappa' N + 2\kappa^3\tau B}{|p|\left(|p|^2 - \kappa^4\right)^{1/2}} \tag{30}$$

Proposition 2.6. *The frame $(D(s), E(s))$ satisfies the equations*

$$\begin{aligned} D' &= -\frac{\langle l, p\rangle |p|}{2\left(|p|^2 - \kappa^4\right)} E, \\ E' &= \frac{\langle l, p\rangle |p|}{2(|p|^2 - \kappa^4)} D, \end{aligned}$$

where $\kappa^2\tau = -\frac{1}{4}\langle l, p\rangle$.

Proof. Equations (22) and (23) give

$$x' \times p = -2\kappa' B - 2\kappa B' = -2\kappa' B + 2\kappa\tau N \tag{31}$$

and

$$(x' \times p) \times p = \langle p, p\rangle x' - \langle x', p\rangle p = (4(\kappa')^2 + 4\kappa^2\tau^2)T - 2\kappa^2\kappa' N - 2\kappa^3\tau B.$$

Differentiation yields

$$\begin{aligned} \frac{d}{ds}(x' \times p) &= x'' \times p = \kappa N \times (-2\kappa' N - 2\kappa\tau B - \kappa^2 T) \\ &= -2\kappa^2\tau T + \kappa^3 B \\ &= -\frac{2\kappa^2\tau}{(4(\kappa')^2 + 4\kappa^2\tau^2)}\{(x' \times p) \times p + 2\kappa^2\kappa' N + 2\kappa^3\tau B\} + \kappa^3 B \\ &= -\frac{2\kappa^2\tau}{(4(\kappa')^2 + 4\kappa^2\tau^2)}(x' \times p) \times p - \frac{4\kappa^4\kappa'\tau}{(4(\kappa')^2 + 4\kappa^2\tau^2)} N + \frac{4(\kappa')^2}{(4(\kappa')^2 + 4\kappa^2\tau^2)}\kappa^3 B \\ &= -\frac{2\kappa^2\tau}{(4(\kappa')^2 + 4\kappa^2\tau^2)}(x' \times p) \times p - \frac{2\kappa^3\kappa'}{(4(\kappa')^2 + 4\kappa^2\tau^2)}(2\kappa\tau N - 2\kappa' B) \\ &= -\frac{2\kappa^2\tau}{(4(\kappa')^2 + 4\kappa^2\tau^2)}(x' \times p) \times p - \frac{2\kappa^3\kappa'}{(4(\kappa')^2 + 4\kappa^2\tau^2)}(x' \times p), \end{aligned}$$

and

$$\begin{aligned}
\frac{d}{ds}(x' \times p) \times p &= (x'' \times p) \times p \\
&= \left(-\frac{2\kappa^2\tau}{(4(\kappa')^2 + 4\kappa^2\tau^2)}(x' \times p) \times p - \frac{2\kappa^3\kappa'}{(4(\kappa')^2 + 4\kappa^2\tau^2)}(x' \times p)\right) \times p \\
&= -\frac{2\kappa^2\tau}{(4(\kappa')^2 + 4\kappa^2\tau^2)}((x' \times p) \times p) \times p - \frac{2\kappa^3\kappa'}{(4(\kappa')^2 + 4\kappa^2\tau^2)}(x' \times p) \times p \\
&= -\frac{2\kappa^2\tau}{(4(\kappa')^2 + 4\kappa^2\tau^2)}(-x'|p|^2 + p\langle x', p\rangle) \times p - \frac{2\kappa^3\kappa'}{(4(\kappa')^2 + 4\kappa^2\tau^2)}(x' \times p) \times p \\
&= \frac{2\kappa^2\tau|p|^2|}{(4(\kappa')^2 + 4\kappa^2\tau^2)}x' \times p - \frac{2\kappa^3\kappa'}{(4(\kappa')^2 + 4\kappa^2\tau^2)}(x' \times p) \times p.
\end{aligned}$$

Now compute for the orthonormal frame

$$\left\{\left|x' \times p\right|^{-1} x' \times p, \left|(x' \times p) \times p\right|^{-1} (x' \times p) \times p\right\}.$$

Since

$$\left|x' \times p\right| = \left|-2\kappa' B + 2\kappa\tau N\right| = \sqrt{4(\kappa')^2 + 4\kappa^2\tau^2} = \sqrt{|p|^2 - \kappa^4},$$

it follows that

$$\frac{d}{ds}\left|x' \times p\right|^{-1} = \frac{d}{ds}(|p|^2 - \kappa^4)^{-1/2} = -\frac{1}{2}(|p|^2 - \kappa^4)^{-3/2}(-4\kappa^3\kappa') = 2\kappa^3\kappa'(|p|^2 - \kappa^4)^{-3/2}.$$

This further implies

$$\begin{aligned}
\frac{d}{ds}(x' \times p) &= -\frac{2\kappa^2\tau}{(4(\kappa')^2 + 4\kappa^2\tau^2)}(x' \times p) \times p - \frac{2\kappa^3\kappa'}{(4(\kappa')^2 + 4\kappa^2\tau^2)}x' \times p \quad (32)\\
&= -\frac{2\kappa^2\tau}{|p|^2 - \kappa^4}(x' \times p) \times p - \frac{2\kappa^3\kappa'}{|p|^2 - \kappa^4}(x' \times p), \\
\frac{d}{ds}(x' \times p) \times p &= \frac{2\kappa^2\tau\left|p^2\right|}{(4(\kappa')^2 + 4\kappa^2\tau^2)}x' \times p - \frac{2\kappa^3\kappa'}{(4(\kappa')^2 + 4\kappa^2\tau^2)}(x' \times p) \times p \\
&= \frac{2\kappa^2\tau\left|p^2\right|}{|p|^2 - \kappa^4}x' \times p - \frac{2\kappa^3\kappa'}{|p|^2 - \kappa^4}(x' \times p) \times p.
\end{aligned}$$

Similarly,

$$\left|(x' \times p) \times p\right|^2 = \left|-x'|p|^2 + p\langle x', p\rangle\right|^2 = |p|^4 - |p|^2\kappa^4,$$

and thus

$$\begin{aligned}
\frac{d}{ds}\left|(x' \times p) \times p\right|^{-1} &= \frac{d}{ds}(|p|^4 - |p|^2\kappa^4)^{-1/2} = |p|^{-1}\frac{d}{ds}(|p|^2 - \kappa^4)^{-1/2} \\
&= \frac{2\kappa^3\kappa'}{|p|}(|p|^2 - \kappa^4)^{-3/2}.
\end{aligned}$$

Therefore,

$$\frac{d}{ds}(\left|x' \times p\right|^{-1} x' \times p) = \frac{d}{ds}(\left|x' \times p\right|^{-1})x' \times p + \left|x' \times p\right|^{-1} \frac{d}{ds}(x' \times p) \tag{33}$$

$$= -\frac{2\kappa^2\tau}{\left(|p|^2 - \kappa^4\right)^{3/2}}(x' \times p) \times p \tag{34}$$

$$= -\frac{2\kappa^2\tau\,|p|}{\left(|p|^2 - \kappa^4\right)}\frac{1}{|(x' \times p) \times p|}(x' \times p) \times p. \tag{35}$$

Similarly,

$$\begin{aligned}
\frac{d}{ds}(\left|(x' \times p) \times p\right|^{-1} (x' \times p) \times p) &= \left(\frac{d}{ds}\left|(x' \times p) \times p\right|^{-1}\right)(x' \times p) \times p + \\
&\quad + \left|(x' \times p) \times p\right| \frac{d}{ds}((x' \times p) \times p) \\
&= (|p|^2 - \kappa^4)^{-3/2} 2\kappa^2\tau\,|p|\, x' \times p \\
&= \frac{2\kappa^2\tau\,|p|}{(|p|^2 - \kappa^4)^{3/2}}\frac{|x' \times p|}{|x' \times p|}x' \times p \\
&= \frac{2\kappa^2\tau\,|p|}{(|p|^2 - \kappa^4)}\frac{1}{|x' \times p|}x' \times p.
\end{aligned}$$

Thus,

$$\begin{aligned}
D' &= -\frac{2\kappa^2\tau\,|p|}{\left(|p|^2 - \kappa^4\right)}E, \\
E' &= \frac{2\kappa^2\tau\,|p|}{(|p|^2 - \kappa^4)}D,
\end{aligned}$$

and the proof is finished since $\kappa^2\tau = -\frac{1}{4}\langle l, p\rangle$. q.e.d.

Define the curve of complex-valued vectors $Z(s)$ by

$$Z(s) = D(s) + iE(s). \tag{36}$$

Proposition 2.6 implies

$$Z' = D' + iE' = -i\frac{\langle l, p\rangle\,|p|}{2(|p|^2 - \kappa^4)}Z.$$

Proposition 2.7. *Define*

$$\phi(s) = \langle l, p\rangle\,|p| \int_{s_0}^{s} \frac{1}{(|p|^2 - \kappa^4)}ds,$$

and set $Z_0 = Z(s_0) = Z_0 = D_0 + iE_0$, *then*

$$Z(s) = e^{-i\phi(s)}Z_0. \tag{37}$$

In particular,

$$D(s) = \cos\phi(s)D_0 + \sin\phi(s)E_0, \quad (38)$$
$$E(s) = -\sin\phi(s)D_0 + \cos\phi(s)E_0. \quad (39)$$

Proof. This follows immediately upon differentiating

$$Z' = \frac{d}{ds}Z = \frac{d}{ds}e^{-i\phi(s)}Z_0 = -i\phi' e^{-i\phi(s)}Z_0 = -i\frac{\langle l, p\rangle\,|p|}{2(|p|^2-\kappa^4)}Z,$$

and $Z(s_0) = e^{-i\phi(s_0)}Z_0 = Z_0$. q.e.d.

Note that

$$(x' \times p) \times p = -\,|p|^2\,x' + \langle x', p\rangle\,p$$

implies that

$$(x')^{\perp}_{p} := -\,|p|^{-2}\,(x' \times p) \times p$$

is the component of x' perpendicular to p.

Theorem 2.8. *The arclength evolution of $(x')^{\perp}_{p}$ is*

$$(x')^{\perp}_{p}(s) = -\frac{(|p|^2 - \kappa(s)^4)^{1/2}}{|p|}(-\sin\phi(s)D_0 + \cos\phi(s)E_0).$$

Hence, the component $s \mapsto x_0 + x^{\perp}_{p}(s)$ of the motion of the elastica in the plane perpendicular to p through $x_0 = x(s_0)$ is

$$x^{\perp}_{p}(s) = -\,|p|^{-2}\,(x_0 \times p)\times p - \frac{1}{|p|}\int_{s_0}^{s}(|p|^2 - \kappa(s)^4)^{1/2}(-\sin\phi(s)D_0 + \cos\phi(s)E_0)\,ds.$$

Proof.

$$\begin{aligned}
(x')^{\perp}_{p} &= \frac{|(x' \times p) \times p|}{|p|^2}E \\
&= \frac{|p|\left(|p|^2 - \kappa^4\right)^{1/2}}{|p|^2}E \\
&= \frac{\left(|p|^2 - \kappa^4\right)^{1/2}}{|p|}(-\sin\phi(s)D_0 + \cos\phi(s)E_0).
\end{aligned}$$

q.e.d.

Corollary 2.9. *The elastica equations in the arclength parametrization*

$$\sigma : I \to R^n : s \mapsto x(s),$$

have a unique solution

$$x(s) = x_0 + \int_{s_0}^{s}\left(-\frac{\kappa(s)^2}{|p|^2}p - \frac{\left(|p|^2 - \kappa(s)^4\right)^{1/2}}{|p|}(-\sin\phi(s)D_0 + \cos\phi(s)E_0)\right)ds$$

for initial data in

$$M^3_0 = \{(s, x, x', x'', x''') \in M^3 \mid \kappa \neq 0,\ \tau \neq 0\}.$$

It remains to consider the special cases when $\kappa \neq 0$ and $\tau = 0$, and when $\kappa = 0$. Equation (25), $\kappa^2\tau = -\frac{1}{4}\langle l, p\rangle$, shows that $\langle l, p\rangle = 0$. Hence, if either κ or τ vanishes at some point s_0, then it vanishes for all s for which the solution exists.

(1) If $\tau \equiv 0$ and $\kappa \neq 0$, the Frenet equations are $T' = \kappa N$, $N' = -\kappa T$, $B' = 0$, and the conservation of the linear momentum p and the angular momentum l are

$$\begin{aligned} p &= -2\kappa' N - \kappa^2 T, \\ l &= x \times p + 2\kappa B. \end{aligned} \tag{40}$$

Thus, there is an additional conserved quantity,

$$B = \frac{1}{\kappa} x' \times x''.$$

As the scalar product $\langle x', p\rangle = -\kappa^2 \langle x', x'\rangle = -\kappa^2$,

$$\langle x(s), p\rangle = \langle x(s_0), p\rangle - \int_{s_0}^{s} \kappa^2(s)\, ds.$$

Taking the cross product of equation (40) with B yields

$$-2\kappa' N \times B - \kappa^2 T \times B = p \times B.$$

Since $N \times B = T$ and $T \times B = -N$,

$$-2\kappa' T + \kappa^2 N = p \times B.$$

Therefore,

$$\langle x', p \times B\rangle = -2\kappa',$$

and

$$\begin{aligned} \langle x(s), p \times B\rangle &= \langle x(s_0), p \times B\rangle + \int_{s_0}^{s} \langle x'(s), p \times B\rangle\, ds \\ &= \langle x(s_0), p \times B\rangle - 2\kappa(s) + 2\kappa(s_0). \end{aligned}$$

Thus, if $p \times B \neq 0$, then

$$x(s) = x(s_0) - \frac{2\kappa(s_0)}{|p \times B|^2} p \times B - \left(\frac{1}{|p|^2} \int_{s_0}^{s} \kappa^2(s)\, ds\right) p - \frac{2\kappa(s)}{|p \times B|^2} p \times B.$$

(2) The special case $p \times B = 0$, $p \neq 0$. If $p \times B = 0$, and $p \neq 0$, then p is parallel to B, and equation (40) implies that $\kappa = 0$, so the solution is a straight line.

(3) If $p = 0$, then equation (40) implies that $\kappa = 0$. If $\kappa = 0$, then x' is constant, and the motion is again a straight line.

2.4. **Closed elastica.** As mentioned in the introduction, a significant motivatation for this work was understanding how symmetry and conservation laws could be systematically exploited to integrate the elastica equations. In the case of closed elastica, as studied by Langer and Singer [13], it is necessary to add an arclength constraint to the variational problem. This results in studying a modified problem with an undetermined Lagrange multiplier. The modifications to our analysis are straightforward insofar as the use of the conserved quantities is concerned. However, since there is also an immediate integration and reduction of order in the problem, which results in a loss of manifest Euclidean invariance, it seemed preferable to avoid the arclength constrained problem and keep the full symmetry in order to see more clearly how the conservation laws enabled the integration, which is the route taken in the previous section.

In more detail, a slicker, but less transparent approach to the Euler-Lagrange equations runs as follows. For a second order Lagrangian $L(x, \dot{x}, \ddot{x}, t)$ the Euler-Lagrange equations are

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) + \frac{d^2}{dt^2}\left(\frac{\partial L}{\partial \ddot{x}}\right) = 0,$$

it follows that if

$$\frac{\partial L}{\partial x} \equiv 0,$$

which is the case in the elastica problem,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}} - \frac{d}{dt}\left(\frac{\partial L}{\partial \ddot{x}}\right)\right) = 0,$$

immediately integrates to

$$\frac{\partial L}{\partial \dot{x}} - \frac{d}{dt}\left(\frac{\partial L}{\partial \ddot{x}}\right) = c,$$

where c is a constant.

Now put $q = \dot{x}$, $\dot{q} = \ddot{x}$, and set

$$l(q, \dot{q}, t) := L(x, \dot{x}, \ddot{x}, t) - c \cdot \dot{x}$$

Then the integrated equations are now the Euler-Lagrange equations for the *first order* Lagrangian l

$$\frac{\partial l}{\partial q} - \frac{d}{dt}\left(\frac{\partial l}{\partial \dot{q}}\right) = 0.$$

2.4.1. *Reduction for elastica.*

Linear momentum. The elastica functional for fixed arclength is

$$\int_\gamma \kappa^2 + \lambda\, ds.$$

Here λ is a constant whose value is *a priori* unknown. The curvature κ is

$$\kappa = \frac{|\dot{x} \times \ddot{x}|}{|\dot{x}|^3}$$

so since $ds = |\dot{x}|\, dt$, the reduced Lagrangian is

$$l(q, \dot{q}) = \frac{|q \times \dot{q}|^2}{|q|^5} + \lambda |q| - c \cdot q$$

It remains to compute the Euler-Lagrange equations and look at them in the Frenet frame. The derivatives are

$$\frac{\partial l}{\partial q} = \left(-5\frac{|q\times\dot q|}{|q|^7} + \frac{\lambda}{|q|}\right)q + \frac{2}{|q|^5}\dot q\times(q\times\dot q) - c,$$

$$\frac{\partial l}{\partial \dot q} = \frac{2}{|q|^5}q\times(\dot q\times q).$$

Define new variables T, N, B and v by setting $v = |q|$, $T = v^{-1}q$,

$$N = \frac{(q\times\dot q)\times q}{|q\times\dot q|\,|q|} = \frac{|q|^2\,\dot q - \langle q,\dot q\rangle\, q}{|q\times\dot q|\,|q|}, \qquad B = T\times N.$$

This implies

$$\dot q = \dot v T + v^2\kappa N, \qquad \frac{\partial l}{\partial\dot q} = \frac{2}{v}\kappa N.$$

If we recall the Frenet equations then it follows that

$$\frac{\partial l}{\partial q} = (-3\kappa^2 + \lambda)T - 2\frac{\kappa\dot v}{v^2}N - c,$$

and

$$\frac{d}{dt}\left(\frac{\partial l}{\partial\dot q}\right) = 2\left(-\kappa^2 T + \left(\frac{\kappa}{v}\right)^{\cdot} N + \kappa\tau B\right).$$

This implies that the Euler-Lagrange equations are

$$(\lambda-\kappa^2)T - 2\frac{\dot\kappa}{v}N - 2\kappa\tau B = c.$$

Taking the inner product of this equation with itself and choosing the arclength parametrization (so $v = 1$) yields

$$4(\kappa')^2 + (\lambda-\kappa^2)^2 + 4\kappa^2\tau^2 = c^2.$$

Angular momentum. The reduced Lagrangian l is not invariant under the rotation group SO(3), but it is invariant under the SO(2) subgroup generated by the vector field

$$X = (c\times q)\frac{\partial}{\partial q}$$

which is rotation about the axis defined by $c \neq 0$. The associated conserved momentum is

$$j = \langle p\,dq, X\rangle = \langle 2v^{-1}\kappa N, c\times q\rangle.$$

The only nonzero component of this contraction is in the N component of $c\times q$, and since the Frenet frame is orthonormal,

$$j = 2\kappa\langle c, B\rangle.$$

Taking the inner product of c with the Euler-Lagrange equations gives

$$\langle c, B\rangle = -2\kappa\tau,$$

and this implies that the conserved angular momentum j is

$$j = -4\kappa^2\tau.$$

Thus $4\kappa^2\tau^2 = j^2/4\kappa^2$, and substituting back into the equation for κ' yields

$$4(\kappa')^2 + (\lambda - \kappa^2)^2 + \frac{j^2}{4\kappa^2} = c^2.$$

This recovers equations (3) and (4) of FOLTINEK [7], together with the interpretation of c as linear momentum and j as angular momentum.

3. Elastica as a constrained Hamiltonian system

3.1. **Range of the Legendre transformation.** In this section we discuss the range of the Legendre transformation

$$\mathscr{L} : J^3_0 \to T^* J^1_0 : (t, x, \dot{x}, \ddot{x}, \dddot{x}) \mapsto (t, x, \dot{x}, p_t, p_x, p_{\dot{x}})$$

for elastica with Lagrangian

$$L(x, \dot{x}, \ddot{x}) = \frac{\langle \ddot{x}^\perp, \ddot{x}^\perp \rangle}{|\dot{x}|^3}. \tag{41}$$

Here, the subscript 0 denotes that $\dot{x} \neq 0$, $\ddot{x}^\perp$ is the component of $\ddot{x}$ that is perpendicular to $\dot{x}$, and

$$\begin{aligned} p_{\dot{x}} &= \frac{\partial L}{\partial \ddot{x}} = 2\frac{\ddot{x}^\perp}{|\dot{x}|^3}, & (42)\\ p_x &= \frac{\partial L}{\partial \dot{x}} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{x}} = -\frac{2}{|\dot{x}|^3}\dddot{x}^\perp + \frac{6}{|\dot{x}|^5}\langle \dot{x}, \ddot{x}\rangle \ddot{x}^\perp - \frac{\langle \ddot{x}^\perp, \ddot{x}^\perp\rangle}{|\dot{x}|^5}\dot{x}, & (43)\\ p_t &= L - \langle p_x, \dot{x}\rangle - \langle p_{\dot{x}}, \ddot{x}\rangle = 0. & (44) \end{aligned}$$

The Diff$_+$ $\mathbf{R}$-invariance of $L\,dt$ is responsible for the vanishing of p_t above, and it implies that the variable $p_{\dot{x}}$ satisfies the equation

$$\langle p_{\dot{x}}, \dot{x}\rangle = 0 \tag{45}$$

(see remark A.32). Note that equations (42) and (43) imply

$$\begin{aligned} \langle p_{\dot{x}}, p_{\dot{x}}\rangle &= 4\frac{\langle \ddot{x}^\perp, \ddot{x}^\perp\rangle}{|\dot{x}|^6} = \frac{4}{|\dot{x}|^3}L, & (46)\\ \langle p_{\dot{x}}, \ddot{x}\rangle &= \left\langle 2\frac{\ddot{x}^\perp}{|\dot{x}|^3}, \ddot{x}\right\rangle = 2\frac{\langle \ddot{x}^\perp, \ddot{x}\rangle}{|\dot{x}|^3} = 2\frac{\langle \ddot{x}^\perp, \ddot{x}^\perp\rangle}{|\dot{x}|^3} = 2L = \frac{1}{2}|\dot{x}|^3\langle p_{\dot{x}}, p_{\dot{x}}\rangle, & (47)\\ \langle p_x, \dot{x}\rangle &= \left\langle -\frac{2}{|\dot{x}|^3}\dddot{x}^\perp + \frac{6}{|\dot{x}|^5}\langle \dot{x}, \ddot{x}\rangle \ddot{x}^\perp - \frac{\langle \ddot{x}^\perp, \ddot{x}^\perp\rangle}{|\dot{x}|^5}\dot{x}, \dot{x}\right\rangle = -\frac{1}{4}|\dot{x}|^3\langle p_{\dot{x}}, p_{\dot{x}}\rangle. & (48) \end{aligned}$$

Equation (44), written in terms of the variables on $T^*J^1_0$, reads

$$H(x, \dot{x}, p_x, p_{\dot{x}}) = -\frac{1}{4}|\dot{x}|^3\langle p_{\dot{x}}, p_{\dot{x}}\rangle + \frac{1}{2}|\dot{x}|^3\langle p_{\dot{x}}, p_{\dot{x}}\rangle - \frac{1}{4}|\dot{x}|^3\langle p_{\dot{x}}, p_{\dot{x}}\rangle = 0.$$

Hence, it does not introduce further restrictions of the variables $(t, x, \dot{x}, p_t, p_x, p_{\dot{x}})$. Equations (47) and (48) lead to the new constraint equation

$$\langle p_{\dot{x}}, \ddot{x}\rangle + 2\langle p_x, \dot{x}\rangle = 0, \tag{49}$$

while equation (49) in terms of the variables on $T^*J^1_0$ is

$$|\dot{x}|^3 \langle p_{\dot{x}}, p_{\dot{x}} \rangle + 4 \langle p_x, \dot{x} \rangle = 0. \tag{50}$$

Theorem 3.1. *The range of the Legendre transformation* $\mathscr{L} : J^3_0 \to T^*J^1_0$ *is the common zero set of the three functions* p_t, $\langle p_{\dot{x}}, \dot{x} \rangle$ *and* $\langle p_{\dot{x}}, p_{\dot{x}} \rangle + \frac{4}{|\dot{x}|^3} \langle p_x, \dot{x} \rangle$. *That is,*

$$\text{range } \mathscr{L} = \{(t, x, \dot{x}, p_t, p_{x,} p_{\dot{x}}) \in T^*J^1_0 \mid p_t = \langle p_{\dot{x}}, \dot{x} \rangle = |\dot{x}|^3 \langle p_{\dot{x}}, p_{\dot{x}} \rangle + 4 \langle p_x, \dot{x} \rangle = 0\}. \tag{51}$$

Proof. Equations (44), (45) and (50) imply that

$$\text{range } \mathscr{L} \subseteq \{(t, x, \dot{x}, p_t, p_{x,} p_{\dot{x}}) \in T^*J^1_0 \mid p_t = \langle p_{\dot{x}}, \dot{x} \rangle = |\dot{x}|^3 \langle p_{\dot{x}}, p_{\dot{x}} \rangle + 4 \langle p_x, \dot{x} \rangle = 0\}.$$

Suppose $(t, x, \dot{x}, p_t, p_{x,} p_{\dot{x}}) \in T^*J^1_0$ is such that $p_t = 0$, $\langle p_{\dot{x}}, \dot{x} \rangle = 0$ and $|\dot{x}|^3 \langle p_{\dot{x}}, p_{\dot{x}} \rangle + 4 \langle p_x, \dot{x} \rangle = 0$. Since $\langle p_{\dot{x}}, \dot{x} \rangle = 0$, it follows that $p_{\dot{x}} = p^{\perp}_{\dot{x}}$, and equation (42) implies $\ddot{x}^{\perp} = \frac{1}{2} |\dot{x}|^3 p_{\dot{x}}$. By definition, the vanishing of $|\dot{x}|^3 \langle p_{\dot{x}}, p_{\dot{x}} \rangle + 4 \langle p_x, \dot{x} \rangle$ is equivalent to equation (49); that is $\langle p_{\dot{x}}, \ddot{x} \rangle + 2 \langle p_x, \dot{x} \rangle = 0$. Splitting equation (43) into its components perpendicular and parallel to $\dot{x}$ gives

$$p^{\perp}_x = -\frac{2}{|\dot{x}|^3} \dddot{x}^{\perp} + \frac{6}{|\dot{x}|^5} \langle \dot{x}, \ddot{x} \rangle \ddot{x}^{\perp}, \tag{52}$$

$$p^{\|}_x = -\frac{\langle \ddot{x}^{\perp}, \ddot{x}^{\perp} \rangle}{|\dot{x}|^5} \dot{x}. \tag{53}$$

Equation (52) gives

$$\begin{aligned} \frac{2}{|\dot{x}|^3} \dddot{x}^{\perp} &= -p^{\perp}_x + \frac{6}{|\dot{x}|^5} \langle \dot{x}, \ddot{x} \rangle \ddot{x}^{\perp} = -p^{\perp}_x + \frac{6}{|\dot{x}|^5} \langle \dot{x}, \ddot{x} \rangle \frac{1}{2} |\dot{x}|^3 p_{\dot{x}} \\ &= -p^{\perp}_x + \frac{3}{|\dot{x}|^2} \langle \dot{x}, \ddot{x} \rangle p_{\dot{x}}, \end{aligned}$$

where $\langle \dot{x}, \ddot{x} \rangle$ is arbitrary. Equation (53) is equivalent to equation (49) because

$$\langle p_x, \dot{x} \rangle = \left\langle p^{\|}_x, \dot{x} \right\rangle = -\frac{\langle \ddot{x}^{\perp}, \ddot{x}^{\perp} \rangle}{|\dot{x}|^5} \langle \dot{x}, \dot{x} \rangle = -\frac{\langle \ddot{x}^{\perp}, \ddot{x}^{\perp} \rangle}{|\dot{x}|^3} = -\frac{1}{2} \langle p_{\dot{x}}, \ddot{x} \rangle$$

by equation (47).

The above argument shows that the fibre of $\mathscr{L}$ over the point $(t, x, \dot{x}, p_t, p_{x,} p_{\dot{x}}) \in T^*J^1_0$ such that $p_t = 0$, $\langle p_{\dot{x}}, \dot{x} \rangle = 0$ and $|\dot{x}|^3 \langle p_{\dot{x}}, p_{\dot{x}} \rangle + 4 \langle p_x, \dot{x} \rangle = 0$ is not empty. In fact,

$$\mathscr{L}^{-1}(t, x, \dot{x}, p_t, p_{x,} p_{\dot{x}}) = \{(t, x, \dot{x}, \ddot{x}^{\perp} + \ddot{x}^{\|}, \dddot{x}^{\perp} + \dddot{x}^{\|}\},$$

where

$$\begin{aligned} \ddot{x}^{\perp} &= \frac{1}{2} |\dot{x}|^3 p_{\dot{x}}, \\ \dddot{x}^{\perp} &= \frac{1}{2} \left(-|\dot{x}|^3 p^{\perp}_x + 3 |\dot{x}| \left\langle \dot{x}, \ddot{x}^{\|} \right\rangle p_{\dot{x}} \right), \end{aligned}$$

and $\ddot{x}^{\|}$ and $\dddot{x}^{\|}$ are arbitrary. q.e.d.

Theorem 3.2. *The range of the Legendre transformation is a submanifold of* $T^*J^1_0$.

Proof. Fix $(t, x, \dot{x}) \in J_0^1$. The constraint equations

$$\begin{aligned} \langle p_{\dot{x}}, \dot{x}\rangle &= 0, \\ \langle \dot{x}, \dot{x}\rangle^{3/2} \langle p_{\dot{x}}, p_{\dot{x}}\rangle + 4 \langle p_x, \dot{x}\rangle &= 0, \end{aligned}$$

give

$$\begin{aligned} p_{\dot{x}}^{\|} &= 0, \\ p_x^{\|} &= -\frac{1}{4} \langle \dot{x}, \dot{x}\rangle^{3/2} \langle p_{\dot{x}}, p_{\dot{x}}\rangle = -\frac{1}{4} \langle \dot{x}, \dot{x}\rangle^{3/2} \left\langle p_{\dot{x}}^{\perp}, p_{\dot{x}}^{\perp}\right\rangle . \end{aligned}$$

By assumption, $\dot{x} \neq 0$, which implies that the splitting of vectors into components parallel and perpendicular to $\dot{x}$ is smooth. Therefore,

$$\{(t, x, \dot{x}, p_t, p_{x,} p_{\dot{x}}) \in T^*J_0^1 \mid p_t = 0, \langle p_{\dot{x}}, \dot{x}\rangle = 0 \text{ and } |\dot{x}|^3 \langle p_{\dot{x}}, p_{\dot{x}}\rangle + 4 \langle p_x, \dot{x}\rangle = 0\}$$

is equal to

$$\{(t, x, \dot{x}, p_t, p_{x,} p_{\dot{x}}) \in T^*J_0^1 \mid p_t = 0,\ p_{\dot{x}}^{\|} = 0 \text{ and } p_x^{\|} = -\frac{1}{4} |\dot{x}|^3 \langle p_{\dot{x}}^{\perp}, p_{\dot{x}}^{\perp}\rangle\}$$

and is a submanifold of $T^*J_0^1$. Hence, range $\mathscr{L}$ is a submanifold of $T^*J_0^1$. q.e.d.

Recall that the Liouville form of $T^*J_0^1$ is

$$\theta = p_t\, dt + p_x\, dx + p_{\dot{x}}\, d\dot{x} \tag{54}$$

with exterior derivative

$$\omega = d\theta = dp_t \wedge dt + dp_x \wedge dx + dp_{\dot{x}} \wedge d\dot{x} \tag{55}$$

the canonical symplectic form of $T^*J_0^1$.

For $f \in C^\infty(T^*J_0^1)$, the Hamiltonian vector field of f is the unique vector field X_f on $T^*J_0^1$ such that

$$X_f \lrcorner\, \omega = -df,$$

where $\lrcorner$ denote the left interior product (contraction). The Poisson bracket of two functions $f_1, f_2 \in C^\infty(T^*J_0^1)$ is given by

$$\{f_1, f_2\} = X_{f_2}(f_1). \tag{56}$$

It is bilinear, antisymmetric, and it satisfies the Jacobi identity

For the sake of future convenience, define the reparametrization-invariant function h by

$$h = \frac{|\dot{x}|^2}{4} \langle p_{\dot{x}}, p_{\dot{x}}\rangle + \frac{\langle p_x, \dot{x}\rangle}{|\dot{x}|}. \tag{57}$$

Note that h is smooth, because $\dot{x} \neq 0$, and we can use the constraint $h = 0$ instead of $|\dot{x}|^3 \langle p_{\dot{x}}, p_{\dot{x}}\rangle + 4 \langle p_x, \dot{x}\rangle = 0$ in describing the range of $\mathscr{L}$. In other words,

$$\text{range } \mathscr{L} = \{(t, x, \dot{x}, p_t, p_{x,} p_{\dot{x}}) \in T^*J_0^1 \mid p_t = 0,\ \langle p_{\dot{x}}, \dot{x}\rangle = 0 \text{ and } h = 0\}. \tag{58}$$

The Hamiltonian vector fields of the constraint functions p_t, $\langle p_{\dot{x}}, \dot{x}\rangle$ and h are

$$\begin{aligned} X_{p_t} &= \frac{\partial}{\partial t}, \\ X_{\langle p_{\dot{x}}, \dot{x}\rangle} &= \dot{x}\frac{\partial}{\partial \dot{x}} - p_{\dot{x}}\frac{\partial}{\partial p_{\dot{x}}}, \\ X_h &= \frac{1}{2}\langle \dot{x}, \dot{x}\rangle p_{\dot{x}}\frac{\partial}{\partial \dot{x}} + \frac{1}{|\dot{x}|}\dot{x}\frac{\partial}{\partial x} - \frac{1}{|\dot{x}|}p_x\frac{\partial}{\partial p_{\dot{x}}} + \left(-\frac{1}{2}\langle p_{\dot{x}}, p_{\dot{x}}\rangle + \frac{\langle p_x, \dot{x}\rangle}{|\dot{x}|^3}\right)\dot{x}\frac{\partial}{\partial p_{\dot{x}}} \end{aligned}$$

Note that all the Poisson brackets of the constraint functions vanish identically

$$\{\langle p_{\dot{x}}, \dot{x}\rangle, p_t\} = \{h, p_t\} = \{h, \langle p_{\dot{x}}, \dot{x}\rangle\} = 0. \tag{59}$$

This implies that range $\mathscr{L}$ is a coisotropic submanifold of $(T^*J^1_0, \omega)$.

3.2. **Action of** $\mathrm{Diff}_+ \mathbf{R}$ **on** $T^*J^1_0$**.** Recall that for $X = \tau\partial_t \in \mathrm{diff}_+ \mathbf{R}$, the action of the one-parameter subgroup $\exp sX$ on J^1_0 is generated by the vector field $X^1 = \tau\frac{\partial}{\partial t} - \dot{\tau}\dot{x}\frac{\partial}{\partial \dot{x}}$. The lifted action of $\exp sX$ on $T^*J^1_0$ is generated by the Hamiltonian vector field $X_{\mathscr{J}_\tau}$, where

$$\mathscr{J}_\tau(t, x, \dot{x}, p_t, p_x, p_{\dot{x}}) = \langle p_t dt + p_x dx + p_{\dot{x}} d\dot{x}, X_\tau(t, x, \dot{x})\rangle = \tau(t)p_t - \dot{\tau}(t)\langle p_{\dot{x}}, \dot{x}\rangle.$$

The map

$$\mathscr{J}_{\mathrm{diff}} : \mathrm{diff}_+\mathbf{R} \to T^*J^1_0 : \tau\frac{\partial}{\partial t} \mapsto \mathscr{J}_\tau = \tau(t)p_t - \dot{\tau}(t)\langle p_{\dot{x}}, \dot{x}\rangle$$

may be interpreted as the momentum map for the action of the group $\mathrm{Diff}_+ \mathbf{R}$ on $T^*J^1_0$. Writing it this way avoids unnecessary discussion about the topology of the dual of the Lie algebra $\mathrm{diff}_+\mathbf{R}$. The constraint equations $p_t = 0$ and $\langle p_{\dot{x}}, \dot{x}\rangle = 0$ imply that $\mathscr{J}_{\mathrm{diff}}$ vanishes on range $\mathscr{L}$. In other words,

$$\text{range } \mathscr{L} \subseteq \mathscr{J}^{-1}_{\mathrm{diff}}(0).$$

Proposition 3.3. *$\mathscr{J}^{-1}_{\mathrm{diff}}(0)$ is a coisotropic submanifold of $T^*J^1_0$. The null distribution of the pullback of ω to $\mathscr{J}^{-1}_{\mathrm{diff}}(0)$ is spanned by the Hamiltonian vector fields X_{p_t} and $X_{\langle p_{\dot{x}}, \dot{x}\rangle}$.*

Proof. This follows from the proof of theorem 3.2 and equation (59). q.e.d.

Integral curves of the Hamiltonian vector field $X_{p_t} = \frac{\partial}{\partial t}$ are lines parallel to the t-axis. Integral curves of $X_{\langle p_{\dot{x}}, \dot{x}\rangle}$ satisfy equations

$$\begin{aligned} \frac{d}{ds}\dot{x}(s) &= \dot{x}(s), \\ \frac{d}{ds}p_{\dot{x}}(s) &= -p_{\dot{x}}(s). \end{aligned}$$

Hence, for each $\boldsymbol{p} = (t, x, \dot{x}, p_t, p_x, p_{\dot{x}}) \in T^*J^1_0$, the integral manifold of the distribution on $T^*J^1_0$ spanned by X_{p_t} and $X_{\langle p_{\dot{x}}, \dot{x}\rangle}$ that passes through $\boldsymbol{p}$ is

$$O_{\boldsymbol{p}} = \{(u, x, e^s\dot{x}, p_t, p_x, e^{-s}p_{\dot{x}}) \mid (u, s) \in \mathbf{R}^2\}. \tag{60}$$

THEOREM 3.4. *For each* $\boldsymbol{p} \in \mathscr{J}_{\mathrm{diff}}^{-1}(0) \subset T^*J_0^1$, *the orbit of the Lie algebra* $\mathrm{diff}_+ \mathbf{R}$ *through* $\boldsymbol{p}$ *and of the reparametrization group* $\mathrm{Diff}_+ \mathbf{R}$ *coincides with the integral manifold* $O_{\boldsymbol{p}}$ *given by equation (60), where* $p_t = 0$.

Proof. Orbits of the action of the Lie algebra $\mathrm{diff}_+ \mathbf{R}$ on $T^*J_0^1$ are orbits (accessible sets) of the family $\{X_{\mathscr{J}_\tau} \mid \tau \frac{\partial}{\partial t} \in \mathrm{diff}_+ \mathbf{R}\}$ of Hamiltonian vector fields on $T^*J_0^1$. Since $X_{\mathscr{J}_\tau} = X_{\tau p_t} - X_{\dot{\tau}\langle p_{\dot{x}}, \dot{x}\rangle}$ and $\mathscr{J}_\tau$ vanishes on $\mathscr{J}_{\mathrm{diff}}^{-1}(0)$, it follows that the restriction of $X_{\mathscr{J}_\tau}$ to $\mathscr{J}_{\mathrm{diff}}^{-1}(0)$ is

$$X_{\mathscr{J}_\tau | \mathscr{J}_{\mathrm{diff}}^{-1}(0)} = \tau X_{p_t | \mathscr{J}_{\mathrm{diff}}^{-1}(0)} - \dot{\tau} X_{\langle p_{\dot{x}}, \dot{x}\rangle | \mathscr{J}_{\mathrm{diff}}^{-1}(0)}.$$

Therefore, $X_{\mathscr{J}_\tau | \mathscr{J}_{\mathrm{diff}}^{-1}(0)}$ is contained in the distribution spanned by $X_{p_t | \mathscr{J}_{\mathrm{diff}}^{-1}(0)}$ and $X_{\langle p_{\dot{x}}, \dot{x}\rangle | \mathscr{J}_{\mathrm{diff}}^{-1}(0)}$. Hence, for each $\boldsymbol{p} \in \mathscr{J}_{\mathrm{diff}}^{-1}(0)$, the orbit of $\mathrm{diff}_+ \mathbf{R}$ through $\boldsymbol{p}$ coincides with the integral manifold $O_{\boldsymbol{p}}$ given by equation (60).

The reparametrization group $\mathrm{Diff}_+ \mathbf{R}$ acts on J_0^1 by

$$\mathrm{Diff}_+ \mathbf{R} \times J_0^1 \to J_0^1 : (\varphi, (t, x, \dot{x})) \mapsto \left(\varphi(t), x, \frac{\dot{x}}{\dot{\varphi}(t)}\right),$$

where φ is a smooth function on $\mathbf{R}$ such that $\dot{\varphi}(t) > 0$ for $t \in \mathbf{R}$. The lift of this action to $T^*J_0^1$ is

$$\mathrm{Diff}_+ \mathbf{R} \times T^*J_0^1 \to T^*J_0^1 : (\varphi, \boldsymbol{p}) \mapsto \left(\varphi(t), x, \frac{\dot{x}}{\dot{\varphi}(t)}, \frac{p_t}{\dot{\varphi}(t)} + \frac{\langle p_{\dot{x}}, \dot{x}\rangle \ddot{\varphi}(t)}{\dot{\varphi}(t)^2}, p_x, \dot{\varphi}(t) p_{\dot{x}}\right).$$

It preserves $\mathscr{J}_{\mathrm{diff}}^{-1}(0)$. Hence, the orbit of $\mathrm{Diff}_+ \mathbf{R}$ is

$$\{(\varphi(t), x, \dot{\varphi}^{-1}(t)\dot{x}, \dot{\varphi}(t)^{-1} p_t, p_x, \dot{\varphi}(t) p_{\dot{x}}) \in T^*J_0^1 \mid \varphi \in C^\infty(R), \dot{\varphi}(t) > 0\}.$$

The action of $\mathrm{Diff}_+ \mathbf{R}$ preserves $\mathscr{J}_{\mathrm{diff}}^{-1}(0)$, given by $p_t = 0$ and $\langle p_{\dot{x}}, \dot{x}\rangle = 0$. Hence, orbits of $\mathrm{Diff}_+ \mathbf{R}$ contained in $\mathscr{J}_{\mathrm{diff}}^{-1}(0)$ are

$$\{(\varphi(t), x, \dot{\varphi}^{-1}(t)\dot{x}, 0, p_x, \dot{\varphi}(t) p_{\dot{x}}) \in T^*J_0^1 \mid \langle p_{\dot{x}}, \dot{x}\rangle = 0, \varphi \in C^\infty(R), \dot{\varphi}(t) > 0\}.$$

For each t, $\varphi(t) = u$ and $\dot{\varphi}(t) = -s$, are independent. Therefore, orbits of $\mathrm{Diff}_+ \mathbf{R}$ contained in $\mathscr{J}_{\mathrm{diff}}^{-1}(0)$ coincide with the corresponding integral manifolds given by equation (60). q.e.d.

3.3. **Reduction of $\mathbf{Diff_+R}$ symmetries.** In this section, we discuss the space

$$R = \mathscr{J}_{\mathrm{diff}}^{-1}(0) / \mathrm{Diff}_+ \mathbf{R}$$

of $\mathrm{Diff}_+ \mathbf{R}$-orbits in $\mathscr{J}_{\mathrm{diff}}^{-1}(0)$. According to Theorem 3.4, the reduced phase space R is the space of integral manifolds in $\mathscr{J}_{\mathrm{diff}}^{-1}(0)$ of the distribution spanned by X_{p_t} and $X_{\langle p_{\dot{x}}, \dot{x}\rangle}$. We have shown that the orbit of the vector fields $\{X_{p_t}$ and $X_{\langle p_{\dot{x}}, \dot{x}\rangle}\}$ through $\boldsymbol{p} = (t, x, \dot{x}, p_t, p_x, p_{\dot{x}}) \in \mathscr{J}_{\mathrm{diff}}^{-1}(0)$ is

$$O_{\boldsymbol{p}} = \{(u, x, e^s \dot{x}, p_t, p_x, e^{-s} p_{\dot{x}}) \mid (u, s) \in \mathbf{R}^2\}. \tag{61}$$

We are going to show R is a quotient manifold of $\mathscr{J}_{\mathrm{diff}}^{-1}(0)$, which will imply that R has a unique symplectic form ω_R such that

$$\rho^* \omega_R = \iota^* \omega, \tag{62}$$

where $\iota : \mathscr{J}_{\rm diff}^{-1}(0) \to T^*J_0^1$ is the inclusion map and $\rho : \mathscr{J}_{\rm diff}^{-1}(0) \to R$ is the reduction map associating to each point in $\mathscr{J}_{\rm diff}^{-1}(0)$ the orbit of $\{X_{p_t}, X_{\langle p_{\dot{x}}, \dot{x}\rangle}\}$ through that point.

In order to parametrize the reduced phase space R, define spherical coordinates $(\dot{r}, \dot{\alpha}, \dot{\beta})$ by

$$\begin{aligned} \dot{x}^1 &= \dot{r}\sin\dot{\beta}\cos\dot{\alpha}, \\ \dot{x}^2 &= \dot{r}\sin\dot{\beta}\sin\dot{\alpha}, \\ \dot{x}^3 &= \dot{r}\cos\dot{\beta}, \end{aligned} \tag{63}$$

together with the dual momentum variables $(p_{\dot{r}}, p_{\dot{\beta}}, p_{\dot{\alpha}})$ defined by

$$p_{\dot{x}} d\dot{x} = p_{\dot{r}} d\dot{r} + p_{\dot{\beta}} d\dot{\beta} + p_{\dot{\alpha}} d\dot{\alpha}. \tag{64}$$

Proposition 3.5. $p_{\dot{r}} = \langle p_{\dot{x}}, \dot{x}\rangle / \dot{r}$.

Proof. This is a simple verification. q.e.d.

Let

$$S = \{(x, \dot{x}) \in T\mathbf{R}^3 \mid |\dot{x}| = 1\}$$

be the unit sphere bundle over $\mathbf{R}^3$ parametrized by coordinates $(x, \dot{\alpha}, \dot{\beta})$. Even though points $(x, \dot{x}) \in S$ correspond to the arclength parametrization, we use $\dot{x}$ instead of x' in order to emphasize that S is embedded in $T\mathbf{R}^3$. The Liouville form of T^*S is

$$\theta_S = p_x dx + p_{\dot{\beta}} d\dot{\beta} + p_{\dot{\alpha}} d\dot{\alpha}, \tag{65}$$

and

$$\omega_S = d\theta_S$$

is the canonical symplectic form of T^*S.

Proposition 3.6. *There is a unique symplectomorphism $\varkappa : (R, \omega_R) \to (T^*S, \omega_S)$ such that*

$$\varkappa \circ \rho : \mathscr{J}_{\rm diff}^{-1}(0) \to T^*S : (t, x, \dot{x}, 0, p_x, p_{\dot{x}}) \mapsto (x, \dot{\beta}, \dot{\alpha}, p_x, p_{\dot{\beta}}, p_{\dot{\alpha}}),$$

where the $(\dot{\beta}, \dot{\alpha}, p_{\dot{\beta}}, p_{\dot{\alpha}})$ are related to $(x, \dot{x}, p_x, p_{\dot{x}})$ by equations (63) and (64).

Proof. Consider first the space $R_1 = p_t^{-1}(0)/X_{p_t}$ of integral curves of X_{p_t} in $p_t^{-1}(0)$. It is a quotient manifold of $p_t^{-1}(0)$ with projection map

$$\rho_1 : p_t^{-1}(0) \to R_1 : (t, x, \dot{x}, 0, p_x, p_{\dot{x}}) \mapsto (x, \dot{x}, p_x, p_{\dot{x}}).$$

Moreover, it is a symplectic manifold with the symplectic form

$$\omega_1 = dp_x \wedge dx + dp_{\dot{x}} \wedge d\dot{x}.$$

The constraint function $\langle p_{\dot{x}}, \dot{x}\rangle$ is left invariant by the action of X_{p_t}, and pushes forward to a function on R_1, denoted by $\langle p_{\dot{x}}, \dot{x}\rangle_1$. That is, $\langle p_{\dot{x}}, \dot{x}\rangle = \rho_1^* \langle p_{\dot{x}}, \dot{x}\rangle_1$. Moreover, the Hamiltonian vector field $X_{\langle p_{\dot{x}}, \dot{x}\rangle}$ restricted to $p_t^{-1}(0)$ pushes forward to the Hamiltonian vector field on R_1 corresponding to the function $\langle p_{\dot{x}}, \dot{x}\rangle_1$ on R_1. Denote this vector field by $X_{\langle p_{\dot{x}}, \dot{x}\rangle_1}$.

By definition, $\dot{r} = |\dot{x}| \neq 0$ on $T^*J^1_0$. Since $p_{\dot{r}}\dot{r} = \langle p_{\dot{x}}, \dot{x}\rangle$, it follows that on $\rho_1\left(\mathscr{J}^{-1}_{\mathrm{diff}}(0)\right)$ the Hamiltonian vector field of $\langle p_{\dot{x}}, \dot{x}\rangle_1$ is proportional to the Hamiltonian vector field $X_{p_{\dot{r}}} = \frac{\partial}{\partial \dot{r}}$. Therefore, the space $R_2 = \langle p_{\dot{x}}, \dot{x}\rangle_1^{-1}(0)/X_{\langle p_{\dot{x}}, \dot{x}\rangle_1}$ of orbits of $X_{\langle p_{\dot{x}}, \dot{x}\rangle_1}$ in $\langle p_{\dot{x}}, \dot{x}\rangle_1^{-1}(0)$ can be parametrized by $(x, \dot{\beta}, \dot{\alpha}, p_x, p_{\dot{\beta}}, p_{\dot{\alpha}})$. It is a symplectic manifold with the symplectic form

$$\omega_2 = dp_x \wedge dx + dp_{\dot{\beta}} \wedge d\dot{\beta} + dp_{\dot{\alpha}} \wedge d\dot{\alpha}.$$

The coordinates $(x, \dot{\beta}, \dot{\alpha}, p_x, p_{\dot{\beta}}, p_{\dot{\alpha}})$ define a symplectomorphism between (R_2, ω_2) and (T^*S, ω_S), where ω_s is the pullback to T^*S of the canonical symplectic form on $T^*(T\mathbf{R}^3)$. However, $R = \mathscr{J}^{-1}_{\mathrm{diff}}(0)/\{X_{p_t}, X_{\langle p_{\dot{x}}, \dot{x}\rangle}\}$ with the symplectic form ω_R is naturally symplectomorphic to (R_2, ω_2). Hence, (R, ω_R) is symplectomorphic to (T^*S, ω_S). q.e.d.

It follows from Proposition 3.6 that we may identify (R, ω_R) with (T^*S, ω_S).

The action of the Euclidean group SE(3) on $\mathbf{R}^3$ induces a Hamiltonian action of SE(3) on $T^*J^1_0$ generated by the Hamiltonian vector fields X_{p_x}, $X_{p_{\dot{\beta}}}$ and $X_{p_{\dot{\alpha}}}$. This action preserves the constraint functions p_t, $\langle p_{\dot{x}}, \dot{x}\rangle$ and h. In particular, it induces an action of SE(3) on the zero level set $\mathscr{J}^{-1}_{\mathrm{diff}}(0)$ of the momentum map for the action of $\mathrm{diff}_+ \mathbf{R}$. On the other hand, the action of SE(3) on $\mathbf{R}^3$ induces a Hamiltonian action of SE(3) on T^*S, presented as

$$T^*S = \{ (x, \dot{x}, p_x, p_{\dot{x}}) \in T^*(T\mathbf{R}^3) \mid |\dot{x}| = 1,\ p_{\dot{r}} = 0 \},$$

which is generated by the Hamiltonian vector fields of p_x, $p_{\dot{\beta}}$, and $p_{\dot{\alpha}}$ considered as functions on T^*S. Moreover, these actions of SE(3) are intertwined by the reduction map $\rho : \mathscr{J}^{-1}_{\mathrm{diff}}(0) \to R$ followed by the identification $R \cong T^*S$.

3.4. **Hamiltonian dynamics.** The range of the Legendre transformation is characterized as

$$\text{range}\, \mathscr{L} = \mathscr{J}^{-1}_{\mathrm{diff}}(0) \cap h^{-1}(0),$$

where

$$h = \frac{|\dot{x}|^2}{4} \langle p_{\dot{x}}, p_{\dot{x}}\rangle + \frac{\langle p_x, \dot{x}\rangle}{|\dot{x}|}.$$

Since the Legendre transformation is not onto, it is not automatically true that all solutions of Hamilton's equations are the image of a solution of the Euler-Lagrange equations via the Legendre transformation or that the second jet extension of the projection of an integral curve of the Hamiltonian vector field is a solution of the Euler-Lagrange equations. It is a consequence of the calculations in this section that both of these assertions hold for the elastica problem when the arclength parametrization is chosen and thus that it is also true for arbitrary parametrizations. This implies that the solutions of the Euler-Lagrange equations are equivalent to the integral curves of a suitable Hamiltonian vector field. Hence, many of the computations on the Hamiltonian side will not look like much more than a change of variables on the Lagrangian side.

The Hamiltonian vector field of h is

$$X_h = \frac{1}{2}|\dot{x}|^2 p_{\dot{x}}\frac{\partial}{\partial \dot{x}} + \frac{1}{|\dot{x}|}\dot{x}\frac{\partial}{\partial x} - \frac{1}{|\dot{x}|}p_x\frac{\partial}{\partial p_{\dot{x}}} + \left(-\frac{1}{2}|p_{\dot{x}}|^2 + \frac{\langle p_x, \dot{x}\rangle}{|\dot{x}|^3}\right)\dot{x}\frac{\partial}{\partial p_{\dot{x}}}.$$

In order to find the integral curves of X_h on the range of $\mathscr{L}$, observe that they satisfy the equations

$$\begin{aligned}
\frac{d}{ds}x &= \frac{1}{|\dot{x}|}\dot{x},\\
\frac{d}{ds}\dot{x} &= \frac{1}{2}|\dot{x}|^2 p_{\dot{x}},\\
\frac{d}{ds}p_{\dot{x}} &= -\frac{1}{|\dot{x}|}p_x + \left(-\frac{1}{2}|p_{\dot{x}}|^2 + \frac{\langle p_x, \dot{x}\rangle}{|\dot{x}|^3}\right)\dot{x},\\
\frac{d}{ds}p_x &= 0,\\
\frac{d}{ds}t &= 0.
\end{aligned}$$

where s stands for the canonical parameter of the flow of the vector field X_h on $T^*J^1_0$. Multiplying the equation $\frac{d}{ds}\dot{x} = \frac{1}{2}|\dot{x}|^2 p_{\dot{x}}$ by $\dot{x}$, and using the constraint equation $\langle p_{\dot{x}}, \dot{x}\rangle = 0$, yields $\frac{d}{ds}|\dot{x}| = 0$, hence $|\dot{x}(s)| = |\dot{x}_0|$. Since h is reparametrization invariant, without loss of generality, we may assume that $|\dot{x}_0| = 1$, that is, the arclength parametrization. Moreover, on the range of $\mathscr{L}$, $h = 0$, which implies that $\frac{1}{2}\langle p_{\dot{x}}, p_{\dot{x}}\rangle = -2\langle p_x, \dot{x}\rangle$. This leads to

$$\frac{d}{ds}x = \dot{x}, \tag{66}$$

$$\frac{d}{ds}\dot{x} = \frac{1}{2}p_{\dot{x}}, \tag{67}$$

$$\frac{d}{ds}p_{\dot{x}} = -p_x + 3\langle p_x, \dot{x}\rangle\dot{x}, \tag{68}$$

$$\frac{d}{ds}p_x = 0, \tag{69}$$

$$\frac{d}{ds}t = 0. \tag{70}$$

Equation (69) implies that p_x is constant. The angular momentum is given by $l = x \times p_x + \dot{x} \times p_{\dot{x}}$ (see equation (20.) Hence,

$$\begin{aligned}
\frac{d}{ds}l &= \left(\frac{d}{ds}x\right) \times p_x + x \times \left(\frac{d}{ds}p_x\right) + \left(\frac{d}{ds}\dot{x}\right) \times p_{\dot{x}} + \dot{x} \times \left(\frac{d}{ds}p_{\dot{x}}\right)\\
&= \dot{x} \times p_x + \frac{1}{2}p_{\dot{x}} \times p_{\dot{x}} + \dot{x} \times (-p_x + 3\langle p_x, \dot{x}\rangle\dot{x})\\
&= 0,
\end{aligned}$$

and so is conserved. Therefore, $\langle p_x, l\rangle = \langle p_x, \dot{x} \times p_{\dot{x}}\rangle$ is also conserved. Multiplying equations (67) and (68) by p_x yields

$$\begin{aligned} \frac{d}{ds}\langle p_x, \dot{x}\rangle &= \langle p_x, p_{\dot{x}}\rangle, \\ \frac{d}{ds}\langle p_x, p_{\dot{x}}\rangle &= -\langle p_x, p_x\rangle + 3\langle p_x, \dot{x}\rangle^2, \end{aligned}$$

or

$$\frac{d^2}{ds^2}\langle p_x, \dot{x}\rangle = -\langle p_x, p_x\rangle + 3\langle p_x, \dot{x}\rangle^2.$$

Multiplying by $\frac{d}{ds}\langle p_x, \dot{x}\rangle$ and integrating gives

$$\frac{1}{2}\left(\frac{d\langle p_x, \dot{x}\rangle}{ds}\right)^2 = -\langle p_x, p_x\rangle\langle p_x, \dot{x}\rangle + \langle p_x, \dot{x}\rangle^3 + \text{ constant},$$

which can be integrated since it is separable. If $p_x \neq 0$, then this equation gives the component

$$\dot{x}^{\|} = \frac{\langle p_x, \dot{x}\rangle}{|p_x|^2} p_x$$

of $\dot{x}$ parallel to p_x. Integrating $\dot{x}^{\|}(s)$ yields the component of the motion in the direction of p_x. Returning to equations (67) and (68) gives

$$\frac{d}{ds}\dot{x} = \frac{1}{2}p_{\dot{x}}, \tag{71}$$

$$\frac{d}{ds}p_{\dot{x}} = -p_x + 3\langle p_x, \dot{x}\rangle \dot{x}, \tag{72}$$

where $\langle p_x, \dot{x}\rangle$ is assumed known from the discussion above. Hence,

$$\frac{d^2}{ds^2}\dot{x} = -\frac{1}{2}p_x + \frac{3}{2}\langle p_x, \dot{x}\rangle \dot{x}. \tag{73}$$

Writing $\dot{x}$ and p_x in terms of their components $\dot{x}^i$ and p_{x^i},

$$\frac{d^2}{ds^2}\dot{x}^i = -\frac{1}{2}p_{x^i} + \frac{3}{2}\langle p_x, \dot{x}\rangle \dot{x}^i. \tag{74}$$

Division by $\dot{x}_i$ implies

$$\frac{d}{ds}\left(\ln\left|\frac{d}{ds}\dot{x}^i\right|\right) = \frac{1}{\frac{d\dot{x}^i}{ds}}\frac{d^2\dot{x}^i}{ds^2} = -\frac{p_{x^i}}{2\frac{d\dot{x}^i}{ds}} + \frac{3}{2}\langle p_x, \dot{x}\rangle,$$

which implies

$$\ln\left|\frac{d}{ds}\dot{x}^i\right| = -\frac{p_{x^i}}{2}\ln\left|\dot{x}^i\right| + \frac{3}{2}\int \langle p_x, \dot{x}\rangle(s)\, ds$$

or

$$\ln\left|(\dot{x}^i)^{p_{x^i}/2}\frac{d}{ds}\dot{x}^i\right| = \frac{3}{2}\int \langle p_x, \dot{x}\rangle(s)\, ds,$$

so that

$$(\dot{x}^i)^{p_{x^i}/2}\frac{d}{ds}\dot{x}^i = c\exp\left(\int f(s)ds\right),$$

where c is a constant dependent on the initial data. Integrating once more yields

$$\frac{1}{p_{x^i}/2+1}(\dot{x}^i)^{1+p_{x^i}/2} = c\int \exp\left(\int f(s)\,ds\right)ds$$

if $p_{x^i}/2 \neq -1$, and

$$\ln\left|\dot{x}^i\right| = \int \exp\left(\int f(s)\,ds\right)ds$$

if $p_{x^i}/2 = -1$.

REMARK 3.7. The Hamiltonian vector field of $p_t + h$ is

$$X_{p_t+h} = X_{p_t} + X_h = \frac{\partial}{\partial t} + \frac{1}{2}p_{\dot{x}}\frac{\partial}{\partial \dot{x}} + \frac{1}{|\dot{x}|^3}\dot{x}\frac{\partial}{\partial x} - \frac{1}{|\dot{x}|^3}p_x\frac{\partial}{\partial p_{\dot{x}}} + 3\frac{\langle p_x, \dot{x}\rangle}{|\dot{x}|^5}\dot{x}\frac{\partial}{\partial p_{\dot{x}}}.$$

In the arclength parametrization, it is

$$\begin{aligned} \frac{d}{ds}t &= 1, \\ \frac{d}{ds}x &= \dot{x}, \\ \frac{d}{ds}\dot{x} &= \frac{1}{2}p_{\dot{x}}, \\ \frac{d}{ds}p_{\dot{x}} &= -p_x + 3\langle p_x, \dot{x}\rangle\,\dot{x}, \\ \frac{d}{ds}p_x &= 0. \end{aligned}$$

Hence, $t = t_0+s$, and the solutions of this system can be obtained from the solutions for integral curves of X_h by replacing s by $t - t_0$.

THEOREM 3.8. *If* $(t_0, x_0, \dot{x}_0, 0, p_{x_0}, p_{\dot{x}_0}) = \mathscr{L}(t_0, x_0, \dot{x}_0, \ddot{x}_0, \dddot{x}_0)$*, the solution of the Euler-Lagrange equation with initial data* $(t_0, x_0, \dot{x}_0, \ddot{x}_0, \dddot{x}_0)$ *is equivalent to the integral curve of* X_{p_t+h} *through* $(t_0, x_0, \dot{x}_0, 0, p_{x_0}, p_{\dot{x}_0})$ *given above.*

Proof. The Euler-Lagrange equations for a second order Lagrangian are

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} + \frac{d^2}{dt^2}\frac{\partial L}{\partial \ddot{x}} = 0.$$

Since $p_{\dot{x}} = \frac{\partial L}{\partial \ddot{x}}$ and $p_x = \frac{\partial L}{\partial \dot{x}} - \frac{d}{dt}p_{\dot{x}}$, they are $\frac{\partial L}{\partial x} - \frac{d}{dt}p_x = 0$. For elastica,

$$L(x, \dot{x}, \ddot{x}) = \frac{|\ddot{x}|^2}{|\dot{x}|^3} - \frac{\langle \dot{x}, \ddot{x}\rangle^2}{|\dot{x}|^5},$$

$\frac{\partial L}{\partial x} = 0$, and the Euler-Lagrange equations reduce to

$$\frac{d}{dt}p_x = 0.$$

Since $p_x = \frac{\partial L}{\partial \dot{x}} - \frac{d}{dt}p_{\dot{x}}$, the evolution of $p_{\dot{x}}$ is given by

$$\frac{d}{dt}p_{\dot{x}} = -p_x + \frac{\partial L}{\partial \dot{x}},$$

where

$$p_{\dot{x}} = \frac{\partial L}{\partial \ddot{x}} = 2\frac{\ddot{x}}{|\dot{x}|^3} - 2\frac{\langle \dot{x}, \ddot{x}\rangle \dot{x}}{|\dot{x}|^5},$$

and

$$\frac{\partial L}{\partial \dot{x}} = \frac{\partial}{\partial \dot{x}}\left(\frac{|\ddot{x}|^2}{|\dot{x}|^3} - \frac{\langle \dot{x}, \ddot{x}\rangle^2}{|\dot{x}|^5}\right) = -3\frac{|\ddot{x}|^2}{|\dot{x}|^5}\dot{x} - 2\frac{\langle \dot{x}, \ddot{x}\rangle}{|\dot{x}|^5}\ddot{x} + 5\frac{\langle \dot{x}, \ddot{x}\rangle^2}{|\dot{x}|^7}\dot{x}.$$

In the arclength parametrization $|\dot{x}| = 1$, $\langle \dot{x}, \ddot{x}\rangle = 0$, $p_{\dot{x}} = 2\ddot{x}$ and

$$\frac{\partial L}{\partial \dot{x}} = -3\,|\ddot{x}|^2\,\dot{x} = -\frac{3}{4}\,|p_{\dot{x}}|^2\,\dot{x} = 3\,\langle p_x, \dot{x}\rangle\,\dot{x}$$

because

$$h = \frac{1}{4}\,\langle p_{\dot{x}}, p_{\dot{x}}\rangle + \langle p_x, \dot{x}\rangle = 0$$

on the range of $\mathscr{L}$. Thus, the Euler-Lagrange equations of elastica in the arclength parametrization are equivalent to

$$\begin{aligned}
\frac{d}{dt}p_x &= 0,\\
\frac{d}{dt}p_{\dot{x}} &= -p_x + 3\,\langle p_x, \dot{x}\rangle\,\dot{x},\\
\frac{d}{dt}\dot{x} &= \frac{1}{2}p_{\dot{x}},\\
\frac{d}{dt}x &= \dot{x}.
\end{aligned}$$

This system of equations, together with the substitution $t = t_0 + s$, leads to the equation for integral curves of the Hamiltonian vector field of X_{p_t+h} given in the remark above. q.e.d.

4. Comments on Gupta-Bleuler quantization

Quantization of elastica is not physically important. However, it is mathematically important because it serves as a model for quantization of gravity. This is because of the appearance of constraints due to the $\mathrm{Diff}_+\,\mathbf{R}$-invariance.

There are two approaches to quantization of a system with constraints: the Gupta-Bleuler quantization of the extended phase space followed by quantum reduction [1], [8], and Dirac's classical reduction followed by quantization of the reduced phase space [4]. Both approaches have to overcome difficulties that do not appear in the quantization of unconstrained systems. In the case of Gupta-Bleuler quantization, one has to decide how to implement the classical constraints on the quantum level (quantum reduction). In elastica, as well as in general relativity, the reduced phase space fails to be a manifold, therefore in Dirac's approach we would have to decide what is meant by quantization of a singular phase space. Here, we concentrate on the quantum reduction in the Gupta-Bleuler approach; that is on the quantum implementation of the classical constraint equations.

4.1. **Quantization of the extended phase space.** The extended phase space $T^*J^1_0$ is the cotangent bundle of an open subset of $\mathbf{R}^7$. The dynamical variables, which are important in the problem are the coordinates t, x and $\dot{x}$, their conjugate momenta p_t, p_x and $p_{\dot{x}}$, the angular momentum

$$l = x \times p_x + \dot{x} \times p_{\dot{x}},$$

and the constraints

$$\begin{aligned} p_t &= 0, \\ \langle p_{\dot{x}}, \dot{x} \rangle &= 0, \\ |\dot{x}|^3 \langle p_{\dot{x}}, p_{\dot{x}} \rangle + 4 \langle p_x, \dot{x} \rangle &= 0. \end{aligned}$$

The first two constraints are equivalent to vanishing of the momentum map

$$\mathscr{J}_{X_\tau}(t, x, \dot{x}, p_t, p_x, p_{\dot{x}}) = \tau(t) p_t - \dot{\tau}(t) \langle p_{\dot{x}}, \dot{x} \rangle$$

for the action of the 1-parameter subgroup of $\mathrm{Diff}_+ \mathbf{R}$ generated by $X_\tau = \tau \frac{\partial}{\partial t} \in \mathrm{diff}_+ \mathbf{R}$. Since $|\dot{x}| \neq 0$ on J^1_0, the third constraint equation is equivalent to the vanishing of

$$f = \langle p_{\dot{x}}, p_{\dot{x}} \rangle + \frac{4}{|\dot{x}|^3} \langle p_x, \dot{x} \rangle .$$

All these functions on $T^*J^1_0$ are at most quadratic in the conjugate momenta, hence it is convenient to use geometric quantization with the vertical polarization (the polarization tangent to fibres of the cotangent bundle projection). This approach leads to Schrödinger quantization, and we can use results on Schrödinger quantization as they are presented in texts on quantum mechanics.

We consider the space $C^\infty_0(J^1_0) \otimes \mathbf{C}$ of complex valued compactly supported smooth functions Ψ on J^1_0 endowed with the scalar product

$$(\Psi_1 \mid \Psi_2) = \int_{J^1_0} \overline{\Psi}_1(t, x, \dot{x}) \Psi_2(t, x, \dot{x}) dt d_3 x d_3 \dot{x}.$$

The completion of $C^\infty_0(J^1_0) \otimes \mathbf{C}$ with respect to the norm given by this scalar product gives rise to the Hilbert space $\mathfrak{H}$ of quantum states of the system. In the Schrödinger theory, quantum operators associated to the position variables t, x and $\dot{x}$ act on functions $\Psi \in C^\infty_0(J^1_0) \otimes \mathbf{C}$ by multiplication:

$$\begin{aligned} \boldsymbol{Q}_t \Psi(t, x, \dot{x}) &= t \Psi(t, x, \dot{x}), \\ \boldsymbol{Q}_x \Psi(t, x, \dot{x}) &= x \Psi(t, x, \dot{x}), \\ \boldsymbol{Q}_{\dot{x}} \Psi(t, x, \dot{x}) &= \dot{x} \Psi(t, x, \dot{x}). \end{aligned}$$

Quantum operators corresponding to the conjugate momenta $p_t,\ p_x$ and $p_{\dot{x}}$ are partial differential operators

$$\begin{aligned}
\boldsymbol{Q}_{p_t}\Psi(t,x,\dot{x}) &= -i\hbar\frac{\partial\Psi}{\partial t}(t,x,\dot{x}),\\
\boldsymbol{Q}_{p_x}\Psi(t,x,\dot{x}) &= -i\hbar\frac{\partial\Psi}{\partial x}(t,x,\dot{x}),\\
\boldsymbol{Q}_{p_{\dot{x}}}\Psi(t,x,\dot{x}) &= -i\hbar\frac{\partial\Psi}{\partial \dot{x}}(t,x,\dot{x}).
\end{aligned}$$

The angular momentum operator is given by

$$\boldsymbol{Q}_l\Psi(t,x,\dot{x}) = -i\hbar\left(x\times\frac{\partial\Psi}{\partial x}+\dot{x}\times\frac{\partial\Psi}{\partial\dot{x}}\right)(t,x,\dot{x}).$$

These operators are self-adjoint on $\mathfrak{H}$.

As far as quantization of the momenta $\mathscr{J}_{X_\tau}$ is concerned, straightforward replacement of the classical variables by the corresponding quantum operators gives the operator

$$\Psi\mapsto -i\hbar\left(\tau\frac{\partial\Psi}{\partial t}-\dot{\tau}\left\langle\frac{\partial\Psi}{\partial\dot{x}},\dot{x}\right\rangle\right),$$

which is not symmetric. Following Dirac, we can symmetrize this operator obtaining a symmetric operator $\boldsymbol{Q}_{\mathscr{J}_{X_\tau}}$ such that

$$\boldsymbol{Q}_{\mathscr{J}_{X_\tau}}\Psi(t,x,\dot{x}) = -i\hbar\left(\tau\frac{\partial\Psi}{\partial t}-\dot{\tau}\left\langle\dot{x},\frac{\partial\Psi}{\partial\dot{x}}\right\rangle-\dot{\tau}\Psi\right)(t,x,\dot{x}).$$

We are going to show that, for each $X_\tau\in\operatorname{diff}_+\mathbf{R}$, the operator $\boldsymbol{Q}_{\mathscr{J}_{X_\tau}}$ generates a unitary representation on $\mathfrak{H}$ of the one parameter subgroup of $\operatorname{Diff}_+\mathbf{R}$ generated by X_τ.

Finally, straightforward replacement in $f=\langle p_{\dot{x}},p_{\dot{x}}\rangle+\frac{4}{|\dot{x}|^3}\langle p_x,\dot{x}\rangle$ of the classical variables by the corresponding quantum operators gives a symmetric operator $\boldsymbol{Q}_f$ such that

$$\boldsymbol{Q}_f\Psi=\left\{-\hbar^2\dot{\Delta}-\frac{4i\hbar}{|\dot{x}|^3}\left\langle\dot{x},\frac{\partial}{\partial x}\right\rangle\right\}\Psi,$$

where $\dot{\Delta}$ denotes the Laplace operator in the variables $\dot{x}$.

4.2. **Quantization representation of** $\operatorname{Diff}_+\mathbf{R}$**.** The reparametrization group $\operatorname{Diff}_+\mathbf{R}$ acts on $C^\infty(J^1_0)\otimes\mathbf{C}$ by the pullback of the inverse of its action on J^1_0

$$\operatorname{Diff}_+\mathbf{R}\times(C^\infty_0(J^1_0)\otimes\mathbf{C})\to C^\infty(J^1_0)\otimes\mathbf{C}:(\varphi,\Psi)\mapsto((\varphi^1)^{-1})^*\Psi,$$

where φ^1 is given by

$$\varphi^1: J^1\to J^1:(t,x,\dot{x})\mapsto\varphi^1(t,x,\dot{x})=\left(\varphi(t),x,\frac{\dot{x}}{\dot{\varphi}(t)}\right)$$

(see equation (96) in the appendix.)

For the diffeomorphism $\varphi_\epsilon(t)=t+\epsilon\tau(t)+\cdots$ generated by $\tau(t)$,

$$((\varphi^1_\epsilon)^{-1})\Psi(t,x,\dot{x}) = \Psi(t,x,\dot{x}) - \epsilon\left(\tau\frac{\partial}{\partial t} - \dot{x}\dot{\tau}\frac{\partial f}{\partial \dot{x}}\right)\Psi + \cdots$$

and so

$$\begin{aligned}\frac{d}{d\epsilon}((\varphi^1_\epsilon)^{-1})^*\Psi(t,x,\dot{x})_{|\epsilon=0} &= \left(\tau\frac{\partial}{\partial t} - \dot{x}\dot{\tau}\frac{\partial f}{\partial \dot{x}}\right)\Psi(t,x,\dot{x}) \\ &= \left(\frac{-i}{\hbar}\boldsymbol{Q}_{\mathscr{J}_{X_\tau}}\Psi\right)(t,x,\dot{x}) - \dot{\tau}\Psi(t,x,\dot{x}).\end{aligned}$$

Therefore,

$$\begin{aligned}\left(\frac{-i}{\hbar}\boldsymbol{Q}_{\mathscr{J}_{X_\tau}}\Psi\right)(t,x,\dot{x}) &= \frac{d}{d\epsilon}((\varphi^1_\epsilon)^{-1})^*\Psi(t,x,\dot{x})_{|\epsilon=0} + \dot{\tau}\Psi(t,x,\dot{x}) \\ &= \frac{d}{d\epsilon}\{((\varphi^1_\epsilon)^{-1})^*\Psi(t,x,\dot{x}) + \dot{\varphi}_\epsilon(t)\Psi(t,x,\dot{x}))\}_{|\epsilon=0} \\ &= \frac{d}{d\epsilon}\{((\varphi^1_\epsilon)^{-1})^*[\dot{\varphi}_\epsilon(t)\Psi(t,x,\dot{x})]\}_{|\epsilon=0}.\end{aligned}$$

We have established the following result.

Proposition 4.1. *The operator*

$$\frac{-i}{\hbar}\boldsymbol{Q}_{\mathscr{J}_{X_\tau}} = \tau\frac{\partial}{\partial t} - i\hbar\dot{\tau}\left\langle \dot{x}, \frac{\partial}{\partial \dot{x}}\right\rangle + \dot{\tau}$$

generates the action on $C^\infty(J^1_0)\otimes\mathbf{C}$ *of the one parameter group* $\varphi_\epsilon = \exp\epsilon\tau$ *given by*

$$\Xi : (\varphi_\epsilon, \Psi) \mapsto \Xi_{\varphi_\epsilon}\Psi,$$

where

$$\Xi_{\varphi_\epsilon}\Psi(t,x,\dot{x}) = ((\varphi^1_\epsilon)^{-1})^*[\dot{\varphi}_\epsilon\Psi(t,x,\dot{x})].$$

Theorem 4.2. *The map*

$$\Xi : \mathrm{Diff}_+\mathbf{R}\times\left(C^\infty(J^1_0)\otimes\mathbf{C}\right) \to C^\infty(J^1_0)\otimes\mathbf{C} : (\varphi,\Psi)\mapsto \Xi_\varphi\Psi,$$

where

$$\Xi_\varphi\Psi(t,x,\dot{x}) = ((\varphi^1)^{-1})^*[\dot{\varphi}(t)\Psi(t,x,\dot{x})],$$

is a linear representation of $\mathrm{Diff}_+\mathbf{R}$ *on* $C^\infty(J^1_0)\otimes\mathbf{C}$ *which preserves the scalar product.*

Proof. Clearly, Ξ_φ acts linearly on $C^\infty(J^1_0)\otimes\mathbf{C}$. For $\varphi_1,\varphi_2 \in \mathrm{Diff}_+\mathbf{R}$ and $\Psi \in C^\infty(J^1_0)\otimes\mathbf{C}$,

$$\begin{aligned}\Xi_{\varphi_2}\Xi_{\varphi_1}\Psi &= \Xi_{\varphi_2}\{((\varphi^1_1)^{-1})^*[\dot{\varphi}_1\Psi]\} \\ &= ((\varphi^1_2)^{-1})^*\{\dot{\varphi}_2(t)((\varphi^1_1)^{-1})^*[\dot{\varphi}_1\Psi]\} \\ &= [(\varphi_2^{-1})^*\dot{\varphi}_2][((\varphi^1_2)^{-1})^*\{((\varphi^1_1)^{-1})^*\dot{\varphi}_1[((\varphi^1_1)^{-1})^*\Psi]\}] \\ &= [(\varphi_2^{-1})^*\dot{\varphi}_2][(\varphi_2^{-1})^*(\varphi_1^{-1})^*\dot{\varphi}_1][((\varphi^1_2)^{-1})^*(((\varphi^1_1)^{-1})^*\Psi)] \\ &= (\varphi_2^{-1})^*[\dot{\varphi}_2(\varphi_1^{-1})^*\dot{\varphi}_1][((\varphi^1_2\circ\varphi^1_1)^{-1})^*\Psi]\end{aligned}$$

But

$$
\begin{aligned}
[(\varphi_2\circ\varphi_1)^{-1}]^*(\varphi_2\circ\varphi_1\dot{)}(t) &= (\varphi_2\circ\varphi_1\dot{)}((\varphi_2\circ\varphi_1)^{-1}(t))\\
&= (\varphi_2\circ\varphi_1\dot{)}((\varphi_1^{-1}\circ\varphi_2^{-1})(t))\\
&= (\varphi_2\circ\varphi_1\dot{)}((\varphi_1^{-1}(\varphi_2^{-1}(t)))\\
&= \dot\varphi_2(\varphi_1(\varphi_1^{-1}(\varphi_2^{-1}(t))))\dot\varphi_1(\varphi_1^{-1}(\varphi_2^{-1}(t)))\\
&= \dot\varphi_2(\varphi_2^{-1}(t))\dot\varphi_1(\varphi_1^{-1}(\varphi_2^{-1}(t)))\\
&= (\varphi_2^{-1})^*[\dot\varphi_2(\varphi_1^{-1})^*\dot\varphi_1].
\end{aligned}
$$

Therefore,

$$(\Xi_{\varphi_2}\Xi_{\varphi_1}\Psi) = \Xi_{\varphi_2\circ\varphi_1}\Psi,$$

as required.

For $\Psi_1,\Psi_2\in C^\infty(J_0^1)\otimes\mathbf{C}$ and $\varphi\in\mathrm{Diff}_+\,\mathbf{R}$,

$$
\begin{aligned}
(\Xi_\varphi\Psi_1 \mid \Xi_\varphi\Psi_2) &= \int_{J_0^1}\overline{\Xi_\varphi\Psi_1(t,x,\dot x)}\Xi_\varphi\Psi_2(t,x,\dot x)dtd_3xd_3\dot x\\
&= \int_{J_0^1}\overline{((\varphi^1)^{-1})^*\{\dot\varphi(t)\Psi_1(t,x,\dot x)\}}((\varphi^1)^{-1})^*\{\dot\varphi(t)\Psi_2(t,x,\dot x)\}dtd_3xd_3\dot x\\
&= \int_{J_0^1}[\dot\varphi(\varphi^{-1}(t))]^2\overline{\Psi_1(\varphi^{-1}(t),x,\dot x/(\varphi^{-1}\dot{)}(t))}\Psi_1(\varphi^{-1}(t),x,\dot x/(\varphi^{-1}\dot{)}(t))dtd_3xd_3\dot x.
\end{aligned}
$$

Note that the inverse function theorem guarantees

$$\dot\varphi(\varphi^{-1}(t)) = \frac{1}{(\varphi^{-1}\dot{)}(t)}.$$

Introducing new variables

$$\bar t = \varphi^{-1}(t),\ \ \bar x = x \text{ and } \bar x' = \dot x/(\varphi^{-1}\dot{)}(t),$$

yields

$$
\begin{aligned}
d\bar t &= (\varphi^{-1}\dot{)}(t)dt,\\
d\bar x &= dx,\\
d\bar x' &= \frac{1}{(\varphi^{-1}\dot{)}(t)}d\dot x - \frac{(\varphi^{-1}\ddot{)}(t)\dot x}{[(\varphi^{-1}\dot{)}(t)]^2}dt,
\end{aligned}
$$

and

$$d\bar t\, d_3\bar x\, d_3\bar x' = \frac{1}{\left((\varphi^{-1}\dot{)}(t)\right)^2}\, dt\, d_3x\, d_3\dot x,$$

so that

$$dt\, d_3x\, d_3\dot x = \left[(\varphi^{-1}\dot{)}(t)\right]^2 d\bar t\, d_3\bar x\, d_3\bar x'.$$

Therefore,

$$
\begin{aligned}
(\Xi_\varphi \Psi_1 \quad | \quad \Xi_\varphi \Psi_2) &= \int_{J_0^1} [\dot{\varphi}(\varphi^{-1}(t))]^2 \overline{\Psi_1(\bar{t}, \bar{x}, \bar{x}')} \Psi_1(\bar{t}, \bar{x}, \bar{x}')\, dt\, d_3x\, d_3\dot{x} \\
&= \int_{J_0^1} \frac{1}{[(\varphi^{-1}\dot{)}(t)]^2} \overline{\Psi_1(\bar{t}, \bar{x}, \bar{x}')} \Psi_1(\bar{t}, \bar{x}, \bar{x}') \left[(\varphi^{-1}\dot{)}(t)\right]^2 d\bar{t}\, d_3\bar{x}\, d_3\bar{x} \\
&= \int_{J_0^1} \overline{\Psi_1(\bar{t}, \bar{x}, \bar{x}')} \Psi_1(\bar{t}, \bar{x}, \bar{x}')\, d\bar{t}\, d_3\bar{x}\, d_3\bar{x}.
\end{aligned}
$$

q.e.d.

Note that, for every $\Psi \in C_0^\infty(J_0^1) \otimes \mathbf{C}$, and $\Psi' \in C^\infty(J_0^1) \otimes \mathbf{C}$, the integral defining the scalar product

$$
(\Psi' \mid \Psi) = \int_{J_0^1} \overline{\Psi'(t, x, \dot{x})} \Psi(t, x, \dot{x})\, dt\, d_3x\, d_3\dot{x}
$$

can be interpreted as the evaluation on Ψ of the generalized function (distribution) $\Psi' \in (C_0^\infty(J_0^1) \otimes \mathbf{C})'$. The representation Ξ of Diff$_+$ $\mathbf{R}$ on $C_0^\infty(J_0^1) \otimes \mathbf{C}$ extends to a representation of Diff$_+$ $\mathbf{R}$ on $(C^\infty(J_0^1) \otimes \mathbf{C})'$, which we also denote by Ξ, such that

$$
(\Xi_\varphi \Psi' \mid \Psi) = (\Psi' \mid \Xi_{\varphi^{-1}} \Psi)
$$

for $\varphi \in$ Diff$_+$ $\mathbf{R}$, $\Psi' \in (C_0^\infty(J_0^1) \otimes \mathbf{C})'$. Note that if $\Psi' \in C^\infty(J_0^1) \otimes \mathbf{C}$, then the definition of the action Ξ_φ on Ψ' coincides with the definition given in theorem (4.2).

It remains to examine the space of Diff$_+$ $\mathbf{R}$-invariant functions. Since the group Diff$_+$ $\mathbf{R}$ is not compact, the only compactly supported Diff$_+$ $\mathbf{R}$-invariant function in $C_0^\infty(J_0^1) \otimes \mathbf{C}$ is identically zero. Hence, Diff$_+$ $\mathbf{R}$-invariant functions are in $C^\infty(J_0^1) \otimes \mathbf{C}$.

Lemma 4.3. *For each* $\boldsymbol{q} = (t_0, x_0, \dot{x}_0) \in J_0^1$, *the orbit* exp(diff$_+$ $\mathbf{R}$)($\boldsymbol{q}$) *of* diff$_+$ $\mathbf{R}$ *through* $\boldsymbol{q}$ *coincides with the orbit* Diff$_+$ $\mathbf{R}(\boldsymbol{q})$ *of* Diff$_+$ $\mathbf{R}$ *through* $\boldsymbol{q}$

$$
\exp(\mathrm{diff}_+\, \mathbf{R})(\boldsymbol{q}) = \mathrm{Diff}_+\, \mathbf{R}(\boldsymbol{q}).
$$

Proof. The prolongation of $X_\tau \in$ diff$_+$ $\mathbf{R}$ to J_0^1 is $X_\tau^1 = \tau \frac{\partial}{\partial t} - \dot{\tau}\dot{x}\frac{\partial}{\partial \dot{x}}$. Its integral curves satisfy the differential equations

$$
\frac{dt}{ds} = \tau, \ \frac{dx}{ds} = 0 \text{ and } \frac{d\dot{x}}{ds} = -\dot{\tau}\dot{x}.
$$

Hence,

$$
\frac{dt}{\tau} = ds
$$

and

$$
\int_{t_0}^{t} \frac{dt'}{\tau(t')} = s.
$$

Choosing $\tau(t) = e^{-t}$ yields

$$
\int_{t_0}^{t} \frac{dt'}{\tau(t')} = \int_{t_0}^{t} e^{t'} dt' = e^t - e^{t_0}.
$$

Hence, $e^t - e^{t_0} = s$, which implies that

$$t = \ln \left| s + e^{-t_0} \right| .$$

Since the range of the logarithm is $(-\infty, \infty)$, the range of values of t on the orbit of $\mathrm{diff}_+ \mathbf{R}$ through $\boldsymbol{q}$ is $(-\infty, \infty)$.

Moreover, for $i = 1, 2, 3$,

$$\frac{d\dot{x}_i}{ds} = -\dot{\tau}\dot{x}_i$$

implies

$$\frac{d\dot{x}_i}{\dot{x}_i} = -\dot{\tau}(t(s))ds$$

so that

$$\dot{x}(s) = \dot{x}_0 \exp\left(-\int_0^s \dot{\tau}(t(s))ds \right).$$

Since $\tau(t)$ is an arbitrary function of t, it follows that the orbit of $\mathrm{diff}_+ \mathbf{R}$ through $q = (t_0, x_0, \dot{x}_0)$ is

$$\exp(\mathrm{diff}_+ \mathbf{R})(q) = \left\{ (u, x_0, e^v \dot{x}) \mid (u, v) \in \mathbf{R}^2 \right\}.$$

The action of $\mathrm{Diff}_+ \mathbf{R}$ on J^1_0 is

$$\mathrm{Diff}_+ \mathbf{R} \times J^1_0 \to J^1_0 : (\varphi, (t, x, \dot{x})) \mapsto \left(\varphi(t), x, \frac{\dot{x}}{\dot{\varphi}(t)} \right),$$

where $\dot{\varphi}(t) > 0$. As $\varphi(t)$ and $\dot{\varphi}(t)$ are independent, it follows that the orbit of $\mathrm{Diff}_+ \mathbf{R}$ through q is

$$\mathrm{Diff}_+ \mathbf{R}(q) = \{(u, x_0, w\dot{x}) \mid (u, w) \in \mathbf{R}^2,\ w > 0\}.$$

Hence, $\exp(\mathrm{diff}_+ \mathbf{R})(q) = \mathrm{Diff}_+ \mathbf{R}(q)$.

q.e.d.

Theorem 4.4. *A function* $\Psi' \in C^\infty(J^1_0) \otimes \mathbf{C}$ *is* $\mathrm{Diff}_+ \mathbf{R}$*-invariant if and only if*

$$\boldsymbol{Q}_{\mathscr{J}_{X_\tau}} \Psi' = 0$$

for all $\tau \in \mathrm{diff}_+ \mathbf{R}$.

Proof. If $\Psi' \in C^\infty(J^1_0) \otimes \mathbf{C}$ is $\mathrm{Diff}_+ \mathbf{R}$-invariant, then it is invariant under the action of every one-parameter subgroup of $\mathrm{Diff}_+ \mathbf{R}$. By Proposition 4.1, actions of one parameter subgroups of $\mathrm{Diff}_+ \mathbf{R}$ on $C^\infty(J^1_0) \otimes \mathbf{C}$ are generated by $\frac{-i}{\hbar}\boldsymbol{Q}_{\mathscr{J}_{X_\tau}}$ for $\tau \in \mathrm{diff}_+ \mathbf{R}$. Hence, $\boldsymbol{Q}_{\mathscr{J}_{X_\tau}} \Psi' = 0$ for all $X_\tau \in \mathrm{diff}_+ \mathbf{R}$.

Conversely, suppose that Ψ' is a function in $C^\infty(J^1_0) \otimes \mathbf{C}$ such that $\boldsymbol{Q}_{\mathscr{J}_\tau} \Psi' = 0$ for all $\tau \in \mathrm{diff}_+ \mathbf{R}$. Hence, $\Xi_{\varphi_\epsilon} \Psi' = \Psi'$ for every one parameter subgroup φ_ϵ of $\mathrm{Diff}_+ \mathbf{R}$. Recall that, for $\varphi \in \mathrm{Diff}_+ \mathbf{R}$,

$$\begin{aligned}
\Xi_\varphi \Psi'(t, x, \dot{x}) &= ((\varphi^1)^{-1})^* [\dot{\varphi}(t) \Psi'(t, x, \dot{x})] = [(\varphi^{-1})^* \dot{\varphi}(t)][((\varphi^1)^{-1})^* \Psi'(t, x, \dot{x})] \\
&= \dot{\varphi}(\varphi^{-1}(t)) \Psi'(\varphi^{-1}(t, x, \dot{x})) = \frac{1}{\frac{d\varphi^{-1}(t)}{dt}} \Psi'(\varphi^{-1}(t, x, \dot{x}))
\end{aligned}$$

by the inverse function theorem. Therefore,

$$\Xi_{\varphi^{-1}}\Psi'(t,x,\dot{x}) = \frac{1}{\dot{\varphi}(t)}\Psi'(\varphi(t,x,\dot{x})).$$

It follows from the lemma above that there exists a finite sequence of one parameter subgroups $\varphi_{\epsilon_1},...,\varphi_{\epsilon_k}$ such that

$$\varphi^1(t,x,\dot{x}) = \varphi_{\epsilon_k}(...(\varphi_{\epsilon_1}((t,x,\dot{x})))).$$

Since

$$\varphi^1(t,x,\dot{x}) = \left(\varphi(t), x, \frac{\dot{x}}{\dot{\varphi}(t)}\right),$$

$$\varphi(t) = \varphi_{\epsilon_k}(...(\varphi_{\epsilon_1}(t))) = \varphi_{\epsilon_k}\circ ...\circ\varphi_{\epsilon_1}(t)$$

and

$$\dot{\varphi}(t) = \dot{\varphi}_{\epsilon_k}(...(\varphi_{\epsilon_1}(t)))...\dot{\varphi}_{\epsilon_1}(t) = \frac{d}{dt}\varphi_{\epsilon_k}\circ ...\circ\varphi_{\epsilon_1}(t).$$

Hence,

$$\begin{aligned}
\Xi_{\varphi^{-1}}\Psi'(t,x,\dot{x}) &= \frac{1}{\dot{\varphi}(t)}\Psi'(\varphi^1(t,x,\dot{x})) = \frac{1}{\dot{\varphi}(t)}\Psi'\left(\varphi(t), x, \frac{\dot{x}}{\dot{\varphi}(t)}\right)\\
&= \frac{1}{\frac{d}{dt}\varphi_{\epsilon_k}\circ ...\circ\varphi_{\epsilon_1}(t)}\Psi'\left(\varphi_{\epsilon_k}\circ ...\circ\varphi_{\epsilon_1}(t), x, \frac{\dot{x}}{\frac{d}{dt}\varphi_{\epsilon_k}\circ ...\circ\varphi_{\epsilon_1}(t)}\right)\\
&= \Xi_{(\varphi_{\epsilon_k}\circ ...\circ\varphi_{\epsilon_1})^{-1}}\Psi'(t,x,\dot{x}) = \Xi_{\varphi_{\epsilon_1}^{-1}\circ ...\circ\varphi_{\epsilon_k}^{-1}}\Psi'(t,x,\dot{x}) = \Xi_{\varphi_{-\epsilon_1}\circ ...\circ\varphi_{-\epsilon_k}}\Psi'(t,x,\dot{x})\\
&= \Xi_{\varphi_{-\epsilon_1}}...\Xi_{\varphi_{-\epsilon_k}}\Psi'(t,x,\dot{x}) = \Psi'(t,x,\dot{x}),
\end{aligned}$$

because Ψ' is invariant under the action of one parameter subgroups of Diff$_+$ **R**.

q.e.d.

4.3. **Quantum implementation of constraints.** In the Gupta-Bleuler approach, we first quantize the extended phase space. This associates to each constraint function the corresponding quantum operator. The next step is to implement the constraint conditions on the quantum level. This is done by placing a restriction on states of the system.

Definition 4.5. Admissible quantum states are the common zero eigenstates of the quantum operators associated to the constraint functions.

For elastica, we have written the classical constraints in the form

$$\begin{aligned}
\mathscr{J}_{X_\tau} &\equiv \tau(t)p_t - \dot{\tau}(t)\langle p_{\dot{x}}, \dot{x}\rangle = 0 \text{ for } X_\tau \in \text{diff}_+\,\mathbf{R},\\
f &\equiv \langle p_{\dot{x}}, p_{\dot{x}}\rangle + \frac{4}{|\dot{x}|^3}\langle p_x, \dot{x}\rangle = 0.
\end{aligned}$$

We have shown in the preceding section that $Q_{\mathscr{J}_{X_\tau}}\Psi = 0$ for all X_τ is equivalent to the Diff$_+$ **R**-invariance of Ψ. Hence, admissible quantum states of the elastica are Diff$_+$ **R**-invariant. In particular, if $\Psi(t,x,\dot{x})$ is admissible then Ψ is independent of t and (positive) homogeneous of degree -1 in the second variable $\dot{x}$. Therefore, we can write $\Psi = \Psi(x,\dot{x})$, and

$$\dot{x}^i\frac{\partial\Psi}{\partial\dot{x}^i} = -\Psi.$$

The second constraint equation $f = 0$ implies that admissible quantum states of elastica satisfy the equation $\boldsymbol{Q}_f\Psi = 0$, that is

$$-\hbar^2\dot{\Delta}\Psi - \frac{4i\hbar}{|\dot{x}|^3}\left\langle \dot{x}, \frac{\partial\Psi}{\partial x}\right\rangle = 0. \tag{75}$$

Since $[\boldsymbol{Q}_f, \boldsymbol{Q}_{\mathscr{J}X_\tau}] = 2(i\hbar\dot{\tau})\boldsymbol{Q}_f$, there are no further quantum constraints.

Conclusion 4.6. The space $\mathfrak{A}$ of admissible states of quantum elastica consists of functions $\Psi \in C^\infty(J_0^1) \otimes \mathbf{C}$ that are invariant under the quantization representation of $\mathrm{Diff}_+\,\mathbf{R}$ and satisfy equation (75).

The operators $\boldsymbol{Q}_{p_x}$ and $\boldsymbol{Q}_l$ commute with $\boldsymbol{Q}_{\mathscr{J}X_\tau}$ and $\boldsymbol{Q}_f$, which implies that they act on the space $\mathfrak{A}$ of admissible states. To have a full quantum theory, we need a scalar product on the space of admissible states such that the actions of $\boldsymbol{Q}_{p_x}$ and $\boldsymbol{Q}_l$ on $\mathfrak{A}$ extend to self-adjoint operators.

Appendix A. Second order variational problems

In this section we list some common notions of second order variational problems including a discussion of symmetries and conservation laws. This will serve to define terms, summarize known results and establish our notation. We pay special attention to the case of parametrization independence and the second Noether theorem, with an emphasis on arclength parametrization. Finally, we discuss the Hamiltonian formalism from this point of view. As almost all of the results are established by straightforward (but sometimes tedious) calculation, we do not give proofs in this section, referring the reader to the literature, the main reference being Grässer [8].

A.1. **Variation of the action integral.** A curve $[t_0, t_1] \to \mathbf{R}^n : t \mapsto x(t)$ can be uniquely described by the corresponding section

$$\sigma : [t_0, t_1] \to [t_0, t_1] \times \mathbf{R}^n : t \mapsto (t, x(t)). \tag{76}$$

We consider $Q = \mathbf{R}\times\mathbf{R}^n$ as a bundle over $\mathbf{R}$ with typical fibre $\mathbf{R}^n$. For each integer $k \geq 0$, we denote by J^k the k-th jet of sections of Q. Furthermore, interpret the section σ given in (76) as a local section of Q and denote by $j^k\sigma : [t_0, t_1] \to J^k$ the k-jet extension of σ. This paper considers variational problems defined by second order Lagrangians. For a Lagrangian $L : J^2 \to \mathbf{R} : (t, x, \dot{x}, \ddot{x}) \mapsto L(t, x, \dot{x}, \ddot{x})$, the corresponding action integral is

$$A(\sigma) = \int_{t_0}^{t_1} (L \circ j^2\sigma)\,dt = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t), \ddot{x}(t))\,dt.$$

A variation of a section (without variation of time) $\sigma \mapsto \sigma + \delta\sigma : t \mapsto x(t) + \delta x(t)$ extends to the second jets as

$$j^2\sigma \mapsto j^2\sigma + \delta j^2\sigma : t \mapsto (x(t) + \delta x(t), \dot{x}(t) + \delta\dot{x}(t), \ddot{x}(t) + \delta\ddot{x}(t)),$$

where

$$\delta\dot{x}(t) = \frac{d}{dt}\delta x(t) \quad\text{and}\quad \delta\ddot{x}(t) = \frac{d}{dt}\delta\dot{x}(t) = \frac{d^2}{dt^2}\delta x(t).$$

Then, integrating by parts twice, it follows that the action varies as

$$\begin{aligned}\delta A(\sigma) &= \int_{t_0}^{t_1} \delta L(t, x(t), \dot{x}(t), \ddot{x}(t))\, dt \\ &= \int_{t_0}^{t_1} \left\{ \frac{\partial L}{\partial x}(t) - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}(t) + \frac{d^2}{dt^2}\frac{\partial L}{\partial \ddot{x}}(t) \right\} \delta x\, dt + \\ &\quad + \left(\frac{\partial L}{\partial \dot{x}}(t) - \frac{d}{dt}\frac{\partial L}{\partial \ddot{x}}(t) \right) \delta x \Big|_{t_0}^{t_1} + \frac{\partial L}{\partial \ddot{x}}(t)\, \delta \dot{x} \Big|_{t_0}^{t_1}.\end{aligned}$$

It follows from the fundamental lemma of the calculus of variations that

Conclusion A.1. The action integral $A(\sigma)$ is stationary with respect to all variations $\sigma \mapsto \sigma + \delta\sigma : t \mapsto x(t) + \delta x(t)$, such that δx and $\delta \dot{x}$ vanish on the boundary, if and only if the section σ satisfies the Euler-Lagrange equations

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} + \frac{d^2}{dt^2}\frac{\partial L}{\partial \ddot{x}} = 0. \tag{77}$$

Consider now the boundary terms in the variation. The partial derivative $\frac{\partial L}{\partial \ddot{x}}$ is a map from J^2 to $\mathbf{R}^n$, and

$$\frac{\partial L}{\partial \ddot{x}}(t)\, \delta \dot{x} = \left\langle \frac{\partial L}{\partial \ddot{x}}(t, x(t), \dot{x}(t), \ddot{x}(t)),\ \delta \dot{x}(t) \right\rangle,$$

where the angle bracket denotes the Euclidean scalar product in $\mathbf{R}^n$. However,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \ddot{x}}(t) \right) = \left(\frac{\partial}{\partial t}\frac{\partial L}{\partial \ddot{x}} + \dot{x}\frac{\partial}{\partial x}\frac{\partial L}{\partial \ddot{x}} + \ddot{x}\frac{\partial}{\partial \dot{x}}\frac{\partial L}{\partial \ddot{x}} + \dddot{x}\frac{\partial}{\partial \ddot{x}}\frac{\partial L}{\partial \ddot{x}} \right)$$

depends on the third derivative $\dddot{x}$ of the section σ, and hence, $\frac{d}{dt}\frac{\partial L}{\partial \ddot{x}}$ can be interpreted as a map from J^3 to $\mathbf{R}^n$. Using the projection map

$$\pi_{32} : J^3 \to J^2 : (t, x, \dot{x}, \ddot{x}, \dddot{x}) \mapsto (t, x, \dot{x}, \ddot{x}),$$

define Ostrogradski's momenta by

$$\begin{aligned} p_{\dot{x}} &= \pi_{32}^* \frac{\partial L}{\partial \ddot{x}}, \\ p_x &= \pi_{32}^* \frac{\partial L}{\partial \dot{x}} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{x}}, \end{aligned}$$

and interpret them as maps from J^3 to $\mathbf{R}^n$. In the following, in order to simplify the notation, the pull-back sign is omitted and an overdot is used to denote the derivative with respect to t. This leads to the usual expressions

$$\begin{aligned} p_{\dot{x}} &= \frac{\partial L}{\partial \ddot{x}}, \\ p_x &= \frac{\partial L}{\partial \dot{x}} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{x}} = \frac{\partial L}{\partial \dot{x}} - \dot{p}_{\dot{x}}. \end{aligned} \tag{78}$$

With this notation, the variation equation is

$$\delta A(\sigma) = \int_{t_0}^{t_1} \left\{ \frac{\partial L}{\partial x}(t) - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}(t) + \frac{d^2}{dt^2}\frac{\partial L}{\partial \ddot{x}}(t) \right\} \delta x \, dt + \, p_x \, \delta x|_{t_0}^{t_1} + \, p_{\dot{x}} \, \delta \dot{x}|_{t_0}^{t_1} . \quad (79)$$

With an eye towards towards the Cartan form, it is convenient to reinterpret a variation as the Lie derivative with respect to a vector field. Let a variation $\sigma \mapsto \sigma + \delta\sigma : t \mapsto x(t) + \delta x(t)$ of σ be given by a vector field X^2 on J^2 that is tangent to the fibres of the source map $J^2 \to [t_0, t_1] : (t, x, \dot{x}, \ddot{x}) \mapsto t$. In other words, the variation $j^2\sigma \mapsto j^2\sigma + \delta j^2\sigma$ is given by the prolongation

$$X^2 = X_x \frac{\partial}{\partial x} + X_{\dot{x}} \frac{\partial}{\partial \dot{x}} + X_{\ddot{x}} \frac{\partial}{\partial \ddot{x}}$$

of X to J^2, where

$$\delta(x)(t) = X_x(\sigma) \text{ and } X_{\dot{x}} = \frac{d}{dt} X_x \text{ and } X_{\ddot{x}} = \frac{d}{dt} X_{\dot{x}} = \frac{d^2}{dt^2} X_x.$$

With this identification,

$$\begin{aligned} \delta A(\sigma) &= \int_{t_0}^{t_1} \left\{ \frac{\partial L}{\partial x}(t)\, \delta x + \frac{\partial L}{\partial \dot{x}}(t)\, \delta \dot{x} + \frac{\partial L}{\partial \ddot{x}}(t)\, \delta \ddot{x} \right\} dt \qquad (80) \\ &= \int_{t_0}^{t_1} \pounds_{X^2}(L\, dt) = \int_{t_0}^{t_1} X^2 \lrcorner\, d(L\, dt) \end{aligned}$$

because

$$\pounds_{X^2}(L\, dt) = X^2 \lrcorner\, d(L\, dt) + d(X^2 \lrcorner\, L\, dt),$$

and the the assumption that X is tangent to the fibres of the source map implies that $X^2 \lrcorner\, L\, dt = 0$. Comparing equations (79) and (80) yields

$$\int_{t_0}^{t_1} X^2 \lrcorner\, d(L\, dt) = \int_{t_0}^{t_1} \left\{ \frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} + \frac{d^2}{dt^2}\frac{\partial L}{\partial \ddot{x}} \right\} X_x \, dt + \left\langle p_x \, dx + p_{\dot{x}} \, d\dot{x}, X^3 \right\rangle \Big|_{t_0}^{t_1}, \quad (81)$$

where $\langle p_x \, dx + p_{\dot{x}} d\dot{x}, X^3 \rangle$ is the evaluation of the one-form $p_x \, dx + p_{\dot{x}} \, d\dot{x}$ on the first jet bundle on the third jet prolongation X^3 of X (see proposition (A.6)).

A.2. **The Cartan form.** The contact forms of the second jet bundle J^2 are

$$\theta_1 = dx - \dot{x}\, dt \;\text{ and }\; \theta_2 = d\dot{x} - \ddot{x}\, dt.$$

Their importance stems from the following

Proposition A.2. *A section* $\sigma : [t_0, t_1] \to J^2 : t \mapsto (t, x(t), \dot{x}(t), \ddot{x}(t))$ *is the jet extension of the section* $[t_0, t_1] \to Q : t \mapsto (t, x(t))$ *by the source map* $J^2 \to Q : (t, x, \dot{x}, \ddot{x}) \mapsto (t, x)$ *if and only if* $\sigma^*\theta_1 = 0$ *and* $\sigma^*\theta_2 = 0$.

Definition A.3. The Cartan form corresponding to a Lagrangian L is the one-form Θ on J^3 given by

$$\Theta = L\, dt + p_x(dx - \dot{x}\, dt) + p_{\dot{x}}(d\dot{x} - \ddot{x}\, dt), \quad (82)$$

where p_x and $p_{\dot{x}}$ are the Ostrogradski momenta (78).

Observe that Θ may by written in the form

$$\Theta = p_x\,dx + p_{\dot{x}}\,d\dot{x} - H\,dt, \tag{83}$$

where

$$H = p_x\dot{x} + p_{\dot{x}}\ddot{x} - L \tag{84}$$

is the Hamiltonian of the theory. Since Θ differs from the Lagrange form $L\,dt$ by terms that are proportional to the contact forms, it follows that for any section σ the action $A(\sigma)$ can be expressed as the integral of Θ over $j^3\sigma$. In other words,

$$A(\sigma) = \int_{t_0}^{t_1} (L \circ j^2\sigma)\,dt = \int_{t_0}^{t_1} (j^2\sigma)^* L\,dt = \int_{t_0}^{t_1} (j^3\sigma)^*\Theta. \tag{85}$$

Therefore, the Cartan form Θ may be used instead of the Lagrange form $L\,dt$ to describe the variational problem under consideration. Other aspects of the Cartan form are discussed in [11].

Definition A.4. A Langrangian L is *regular* if the matrix

$$\frac{\partial^2 L}{\partial\ddot{x}^j \partial\ddot{x}^i}$$

is non-singular.

Theorem A.5. *Let γ be a section of the source map $J^3 \to [t_0, t_1]$ projecting to a section σ of $[t_0, t_1] \times \mathbf{R}^n \to [t_0, t_1]$ and let $j^3\sigma$ be the third jet extension of σ.*

(1) *If $\gamma = j^3\sigma$, then σ satisfies the Euler-Lagrange equations if and only if the tangent bundle of the range of γ is contained in the kernel of $d\Theta$.*

(2) *If the Lagrangian L is regular and the tangent bundle of the range of γ is contained in the kernel of $d\Theta$, then $\gamma = j^3\sigma$ and σ satisfies the the Euler-Lagrange equations.*

A.3. **Symmetries and conservation laws.**

A.3.1. *Symmetries of the Lagrange form.* Consider an infinitesimal transformation in $(t_0, t_1) \times \mathbf{R}^n$ generated by the vector field

$$X = \tau\frac{\partial}{\partial t} + \xi^i \frac{\partial}{\partial x^i}. \tag{86}$$

Proposition A.6. *The prolongations X^1, X^2 and X^3 of the vector field X in equation (86) to the jet bundles J^1, J^2 and J^3, respectively, are*

$$\begin{aligned}
X^1 &= \tau\frac{\partial}{\partial t} + \xi^i\frac{\partial}{\partial x^i} + (\dot{\xi} - \dot{x}\dot{\tau})\frac{\partial}{\partial \dot{x}},\\
X^2 &= \tau\frac{\partial}{\partial t} + \xi\frac{\partial}{\partial x} + (\dot{\xi} - \dot{x}\dot{\tau})\frac{\partial}{\partial \dot{x}} + (\ddot{\xi} - 2\ddot{x}\dot{\tau} - \dot{x}\ddot{\tau})\frac{\partial}{\partial \ddot{x}},\\
X^3 &= \tau\frac{\partial}{\partial t} + \xi\frac{\partial}{\partial x} + (\dot{\xi} - \dot{x}\dot{\tau})\frac{\partial}{\partial \dot{x}} + (\ddot{\xi} - 2\ddot{x}\dot{\tau} - \dot{x}\ddot{\tau})\frac{\partial}{\partial \ddot{x}} + (\dddot{\xi} - 3\dddot{x}\dot{\tau} - 3\ddot{x}\ddot{\tau} - \dot{x}\dddot{\tau})\frac{\partial}{\partial \dddot{x}}.
\end{aligned}$$

It remains to relate the prolongations of X to the contact forms $\theta_1 = dx - \dot{x}\,dt$ and $\theta_2 = d\dot{x} - \ddot{x}\,dt$.

PROPOSITION A.7. *Let $I = [t_0, t_1]$. For a section σ of $I \times \mathbf{R}^n \to I$,*

$$j^1\sigma^* £_{X^1}\theta_1 = 0 \text{ and } j^2\sigma^* £_{X^2}\theta_2 = 0.$$

PROPOSITION A.8. *The action integral*

$$A(\sigma) = \int_I (L \circ j^2\sigma)\,dt = \int_I j^2\sigma^*(L\,dt) = \int_{j^2\sigma(I)} L\,dt. \tag{87}$$

is invariant under the one-parameter local group $\exp tX^2$ *of local diffeomorphisms of J^2 generated by X^2 if*

$$j^2\sigma^*(£_{X^2}(L\,dt)) = 0.$$

Moreover, $\frac{d}{dt}A(\exp tX(\sigma))|_{t=0}$ if and only if $£_{X^2}(L\,dt) = 0$.

DEFINITION A.9. A vector field X on $I \times \mathbf{R}^n$ is an *infinitesimal symmetry* of the Lagrange form $L\,dt$ if $£_{X^2}(L\,dt) = 0$.

LEMMA A.10. *The Lie derivative of the Lagrange form $L\,dt$ with respect to X^2 is*

$$£_{X^2}(L\,dt) = \left(\tau\frac{\partial L}{\partial t} + \xi^i\frac{\partial L}{\partial x^i} + (\dot{\xi} - \dot{x}\dot{\tau})\frac{\partial L}{\partial \dot{x}} + (\ddot{\xi} - 2\ddot{x}\dot{\tau} - \dot{x}\ddot{\tau})\frac{\partial L}{\partial \ddot{x}}\right)dt + L\,d\tau.$$

Hence, for every section σ of $I \times \mathbf{R}^n \to I$,

$$j^2\sigma^*(£_{X^2}(L\,dt)) = \left(\tau\frac{\partial L}{\partial t} + \xi^i\frac{\partial L}{\partial x^i} + (\dot{\xi} - \dot{x}\dot{\tau})\frac{\partial L}{\partial \dot{x}} + (\ddot{\xi} - 2\ddot{x}\dot{\tau} - \dot{x}\ddot{\tau})\frac{\partial L}{\partial \ddot{x}} + L\dot{\tau}\right)dt.$$

LEMMA A.11. *(Noether identities.) The equation $j^2\sigma^*(£_{X^2}(L\,dt)) = 0$ is equivalent to*

$$\begin{aligned} \frac{d}{dt}\left(L\tau + \left(\frac{\partial L}{\partial \dot{x}} - \frac{d}{dt}\left(\frac{\partial L}{\partial \ddot{x}}\right)\right)(\xi - \tau\dot{x}) + \frac{\partial L}{\partial \ddot{x}}\frac{d}{dt}(\xi - \tau\dot{x})\right) = \\ = \quad -\left(\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) + \frac{d^2}{dt^2}\left(\frac{\partial L}{\partial \ddot{x}}\right)\right)(\xi - \tau\dot{x}). \end{aligned} \tag{88}$$

Proof. The details of this routine but lengthy calculation may be found in [17]. q.e.d.

REMARK A.12. The Noether identity is essentially the extension of the equation $j^2\sigma^*(£_{X^2}(L\,dt)) = 0$ to the fourth jet bundle. More precisely, if $\pi_{4,2} : J^4 \to J^2$ is the natural projection and X^4 is the prolongation of X to J^4, then

$$\begin{aligned} j^4\sigma^*(\pi_{4,2}^*(£_{X^2}(L\,dt))) \quad = \quad & j^4\sigma^*\left(\left(\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) + \frac{d^2}{dt^2}\left(\frac{\partial L}{\partial \ddot{x}}\right)\right)(\xi - \tau\dot{x})\right) + \\ & + j^4\sigma^*\left(\frac{d}{dt}\left(L\tau + \left(\frac{\partial L}{\partial \dot{x}} - \frac{d}{dt}\left(\frac{\partial L}{\partial \ddot{x}}\right)\right)(\xi - \tau\dot{x}) + \frac{\partial L}{\partial \ddot{x}}\frac{d}{dt}(\xi - \tau\dot{x})\right)\right). \end{aligned} \tag{89}$$

An immediate corollary of the Noether identity is the following conservation law.

THEOREM A.13. *(First Noether theorem.) To every infinitesimal symmetry* $X = \tau\frac{\partial}{\partial t} + \xi\frac{\partial}{\partial x}$ *of the Lagrange form* $L\,dt$, *there corresponds a conserved quantity*

$$\mathscr{J}_X = L\tau + \left(\frac{\partial L}{\partial \dot{x}} - \frac{d}{dt}\left(\frac{\partial L}{\partial \ddot{x}}\right)\right)(\xi - \tau\dot{x}) + \frac{\partial L}{\partial \ddot{x}}\frac{d}{dt}(\xi - \tau\dot{x}). \tag{90}$$

That is, $\mathscr{J}_X$ *is constant along solutions of the Euler-Lagrange equations*

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) + \frac{d^2}{dt^2}\left(\frac{\partial L}{\partial \ddot{x}}\right) = 0.$$

In other words, if σ *satisfies the Euler-Lagrange equations, then* $j^3\sigma^*(X^3 \lrcorner\, \Theta)$ *is constant.*

EXAMPLE A.14. (Conservation of linear momentum.) If the Lagrangian L does not depend on the coordinate x^i, then $X = \frac{\partial}{\partial x^i}$ is an infinitesimal symmetry and the momentum

$$\mathscr{J}_{\frac{\partial}{\partial x^i}} = \frac{\partial L}{\partial \dot{x}^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \ddot{x}^i}\right)$$

is conserved.

EXAMPLE A.15. (Conservation of energy) If the Lagrangian L does not depend on the parameter t, then $X = \frac{\partial}{\partial t}$ is an infinitesimal symmetry and the energy

$$H = p_x\dot{x} + p_{\dot{x}}\ddot{x} - L = -\mathscr{J}_{\frac{\partial}{\partial t}}$$

is conserved.

Other conserved quantities will be discussed later.

REMARK A.16. There is a vast amount of work on symmetry principles and conservation laws following Noether's fundamental paper [19]. Three works in particular are noteworthy: the monographs by Logan [17] and Kosmann-Schwarzbach [10], and a review by Krupkova [12].

A.3.2. *The Cartan form approach.* Recall that the Cartan form Θ is

$$\begin{aligned} \Theta &= L\,dt + p_x\theta_1 + p_{\dot{x}}\theta_2 \\ &= L\,dt + \left(\frac{\partial L}{\partial \dot{x}} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{x}}\right)(dx - \dot{x}\,dt) + \frac{\partial L}{\partial \ddot{x}}(d\dot{x} - \ddot{x}\,dt). \end{aligned} \tag{91}$$

LEMMA A.17. *For every section* $\sigma : I \to I\times\mathbf{R}^n$, *and each vector field* $X = \tau\frac{\partial}{\partial t} + \xi\frac{\partial}{\partial x}$ *on* $I\times\mathbf{R}^n$,

$$j^3\sigma^*(\pounds_{X^3}\Theta) = j^2\sigma^*(\pounds_{X^2}(L\,dt)).$$

PROPOSITION A.18. *If* $X = \tau\frac{\partial}{\partial t} + \xi\frac{\partial}{\partial x}$ *is an infinitesimal symmetry of the Lagrange form* $L\,dt$ *then, for every section* σ *of* $I\times\mathbf{R}^n \to I$,

$$j^3\sigma^*\mathscr{J}_X = j^3\sigma^*(X^3 \lrcorner\, \Theta), \tag{92}$$

where X^3 *is the prolongation of* X *to* J^3. *If* σ *satisfies the Euler-Lagrange equations for* L, *then* $j^3\sigma^*(X^3 \lrcorner\, \Theta)$ *is constant.*

Equation (92) gives a simple way of finding constants of motion corresponding to symmetries of the Lagrange form.

Example A.19. If $x = (x^i)$ are Cartesian coordinates in $\mathbf{R}^n$, then the action of SO(n) on J^3 is generated by vector fields

$$X^3_{ij} = x^i \frac{\partial}{\partial x^j} - x^j \frac{\partial}{\partial x^i} + \dot{x}^i \frac{\partial}{\partial \dot{x}^j} - \dot{x}^j \frac{\partial}{\partial \dot{x}^i} + \ddot{x}^i \frac{\partial}{\partial \ddot{x}^j} - \ddot{x}^j \frac{\partial}{\partial \ddot{x}^i} + \dddot{x}^i \frac{\partial}{\partial \dddot{x}^j} - \dddot{x}^j \frac{\partial}{\partial \dddot{x}^i}.$$

Hence, for a section σ of $I \times \mathbf{R}^n \to I$,

$$\begin{aligned} j^3\sigma^*(\mathscr{J}_{X_{ij}}) &= j^3\sigma^*(X^3_{ij} \lrcorner\, \Theta) = \\ &= j^3\sigma^*(x^i p_{x^j} - x^j p_{x^i} + \dot{x}^i p_{\dot{x}^j} - \dot{x}^j p_{\dot{x}^i}). \end{aligned}$$

In the following we omit the symbol $j^3\sigma^*$, and write

$$\mathscr{J}_{X_{ij}} = x^i p_{x^j} - x^j p_{x^i} + \dot{x}^i p_{\dot{x}^j} - \dot{x}^j p_{\dot{x}^i}.$$

If L is invariant under the action of SO(3) on J^2, then $\mathscr{J}_{X_{ij}}$ is constant on solutions of the Euler-Lagrange equations.

There may be additional conserved quantities coming from symmetries of the Cartan form that are not symmetries of the Lagrange form.

Definition A.20. An infinitesimal symmetry of the Cartan form Θ is a vector field Z on J^3 such that $\pounds_Z \Theta = 0$.

Set

$$\mathscr{J}_Z = Z \lrcorner\, \Theta$$

for each infinitesimal symmetry Z of the Cartan form Θ.

Theorem A.21. *Let Z be an infinitesimal symmetry of the Cartan form and let $\gamma : I \to J^3$ be a section of the source map. If $T\gamma(I)$ is in* $\ker d\Theta$*, then*

$$\gamma^* \mathscr{J}_Z = \gamma^*(Z \lrcorner\, \Theta)$$

is constant. In particular, if σ satisfies the Euler-Lagrange equations, then $j^3\sigma\, \mathscr{J}_Z$ is a constant.

Proof. Since

$$\pounds_Z \Theta = Z \lrcorner\, d\Theta + d(Z \lrcorner\, \Theta),$$

it follows that for any infinitesimal symmetry Z of Θ, and any section $\gamma : I \to J^3$ of the source map, that

$$d\gamma^*(Z \lrcorner\, \Theta) = -\gamma^*(Z \lrcorner\, d\Theta).$$

If $T\gamma(I)$ is contained in $\ker d\Theta$, then $\gamma^*(Z \lrcorner\, d\Theta) = 0$ and $\gamma^*(Z \lrcorner\, \Theta)$ is constant. q.e.d.

A.3.3. *Symmetries up to a differential.* The Cartan form may yield more conserved quantities than those that follow directly from the Lagrangian approach. However, even more conserved quantities may arise if the notion of symmetry is relaxed somewhat.

Definition A.22. A vector field X on $I \times \mathbf{R}$ is a *symmetry up to a differential* of the Lagrange form $L\,dt$ if there exists a function F on J^2 such that

$$\pounds_{X^2}(L\,dt) = -dF,$$

where X^2 is the prolongation of X to J^2.

Proposition A.23. *If a vector field $X = \tau\frac{\partial}{\partial t} + \xi\frac{\partial}{\partial x}$ on $I \times \mathbf{R}^n$ satisfies the condition*

$$\pounds_{X^2}(L\,dt) = -dF, \tag{93}$$

where X^2 is the prolongation of X to J^2, and F is a function on J^2, then

$$\mathscr{J}_X + F = F + L\tau + \left(\frac{\partial L}{\partial \dot{x}} - \frac{d}{dt}\left(\frac{\partial L}{\partial \ddot{x}}\right)\right)(\xi - \tau\dot{x}) + \frac{\partial L}{\partial \ddot{x}}\frac{d}{dt}(\xi - \tau\dot{x}) \tag{94}$$

is constant along solutions of the Euler-Lagrange equations

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) + \frac{d^2}{dt^2}\left(\frac{\partial L}{\partial \ddot{x}}\right) = 0.$$

Example A.24. Probably the most well-known example of this sort of behaviour occurs in the first-order theory as the Runge-Lenz vector in the Kepler problem. In this case there is a symmetry of the dynamical system that is not lifted from the configuration space, and this implies that the Lagrangian is not invariant under the action of the symmetry group, but changes by a total derivative. This is discussed in [16].

In a similar way, infinitesimal symmetries up to a differential of the Cartan form are defined as

Definition A.25. A vector field Z on J^3 is a *symmetry up to a differential* of the Cartan form Θ if there exists a function F on J^3 such that

$$\pounds_Z\Theta = -dF. \tag{95}$$

Proposition A.26. *If $\pounds_Z\Theta = -dF$, then for a section $\gamma : I \to J^3$ such that $T(\gamma(I))$ is contained in* $\ker d\Theta$*, the function $F + \langle\Theta, Z\rangle$ is constant along γ. In particular, $F + \langle\Theta, Z\rangle$ is constant along the jet extensions of sections σ of $I \times \mathbf{R}^n$ that satisfy the Euler-Lagrange equations.*

Remark A.27. It is an interesting exercise to compute the effect of the dilation group in the elastica problem to see to what extent it is a symmetry in one of these extended senses, as it is clearly not a symmetry of the Lagrange form.

A.4. **Parametrization invariance.** Let $\mathrm{Diff}_+\,\mathbf{R}$ be the group of orientation preserving diffeomorphisms of the real line $\mathbf{R}$. Then, for every $\varphi \in \mathrm{Diff}_+\,\mathbf{R}$, $\dot{\varphi}(t) > 0$ for all t. Each $\varphi \in \mathrm{Diff}_+\,\mathbf{R}$ gives rise to another diffeomorphism

$$\varphi^0 : \mathbf{R}\times\mathbf{R}^n \to \mathbf{R}\times\mathbf{R}^n : (t,x) \mapsto \varphi^0(t,x) = (\varphi(t), x).$$

The prolongations of φ^0 to jet bundles can be written as follows

$$\begin{aligned}
\varphi^1 \quad &: \quad J^1 \to J^1 : (t,x,\dot{x}) \mapsto \varphi^1(t,x,\dot{x}) = \left(\varphi(t), x, \frac{\dot{x}}{\dot{\varphi}(t)}\right), \qquad (96)\\
\varphi^2 \quad &: \quad J^2 \to J^2 : (t,x,\dot{x},\ddot{x}) \mapsto \varphi^2(t,x,\dot{x},\ddot{x}) = \left(\varphi(t), x, \frac{\dot{x}}{\dot{\varphi}(t)}, \frac{\ddot{x}}{\dot{\varphi}(t)^2} - \dot{x}\frac{\ddot{\varphi}}{\dot{\varphi}(t)^3}\right),\\
\varphi^3 \quad &: \quad J^3 \to J^3 : (t,x,\dot{x},\ddot{x},\dddot{x}) \mapsto \varphi^3(t,x,\dot{x},\ddot{x}) =\\
&= \left(\varphi(t), x, \frac{\dot{x}}{\dot{\varphi}(t)}, \frac{\ddot{x}}{\dot{\varphi}(t)^2} - \dot{x}\frac{\ddot{\varphi}}{\dot{\varphi}(t)^3}, \frac{\dddot{x}}{\dot{\varphi}(t)^3} - 3\ddot{x}\frac{\ddot{\varphi}(t)}{\dot{\varphi}(t)^4} - \dot{x}\frac{\dddot{\varphi}(t)}{\dot{\varphi}(t)^4} + 3\dot{x}\frac{\ddot{\varphi}(t)^2}{\dot{\varphi}(t)^5}\right).
\end{aligned}$$

Proposition A.28. *For $\varphi \in \mathrm{Diff}_+\,\mathbf{R}$,*

$$\begin{aligned}
\varphi^{1*}\theta_1 &= \theta_1,\\
\varphi^{2*}\theta_2 &= \frac{1}{\dot{\varphi}}\,\theta_2.
\end{aligned}$$

A one-parameter subgroup $\varphi_\varepsilon : t \mapsto \bar{t} = \varphi_\varepsilon(t)$ of $\mathrm{Diff}_+\,\mathbf{R}$ is generated by a vector field $X_\tau = \tau(t)\partial_t$, where

$$\tau(t) = \left.\frac{\partial\varphi_\varepsilon(t)}{\partial\varepsilon}\right|_{\varepsilon=0}$$

is an arbitrary smooth function. The Lie algebra of the group $\mathrm{Diff}_+\,\mathbf{R}$ is the collection of vector fields

$$\mathrm{diff}_+\,\mathbf{R} = \{X_\tau = \tau(t)\partial_t \mid \tau \in C^\infty(\mathbf{R}), \text{ and } \dot{\tau}(t) \neq 0 \text{ for all } t\}$$

with the Lie bracket

$$[\tau_1(t)\partial_t, \tau_2(t)\partial_t] = (\tau_1(t)\dot{\tau}_2(t) - \tau_2(t)\dot{\tau}_1(t))\partial_t.$$

Definition A.29. The variational problem with the Lagrangian L is *parametrization invariant* if the Lagrange form $L\,dt$ is invariant under the action of $\mathrm{Diff}_+\,\mathbf{R}$ on J^2.

Remark A.30. Suppose that the Lagrange form $L\,dt$ is $\mathrm{Diff}_+\,\mathbf{R}$-invariant. This implies that for $X_\tau = \tau(t)\partial_t$, the Lagrange form $L\,dt$ is invariant under the one-parameter subgroup of $\mathrm{Diff}_+\,\mathbf{R}$ generated by X_τ. By Theorem A.18, $\mathscr{J}_{X_\tau} = \langle\Theta, X^3_\tau\rangle$ is constant on solutions of the Euler-Lagrange equations.

Theorem A.31. *If the Lagrange form $L\,dt$ is* $\mathrm{Diff}_+\,\mathbf{R}$*-invariant, then*

$$j^3\sigma^*\mathscr{J}_{X_\tau} = 0 \qquad (97)$$

for all $X_\tau \in \mathrm{diff}_+\,\mathbf{R}$ and all solutions σ of the Euler-Lagrange equations.

Proof. Recall that $\Theta = p_x\, dx + p_{\dot{x}}\, d\dot{x} - H\, dt$. Omitting pull-backs by $j^3\sigma$ for the sake of cleanliness,

$$\begin{aligned} \mathscr{J}_{X_\tau} = \langle \Theta, X^3_\tau \rangle &= \left\langle pdx + p_{\dot{x}} d\dot{x} - Hdt, \tau\frac{\partial}{\partial t} - \dot{x}\dot{\tau}\frac{\partial}{\partial \dot{x}} \right\rangle \\ &= -\langle p_{\dot{x}}, \dot{x}\rangle\, \dot{\tau} - H\tau. \end{aligned} \tag{98}$$

If σ satisfies the Euler-Lagrange equations, then $j^3\sigma^* \mathscr{J}_{X_\tau}$ is constant.

Take two points, $t_0 < t_1$ in I, and consider two other vector fields X_{τ_1} and X_{τ_2} in $\mathrm{diff}_+ \mathbf{R}$ such that

$$\begin{aligned} \tau(t_0) &= \tau_1(t_0) = \tau_2(t_0) \text{ and } \dot{\tau}(t_0) = \dot{\tau}_1(t_0) = \dot{\tau}_2(t_0), \\ \tau(t_1) &\neq \tau_1(t_1) = \tau_2(t_1) \text{ and } \dot{\tau}(t_1) = \dot{\tau}_1(t_1) \neq \dot{\tau}_2(t_1), \\ \tau(t_2) &\neq \tau_1(t_2) = \tau_2(t_2) \text{ and } \dot{\tau}(t_2) = \dot{\tau}_1(t_2) \neq \dot{\tau}_2(t_2). \end{aligned}$$

Then, $\mathscr{J}_{X_\tau}(t)$, $\mathscr{J}_{X_{\tau_1}}(t)$ and $\mathscr{J}_{X_{\tau_2}}(t)$ are constant along $j^3\sigma$. Moreover, the assumption that $\tau(t_0) = \tau_1(t_0) = \tau_2(t_0)$, $\dot{\tau}(t_0) = \dot{\tau}_1(t_0) = \dot{\tau}_2(t_0)$ and equation (98) imply that $\mathscr{J}_{X_\tau}(t) = \mathscr{J}_{X_{\tau_1}}(t) = \mathscr{J}_{X_{\tau_2}}(t)$ for all t. Therefore, $\mathscr{J}_{X_{\tau_1}}(t) - \mathscr{J}_{X_\tau}(t) = 0$ and $\mathscr{J}_{X_{\tau_2}}(t) - \mathscr{J}_{X_\tau}(t) = 0$ for all t. Using equation (98) and setting $t = t_1$ yields

$$\begin{aligned} p_x(t_1)\dot{x}(t_1)\dot{\tau}_1(t_1) + H(t_1)\tau_1(t_1) - p_x(t_1)\dot{x}(t_1)\dot{\tau}(t_1) - H(t_1)\tau(t_1) &= 0 \\ p_x(t_1)\dot{x}(t_1)\dot{\tau}_2(t_1) + H(t_1)\tau_2(t_1) - p_x(t_1)\dot{x}(t_1)\dot{\tau}_1(t_1) - H(t_1)\tau_1(t_1) &= 0. \end{aligned}$$

Since $\tau(t_1) \neq \tau_1(t_1)$ and $\dot{\tau}(t_1) = \dot{\tau}_1(t_1)$, the first equation above yields $H(t_1)(\tau(t_1) - \tau_1(t_1)) = 0$, which implies that $H(t_1) = 0$. Similarly, the assumption that $\tau_1(t_1) = \tau_2(t_1)$ and $\dot{\tau}_1(t_1) \neq \dot{\tau}_2(t_1)$ together with the second equation above yield $p_x(t_1)\dot{x}(t_1) = 0$. Since t_1 is an arbitrary point in I different from t_0, it follows that

$$H(t) = 0 \;\text{ and }\; p_x(t)\dot{x}(t) = 0 \text{ for all } t \in I. \tag{99}$$

Substituting this result into equation 98 gives (97). q.e.d.

Remark A.32. Equations (99), rewritten in terms of the configuration variables read

$$\begin{aligned} \dot{x}\frac{\partial L}{\partial \ddot{x}} &= 0, \\ \dot{x}\left(\frac{\partial L}{\partial \dot{x}} - \frac{d}{dt}p_{\dot{x}}\right) + \ddot{x}\frac{\partial L}{\partial \ddot{x}} - L &= 0. \end{aligned} \tag{100}$$

These are the identities for our reparametrization invariant Lagrangian that follow from the second Noether theorem ([19]). The proof of Theorem A.31 establishes the equivalence between the Noether identities (100) and the vanishing of the constant of motion $\mathscr{J}_{X_\tau}$ corresponding to every $X_\tau \in \mathrm{diff}_+ \mathbf{R}$.

A.5. **Arclength parametrization.** Denote by $\langle x, x'\rangle$ the Euclidean scalar product and by $|x| = \sqrt{\langle x, x\rangle}$, the corresponding norm in $\mathbf{R}^n$. For a curve $c : I \to \mathbf{R}^n : t \mapsto x(t)$, where $I = [t_0, t_1]$, the arclength of the section of c from t_0 to t is

$$s(t) = \int_{t_0}^{t} \left|\dot{x}(t')\right| dt'. \tag{101}$$

In geometric problems it is often convenient to parametrize a curve in terms of its arclength. If t is the arclength of c, then along c

$$\langle \dot{x}, \dot{x} \rangle = |\dot{x}|^2 = 1, \tag{102}$$

$$\langle \dot{x}, \ddot{x} \rangle = 0, \tag{103}$$

$$\langle \dot{x}, \dddot{x} \rangle + \langle \ddot{x}, \ddot{x} \rangle = 0, \tag{104}$$

$$\langle \dot{x}, \ddddot{x} \rangle + 3 \langle \ddot{x}, \dddot{x} \rangle = 0. \tag{105}$$

These equations determine submanifolds M^1, M^2, M^3 and M^4 of J^1, J^2, J^3 and J^4, respectively.

Proposition A.33. *Let $X = \tau \frac{\partial}{\partial t}$ be a vector field on the configuration space Q. The necessary and sufficient condition for its prolongations X^1, X^2, X^3 and X^4 to be tangent to M^1, M^2, M^3 and M^4, respectively, is that the restriction of τ to M^1, M^2, M^3 and M^4, respectively, is constant.*

Suppose a local section σ of Q with domain $I \subset \mathbf{R}$ and with $j^1\sigma(I)$ not in M^1, satisfies $\dot{x}(t) \neq 0$ for all $t \in I$.

Lemma A.34. *There exists $\varphi \in \mathrm{Diff}_+ \mathbf{R}$ such that*

$$\frac{d\varphi}{dt} = |\dot{x}(t)|$$

for all $t \in I$.

Then,

$$\frac{dx}{d\varphi} = \frac{dx}{dt}\frac{dt}{d\varphi} = \frac{dx}{dt}\frac{1}{|\dot{x}(t)|} = \frac{\dot{x}(t)}{|\dot{x}(t)|},$$

and it follows that

$$\left| \frac{dx}{d\varphi} \right| = 1.$$

Thus the new parametrization gives rise to a section $\varphi^*\sigma$ with its first jet in M^1. Similarly, the k-jet of $\varphi^*\sigma$ is in M^k.

A.6. **Hamiltonian formulation.** The Liouville form on the cotangent bundle T^*J^1 with variables $(t, x, \dot{x}, p_t, p_x, p_{\dot{x}})$ is

$$\theta = p_t\, dt + p_x\, dx + p_{\dot{x}}\, d\dot{x}. \tag{106}$$

The exterior derivative

$$\omega = d\theta \tag{107}$$

is the canonical symplectic form of T^*J^1.

Lemma A.35. *The action*

$$\mathrm{Diff}_+ \mathbf{R} \times J^1 \to J^1 : (\varphi, (t, x, \dot{x})) \mapsto \left(\varphi(t), x, \frac{\dot{x}}{\dot{\varphi}(t)} \right)$$

lifts to an action

$$\mathrm{Diff}_+ \mathbf{R} \times T^*J^1 \to T^*J^1 : \tag{108}$$

$$(\varphi, (t, x, \dot{x}, p_t, p_x, p_{\dot{x}})) \mapsto \left(\varphi(t), x, \dot{\varphi}(t)^{-1}\dot{x}, p_t\dot{\varphi}(t)^{-1} + \langle p_{\dot{x}}, \dot{x} \rangle\, \dot{\varphi}(t)^{-2}\ddot{\varphi}(t), p_x, \dot{\varphi}(t)p_{\dot{x}} \right).$$

The lifted action (108) is Hamiltonian with momentum map $\mathscr{J} : T^*J^1 \to \mathrm{diff}_+ \mathbf{R}^*$ *such that, for* $X = \tau(t)\partial_t \in \mathrm{diff}_+ \mathbf{R}$*,*

$$\mathscr{J}_X(t, x, \dot{x}, p_t, p_x, p_{\dot{x}}) = \tau(t)p_t - \dot{\tau}(t) \langle p_{\dot{x}}, \dot{x}\rangle . \tag{109}$$

Definition A.36. The Legendre-Ostrogradski transformation

$$\mathscr{L} : J^3 \to T^*J^1 : (t, x, \dot{x}, \ddot{x}, \dddot{x}) \mapsto (t, x, \dot{x}, p_t, p_x, p_{\dot{x}}), \tag{110}$$

is given by

$$\begin{aligned} p_t &= -H = -\dot{x}p_x - \ddot{x}p_{\dot{x}} + L \\ p_{\dot{x}} &= \frac{\partial L}{\partial \ddot{x}}, \\ p_x &= \frac{\partial L}{\partial \dot{x}} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{x}} = \frac{\partial L}{\partial \dot{x}} - \frac{d}{dt}p_{\dot{x}}. \end{aligned}$$

The Legendre transformation was extended by Ostragradski in [20]. However, for brevity, from now on we shall just refer to it as the Legendre transformation.

Clearly, $\mathscr{L}$ is a smooth map of J^3 into T^*J^1. If $\mathscr{L}$ is a diffeomorphism, Ostrogradski's approach leads to a regular time-dependent Hamiltonian theory with the Hamiltonian

$$H = -p_t = \dot{x}p_x + \ddot{x}p_{\dot{x}} - L.$$

In geometric problems, the Lagrangian is often reparametrization invariant. In this case L does not depend on t and the range of the Legendre transformation is restricted by the equations (see (A.31))

$$H = 0 \text{ and } \langle p_{\dot{x}}, \dot{x}\rangle = 0.$$

Theorem A.37. *If the Lagrange form* $L\,dt$ *is invariant under the action of* $\mathrm{Diff}_+ \mathbf{R}$*, then the Legendre transformation* $\mathscr{L}$ *intertwines the actions of* $\mathrm{Diff}_+ \mathbf{R}$ *on* J^3 *and on* T^*J^1*.*

Corollary A.38. *If the Lagrange form* $L\,dt$ *is* $\mathrm{Diff}_+ \mathbf{R}$*-invariant, then the Cartan form* Θ *is* $\mathrm{Diff}_+ \mathbf{R}$*-invariant.*

Proof. Since $L\,dt$ is $\mathrm{Diff}_+ \mathbf{R}$-invariant, theorem (A.37) implies that for $\varphi \in \mathrm{Diff}_+ \mathbf{R}$, $\varphi^{3*}\mathscr{L}^* = \mathscr{L}^*\tilde{\varphi}^{1*}$, where $\tilde{\varphi}^1$ is the lift of φ^1 to the cotangent bundle T^*J^1. But, $\Theta = \mathscr{L}^*\theta$, and the Liouville form θ is invariant under the the lifted action $\tilde{\varphi}^{1.}$ Therefore,

$$\varphi^{3*}\Theta = \varphi^{3*}\mathscr{L}^*\theta = \mathscr{L}^*\tilde{\varphi}^{1*}\theta = \mathscr{L}^*\theta = \Theta.$$

q.e.d.

Corollary A.39. *The range of the Legendre transformation* $\mathscr{L} : J^3 \to T^*J^1$ *is contained in the zero set of the momentum map* $\mathscr{J} : T^*J^1 \to \mathrm{diff}_+ \mathbf{R}^*$*. That is,*

$$\mathscr{L}(J^3) \subseteq \mathscr{J}^{-1}(0).$$

References

[1] K. Bleuler. Eine neue Methode sur Behandlung der longitudinalen und skaleren Photonen. *Helv. Phys. Acta*, 23:567, 1950.

[2] M. Born. *Untersuchungen über die Stäbilitat der elastichen Linie in Ebene und Raum, under verschiedenen Grenzbedingungen*. PhD thesis, University of Göttingen, 1906.

[3] R. Bryant and P. Griffiths. Reduction for constrained variational problems and $\int(k^2/2)\,ds$. *American journal of mathematics*, 108:525–570, 1986.

[4] P. Coronado. Hamilton equations for elasticae in the Euclidean 3-space. *Physica D*, 141(3-4):248-260, 2000.

[5] P. Dirac. Hamiltonian methods and quantum mechanics. *Proceedings of the Royal Irish Academy*, 63(3):49–59, 1963.

[6] L. Euler. *Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes, sive solutio problematis isoperimetrici lattissimo sensu accepti*, volume E065. chapter Additamentum 1. eulerarchive.org, 1744.

[7] K. Foltinek. The Hamiltonian theory of elastica. *American Journal of Mathematics*, 116:1479–1488, 1994.

[8] H. Grässer. *A monograph on the general theory of second order parameter-invariant problems in the calculus of variations*. Number M5 in Mathematical communications of the University of South Africa. Pretoria, 1967.

[9] S. Gupta. Theory of longitudinal photons in quantum electrodynamics. *Proc. Phys. Soc.*, A63:681, 1950.

[10] Y. Kosmann-Schwarzbach. *The Noether theorems*. Springer, New York, 2011.

[11] D. Krupka, O. Krupkova, and D. Saunders. Cartan - Lepage forms in geometric mechanics. *International Journal of Non-Linear Mechanics*, 47:1154–1160, 2012.

[12] Olga Krupkova. Noether theorem, 90 years on. *AIP Conf.Proc.*, 1130:159–170, 2009.

[13] J. Langer and D. Singer. Knotted elastic curves in $\mathbf{R}^3$. *Journal of the London Mathematical Society*, 30(2):512–520, 1984.

[14] R. Levien. The elastica: a mathematical history. Technical Report UCB/EECS-2008-103, EECS Department, University of California, Berkeley, Aug 2008.

[15] R. Levien. *From Spiral to Spline: Optimal Techniques in Interactive Curve Design*. PhD thesis, University of California, Berkeley, 2009.

[16] J.M. Lévy-Leblond. Conservation laws for gauge-variant Lagrangians in classical mechanics. *American journal of physics*, 39(5):502–506, 1971.

[17] J. Logan. *Invariant variational principles*. Academic Press, 1977.

[18] V. Nesterenko and G. Scarpetta. Singular Lagrangians with higher derivatives, in *Geometry of constrained dynamical systems*, edited by J. Charap, Cambridge University Press, 1995.

[19] E. Noether. Invariante variationsprobleme. *Nachtrichten von der Gesellschaft der Wissenshaften zu Göttingen Mathematisch-physickalische Klasse*, pages 235–258, 1918.

[20] M. Ostrogradski. Memoires sur les equations differentielles relatives au probleme des isoperimetres. *Mem. Ac. St. Petersbourg*, VI:385, 1850.

[21] J. Śniatycki. *Geometric quantization and quantum mechanics*. Number 30 in Applied mathematical sciences. Springer-Verlag, 1980.

[22] W. Wunderlich. über ein abwickelbares möbiusband. *Monatshefte für Mathematik*, 66:276–289, 1962.

Larry M. Bates
Department of Mathematics
University of Calgary
Calgary, Alberta

Canada T2N 1N4
bates@ucalgary.ca

Robin Chhabra
Guidance, Navigation and Control Department
MacDonald, Dettwiler and Associates Ltd.
Brampton, Ontario
Canada L6S 4J3
robin.chhabra@mdacorporation.com

Jędrzej Śniatycki
Department of Mathematics
University of Calgary
Calgary, Alberta
Canada T2N 1N4
sniatyck@ucalgary.ca