ANALYSIS OF FUEL-OPTIMAL ORBITAL TRANSFER TO GEOSYNCHRONOUS ORBIT USING ELECTRIC PROPULSION

Rui Zhang,* Ran Zhang†, and Chao Han‡

Electric propulsion has been applied to orbital transfer on account of its high specific impulse. To minimize the fuel consumption, the problem of fuel-optimal orbital transfer to geosynchronous orbit is calculated and analyzed in this paper. The indirect approach is applied to the optimization of the orbital transfer to geosynchronous orbit. The calculus of variations and the Pontryagin’s maximum principal are applied to the transformation from an optimal control problem to a two-point-boundary-value-problem of differential equations. The two-point-boundary-value-problem is solved by the unscented Kalman filter parameter estimation algorithm, which is of larger convergence domain and gradient-matrix free. In this paper, the influence of the initial orbit states and the J2 perturbation on the fuel consumption is analyzed through multiple numerical simulations. The results show that the optimization method is effective and the influence on the fuel consumption provide some valuable engineering guidance.

INTRODUCTION

Electric propulsion is gradually applied to the orbital transfer of the spacecrafts because its high specific impulse will enhance the payload mass ratio. However, the thrust of the electric propulsion is relative low comparing with the traditional chemical propulsion. On account of the characteristics of low-thrust, the traditional solution to the orbital transfer problem with pulse thrust assumption is not suitable now. In order to realize the orbital transfer, electric propulsion must keep working for a relatively long time which means the orbital dynamics equations is hard to solve due to high nonlinearity.

Nowadays, there are methods including analytical method, semi-analytical method and numerical method for solving the orbital transfer problem with electric propulsion. Analytical method, as the name suggests, is a process of theoretical derivation under some assumptions. So an approximate solution can be obtained which is not enough for the detailed orbital transfer design.1–3 Semi-analytical method is a process which combines the theoretical derivation and the numerical calculation. The solution is not accurate either.4–7 For the numerical method, it is about solving an optimal control problem. In order to realize the process by numerical calculations, the numerical method can be divided into three approaches: direct approach, indirect approach and hybrid approach. The direct approach introducing sorts of discrete transformation to transform the problem into a finite non-linear programming problem. There are some advantages such as a large radius of convergence and a good algorithm robustness.8,9 However, this approach relies on the non-linear programming software and the solutions got from it are always approximate optimal which can not
be used for detailed calculations.\textsuperscript{10,11} The hybrid approach combines the direct approach and the indirect approach. A common solution of the hybrid approach abandons the transversality condition which leads to a loss of optimality.\textsuperscript{12–15} The indirect approach uses calculus of variations and Pontryagin’s maximum principle to derive the first order necessary condition so the optimal control problem can be transformed to a two-point-boundary-value-problem (TPBVP) of differential equations.\textsuperscript{16–19} Then the critical process is to solve TPBVP and three problems are presented. Firstly, the initial values of the costate variables is hard to predict. The basic method to solve TPBVP is single shooting method, namely a set of initial values should be given to integrate numerically for the final value. Then compare the final value with the given target. If the error is intolerable, the initial values must be adjusted according to the gradient matrix of error. On account of the high sensitivity of the nonlinear equations, a little change of the initial values will bring about a large change of the final values which makes it hard to predict the initial values.\textsuperscript{20–22} Secondly, the control switching point is hard to determine. For the fuel optimal problem, frequent switching between power on and off makes it hard to calculate.\textsuperscript{23} Thirdly, the gradient matrix is hard to calculate and always ill-conditioned.\textsuperscript{24}

In order to solve the problems mentioned above, unscented Kalman filter parameter estimation (UKFPE) method was proposed by Jian Li and Chao Han to solve TVBVP.\textsuperscript{25} This method does not rely on gradient information and is of low sensitivity to initial values. Meanwhile, the optimality of solution is guaranteed strictly. In this paper, the fuel-optimal orbital transfer from geosynchronous transfer orbit (GTO) to geosynchronous orbit (GEO) is calculated. UKFPE algorithm is adopted to solve TPBVP transformed from the fuel-optimal control problem. The influence of different initial orbit states on the consumption of fuel is analyzed. J2 perturbation is considered in the analysis of the process of fuel-optimal orbital transfer.

The remainder of the paper is briefly outlined as follows: In Section 2, the original optimization problem is described in detail and the optimal control TPBVP is formulated. In Section 3, the UKFPE algorithm is introduced to solve the TPBVP. In Section 4, numerical simulations are conducted and the influence of different initial orbit states is analyzed. The influence of J2 perturbation is considered in the analysis at the same time.

**BASIC FORMULATION**

**Description of Optimal Control Problem**

The orbit motion can be described by classical orbital elements or position vector and velocity vector under the Cartesian coordinate system. Classical orbital elements is not suitable for the calculation of GEO insertion because there will be singularity when \( e = 0 \) or \( i = 0^\circ \) or \( 90^\circ \). Position and velocity are not suitable either because they are fast variables which is bad for long step numerical integral. For the reasons mentioned above, a set of improved vernal equinox elements is adopted to well describe the orbit which is as follows

\[
\begin{align*}
  p &= a(1 - e^2) \\
  e_x &= e \cos(\Omega + \omega) \\
  e_y &= e \sin(\Omega + \omega) \\
  h_x &= \tan(i/2) \cos \Omega \\
  h_y &= \tan(i/2) \sin \Omega \\
  L &= \Omega + \omega + \theta
\end{align*}
\] (1)
in which $p$ is the semi-latus rectum of orbit, $a$ is semi-major axis. $e$ is eccentricity. $\Omega$ is right ascension of ascending node. $\omega$ is argument of perigee. $i$ is inclination and $\theta$ is true anomaly.

For the convenience of calculation, three instrumental variables are defined as follows

\[
\begin{align*}
W &= 1 + e_x \cos L + e_y \sin L \\
Z &= h_x \sin L - h_y \cos L \\
C &= 1 + h_x^2 + h_y^2
\end{align*}
\]  

(2)

So the orbital dynamics equations can be formulated as follows

\[
\begin{align*}
\dot{x} &= A + Ba \\
x &= [p, e_x, e_y, h_x, h_y, L]^T
\end{align*}
\]  

(3)

in which

\[
A = \sqrt{\frac{p}{\mu}} W [0 0 0 0 0 \frac{W^3 \mu}{p^3}]^T
\]  

(4)

\[
B = \sqrt{\frac{p}{\mu}} W
\begin{bmatrix}
0 & 2p & 0 \\
W \sin L & (W + 1) \cos L + e_x & -Ze_y \\
-W \cos L & (W + 1) \sin L + e_y & Ze_x \\
0 & 0 & C \cos L/2 \\
0 & 0 & C \sin L/2 \\
0 & 0 & Z
\end{bmatrix}
\]  

(5)

\[
a = [a_r, a_u, a_h]^T = a_{eng} + a_{dis}
\]  

(6)

In Equation 5, $a$ is the acceleration vector, including two parts: thrust acceleration and perturbation acceleration. If only J2 perturbation considered, then the perturbation acceleration can be expressed as follows

\[
a_{dis} = a_{J2} = \begin{bmatrix}
(a_{J2})_r \\
(a_{J2})_u \\
(a_{J2})_h
\end{bmatrix}
= -\frac{\delta \mu}{r^4} \begin{bmatrix}
1 - 3 \sin^2 i \sin^2 u \\
\sin^2 i \sin 2u \\
\sin 2i \sin u
\end{bmatrix}
\]  

\[
\delta = 1.5 J_2 R_e^2 \approx 66.0632 \times 10^9 \text{m}^2
\]  

(7)

in which $R_e$ is the radius of the Earth.

Then the orbital dynamics equation can be rewritten as follows

\[
\begin{align*}
\dot{x} &= A + \frac{T_{max}}{m} Bu + Ba_{J2} \\
\dot{m} &= -\frac{T_{max}}{I_{sp} \theta_0} \| u \| \\
u &= \frac{a_{eng}}{T_{max}/m} \| u \| \leq 1
\end{align*}
\]  

(8)

in which $m$ is total mass of the spacecraft. $T_{max}$ and $I_{sp}$ are the maximum thrust and the specific impulse of the electric propulsion.

For the time-optimal problem, the performance indicator is $\min \int_{t_0}^{t_f} dt$ and the time of transfer is not fixed. For the fuel-optimal problem, the performance indicator is $\min \int_{t_0}^{t_f} \| u \| dt$ and the time of transfer is given a value in advance.
Solution of Fuel-optimal Control Problem

Before solving fuel-optimal control problem, time-optimal control problem must be solved in order to provide a minimum time of flight. Here in this paper, only the solution of fuel-optimal problem is given because the solution of time-optimal problem is similar to fuel-optimal problem formally.

According to the direction of indirect approach, costate variables are defined as \( \lambda = [\lambda_p, \lambda_{ex}, \lambda_{ey}, \lambda_{hx}, \lambda_{hy}, \lambda_L]^T \) which corresponds to the improved vernal equinox orbital elements and \( \lambda_m \) which corresponds to the mass of the spacecraft. Then the Hamilton function can be expressed as follows

For fuel-optimal problem, the Hamilton function is

\[
H = \|u\| + \lambda^T \dot{x} + \lambda_m \dot{m} = \lambda^T A + \|u\| \left(1 + \frac{T_{\text{max}}}{m} \right) \|B^T u\| \cos \alpha - \frac{T_{\text{max}}}{I_{sp} g_0} \lambda_m
\]

in which \( \alpha \) is the angle between \( u \) and \( B^T u \).

Performance indicator which is related to \( x(t_f) \) and \( t_f \) is as follows

\[
K(x(t_f), t_f) = 0
\]

State constraint is

\[
g = x(t_f) - \begin{bmatrix} p_f e_{xf} e_{yf} h_{xf} h_{yf} \end{bmatrix}^T = 0
\]

According to Pontryagin’s Maximum Principle, state equations and costate equations are

\[
\dot{x} = \frac{\partial H}{\partial \lambda}, \quad \dot{\lambda} = -\frac{\partial H}{\partial x}
\]

\[
\dot{m} = \frac{\partial H}{\partial \lambda_m}, \quad \dot{\lambda}_m = -\frac{\partial H}{\partial m}
\]

Initial value conditions are

\[
x(t_0) = \begin{bmatrix} p_0 e_{x0} e_{y0} h_{x0} h_{y0} L_0 \end{bmatrix}^T, \quad m(t_0) = m_0
\]

Final value conditions are

\[
\lambda(t_f) = \frac{\partial K}{\partial x_f} + \sum_{i=1}^n \nu_i \frac{\partial g_i}{\partial x_f} \quad \Rightarrow \lambda_L(t_f) = 0, \lambda_m(t_f) = 0
\]

Optimal control minimizes the Hamilton function value as follows

\[
H(x^*, \lambda^*, u^*) = \min H(x^*, \lambda^*, u)
\]
According to Eq.(9) and Eq.(15), the control is optimal when $\alpha = \pi$ and Hamilton function is a constant because it does not significantly include time.

Considering the function of $\lambda_m$ in Eq.(12) as follows

$$\dot{\lambda}_m = - \frac{\partial H}{\partial m} = \frac{T_{\text{max}} \| \lambda^T B \| \| u \| \cos \alpha}{m^2}$$

when $\alpha = \pi$, the derivative of $\lambda_m$ is positive. Considering Eq.(14), when the control is optimal, $\lambda_m(t_f) = 0$. So $\lambda_m$ is not negative from $t_0$ to $t_f$.

Switching function is set as follows

$$\psi = 1 - T_{\text{max}} \left( \frac{\| \lambda^T B \|}{m} + \frac{\lambda_m}{I_{sp0}} \right)$$

The optimal control has the following form

if $\| \lambda^T B \| \neq 0$

$$\begin{cases}
  u^* = - \frac{\lambda^T B}{\| \lambda^T B \|}, & \text{if } \psi < 0 \\
  u^* = - \frac{\alpha \lambda^T B}{\| \lambda^T B \|}, & \alpha \in [0, 1] \text{ if } \psi = 0 \\
  u^* = [0, 0, 0]^T, & \text{if } \psi > 0
\end{cases}$$

if $\| \lambda^T B \| = 0$

$$\begin{cases}
  u^* = \{ u \| u \| = 1 \} \text{ if } 1 - \frac{T_{\text{max}}}{I_{sp0}} \lambda_m < 0 \\
  u^* = \{ u \| u \| \leq 1 \} \text{ if } 1 - \frac{T_{\text{max}}}{I_{sp0}} \lambda_m = 0 \\
  u^* = [0, 0, 0]^T \text{ if } 1 - \frac{T_{\text{max}}}{I_{sp0}} \lambda_m > 0
\end{cases}$$

Then the fuel-optimal control problem can be formulated in the form of TPBVP as follows

$$\begin{cases}
  \dot{x} = A + \frac{T_{\text{max}}}{m} B u, \dot{r} = -\frac{T_{\text{max}}}{I_{sp0}} \| u \| \\
  H = \| u \| + \lambda^T \dot{x} + \lambda_m \dot{r} \\
  \lambda = - \frac{\partial H}{\partial x}, \lambda_m = - \frac{\partial H}{\partial m} \\
  x(t_0) = [p_0 \ e_{x0} \ e_{y0} \ h_{x0} \ h_{y0} \ L_0]^T \\
  x(t_f) = [p_f \ e_{xf} \ e_{yf} \ h_{xf} \ h_{yf}]^T \\
  \lambda_L(t_f) = 0, \lambda_m(t_f) = 0, m(t_0) = m_0 \\
  t_0 = 0, t_f = t_{f\text{min}} c_{tf}, c_{tf} \geq 1
\end{cases}$$

So far, the description and solution of the fuel-optimal problem has been formulated completely.

**UKFPE ALGORITHM FOR SOLVING TPBVP**

Parameter estimation is to determine the internal parameters of an unknown system and the input and output are required in this process. Promise there is non-linear map $y_k = G(x_k, w)$, in which $x_k$ is the input and $y_k$ is the output. $w$ is the parameter of the nonlinear map $G$. $e_k$ is defined as output error. In order to solve the parameter $w$, the state-space equations are written as follows

$$w_{k+1} = w_k + r_k$$

$$d_k = G(x_k, w_k) + e_k$$

(20)
Eq.(20) shows that this is a static state process in which $e_k$ is a measurement noise and $r_k$ is the process noise. If we define the measurement error as $q_k = d_k - G(x_k, w_k)$, Eq.(20) can be rewritten in the following form

$$w_{k+1} = w_k + r_k$$
$$0 = -q_k + e_k \quad (21)$$

For the fuel-optimal control problem, the parameter to be estimated is expressed as follows

$$w = [\lambda(t_0), \lambda_m(t_0)]^T \quad (22)$$

Measurement error is

$$q = [x(t_f), \lambda_L(t_f), \lambda_m(t_f)]^T - G[x(t_0), m(t_0), w]$$
$$x(t_f) = [p_f, e_{xf}, e_yf, h_{xf}, h_{yf}]^T$$
$$\lambda_L(t_f) = 0, \lambda_m(t_f) = 0 \quad (23)$$

in which $G$ is the solution to the following initial value problem of differential equations

$$\begin{cases}
\dot{x} = A + \frac{T_{max}}{m} Bu, \dot{m} = -\frac{T_{max}}{m} ||u|| \\
H = ||u|| + \lambda^T \dot{x} + \lambda m \dot{m} \\
\dot{\lambda} = -\frac{\partial H}{\partial x}, \quad \dot{\lambda}_m = -\frac{\partial H}{\partial m} \\
x(t_0) = [p_0, e_{x0}, e_{y0}, h_{x0}, h_{y0}, h_{o0}]^T \\
\lambda(t_0) = \lambda(t_0) \\
m(t_0) = m_0, \lambda_m(t_0) = \lambda_m(t_0) \\
t_0 = 0
\end{cases} \quad (24)$$

The UKFPE algorithm is shown in Appendix, in which

$$\gamma = \sqrt{N + 1}$$
$$\eta = \alpha^2(N + \kappa) - N$$
$$W^{(m)}_0 = \frac{\eta}{N + \eta}$$
$$W^{(c)}_0 = \frac{\eta}{N + \eta} + (1 - \alpha^2 + \beta)$$
$$W^{(m)}_i = W^{(c)}_i = 0.5/(N + \eta) \quad i = 1, 2, ..., 2N \quad (25)$$

$\alpha$ determines the distribution range of sigma point relative to the current mean value of $w$. $\beta$ is related to the distribution of $w$, for Gaussian distribution, $\beta = 2$ is optimal. $N$ is the dimension of $w$. $\kappa$ is always set as 0 or 3 $- N$. $\lambda_{RLS}$ is forgetting factor which is always set in $(0, 1]$ in case of divergence. $\eta$ is the scale parameter.

**ANALYSIS OF CALCULATION RESULT**

Based on the formulas in the previous sections, numerical simulations are conducted in order to analyze the influence of different initial orbit states. Before calculating the fuel-optimal problem, time-optimal problem must be calculated in order to provide the optimal orbital transfer time under the current initial orbit conditions. Then we set the time of fuel-optimal problem to a certain value which is larger than the optimal orbital transfer time. In this paper, orbital transfer time of all the simulations is set to 1500$h$. If there is no special statement, the initial orbital elements are
\[ h_a = 35794 \text{km}, \, h_p = 200 \text{km}, \, i = 20^\circ, \, \omega = 0^\circ, \, \Omega = 0^\circ, \, \theta = 0^\circ, \] the maximum thrust of the electric propulsion is 1 N and the specific impulse is 2000 s.

Figure 1 shows that as the height of perigee increases, the optimal fuel mass residual increases linearly. It means when the height of perigee of the initial orbit becomes larger, less fuel will be consumed. Therefore, if more fuel is expected to be saved, the launch vehicle of the spacecraft should increase the height of perigee moderately. This is a well-known conclusion among the space engineering field.

![Figure 1. Mass Residual Varies with Height of Perigee](image)

Figure 2 shows that when the inclination is large like \(20^\circ\), as the height of apogee increases, the optimal fuel mass residual increases and the rate of increasing is getting smaller. However, when the inclination is small like \(5^\circ\), as the height of apogee increases, the optimal fuel mass residual doesn’t vary obviously. It means when the height of apogee of the initial orbit becomes larger, less fuel will be consumed. However, when the height of apogee reaches a certain large value or the initial inclination is not so large, its effect on fuel-saving will become not so significant. This can be explained easily that when the initial inclination is large, the increase of height of apogee which will decrease the velocity at the ascending node or descending node is able to save more fuel for the inclination control. And when the initial inclination is small, the effect will not be so obvious. Hence if more fuel is expected to be save, the launch vehicle which is launched at high latitudes should increase the height of apogee moderately. That is to say the super geosynchronous transfer orbit (SGTO) should be used in real space engineering practice. A proper initial height of apogee should be selected by considering some factors such as the carrying capability of the launch vehicle.

Figure 3 shows that as the inclination increased, the optimal fuel mass residual decreases rapidly and the rate of decreasing is getting larger. It means when the inclination of the initial orbit becomes larger, more fuel will be consumed. Hence if more fuel is expected to be saved, the launch site should be chosen at low latitudes in order to provide a minimum inclination for the spacecraft.

In figure 4, the blue line shows that as the argument of perigee varies between \(0^\circ\) to \(360^\circ\), the optimal fuel mass residual varies periodically. When the argument of perigee is \(90^\circ\) or \(270^\circ\), the mass residual is minimum. When the argument of perigee is \(0^\circ\) or \(180^\circ\), the mass residual is maximum. It means when the perigee coincide with the ascending node or descending node, less fuel will be
consumed. Figure 5 shows obviously that optimal control is always near the apogee of the orbit in the early period of orbital transfer and optimal control is always near the apogee or perigee in the later period of orbital transfer. According to the basic orbit dynamics, the efficiency of inclination control is highest when the control is at the ascending node or descending node. Hence when the argument of perigee is $0^\circ$ or $180^\circ$, the ascending node or descending node coincides with the perigee and then the inclination can be reduced effectively with less fuel cost.

As is known to all, J2 perturbation has long-term influence on the right ascension of ascending node (RAAN) and the argument of perigee. The influence of the argument of perigee has been discussed in the previous part. Here J2 perturbation is considered and the result is drawn as a red line in Figure 4. An obvious phase drift of the curve can be observed. This is because when J2 perturbation is considered, the argument of perigee will increase slowly in the orbital transfer process ($i = 20^\circ$). Moreover, it can be noticed that the maximum value of the mass residual
decreased and the minimum value of the mass residual increased. Combining the discussion of the
influence of argument of perigee, we know that when J2 perturbation is considered, the perigee will
not always coincide with the ascending node and descending node (initial argument of perigee is 0°
or 180°). So the efficiency of inclination control will go down and more fuel will be cost. Similarly,
when the initial argument of perigee is 90° or 270°, the efficiency of inclination control will go up
and more fuel will be saved.

CONCLUSION

In order to minimize the orbital transfer fuel consumption of the spacecraft heading for GEO with
electric propulsion, the fuel-optimal problem is formulated with indirect approach. A transformation
from an optimal control problem to a TPBVP is conducted by using the calculus of variations and
the Pontryagin’s maximum principal. TPBVP is solved by using UKFPE algorithm. Numerical
simulations are conducted and the influence of the initial orbit states and the J2 perturbation on the fuel consumption is analyzed. The results show that increasing the initial height of perigee or apogee is effective for fuel-savings. However, when the inclination is small, the fuel-saving effect of increasing the initial height of apogee is not obvious. That is to say, for the spacecraft launched at high latitudes, SGTO should be considered. The decreasing of inclination is effective for fuel-saving which explains that a low latitude launch site should be selected for GEO spacecraft. The initial argument of perigee has influence on the fuel-optimal orbital transfer. According to the result of optimization in this paper, fuel-optimal control is always switched on at the apogee or perigee. The efficiency of inclination control is highest when the control is near the ascending node or descending node according to the basic orbital dynamics, so when the initial argument of perigee changes, the optimal cost of fuel will change. J2 perturbation has influence on fuel-optimal orbital transfer because it causes a drift of the argument of perigee. The optimal cost of fuel will change because the relative position between perigee and ascending node is always changing due to the long-term change of argument of perigee.

**APPENDIX: UKFPE ALGORITHM**

The specific unscented Kalman filter parameter estimation (UKFPE) algorithm is shown in Figure 6.

**REFERENCES**


Figure 6. UKF Parameter Estimation Algorithm

Initialization
\[ \hat{\mathbf{w}}_0 = \mathcal{E}[\mathbf{w}] \]
\[ \mathbf{P}_{\mathbf{w}_0} = \mathcal{E}[(\mathbf{w} - \hat{\mathbf{w}}_0)(\mathbf{w} - \hat{\mathbf{w}}_0)^\top] \]

Time update and generate sigma point
\[ \hat{\mathbf{w}}_k = \hat{\mathbf{w}}_{k-1} \]
\[ R'_k = (\lambda_{\text{RLS}}^{-1} - 1)\mathbf{P}_{\mathbf{w}_k} \]
\[ \mathbf{P}_{\mathbf{w}_k} = \mathbf{P}_{\mathbf{w}_{k-1}} + R'_k \]
\[ W_{k|k-1} = [\hat{\mathbf{w}}_k, \hat{\mathbf{w}}_k + \gamma\sqrt{\mathbf{P}_{\mathbf{w}_k}}, \hat{\mathbf{w}}_k - \gamma\sqrt{\mathbf{P}_{\mathbf{w}_k}}] \]
\[ D_{k|k-1} = G(x_k, W_{k|k-1}) \]
(option 1): \[ \hat{d}_k = \sum_{i=0}^{2L} W_i^{(m)} D_{i,k|k-1} \]
(option 2): \[ \hat{d}_k = G(x_k, \hat{\mathbf{w}}_k) \]

Measurement update
\[ \mathbf{P}_{d_k d_k} = \sum_{i=0}^{2L} W_i^{(c)} [D_{i,k|k-1} - \hat{d}_k] [D_{i,k|k-1} - \hat{d}_k]^\top + R_i^r \]
\[ \mathbf{P}_{w_k d_k} = \sum_{i=0}^{2L} W_i^{(c)} [W_{i,k|k-1} - \hat{\mathbf{w}}_k] [D_{i,k|k-1} - \hat{d}_k]^\top \]
\[ K_k = \mathbf{P}_{w_k d_k} \mathbf{P}_{d_k d_k}^{-1} \]
\[ \hat{\mathbf{w}}_k = \hat{\mathbf{w}}_k + K_k (d_k - \hat{d}_k) \]
\[ \mathbf{P}_{w_k} = \mathbf{P}_{w_k} - K_k \mathbf{P}_{d_k d_k} K_k^\top \]

Converge?
\[ \text{N} \]
\[ \text{Y} \]
End