Solving two-point boundary value problems using the Hamilton-Jacobi theory

V. M. Guibout
MBDA
Chief System Engineer
Avenue Louis Breguet
MBDA\

D. J. Scheeres
The University of Michigan
Aerospace department
Beal Avenue
USA\

Abstract: This paper focuses on applications of a general methodology that we developed for solving two-point boundary value problems. Using the Hamilton-Jacobi theory in conjunction with canonical transformation induced by the phase flow, we proved that the generating functions for this transformation solve any two-point boundary value problem in phase space. This method has implications in different fields as boundary value problems are widely spread and we succeeded to develop an efficient algorithm. We illustrate its use to solve optimal control problems without an initial guess, study the phase structure and plan spacecraft formation flights.

Key–Words: Boundary value problems, Halmilton-Jaccobi theory, Optimal control, Phase space structure

1 Introduction

One of the most famous two-point boundary value problems in astrodynamics is Lambert's problem, which consists of finding a trajectory in the two-body problem which goes through two given points in a given lapse of time. Even though the two-body problem is integrable, no analytical solution has been found to this problem so far, and solving Lambert's problem still requires one to solve Kepler's equation, which has motivated many papers since 1650 [3]. For a general Hamiltonian dynamical system, a two-point boundary value problem is solved using iterative techniques such as shooting and relaxation methods. The shooting method [9] consists of choosing values for all of the dependent variables at one boundary. These values must be consistent with any boundary conditions for that boundary, but otherwise are initially guessed randomly. After integration of the differential equations, we in general find discrepancies between the desired boundary values at the other boundary. Then, we adjust the initial guess to reduce these discrepancies and reiterate this procedure again. The method provides a systematic approach to solving boundary value problems, but suffers several inherent limitations. As summarized by Bryson and Ho ([2] p 214),

> The main difficulty with this method is getting started; i.e., finding a first estimate of the unspecified conditions at one end that produces a solution reasonably close to the specified conditions at the other end. The reason for this peculiar difficulty is that the extremal solutions are often very sensitive to small changes in the unspecified boundary conditions.

To get rid of the sensitivity to small changes in initial guesses, techniques such as the multiple shooting method [8] were developed. They consist of breaking the time domain into segments and solving a boundary value problem on each of these segments. In this manner, nonlinear effects are limited over each segment, but on the other hand the size of the problem is increased considerably. However, the choice of the initial conditions still remains as the main hurdle to successfully apply shooting methods to any kind of problems.

Relaxation methods [10] use a different approach. The differential equations are replaced by finite-difference equations on a mesh of points that covers the range of the integration. A trial solution consists of values for the dependent variables at each mesh point, not satisfying the desired finite-difference equations, nor necessarily even satisfying the required boundary conditions. The iteration, now called relaxation, consists of adjusting all the values on the mesh so as to bring them into successively closer agreement with the finite-difference equations and simultaneously with the boundary conditions. In general, relaxation works better than shooting when the boundary conditions are especially delicate or subtle. However, if the solution is highly oscillatory then many grid points are required for accurate representation.

Also, the number and positions of the required mesh points are not known *a priori* and must be adjusted manually for each problem. In addition, if solutions to the differential equations develop singularities, attempts to refine the mesh to improve accuracy may fail.

With the advent of computers, these two methods are able to solve most of the two-point boundary value problems. They may require substantial time to find an appropriate initial guess and/or computer memory to refine the mesh, but they often succeed. However, there are problems for which these methods reach their limits. For instance, the design of space missions involving several spacecraft in formation requires one to solve a large number of boundary value problems for which the boundary conditions may in turn depend on parameters. To reconfigure a formation of N spacecraft, there are $N!$ possibilities in general, that is, $N!$ boundary value problems need to be solved. As N increases, the number of boundary value problems dramatically grows. Similarly, suppose that we plan to reconfigure a spacecraft formation to achieve an interferometry mission. We may require the spacecraft to be equally spaced on a circle perpendicular to the line of sight they should observe. In that case, the final positions are specified in terms of the angle that indicates the position of the spacecraft on the circle. In order to find the value of the angle that minimizes fuel expenditure, infinitely many boundary value problems may need to be solved. As a result, the algorithms mentioned above are no longer appropriate as they require excessive computation and time. To address these complex problems arising in spacecraft formation design, Guibout and Scheeres [6] developed a novel approach to solving boundary value problems. This approach outperforms traditional methods for spacecraft formation design and was generalized in [7]. They proved that it allows to formally solve any kind of two-point boundary value problem at a cost of a single function evaluation once generating functions are known.

In this paper we analyze a variety of problems in several fields and show that they can all be formulated as Hamiltonian two-point boundary value problems. Then we review the general methodology that we developed for solving two-point boundary value problems and explain in details how we build up a software tool from it.

2 Hamiltonian two-point boundary value problems

In this section, we explore a variety of problem in several fields and show that they can formulate as Hamiltonian two-point boundary value problems that we solve using our algorithms.

2.1 Defintion

A Hamiltonian two-point boundary value problem is a problem that can be formulated as a Hamiltonian dynamical system whose initial conditions are not completely known. For instance, we may know the position at an initial time and at a later time but have no knowledge of the initial velocity. In that case, we cannot integrate the differential equations describing the system although the knowledge of the position at two different times may be enough to uniquely define the initial velocity. Finding the missing initial conditions is the key issue addressed by two-point boundary value problems.

We now provide a more formal definition.

Definition 1 (Hamiltonian system). *A system is called Hamiltonian if there exists a smooth function $H(q,p,t)$ from $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ to $\mathbb{R}$ such that its dynamics can be described by equations of the form:*

$$\dot{q_i} = \frac{\partial H}{\partial p_i}\,,\ \dot{p_i} = -\frac{\partial H}{\partial q_i}\,. \tag{1}$$

H is called the Hamiltonian function, (q,p) is the state vector that describes the phase space and Eqns. (1) *are known as Hamilton's equations.*

The formalism of Hamiltonian systems covers a wide class of fields, from dynamical problems with conservative forces to optimal control problems. Indeed, any system with conservative forces may be formulated as an Hamiltonian system (often q is the position and p the momentum) and the necessary conditions for optimality transform an optimal control problem into an Hamiltonian problem (q is the state and p is the co-state).

Consider a Hamiltonian function H, two points in phase space, $X_0 = (q_0, p_0)$ and $X_1 = (q,p)$, and two partitions of $(1,\cdots,n)$ into two non-intersecting parts, $(i_1,\cdots,i_p)(i_{p+1},\cdots,i_n)$ and $(k_1,\cdots,k_r)(k_{r+1},\cdots,k_n)$. A Hamiltonian two-point boundary value problem is formulated as follows:

Definition 2. *Given $2n$ coordinates $(q_{i_1},\cdots,q_{i_p},p_{i_{p+1}},\cdots,p_{i_n})$ and $(q_{0_{k_1}},\cdots,q_{0_{k_r}},p_{0_{k_{r+1}}},\cdots,p_{0_{k_n}})$, find the remaining $2n$ variables such that a particle starting at X_0 will reach X_1 in T units of time.*

The methodology that we developed allows us to compute an analytic approximation of the solution of any kind of two-point boundary value problems. In

other words, we are able to find two functions such that:

$$\begin{aligned}((\quad & p_{i_1} \quad ,\cdots,p_{i_p},q_{i_{p+1}},\cdots,q_{i_n}),\\ & (p_{0_{k_1}} \quad ,\cdots,p_{0_{k_r}},q_{0_{k_{r+1}}},\cdots,q_{0_{k_n}})) =\\ (f(\quad & q_{i_1} \quad ,\cdots,q_{i_p},p_{i_{p+1}},\cdots,p_{i_n}),\\ & g(q_{0_{k_1}} \quad ,\cdots,q_{0_{k_r}},p_{0_{k_{r+1}}},\cdots,p_{0_{k_n}})) \end{aligned} \quad (2)$$

We will see later that f and g are (up to a sign) the gradient of a generating function associated with the phase flow transformation. But first, let us look at some examples to illustrate the use of our approach in several fields.

2.2 Finding periodic orbits

Although the search for periodic orbits is naturally formulated as an initial value problem, this problem may be posed as a two-point boundary value problem.

Periodic orbits in a $2n$-dimensional Hamiltonian dynamical system are characterized by the following equations:

$$q(T) = q_0\,, \quad (3)$$
$$p(T) = p_0\,, \quad (4)$$

where T is the period of the orbit, (q_0, p_0) are the initial conditions at time $t_0 = 0$ and $(q(t), p(t))$ verifies Hamilton's equations:

$$\dot{q} = \frac{\partial H}{\partial p}(q,p,t)\,,\ \dot{p} = -\frac{\partial H}{\partial q}(q,p,t)\,. \quad (5)$$

In the most general case, the search for periodic orbits consists of solving the $2n$ equations (3) and (4) for the $2n+1$ unknowns (q_0, p_0, T). Simple methods that solve this problem take a set of initial conditions (q_0, p_0), and integrate Hamilton's equations. If there exists a time $t = T$ such that Eqns. (3) and (4) are verified, then a periodic orbit is found. Else, other initial conditions need to be guessed. In the approach we propose, instead of looking at the initial conditions and the period as the only variables of the problem, we suppose that the period, n initial conditions as well as n components of the state vector at time T are unknowns. Then the search for periodic orbits reduces to solving the $2n$ equations (3) - (4) for these $2n+1$ unknowns.

If $(q(T), q_0, T)$ are taken to be the $2n+1$ unknowns, then the search for periodic orbits consists of solving the $2n$ equations (3)-(4) for $(q(T), q_0, T)$. For instance, let us find all periodic orbits of a given period. In other words, T is given and we need to find $(q(T), q_0)$ such that $q(T) = q_0$ and $p(T) = p(0)$. This is a two-point boundary value problem with constraints that can be solved using the approach presented in this paper. Combining the assumptions that $q(T) = q_0$ and $p(T) = p_0$ together with the knowledge of the functions f and g defined as $p(T) = f(q, q_0, T)$ and $p_0 = g(q, q_0, T)$, we obtain:

$$f(q(T) = q_0, q_0, T) - g(q(T) = q_0, q_0, T) = 0\,, \quad (6)$$
$$p = p_0 = f(q(T) = q_0, q_0, T)\,. \quad (7)$$

Hence, the search for all periodic orbits of a given period is reduced to solving n equations (6) for n variables, the q_0's, and then evaluate n equations (7) to compute the corresponding momenta. $2n$ equations still need to be solved, but now n of them are decoupled. Most importantly, once f and g are known analytically, no additional integration is required.

We now illustrate the power of the proposed method to find periodic orbits of nonlinear systems. We address a non-trivial example : we study periodic orbits about the Libration point L_2 in the normalized Hill three-body problem.

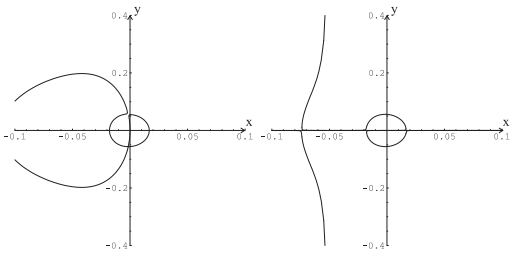
The Hill three-body problem is a three-body problem in which three main assumptions are made: 1) One of the three bodies has negligible mass compared to the other two-bodies. 2) One of the two massive bodies is in circular orbit about the other one. 3) One of the two massive bodies has larger mass than the other one. These hypothesis hold to study the motion of a spacecraft under the influence of the Sun and the Earth for example. Under these assumptions, the normalized Lagrangian for this system is

$$L(q,\dot{q}) = \frac{1}{2}(\dot{q}_x^2+\dot{q}_y^2)+\frac{1}{\sqrt{q_x^2+q_y^2}}+\frac{3}{2}q_x^2-(\dot{q}_x q_y-\dot{q}_y q_x)$$

where $(q_x, q_y) = (x, y)$. This problem has 2 equilibrium points, L_1 and L_2 whose coordinates are $L_1(-\left(\frac{1}{3}\right)^{1/3}, 0)$ and $L_2(\left(\frac{1}{3}\right)^{1/3}, 0)$.

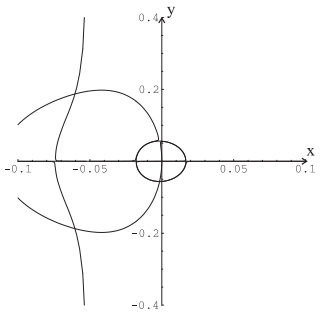
In order to apply the method presented above to find periodic orbits we need to compute the generating functions (the f and g functions). Using our algorithm we are able to find a polynomial approximation of these functions up to order 5 (the meaning of this approximation will be given later). We now search all periodic orbits of a given period $T = 3.0345$. To solve this problem we use the generating functions which defines two equations with two unknowns (the phase space is of dimension 4) that can be solved graphically. In Fig. 1, we plot the solutions to each of these two equations and then superimpose them to find their intersection. The intersection corresponds to the set of points that belongs to periodic orbits of period T. We observe that the intersection is composed of a circle and two points whose coordinates

are $(q_x, q_y) = (-0.0603795, \pm 0.187281)$. The circle is obviously a periodic orbit but the two points are not equilibrium points, and rather correspond to out-of-plane periodic orbits[1].



(a) Plot of the solution to the first equation defined by Eq. (6)

(b) Plot of the solution to the second equation defined by Eq. (6)



(c) Superposition of the two sets of solutions

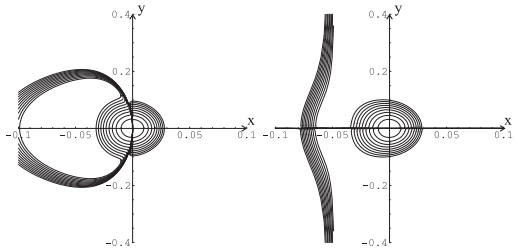
Figure 1: Periodic orbits for the nonlinear motion about a Libration point

By plotting the intersection for different periods T, we generate a map of a family of periodic orbits around the Libration point. In Fig. 2 we represent the solutions to Eq. (6) for $t = 3.033 + 0.0005n\,,\ n \in \{1 \cdots 10\}$. For $t = 3.033$ (which is less than the period of periodic orbits in the linearized system), the intersection only contains the origin, which is why there are only 9 periodic orbits shown around the origin. We note that at larger values of $x^2 + y^2$ the curves do not overlay precisely, indicating that higher order terms are needed in the analytic approximation of the solution to the two-point boundary value problem.

2.3 A multi-task mission about the Earth

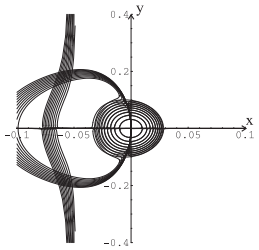
Another interesting problem to solve is a multi-task mission. We consider a spacecraft formation about an oblate earth that must achieve five missions over a one month period. For each mission the formation must be in a given configuration C_i that has been specified beforehand, and we wish to minimize the overall fuel expenditure. We will see that this problems naturally formulates as a two-point boundary value problem that can be easily solved using the approach developped in this paper.

[1]We point out that these points do not lie in the domain of convergence of our method and are only consequences of our approximation of the dynamics.



(a) Plot of the solution to the first equation defined by Eq. (6) for $t = 3.033 + 0.0005n \quad n \in \{1 \cdots 10\}$

(b) Plot of the solution to the second equation defined by Eq. (6) for $t = 3.033 + 0.0005n \quad n \in \{1 \cdots 10\}$



(c) Superposition of the two sets of solutions

Figure 2: Periodic orbits for the nonlinear motion about a Libration point

The motion of a satellite under the influence of the Earth modeled by an oblate sphere (J_2 and J_3 gravity coefficients are taken into account) in the fixed coordinate system (x, y, z) whose origin is the Earth center of mass is described by the following Hamiltonian:

$$H = \frac{1}{2}(p_x^2 + p_y^2 + p_z^2) - \frac{1}{\sqrt{x^2+y^2+z^2}}\left[1 - \frac{R^2}{2r_0^2(x^2+y^2+z^2)}\left(3\frac{z^2}{x^2+y^2+z^2} - 1\right)J_2 - \frac{R^3}{2r_0^3(x^2+y^2+z^2)^2}\left(5\frac{z^3}{x^2+y^2+z^2} - 3z\right)J_3\right],$$

where

$$GM = 398600.4405\ km^3 s^{-2}\,,\ R = 6378.137\ km\,,$$
$$J_2 = 1.082626675 \cdot 10^{-3}\,,\ J_3 = 2.532436 \cdot 10^{-6}\,,$$

and all the variables are normalized. Distances are normalized by r_0, the radius of the trajectory at the initial time, and the time is normalized by $\sqrt{r_0^3/GM}$. We consider a "reference" trajectory whose state is designated by (q^0, p^0) and study the relative motion of spacecraft with respect to it. The reference trajectory is chosen to be highly eccentric and inclined, but

any other choice could have been considered. At the initial time its state is:

$$\begin{array}{ll} q_x^0 = r_p\,, & p_x^0 = 0\ kms^{-1}\,, \\ q_y^0 = 0\ km\,, & p_y^0 = \sqrt{\frac{GM}{\frac{1}{2}(r_a+r_p)}}\sqrt{\frac{r_a}{r_p}}\cos(\alpha)\ km\,s^{-1}\,, \\ q_z^0 = 0\ km\,, & p_z^0 = \sqrt{\frac{GM}{\frac{1}{2}(r_a+r_p)}}\sqrt{\frac{r_a}{r_p}}\sin(\alpha)\ km\,s^{-1}\,, \\ \alpha = \frac{\pi}{3} rad\,, & r_p = 7,000\ km\,,\ r_a = 13,000\ km\,. \end{array}$$

Without the J_2 and J_3 gravity coefficients the reference trajectory would be an elliptic orbit with eccentricity $e = 0.3$, inclination $i = \pi/3\ rad$, argument of perigee $\omega = 0$, longitude of the ascending node $\Omega = 0$, semi-minor axis $r_p = 7,000\ km$, semi-major axis $r_a = 13,000\ km$ and of period $t_p = 2\pi\sqrt{\frac{1}{2^3}\frac{(r_a+r_p)^3}{r_p^3}}\ sec \approx 2\ hours\ 45\ min$. The Earth oblateness perturbation causes secular drifts and short terms oscillations in the orbital elements.
We consider four imaging satellites flying in formation about the reference trajectory. We want to plan spacecraft maneuvers over the next month knowing that they must observe the Earth, i.e., must be in a given configuration C_i at the following instants (chosen arbitrarily for our study):

$$\begin{array}{lll} t_0 = 0\,, & t_1 = 5\ d\ 22\ h\,, & t_2 = 10\ d\ 20\ h\,, \\ t_3 = 16\ d\ 2\ h\,, & t_4 = 21\ d\ 14\ h\,, & t_5 = 26\ d\ 20\ h\,. \end{array}$$

Define the local horizontal by the unit vectors $(\hat{e}_1, \hat{e}_2)$ such that $\hat{e}_2$ is along $r^0 \times v^0$ and $\hat{e}_1$ is along $\hat{e}_2 \times r^0$. At every t_i, the configuration C_i is defined by the four following relative positions (or slots):

$$\begin{array}{ll} q^1 = 700\ m\ \hat{e}_1\,, & q^2 = -700\ m\ \hat{e}_1\,, \\ q^3 = 700\ m\ \hat{e}_2\,, & q^4 = -700\ m\ \hat{e}_2\,. \end{array} \qquad (8)$$

Note that at t_i, q^1 is in front of the reference state (in the local horizontal plane), q^2 is behind, q^3 is on the left and q^4 is on the right (see Fig. 3). At each t_i, there must be one spacecraft per slot and we want to determine the sequence of reconfigurations that minimizes the total fuel expenditure (other cost functions such as equal fuel consumption for each spacecraft may be considered as well). For the first mission, there are $4!$ configurations (number of permutation of the set $\{1,2,3,4\}$). For the second mission, for each of the previous $4!$ configurations, there are again $4!$ configurations, that is a total of $4!^2$ possibilities. Thus for 5 missions, there are $4!^5 = 7,962,624$ possible configurations.

In this example, we assume impulsive controls that consist of impulsive thrusts applied at $t_{i\in[0,5]}$. For each of the four spacecraft, we need to compute the velocity at t_i so that the spacecraft moves to its position specified at t_{i+1} under gravitational forces only. As a result, we must solve $5 \cdot 4! = 120$ position to position boundary value problems (given two positions at t_i and t_{i+1}, we need to compute the associated velocity). Using the knowledge of the functions f and g, this problem can be handled at the cost of only 120 function evaluations. Then, we need to evaluate the fuel expenditure (sum of the norm of all the required impulses, assuming zero relative velocities at the initial and final times) for all the permutations (there are $7,962,624$ combinations) to find the sequence that minimizes the cost function. Fig. 4 represents the number of configurations as a function of the values of the cost function. We notice that most of the configurations require at least three times more fuel than the best configuration, and less than 6% yield values of the cost function that are less than twice the value associated with the best configuration. The cost function for the optimal sequence of reconfigurations is $0.00644\ km \cdot s^{-1}$ whereas it is $0.0396\ km \cdot s^{-1}$ in the least optimal design.

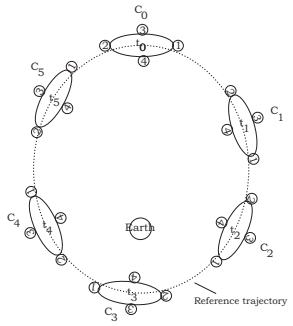


Figure 3: At each t_i, spacecraft must be in the configuration C_i

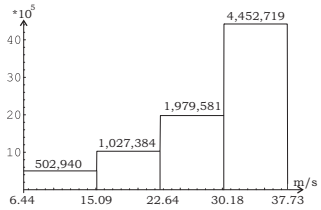


Figure 4: Number of configurations as a function of the value of the cost function

We may verify, *a posteriori*, if the solutions found meet the mission goals, i.e., if the order 4 approximation of the dynamics is sufficient to simulate the true dynamics. Explicitly comparing the analytical solution with numerically integrated results shows that the spacecraft are at the desired positions at every t_i with

a maximum error of $1.5 \cdot 10^{-8}\ km$.

2.4 The deployment problem in the Hill three-body problem

We come back to the Hill problem considered before for the study of periodic orbits. We now consider several spacecraft at L_2 at the initial time and solve the deployment problem. In other words, we want to find the optimal control laws that drive the formation from L_2 to a given configuration at $t = T$. We assume continuous thrusts and no thrust constraints. Thus, for each of the spacecraft, we need to solve an optimal control problem formulated as:

$$\min_{U=(u_x,u_y)} J = \min_{U=(u_x,u_y)} \frac{1}{2}\int_{t=0}^{t=T} (u_x^2 + u_y^2)dt\,, \quad (9)$$

subject to the dynamics:

$$\frac{\partial L}{\partial q}(q,\dot{q}) - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}(q,\dot{q}) = U\,, \quad (10)$$

and the boundary conditions:

$$X(t=0) = X_{L_2} = (3^{-1/3}, 0, 0, 0)\,,\ X(t=T) = X_T\,,$$

where $X = (q_x, q_y, \dot{q}_x, \dot{q}_y)$ and $U = (u_x, u_y)$. Necessary conditions for optimality can be found from Pontryagin's maximum principle:

$$\dot{X} = \frac{\partial H}{\partial P}\,,\ \dot{P} = -\frac{\partial H}{\partial X}\,,\ \frac{\partial H}{\partial U} = 0\,, \quad (11)$$

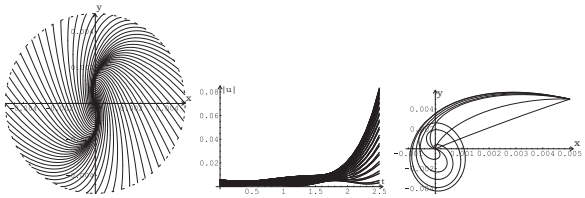
where $P = (p_1, p_2, p_3, p_4)$ and

$$H(X,P,U) = P^T\dot{X} + \frac{1}{2}u_x^2 + \frac{1}{2}u_y^2\,.$$

Then, from $\frac{\partial H}{\partial U} = 0$, we find the optimal control feedback law: $u_x = -p_3\,,\ u_y = -p_4$. We substitute $U = (u_x, u_y)$ into H to obtain $\bar{H}(X,P) = H(X,P,U(X,P))$. Thus, the necessary conditions for optimality now define a *Hamiltonian* position to position boundary value problem that can be solved using F_1:

$$\dot{X} = \frac{\partial \bar{H}}{\partial P}, \dot{P} = -\frac{\partial \bar{H}}{\partial X}, X(0) = X_{L_2}, X(T) = X_T \quad (12)$$

In Fig. 5(a) and 5(b), we plot the optimal control trajectories and the norm of the optimal control laws for different final positions X_T on a circle of radius $r = 0.05$ (10,700 km in the Earth-Sun system) and a transfer time of $t = 2.5$ (i.e., about 145 days in the Earth-Sun system). We observe that some values of the final position requires less fuel, they correspond to $X_T = r\cos(\theta) + r\sin(\theta)$ where $\theta = \{19\pi/32, 51\pi/32\}$. Similarly, we may vary the transfer time. In Fig. 5(c), we plot the optimal trajectories for $T \in \{0.1, 1.1, 2.1, 3.1, 4.1, 5.1, 6.1\}$ (i.e., from 6 to 290 days). As T increases, the trajectory wraps around L_2 so that the spacecraft takes advantage of the geometry of the Libration point.



(a) Optimal trajectories (b) Norm of the optimal control laws (c) Optimal trajectories

Figure 5: The deployment problem

In this manner, we can explore the best deployment sequence. Depending on the final configuration geometry (e.g., the spacecraft must be equally spaced on a circle of radius r) and mission specifications, we are able to choose the optimal transfer time and final configuration to minimize the fuel expenditure by evaluating a set of functions.

3 General methodology to solve Hamiltonian two-point boundary value problems

In this section, we review the method developed by Guibout and Scheeres [5, 7] for solving Hamiltonian two-point boundary value problems. Then we will focus on the algorithm that effectively allows to compute the analytic solutions to two-points boundary value problems. Trajectories of Hamiltonian systems verifies Hamilton's equations 1 but can also be characterized by the following variational principle:

Theorem 3 (Modified Hamilton's principle). *Critical points of $\int_{t_0}^{t_1}(\langle p,\dot{q}\rangle - H)dt$ in the class of paths $\gamma : \mathbb{R} \to \mathbb{R}^n \times \mathbb{R}^n$ whose ends lie in the n-dimensional subspaces $q = q_0$ at $t = t_0$ and $q = q_1$ at $t = t_1$ correspond to trajectories of the Hamiltonian system whose ends are q_0 at t_0 and q_1 at t_1.*

Proof. We proceed to the computation of the variation.

$$\delta\int_\gamma (\langle p,\dot{q}\rangle - H)dt =$$
$$\int_\gamma \left(\dot{q}_i\delta p_i + p_i\delta\dot{q}_i - \frac{\partial H}{\partial q_i}\delta q_i - \frac{\partial H}{\partial p_i}\delta p_i\right)dt = [p_i\delta q_i]_{t_0}^{t_1}$$
$$+\int_\gamma \left[\left(\dot{q}_i - \frac{\partial H}{\partial p_i}\right)\delta p_i - \left(\dot{p}_i + \frac{\partial H}{\partial q_i}\right)\delta q_i\right]dt\,.$$

Therefore, since the variation vanishes at the end points, the integral curves of Hamilton's equations are the only extrema. □

We now introduce the concept of canonical transformation, a class of coordinate transformations that preserves the Hamiltonian structure of the system.

Definition 4. *A smooth map $f : \mathbb{R}^{2n}\times\mathbb{R} \to \mathbb{R}^{2n}\times\mathbb{R}$ is a canonical transformation from (q, p, t) to (Q, P, t) if and only if:*

1. *f is a diffeomorphism,*
2. *f preserves the time, i.e., there exists a function g_t such that $f(x, t) = (g_t(x), t)$,*
3. *Critical points of $\int_{t_0}^{t_1} \left(\langle P, \dot{Q}\rangle - K(Q, P, t)\right) dt$ correspond to trajectories of the Hamiltonian system, where $K(Q, P, t)$ is the Hamiltonian function expressed in the new set of coordinates.*

We consider a canonical transformation $f : (q, p, t) \mapsto (Q, P, t)$ and a Hamiltonian system defined by H. Along trajectories, we have by definition:

$$\delta \int_{t_0}^{t_1} \left(\sum_{i=1}^{n} p_i\dot{q}_i - H(q, p, t)\right) dt = 0\,, \qquad (13)$$

$$\delta \int_{t_0}^{t_1} \left(\sum_{i=1}^{n} P_i\dot{Q}_i - K(Q, P, t)\right) dt = 0\,. \qquad (14)$$

From Eqns. (13) - (14), we conclude that the integrands of the two integrals differ at most by a total time derivative of an arbitrary function F:

$$\sum_i p_i dq_i - H dt = \sum_j P_j dQ_j - K dt + dF\,. \quad (15)$$

Such a function is called a generating function for the canonical transformation f and is, *a priori*, a function of both the "old" and the "new" variables and time. The two sets of coordinates being connected by the $2n$ equations, namely, $f(q, p, t) = (Q, P, t)$, F can be reduced to a function of $2n + 1$ variables among the $4n + 1$. Hence, we can define 4^n generating functions that have n "old" variables and n "new". Among these are the four kinds defined by Goldstein [4], $F_1(q, Q, t)$, $F_2(q, P, t)$, $F_3(p, Q, t)$ and $F_4(p, P, t)$. In this paper, we focus on the generating function of the first kind, F_1 as it is the one used to solve the previous examples. In other words, we assume that (q, Q) are independent variables. Then, from $dF_1 = \frac{\partial F_1}{\partial q}dq + \frac{\partial F_1}{\partial Q}dQ + \frac{\partial F_1}{\partial t}$, Eqn. (15) simplifies to the following vector equation:

$$(p - \frac{\partial F_1}{\partial q})dq - H dt = (P + \frac{\partial F_1}{\partial Q})dQ - K dt + \frac{\partial F_1}{\partial t}$$

Hence, since (q, Q, t) are independent variables, we obtain:

$$\begin{array}{l} q = \frac{\partial F_1}{\partial p}(q, Q, t)\,,\ Q = -\frac{\partial F_1}{\partial Q}(q, Q, t)\,, \\ \frac{\partial F_1}{\partial t} + H(q, \frac{\partial F_1}{\partial p}, t) = K(Q, -\frac{\partial F_1}{\partial Q}, t)\,. \end{array} \qquad (16)$$

Let us particularize Eqns. (16) for the canonical transformation induced by the inverse of the phase flow (a proof that this transformation is canonical can be found in [1]). Such a transformation maps the state of the system at time t to its state at the initial time while preserving the time. Thus, it maps the system to a trivial one with constant Hamiltonian function that can be chosen to be 0. The associated generating function F_1 verifies Eqns. (16) where (Q, P) now denotes the initial state (q_0, p_0) and $K = 0$:

$$p = \frac{\partial F_1}{\partial q}(q, q_0, t)\,,\ p_0 = -\frac{\partial F_1}{\partial q_0}(q, q_0, t)\,, \qquad (17)$$

$$\frac{\partial F_1}{\partial t} + H(q, \frac{\partial F_1}{\partial p}, t) = 0\,. \qquad (18)$$

Given two positions q_0 and q, and a transfer time T, we immediately notice that Eqns. (17) solves the two-point boundary value problem that consists of going from q_0 to q in T units of time. Comparing the definition of f and g in the above examples with Eq. 17 yields:

$$f = \frac{\partial F_1}{\partial q}(q, q_0, t) \qquad (19)$$

$$g = -\frac{\partial F_1}{\partial q_0}(q, q_0, t) \qquad (20)$$

This remark is of prime importance since it provides us with a very general technique for solving Hamiltonian position to position boundary value problems. However, this approach relies on knowledge of F_1. In the next section, we develop an algorithm for computing this function.

4 Computing the generating functions

The Hamilton-Jacobi theory provides us with a direct approach for computing the generating functions. Indeed, it tells us that they are solutions of the Hamilton-Jacobi equation (Eqn. (18)). This is a partial differential equation that is difficult to solve in general. However, the Hamiltonian function for describing the relative motion has a particular structure that enables us to solve this differential equation.

4.1 Relative motion

Consider a Hamiltonian system with Hamiltonian function $H(q,p,t)$. Let (q_0^0, p_0^0) and (q_0^1, p_0^1) be two points in phase space such that $q_0^1 = q_0^0 + \Delta q_0$, $p_0^1 = p_0^0 + \Delta p_0$, where $(\Delta q_0, \Delta p_0)$ is small enough to guaranty the convergence of the Taylor series in Eqn. (25). We denote by (q^i, p^i) the trajectory with initial conditions (q_0^i, p_0^i), i.e.,

$$q^1 = q(q_0^1, p_0^1, t)\,,\, p^1 = p(q_0^1, p_0^1, t)\,, \qquad (21)$$

$$q^0 = q(q_0^0, p_0^0, t)\,,\, p^0 = p(q_0^0, p_0^0, t)\,. \qquad (22)$$

and we define $X^h = \begin{pmatrix} \Delta q \\ \Delta p \end{pmatrix}$ the relative state vector by:

$$X^1 = X^0 + X^h\,, \qquad (23)$$

where $X^i = \begin{pmatrix} q^i \\ p^i \end{pmatrix}$. For convenience we shall call (q^0, p^0) the reference trajectory and (q^1, p^1) the displaced trajectory .

Using our previous notation, Hamilton's equations for the displaced trajectory reads:

$$\dot{X}^0 + \dot{X}^h = J\nabla H^1\,. \qquad (24)$$

We expand the right hand side of Eqn. (24) about the reference trajectory X^0, assuming $(\Delta q, \Delta p)$ small enough for convergence of the series:

$$\nabla H(q^1,p^1,t) = \nabla H(q^0,p^0,t) + \begin{pmatrix} \frac{\partial^2 H}{\partial q^2}(q^0,p^0,t)\Delta q + \frac{\partial^2 H}{\partial q \partial p}(q^0,p^0,t)\Delta p \\ \frac{\partial^2 H}{\partial q \partial p}(q^0,p^0,t)\Delta q + \frac{\partial^2 H}{\partial p^2}(q^0,p^0,t)\Delta p \end{pmatrix} + \cdots$$

Substituting this into Eqn. (24) yields $\dot{X}^h = J\nabla H^h$, where

$$H^h(X^h,t) = \sum_{p=2}^{\infty} \sum_{\substack{i_1,\cdots,i_{2n}=0 \\ i_1+\cdots+i_{2n}=p}}^{p} \frac{1}{i_1!\cdots i_{2n}!} \frac{\partial^p H}{\partial q_1^{i_1}\cdots\partial q_n^{i_n}\partial p_1^{i_{n+1}}\cdots\partial p_n^{i_{2n}}}(q^0,p^0,t) X_1^{h\,i_1}\cdots X_{2n}^{h\,i_{2n}}$$

Thus, the dynamics of a particle relative to a known trajectory is Hamiltonian with a Hamiltonian function $H^h(X^h,t) = H^h(\Delta q, \Delta p, t)$. The coefficients of the Taylor series $\frac{1}{i!j!}\frac{\partial^{i+j}H}{\partial q^i \partial p^j}(q^0,p^0,t)$ are time varying quantities and are easily evaluated for any Hamiltonian once the reference trajectory is known.

4.2 Algorithm

We found that the Hamiltonian describing the dynamics of two particles relative to each other is a power series in its spatial variables, with time-dependent coefficients. At first glance, the associated Hamilton-Jacobi equation may appear impractical. However, if we truncate H^h, a closed-form solution for the generating functions can be found. In this section we review the solution procedure. We refer to [5] for additional details and a study of the convergence properties of our algorithm.

We assume that F_1 can be expressed as a Taylor series about the reference trajectory in its spatial variables.

$$F_1(y,t) = \sum_{q=2}^{\infty} \sum_{\substack{i_1,\cdots,i_{2n}=0 \\ i_1+\cdots+i_{2n}=q}}^{q} \frac{1}{i_1!\cdots i_{2n}!} f_{q,i_1,\cdots,i_{2n}}(t) y_1^{i_1}\cdots y_{2n}^{i_{2n}}\,, \qquad (25)$$

where $y = (\Delta q, \Delta q_0)$. We substitute this expression in the Hamilton-Jacobi equation (Eqn. (18), with $H = H^h$). The resulting equation is an ordinary differential equation that has the following structure:

$$P(y, f^{p,r}_{q,i_1,\cdots,i_{2n}}(t), \dot{f}^{p,r}_{q,i_1,\cdots,i_{2n}}(t)) = 0\,, \qquad (26)$$

where P is a series in y with time dependent coefficients that are functions of $f_{q,i_1,\cdots,i_{2n}}(t)$ and $\dot{f}_{q,i_1,\cdots,i_{2n}}(t)$. Eqn. (26) holds for all y if and only if all the coefficients of P are zero. In this manner, we transform the ordinary differential equation (26) into a set of ordinary differential equations whose solutions are the coefficients of the generating function F_1.

This approach provides us with a closed form approximation of the generating functions. However, there are inherent difficulties as generating functions may develop singularities which prevent the integration from going further (see [1, 7] for more details on singularities). Techniques that rely on the Legendre transformation have been developed [5] to bypass this problem but have a cost in terms of computation.

An alternative approach for computing F_1 has been explored in [5]. We present the main ideas of this approach as in general we combine both methods to increase performance. We suppose that $\Delta q(\Delta q_0, \Delta p_0, t)$ and $\Delta p(\Delta q_0, \Delta p_0, t)$ can be expressed as series in the initial conditions $(\Delta q_0, \Delta p_0)$ with time dependent coefficients. We truncate the series to order N and insert these into Eqn. (1). Hamilton's equations reduce to a series in $(\Delta q_0, \Delta p_0)$ whose coefficients depend on the coefficients of the series $\Delta q(\Delta q_0, \Delta p_0, t)$ and $\Delta p(\Delta q_0, \Delta p_0, t)$ and their time derivatives. By balancing terms of the same

order, we transform Hamilton's equations into a set of ordinary differential equations whose variables are the coefficients defining Δq and Δp as a series in Δq_0 and Δp_0. For linear systems, this approach recovers the state transition matrix. Then, a series inversion of the phase flow provides us with the gradient of the generating functions that can be integrated to find the generating functions.

The main advantage of this approach is that the phase flow is never singular, therefore the ordinary differential equations are always well-defined. However, this method requires us to solve more equations than the previous method and provides us with the value of F_1 at a given time only (the time at which we perform the series inversion).

In practice, we use a "combined" algorithm. The alternative approach is used to compute the phase flow over a long time span. Then, we compute the value of F_1 at a time of interest, t_1, and solve the Hamilton-Jacobi equation around t_1. For both examples, we compute the first four terms in the series expansion of F_1. We will see that they provide an accurate picture of the nonlinear dynamics about the reference trajectory.

5 Conclusion

We have shown that the knowledge of an analytical solution to two-point boundary value problem allows us to solve difficult problems in a wide range of fields. We linked such a solution with the gradient of the generating functions associated with a specific canonical transformation, the phase flow. Such a description of the phase space is superior in many ways to the traditional approach based on the initial value problem. Then we showed how to implement this novel approach under $Mathmatica^{©}$. The algorithm we developped is efficient and robust. It allows to get around singularities and to obtain a semi-analytic description of the phase flow of complex dynamical systems.

References:

[1] Vladimir I. Arnold. *Mathematical Methods of Classical Mechanics*. Springer-Verlag, 2nd edition, 1988.

[2] Arthur E. Bryson and Yu-Chi Ho. *Applied optimal control : optimization, estimation, and control*. Halsted Press, revised edition, 1975.

[3] Peter Colwell. *Solving Kepler's equation over three centuries*. Richmond, Va. : Willmann-Bell, 1993.

[4] Herbert Goldstein. *Classical Mechanics*. Addison-Wesley, 2nd edition, 1980.

[5] Vincent M. Guibout. *The Hamilton-Jacobi theory for solving two-point boundary value problems: Theory and Numerics with application to spacecraft formation flight, optimal control and the study of phase space structure*. PhD thesis, University of Michigan, 2004.

[6] Vincent M. Guibout and Daniel J. Scheeres. Solving relative two-point boundary value problems: Spacecraft formation flight transfers application. *AIAA, Journal of Control, Guidance and Dynamics*, 27(4):693–704, 2003.

[7] Vincent M. Guibout and Daniel J. Scheeres. Solving two-point boundary value problems using generating functions: Theory and applications to optimal control and the study of Hamiltonian dynamical systems. In *Modern Astrodynamics, P. Gurfil*, volume 1. Elsevier Astrodynamics series, 2006.

[8] H. B. Keller. *Numerical methods for two-point boundary value problems*. Blaisdell, 1968.

[9] David L. Powers. *Boundary value problems*. San Diego : Harcourt Brace Jovanovich, 1987.

[10] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical recipes in C, the art of scientific computing*. CAMBRIDGE UNIVERSITY PRESS, second edition, 1992.