

DESIGN OF EARTH-MARS TRANSFER TRAJECTORIES USING EVOLUTION-BRANCHING TECHNIQUE

Massimiliano Vasile, Leopold Summerer
ESA – ESTEC, Advanced Concepts Team
Keplerlaan 1, 2200 AZ, Noordwijk ZH, The Netherlands
Email: Massimiliano.Vasile@esa.int

Paolo De Pascale
Politecnico di Milano, via La Masa 34, 2158, Milano Italy

ABSTRACT

In this paper a novel global optimisation approach is used to look for potentially interesting solutions for a mission to Mars. The approach blends the characteristics of evolution programming with the systematic search, typical of branching techniques. Solutions for a roundtrip to Mars, either direct or via Venus, considering long and short stays on Mars or free-return trajectories are considered providing a comprehensive view of all the opportunities in a wide range of launch dates. Finally electric propulsion options are investigated including the possibility of using Mars' lagrangian points for a low cost capture.

INTRODUCTION

In recent years the space community has demonstrated a growing interest in global optimisation techniques as a viable tool for the design of space trajectories^{1,2,3,4,5}. However most of the global methods used so far can be classified as stochastic or heuristic approaches and in particular of the evolution-programming category, which represents only a portion of all available global methods⁶. In many cases they have proven to explore efficiently the solution space providing even unexpected optimal solutions or a number of good initial guesses useful for a further optimisation with more accurate local optimisation techniques.

Other classes of global methods like deterministic ones, as branch and bounds approaches, have received less attention even though they have demonstrated to be extremely effective in many other fields. A hybridisation of both stochastic and deterministic approaches could be beneficial to improve the effectiveness of both at solving space related problems. In this paper an analysis of a variety of Earth-Mars transfer trajectories has been performed using a particular global optimisation approach which combines a stochastic and a deterministic method.

The basic idea of this novel approach is to use a limited set of potential solutions evolving for a small number of generations, according to some specific evolution programming rules (the stochastic step), in subregions of the solution space defined by a branching procedure (the deterministic step). On the other hand the branching rules, i.e. the rules used to partition the solutions space, are functions of the outcome from the evolution step. This technique has been used to conduct an extensive search for families of potentially interesting transfer trajectories from Earth to Mars, and return, in view of future exploration and colonisation missions envisaged by the Aurora program. Several types of trajectories have been modelled including: multiple impulsive transfers, low thrust propulsion trajectories and indirect transfers, exploiting gravity assisted manoeuvres or weak stability regions in order to reduce either the transfer time or propellant consumption. Some interesting results will be presented.

GENERAL PROBLEM FORMULATION

Optimisation problems in trajectory design can be formulated, in their general form, as:

$$\begin{aligned} \min f(\mathbf{y}) \\ \mathbf{b}^l \leq \mathbf{C}(\mathbf{y}) \leq \mathbf{b}^u \\ \text{with } \mathbf{y} \in D \end{aligned} \quad (1)$$

where f is a scalar nonlinear function of a multidimensional vector \mathbf{y} defined within the domain D . The domain D is a hypercube defined by the upper and lower bounds on the components of the vector \mathbf{y} :

$$y_i \in [b_i^l, b_i^u] \quad (2)$$

The vector $\mathbf{C}(\mathbf{y})$ is formed by all nonlinear constraint functions of the vector \mathbf{y} . If problem (1) is twice continuously differentiable and presents a single solution, i.e. only one vector \mathbf{y} in the domain D minimises f and satisfies \mathbf{C} , a nonlinear programming method like sequential quadratic programming (SQP) can be efficiently used. This means implicitly that the problem must be formulated properly and cannot contain non-differentiable quantities. However even in

this case the problem may present more than one solution within the required domain D .

If the problem is either non-differentiable, i.e. no gradient method can be applied, or more than a solution is expected, a global optimisation method must be considered. The idea is to perform an extensive search of the solution space D looking for possible solutions to problem (1). In this respect the interest could be more to find a number of good initial guesses for the nonlinear programming solver, than finding the global optimum with a high level of accuracy.

Among all global methods two categories are here considered: heuristic methods and systematic methods.

Heuristic methods contain all methods that cannot be proven to find a global optimum with a predictable amount of work. Most stochastic methods are in this class. For them, it is sometimes possible to prove convergence with probability arbitrarily close to 1 but with a number arbitrarily large of function evaluations.

Systematic methods contain all methods that (in exact arithmetic) are guaranteed to find the global optimum with a predictable (deterministic) amount of work. The bound on the amount of work is anyway quite high: exponential in the problem characteristics. The simplest systematic method for bound constrained problems is grid search where all points on finer and finer grids are tested and the best point on each grid is used as a starting point for local optimisation. The number of grid points grows exponentially with the dimensions of the problem and so does the amount of work. Even though systematic methods are generally more reliable than heuristic ones they need some level of insight into the problem and the structure of the objective function, to be efficient (an exception can be made for methods based on interval analysis⁹). If the problem is represented by a black box then they may not find the global optimum in a reasonable amount of time. This is understandable if we look at the density theorem (Törn and Zilinskas 1989), which states that any method based on local information only, that converges for every continuous f to a global minimizer of f in a feasible domain D must produce a sequence of points y^1, y^2, y^3, \dots that is dense in D . A well known stochastic method is represented by Genetic Algorithms (GA) that make use of analogies to biological evolution by allowing mutations and crossing over among candidates for good local optima in the hope to derive even better ones. The original concept of Genetic Algorithms is to encode a potential solution (individual) of the problem under study, in the form of a binary string in which each binary number represents a chromosome of the "DNA" (or genotype) of the solution (or phenotype). More sophisticated genetic algorithms make use of the data structure of the

problem to encode the individual in the more appropriate way⁷.

In general all methods that resort to some heuristic ideas derived from biological evolution can be defined as evolution programming methods. An interesting concept is represented by niching-GA⁸. The basic concept is that in nature different species can exploit different niches in the environment. This translates in the formation of subpopulations with each subpopulation specialized at a subtask of the problem or exploring a subregion. Subpopulations can compete as in pure GA or cooperate.

Among systematic methods there are some that split the solution domain on the base of some local information. Each time the domain is split, a number of new branches are created, each branch corresponds to a further exploration of the solution space and each subdomain represents a node that can be expanded and explored further. If the diameter of all the nodes converges to zero, convergence of the algorithm is straightforward.

The proposed optimisation approach is composed of a stochastic and a systematic step. The stochastic step is performed using evolution programming and is meant to obtain information on the possible presence of optima in a subdomain $D_l \subseteq D$. The systematic step is performed through a branching approach and is used to partition the domain D into subdomains D_l , where the presence of an optimum is expected. Each subdomain may or may not contain the global optimum but the systematic exploration and the qualification of each subdomain on the base of the best solution found and the volume of the subdomain, allows finding a number of optima and eventually the global one. This particular hybridization can be seen as a form of *forced* niching since populations evolving in subregions form different species.

EVOLUTION-BRANCHING APPROACH

Evolution Step

Each solution y is represented by a string containing in the first m components integer values and in the remaining s components real values. This particular encoding allows the treatment of problems with a mixed integer-real data structure. A hypercube S is associated to each individual y , the hypercube, enclosing a region of the solution space surrounding the individual, is defined by a set of intervals $S = S_1 \times S_2 \times \dots \times S_n \subseteq D_l$, where S_i contains the value of the component y_i . The solution space is then explored locally by acquiring information about the landscape within each region S and globally using a population of individuals y with their associated intervals. Each individual can communicate its findings to the others in

order to evolve the entire population toward a better status.

Evolution is governed by four fundamental operators: mutation, migration, mating and filtering. The mutation operator generates a new individual perturbing randomly an old one. The mating procedure takes two individuals and generates one or two children mixing the genotypes of the two parents. Four schemes are used to mate individuals:

- Single point crossover exchanges part of the genes between the two parents;
- Arithmetic crossover generates a new individual with an interpolation of the two parents;
- Extrapolation generates a new individual on the side of the best individual between the two parents \mathbf{y}^1 and \mathbf{y}^2 at a distance from the best parents proportional to the vector connecting the two parents:

$$\mathbf{y}^3 = \alpha(\mathbf{y}^2 - \mathbf{y}^1) + \mathbf{y}^2; \quad (3)$$

- Second order extrapolation mating generates a child using two parents and the child generated with an extrapolation mating. If \mathbf{p} is the vector difference between \mathbf{y}^1 and \mathbf{y}^3 and f^1, f^2, f^3 are the fitness values for the three individuals $\mathbf{y}^1, \mathbf{y}^2, \mathbf{y}^3$ respectively, then a second order one-dimensional model of the fitness function is built and the new child is generated taking the minimum of the resulting parabola (see Figs. 1):

$$\mathbf{y}^4 = \mathbf{y}^1 + \mathbf{p}\chi_{\min} \quad (4)$$

$$f_{\min} = a(\mathbf{y}^1, \mathbf{y}^2, \mathbf{y}^3)\chi_{\min}^2 + b(\mathbf{y}^1, \mathbf{y}^2, \mathbf{y}^3)\chi_{\min} + f(\mathbf{y}^1) \quad (5)$$

The mating operator is used also to prevent crowding of more than one principal individual in the basin of attraction of the same solution: if two or more principal individuals are colliding (intersecting their migration regions) they are automatically mated but if they reciprocal distance falls down below a given threshold, a repelling mechanism is activated which mates the worse individual (between the two colliding) with the boundaries of the subdomain D_i : each component of the selected individual is blended with the value of the furthest bound, projecting the individual into a random point within D_i , according to the following relation:

$$y_i^2 = \alpha b_i + (1 - \alpha)y_i^1 \quad (6)$$

Environment perception and Migrations

Each region \mathbf{S} is evaluated using two mechanisms: breeding or *perception* and learning. Breeding generates a subpopulation and selects the best child, if better than the parent. A new region \mathbf{S} is then associated to the child generating a migration of the entire subpopulation toward a place where better resources are expected. For this reason each hypercube \mathbf{S} is here called *migration region*. The subpopulation is generated with the following procedure: a first child is generated, within \mathbf{S} , mutating the parent, then an

extrapolation mating is performed. The two resulting children and the parent are then used to generate a third child using second order extrapolation mating. The procedure is repeated until a number of children equal to the number of coordinates have been generated (see Fig. 1). This procedure can be seen as a way for the individual \mathbf{y} to *perceive* or sense locally the environment to obtain clues about where to proceed with the exploration of the solution space.

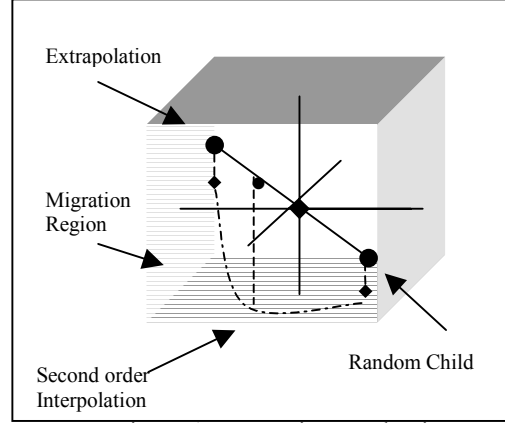


Figure 1. Perception mechanism

Learning is performed using two mechanisms: interval analysis and gradient methods. In the first case the inferior value of f within the migration region is associated to the individual \mathbf{y} and perception is then not necessary. In the second case, if the function f is locally continuous and differentiable a number of SQP steps are taken if perception gives no results and the migration region shrinks down to a small value.

The contraction or expansion of each region \mathbf{S} is regulated through a migration radius ρ and depends on the use of learning by interval analysis or through perception. The migration radius is defined as the ratio between the value of the distance from the boundary \mathbf{b}^j of the migration region of the j -th individual and the value of the distance from the correspondent boundary \mathbf{b} of the domain D :

$$\rho^j = \frac{b_i^j - y_i^j}{b_i - y_i^j} \quad (7)$$

If none of the children of the subpopulation is better than the parent the radius is reduced according to:

$$\rho^j = \begin{cases} \max([1e-8, \delta y_{\min}]) & \text{if } \delta y_{\min} \geq \varepsilon \rho^j \\ \varepsilon \rho^j & \text{if } \delta y_{\min} < \varepsilon \rho^j \end{cases} \quad (8)$$

where ε has been set to 0.5 and δy_{\min} is the distance of the best child \mathbf{y}^* , among the ones in the migration region, from the parent j , normalized with respect to the dimensions of the migration region:

$$\delta y_{\min} = \sqrt{\sum_{i=1}^n \left(\frac{y_i^* - y_i^j}{S_i^j} \right)^2} \quad (9)$$

where, for individual j and for dimension i , S_i^j is the difference between the value of the upper bound and of the lower bound and the summation is over non-zero dimensions. Now, if from generation k to generation $k+1$ the differential improvement increases, then the migration radius is recomputed according to the prediction:

$$\rho_{k+1}^j = \rho_k^j \eta \log(e-1+j) \quad (10)$$

where η is equal to 2 in this implementation. For integer numbers migration operates in the same way but the migration regions and migration radius are generated and treated differently. In particular if ρ_{min} is 1 and ρ is defined as:

$$\rho_{k+2}^j = \min[\text{int}(\log(2+j)\Delta f^j), \rho_{min}] \quad (11)$$

where Δf^j is the difference between the function f^j at generation $k+1$ minus f^j at generation k . The migration region is therefore contracted differently for real and for integer variables allowing a better spatial exploration.

Filtering

Instead of traditional selection mechanisms based on fitness here a permanent population of n individuals is maintained from one generation to another. Each individual has a chance to survive provided that it remains inside the filter. The filter ranks all the individuals on the basis of their fitness from the best to the worst. All the individuals with a fitness worse than a given threshold are either hibernated (i.e. no operator is applied) or mutated while migration is applied to all individuals within the filter. This allows each of the individuals within the filter to evolve toward a different local optimum. Mating is operated on all the individuals in the upper part of the filter combining them with individuals in the lower part of the filter. After mating, the resulting children survive according to their predicted position in the filter.

Branching Step

Even though the filter increases the chances of finding several optima and eventually the global one, convergence is not guaranteed due to the stochastic nature of the process. Therefore, a systematic step is taken on the basis of the output of the evolution algorithm. The initial domain $D_0 \equiv D$ is partitioned generating a number of subdomains D_l . Each subdomain is then qualified and explored further according to its qualification.

The partitioning, or branching, process begins taking the worst individual, which is out of the filter, and cutting D_0 into L subdomains, corresponding to L potentially new branches (or nodes). Each one of the L nodes may or may not contain an individual coming from the previous step of evolution and the volume of

the node depends on the position of the cutting point (a special mechanism prevents cuts too close to a boundary moving the cutting point to the middle of the interval). For each node D_l the ratio between the relative number of individuals and the relative volume is computed and the resulting quantity defines how necessary further exploration of the node is:

$$\sigma_{D_l} = \frac{\sum_{D_l} j}{\sum_D j} \sqrt[n]{\frac{V_{D_l}}{V_D}} \quad l=1, \dots, L \quad (12)$$

where the volumes V_{D_l} and V_D are computed taking only edges with a non-zero dimension. This quantity is then added to a fitness ϕ_{D_l} defined as:

$$\phi_{D_l} = \begin{cases} \frac{1}{J} \sum_{j=1}^J f_j - f_{best} & \text{if } i \neq 0 \\ f_{worst} - f_{best} & \\ 1 & \text{otherwise} \end{cases} \quad (13)$$

where J is the number of individuals in domain D_l . If interval analysis is available each subregion is evaluated taking the inferior and superior values, the quantity ϕ_{D_l} is then defined as:

$$\phi_{D_l} = i \eta f_{D_l}(f) \quad (14)$$

The node is then qualified by the quantity:

$$\psi_{D_l} = \sigma \sigma_{D_l} + (1-\sigma)\phi_{D_l} \quad (15)$$

where σ is the weighting factor that weights how reliable the result coming from the evolution step is considered. If σ is 0, only the nodes with low fitness are explored because the EP algorithm is considered reliable enough to explore exhaustively the domain D without leaving any region unexplored. On the other hand if σ is 1 the result from the EP algorithm is considered to be not reliable due to a premature convergence or to a poor exploration of the solution space. Now every time a node D_l is subdivided into other Q subnodes only the most promising pair is taken into account. If ψ_{D_l} is used to select the most promising ones among all L subdomains, the best pair out of the Q subnodes generated for each subdomains is selected using the following slightly different quantity:

$$\tilde{\psi}_{D_q} = \sigma \frac{\sum_{D_q} j}{\sum_{D_l} j} \gamma_q + (1-\sigma)\phi_{D_q}, \quad q=1, \dots, Q \quad (16)$$

where γ_q is, for each of the subnodes q , the ratio between the length of the edge along which the subdomain D_l is cut and the corresponding edge of D_q . Once a D_q is selected the other subnode of the pair will be the complement $D_{q+1} = D_l - D_q$. For a fast search only nodes presenting a high fitness and large volume are explored further.

In order to avoid the rediscovery of minima already found, the original domain is partitioned using more than one individual. If the worst individual is useful to

determine an upper bound on the objective function, converged individuals suggest where a further exploration is unnecessary. Therefore, in the general scheme, all converged individuals are ranked depending on the value of their fitness function. The principal cut is then, as stated above, performed using coordinates of the worst individual, the second cut takes the worst converged individual and so on up to the best converged individual.

Constraint satisfaction

The algorithm described solves bound-constrained problems but since in most of the cases constraints are nonlinear an extension of the algorithm has been developed that takes into account nonlinear inequality constraints. In this case the algorithm is run at first in order to look for a feasible set and solving, therefore, the problem:

$$\min_{\mathbf{y} \in D} f = \sum_{j=1}^m e^{C_j} \quad (17)$$

where m is the number of constraints. Once a feasible set has been found the perception mechanism is used to ensure that every move maintains the population inside the feasible set. If f^* is the value of the objective function of an individual \mathbf{y} inside the feasible set, the objective function of a new individual generated from \mathbf{y} is then augmented in the following way:

$$\min_{\mathbf{y} \in D} f = \begin{cases} f^* & \text{if every } C_j \leq 0 \\ f^* + \max C & \text{if any } C_j > 0 \end{cases} \quad (18)$$

Interval Analysis and Stopping Criteria

There are two combined stopping criteria: one for local convergence and one for global convergence. Both are based on some heuristics and not on any rigorous proof of global converge. Local convergence of each subpopulation is determined by the improvement of each individual and by the migration radius. In a convex problem, both should tend to zero in a neighborhood of the solution. Since each individual is supposed either to converge to a different minimum or not to converge (letting just the individual with highest rank in the filter to converge) a global stopping criterion for the EP is the convergence of the filter. The convergence of the filter is determined by the convergence of all the individuals if they are not clustered, i.e. if their migration regions are not intersecting, and, otherwise, by the convergence of the best individual. It must be noticed that when EP is used in conjunction with branching the convergence of the filter is not usually necessary since the branching takes care of the global exploration of the solution space. The global convergence of the branching is reached either if all the nodes reduce below a given tolerance or if EP

have converged in all subdomains and no improvement is reported after branching, i.e. no new local minima are discovered. Interval analysis, when used, guarantees that the node containing the global minimum is always in the list of explored nodes therefore if the difference between the inferior value of the best node and the best individual contained in that node is below a given tolerance, convergence to the global optimum is achieved.

CHARACTERISATION OF EARTH-MARS ROUNDTRIPS

The global search algorithm presented in the previous chapters is now applied to the problem of characterising Earth-Mars transfers.

The first analysis looks for roundtrips from Earth to Mars and back with minimal total Δv . Roundtrip trajectories are made of an Earth-Mars transfer, departing from either a circular or an elliptical orbit around the Earth and aiming at either a circular or an elliptical orbit about Mars, a certain stay time around Mars and a return transfer to either a circular or an elliptical orbit around the Earth. Depending on the launch date and on the transfer time, for each leg, different families of roundtrips can be envisaged. In order to include even free return trajectories, instead of a braking and a departure manoeuvre, a swing-by of Mars is performed every time the stay time drops below 1 day. Then this problem can be formulated as:

$$\min_D f = \Delta v_1 + \Delta v_2 + \Delta v_3 + \Delta v_4 \quad (19)$$

where Δv_2 and Δv_3 represent respectively the braking manoeuvre and the departure manoeuvre at Mars, while Δv_1 and Δv_4 are the departure manoeuvre and the braking manoeuvre at Earth respectively. Each Δv is a function of the departure date t_0 , the time of flight for the outbound leg T_1 and for the inbound leg T_2 , the stay time t_s and of course of the pericentre and apocentre of the departure (r_p^E, r_a^E) and arrival (r_p^M, r_a^M) orbits. Each solution is therefore defined by the following vector:

$$\mathbf{y} = [t_0, T_1, T_2, t_s, r_p^E, r_a^E, r_p^M, r_a^M]^T \quad (20)$$

And the solution space D contains all possible values of \mathbf{y} . In case of swing-by of Mars the objective function becomes:

$$f = \Delta v_1 + 10C_1^2 + 20C_2^2 + \Delta v_4 \quad (21)$$

where the two constraint violations for the swing-by manoeuvre, C_1 and C_2 are defined as:

$$C_1 = v_i^2 - v_o^2; C_2 = \langle \mathbf{v}_1, \mathbf{v}_o \rangle + \cos(2\beta)v_i v_o \quad (22)$$

where \mathbf{v}_i and \mathbf{v}_o are the incoming and outgoing velocity vectors relative to Mars and β is the deviation angle, function of the modulus of the incoming velocity and of the radius r_p^M of the pericentre at Mars. A further analysis of return trajectories via Venus has been done, introducing an additional swingby in the model and

extending the solution vector (and therefore the domain D) as follows:

$$\mathbf{y} = [t_0, T_1, T_2, t_s, r_p^E, r_a^E, r_p^M, r_a^M, \omega, h, T_3, T_4]^T \quad (23)$$

where now T_2 is the time of flight from Mars to Venus, h is the pericentre altitude at Venus, ω is the rotation angle of the plane of the hyperbola around the incoming vector with respect to the ecliptic plane¹⁰, T_3 is the time of flight after the swingby, up to a deep space manoeuvre and T_4 is the time of flight from the deep space manoeuvre up to the Earth. The new objective function must include the deep space manoeuvre and therefore becomes:

$$\min_D f = \Delta v_1 + \Delta v_2 + \Delta v_3 + \Delta v_4 + \Delta v_5 \quad (24)$$

Short and Long Stay Options

At first problem (19) was solved looking for roundtrips with a variable stay time from 0 to 600 days and evaluating the total Δv necessary for each launch opportunity. Then, the search was focused on short stay opportunities including returns via Venus (commonly called opposition class missions). The solution space D for problem (19) and (21) is defined in Tab.1 and comprises all the possibilities including free return trajectories. Launch windows from 2015 to 2043 were explored and all the solutions with a total Δv less than 13 km/s have been collected and plotted in Figs. 2 and 3, where diamonds represent short stay and free return options.

Then if the upper limit on total Δv is extended to 15 km/s several short stay trajectories with a return via Venus become feasible. The result has been plotted in Fig.4. It should be noticed that a return via Venus is not always available and for each solution via Venus it is often possible to find a direct return with a comparable level of Δv . This is true apart from two launch windows in which only a return via Venus allows, for a short stay, a total Δv less than 15 km/s.. Furthermore, for two particular launch opportunities (15 years away one from the other) an almost continuous range of short stay periods are allowed.

The latter one of the two comprises almost all the returns via Venus since for this date Venus is in a particularly favourable position.

Very short stays (less than 10 days) are also possible both via Venus and via direct return. This class of trajectories can become interesting as abort options in case an immediate return is necessary and a manoeuvre at Mars can still be executed. The most interesting options for a short stay either via Venus are summarised in Tab. 2 for the period from 2028 to 2037 while optimal solutions for each launch date, in the same period, are summarised in Tab. 3.

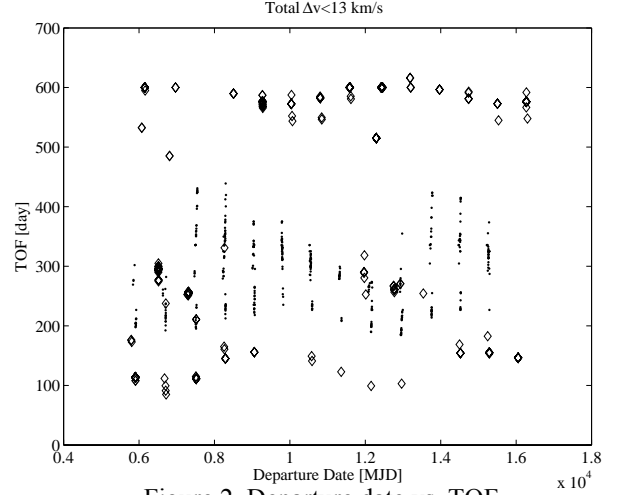


Figure 2. Departure date vs. TOF

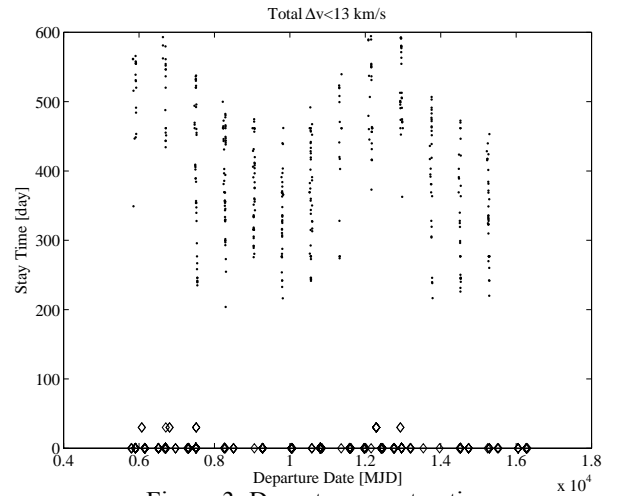


Figure 3. Departure vs. stay time

Table 1. Domain D for the roundtrip problem

| T_0 (MJD) | T_1 (DAY) | T_2 (DAY) | T_S (DAY) | R_P^E (KM) | R_A^E (KM) | R_P^M (KM) | R_A^M (KM) |
|----------------|----------------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|
| 5479 | 50 | 50 | 0 | 6778 | 6778 | 3789 | 3789 |
| 15775 | 700 | 700 | 600 | 6778 | 6778 | 5.7e5 | 5.7e5 |

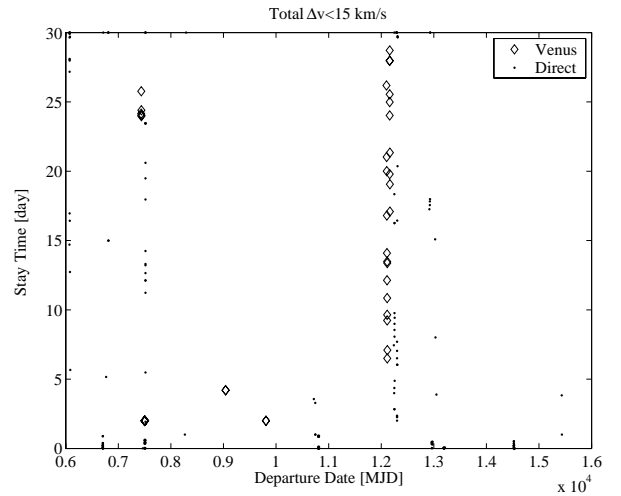


Figure 4. Return via Venus vs. direct return

Since total absorbed radiation dose is one of the key issues for human Mars mission design and since the trajectory choices are the most influential parameters, the total dose due to galactic cosmic radiation (GCR) were calculated for all presented options in Tab. 3. The

values are equivalent BFO dose values in sievert (Sv). Interplanetary dose values behind 10 g/cm² Al are assumed at 0.24 Sv/a. Mars surface levels behind 5 g/cm² Al are assumed 0.15 Sv/a. Dose levels for equivalent PE shielding materials are given in brackets.

Table 2 Best short stay options via Venus from 2028 to 2037

| LAUNCH DATE | E-M(DAY) | M-E(DAY) | T _S (DAY) | ΔV ₁ (KM/S) | ΔV ₂ (KM/S) | ΔV ₃ (KM/S) | ΔV ₄ (KM/S) | ΔV ₅ (KM/S) |
|-------------|----------|----------|----------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
| 21/10/2028 | 252 | 251 | 2 | 4.16 | 3.63 | 4.80 | 5.98 | 0.193 |
| 04/07/2029 | 529 | 404 | 2 | 4.28 | 5.28 | 3.27 | 4.72 | 1.579 |
| 30/06/2031 | 460 | 490 | 2 | 4.12 | 4.56 | 3.28 | 6.53 | 0.644 |
| 16/04/2033 | 199 | 352 | 28 | 3.59 | 2.43 | 4.21 | 3.94 | 6.15e-4 |
| 17/02/2035 | 236 | 318 | 2 | 4.97 | 3.02 | 4.29 | 3.78 | 3.13e-4 |
| 02/09/2037 | 225 | 475 | 30 | 4.06 | 2.03 | 3.33 | 4.36 | 1.369 |

Table3. Optimal solutions for direct Earth-Mars roundtrip from 2028 to 2037

| LAUNCH DATE | E-M(DAY) | M-E(DAY) | T _S (DAY) | ΔV ₁ (KM/S) | ΔV ₂ (KM/S) | ΔV ₃ (KM/S) | ΔV ₄ (KM/S) | GCR(SV) | M _F /M ₀ |
|-------------|----------|----------|----------------------|------------------------|------------------------|------------------------|------------------------|------------|--------------------------------|
| 23/11/2028 | 300.1 | 353.5 | 344.1 | 3.589 | 2.244 | 2.077 | 3.9348 | 0.57(0.43) | 0.035 |
| 09/08/2029 | 596.6 | 320.9 | 600 | 4.366 | 5.067 | 2.1498 | 4.123 | 0.85(0.63) | 0.01 |
| 24/12/2030 | 283.1 | 217.7 | 497.8 | 3.663 | 2.53 | 1.992 | 3.749 | 0.53(0.38) | 0.035 |
| 17/02/2031 | 210.7 | 217.7 | 514.9 | 3.806 | 2.776 | 1.992 | 3.749 | 0.49(0.35) | 0.031 |
| 16/04/2033 | 199.7 | 198.0 | 553.0 | 3.587 | 2.435 | 2.239 | 3.589 | 0.49(0.34) | 0.036 |
| 26/11/2034 | 248.1 | 251.0 | 30 | 6.1361 | 3.9151 | 3.498 | 5.1868 | 0.34(0.28) | 0.003 |
| 27/06/2035 | 201.9 | 267.5 | 535.6 | 3.6447 | 2.07 | 2.5836 | 3.689 | 0.53(0.37) | 0.034 |
| 02/06/2036 | 700 | 284.9 | 444.0 | 4.7548 | 8.347 | 2.313 | 3.5439 | 0.83(0.64) | 0.001 |
| 15/08/2037 | 347.5 | 282.9 | 358.1 | 3.9298 | 2.131 | 2.313 | 3.5439 | 0.56(0.42) | 0.035 |

Free Return Trajectories and Cyclers

As can be noticed from Figure 3, for each launch window it is possible to find a solution with a stay period below one day which in fact corresponds to a trajectory that departs from Earth, flybys Mars and comes back to Earth ballistically, i.e. without manoeuvres. These trajectories, known as free return trajectories, can be grouped in three main categories (see Fig. 8).

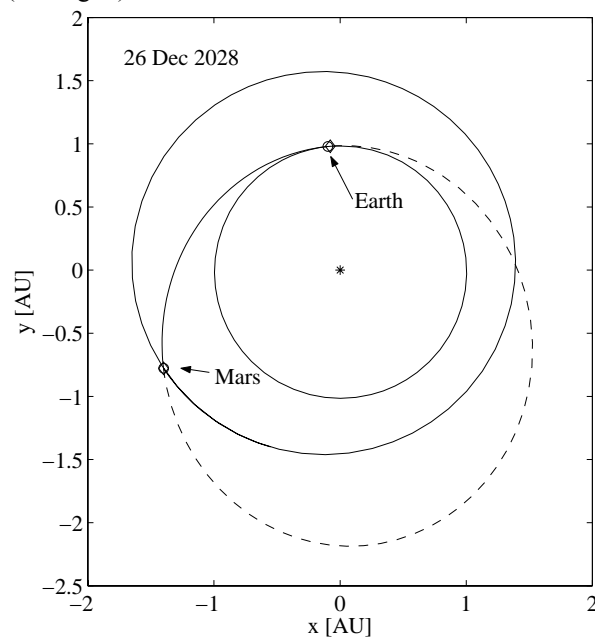


Figure 5. Up escalator

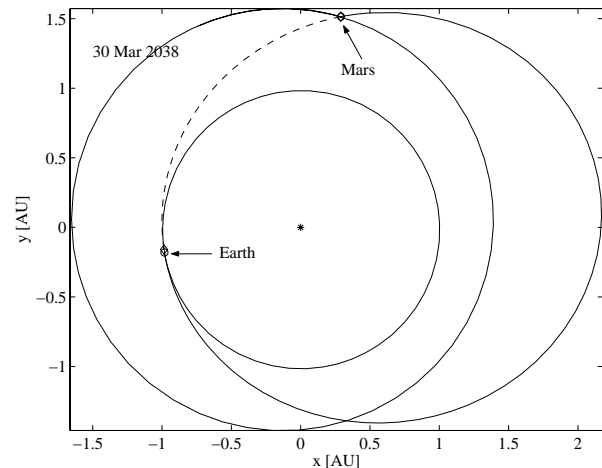


Figure 6. Down escalator

The first one comprises all transfers with a low departure velocity from Earth, a correspondent short transfer arc to Mars and a long arc leading the spacecraft back to Earth again with a low arrival velocity (see Fig.5). The overall period in space is about two terrestrial years, therefore, in analogy with Earth-Mars one synodic period cyclers, these free return trajectories are here called up escalators. The second category comprises all free return trajectories with an initial long transfer to Mars and a short return leg to Earth, the total time in space is again about two years and therefore these trajectories are here called down escalators (see Fig.6). The third category of free

return trajectories presents a relatively short transfer time on both legs either to go or to come back (see Fig.7).

As can be seen in Fig. 8 short free return trajectories can be subdivided further in three groups depending on the length of each leg. All best free return opportunities for the period from 2028 to 2037 have been summarised in Tab. 4 where transfer time, infinite velocity at Mars and Δv s at Earth are reported. As can be read, although up and down escalators are appealing for their relatively low Δv at departure they can become prohibitive if the spacecraft has to be inserted in orbit around Mars with a propulsive manoeuvre, due to the high infinite velocity.

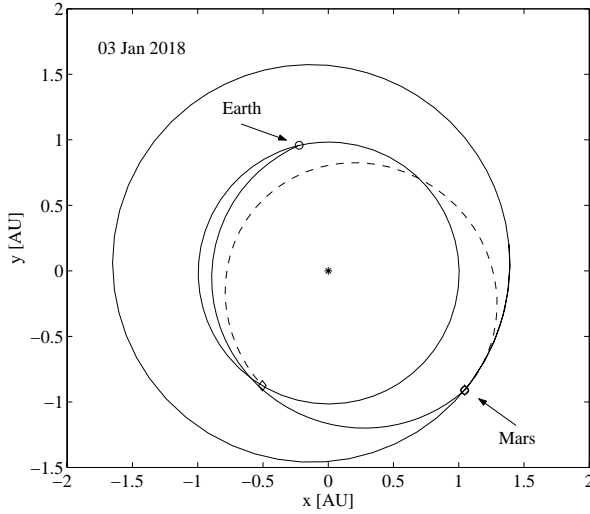


Figure 7. Short roundtrip

On the other hand some short free return options, although more demanding in terms of Δv at departure have in general a lower velocity at Mars and could be interesting either as nominal trajectories in order to increase safety for manned missions or as abort options. In fact if a failure, not affecting the propulsion system, forces the mission to be aborted on the way to Mars a deep space manoeuvre, exploiting the whole remaining propellant, could be used to inject the spacecraft on a short free return transfer back to Earth

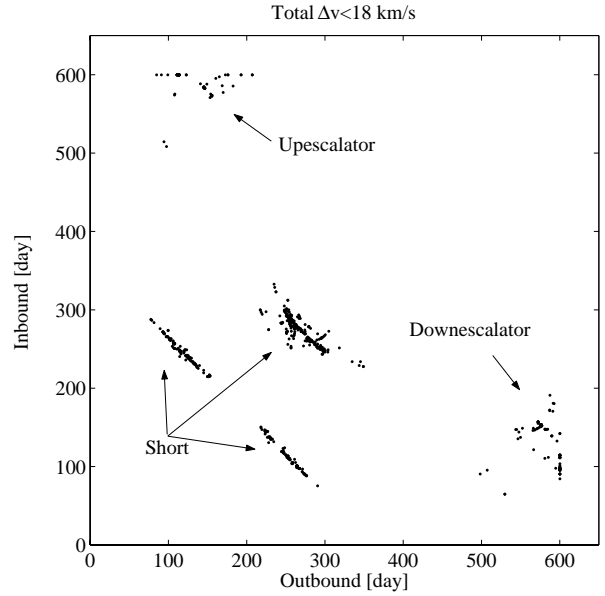


Figure 8. Free return trajectories

Table 4. Best free return options found from 2028 to 2037

| LAUNCH DATE | E-M (DAY) | M-E (DAY) | ΔV_1 (KM/S) | ΔV_4 (KM/S) | V_∞ (KM/S) | R_p (KM) |
|-------------|-----------|-----------|---------------------|---------------------|-------------------|------------|
| 26/12/2028 | 141.1 | 588.4 | 4.281 | 4.283 | 10.77 | 43389 |
| 01/08/2029 | 582.49 | 147.4 | 4.349 | 4.347 | 6.961 | 43301 |
| 18/09/2030 | 272.5 | 253.82 | 5.827 | 8.11 | 5.253 | 3789.0 |
| 09/02/2031 | 122.87 | 600 | 4.285 | 4.319 | 11.653 | 43389 |
| 21/11/2032 | 252.67 | 255.98 | 4.977 | 6.38 | 5.1056 | 3790.7 |
| 12/04/2033 | 99.05 | 600 | 4.446 | 4.876 | 11.44 | 13732 |
| 02/01/2034 | 600 | 90.48 | 5.103 | 4.499 | 10.488 | 9840.1 |
| 25/02/2036 | 600 | 11.35 | 4.504 | 4.311 | 11.776 | 36714 |
| 24/01/2037 | 254.32 | 262.8 | 7.2679 | 5.2115 | 5.252 | 3789.0 |

Optimal Staging

Even in case cryogenic propellants are considered for propulsion (with an $I_{sp}=450s$) the mass budget for a roundtrip to Mars could be prohibitive even for minimum Δv transfers (see last column of Tab.3). A solution could be to resort to staging in order to improve the payload returned to Earth. Therefore, the natural extension of problem (19) is to introduce staging sequences in the model and to optimise for the final mass m_f instead of the total Δv . The staging model assumed here is fairly simple and does not take into account gravity losses. Furthermore, a constant specific

impulse and a constant structural factor of 0.15 has been considered for each stage.

$$m_{i+1} = m_i e^{-\frac{\Delta v}{g_0 I_{sp}}} \quad (25)$$

$$m_{i+1} = m_{pl} + m_{p_{i+1}} + m_{s_{i+1}} + m_{s_i} \quad (26)$$

$$m_s = 0.15 m_p \quad (27)$$

the objective function then becomes:

$$\min_{y \in D} f = -m_f / m_0 \quad (28)$$

where m_0 is the initial mass and each solution is defined by the vector:

$$\mathbf{y} = [t_0, T_1, T_2, t_s, r_p^E, r_a^E, r_p^M, r_a^M, \epsilon_i^E, \epsilon_i^M]^T \quad (29)$$

where ϵ_i^E and ϵ_i^M represent for Earth departure and for Mars insertion respectively the ratio between the apocentre of the departure (arrival) orbit r_a^E (r_a^M) and the apocentre radius of the intermediate orbits. The pericentre altitude of the departure (arrival) is constrained to be at 400 km and the initial and final orbits are assumed to be circular with the same altitude. Therefore, if a 2-stage strategy is used for Earth escape the first stage injects the spacecraft from the 400x400km circular orbit into an intermediate orbit with apocentre $\epsilon_1^E r_a^E$ and the second stage injects the spacecraft into a departure orbit with apocentre r_a^E . Furthermore, the number of stages is fixed and equal for each escape or capture manoeuvre. In Table 5 some optimal solutions for the interval [2028,2037] are reported for a two stage strategy.

Table 5 Mass fractions at arrival at Earth with 2 stages

| DATE | R_A^E (KM) | ϵ_1^E | R_A^M (KM) | ϵ_1^M | MF/M_0 |
|------------|-----------------|----------------|-----------------|----------------|----------|
| 23/11/2028 | 5.4e4 | 0.162 | 1.9e5 | 0.03 | 0.168 |
| 24/12/2030 | 4.3e4 | 0.195 | 1.0e5 | 0.06 | 0.167 |
| 16/04/2033 | 4.1e4 | 0.197 | 4.1e5 | 0.015 | 0.168 |
| 09/07/2035 | 4.8e4 | 0.178 | 5.0e5 | 0.01 | 0.166 |

LOW-THRUST TRANSFERS

All the analysis of the previous chapter assumed the use a high thrust engines, anyway low-thrust propulsion systems may become interesting both for manned and unmanned missions. Therefore, an analysis of direct low-thrust Earth-Mars transfer will follow. A low-thrust trajectory is here modelled using an inverse method: the Cartesian coordinates of the spacecraft are described by means of a set of *pseudo-equinoctial elements* α . The set of elements used to parameterise the Cartesian coordinates are here called *pseudo-equinoctial* because they do not satisfy exactly the Gauss'planetary equations unless the thrust is zero. Each element is then developed in form of a parameterised function of the anomaly L. This function is the shape of the pseudo-element.

Once position is defined in terms of the pseudo-elements velocity and accelerations can be computed by differentiation:

$$v = \frac{dr}{dt} = \frac{dr}{dL} \frac{dL}{dt}; \quad a = \frac{dv}{dt} = \frac{dv}{dL} \frac{dL}{dt} \quad (30)$$

$$\frac{dr}{dL} = \sum_{i=1}^5 \frac{\partial r}{\partial \alpha_i} \frac{\partial \alpha_i}{\partial L} + \frac{\partial r}{\partial L}$$

In order to obtain the set of pseudo-elements that satisfies exactly the conditions at boundaries, the following nonlinear programming problem must be solved:

$$\mathbf{r}(\alpha(L_0), L_0) = \mathbf{r}_0; \quad \mathbf{v}(\alpha(L_0), L_0) = \mathbf{v}_0; \quad (31)$$

$$\mathbf{r}(\alpha(L_f), L_f) = \mathbf{r}_f; \quad \mathbf{v}(\alpha(L_f), L_f) = \mathbf{v}_f;$$

anyway for low values of the acceleration it is sufficient to solve the easier linear problem:

$$\alpha(L_0) = \alpha_0; \quad \alpha(L_f) = \alpha_f; \quad (32)$$

For each set of pseudo-elements a different trajectory can be generated, connecting two points in the state space. The controls necessary to achieve the imposed shape of the trajectory can then be obtained by solving the following system:

$$\mathbf{a}_c = \mathbf{a} - \frac{\mu}{r^3} \mathbf{r}; \quad m_f = m_0 e^{\int_0^{t_f} -c|a_c| dt} \quad (33)$$

with the additional constraint:

$$t_f - t_0 = \int_{L_0}^{L_f} \frac{dt}{dL} dL \quad (34)$$

This approach is extremely fast and the computational cost extremely low since no propagation or collocation is necessary.

Of course the thrust profile, though constrainable, is a direct consequence of the shape and must be considered only a first guess useful for further, more refined optimisation. However, the attempt here is to widely explore the solution space rather than to find an accurate solution. For this reason the design of the low-thrust trajectory has not been written either in the optimal control form (with adjoint equations) or in any direct transcription form (collocation or shorting).

It is anyway expected that as the shape of the pseudo-elements approaches the solution of the corresponding optimal control problem the inverse method will yield the associated optimal control for the thrust. For the analysis conducted in this paper the following shape has been used:

$$\alpha(L) = \alpha_0 + \alpha_f(L - L_0) + \mathbf{p} \sin(L - L_0) \quad (35)$$

where $\mathbf{p} = [p_1, p_2, p_3, p_4, p_5]^T$ is a set of parameters shaping each pseudo-element.

The optimality of the solution found can be seen from the comparison with the optimal solution computed for mars Exobiology. In the optimised solution the trajectory is characterised by two cost arcs and three thrust arcs and a maximum thrust dependent on the distance from the Sun. On the other hand the first estimate obtained with the inverse approach does not contain any model for power and thrust and no coast arcs are introduced a priori.

The optimisation problem then becomes:

$$\min_D -m_f \quad (36)$$

$$t_f - t_0 = \int_{L_0}^{L_f} \frac{dt}{dL} dL; \quad a_c \leq a_{\max}$$

and the solution vector is:

$$\mathbf{y} = [n, t_0, T_1, \theta, \phi, p_1, p_2, p_3, p_4, p_5]^T \quad (37)$$

where n is an integer number representing the number of revolutions, t_0 is the departure date and T_1 the transfer time. The domain D is specified in Tab. 6.

Table 6. Domain D for the low-thrust problem

| N | T_0 | T_1 | θ | ϕ | P_1 | P_2 | P_3 | P_4 | P_5 |
|---|-------|-------|----------|--------|-------|-------|-------|-------|-------|
| 0 | 2500 | 500 | -pi | -pi | -0.1 | -0.1 | -0.1 | -0.1 | -0.1 |
| 2 | 3000 | 1e3 | pi | pi | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |

Table 7. Comparison with Mars Exobiology

| SOL. | DEPARTURE | TOF(DAY) | M_F/M_0 | $V_{INF}(M/S)$ |
|------|------------|----------|-----------|----------------|
| Opt. | 18/03/2007 | 873 | 0.16 | 622 |
| FG | 09/03/2007 | 772.1 | 0.162 | 622 |

The resulting trajectory is represented in Fig. 9 with the associated thrust profile plot in Fig. 10. The main characteristics of the two trajectories are summarised and compared in Tab.7 where FG stands for First Guess and represents the solution computed with the inverse approach.

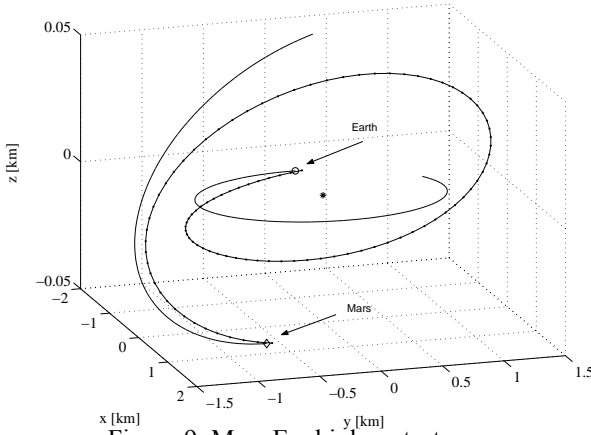


Figure 9. Mars Exobiology test case

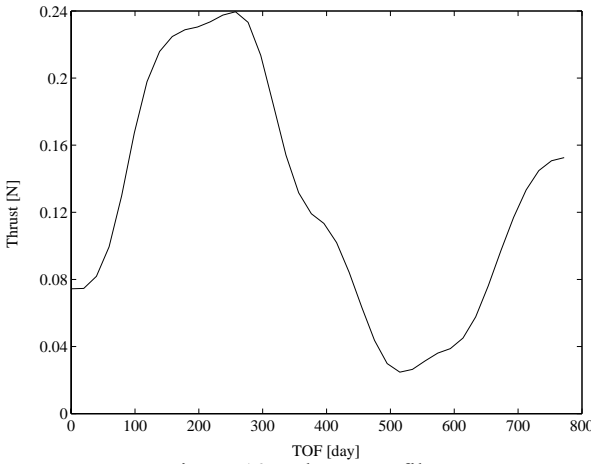


Figure 10 . Thrust profile

The algorithm successfully identified a solution with a low mass consumption comparable with the optimised solution for Mars Exobiology. The transfer time however is quite different due to the selected shape of the orbit. The solution found is expected to have an error in the velocity due to the not exact solution of problem (31). However this error has been verified to

lead, in general, to an error in propellant mass consumption which is below 15%. Therefore this solution is considered acceptable as a first estimate of a possible transfer with low-thrust propulsion, since the error is within the usual margin taken in preliminary mission design.

Low-thrust transfers with ballistic capture at Mars

The propellant consumption to reach a Mars with a low excess velocity, provided by low thrust transfers opens the interesting possibility to exploit lagrangian points of the Mars-Sun system to attempt a low-energy capture in Martian orbit. Maintaining the previous model for low-thrust arcs now the dynamics at arrival is modelled taking into account third body effects. Final conditions at Mars are taken perturbing the state vector at L_1 and propagating backward for a time Δt .

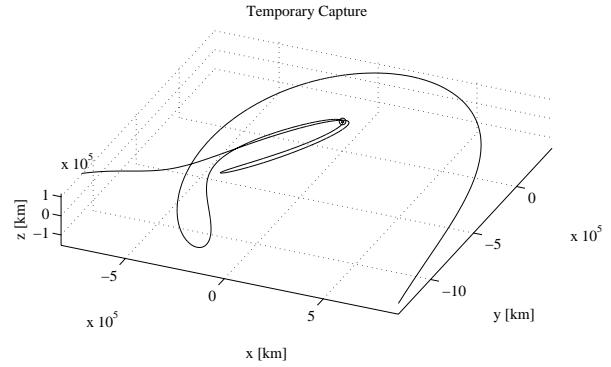


Figure 11. Temporary capture

Table 8. Low-thrust transfers with ballistic capture

| DEPARTURE | $A_{MAX.}(M/S^2)$ | TOF(DAY) | M_F/M_0 | $V_{\infty}(KM/S)$ |
|------------|-------------------|----------|-----------|--------------------|
| 15/08/2030 | 2.3e-4 | 721 | 0.165 | 0.2 |

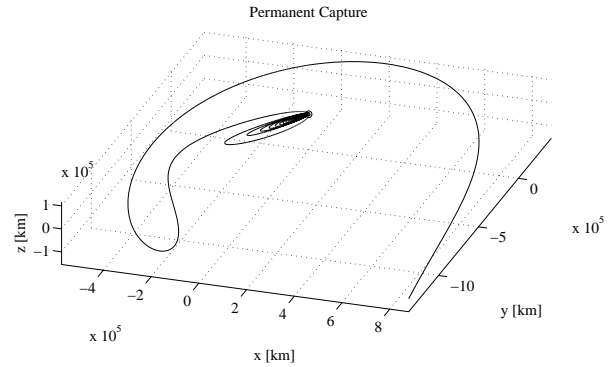


Figure 12. Permanent capture with low-thrust manoeuvre at periarees

The solution vector has been extended as follows:

$$y = [n, t_0, T_1, \Delta t, \delta v_1, \delta v_2, \delta v_3, p_1, p_2, p_3, p_4, p_5]^T \quad (38)$$

Where now $\delta v_1, \delta v_2$ and δv_3 are the three components of the velocity vector at the lagrangian point L_1 defined in the local radial, transversal, normal

martian reference frame. The value of first component belongs to the interval $[-0.12, 0.0]$ km/s while the values of the others belong to the interval $[-0.12, 0.12]$ km/s.

The resulting point in deep space represents the target of the electric propulsion arc. An example of low-thrust transfer of this kind is reported in reported in Tab. 9. The arrival at Mars has been plotted in Fig. 11 and 12 for an unpowered and for a powered capture. In the first case no manoeuvres are performed at periapsis and the resulting capture by Mars is only temporary and lasts less than 500 days. In the second case a low-thrust manoeuvre is inserted when the distance from Mars falls down below $1e5$ km, the resulting capture is now permanent.

CONCLUSIONS

In this paper a combined systematic-heuristic approach is proposed to solve trajectory design problems in which more than one solution is expected and where not just the global optimum should be obtained. The proposed combination of evolution programming and branching is suitable for problems characterised by differentiable and non-differentiable functions combining integer and real variables and have demonstrated to be an interesting tool for preliminary mission analysis especially when the objective function is a black-box. In fact in this respect an ad hoc systematic approach specifically dedicated to solve a certain problem is expected to be more efficient.

Even though the obtained results must be considered preliminary, the proposed algorithm appears to be promising even for more complex space trajectory design problems. In particular the proposed use of interval analysis represents a promising way of guaranteeing and controlling convergence.

The capabilities of this approach have been demonstrated solving the complex problem of identifying all optimal solutions for a Mars roundtrip in a given time frame. This first analysis has revealed that free return trajectories are always available for each launch window and can be classified in three major groups depending on the length of each transfer leg. Although all of them present the significant drawback of having high velocity either at Earth or at Mars they could represent an option for high specific impulse engines or as abort options. Among nominal transfers in the time frame 2028, 2037 the 2033 launch window seems to offer interesting features since the transfer time for both legs is relative low with an associated low total Δv and a low cost return via Venus is possible for a short stay. For nominal transfers the analysis of optimal staging sequences has shown how the optimal orbit for departure from Earth is elliptical but with the apocentre almost at the altitude of a geostationary orbit while for Mars the apocentre is much closer to the

sphere of influence. Finally the last analysis presented has opened the interesting possibility to use low-thrust transfers for low-energy planetary capture at Mars. However this problem and the inverse approach used to design low-thrust arcs, are the subjects of an ongoing more detailed analysis and therefore, the results presented in this paper must be considered preliminary.

ACKNOWLEDGMENTS

The authors are grateful to Guy Janin of ESOC/ESA for Mars Exobiology trajectory with electric propulsion.

REFERENCES

- [1] Gage P.J., Braun R.D., Kroo I.M. Interplanetary Trajectory Optimisation Using a Genetic Algorithm. The Journal of the Astronautical Sciences, Vol. 43, No. 1, January-March 1995, pp. 59-75.
- [2] Rauwolf G., Coverstone-Carroll V. Near-optimal low-thrust orbit transfers generated by a genetic algorithm. Journal of Spacecraft and Rockets 33(6) 1996 859-862.
- [3] Coverstone-Carroll V., Hartmann J.W., Mason W.J. Optimal multi-objective low-thrust spacecraft trajectories. Computer methods in applied mechanics and engineering 186(2000) 387-402.
- [4] Vasile M., Comoretto G., Finzi A.E. *A Combination of Evolution Programming and SQP for WSB Transfer Optimisation*. AIRO2000, September 18-21 2000 Milano, Italy.
- [5] Gurfil, P. and Kasdin, N. J., Niching Genetic Algorithms-Based Characterization of Geocentric Orbits in the 3D Elliptic Restricted Three-Body Problem, Computer Methods in Applied Mechanics and Engineering, Vol. 191, Issue 49-50, Dec. 2002, pp. 5673-5696.
- [6] C. Blied, P. Spellucci, L.N. Vicente, A. Neumaier, L. Granvilliers, E. Huens, P. Van Hentenryck, D. Sam-Haroud, B. Faltings. *COCONUT Deliverable D1. Algorithms for Solving Nonlinear Constrained and Optimisation Problems: The State of the Art*, June 8, 2001. The Coconut Project.
- [7] Michalewicz Z. *Genetic Algorithms + Data Structures = Evolution Programs*. Springer-Verlag, Telos, 1996, Third Edition.
- [8] Horn J. The Nature of Niching: Genetic Algorithms and the Evolution of Optimal, Cooperative Populations. Ph.D. Thesis, University of Illinois at Urbana-Champaign, 1997.
- [9] Hansen E. *Global Optimization Using Interval Analysis*. Marcel Dekker Inc. 1992.
- [10] Vasile M. A Global Approach To Optimal Space Trajectory Design. AAS-03-141, 13th AAS/AIAA Space Flight Mechanics Meeting, 9-13 February 2003 Puerto Rico.