

# Bifurcation preserving discretisations of optimal control problems

Christian Offen\* and Sina Ober-Blöbaum\*

\* Department of Mathematics, Paderborn University, Germany  
(e-mail: christian.offen@uni-paderborn.de)

**Abstract:** The first order optimality conditions of optimal control problems (OCPs) can be regarded as boundary value problems for Hamiltonian systems. Variational or symplectic discretisation methods are classically known for their excellent long term behaviour. As boundary value problems are posed on intervals of fixed, moderate length, it is not immediately clear whether methods can profit from structure preservation in this context. When parameters are present, solutions can undergo bifurcations, for instance, two solutions can merge and annihilate one another as parameters are varied. We will show that generic bifurcations of an OCP are preserved under discretisation when the OCP is either directly discretised to a discrete OCP (direct method) or translated into a Hamiltonian boundary value problem using first order necessary conditions of optimality which is then solved using a symplectic integrator (indirect method). Moreover, certain bifurcations break when a non-symplectic scheme is used. The general phenomenon is illustrated on the example of a cut locus of an ellipsoid.

Copyright © 2021 The Authors. This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

**Keywords:** catastrophe theory, bifurcations, variational methods, symplectic integrators

## 1. INTRODUCTION

There are two main strategies to discretise optimal control problems (OCPs): *direct* and *indirect* methods. In direct methods the OCP is approximated by a discrete optimisation problem, which is then solved using techniques from nonlinear programming. In indirect methods first order necessary conditions for optimality are calculated for the OCP. These have the structure of a boundary value problem for a Hamiltonian system (Pontryagin's principle). The boundary value problem is solved numerically using methods such as shooting algorithms or implicit solvers for the fully discretised problem. For this, Hamilton's equations need to be discretised. If a symplectic partitioned Runge-Kutta method is used, then the scheme is mathematically equivalent to a direct method, in which the state equation is integrated with the underlying Runge-Kutta method. If, on the other hand, a non-symplectic integrator is used, then the scheme cannot be obtained as a direct discretisation method (Ober-Blöbaum et al., 2011).

While there is some mathematical beauty in the fact that forming first order necessary conditions for optimality and discretisation commutes, provided that a symplectic integration scheme is used, any practical relevance of structure preservation in this context may not be immediately clear. Indeed, it has been argued that for OCPs symplectic integrators have no advantages over non-symplectic schemes with the exception of some special cases (Chyba et al., 2009). However, in this paper we show that using structure preserving integration schemes can be crucial when *bifurcation phenomena* occur.

Solutions to first order necessary conditions of parameter-dependent OCPs may not be unique but bifurcate as

parameters are varied: for instance, two solutions can merge and annihilate one another or three solutions can interact. More specifically, we will focus on families of OCPs where a cost function

$$S(u; \mu) = \int_{t_0}^{t_N} L(q(t), u(t); \mu) dt \quad (1)$$

is extremised subject to a state equation

$$\dot{q} = f(q, u; \mu), \quad q(t_0) = q_0(\mu), \quad q(t_N) = q_N(\mu) \quad (2)$$

among all admissible controls  $u$ . Here  $\mu \in \Lambda$  is the parameter of the family of OCPs and  $\Lambda$  is the parameter space. The parameter is fixed during optimisation.

An analysis of the bifurcation behaviour of solutions to OCPs helps to determine for which parameter values a unique optimal solution exists and in which parameter ranges there are several solutions which fulfil first order optimality conditions. These bifurcation phenomena should be contrasted to bifurcations analysed in the literature related to branching due to low regularity (Kogan, 1986). Here, we restrict to a description of bifurcation phenomena of regular solutions which do not interact with boundaries of the state or control space.

The main example of the paper is the computation of shortest paths connecting two points  $q_0$  and  $q_N$  on an ellipsoid. There can be several connecting curves which extremise the length functional on an ellipsoid. These are geodesics. We will formulate the problem of finding geodesics starting at  $q_0$  as a family of optimal control problems, where the parameter  $\mu$  is given by the end point  $q_N$ . As  $q_N$  is varied while  $q_0$  is kept fix, the solutions bifurcate. The set of  $q_N$ , for which a bifurcation takes place, forms the *cut locus* to  $q_0$ . We will use popular discretisation methods from optimal control theory to compute cut loci and demonstrate that direct methods as

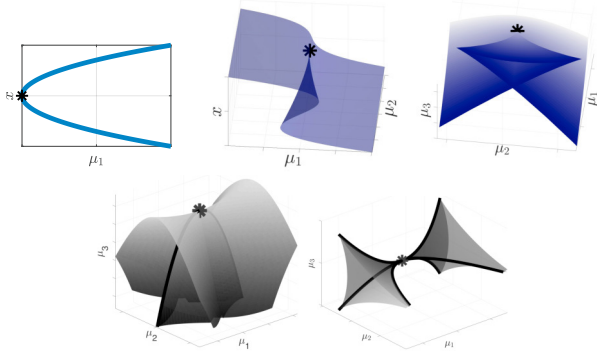


Fig. 1. The plots show generic behaviour of critical points of functions  $r_\mu: \mathbb{R}^k \rightarrow \mathbb{R}$  when parameters  $\mu$  are varied. From left to right, top to bottom we see models for fold, cusp, swallowtail, hyperbolic umbilic, and elliptic umbilic bifurcations. The most singular points are marked by \*. They are persistent under small perturbations of  $r_\mu$ . See Offen (2019) for animations.

well as indirect methods in combination with symplectic integrators resolve the loci correctly while non-symplectic integrators make qualitative errors.

Classical catastrophe theory considers the bifurcation behaviour of critical points of smooth, parameter dependent functions  $r_\mu: \mathbb{R}^k \rightarrow \mathbb{R}$ , where  $\mu \in \mathbb{R}^l$  is a parameter. Stable bifurcations have been classified, see, for instance (Arnold et al., 1998). Figure 1 shows the first five elementary catastrophes. We relate bifurcations of solutions to first order necessary conditions of OCPs to classical catastrophe theory and explain why only direct methods and indirect methods in combination with symplectic integrators can preserve all stable bifurcations under discretisation.

The article is structured as follows. To exemplify the theoretical aspects of this work and to show their significance for optimal control theory, we introduce the main example of the paper, the computation of cut loci, and provide an optimal control formulation of the problem in section 2. In section 3 we apply typical discretisation schemes to the optimal control formulation. Section 4 contains a numerical experiment, in which the conjugate locus of an ellipsoid is computed with the different discretisation schemes. The experiment demonstrates that some discretisation approaches preserve important qualitative aspects of the conjugate locus, while others break them. That the observed behaviour is prototypical for a large class of OCPs and discretisation schemes is proved in section 5. The section, furthermore, connects bifurcations in OCPs to one of the authors' classification results for bifurcations in Hamiltonian boundary value problems (McLachlan and Offen, 2018a, 2019, 2020; Offen, 2020), on which this work is based. Section 6 summarises the findings.

## 2. GEODESICS ON SUBMANIFOLDS AS OPTIMAL CONTROL PROBLEMS

To prepare the computation of cut loci, we formulate the geodesic equation on Riemannian submanifolds as variational problems and OCPs.

### 2.1 Variational formulation

Consider a submanifold  $M \subset \mathbb{R}^n$  given as the zero level set  $M = g(0)$  of a smooth function  $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ , where  $m \leq n$  and the Jacobian matrix of  $g$  at each point in  $M$  is of maximal rank. Let  $\|\cdot\|$  denote the Euclidean norm in  $\mathbb{R}^n$ . A curve  $q \in \mathcal{C}^\infty([0, 1], M) \subset \mathcal{C}^\infty([0, 1], \mathbb{R}^n)$  is a geodesic on  $M$  that connects  $q(0) = q_0$ ,  $q(1) = q_N$  for  $q_0, q_N \in M$  if the length functional

$$S(q) = \frac{1}{2} \int_0^1 \|\dot{q}(t)\|^2 dt$$

is stationary at  $q$  among all curves of  $\mathcal{C}^\infty([0, 1], M)$  connecting  $q_0$  and  $q_N$ . More precisely,

$$\delta S(q)(v) = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (S(q + \epsilon v) - S(q)) = 0$$

for all  $v \in \mathcal{C}^\infty([0, 1], M)$  with  $v(0) = 0 = v(1)$ . Equivalently,  $q \in \mathcal{C}^\infty([0, 1], \mathbb{R}^n)$  with  $q(0) = q_0$ ,  $q(1) = q_N$  is a geodesic on  $M$  if there exists a Lagrangian multiplier  $\lambda \in \mathcal{C}^\infty([0, 1], \mathbb{R}^m)$  such that  $(q, \lambda)$  is a stationary point of the extended functional

$$\bar{S}(q, \lambda) = \int_0^1 \left( \frac{1}{2} \|\dot{q}(t)\|^2 - g(q(t))^\top \lambda(t) \right) dt,$$

i.e.  $\delta \bar{S}(q, \lambda)(v, w) = 0$  for all variations  $v \in \{v \in \mathcal{C}^\infty([0, 1], \mathbb{R}^n) \mid v(0) = 0 = v(1)\}$  and  $w \in \mathcal{C}^\infty([0, 1], \mathbb{R}^m)$ . Here,  $g(q(t))^\top$  denotes the transposition of  $g(q(t))$ .

Using partial integration and the fundamental theorem of variational calculus on the condition  $\delta \bar{S}(q, \lambda)(v, w) = 0$  shows that state and Lagrangian multiplier  $(q, \lambda)$  constitute a stationary point of  $\bar{S}$  if and only if the boundary conditions and constrained Euler-Lagrange equations

$$\ddot{q} = -g'(q)^\top \lambda, \quad g(q) = 0, \quad q(0) = q_0, \quad q(1) = q_N \quad (3)$$

are fulfilled.

### 2.2 Optimal control formulation

The variational principle  $\delta \bar{S} = 0$  or the constrained equations (3) could be discretised directly. However, we would like to analyse which discretisation strategies for OCPs preserve qualitative aspects of the solution space to OCPs, which admit several extremal solutions. To generate an example which is simple on the one hand but rich enough on the other hand to exhibit the considered phenomena, we will cast the problem of finding geodesics as an OCP as follows.

A cost functional

$$J(u, \lambda) = \int_0^1 \left( \frac{1}{2} \|u(t)\|^2 - g(q(t))^\top \lambda(t) \right) dt$$

is extremised among all controls  $(u, \lambda) \in \mathcal{C}^\infty([0, 1], \mathbb{R}^n \times \mathbb{R}^m)$  with associated states  $q \in \mathcal{C}^\infty([0, 1], \mathbb{R}^n)$  that fulfil the state equation

$$\dot{q} = u, \quad \text{subject to} \quad q(0) = q_0, \quad q(1) = q_N.$$

*Proposition 1.* To any optimal control  $(u, \lambda) \in \mathcal{C}^\infty([0, 1], \mathbb{R}^n \times \mathbb{R}^m)$  and its associated state trajectory  $q$  there exists a costate trajectory  $p$  such that

$$\dot{q} = u, \quad \dot{p} = -g'(q)^\top \lambda \quad (4)$$

subject to the algebraic constraints

$$g(q) = 0, \quad p = u \quad (5)$$

and the boundary conditions  $q(0) = q_0$ ,  $q(1) = q_N$ .

**Proof.** By Pontryagin's principle (Liberzon, 2012) there exists a scalar  $p_0 \leq 0$  and a costate trajectory  $p: [0, 1] \rightarrow \mathbb{R}^n$  with  $(p_0, p) \neq (0, 0)$  such that the optimal control  $(u, \lambda)$  and its associated state trajectory  $q$  fulfil Hamilton's equations

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q} \quad (6)$$

for the Hamiltonian

$$H(q, p, u, \lambda) = p^\top u + p_0 \left( \frac{1}{2} \|u\|^2 - g(q)^\top \lambda \right)$$

and, as the optimal control variables  $(u, \lambda)$  are assumed to take values in the open set  $\mathbb{R}^n \times \mathbb{R}^m$ , the optimality condition  $\frac{\partial H}{\partial(u, \lambda)} = 0$  holds. In particular

$$0 = \frac{\partial H}{\partial u} = p + p_0 u.$$

If the abnormal multiplier  $p_0$  is zero, then  $p \equiv 0$  which contradicts the non-triviality condition  $(p_0, p) \neq (0, 0)$ . After rescaling, if necessary, we can assume  $p_0 = -1$ . Now (4) and (5) are obtained from (6) and the optimality condition  $\frac{\partial H}{\partial(u, \lambda)} = 0$ .  $\square$

Relations (4) and (5) constitute first order necessary conditions for optimal controls  $(u, \lambda) \in C^\infty([0, 1], \mathbb{R}^n \times \mathbb{R}^m)$ . This yields the following constrained boundary value problem for optimal state and costate trajectories:

$$\begin{aligned} \dot{q} &= p, \quad \dot{p} = -g'(q)^\top \lambda \\ g(q) &= 0, \quad q(0) = q_0, \quad q(1) = q_N. \end{aligned} \quad (7)$$

Notice that (7) recovers (3).

### 3. DISCRETISATION

We now apply several popular discretisation schemes to the different formulations of the geodesic problem.

#### 3.1 Discrete Euler-Lagrange equations

We discretise the action  $\bar{S}$  to

$$\bar{S}_\Delta = \frac{1}{2} \frac{\|q_1 - q_0\|^2}{\Delta t} + \sum_{k=1}^{N-1} \left( \frac{1}{2} \frac{\|q_{k+1} - q_k\|^2}{\Delta t} - \Delta t g(q_k)^\top \lambda_k \right), \quad (8)$$

where  $\Delta t > 0$  is a discretisation parameter. The values  $(q_k, \lambda_k) \in \mathbb{R}^n \times \mathbb{R}^m$  are interpreted as an approximation to  $(q(k\Delta t), \lambda(k\Delta t))$  for  $k = 1, \dots, N-1$ . The values  $q_0$  and  $q_N$  are determined by the boundary condition. An extremum  $\{(q_k, \lambda_k)\}_{k=1}^{N-1}$  fulfils

$$\begin{aligned} 0 &= \frac{\partial \bar{S}_\Delta}{\partial q_k} = -\frac{q_{k+1} - 2q_k + q_{k-1}}{\Delta t} - \Delta t g'(q_k)^\top \lambda_k \\ 0 &= \frac{\partial \bar{S}_\Delta}{\partial \lambda_k} = -g(q_k) \end{aligned}$$

with  $k = 1, \dots, N-1$ . This induces a scheme in which  $q_{k+1}$  can be computed from  $q_k$  and  $q_{k-1}$  by solving the  $n+m$ -dimensional system

$$0 = \frac{q_{k+1} - 2q_k + q_{k-1}}{\Delta t} + \Delta t g'(q_k)^\top \lambda_k, \quad 0 = g(q_{k+1}) \quad (9)$$

for  $q_{k+1}$  and  $\lambda_k$ . If  $q_0$  and  $q_N$  are given, then a collection of the formulas (9) with  $k = 1, \dots, N-1$  together with  $g(q_1) = 0$  constitutes an  $(N-1)(n+m)$ -dimensional

system of nonlinear equations<sup>1</sup> which can be solved numerically with an iterative method. However, to reduce dimensionality *shooting methods* are usually preferred: for fixed  $q_0$ , an iteration of (9) for  $k = 1, \dots, N-2$  yields a map  $q_1 \mapsto (q_{N-2}, q_{N-1})$ . A composition with the map  $(q_{N-2}, q_{N-1}, \lambda_{N-1}) \mapsto q_N$  with

$$q_N = 2q_{N-1} - q_{N-2} - \Delta t^2 g'(q_{N-1})^\top \lambda_{N-1}$$

yields a map  $\psi: (q_1, \lambda_{N-1}) \mapsto q_N$ . To given  $q_N$ , the value  $q_1$  can be computed numerically from the  $n+m$ -dimensional system  $\psi(q_1, \lambda_{N-1}) - q_N = 0$ ,  $g(q_1) = 0$ . Finally, obtain  $\{q_k\}_{k=1}^{N-1}$  from (9).

If, on the other hand, instead of the boundary values  $q_0$  and  $q_N$  an initial state  $q_0$  and momentum  $p_0$  (corresponding to tangential velocity) are given, then  $q_1$  can be obtained by a discrete Legendre transformation (Marsden and West, 2001) by solving

$$p_0 = \frac{q_1 - q_0}{\Delta t} + g'(q_0)^\top \tilde{\lambda}_0, \quad g(q_1) = 0$$

for  $q_1$  and the Lagrange multiplier  $\tilde{\lambda}_0$ . Afterwards,  $\{q_k\}_{k=1}^N$  are obtained from (9).

#### 3.2 Indirect method with symplectic discretisation

We discretise (7) by a symplectic integrator, for instance, the symplectic Euler-Method:

$$\begin{aligned} q_{k+1} &= q_k + \Delta t p_k \\ p_{k+1} &= p_k - \Delta t g'(q_{k+1})^\top \lambda_{k+1}, \quad 0 = g(q_{k+1}) \end{aligned}$$

In the above scheme, the variables  $p_k$  can be eliminated such that

$$\frac{q_{k+1} - 2q_k + q_{k-1}}{\Delta t} + \Delta t g'(q_k)^\top \lambda_k = 0.$$

The scheme is, therefore, equivalent to (9). The values for  $p_k$  can be computed in a post-processing step, if required.

#### 3.3 Indirect method with non-symplectic discretisation

We discretise (7) by the (non-symplectic) explicit midpoint rule:

$$\begin{aligned} q_{k+1} &= q_k + \Delta t \left( p_k - \frac{\Delta t}{2} g'(q_k)^\top \lambda_k \right), \quad 0 = g(q_{k+1}) \\ p_{k+1} &= p_k - \Delta t g' \left( q_k + \frac{\Delta t}{2} p_k \right)^\top \lambda_k \end{aligned} \quad (10)$$

We will see later that this scheme is *not* equivalent to (9) because it has different preservation properties.

#### 3.4 Karush–Kuhn–Tucker condition (direct method)

When using the Karush–Kuhn–Tucker condition (KKT) for discretisation, we do *not* discretise the first order optimality conditions (7) but discretise the cost functional  $J$  first and then derive first order optimality conditions for the discrete OCP (direct method). For this, we first apply a numerical method to the state equation  $\dot{q} = u$ . The Euler

<sup>1</sup> The condition  $g(q_{k+1}) = 0$  is removed from the last instance  $k = N$  of (9) as it is fulfilled by assumption.

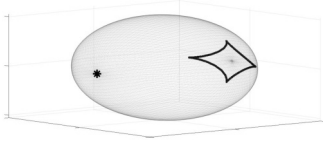


Fig. 2. The cut locus  $C$  on a 2-dimensional ellipsoid  $E$  with respect to a point  $q_0 \in E$  marked by  $*$ . Denote the connected components of  $E \setminus C$  by  $E_0$  and  $E_1$  such that  $q_0 \in E_0$ . To any point  $q_1 \in E_1$  there are three geodesics connecting  $q_0$  with  $q_1$  with length at most by  $\frac{3}{2}\pi$ , while there is only one such geodesic if  $q_1 \in E_0$ . As  $q_1$  is moved from  $E_1$  to  $E_0$  through a line of  $C$ , two of the geodesics merge and annihilate (fold bifurcation). If  $q_1$  is moved through one of the four corners of  $C$ , three of the geodesics merge into one in a cusp bifurcation. If  $q_0$  is in general position,  $C$  always consists of four cusps connected by fold lines (Itoh and Kiyohara, 2004).

method yields  $q_{k+1} = q_k + \Delta t u_k$ . Next, the cost functional  $J$  is discretised to

$$J_\Delta = \sum_{k=0}^{N-1} \frac{1}{2} \|u_k\|^2 - g(q_k)^\top \lambda_k + \mu_k^\top (q_{k+1} - q_k - \Delta t u_k),$$

where the discretised state equation has been incorporated into the discrete action using Lagrange multipliers  $\mu_k$ . Then  $J_\Delta$  is extremised. We obtain

$$0 = \frac{\partial J_\Delta}{\partial \mu_k} = q_{k+1} - q_k - \Delta t u_k, \quad 0 = \frac{\partial J_\Delta}{\partial u_k} = u_k - \Delta t \mu_k$$

$$0 = \frac{\partial J_\Delta}{\partial q_k} = -g'(q_k)^\top \lambda_k - \mu_k + \mu_{k-1}, \quad 0 = \frac{\partial J_\Delta}{\partial \lambda_k} = -g(q_k)^\top.$$

for  $k = 1, \dots, N-1$ . Eliminating  $u_k = \Delta t \mu_k$  and  $\mu_k = (q_{k+1} - q_k)/\Delta t^2$  we can recover the scheme (9).

#### 4. NUMERICAL EXPERIMENT

The shortest path that connects two points on a (complete and connected) Riemannian manifold is always a geodesic. Moreover, each geodesic  $\gamma: [0, T] \rightarrow M$  is locally length minimising, i.e. there exists an  $\epsilon > 0$  such that for all  $t \in [0, \epsilon]$  the geodesic  $\gamma$  is the shortest path connecting  $\gamma(0)$  with  $\gamma(t)$ . If  $\epsilon$  is maximal with the property that  $\gamma|_{[0, \epsilon]}$  is length minimising, then  $\gamma(\epsilon)$  is a *conjugate point* to  $\gamma(0)$ . The set of all conjugate points to a reference point  $q_0$  is referred to as the *cut locus* to  $q_0$  (see (Flaherty and do Carmo, 1992) for exact definitions). The cut locus of a 2-dimensional ellipsoid in  $\mathbb{R}^3$  is displayed in figure 2.

A computation of a cut locus to a point  $q_0$  on a Riemannian manifold  $M$  given as a level set of a smooth function  $g$  proceeds as follows.

- Given a step-size  $\Delta t > 0$  and  $N \in \mathbb{N}$  such that  $1 = \Delta t N$  the scheme (9) provides a map  $M \ni q_1 \mapsto q_N \in M$ . An open neighbourhood of  $q_0 \in M$  can be identified with an open neighbourhood  $\mathcal{O}$  of  $0 \in \mathbb{R}^{m-n}$  such that the scheme induces a computable map  $\phi: \mathcal{O} \rightarrow M$ . The map  $\phi$  corresponds to an expression of the classical geodesic exponential map in a chart.
- We compute the set of critical points  $C_0 \subset \mathcal{O}$  of  $\phi$ . For this, we consider the determinant  $\det \phi'$  of the

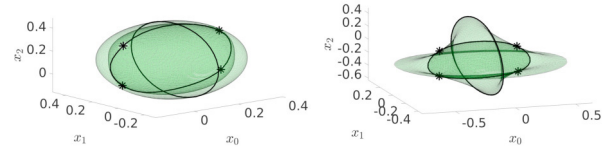


Fig. 3. The first figure shows the critical set  $C_0$  and the second figure the conjugate locus of a 3-dimensional ellipsoid. Each point on the sheets corresponds to a (preimage of a) fold singularity. The solid lines correspond to (preimages of) lines of cusp singularities. At the points marked by  $*$  there are umbilic singularities which we investigate closer in figure 4. The cusp lines and locations of umbilic singularities have been computed using techniques from (Kreusser et al., 2020; McLachlan and Offen, 2018b)

Jacobian  $\phi'$  of  $\phi$ , whereas the Jacobian of  $\phi$  is computed using automatic differentiation. The critical set  $C_0$  can be computed by evaluating  $\phi$  on a mesh on  $\mathcal{O}$  and then using a level-set method (such as `contour` or `isosurface` from Python's `plotly` package or MATLABs `contourf` or `isosurface`). Alternatively,  $C_0$  can be computed by a level set continuation method such as pseudo-arclength continuation, if  $C_0$  is one-dimensional, or by manifold continuation methods (Krauskopf et al., 2007).

- The critical set  $C_0$  is mapped with  $\phi$  to the set of critical values  $C$ , which is the cut locus.
- Additionally, highly degenerate points within  $C_0$  (such as cusps, swallowtail points, etc.) can be computed using techniques such as those developed by one of the authors in (Kreusser et al., 2020).

Figure 3 shows the conjugate locus with respect to a typical point of a three-dimensional ellipsoid considered as a submanifold of  $\mathbb{R}^4$ . Only the first three components  $x_0, x_1, x_2$  are plotted. The last component  $x_3$  can be recovered from  $x_0, x_1, x_2$ . A theoretical description of the structure of loci of high-dimensional ellipsoids has recently been given in (Itoh and Kiyohara, 2020). Numerical computations on a normal form of a high-dimensional ellipsoid can be found in (Joets and Ribotta, 1999).

Figure 4 shows (rotated) close-ups of the cut locus near one of the hyperbolic umbilic points. Only the variational scheme (9) correctly captures the hyperbolic umbilic bifurcation. The non-variational method (10) breaks the bifurcation, contains an artificial hole, and two spurious highly degenerate points at the locations where the line of cusp bifurcations touches the sheet of folds. The structural error of the non-variational scheme can also be seen in the computed critical sets  $C_0$ , i.e. the preimage on the locus. While the first plot of figure 5 corresponding to the variational scheme correctly shows two sheets intersecting in one point, in the second plot two sheets connect along a circle. This demonstrates the significance of structure preservation for the computation of solutions to variational problems such as OCPs when bifurcations occur. Source code is available in (Offen, 2021).



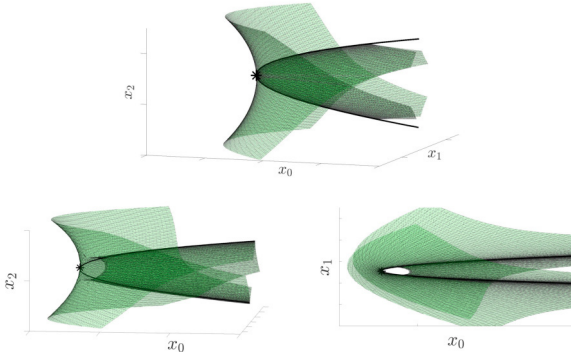


Fig. 4. The plots show the conjugate locus close to an umbilic bifurcation point. While the top figure was obtained by the variational scheme (9), the other two plots correspond to an experiment with (10). While the locus to (9) correctly shows a hyperbolic umbilic bifurcation, the locus to (10) is broken and contains an artificial hole and does *not* contain a hyperbolic umbilic singularity. Indeed, the most singular point marked by  $*$  corresponds to a simple fold singularity rather than an umbilic singularity. See Offen (2019) for animations that have been obtained by investigating normal forms of classical catastrophes.

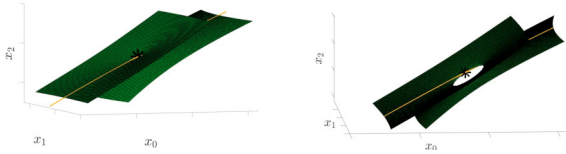


Fig. 5. The plots show the preimage  $C_0$  of the conjugate locus close to an umbilic bifurcation point. The left figure was obtained by the variational scheme (9) and correctly shows two tangentially intersecting sheets. The right figure was obtained by the non-variational scheme (10) and does not capture the situation well.

## 5. THEORETICAL CONSIDERATIONS

We will now develop a theoretical framework to predict which local bifurcations occur generically in parameter dependent OCPs and to explain the different behaviour of symplectic and non-symplectic discretisation schemes. Using Pontryagin's principle, the local bifurcation behaviour of regular solutions to necessary conditions of OCPs will be translated to the bifurcation behaviour of solutions to Hamiltonian boundary value problems, which are related to catastrophe theory (McLachlan and Offen, 2018a; Offen, 2020). In a neighbourhood of a smooth solution the infinite-dimensional setting of OCPs is, thus, reduced to a setting to which finite-dimensional theory applies.

Let the state space be given by a smooth manifold  $M$  without boundary and let the control space  $U$  and parameter space  $\Lambda$  be open subsets  $U \subset \mathbb{R}^l$ ,  $\Lambda \subset \mathbb{R}^k$ . For a given parameter  $\mu \in \Lambda$  we consider the extremisation of the cost function  $S$  from (1) among smooth controls  $u \in \mathcal{C}^\infty([t_0, t_N], U)$  subject to the state equation (2). The parameter  $\mu$  is fixed during the extremisation process. Here  $q_0(\mu), q_N(\mu) \in M$  and for each  $u \in U$  the function  $f(\cdot, u; \mu): M \rightarrow TM$  is a smooth vectorfield on  $M$ .

Moreover,  $L$  is smooth and all data depends smoothly on the parameter  $\mu$ . Let  $\pi: T^*M \rightarrow M$  denote the cotangent bundle projection and let  $\langle \cdot, \cdot \rangle$  be the pairing of  $T^*M$  and  $TM$ . Assume that for  $\mu^* \in \Lambda$  there exists a control  $u^* \in \mathcal{C}^\infty([t_0, t_N], U)$  with associated state space trajectory  $q^* \in \mathcal{C}^\infty([t_0, t_N], M)$  which extremises  $S(\cdot; \mu^*)$ . By Pontryagin's principle (Barbero-Liñán and Muñoz-Lecanda, 2008) there exists an abnormal multiplier  $p_0^* \leq 0$  and a costate trajectory  $\alpha^*: [t_0, t_N] \rightarrow T^*M$  with  $\pi(\alpha^*) = q^*$  and  $(p_0^*, \alpha^*) \neq (0, 0_{T^*M})$ , where  $0_{T^*M}$  is the zero section in  $T^*M$ , such that  $\alpha^*$  is a trajectory of the Hamiltonian vectorfield  $X_{H_{u^*, \mu^*}}$  defined by the Hamiltonian  $H: T^*M \times U \times \Lambda \rightarrow \mathbb{R}$ ,  $H(\alpha, u; \mu) = \langle \alpha, f(q, u; \mu) \rangle + p_0 L(q, u; \mu)$ . In other words, the Hamiltonian boundary value problem

$$\begin{aligned} \frac{d}{dt} \alpha(t) &= X_{H_{u, \mu}}(\alpha(t)) \\ \pi(\alpha(t_0)) &= q_0(\mu), \quad \pi(\alpha(t_N)) = q_N(\mu) \end{aligned} \quad (11)$$

is fulfilled for  $(\alpha, u; \mu) = (\alpha^*, u^*; \mu^*)$ . Here  $X_{H_{u, \mu}}$  denotes the Hamiltonian vectorfield to the Hamiltonian  $H_{u, \mu}(\alpha) = H(\alpha, u; \mu)$ . Moreover, since the controls  $u$  take values in the open set  $U \subset \mathbb{R}^l$ , the optimality condition

$$\frac{\partial H}{\partial u}(\alpha(t), u(t); \mu) = 0, \quad \forall t \in [t_0, t_N] \quad (12)$$

holds true for  $(\alpha, u; \mu) = (\alpha^*, u^*; \mu^*)$ . The Hamiltonian boundary value problem (11) and the optimality condition (12) constitute first order necessary conditions for the OCP (1).

**Theorem 2.** Assume that the Hessian matrix  $\frac{\partial^2 H}{\partial u^i \partial u^j}$  is invertible. All stable, local bifurcations of smooth solutions to the first order necessary conditions with an abnormal multiplier  $p_0 \neq 0$  of the OCP (1) are governed by catastrophe theory. Direct discretisation methods for OCPs as well as symplectic discretisation methods for the first order necessary conditions preserve all stable bifurcations.

**Remark 3.** • As the constraint considered in the OCP for the computation of a conjugate locus of a 3-dimensional ellipsoid  $M$  is holonomic, the problem is equivalent to a problem covered by theorem 2, where the ellipsoid is the manifold  $M$ . The parameters enters only in the boundary condition  $q(t_N) = q_N(\mu)$ , where  $q_N: \Lambda \rightarrow M$  is a local chart of  $M$  with  $\Lambda \subset \mathbb{R}^3$ .

- Stability in this context means that the bifurcations are persistent under small perturbations of the necessary condition (11) within the class of Hamiltonian boundary value problems, or, slightly more generally, the class of boundary value problems for symplectic maps. Admissible perturbations of the OCP (1) include small perturbations of the state equation, the Lagrangian  $L$  as well as the boundary condition for  $q$  (whereas care needs to be taken if the perturbed boundary condition is allowed to involve  $\dot{q}$ , which leads to the notion of Lagrangian boundary conditions (Offen, 2020)).
- The non-degeneracy assumption on  $\frac{\partial^2 H}{\partial u^i \partial u^j}$  only needs to hold in a tubular neighbourhood of the graph of a solution  $(\alpha^*, u^*; \mu^*)$  for the theorem to hold close to  $(q^*, u^*; \mu^*)$ .
- If the state space  $M$  or the control space  $U$  contain boundaries, then the theorem can be applied locally by restricting to a tubular neighbourhood of the graph of  $(q^*, u^*; \mu^*)$  if the image of the curves  $q^*, u^*$  do not intersect with the boundaries.

**Proof.** Let  $(\alpha^*, u^*; \mu^*)$  be a solution to the first order necessary conditions (11) and (12). Applying the implicit function theorem to  $\frac{\partial H}{\partial u}(\alpha, u; \mu) = 0$ , there exists a unique function  $\bar{u}$  depending on  $(t, \alpha; \mu)$  such that  $\bar{u}(t, \alpha^*(t); \mu^*) = u^*(t)$  and  $\frac{\partial H}{\partial u}(\alpha, u(t, \alpha; \mu); \mu) = 0$ . The function  $\bar{u}$  is defined on the Cartesian product of a tubular neighbourhood of the graph of  $(\alpha^*, u^*)$  and an open neighbourhood of  $\mu^*$  in  $\Lambda$ . Close to  $(\alpha^*, u^*; \mu^*)$  the first order necessary conditions (11) and (12) are, therefore, equivalent to the (time-dependent) Hamiltonian boundary value problem  $\frac{d}{dt}\alpha(t) = X_{H_{\bar{u}, \mu}}(\alpha(t))$ ,  $\pi(\alpha(t_0)) = q_0(\mu)$ ,  $\pi(\alpha(t_N)) = q_N(\mu)$ . The flow map of  $X_{H_{\bar{u}, \mu}}$  is a symplectic map such that we obtain a family of boundary value problems for symplectic maps  $\{\phi_\mu\}_{\mu \in \Lambda^*}$ , where  $\Lambda^* \subset \Lambda$  is an open neighbourhood of  $\mu^*$ . The local bifurcation behaviour of solutions to such systems was related to catastrophe theory in (McLachlan and Offen, 2018a; Offen, 2020). In an indirect discretisation method of (1), Hamilton's equations (11) are discretised and the control is (as before) obtained from the optimality condition (12). If a symplectic integrator is used in (11), then we obtain a family of boundary value problems for symplectic maps  $\bar{\phi}_\mu$  close to  $\phi_\mu$ . A stable, catastrophe type bifurcation of the boundary value problem for  $\phi_\mu$  is present in the nearby problem for  $\bar{\phi}$  as well for sufficiently small discretisation parameters by the definition of stability. As direct discretisation methods for OCPs correspond to indirect methods with symplectic integration (Ober-Blöbaum et al., 2011), the conclusion also holds for direct methods.  $\square$

As shown in (McLachlan and Offen, 2020),  $D$ -series bifurcations, such as hyperbolic umbilic bifurcations, are stable bifurcations in families of boundary value problems for symplectic maps but unstable in more general classes of boundary value problems. Other bifurcations, such as fold, cusp, which belong to the  $A$ -series, are also stable in wider classes of boundary value problems. Therefore, the correspondence of regular OCP and symplectic boundary value problems, provided by the proof of theorem 2, explains our observations from the numerical example that hyperbolic umbilic bifurcations are preserved when using a structure preserving discretisation schemes, while the sheets of fold bifurcations and lines of cusp singularities persist even if variational or symplectic structure is destroyed under discretisation.

## 6. CONCLUSIONS AND DISCUSSION

Solutions to first order necessary conditions of families of optimal control problems can undergo bifurcations as parameters are varied. Under regularity assumptions we showed that local bifurcations which are persistent under small perturbations of the family of OCPs are exactly the classical catastrophes. Moreover, to preserve all stable bifurcations under discretisation, either direct discretisation methods for OCPs or indirect methods in combination with symplectic integrators can be used. If, however, discretisation methods are used which are not structure preserving, then certain bifurcations, such as  $D$ -series bifurcations, break. A preservation of bifurcations is necessary when computing bifurcation diagrams to determine in which parameter ranges how many first order optimal solutions exist.

## REFERENCES

- Arnold, V.I., Goryunov, V.V., Lyashko, O.V., and Vasil'ev, V.A. (1998). *Critical Points of Functions*, 10–50. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Barbero-Liñán, M. and Muñoz-Lecanda, M. (2008). Geometric approach to Pontryagin's maximum principle. *Acta Applicandae Mathematicae*, 108(2), 429–485.
- Chyba, M., Hairer, E., and Vilmart, G. (2009). The role of symplectic integrators in optimal control. *Optimal Control Applications and Methods*, 30(4), 367–382.
- Flaherty, F. and do Carmo, M. (1992). *Riemannian Geometry*. Mathematics: Theory & Applications. Birkhäuser Boston.
- Itoh, J. and Kiyohara, K. (2004). The cut loci and the conjugate loci on ellipsoids. *manuscripta mathematica*, 114(2), 247–264.
- Itoh, J. and Kiyohara, K. (2020). The structure of the conjugate locus of a general point on ellipsoids and certain liouville manifolds. *Arnold Mathematical Journal*.
- Joets, A. and Ribotta, R. (1999). Caustique de la surface ellipsoïdale à trois dimensions. *Experimental Mathematics*, 8(1), 49 – 55.
- Kogan, J. (1986). *Bifurcation of Extremals in Optimal Control*. Springer Berlin Heidelberg.
- Krauskopf, B., Osinga, H.M., and Galán-Vioque, J. (eds.) (2007). *Numerical Continuation Methods for Dynamical Systems*. Springer Netherlands.
- Kreusser, L.M., McLachlan, R.I., and Offen, C. (2020). Detection of high codimensional bifurcations in variational PDEs. *Nonlinearity*, 33(5), 2335–2363.
- Liberzon, D. (2012). *The Maximum Principle*, 102–155. Princeton University Press.
- Marsden, J.E. and West, M. (2001). Discrete mechanics and variational integrators. *Acta Numerica*, 10, 357–514.
- McLachlan, R.I. and Offen, C. (2018a). Bifurcation of solutions to Hamiltonian boundary value problems. *Nonlinearity*, 31(6), 2895–2927.
- McLachlan, R.I. and Offen, C. (2018b). Hamiltonian boundary value problems, conformal symplectic symmetries, and conjugate loci. *New Zealand Journal of Mathematics (NZJM)*, 48, 83–99.
- McLachlan, R.I. and Offen, C. (2019). Symplectic integration of boundary value problems. *Numerical Algorithms*, 81(4), 1219–1233.
- McLachlan, R.I. and Offen, C. (2020). Preservation of bifurcations of Hamiltonian boundary value problems under discretisation. *Foundations of Computational Mathematics (FoCM)*, 20, 1363–1400.
- Ober-Blöbaum, S., Junge, O., and Marsden, J.E. (2011). Discrete mechanics and optimal control: an analysis. *ESAIM: Control, Optimisation and Calculus of Variations*, 17(2), 322–352.
- Offen, C. (2019). Singularities animations. [https://www.youtube.com/playlist?list=PLIp-UrjLTJ5m-3ZASHPurIkehiBuW\\_s0](https://www.youtube.com/playlist?list=PLIp-UrjLTJ5m-3ZASHPurIkehiBuW_s0). Accessed 2021-05-02.
- Offen, C. (2020). *Analysis of Hamiltonian boundary value problems and symplectic integration (Doctoral Thesis)*. Massey University.
- Offen, C. (2021). Release v1.0 of GitHub repository Christian-Offen/ConjugateLocus. URL <https://doi.org/10.5281/zenodo.4562664>.