

CONTROL PROBLEMS WITH DIFFERENTIAL CONSTRAINTS OF HIGHER ORDER

FRANCO CARDIN CRISTINA GIANNOTTI ANDREA SPIRO

ABSTRACT. We consider cost minimising control problems, in which the dynamical system is constrained by higher order differential equations of Euler-Lagrange type. Following ideas from a previous paper, we prove that a curve of controls $u_o(t)$ and a set of initial conditions σ_o gives an optimal solution for a control problem of the considered type if and only if an appropriate double integral is greater than or equal to zero along any homotopy $(u(t, s), \sigma(s))$ of control curves and initial data starting from $u_o(t) = u(t, 0)$ and $\sigma_o = \sigma(0)$. This property is called *Principle of Minimal Labour*. From this principle we derive a generalisation of the classical Pontryagin Maximum Principle that holds under higher order differential constraints of Euler-Lagrange type and without the hypothesis of fixed initial data.

1. INTRODUCTION

In [3] the first and the third author presented a new proof of the classical Pontryagin Maximum Principle (PMP) for controlled systems, which was crucially based on the observation that the first order differential constraints of the system can be considered as the Euler-Lagrange equations determined by an appropriate controlled first order Lagrangian. Following the same ideas of that proof, we give here a generalisation of the PMP to the control problems, in which the differential constraints are given by Euler-Lagrange equations of higher order.

More precisely, we consider cost minimising problems for dynamical system which are controlled through the Euler-Lagrange equations determined by higher order Lagrangians with controlling parameters, i.e. by Lagrangians depending on appropriate controls $u^a(t)$

$$L^{u(\cdot)} \left(t, q^i, \frac{dq^i}{dt}, \dots, \frac{d^k q^i}{dt^k} \right) := L \left(t, q^i, \frac{dq^i}{dt}, \dots, \frac{d^k q^i}{dt^k}, u^a(t) \right) .$$

We do not impose any particular assumption on the initial conditions of the solutions of the differential constraints, nor on the control curves $u(t) = (u^a(t))$, besides merely technical requirements of smoothness and constant rank conditions on the constraints. In fact, in this paper we restrict our discussion just to control problems satisfying strong regularity assumptions: this choice is only for the purpose of making as much as possible

2010 *Mathematics Subject Classification.* 34H05, 49K15.

Key words and phrases. Pontryagin Maximum Principle; Mayer Problem; Higher Order Differential Constraint; Geometric Optimal Control; Poincaré-Cartan form.

Acknowledgments. This research was partially supported by the Projects MIUR “Regular and stochastic behaviour in dynamical systems” and “Real and Complex Manifolds: Geometry, Topology and Harmonic Analysis” and by GNFM and GNSAGA of INdAM.

transparent the main ideas of our approach. An illustration on how the main results of this paper can be generalised under much weaker regularity assumptions is given in [4].

Considering the proof of the PMP presented in [3] as a model, we are able to prove that a curve of controls $u_o(t)$ and a set of initial conditions σ_o for the evolution of the controlled dynamical system corresponds to an optimal solution if and only if a particular double integral is greater than or equal to zero for any homotopy $(u(t, s), \sigma(s))$ of control curves and initial conditions, having $u_o(t) = u(t, 0)$ and $\sigma_o = \sigma(0)$ as starting point. We called this property *Principle of Minimal Labour*. Using this and an appropriate formalisation of Pontryagin's notion of needle variation of control curves, we derive a generalisation of the classical PMP for the control problems that are subjected to higher order Euler-Lagrange constraints *of normal type*. This is a very large class of constraints which naturally includes the classical Mayer problems to which the usual version of the PMP applies. Actually our main result provides additional information also for the classical first order differential constraints since it allows variations of the initial data and establishes an innovative shortcut between the proofs of the classical PMP and the Noether Theorem about conservation laws (see for instance the approach to Noether Theorem based on Stokes Theorem given in [6], and compare it with the use of Stokes Theorem in the proof of the PMP given here and in [3], as illustrated in Sect. 2 below). A discussion in greater detail of our main results and some simple illustrations of how they can be exploited are given in Sect. 2 and Sect. 8.

Before concluding this introduction, we would like to recall that, considering an appropriate set of auxiliary variables, any control problem with higher order differential constraints – even those of variational type for which we establish our generalised PMP – can be reduced to an equivalent one with only first order constraints (but, in general, no longer of variational type). By introducing other auxiliary variables, the Pontryagin variables p_j , the original higher order problem is finally transformed into an equivalent one, which is now variational and to which the classical PMP applies. This kind of reduction procedure demands the introduction of a (in general very large) collection of auxiliary variables. In contrast with this, our generalised PMP directly applies to any set of higher order constraints, with almost no need of auxiliary variables. This is a valuable feature, which we briefly illustrate with an elementary example in Sect. 8. More substantial examples and applications will be discussed in detail in a future work.

This paper is structured as follows. In Sect. 2, we review the main results of [3] and give an informal presentation of the main ideas on which our results are based. A detailed description of the Principle of Minimal Labour and of our generalisation of the PMP are also given. In Sect. 3 and Sect. 4 the needed preliminaries and a rigorous definition of the class of control problem we are considering are given. In Sect. 5, the considered control problems are transformed into an equivalent form, which is more appropriate for the subsequent manipulations. The proofs of the Principle of Minimal Labour and of our generalisation of the PMP are given in Sect. 6 and Sect. 7, respectively. In Sect. 8, we illustrate some of the main features of our approach by discussing a couple of elementary problems.

2. AN OVERVIEW OF OUR APPROACH AND MAIN RESULTS

Since the results of the present paper can be considered as natural developments of the ideas of [3], we decided to precede our discussion with a short overview of the contents of that paper. We then briefly indicate how the scheme of such a previous paper is here implemented to obtain our new results.

2.1. Pontryagin Maximum Principle and Stokes Theorem in a nutshell.

2.1.1. *The basic scheme of a classical Mayer problem.* Consider a dynamical system, whose evolutions on a fixed time interval $[0, T]$ are represented by curves $x(t) = (x^1(t), \dots, x^N(t))$ in \mathbb{R}^N satisfying the differential problem

$$\frac{dx^i}{dt}(t) = f^i(t, x^i(t), u^a(t)) , \quad x^i(0) = x_o^i \in \mathbb{R}^N . \quad (2.1)$$

Here $u(t)$ is a (measurable) function with values in a fixed subset K of \mathbb{R}^M

$$u(t) = (u^1(t), \dots, u^M(t)) \in K \subset \mathbb{R}^M$$

and represents the evolution in time of control parameters. The $f^i(t, x^i, u^a)$ are continuous functions on $[0, T] \times \mathbb{R}^N \times K$ and differentiable in the x . The initial value x_o is fixed and is the same for all of the evolutions of the system. Given a terminal cost function $C : \mathbb{R}^N \rightarrow \mathbb{R}$, the corresponding *Mayer problem* consists of looking for a curve $\bar{u}(t)$ of control parameters, for which the following holds: *the terminal cost $C(\bar{x}(T))$ of the solution $\bar{x}(t)$ to (2.1) with $u(t) = \bar{u}(t)$ is less than or equal to the terminal cost of the solution determined by any other choice of the curve $u(t)$.*

Of course, this is only one of the many variants of the classical Mayer problem. But in what follows we limit ourselves to such a basic version. Moreover, we constantly assume that all the data satisfy much higher regularity assumptions than those mentioned above. Take this restriction as a sort of blanket assumption, which is adopted to easily allow manipulations and prevent the risk of diverting the attention of the reader towards inessential technical issues.

2.1.2. *The auxiliary variables p_i .* The classical approach to a Mayer problem is usually based on the introduction of a set of auxiliary variables p_1, \dots, p_N and on the representation of the dynamical system through curves $(x(t), p(t)) = (x^i(t), p_j(t))$ in \mathbb{R}^{2N} (of which, however, only the $x^i(t)$ are the interesting ones), constrained by the (2.1) and the auxiliary equations

$$\frac{dp_j}{dt} := - \sum_{i=1}^N p_i(t) \frac{\partial f^i}{\partial x^j} \Big|_{(t, x^i(t), u^a(t))} . \quad (2.2)$$

The introduction of the auxiliary variables p_j and of the new constraints (2.2) has the following effect: *for each curve of control parameters $u(t) = (u^a(t))$, the constraints on the corresponding curve $(x(t), p(t))$ are the Euler-Lagrange equations of the variational principle determined by the (controlled) Lagrangian*

$$L^{u(\cdot)}(t, p, x, \dot{x}) = \sum_{j=1}^N p_j(\dot{x}^j - f^j(t, x, u(t))) . \quad (2.3)$$

This Lagrangian has also the special feature of being identically vanishing along the solutions of the corresponding Euler-Lagrange equations (in fact the $x^i(t)$ satisfy the (2.1)).

All this has the consequence that the original cost minimising problem is equivalent to the following one. By subtracting a constant, with no loss of generality we may assume that $C(x_o) = 0$. Hence by the vanishing of $L^{u(t)}$ along the constrained curves, the value on such curves of the functional

$$I^{u(\cdot)} := \int_0^T \left(L^{u(\cdot)}(t, p, x, \dot{x}) + \sum_{i=1}^N \frac{\partial C}{\partial x^i} \Big|_{(t, x(t))} \dot{x}^i(t) \right) dt \quad (2.4)$$

appears to be equal to

$$I^{u(\cdot)} = \int_0^T \frac{dC(x(t))}{dt} \Big|_{(t, x(t))} dt = C(x(T)) - C(x(0)) \stackrel{C(x_o)=0}{=} C(x(T)). \quad (2.5)$$

Thus the original problem turns out to be the same of looking for a curve of controls $u(t) = (u^a(t))$ such that, along the solution of the Euler-Lagrange equations of $L^{u(\cdot)}$, the functional $I^{u(\cdot)}$ takes the minimum possible value.

We also observe that, along the solutions $(x(t), p(t))$, the value of the above action is independent of any boundary (initial or final) conditions for the $p_j(t)$. This means that any value at $t = 0$ (or at $t = T$) can be imposed on the $p_j(t)$, having absolutely no consequences on the minimising problem. As we will see in the next two subsections, among all of the possible choices for such boundary conditions, some are much more convenient than the others.

2.1.3. *A smart choice for the boundary values of the p_i .* Consider a homotopy of control curves $u^{(s)}(\cdot) : [0, T] \rightarrow K$, $s \in [0, 1]$, and denote by $(x^{(s)i}(t), p_j^{(s)}(t))$ the corresponding homotopy of curves satisfying the constraints (2.1) and (2.2). Exploiting our blanket assumptions on the regularity of the data, we may say that the first order jets $(x^{(s)i}(t), \dot{x}^{(s)i}(t), p_j^{(s)}(t), \dot{p}_j^{(s)}(t))$, determined by the pairs $(t, s) \in [0, T] \times [0, 1]$, span a smooth surface \mathcal{S} in the first order jet space of curves in \mathbb{R}^{2N} , whose boundary is formed by four smooth curves. We denote them by $\gamma^{(s=0)}(t)$, $\gamma^{(s=1)}(t)$, $\eta^{(0)}(s)$ and $\eta^{(T)}(s)$.

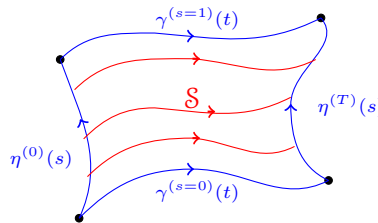


Fig. 1

The first two curves correspond to the sides $[0, T] \times \{0\}$, $[0, T] \times \{1\}$ of $\partial([0, T] \times [0, 1])$:

$$\gamma^{(s=0)}(t) = (x^{(0)i}(t), \dot{x}^{(0)i}(t), p_j^{(0)}(t), \dot{p}_j^{(0)}(t)) ,$$

$$\gamma^{(s=1)}(t) = (x^{(1)i}(t), \dot{x}^{(1)i}(t), p_j^{(1)}(t), \dot{p}_j^{(1)}(t)) .$$

The others are the curves that correspond to the remaining two sides $\{0\} \times [0, 1]$, $\{T\} \times [0, 1]$ of the boundary of $[0, T] \times [0, 1]$. We call them the *vertical sides* of \mathcal{S} . In [3] it was

observed that *if one imposes that the $p_j(t)$ satisfy the terminal values conditions*

$$p_j(T) = - \frac{\partial C}{\partial x^j} \Big|_{(x(T))} , \quad (2.6)$$

then the integrals of the 1-form $\beta = \left(L^{u(\cdot)}(t, p, x, \dot{x}) + \sum_{i=1}^N \frac{\partial C}{\partial x^i} \Big|_{(t, x(t))} \dot{x}^i(t) \right) dt$ along the two vertical sides $\eta^{(0)}(s)$, $\eta^{(T)}(s)$ of \mathcal{S} (or, more precisely, of an appropriate modification of β , modelled on the classical Poincaré-Cartan 1-form) are equal to 0.

This is an important property, because in combination with (2.5) it implies that the integral of (the Poincaré-Cartan type modification of) β along the anti-clockwise oriented boundary $\partial\mathcal{S}$ is equal to $-C(x^{(1)}(T)) + C(x^{(0)}(T))$. On the other hand, by Stokes Theorem, such an integral is equal to the integral of the exterior differential of the Poincaré-Cartan type modification of β on \mathcal{S} . By exploiting certain properties of the actions and of the 1-forms of Poincaré-Cartan type (it is not essential to recall them here - in Sect. 5.2 we discuss them in greater detail), such an integral reduces to a very simple form, namely to

$$C(x^{(1)}(T)) - C(x^{(0)}(T)) = - \iint_{t \in [0, T], s \in [0, 1]} \frac{\partial \mathcal{H}}{\partial u^a} \frac{\partial u^a}{\partial s} \Big|_{(t, x^{(s)i}(t), p_j^{(s)}(t), u^{(s)a}(t))} dt ds , \quad (2.7)$$

where $\mathcal{H}(t, x^i, p_j, u^a) := - \sum_{j=1}^N p_j f^j(t, x^i, u^a)$.

The main reason of interest for this identity comes from the fact that *it expresses the difference between the two terminal costs as a double integral of an appropriate function of the parameters (t, s) of the homotopy.* From our point of view, this is a cornerstone in the proof of the PMP.

2.1.4. *The Principle of Minimal Labour and the Pontryagin Maximum Principle.* An immediate consequence of (2.7) is the following:

Principle of Minimal Labour. *If the curve $\bar{u}(t) \in K$ gives a solution to the considered Mayer problem, then for any other curve $u(t)$ which is connected to $\bar{u}(t)$ through an homotopy of curves $u^{(s)}(t)$ in K , the double integral on the left hand side of (2.7) is less than or equal to 0.*

By considering appropriate highly localised deformations of the curve $\bar{u}(t)$ (the so-called *Pontryagin's needle variations* – see [3] or Sect. 7.2 below for details) and associated interpolating homotopies, the classical PMP can be derived as if a ‘pointwise version’ of the above principle. Indeed, adopting a very informal language, we may state the **Pontryagin Maximum Principle** as follows (the literature on such a classical principle is vast – for extensive and fundamental presentations we refer to [8, 11, 7, 1, 2, 5] and references therein):

If $\bar{u}(t) \in K$ gives a solution $(x^i(t))$ to the considered Mayer problem, then for any $t_o \in [0, T]$ the value of \mathcal{H} at the point $(t_o, x^i(t_o), p_j(t_o), \bar{u}^a(t_o))$ is maximal among all the values that it assumes at the points $(t_o, x^{(\omega)i}(t_o), p_j^{(\omega)}(t_o), \omega^a)$ determined by

(i) *replacing $(\bar{u}^a(t_o))$ by some other value $(\omega^a) \in K$,*

- (ii) substituting the values $x^i(t_o)$ and $p_j(t_o)$ by the values $x^{(\omega)i}(t_o)$ and $p_j^{(\omega)}(t_o)$, which are assumed by the solution $(x^{(\omega)i}(t), p_j^{(\omega)}(t))$ at t_o of the constraints determined by a control curve $u^{(\omega)}(t)$ taking the value $u^{(\omega)}(t_o) = \omega$ in an infinitesimal neighbourhood of $t = t_o$ and coinciding with the original $\bar{u}(t)$ at all other points.

2.2. Our road map towards the main results.

Let us now focus on the following two facts, pointed out in the above summary of [3].

- (1) The problems, to which the classical PMP applies, are costs minimising problems on curves $x(t) = (x^i(t))$ that are controlled by means of *first order differential equations with parameters*.
- (2) By introducing auxiliary variables p_j and an appropriate family of controlled Lagrangians $L^{u^{(\cdot)}}(t, p, x, \dot{x})$, the differential constraints of the original control problem are replaced by the Euler-Lagrange equations of such Lagrangians.
- (3) The Lagrangians $L^{u^{(\cdot)}}(t, p, x, \dot{x})$ have the following peculiar property: *they vanish identically along the solutions of their corresponding Euler-Lagrange equations*. This is a crucial fact that leads to the identity (2.5).
- (4) If appropriate boundary conditions are imposed on the $p_j(t)$, then for any given homotopy $u^{(s)}(t)$ of control curves, the integrals of the actions of Poincaré-Cartan type along the “vertical sides” $\eta^{(0)}(s)$, $\eta^{(T)}(s)$ of the surface in Fig. 1 are identically vanishing. This property together with (3) leads to the identity (2.7), which expresses the difference between two terminal costs as a double integral of an appropriate function of the parameters of the homotopy.
- (5) The identity (2.7) immediately gives the Principle of Minimal Labour, from which the PMP is derived using highly localised (needle) variations. In a sense, the Principle of Minimal Labour can be taken as an underlying substratum for the PMP.

In this paper we consider a special class of cost minimising problems, in which the curves are constrained by *differential equations of higher order and of variational type*, that is by Euler-Lagrange equations of controlled Lagrangians of higher order (Sect. 4). For such problems we are able to follow the same circle of ideas described above and, at the end, we reach a generalised version of the PMP that works for this wider class of cost minimising problems.

We remark that, since our differential constraints are assumed to be of Euler-Lagrange type, there is no need to introduce auxiliary variables and new constraints in order to obtain the property described in the point (2) above: It is granted from the very beginning. However, since we are no longer requiring that the controlled Lagrangian is of a very special form, the crucial phenomena described in the points (3) and (4) are in general not occurring. We nonetheless manage to overcome this difficulty through the following two steps.

- (a) We consider a special set of functions, denoted by $\mathbf{h}_\beta^i(t)$, $\mathbf{h}'_\beta^i(t)$ and $\mathbf{h}''_\beta^i(t)$, which are completely determined by the initial and the terminal points of each controlled curve. Such new functions are used in a convenient way to modify the controlled Lagrangian and obtain an analogue of the phenomenon described in (4) (Sect. 5.1).

- (b) We introduce two auxiliary variables, called λ and μ , and we further adjust the Lagrangian in order to obtain a new controlled Lagrangian. The purpose of this is to get an additional property, which is analogous to the one described in (3). As a consequence, we get an identity of the form (2.5) (Sect. 5.2). We stress the fact that the introduction of the new variables λ and μ is a merely technical expedient and that such variables do not occur in the statements of the final results.

The modifications described in (a) and (b) have been found by heuristic arguments. We do not know whether these are the only possible ones and/or there are deep reasons for why they work ⁽¹⁾. Nonetheless for the purposes of the present paper we only need to know that they work. In fact, we would like to stress that their main use is basically to re-write the sum of a particular surface integral and of two boundary line integrals into a single surface integral. This yields an elegant expression for the Principle of Minimal Labour, which nonetheless is by no means the only possible one. Other different equivalent statements are admissible, which can be proven with no need of the above mentioned modifications, but have the disadvantage of being much more involved.

Performing the steps (a) and (b) and following essentially the same ideas used in [3] we finally get the desired analogues of the Principle of Minimal Labour and of the PMP (Sect. 7.2). An informal description of such results is given in the next subsection.

We conclude inviting the reader to consider the outline of this section as a road map for the following constructions and to constantly keep it in mind while going through the rest of the paper.

2.3. Main results.

The outcomes of the above described approach are basically two. The first is a principle (Theorem 6.6) that can be considered as a generalisation of the first version of the Principle of Minimal Labour proved in [3]. It can be described as follows.

Principle of Minimal Labour. *Consider a terminal cost minimising problem for evolutions $q(t) = (q^i(t))$, $t \in [0, T]$, which are constrained by a set of smooth equations on the initial values and by a system of ordinary differential equations of Euler-Lagrange form*

$$\frac{\partial L^{u(\cdot)}}{\partial q^i} \Big|_{q(t)} + \sum_{\beta=1}^r (-1)^\beta \left(\frac{d}{dt} \right)^\beta \left(\frac{\partial L^{u(\cdot)}}{\partial \frac{d^\beta q^i}{dt^\beta}} \right) \Big|_{q(t)} = 0 .$$

Here, $L^{u(\cdot)} = L^{u(\cdot)}(t, q^i, \frac{dq^i}{dt}, \dots, \frac{d^r q^i}{dt^r})$ is a family of Lagrangians of order $r \geq 1$, which smoothly depends on the values of a curve $u(t) = (u^a(t))$ of control parameters.

A curve $\bar{u}(t)$ and admissible initial conditions determine a solution for the considered cost minimising problem only if for any other curve $u(t)$ and any other set of admissible initial conditions, which can be joint to the previous by a smooth one-parameter

¹We guess that a more elegant approach should exist. For instance, the idea we used for the modifications described in (b) calls to mind a well known trick, which is usually exploited to translate a Bolza problem into a Mayer problem (see e.g. [2], p. 116).

deformation, the following inequality holds

$$\int_0^T \left(\int_0^1 \left(\frac{\partial \mathcal{P}_{(t,s)}}{\partial u^a} \frac{\partial u^{(s)a}}{\partial s} \Big|_{u^{(s)}(t)} - \frac{\partial^2 \mu'}{\partial t \partial s} \Big|_{(t,s)} \right) ds \right) dt \leq 0. \quad (2.8)$$

Here:

- $s \in [0, 1]$ is the parameter of the smooth deformation of initial data and control curves and $\frac{\partial u^{(s)a}}{\partial s}$ are the derivatives with respect to s of the components of the curves $u^{(s)}(t) = (u^{(s)a}(t))$ of such deformation;
- $\mathcal{P}_{(t,s)}$ is the two-parameters family of functions of the values $u = (u^a)$ defined by

$$\mathcal{P}_{(t,s)}(u) = -L^{u^{(\cdot)} \equiv u} \left(t, q^{(s)i}(t), \frac{dq^{(s)i}}{dt} \Big|_t, \dots, \frac{d^r q^{(s)i}}{dt^r} \Big|_t \right),$$

where $(q^{(s)i}(t))$ denotes the evolution of the system, which is determined by the control curve $u^{(s)}(t)$ and the initial data associated with the deformation parameter s ;

- $\mu' = \mu'(t, s)$ is the function which is defined in (6.19); it is indeed the time integral between 0 and t of a certain function, which is explicitly given in that formula and it is determined by the following three sets of objects:
 - (A) the curves $u^{(s)}(t)$ of the homotopy of the control curves;
 - (B) the values and the derivatives up to the order $2r$ of the curves $(q^{(s)i}(t))$;
 - (C) the infinitesimal variations of the terminal costs of these curves w.r.t. s .

For a classical Mayer problem, where the dynamical system is described by curves

$$q(t) = (x^i(t), p_j(t)) \quad \text{with} \quad p_j(T) = -\frac{\partial C}{\partial x^j} \Big|_{(x(T))}, \quad (2.9)$$

the above principle radically simplifies. More precisely one has that:

- (1) For any homotopy $q^{(s)}(t)$ of controlled curves of the above type, the double integral $\int_0^T \int_0^1 \frac{\partial^2 \mu'}{\partial t \partial s} \Big|_{(t,s)} ds dt$ in (2.8) vanishes. In fact, the end-point constraints on the p_j force the components $p_j(t, s) := p_j^{(s)}(t)$ of each homotopy $q^{(s)}(t) = (x^{(s)i}(t), p_j^{(s)}(t))$ to play the role of surrogates for the $\mu'(t, s)$. Indeed each $p_j(t, s)$ is uniquely determined by the curves $u^{(s)}(t)$ and $x^{(s)}(t) = (x^{(s)i}(t))$, and by the infinitesimal variations of the terminal costs at $t = T$. This occurs in perfect analogy with the properties (A), (B), (C) of the function $\mu'(t, s)$.
- (2) The partial derivatives $\frac{\partial \mathcal{P}_{(t,s)}}{\partial u^a}$ coincide with the partial derivatives $\frac{\partial \mathcal{H}}{\partial u^a}$.
- (3) The principle we just mentioned reduces to the Principle of Minimal Labour presented in [3] (see Sect. 7.4 below for details).

In contrast with all this, for other kinds of cost minimising problems, no analogues of the auxiliary variables p_j are involved and the term $\int_0^T \int_0^1 \frac{\partial^2 \mu'}{\partial t \partial s} \Big|_{(t,s)} ds dt$ in (2.8) cannot be expected to be zero.

The second main result of our paper is obtained by applying the above Generalised Principle of Minimal Labour to the case of highly localised (“needle”) variations. Indeed, what we obtain can be considered as an analogue of the classical PMP for the above mentioned large class of the higher order Euler-Lagrange constraints of *normal type* (see Sect. 7.1 for the precise definition). It consists of a necessary condition for a control

curve $u_o(t)$ in order to determine an optimal solution γ_o and can be roughly described as follows. Let \mathcal{P} be the function on K defined by

$$\mathcal{P} : K \longrightarrow \mathbb{R} , \quad \mathcal{P}(u^a) := -L\left(t, \gamma_o(t), \dots, \frac{d^r \gamma}{dt}, u^a\right) .$$

Then, $u(t)$ gives an optimal solution γ_o only if, for any $\tau \in [0, T]$ and any $\omega \in K$, for which the curve $u_o(t)$ admits a smooth deformation $u_{(s)}(t)$ of needle type around $t = \tau$ with $u_{s=1}(\tau) = \omega$, the following inequality holds

$$\mathcal{P}(\omega) - \liminf_{\varepsilon \rightarrow 0^+} \frac{\mu'(\tau, \omega, \varepsilon, \Sigma)(T, 1) - \mu'(\tau, \omega, \varepsilon, \Sigma)(T, 0)}{\varepsilon} \leq \mathcal{P}(u_o(\tau)) . \quad (2.10)$$

Here, $\liminf_{\varepsilon \rightarrow 0^+} \frac{\mu'(\tau, \omega, \varepsilon, \Sigma)(T, 1) - \mu'(\tau, \omega, \varepsilon, \Sigma)(T, 0)}{\varepsilon}$ is a corrective term, which can be determined by means of the data of the needle variation, namely: (i) the point $t = \tau$ where it is applied, (ii) the width ε , (iii) the top value ω and (iv) the 1-parameter family Σ of initial or terminal values for some variables, as e.g. the conditions (2.6) in the classical Pontryagin setting (see Sect. 7 for details). For the classical Mayer problems with first order constraints, the above corrective term is zero and the resulting condition (2.10) on the function \mathcal{P} reduces to the usual PMP on the Pontryagin function \mathcal{H} . For what concerns more general cases with higher order constraints, we offer a characterisation of the needle variations, which allow to neglect such corrective term. This yields to an alternative version of the usual PMP, which we briefly illustrate and compare with the classical PMP through the discussion of an elementary example in Sect. 8.

3. PRELIMINARIES

3.1. Notational issues.

Throughout this paper, we consider a dynamical system, whose states are represented by the points of an appropriate N -dimensional *manifold* \mathcal{Q} , which might be for instance a configuration space or a phase space for the system. A generic set of local coordinates on \mathcal{Q} will be usually denoted by $(q^i)_{i=1, \dots, N}$, so that the evolutions in time of our system correspond to parameterised curves with coordinate expressions of the form $q(t) = (q^1(t), \dots, q^N(t))$ for t in a fixed interval $I \subset \mathbb{R}$.

Any such curve is uniquely associated with the corresponding parameterised graph

$$\gamma : I \subset \mathbb{R} \longrightarrow \mathbb{R} \times \mathcal{Q} , \quad \gamma(t) = (t, q(t)) , \quad (3.1)$$

i.e. the associated local section of the trivial bundle $\mathbb{R} \times \mathcal{Q}$ over \mathbb{R} . Due to this, from now on we identify any evolution of our system with the associated map $t \mapsto \gamma(t) = (t, q(t))$.

For any such a map $\gamma : I \rightarrow \mathbb{R} \times \mathcal{Q}$ of class \mathcal{C}^n , $n \geq 1$, we denote by $j_{t_o}^n(\gamma)$ its n -th order jet at the time t_o . For a classical reference on jets, see for instance [9] ⁽²⁾. The collection of all these n -th order jets has a natural structure of smooth manifold and it is called the *jet bundle of order n* of the (trivial) bundle $\mathbb{R} \times \mathcal{Q}$ over \mathbb{R} . We denote it by $J^n(\mathcal{Q}|\mathbb{R})$.

²For convenience of the reader, it might be convenient to briefly mention what we are going to adopt as definition of an n -th order jet of a curve γ in $\mathbb{R} \times \mathcal{Q}$ at a point t_o . It is the equivalence class of all the curves of the form $\eta :] - t_o - \delta, t_o + \delta[\rightarrow \mathbb{R} \times \mathcal{Q}$, whose components in some (thus, in any) set of coordinates have values and derivatives up to order n at t_o equal to those of γ at such a point.

For a fixed chart $(q^i) : \mathcal{U} \rightarrow \mathbb{R}^N$ on some open set $\mathcal{U} \subset \mathcal{Q}$, we may consider the map which sends each jet $j_t^n(\gamma)$ into the \tilde{N} -tuple, $\tilde{N} := N(n+1) + 1$,

$$j_t^n(\gamma) \mapsto \left(t, q_{(0)}^i(t) := q^i(t), q_{(1)}^i(t) := \frac{dq^i}{dt} \Big|_t, q_{(2)}^i(t) := \frac{d^2q^i}{dt^2} \Big|_t, \dots, q_{(n)}^i(t) := \frac{d^n q^i}{dt^n} \Big|_t \right).$$

This map is well known to be a (locally defined) system of coordinates for $J^n(\mathcal{Q}|\mathbb{R})$. In a short notation, we denote such coordinates by

$$(t, q_{(\beta)}^i) = (t, q_{(0)}^i, q_{(1)}^i, \dots, q_{(n)}^i)$$

and we call them *canonical jet coordinates* determined by the coordinates (q^i) of \mathcal{Q} .

For any $\gamma(t) = (t, q^i(t))$, the n -th order lift of γ is the corresponding curve of jets

$$\gamma^{(n)} : I \longrightarrow J^n(\mathcal{Q}|\mathbb{R}), \quad \gamma^{(n)}(t) := j_t^n(\gamma) = (t, q_{(\beta)}^i(t)) = \left(t, q^i(t), \frac{dq^i}{dt} \Big|_t, \dots, \frac{d^n q^i}{dt^n} \Big|_t \right).$$

We denote by $K \subset \mathbb{R}^M$ a fixed set of real M -tuples $u = (u^a)$. In what follows, a (continuous or k -differentiable) curve $t \mapsto u(t)$ with values in K plays the role of a *control* for our system. We stress the fact that, in the literature, the terms “control”, “control parameter” or “control value” is usually adopted to refer just to a single value of the curve $u(t)$, not to the curve $u(\cdot)$ as a whole, in contrast with what we do in this paper. We hope that this will not be a source of confusion.

We always assume that K is the closure of a bounded open subset of \mathbb{R}^M and that the boundary ∂K is smooth. This assumption is mainly made for the sake of simplicity, since most of our arguments can be generalised to a large class of more general situations and under weaker regularity assumptions.

3.2. Controlled Lagrangians and controlled Euler-Lagrange operators.

Let us consider a (*smooth*) *Lagrangian with controls*, that is a \mathcal{C}^∞ function

$$L = L(t, q_{(\beta)}^i, u^a) : J^n(\mathcal{Q}|\mathbb{R}) \times K \rightarrow \mathbb{R}$$

depending on

- the coordinates $(t, q_{(0)}^i, q_{(1)}^i, \dots, q_{(n)}^i)$ of the n -th jets in $J^n(\mathcal{Q}|\mathbb{R})$,
- the parameters $u = (u^a)$ arbitrarily varying in $K \subset \mathbb{R}^M$.

If there is an integer r such that L is independent on all of the jets coordinates $q_{(\beta)}^i$ with $\beta \geq r + 1$, we say that r is the *actual order of L* . For example, the function

$$L(t, q_{(\beta)}^i, u^a) = \left(\sum_{a=1}^M u^a \right) \frac{1}{2} \sum_{i=1}^N (q_{(1)}^i)^2 - \frac{1}{2} \sum_{i=1}^N (q^i)^2$$

can be surely considered as a Lagrangian with controls on any controlled jet space $J^n(\mathcal{Q}|\mathbb{R}) \times K$ of order $n \geq 1$. If we decide that our working ambient is $J^n(\mathcal{Q}|\mathbb{R})$ for some n which is strictly larger than 1, the property that L is independent of the jet coordinates $q_{(2)}^i, q_{(3)}^i, \dots$, is synthetically expressed by saying that $r = 1$ is the *actual order of L* .

In our discussions, we will always assume that the order n of the controlled jet space $J^n(\mathcal{Q}|\mathbb{R}) \times K$ is sufficiently larger than the actual order r of the considered Lagrangian L .

This is needed for letting all of the operators considered in this paper (as, for instance, the Euler-Lagrange operator described below) to be meaningful. As we will shortly see, we just need that the order of the jet space $J^n(\mathcal{Q}|\mathbb{R})$ satisfies the following inequality

$$2r + 1 \leq n , \tag{3.2}$$

which, from now on, we constantly and tacitly assume to be satisfied ⁽³⁾.

The *controlled Euler-Lagrange operator* is the N -tuple $E = (E_i)_{i=1}^N$ of differential operators, acting on any controlled Lagrangian L of actual order r , defined by

$$E_i(L) := \frac{\partial L}{\partial q^i} + \sum_{\beta=1}^r (-1)^\beta \left(\frac{d}{dt}\right)^\beta \left(\frac{\partial L}{\partial q_{(\beta)}^i}\right) , \quad i = 1, \dots, N . \tag{3.3}$$

Here, the symbol $\frac{d}{dt}$ denotes the *total derivative operator*, that is the operator that transforms any function $f : J^n(\mathcal{Q}|\mathbb{R}) \times K \rightarrow \mathbb{R}$ of actual order $r' \leq n - 1$ into the function

$$\frac{df}{dt} : J^n(\mathcal{Q}|\mathbb{R}) \times K \longrightarrow \mathbb{R} , \quad \frac{df}{dt} \Big|_{(t, q_{(\beta)}^i)} := \frac{\partial f}{\partial t} \Big|_{(t, q_{(\beta)}^i)} + \sum_{\substack{1 \leq j \leq N \\ 0 \leq \delta \leq r'}} \frac{\partial f}{\partial q_{(\delta)}^j} q_{(\delta+1)}^j \Big|_{(t, q_{(\beta)}^i)} . \tag{3.4}$$

Such an operator is called “total derivative” simply because, for any curve $(\gamma(t), u(t)) \in (\mathbb{R} \times \mathcal{Q} \times K)^{[0, T]}$ with constant $u(t) \equiv u_o$, the evaluation of $\frac{df}{dt}$ at the points of the curve of jets $(\gamma^{(n)}(t), u(t) = u_o)$ is equal to

$$\frac{df}{dt} \Big|_{(\gamma^{(n)}(t), u_o)} = \frac{d}{dt} \left(f(\gamma^{(n)}(t), u_o) \right) \Big|_t \quad \text{for each } t \in [0, T] , \tag{3.5}$$

i.e. it coincides with the derivative with respect to t of the map $t \mapsto f(\gamma^{(n)}(t), u_o)$.

Notice that:

- The total derivative raises the actual order of a function of at most one unit and the iterated total derivatives $\left(\frac{d}{dt}\right)^\beta$, $1 \leq \beta \leq r$, raise the actual orders of at most r units. This is one of the reasons why we assume (3.2). Other reasons for this come from the fact that certain arguments of the proof of Lemma 5.2 below work nicely only if $2r$ is actually *strictly* less than n .
- What we call “controlled Euler-Lagrange operator” is *almost* the same of the Euler-Lagrange operator of the classical theory of variations. The only difference with respect to the usual one is that the operators $E_a = \frac{\partial}{\partial u^a} + \sum_{\beta=1}^r (-1)^\beta \left(\frac{d}{dt}\right)^\beta \left(\frac{\partial}{\partial u_{(\beta)}^a}\right)$, $1 \leq a \leq M$, corresponding to the infinitesimal variations of the coordinates u^a , are here missing.

4. DEFINING TRIPLES AND GENERALISED MAYER PROBLEMS

We are now able to delineate in detail the particular class of control problems, which is the main object of study of this paper. As we already mentioned, the dynamical systems we are dealing with evolve according to curves, whose parameterised graphs are

³This assumption could have been safely omitted if we considered the infinite jet spaces $J^\infty(\mathcal{Q}|\mathbb{R})$ in place of finite order jet spaces $J^n(\mathcal{Q}|\mathbb{R})$. However this would have forced us to work with infinite-dimensional manifolds, a category that, for simplicity of the exposition, we prefer to leave undisturbed.

the curves (3.1) in $\mathbb{R} \times \Omega$. The independent variable t (the “time”) of these evolutions is from now on always assumed to be varying in a fixed interval $[0, T]$.

The control problems on which we focus are those given by the following ingredients.

- A *set of control parameters* \mathcal{K} , i.e. a set of pairs $U = (u(t), \sigma)$, in which:
 - a) the first element $u(t)$ is a smooth curve $u : [0, T] \rightarrow K$ in the above fixed ambient space $K \subset \mathbb{R}^M$;
 - b) the second element is a jet $\sigma = j_{t=0}^{2r-1}(\gamma)$ where r is the actual order of the controlled Lagrangian considered below; the jet σ is possibly subjected to some constraints (as, for instance, that the 0-th component $j_{t=0}^0(\gamma) = \gamma(0)$ is equal to a fixed point $q_o \in \Omega$) and is later used as the initial datum for a curve $\gamma(t) = (t, q(t))$, described in the next point.
- A *Lagrangian with controls* $L = L(t, q_{(\beta)}^i, u^a)$ of actual order r satisfying (3.2), which gives the *system of controlled Euler-Lagrange equations* of order $2r$

$$E_i(L)|_{(j^n(\gamma(t)), u(t))} = 0 \quad (4.1)$$

for each smooth curve $u : [0, T] \rightarrow K$. We also assume that L satisfies all needed regularity and maximal rank conditions that guarantee the following: *for each pair $U = (u(t), \sigma)$ in the set \mathcal{K} , there exists a unique solution $\gamma^{(U)}(t)$ to the initial value problem formed by the equations (4.1) and the initial condition*

$$j_{t=0}^{2r-1}(\gamma(t)) = \sigma . \quad (4.2)$$

The curves $\gamma^{(U)}$ determined in this way are called \mathcal{K} -controlled.

- A *terminal cost function*, that is a real function of the jets at the time $t = T$ of some fixed order $\tilde{r} \leq n - 1$ ⁽⁴⁾. We assume that such a terminal cost function is actually the restriction $C|_{J^n(\mathcal{Q}|\mathbb{R})|_{t=T}}$ of a smooth real function of actual order \tilde{r} on the whole jet space that vanishes identically on $J^n(\mathcal{Q}|\mathbb{R})|_{t=0}$.

For any terminal cost function on the jets at $t = T$, there are clearly infinite possibilities for a smooth function C on $J^n(\mathcal{Q}|\mathbb{R})$ that vanishes at the jets at $t = 0$ and that gives the desired cost function at $t = T$. But in what follows we select just one of such globally defined functions and we call it *the (extended) cost function* of our problem.

Any triple (\mathcal{K}, L, C) , formed by three ingredients of the above form, is called a *defining triple*. Given such a triple we may consider the following

Definition 4.1. The *generalised Mayer problem determined by (\mathcal{K}, L, C)* is the problem of looking for all \mathcal{K} -controlled evolutions $\gamma^{(U)}$ for which the value of the terminal cost $C(j_{t=T}^n(\gamma^{(U)}))$ is minimal among the terminal costs of all \mathcal{K} -controlled evolutions. For such curves, the corresponding pairs $U = (u(t), \sigma) \in \mathcal{K}$ are called *optimal controls*.

⁴This condition on the actual order \tilde{r} is just a convenient technical requirement and is used only in the proof of Lemma 5.2 below. As observed before, if we work in jets spaces of sufficiently high order, this requirement is always easily satisfied.

5. MODIFIED TRIPLES AND MAYER PROBLEMS IN INTEGRAL FORM

5.1. Performing the step (a) of the road map.

We now proceed according to the step (a) described in Sect. 2.2. More precisely, we canonically associate with any given curve $(t, q(t))$ in $[0, T] \times \Omega$ a special set of real functions. They are

$$\mathbf{h}_\beta^i(t) = A_\beta^i e^t + B_\beta^i e^{-t}, \quad (5.1)$$

$$\mathbf{h}'_\beta^i(t) = A'^i_\beta e^{\frac{\pi}{2T}t} + B'^i_\beta e^{-\frac{\pi}{2T}t} + C'^i_\beta \cos\left(\frac{\pi}{2T}t\right) + D'^i_\beta \sin\left(\frac{\pi}{2T}t\right), \quad (5.2)$$

$$\mathbf{h}''_\beta^i(t) = A''^i_\beta e^{\frac{\pi}{2T}t} + B''^i_\beta e^{-\frac{\pi}{2T}t} + C''^i_\beta \cos\left(\frac{\pi}{2T}t\right) + D''^i_\beta \sin\left(\frac{\pi}{2T}t\right), \quad (5.3)$$

where the indices i and β run between $1 \leq i \leq N$ and $0 \leq \beta \leq r - 1$ (here, r is the actual order of L) and the $A_\beta^i, B_\beta^i, A'^i_\beta, B'^i_\beta$ etc., are the constants that are uniquely determined by the following conditions on the initial and terminal data of the $q^i(t)$. The A_β^i and B_β^i are determined by solving the linear equations

$$A_\beta^i + B_\beta^i (= \mathbf{h}_\beta^i|_{t=0}) = q_{(\beta)}^i|_{t=0}, \quad A_\beta^i - B_\beta^i (= \left. \frac{d\mathbf{h}_\beta^i}{dt} \right|_{t=0}) = - \sum_{\substack{1 \leq \delta \leq r \\ 0 \leq \varepsilon \leq \delta \\ \delta - \varepsilon - 1 = \beta}} (-1)^\varepsilon \frac{d^\varepsilon}{dt^\varepsilon} \left(\frac{\partial L}{\partial q_{(\delta)}^i} \right) \Big|_{j_{t=0}^{n-1}(q(t))}. \quad (5.4)$$

The remaining constants are set to

$$\begin{pmatrix} A_\beta^i \\ B_\beta^i \\ C_\beta^i \\ D_\beta^i \end{pmatrix} = \mathcal{A}^{-1} \begin{pmatrix} 0 \\ 0 \\ q_{(\beta)}^i(T) \\ \sum_{\substack{1 \leq \delta \leq r \\ 0 \leq \varepsilon \leq \delta \\ \delta - \varepsilon - 1 = \beta}} (-1)^\varepsilon \frac{d^\varepsilon}{dt^\varepsilon} \left(\frac{\partial L}{\partial q_{(\delta)}^i} + \frac{\partial}{\partial q_{(\delta)}^i} \left(\frac{dC}{dt} \right) \right) \Big|_{j_{t=T}^{n-1}(q(t))} \end{pmatrix}, \quad (5.5)$$

$$\begin{pmatrix} A''_\beta^i \\ B''_\beta^i \\ C''_\beta^i \\ D''_\beta^i \end{pmatrix} = \mathcal{A}^{-1} \begin{pmatrix} 0 \\ 0 \\ \mathbf{h}_\beta^i(T) \\ \mathbf{h}_{\beta(1)}^i(T) \end{pmatrix}, \quad (5.6)$$

where \mathcal{A} is the real matrix

$$\mathcal{A} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ \frac{\pi}{2T} & -\frac{\pi}{2T} & 0 & \frac{\pi}{2T} \\ \frac{\pi}{2T} e^{\frac{\pi}{2}} & -\frac{\pi}{2T} e^{-\frac{\pi}{2}} & -\frac{\pi}{2T} & 0 \\ \left(\frac{\pi}{2T}\right)^2 e^{\frac{\pi}{2}} & \left(\frac{\pi}{2T}\right)^2 e^{-\frac{\pi}{2}} & 0 & -\left(\frac{\pi}{2T}\right)^2 \end{pmatrix}.$$

A tedious but straightforward check shows that the functions (5.1) – (5.3) are precisely the unique solutions to the system of differential equations

$$\frac{d^2 \mathbf{h}_\beta^i}{dt^2} - \mathbf{h}_\beta^i = 0, \quad \frac{d^4 \mathbf{h}'_\beta^i}{dt^4} - \left(\frac{\pi}{2T}\right)^4 \mathbf{h}'_\beta^i = 0, \quad \frac{d^4 \mathbf{h}''_\beta^i}{dt^4} - \left(\frac{\pi}{2T}\right)^4 \mathbf{h}''_\beta^i = 0 \quad (5.7)$$

together with the set of the boundary conditions formed by the (5.4) and by

$$\mathbf{h}'_{\beta}(0) = 0, \quad \left. \frac{d\mathbf{h}'_{\beta}}{dt} \right|_{t=0} = 0, \quad (5.8)$$

$$\left. \frac{d\mathbf{h}'_{\beta}}{dt} \right|_{t=T} = \dot{q}_{(\beta)}^i(T), \quad \left. \frac{d^2\mathbf{h}'_{\beta}}{dt^2} \right|_{t=T} = \sum_{\substack{1 \leq \delta \leq r \\ 0 \leq \varepsilon \leq \delta \\ \delta - \varepsilon - 1 = \beta}} (-1)^{\varepsilon} \frac{d^{\varepsilon}}{dt^{\varepsilon}} \left(\frac{\partial (L + \frac{dC}{dt})}{\partial q_{(\delta)}^i} \right) \Big|_{j_{t=T}^{n-1}(q(t))} \quad (5.9)$$

$$\mathbf{h}''_{\beta}(0) = 0, \quad \left. \frac{d\mathbf{h}''_{\beta}}{dt} \right|_{t=0} = 0, \quad \left. \frac{d\mathbf{h}''_{\beta}}{dt} \right|_{t=T} = \mathbf{h}_{\beta}^i(T), \quad \left. \frac{d^2\mathbf{h}''_{\beta}}{dt^2} \right|_{t=T} = \mathbf{h}_{\beta(1)}^i(T). \quad (5.10)$$

As a matter of fact, we consider these functions just because we want them to satisfy such a differential problem. The motivation for this requirement has been very roughly indicated in our road map and it will be definitely clarified in the next section.

Using the functions (5.1) – (5.3), with any given curve $(t, q^i(t))$ in $[0, T] \times \mathcal{Q}$ we may associate a curve $(t, q^i(t), \mathbf{h}_{\beta}^i(t), \mathbf{h}'_{\beta}(t), \mathbf{h}''_{\beta}(t))$ in the enlarged manifold $[0, T] \times \tilde{\mathcal{Q}}$, $\tilde{\mathcal{Q}} := \mathcal{Q} \times \mathbb{R}^{3Nr}$. Such a bijective correspondence between curves in $[0, T] \times \mathcal{Q}$ and in $[0, T] \times \tilde{\mathcal{Q}}$ establishes a natural equivalence between our original control problem, determined by the triple (\mathcal{K}, L, C) , and a new control problem, determined by an appropriate modified defining triple $(\tilde{\mathcal{K}}, \tilde{L}, \tilde{C})$ which is defined as follows.

- $\tilde{\mathcal{K}}$ is the collection of pairs $\tilde{U} = (u(t), \tilde{\sigma})$ in which: (a) $u(t)$ is precisely as it occurs in the pairs $(u(t), \sigma) \in \mathcal{K}$ and (b) $\tilde{\sigma}$ is a jet in $J^n(\tilde{\mathcal{Q}}|\mathbb{R})|_{t=0}$ of actual order $2 \max\{r, 2\} - 1$, playing the role of the initial datum of a curve $\tilde{\gamma}(t) = (t, q^i(t), \mathbf{h}_{\beta}^i(t), \mathbf{h}'_{\beta}(t), \mathbf{h}''_{\beta}(t))$, constrained by the following conditions:

- (i) the part of $\tilde{\sigma}$, corresponding to the initial datum of $\gamma(t) = (t, q^i(t))$, satisfies the same constraints that are imposed on the data σ in the pair $(u(t), \sigma)$ in \mathcal{K} ;
- (ii) the initial values for the curves $\mathbf{h}_{\beta}^i(t)$, $\mathbf{h}'_{\beta}(t)$ and $\mathbf{h}''_{\beta}(t)$ are required to satisfy the conditions given by the (5.4), (5.9) and (5.10); no other condition is imposed besides those which are naturally requested in order to be initial conditions that are fully compatible with the Euler-Lagrange equations (5.13) and (5.14) below ⁽⁵⁾.

- \tilde{L} is the controlled Lagrangian of actual order $\tilde{r} = \max\{r, 2\}$

$$\begin{aligned} \tilde{L}(t, q_{(\delta)}^i, \mathbf{h}_{\beta(\delta)}^j, \dots, u^a) := & L + \frac{1}{2} \sum_{\substack{1 \leq j \leq N \\ 0 \leq \beta \leq r-1}} \left((\mathbf{h}_{\beta(1)}^j)^2 - (\mathbf{h}'_{\beta(2)}^j)^2 - (\mathbf{h}''_{\beta(2)}^j)^2 \right) + \\ & + \sum_{\substack{1 \leq j \leq N \\ 0 \leq \beta \leq r-1}} \left(\frac{1}{2} (\mathbf{h}_{\beta}^j)^2 + \frac{\pi^4}{32T^4} (\mathbf{h}'_{\beta}^j)^2 + \frac{\pi^4}{32T^4} (\mathbf{h}''_{\beta}^j)^2 \right). \quad (5.11) \end{aligned}$$

- \tilde{C} is equal to $\tilde{C} = C$. The only difference between C and \tilde{C} is just that its formal domain is now $J^n(\tilde{\mathcal{Q}}|\mathbb{R})$ (and no longer $J^n(\mathcal{Q}|\mathbb{R})$).

⁵As we will shortly see, such Euler-Lagrange equations are nothing but the (5.7).

If we now replace the triple (\mathcal{K}, L, C) by $(\tilde{\mathcal{K}}, \tilde{L}, \tilde{C} = C)$, we have a new Euler-Lagrange operator, consisting of the differential operators $E_i(\cdot)$, $E_i^\beta(\cdot)$, $E_i^{\prime\beta}(\cdot)$, $E_i^{\prime\prime\beta}(\cdot)$, corresponding to the variables q^i , \mathbf{h}_β^i , \mathbf{h}'_β^i and \mathbf{h}''_β^i , respectively. The new set of Euler-Lagrange equations is

$$E(\tilde{L})_i = \frac{\partial L}{\partial q^i} + \sum_{\ell=1}^r (-1)^\ell \frac{d^\ell}{dt^\ell} \left(\frac{\partial L}{\partial q^{i(\ell)}} \right) = 0, \tag{5.12}$$

$$E(\tilde{L})_i^\beta = -\frac{d^2 \mathbf{h}_\beta^i}{dt^2} + \mathbf{h}_\beta^i = 0, \quad E(\tilde{L})_i^{\prime\beta} = -\frac{d^4 \mathbf{h}'_\beta^i}{dt^4} + \left(\frac{\pi}{2T}\right)^4 \mathbf{h}'_\beta^i = 0, \tag{5.13}$$

$$E(\tilde{L})_i^{\prime\prime\beta} = -\frac{d^4 \mathbf{h}''_\beta^i}{dt^4} + \left(\frac{\pi}{2T}\right)^4 \mathbf{h}''_\beta^i = 0. \tag{5.14}$$

From these equations, we directly see that any $\tilde{\mathcal{K}}$ -controlled curve $\tilde{\gamma}^{\tilde{U}}(t) = (t, q^i(t), \mathbf{h}_\beta^i(t), \mathbf{h}'_\beta^i(t), \mathbf{h}''_\beta^i(t))$ has the following two crucial properties.

- (1) Being solutions to the same differential problem, the components $q^i(t)$ of a $\tilde{\mathcal{K}}$ -controlled curve $\tilde{\gamma}^{\tilde{U}}$ and of the corresponding \mathcal{K} -controlled curve γ^U are the same. Therefore the terminal costs $\tilde{C}|_{j_{t=T}^n(\tilde{\gamma}^{\tilde{U}})}$ and $C|_{j_{t=T}^n(\gamma^U)}$ coincide.
- (2) The functions $\mathbf{h}_\beta^i(t)$, $\mathbf{h}'_\beta^i(t)$, $\mathbf{h}''_\beta^i(t)$ have necessarily the forms (5.1) – (5.3).

Thus the generalised Mayer problem defined by $(\tilde{\mathcal{K}}, \tilde{L}, \tilde{C})$ is perfectly equivalent to the original one, given by (\mathcal{K}, L, C) . The bijection between the two families of controlled curves is established by simply considering the functions defined in (5.1) – (5.3) as the last components of a curve in $\mathbb{R} \times \tilde{\mathcal{Q}} = (\mathbb{R} \times \mathcal{Q}) \times \mathbb{R}^{3Nr}$.

5.2. Performing the step (b) of the road map.

We now go into the step (b) of Sect. 2.2. Namely, we introduce two auxiliary variables λ , μ and further modify the defining triple of the problem, so that an analogue of (2.5) holds for any solution of the new controlled Euler-Lagrange equations.

Introducing two new variables corresponds to enlarging the manifold $\tilde{\mathcal{Q}} = \mathcal{Q} \times \mathbb{R}^{3Nr}$ of the previous section into the new manifold $\hat{\mathcal{Q}} = \tilde{\mathcal{Q}} \times \mathbb{R}^2$ parameterised by the coordinates $(t, q^i, \mathbf{h}_\beta^i, \mathbf{h}'_\beta^i, \mathbf{h}''_\beta^i, \lambda, \mu)$. After considering such new enlarged manifold $\hat{\mathcal{Q}}$, we have to introduce the *further modified triple* $(\hat{\mathcal{K}}, \hat{L}, \hat{C})$ defined as follows.

- $\hat{\mathcal{K}}$ is the collection of pairs $\hat{U} = (u(t), \hat{\sigma})$ in which: (a) $u(t)$ is precisely as it occurs in the pairs in \mathcal{K} and $\tilde{\mathcal{K}}$ and (b) $\hat{\sigma}$ is a jet in $J^n(\hat{\mathcal{Q}}|\mathbb{R})|_{t=0}$ of actual order $2 \max\{r, 2\} - 1$, playing the role of the initial datum of a curve $\hat{\gamma}(t) = (t, q^i(t), \mathbf{h}_\beta^i(t), \mathbf{h}'_\beta^i(t), \mathbf{h}''_\beta^i(t), \lambda(t), \mu(t))$, constrained by the following conditions:

- (i) the part of $\hat{\sigma}$, corresponding to the initial datum of $\tilde{\gamma}(t) = (t, q^i(t), \mathbf{h}_\beta^i(t), \mathbf{h}'_\beta^i(t), \mathbf{h}''_\beta^i(t))$, satisfies the same conditions that are imposed on the pairs $(u(t), \tilde{\sigma}) \in \tilde{\mathcal{K}}$;
- (ii) the component $\lambda|_{t=0}$ of $\hat{\sigma}$ (i.e. the initial value of $\lambda(t)$) is always set to be $\lambda|_{t=0} = 1$; the values of all other components of the jet giving the initial datum for $\lambda(t)$ are required to be just compatible with the Euler-Lagrange equations (5.16) below ⁽⁶⁾;

⁶By looking at those equations, one can see that this compatibility requirement simply means that all derivatives of $\lambda(t)$ at $t = 0$ must be equal to 0.

(iii) the component $\mu|_{t=0}$ of $\widehat{\sigma}$ (i.e. the initial value of $\mu(t)$) must be $\mu|_{t=0} = 0$; the other components of that initial datum for $\mu(t)$ have just to be compatible with the (5.16).

• \widehat{L} is the controlled Lagrangian of actual order $\widehat{r} = \max\{r, 2\}$ defined by

$$\begin{aligned} \widehat{L}(t, q_{(\delta)}^i, \mathbf{h}_{\beta(\delta)}^i, \mathbf{h}'_{\beta(\delta)}{}^i, \mathbf{h}''_{\beta(\delta)}{}^i, \lambda_{(\delta)}, \mu_{(\delta)}, u^a) &:= \lambda(\mu_{(1)} + \widetilde{L}) + \frac{dC}{dt} = \\ &= \lambda \left(\mu_{(1)} + L + \frac{1}{2} \sum_{\substack{1 \leq i \leq N \\ 0 \leq \beta \leq r-1}} \left((\mathbf{h}_{\beta(1)}^i)^2 - (\mathbf{h}'_{\beta(2)}{}^i)^2 - (\mathbf{h}''_{\beta(2)}{}^i)^2 \right) + \right. \\ &\quad \left. + \frac{1}{2} \sum_{\substack{1 \leq i \leq N \\ 0 \leq \beta \leq r-1}} \left((\mathbf{h}_{\beta}^i)^2 + \left(\frac{\pi}{2T}\right)^4 (\mathbf{h}'_{\beta}{}^i)^2 + \left(\frac{\pi}{2T}\right)^4 (\mathbf{h}''_{\beta}{}^i)^2 \right) \right) + \frac{dC}{dt}. \end{aligned} \quad (5.15)$$

Here, according to (3.4), $\frac{dC}{dt}$ denotes the total derivative of the function $C = C(t, q_{(\beta)}^i)$.

• \widehat{C} is the same of the original cost function $\widehat{C} = C$. As before, the difference between \widehat{C} and C is just that we are now considering it as a function on the new jet space $J^n(\widehat{\mathcal{Q}}|\mathbb{R})$.

If we now replace the defining triple $(\widehat{\mathcal{K}}, \widetilde{L}, \widetilde{C})$ of the previous section by $(\widehat{\mathcal{K}}, \widehat{L}, \widehat{C} = C)$, we have to consider another Euler-Lagrange operator, formed not only by the previous operators $E_i(\cdot)$, $E_i^\beta(\cdot)$, $E_i''^\beta(\cdot)$, $E_i''''^\beta(\cdot)$, but also by the operators $E_{\{\lambda\}}(\cdot)$ and $E_{\{\mu\}}(\cdot)$, corresponding to the new variables λ and μ , respectively. The new set of Euler-Lagrange equations is:

$$\begin{aligned} E(\widehat{L})_{\{\mu\}} &= -\frac{d\lambda}{dt} = 0 \quad (\implies \lambda \equiv 1), \\ E(\widehat{L})_i &= E(\lambda \widetilde{L})_i = \frac{\partial L}{\partial q^i} + \sum_{\ell=1}^r (-1)^\ell \frac{d^\ell}{dt^\ell} \left(\frac{\partial L}{\partial q^{i(\ell)}} \right) = 0, \\ E(\widehat{L})_i^\beta &= E(\lambda \widetilde{L})_i^\beta = -\frac{d^2 \mathbf{h}_\beta^i}{dt^2} + \mathbf{h}_\beta^i = 0, \\ E(\widehat{L})_i''^\beta &= E(\lambda \widetilde{L})_i''^\beta = -\frac{d^4 \mathbf{h}'_\beta{}^i}{dt^4} + \left(\frac{\pi}{2T}\right)^4 \mathbf{h}'_\beta{}^i = 0, \\ E(\widehat{L})_i''''^\beta &= E(\lambda \widetilde{L})_i''''^\beta = -\frac{d^4 \mathbf{h}''_\beta{}^i}{dt^4} + \left(\frac{\pi}{2T}\right)^4 \mathbf{h}''_\beta{}^i = 0, \\ E(\widehat{L})_{\{\lambda\}} &= \frac{d\mu}{dt} + \widetilde{L} = \\ &= \frac{d\mu}{dt} + L + \frac{1}{2} \sum_{\substack{1 \leq i \leq N \\ 0 \leq \beta \leq r-1}} \left((\mathbf{h}_{\beta(1)}^i)^2 - (\mathbf{h}'_{\beta(2)}{}^i)^2 - (\mathbf{h}''_{\beta(2)}{}^i)^2 \right) + \\ &\quad + \sum_{\substack{1 \leq i \leq N \\ 0 \leq \beta \leq r-1}} \left(\frac{1}{2} (\mathbf{h}_\beta^i)^2 + \frac{\pi^4}{32T^4} (\mathbf{h}'_\beta{}^i)^2 + \frac{\pi^4}{32T^4} (\mathbf{h}''_\beta{}^i)^2 \right) = 0, \end{aligned} \quad (5.16)$$

By just looking at these equations, we see that any $\widehat{\mathcal{K}}$ -controlled curve

$$\widehat{\gamma}^{(\widehat{U})}(t) = (t, q^i(t), \mathbf{h}_\beta^i(t), \mathbf{h}'_\beta{}^i(t), \mathbf{h}''_\beta{}^i(t), \lambda(t), \mu(t)) \quad (5.17)$$

has the following properties.

- (1) The value $\lambda(t)$ is constant and equal to the prescribed initial value, i.e. $\lambda(t) = 1$. It follows that *the new differential constraints on the curve $\tilde{\gamma}(t) = (t, q^i(t), \mathbf{h}_\beta^i(t), \mathbf{h}'_\beta^i(t), \mathbf{h}''_\beta^i(t))$ are identical with the original constraints (5.12) – (5.14)*. This, together with the uniqueness of the solution to the differential problem for $\mu(t)$, imply that there exists a natural bijection between the class of $\tilde{\mathcal{K}}$ -controlled curves $\tilde{\gamma}^{(\tilde{U})}$ and the class of the $\hat{\mathcal{K}}$ -controlled curves $\hat{\gamma}^{(\hat{U})}$.
- (2) Due to (1), the cost $\hat{C}|_{j_{t=T}^n(\hat{\gamma}^{(\hat{U})})}$ for a $\hat{\mathcal{K}}$ -controlled curve is always equal to the cost $\tilde{C}|_{j_{t=T}^n(\tilde{\gamma}^{(\tilde{U})})} = C|_{j_{t=T}^n(\gamma^{(U)})}$ of the corresponding $\tilde{\mathcal{K}}$ -controlled curve $\tilde{\gamma}^{(\tilde{U})}$.
- (3) The last equation in (5.16) implies that the 1-form $\lambda(\mu_{(1)} + \tilde{L})dt$ vanishes identically along any curve of jets $\hat{\gamma}^{(\hat{U})^{(n)}}(t) = j_t^n(\hat{\gamma}^{(\hat{U})})$ of a $\hat{\mathcal{K}}$ -controlled curve $\hat{\gamma}^{(\hat{U})}$.

From (1) and (2) and previous discussion, we see that *the generalised Mayer problem defined by the triple (\mathcal{K}, L, C) is not only equivalent to the problem of Sect. 5.1, determined by the triple $(\tilde{\mathcal{K}}, \tilde{L}, \tilde{C})$, but also equivalent to this new problem, determined by the triple $(\hat{\mathcal{K}}, \hat{L}, \hat{C})$* . Furthermore (3) shows that the improvement, which was mentioned in the step (b) of the road map, is now reached, In fact, if we consider the 1-form

$$\alpha := \hat{L}dt = \left(\lambda(\mu_{(1)} + \tilde{L}) + \frac{dC}{dt} \right) dt \quad (5.18)$$

we may observe that, for any curve of jets $\hat{\gamma}^{(\hat{U})^{(n)}}(t)$ of a $\hat{\mathcal{K}}$ -controlled curve $\hat{\gamma}^{(U)}$

$$\begin{aligned} \int_0^T \alpha \left(\frac{d\hat{\gamma}^{(\hat{U})^{(n)}}}{dt} \right) \Big|_{\hat{\gamma}^{(\hat{U})^{(n)}}(t)} dt &= \int_0^T \hat{L}|_{\hat{\gamma}^{(\hat{U})^{(n)}}(t)} dt \stackrel{\text{Property (3)}}{=} \int_0^T \frac{dC}{dt} \Big|_{\hat{\gamma}^{(\hat{U})^{(n)}}(t)} dt \stackrel{(3.5)}{=} \\ &= C(\gamma^{(U)^{(n)}}(T)) - C(\gamma^{(U)^{(n)}}(0)) \stackrel{C|_{J^n(\mathbb{Q}|\mathbb{R})|_{t=0}=0}}{=} C(j_{t=T}^n(\gamma^{(U)})) . \end{aligned} \quad (5.19)$$

This means that for each $\hat{\mathcal{K}}$ -controlled curve, the integral (5.19) is just equal to the terminal cost $C|_{j_{t=T}^n(\gamma^{(U)})}$ and that *looking for a solution to the original problem is perfectly equivalent to looking for a $\hat{\mathcal{K}}$ -controlled evolution, for which the integral (5.19) is minimal among those of all other $\hat{\mathcal{K}}$ -controlled evolutions*.

5.3. A convenient replacement of the integrand in (5.19).

We now want to show that we may safely substitute the 1-form (5.18) by a different one, which turns out to be much more convenient for our further developments. In order to introduce such convenient replacement, we first need to recall that on the jet bundle $J^n(\hat{\mathcal{Q}}|\mathbb{R})$ there exist an important class of distinguished 1-forms, namely the family of *the 1-forms that vanish identically on the tangent vectors of the curves of jets $\hat{\gamma}^{(n)}(t)$ of the parameterised graphs $\hat{\gamma}(t) = (t, q^i(t), \mathbf{h}_\beta^i(t), \mathbf{h}'_\beta^i(t), \mathbf{h}''_\beta^i(t), \lambda(t), \mu(t))$* .

It is known that such distinguished 1-forms are precisely those that are pointwise linear combinations of the 1-forms

$$\begin{aligned}
\omega_{(\delta)}^i &:= dq_{(\delta)}^i - q_{(\delta+1)}^i dt, \\
\varpi_{\beta(\delta)}^i &:= d\mathbf{h}_{\beta(\delta)}^i - \mathbf{h}_{\beta(\delta+1)}^i dt, \quad \varpi'_{\beta(\delta)}{}^i := d\mathbf{h}'_{\beta(\delta)}{}^i - \mathbf{h}'_{\beta(\delta+1)}{}^i dt, \quad i = 1, \dots, N, \\
\varpi''_{\beta(\delta)}{}^i &:= d\mathbf{h}''_{\beta(\delta)}{}^i - \mathbf{h}''_{\beta(\delta+1)}{}^i dt, \quad \delta = 0, \dots, n-1, \\
\varpi_{(\delta)}^{\{\lambda\}} &:= d\lambda_{(\delta)} - \lambda_{(\delta+1)} dt, \quad \varpi_{(\delta)}^{\{\mu\}} := d\mu_{(\delta)} - \mu_{(\delta+1)} dt.
\end{aligned} \tag{5.20}$$

Since these 1-forms vanish identically on the tangent vectors of the curves $\widehat{\gamma}^{(n)}(t)$, the value of the integral (5.19) does not change if α is replaced by any other 1-form

$$\alpha' = \alpha + \mathbb{P}_i^{(\delta)} \omega_{(\delta)}^i + \mathbb{Q}_i^{(\delta)\beta} \varpi_{\beta(\delta)}^i + \mathbb{Q}'_i^{(\delta)\beta} \varpi'_{\beta(\delta)}{}^i + \mathbb{Q}''_i^{(\delta)\beta} \varpi''_{\beta(\delta)}{}^i + \mathbb{L}^{(\delta)} \varpi_{(\delta)}^{\{\lambda\}} + \mathbb{M}^{(\delta)} \varpi_{(\delta)}^{\{\mu\}} \tag{5.21}$$

for some arbitrary choices of smooth functions $\mathbb{P}_i^{(\delta)}, \mathbb{Q}_i^{(\delta)\beta}, \mathbb{Q}'_i^{(\delta)\beta}, \mathbb{Q}''_i^{(\delta)\beta}, \mathbb{L}^{(\delta)}, \mathbb{M}^{(\delta)}$ of the points of $J^n(\widehat{\mathcal{Q}}|\mathbb{R}) \times K$. Following the terminology used in [10], we say that any such α' is *variationally equivalent* to α . The invariance of (5.19) under replacements with variationally equivalent 1-forms might be considered as a sort of “invariance under gauge transformations” of the cost functional.

A particular choice for the $\mathbb{P}_i^{(\delta)}, \mathbb{Q}_i^{(\delta)\beta}, \mathbb{Q}'_i^{(\delta)\beta}$, etc. yields to the following 1-form.

Definition 5.1. The *controlled Poincaré-Cartan form associated with \widehat{L}* is the 1-form on $J^n(\widehat{\mathcal{Q}}|\mathbb{R}) \times K$ defined by

$$\begin{aligned}
\alpha^{PC} &= \widehat{L} dt + \sum_{\delta=1}^r \sum_{\varepsilon=0}^{\delta-1} (-1)^\varepsilon \frac{d^\varepsilon}{dt^\varepsilon} \left(\frac{\partial \widehat{L}}{\partial q_{(\delta)}^i} \right) \omega_{(\delta-(\varepsilon+1))}^i + \\
&\quad + \lambda \sum_{\substack{1 \leq i \leq N \\ 0 \leq \beta \leq r-1}} \left(\mathbf{h}_{\beta(1)}^i \varpi_{\beta(0)}^i - \mathbf{h}_{\beta(2)}^i \varpi_{\beta(1)}^i - \mathbf{h}''_{\beta(2)}{}^i \varpi''_{\beta(1)}{}^i \right) + \lambda \varpi_{(0)}^{\{\mu\}} = \\
&= \left(\lambda \left(\mu_{(1)} + L + \frac{1}{2} \sum_{\substack{1 \leq i \leq N \\ 0 \leq \beta \leq r-1}} \left((\mathbf{h}_{\beta(1)}^i)^2 - (\mathbf{h}'_{\beta(2)}{}^i)^2 - (\mathbf{h}''_{\beta(2)}{}^i)^2 \right) \right) + \right. \\
&\quad \left. + \sum_{\substack{1 \leq i \leq N \\ 0 \leq \beta \leq r-1}} \left(\frac{1}{2} (\mathbf{h}_\beta^i)^2 + \frac{\pi^4}{32T^4} (\mathbf{h}'_\beta{}^i)^2 + \frac{\pi^4}{32T^4} (\mathbf{h}''_\beta{}^i)^2 \right) + \frac{dC}{dt} \right) dt + \tag{5.22} \\
&\quad + \lambda \sum_{\delta=1}^r \sum_{\varepsilon=0}^{\delta-1} (-1)^\varepsilon \frac{d^\varepsilon}{dt^\varepsilon} \left(\frac{\partial \left(L + \frac{dC}{dt} \right)}{\partial q_{(\delta)}^i} \right) \omega_{(\delta-(\varepsilon+1))}^i + \\
&\quad + \lambda \sum_{\substack{1 \leq i \leq N \\ 0 \leq \beta \leq r-1}} \left(\mathbf{h}_{\beta(1)}^i \varpi_{\beta(0)}^i - \mathbf{h}'_{\beta(2)}{}^i \varpi_{\beta(1)}^i - \mathbf{h}''_{\beta(2)}{}^i \varpi''_{\beta(1)}{}^i + \right. \\
&\quad \left. + \mathbf{h}'_{\beta(3)}{}^i \varpi_{\beta(0)}^i + \mathbf{h}''_{\beta(3)}{}^i \varpi''_{\beta(0)}{}^i \right) + \lambda \varpi_{(0)}^{\{\mu\}}.
\end{aligned}$$

Since (5.22) is variationally equivalent to α , we may safely replace α by α^{PC} in (5.19). And, in fact, such a substitution is the analogue of what is done in [3] for the classical

Mayer problems, where the 1-form

$$\alpha = \left(\sum_{i=1}^N p_i (\dot{x}^i - f^i(t, x, u(t))) + \sum_{i=1}^N \frac{\partial C}{\partial x^i} \Big|_{(t, x(t))} \dot{x}^i(t) \right) dt$$

is replaced by the 1-form

$$\begin{aligned} \alpha^{PC} &= \left(\sum_{i=1}^N p_i (\dot{x}^i - f^i(t, x, u(t))) + \sum_{i=1}^N \frac{\partial C}{\partial x^i} \Big|_{(t, x(t))} \dot{x}^i(t) \right) dt + \sum_{i=1}^N p_i (dx^i - \dot{x}^i dt) = \\ &= \sum_{i=1}^N p_i dx^i - p_i f^i(t, x, u(t)) dt + \sum_{i=1}^N \frac{\partial C}{\partial x^i} \Big|_{(t, x(t))} \dot{x}^i(t) dt . \end{aligned} \quad (5.23)$$

The convenience of considering the 1-form (5.23) and, in more general situations, the controlled Poincaré-Cartan forms (5.22) comes from the special feature that is described in the next lemma and which will be exploited in the proof of our first main result.

Lemma 5.2. *The differential $d\alpha^{PC}$ of the controlled Poincaré-Cartan form has the form*

$$\begin{aligned} d\alpha^{PC} &= E(\widehat{L})_i \omega_{\beta(0)}^i \wedge dt + \\ &+ E(\widehat{L})_i^\beta \varpi_{\beta(0)}^i \wedge dt + E(\widehat{L})_i^{\beta} \varpi_{\beta(0)}^i \wedge dt + E(\widehat{L})_i^{\beta} \varpi_{\beta(0)}^i \wedge dt + \\ &+ E(\widehat{L})_{\{\lambda\}} \varpi_{(0)}^{\{\lambda\}} \wedge dt + E(\widehat{L})_{\{\mu\}} \varpi_{(0)}^{\{\mu\}} \wedge dt + \frac{\partial \widehat{L}}{\partial u^a} du^a \wedge dt + \\ &+ \text{linear combinations of wedges of pairs of 1-forms of the list (5.20)} \end{aligned} \quad (5.24)$$

Proof. The claim is an immediate consequence of a general fact concerning the classes of variationally equivalent 1-forms on jets spaces (see e.g. [10, Prop. A2]). For reader's convenience, we present here a direct proof. For simplicity of notation, from now on we denote any tuple of coordinates $(q^i, \mathbf{h}_\beta^i, \mathbf{h}''_\beta^i, \lambda, \mu)$ for $\widehat{\mathcal{Q}}$ just by $y = (y^\ell)$, where the index ℓ ranges between 1 and $\widehat{N} = N(3r + 1) + 2$. Accordingly, the coordinates of the whole jet space are denoted by t and by $y_{(\delta)}^\ell$, $0 \leq \delta \leq n$, and the associated 1-forms of the list (5.20) are briefly indicated as

$$\omega_{(\delta)}^\ell = dy_{(\delta)}^\ell - y_{(\delta+1)}^\ell dt \quad \text{with } 1 \leq \ell \leq \widehat{N} \text{ and } 1 \leq \delta \leq n-1 . \quad (5.25)$$

Note that, being $dt \wedge dt = 0$ and $d(dy_{(\delta)}^\ell) = 0$, for each integer $1 \leq \delta \leq n-1$,

$$dy_{(\delta)}^\ell \wedge dt = \omega_{(\delta)}^\ell \wedge dt = -d\omega_{(\delta-1)}^\ell . \quad (5.26)$$

Hence for any function $f : J^n(\widehat{\mathcal{Q}}|\mathbb{R}) \times K \rightarrow \mathbb{R}$ of actual order $\widehat{r} \leq n-1$

$$\begin{aligned} f dy_{(\delta)}^\ell \wedge dt &= -f d\omega_{(\delta-1)}^\ell = \\ &= -d(f\omega_{(\delta-1)}^\ell) + \frac{\partial f}{\partial t} dt \wedge \omega_{(\delta-1)}^\ell + \sum_{\delta'=0}^{\widehat{r}} \frac{\partial f}{\partial y_{(\delta')}^m} dy_{(\delta')}^m \wedge \omega_{(\delta-1)}^\ell = \\ &= -d(f\omega_{(\delta-1)}^\ell) + \frac{\partial f}{\partial t} dt \wedge \omega_{(\delta-1)}^\ell + \sum_{\delta'=0}^{\widehat{r}} \frac{\partial f}{\partial y_{(\delta')}^m} \omega_{(\delta')}^m \wedge \omega_{(\delta-1)}^\ell + \sum_{\delta'=0}^{\widehat{r}} \frac{\partial f}{\partial y_{(\delta')}^m} y_{(\delta')}^m dt \wedge \omega_{(\delta-1)}^\ell = \\ &= -d(f\omega_{(\delta-1)}^\ell) - \frac{df}{dt} dy_{(\delta-1)}^\ell \wedge dt \quad \text{modulo terms of the form } \omega_{(\eta)}^r \wedge \omega_{(\zeta)}^s \end{aligned} \quad (5.27)$$

If $\delta - 1 \geq 1$, we may iterate and apply this identity to the term $-\frac{df}{dt} dy_{(\delta-1)}^\ell \wedge dt$ which appear in the right hand side of such identity. In this way we get that

$$f dy_{(\delta)}^\ell \wedge dt = -d \left(f \omega_{(\delta-1)}^\ell \right) + d \left(\frac{df}{dt} \omega_{(\delta-2)}^\ell \right) + \frac{d^2 f}{dt^2} dy_{(\delta-2)}^\ell \wedge dt \pmod{\omega_{(\eta)}^r \wedge \omega_{(\zeta)}^s}. \quad (5.28)$$

If $\delta - 2 \geq 1$, we may again apply (5.27) to the term $\frac{d^2 f}{dt^2} y_{(\delta-2)}^\ell \wedge dt$ and so on. After δ iterations of such use of (5.27), we end up with

$$f dy_{(\delta)}^\ell \wedge dt = -d \left(\sum_{\varepsilon=0}^{\delta-1} (-1)^\varepsilon \frac{d^\varepsilon f}{dt^\varepsilon} \omega_{(\delta-1-\varepsilon)}^\ell \right) + (-1)^\delta \frac{d^\delta f}{dt^\delta} \omega_{(0)}^\ell \wedge dt \pmod{\omega_{(\eta)}^r \wedge \omega_{(\zeta)}^s} \quad (5.29)$$

Applying (5.29) to each term $\frac{\partial \widehat{L}}{\partial y_\delta^\ell} dy_{(\delta)}^\ell \wedge dt$, $\delta \geq 1$, appearing in the exterior differential $d(\widehat{L}dt)$, we obtain

$$\begin{aligned} d(\widehat{L}dt) &= \frac{\partial \widehat{L}}{\partial y_{(0)}^\ell} dy_{(0)}^\ell \wedge dt + \sum_{\delta=1}^r \frac{\partial \widehat{L}}{\partial y_{(\delta)}^\ell} dy_{(\delta)}^\ell \wedge dt = \\ &= \frac{\partial \widehat{L}}{\partial y_{(0)}^\ell} \omega_{(0)}^\ell \wedge dt - \sum_{\delta=1}^r d \left(\sum_{\varepsilon=0}^{\delta-1} (-1)^\varepsilon \frac{d^\varepsilon}{dt^\varepsilon} \frac{\partial \widehat{L}}{\partial y_{(\delta)}^\ell} \omega_{(\delta-1-\varepsilon)}^\ell \right) + \\ &\quad + \sum_{\delta=1}^r (-1)^{\delta+1} \frac{d^\delta}{dt^\delta} \frac{\partial \widehat{L}}{\partial y_{(\delta)}^\ell} \omega_{(0)}^\ell \wedge dt \pmod{\omega_{(\eta)}^r \wedge \omega_{(\zeta)}^s} = \\ &= -d \left(\sum_{\delta=1}^r \sum_{\varepsilon=0}^{\delta-1} (-1)^\varepsilon \frac{d^\varepsilon}{dt^\varepsilon} \frac{\partial \widehat{L}}{\partial y_{(\delta)}^\ell} \omega_{(\delta-(\varepsilon+1))}^\ell \right) + E(\widehat{L})_\ell \omega_{(0)}^\ell \wedge dt \pmod{\omega_{(\eta)}^r \wedge \omega_{(\zeta)}^s} \end{aligned} \quad (5.30)$$

Note that, in the simplified notation used in this proof, the 1-form α^{PC} is nothing but

$$\alpha^{PC} = \widehat{L}dt + \sum_{\delta=1}^r \left(\sum_{\varepsilon=0}^{\delta-1} (-1)^\varepsilon \frac{d^\varepsilon}{dt^\varepsilon} \frac{\partial \widehat{L}}{\partial y_{(\delta)}^\ell} \omega_{(\delta-(\varepsilon+1))}^\ell \right).$$

From this and (5.30), the lemma follows immediately. \square

6. THE PRINCIPLE OF MINIMAL LABOUR

We are now ready to prove our first main result, the Principle of Minimal Labour for generalised Mayer problems. As we mentioned in Sect. 2.2, this is reached by: (i) first proving a generalisation of the identity (2.7) for the problem associated with the modified defining triple $(\widehat{\mathcal{K}}, \widehat{L}, \widehat{C})$ and then (ii) deriving a corresponding identity for the original problem, determined by (\mathcal{K}, L, C) . These two identities are proven in Sect. 6.1 and Sect. 6.2, respectively.

6.1. The homotopy formula for the problem defined by the triple $(\widehat{\mathcal{K}}, \widehat{L}, \widehat{C})$.

Let $\widehat{U}_o = (u_o(t), \widehat{\sigma}_o(t))$ be a fixed element in $\widehat{\mathcal{K}}$ and $\widehat{\gamma}^{(\widehat{U}_o)} : [0, T] \rightarrow [0, T] \times \widehat{\mathcal{Q}}$ the corresponding $\widehat{\mathcal{K}}$ -controlled curve. We call a smooth 1-parameter family $\widehat{F}(\cdot, s)$, $s \in [0, 1]$, of $\widehat{\mathcal{K}}$ -controlled curves with initial curve $\widehat{F}(\cdot, 0) = \widehat{\gamma}^{(\widehat{U}_o)}$ a $\widehat{\mathcal{K}}$ -controlled variation of $\widehat{\gamma}^{(\widehat{U}_o)}$. More precisely, a $\widehat{\mathcal{K}}$ -controlled variation \widehat{F} is a smooth homotopy of the form

$$\widehat{F} : [0, T] \times [0, 1] \longrightarrow \widehat{\mathcal{Q}}, \quad \widehat{F}(t, s) = \widehat{\gamma}^{(\widehat{U}(s))}(t), \quad (6.1)$$

where $\widehat{U}(s) = (u(\cdot, s), \widehat{\sigma}(s))$, $s \in [0, 1]$, is a smooth curve in $\widehat{\mathcal{K}}$ starting from $\widehat{U}(0) = \widehat{U}_o$. For each \widehat{F} , we consider the corresponding homotopy $\widehat{F}^{(n)}$ in $J^n(\widehat{\mathcal{Q}}|\mathbb{R}) \times K$ defined by

$$\widehat{F}^{(n)} : [0, T] \times [0, 1] \longrightarrow J^n(\widehat{\mathcal{Q}}|\mathbb{R}) \times K, \quad \widehat{F}^{(n)}(t, s) := (j_t^n(\widehat{F}(\cdot, s)), u(t, s)).$$

Note that, for each fixed s_o , the curve $t \mapsto j_t^n(\widehat{F}(\cdot, s_o))$ is nothing but the curve $\widehat{\gamma}^{(\widehat{U}(s_o))^{(n)}}$ of the n -th order jets of the curve $\widehat{\gamma}^{(\widehat{U}(s_o))}$.

Given a $\widehat{\mathcal{K}}$ -controlled variation \widehat{F} , we denote by $X_{\widehat{F}}$ and $Y_{\widehat{F}}$ the vector fields – defined just at the points of the surface $\widehat{\mathcal{S}} := \widehat{F}^{(n)}([0, T] \times [0, 1])$ – which are determined by considering the infinitesimal variations of the first parameter t and of the second parameter s , respectively. More precisely, $X_{\widehat{F}}$ and $Y_{\widehat{F}}$ are the vector fields at the points of $\widehat{\mathcal{S}}$

$$X_{\widehat{F}}|_{\widehat{F}^{(n)}(t,s)} := \widehat{F}_*^{(n)} \left(\left. \frac{\partial}{\partial t} \right|_{(t,s)} \right) = \left. \frac{\partial \widehat{F}^{(n)}}{\partial t} \right|_{(t,s)}, \quad (6.2)$$

$$Y_{\widehat{F}}|_{\widehat{F}^{(n)}(t,s)} := \widehat{F}_*^{(n)} \left(\left. \frac{\partial}{\partial s} \right|_{(t,s)} \right) = \left. \frac{\partial \widehat{F}^{(n)}}{\partial s} \right|_{(t,s)}. \quad (6.3)$$

We remark that, for each fixed $s_o \in [0, 1]$,

- (1) The restriction of $X_{\widehat{F}}$ to the trace of the curve

$$t \mapsto \widehat{F}^{(n)}(t, s_o) = \left(\widehat{\gamma}^{(\widehat{U}(s_o))^{(n)}}(t), u(t, s_o) \right) = \left(j_t^n \widehat{F}(\cdot, s_o), u(t, s_o) \right)$$

coincides with the family of the tangent vectors of such a curve.

- (2) The restriction of $Y_{\widehat{F}}$ to the trace of the same curve is the Jacobi vector field corresponding to the (infinitesimal) variation $\varepsilon \rightarrow \widehat{F}^{(n)}(\cdot, s_o + \varepsilon)$ of $\widehat{F}^{(n)}(\cdot, s_o)$.

From (1), (2) and (5.16), we have that the vector fields $X_{\widehat{F}}$ and $Y_{\widehat{F}}$ must have the form

$$\begin{aligned} X_{\widehat{F}} &= \frac{\partial}{\partial t} + X_{(\beta)}^i \frac{\partial}{\partial q_{(\beta)}^i} + X_{(\beta)}^{\{\mu\}} \frac{\partial}{\partial \mu_{(\beta)}} + X^a \frac{\partial}{\partial u^a} + \\ &\quad + X_{\alpha(\beta)}^i \frac{\partial}{\partial \mathbf{h}_{\alpha(\beta)}^i} + X_{\alpha(\beta)}^{\prime i} \frac{\partial}{\partial \mathbf{h}_{\alpha(\beta)}^{\prime i}} + X_{\alpha(\beta)}^{\prime\prime i} \frac{\partial}{\partial \mathbf{h}_{\alpha(\beta)}^{\prime\prime i}}, \\ Y_{\widehat{F}} &= Y_{(\beta)}^i \frac{\partial}{\partial q_{(\beta)}^i} + Y_{(\beta)}^{\{\mu\}} \frac{\partial}{\partial \mu_{(\beta)}} + Y^a \frac{\partial}{\partial u^a} + \\ &\quad + Y_{\alpha(\beta)}^i \frac{\partial}{\partial \mathbf{h}_{\alpha(\beta)}^i} + Y_{\alpha(\beta)}^{\prime i} \frac{\partial}{\partial \mathbf{h}_{\alpha(\beta)}^{\prime i}} + Y_{\alpha(\beta)}^{\prime\prime i} \frac{\partial}{\partial \mathbf{h}_{\alpha(\beta)}^{\prime\prime i}} \end{aligned} \quad (6.4)$$

for appropriate smooth real functions $X_{(\beta)}^i$, $X_{(\beta)}^{\{\mu\}}$, etc., defined *only* at the points of $\widehat{\mathcal{S}}$.

Theorem 6.1 (Homotopy Formula - First Version). *Let $\widehat{U}_0, \widehat{U}_1 \in \widehat{\mathcal{K}}$ be the endpoints of a smooth curve $\widehat{U}(s) \in \widehat{\mathcal{K}}$, $s \in [0, 1]$, and $\widehat{\gamma} := \widehat{\gamma}^{(\widehat{U}_0)}$, $\widehat{\gamma}' := \widehat{\gamma}^{(\widehat{U}_1)}$ the $\widehat{\mathcal{K}}$ -controlled curves corresponding to $\widehat{U}_0, \widehat{U}_1$, with terminal costs $C_0 := C(j_{t=T}^{n-1}(\widehat{\gamma}))$ and $C_1 := C(j_{t=T}^{n-1}(\widehat{\gamma}'))$, respectively. Furthermore*

(i) *for any jet $j_t^n(\widehat{\gamma}) \in J^n(\widehat{\mathcal{Q}}|\mathbb{R})$, let $\mathcal{P}_{j_t^n(\widehat{\gamma})}$ be the function on K defined by*

$$\mathcal{P}_{j_t^n(\widehat{\gamma})} : K \longrightarrow \mathbb{R}, \quad \mathcal{P}_{j_t^n(\widehat{\gamma})}(u^a) := -L(j_t^n(\widehat{\gamma}), u^a); \quad (6.5)$$

(ii) *let $\widehat{\mu} : [0, T] \times [0, 1] \rightarrow \mathbb{R}$ be the function defined by $\widehat{\mu}(t, s) := \mu^{(\widehat{U}(s))}(t)$, where $\mu^{(\widehat{U}(s))}(t)$ is the value at t of the μ -component of the $\widehat{\mathcal{K}}$ -controlled curve*

$$\widehat{\gamma}^{(\widehat{U}(s))}(t) = (t, q^{(\widehat{U}(s))i}(t), \mathbf{h}_\beta^{(\widehat{U}(s))i}(t), \dots, \lambda^{(\widehat{U}(s))}(t) = 1, \mu^{(\widehat{U}(s))}(t)).$$

Then,

$$\begin{aligned} C_1 - C_0 = & - \int_0^T \left(\int_0^1 \left(Y^a \frac{\partial \mathcal{P}_{\widehat{\gamma}^{(\widehat{U}(s))}(n)}}{\partial u^a} \Big|_{u(s)(t)} - \frac{\partial^2 \widehat{\mu}}{\partial t \partial s} \Big|_{(t,s)} - \right. \right. \\ & \left. \left. - \sum_{\beta=1}^{r-1} \sum_{i=1}^N \frac{\partial^2}{\partial t \partial s} \left(\int_0^s \mathbf{h}_{\beta(3)}^i Y_{\beta(0)}^i + \mathbf{h}_{\beta(3)}''^i Y_{\beta(0)}''^i \right) \Big|_{\widehat{F}^{(n)}(t,v)} \right) dv \Big) dt \end{aligned} \quad (6.6)$$

where Y^a , $Y_{\beta(0)}^i$, $Y_{\beta(0)}''^i$ are the u^a -, \mathbf{h}_β^i - and $\mathbf{h}_\beta''^i$ - components, respectively, of the vector field $Y_{\widehat{F}}$ defined in (6.4), associated with the $\widehat{\mathcal{K}}$ -controlled variation \widehat{F} determined by the $\widehat{U}(s)$, $s \in [0, 1]$,

Remark 6.2. From the definition of the functions $\mathbf{h}_\beta^i(t)$, $\mathbf{h}_\beta^i(t)$ and $\mathbf{h}_\beta''^i(t)$ and the Euler-Lagrange equation $E(\widehat{L})_{\{\lambda\}} = 0$, it follows immediately that the second summand in the right hand side of (6.6) is uniquely determined by the $\widehat{\mathcal{K}}$ -controlled curves $t \mapsto \gamma^{(U(s))}(t) = (t, q^{(U(s))i}(t))$ in $[0, T] \times \mathcal{Q}$. It is also simple to check that the same is true for the first summand as well. These two facts will be used in the next subsection.

Proof. Consider the embedded surface $\widehat{\mathcal{S}} := \widehat{F}^{(n)}([0, T] \times [0, 1])$ and the vector fields $X_{\widehat{F}}$, $Y_{\widehat{F}}$ defined in (6.2) and (6.3). From (5.19), the property (1) of $X_{\widehat{F}}$ and the fact that α^{PC} is a 1-form which is variationally equivalent to α , we have that for each curve $\widehat{\gamma}^{(\widehat{U}(s))}$, $s \in [0, 1]$,

$$\begin{aligned} \int_0^T \left(\widehat{F}^{(n)*}(\alpha^{PC}) \right) \Big|_{(t,s)} \left(\frac{\partial}{\partial t} \right) dt &= \int_0^T \alpha^{PC} (X_{\widehat{F}}) \Big|_{\widehat{F}^{(n)}(t,s)} dt = \\ &= \int_0^T \alpha (X_{\widehat{F}}) \Big|_{\widehat{F}^{(n)}(t,s)} dt = C(j_{t=T}^n(\widehat{\gamma}^{(\widehat{U}(s))})). \end{aligned}$$

This implies that

$$\begin{aligned} C_0 - C_1 &= C(j_{t=T}^n(\widehat{\gamma}^{(\widehat{U}(0))})) - C(j_{t=T}^n(\widehat{\gamma}^{(\widehat{U}(1))})) = \\ &= \int_0^T \left(\widehat{F}^{(n)*}(\alpha^{PC}) \right) \left(\frac{\partial}{\partial t} \right) \Big|_{(t,0)} dt - \int_0^T \left(\widehat{F}^{(n)*}(\alpha^{PC}) \right) \left(\frac{\partial}{\partial t} \right) \Big|_{(t,1)} dt. \end{aligned} \quad (6.7)$$

On the other hand, by property (2) of $Y_{\widehat{F}}$ and the assumptions on the initial data of the $\mathbf{h}_{\beta}^i(t)$, $\mathbf{h}'_{\beta}(t)$ and $\mathbf{h}''_{\beta}(t)$, described in Sect. 5.1, we have that

$$Y_{(0)}^{\{\mu\}}|_{\widehat{F}^{(n)}(0,s)} = Y_{\beta(0)}^{\prime i}|_{\widehat{F}^{(n)}(0,s)} = Y_{\beta(0)}^{\prime\prime i}|_{\widehat{F}^{(n)}(0,s)} = Y_{\beta(1)}^{\prime i}|_{\widehat{F}^{(n)}(0,s)} = Y_{\beta(1)}^{\prime\prime i}|_{\widehat{F}^{(n)}(0,s)} = 0 ,$$

$$Y_{\beta(0)}^i|_{\widehat{F}^{(n)}(0,s)} = Y_{(\beta)}^i|_{\widehat{F}^{(n)}(0,s)} .$$

From this we obtain

$$\begin{aligned} & \int_0^1 \iota_{\frac{\partial}{\partial s}} \left(\widehat{F}^{(n)*}(\alpha^{PC}) \right) \Big|_{(0,s)} ds = \int_0^1 \alpha^{PC} (Y_{\widehat{F}}) |_{\widehat{F}^{(n)}(0,s)} ds = \\ & = \sum_{\beta=1}^{r-1} \sum_{i=1}^N \int_0^1 \left(\left(\sum_{\substack{1 \leq \delta \leq r \\ 0 \leq \varepsilon \leq \delta \\ \delta - \varepsilon - 1 = \beta}} (-1)^\varepsilon \frac{d^\varepsilon}{dt^\varepsilon} \left(\frac{\partial}{\partial q_{(\delta)}^i} \left(L + \frac{dC}{dt} \right) \right) \right) Y_{(\beta)}^i + \mathbf{h}_{\beta(1)}^i Y_{\beta(0)}^i \right) \Big|_{\widehat{F}^{(n)}(0,s)} ds = \\ & = \sum_{\beta=1}^{r-1} \sum_{i=1}^N \int_0^1 \left(\left(\sum_{\substack{1 \leq \delta \leq r \\ 0 \leq \varepsilon \leq \delta \\ \delta - \varepsilon - 1 = \beta}} (-1)^\varepsilon \frac{d^\varepsilon}{dt^\varepsilon} \left(\frac{\partial}{\partial q_{(\delta)}^i} \left(L + \frac{dC}{dt} \right) \right) + \mathbf{h}_{\beta(1)}^i \right) Y_{(\beta)}^i \right) \Big|_{\widehat{F}^{(n)}(0,s)} ds \stackrel{(5.4)}{=} 0 . \end{aligned} \quad (6.8)$$

Remark 6.3. This identity is precisely what we planned to obtain by introducing the functions $\mathbf{h}_{\alpha}^i(t)$ and the corresponding terms in \widetilde{L} and \widehat{L} . In fact, the integral (6.8) is precisely the integral of α^{PC} along the first “vertical side” of \mathcal{S} (see Sect. 2.1.3 and 2.2) and the functions $\mathbf{h}_{\alpha}^i(t)$ have been chosen so that such an integral is always vanishing.

Let us now consider the similar line integral for $t = T$. We have

$$\begin{aligned} & \int_0^1 \iota_{\frac{\partial}{\partial s}} \left(\widehat{F}^{(n)*}(\alpha^{PC}) \right) \Big|_{(T,s)} ds = \int_0^1 \alpha^{PC} (Y_{\widehat{F}}) |_{\widehat{F}^{(n)}(T,s)} ds = \\ & = \sum_{\beta=1}^{r-1} \sum_{i=1}^N \int_0^1 \left\{ \left(\sum_{\substack{1 \leq \delta \leq r \\ 0 \leq \varepsilon \leq \delta \\ \delta - \varepsilon - 1 = \beta}} (-1)^\varepsilon \frac{d^\varepsilon}{dt^\varepsilon} \left(\frac{\partial (L + \frac{dC}{dt})}{\partial q_{(\delta)}^i} \right) \right) Y_{(\beta)}^i + \right. \\ & \left. + \left(\mathbf{h}_{\beta(1)}^i Y_{\beta(0)}^i - \mathbf{h}'_{\beta(2)} Y'_{\beta(1)} - \mathbf{h}''_{\beta(2)} Y''_{\beta(1)} + \mathbf{h}'_{\beta(3)} Y'_{\beta(0)} + \mathbf{h}''_{\beta(3)} Y''_{\beta(0)} \right) \right\} \Big|_{\widehat{F}^{(n)}(T,s)} ds + \\ & \quad + \int_0^1 Y_{(0)}^{\{\mu\}} \Big|_{\widehat{F}^{(n)}(T,s)} ds . \end{aligned} \quad (6.9)$$

By property (2) of $Y_{\widehat{F}}$ and the boundary values (5.9) and (5.10) of the curves $\mathbf{h}_{\beta}^i(t)$ and $\mathbf{h}''_{\beta}(t)$, we have that

$$\mathbf{h}_{\beta(2)}^{\prime i} \Big|_{\widehat{F}^{(n)}(T,s)} = \sum_{\substack{1 \leq \delta \leq r \\ 0 \leq \varepsilon \leq \delta \\ \delta - \varepsilon - 1 = \beta}} (-1)^\varepsilon \frac{d^\varepsilon}{dt^\varepsilon} \left(\frac{\partial (L + \frac{dC}{dt})}{\partial q_{(\delta)}^i} \right) \Big|_{j_{t=T}^{n-1}(q(t))} , \quad Y_{\beta(1)}^{\prime i} \Big|_{\widehat{F}^{(n)}(T,s)} = Y_{(\beta)}^i \Big|_{\widehat{F}^{(n)}(T,s)} ,$$

$$\mathbf{h}_{\beta(2)}^{\prime\prime i} \Big|_{\widehat{F}^{(n)}(T,s)} = \mathbf{h}_{\beta(1)}^i \Big|_{\widehat{F}^{(n)}(T,s)} , \quad Y_{\beta(1)}^{\prime\prime i} \Big|_{\widehat{F}^{(n)}(T,s)} = Y_{\beta(0)}^i \Big|_{\widehat{F}^{(n)}(T,s)} .$$

From this and (6.9), we get that

$$\begin{aligned} \int_0^1 \iota_{\frac{\partial}{\partial s}} \left(\widehat{F}^{(n)*}(\alpha^{PC}) \right) \Big|_{(T,s)} ds &= \\ &= \int_0^1 \left(\sum_{\beta=1}^{r-1} \sum_{i=1}^N \left(\mathbf{h}'_{\beta(3)} Y'_{\beta(0)} + \mathbf{h}''_{\beta(3)} Y''_{\beta(0)} \right) + Y_{(0)}^{\{\mu\}} \right) \Big|_{\widehat{F}^{(n)}(T,s)} ds . \end{aligned} \quad (6.10)$$

On the other hand, by the definition of the vector field $Y_{\widehat{F}}$,

$$Y_{(0)}^{\{\mu\}} \Big|_{\widehat{F}^{(n)}(T,s)} = \frac{\partial \widehat{F}^{\{\mu\}}}{\partial s} \Big|_{(T,s)} = \frac{\partial (\widehat{\gamma}(\widehat{U}(s)))^{\{\mu\}}}{\partial s} \Big|_{(T,s)} = \frac{\partial \widehat{\mu}(T,s)}{\partial s} .$$

Hence

$$\int_0^1 Y_{(0)}^{\{\mu\}} \Big|_{\widehat{F}^{(n)}(T,s)} ds = \int_0^1 \frac{\partial}{\partial s} \left(\int_0^T \frac{d\widehat{\mu}(t,s)}{dt} dt \right) = \iint_{[0,T] \times [0,1]} \frac{\partial^2 \widehat{\mu}}{\partial t \partial s} ds dt . \quad (6.11)$$

Remark 6.4. In analogy with what we mentioned in Remark 6.3, the purpose of the functions $\mathbf{h}'_{\beta}(t)$, $\mathbf{h}''_{\beta}(t)$ and of the corresponding terms in \widehat{L} was precisely to get (6.10), which simplifies the integral of α^{PC} along the second “vertical side” of \mathcal{S} .

From (6.7), (6.8), (6.10), (6.11) and the Stokes Theorem, we obtain that

$$\begin{aligned} C_0 - C_1 + \int_0^T \left(\int_0^1 \frac{\partial^2 \widehat{\mu}}{\partial t \partial s} \Big|_{(t,s)} + \right. \\ \left. + \sum_{\beta=1}^{r-1} \sum_{i=1}^N \frac{\partial}{\partial t} \left(\mathbf{h}'_{\beta(3)} Y'_{\beta(0)} + \mathbf{h}''_{\beta(3)} Y''_{\beta(0)} \right) \Big|_{\widehat{F}^{(n)}(t,s)} ds \right) dt = \\ = \int_{\partial([0,T] \times [0,1])} \left(\widehat{F}^{(n)*}(\alpha^{PC}) \right) \stackrel{\text{Stokes Thm.}}{=} \\ = \int_{[0,T] \times [0,1]} d \left(\widehat{F}^{(n)*}(\alpha^{PC}) \right) \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial s} \right) dt ds = \\ = \int_{[0,T] \times [0,1]} \left(\widehat{F}^{(n)*}(d\alpha^{PC}) \right) \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial s} \right) dt ds = \\ = \int_{[0,T] \times [0,1]} d\alpha^{PC}(X_{\widehat{F}}, Y_{\widehat{F}}) \Big|_{\widehat{F}^{(n)}(t,s)} dt ds , \end{aligned} \quad (6.12)$$

where, in the integral of the first line, the integration of $\widehat{F}^{(n)*}(\alpha^{PC})$ is performed along the usual counterclockwise parameterisation of $\partial([0, T] \times [0, 1])$.

At this point, it suffices to recall that the controlled Euler-Lagrange equations (5.16) are satisfied at all points of $\widehat{\mathcal{S}}$ and that the vectors $X_{\widehat{F}}$ are tangent vectors to curves of jets, determined by solutions to the controlled Euler-Lagrange equations. Indeed, due to this, (5.24) and the fact that each of the 1-forms (5.20) vanish identically on the vectors

$X_{\widehat{F}}$, we immediately have that ⁽⁷⁾

$$\begin{aligned} d\alpha^{PC}(X_{\widehat{F}}, Y_{\widehat{F}})|_{\widehat{F}^{(n)}(t,s)} &= \left(\frac{\partial \widehat{L}}{\partial u^a} du^a \wedge dt \right) (X_{\widehat{F}}, Y_{\widehat{F}})|_{\widehat{F}^{(n)}(t,s)} = \\ &= -\frac{\partial \widehat{L}}{\partial u^a} Y^a|_{\widehat{F}^{(n)}(t,s)} = -\frac{\partial L}{\partial u^a} Y^a|_{\widehat{F}^{(n)}(t,s)}. \end{aligned} \tag{6.13}$$

From this and (6.12), the conclusion follows. □

6.2. The Principle of Minimal Labour.

Let us now go back to our original triple (\mathcal{K}, L, C) and to the parameterised graphs $\gamma(t)$ in $\mathbb{R} \times \mathcal{Q}$. We have already observed that for any \mathcal{K} -controlled curve $\gamma^{(U)}(t) = (t, q^i(t))$ there is a uniquely associated element $\widehat{U} \in \widehat{\mathcal{K}}$ and a uniquely associated $\widehat{\mathcal{K}}$ -controlled curve of the form

$$\widehat{\gamma}^{(\widehat{U})} = (t, q^i(t), \mathbf{h}_\beta^i(t), \mathbf{h}''_{\beta}{}^i(t), \lambda = 1, \mu^{(\widehat{U})}(t)) ,$$

i.e. with the same components $q^i(t)$ of $\gamma^{(U)}$. On the basis of this, in what follows for any given $U \in \mathcal{K}$ we use the symbols \widehat{U} and $\widehat{\gamma}^{(\widehat{U})}$ to denote the uniquely associated element in $\widehat{\mathcal{K}}$ and the corresponding $\widehat{\mathcal{K}}$ -controlled curve, respectively.

Let $U_o = (u_o(t), \sigma_o(t))$ be a fixed element in \mathcal{K} and $\gamma^{(U_o)} : [0, T] \rightarrow [0, T] \times \mathcal{Q}$ the corresponding \mathcal{K} -controlled curve. Exactly as in the previous section, we may now consider a \mathcal{K} -controlled variation of this curve, i.e. a smooth 1-parameter family $F(\cdot, s)$, $s \in [0, 1]$, of \mathcal{K} -controlled curves with initial curve $F(\cdot, 0) = \gamma^{(U_o)}$. As we did in dealing with $\widehat{\mathcal{K}}$ -controlled variations, we denote by $F^{(n)}$ the homotopy in $J^n(\mathcal{Q}|\mathbb{R}) \times K$

$$F^{(n)} : [0, T] \times [0, 1] \longrightarrow J^n(\mathcal{Q}|\mathbb{R}) \times K , \quad F^{(n)}(t, s) := (j_t^{(n)}(F(\cdot, s)), u(t, s)) ,$$

determined by the homotopy of the n -th order jets of the curves $\gamma^{(U(s))} = F(\cdot, s)$ and the homotopy $u(t, s) = u^{(s)}(t)$ of control curves. We may also consider the vector fields

$$X_F|_{F^{(n)}(t,s)} := F_*^{(n)} \left(\frac{\partial}{\partial t} \Big|_{(t,s)} \right) = \frac{\partial F^{(n)}}{\partial t} \Big|_{(t,s)} , \tag{6.14}$$

$$Y_F|_{F^{(n)}(t,s)} := F_*^{(n)} \left(\frac{\partial}{\partial s} \Big|_{(t,s)} \right) = \frac{\partial F^{(n)}}{\partial s} \Big|_{(t,s)} , \tag{6.15}$$

defined at the points of the surface $\mathcal{S} := F^{(n)}([0, T] \times [0, 1]) \subset J^n(\mathcal{Q}|\mathbb{R}) \times K$. By construction, for each fixed $s_o \in [0, 1]$ we have that: (a) the restriction of X_F to the trace of the curve

$$t \mapsto \left(\gamma^{(U(s_o))^{(n)}}(t), u(t, s_o) \right) = (j_t^n(F(\cdot s_o)), u(t, s_o))$$

is the family of the tangent vectors of such a curve; (b) the restriction of Y_F to the trace of the same curve is the Jacobi vector field corresponding to the (infinitesimal) variation $\varepsilon \rightarrow F^{(n)}(\cdot, s_o + \varepsilon)$ of $F^{(n)}(\cdot, s_o)$; (c) the vector fields X_F and Y_F have the form

$$X_F = \frac{\partial}{\partial t} + X_{(\beta)}^i \frac{\partial}{\partial q_{(\beta)}^i} + X^a \frac{\partial}{\partial u^a} , \quad Y_F = Y_{(\beta)}^i \frac{\partial}{\partial q_{(\beta)}^i} + Y^a \frac{\partial}{\partial u^a} . \tag{6.16}$$

⁷This is precisely the point that motivated the substitution of the 1-form α by α^{PC} .

Finally, we have that the curve $U(s) \in \mathcal{K}$ determines uniquely a curve $\widehat{U}(s)$ in $\widehat{\mathcal{K}}$ and a corresponding $\widehat{\mathcal{K}}$ -controlled variation \widehat{F} starting from the curve $\widehat{\gamma} := \widehat{\gamma}^{(\widehat{U}_0)}$ and ending with the curve $\widehat{\gamma}' := \widehat{\gamma}^{(\widehat{U}_1)}$.

We are now ready to establish the following corollary of Theorem 6.1 and Remark 6.2.

Corollary 6.5 (Homotopy Formula). *Let $U_0, U_1 \in \mathcal{K}$ be the endpoints of a smooth curve $U(s) \in \mathcal{K}$, $s \in [0, 1]$, and $\gamma := \gamma^{(U_0)}$, $\gamma' := \gamma^{(U_1)}$ the \mathcal{K} -controlled curves corresponding to U_0, U_1 , with terminal costs $C_0 := C(j_{t=T}^{n-1}(\gamma))$ and $C_1 := C(j_{t=T}^{n-1}(\gamma'))$, respectively. Let also $\mathcal{P}_{j_t^n(\gamma)} : K \rightarrow \mathbb{R}$ and $\mu, \mu' : [0, T] \times [0, 1] \rightarrow \mathbb{R}$ be the real functions defined by*

$$\mathcal{P}_{j_t^n(\gamma)}(u^a) := -L(j_t^n(\gamma), u^a), \quad (6.17)$$

$$\begin{aligned} \mu(t, s) := & - \int_0^t \left(L + \frac{1}{2} \sum_{\substack{1 \leq i \leq N \\ 0 \leq \beta \leq r-1}} \left((\mathbf{h}_{\beta(1)}^i)^2 - (\mathbf{h}_{\beta(2)}^i)^2 - (\mathbf{h}_{\beta(2)}^{\prime i})^2 \right) \right. \\ & \left. + \sum_{\substack{1 \leq i \leq N \\ 0 \leq \beta \leq r-1}} \left(\frac{1}{2} (\mathbf{h}_{\beta}^i)^2 + \frac{\pi^4}{32T^4} (\mathbf{h}'_{\beta}^i)^2 + \frac{\pi^4}{32T^4} (\mathbf{h}''_{\beta}^i)^2 \right) \right) \Big|_{\gamma^{(U(s))}(\tau)} d\tau \end{aligned} \quad (6.18)$$

$$\mu'(t, s) := \mu(t, s) + \sum_{\beta=1}^{r-1} \sum_{i=1}^N \int_0^s \left(\mathbf{h}_{\beta(3)}^i Y_{\beta(0)}^i + \mathbf{h}_{\beta(3)}^{\prime i} Y_{\beta(0)}^{\prime i} \right) \Big|_{\widehat{F}^{(n)}(t, v)} dv \quad (6.19)$$

where we denoted by \widehat{F} the uniquely determined $\widehat{\mathcal{K}}$ -controlled variation described above and by $Y_{\beta(0)}^i, Y_{\beta(0)}^{\prime i}$ the components in the directions of the coordinates $\mathbf{h}_{\beta(0)}^i, \mathbf{h}_{\beta(0)}^{\prime i}$ of the vector field $Y_{\widehat{F}}$ defined in (6.3). Let us finally denote by Y_F the vector field in (6.16), associated with the \mathcal{K} -controlled variation F determined by the $U(s)$, $s \in [0, 1]$, Then:

(i) The difference between the terminal costs C_0 and C_1 is equal to

$$C_1 - C_0 = - \int_0^T \left(\int_0^1 \left(Y^a \frac{\partial \mathcal{P}_{\sigma_t^{(s)}}}{\partial u^a} \Big|_{u^{(s)}(t)} - \frac{\partial^2 \mu'}{\partial t \partial s} \Big|_{(t, s)} \right) ds \right) dt, \quad \sigma_t^{(s)} := j_t^n(\gamma^{(U(s))}); \quad (6.20)$$

(ii) For any $s_o \in [0, 1]$,

$$\alpha^{PC}(Y_{\widehat{F}}) \Big|_{\widehat{F}^{(n)}(T, s_o)} = \alpha^{PC}(Y_{\widehat{F}}) \Big|_{\widehat{F}^{(n)}(0, s_o)} + \int_0^T \left(\frac{\partial^2 \mu'}{\partial t \partial s} \Big|_{s=s_o} \right) dt. \quad (6.21)$$

Proof. By Theorem 6.1, Remark 6.2 and the fact that, due to the Euler-Lagrange equations, the (6.18) is nothing but the explicit expression for the μ -components of the curve $\widehat{\gamma}^{(\widehat{U}(s))}(t)$ (thus, of the function $\widehat{\mu}(t, s)$ of Theorem 6.1), claim (i) follows immediately.

We now prove (ii). Consider the embedded surface with boundary $\widehat{\mathcal{S}} := \widehat{F}^{(n)}([0, T] \times [0, 1])$ and the vector fields $X_{\widehat{F}}, Y_{\widehat{F}}$, which we defined above at the points of $\widehat{\mathcal{S}}$. From (5.19) and the fact that α^{PC} is variationally equivalent to α , we have that

$$\begin{aligned} C_0 - C_1 &= C(j_{t=T}^n(\gamma^{(U_0)})) - C(j_{t=T}^n(\gamma^{(U_1)})) = \\ &= \int_0^T \iota_{\frac{\partial}{\partial t}} \left(\widehat{F}^{(n)*}(\alpha^{PC}) \right) \Big|_{(t, 0)} dt - \int_0^T \iota_{\frac{\partial}{\partial t}} \left(\widehat{F}^{(n)*}(\alpha^{PC}) \right) \Big|_{(t, 1)} dt. \end{aligned} \quad (6.22)$$

On the other hand,

$$\begin{aligned} \int_0^1 \iota_{\frac{\partial}{\partial s}} \left(\widehat{F}^{(n)*}(\alpha^{PC}) \right) \Big|_{(0,s)} ds &= \int_0^1 \alpha^{PC} (Y_{\widehat{F}}) |_{\widehat{F}^{(n)}(0,s)} ds , \\ \int_0^1 \iota_{\frac{\partial}{\partial s}} \left(\widehat{F}^{(n)*}(\alpha^{PC}) \right) \Big|_{(T,s)} ds &= \int_0^1 \alpha^{PC} (Y_{\widehat{F}}) |_{\widehat{F}^{(n)}(T,s)} ds . \end{aligned} \quad (6.23)$$

From (6.22), (6.23) and the Stokes Theorem, we get that

$$\begin{aligned} C_0 - C_1 &= \int_{\partial([0,T] \times [0,1])} \left(\widehat{F}^{(n)*}(\alpha^{PC}) \right) \Big|_{(t,s)} + \\ &\quad + \int_0^1 \iota_{\frac{\partial}{\partial s}} \left(\widehat{F}^{(n)*}(\alpha^{PC}) \right) \Big|_{(0,s)} ds - \int_0^1 \iota_{\frac{\partial}{\partial s}} \left(\widehat{F}^{(n)*}(\alpha^{PC}) \right) \Big|_{(T,s)} ds = \\ &= \int_{[0,T] \times [0,1]} \left(\widehat{F}^{(n)*}(d\alpha^{PC}) \right) \Big|_{(t,s)} + \\ &\quad + \int_0^1 \alpha^{PC} (Y_{\widehat{F}}) |_{\widehat{F}^{(n)}(0,s)} - \int_0^1 \alpha^{PC} (Y_{\widehat{F}}) |_{\widehat{F}^{(n)}(T,s)} ds = \\ &= \int_{[0,T] \times [0,1]} d\alpha^{PC} (X_{\widehat{F}}, Y_{\widehat{F}}) |_{\widehat{F}^{(n)}(t,s)} dt ds + \\ &\quad + \int_0^1 \alpha^{PC} (Y_{\widehat{F}}) |_{\widehat{F}^{(n)}(0,s)} - \int_0^1 \alpha^{PC} (Y_{\widehat{F}}) |_{\widehat{F}^{(n)}(T,s)} ds . \end{aligned} \quad (6.24)$$

As in the proof of Theorem 6.1, we now observe that the vectors $X_{\widehat{F}}$ are the tangent vectors to curves of jets that are solutions to the Euler-Lagrange equations (5.16). Hence, by Lemma 5.2, at any point of $\mathcal{S} = \widehat{F}^{(n)}([0, T] \times [0, 1])$, we have that

$$d\alpha^{PC} (X_{\widehat{F}}, Y_{\widehat{F}}) = -Y^a \frac{\partial L}{\partial u^a} = Y^a \frac{\partial \mathcal{P}_{\sigma_t^{(s)}}}{\partial u^a} \Big|_{u^{(s)}(t)} . \quad (6.25)$$

From this and (6.24), it follows that

$$\begin{aligned} C_0 - C_1 &= \int_0^T \left(\int_0^1 \left(Y^a \frac{\partial \mathcal{P}_{\sigma_t^{(s)}}}{\partial u^a} \right) \Big|_{u^{(s)}(t)} ds \right) dt + \\ &\quad + \int_0^1 \alpha^{PC} (Y_{\widehat{F}}) |_{\widehat{F}^{(n)}(0,s)} - \int_0^1 \alpha^{PC} (Y_{\widehat{F}}) |_{\widehat{F}^{(n)}(T,s)} ds . \end{aligned} \quad (6.26)$$

This and (i) imply that

$$\int_0^1 \alpha^{PC} (Y_{\widehat{F}}) |_{\widehat{F}^{(n)}(T,s)} ds = \int_0^1 \alpha^{PC} (Y_{\widehat{F}}) |_{\widehat{F}^{(n)}(0,s)} ds + \int_0^1 \left(\int_0^T \frac{\partial^2 \mu'}{\partial t \partial s} dt \right) ds .$$

This identity holds not only for the \mathcal{K} -controlled homotopy F , but, for a fixed $s_o \in [0, 1]$ and a fixed sufficiently small $\varepsilon > 0$, it holds also for any other \mathcal{K} -controlled homotopy of the form

$$F^{(s_o, \varepsilon)}(t, \tau) := F(t, s_o(1 - \tau) + (s_o + \varepsilon)\tau) , \quad \tau \in [0, 1] ,$$

which interpolates between the curve $\gamma^{(U(s_o))}$ to the curve $\gamma^{(U(s_o+\varepsilon))}$. This implies that for any fixed choice of $s_o \in [0, 1)$ and any small $\varepsilon > 0$

$$\int_{s_o}^{s_o+\varepsilon} \alpha^{PC} (Y_{\widehat{F}}) |_{\widehat{F}^{(n)}(T,s)} ds = \int_{s_o}^{s_o+\varepsilon} \alpha^{PC} (Y_{\widehat{F}}) |_{\widehat{F}^{(n)}(0,s)} ds + \int_{s_o}^{s_o+\varepsilon} \left(\int_0^T \frac{\partial^2 \mu'}{\partial t \partial s} dt \right) ds.$$

From this and the continuity of all involved functions, claim (ii) follows. \square

The homotopy formula has the following immediate consequence, which is the first main result that was expected according to the road map.

Theorem 6.6 (Generalised Principle of Minimal Labour). *A necessary condition for a \mathcal{K} -controlled curve $\gamma_o := \gamma^{(U_o)}$ to be a solution to the considered generalised Mayer problem is that for any \mathcal{K} -controlled variation F with $F(t, 0) = \gamma_o$ one has that*

$$\int_0^T \left(\int_0^1 \left(Y^a \frac{\partial \mathcal{P}_{\sigma_t^{(s)}}}{\partial u^a} \Big|_{u^{(s)}(t)} - \frac{\partial^2 \mu'}{\partial t \partial s} \Big|_{(t,s)} \right) ds \right) dt \leq 0. \quad (6.27)$$

In case the considered problem is such that for any two \mathcal{K} -controlled curves $\gamma := \gamma^{(U)}$ and $\gamma' := \gamma^{(U')}$ there is a \mathcal{K} -controlled variation which has them as endpoints, the above is also a sufficient condition.

Notice that the above principle admits the following equivalent formulation.

Let $\gamma_o := \gamma^{(U_o)}$ be a fixed \mathcal{K} -controlled curve and for any \mathcal{K} -controlled variation F with $F(t, 0) = \gamma_o$ consider the real function $\mathcal{W}^F : [0, 1] \rightarrow \mathbb{R}$ defined by

$$\mathcal{W}^F(\delta) := \int_0^T \left(\int_0^\delta \left(Y^a \frac{\partial \mathcal{P}_{\sigma_t^{(s)}}}{\partial u^a} \Big|_{u^{(s)}(t)} - \frac{\partial^2 \mu'}{\partial t \partial s} \Big|_s \right) ds \right) dt. \quad (6.28)$$

Then $\gamma_o := \gamma^{(U_o)}$ is a solution to the Mayer problem determined by (\mathcal{K}, L, C) only if for any \mathcal{K} -variation as above

$$\mathcal{W}^F(\delta) \leq 0 \quad \text{for each } \delta \in [0, 1]. \quad (6.29)$$

It should be pointed out that an explicit check of (6.29) is expected to be quite hard in generic situations: it demands the study of the sign behaviour of the functions \mathcal{W}^F for *any* \mathcal{K} -controlled variation F of a candidate γ_o . Nonetheless it has a number of consequences, the most elementary one represented by an infinitesimal version of Theorem 6.6, the concluding result of this section.

Let $\gamma_o := \gamma^{(U_o)}$ be a \mathcal{K} -controlled curve associated with the pair $U_o = (u_o(t), \sigma_o)$ and denote by \mathcal{F} the full collection of the \mathcal{K} -controlled variations of γ_o . Let also denote by $\mathcal{J}ac^{(n)}$ the class of all vectors fields V in $TJ^n(\mathcal{Q}|\mathbb{R})$, defined just at the points of the trace of $\gamma_o^{(n)}(t)$, of the form $V|_{\gamma_o^{(n)}(t)} := Y_F|_{\gamma_o^{(n)}(t)}$ for some $F \in \mathcal{F}$. In other words, $\mathcal{J}ac^{(n)}$ is the family of all Jacobi vector fields of $\gamma_o^{(n)}(t)$, which are determined by the variations of curves of jets determined by the \mathcal{K} -controlled variations of γ_o .

Theorem 6.7. *A \mathcal{K} -controlled curve $\gamma_o := \gamma^{(U_o)}$ is a solution to the Mayer problem of (\mathcal{K}, L, C) only if the following two conditions hold for any $V \in \mathcal{J}ac^{(n)}$:*

$$(1) \quad dC \left(V|_{\gamma_o^{(n)}(T)} \right) \geq 0;$$

$$(2) \quad \int_0^T \left(Y^a \frac{\partial \mathcal{P}_{\sigma_i^{(s)}}}{\partial u^a} \Big|_{u_o(t)} - \frac{\partial^2 \mu'}{\partial t \partial s} \Big|_{s=0} \right) dt \leq 0.$$

Proof. (2) is an immediate consequence of (6.29). For (1), one can obtain it by just observing the function W^F coincides with the map $\delta \xrightarrow{W^F} C|_{F^{(n)}(T,0)} - C|_{F^{(n)}(T,\delta)}$. The inequality is then obtained taking the derivative of this expression at $\delta = 0$. \square

7. THE GENERALISED PONTRYAGIN MAXIMUM PRINCIPLE

In this section we want to show that the Principle of Minimal Labour yields a strict analogue of the classical PMP for a wide class of generalised Mayer problems of higher order differential constraints of variational type and without the usual restriction of fixed initial values. The class we consider is characterised by *differential constraints of normal type*, which are defined in the next subsection. Immediately after this, we introduce a generalised version of the classical Pontryagin needle variations and we finally prove the advertised result.

7.1. Differential constraints of normal type.

As usual, we consider a Mayer problem determined by a triple (\mathcal{K}, L, C) with controlled Lagrangian L of actual order $r \geq 1$ and associated controlled Euler-Lagrange equations of order at most $2r$. Let us indicate such a system of differential constraints as

$$\mathcal{E}^j \left(t, q^i, \frac{dq^i}{dt}, \dots, \frac{d^{2r} q^i}{dt^{2r}}, u^a(t) \right) = 0, \quad 1 \leq i \leq N. \quad (7.1)$$

By considering appropriate auxiliary variables, say

$$y^1(t) = q^1(t), \quad y^2(t) = q^2(t), \quad \dots, \quad y^{N+1}(t) = \frac{dq^1}{dt}, \quad y^{N+2}(t) = \frac{dq^2}{dt}, \quad \dots,$$

the original system (7.1) can be always transformed into an equivalent system on the new functions $y^A(t)$, which consists only of *first* order differential equations of the form

$$\mathcal{G}^B \left(t, y^1, y^2, \dots, y^{\tilde{N}}, \frac{dy^1}{dt}, \frac{dy^2}{dt}, \dots, \frac{dy^{\tilde{N}}}{dt}, u^a(t) \right) = 0, \quad 1 \leq B \leq \tilde{N}' \quad (7.2)$$

for some appropriate \tilde{N} -tuples of variables and \tilde{N}' -tuples of equations, with both \tilde{N} and \tilde{N}' greater than or equal to N . In general, there is not a unique way to reduce the constraints (7.1) into a form (7.2). For instance, the second order equation on curves $q(t) \in \mathcal{Q} = (0, +\infty)$

$$q \frac{d^2 q}{dt^2} + \left(\frac{dq}{dt} \right)^2 - u(t) = 0 \quad (7.3)$$

can be reduced not only to the equivalent first order system

$$y^1 \frac{dy^2}{dt} + (y^2)^2 - u = 0, \quad \frac{dy^1}{dt} - y^2 = 0, \quad (7.4)$$

but also to the system (via the change of variable $\tilde{y}^1 = q^2$)

$$\frac{d\tilde{y}^2}{dt} - 2u = 0, \quad \frac{d\tilde{y}^1}{dt} - \tilde{y}^2 = 0. \quad (7.5)$$

We say that *the differential constraints (7.1) are of normal type* if they are equivalent to at least one first order differential system (7.2), consisting only of first order equations in normal form, that is having the form

$$\frac{dy^A}{dt} = g^A(t, y^B, u^a(t)), \quad (7.6)$$

with some smooth function $g = (g^A) : \mathbb{R} \times \mathbb{R}^{\tilde{N}} \times K \subset \mathbb{R}^{\tilde{N}+M+1} \rightarrow \mathbb{R}^{\tilde{N}}$. Note that, in the previous example, the system (7.4) does not satisfy such a regularity assumption, but (7.5) does. This is good enough for us to consider the (7.3) as a constraint of normal type.

In case it is possible to transform the original constraints into some of the form (7.6) in which the g^A and the partial derivatives $\frac{\partial g^A}{\partial y^B} \Big|_{(t, y^B, u)}$ have uniform bounds we say that they are of *bounded normal type*.

Remark 7.1. In dealing with localised properties of solutions of a set of differential constraints, using cut-off functions it is in general possible to replace some (possibly unbounded) smooth functions g^A by other functions g'^A , coinciding with the g^A on an appropriate compact set and satisfying global bounds for their values and the partial derivatives $\frac{\partial g^A}{\partial y^B}$. This idea is a common tool for extending estimates on solutions of constraints of *bounded normal type* to those of more general constraints of normal type.

Consider now the distance function on the family of control curves $u, u' : [0, T] \rightarrow K \subset \mathbb{R}^M$ defined by

$$\text{dist}(u, u') := \text{measure} \{ t \in [0, T] : u(t) \neq u'(t) \}$$

and denote by ρ a fixed metric on the jet space $J^{n-1}(\mathbb{Q}|\mathbb{R})|_{t=0}$ of initial values, which generates the standard topology of such a space. The next lemma is an immediate consequence of a classical fact on systems of controlled first order differential equations (see e.g. [2, Prop. 3.2.2]). It gives the main motivation for considering the class of differential constraints of normal type.

Lemma 7.2. *Let (\mathcal{K}, L, C) be a defining triple of a Mayer problem with differential constraints of order $2r$ of bounded normal type and $K \subset J^{n-1}(\mathbb{Q}|\mathbb{R})|_{t=0}$ a compact subset in the space of the $(n-1)$ -jets at $t=0$. There exists a constant \mathfrak{c} , depending only on L , such that for any two \mathcal{K} -controlled curves*

$$\gamma^{(U)}, \gamma^{(U')} : [0, T] \longrightarrow [0, T] \times \mathcal{Q},$$

determined by pairs $U = (u(\cdot), \sigma)$, $U' = (u'(\cdot), \sigma')$ with σ, σ' in K , one has

$$\|\gamma^{(U)} - \gamma^{(U')}\|_{\mathcal{C}^{2r-1}} \leq \mathfrak{c} (\text{dist}(u, u') + \rho(\sigma, \sigma')). \quad (7.7)$$

7.2. Generalised needle variations.

Let $U_o = (u_o(\cdot), \sigma_o)$ be a fixed element in \mathcal{K} and $\gamma^{(U_o)} : [0, T] \rightarrow [0, T] \times \mathcal{Q}$ the uniquely corresponding \mathcal{K} -controlled curve. Pick also a strictly positive time $\tau \in (0, T)$, a point $\omega \in K$ and two real numbers $0 < k \ll 1$ and $0 < \varepsilon_o$ small enough so that $[\tau - \varepsilon_o - k\varepsilon_o^2, \tau + k\varepsilon_o^2] \subset (0, T)$. After fixing k and the triple $(\tau, \omega, \varepsilon_o)$, we may consider the piecewise continuous map

$$u^{(\tau, \omega, \varepsilon_o)} : [0, T] \rightarrow K, \quad u^{(\tau, \omega, \varepsilon_o)}(t) := \begin{cases} u_o(t) & \text{if } t \in [0, \tau - \varepsilon_o), \\ \omega & \text{if } t \in [\tau - \varepsilon_o, \tau), \\ u_o(t) & \text{if } t \in [\tau, T] \end{cases} \quad (7.8)$$

and an appropriate associated smooth map $\check{u}^{(\tau, \omega, \varepsilon_o)} : [0, T] \rightarrow \check{K}$ with values in the convex hull $\check{K} \subset \mathbb{R}^M$ of a slightly larger open neighbourhood of K , described as follows. We assume that $\check{u}^{(\tau, \omega, \varepsilon_o)}(t)$ is equal to $u^{(\tau, \omega, \varepsilon_o)}(t)$ for all points t in $[0, T]$ with the only exception of those in two small intervals of the form $[\tau - \varepsilon_o - k\varepsilon_o^2, \tau - \varepsilon_o]$ and $[\tau, \tau + k\varepsilon_o^2]$ with $k \ll 1$, in which the function $\check{u}^{(\tau, \omega, \varepsilon_o)}(t)$ is required just to take values in \check{K} , with no further restrictions. We call $u^{(\tau, \omega, \varepsilon_o)}$ the *needle modification of u_o at $t = \tau$ of ceiling value ω and width ε_o* . Any associated smooth approximation $\check{u}^{(\tau, \omega, \varepsilon_o)}$ will be called *smoothed needle modification* (see Fig. 2 and Fig. 3). There are of course several ways to build up a smoothed needle modification for a given discontinuous one. Nonetheless, we assume that a fixed algorithm has been chosen and that each discontinuous needle modification has a uniquely associated smoothed one.

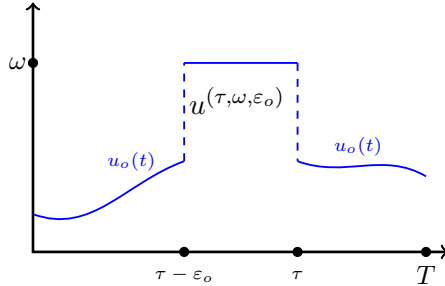


Fig. 2 Needle modification

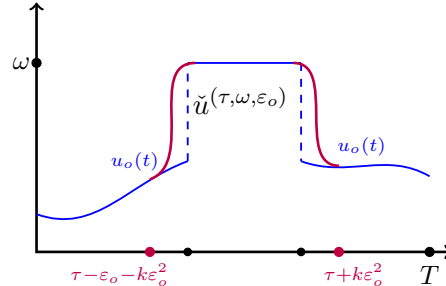


Fig. 3 Smoothed needle modification

From now on, we also assume the following convenient assumption:

The differential problems that determine the \mathcal{K} -controlled curves are well defined and with unique solutions also for the pairs $(u(t), \sigma)$, in which the initial datum σ is as usual and $u(t)$ takes values in the convex hull \check{K} of some open neighbourhood of K .

On the basis of this assumption, we may speak of a \mathcal{K} -controlled curve $\gamma^{(U)}$, $U = (u(t), \sigma)$, not only when $u([0, T])$ is entirely included in K , but also when it is within an appropriate convex set $\check{K} \supset K$.

In the next definition, the function $\Sigma = \Sigma(\varepsilon, s)$ is a continuous two-parameters family of initial data in $J^{n-1}(\mathcal{Q}|\mathbb{R})|_{t=0}$ satisfying the condition $\Sigma(\varepsilon, 0) = \sigma_o$ for each $\varepsilon \in [0, \varepsilon_o]$.

Definition 7.3. The (generalised) needle variation of the \mathcal{K} -controlled curve $\gamma^{(U_o)}$, corresponding to the triple $(\tau, \omega, \varepsilon_o)$ and led by the family $\Sigma = \Sigma(\varepsilon, s)$ is the one-parameter

family of \mathcal{K} -controlled variations

$$\text{Needle}^{(\tau, \omega, \varepsilon_o, \Sigma)}(\gamma^{(U_o)}) := \{ F^{(\tau, \omega, \varepsilon, \Sigma)} : [0, T] \times [0, 1] \rightarrow [0, T] \times \mathcal{Q}, \varepsilon \in (0, \varepsilon_o] \}, \quad (7.9)$$

in which the variations $F^{(\tau, \omega, \varepsilon, \Sigma)}$ are determined as follows. For each $\varepsilon \in (0, \varepsilon_o]$, let $U^{(\varepsilon, s)} = (u^{(\varepsilon, s)}(t), \sigma^{(\varepsilon, s)} = \sigma_o)$ be the curve in \mathcal{K} , in which:

- the initial data $\sigma^{(\varepsilon, s)}$ are given by $\sigma^{(\varepsilon, s)} = \Sigma(\varepsilon, s)$;
- the one-parameter family of curves $u^{(\varepsilon, s)}(t)$ takes values in the convex set \tilde{K} and is defined by

$$u^{(\varepsilon, s)}(t) = (1 - s)u_o(t) + s\check{u}^{(\tau, \omega, \varepsilon)}(t), \quad s \in [0, 1]. \quad (7.10)$$

Then, we define $F^{(\tau, \omega, \varepsilon, \Sigma)}$ as the \mathcal{K} -controlled variation

$$F^{(\tau, \omega, \varepsilon, \Sigma)}(t, s) := \gamma^{(U^{(\varepsilon, s)})}(t).$$

In the following, given a needle variation $\text{Needle}^{(\tau, \omega, \varepsilon_o, \Sigma)}$, for each $\varepsilon \in (0, \varepsilon_o]$ we denote by $\mu^{(\tau, \omega, \varepsilon, \Sigma)}, \mu'^{(\tau, \omega, \varepsilon, \Sigma)} : [0, T] \times [0, 1] \rightarrow \mathbb{R}$ the functions defined in (6.18), (6.19), by means of the homotopy $F^{(\tau, \omega, \varepsilon, \Sigma)}$.

7.3. The generalised Pontryagin Maximum Principle.

We are now ready to prove our announced analogue of the PMP for the class of generalised Mayer problems with differential constraints of normal type. As in all previous sections, we still consider a fixed generalised Mayer problem determined by one of the triples (\mathcal{K}, L, C) described in Sect. 4. We start with a preparatory lemma.

Lemma 7.4. *Let $\gamma_o := \gamma^{(U_o)}$ be a \mathcal{K} -controlled curve and $\text{Needle}^{(\tau, \omega, \varepsilon_o, \Sigma)}(\gamma_o) = \{F^{(\tau, \omega, \varepsilon, \Sigma)}, 0 < \varepsilon \leq \varepsilon_o\}$ a needle variation of γ_o with associated function $\mu'^{(\tau, \omega, \varepsilon, \Sigma)} : [0, T] \times [0, 1] \rightarrow \mathbb{R}$ as defined in § 7.2. If the Mayer problem has differential constraints of normal type, the limit $\lim_{\varepsilon \rightarrow 0} \mathcal{P}_{j_\tau^r(\gamma^{(U^{(\varepsilon, s)}))}(\omega)$ exist and is equal to*

$$\lim_{\varepsilon \rightarrow 0} \mathcal{P}_{j_\tau^r(\gamma^{(U^{(\varepsilon, s)}))}(\omega) = \mathcal{P}_{j_\tau^r(\gamma^{(U_o)})}(\omega). \quad (7.11)$$

Proof. First of all, we recall that, for any $\varepsilon \in (0, \varepsilon_o]$, the curve of controls $U^{(\varepsilon, s)} = (u^{(\varepsilon, s)}(t), \sigma^{(\varepsilon, s)} = \Sigma(\varepsilon, s)), s \in [0, 1]$, have all initial data $\sigma^{(\varepsilon, s)}$ in the compact set $\Sigma([0, \varepsilon_o] \times [0, 1])$ and all control curves $u^{(\varepsilon, s)}$ differ from the control curve $u_o(t)$ only at the points of an interval of measure ε (more precisely, of measure $\varepsilon + o(\varepsilon)$ since we are considering smoothed needle variations). Thus, since all of the $\gamma^{(U^{(\varepsilon, s)})}$ and their jets lie in a compact subset of $J^{2r+1}(\mathcal{Q}|\mathbb{R})$, by the standard use of cut-off functions described Remark 7.1 and Lemma 7.2, we have that $\|\gamma^{(U^{(\varepsilon, s)})} - \gamma^{(U_o)}\|_{\mathcal{C}^{2r-1}}$ goes to 0 for $\varepsilon \rightarrow 0$. From this the function $\mathcal{P}_{j_\tau^r(\gamma^{(U^{(\varepsilon, s)}))} : K \rightarrow \mathbb{R}$ tends to the function $\mathcal{P}_{j_\tau^r(\gamma^{(U_o)})} : K \rightarrow \mathbb{R}$. \square

Theorem 7.5 (Generalised Pontryagin Maximum Principle). *Let $\gamma_o := \gamma^{(U_o)}$ be a \mathcal{K} -controlled curve for a generalised Mayer problem with differential constraints of normal type determined by a triple (\mathcal{K}, L, C) . A necessary condition for γ_o to be a solution to the Mayer problem is that for any $\text{Needle}^{(\tau, \omega, \varepsilon_o, \Sigma)}(\gamma_o) = \{F^{(\tau, \omega, \varepsilon, \Sigma)}, 0 < \varepsilon \leq \varepsilon_o\}$*

$$\mathcal{P}_{j_\tau^r(\gamma_o)}(\omega) - \liminf_{\varepsilon \rightarrow 0^+} \frac{\mu'^{(\tau, \omega, \varepsilon, \Sigma)}(T, 1) - \mu'^{(\tau, \omega, \varepsilon, \Sigma)}(T, 0)}{\varepsilon} \leq \mathcal{P}_{j_\tau^r(\gamma_o)}(u_o(\tau)). \quad (7.12)$$

Proof. Let $Needle^{(\tau,\omega,\varepsilon_o,\Sigma)}(\gamma_o)$ be a fixed needle variation and $\mathcal{Z} : (0, \varepsilon_o] \rightarrow \mathbb{R}$ the function defined by

$$\mathcal{Z}(\varepsilon) := \int_0^T \left(\int_0^1 \left(Y^{(\varepsilon)a} \frac{\partial \mathcal{P}_{\sigma_t^{(s)}}}{\partial u^a} \Big|_{u^{(\varepsilon,s)}(t)} - \frac{\partial^2 \mu^{(\tau,\omega,\varepsilon,\Sigma)}}{\partial t \partial s} \Big|_{(t,s)} \right) ds \right) dt, \quad (7.13)$$

where, as usual, $\sigma_t^{(s)}$ denotes the curve of jets $\sigma_t^{(s)} = j_t^n(\gamma^{(U(\varepsilon,s))})$ and the $Y^{(\varepsilon)a}$ are the components of the vector field $Y_{F^{(\tau,\omega,\varepsilon,\Sigma)}}$ in the directions of the u^a -axes. We recall that, due to the particular construction of the homotopies $F^{(\tau,\omega,\varepsilon,\Sigma)}$,

$$Y^{(\varepsilon)a}(t, s) = \frac{\partial u^{(\varepsilon,s)a}(t)}{\partial s} = \frac{\partial \left((1-s)u_o^a(t) + s\check{u}^{(\tau,\omega,\varepsilon)a}(t) \right)}{\partial s} = \check{u}^{(\tau,\omega,\varepsilon)a}(t) - u_o^a(t) \quad (7.14)$$

and $(\check{u}^{(\tau,\omega,\varepsilon,\Sigma)} - u_o) \Big|_{[0,T] \setminus [\tau-\varepsilon-k\varepsilon^2, \tau+k\varepsilon^2]} = 0$. Thus the functions $Y^{(\varepsilon)a} \Big|_{u^{(\varepsilon,s)}(t)}$ are equal to 0 at the points outside of the rectangle $[\tau - \varepsilon - \frac{k\varepsilon^2}{2}, \tau + \frac{k\varepsilon^2}{2}] \times [0, 1]$. By Theorem 6.6, a necessary condition for γ_o to be a solution to the Mayer problem is that $\mathcal{Z}(\varepsilon) \leq 0$ for any $\varepsilon \in (0, \varepsilon_o]$. Hence we have that for any such ε

$$0 \geq \frac{1}{\varepsilon} \mathcal{Z}(\varepsilon) = \frac{1}{\varepsilon} \left(\iint_{[0,T] \times [0,1]} Y^{(\varepsilon)a} \frac{\partial \mathcal{P}_{\sigma_t^{(s)}}}{\partial u^a} \Big|_{u^{(\varepsilon,s)}(t)} dt ds - \iint_{[0,T] \times [0,1]} \frac{\partial^2 \mu^{(\tau,\omega,\varepsilon,\Sigma)}}{\partial t \partial s} \Big|_{(t,s)} dt ds \right). \quad (7.15)$$

From (7.14) the first double integral in (7.15) reduces to the sum

$$\begin{aligned} & \int_0^1 \int_{\tau-\varepsilon-k\varepsilon^2}^{\tau-\varepsilon} Y^{(\varepsilon)a} \frac{\partial \mathcal{P}_{\sigma_t^{(s)}}}{\partial u^a} \Big|_{u^{(\varepsilon,s)}(t)} dt ds + \int_0^1 \int_{\tau-\varepsilon}^{\tau} Y^{(\varepsilon)a} \frac{\partial \mathcal{P}_{\sigma_t^{(s)}}}{\partial u^a} \Big|_{u^{(\varepsilon,s)}(t)} dt ds + \\ & + \int_0^1 \int_{\tau}^{\tau+k\varepsilon^2} Y^{(\varepsilon)a} \frac{\partial \mathcal{P}_{\sigma_t^{(s)}}}{\partial u^a} \Big|_{u^{(\varepsilon,s)}(t)} dt ds = \int_0^1 \int_{\tau-\varepsilon}^{\tau} (\omega^a - u_o^a(t)) \frac{\partial \mathcal{P}_{\sigma_t^{(s)}}}{\partial u^a} \Big|_{u^{(\varepsilon,s)}(t)} dt ds + O(\varepsilon^2). \end{aligned}$$

From this, we obtain

$$\begin{aligned} 0 \geq \frac{1}{\varepsilon} \mathcal{Z}(\varepsilon) &= \int_0^1 \frac{1}{\varepsilon} \left(\int_{\tau-\varepsilon}^{\tau} (\omega^a - u_o^a(t)) \frac{\partial \mathcal{P}_{\sigma_t^{(s)}}}{\partial u^a} \Big|_{u^{(\varepsilon,s)}(t)} dt \right) ds - \\ & \quad - \frac{\mu^{(\tau,\omega,\varepsilon,\Sigma)}(T, 1) - \mu^{(\tau,\omega,\varepsilon,\Sigma)}(T, 0)}{\varepsilon} + O(\varepsilon) = \\ &= \int_0^1 (\omega^a - u_o^a(\tau)) \frac{\partial \mathcal{P}_{\sigma_\tau^{(s)}}}{\partial u^a} \Big|_{u^{(\varepsilon,s)}(\tau)} ds - \frac{\mu^{(\tau,\omega,\varepsilon,\Sigma)}(T, 1) - \mu^{(\tau,\omega,\varepsilon,\Sigma)}(T, 0)}{\varepsilon} + O(\varepsilon) = \\ &= \int_0^1 \frac{\partial u^{(\varepsilon,s)a}}{\partial s} \Big|_{t=\tau} \frac{\partial \mathcal{P}_{\sigma_\tau^{(s)}}}{\partial u^a} \Big|_{u^{(\varepsilon,s)}(\tau)} ds - \frac{\mu^{(\tau,\omega,\varepsilon,\Sigma)}(T, 1) - \mu^{(\tau,\omega,\varepsilon,\Sigma)}(T, 0)}{\varepsilon} + O(\varepsilon) = \\ &= \mathcal{P}_{j_\tau^n(\gamma^{(U(\tau,\omega,\varepsilon))})}(\omega) - \mathcal{P}_{j_\tau^n(\gamma^{(U_o)})}(u_o(\tau)) - \frac{\mu^{(\tau,\omega,\varepsilon,\Sigma)}(T, 1) - \mu^{(\tau,\omega,\varepsilon,\Sigma)}(T, 0)}{\varepsilon} + O(\varepsilon). \end{aligned} \quad (7.16)$$

From this and (7.11) the claim follows. \square

In order to have a truly helpful theorem, the previous result should be combined with some efficient way to determine the sign of the corrective term $\liminf_{\varepsilon \rightarrow 0^+} \frac{\mu'(\tau, \omega, \varepsilon, \Sigma)(T, 1) - \mu'(\tau, \omega, \varepsilon, \Sigma)(T, 0)}{\varepsilon}$ appearing in (7.12). This can be reached exploiting the next technical lemma.

Lemma 7.6. *Let the triple (\mathcal{K}, L, C) and the curve $\gamma_o := \gamma^{(U_o)}$ be as in Theorem 7.5. For any Needle $^{(\tau, \omega, \varepsilon_o, \Sigma)}(\gamma_o) = \{F^{(\tau, \omega, \varepsilon, \Sigma)}, 0 < \varepsilon \leq \varepsilon_o\}$ of γ_o , one has that*

$$\begin{aligned} \mu'(\tau, \omega, \varepsilon, \Sigma)(T, 1) - \mu'(\tau, \omega, \varepsilon, \Sigma)(T, 0) &= C|_{F^{(\tau, \omega, \varepsilon, \Sigma)(n)}(T, 1)} - C|_{\gamma_o^{(n)}(T)} - \\ &\quad - \int_0^T L|_{F^{(\tau, \omega, \varepsilon, \Sigma)(n)}(t, 1)} dt + \int_0^T L|_{\gamma_o^{(n)}(t)} dt + \\ &\quad + \int_0^1 \sum_{\delta=1}^r \sum_{\varepsilon=0}^{\delta-1} (-1)^\varepsilon \frac{d^\varepsilon}{dt^\varepsilon} \left(\frac{\partial L}{\partial q_{(\delta)}^i} \right) (Y_F)^i_{(\delta - (\varepsilon + 1))} |_{F^{(\tau, \omega, \varepsilon, \Sigma)(n)}(T, s)} ds - \\ &\quad - \int_0^1 \sum_{\delta=1}^r \sum_{\varepsilon=0}^{\delta-1} (-1)^\varepsilon \frac{d^\varepsilon}{dt^\varepsilon} \left(\frac{\partial L}{\partial q_{(\delta)}^i} \right) (Y_F)^i_{(\delta - (\varepsilon + 1))} |_{F^{(\tau, \omega, \varepsilon, \Sigma)(n)}(0, s)} ds. \end{aligned} \quad (7.17)$$

Proof. Let Π and Π^{PC} be the 1-forms on $J^n(\widehat{\mathcal{Q}}|\mathbb{R})$ defined by

$$\begin{aligned} \Pi := &\left\{ \frac{1}{2} \sum_{\substack{1 \leq i \leq N \\ 0 \leq \beta \leq r-1}} \left((\mathbf{h}_{\beta(1)}^i)^2 - (\mathbf{h}'_{\beta(2)}^i)^2 - (\mathbf{h}''_{\beta(2)}^i)^2 \right) + \right. \\ &\left. + \sum_{\substack{1 \leq i \leq N \\ 0 \leq \beta \leq r-1}} \left(\frac{1}{2} (\mathbf{h}_\beta^i)^2 + \frac{\pi^4}{32T^4} (\mathbf{h}'_\beta^i)^2 + \frac{\pi^4}{32T^4} (\mathbf{h}''_\beta^i)^2 \right) \right\} dt, \end{aligned} \quad (7.18)$$

$$\begin{aligned} \Pi^{PC} := &\Pi + \lambda \sum_{\substack{1 \leq i \leq N \\ 0 \leq \beta \leq r-1}} \left(\mathbf{h}_{\beta(1)}^i \varpi_{\beta(0)}^i - \mathbf{h}'_{\beta(2)}^i \varpi'_{\beta(1)}^i - \mathbf{h}''_{\beta(2)}^i \varpi''_{\beta(1)}^i + \right. \\ &\left. + \mathbf{h}'_{\beta(3)}^i \varpi'^i_{\beta(0)} + \mathbf{h}''_{\beta(3)}^i \varpi''^i_{\beta(0)} \right) \end{aligned} \quad (7.19)$$

(for the definitions of the 1-forms $\varpi_{\beta(\delta)}^i$ etc., see (5.20)). Consider the variation $\widehat{F} = \widehat{F}^{(\tau, \omega, \varepsilon, \Sigma)}$ in the extended space $\mathbb{R} \times \widehat{\mathcal{Q}}$, which is uniquely associated with one of the $F = F^{(\tau, \omega, \varepsilon, \Sigma)}$ belonging to the considered needle variation, as we indicated at the beginning of Sect. 6.2. Let also denote by $X_{\widehat{F}}$ and $Y_{\widehat{F}}$, the corresponding vector fields in $\widehat{\mathcal{S}} = \widehat{F}([0, T] \times [0, 1])$, as defined in (6.14) and (6.15). Since Π and Π^{PC} are variationally equivalent (see definition in Sect. 5.3) and, at each point of $\widehat{\mathcal{S}}$, the vector field $X_{\widehat{F}}$ is tangent to curves of jets of one of the curves in $\widehat{\mathcal{Q}}$ determined by the homotopy \widehat{F} , we have

$$\begin{aligned} & - \int_0^T \Pi^{PC}(X_{\widehat{F}})|_{\widehat{F}^{(n)}(t, 1)} dt + \int_0^T \Pi^{PC}(X_{\widehat{F}})|_{\widehat{F}^{(n)}(t, 0)} dt = \\ & = - \int_0^T \Pi(X_{\widehat{F}})|_{\widehat{F}^{(n)}(t, 1)} dt + \int_0^T \Pi(X_{\widehat{F}})|_{\widehat{F}^{(n)}(t, 0)} dt = \\ & = \mu^{(\tau, \omega, \varepsilon, \Sigma)}(T, 1) - \mu^{(\tau, \omega, \varepsilon, \Sigma)}(T, 0) + \int_0^T L|_{\widehat{F}^{(n)}(t, 1)} dt - \int_0^T L|_{\gamma_o^{(n)}(t)} dt. \end{aligned} \quad (7.20)$$

On the other hand, by the Stokes Theorem

$$\begin{aligned} & - \int_0^T \Pi^{PC} (X_{\widehat{F}}) |_{\widehat{F}^{(n)}(t,1)} dt + \int_0^T \Pi^{PC} (X_{\widehat{F}}) |_{\widehat{F}^{(n)}(t,0)} dt = \int_0^1 \Pi^{PC} (Y_{\widehat{F}}) |_{\widehat{F}^{(n)}(0,s)} ds - \\ & - \int_0^1 \Pi^{PC} (Y_{\widehat{F}}) |_{\widehat{F}^{(n)}(T,s)} ds + \iint_{[0,T] \times [0,1]} d\Pi^{PC} (X_{\widehat{F}}, Y_{\widehat{F}}) |_{\widehat{F}^{(n)}(t,s)} dt ds \quad (7.21) \end{aligned}$$

We claim that the third summand in the right hand side of (7.21) is 0. Indeed, by the same arguments of the proof of Lemma 5.2, the differential $d\Pi^{PC}$ is equal to a sum of 2-forms that are (a) either identically vanishing on the vector field $X_{\widehat{F}}$ or (b) have coefficients that vanish identically along the solutions of the Euler-Lagrange equations determined by

$$\begin{aligned} L' = & \left\{ \frac{1}{2} \sum_{\substack{1 \leq i \leq N \\ 0 \leq \beta \leq r-1}} \left((\mathbf{h}_{\beta(1)}^i)^2 - (\mathbf{h}_{\beta(2)}^i)^2 - (\mathbf{h}''_{\beta(2)}^i)^2 \right) + \right. \\ & \left. + \sum_{\substack{1 \leq i \leq N \\ 0 \leq \beta \leq r-1}} \left(\frac{1}{2} (\mathbf{h}_{\beta}^i)^2 + \frac{\pi^4}{32T^4} (\mathbf{h}'_{\beta}^i)^2 + \frac{\pi^4}{32T^4} (\mathbf{h}''_{\beta}^i)^2 \right) \right\}. \end{aligned}$$

Since we are integrating along the points of the surface $\widehat{\mathcal{S}}$ (whose components $\mathbf{h}_{\beta}^i(t)$, $\mathbf{h}'_{\beta}^i(t)$, \mathbf{h}''_{β}^i are solutions precisely to such Euler-Lagrange equations), the claim follows. From this, (7.20) and (7.21), we obtain

$$\begin{aligned} & \mu^{(\tau,\omega,\varepsilon,\Sigma)}(T, 1) - \mu^{(\tau,\omega,\varepsilon,\Sigma)}(T, 0) + \int_0^T L |_{\widehat{F}^{(n)}(t,1)} dt - \int_0^T L |_{\gamma_o^{(n)}(t)} dt = \\ & = \int_0^1 \Pi^{PC} (Y_{\widehat{F}}) |_{\widehat{F}^{(n)}(0,s)} ds - \int_0^1 \Pi^{PC} (Y_{\widehat{F}}) |_{\widehat{F}^{(n)}(T,s)} ds + \\ & + \sum_{\beta=1}^{r-1} \sum_{i=1}^N \int_0^1 \left(\mathbf{h}'_{\beta(3)}^i Y_{\beta(0)}^i + \mathbf{h}''_{\beta(3)}^i Y_{\beta(0)}^i \right) \Big|_{\widehat{F}^{(n)}(T,v)} dv. \quad (7.22) \end{aligned}$$

We now recall that the initial and terminal conditions on the functions $\mathbf{h}_{\beta}^i(t)$, $\mathbf{h}'_{\beta}^i(t)$, \mathbf{h}''_{β}^i have been selected such a way that the three terms of the right hand side of (7.22) are equal to minus the corresponding integrals along the two “vertical sides” of $\widehat{\partial\mathcal{S}}$ of the 1-form

$$\sum_{\delta=1}^r \sum_{\zeta=0}^{\delta-1} (-1)^{\zeta} \frac{d^{\zeta}}{dt^{\zeta}} \left(\frac{\partial (L + \frac{dC}{dt})}{\partial q_{(\delta)}^i} \right) \omega_{(\delta-(\zeta+1))}^i \quad (7.23)$$

(see Remarks 6.3 and 6.4). From this and the fact that the cost function C vanishes identically on $J^n(\mathcal{Q}|\mathbb{R})_{t=0}$, it follows that

$$\begin{aligned}
\mu'^{(\tau,\omega,\varepsilon,\Sigma)}(T, 1) - \mu'^{(\tau,\omega,\varepsilon,\Sigma)}(T, 0) &= - \int_0^T L|_{\widehat{F}^{(n)}(t,1)} dt + \int_0^T L|_{\gamma_o^{(n)}(t)} dt + \\
&+ \int_0^1 \sum_{\delta=1}^r \sum_{\zeta=0}^{\delta-1} (-1)^\zeta \frac{d^\zeta}{dt^\zeta} \left(\frac{\partial L}{\partial q^i_{(\delta)}} \right) Y_{\widehat{F}^{(n)}(\delta-(\zeta+1))} |_{\widehat{F}^{(n)}(T,s)} ds - \\
&- \int_0^1 \sum_{\delta=1}^r \sum_{\zeta=0}^{\delta-1} (-1)^\zeta \frac{d^\zeta}{dt^\zeta} \left(\frac{\partial L}{\partial q^i_{(\delta)}} \right) Y_{\widehat{F}^{(n)}(\delta-(\zeta+1))} |_{\widehat{F}^{(n)}(0,s)} ds + \\
&+ \int_0^1 \sum_{\delta=1}^r \sum_{\zeta=0}^{\delta-1} (-1)^\zeta \frac{d^\zeta}{dt^\zeta} \frac{\partial}{\partial q^i_{(\delta)}} \left(\frac{dC}{dt} \right) Y_{\widehat{F}^{(n)}(\delta-(\zeta+1))} |_{\widehat{F}^{(n)}(T,s)} ds.
\end{aligned} \tag{7.24}$$

We now observe that the fifth summand in (7.24) is equal to the sum along the two “vertical sides” of $\widehat{\mathcal{D}}\widehat{\mathcal{S}}$ of the 1-form

$$\beta^{PC} := \frac{dC}{dt} dt + \sum_{\delta=1}^r \sum_{\zeta=0}^{\delta-1} (-1)^\zeta \frac{d^\zeta}{dt^\zeta} \frac{\partial}{\partial q^i_{(\delta)}} \left(\frac{dC}{dt} \right) \omega_{\delta-(\zeta+1)}^i.$$

This 1-form is variationally equivalent to the 1-form $\beta := \frac{dC}{dt} dt$. Hence, by Stokes Theorem and the properties of the variationally equivalent 1-forms

$$\begin{aligned}
&\int_0^1 \sum_{\delta=1}^r \sum_{\zeta=0}^{\delta-1} (-1)^\zeta \frac{d^\zeta}{dt^\zeta} \frac{\partial}{\partial q^i_{(\delta)}} \left(\frac{dC}{dt} \right) Y_{\widehat{F}^{(n)}(\delta-(\zeta+1))} |_{\widehat{F}^{(n)}(T,s)} ds = \\
&= \int_0^T \left(\frac{dC}{dt} \Big|_{\widehat{F}^{(n)}(t,1)} - \frac{dC}{dt} \Big|_{\widehat{F}^{(n)}(t,0)} \right) dt + \iint_{[0,T] \times [0,1]} d\beta^{PC}(X_{\widehat{F}}, Y_{\widehat{F}}) \Big|_{\widehat{F}^{(n)}(t,s)} dt ds = \\
&= C|_{\widehat{F}^{(n)}(T,1)} - C|_{\gamma_o^{(n)}(T)} + \iint_{[0,T] \times [0,1]} d\beta^{PC}(X_{\widehat{F}}, Y_{\widehat{F}}) \Big|_{\widehat{F}^{(n)}(t,s)} dt ds. \tag{7.25}
\end{aligned}$$

Using once again the proof of Lemma 5.2, and that the curves $\widehat{F}^{(n)}(\cdot, s)$, $s \in [0, 1]$, have tangent vectors on which the holonomic 1-forms vanish identically, we obtain that the double integral in (7.25) reduces to the integral of a linear combination of 2-forms with coefficients given by the Euler-Lagrange operator applied to the Lagrangian $\frac{dC}{dt}$. By the well-known property that an Euler-Lagrange operator on a total differential gives an identically vanishing function, we conclude that $\iint_{[0,T] \times [0,1]} d\beta^{PC}(X_{\widehat{F}}, Y_{\widehat{F}}) \Big|_{\widehat{F}^{(n)}(t,s)} dt ds = 0$. From (7.25) and (7.24) the lemma follows. \square

Corollary 7.7 (Generalised Pontryagin Maximum Principle – II Version). *Let (\mathcal{K}, L, C) be as in Theorem 7.5 and for any \mathcal{K} -controlled curve $\gamma := \gamma^{(U)}$ denote by $\text{Good}\mathcal{N}(\gamma)$ the class of needle variations $\text{Needle}^{(\tau,\omega,\varepsilon_o,\Sigma)}(\gamma_o) = \{F^{(\tau,\omega,\varepsilon,\Sigma)}, 0 < \varepsilon \leq \varepsilon_o\}$ of γ_o , which*

satisfy the following inequality for any $\varepsilon \in (0, \varepsilon_o]$

$$\begin{aligned} & \int_0^T \left(L|_{F^{(\tau, \omega, \varepsilon, \Sigma)(n)}(t, 1)} - L|_{\gamma_o^{(n)}(t)} \right) dt + \int_0^1 \left(- \frac{\partial C}{\partial q_{(\beta)}^i} (Y_F)_{(\beta)}^i \Big|_{F^{(\tau, \omega, \varepsilon, \Sigma)(n)}(T, s)} - \right. \\ & \quad \left. - \sum_{\delta=1}^r \sum_{\varepsilon=0}^{\delta-1} (-1)^\varepsilon \frac{d^\varepsilon}{dt^\varepsilon} \left(\frac{\partial L}{\partial q_{(\delta)}^i} \right) (Y_F)_{(\delta-(\varepsilon+1))}^i \Big|_{F^{(\tau, \omega, \varepsilon, \Sigma)(n)}(T, s)} \right) ds + \\ & \quad + \int_0^1 \sum_{\delta=1}^r \sum_{\varepsilon=0}^{\delta-1} (-1)^\varepsilon \frac{d^\varepsilon}{dt^\varepsilon} \left(\frac{\partial L}{\partial q_{(\delta)}^i} \right) (Y_F)_{(\delta-(\varepsilon+1))}^i \Big|_{F^{(\tau, \omega, \varepsilon, \Sigma)(n)}(0, s)} ds \geq 0 . \end{aligned} \quad (7.26)$$

A \mathcal{K} -controlled curve $\gamma_o = \gamma^{(U_o)}$ is a solution to the Mayer problem only if for any needle variation in $\text{GoodN}(\gamma_o)$

$$\mathcal{P}_{j_\tau^r(\gamma_o)}(\omega) \leq \mathcal{P}_{j_\tau^n(\gamma_o)}(u_o(\tau)) . \quad (7.27)$$

Proof. From Lemma 7.6, if a needle variation satisfies (7.26), then the expression $-\frac{\mu'(\tau, \omega, \varepsilon, \Sigma)(T, 1) - \mu'(\tau, \omega, \varepsilon, \Sigma)(T, 0)}{\varepsilon}$ is non-negative for any $\varepsilon \in (0, \varepsilon_o]$. From this and Theorem 7.5, the necessary condition (7.27) holds. \square

7.4. The classical Pontryagin Maximum Principle.

Consider now a classical Mayer problem, i.e. a problem as described in Sect. 2.1.1. Let us represent the controlled evolutions of the system by curves $x(t) = (x^1(t), \dots, x^{N'}(t))$, $t \in [0, T]$, in some $\mathbb{R}^{N'}$. They are constrained by the conditions:

- (a) $x(0) = x_o$ for a fixed initial value x_o ;
- (b) they satisfy the differential constraints $\frac{dx^i}{dt} = f^i(t, x(t), u(t))$.

As pointed out in Sect. 2.1.2, if we add the auxiliary variables $p = (p_1, \dots, p_{N'})$, impose that they are solutions to the equations $\frac{dp_i}{dt} = -p_\ell \frac{\partial f^\ell}{\partial x^i}(t, x(t), u(t))$ and set

$$q^1 := x^1, \dots, q^{N'} := x^{N'}, q^{N'+1} := p_1, \dots, q^{2N'} := p_{N'},$$

such a classical problem can be considered as a generalised Mayer problem on $J^3(\mathbb{R}^N|\mathbb{R})$, $N = 2N'$ ⁽⁸⁾ with the defining triple given by:

- the set \mathcal{K} of the pairs $U = (u(t), \sigma)$, in which $u(t)$ is a smooth curve $u : [0, T] \rightarrow K \subset \mathbb{R}^M$ and $\sigma = (A^i = x^i(0), B_\ell = p_\ell(0))$ is a 0-th order jet where $x(0) = x_o$ and $p(0)$ is (provisionally) unconstrained.
- the controlled Lagrangian

$$L(t, q_{(\beta)}^j, u^a) := p_i \left(x_{(1)}^i - f^i(t, x^i, u^a) \right) .$$

- a cost function $C : J^3(\mathbb{R}^N|\mathbb{R}) \rightarrow \mathbb{R}$ which is of actual order $r = 0$ and coincides on $J^3(\mathbb{R}^N|\mathbb{R})|_{t=T}$ with a classical terminal cost function, depending just on the coordinates x^i . With no loss of generality, we assume that C depends just on the x^i at all points.

⁸We consider a jet bundle of order 3 just to be sure that (3.2) is satisfied by the actual order $r = 1$ of the controlled Lagrangian defined below.

Such a (generalised) Mayer problem is manifestly of normal type, as defined in Sect. 7.1, and Corollary 7.7 applies. Let us therefore determine what are the needle variations of the class $\mathcal{G}ood\mathcal{N}$ for this setting. First of all, we observe that along any solution of the controlled Euler-Lagrange equations of this problem, the function L vanishes identically. Hence, since the actual order r of the Lagrangian is $r = 1$, the characterising inequality (7.26) reduces to

$$C|_{F^{(\tau,\omega,\varepsilon,\Sigma)}(n)(T,1)} - C|_{\gamma_o^{(n)}(T)} \leq - \int_0^1 p_i \frac{\partial x^i}{\partial s} \Big|_{F^{(\tau,\omega,\varepsilon,\Sigma)}(T,s)} ds + \int_0^1 p_i \frac{\partial x^i}{\partial s} \Big|_{F^{(\tau,\omega,\varepsilon,\Sigma)}(0,s)} ds . \quad (7.28)$$

Here, we denoted by $x^i = x^i(t, s)$ the component in the x^i -direction of the \mathcal{K} -controlled variation $F(t, s) = F^{(\tau,\omega,\varepsilon,\Sigma)}(t, s)$ of a considered generalised needle variation. We now observe that

$$C|_{F^{(\tau,\omega,\varepsilon,\Sigma)}(n)(T,1)} - C|_{\gamma_o^{(n)}(T)} = \int_0^1 \frac{\partial C}{\partial x^i} \frac{\partial x^i}{\partial s} \Big|_{F^{(\tau,\omega,\varepsilon,\Sigma)}(T,s)} ds$$

and $x^i(0, s) \equiv x_o^i$ so that $\frac{\partial x^i}{\partial s} \Big|_{F^{(\tau,\omega,\varepsilon,\Sigma)}(0,s)} = 0$. Thus, (7.28) is equivalent to

$$- \int_0^1 \left(\frac{\partial C}{\partial x^i} + p_i \right) \frac{\partial x^i}{\partial s} \Big|_{F^{(\tau,\omega,\varepsilon,\Sigma)}(T,s)} ds \geq 0 . \quad (7.29)$$

We now recall that we are free to impose any initial condition on the auxiliary variables p_j . Furthermore, by the particular form of the differential constraints on the curves $p_j(t)$, it is certainly possible to determine a family $\Sigma = \Sigma(\varepsilon, s)$ of initial data for a needle variation, with the property that the corresponding functions $p_j(t, s)$ of the \mathcal{K} -controlled curves of a homotopy $F^{(\tau,\omega,\varepsilon,\Sigma)}$ satisfy the terminal conditions

$$p_i(T, s) = - \frac{\partial C}{\partial x^i} \Big|_{x(T,s)} , \quad (7.30)$$

provided, of course, that such terminal conditions are satisfied in the first place by the components $p_{io}(T) = p_i(T, 0)$ of $\gamma_o(T)$. (see also [3, Sect. 5.1]). This and (7.29) has the following crucial consequence: *if γ_o satisfies $p_i(T) = -\frac{\partial C}{\partial x^i} \Big|_{x(T)}$, all of the needle variations of γ_o , which are led by a Σ forcing (7.30), are in the class $\mathcal{G}ood\mathcal{N}(\gamma_o)$ described in Corollary 7.7.* In particular, there is a needle variation $\mathcal{N}eedle^{(\tau,\omega,\varepsilon_o,\Sigma)}(\gamma_o)$ in the class $\mathcal{G}ood\mathcal{N}(\gamma_o)$ for any choice of $\tau \in (0, T]$, $\omega \in K$ and ε_o sufficiently small. Thus, by Corollary 7.7, a necessary condition for such a curve $\gamma_o = \gamma^{(U_o)}$ to be a solution of the Mayer problem is that the inequality (7.27) holds for *any* needle variation as above.

We finally observe that, for the classical Mayer problem considered in this section, we have $\mathcal{P}_{j_t^1(\gamma_o)}(u^a) = \left(p_i f^i(t, x^i, u^a) - p_i x_{(1)}^i \right) \Big|_{j_t^1(\gamma_o)}$. Hence, if for each $(t, x^i, p_i) \in \mathbb{R} \times \mathbb{R}^{2N'}$, we denote by $\mathcal{H}(t, x^i, p_i) : K \rightarrow \mathbb{R}$ the classical Pontryagin function

$$\mathcal{H}(t, x^i, p_i)(u^a) := \sum_{i=1}^{N'} p_i f^i(t, x^i, u^a) ,$$

for any needle variation we have $\mathcal{P}_{j_t^1(\gamma_o)}(\omega^a) = \mathcal{H}(t, x^i, p_i)|_{\gamma_o(t)}(\omega^a) - \left(p_i x_{(1)}^i \right) \Big|_{j_t^1(\gamma_o)}$. From this, we immediately derive the following version of the classical Pontryagin Maximum Principle.

Corollary 7.8 (Pontryagin Maximum Principle). *A $\tilde{\mathcal{K}}$ -controlled curve $\gamma_o := \gamma^{(U_o)}$, with components $p_j(t)$ satisfying $p_i(T) = -\frac{\partial C}{\partial x^i}|_{x(T)}$, is a solution to the Mayer problem determined by the above described triple $(\tilde{\mathcal{K}}, L, C)$ only if for any $\tau \in (0, T]$ and $\omega \in K \subset \mathbb{R}^M$ the following inequality holds:*

$$\mathcal{H}|_{\gamma_o(\tau)}(\omega) \leq \mathcal{H}|_{\gamma_o(\tau)}(u_o) . \tag{7.31}$$

8. A DISCUSSION OF A BASIC EXAMPLE AND SOME CONCLUDING REMARKS

8.1. Comparison of different approaches to an elementary problem: the controlled linearised pendulum.

Let $x(t) \in \mathbb{R}$, $t \in [0, T]$, be the coordinate which describes the position in time of a linearised pendulum controlled by a force $u(t)$, that is a dynamical system subjected to the differential constraint $\ddot{x}(t) = -x(t) + u(t)$. Assume also that the force $u(t)$ is bound to take values in $K = [-1, 1] \subset \mathbb{R}$ and that the initial position for the pendulum is constrained to be $x(0) = 0$ and the initial velocity $\dot{x}(0)$ is bound to be in $[-v_{\max}, v_{\max}]$ for some $v_{\max} > 0$. We want to discuss the Mayer problem corresponding to finding a time-dependent force $u(t)$, under which the position $x(t)$ at $t = T$ is maximised or, equivalently, the terminal cost $C(x(T)) := -x(T)$ is minimised.

The classical approach to such a problem is the following. First, let us reduce the differential constraint to a system of first order. This can be done by introducing an auxiliary variable, i.e. by representing the dynamical system with curves $(x^1(t), x^2(t)) \in \mathbb{R}^2$ with $x^1(t) = x(t)$ and $x^2(t) = \dot{x}(t)$. In this way the evolution of the system is described by curves $(t, x^1(t), x^2(t))$, subjected to the differential constraints

$$\begin{aligned} \dot{x}^1(t) &= x^2(t) , & x^1(0) &= 0 . \\ \dot{x}^2(t) &= -x^1(t) + u(t) , \end{aligned} \tag{8.1}$$

In these coordinates the terminal cost is determined by the function $C(x^1, x^2) = -x^1$.

Second, the Pontryagin auxiliary variables $p_1(t)$ and $p_2(t)$ are introduced and the evolutions of the system is now described by curves $\gamma(t) = (t, x^1(t), x^2(t), p_1(t), p_2(t))$ in $\mathbb{R} \times \mathcal{Q}$, $\mathcal{Q} = \mathbb{R}^4$, constrained by the (8.1) and, at the same time, by

$$\begin{aligned} \dot{p}_1(t) &= p_2(t) , & p_1(T) &= -\frac{\partial C}{\partial x^1}|_{x^1(T)} = 1 , \\ \dot{p}_2(t) &= -p_1(t) , & p_2(T) &= -\frac{\partial C}{\partial x^2}|_{x^1(T)} = 0 . \end{aligned} \tag{8.2}$$

The (8.2) and the (8.1) are uncoupled. This allows to determine explicitly the components $p_i(t)$ for each curve $\gamma(t) = (t, x^1(t), x^2(t), p_1(t), p_2(t))$. They are

$$p_1(t) = \cos(T - t) , \quad p_2(t) = \sin(T - t) .$$

Due to this, for each given $\gamma_o(t) = (t, x_o^1(t), x_o^2(t), p_{o1}(t), p_{o2}(t))$ satisfying the above constraints, the associated Pontryagin function $\mathcal{H} : K \rightarrow \mathbb{R}$ takes the form

$$\mathcal{H}(\omega) = \cos(T - t)x_o^2(t) + \sin(T - t)(-x_o^1(t) + \omega) .$$

Then, the usual PMP implies that an optimal control curve $u_o(t)$ must satisfy the following conditions:

$$u_o(t) = +1 \quad \text{when } \sin(T - t) > 0 , \quad (8.3)$$

$$u_o(t) = -1 \quad \text{when } \sin(T - t) < 0 .$$

The meaning of this result is the following: *for any initial velocity $v \in [-v_{max}, v_{max}]$, if $u(t)$ is different from (8.3), there exists at least one needle modification $\tilde{u}(t)$ of $u(t)$, such that the controlled curve determined by the pair $\tilde{U} = (\tilde{u}(t), \sigma = (x(0) = 0, \dot{x}(0) = v))$ has a cost which is smaller than the cost of the curve determined by $U = (u(t), \sigma = (x(0) = 0, \dot{x}(0) = v))$. Thus, if an optimal control $(u_o(t), \sigma)$ exists, it must have $u_o(t)$ as in (8.3). After establishing this, by varying only the initial datum σ (i.e. the initial velocity v_o), it is quite straightforward to find that an optimal control exists and it is $U = (u_o(t), \sigma = (x(0) = 0, \dot{x}(0) = v_{max}))$.*

Let us now see how our results offer an alternative way to solve the same problem. Let us consider just one auxiliary variable, say p , and the controlled Lagrangian on $J^5(\mathcal{Q}|\mathbb{R}) \times K$, $\mathcal{Q} = \mathbb{R}$, $K = [-1, 1]$, of actual order $r = 2$:

$$L(t, x, \dot{x}, \ddot{x}, p, u) := p(\ddot{x} + x - u) . \quad (8.4)$$

The corresponding controlled Euler-Lagrange equations are only two and are

$$\ddot{x} + x - u = 0 , \quad \ddot{p} + p = 0 . \quad (8.5)$$

Our problem can be now recognised as the generalised Mayer problem determined by the triple (\mathcal{K}, L, C) , where $C : J^5(\mathcal{Q}|\mathbb{R}) \rightarrow \mathbb{R}$ is the smooth function $C(j_t^5(x)) = -x(t)$ and \mathcal{K} is the set of pairs $U = (u(t), \sigma)$, given by a control curve taking values in $K = [-1, 1]$ and initial conditions $\sigma = j_{t=0}^5(\gamma) = (x(0), \dot{x}(0), \ddot{x}(0), \dots, p(0), \dot{p}(0), \ddot{p}(0), \dots)$ subjected to no restrictions except for the value $x(0) = 0$.

In this setting, given a controlled curve $\gamma_o = \gamma^{(U_o)}$, the condition (7.26), which characterises the needle variations $\text{Needle}^{(\tau, \omega, \varepsilon_o, \Sigma)}(\gamma_o) = \{F^{(\tau, \omega, \varepsilon, \Sigma)}, 0 < \varepsilon \leq \varepsilon_o\}$ in the class $\text{GoodN}(\gamma_o)$ is very simple and we give it in the next formula (8.6), where we denote by $(x(t, s), p(t, s), \dot{x}(t, s), \dot{p}(t, s), \dots)$ the components of the jets homotopy $F^{(\tau, \omega, \varepsilon, \Sigma)^{(5)}}(t, s)$, associated with the curve $U(s) \in \mathcal{K}$

$$U(s) = (u(t, s), (x(0, s) = 0, p(0, s), \dot{x}(0, s) = 0, \dot{p}(0, s), \dots)) .$$

It is

$$\int_0^1 \left(\frac{\partial x}{\partial s} \Big|_{(t=T, s)} - p(T, s) \frac{\partial \dot{x}}{\partial s} \Big|_{(t=T, s)} + \dot{p}(T, s) \frac{\partial x}{\partial s} \Big|_{(t=T, s)} \right) ds \geq 0 . \quad (8.6)$$

Since we are free to choose any initial condition for the variable p , we may consider a family $\Sigma = \Sigma(\varepsilon, s)$ of initial data so that the corresponding functions $p(t, s)$ satisfies the terminal conditions

$$p(T, s) = 0 , \quad \dot{p}(T, s) = -1 . \quad (8.7)$$

For such Σ , the condition (8.6) is automatically satisfied. This means that, for any choice of τ , ω and ε_o , the class $\text{GoodN}(\gamma_o)$ is not empty and contains all needle variations satisfying (8.7). Note also that (8.7) together with the differential constraint $\ddot{p} + p = 0$ completely determines the function $t \mapsto p(t, s)$ which is

$$p(t, s) = \sin(T - t) \quad \text{for any } s . \quad (8.8)$$

On the other hand, given a controlled curve $\gamma_o(t) = (t, x_o(t), p_o(t))$ with $p_o(t)$ as in (8.8), the corresponding function $\mathcal{P}_{j^2(\gamma_o)} : K \rightarrow \mathbb{R}$ is

$$\mathcal{P}_{j^2(\gamma_o)}(\omega) = -\sin(T - t)(\ddot{x} + x - \omega) . \tag{8.9}$$

From the above observations and our Generalised Pontryagin Maximum Principle (Corollary 7.7), it follows that a curve $\gamma_o(t)$, determined by a control curve $u_o(t)$ is a solution to our Mayer problem only if it is as in (8.3), as prescribed by the classical approach a la Pontryagin. As before, this condition completely determines a (non-smooth) optimal control curve $u_o(t)$ and, by the same argument of before, it follows that there is an optimal control and it is the pair $(u_o(t), \sigma_o)$ with $\sigma_o := (x(0) = 0, \dot{x}(0) = v_{\max})$.

Remark 8.1. The control curve $u_o(t)$ determined via (8.9) is clearly non smooth and hence not fitting the simplifying assumptions adopted throughout this paper. However, using the results of this paper and standard approximation techniques, one can rigorously prove that the above (discontinuous) curve $u_o(t)$ is the only measurable control curve, for which there exists no needle modification corresponding to a curve with a strictly smaller terminal cost. A detailed proof of this claim is given in [4].

We stress the fact that this second (new) approach to the addressed Mayer problem involves *just one* auxiliary variable, instead of the three used in the classical approach. In fact, the same circle of ideas can easily find solutions to the large class of similar Mayer problems with one control variable $u \in [-1, 1]$, one dependent variable $x(t)$, $t \in [0, T]$, constrained by a differential problem of order m of the form

$$\sum_{\ell=0}^m a_\ell \frac{d^\ell x}{dt^\ell} = u , \quad x(0) = \frac{dx}{dt} \Big|_{t=0} = \dots = \frac{d^{m-1}x}{dt^{m-1}} \Big|_{t=0} = 0$$

with constant coefficients a_ℓ , and the cost function $C = -x$. By considering just one auxiliary variable p and the controlled Lagrangian $L(t, x(\ell), p) := p(\sum_{\ell=0}^m a_\ell x(\ell) - u)$, one can find the optimal control $u_o(t)$, $t \in [0, T]$, of such a Mayer problem with the same arguments of before. It is $u_o(t) = \text{sign}(p_o(t))$, where $p_o(t)$ is the unique solution to the differential problem

$$\sum_{\ell=0}^m (-1)^\ell a_\ell \frac{d^\ell p}{dt^\ell} = 0 \quad \text{in } [0, T]$$

with terminal conditions $p(T) = \frac{dp}{dt} \Big|_{t=T} = \dots = \frac{d^{N-2}p}{dt^{N-2}} \Big|_{t=T} = 0 , \quad \frac{d^{N-1}p}{dt^{N-1}} \Big|_{t=T} = -1 .$ (8.10)

Of course, the same solution can be easily found also using the classical approach and the classical PMP, provided that, instead of the above single auxiliary variable p , one introduces and handles $2m - 1$ auxiliary variables: In fact, one needs $m - 1$ auxiliary variables to reduce the constraint to a system of first order equations, and the m Pontryagin auxiliary variables p_j .

We conclude showing that there exists also a third different approach to the above linearised pendulum problem. In fact, with no need of auxiliary variables, this problem

can be immediately seen to be a generalised Mayer problem (in the sense of this paper) with differential constraints given by the controlled Lagrangian

$$L : J^3(\mathbb{R}|\mathbb{R}) \times K \rightarrow \mathbb{R} , \quad L(t, x, \dot{x}, u) := \frac{1}{2} \dot{x}^2 - \frac{1}{2} x^2 + ux , \quad K = [-1, 1] . \quad (8.11)$$

We now restrict to the cases in which $T \neq \pi k$ for any $k \in \mathbb{N}$. Under this assumption, a direct inspection of the general solutions of the differential constraints shows that, for a given control map $u : [0, T] \rightarrow K$, the map

$$\varphi : \mathbb{R} \rightarrow \mathbb{R} , \quad \varphi(v) := x^{(U)}(T) , \quad U = (u(t), \sigma = (x(0) = 0, \dot{x}(0) = v)) , \quad (8.12)$$

is surjective (here, as usual, we denote by $x^{(U)}(t)$ the unique solution of the differential constraints with initial datum σ and control curve $u(t)$). Due to the particularly simple form of the cost function, this information immediately leads to the conclusion that *if there exists an optimal control $U_o = (u_o(t), \sigma_o)$, then the initial velocity must take an extremal value*. However, there is also an enlightening (but longer) argument based on the generalised PMP yielding to the same conclusion. Using the surjectivity of the map (8.12), one can see that for $v \neq \pm v_{\max}$ and $u(t)$ arbitrary, there exist *good* needle variation for any pair $(\tau, \omega) \in (0, T) \times K$, provided that ω is sufficiently close to $u(\tau)$. From this and the explicit expression of the map $\mathcal{P}_{j_{\tau}^5(\gamma(v))} : K \rightarrow \mathbb{R}$, it follows that if $v \in (0, v_{\max})$ and if $u(\tau) \neq -\text{sign}(x(\tau))$ at some $\tau \in (0, T)$, then there is *an appropriate needle modification and an appropriate change of the initial velocity*, which gives a curve with a strictly lower terminal cost. By studying the terminal costs determined by the control curves of the form $u(t) = -\text{sign}(x(t))$, one concludes (once again) that the desired optimal control $U_o = (u_o(t), \sigma_o)$ must have initial velocity $v_o = \pm v_{\max}$.

It now remains to look for an optimal control U_o in the restricted family of pairs $U = (u(t), \sigma_o)$ in which the initial value is completely fixed $\sigma_o = (x(0) = 0, \dot{x}(0) = \pm v_{\max})$. In this subclass, $\dot{x}(0) = v_{\max}$ is no longer modifiable and there is no way of constructing *good* needle variations. Our Corollary 7.7 gives no exploitable condition in this situation. Thus, the easiest way to conclude is now to go back to one of the previous approaches and determine explicitly the control curve $u_o(t)$ using auxiliary variables. Nonetheless, with a not so big effort (the details are omitted just for not weighing the reader down), all terms in (7.12) can be made fully explicit for any needle variation and, from Theorem 7.5, the previously determined optimal control is once again derived.

8.2. Concluding remarks.

In Sect. 8.1 we showed how a very simple case of a control problem with a differential constraint of higher order can be studied with no need of reducing it to an equivalent system of first order constraints. The same approach can be adopted in much less trivial situations. Just to fix the ideas, consider a control problem determined by a differential constraint of the third order

$$\ddot{x} = f(t, x, \dot{x}, \ddot{x}, u) \quad (8.13)$$

with $u = (u^a)$ varying in a compact set $K \subset \mathbb{R}^m$ and terminal cost function $C = C(x(T), \dot{x}(T), \ddot{x}(T))$. By introducing a dual variable p , the constraint (8.13) can be

identified as one of the Euler-Lagrange equations of the controlled Lagrangian

$$L(t, x, \dot{x}, \ddot{x}, p, u) := p(\ddot{x} - f(t, x, \dot{x}, \ddot{x}, u)) . \quad (8.14)$$

Indeed, the Euler-Lagrange equation corresponding to p is precisely (8.13), while the Euler-Lagrange equation corresponding to x is

$$\begin{aligned} \frac{\partial L}{\partial x} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) + \frac{d^2}{dt^2} \left(\frac{\partial L}{\partial \ddot{x}} \right) - \frac{d^3}{dt^3} \left(\frac{\partial L}{\partial \ddot{\ddot{x}}} \right) = \\ = p \left(-\frac{\partial f}{\partial x} + \frac{d}{dt} \left(\frac{\partial f}{\partial \dot{x}} \right) - \frac{d^2}{dt^2} \left(\frac{\partial f}{\partial \ddot{x}} \right) \right) + \dot{p} \left(\frac{\partial f}{\partial \dot{x}} - 2 \frac{d}{dt} \left(\frac{\partial f}{\partial \ddot{x}} \right) \right) - \ddot{p} \frac{\partial f}{\partial \ddot{x}} - \ddot{\ddot{p}} = 0 . \end{aligned} \quad (8.15)$$

Given a control curve $u_o(t)$ and an initial condition σ_o , there is a uniquely associated controlled curve $\gamma(t) = (t, x(t), p(t))$, in which $x(t)$ satisfies (8.13) and the initial condition $j_{t=0}^2(x) = \sigma_o$, and $p(t)$ satisfies (8.15) and the terminal conditions

$$\begin{aligned} p(T) &= \frac{\partial C}{\partial \ddot{x}} \Big|_{j_T^2(x)} , & \dot{p}(T) &= -\frac{\partial C}{\partial \dot{x}} \Big|_{j_T^2(x)} - p(T) \frac{\partial f}{\partial \ddot{x}} \Big|_{(j_T^2(x), u(T))} , \\ \ddot{p}(T) &= \frac{\partial C}{\partial x} \Big|_{j_T^2(x)} + p(T) \frac{\partial f}{\partial \dot{x}} \Big|_{(j_T^2(x), u(T))} - \dot{p}(T) \frac{\partial f}{\partial \ddot{x}} \Big|_{(j_T^2(x), u(T))} - p(T) \frac{d}{dt} \left(\frac{\partial f}{\partial \ddot{x}} \right) \Big|_{(j_T^2(x), u(T))} . \end{aligned} \quad (8.16)$$

These conditions are the higher order analogues of the terminal condition (7.30) used in the classical case of first order differential constraints. In fact, they are specially designed to make the left hand side of (7.26) equal to 0 for any generalised needle variation that consists of homotopies of curves satisfying the above two-point boundary problem (see [4]). This means that any such generalised needle variation is in $Good\mathcal{N}(\gamma)$ and that the generalised PMP given by Corollary 7.7 holds at all points of a solution $\gamma(t) = (t, x(t), p(t))$ of the above problem, in perfect analogy with the case of differential constraints of first order.

We stress the fact that, for each control curve $u(t)$ and associated solution $x(t)$ to (8.14), the corresponding curve $p(t)$ is completely determined by the terminal conditions (8.16) and the *linear* equation (8.15) with time-dependent coefficients. Such linearity of the equation for p is a general and quite useful fact. Indeed, it holds *for any control problem determined by an r -th order differential system of n equations the form*

$$\frac{d^r x^i}{dt^r} = f^i \left(t, x^j(t), \dots, \frac{d^{r-1} x^j}{dt^{r-1}}, u^a(t) \right) , \quad (8.17)$$

provided that such constraints are coupled with all other Euler-Lagrange equations of the controlled Lagrangian

$$L \left(t, x^j(t), \dots, \frac{d^{r-1} x^j}{dt^{r-1}}, p_j, u^a \right) := p_i \left(x_{(r)}^i - f^i \left(t, x^j(t), \dots, \frac{d^{r-1} x^j}{dt^{r-1}}, u^a \right) \right) . \quad (8.18)$$

On the other hand, as we mentioned in Remark 8.1, standard approximation techniques yield to versions of the above generalised PMP that hold under much weaker regularity assumptions ([4]). We expect that the above described two-point boundary problems together with such extended versions of the PMP can be helpful also for the direct computations of the optimal controls under higher order constraints. Investigations on this aspect are left to future work.

REFERENCES

- [1] A. A. Agrachev and Yu. L. Sachkov, *Control theory from the geometric viewpoint* in “Encyclopaedia of Mathematical Sciences, 87. Control Theory and Optimization, II.”, Springer-Verlag, Berlin, 2004.
- [2] A. Bressan and B. Piccoli, Introduction to the mathematical theory of control, *American Institute of Mathematical Sciences (AIMS), Springfield, MO*, 2007.
- [3] F. Cardin and A. Spiro, *Pontryagin maximum principle and Stokes theorem*, *J. Geom. Phys.* **142**, (2019), 274–286.
- [4] F. Cardin, C. Giannotti and A. Spiro, *On the Pontryagin Maximum Principle under differential constraints of higher order*, in preparation.
- [5] L. Cesari, Optimization—theory and applications, *Springer-Verlag, New York*, 1983.
- [6] E. Fiorani and A. Spiro *Lie algebras of conservation laws of variational ordinary differential equations*, *J. Geom. Phys.*, **88** (2015), 56–75.
- [7] V. Jurdjevic, Geometric control theory, *Cambridge University Press, Cambridge*, 1997.
- [8] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mishchenko, The mathematical theory of optimal processes (translated from Russian by D. E. Brown), *A Pergamon Press Book, The Macmillan Co., New York*, 1964.
- [9] D.J. Saunders, The geometry of jet bundles, *Cambridge University Press, Cambridge*, 1989.
- [10] A. Spiro, *Cohomology of Lagrange complexes invariant under pseudogroups of local transformations*, *Int. J. Geom. Methods Mod. Phys.*, **4** (2007), 669–705.
- [11] H.J. Sussmann and J.C. Willems, Three Centuries of Curve Minimization: From the Brachistochrone to Modern Optimal Control Theory, 2003 (*a book downloadable from www.math.rutgers.edu/~sussmann/papers/main-draft.ps.gz*).

FRANCO CARDIN
 DIPARTIMENTO DI MATEMATICA TULLIO LEVI-CIVITA
 UNIVERSITÀ DEGLI STUDI DI PADOVA,
 VIA TRIESTE, 63,
 I-35121 PADUA
 ITALY

CRISTINA GIANNOTTI & ANDREA SPIRO
 SCUOLA DI SCIENZE E TECNOLOGIE
 UNIVERSITÀ DI CAMERINO
 VIA MADONNA DELLE CARCERI
 I-62032 CAMERINO (MACERATA)
 ITALY

E-mail: cardin@math.unipd.it

E-mail: cristina.giannotti@unicam.it
E-mail: andrea.spiro@unicam.it