Publ. Obs. Astron. Belgrade No. 49 (1995), 19-39

AN ANALYTICAL PERTURBATION THEORY FOR THE LOW LUNAR POLAR ORBITER

Z. KNEŽEVIĆ
Astronomical Observatory, Volgina 7, 11050 Belgrade, Yugoslavia
E-mail zoran@aob.aob.bg.ac.yu

A. MILANI
Department of Mathematics, Pisa, Italy
E-mail milani@adams.dm.unipi.it

Abstract. We present in this paper a fully analytical theory of motion for the low lunar polar orbiter, intended for the mission analysis i.e. for the preliminary study of the satellite orbit. We first explain how in general one develops an analytic perturbation theory, and describe the requirements to be met and choices to be done to build the theory and compute the solutions. Then we give the basic equations for the theory in question and discuss some particularities and technicalities regarding the methods and procedures employed. The results achieved with the current version of the theory (rms's in semimajor axis, eccentricity and inclination are $\approx 13\,m$, < 0.001 and $< 0.001\,rad$, respectively, in the course of 6 months) are more than enough for the present purpose. Finally, we briefly discuss the limitations of the theory, and ways to improve it if and when the need arises.

1. BUILDING AN ANALYTICAL PERTURBATION THEORY

In order to build an analytical perturbation theory of motion of any natural or artificial celestial body, one has to know precise answers to several apparently simple questions. So, for example, it is necessary to know what kind of motion the theory is supposed to describe, which part of the phase space of orbital elements will be covered, what is the minimal/optimal dynamical model needed to meet the requirements on the accuracy and the time stability of the solutions, what kinds of forces are involved, which analytical methods best suite the purpose, etc. The success of any pertubation theory critically depends on the right answers to the above questions (and many other more specific and technical ones), and it is quite a complex task to develop and correctly tune a theory which will perform in the best way, while still keeping the effort and the time needed to complete and apply it within some reasonable limits.

We shall try to show in the following how we have fulfilled such a task in a particular case of building the analytical theory of motion of a low-altitude, nearly-polar lunar orbiter. The purpose of the theory is the mission analysis, i.e. the preliminary study of the satellite orbit which must be at the same time suitable for the experiments to be performed in the framework of the mission and such that the life-time of the satellite is long enough and manoeuvering as seldom and as little fuel-consuming as possible. In a way, this theory bears some resemblance to the theory of computation of asteroid proper elements (in particular from the viewpoints of basic analytical tools used to construct the theory, and the forms of the solutions and procedures of their analysis), so that one can also refer to it as "the theory of selenocentric proper elements".

Let us first define our problem along the lines described above.

A spacecraft orbiting the Moon is subject to perturbation of its osculating two-body orbital elements, due to the harmonics of the lunar gravity field, to the differential attraction from the Earth and from the Sun, to non gravitational perturbations, and others. The effects of these perturbations belong to three main classes : very short periodic (with periods of the order of the satellite orbital period, i.e. a few hours, or less), medium periodic (with periods longer than one orbital period of the satellite, but shorter than one lunar month), and long periodic. For the sake of simplicity, very short and medium periodic perturbations will be collectively called short periodic. The long term evolution of the orbit, in particular the time series of the periselenium altitude which determines the safety of the mission, depends mostly on the long periodic effects. Thus a numerical computation of the orbit is extremely ineffective as a tool for the study of the qualitative behaviour of the orbit. The computation of a single orbit by brute force numerical integration is not a problem, but a systematic exploration of the phase space to define the safe region and the optimum orbit maintenance strategy is almost impossible.

There are three types of elements involved : osculating, mean and proper. Osculating elements are the instantaneous ones; they can be expressed as a set of orbital elements (e.g. keplerian), and usually they are obtained by a time independent coordinate transformation from the state vector (cartesian position and velocity). Mean elements are obtained from the osculating ones by removing all the perturbations (due to the non spherically symmetric part of the gravity field of the Moon) with short periods. As aforementioned, short periods are by definition shorter than one lunar month, which is also one rotation period of the Moon. Proper elements are obtained by removing from mean elements the long periodic perturbations. Proper elements are therefore solution of an integrable problem, whose time evolution can be computed analytically (and with a comparatively simple formula). However, it has to be remembered that transformation of a non integrable problem into an integrable one cannot be performed in an exact way, but only by neglecting some higher degree and order terms; in practice, this means that the proper elements which should be constant in the trivial dynamics, such as the proper eccentricity, are not exactly constant when computed from a time series of state vectors. Following a procedure well established for asteroid proper elements, we use the standard deviation of these proper elements with respect to their long term mean as a measure of the accuracy of the proper elements theory (see Milani and Knežević, 1990, 1992, 1994).

Intrinsic to any analytic theory is the set of rules governing necessary truncations of the gravity field potential, and of the order of theory in agreement with accuracy requirements. In the particular case, the main choice to be done is the rule to be used to truncate R, the potential of the lunar gravity field. For a low lunar orbiter, the eccentricity can not be large, while the inclination can be large (and indeed we are here interested mostly in polar orbits). Hence, we are using a rule based upon the eccentricity, such that we truncate all the perturbations to degree 1 in eccentricity; this requires to expand the perturbing function to degree 2 in eccentricity, since some perturbations contain derivatives such as $\partial R/\partial e$. Note that we also perform some truncation which takes into account that the orbit is nearly polar, that is $\cos I$ is small.

A second truncation is with respect to the degree l in the spherical harmonics expansion (see later). This is justified by the fact that the harmonic coefficients C_{lm}, S_{lm} are decreasing with l, roughly proportionally to $1/l^2$, i.e. according to the well known Kaula's (1966) rule. Our theory has no *a priori* upper limit to l, but of course some limitation has to be chosen to control the computational cost, and also to avoid numerical instabilities. Moreover the actual values of the high degree and order harmonic coefficients are highly uncertain, and there is no point in doing very long computations based on unreliable input data.

The third truncation is a truncation to some order in the small parameters appearing in the perturbations. For the computation of the short periodic perturbations, a first order theory is accurate enough. On the contrary, for the long periodic perturbations, if the accuracy required is very high and the time span is very long, some terms belonging to the second order in the small parameters should be added. The current version of our theory does not include these second order terms, also because the uncertainty of the harmonic coefficients results in a larger error in the solution.

Finally, we should define our dynamical model (the current version of our theory does not include the effects of the perturbations due to the Earth and to the Sun, neither the non-gravitational effects), decide on the choice of variables (non-singular, canonical) and perturbation methods (Lagrangian, Hamiltonian), coordinate system (inertial, body-fixed), etc.

2. BASIC EQUATIONS

Let us quickly browse through the basic equations that serve to set up the problem.

2. 1. EQUATIONS OF MOTION

The potential of the lunar gravity field is given in terms of the development into spherical harmonics (Kaula 1966) as a sum of the monopol term (potential of a sphere) and the perturbation (accounting for all the deviations of a real body from a sphere) :

$$
\begin{aligned}
U &= \frac{GM}{r} + R \\
R &= \frac{GM}{r} \sum_{l=2}^{+\infty} \left(\frac{R_M}{r}\right)^l \sum_{m=0}^{l} P_{lm}(\phi)\left[C_{lm}\cos m\lambda + S_{lm}\sin m\lambda\right]
\end{aligned}
\tag{1}
$$

Since the $l = 1$ terms are removed by translation of the origin of the reference system to the centre of mass of the Moon, the *perturbing function* R contains only the terms of degree $l \geq 2$. The perturbing function can be expressed as a function of the usual keplerian orbital elements $(a, e, I, \Omega, \omega, \ell)$ (semimajor axis, eccentricity, inclination to the lunar equator, longitude of node, argument of periselenium, mean anomaly), and expanded as follows :

$$R = \frac{GM}{a} \sum_{l=2}^{+\infty} \left(\frac{R_M}{a}\right)^l \sum_{m=0}^{l} \sum_{p=0}^{l} F_{lmp}(I) \sum_{q=-\infty}^{+\infty} G_{lpq}(e) S_{lmpq}(\omega, \ell, \Omega, \theta) \tag{2}$$

$$S_{lmpq} = \begin{cases} C_{lm} \cos \Psi_{lmpq} + S_{lm} \sin \Psi_{lmpq} & (l-m \quad \text{even}) \\ -S_{lm} \cos \Psi_{lmpq} + C_{lm} \sin \Psi_{lmpq} & (l-m \quad \text{odd}) \end{cases}$$

$$\Psi_{lmpq} = (l-2p)\,\omega + (l-2p+q)\ell + m(\Omega - \theta)$$

where θ is the phase of the lunar rotation, namely the angle between some body fixed direction along the equator (Davies et al. 1992) and some inertial direction along the equator; precession of the lunar pole and physical librations can be neglected. The *inclination functions* F_{lmp} and the *eccentricity functions* G_{lpq} can be explicitly computed.

We can now define very short periodic terms in R as those with $l - 2p + q \neq 0$ (i.e., those containing the mean anomaly ℓ); medium period terms are those with $l - 2p + q = 0$ but $m \neq 0$ (i.e. those containing $m\theta$); long periodic terms have both $l - 2p + q = 0$ and $m = 0$. We shall use the following notation :

$$R = \overline{R} + \hat{R} + \tilde{R} \tag{3}$$

where $\overline{R}$ contains only the long periodic terms, $\hat{R}$ only the medium periodic terms, $\tilde{R}$ only the short periodic ones.

Thus we can formally define the mean elements by saying that they are such that the equations of motion for them contain only the derivatives of $\overline{R}$. The algorithm we want to define contains two stages : first the short periodic perturbations (containing the derivatives of $\hat{R} + \tilde{R}$) are removed, second we truncate $\overline{R}$ in such a way to obtain an integrable system, which we can solve in closed form.

As already explained, the theory is to be used for the mission analysis of low lunar polar orbiters. Low orbits imply low eccentricities, e.g. for an orbit with a mean altitude of $100\,Km$ eccentricities larger than about 0.06 result in crash against the Moon in the periselenium. Low eccentricities, in turn, require using of the variables which are not singular for $e = 0$. Therefore we switched to the nonsingular variables :

$$h = e \sin \omega; \qquad k = e \cos \omega \tag{4}$$

For other elements such a switch was not necessary, and we eventually started with Lagrangian equations of motion in mixed variables :

$$\frac{da}{dt} = \frac{2}{na}\frac{\partial R}{\partial \lambda}$$

$$\frac{dh}{dt} = \frac{\beta}{na^2}\frac{\partial R}{\partial k} - k\frac{\cot I}{na^2\beta}\frac{\partial R}{\partial I}$$

$$\frac{dk}{dt} = -\frac{\beta}{na^2}\frac{\partial R}{\partial h} + h\frac{\cot I}{na^2\beta}\frac{\partial R}{\partial I}$$

$$\frac{dI}{dt} = \frac{\cot I}{na^2\beta}\frac{\partial R}{\partial \omega} - \frac{1}{na^2\beta\sin I}\frac{\partial R}{\partial \Omega}$$

$$\frac{d\Omega}{dt} = \frac{1}{na^2\beta\sin I}\frac{\partial R}{\partial I} \tag{5}$$

where $\beta = \sqrt{1-e^2}$

2. 2. DYNAMICS OF THE MEAN ELEMENTS

Now let us start by discussing long periodic perturbations, that is the effect of the perturbing potential $\overline{R}$, and by ignoring $\hat{R}+\tilde{R}$; hence, we are considering the dynamics in the phase space of the mean elements.

The perturbing potential $\overline{R}$ contains only the terms of the expansion (2) in which $l-2p+q=0$ and $m=0$, therefore only the so called "zonal" harmonics of the gravity field :

$$\overline{R} = \frac{GM}{a}\sum_{l=2}^{+\infty}\left(\frac{R_M}{a}\right)^l C_{l0}\sum_{p=0}^{l} F_{l0p}(I)G_{lpq}(e)S_q(\omega) \tag{6}$$

$$S_q(\omega) = \begin{cases} \cos(-q\omega) & (l \ \text{ even}) \\ \sin(-q\omega) & (l \ \text{ odd}) \end{cases}$$

$$q = 2p - l$$

Let us next expand the eccentricity function in powers of eccentricity $G_{lpq}(e) = g^0_{lpq}e^{|q|} + g^1_{lpq}e^{|q|+2} + \ldots$, where the upper index of g denotes merely that this is the coefficient of the first (g^0), second (g^1), etc. term in the development. The truncation to degree 2 in eccentricity e (that is in h, k) implies that the value of the index q must be between -2 and $+2$; since in the long periodic terms $q = 2p - l$, then for a given value of l there are only few possible values of p. For even $l = 2s$, $p = s, s+1, s-1$ are the only admissible values; for odd $l = 2s+1$, p can only be $s, s+1$. As a result for each value of s there are only 5 terms to be computed. The final result can be expressed by means of only four quantities A, C, D, W, each a function of a and of the inclination I only :

$$\sigma = \frac{GM}{na^3}\beta\left(\frac{R_M}{a}\right)^2 = n\beta\left(\frac{R_M}{a}\right)^2 \approx n\left(\frac{R_M}{a}\right)^2$$

$$A = \sigma\sum_{s=1}\left(\frac{R_M}{a}\right)^{2s-2} C_{2s,0}F_{2s,0,s}(I)g^1_{2s,s,0}\,2$$

$$C = \sigma\sum_{s=1}\left(\frac{R_M}{a}\right)^{2s-2} C_{2s,0}\left[F_{2s,0,s+1}(I)g^0_{2s,s+1,2} + F_{2s,0,s-1}(I)g^0_{2s,s-1,-2}\right]2$$

$$D = \sigma\sum_{s=1}\left(\frac{R_M}{a}\right)^{2s-1} C_{2s+1,0}\left[F_{2s+1,0,s}(I)g^0_{2s+1,s,-1} - F_{2s+1,0,s+1}(I)g^0_{2s+1,s+1,1}\right]$$

β disappearing since $\beta \approx 1 + O(e^2)$. The equations of motion (5) for the nonsingular variables h, k, when truncated to degree 1, then become :

$$\frac{dh}{dt} = (A+C)k - Wk + O(h^2+k^2)$$

$$\frac{dk}{dt} = -(A-C)h + Wh - D + O(h^2+k^2) \qquad (7)$$

If $R = \overline{R}$, the component of the angular momentum along the lunar polar axis is an integral of motion : $H = \sqrt{GMa(1-e^2)}\cos I = const$, thus a separate equation of motion for I is not needed; the semimajor axis a is also constant, since the mean anomaly ℓ does not appear in the potential. If the orbit is nearly polar, then the changes in inclination are small : this can be deduced by differentiating the integral H, from which we obtain, by using (5) :

$$\frac{dI}{dt} = -\frac{\cos I}{\sin I\,\beta^2}\left(h\frac{dh}{dt} + k\frac{dk}{dt}\right) = \frac{\cos I}{\sin I}\,k\,D + O(h^2+k^2) \qquad (8)$$

While equation (8) is applicable for every inclination, if the orbit is nearly polar, with $\cos I = O(e)$, then all the right hand side is $O(e^2)$ and can be neglected in our truncation. Thus we can assume that the coefficients A, C, D, W in (7) are constant. Equation (7), once $O(e^2)$ and $O(e \cos I)$ have been neglected, is a system of linear differential equations with constant coefficients.

Geometrically, it is clear that (7) has a single equilibrium point for $k = 0, h = h_F$ with :

$$h_F = \frac{D}{C - A + W} \qquad (9)$$

Thus there is a particular solution of the long periodic equation which has constant mean h and k; this *frozen orbit* has eccentricity $e_F = |h_F|$. The existence of a frozen orbit with nonzero eccentricity results from $D \neq 0$, that is from the presence of odd zonals, i.e. from the asymmetry between the northern and southern hemispheres of the Moon. The other solutions describe ellipses in the mean h, k plane (see Figure 1).

We define the proper eccentricity e_P as the length of the semiaxis of these ellipses in the h direction. $e_P = 0$ for the frozen orbit. The proper argument of periselenium ω_P is the phase of the solution of the linear equation (7), thus it is –within our approximation– a linear function of time. The frequency Λ_1 of ω_P is small, corresponding to periods of a few years; for this reason the separation of the long periodic perturbations from the short periodic ones is justified.

The proper inclination I_P is such that

$$\cos I_P = \sqrt{1-e^2}\cos I \tag{10}$$

I_P can be described as the value of I_P corresponding to the origin in the h, k plane on the $\sqrt{1-e^2}\cos I = const$ surface, that is $\cos I_P$ is the minimum value compatible with this integral. For the longitude of the node Ω, starting from the usual Lagrange equation :

$$\frac{d\Omega}{dt} = \frac{1}{na^2\beta\sin I}\frac{\partial R}{\partial I} \tag{11}$$

and by the same method we obtain :

$$\frac{d\Omega}{dt} = V + Zh + O(h^2+k^2) \tag{12}$$

and V and Z can be considered as constants by the same argument as above. In this way we can derive a solution for the long periodic perturbations on Ω, which revolves with average frequency Λ_2, also very small.

2. 3. DYNAMICS OF THE OSCULATING ELEMENTS

To build a theory of short periodic perturbations let us recall the distinction between medium and very short period perturbations, that is the splitting of the perturbing potential R into three parts $R = \overline{R} + \hat{R} + \tilde{R}$.

Since the medium period perturbations on the eccentricity are much larger than the very short periodic ones, let us first handle $\hat{R}$.

2.3.1 *Medium periodic perturbations*

The medium period part of the perturbing function is :

$$\hat{R} = \frac{GM}{a}\sum_{l=2}^{+\infty}\left(\frac{R_M}{a}\right)^l \sum_{m=1}^{l}\sum_{p=0}^{l} F_{lmp}(I)G_{lpq}(e)J_{lm}S_{lmp}(\omega,\Omega,\theta) \tag{13}$$

where $q = l - 2p$ and $m \neq 0$ to isolate the medium periodic terms. The trigonometric part is defined as follows :

$$S_{lmp} = \begin{Bmatrix}\cos\\ \sin\end{Bmatrix} \left[(-q)\omega + m(\Omega-\theta) - \delta_{lm}\right] \begin{Bmatrix} l-m \quad \text{even}\\ l-m \quad \text{odd}\end{Bmatrix} \tag{14}$$

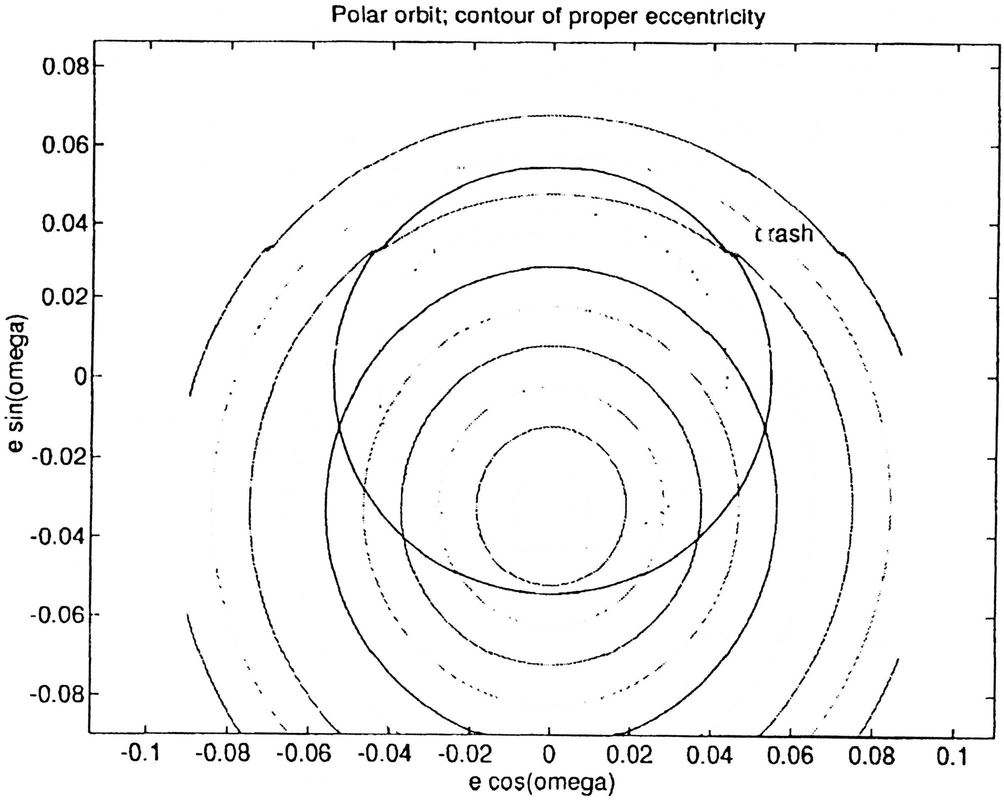


Fig. 1. Contour curves of the proper eccentricity in the mean k, h plane. The solutions of the long periodic perturbations equations move clockwise along these curves. When the distance from the origin in this plane, that is the mean eccentricity, grows too much, hard landing occurs. This plot is for a mean inclination of exactly 90° and for the Konopliv et al. (1993) model of the lunar potential.

In these formulas the harmonic coefficients C_{lm}, S_{lm} have been replaced by the corresponding amplitude and phase J_{lm}, δ_{lm} :

$$J_{lm} = \sqrt{C_{lm}^2 + S_{lm}^2}; \qquad \delta_{lm} = \arctan \frac{-S_{lm}}{C_{lm}} \tag{15}$$

To compute the medium period perturbations we use a canonical transformation method. To do this, we need to adopt a canonical coordinate system which also removes the $e = 0$ singularity (but not the $I = 0$ singularity); this is accomplished by a small modification of the coordinate system introduced by Poincaré :

$$\left\{\begin{array}{l}\omega\\ \ell\\ \Omega\\ G\\ L\\ H\end{array}\right\} \longrightarrow \left\{\begin{array}{l}\xi = \sqrt{2L-2G}\cos\omega\\ \lambda = \ell + \omega\\ \Omega\\ \eta = \sqrt{2L-2G}\sin\omega\\ L\\ H\end{array}\right\} \tag{16}$$

where :

$$L = \sqrt{\mu a}; \qquad G = \sqrt{\mu a(1-e^2)}; \qquad H = \sqrt{\mu a(1-e^2)}\cos I\,. \tag{17}$$

The Hamilton function in these variables is :

$$F(\lambda, \xi, \Omega, L, \eta, H; \theta) = \frac{GM}{2L^2} + R \tag{18}$$

where θ, the phase of the lunar rotation, is a known linear function of time :

$$\frac{d\theta}{dt} = \nu \tag{19}$$

Note again that for the medium period perturbations $\hat{R}$ does not depend on the fast variable λ (ω being incorporated into ξ, η) :

$$\hat{F}(-, \xi, \Omega, L, \eta, H; \theta) = \frac{GM}{2L^2} + \overline{R} + \hat{R}; \quad \text{and} \quad \frac{dL}{dt} = 0 \tag{20}$$

In order to remove terms with $m\theta$ we perform the canonical transformation :

$$\hat{F}(\xi, \Omega, \eta, H; L, \theta) = \hat{F}'(\xi', \Omega', \eta', H'; L) \tag{21}$$

Notice that the variable L does not play any rôle as a dynamical parameter and does not need to be transformed; the purpose of the transformation is to eliminate the dependency on θ, that is all the medium period terms from the Hamiltonian. This transformation is performed by means of the generating function in mixed variables $\mathbf{q} = (\xi, \Omega)$ being the coordinates, and $\mathbf{p} = (\eta, H)$ being the momenta :

$$\begin{aligned} S &= S(\mathbf{q}, \mathbf{p}') = \xi\eta' + \Omega H' + \hat{S}_1 \\ \hat{S}_1 &= \int \hat{R}\, dt \end{aligned} \tag{22}$$

so that :

$$\begin{aligned} p_i &= p_i' + \frac{\partial \hat{S}_1}{\partial q_i} \\ q_i &= q_i' - \frac{\partial \hat{S}_1}{\partial p_i'} \end{aligned} \tag{23}$$

We would like to stress that the need to use this method, rather than the simpler non canonical perturbation theory, arises from the particular properties of the problem. In the low altitude lunar satellite case, the medium period perturbations on h, k are comparatively large with respect to the initial values of the same variables; actually, in many mission analysis simulations, the initial value is $e = 0$, that is $h = k = 0$. Therefore it is not consistent to compute a first order perturbation by setting the values of the variables h, k in the right hand side at their initial values, otherwise most of the right hand side just disappears when initial $e = 0$. Truncation to degree 1 in eccentricity can be performed consistently also with the non canonical formalism, but truncation to order 1 in the perturbations to h, k result in an inconsistent approximation. We have therefore adopted a method which does not truncate the transformation, but fully accounts for the implicit nature of the transformation equations, without replacing old variables for the new ones in the right hand side. In essence, since the equations are linearised, our method is a form of the Newton's method normally used to prove the results of the KAM (Kolmogorov, Arnold, Moser) theory.

Having defined the formalism, the detailed computations are very similar to the ones performed for the long periodic perturbations. The generating function $\hat{S}_1$ is truncated to degree 2 in e, and again the number of terms is reduced to three for each harmonic l, m with l even, and two for each harmonic l, m with l odd. Then the derivatives of $\hat{S}_1$ are computed with respect to the non singular variables h', k and transformed into the derivatives with respect to η', ξ by means of the equations :

$$\frac{\partial \hat{S}_1}{\partial \eta} = \frac{1}{\sqrt{L}} \frac{\partial \hat{S}_1}{\partial h} - h \frac{\cot I}{\sqrt{L}\beta} \frac{\partial \hat{S}_1}{\partial I}$$

$$\frac{\partial \hat{S}_1}{\partial \xi} = \frac{1}{\sqrt{L}} \frac{\partial \hat{S}_1}{\partial k} - k \frac{\cot I}{\sqrt{L}\beta} \frac{\partial \hat{S}_1}{\partial I} \tag{24}$$

Then we can switch back to the usual non singular variables h, k by using :

$$k = \frac{\xi}{\sqrt{L}} + O(e^3); \qquad h = \frac{\eta}{\sqrt{L}} + O(e^3) \tag{25}$$

and the final transformation equations have the form :

$$A \begin{pmatrix} h' \\ k' \end{pmatrix} = B \begin{pmatrix} h \\ k \end{pmatrix} + C \tag{26}$$

which are implicit equations (because of the mixed variables appearing in the generating function $\hat{S}_1$), but linear (because of the truncation to degree 1 in h, k). Therefore they can be solved simply by :

$$\begin{pmatrix} h' \\ k' \end{pmatrix} = A^{-1} \left[B \begin{pmatrix} h \\ k \end{pmatrix} + C \right] \tag{27}$$

The 2x2 matrices A, B have as entries linear combinations of even harmonic coefficients, while the vector C contains linear combinations of odd harmonics :

$$A = \begin{pmatrix} 1 + \sum_{s=1} \sum_{m=1}^{2s} \frac{1}{-m\nu} C^-_{2s,m} \begin{bmatrix} \cos \\ \sin \end{bmatrix} \Psi_{2s,m} & 0 \\ -\sum_{s=1} \sum_{m=1}^{2s} \frac{1}{-m\nu} (A_{2s,m} - C^+_{2s,m} - W_{2s,m}) \begin{bmatrix} \sin \\ -\cos \end{bmatrix} \Psi_{2s,m} & 1 \end{pmatrix}$$

$$B = \begin{pmatrix} 1 & -\sum_{s=1} \sum_{m=1}^{2s} \frac{1}{-m\nu} (A_{2s,m} + C^+_{2s,m} - W_{2s,m}) \begin{bmatrix} \sin \\ -\cos \end{bmatrix} \Psi_{2s,m} \\ 0 & 1 + \sum_{s=1} \sum_{m=1}^{2s} \frac{1}{-m\nu} C^-_{2s,m} \begin{bmatrix} \cos \\ \sin \end{bmatrix} \Psi_{2s,m} \end{pmatrix}$$

$$C = \begin{pmatrix} \sum_{s=1} \sum_{m=1}^{2s+1} \frac{1}{-m\nu} D^+_{2s+1,m} \begin{bmatrix} -\sin \\ \cos \end{bmatrix} \Psi_{2s+1,m} \\ \sum_{s=1} \sum_{m=1}^{2s+1} \frac{1}{-m\nu} D^-_{2s+1,m} \begin{bmatrix} \cos \\ \sin \end{bmatrix} \Psi_{2s+1,m} \end{pmatrix}$$

where (introducing again compact notation) :

$$A_{2s,m} = \sigma \left(\frac{R_M}{a} \right)^{2s-2} J_{2s,m} F_{2s,m,s}(I) g^1_{2s,s,0} 2$$

$$C^+_{2s,m} = \sigma \left(\frac{R_M}{a} \right)^{2s-2} J_{2s,m} \left[F_{2s,m,s-1}(I) g^0_{2s,s-1,-2} + F_{2s,m,s+1}(I) g^0_{2s,s+1,2} \right] 2$$

$$C^-_{2s,m} = \sigma \left(\frac{R_M}{a} \right)^{2s-2} J_{2s,m} \left[F_{2s,m,s-1}(I) g^0_{2s,s-1,-2} - F_{2s,m,s+1}(I) g^0_{2s,s+1,2} \right] 2$$

$$D^+_{2s+1,m} = \sigma \left(\frac{R_M}{a} \right)^{2s-1} J_{2s+1,m} \left[F_{2s+1,m,s}(I) g^0_{2s+1,s,-1} + F_{2s+1,m,s+1}(I) g^0_{2s+1,s+1,1} \right]$$

$$D^-_{2s+1,m} = \sigma \left(\frac{R_M}{a} \right)^{2s-1} J_{2s+1,m} \left[F_{2s+1,m,s}(I) g^0_{2s+1,s,-1} - F_{2s+1,m,s+1}(I) g^0_{2s+1,s+1,1} \right]$$

$$W_{2s,m} = \sigma \left(\frac{R_M}{a} \right)^{2s-2} J_{2s,m} \cot I \, \frac{dF_{2s,m,s}(I)}{dI} g^0_{2s,s,0}$$

The medium periodic perturbations on H (hence I) and Ω are computed directly from :

$$H = H' + \frac{\partial \hat{S}_1}{\partial \Omega}; \qquad \Omega = \Omega' - \frac{\partial \hat{S}_1}{\partial H'} \tag{28}$$

In this case we do not need, to the level of approximation we are using, to account for the mixed variables appearing in $\hat{S}_1$, and therefore for the implicit nature of equations (28), because our theory is meant to be used for mission analysis, thus the accuracy in inclination does not need to be as high as in eccentricity (changes in $\sin I$ of a few 0.01 cannot result in hard landing). The theory could be improved by using the same Newton's method formalism –used for k, h'– also for the variables Ω, H', if this is required.

2.3.2 *Very short periodic perturbations*

The most important perturbations with a very short period are the ones on the semimajor axis; on the semimajor axis there are perturbations neither with long nor with medium periods. The very short periodic perturbations on h, k, I, Ω are small and therefore less important, but they can be accounted for by an analogous procedure if the accuracy requirements are strict.

The relevant very short periodic part of the perturbing function can be expanded as follows. Let us first introduce :

$$\Theta_{lmr} = r\lambda + m(\Omega - \theta) - \delta_{lm}; \quad \lambda = l + \omega; \quad r = l - 2p + q; \quad \frac{d\Theta_{lmr}}{dt} = rn - m\nu \tag{29}$$

Then (always truncating at the same level of approximation) :

$$\tilde{R} = \frac{GM}{a} \sum_{l=2} \left(\frac{R_M}{a}\right)^l \sum_{m=0}^{l} \sum_{p=0}^{l} F_{lmp}(I) \sum_{q=-1}^{1} g^0_{lpq} e^{|q|} J_{lm} \begin{Bmatrix} \cos \\ \sin \end{Bmatrix} (\Theta_{lmr} - q\omega) + O(e^2) \tag{30}$$

where the summation should be performed only for the terms with $r = l - 2p + q \neq 0$. Then, by using the usual Lagrange equation for the semimajor axis :

$$\frac{1}{a}\frac{da}{dt} = \frac{2}{na^2}\frac{\partial \tilde{R}}{\partial \lambda} \tag{31}$$

the very short periodic perturbations can be directly computed.

3. RESULTS

In order to illustrate the behaviour of a low lunar polar orbit, we have plotted in Fig. 2 the time evolution (for about one lunar month) of the osculating values of non-singular orbital elements h, k. All the three periodicities involved are clearly visible in the plot : the very short period changes (of the amplitude of $\approx 2 - 3 \times 10^{-4}$) are superimposed onto the medium period variations (amplitude $\approx 5 \times 10^{-3}$), while the long period variation appears as an overall trend roughly in the direction upper left to lower right corner of the plot. The data for this plot come from numerical integration in which use was made of the GEODYNE software system, with Konopliv et al. (1993) 60×60 model of the lunar gravity field (R. Floberghagen, private communication).

In order to assess the accuracy, reliability and efficiency of the algorithm we developed on the basis of the described theory, we have performed a great number of different tests. Here, however, we shall report only on two kinds of tests, those which in the most straightforward and compact way show the quality of our results. The first test consists of the computation of proper elements for each input record, containing osculating elements output from the numerical integrator. The result of the test can be assessed in the simplest way by computing the RMS of the deviations of the proper elements, as computed, from their average value. Since a perfect proper element should be exactly constant, this RMS measures the inaccuracy resulting from the truncations and approximations performed in the computation.

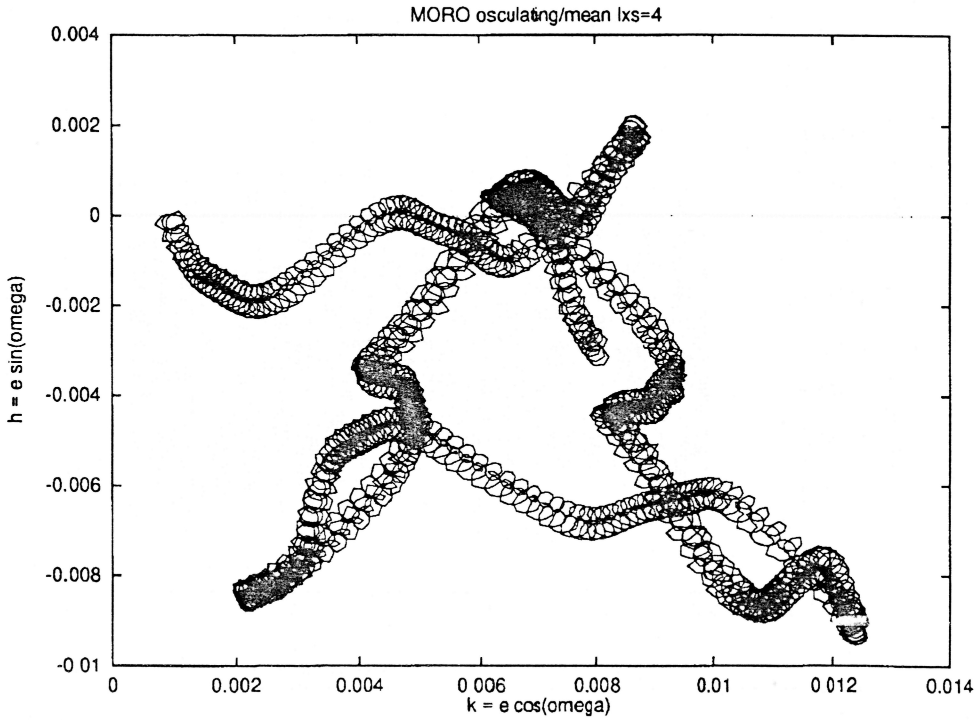


Fig. 2. Osculating h vs. k plane. Variations due to perturbations of three different periodicities are clearly distinguished in the plot. Numerical integration from GEODYNE; time span about one lunar month.

In Figure 3, we thus compare osculating and proper semimajor axis time variations; in Figure 4, the corresponding variations of the inclination are given, while in Figure 5, the same is shown for the eccentricity. Numerical integrations used in these plots are made with the USOC software system (G. Lecohier, private communication), and pertains to a polar orbiter with a mean altitude of 100 Km, initial eccentricity 0.02 and initial $\omega = 270°$. The lunar gravity potential used was the Lemoine et al. (1994) 70x70 model, which includes the Clementine tracking data. Integration included only the effects of the Moon (remember that our theory in the present version does not include the other effects, in particular the Earth). The selenographic longitude of the node corresponds to nearly $0°$ at date 2000/1/1/ 00 : 00 : 00. This orbit "decays" after 275 days (perilune goes below 20 km).

Although there are still some unremoved oscillations of the proper values left (in particular the trend of increase with time of the amplitude of proper eccentricity; see later), the overall conclusion is that the results are very good. The proper semimajor axis variations are astonishingly small ($\leq$ 13 meters only), while the rms's of both

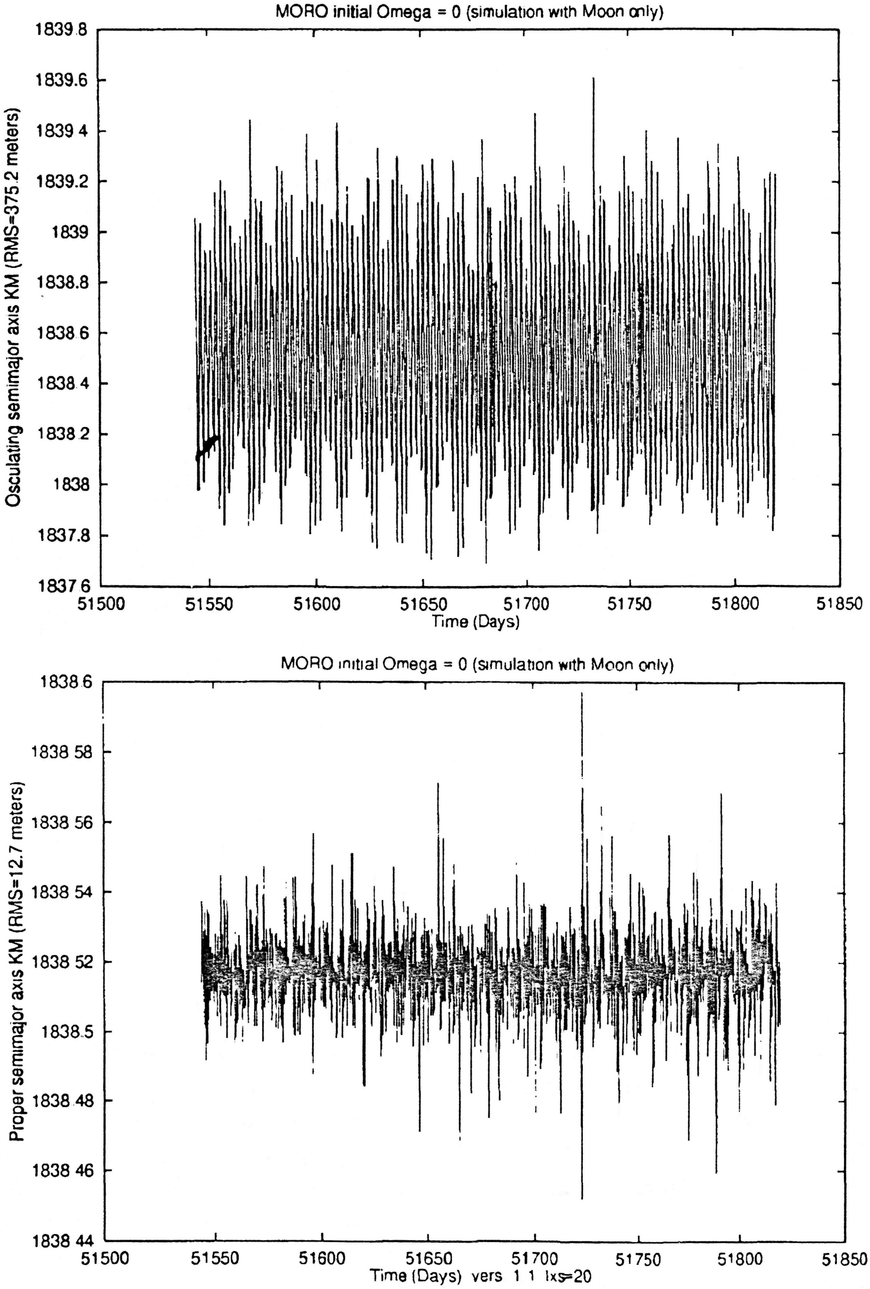
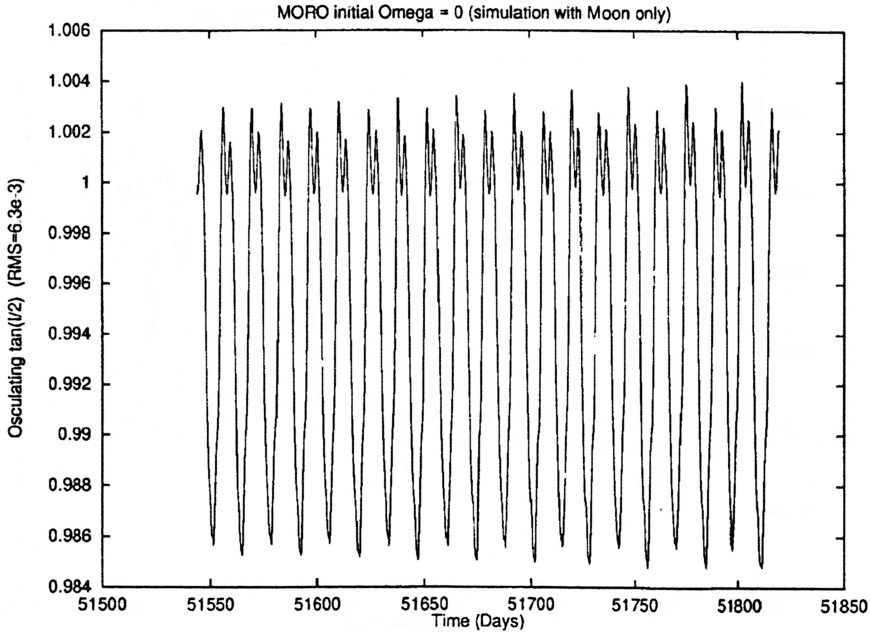


Fig. 3. Osculating (top), against proper (bottom) semimajor axis. The corresponding RMS's are given in the labels of y-axes; note the difference of the y-scales. Numerical integration from USOC, period 275 days.



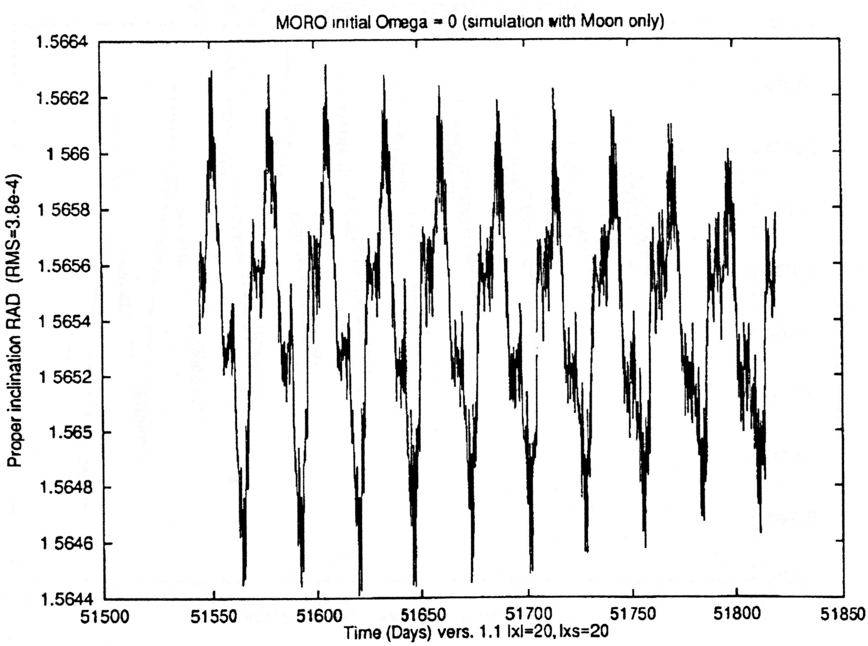
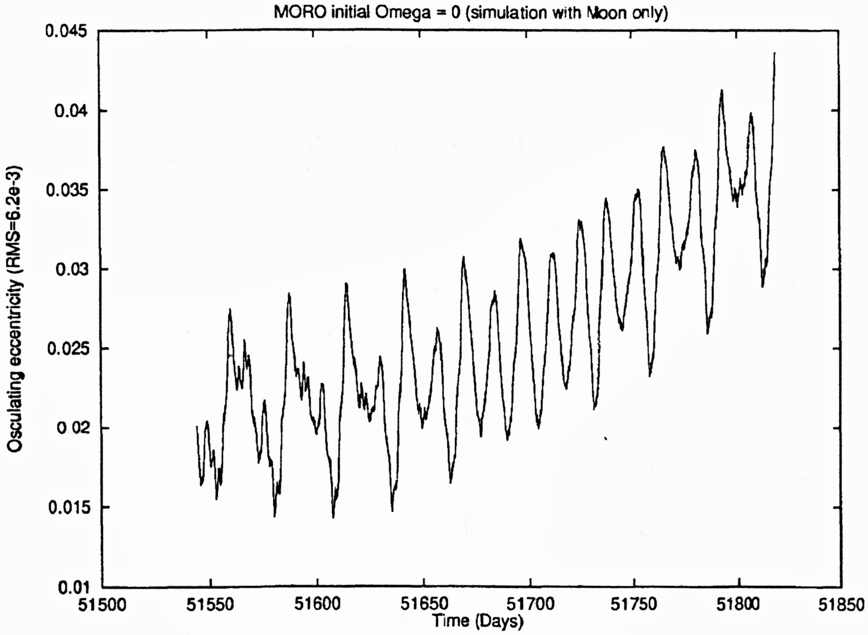


Fig. 4. The same as Figure 3, but for the inclination. Note that osculating values (top) are given in terms of the tan $I/2$, while proper ones (bottom) are in terms of the inclination itself.



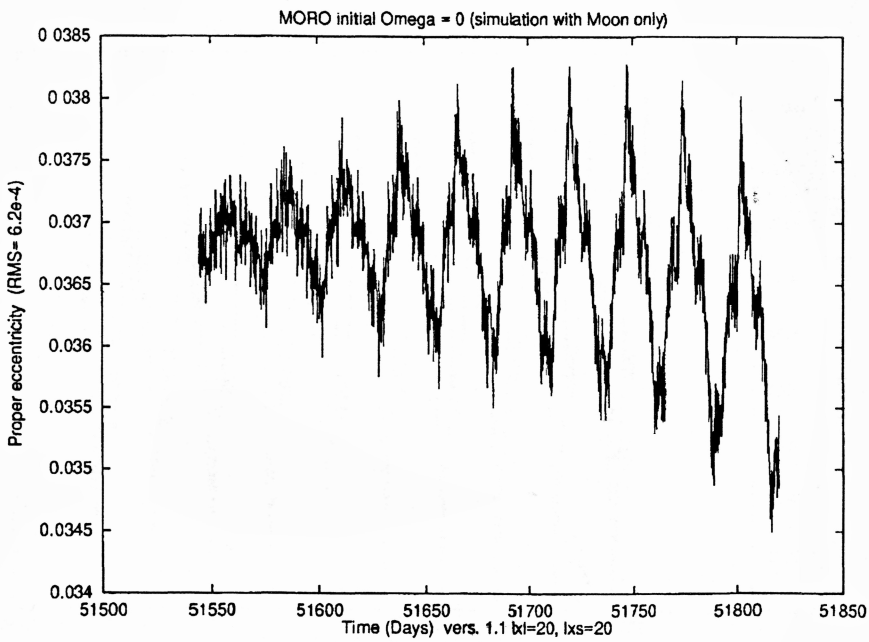


Fig. 5. The same as Figure 3, but for the eccentricity. The trend of increase with time of the amplitude of proper eccentricity is due to second order effects and/or to effects from harmonics of degree higher than 20.

proper inclination (in radians) and (more important) of proper eccentricity are less than 10^{-3}, which is more than enough for the mission analysis purposes.

The second test is even more demanding : we have computed analytically a solution for the same time span of the numerical test. What we do in fact is to compute the proper elements for the initial instant of the numerical integration, propagate them analytically for a span of time covered by integration, and recompute the osculating values for these instants of time for which the values are sampled in the numerical integration. Then we compute the difference between the analytical and the numerical solution.

In Figures 6 and 7 differences of the analytically propagated and the osculating values of h, k and $\tan I/2$ are shown, for the same USOC integration decribed above and lasting 275 days. Harmonics up to 20×20 are used in the computation, with coefficients taken again from the lunar gravity field model by Lemoine et al. (1994). The trend apparent in the differences of h, k is due to second order effects and/or to effects from harmonics of degree higher than 20. However, for about 6 months (more than enough with respect to typical duration of a mission) theory provides solution at the entirely satisfactory 0.001 level of accuracy in eccentricity. Hence, one can conclude that, although the analytical theory is not meant to provide precise ephemerides of the satellite, but only to study the qualitative long term behaviour of the orbit (e.g. for manoeuver planning purposes), this test shows its capability to actually predict the orbit in a qualitatively correct way and even quantitatively with a reasonable accuracy, more than presumably it is needed for the preliminary mission analysis required in the mission definition phase.

4. LIMITATIONS OF THE CURRENT THEORY

Let us in conclusion briefly discuss limitations of the current theory. We have already mentioned afore some open problems, which are responsible for the remaining errors and uncertainties of the presently achievable results, but represent at the same time the possibilities for future developments and improvement.

The theory described in this paper employes some assumptions and performs some truncations and simplifications with respect to a complete problem. The choices we have made correspond to the requirements arising from its use for the purpose of the preliminary mission analysis of a low lunar polar orbiter. However, these assumptions and simplifications ought to be explicitly stated, to be able to remove them if later the need arises for a theory capable of higher accuracy and/or more general applicability. In this section we list all these limitations, together with a few comments on what should be done to remove each one of them.

As for the effects not included in the current version, perturbations by the Earth are not accounted for, but are of a size relevant for a more accurate solution. Given the theory available from Kaula (1962), we anticipate no special difficulty in including these effects both in the long and in the short periodic perturbations. Gravitational perturbations from the Sun are smaller (by a factor $\simeq 200$) than those of the Earth, therefore they will not be needed, unless an extremely accurate theory is required for real time operations. On the contrary the effects of radiation pressure can be relevant,

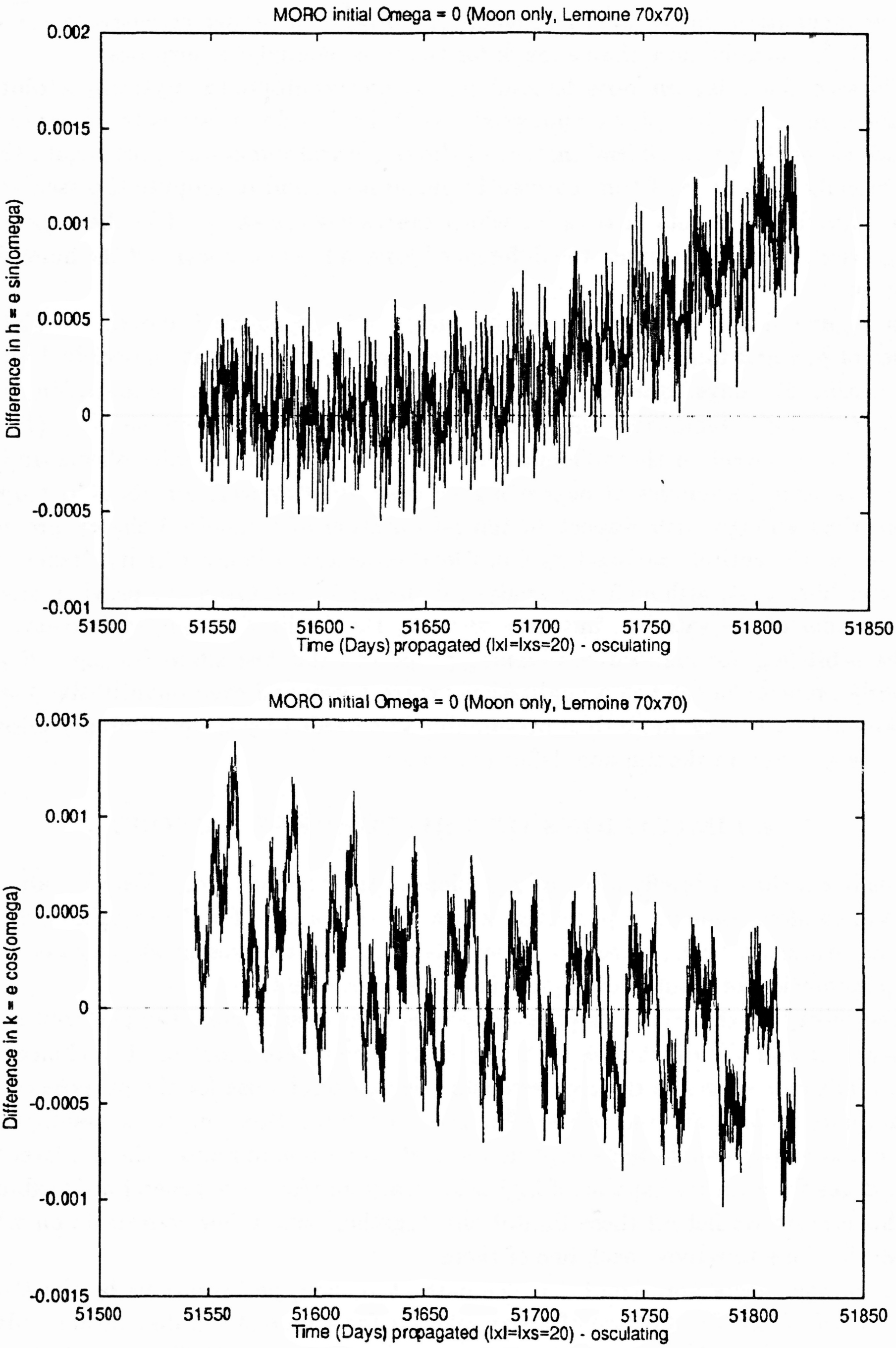


Fig. 6. Differences of the analytically propagated values of h (top) and k (bottom) and their counterparts coming from the numerical integration.

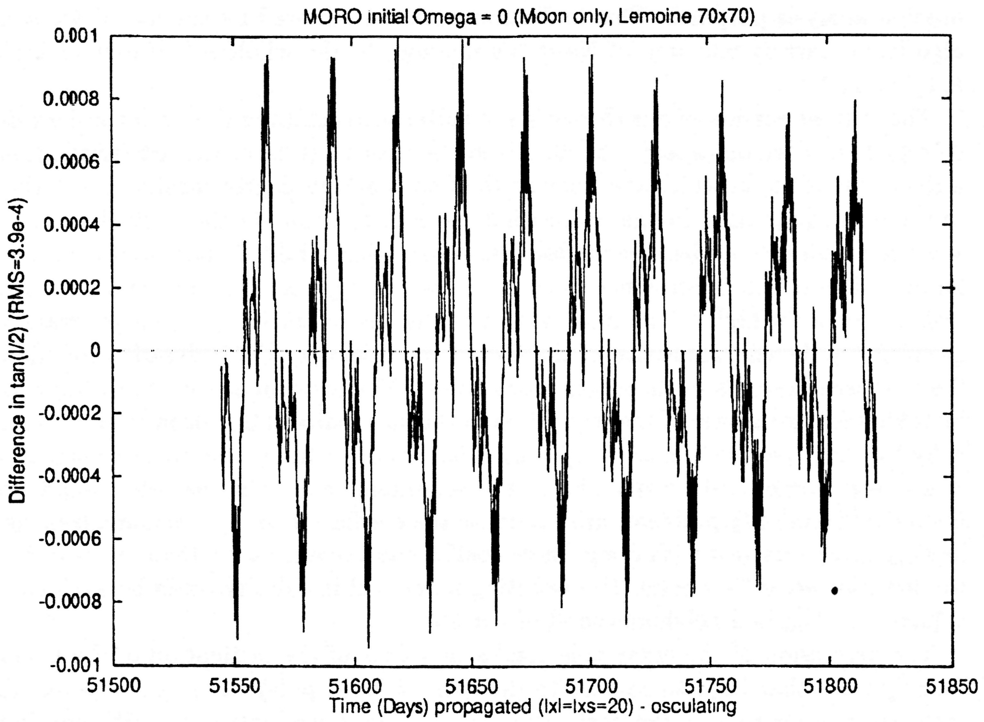


Fig. 7. The same as Figure 6, but for the tan $I/2$.

at least when the lunar satellite undergoes eclipses (Milani et al., 1987); in one of the test we performed, the change in proper semimajor axis due to the long periodic effects of radiation pressure accumulates to $\simeq 50\,m$. An analytic theory of radiation pressure would be possible, but requires additional effort.

By using the semianalytical integration (that is, numerical integration of the analytical perturbation equations), and comparing it with the results of a purely analytical propagation, we have been able to measure the size of the neglected second order effects. This difference has a medium periodic component (mostly $m = 1, 2, 3$) and a long periodic component, which accumulates to about 0.001 in h, k over 275 days. This source of error was considered unimportant at this stage of development of mission analysis tools, because the uncertainty in the lunar potential results in a much larger uncertainty in the long term behaviour. However, the inclusion of the main second order long periodic effects is certainly a worthwhile upgrade of our theory, which would become necessary when a better model of the lunar gravity field will be available.

As pointed out in Section 2 (see comment with eq. (28)), we have chosen to replace I' with I in the mixed variable generating function, when the perturbations on I and

Ω are computed. This results in lower accuracy with respect to the full solution of the implicit equations, which was adopted for the h, k variables. This choice is justified by the lower accuracy required in the inclination, with respect to the eccentricity, for mission analysis purposes. This limitation could be removed by the use of the same algorithm, namely one step of Newton's method, to the whole set of four variables $h, k, \Delta I, \Delta\Omega$.

The current version of our theory is not suitable to compute the perturbations due to very high harmonics, e.g. $l > 40$. Given the present state of the art for the lunar potential models, in which the harmonics of such a high degree mostly reflect the a priori constraints used in the collocation process, to compute the perturbations up to such a high l would be meaningless. However, when a reliable potential model will be available, it will become necessary to ensure enough performance and numerical stability even for high l. The main source of the loss of efficiency of our program for growing l is the enormous size of the files of coefficients of the inclination functions; the number of records in these files grows like l^4. The performance could be improved by taking full advantage of the truncation in the harmonics of the mean anomaly (that is by limiting the size of $r = l - 2p + q$). The loss of accuracy due to this truncation is not very large, and mostly affects the semimajor axis. On the orher hand, the numerical instability problems arise because the coefficients of the monomials in $\sin I$ in F_{lmp} grow very fast with l; e.g. these coefficients become larger than the inverse of the machine error for $l \simeq 50$. The resulting numerical instability could be avoided by expanding F_{lmp} in a neighbourhood of $I = 90°$.

The precession of the lunar pole results in a drift of the inclination in the true of date system (that is, with respect to the current lunar pole). After a few years, the inclination appearing in the coefficients $F_{lmp}(I)$ becomes noticeably different from the one in the true of epoch system we are using, and this results in a degradation of the solutions, because of a less accurate removal of medium periodic perturbations (mostly $m = 1$). This could be improved, if the need arises for a longer time span to be covered by mission analysis software, by explicitly accounting for this effect.

At the moment we are computing the very short periodic perturbations only for the semimajor axis. The very short periodic perturbations on h, k are less than 0.0005, thus their computation is not needed for the present requirements. However, the theory is available and these terms can be easily added if the need arises. We do not compute the perturbations on the mean anomaly, because we think that this element is not really used at all in mission analysis; again, this computation could be done if needed.

The approximation by which the coefficients of the long periodic equations (7) can be considered constant is not consistent for non polar orbits; the changes are of the order of $e \cos I$, hence they are of the second order if $\cos I$ is of the order of e. The very notion of a frozen orbit must be taken with some reserve for an orbit with low inclination.

As a matter of principle, most of our theory could be applicable to satellites of other bodies, including asteroids and comets. However, the modifications required are far from trivial, and they depend essentially on the ratios between the orbital and the rotation frequencies.

References

Davies, M. E., Abalakin, V. K., Brahic, A., Bursa, M., Chovitz, B. H., Lieske, J. H., Seidelmann, P. K., Sinclair, A. T. and Tjuflin, Y. S. : 1992, Report of the IAU/IAG/COSPAR working group on cartographic coordinates and rotational elements of the planets and satellites, *Celestial Mechanics*, **53**, 377–397.

Kaula, W. M. : 1962, Development of the lunar and solar disturbing functions for a close satellite, *A*stron. J. **67**, 300-303.

Kaula, W. M. : 1966, *Theory of Satellite Geodesy*, Blaisdell Publishing Company, Waltham, Toronto and London.

Konopliv, A. S., Sjogren, W. L., Wimberly, R. N., Cook, R. A. and Vijaraghavan, A. : 1993, A High Resolution Lunar Gravity Field and Predicted Orbit Behaviour *A*AS Paper, 93-622. Presented at the 1993 AAS/AIAA Astrodynamics Specialist Conference, Victoria, B.C..

Lemoine, F. G., Smith, D. E. and Zuber, M.T . : 1994, A 70th degree and order gravity model for the Moon, P11A-9, EOS, *T*rans. American Geophys. Union **75**, No. 44.

Milani, A. and Knežević, Z. : 1990, Secular Perturbation Theory and Computation of Asteroid Proper Elements. *Celestial Mechanics*, 49, 347–411.

Milani, A. and Knežević, Z. : 1992, Asteroid Proper Elements and Secular Resonances. *I*carus, **98**, 211–232.

Milani, A. and Knežević, Z. : 1994, Asteroid Proper Elements and the Dynamical Structure of the Asteroid Main Belt, *I*carus, 107, 219–254.

Milani, A., Nobili, A. M., and Farinella, P. : 1987, *Non gravitational perturbations and satellite geodesy*, A. Hilger, Bristol.